Evolving Robot Empathy through the Generation of Artificial Pain in an Adaptive Self-Awareness Framework for Human-Robot Collaborative Tasks

Muh Anshar

Faculty of Engineering and Information Technology
University of Technology Sydney

This dissertation is submitted for the degree of
Doctor of Philosophy

March 2017

Bismillahirrahmanirrahim

All Praise and Gratitude to the Almighty God, **Allah SWT**, for His Mercy and Guidance which have given me strength and tremendous support to maintain my motivation from the very beginning of my life journey and into the far future.

I would like to dedicate this thesis to my love ones, my wife and my son,

Nor Faizah & Abdurrahman Khalid Hafidz

for always being beside me which has been a great and undeniable support throughout my study…

CERTIFICATE OF ORIGINAL AUTHORSHIP

This thesis is the result of a research candidature conducted jointly with another University as part of a collaborative Doctoral degree. I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

Date: 13 March 2017

Muh Anshar
March 2017

Acknowledgements

I would like to acknowledge and thank my Principal Supervisor, Professor Mary-Anne Williams for her great dedication, support and supervision throughout my PhD journey. I would also like to thank the members of the Magic Lab for being supportive colleagues during my study. Many Thanks also to my proofreader, Sue Felix, for the fruitful comments and constructive suggestions. In addition, I acknowledge the support of the Advanced Artificial Research Community (A2RC), Electrical Engineering, University of Hasanuddin - UNHAS Makassar Indonesia, which was established in early 2009 as a manifestation of the research collaboration commitment between the UTS Magic Lab and the UNHAS A2RC Community.

Abstract

The application and use of robots in various areas of human life have been growing since the advent of robotics, and as a result, an increasing number of collaboration tasks are taking place. During a collaboration, humans and robots typically interact through a physical medium and it is likely that as more interactions occur, the possibility for humans to experience pain will increase. It is therefore of primary importance that robots should be capable of understanding the human concept of pain and to react to that understanding. However, studies reveal that the concept of human pain is strongly related to the complex structure of the human nervous system and the concept of *Mind* which includes concepts of *Self-Awareness* and *Consciousness*. Thus, developing an appropriate concept of pain for robots must incorporate the concepts of *Self-Awareness* and *Consciousness*.

Our approach is firstly to acquire an appropriate concept of self-awareness as the basis for a robot framework. Secondly, it is to develop an internal capability for a framework for the the internal state of the mechanism by inferring information captured through internal and external perceptions. Thirdly, to conceptualise an artificially created pain classification in the form of synthetic pain which mimics the human concept of pain. Fourthly, to demonstrate the implementation of synthetic pain activation on top of the robot framework, using a reasoning approach in relation to past, current and future predicted conditions. Lastly, our aim is to develop and demonstrate an empathy function as a counter action to the kinds of synthetic pain being generated.

The framework allows robots to develop "self-consciousness" by focusing attention on two primary levels of self, namely subjective and objective. Once implemented, we report the results and provide insights from novel experiments designed to measure whether a robot is capable of shifting its "self-consciousness" using information obtained from exteroceptive and proprioceptive sensory perceptions. We consider whether the framework can support reasoning skills that allow the robot to predict and generate an accurate "pain" acknowledgement, and at the same time, develop appropriate counter responses.

Our experiments are designed to evaluate synthetic pain classification, and the results show that the robot is aware of its internal state through the ability to predict its joint motion and produce appropriate artificial pain generation. The robot is also capable of

alerting humans when a task will generate artificial pain, and if this fails, the robot can take considerable preventive actions through joint stiffness adjustment. In addition, an experiment scenario also includes the projection of another robot as an object of observation into an observer robot. The main condition to be met for this scenario is that the two robots must share a similar shoulder structure. The results suggest that the observer robot is capable of reacting to any detected synthetic pain occurring in the other robot, which is captured through visual perception. We find that integrating this awareness conceptualisation into a robot architecture will enhance the robot's performance, and at the same time, develop a self-awareness capability which is highly advantageous in human-robot interaction.

Building on this implementation and proof-of-concept work, future research will extend the pain acknowledgement and responses by integrating sensor data across more than one sensor using more sophisticated sensory mechanisms. In addition, the reasoning will be developed further by utilising and comparing the performance with different learning approaches and different collaboration tasks. The evaluation concept also needs to be extended to incorporate human-centred experiments. A major possible application of the proposal to be put forward in this thesis is in the area of assistive care robots, particularly robots which are used for the purpose of shoulder therapy.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter presents an overview of the background to the study followed by the currently identified issues in the field of human-robot interactions and related fields. The chapter then provides a brief introduction to the proposed means of addressing these issues, together with the experimental setup, followed by the analysis and outcomes of the findings. The significance and contribution of the work are given, together with a short description of future related work, followed by the overall structure of the thesis.

1.1 Overview of the Study Background

As the number of robots applications in various areas of human life increases, it is inevitable that more collaborative tasks will take place. During an interaction, humans and robots commonly utilise a physical medium to engage, and the more physical the interaction is, the greater the possibility that robots will cause humans to experience pain. This possibility may arise from human fatigue, robot failure, the working environment or other contingencies that may contribute to accidents. For instance, take the scenario in which robots and humans work together to lift a heavy cinder block. Humans may experience fatigue due to constraints placed on certain body muscles, and over time, this muscle constraint may extend beyond its limit. An overload constraint on muscle degrades the muscle strength and in time introduces damage to internal tissue, leading to the experience of pain. Humans occasionally communicate this internal state verbally or through facial expression. It is of primary importance for robots to consider these sophisticated social cues, capture them and translate them into useful information. Robots can then provide appropriate counter-responses that will prevent humans from experiencing an increase in the severity of pain. Furthermore, robots may play a significant role in anticipating and preventing work accidents from happening.

Having the capability to acknowledge pain and develop appropriate counter responses to the pain experience by the human peer will improve the success of the collaboration. Failure to acknowledge this important human social cue may cause the quality of the interaction to deteriorate and negatively affect the acceptance of future robot applications in the human environment.

1.2 Current Issues

Literature studies show that there are a considerable number of works that have investigated the emergence of robot cognition and have proposed concepts of the creation of conscious robots. However, there are very few studies that acknowledge pain and those studies only use the terminology to refer to robot hardware failure without real conceptualisation of pain. The studies do not correlate the importance of evolving a concept of pain within the robot framework with developing reactions in response to the identified pain. At lower levels of perception, robots rely only on their proprioceptive and exteroceptive sensors, which are limited to building their external and internal representations. Not all robots have uniform sensory and body mechanisms, which consequently, it affects the quality of pain information retrieval and processing. In contrast, humans have a rich and complex sensory system which allows robust pain recognition and the generation of empathic responses. Studies reveal that concepts of self-awareness, pain identification and empathy with pain are strongly attached to the cognitive aspect of humans, who have vast and complex nerve mechanisms (Goubert et al., 2005; Hsu et al., 2010; Lamm et al., 2011; Steen and Haugli, 2001).

These factors present huge challenges to the notion of developing robots with social skills that can recognise human pain and develop empathic responses. Thus, it is of key importance to develop an appropriate concept of self and pain to incorporate in a robot's framework that will allow the development of human pain recognition.

1.3 Description of Proposed Approach

There are five main objectives of this work. The first is to develop an appropriate concept of self-awareness as the basis of a robot framework. The proposed robot self-awareness framework is implemented on robot cognition, which focuses attention on the two primary levels of self, namely subjectivity and objectivity, derived from the human concept of self proposed by Lewis (1991). It should be pointed out that robot cognition in this work refers to the change in the focus of attention between these levels, and does not necessarily refer to 'human consciousness'. The second is to develop the internal state of the mechanism over time by inferring information captured through internal and external perceptions. The construction of internal process is based on current and future predicted states of the robot that are captured through the robot's proprioceptive perception. When an interaction takes place, the information captured by the robot's exteroceptive perception is also used to determine the internal state. The third is to conceptualise artificial pain for the robot through a set of synthetic pain categories, mimicking human conceptualisation of pain. Fault detection provides the stimulus function and defines classified magnitude values which constitute the generation of artificial pain, which is recorded in a dictionary of synthetic pain. The fourth is to demonstrate the generation of synthetic pain through a reasoning process of the robot's internal state with respect to the current and predicted robot information captured from proprioceptive perception and the aim of the overall task. The final objective is to develop an appropriate counter-response, mimicking the empathy function, to the generated synthetic pain experienced by the robot.

To briefly describe how the robot mind functions: the framework develops a planning scheme by reasoning the correlation of the robot's current internal states with the robot's belief, desire and intention framework. The robot framework determines the type of synthetic pain to be generated, which the robot experiences. Whenever the pain intensity increases, the framework switches its attention to the subjective level, giving priority to the generation of empathy responses to the synthetic pain and disregarding the objective level of the task. In other words, the robot framework manifests the concept of self by actively monitoring its internal states and external world, while awareness is implemented by shifting the focus of attention to either the subjective or the objective level. At the same time, the reasoning process analyses the information captured by the robot's perceptions with respect to the dictionary of synthetic pain embedded in the framework.

Embedding this ability into the robot's mechanism will enhance the robot's understanding of pain, which will be a useful stepping stone in developing the robot's social skills for recognising human pain. This ability will allow robots to work robustly to understand human expressions during collaborative tasks, particularly when the interaction might lead to painful experiences. This framework will equip the robot with the ability to reconfigure its focus of attention during collaboration, while actively monitoring the condition of its internal state. At the same time, the robot will be capable of generating appropriate synthetic pain and generating associated empathic responses. These empathic responses are designed to prevent robots from suffering catastrophic hardware failure, which is equivalent to an increase in the intensity of the pain level.

1.4 Brief Description of Experiments

Two types of experiment are designed to demonstrate the performance of the robot framework. The first involves one robot and a human partner interacting with each other in a hand pushing task which produces a sequence of arm joint motion data. This type has two scenarios, namely, offline and online scenarios. In the offline scenario, two experiments are carried out in which the first stage is dedicated to recording the arm joint motion data, which will be going to be stored in a database. In the second stage, the data are taken from the database and fed into the robot's mind in the second stage (i.e., as a simulation in the robot's mind). In the online scenario, the data are obtained directly from the hand pushing task and fed to the robot's mind for further processing.

The second type of experiment involves two robots and a human partner. An observer robot is assigned a task to observe another robot, acting as a mediator robot, which is involved in an interaction with the human partner. There are two stages in this experiment: stage one serves as an initiation or calibration stage, and stage two is the interaction stage. The initiation stage sets the awareness region of the mind of the observer robot and the joint restriction regions for both robots that should be avoided. These joint restriction regions contain robot joint position values which correspond to the faulty joint settings. This stage is also dedicated to calibrating the camera position of the observer robot towards the right arm position of the mediator robot. A red circular shape attached to the back of the right hand of the mediator robot is used as a marker throughout the experiments. The second stage comprises two experiments, robot self-reflection and robot empathy. During the self-reflection experiment, both robots are equipped with an awareness framework, with the exception that the mediator robot does not have an activated consciousness direction function. The final experiment applies the same settings, with the addition of the activation of counter-response actions that simulate the function of the empathy response.

1.5 Contributions and Significance

There are a minimum of four contributions identified by this study:

1. The conceptualisation of robot self-awareness by shifting the focus of attention between two levels of self, namely subjective and objective.

2. A dictionary of artificial robot pain containing a set of synthetic pain class categories.

3. The integration of high reasoning skills within the internal state framework of the robot.

4. The derivation of a novel concept of empathy responses towards synthetic pain for a robot, which is essential for engaging in collaborative tasks with humans.

The significance of the study is that it mostly affects the creation of a cognitive robot and the future coexistence of humans and robots through:

1. Proposing a concept of robot self-awareness, by utilising a high reasoning-based framework.

2. Promoting the importance of self-development within robot internal state representation.

3. Promote a better acceptance of robots in a human-friendly environment, particularly in collaborative tasks.

1.6 Future Development

Four aspects of development will be addressed in respect of current achievements. The first is various sensor utilization which provides complex information for the framework to handle, and the implementation of machine learning approaches to increase the framework reasoning capability. The second addresses awareness regions of the framework and other kinds of synthetic pain, which have not previously been explored. The third highlights the proof-of-concept with the focus on human-centred experiments which serve as task performance assessment. The assessment sets a predefined scenario of human-robot interaction, and human volunteers are involved in assessing the robot's performance. The last aspect is to look into possible real implementation in health care services.

1.7 Structure of Thesis

The structure of the thesis is as follows: Chapter 2 presents a review of the literature that forms the foundation of the work, divided into two main categories. Literature in the the first category discusses motion planning for robots, which focuses on lower level planning and higher level planning. Studies in the second category deal with the metaphysical aspect of the robot, which centres on human cognition, covering the concept of mind, self-awareness, pain and empathy, and the development of the robot empathy concept.

The conceptual foundation of the proposal, which discusses the elements of perception, artificial pain and empathic response, is presented in Chapter 3. The description of perception is divided according to the origin of the sensory data followed by the artificial pain proposal for robots. This chapter also presents how pain levels can be designed, along with the activation procedures and mathematical representation, regardless of whether a simplified method or a more complex approach is used. The concept of robot empathy generation is presented and includes details of how this approach can be implemented, and the mathematical analysis.

Chapter 4 discusses the Adaptive Self-Awareness Framework for Robot, together with several key elements of the framework. The discussion covers a wide range of aspects of each element, including the mathematical representations of retrieved perception data which are arranged into pattern data sequences.

A practical implementation as a proof of concept is highlighted in Chapter 5 which focuses on description of the robot hardware and experimental settings. A humanoid-based robot is used as the experiment platform and a human-robot-interaction as the medium for assessing the technical performance of the robot system.

Chapter 6 provides the outcomes of the experiments conducted in the previous chapter, followed by analysis and discussion of the results. All data are obtained from the module in the framework which is responsible for retaining all incoming data from the sensory mechanisms, pre-recorded synthetic pain values, processed data and output of the robot mind analyses.

Chapter 7 concludes the thesis. It highlights the fundamental achievements of the experiments. It also previews future work, which might include such aspects as more sophisticated data integration from different sensors and possible future implementation in assistive care robots for aiding people with disability.

Chapter 2

Robot Planning and Robot Cognition

This chapter discusses two aspects of robot development covered in the literature in the field of robot planning, particularly in motion planning, and robot cognition, and presents a thorough discussion of the cognitive element of the robot.

2.1 Motion Planning

The discussion of robot motion planning falls into two major categories, stimulus-based planning and reasoning-based planning. Stimulus-based planning concerns planning approaches that originate from the stimulus generated at the low level of robot hardware, while reasoning-based planning focuses on the higher level of data processing.

2.1.1 Stimulus-based Planning

Stimulus-based planning centres on fault detection in robot hardware, which utilises robot proprioceptive and exteroceptive sensors to detect and localise a fault when it occurs. Early studies reported in Elliott Fahlman (1974), Firby (1987), Koditschek (1992) promote the importance of incorporating a failure recovery detection system into robot planning mechanisms. Firby (1987) proposed the very first planner for a robot, embedded in the reactive action package. The proposal does not give an adequate representation of the robot's internal state; rather, the planner centres more on the stimuli from the robot's environment or reactive basis. Further study on failure recovery planning is reported in Tosunoglu (1995); however, this work proposes a planning scheme that relies only on the stimuli received from fault tolerant architecture, which is still a reaction-based approach. A small development was then proposed by Paredis and Khosla (1995). The authors developed a manipulator trajectory plan for the global detection of kinematic fault tolerance which is capable of avoiding violations of secondary kinematics requirements. The planning algorithm is designed to eliminate unfavourable joint positions. However, it is a pre-defined plan and does not include the current state of the manipulator. Ralph and Pai (1997) proposed fault tolerant motion planning utilising the least constraints approach, which measures motion performance based on given faults obtained from sensor readings. The proposal is processed when a fault is detected and the longevity measure constructs a recovery action based on feasible configurations. Soika (1997) further examined the feasibility of sensor failure, which may impair a robot's ability to accurately develop a world model of the environment.

In terms of multi-robot cooperation, addressing the issues mentioned above are extremely important. If the internal robot states are not monitored and are disregarded in the process of adjusting robot actions for given task, replanning when faults occur will result in time delay. This situation will eventually raise issues which may deter robot coordination. A multi-robot cooperation in Alami et al. (1998) failed to consider this problem. According to Kaminka and Tambe (1998), any failure in multi-agent cooperation will cause a complex explosion of state space. Planning and coordination will be severely affected by countless possibilities of failure. Studies conducted in Hashimoto et al. (2001) and Jung-Min (2003) focus on the reactive level; the former authors address fault detection and identification, while the latter stresses the need for recovery action after a locked joint failure occurs. Another work reported in Hummel et al. (2006) also focuses on building robot planning on vision sensors, to develop a world model of the robot environment. Fagiolini et al. (2007) in multi-agent systems-based studies proposed a decentralised intrusion approach to identify possible robot misbehaviour by using local information obtained from each robot, and reacted to this information by proposing a new shared cooperation protocol. The physical aspect of human-robot interaction is very important as it concerns safety procedures. A review by De Santis et al. (2008) mentions that safety is a predominant factor that should be considered in building physical human-robot interaction. Monitoring possible hardware failure is made achieveable by the ability of the planning process to integrate the proprioceptive state of robots during interactions. By having updated information, robots are able to accurately configure and adjust their actions in given tasks, and at the same time, to communicate adjustment actions to their human counterparts. Hence, both parties are aware of the progress of the interaction. A study by Scheutz and Kramer (2007) proposed a robust architecture for human-robot interaction. This study signifies the importance of detecting hardware failure and immediately generating post recovery actions. A probabilistic reasoning for robot capabilities was proposed in Jain et al. (2009). The proposal targeted the achievement of capability to anticipate possible failures and generate a set of plausible actions which would have a greater chance of success. Ehrenfeld and Butz (2012) discussed sensor management in the sensor fusion area in relation to fusion detection. Their paper focuses on detecting sensor failure that is due to hardware problems or changes within the environment. A recent study reported by Yi et al. (2012) proposes a geometric planner which focuses on detecting failure and replanning online. The planner functionality is still a reaction-based failure detection.

2.1.2 Reasoning-based Planning

Reasoning-based planning is higher level planning. In this sub-section, we discuss the internal state representation of robots and artificial intelligence planning in general.

Internal State Representation Framework

In higher level planning, robots are considered to be agents, and to represent an agent's internal state requires rationality. One of the most well-recognised approaches to representing an agent's internal is the Belief (B), Desire (D) and Intention (I) framework. Georgeff et al. (1999) refer to Belief as the agent's knowledge which contains information about the world, Desire sets the goals that the agent wants to achieve, and Intention represents a set of executable actions. According to Rao and Georgeff (1991), the Belief-Desire-intention (BDI) architecture has been developed since 1987 by the work of Bratman (1987), Bratman et al. (1988) and Georgeff and Pell (1989). The latter's paper presents the formalised theory of BDI semantics by utilising the Computation Tree Logic form proposed by Emerson and Srinivasan (1988). However, this earlier development of intelligence has received criticism as reported in Kowalski and Sadri (1996) which quotes the argument by Brooks (1991) that an agent needs to react to the changes within that agent's environment. Kowalski and Sadri (1996) proposed a unification approach which incorporates elements of rationality and reactivity into the agent architecture. Busetta et al. (1999) proposed an intelligent agent framework based on the BDI model JACK, which integrates reactive behaviours such as failure management into its modular-based mechanism. Braubach et al. (2005) claimed that the available BDI platforms tend only to abstract the goal without explicit representation. The authors point out several key points that are not well addressed in BDI architecture planning, which is the explicit mapping of a goal from analysis and design to the implementation stage. The important feature of the proposal is the creation of context which determines whether a goal action is to be adopted or suspended. In the same year, Padgham and Lambrix (2005) formalised the BDI framework with the ability to influence the intentions element of the agent. This extension of the BDI theoretical framework has been implemented in the updated version of the JACK framework. Another development platform, named JASON, presented in Bordini and Hübner (2006), utilises an extended version of agent-oriented logic programming language inspired by the BDI architecture. The paper provides an overview of several features of JASON, one of which is failure handling. However, it does not involve the semantics implementation of a failure recovery system. Still within the same BDI agent framework, Sudeikat et al. (2007) highlighted the validation criterion for BDI-based agents and proposed an evaluation mechanism for asserting the internal action of an agent and the communication of events between the involved agents. The assertion of internal action of an agent relies only on agent performance. Gottifredi et al. (2008, 2010) reported an implementation of BDI architecture on the robot soccer platform. The authors addressed the importance of a recovery failure capability integrated into their BDI-based high level mobile robot control system to tackle adverse situations. Error recovery planning was further investigated by Zhao and Son (2008) who proposed an extended BDI framework. This framework was developed to mitigate improper corrective actions proposed by humans as a result of inconsistency in human cognitive functions resulting from increased automation that introduces complexity into tracking activity. An intelligent agent should have learning capabilities and this is not addressed in the BDI paradigm. Singh et al. (2010) conducted a study, later known to be the earliest study to address the issue, that introduced decision tree-based learning into the BDI framework. This proposal targeted planning selection, which is influenced by the success probability of executed experiences. Any failure is recorded and used to shape the confidence level of the agent within its planning selection. A further study in Singh et al. (2011) integrates dynamic aspects of the environment into the plan-selection learning of a BDI agent. The study demonstrates the implementation of the proposed dynamic confidence measure in plan-selection learning on an embedded battery system control mechanism which monitors changes in battery performance. A recent study carried out by Thangarajah et al. (2011) focuses on the behaviour analysis of the BDI-based framework. This analysis considers the execution, suspension and abortion of goal behaviour which have been addressed in the earlier study reported in Braubach et al. (2005). Cossentino et al. (2012) developed a notation which covers the whole cycle process from analysis to implementation by utilising the Jason interpreter for agent model development. The proposed notation does not address issues of failure recovery; rather, it focuses on the meta-level of agent modelling.

Artificial Intelligent - AI Planning

According to McDermott (1992), robot planning consists of three major elements, namely automatic robot plan generation, the debugging process and planning optimisation. The author points out that constraints play an important role by actively acting as violation monitoring agents during execution. Planning transformation and learning are also crucial elements to include in robot planning. Two of the earliest studies conducted on AI-based task planning, which have become the best-known methods, are reported in Fikes and Nilsson (1972) and Erol et al. (1994). Fikes and Nilsson (1972) proposed the STandford Research Institute Problem Solver (STRIPS) and the study reported in Erol et al. (1994) classifies several different works as the Hierarchical Task Network (HTN), which is decomposition-based. The STRIPS develops its planning linearly with respect to the distance measurement of the current world model from the target. The drawback of this method is that state space explosions occur as more complicated tasks are involved, which is counter-productive. Sacerdoti (1975) argued that regardless of the linearity of execution, the plan itself by nature has a non-linear aspect. The author instead proposed the Nets of Action Hierarchies (NOAH), which are categorised according to the family of HTN-based approaches. The development of a plan in NOAH keeps repeating in the simulation phase in order to generate a more detailed plan, and is followed by a criticising or reassesment phase through processes of reordering or eliminating redundant operations. This work is an advancement of the work on the HACKER model, developed by Sussman (1973), which replaces destructive criticism with constructive criticism to remove the constraints on plan development. Another comparison made by Erol et al. (1996) points out that STRIP-based planners maximise the search of action sequences to produce a world state that satisfies the required conditions. As a result, actions are considered as a set of state transition mapping. HTN planners, in contrast, consider actions as primitive tasks and optimise the network task through task decomposition and conflict resolution. The HTN-style planner NONLIN introduced by Tate (1977) incorporates a task formalism that allows descriptive details to be added during node linking and expansions. In contrast to NOAH, the NONLIN planner has the ability to perform backtracking operations.

Current advancement in AI planning has been directed towards utilisation of proportional methods (Weld, 1999), which generalizes the classical AI planning into three descriptions:

1. Descriptions of initial states
2. Descriptions of goals
3. Descriptions of possible available actions - domain theory

One major AI planning achievement was a proposal made by Blum and Furst (1997), the two-phase GRAPHPLAN planning algorithm, which is a planning method in STRIPS-like domains. The GRAPHPLAN approaches a planning problem by alternating graph expansion and solution extraction. When solution extraction occurs, it performs a backtracking search on the graph until it finds a solution to the problem, otherwise, the cycle of expanding the existing graph is repeated. An extension to this planner was proposed by Koehler et al. (1997), IPP with three main features which differ from the original GRAPHPLAN approach.

1. The input is a form of a pair of sets;

2. The selection procedure for actions takes into consideration that an action can obtain the same goal atom even under different effect conditions;

3. The resolution of conflicts occurs as a result of conditional effects.

In similar STRIP-based domain, Long and Fox (1999) developed a GRAPHPLAN-style planner, STAN, which performs a number of preprocessing analyses on the domain before executing planning processes. The approach firstly observes pre- and post-conditions of actions and represent those actions bit vectors form. Logical operators are applied on these bit vectors in order to check mutual exclusion between pairs of actions which directly interact. Similarly, mutual exclusion (mutex relations) is implemented between facts. A two-layer graph construction (spike) is used to represent the best exploited bit vector, which is useful to avoid unnecessary copying of data and to allow a clear separation on layer-dependent information about a node. The spike construction allow mutex relations recording for efficient mutex testing in indirect interactions. Secondly, there is no advantage in explicit construction of the graph beyond the stage at which the fix point is reached. Overall, the plan graph maintains a wave front which keeps track of all of the goal sets remaining to be considered during search.

A study reported in Kautz and Selman (1992) proposes a SAT-based plan (SATPLAN), which considers planning as satisfiability. The planning is further developed to BLACKBOX planner, which is a unification of SATPLAN and GRAPHPLAN (Kautz and Selman, 1999). The BLACKBOX planner solves a planning problem by translating the plan graph into SAT and applying a general SAT solver to boost the performance. A report in Silva et al. (2000) further develops the GRAPHPLAN-style by translating the plan graph obtained in the first phase of Graphplan into an acyclic Petrinet. Kautz and Selman (2006) later develop SATPLAN04 planner, which shares a unified framework with the old version of SATPLAN. The SATPLAN04 requires several stages when solving planning problems, which can be described as follows:

- Generating planning graph in a graphplan-style;

- Generating a set of clauses which derived from constraints implied by the graph, where each specific instance of an action or fact at a point in time is a proposition;

- Finding a satisfying truth assignment for the formula by utilizing general SAT problem solver;

- Extending the graph if there is no satisfactory solution or it reaches a time-out, otherwise, translating the solution to the SAT problem to a solution to the original planning problem;

- Post processing to remove unnecessary actions. actions.

Another planner such as HSP, which was developed by Bonet and Geffner (1999, 2001), is built based on the ideas of heuristic search. Vidal (2004) proposes a lookahead strategy for extracting information from generated plan in heuristic search domain. A later study by Vidal and Geffner (2006) further develop a branching and pruning method to optimise the heuristic search planning approach. The method allows the reasoning supports, precedences, and causal links involving actions that are not in the plan. Similar author later proposes an approach to automate planning which utilises a Fast-Downward approach as the base planner in exploring a plan tree. This approach estimates which propositions are more likely to be obtained together with some solution plans and uses that estimation as a bias, to sample more relevant intermediates states. A message passing algorithm is applied on the planning graph with landmark support in order to compute the bias (Vidal, 2011).

A different approach proposed in AI planning domain theory utilises heuristic pattern databases (PDBs), for example a study reported in Edelkamp (2000, 2002, 2014). Sievers et al. (2010) further assess that PDBs is lack of efficient implementation as the construction time must be amortized within a single planner run, which requires separate evaluation according to its own state space, set of actions and goal. Hence, it is impossible to perform computation processes at one time and reuse it for multiple inputs. The authors propose and efficient way to implement pattern database heuristics by utilising the Fast Downward planner (Helmert, 2006).

2.2 Robot Cognition

Studies by Franklin and Graesser (1997) and Barandiaran et al. (2009) point out that robots are real world agents, and consequently, the terms ‘robot’and ‘agent’are used interchangeably throughout this thesis.

Discussions on robot cognition can be traced back to the early development of human mind and consciousness theories. A study by Shear (1995) suggests that there is a direct correspondence between consciousness and awareness. We elaborate on these notions of consciousness and awareness in the following subsections.

2.2.1 Discussion on Theories of Mind

The mind is a collection of concepts that cover aspects of cognition which may or may not refer to an existing single entity or substance (Haikonen, 2012). In other words, the discussion of mind is restricted to perceptions, thoughts, feelings and memories within the framework of self. A large number of studies have addressed this field, and there are several important theories, described as follows:

- Traditional Approach
  A number of theoretical approaches identified throughout the history of human mind studies and their key points are described below.

  - Cartesian Dualism
    This theory, proposed by Rene Descartes, is based on the work of the Greek philosopher Plato (Descartes and Olscamp, 2001). The theory divides existence into two distinct worlds: the body, which is a material world, and the soul, which is an immaterial world. Descartes claimed that the body as a material machine follows the laws of physics, while the mind as an immaterial thing connected to the brain does not follow physical law. However, they interact with each other; the mind is capable of controlling the body but at the same time, the body may influence the mind.
  - Property Dualism
    This theory counters the Cartesian Dualism theory by suggesting that the world consists of only one physical material but that it has two different kinds of properties, physical and mental. Mental properties may emerge from physical properties, and can change whenever a change occurs in the physical properties, but mental properties may not be present all the time (Haikonen, 2012).
  - Identity Theory
    This theory is based on the concept of human nerve mechanisms which contain the various actions of nerve cells and their connections which structurise the neural process of the brain. Crick (1994) concluded that the human mind is the result of the behaviour of human nerve cells.

- Modern Studies
  Currently, studies of the mind focus on the neural pathways inside the human brain. A vast assembly of neurons, synapses and glial cells in the brain allow subjective experiences to take place (Haikonen, 2012, p.12). Studies on the nerve cells have led

to neural network and mirror neuron investigations, and these studies have made a large contribution to the concept of human mind and consciousness.

Consciousness

Since the early studies of consciousness, there has been no unanimous and uniform definition of consciousness. This thesis highlights a few important studies related to consciousness and robot cognition.

According to Gamez (2008), various terms are used to refer to the studies on consciousness theories using computer models to create intelligent machines, and the term ‘machine consciousness’, is typically the standardised terminology used in this field. According to Chalmers (1995), the consciousness problem can be divided into easy problems and hard problems. The easy problems assume that this consciousness phenomenon is directly susceptible to standardised explanation methods, which focus on computational or neural-based mechanisms (a functional explanation). In contrast, hard problems are related to experience, and appear to oppose the approaches used in the easy problems to explain consciousness. The author lists the phenomena associated with the consciousness notion as follows:

- Ability to discriminate, categorise and react to external stimuli
- Information integration by a cognitive system
- Reportability of mental states
- Ability to access one’s own internal state
- Focus of attention
- Deliberate control of behaviour
- Differentiation between wakefulness and sleep

Several studies have attempted to derive machine consciousness by capturing the phenomenal aspects of consciousness. Husserlian phenomenology refers to consciousness giving meaning to an object through feedback processes (Kitamura et al., 2000, p.265). Any system to be considered conscious should be assessed through the nine features of consciousness functions and Kitamura et al. (2000) further developed these nine characteristics form a technical view point as listed below:

1. First person preference: self-preference

2. Feedback process: shift attention until the essence of the object and its connection are obtained

3. Intentionality: directing self towards an object

4. Anticipation: a reference is derived for which objective meaning is to be discarded, and it becomes a belief with the property of an abstract object whenever the anticipation is unsatisfied.

5. Embodiment: related to the consciousness of events, which are the inhibition of perception and body action

6. Certainty: the degree of certainty in each feedback process of understanding

7. Consciousness of others: the belief that others have similar beliefs to our own

8. Emotion: qualia of consciousness which relies on elements of perception and corporeality

9. Chaotic performance: an unbalanced situation resulting from randomly generated mental events, which perturb the feedback process and intentionality.

Based on these features, Kitamura (1998) and Kitamura et al. (2000) proposed Consciousness-based Architecture (CBA) which is a software architecture with an evolutionary hierarchy to map animal-like behaviours to symbolic behaviours. These symbolic behaviours are a reduced model of the mind-behaviour relationship of the human. The architecture deploys a five-layer-hierarchy principle, which corresponds to the relationship between consciousness and behaviour. The foundation of the work is built on the principle of the conceptual hierarchical model proposed by Tran (1951, cited in Kitamura, 1998, pp.291-292) which is shown in Table 2.1. In a similar approach, Takeno (2012) proposed a new architecture which originated from

Table 2.1 Hierarchical Model of Consciousness and Behaviour

Level	Subjective Field	Category of Behaviours
0	Basic consciousness of awakening	Basic reaction of survival
1	Primitive sensation - likes and dislikes	Reflective actions, displacement and feeding
2	Valued sensation field of likes and dislikes (two dimensional environment)	Body localisation
3	Temporary emotions of likes and dislikes	Capture, approach, attack, posture, escape
4	Stable emotions towards present and unrepresented objects	Detour, search, body manipulation, pursuit, evasion
5	Temporal and spatial-based symbolic relation	Media usage, geography, mates, motion, ambush

Husserlian phenomenology and Minsky's idea which postulates that there are higher-level areas that constitute newly evolved areas which supervise the functionality of the old areas.

This new architecture conceptualisation of robot consciousness is achieved through a model-based computation that utilises a complex structure of artificial neural networks, named MoNAD. However, this model only conceptualises the functional consciousness category and studies have shown that understanding conciousness also involves the explanation of feeling, which is known as qualia. It is a physical subjective experience and, since it is a cognitive ability, its study can only be investigated through indirect observation (Haikonen, 2012, p.17).

Gamez (2008) divided studies on machine consciousness into four major categories:

1. External behaviour of machines that are associated with consciousness
2. Cognitive characteristics of machines that are associated with consciousness
3. An architecture of machines that is considered to be associated with human consciousness
4. Phenomenal experience of machines which are conscious by themselves

External behaviour, cognitive characteristics and machine architecture, associated with consciousness, are areas about which there is no controversy. Phenomenally conscious machines, on the the other hand, that have real phenomenal experiences, have been philosophically problematic. However, Reggia (2013) points out that computational modelling has been scientifically well accepted in consciousness studies involving cognitive science and neuroscience. Furthermore, computer modelling has successfully captured several conscious forms of information processing in the form of machine simulations, such as neurobiological, cognitive, and behavioural information.

2.2.2 Self-Awareness

In broad terminology, self-awareness can be defined as the state of being alert and knowledgeable about one's personality, including characteristics, feelings and desires (Dictionary.com Online Dictionary, 2015; Merriam-Webster Online Dictionary, 2015; Oxford Online Dictionary, 2015). In the field of developmental study, a report by Lewis (1991) postulates that there are two primary elements of self-awareness: subjective self-awareness, i.e. concerning the machinery of the body, and objective self-awareness, i.e. concerning the focus of attention on one's own self, thoughts, actions and feelings.

In order to be aware, particularly at the body level, sensory perception plays an important role in determining the state of self. This perception involves two different kinds of sensory mechanisms: proprioceptive sensors, which function to monitor the internal state, and exteroceptive sensors, which are used to sense the outside environment. Numerous studies on this sensory perception level have been carried out, and the earliest paper (Siegel, 2001) discusses the dimension aspect of the sensors to be incorporated into the robot. The author states that proprioception allows the robot to sense its personal configuration associated with the surrounding environment. Scassellati (2002) further correlates self-awareness with a framework of beliefs, goals and percepts attributes which refer to a mind theory. Within a goal-directed framework, this mind theory enables a person to understand the actions and expressions of others. The study implements animate and inanimate motion models together with gaze direction identification. A study conducted by Michel et al. (2004) reports the implementation of self-recognition onto a robot mechanism named NICO. The authors present a self-recognition mechanism through a visual field that utilises a learning approach to identify the characteristic time delay inherent in the action-perception loop. The learning observes the robot arm motion through visual detection within a designated time marked by timestamps. Two timestamp markings are initiated; one at the state when movement commands are sent to the arm motors, and one at the state in which no motion is detected. Within the same robot platform and research topic, a study was carried out by Gold and Scassellati (2009) which utilises Bayesian network-based probabilistic approach. The approach compares three models of every object that exists in the visual field of the robot. It then determines whether the object is the robot itself (self model), another object (animate model), or something else (inanimate model) which is possibly caused by sensor noise or a falling object. The likelihood calculation involves the given evidence for each of these objects and models. Within the same stochastic optimisation-based approach, a study conducted by Bongard et al. (2006) proposed a continuous monitoring system to generate the current self-modelling of the robot. The system is capable of generating compensatory behaviours for any morphological alterations due to the impact of damage, the introduction of new tools or environmental changes. On a lesser conceptual level, a study presented in Jaerock and Yoonsuck (2008) proposed prediction of the dynamic internal state of an agent through neuron activities. Each neuron prediction process is handled by a supervised learning predictor that utilises previous activation values for quantification purposes. Novianto and Williams (2009) proposed a robot architecture which focuses on attention as an important aspect of robot self-awareness. The study proposes an architecture in which all requests compete and the winning request takes control of the robot's attention for further precessing. Further research was conducted in Zagal and Lipson (2009), who proposed an approach which minimises physical exploration to achieve resilient adaptation. The minimisation of physical exploration is obtained by implementing a self-reflection method that consists of an innate controller for lower level control and a meta-controller, which governs the innate controller's activities. Golombek et al. (2010) proposed fault detection based on the self-awareness model. The authors focused on is the internal exchange of the system and the inter-correlative communication between inherent dynamics detected through anomalies generated as a result of environmental changes caused by system failures. At a meta-cognitive level, Birlo and Tapus (2011) presented their preliminary study which reflects a robot's awareness of object preference based on its available information in the context of human and robot interaction. Their meta-concept regenerates the robot's attention behaviour based on the robot's reflection of what the human counterpart is referring to during collaboration. The implementation of self-awareness in other areas, such as health services, has been highlighted in Marier et al. (2013), who proposed an additional method to their earlier study which adapts coverage to variable sensor health by adjusting the cells online. The objective is to achieve equal cost across all cells by adding an algorithm that detects the active state of the vehicless as the mission unfolds. Agha-Mohammad et al. (2014) also proposed a framework that has a health-aware planning capability. The framework is capable of minimising the computational cost of the online forward search by decreasing the dimension of the belief subset of the potential solution that requires an online forward search.

Much of the literature also identifies the lack of a concept of 'self'. This paper proposes a self-awareness framework for robots which uses a concept of self-awareness as proposed by Lewis (1991). The author postulates that in self-awareness, the concept of self is divided into two levels, subjective awareness and objective awareness. The author shows that human adults have the ability to function at both levels, under certain conditions, and that human adults utilise one level of self-awareness at a time. It can be inferred, however, that these two primary levels of self-awareness coexist and that human adults utilise them by switching the focus of attention between them. The change of direction in robot awareness mimics the principle of attention, which corresponds to processes of mental selection. During switching time, the attention process occurs in three phase sequences: the engagement phase, the sustainment phase and the disengagement phase (Haikonen, 2012). Haikonen (2012) also mentions two types of attention: inner attention and sensory attention. Sensory attention refers specifically to a sensor mechanism, which is designated to monitor a specific part of the body, such as joint attention or visual attention. We utilise this insight, particularly the ability to switch between both levels via attention phases, and through this action, a new framework can be used to change the robot's awareness from subjective to objective, and vice versa. In this framework, we refer to the physical parts of a robot, such as motors and joints (joint attention) as the subjective element, and the metaphysical aspects of the robot, such as the robot's representation of its position in relation to the external object or the robot's success in task performances (inner attention) as the objective elements.

2.2.3 Empathy with the Experience of Pain

This subsection comprehensively reviews literature studies on pain, the correlation of pain with self-awareness, the concept of empathy with pain and the evolving concept of robot empathy.

Pain

Various definitions have appeared throughout the history of human pain, such as the belief in early civilisations that pain is a penalty for sin and the correlation in the first century CE of the four humors and pain in Galen's theory (Finger, 1994). In the second century CE, Avicenna's postulate on a sudden change in stimulus for pain or pleasure generation was formulated (Tashani and Johnson, 2010). In modern times, concepts of pain are framed within the theory of functional neuroanatomy and the notion that pain is a somatic sensation transmitted through neural pathways (Perl, 2007). The culmination of the enormous number of works that have explored the concept of pain is the establishment of the following definition of pain as *"an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of tissue damage or both"* (The International Association for the Study of Pain, IASP 1986, cited in Harold Merskey and Bogduk, 1994).

Pain plays a pivotal role in the lives of humans, serving as an early sensory-based detection system and also facilitating the healing of injuries (Chen, 2011). In general, there are four theories of pain perception that have been most influential throughout history, reported in Moayedi and Davis (2013):

1. Specificity Pain Theory. This theory acknowledges that each somatosensory modality has its own dedicated pathway. Somatosensory systems are part of human sensory systems that provide information about objects that exist in the external environment through physical contact with the skin. They also identify the position and motion of body parts through the stimulation of muscles and joints, and at the same time, monitor body temperature (Byrne and Dafny, 1997). Details of the modalities are shown in Table 2.2.

2. Intensity Pain Theory. This theory develops the notion that pain results from the detection of the intense application of stimuli, and occurs when an intensity threshold is reached. Woolf and Ma (2007) proposed a framework for the specificity theory for pain and postulated that noxious stimuli respond to sensory perceptors known as *nociceptors*. When the intensity of the nociceptive information exceeds the inhibition threshold, the gate switches to open, allowing the activation of pathways and leading

Table 2.2 Modalities of Somatosensory Systems
(Source: Byrne and Dafny, 1997)

Modality	Sub Modality	Sub-Sub Modality
Pain	Sharp cutting pain dull burning pain deep aching pain	
Temperature	warm/hot cool/cold	
Touch	itch/tickle & crude touch	
	Discriminative Touch	Touch Pressure Flutter Vibration
Proprioception	Position:Static Forces	Muscle Length Muscle Tension Joint Pressure
	Movement: Dynamic Forces	Muscle Length Muscle Tension Joint Pressure Joint Angle

to the generation of the pain experience and associated response behaviours. Studies related to noxious stimulus and nociceptor are presented in Cervero and Merskey (1996) and Moseley and Arntz (2007).

3. Pattern Pain Theory. This theory postulates that somaesthetic sensation takes place as the result of a neural firing pattern of the spatial and temporal peripheral nerves, which are encoded in stimulus type and intensity. Garcia-Larrea and Peyron (2013) provided a review on pain matrices which asserts that painful stimuli activate parts of the brain's structure.

4. Gate Control Pain Theory. This theory, proposed by Melzack and Wall (1996), postulates that whenever stimulation is applied on the skin, it generates signals that are transmitted through a gate which is controlled by the activity of large and small fibres.

It can be seen that humans possess a complex structure of interconnected networks within the nervous system which permits a number of robust pain mechanisms, from detection, signal activation, and transmission to the inhibition of behaviours. However as Haikonen (2012) points out, artificial pain can be generated on a machine without involving any real feeling of pain. In other words, artificial pain can be evolved by realising the functional aspects of pain which is focused on a technical and practical way on how pain works and operates.

Pain and Self-Awareness Association in Human and Robot

Evolving pain mechanisms as an integrated element of awareness within a robot is a topic that has barely been addressed. One key reason is that self-awareness is a new area of research in human health, so few insights have been translated into the robot realm. A small number of papers have correlated pain with the self-awareness concept in robots and humans. The earliest study, conducted by Steen and Haugli (2001), investigates the correlation of musculoskeletal pain and the increase in self-awareness in people. This study suggests that awareness of the internal relationship between body, mind and emotions enables a person to understand and respond to neurological messages generated by the perception of musculoskeletal pain. A different study carried out by Hsu et al. (2010) investigates the correlation between self-awareness and pain, and proposes that the development of affective self-awareness has a strong association with the severity level of pain. The study utilises a self-reporting assessment mechanism in which reports were collected from people who suffer from *fibromyalgia*[1]. Steen and Haugli (2001) used pain acknowledgement to generate self-awareness, while Hsu et al. (2010) focused on the opposite phenomenon, namely, the measurement of affective self-awareness to accurately acknowledge pain. A recent study on self-awareness in robotics in relation to pain has been reported in Koos et al. (2013); this study uses the concept of pain to develop a fast recovery approach from physical robot damage. This work was also used in earlier studies including those of Bongard et al. (2006) and Jain et al. (2009). The study by Koos et al. (2013) is extended in Ackerman (2013) to produce a recovery model which does not require any information about hardware faults or malfunctioning parts. In fact, this approach demonstrates that the recovery model proposal disregards the importance of acquiring self-awareness in detecting pain that results from the faults generated by robot joints.

Empathy

The term *empathy* was introduced by the psychologist Edward Titchener in 1909 and is a translation of the German word *Einfühlung* (Stueber, 2014). Notwithstanding the extensive studies on empathy, the definition of this notion has remained ambigous since its introduction, and there is no consensus on how this phenomenon exists. Preston and De Waal (2002) mention that early definitions tend to be abstract and do not include an understanding of the neuronal systems that instantiate empathy. For instance, Goldie (1999) defines empathy as a process whereby the narrative of another person is centrally imagined by projecting that

[1] widespread pain and tenderness in the human body, sometimes accompanied by fatigue, cognitive disturbance and emotional distress.

narrative onto oneself. The author specifies that it is necessary for the individual to have the awareness that they are distinct from the other person. It is important to acquire substantial characterization which is derivable and necessary to build an appropriate narrative. Preston and De Waal (2002) discuss discrepancies in the literature and present an overview of the Perception-Action Model (PAM) of empathy, which focuses on how empathy is processed. The PAM states that attending to perception of oneself activates a subjective representation of the other person, which includes the state of the person, the situation, and the object. This subjective representation, if not controlled, creates correlated autonomic and somatic responses. A discussion of the functional architecture of human empathy presented by Decety and Jackson (2004) mentions that empathy is not only about inferring another's emotional state through the cognitive process, known as cognitive empathy, but is also about the recognition and understanding of another's emotional state, which is known as affective empathy. This is verified by the work in Cuff et al. (2014) in a review of the empathy concept, which discusses differences in the conceptualisation of empathy and proposes a summary of the empathy concept formulation as follows:

Empathy is an emotional response (affective), dependent upon the interaction between trait capacities and state influences. The processes are elicited automatically, and at the same time, shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined). In other words, the understanding (cognitive empathy) of stimulus emotion, with the recognition of the emotion source, is not from one's own. (Cuff et al., 2014, p.7)

Two common approaches are used to study human brain function: functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS). After Rizzolatti et al. (1996) introduced the mirror neuron concept, studies on empathy focused on the neural basis of the human brain structure and testing using fMRI and TMS. Discussions on the fMRI approach are presented in Jackson et al. (2005) and Banissy et al. (2012), and on TMS in Avenanti et al. (2006). Krings et al. (1997) mention that both fMRI and TMS are used to map the motor cortex which functions to generate nerve impulses for the initiation of muscular activities. The authors identify that fMRI is specifically utilised for identifying hemodynamic areas, which change during an action, while TMS is used for collecting information about the localisation and density of motoneurons, which are efferent neurons responsible for conveying impulses. De Vignemont and Singer (2006) remark on the common suggestion that shared affective neural networks exist that affect the reflection of emotional feelings of oneself towards others. According to the authors, these networks are automatically triggered whenever the other objects being observed deliver emotional displays. The authors propose two major functions of empathy:

1. Epistemology role. This means that empathy is used as an indicator to detect increased accuracy in the future prediction of the actions of the other people that are being observed. It serves to share emotional networks, which provides the associated motivation for others to perform actions. It also functions as a source of information about environmental properties.

2. Social role. This provides a basis for cooperation and prosocial behaviour motivation, and at the same time, promotes effective social communication.

An experimental work by Lamm et al. (2011) presents more quantitative evidence for the neural structures in the brain, involving the elicitation of pain experiences that originate either from direct experiences or indirect or empathic experiences. The study corroborates the findings in the literature mentioned earlier, that is, that there are shared neural structures and an overlapping activation between direct pain experiences and empathic pain experiences. The results also indicate that these shared neural structures overlap each other.

Empathy with Pain

A characteristic of human empathy is the ability to experience the feelings of others when they suffer (Singer et al., 2004). Singer et al. (2004) conducted an experiment on pain empathy by imaging the neural stimulation of the brain using fMRI. The authors reported that some regions of the brain form a pain-related network, known as a pain matrix. The study confirms that only that region of the pain matrix which is associated with the affective dimension is activated during the expression of an empathic pain experience. It also mentions that an empathic response can still be elicited in the absence of facial expression. These findings were confirmed by Jackson et al. (2005), who investigated perceptions of the pain of others through the medium of photographs. The study's experiment focused on the hemodynamic[2] changes in the cerebral network related to the pain matrix. Goubert et al. (2005) asserted that the following important points need to be considered:(i) The experience of pain distress captured by the observer may be related to contextual factors, such as an interpersonal relationship. (ii) The level of empathy is affected by bottom-up or stimulus-based processes and by top-down processes or observer knowledge and disposition. The common media used to communicate a distress level in bottom-up processes are social cues such as facial expressions, verbal or non-verbal behaviours and involuntary actions. In top-down processes, personal and interpersonal knowledge may affect the elicited pain response. Observer judgement, which includes beliefs and the context of others' pain experiences, also affect the empathic experience. (iii) Empathic accuracy, which concerns the problem of correctly

[2]factors involved in the circulation of blood, including pressure, flow and resistance.

estimating risk, plays an important role in the care of people who suffer from pain. If a situation is underestimated, people receive inadequate treatment, while overestimation may elicit a false diagnosis, leading to over-treatment. All these factors may have a devastating impact on a person's health. A topical review presented in Jackson et al. (2006) reports that mental representation is used as a medium to relate one's own pain experiences to the perception of the pain of others. The authors remark that experience of one's pain may be prolonged as one's self-persepection influences internal pain elicitation regardless of the absence of nociceptive invocation. The authors corroborate the work of Goubert et al. (2005) which suggests that the interpretation of pain representation, captured through pain communication, may not overlap with the exact pain experienced by the other person. This argument reflects the incompleteness of the mapping of the pain of others to oneself. In other words, the perception of one's own pain in relation to the pain of another shares only a limited level of similarity, and this enables the generation of controlled empathic responses. Loggia et al. (2008) extended this study and proposed that a compassionate interpersonal relationship between oneself and others affects the perception of pain. With the element of compassion, empathy-evoked activation tends to increase the magnitude of the empathic response. Hence, one's perception of pain in relation to other can be over-estimated regardless of the observation of pain behaviours. Another technique that has been utilised to disclose aspects that underlie human thought and behaviour, such as sensory, cognitive, and motor processes, is the event-related potential (ERP) technique, as described in Kappenman and Luck (2011). This technique, combined with a photograph-based experiment, was used in a study conducted by Meng et al. (2013). The authors investigated whether priming an external heat stimulus on oneself would affect one's perception in relation to another's pain. The paper concludes that a shared-representation of a pain model is affected by painful primes through an increased response in reaction time (RT).

2.2.4 Robot Empathy

This subsection reviews the literature that focuses on how the empathic element can be assessed and the possibility of its successful implementation in robot applications.

Empathic Robot Assessment

To justify the extent to which the empathic robot has been successfully achieved, it is important to establish measurement and assessment criteria. The assessment process can be divided into two major categories: robot-centred experiments and human-centred experiments.

In robot-centred experiments, robot performance is assessed by the robot's ability to function according to a predetermined empathic criterion, such as the ability to monitor its internal state by identifying body parts, the ability to direct its attention between the two levels of self, subjective awareness and objective awareness, and the ability to communicate through either verbal or pysical gestures (hand movements or facial expression) with its robot peers. Assessment is generally conducted according to machine performance, such as the speed of the robot's joints, the accuracy and effectiveness of the medium of communication being used, and response times. Gold and Scassellati (2009) carried out an assessment of their robot experimentations by measuring the time basis of the robot arm movements. Specific time allocations were determined to measure the robot's performance by observation of the robot's own unreflected arm. Time basis assessment was also used in a study on the self-awareness model proposed by Golombek et al. (2010). This study detects data pattern anomalies by generating training data models for anomaly threshold and training purposes. The approach splits all data into data sequences with a unified time length, and when an error occurs, an amount of time is dedicated to create the error plots for each occurence. In an experiment conducted by Hart and Scassellati (2011), the distance of an end effector of a robot right arm was measured from the predicted position to the recent position of the end effector. A recent study in Anshar and Williams (2015) assessed the performance of a robot awareness framework by measuring the predicted sequence of robot arm joint positions with the joint sensor position reading. The overall performance of the robot framework was reflected in low standard deviation values.

In contrast to the robot-centred experiments, where robot performance is measured according to proprioceptive and exteroceptive sensor data, human-centred experiments are concerned with task achievement from a human perspective. Humans are involved in assessing the performance of the robot within a predefined series of human-robot collaboration tasks. Several empathy measurement techniques are commonly used, such as the Hogan Empathy Scale (HES), updated to the Balanced Empathy Emotional Scale (BEES), the Interpersonal Reaction Index, the Basic Empathy Scale (BES) and the Barrett-Lennard Relationship Inventory (BLRI). The HES technique proposed by Hogan (1969) is utilised to measure cognitive elements, and its measurement process has evolved into four key stages. First is the generation of criteria for the rating assessment, followed by the evaluation of those rating criteria. The rating criteria are then used to define the highly empathic and non-empathic groups. Lastly, analyses are carried out to select the items for each scale, which function as discriminative tools between the nominated groups. The BEES was proposed by Mehrabian (1996), and is an updated version of the Questionnaire Measure of Emotional Empathy (QMEE) reported in Mehrabian and Epstein (1972). These techniques are designed to explore two social situations featuring emotional empathy, namely aggression and helping behaviour. QMEE utilises a 33-item scale that contains intercorrelated subscales, mapping the aspects of emotional empathy into a 4-point scale, while BEES utilises 30 items with a 9-point agreement-disagreement scale. In the IRI method, introduced by Davis (1983), the rationality assessment of empathy is constructed according to four subscales. Each subscale correlates to four constructs: Perspective Taking (PT), Fantasy Scale (FS), Empathic Concern (ES) and Personal Distress (PD). This method is considered to evaluate both cognitive and emotional empathy. A discussion of these three techniques is presented in Jolliffe and Farrington (2006), in which the authors propose the BES approach. This technique maps the empathy elements into 40 items which are used in the assessment of affective and cognitive empathy. Barrett-Lennard (1986) proposed the BLRI technique, which is particularly used in the study of interpersonal relationships, such as a helping relationship for therapeutic purposes. This technique measures and represents aspects of experience in a relationship on a quantity scale basis.

Current Achievement of Empathy Concept Implementation in the Field of Robotics

A report in Tapus and Mataric (2007) investigated the possible implementation of empathy in socially assistive robotics. The study gave descriptions of a specific empathic modelling, emulation and empathic measurement derived from the literature. The paper corroborates the significance of emulating empathy into robotics, particularly in robot assistive care, as a forward step towards the notion of the integration of robots into the daily lives of humans. A case study by Leite et al. (2011) investigates scenario-dependent user affective states through interaction between children and robots in a chess game. This study was extended by Pereira et al. (2011) and involved two people in a chess game in which a robot functioned as a companion robot to one player and remained neutral against the other player. The robot communicated through facial expression on every movement of the player, whether it was agreed, disagreed or was neutral. It was found that the user with whom the robot behaved empathetically perceived the robot's companionship as friendly.

An early study that investigated the neurological basis of human empathy in the field of robotics was reported in Pütten et al. (2013). A human observer was shown videos of a human actor treating a human participant, a robot and an inanimate object in affectionate (positive) and violent (negative) ways. fMRI was used to monitor parts of the brain which are active when an empathic response is elicited in humans. An important finding of this study is that in positive interaction in particular, there are no significant differences in the neural activation in the brain of the observer when empathic reactions are stimulated during human-human interaction or during human-robot interactions, whereas in negative situations, neural activation towards humans is higher than it is towards robots. The study was extended in Pütten et al. (2014), which investigates the emotional effect, the neural basis of human empathy towards humans, and the neural basis of generating the notion of human empathy towards robots. It was reported that the participants' reactions included emotional attitudes during positive and negative interactions. During positive interactions, there was no differences in neural activation patterns were found in the human observer's reactions either during empathy towards human experiments or in empathy towards robots. However, during negative interactions, when participants were shown abusive and violent videos, neural activity increased, leading to more emotional distress for the participants and a higher negative empathic concern for humans than for robots.

A new issue has arisen in the literature, which is the emerging notion of empathic care robots. It is reported in Stahl et al. (2014) that such technology will potentially create ethical problems, and there is a need to initiate a new scope of research to identify possible challenges that will need to be addressed.

Chapter 3

Perceptions, Artificial Pain and the Generation of Robot Empathy

This chapter discusses the elements that play a dominant role in artificial pain and the generation of empathic actions. Artificial pain generation is implemented in the pain activation mechanisms that serve as a pain generator. This pain generator precipitates the kinds of synthetic pain associated with the information obtained through the sensory mechanisms. Empathic actions are then generated as counter reactions based on proposals made by the pain generator.

Overall, there are few aspects derived from literature studies in Chapter 2 described as follows.

1. At lower level, the proposal should cover the ability to monitor the internal state of the robot by optimizing information derived from the robot perception. Robot perception as the gateway to obtain information could be derived from proprioceptive sensors (drawing information internally) and exteroceptive sensors (acquiring information from surrounding). These stimulus are used as the main building block for the robot to build and structure plans of actions, including anticipation possible failures.

2. At higher level, the proposal should consider the robot internal state representation in building the planning mechanism. In terms of representation, a possible choice is by looking into the BDI-based representation model, and for the planning itself should include three major elements, which are:

    - Automatic robot plan generation
    - Debugging process
    - Planning optimisation

3. At cognitive level, the approach should utilises a model which is scientifically well accepted, such as using computational modelling. Through computational modelling, cognitive element is directed towards the element. The term of consciousness is to signify the cognitive focus (the focus of attention), and should not be understood to mean human consciousness.

4. Concept of self-awareness could be derived by switching focus of attention from subjective elements to objective elements.

5. Proposed concept of artificial pain or synthetic pain could be originated from health studies by considering appropriate mapping into the embodiment element of robot. Identification process could be combined with the approach at Point 1 above. Pain activation approach could utilise *pain matrix*

6. Decision approach which utilises reasoning mechanisms should allow robust analysis within a shorter time.

7. Empathy concept could be generated by considering projection of another robot internal state onto a robot which precipitates empathic actions.

The following sections cover more details on the aspects of perception, artificial pain classification, pain activation and the implementation of the empathy concept in robots.

3.1 Perceptions

Perception, from the human perspective, concerns the ability to perceive objects through the senses. As as result of this ability, humans build interpretation and understanding, and later, become aware of the object of their senses. Mesulam (1998) points out that the human central nervous system (CNS) is responsible for handling the link configuration of sensory information to produce adaptive responses and meaningful experiences. In the field of somatics [1], Hanna (1991) states that an internally perceived soma is an immediate proprioception, which is unique data that originates at a sensory level. In terms of visual perception, there are five kinds of visual difference that contribute to image segregation: luminance, texture, motion, colour and binocular disparity, and visual perceivablity (Regan, 2000, p.3). Perception plays a crucial role in robotics and is one of the most important and necessary abilities in human-robot interaction (HRI) for enabling intelligent behaviours to emerge (Fitzpatrick, 2003; Yan et al., 2014). Yan et al. (2014) refer to perception as

[1] the field of study about the human body (soma) as it is perceived by the first person perception

an acquisition function of environmental information and analysis modules. This function divides robot perception into a lower level, concerned with the hardware and raw data, and a higher level, which focuses on data acquisition and analysis. The authors list three methods related to the perception in HRI, namely, feature extraction, dimensionality reduction and semantic understanding. Feature extraction concerns the lower level while the other two methods focus on the higher level of data extraction. Similarly, in the field of robot fault detection, perception is associated with sensory mechanisms, which are of primary importance as upfront error detection mechanisms. In other words, sensory mechanisms function as the gateway for robots to capture and retrieve information about their environment.

3.1.1 Proprioception and Exteroception

Robots are enabled to capture information originating from their internal systems (proprioception) or external environment (exteroception). An early study by Watanabe and Yuta (1990) presented the utilisation of proprioceptive and exteroceptive sensors to estimate mobile robot positions. A self-reconfigurable robot presented in Jorgensen et al. (2004) utilises several robot cells equipped with accelerometers and infrared sensors. The accelerometers are responsible for monitoring the tilt angles of each robot cell while the infrared sensors gather information about the connectivity and distance of neighbouring cells. Proprioceptive and exteroceptive sensors were also introduced in an experimental robot used in a study reported in Hyun et al. (2014). In this study, the external sensory information is obtained from the feet force sensors, while the internal kinematic changes are monitored by the joint encoders. A study by Salter et al. (2007) implemented accelerometers and tilt sensors as proprioceptive sensors in their rolling robot experiment. Accelerometers handle robot acceleration while tilt sensors detect the direction of tilt. Several other studies such as Anshar and Williams (2007) and Ziemke et al. (2005) utilise exteroceptive sensors to detect the experimental environments of robots, a vision sensor to detect environment landmarks, and a long-range proximity sensor to detect an object on the robot pathway.

Similarly, our approach to sensory perception utilises the proprioceptive and exteroceptive sensors which already exist in the robot mechanism. Each sensor category is used as a driving source of pain activation, which will be further explained in the following sections.

3.2 Faulty Joint Setting Region and Artificial Pain

The literature-based study in Section 2.2.3 mentions that the thesis proposal on the evolution of artificial pain for robots emphasises the aspect of functional pain. Stimuli are generated from the proprioceptive mechanisms of the robot body parts and this process mimics subjective awareness, which reflects the element of embodiment, known to be one of the features of consciousness (Takeno, 2012). The proprioceptive mechanisms detect and capture any fault occurrences, and then assign a specific intensity value to them to determine the level of pain to be invoked. At the same time, these mechanisms generate a reactive behaviour as a counter response which is relevant to the pain experience. Our artificial pain concept is inspired by the definition of pain proposed by Woolf (2010), and our proposal for the classification of artificial pain is developed accordingly. Three classifications of artificial pain are derived from the pain definition in Woolf (2010), and for each class, we assign a designated pain intensity level derived from Zhu (2014). The term *synthetic pain* is introduced whenever the kinds of pain classification are referred to. Descriptions of the proposal are presented in Table 3.1, and details of each category, which relates to the varieties of synthetic pain and their causes, are discussed in the following subsections.

Table 3.1 Artificial Pain for Robots

Category	Synthetic Pain	Description	Definition	Intensity Level
1	Proprioceptive Pain	1.0	Potential hardware damage, as an alert signal	"None", "Slight"
2	Inflammatory Pain	2.1	Predicted robot hardware damage	"None", "Slight"
		2.2	Real robot hardware damage	"Moderate", "Severe"
3	Sensory Malfunction Pain	3.1	Abnormal function of internal sensors	"None", "Slight"
		3.2	Damage to internal sensors	"Moderate", "Severe"

3.2.1 Proprioceptive Pain (PP)

This class of synthetic pain is instigated by stimuli from either internal proprioceptive sensors or from exteroceptive sensors in the form of an empathic response. The pain serves as an alert signal to plausible actual damage as a result of the stimuli received from the environment where the body parts being monitored are involved in an interaction. The type of response to be generated is associated with the sensitivity of the current stimuli and future prediction. It may directly influence an element that will boost the activation process (booster), but it is less likely to activate the pain generator. This kind of pain typically occurs as the robot mind predicts changes in the environmental stimuli, and the robot is required to pay attention to the possibility of future pain. Hence, no true counter actions result from the activation of this type of pain. In other words, these counter reactions simply reside in the robot's memory for future reference.

3.2.2 Inflammatory Pain (IP)

As the robot experiences the PP up to a level that the robot can endure, the robot mind keeps the reasoning process going while continuing to monitor the affected joints. If there is an increased level of stimulus and the alert signals do not subside, the robot evokes the IP and triggers the generation of counter actions as a response to the IP. Counter responses may involve the generation of new joint movements dedicated to alleviating or reducing the severity of the pain's impact. For example, a six-legged robot that suffers from a broken leg could counteract by generating an alternative walking gait. Evoking this kind of pain will directly overrule the booster and cause changes in the robot's consciousness by switching robot awareness into the subjective element. The selection of the region of awareness is determined by the level of pain being evoked. Whenever the reasoning process predicts that the proposed alternative actions could lead to further damage (the PP is activated), the robot mind prepares counter reactions, such as stopping the robot from walking. However, if the change in stimuli is very rapid, the robot immediately generates the IP without invoking the PP.

3.2.3 Sensory Malfunction Pain (SMP)

This kind of pain is related to an internal sensor which may create alarm signals that are false-positive or false-negative. A false-positive alarm means that the sensory malfunction affects the mind and generates an overestimation of the incoming sensory information. This situation may lead to the generation of an unnecessary counter response at the time of detection. By contrast, false-negative alarms are generated as a result of underestimated detections. This kind of pain is originated from physical damages to the internal hardware of the robot's sensory mechanism. The robot has a higher chance of suffering from an increase in the severity of the pain as the robot mind does not produce appropriate counter responses to the pain. A prolonged experience of this kind of pain may lead to a catastrophic impact on the robot hardware. In this situation, the robot reasoning system plays a crucial role in detecting and justifying any hardware issue related to the internal sensor functionalities. Furthermore, the mind may provide a possible diagnosis if the abnormality function occurs as the result of internal damage to the sensor.

The activation procedure for each synthetic pain category is depicted in Figure 3.1. The horizontal axis represents the activation time measured in cycles of data sequence and the vertical axis represents the pain level for each synthetic pain category with respect to the time of activation. At time t_1, the kind of PP is activated at the *Slight* level and as the level increases to *Moderate* ($t_2 == t_5$), the IP is evoked at the *Slight* level. In this situation,

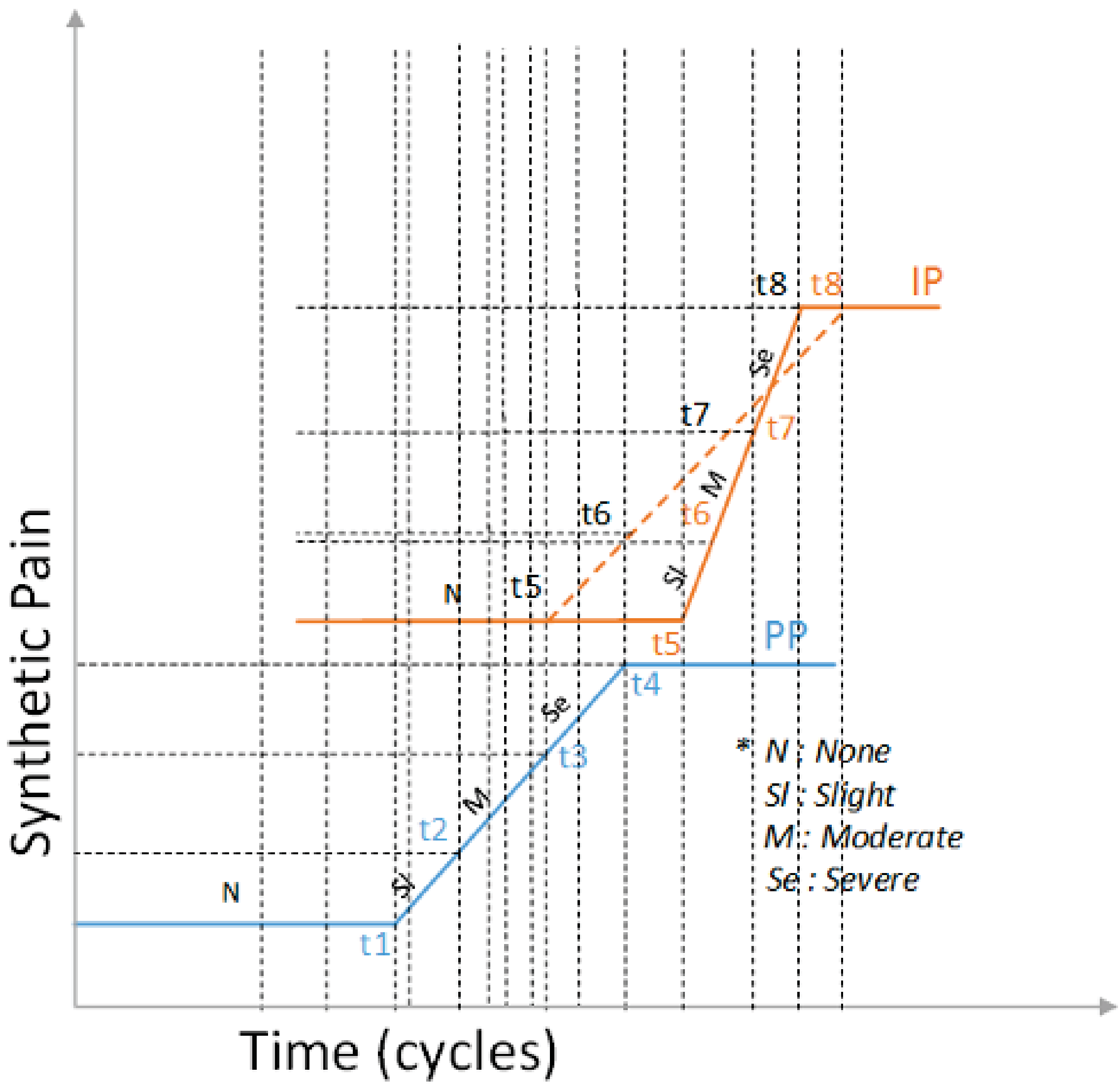

Fig. 3.1 Synthetic Pain Activation PP and IP

robot reasoning can still follow the change in stimuli obtained from the sensory mechanisms. However, if the change in stimuli occurs rapidly, to an extent that the mind cannot cope, the IP will be generated regardless of the PP results (shown at time t_5). In contrast, the SMP activation occurs independently as the robot mind continues to monitor its own sensory mechanisms (see Figure 3.2).

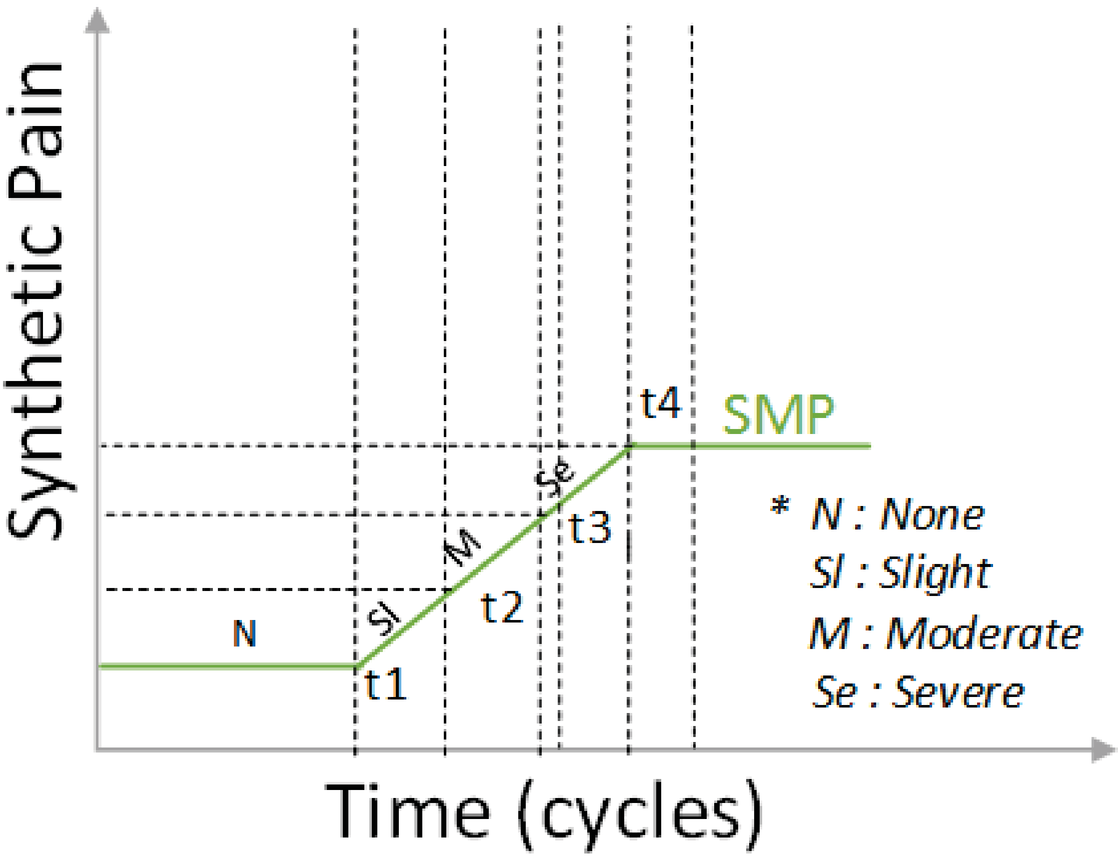

Fig. 3.2 Synthetic Pain Activation SMP

3.3 Pain Level Assignment

The region in which each body part motion occurs determines the pain level. The motion of robot joints, typically, could be divided into two motions (Spong et al., 2006):

1. Rotational where the motions are measured in radian or degree of revolution. This motions cover several types of robot joints, such as rotational (rotary), twisting, orthogonal and revolving joints.

2. Lateral where the motions are measured in length of displacement. This motions refer to the linear or prismatic joints

The pain level is assigned by measuring the distance between the position of the respective body part in the region and the threshold values assigned by the robot awareness framework (see Figure 3.3). The physical motions associated with the joint movements of the robot hardware are actively monitored by the sensory mechanisms, which contain proprioceptive and exteroceptive sensors. The further the distance from the threshold value, the higher the pain level to be assigned. The threshold values can be manually designed by the human user and placed in the database as a reference (static threshold), or they can be generated and configured autonomously by the robot framework itself (self-generated).

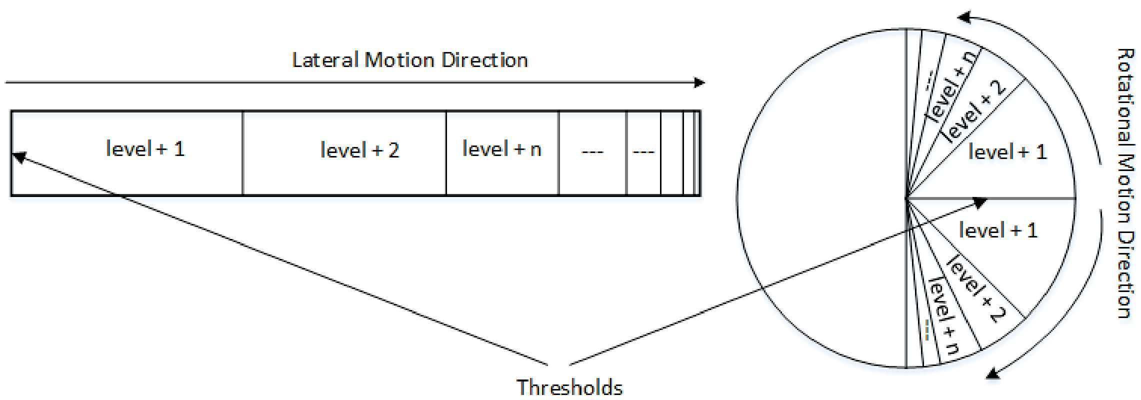

Fig. 3.3 Pain Region Assignment

3.4 Synthetic Pain Activation in Robots

To generate synthetic pain in robots, we set the joint restriction regions to specified values and each region determines the level of pain and the kinds of synthetic pain that the regions will invoke. These joint restriction regions are referred to as threshold values, and represent the areas in which the robot joints should not move. This concept simulates human movement; for example, people who suffer from shoulder pain, in which the pain occurs when the arm is moved into specific positions. Patients with this type of musculoskeletal problem tend to avoid moving the arm attached to the affected shoulder into those positions. Hence, restrictions are introduced to the affordability space of the body part, such as the rotation of the shoulder.

Two approaches are introduced in order to generate synthetic pain: a simplified fault detection-based model (simplified pain detection) and a pain matrix model, as described in the following subsections.

3.4.1 Simplified Pain Detection (SPD)

The assessment criterion for the Simplified Pain Detection (SPD) model is whether the current arm position, which is obtained either from proprioceptive or from exteroceptive sensors, is higher than any of the joint restriction values. If this condition is satisfied, the SPD model generates a set of recommendations to the Robot Mind for further reasoning. These recommendations are shown in Table 3.2.

Based on aspects derived from literature studies, as mentioned early in 3, *Belief* terminology is used to represent the internal state of the Robot Mind. Details of all information that form the *Belief* of the robot is explained in Chapter 4. For early development, the *Belief* is divided into several states, which is called Belief State, as described below:

1. *Current* which refers to the result of the reasoning process of the Robot Mind with information obtained from perception
2. *Prediction* which refers to the result of the reasoning process of the Robot Mind with information derived internally from prediction processes
3. *Special* which refers to the result of the reasoning process of the Robot Mind with special conditions, such as anomaly data from sensory system. The Robot Mind should treat this information differently as it might cause the reasoning proposes a false diagnosis. This state is strongly related to generation of the synthetic pain type *Sensory Malfunction Pain (SMP)*

It can be seen that whenever the Belief State of the framework is *Current*, only one recommendation is activated, namely whether the Mind State is *Constrained*, and other recommendations are disabled. When the Belief State is *Prediction*, all the recommendation elements are activated, giving more information about the occurrence of future pain. In the *SpecialCases* condition, the reasoning process makes more critical analyses of the incoming data and establishes whether the sensor has a temporary faulty function, which means there is no problem with the sensor hardware, or that the fault readings have occurred as the result of defective/broken sensor hardware which requires extra attention, such as the replacement of permanent hardware.

Table 3.2 SPD Recommendation

No	Belief State	Recommendation				
		Mind State	Initiation time	Alert time	Data Alert	Time Details
1	**Current**	Constrained/ Unconstrained	Disabled	Disabled	Disabled	Disabled
2	**Prediction**	Constrained/ Unconstrained	Activated	Activated	Activated	Activated
3	**Special Cases**	Constrained/ Unconstrained	Depend	Dependent	Dependent	Dependent

Data Representation

The discussion of pain generation analysis is presented in a functional or mathematical model, which is by nature a psychophysical model (Regan, 2000, pp.26-27). The functional property of the model follows the assessment criteria mentioned previously (Subsection 3.4.1). The collection time of information from the sensory mechanism is represented as T. The representation of data which is sampled at a time of t_i is d_{t_i}. This data originates from proprioceptive or exteroceptive sensors. The whole collection of data sequences is represented as

$$\prod_{t_i,i=0}^{i<T} d_{t_i}$$

The value of t_i is collected from the initiation of the detection time, and the time span of the data collection follows Criterion 1 below.

$$d_{t_i} = \begin{cases} i = m = 0, \text{ initiation of detection time} \\ i < T, \text{ time span of data collection} \\ i = m, \text{ sampling data length} \end{cases}$$

The kinds of synthetic pain to be invoked are derived from the data obtained, whether the Belief State categorises those data as *Current*, *Prediction* or *Special Cases*, for which sensory assessment is required. Whenever the Belief State is *SpecialCases* category, the data is considered to be noisy, due to faulty readings or defective sensors. The pain assignment guideline for each belief state category follows Criterion 2.

$$pain\,class_{belief\,state} = \begin{cases} belief\,state = current, \text{ pain class: IP} \\ belief\,state = prediction, \text{ pain class: PP} \\ belief\,state = sensory\,assessment, \text{ pain class: SMP} \end{cases}$$

The corresponding pain level to be generated follows Criterion 3, which is derived only from the comparison between the assessed data and the joint restriction values jt_i.

$$painlev_{i(where:i<=3)} \begin{cases} d_{t_i} < jt_1, & \text{pain level: } \textit{None} \\ d_{t_i} > jt_1, & \text{pain level: } \textit{Slight} \\ d_{t_i} > jt_2, & \text{pain level: } \textit{Moderate} \\ d_{t_i} > jt_3, & \text{pain level: } \textit{Severe} \end{cases}$$

3.4.2 Pain Matrix (PM)

Unlike the pain activation mechanism in the previous model, the Pain Matrix (PM) model uses a more sophisticated approach by introducing system properties which are formed by the interconnectivity between several modules integrated into a matrix (as shown in Figure 3.4). Four major modules work together to form the framework of the Pain Matrix, described

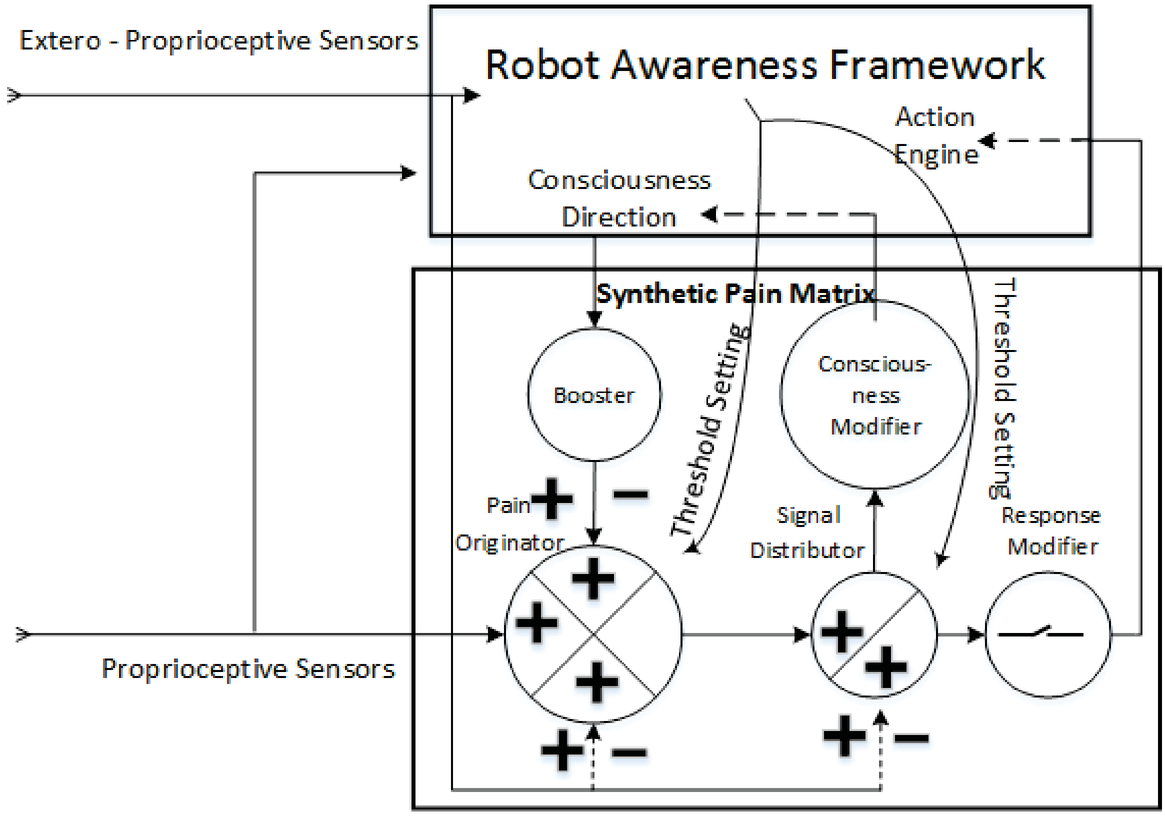

Fig. 3.4 Pain Matrix Diagram

as follows:

1. Pain Originator (PO). This module works by combining information derived from external source, which is from sensory mechanism, and from internal source, which is the Booster. Whenever value resulting from Pain Originator is higher than the internal threshold value, which is set by the Robot Mind, then it will fire the next module, Signal Distributor.

2. Signal Distributor (SD). Taking the firing data from the PO module and comparing it with the data derived from the exteroceptive sensor, the Signal Distributor module further modifies the Robot counter reactions whether internally or externally. Internal reaction will affect the consciousness direction (through the medium of Consciousness Modifier) and external reaction will activate the Response Modifier module. By taking information from exteroceptive directly, the SD module has the ability to guarantee that the recommendation for the PO module is proportional to the current situation facing by the robot.

3. Booster (Bo). This module influences the PO module by taking recommendations from the Robot Mind, whether the changes in Consciousness Direction directed by the Consciousness Modifier module or by the reasoning process run internally by the Robot Mind. This influence may further boost the generation of pain level or alleviate pain generation.

4. Response Modifier (RM). This module selects the most appropriate actions taken with respect to the kind of synthetic pain and the level of pain.

The robot awareness status plays a crucial role in determining the Pain Originator module by influencing the activation of the Booster module. During empathic actions, the Pain Originator disregards data from the proprioceptive sensor and sets the focus of attention on the object of the robot's exteroceptive sensors. When no information is retrieved from the sensory mechanisms, the framework initiates internally, which means that no pain is generated. Empathic actions are generated by taking only the information from the exteroceptive sensors. The Consciousness Modifier and the Response Modifier modules may affect the Consciousness Direction of the framework. The overall functionality of the Pain Matrix is shown in Table 3.3. When the initiation of consciousness direction occurs internally, only the Booster, as the element of the Pain Matrix, is activated, hence the sensory mechanisms and other elements of the Pain Matrix are eliminated from taking part in determining the internal state of the robot; that is, the Booster will not be activated. In this situation, only information retrieved from the proprioceptive sensors drives the Pain Originator module. If the signal from the Pain Originator is below a certain threshold defined

Table 3.3 Pain Matrix Functionality

Robot Mind Element	Initiation			Consciousness Direction
	Internally	Externally		
		Proprioceptive	Exteroceptive	
Awereness Framework	√			√
Paint Matrix				
Booster	√	√	√ or ×	
Pain Originator	Ignored	√	√ or ×	
Signal Distributor	Ignored	√	√ or ×	
Consciousness Modifier	Ignored	√ or ×	√ or ×	Framework + Pain Matrix
Response Modifier	Ignored	√ or ×	√ or ×	Framework + Pain Matrix
Activated Pain				
Proprioceptive	None	√	√ or ×	
Inflammatory Reduction	None	√ or ×	√ or ×	
Sensory Malfunction	None	√ or ×	√ or ×	
Responses				
Self Response		√		√ or ×
Empathy Response			√	√ or ×

by the Robot Mind, the Signal Distributor deactivates the Consciousness Modifier and the Response Modifier. As the robot joint moves and is monitored by the awareness framework, the Pain Originator accumulates information. If the information obtained contains false information, the Consciousness Direction will activate the Booster and provide counter feedback, reducing the values of the accumulated information in the Pain Originator. In this way, the PM prevents the activation of the Consciousness Modifier and the Response Modifier. The focus of attention is thus still fully governed by the internal awareness framework with no influence from the PM, and the robot does not deliver and experience any synthetic pain. When joint motions approach the faulty joint regions, the awareness framework detects and predicts the incoming information. In this situation, the Booster is set to activate and modify the accumulated information obtained from the proprioceptive sensors. The pattern of the accumulation data changes may differ from time to time, producing either a gradual or dramatic increase. The distance from the thresholds will justify the activation of the Consciousness Modifier and the Response Modifier. Once the threshold values have been exceeded, the two modifiers will play their roles in influencing the Robot Mind. This action may change the focus of attention of the Robot Mind through Consciousness Direction modification and the generation of action responses to the synthetic pain the robot is experiencing. In the case of empathy generation, the robot's exteroceptive sensors may affect the accumulation values of the Pain Originator. Similarly, they may also modify the accumulation values of the Signal Distributor to determine whether the Response Modifier should influence the Action Engine to provide empathy responses to the object of empathy. These empathy responses may include approaching the object and providing assistance.

Pain Generation Analysis

The proposal contains functional system properties that are formed by the interconnectivity between the elements of the Pain Matrix. The Pain Originator calculates the overall data of the proprioceptive sensor and the Booster following Equation 3.1) below:

$$painorg_{t_i} = \sum_{t_i, i=0}^{i<m<T} (prio_{t_i} + (\pm boost_{t_i})) \tag{3.1}$$

where $prio_{t_i}$ refers to data being collected from the proprioceptive sensor at a specified time t_i, and $boost_{t_i}$ represents the value of the Booster being injected into the Pain Originator at the time of data being gathered from the proprioceptive sensor. The value of $boost_{t_i}$ could be either to amplify or to attenuate the impact of data from proprioceptive sensor in the pain level generation of the *painorg*.

The Pain Originator will only prime the Signal Distributor if the accumulated data is greater than the threshold value assigned by the robot awareness framework (Criterion 4).

$$\Delta painorg_{t_i} > (painorg_{t_i} - painorgthreshold_{t_i})$$

The higher the value of $\Delta painorg_{t_i}$ the higher the pain level generated by the Pain Originator. This value corresponds to the activation of the Consciousness Modifier as determined by Criterion 5 below.

$$\Delta sigdist_{t_i} > (\Delta painorg_{t_i} - sigdistthreshold_{t_i})$$

3.5 Generation of Robot Empathy

A key point in the realisation of robot empathy is the projection into the robot of the internal state of an external object as an object of attention. This approach is inspired by the work of Goldie (1999), which emphasises that the process of a centralised imagination of another person's narration occurs through the projection of an object into oneself, and that this corresponds to the empathy process. Thus, there are three major aspects of our robot empathy generation, which are:

1. Robot Embodiment. Embodiment, which is considered to be a feature of consciousness, will allow any physical part of the robot to be an object of the robot's own attention. This condition simulates the conceptualisation of the subjective element of robot self-awareness. The state of the embodiment is actively monitored through the robot proprioceptive sensor. When the focus of the robot's attention is directed towards a specific robot body part, the information retrieved from the proprioceptive sensor becomes highly prioritised for thorough assessment.

2. Internal State Projection. By utilising its exteroceptive sensors, a robot observes the body motion of another external object over time. The projection of the internal state of the target object commences by capturing the body motion information of the observed object. This information is assessed by projecting the motion data space into a data coordinate space. This projection corresponds to the fusion process between the observer robot and the object being observed.

3. Synthetic Pain Assessment. Conversion of the data coordinate space into a joint robot space.

3.5.1 Empathy Analysis

During empathy activation, the Pain Originator includes the information from the exteroceptive sensor, and the result is Equation 3.1 modified to Equation 3.2.

$$painorg_{t_i} = \sum_{t_i,i=0}^{i<m<T} (prio_{t_i} + (\pm boost_{t_i}) + (\pm extero_{t_i})) \tag{3.2}$$

where $extero_{t_i}$ represents data being collected from the exteroceptive sensor at the time of *painorg* generates pain level. Similar to the $boost_{t_i}$, its value could be either amplify or to reduce the effect of information gathered from the exteroceptive sensor.

It can be seen that information captured from the exteroceptive sensors of the observer robot, such as the vision sensor, plays an active role in determining the internal projection of the robot being observed into the observer robot. When this process yields to the generation of synthetic pain, the priming of the Signal Distributor also considers the additional data from the same external sensor. This mechanism is designed to keep the external source of information as the basis of the Pain Matrix functionality (see Equation 4.1).

$$sigdist_{t_i} = \sum_{t_i,i=0}^{i<m<T} (\Delta painorg_{t_i} + (\pm extero2_{t_i})) \tag{3.3}$$

The value of $\Delta painorg_{t_i}$ is derived from Criterion 4 and Consciousness Modifier activation follows Criterion 5.

Chapter 4

Adaptive Self-Awareness Framework for Robots

This chapter presents the proposed framework which is used as a benchmark for integrating the conceptualisation of artificial pain and empathy generation with the robot mechanism. An overview of the structure of the framework and outline of its key elements are discussed in the sections that follow.

4.1 Overview of Adaptive Self-Awareness Framework for Robots

The adaptive self-awareness framework for robots, known as ASAF, is comprised of several elements, as shown in Figure 4.1. There are a number of predefined values which are constant values determined by an expert user and these values remain the same throughout the application. They are subject to redefinition by the expert user for different applications. Important elements of the ASAF, that is, Consciousness Direction, Synthetic Pain Description, Robot Mind, Action Execution and Database, are discussed briefly in the following subsections.

4.1.1 Consciousness Direction

We utilise the concept of consciousness as the ability to redirect attention between the two levels of awareness, as proposed by Lewis (1991). Our robot consciousness, therefore, refers to the cognitive aspect of the robot that is used to specifically signify the focus of the robot's attention. There are two predominant factors in directing robot consciousness:

(i) the ability to focus attention on a specified physical aspect of self, and

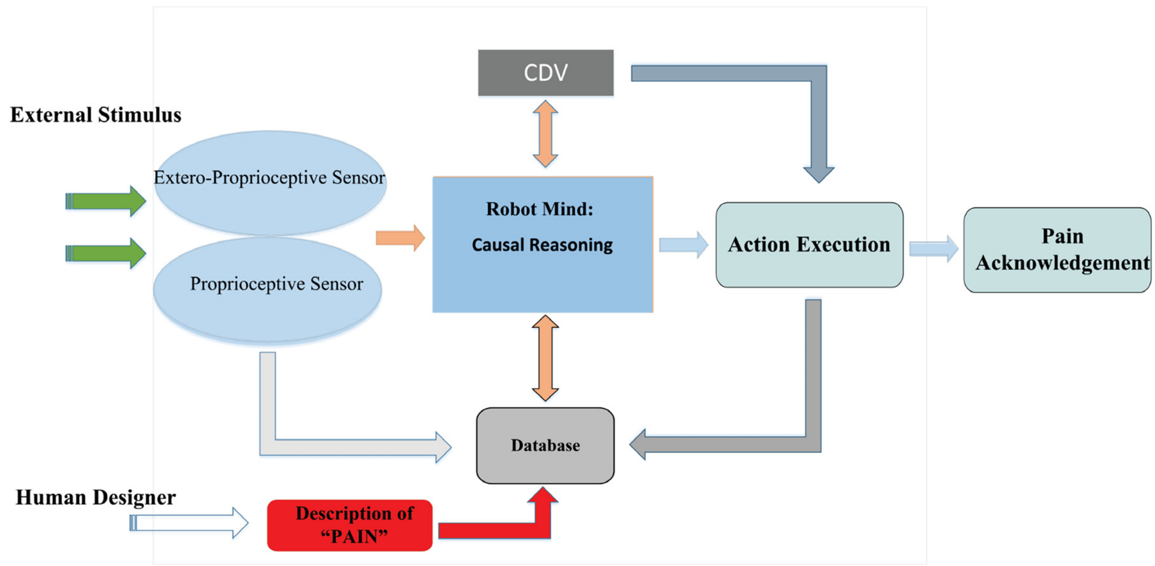

Fig. 4.1 Adaptive Robot Self-Awareness Framework (ASAF)

(ii) the ability to foresee, and at the same time to be aware of the consequences of predicted actions.

Our proposal formulates how to address these two aspects so that they can be developed and built into a robot self-awareness framework, and so that the detection of synthetic pain can be acknowledged and responded to in an appropriate way. Robot awareness is mapped to a discrete range 1 - 3 for subjective elements and 4 - 6 for objective elements. In other words, the robot's cognitive focus is permutable around these predetermined regions. Changing the value of Consciousness Direction (CDV) allows the exploration of these regions, and at the same time, changes the focus of the robot's attention. It is important to keep in mind that our subjective elements specify the physical parts of a robot, such as robot motors and joints, and that the objective elements signify the metaphysical aspects of the robot, such as the robot's representation of its position in relation to an external reference. The Robot Mind sets the CDV and determines the conditions for the exploration of robot awareness regions, whether these conditions are constrained or unconstrained. The structure of the robot awareness regions and CDV are shown in Figure 4.2.

4.1.2 Synthetic Pain Description

To generate synthetic pain in the robot, we set the robot joint restriction regions that are to be avoided. These joint restriction regions contain values of joint robot positions that are considered to be faulty joint values. Synthetic pain can then be generated when the robot joint moves into this region, as described in the previous chapter. Joint movement is monitored

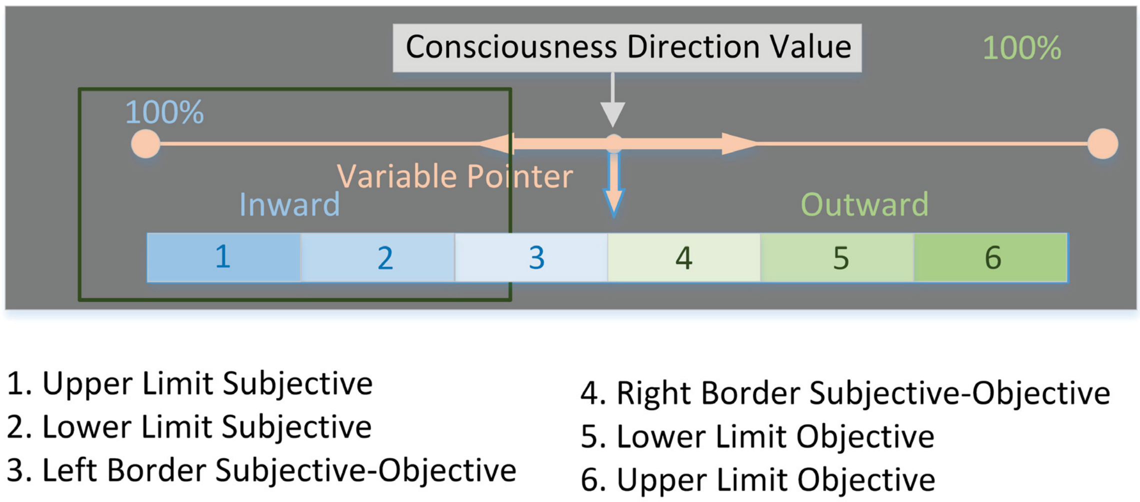

Fig. 4.2 Robot Awareness Region and CDV

by the proprioceptive sensor of the robot, and this information can subsequently be used by the Robot Mind to reason and determine the kinds of pain to be evoked. The method of determining the pain category to be evoked is implemented in the SPD and Pain Matrix models.

4.1.3 Robot Mind

Once the reasoning of the Robot Mind indicates that the joint movements are tending towards, or have fallen into, these restricted joint regions, the Robot Mind performs three consecutive actions:

- Setting the robot awareness into a condition of constraint.
- Modifying the CDV, which will shift the robot's focus of attention to the subjective element of its awareness.
- Providing counter response actions by collecting available pre-defined sets of counter response actions (Event-Goal Pairs stored in the Database), such as alerting human peers through verbal expressions and increasing robot joint stiffness.

The components and pathways of the overall reasoning of the Robot Mind are illustrated in Figure 4.3.

It can be seen from the figure that the Robot Mind is divided into two levels: (1) Body, which concerns with physical elements; and (2) Mind, which lies on the meta-physical level. The agent's motoric and perceptive systems are the two main factors affecting the functionalities of the Body. These two factors serve as the gateway for the robot to interact with the environment, either changing robot's spatial positions with respect to its environment (locomotion purposes) or to gather information from the environment(sensing purposes). On the Mind level, several elements together form a framework which constitutes to the Mind's performances. A Belief set which contains current values of beliefs that the Mind have, including conditions that satisfy each belief to occur. This Belief set is sent and stored as a history in the Database along with associated conditions and other previous data. From the current Belief set, the Mind originates the Event-Goal Pair Queue. A Plan Library is formed by utilizing the data kept in the Database. Three data, Event Goal Pair Queue, Plan Library, Database, are sent to the Causal Reasoning process for an assessment process. This reasoning process analyses those three data and compared with the data from the Belief set. This phase produces first level of recommendation and to be propagated back to the Event-Goal Pair Queue and the Database for updating purposes. The first level recommendation sets the goals of the Intention Engine, producing the second level of recommendation. The logic engine which contains the AND - OR Functions further reformulate the recommendation and send them to the Intention Execution Engine. This recommendation activates corresponding Primitive Actions which affect the Motoric Systems of the robot. This cycle then repeats for every new incoming Belief set.

Overall, the behaviour of the Robot Mind can be explained as the following: the values of faulty joint settings and the limit of the consciousness region areas are defined and placed in the Database. Once the collaborative task involving human and robot has taken place, the Robot Mind sets the robot's awareness to a random state. This means that the robot's attention may be focused in one of six regions by random selection of the CDV. Once selected, the Robot Mind is set in an unconstrained condition, allowing task execution and collaboration to proceed. Although the awareness is focused on the previously selected region, the Robot Mind at the same time monitors its proprioceptive sensor, that is, the arm joint sensors which are physically involved in the interaction with a human peer. Changes in the joint sensor readings produce changes in the pattern, and these changes are captured and used as the reasoning element of the Robot Mind. As the joint moves, the robot's *Belief*, *Desire* and *Intention* are subjected to change and the Action Executions transform the results into primitive actions for execution. For every prediction that may introduce higher risk of the arm joint experiencing faulty joint settings, the Robot Mind alters the CDV, causing awareness to to be focused on the robot arm (Subjective Awareness) and at the same time, the robot's internal state is set to constrained. Once this situation has been reached, the robot's joint stiffness is set to a maximum value and the human peers are alerted by verbal notification. As the Robot Mind's working domain is part of the internal state of the robot, we utilise the terminologies *Beliefs*, *Goals*, and *Intentions* (BDI) to represent the internal processes of the mind. All the elements in BDI reside in the database of the framework which are accessible during the activation of framework.

A simple scenario when the robot moves and finds obstacle in its path. The robot perception (Body level) senses the existence of external object in its path and forms the current belief that obstacles is detected. The Mind also structures conditions that satisfy the criteria as obstacles detected, such as the spatial information of the obstacle with respect to the robot position or other information. Details of information gathered can be summarised as the following:

- Spatial information:
  - Distance to the current position of the robot.
  - Position whether it is on the left or right side of the robot.
- Current states before obstacle detected:
  - Beliefs state.
  - Goals state.
  - Intentions state.
  - Logic state (AND - OR Functions).
  - Active Plan Library.
  - Active Event-Goal Pair Queue.
  - First and second recommendation states.
  - Current pointer of the Database which informs element of data being accessed.
- Miscellaneous information such as visual information captured at the time obstacle being detected.

This set of information forms the current Belief state of the Mind which will be further processed. The new Event-Goal Pair Queue is then constructed along with the Plan Library which maps the manifestation of how the event-goal pairs are achieved (first recommendation). The Causal Reasoning assesses the validity of the first recommendation whenever the Belief state changes again. If there is no changes occurred, the Causal Reasoning proceeds the reasoning process and produces the second recommendation to be passed to the Logic - AND OR Functions and to the Intention Engine. The AND - OR Functions then govern the Intention Execution Engine which activates the Primitive Actions to be executed to avoid the obstacle. If the Perception Systems detect that the changes occur instantly, then the Logic element, AND-OR Function will over-write the reasoning process and decide the action of the Intention Execution Engine by only analysing the current Belief state (obstacle detected) and previous state of the Intentions. This situation occurs when the reasoning process could cause the robot being late in taking proper and accurate actions which could lead the robot to bump into the obstacle.

4.1.4 Database

The Robot Database contains a set of predefined Consciousness Regions, a set of faulty joint settings corresponding to areas of joint pain, pre-recorded sequences of arm joint position movements, Event-Goal pairs and temporary arm joint position readings. Elements of this database are shown in Table 4.1.

Table 4.1 Elements of the Database

No	Elements ***Belief***	Descriptions
1	Pain Definition	Pre defined joint values (Permanent)
2	Primitive Actions	Predefined (Permanent)
3	Current Joint Values	Subject to change (Temporary)
4	Time of Collection	Subject to change (Temporary)
5	Predicted Joint	Subject to change (Temporary)
6	Time of Occurrence	Subject to change (Temporary)
7	Pain Classification	Subject to change (Temporary)
	Desires / Goals	
8	Pain Evocation	Subject to change (Temporary)
9	Empathy Activation	Subject to change (Temporary)
10	Responses	Subject to change - Event-Goal Pairs (Temporary)
	Intentions	
11	Verbal Warning	Subject to change (Temporary)
12	Responses to Actions	Subject to change (Temporary)

4.1.5 Atomic Actions

The Action Execution module is responsible for translating each decision into one of three intentions: (i) Send alert, (ii) Shift the awareness level through CDV, or (iii) Modify joint stiffness values in the robot's body. If the decision is to maximise joint stiffness, the robot will disregard any external physical interaction, e.g., interaction with a human. By increasing stiffness, the robot joint will resist any force generated by physical interaction, and as a result, the robot will be prevented from experiencing the faulty joint settings. Sensing the resistance of the robot joint, the human will realise that the robot is no longer willing to be involved in the interaction.

4.2 Reasoning Mechanism

The Robot Mind can utilise causal reasoning, as reported in Morgenstern and Stein (1988), Schwind (1999), and Stein and Morgenstern (1994), to draw conclusions from its perceptions. Our idea of reasoning is derived from human cognitive competencies that incorporate the cause and effect relationship (Reisberg, 2013). This enables our framework to allow robots to adapt to the world by predicting their own future states through reasoning about perceived or detected facts. We integrate our approach with sequential pattern prediction (Agrawal and Srikant, 1995; Laird, 1993) to capture the behaviour of the observed facts and then use them to predict possible future conditions.

In ASAF, a robot's decision making is built on associative theory (Schwind, 1999), which utilises covariance information obtained from data sequences to facilitate the causal reasoning process. The Robot Mind analyses the relationships in the covariance of the data obtained from the robot's proprioceptive sensor, that is, the joint position sensor, and derives the sequence data pattern. The prediction process only takes place after several sequences of data have been generated to reduce analysis bias. Any decisions made as a result of previous sequence predictions are reassessed according to the current state, and the results are either kept as history for future prediction, or amendment actions are implemented before the decision is executed. This cycle repeats only if the current data and predicted values in the restricted region that refers to the painful joint settings are not classified.

4.2.1 Pattern Data Acquisition

Raw data from sensory mechanisms are collected and arranged according to retrieval time, and these data are analysed to determine the covariance data. By substituting the data covariance into the latest raw data obtained, the prediction data can obtained. This process is discussed in the following subsubsections, and mathematical representations are derived from the previous chapter, Chapter 3.

Raw Proprioceptive Data

The interaction occurs within a specified constant time span, T. The representation of the data collected at a specified time t_i is $\prod_{t_i,i=0}^{i<T} d_{t_i}$, d_{t_i} represents a joint value at a specified time t_i, where the value of t_i is determined by:

$$t_i = \begin{cases} i = 0, \text{ initiating experiment} \\ i < T, \text{ time span of experiment} \end{cases}$$

Data Covariance

Data covariance is derived from the difference between the last joint values obtained and the previous values, as depicted in Equation 4.1:

$$\Delta int = d_{t_T} - d_{t_{T-1}} \tag{4.1}$$

Prediction Data

Data covariance is used during the process of analysis to formulate a sequence of prediction data, allowing the system to reproduce a new set of prediction data sequences. By substituting Equation 4.1 into the obtained data, d_{t_i}, we can obtain the sequence of the prediction data shown in Equation 4.2.

$$\prod_{t_i,i=m}^{i<\bar{T}} \bar{d}_{t_i} = \prod_{t_i,i=m}^{i<\bar{T}} (d_{t_i} + \Delta int) \tag{4.2}$$

$\bar{d}_{t_i}$ represents the prediction data at sequence time t_i, where the values of t_i are determined by (Criterion 6):

$$t_i = \begin{cases} i = m, \text{ data at time m analysing process is initiated} \\ i < \bar{T}, \text{ discrete time of prediction} \end{cases}$$

where $\bar{T}$ refers to the total number of prediction sequences, and the value of m must satisfy the following conditions (Criteria 7):

$$t_i = \begin{cases} c_s > 0, \text{ total similarity of the obtained joint values reference} \\ c_d > 0, \text{ total difference of the obtained joint values reference} \\ c_u >> c_d; c_u >> c_s, \text{unique data} \end{cases}$$

4.2.2 Causal Reasoning

The overall decision-making process of a robot using the ASAF with the synthetic pain activation mechanism is illustrated in Figure 4.4.

After prediction process takes place, the Mind originates the Event-Goal Pair Queue. A Plan Library is formed by utilizing the data kept in the Database, then the Causal Reasoning process further assesses the Event-Goal Pair Queue which produces first level of recommendation and to be propagated back to the Event-Goal Pair Queue and the Database for updating purposes. In a case of first level recommendation suggesting to modify consciousness level as a result of violation to the restricted joint values, the Robot Mind will constrain the conditions of the exploration of robot awareness regions followed by changing consciousness level to the highest level of subjective awareness region. Updating consciousness is achieved by changing the value of CDV which allows the exploration of these regions, and at the same time, changes the focus of the robot's attention.

Before running an experiment, an expert user sets the Robot Mind as online or offline and specifies whether an SPD based-model or Pain Matrix-model is used. The Robot Mind initially sets the CDV to a random state (this can also be pre-set by the user) enabling the consciousness to select an awareness under unconstrained type. The incoming data from the elbow joint of the robot feeds the reasoning process. The prediction process takes places when the quantity of incoming data satisfies a minimum amount of data collected from the sensory mechanism, which remains the same throughout the process.

Criterion 6 is followed, t_i where $i = m$ and m equals $c - constant\ number\ of\ data$ and the value of c is a constant value defined by the expert user.

Once the quantity criterion has been met, the incoming data is assessed to determine whether the pattern of the Joint Data is similar or different from the pattern of the previous data, otherwise it is categorised as unique data. The reasoning and prediction processes then take place by modifying the Beliefs and updating the Database for any changes. The Robot Mind chooses the most suitable recommendation based on the current Beliefs and passes this recommendation to the Goals. This recommendation covers the interval time of pain occurrence, type of warning to be generated, the state of awareness and the kind of synthetic pain to be evoked. Based on this recommendation, the Intentions are derived and sent to the Action Execution Engine. There are three possible actions to be performed by the Action Execution Engine: activating the alert system, setting the joint stiffness and updating the consciousness region.

In practical, causal reasoning is performed in the following manner. As the robot hand moves, the perception generates a sequence of joint positions and the total number of sequences is set to a specific value. When this value is achieved, the reasoning process proceeds by firstly determining the pattern of the joint position sequence data. There are three types of pattern which are defined in this experiment:

1. Different values:
    - Uniform increased values
    - Uniform decreased values
2. Similar values
3. Unique values

In most cases, if the pattern is categorised as *Similar values*, then the reasoning process most likely recommends that there is no changes occur in the robot hand position. This means that the Robot Mind is aware that there is no physical pushing to the robot hand, and as a result, there is no possibility that synthetic pain is generated. If the pattern matches the *Different values*, the reasoning process commences only after awaiting another additional number of sequence of joint position values. When this additional number of sequence is obtained, the Mind starts generating a set of possible future joint position values by taking the different of the current joint value and the previous joint value then accumulating them. From this set of predicted joint position values, the reasoning process maps into the restricted joint values and assesses the validity of the synthetic pain recommendation.

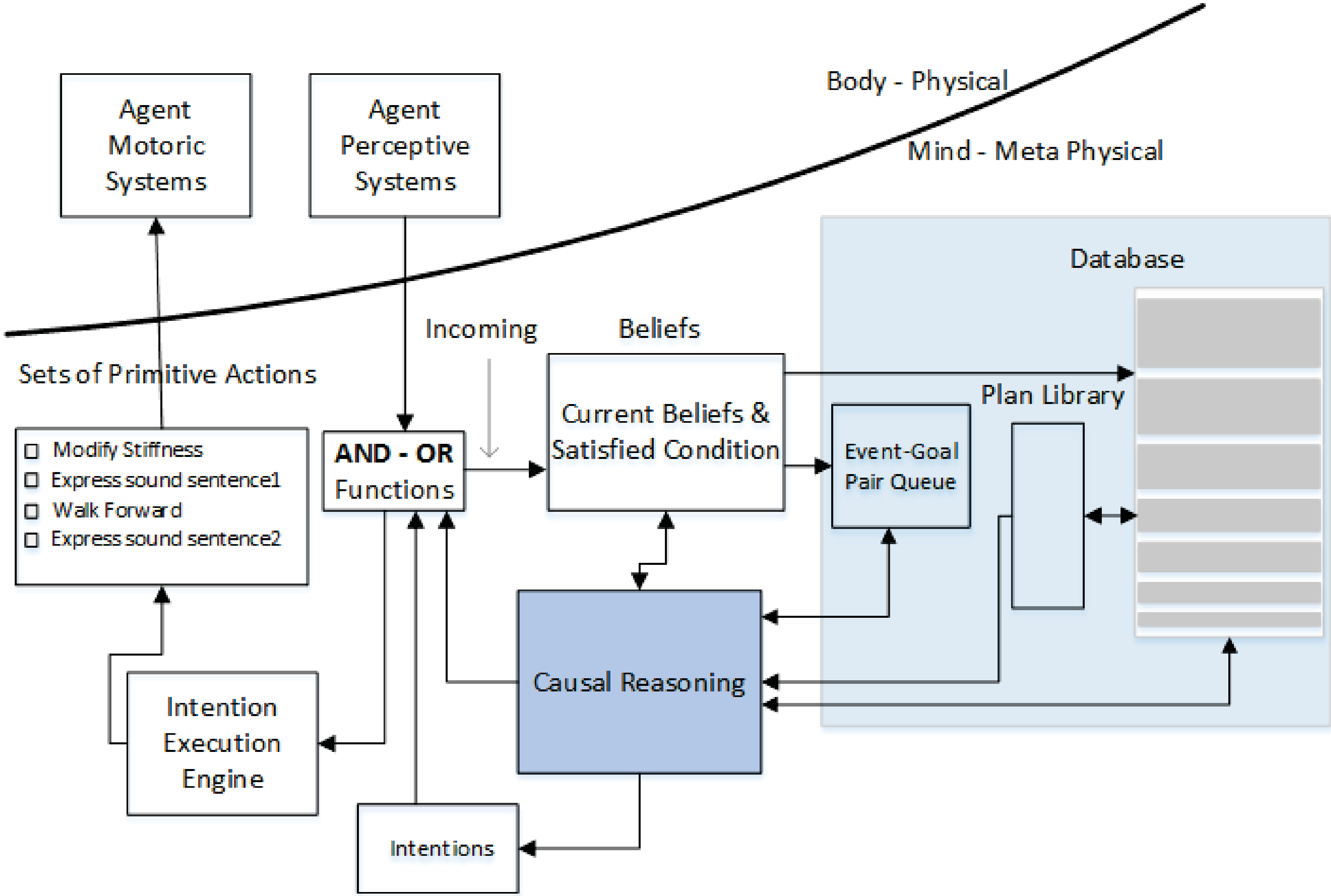

Fig. 4.3 Robot Mind Structure

```
1: database ← (on real time or pre-recorded)
2: CDV region ← random
3: Consciousness ← (CDV)
4: Input  ← joint data; Input ← sensing time
5: While Assessing quantity of data retrieve
       If Similarity Count == Limit set:
          same_count ← Limit
       If Different Count   == Limit set:
          differ_count ← Limit
       Else:
           uni_count ← Limit
6: for each incoming data, DO reasoning
     data_Interval ← same_count OR differ_count OR  uni_count
     update (current_data, current decision, current state)
      Belief ← update
     start predicting ← (current, past, future)
     faulty_joint assessment limit:
        faulty_prediction ← update_prediction(Belief)
        update(Belief)
     P ← update
     database:  update( prediction)
     end for
7: get event-goal pair (e)
     e ← database
     Goal ← Update
8: execute Intention
     Consciousness ← update
     Alert (activation)
     Stiffness (setting)
9: return Input
```

Fig. 4.4 Robot Mind Reasoning Process

Chapter 5

Integration and Implementation

This chapter provides details of the integration of the proposal and the concept of synthetic pain and empathy with pain into the Adaptive Self-Awareness Framework robot framework, the Adaptive Self-Awareness Framework (ASAF).

5.1 Hardware Description

As a proof of concept, the experiment utilises a humanoid robot platform, NAO Aldebaran, and the right arm joint is the preferred joint for artificial pain implementation (depicted in Figure 5.1). Several important features are described in Appendix B and for the complete description of the hardware, see Aldebaran (2006).

5.2 Experiment

The implementation of the ASAF, as mentioned in Chapter 4, is summarised in five key issues as follows.

1. The realisation of *self-consciousness* has two elements:

    i. the ability to focus attention on specific physical aspects of self

    ii. the ability to foresee and consequently, to generate, counter responses as empathic actions.

2. The elements of the position of the right arm joint d_{t_i} and time corresponding to the collection time t_i are obtained by the joint position sensor (proprioceptive). These two data constitute the joint data and the time data.

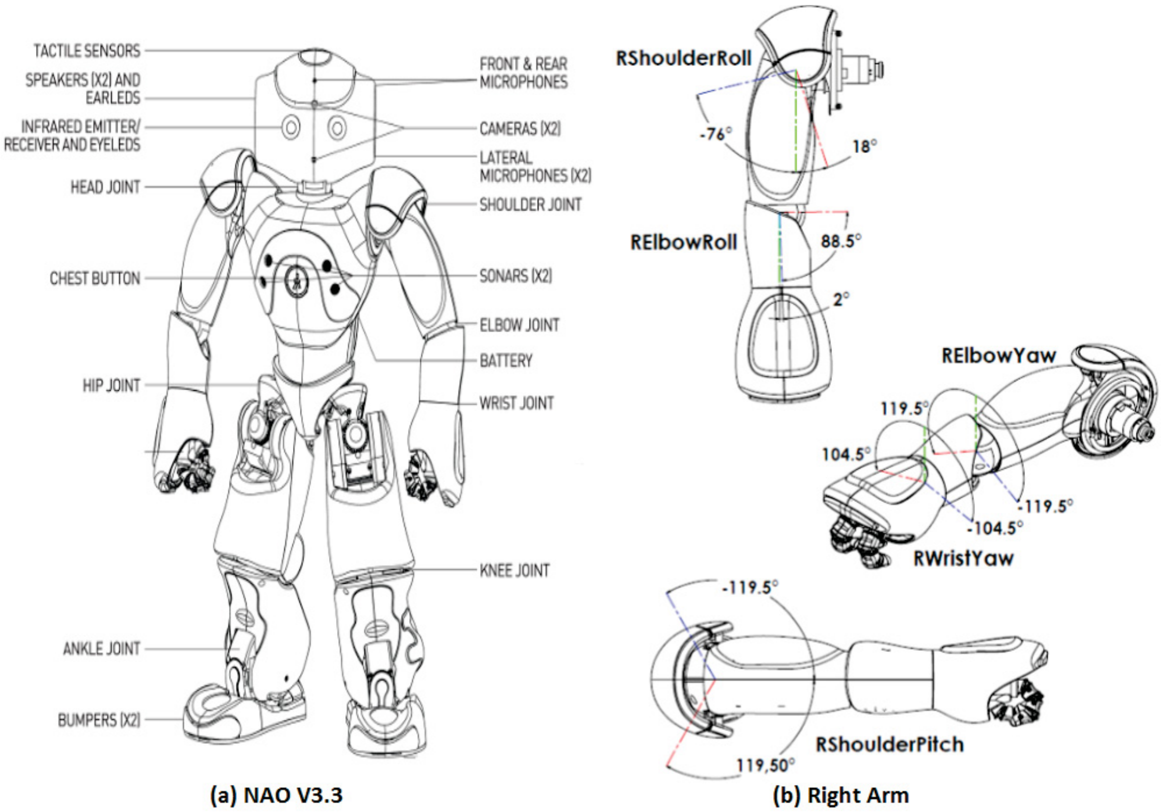

Fig. 5.1 NAO Humanoid Robot (Aldebaran, 2006)

3. The reasoning process produces response times which are derived from the time data prediction for the data of a specified arm joint motion.

4. The Robot Mind states are divided into two conditions:

   i. *unconstrained*, where the Robot Mind is allowed to explore its entire *consciousness* region, e.g. Region 1 to Region 6. This condition occurs by default and it may change throughout the interaction process.

   ii. *constrained*, where the Robot Mind is limited to the highest level of subjective *consciousness*, i.e. Region 1.

   Overall, the change in the state of the Robot Mind, subsequently, affects the awareness of the robot. Hence,the terms *constrained* and *unconstrained* also apply on the Awareness Type of the robot.

5. Empathic experiments are specifically designed to evolve empathic reactions with human shoulder pain as an object of observation.

Two experimental set-ups are prepared which cover the implementation of a non-empathic experiment and an empathic experiment, each of which involves SPD-based and Pain Matrix-based pain activation methods.

5.2.1 Non-empathic Experiment

During the non-empathic experiment, two agents, the NAO robot and a human peer, interact in a shared-task in a static environment; in this case, a hand pushing task and the experiment are divided into offline and online scenarios (the robot set-up is shown in Figure 5.2). In the

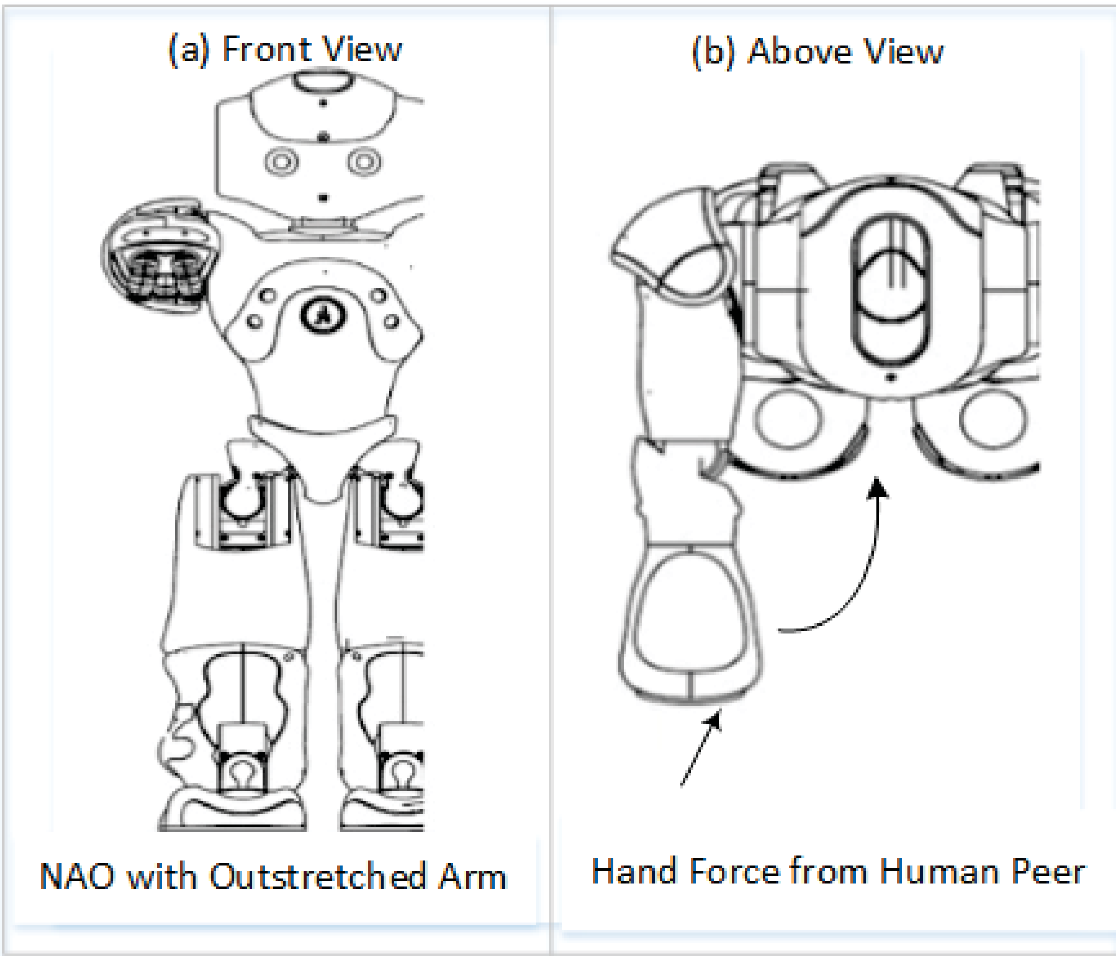

Fig. 5.2 Non Empathic Experiment

offline scenario, the experiment has two stages. In stage one, the robot does not have the awareness framework in the interaction between the robot and the human peer. The purpose of the stage one experiment is to collect a set of elbow joint data - Joint Data and Time Data

- and to place them in the robot database. Two types of action are used to collect the data sets: (i) without physical interaction (phase 1), and (ii) with physical interaction (phase 2). With physical interaction means that the human peer reacts by pushing the arm of the robot, while without physical interaction means that the human peer remains standing in front of the robot without performing a pushing action. Each phase contains five trials that make up a set of ten data sets in total. In the next stage, only the robot with an activated awareness framework performs the actions, without the involvement of a human peer. The experiment is simulated in the robot's mind, and the data for interaction is injected from the datasets obtained from the previous stage and stored in the agent database. This experiment produces an additional set of six datasets, containing data predictions. This stage is designed to first evaluate the mind simulation of the robot's reasoning performance through its ability to shift its consciousness using pre-recorded elbow joint datasets. Second, it is designed to measure the accuracy of the agent's reasoning skills through the ability to predict and generate accurate pain acknowledgement, and the counter-responses carried out by the intention execution engine. In the online scenario, the robot and the human peer perform an interaction; however this time, the robot performs with an activated self-awareness framework. The interaction with the human peer therefore provides the joint data straight away for further processing. This experiment is divided into two phases: phase one without physical interaction and phase two with physical interaction. The objectives of this experiment scenario are to measure the overall performance of the agent with the self-awareness framework embedded in its mechanism, including the robustness of the framework in a real world environment. All the data collected in these two scenarios were ordered according to their reading sequences unless stated otherwise.

5.2.2 Empathic Experiment

The concept of empathy with pain is generated by the projection of the shoulder movements of humans who suffer from a motor injury onto a robot observer's shoulder. The observer robot visually captures (exteroceptively) the shoulder motions and reflects them on its own arm, while also analysing the kinds of of synthetic pain to generate. Three agents are involved: two NAO robots and a human peer. One robot acts as an observer while the other acts as a mediator and helps the human peer (see Figure 5.3 for the initial pose for the robots). The pilot experiment only considers up- and down-rotational direction motions of the human peer's right shoulder. As the human peer shoulder dimension is different from that of the NAO observer (Observer), another NAO robot is introduced as a mediator robot (Mediator). Through the Mediator, the length of the rotation movement of the human shoulder is adjusted

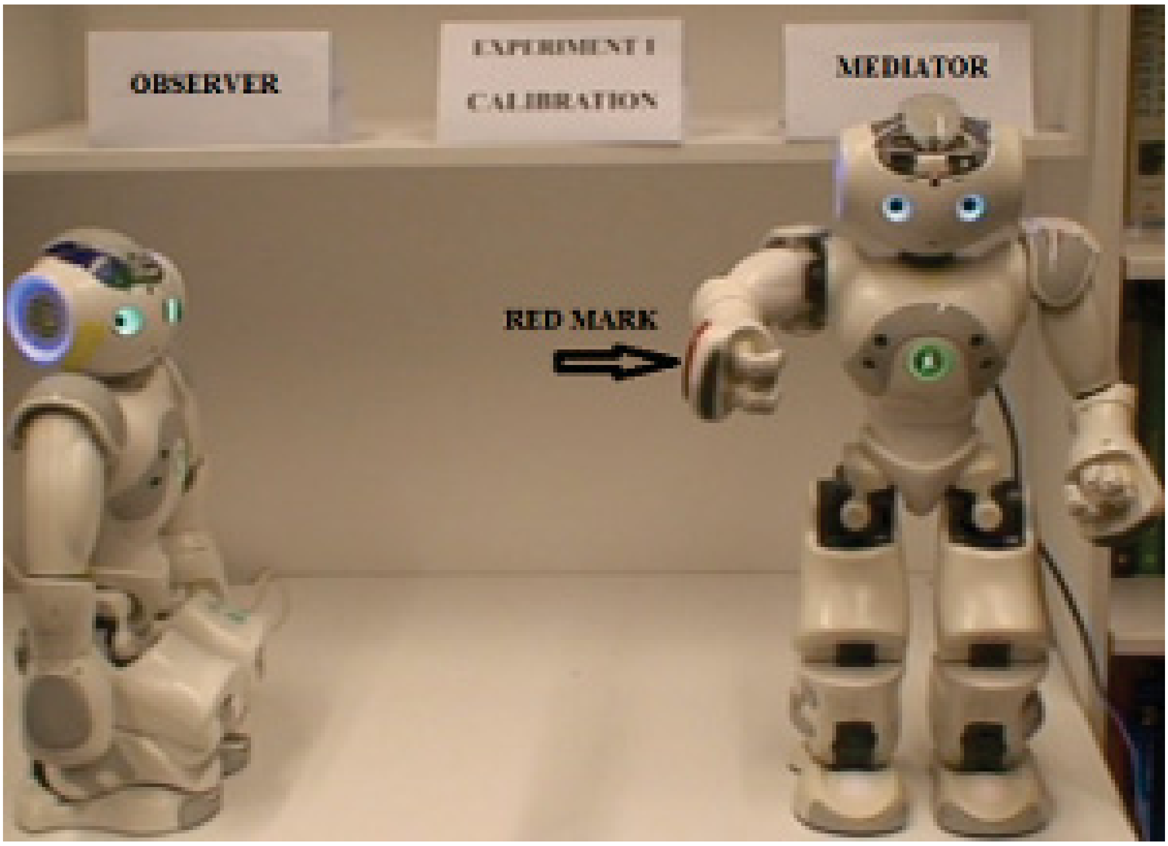

Fig. 5.3 Initial Pose for Robot Experiments

to the length of the shoulder rotation of the Mediator. A red circular mark is attached to the back of the Mediator's hand which will be recognised by the Observer via its camera sensor.

During the experiment, the human peer moves his hand in vertical up- and down-ward motions. The human's hand holds the finger tips of the Mediator's hand which allows both hands to move in parallel. Each hand motion of the Mediator's shoulder joint produces joint position values obtained from the joint position sensor. The Observer converts the visual representation of the Mediator's hand position using a standard geometric-based transformation (see Figure 5.4).

5.3 Pre-defined Values

All the experiments require the interaction between robots and a human peers to take place within a pre-defined environment setting. Several data are defined by an expert user and placed in the Database (see Table 5.1 for the list of pre-defined values). For the SPD model, the faulty joint settings that correspond to the pain region to be avoided have only three

Table 5.1 Pre-Defined Values in the Database

No	Data	Details					
1	**Faulty Joint Setting - SPD Model**	**Level**					
		High		Medium			Low
		1.5621		1.5521			1.5421
2	**Faulty Joint Setting - Pain Matrix Model**	**Level**					
		Upward			**Downward**		
		High	Medium	Low	Low	Medium	High
		-2.08313	-1.58313	-1.38313	1.385669	1.585669	2.085669
3	**Awareness Regions**	**Awareness**			**Value Limit**		**Region Width**
		Upper Subjective			1 -25		24
		Lower Subjective			26 - 50		
		Left Subjective-Objective			51 - 75		
		Right Subjective-Objective			76 - 100		
		Lower Objective			101 - 125		
		Upper Objective			126 - 150		

levels, while in the Pain Matrix model, there are three upward levels and three downward levels. The width of the awareness region remains the same throughout the experiments. The states of robot awareness during the non-empathic experiments are shown in Table 5.2 and the actual kinds of pain to be generated are shown in Table 5.3.

Table 5.2 Awareness State

Consciousness Region		**Robot Action During Visitation**	
		***Awareness Type*: Unconstrained**	***Awareness Type* : Constrained**
Subjective Awareness	Upper Limit	Low Stiffness on Arm Joint	Increased Stiffness and Alert human peer
	Lower Limit	Not Modelled	Not Available
Subjective-Objective Awareness	Left Limit	Not Modelled	Not Available
	Right Limit	Not Modelled	Not Available
Objective Awareness	Lower Limit	Not Modelled	Not Available
	Upper Limit	Not Modelled	Not Available

Table 5.3 Synthetic Pain Experiment

Synthetic Pain	**Description**	**Intensity Level**	**Experiments**	
			SPD Model	**Pain Matrix Model**
Proprioceptive	1.1	Slight	Modelled	Modelled
	2.0	None	Modelled	Modelled
	2.1	Slight	Modelled	Modelled
Inflammatory Reduction	2.2	Moderate	-	Modelled
	2.3	Severe	-	-
	3.0	None	-	Modelled
	3.1	Slight	-	Modelled
Sensory Malfunctions	3.2	Moderate	-	Modelled
	3.3	Severe	-	-

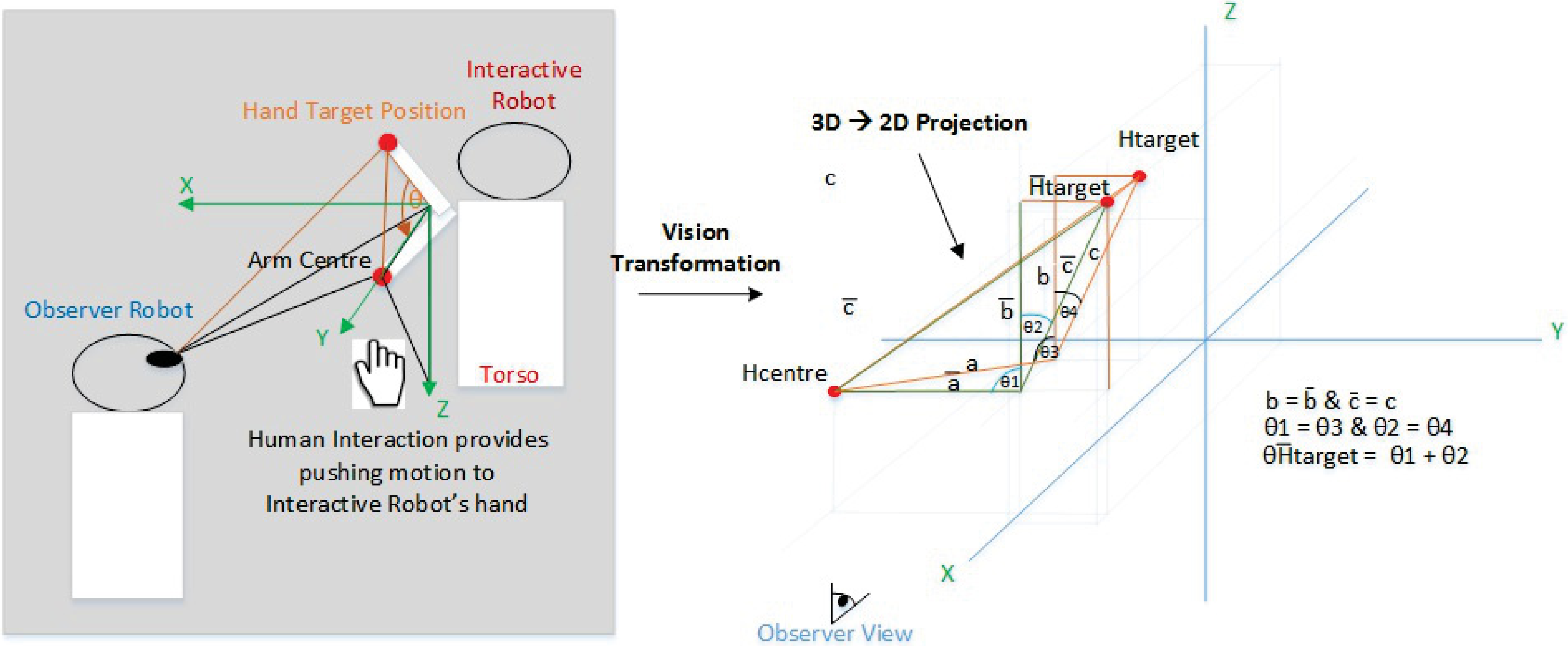

Fig. 5.4 Geometrical Transformation

Chapter 6

Results, Analysis and Discussion

This chapter provides an overview of experiments, the results collected during the experiments reported in Chapter 5 and presents analysis and discussion of the outcomes.

6.1 Experiment Overview

Two groups of experiments were carried out: non empathy-based experiments and empathy-based experiments. The data are divided into the Joint Data, which refers to the current data originating from the Sensory Mechanism, and the prediction data, which represents a set of data predictions internally proposed by the reasoning process. We use the terminology sequence when referring to the Joint Data originating from the sensor readings, and each sequence contains four prediction cycles. Each prediction cycle predicts one future step sequence of the Joint Data. The kinds of data to be collected for each experiment are shown in

Table 6.1 Experiment Overview

		Scenario			
		Offline		Online	
Experiments	**Pain Activation**	**No Shared Task**	**Shared Task Hand Push Direction**	**No Shared Task**	**Shared Task Hand Push Direction**
Non-empathy	SPD-based	✓	Horizontal	✓	Horizontal
	Pain Matrix-based	x	x	✓	Vertical
Empathy	SPD-based	x	x	x	Vertical
	Pain Matrix-based	x	x	x	Vertical

Table 6.1 below. All important experiment results will be presented and discussed thoroughly in this chapter; a number of supporting results are shown in Appendix C.

Overall, the experiments are classified according to Non-empathy based experiments and Empathy-based experiments as described below.

1. Non-empathy based experiments:

   (a) SPD-based Model:

      i. Offline pre-recorded without physical interaction
         - Trial 1
         - Trial 2
         - Trial 3
         - Trial 4
         - Trial 5

      ii. Offline pre-recorded with physical interaction
         - Trial 1
         - Trial 2
         - Trial 3
         - Trial 4
         - Trial 5

      iii. Online pre-recorded without physical interaction
         - Trial 1
         - Trial 2
         - Trial 3
         - Trial 4
         - Trial 5

      iv. Online pre-recorded with physical interaction
         - Trial 1
         - Trial 2
         - Trial 3
         - Trial 4
         - Trial 5

   (b) Pain Matrix-based Model:

      i. Upward hand movement

      ii. Downward hand movement

2. Empathy-based experiments:

   (a) SPD-based Model:

i. Upward hand movement
ii. Downward hand movement

(b) Pain Matrix-based Model:

i. Upward hand movement
ii. Downward hand movement

6.2 Non-empathy based Experiments

The experiment results and discussion are divided into two subsections: the SPD-based model and the Pain Matrix-based model.

6.2.1 SPD-based Model

The sequence of joint positions data used in the offline scenario, without physical interaction and with physical interaction, is shown in Table 6.2, Table 6.3, Table 6.4 and Table 6.5 respectively. The data used for the online scenario, without physical interaction and with physical interaction, are shown in Table 6.6, Table 6.7, Table 6.8 and Table 6.9 respectively.

SPD General Discussion

During the offline scenario (without physical interaction) in Trial 1 to Trial 5, the Joint Data is 0.00873 and this value remains uniform throughout the trials (each trial consists of 21th sequences). The Joint Data (similar to Elbow Data) is retrieved by the robot sensor position at an average interval time of 0.56 across all 21 sequences. In the early sequences, the interval time is 0.52 and lasts until the 4th sequence. From the 5th sequence to the 21st sequence, the interval time is more or less the same, around 0.56. This small fluctuation may introduce noise to the Joint Data and it will degrade the quality of prediction results. The overall interval data retrieval time is shown in 6.10. Joint prediction and reasoning reach the margin of zero per cent error, where the standard deviation (σ) is zero, and the standard deviation for time prediction reaches the maximum deviation at 0.05 (see Table 6.11). The robot commences its prediction at the fourth sequence and the reasoning is capable of maintaining prediction accuracy, which allows the agent to identify the scenario. See the graphical comparisons of the predication data for each trial in Figure 6.1, Figure 6.2, Figure 6.3, Figure 6.4 and Figure 6.5.

During the offline period without physical interaction, the hand interaction with the human partner produces motions on the robot hand. These motions constitute different values in

Table 6.2 Offline Pre-Recorded without Physical Interaction Trial 1 to Trial 3

No	Trial 1		Trial 2		Trial 3	
	Elbow Data	Time	ElbowData	Time	Elbow Data	Time
1	0.00873	221.72	0.00873	316.95	0.00873	385.60
2	0.00873	222.24	0.00873	317.47	0.00873	386.12
3	0.00873	222.75	0.00873	317.98	0.00873	386.64
4	0.00873	223.28	0.00873	318.52	0.00873	387.16
5	0.00873	223.84	0.00873	319.07	0.00873	387.72
6	0.00873	224.40	0.00873	319.63	0.00873	388.30
7	0.00873	224.96	0.00873	320.19	0.00873	388.85
8	0.00873	225.51	0.00873	320.75	0.00873	389.41
9	0.00873	226.07	0.00873	321.31	0.00873	389.97
10	0.00873	226.63	0.00873	321.86	0.00873	390.54
11	0.00873	227.19	0.00873	322.42	0.00873	391.10
12	0.00873	227.76	0.00873	322.98	0.00873	391.66
13	0.00873	228.32	0.00873	323.54	0.00873	392.22
14	0.00873	228.88	0.00873	324.11	0.00873	392.78
15	0.00873	229.44	0.00873	324.67	0.00873	393.35
16	0.00873	230.01	0.00873	325.24	0.00873	393.92
17	0.00873	230.58	0.00873	325.80	0.00873	394.49
18	0.00873	231.14	0.00873	326.37	0.00873	395.05
19	0.00873	231.71	0.00873	326.95	0.00873	395.62
20	0.00873	232.28	0.00873	327.52	0.00873	396.19
21	0.00873	232.85	0.00873	328.10	0.00873	396.76

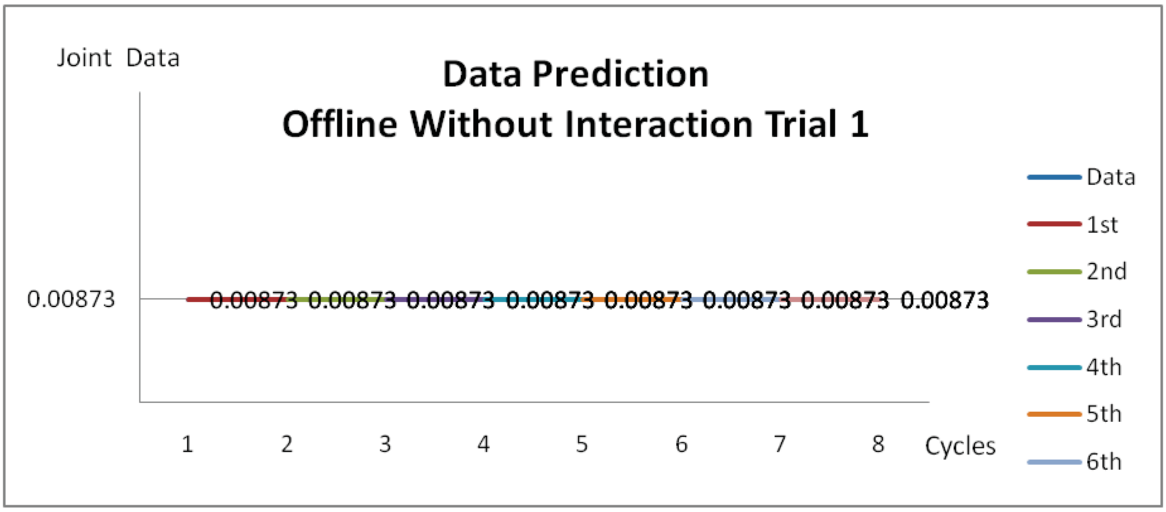

Fig. 6.1 Offline without Human Interaction Trial 1

Table 6.3 Offline Pre-Recorded without Physical Interaction Trial 4 and Trial 5

No	Trial 4		Trial 5	
	Elbow Data	Time	Elbow Data	Time
1	0.00873	449.60	0.00873	514.15
2	0.00873	450.12	0.00873	514.67
3	0.00873	450.64	0.00873	515.18
4	0.00873	451.16	0.00873	515.71
5	0.00873	451.72	0.00873	516.28
6	0.00873	452.28	0.00873	516.84
7	0.00873	452.84	0.00873	517.39
8	0.00873	453.40	0.00873	517.96
9	0.00873	453.95	0.00873	518.51
10	0.00873	454.51	0.00873	519.07
11	0.00873	455.08	0.00873	519.64
12	0.00873	455.65	0.00873	520.21
13	0.00873	456.21	0.00873	520.77
14	0.00873	456.78	0.00873	521.33
15	0.00873	457.35	0.00873	521.89
16	0.00873	457.91	0.00873	522.46
17	0.00873	458.48	0.00873	523.02
18	0.00873	459.04	0.00873	523.60
19	0.00873	459.60	0.00873	524.16
20	0.00873	460.18	0.00873	524.73
21	0.00873	460.74	0.00873	525.30

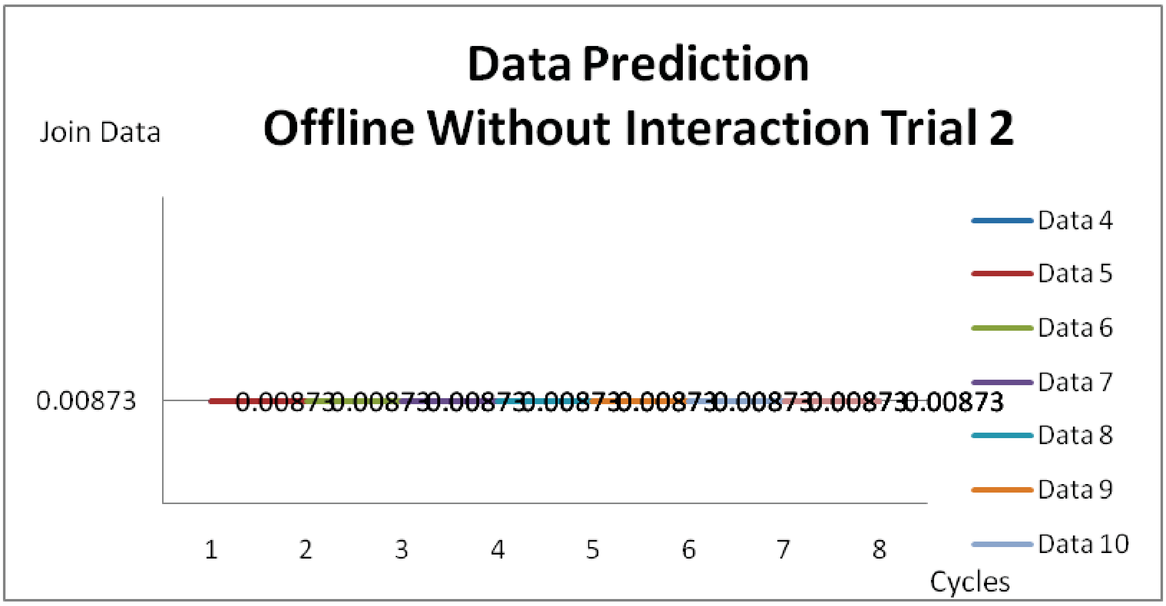

Fig. 6.2 Offline without Human Interaction Trial 2

Table 6.4 Offline Pre-Recorded with Physical Interaction Trial 1 to Trial 3

No	Trial 1		Trial 2		Trial 3	
	Elbow Data	Time	Elbow Data	Time	Elbow Data	Time
1	0.00873	583.66	0.02919	644.54	0.02919	726.37
2	0.02765	584.17	0.02919	645.05	0.02919	726.89
3	0.14884	584.69	0.03072	645.56	0.04606	727.40
4	0.34519	585.22	0.03072	646.09	0.22861	727.93
5	0.57836	585.76	0.13043	646.62	0.40348	728.46
6	0.78238	586.31	0.45871	647.14	0.60444	728.98
7	1.02782	586.85	0.73023	647.66	0.88669	729.50
8	1.30701	587.40	0.94959	648.19	1.08765	730.03
9	1.51870	587.94	1.14441	648.72	1.25485	730.55
10	1.56207	588.48	1.37297	649.25	1.42359	731.08
11	1.56207	589.03	1.53251	649.77	1.56012	731.60
12	1.56207	589.56	1.56207	650.30	1.56207	732.14
13	1.56207	590.08	1.56207	650.83	1.56207	732.66
14	1.56207	590.63	1.56207	651.38	1.56207	733.19
15	1.56207	591.18	1.56207	651.93	1.56207	733.74
16	1.56207	591.74	1.56207	652.48	1.56207	734.29
17	1.56207	592.30	1.56207	653.04	1.56207	734.85
18	1.56207	592.85	1.56207	653.59	1.56207	735.40
19	1.56207	593.40	1.56207	654.14	1.56207	735.96
20	1.56207	593.95	1.56207	654.70	1.56207	736.51

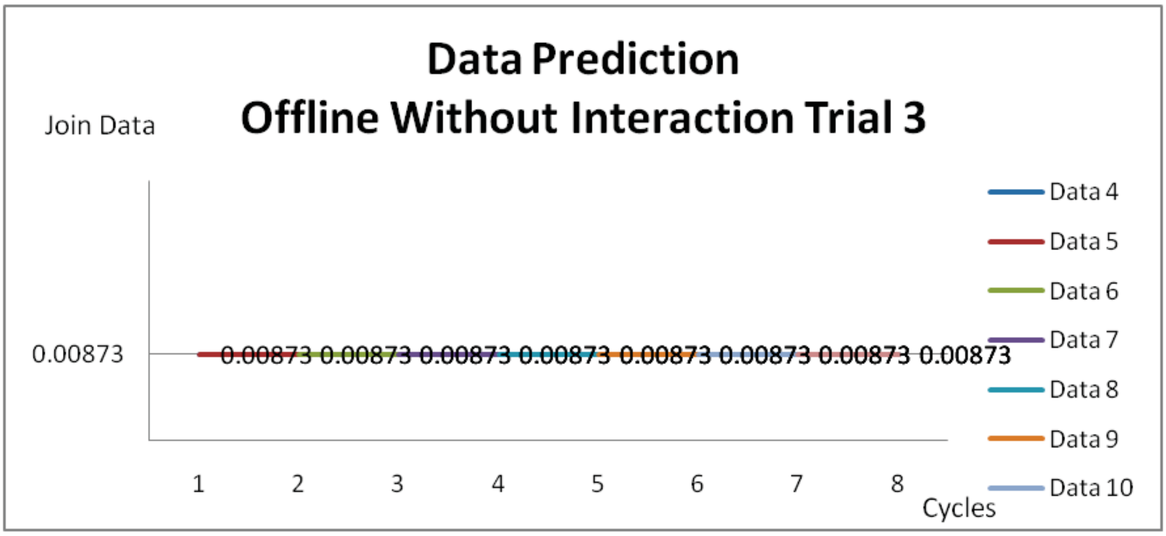

Fig. 6.3 Offline without Human Interaction Trial 3

Table 6.5 Offline Pre-Recorded with Physical Interaction Trial 4 and Trial 5

No	Trial 4		Trial 5	
	Elbow Data	Time	Elbow Data	Time
1	0.02919	773.30	0.02919	823.77
2	0.02765	773.82	0.02765	824.29
3	0.02919	774.34	0.03072	824.81
4	0.02919	774.86	0.14884	825.33
5	0.07214	775.38	0.28997	825.88
6	0.36974	775.92	0.44183	826.42
7	0.74096	776.44	0.59370	826.96
8	1.06924	776.96	0.73176	827.51
9	1.50643	777.49	0.87289	828.06
10	1.56207	778.02	1.01402	828.60
11	1.56207	778.54	1.13980	829.15
12	1.56207	779.07	1.27480	829.69
13	1.56207	779.62	1.39445	830.24
14	1.56207	780.17	1.50643	830.79
15	1.56207	780.72	1.56207	831.34
16	1.56207	781.28	1.56207	831.89
17	1.56207	781.83	1.56207	832.42
18	1.56207	782.38	1.56207	832.94
19	1.56207	782.93	1.56207	833.50
20	1.56207	783.49	1.56207	834.06

Table 6.6 Online without Physical Interaction Trial 1 to Trial 3

No	**Trial 1**		**Trial 2**		**Trial 3**	
	Elbow Data	**Time**	**Elbow Data**	**Time**	**Elbow Data**	**Time**
1	0.02765	367.27	0.02765	551.84	0.02765	793.73
2	0.02765	367.79	0.02765	552.36	0.02765	794.25
3	0.02765	368.30	0.02612	552.87	0.02765	794.76
4	0.02765	368.83	0.02765	553.4	0.02765	795.28
5	0.02765	369.38	0.02765	553.93	0.02612	795.84
6	0.02765	369.94	0.02765	554.46	0.02765	796.38
7					0.02612	796.94

Table 6.7 Online without Physical Interaction Trial 4 and Trial 5

No	Trial 4		Trial 5	
	Elbow Data	**Time**	**Elbow Data**	**Time**
1	0.02765	971.52	0.02612	354.21
2	0.02765	972.04	0.02612	354.73
3	0.02765	972.55	0.02765	355.24
4	0.02765	973.08	0.02765	355.77
5	0.02765	973.64	0.02612	356.3
6	0.02765	974.2	0.02612	356.82
7			0.02612	357.38

Table 6.8 Online with Physical Interaction Trial 1 to Trial 3

No	Trial 1		Trial 2		Trial 3	
	Elbow Data	**Time**	**Elbow Data**	**Time**	**Elbow Data**	**Time**
1	0.22247	38.88	0.02765	776.46	0.02765	267.12
2	0.26696	39.40	0.02765	776.98	0.02765	267.64
3	0.37127	39.91	0.02919	777.49	0.02765	268.15
4	0.49246	40.44	0.21940	778.01	0.06907	268.68
5	0.63205	41.00	0.39735	778.54	0.29917	269.20
6	0.78852	41.54	0.68421	779.11	0.52774	269.73
7	0.95572	42.10	1.30548	781.36	0.71642	270.28
8	1.32695	44.37			0.87902	270.84
9					1.04470	271.39
10					1.41132	273.68

Table 6.9 Online with Physical Interaction Trial 4 and Trial 5

No	Trial 4		Trial 5	
	Elbow Data	**Time**	**Elbow Data**	**Time**
1	0.02612	550.05	0.02765	855.20
2	0.02765	550.57	0.03839	856.34
3	0.02919	551.08	0.29150	856.86
4	0.22861	551.61	0.53234	857.39
5	0.46945	552.16	0.74403	857.94
6	0.69188	552.71	1.25639	860.23
7	1.22878	554.99		

Table 6.10 Offline without Physical Interaction - Interval Time

No	**Interval Time**				
	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**	**Trial 5**
1	-	-	-	-	-
2	0.52	0.52	0.52	0.52	0.52
3	0.51	0.51	0.52	0.52	0.51
4	0.53	0.54	0.52	0.52	0.53
5	0.56	0.55	0.56	0.56	0.57
6	0.56	0.56	0.58	0.56	0.56
7	0.56	0.56	0.55	0.56	0.55
8	0.55	0.56	0.56	0.56	0.57
9	0.56	0.56	0.56	0.55	0.55
10	0.56	0.55	0.57	0.56	0.56
11	0.56	0.56	0.56	0.57	0.57
12	0.57	0.56	0.56	0.57	0.57
13	0.56	0.56	0.56	0.56	0.56
14	0.56	0.57	0.56	0.57	0.56
15	0.56	0.56	0.57	0.57	0.56
16	0.57	0.57	0.57	0.56	0.57
17	0.57	0.56	0.57	0.57	0.56
18	0.56	0.57	0.56	0.56	0.58
19	0.57	0.58	0.57	0.56	0.56
20	0.57	0.57	0.57	0.58	0.57
21	0.57	0.58	0.57	0.56	0.57
Average	**0.56**	**0.56**	**0.56**	**0.56**	**0.56**

Table 6.11 Prediction Error - Offline No Interaction

Data	Prediction Cycles						Std D(σ)	Time
	4	5	6	7	8	9		Std D(σ)
4	0.00%							
5	0.00%	0.00%					0.00	0.02
6	0.00%	0.00%	0.00%				0.00	0.03
7	0.00%	0.00%	0.00%	0.00%			0.00	0.05
8	0.00%	0.00%	0.00%	0.00%	0.00%		0.00	0.05
9		0.00%	0.00%	0.00%	0.00%	0.00%	0.00	0.01
10			0.00%	0.00%	0.00%	0.00%	0.00	0.01
11				0.00%	0.00%	0.00%	0.00	0.02

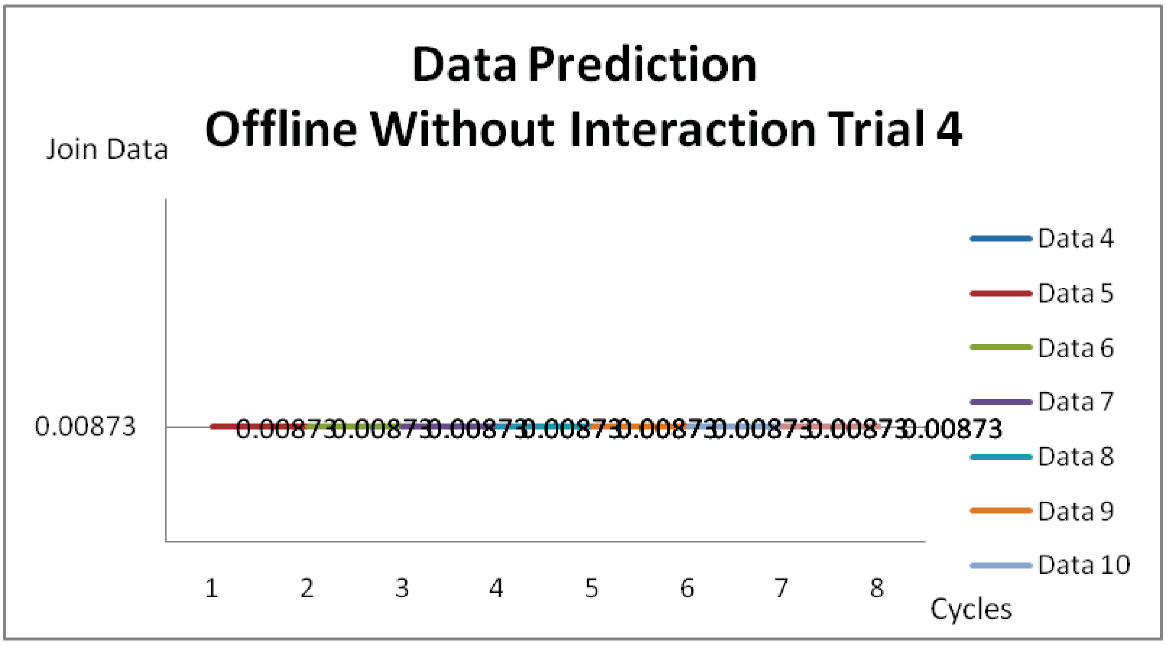

Fig. 6.4 Offline without Human Interaction Trial 4

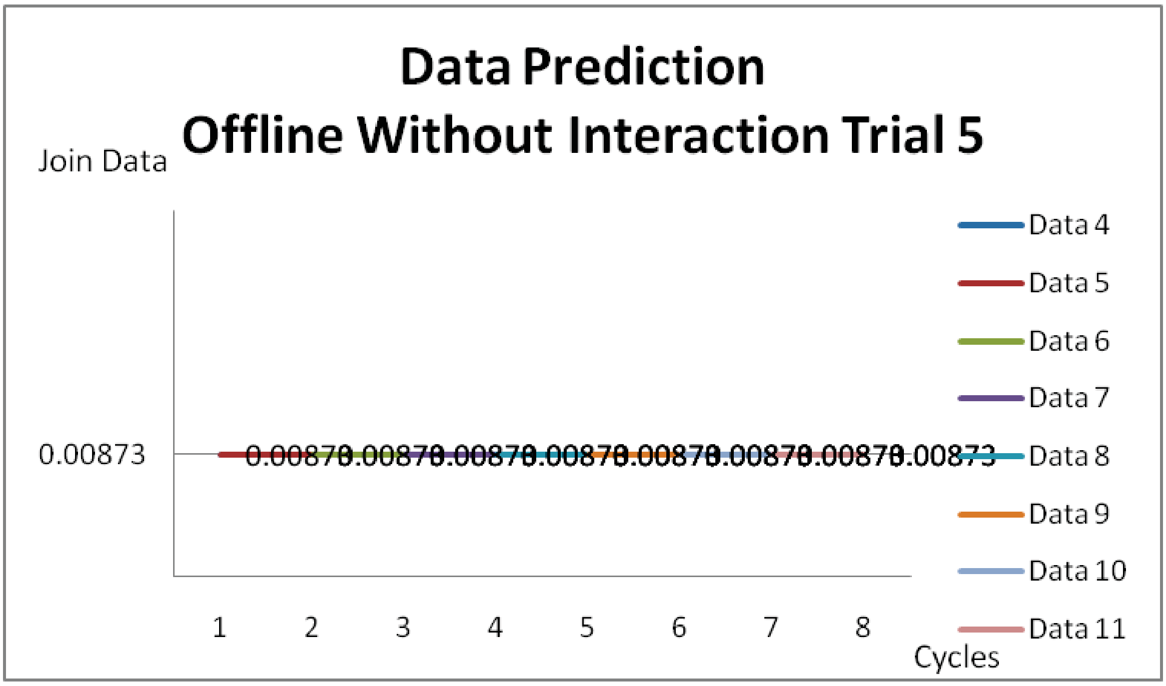

Fig. 6.5 Offline without Human Interaction Trial 5

the Joint Data (refer to Table 6.4 and Table 6.5). In Trial 1, the 1st sequence obtain a Joint Data value of 0.00873. There is a slight increase in this value in each sequence as the human partner pushes the robot's hand: the value in the 2nd is 0.02 and in the 3rd is 0.012. From the 4th sequence to the 9th sequence, there is a similar slight increase of 0.2. At the 10th sequence, the value decreases as the joint movement approaches the faulty joint regions, and the Joint Data value halts at 1.56207 in the 11th sequence. From the 12th sequence to the 20th sequence, the value remains the same, which indicates that the robot has increased its joint stiffness to counteract the force on the robot's hand. The interval time of the sensor data readings is largely the same throughout the sequences at 0.5. In Trial 2, the motions are detected in the 5th sequence with an increase of 0.10, from 0.03072 to 0.12043. A variable increase occurs in the Joint Data from the 5th sequence to the 10th sequence and slowly drops to 0.03 in the 12th sequence. The Joint Data remains the same in the 13th sequence as the robot increases the stiffness of its arm joint. The interval times of sensor readings are the same as those in Trial 1, and are relatively stable at 0.5, which is the same as in Trial 3. However, the Joint Data in Trial 3 fluctuates and halts at the 12th sequence (the joint stiffness is activated). Details of the interval Joint Data and Time for Trials 1 to 3 are depicted in Table 6.12. In Trial 4, the changes in the Joint Data occur abruptly. From the 1st sequence to the 5th sequence, the value of the Joint Data is less than 0.075, which increases to 1.06924 in only two sequences of data. The robot activates its stiffness on the arm joint at the 10th sequence. By contrast, Trial 5 produces a slow increase in the Joint Data with an interval change of around 0.1. This situation lasts until the 14th sequence when the Joint Data starts to reach the maximum value, which triggers the robot to maximise its arm joint stiffness. The interval times of the sensor readings in both trials share the same pattern, with an interval of change around 0.5. Table 6.13 depicts the overall interval values of the Joint Data and sensor reading times.

An analysis of prediction accuracy follows in respect of the offline experiment with physical interaction involving a human partner, online experiment without physical interaction, and online with physical interaction. Each Joint Data sequence contains a minimum of 4 prediction cycles, but in some cases we disregard some cycles as they do not influence the process.

During the offline experiment with physical interaction, Trial 1, the robot deciphers the incoming interaction data at the 4th sequence, with a standard deviation that is relatively low, 0.03 at the 5th sequence of Joint Data. The first prediction at the 4th sequence of Joint Data generates a low prediction error, less than 10%, while the fourth sequence produces a slightly higher prediction error, 17.64%. At the 5th sequence of Joint Data, the prediction error remains very low: the first prediction error is 2.92%, and the fourth prediction error decreases

Table 6.12 Interval Joint Data and Time Offline with Physical Interaction Trial 1 to Trial 3

No	Trial 1		Trial 2		Trial 3	
	Elbow Data	**Time**	**Elbow Data**	**Time**	**Elbow Data**	**Time**
1	-	-	-	-	-	-
2	0.02	0.51	0.00	0.51	0.00	0.52
3	0.12	0.52	0.00	0.51	0.02	0.51
4	0.20	0.53	0.00	0.53	0.18	0.53
5	0.23	0.54	0.10	0.53	0.17	0.53
6	0.20	0.55	0.33	0.52	0.20	0.52
7	0.25	0.54	0.27	0.52	0.28	0.52
8	0.28	0.55	0.22	0.53	0.20	0.53
9	0.21	0.54	0.19	0.53	0.17	0.52
10	0.04	0.54	0.23	0.53	0.17	0.53
11	0.00	0.55	0.16	0.52	0.14	0.52
12	0.00	0.53	0.03	0.53	0.00	0.54
13	0.00	0.52	0.00	0.53	0.00	0.52
14	0.00	0.55	0.00	0.55	0.00	0.53
15	0.00	0.55	0.00	0.55	0.00	0.55
16	0.00	0.56	0.00	0.55	0.00	0.55
17	0.00	0.56	0.00	0.56	0.00	0.56
18	0.00	0.55	0.00	0.55	0.00	0.55
19	0.00	0.55	0.00	0.55	0.00	0.56
20	0.00	0.55	0.00	0.56	0.00	0.55

Table 6.13 Interval Joint Data and Time Offline with Physical Interaction Trial 4 and Trial 5

No	Trial 4 Elbow Data	Trial 4 Time	Trial 5 Elbow Data	Trial 5 Time
1	-	-	-	-
2	0.00	0.52	0.00	0.52
3	0.00	0.52	0.00	0.52
4	0.00	0.52	0.12	0.52
5	0.04	0.52	0.14	0.55
6	0.30	0.54	0.15	0.54
7	0.37	0.52	0.15	0.54
8	0.33	0.52	0.14	0.55
9	0.44	0.53	0.14	0.55
10	0.06	0.53	0.14	0.54
11	0.00	0.52	0.13	0.55
12	0.00	0.53	0.14	0.54
13	0.00	0.55	0.12	0.55
14	0.00	0.55	0.11	0.55
15	0.00	0.55	0.06	0.55
16	0.00	0.56	0.00	0.55
17	0.00	0.55	0.00	0.53
18	0.00	0.55	0.00	0.52
19	0.00	0.55	0.00	0.56
20	0.00	0.56	0.00	0.56

to 0.77%. In the 6th sequence, the prediction error at the second and the third cycles are relatively high, 11.66% and 12.43% respectively. However, it drops dramatically to 3.64% on the final prediction. In the 7th and 8th sequences, prediction error is low, 3.37% and 6.75% respectively, but produces the highest error rate at 20.21% and 30.33% respectively. The 9th sequence produces a relatively high prediction error which of 16.83%. Hence, it can be seen that the 5th sequence of Joint Data, prediction error is the lowest throughout Trial 1, while the highest prediction error occurs in the second cycle of prediction during the 8th Joint Data sequence. In spite of the high prediction error, the overall standard deviation remains low at 0.11 (see Table 6.14). A graphical comparison is shown in Figure 6.6.

Table 6.14 Prediction Error - Offline Physical Interaction Trial 1

Data	Prediction Cycles						Std D(σ)
	4	5	6	7	8	9	
4	0.00%						
5	3.68%	0.00%					0.03
6	4.45%	2.92%	0.00%				0.04
7	9.36%	1.69%	4.14%	0.00%			0.05
8	17.64%	2.91%	11.66%	3.37%	0.00%		0.07
9		0.77%	12.43%	0.00%	6.75%	0.00%	0.07
10			3.64%	20.21%	30.33%	16.83%	0.11

In Trial 2, the process of deciphering incoming data commences at the 6th sequence when the first prediction error is 5.68%. During this sequence, the prediction error increases to 39.89% at the fourth prediction, which is considerably higher. Similarly, the 7th sequence produces a relatively high prediction error with slightly lower error values compared to the previous sequence: 5.22% in first prediction, which increases to 28.38% in the fourth prediction. Three prediction values remain low in the 8th sequence. However, the fourth value still generates a high prediction error of 26.50%. This situation also occurs in the next sequence with a low prediction error in the first and the second values, 3.37% and 0.15% respectively. The third prediction error increases dramatically to 16.68%. The 10th sequence generates a considerably higher prediction error in the final value at 26.80%. During this trial, the overall prediction error results in higher prediction error on the final prediction value throughout the Joint Data sequences, producing the highest error of 39.89% in the 6th sequence. Table 6.15 and Figure 6.7 show the overall data comparisons. Table 6.16 and Figure 6.8 depict the prediction error of Joint Data obtained during the offline physical interaction Trial 3. It can be seen that the first data prediction process occurs in $5th$ sequence with a low prediction error, 2.61%. The next three prediction values increase and are relatively

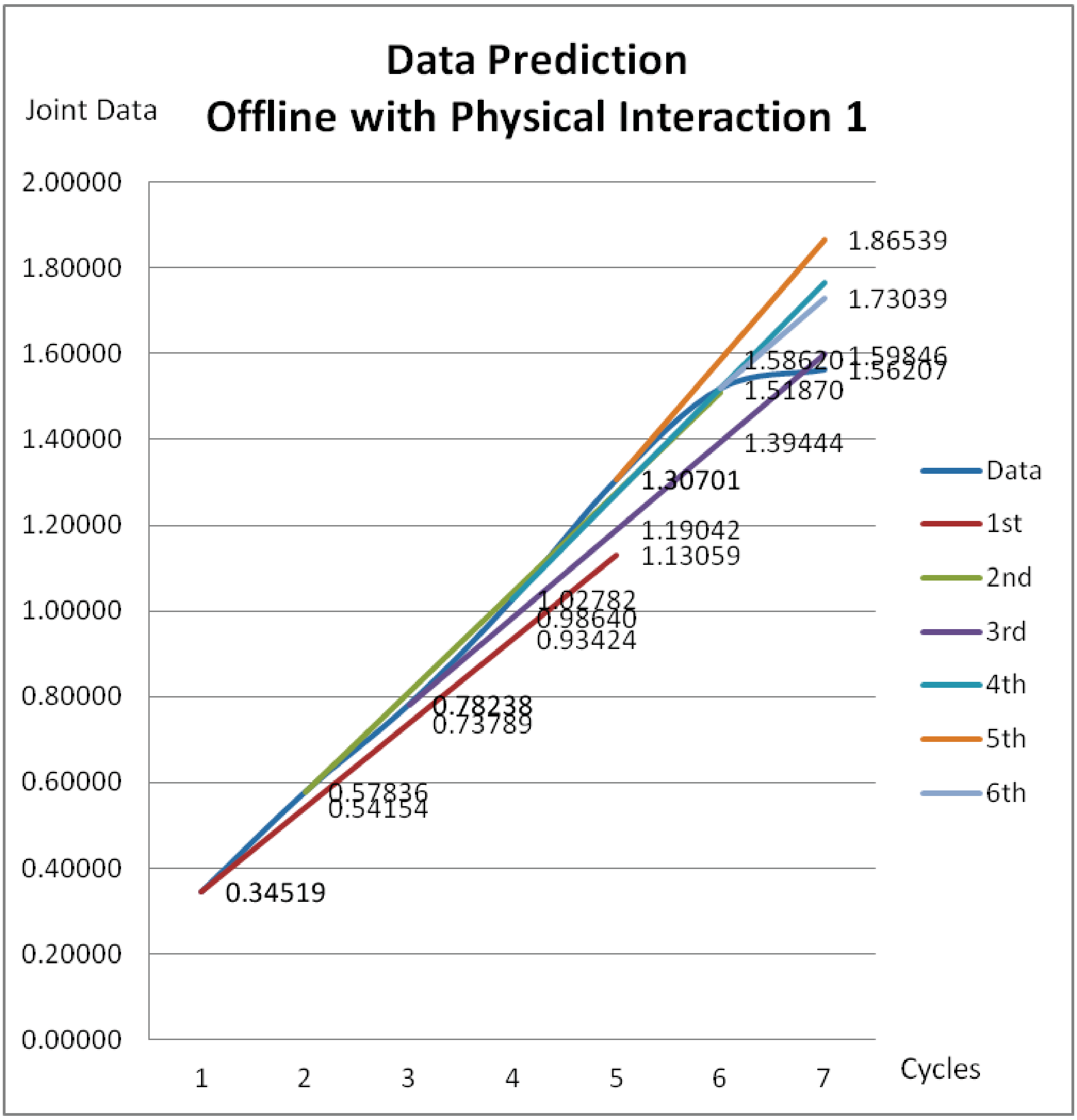

Fig. 6.6 Offline with Human Interaction Trial 1

Table 6.15 Prediction Error - Offline Physical Interaction Trial 2

Data	Prediction Cycles 6	7	8	9	10	11	Std D(σ)
6	0.00%						
7	5.68%	0.00%					0.04
8	16.57%	5.22%	0.00%				0.08
9	29.91%	12.89%	2.45%	0.00%			0.07
10	39.89%	17.18%	1.53%	3.37%	0.00%		0.03
11		28.38%	7.52%	0.15%	6.90%	0.00%	0.12
12			26.50%	16.68%	26.80%	13.00%	0.07

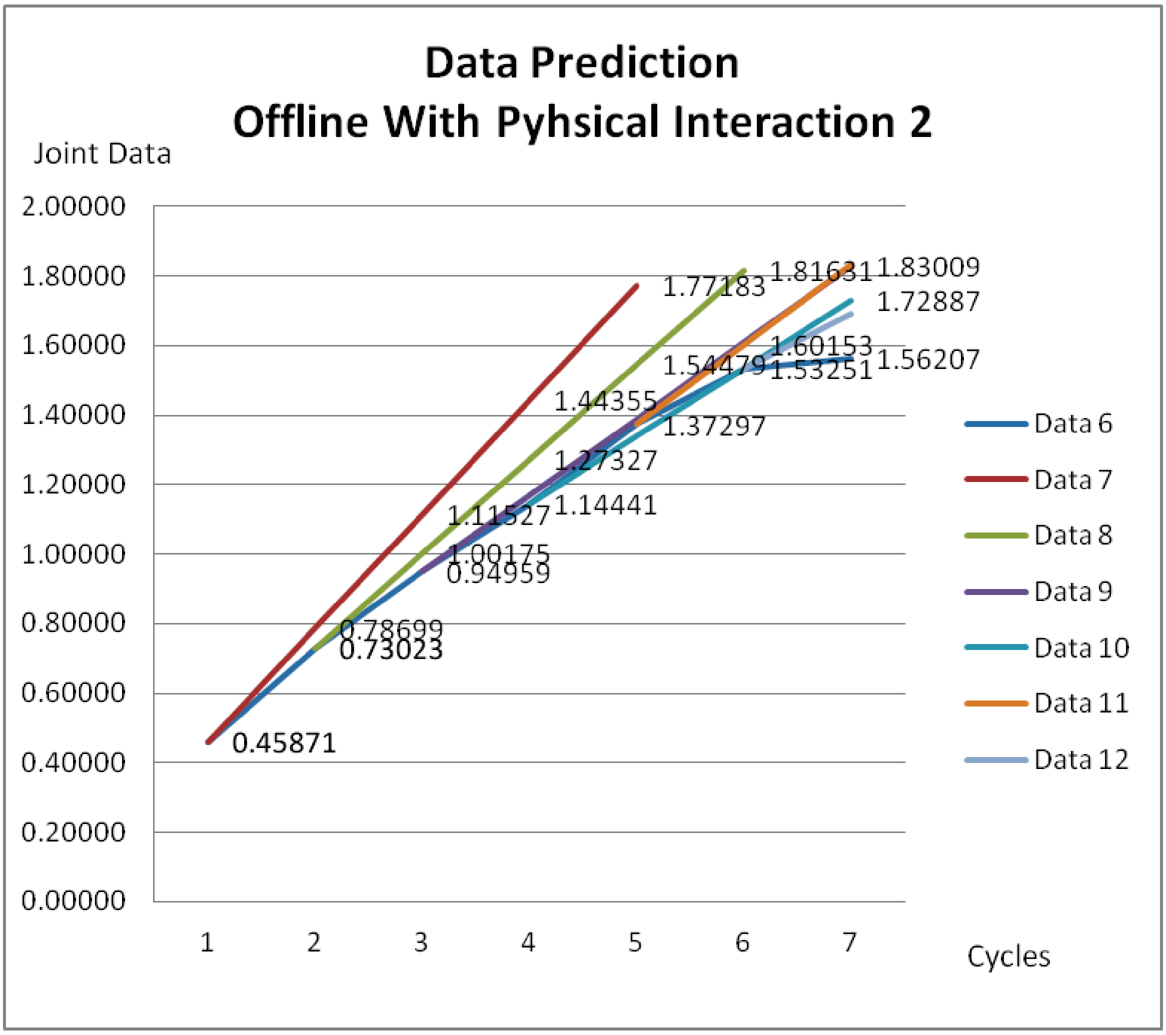

Fig. 6.7 Offline with Human Interaction Trial 2

Table 6.16 Prediction Error - Offline Physical Interaction Trial 3

Data	Prediction Cycles 5	6	7	8	9	10	Std D(σ)
5	0.00%						
6	2.61%	0.00%					0.02
7	13.35%	8.13%	0.00%				0.07
8	15.96%	8.13%	8.13%	0.00%			0.10
9	15.19%	4.75%	19.63%	3.38%	0.00%		0.11
10		1.53%	30.98%	6.60%	0.15%	0.00%	0.15
11			45.56%	13.04%	2.91%	3.22%	0.20

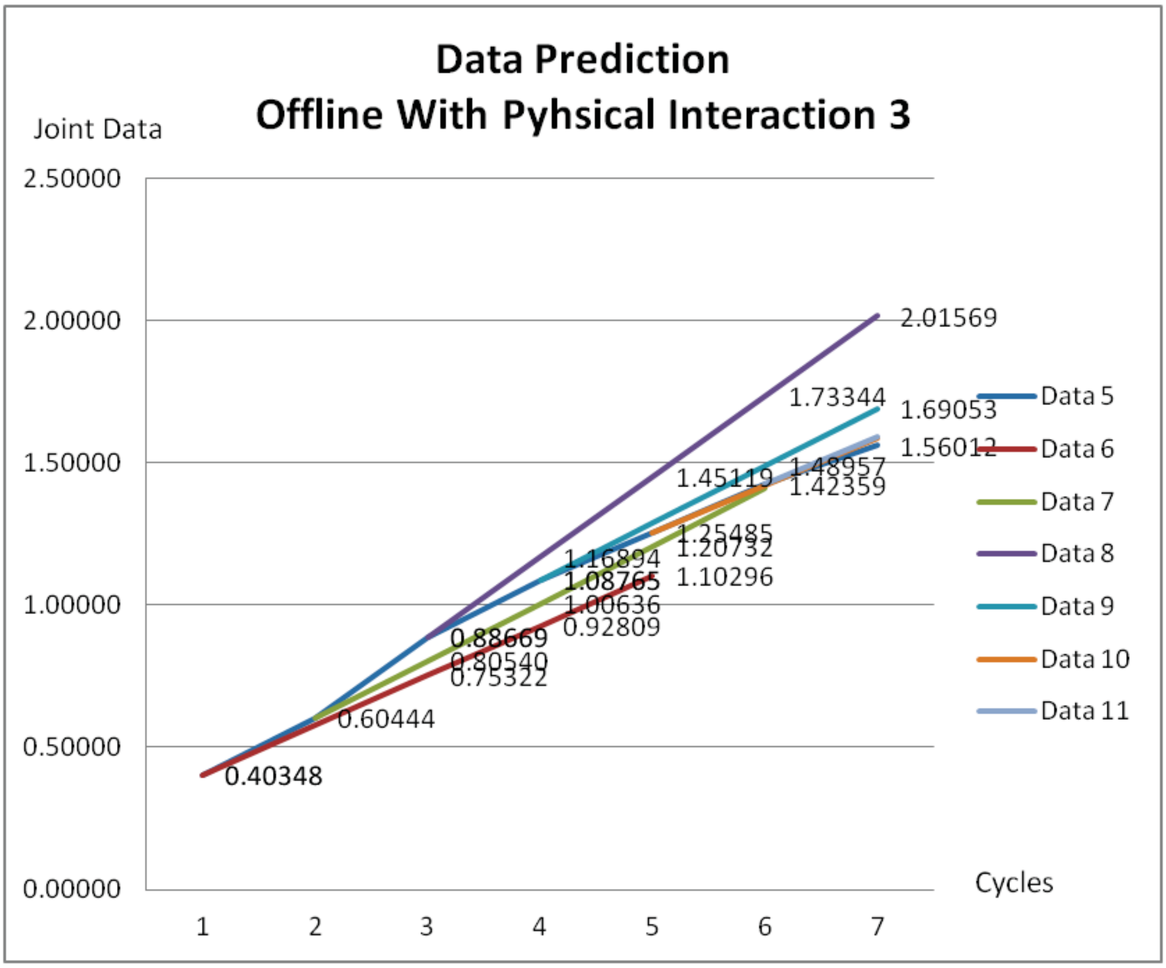

Fig. 6.8 Offline with Human Interaction Trial 3

high at 13.35%, 15.96% and 15.19%. The 6th sequence produces the first prediction error of 8.13% which drops to 1.53% in the final prediction. The highest prediction error occurs in the 7th sequence, 8.13% in the first prediction, and increases dramatically to 45.56% in the final prediction. The 8th sequences produces low prediction error early in the process, with 3.38% for the first prediction and 6.60% for the third prediction. The final prediction increases the error value to 13.04%. Prediction error values remain low, under 5% in two sequences of Joint Data: 0.15% and 2.91% for the 9th sequence and 3.22% for the 10th sequence.

In Trial 4, the prediction process is initiated at the 5th sequence with very high error values. The first prediction has a 25.47% error value which increases throughout the prediction cycles, reaching a final prediction error of 126.25%. The first prediction error value in the next sequence of Joint Data drops to 7.36%. It then increases throughout the next two prediction values: 10.43% for the second prediction and 24.39% for the third. The fourth prediction error drops dramatically to 0.19%. The first and the second prediction error remains low, under 5% during the 7th sequence. The error increases outstandingly in the final prediction, 29.25%. The last two sequences of Joint Data unfortunately suffer relatively high prediction error values: 10.89% and 16.37% in the 8th sequence; and 38.16% in the 9th sequence (see Table 6.17 and Figure 6.9). Table 6.18 and Figure 6.10 show the overall prediction

Table 6.17 Prediction Error - Offline Physical Interaction Trial 4

Data	**Prediction Cycles**						**Std D(σ)**
	5	**6**	**7**	**8**	**9**	**10**	
5	0.00%						
6	25.47%	0.00%					0.18
7	58.29%	7.36%	0.00%				0.32
8	86.83%	10.43%	4.29%	0.00%			0.43
9	126.25%	24.39%	2.30%	10.89%	0.00%		0.53
10		0.19%	29.25%	16.37%	38.16%		0.17

error incurred throughout Trial 5. It can be seen that the robot starts to decipher incoming Joint Data at the 4th sequence with the first prediction error of 2.30%. The second and third prediction error values remain below 10% and the final prediction error increases slightly to 11.04%. These prediction error values drop considerably below 3% in the 5th sequence with 1.07% in the first prediction. The second prediction error increases, doubles the first prediction error, then drops back to 1.84% in the third and fourth predictions. In the 6th sequence, no prediction error occurs in the first prediction, but there is a subsequent regular increases to 1.38%, 2.45%, and 3.53% in the final prediction. The same occurs in the 7th

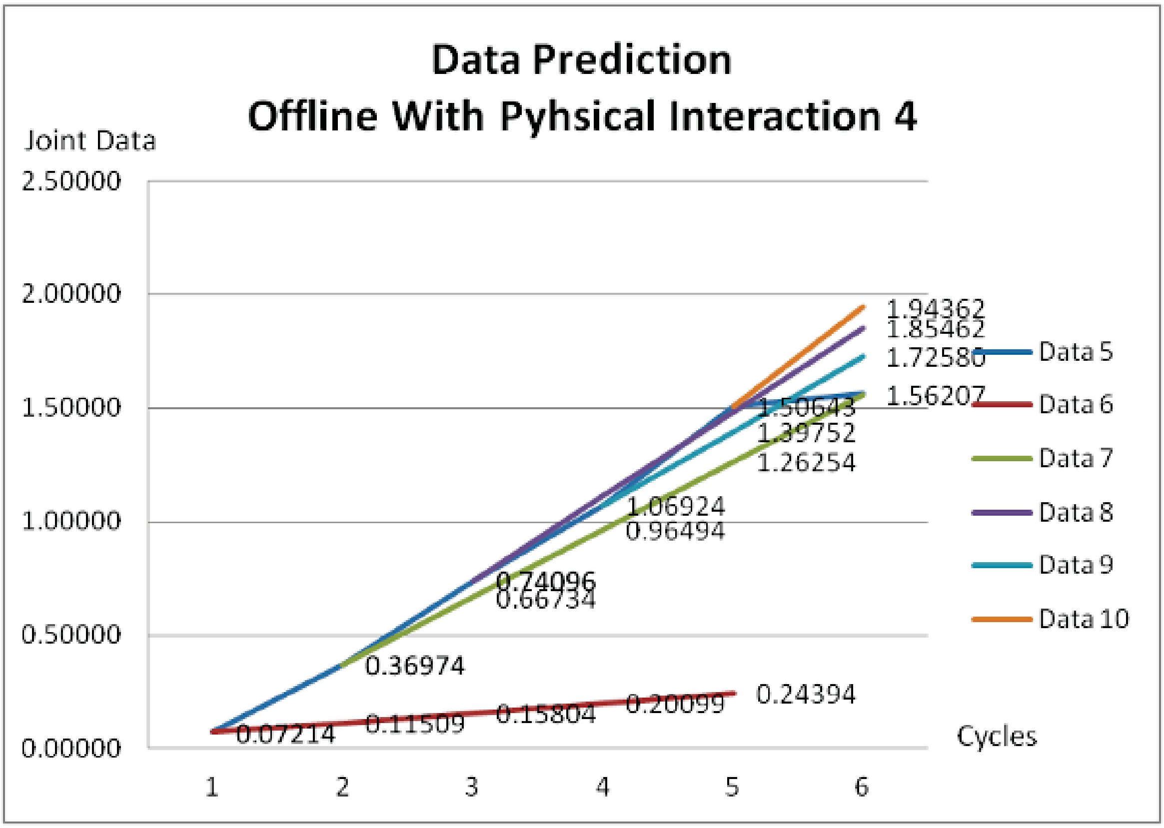

Fig. 6.9 Offline with Human Interaction Trial 4

Table 6.18 Prediction Error - Offline Physical Interaction Trial 5

Data	Prediction Cycles											Std D(σ)
	4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	
4	0.00%											
5	2.30%	0.00%										0.02
6	5.68%	1.07%	0.00%									0.03
7	9.05%	2.15%	0.00%	0.00%								0.04
8	11.04%	1.84%	1.38%	1.38%	0.00%							0.05
9		1.84%	2.45%	2.46%	0.31%	0.00%						0.02
10			3.53%	3.53%	0.61%	0.00%	0.00%					0.02
11				6.14%	0.61%	1.54%	1.54%	0.00%				0.02
12					0.92%	2.15%	2.15%	0.92%	0.00%			0.01
13						4.30%	4.30%	0.31%	1.54%	0.00%		0.02
14							7.21%	1.07%	3.84%	0.77%	0.00%	0.03
15								8.08%	11.77%	7.17%	5.63%	0.03

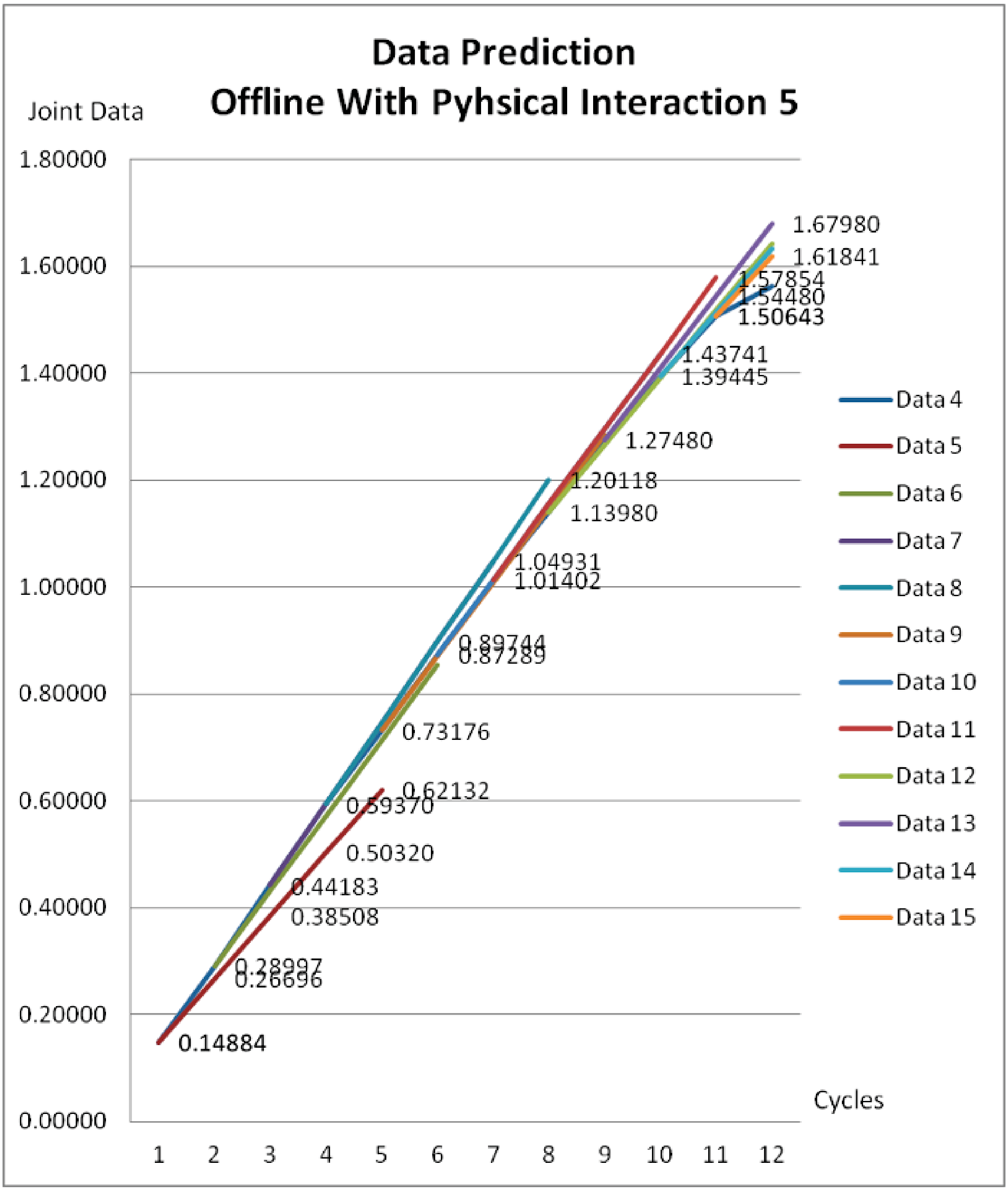

Fig. 6.10 Offline with Human Interaction Trial 5

sequence, with the exception that the final prediction error is double that of the previous sequence at 6.14%. The 8th sequence produces the lowest prediction error, and the highest error value of 0.92% in the final prediction cycle. Prediction error increases slightly in the next sequence with the highest prediction error of 4.30% in the fourth prediction cycle. This slight increase remains in the 10th sequence, staring with a first prediction error of 1.54% and ending with an error of 7.21%. In the 11th sequence, three prediction cycles produce lower prediction error values, 0.92%, 0.31% and 1.07%; the prediction error value then increases to 8.08%. Similarly, prediction error values in the 12th cycle remain low in first two cycles of prediction, 1.54% and 3.84%, then increase to 11.77%. The last two Joint Data sequences, 13th and 14th, produce prediction error values of 7.17% and 5.63% in the final prediction cycles. The final sequence, trial data 14, serves as the primary source for determining the robot's decision, that is, whether to proceed to the next sequence or to constrain robot awareness through joint stiffness, resulting in resistance in the robot's elbow. With a prediction error of 7.17%, the robot is able to deliver accurate decisions in the given situation and predict the consequences of the data in a timely manner. Overall, the 5th to final sequences of the Joint Data produce relatively low prediction error values, and the lowest prediction error on average occurs in the 8th sequence, in which the error values are less than 1%. In addition, there are more sequences of Joint Data to be processed in Trial 5 compared to those of other trials.

In contrast to the offline scenario, the experiment cycles in the online scenario experiments are much shorter. For example, in two trials (without physical interaction), the data sequence in the robot's arm resulting from the hand pushing interaction with the human peer in Trial 1 and Trial 5 takes place in a strikingly short time. The decisions made by Agent 1 were reliable with a very low standard deviation of 0.00146, where the highest prediction error was 0.31% and the deviation level of the prediction time was very low at 0.01 (see Table 6.19). In Trial 2, by contrast, the joint sensor readings were volatile even though they were

Table 6.19 Prediction Error - Online without Physical Interaction

Data	Prediction Cycles			Std D(σ)	Time
	4	5	6		Std D(σ)
4	0.00%				0.01
5	0.00%	0.00%		0.00	0.03
6	0.00%	0.00%	0.00%	0.00	0.04

supposed to remain steady for each specific data reading. The robot is capable of identifying this anomalous situation and correctly maps the behaviour of the robot's arm. It can be seen that the standard deviation for elbow joint prediction remains very low at 0.00146 by the end of the experiment. Details of the prediction data for each experiment are explained in this section.

Table 6.20 and Figure 6.11 depict the prediction error which occurs during Trial 1 of the online experiment without physical interaction. It can be inferred that the robot deciphers

Table 6.20 Prediction Error - Online without Physical Interaction Trial 1

Data	**Prediction Cycles**			**Std D(σ)**
	4	**5**	**6**	
4	0.00%			
5	0.00%	0.00%		0.00
6	0.00%	0.00%	0.00%	0.00

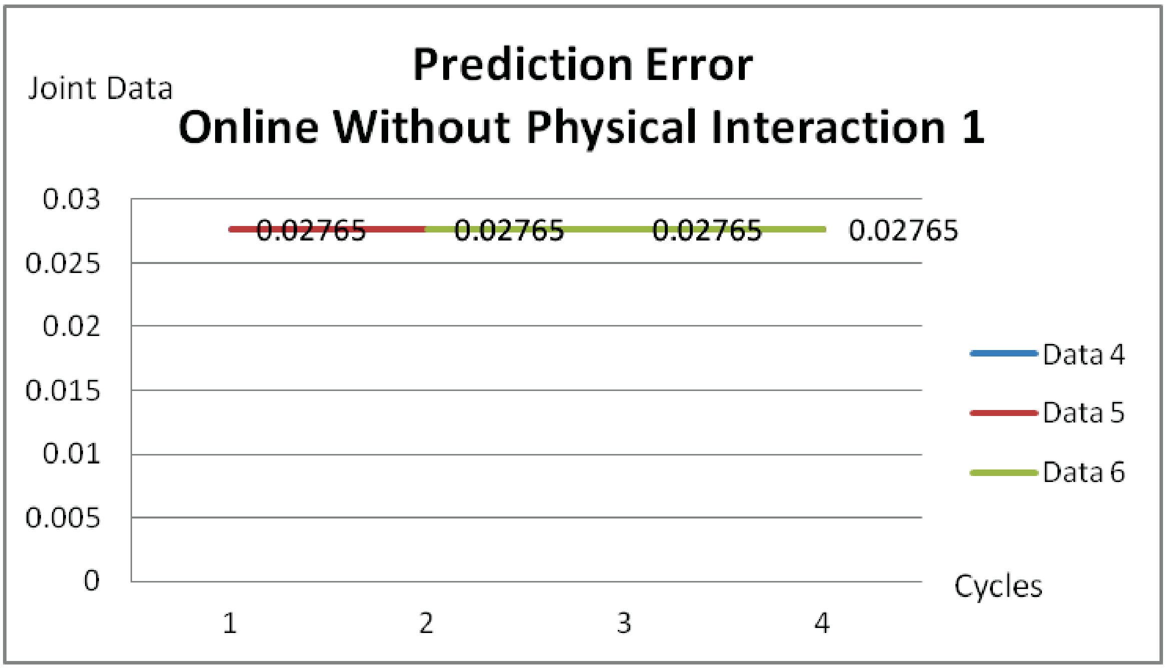

Fig. 6.11 Online without Human Interaction Trial 1

the incoming Joint Data at the 4th sequence and the prediction data accurately converges to the Joint Data. Trial 1 lasts for six Joint Data sequences, and overall, prediction error values remain at 0%. Similarly, the prediction error remains zero and the prediction process takes place in a single prediction sequence, which is in the 5th sequence. All the predictions cycles converge to an accurate Joint Data prediction, 0.02765 without generating any deviation value. Data analyses of the prediction data during Trial 2 of the online experiment without physical interaction are shown in Table 6.21 and the results are illustrated in Figure 6.12. In Trial 3,

Table 6.21 Prediction Error - Online without Physical Interaction Trial 2

Data	**Prediction Cycles** **5**	**6**	**Std D(σ)**
5	0.00%		
6	0.00%		-
7			

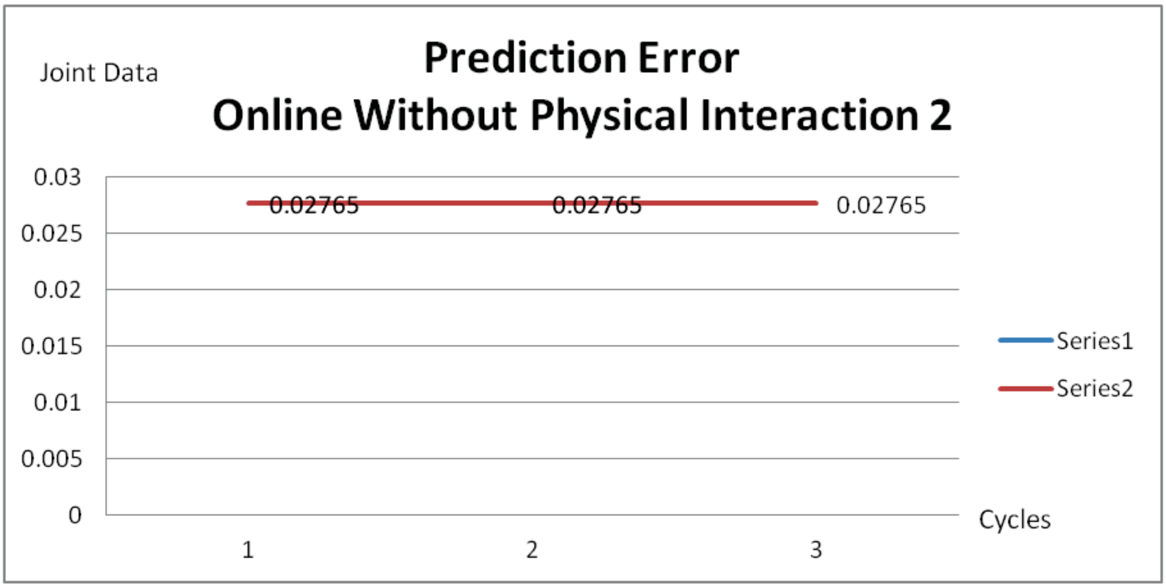

Fig. 6.12 Online without Human Interaction Trial 2

the process lasts longer than the previous trial and the robot deciphers the incoming data in the 4th Joint Data sequence. Prediction error strikingly occurs during this trial at 0.15% for the second and fourth predictions. During the 5th sequence, prediction error of 0.31% occurs in the third prediction. The same prediction error, 0.31%, occurs in the second prediction of the 6th sequence. It can be seen that no changes in the Joint Data could be detected during the non-interaction experiment, and as a result this noisy data could deteriorate the accuracy of the prediction. However, the prediction error remains low throughout the trial (see Table 6.22 and Figure 6.13). Table 6.23 and Figure 6.14 show prediction error of the online experiment

Table 6.22 Prediction Error - Online without Physical Interaction Trial 3

Data	Prediction Cycles 4	5	6	7	Std D(σ)
4	0.00%				
5	0.15%	0.00%			0.00108
6	0.00%	0.00%	0.00%		0.00000
7	0.15%	0.31%	0.31%	0.00%	0.00146
8					0.00177

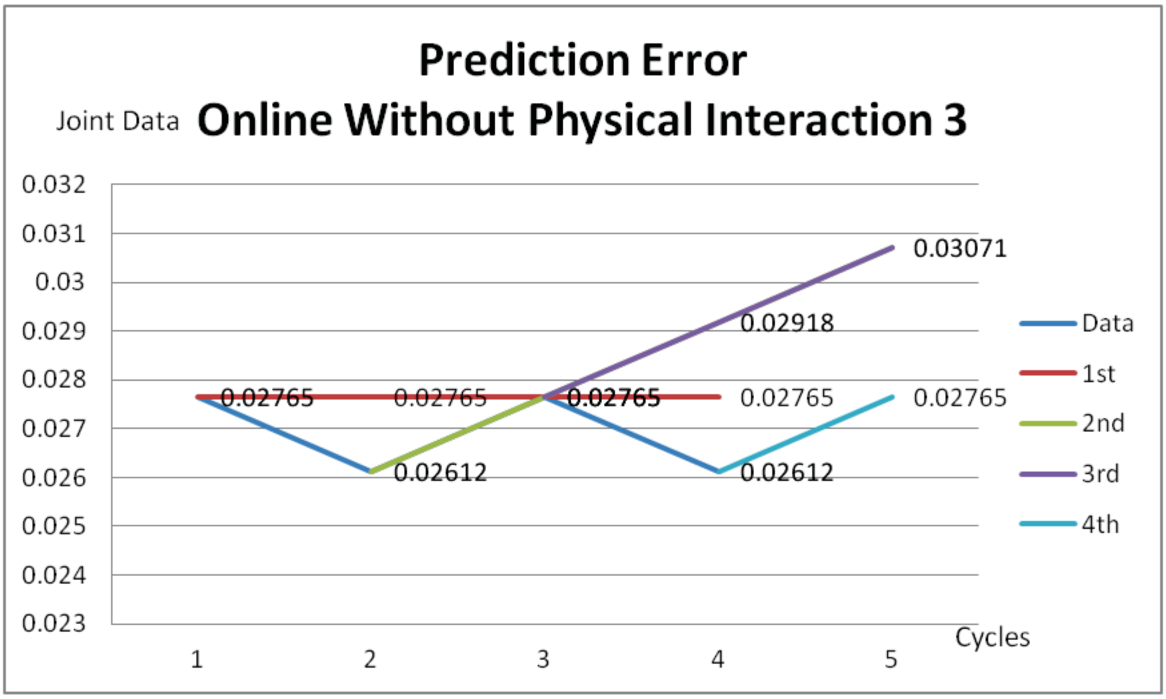

Fig. 6.13 Online without Human Interaction Trial 3

without physical interaction, Trial 4. It can be seen from the table that the robot commences

Table 6.23 Prediction Error - Online without Physical Interaction Trial 4

Data	**Prediction Cycles** 1	2	3	**Std D(σ)**
4	0.00%			
5	0.00%	0.00%		0.00
6	0.00%	0.00%	0.00%	0.00
7				

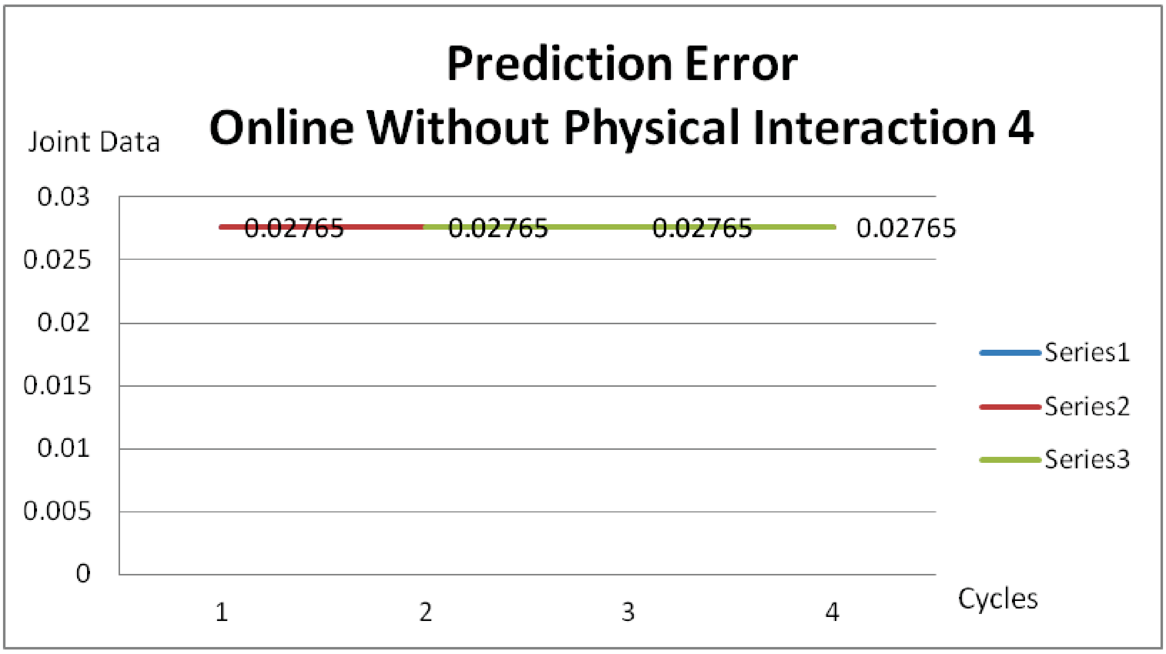

Fig. 6.14 Online without Human Interaction Trial 4

its prediction at the 4th sequence and all Joint Data prediction results converge at 0.02765. Prediction cycles occur until the 6th sequence with an overall prediction error of 0%. The figure depicts the situation during Trial 4 in a straight line along the Cycles axis.

In Trial 5 prediction process occurs in three cycles, as shown in Table 6.24 and Figure 6.15. The first cycle takes place in the 5th sequence, followed by the second in the 6th sequence, and the final cycle occurs in the 7th Joint Data sequence. All Joint Data prediction results converge at 0.02612. This pattern is similar to that in Trial 4 and is illustrated in the figure by a straight line.

Table 6.24 Prediction Error - Online without Physical Interaction Trial 5

Data	Prediction Cycles 1	2	3	Std D(σ)
5	0.00%			
6	0.00%	0.00%		0.00
7	0.00%	0.00%	0.00%	0.00

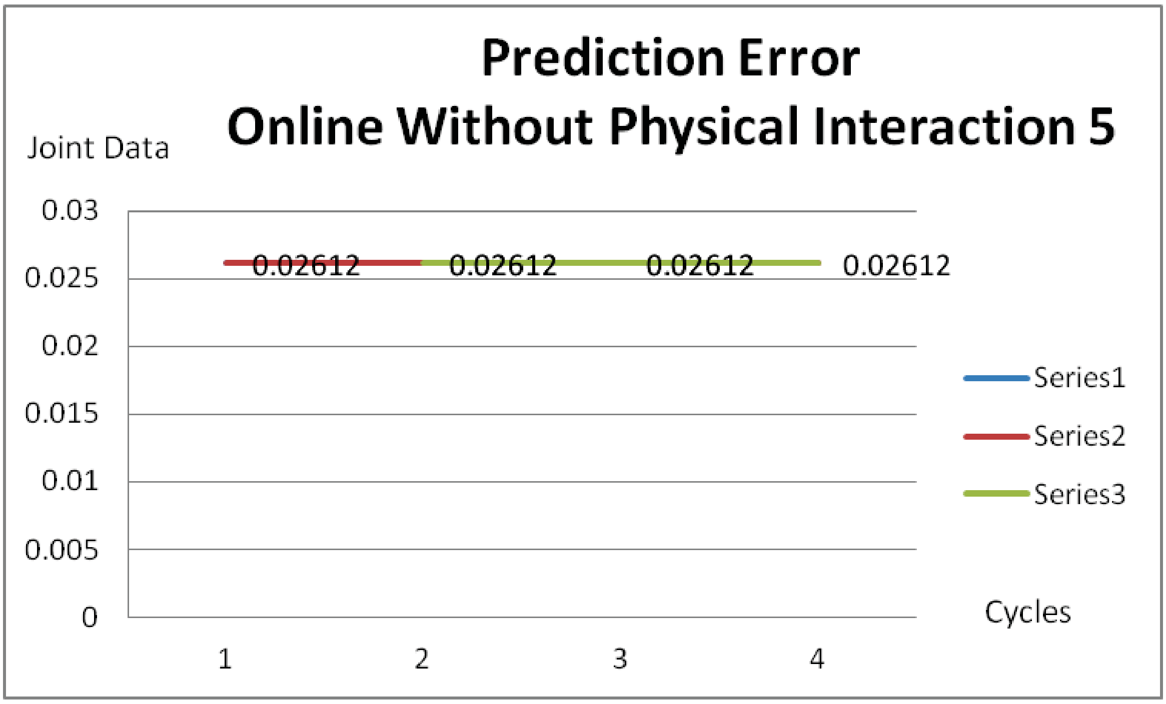

Fig. 6.15 Online without Human Interaction Trial 5

Discussions on the prediction error for the online experiment with physical interaction covering Trial 1 to Trial 5 will be presented below.

Table 6.25 and Figure 6.16 show the results of the prediction error during the online experiment with physical interaction Trial 1. The prediction process commences in the 4th

Table 6.25 Prediction Error - Online with Physical Interaction Trial 1

Data	**Prediction Cycles**						**Std D(σ)**
	1	**2**	**3**	**4**	**5**	**6**	
4	0.00%						
5	1.84%	0.00%					0.01
6	5.37%	1.69%	0.00%				0.03
7	9.97%	4.45%	1.07%	0.00%			0.04
8	34.97%	27.61%	22.55%	20.40%	0.00%		0.13
9							

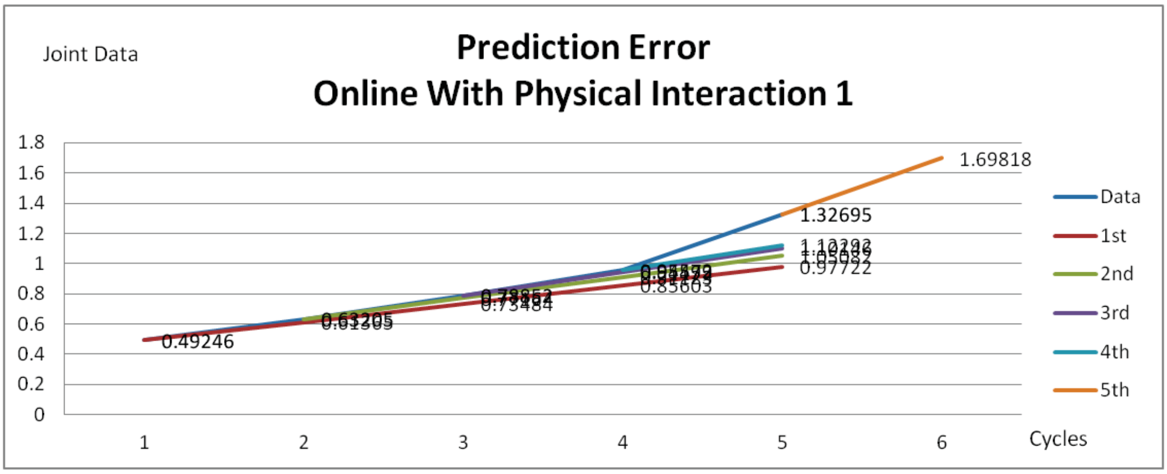

Fig. 6.16 Online with Human Interaction Trial 1

Joint Data sequence with a low prediction error in the first prediction of 1.84% (the data prediction is 0.61365) which increases to 5.37% (data prediction is 0.73484) in the second prediction. The prediction error increases to 9.97% (data prediction is 0.85603) and the final prediction produces a relatively high prediction error at 34.97% (the data prediction is 0.97722 while the real Joint Data at 8th sequence is 1.32695). In the 5th sequence, prediction error values are lower than those in the previous sequence. The first cycle produces 1.69% prediction error (0.77164) and increases to 4.45% (0.91123). The final prediction cycle still produces a high prediction error of 27.61% (1.05802) is slightly lower than the previous sequence. The prediction error during the 6th Joint Data sequence decreases with the final prediction error to 22.55% (the data prediction for the 8th sequence is 1.10146). In the 7th sequence, the prediction error remains slightly high at 20.40% (the joint data prediction is

1.12292 for the next sequence). As this prediction is still below the restricted joint value, the robot continues to process incoming data. In the 8th sequence, the first cycle predicts that the 9th Joint Data sequence will be 1.69818. This suggests that the robot should halt the process and increase the arm joint stiffness to prevent the possibility of the robot experiencing pain. From the figure, it can be seen that the deviation values of the prediction data compared to the real Joint Data throughout the trial are relatively low. The prediction process still allows the robot to interact with the human partner while accurately generating counter actions, as the joint prediction suggests the possibility of a pain experience to the robot.

Trial 2 lasts for seven Joint Data sequences, and the robot starts to decipher incoming data at the 5th sequence. The first and second prediction error values are relatively high, 10.98% and 55.22% respectively. This means that by prediction, the Joint Data at the 6th sequence should be 0.57530 and 0.75325 at the 7th sequence, whereas in fact, the Joint Data at the 6th and 7th sequences are 0.68421 and 1.030548 respectively. The next sequence produces a lesser prediction error 33.44% with a Joint Data prediction value of 0.97107. The process repeats as the predicted Joint Data remains outside the faulty joint region, and at the 7th Joint Data sequence, the robot predicts that the Joint Data will be 1.92675, which is well into the faulty joint region. Hence, the robot stops the interaction and sends out counter actions by alerting the human partner and increasing the arm joint stiffness value. Table 6.26 shows the overall prediction error throughout the trial along with the standard deviation values for each data prediction. Figure 6.17 illustrates each data prediction compared to the real Joint Data

Table 6.26 Prediction Error - Online with Physical Interaction Trial 2

Data	**Prediction Cycles**			**Std D(σ)**
	1	**2**	**3**	
5	0.00%			
6	10.89%	0.00%		0.08
7	55.22%	33.44%	0.00%	0.28
8				

originating from the robot's arm joint sensor.

Table 6.27 and Figure 6.18 show prediction error of the online with physical interaction Trial 3 which lasts for 10 sequences.// The robot initiates its data prediction at the 6th Joint Data sequence with 3.99% prediction error of 0.75631, which increases gradually in the next two prediction cycles: 10.50% for the 7th sequence prediction and 16.87% for the 8th sequence prediction. In the final prediction cycle constitutes the 10th sequence, the prediction error drops to 3.07% (the Joint Data is 1.44202). When the 7th Joint Data sequence arrives, the robot predicts the next three Joint Data sequences. The first prediction

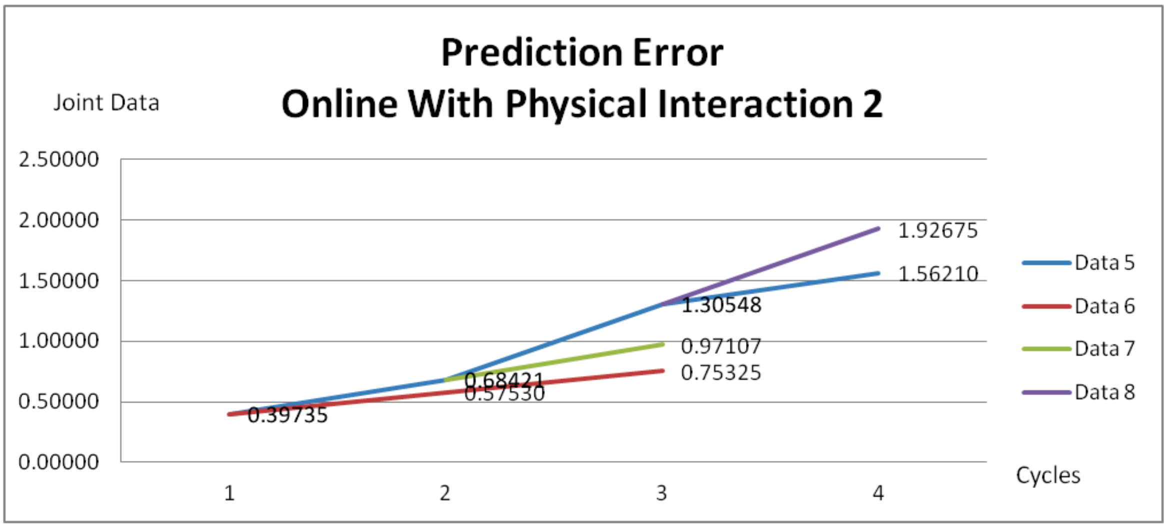

Fig. 6.17 Online with Human Interaction Trial 2

Table 6.27 Prediction Error - Online with Physical Interaction Trial 3

Data	**Prediction Cycles**						**Std D(σ)**
	1	**2**	**3**	**4**	**5**	**6**	
6	0.00%						
7	3.99%	0.00%					0.03
8	10.59%	2.61%	0.00%				0.06
9	16.87%	4.91%	0.31%	0.00%			0.08
10	3.07%	12.89%	20.71%	20.09%	0.00%		0.11
11							

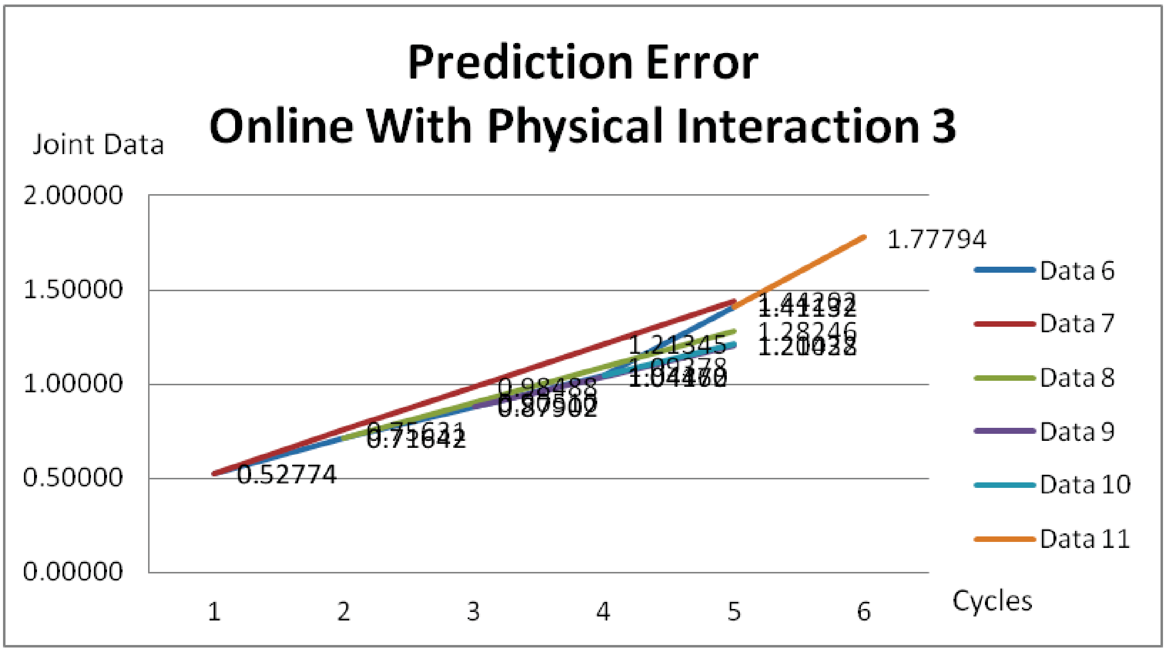

Fig. 6.18 Online with Human Interaction Trial 3

cycle constitutes the 8th sequence with a relatively low prediction error of 2.62% (0.90510) and the second prediction cycle constitutes the 9th sequence and produces prediction error of 4.91% (1.09378). Prediction error increases for the final prediction cycle which relates to the 10th Joint Data sequence at 12.89% (1.28246). During the next data sequence, first cycle of prediction error is relatively very low, 0.31% but it increases dramatically to 20.71% in the next sequence. In the 9th Joint Data sequence, the prediction error is considerably higher at 20.09% and the robot halts the process in the 10th Joint Data sequence. The robot decides to stop the process as a result of the prediction data, which exceeds the faulty joint values with a value of 1.77794. Overall, the prediction data of the 7th sequence produces the lowest prediction error with a standard deviation of 0.03 while the highest prediction error occurs in the 10th Joint Data sequence and has a standard deviation of 0.11. It can be seen from Figure 6.18 that the data moves in a vertical direction, which means that the robot's hand moves towards its chest. Stopping the movement before the hand collides with the robot's own physical limb is therefore of primary importance to prevent possible hardware damage.

Trial 4, as depicted in Table 6.28, lasts for seven Joint Data sequences. The prediction process commences in the 4th Joint Data sequence with prediction error for the next sequence data is 4.14% (0.42803). The second prediction, which constitutes the 6th sequence, produces a prediction error of 6.44%. The final prediction for the 7th sequence is 0.826878 (40.19%) which is far lower than the exact value of the Joint Data in the same sequence. When the 5th Joint Data sequence arrives, the prediction error drops to 1.84% for the predicted Joint

Data on the 6th sequence. However, the prediction error for the 7th sequence increases dramatically to 27.77%. During the 6th data sequence, the prediction value for the 7th Joint Data sequence is relatively high at 31.45%, but is lower than the exact value of the Joint Data in the 7th sequence. However, the robot manages to decipher the incoming data at the 7th sequence and predicts that the future sequence will be higher than the limit of the joint faulty region, 1.76568. This prediction result recommends that the robot should halt the interaction and that the joint stiffness of the robot's arm should be activated. Figure 6.19 shows that the majority of prediction processes generate lower Joint Data predictions. However, in the 7th Joint Data sequence, the prediction process recovers and suggests that in the future sequence, the Joint Data will exceed the faulty joint limit.

Table 6.28 Prediction Error - Online with Physical Interaction Trial 4

Data	Prediction Cycles				Std D(σ)
	1	2	3	4	
4	0.00%				
5	4.14%	0.00%			0.03
6	6.44%	1.84%	0.00%		0.04
7	40.19%	27.77%	31.45%	0.00	0.17
8					

The prediction data for the final trial of the online experiment with physical interaction is shown in Table 6.29 and the graphical analysis is depicted in Figure 6.20.

It can be seen that in Trial 5, the robot commences data prediction in the 4th Joint Data

Table 6.29 Prediction Error - Online with Physical Interaction Trial 5

Data	Prediction Cycles				Std D(σ)
	1	2	3	4	
4	0.00%				
5	2.92%	0.00%			0.02
6	24.24%	30.07%	0.00%		0.16
7					

sequence. The first prediction constitute the 5th sequence with a prediction error of 2.92% then increases in the 6th sequence to 24.24%. In the 5th sequence, the first prediction produces the highest prediction error, 30.07%. The final sequence, with the real Joint Data figure of 1.25639, generates prediction data which falls into the faulty joint region. As a result, the robot halts the process and commences the routine to increase joint stiffness in the

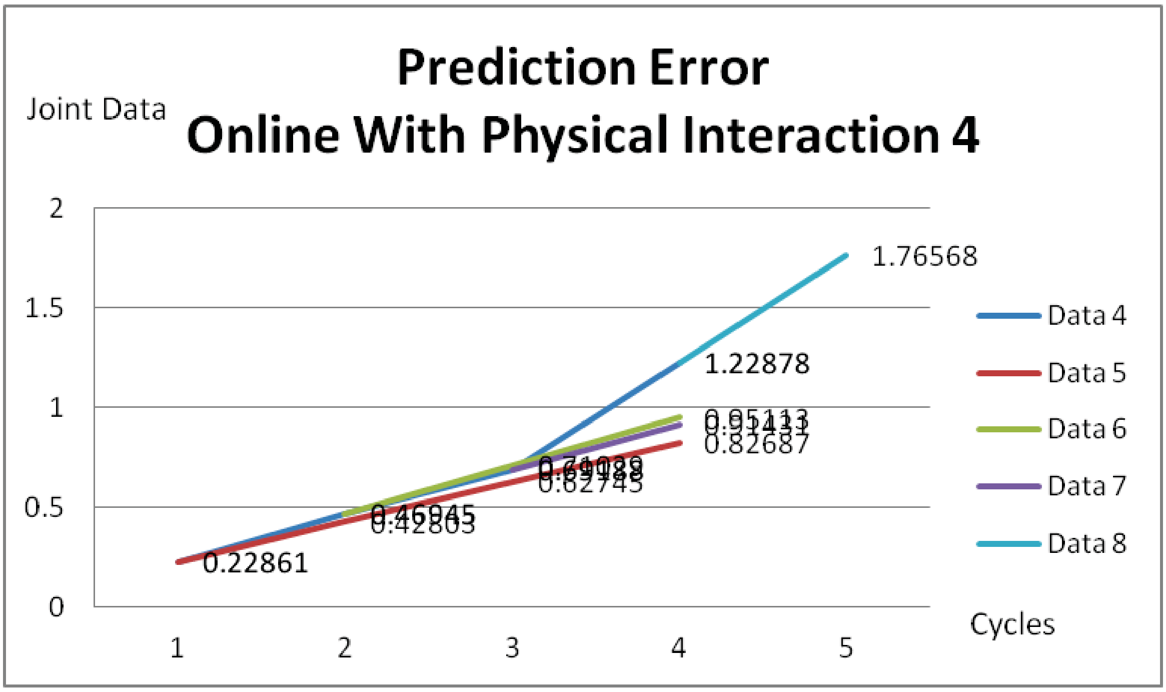

Fig. 6.19 Online with Human Interaction Trial 4

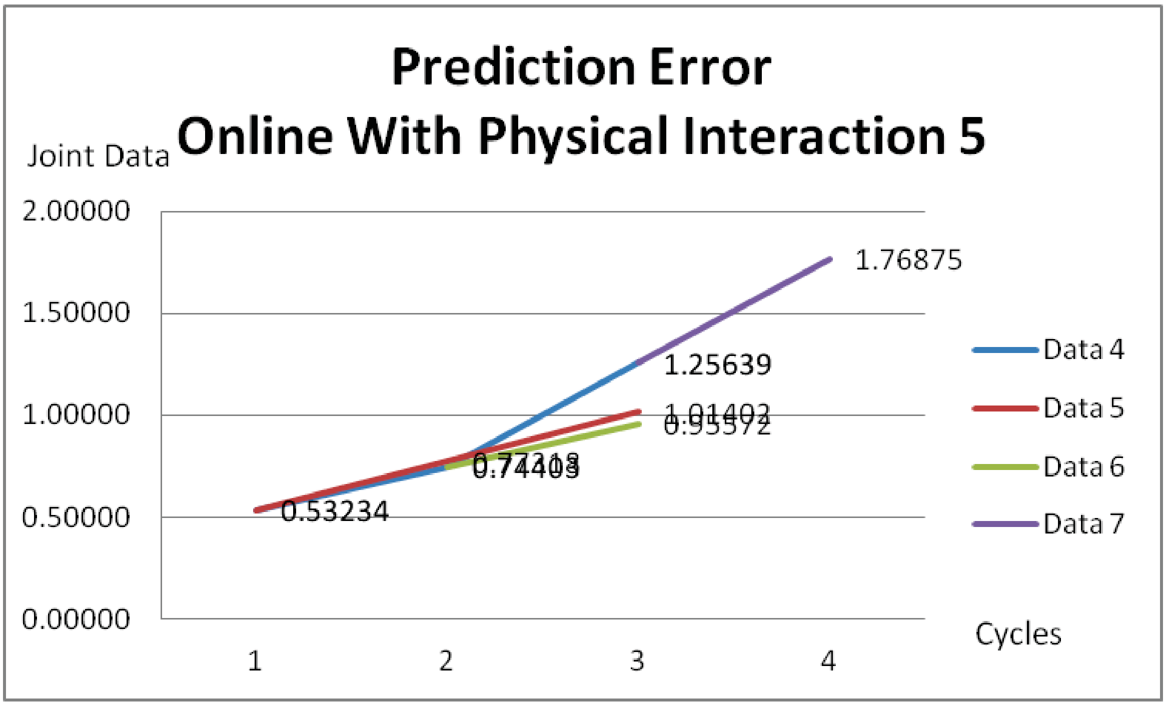

Fig. 6.20 Online with Human Interaction Trial 5

robot's arm. The figure clearly shows that the Joint Data prediction inclines gradually until the 6th data sequence. It then increases dramatically in the 7th sequence with the predicted Joint Data at 1.76875, which forces the robot to stop.

It can be seen that the reasoning behaviour of the robot for the experiments in the online scenario with physical interaction has a similar pattern of error predictions, and the Robot Mind manages to accurately predict the elbow data from the pattern of the robot's arm movement. For instance, in Trial 1, the prediction error made in the first cycle is relatively high at 34.97%. The standard deviation prediction for the eight sequences of data reading, however, is still relatively small at 0.13, and at the same time, prediction reaches its highest standard deviation value at 1.33. Overall, the agent is still capable of identifying future consequences and preventing the agent from experiencing synthetic pain.

Awareness during the early stage is freely explored, but the attention to awareness is subject to change at the Robot Mind's recommendation. The various states of robot awareness during the experiment are shown in Table 6.30. The offline experiments equate to Internal States 1 to 10, and the online experiments equate to Internal States 11 to 20. When no interaction is involved, the robot's early awareness type remains the same throughout the experiments (Internal State 1 to Internal State 5 during the offline experiment and Internal State 11 to Internal State 15 during the online experiment). In contrast, when any physical interaction is involved, the Robot Mind alters the awareness type, resulting in the final awareness being High Priority Subjective Awareness in constrained type (Internal State 6 to Internal States 10 during the offline experiment and Internal State 16 to Internal State 20 during the online experiment).

The robot's internal states following the reasoning process are shown in Table 6.31. It can be seen that the Robot Mind generates the PP, which according to our definition functions as an alert signal. Apart from sending a voice alert to its human partner, the robot also takes preventive action by increasing its elbow joint stiffness to such a degree that the human partner can detect resistance. The robot is still capable of identifying future consequences, thus preventing it from entering the faulty joint region. During Internal States 1 to 5, the real Joint Data and prediction data converge, inferring that the proprioceptive sensor of the robot is reliable which reads the joint arm position at 0.00873. The reasoning process can accurately reasons this Joint Data and produces a uniform Joint Data prediction of 0.00873. As no movement is detected on the robot arm, no change occurs in the awareness type; hence no pain is generated throughout these states. This situation should also occur during Internal States 11 to 15, but data discrepancies are obtained from the proprioceptive sensor. For instance, in Internal State 13, the real Joint Data being processed is 0.02612 and prediction produces 0.02765, which is higher than the real data. However, prediction Internal State 14

Table 6.30 State of Awareness

Feeding Data	Experiment	Data	Early Awareness Type	Final Awareness Type	CDV Value	Awareness Region	Internal States
Offline	No Physical Interaction	1	High Priority Objective	Constrained High Priority Subjective	131	6	1
		2	Low Priority Subjective		33	2	2
		3	High Priority Subjective		17	1	3
		4	Left Border Subjective-Objective		55	3	4
		5	Right Border Subjective-Objective		93	4	5
	Pushing Arm	1	Right Border Subjective-Objective		80	4	6
		2	Low Priority Objective		110	5	7
		3	High Priority Subjective		3	1	8
		4	Low Priority Subjective		38	2	9
		5	Low Priority Objective		105	5	10
On Robot	No Physical Interaction	1	High Priority Subjective	Constrained High Priority Subjective	14	1	11
		2	Left Border Subjective-Objective		62	3	12
		3	Low Priority Objective		116	5	13
		4	High Priority Objective		144	6	14
		5	Low Priority Objective		111	5	15
	Pushing Arm	1	High Priority Subjective		6	1	16
		2	High Priority Objective		126	6	17
		3	High Priority Objective		138	6	18
		4	Left Border Subjective-Objective		65	3	19
		5	High Priority Objective		109	5	20

Table 6.31 Internal States after Reasoning Process

After Reasoning					
Internal State	**Joint Data Real Data**	**Final Prediction**	**Awareness Type**	**Synthetic Pain Categories**	**Intensity**
1	0.00873	0.00873			
2	0.00873	0.00873			
3	0.00873	0.00873	Unconstrained	No Pain	-
4	0.00873	0.00873			
5	0.00873	0.00873			
6	1.51870	1.73039			Slight
7	1.53251	1.69205			Slight
8	1.42359	1.59233	Constrained	1:0 Proprioceptive (PP)	Slight
9	1.50643	1.94362			Slight
10	1.50643	1.61841			Slight
11	0.02765	0.02765			
12	0.02765	0.02765			
13	0.02612	0.02765	Un constrained	No Pain	-
14	0.02765	0.02765			
15	0.02612	0.02612			
16	1.32695	1.69818			Slight
17	1.30548	1.92675			
18	1.41132	1.77794	Constrained	1:0 Proprioceptive (PP)	
19	1.22878	1.76568			Slight
20	1.25639	1.76875			

achieves 100% accuracy, 0.02765. Similarly, accurate prediction occurs for Internal State 15, with predicted Joint Data 0.02765 while the real Joint Data is 0.02612. It can be inferred that the Joint Data should remain the same throughout the experiment as a result of the non-physical interaction. Regardless of the noisy data occurring in Internal State 13 and Internal State 15, the robot awareness still accurately converges to the unconstrained type without pain invocation. The awareness is constrained only during the Internal States 6 to 10 (offline experiment) and Internal States 16 to 20 (online experiment). During Internal State 6, the prediction data suggests that the next incoming data will cause the joint to move into the faulty joint region, and as a result, the mind constrains the awareness type and invokes the PP type whose Intensity is *Slight*. This intensity is measured from the distance of the Joint Data to the faulty joint reference and in this case, it is a slightly further from the limit value of 1.5621. A similar situation occurs in other Internal States, such as Internal States 7, 8 and 10. For Internal State 9, the distance prediction for the faulty joint limit is the furthest, however, our approach only defines the intensity at two levels: *None* and *Slight*. We will include the possible development of this approach in our future work to increase the robustness of our synthetic pain definition. A similar situation also appears in Internal States 16 to 20. In all Internal States, the Joint Data prediction values fall into the faulty joint region, resulting in the Robot Mind invoking the PP type with *Slight* intensity. Internal State 17 produces the highest Joint Data prediction, 1.92675, and it evokes the PP type with the Intensity level *Slight*.

In the early stages, the Robot Mind is at an unconstrained state which indicates that no synthetic pain has been generated. The awareness is forced to switch to a High Priority Subjective type once the reasoning process predicts that the arm joint will move into the faulty joint region. In addition, the High Priority Subjective awareness type may be revisited at a constrained state without generating any synthetic pain (Intensity is *None*).

SPD Synthetic Pain Activation

The Joint Data and prediction data Trial 1 for the SPD-based model can be seen in Table 6.32. This table shows that the Robot Mind deciphers incoming information at sequence 4, when the real Joint Data is 0.49246. First prediction constitutes the 5th sequence, 0.61365; second prediction constitutes the 6th sequence, 0,73484; the third prediction for the 7th sequence is 0.85603 and the fourth prediction for the 8th sequence is 0.97722. The first prediction produces a relatively low prediction error of 1.84%, which then increases gradually to 5.37% for the second prediction and 9.97% for the third prediction. The final prediction generates a considerably high prediction error of 34.97%. However, the prediction error in this sequence remains low overall with a standard deviation of 0.01. (see Table 6.33 for prediction error

Table 6.32 Joint Data and Prediction Data SPD-based Model Trial 1

No	Data	Prediction Data 4	5	6	7	8
1	0.22247					
2	0.26696					
3	0.37127					
4	0.49246					
5	0.63205	0.61365				
6	0.78852	0.73484	0.77164			
7	0.95572	0.85603	0.91123	0.94499		
8	1.32695	0.97722	1.05082	1.10146	1.12292	
			1.19041	1.25793	1.29012	1.69818
				1.41440	1.45732	2.06941
					1.62452	2.44064
						2.81187

during the SPD-based model Trial 1 experiment). When the robot processes the incoming data in the 5th sequence, the prediction error of the first cycle is 1.69% and of the second one is 4.45% (Joint Data predictions being 0.77164 and 0.91123 respectively). The final prediction error is high at 27.61% (Joint Data prediction is 1.05082). A similar pattern occurs in the next two sequences, 6 and 7, with prediction error values for the final cycles for each sequence being 22.55% and 20.40% respectively. In the 8th sequence, the first prediction suggests that the 9th sequence obtains the Joint Data of 1.69818 which recommends that the robot should halt the experiment for Trial 1 as it will exceed the faulty joint value limit. Figure 6.21 depicts the prediction data pattern of the SPD-based model Trial 1 experiment. The Joint Data prediction inclines vertically, with the final prediction data during 8th sequence forcing the robot to halt.

Table 6.34, Table 6.35 and Table 6.36 show the initial state of the Robot Mind, the pain activation and the Robot Mind recommendation during Trial 1. In all situations, the robot's consciousness state is unconstrained and the awareness level is High Subjective, which is located in Region 1 (CDV is 6).

It can be seen from the tables that no SPD recommendation passes to the Robot Mind during the 1st to 4th sequences. As a result, the Mind State remains unconstrained and no action is taken by the Engine-Intention. A similar situation occurs in the next two sequences and the SPD start suggesting pain activation in 6th sequence. In this sequence, the prediction for the 10th sequence (four interval) suggests the kind of PP to be activated and the Robot Mind sends out a warning about future pain. However, a counter-physical action is sent to

Table 6.33 Prediction Error SPD-based Model Trial 1

Data	Prediction Cycles 4	5	6	7	8	Std D(σ)
4	0.00%					
5	1.84%	0.00%				0.01
6	5.37%	1.69%	0.00%			0.03
7	9.97%	4.45%	1.07%	0.00%		0.04
8	34.97%	27.61%	22.55%	20.40%	0.00%	0.06
9						

Table 6.34 SPD Initial State Trial 1

No	Data Sensory	Internally	CDV	Region	Incoming Belief	Awareness	Consciousness State
1	0.22247		6	1	Current	High Subjective	Unconstrained
2	0.26696		6	1	Current	High Subjective	Unconstrained
3	0.37127		6	1	Current	High Subjective	Unconstrained
4	0.49246		6	1	Current	High Subjective	Unconstrained
		0.61365	6	1	Prediction	High Subjective	Unconstrained
		0.73484	6	1	Prediction	High Subjective	Unconstrained
		0.85603	6	1	Prediction	High Subjective	Unconstrained
		0.97722	6	1	Prediction	High Subjective	Unconstrained
5	0.63205		6	1	Current	High Subjective	Unconstrained
		0.77164	6	1	Prediction	High Subjective	Unconstrained
		0.91123	6	1	Prediction	High Subjective	Unconstrained
		1.05082	6	1	Prediction	High Subjective	Unconstrained
		1.19041	6	1	Prediction	High Subjective	Unconstrained
6	0.78852		6	1	Current	High Subjective	Unconstrained
		0.94499	6	1	Prediction	High Subjective	Unconstrained
		1.10146	6	1	Prediction	High Subjective	Unconstrained
		1.25793	6	1	Prediction	High Subjective	Unconstrained
		1.41440	6	1	Prediction	High Subjective	Unconstrained
7	0.95572		6	1	Current	High Subjective	Unconstrained
		1.12292	6	1	Prediction	High Subjective	Unconstrained
		1.29012	6	1	Prediction	High Subjective	Unconstrained
		1.45732	6	1	Prediction	High Subjective	Unconstrained
		1.62452	6	1	Prediction	High Subjective	Unconstrained
8	1.32695		6	1	Current	High Subjective	Unconstrained
		1.69818	6	1	Prediction	High Subjective	Unconstrained
		2.06941	6	1	Prediction	High Subjective	Unconstrained
		2.44064	6	1	Prediction	High Subjective	Unconstrained
		2.81187	6	1	Prediction	High Subjective	Unconstrained

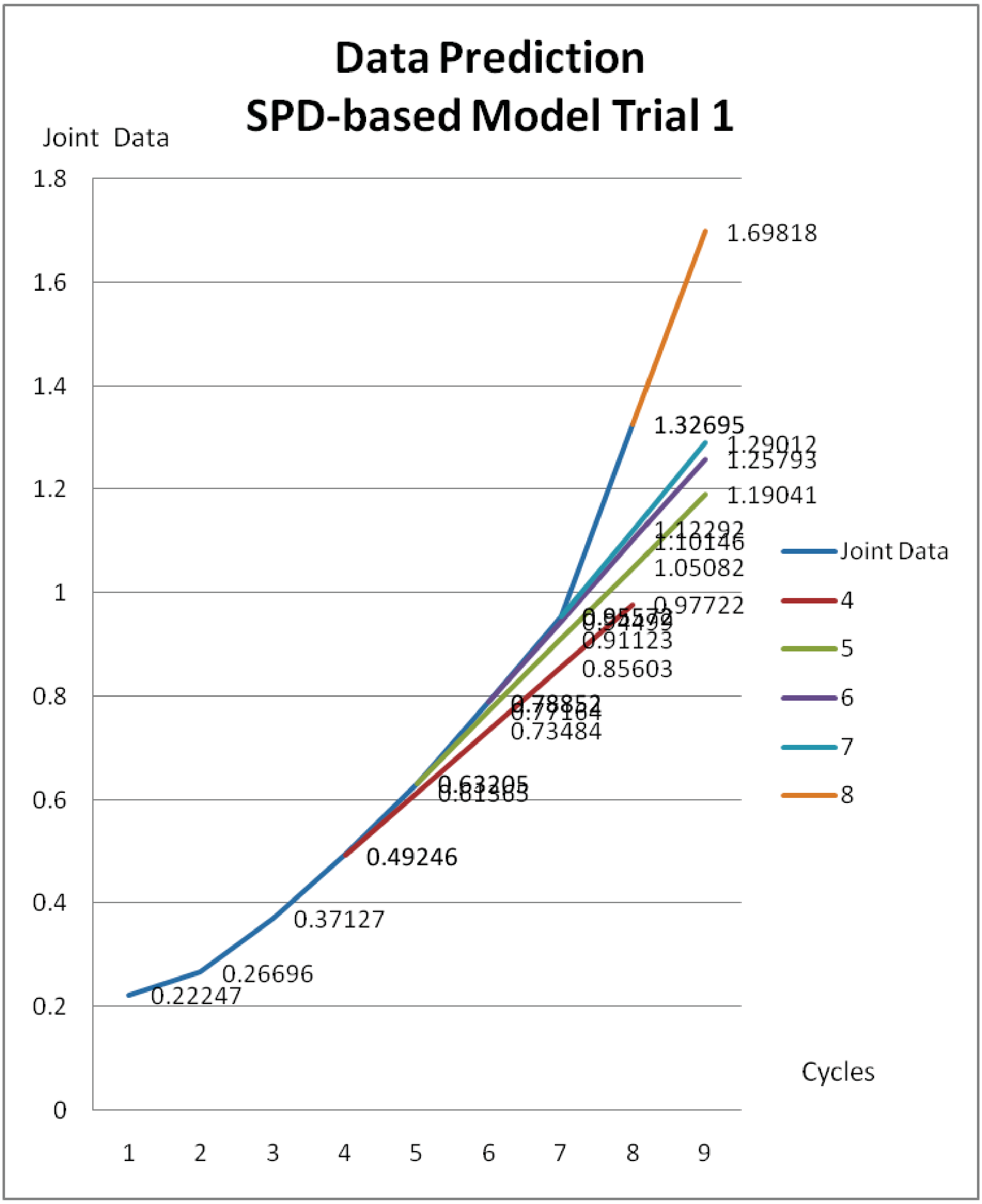

Fig. 6.21 Prediction Data SPD-based Model Trial 1

Table 6.35 SPD Pain Activation Trial 1

SPD Recommendation - Region Mapping						
Current		Prediction				
Pain (Predefined)	**Pain Region**	**Pain (Predefined)**	**Pain Region**	**Danger Interval**	**Consciousness State**	**Warning**
None						
None						
None						
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	Proprioceptive	3	4	Constrained	Future Pain
None	1	N/A	N/A	N/A	None	None
N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Table 6.36 Robot Mind Recommendation Trial 1

Mind Recommendation - Goals			**Action**
Awareness	**Consciousness State**	**Warning**	**Engine-Intention**
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	Future Pain	Alert Future
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
None	None	No danger	None
High Subjective	Constrained	Stop Now	Arm Resist

the Engine-Intention. The situation changes when the 7th Joint Data sequence is executed, in which all prediction cycles suggest that the predicted Joint Data will fall into faulty joint regions. SPD recommends the Robot Mind to exercise constraint after the final prediction has been processed. The Robot Mind then interprets this recommendation and constrains the Mind State, forces the Awareness to High Subjective type, sends out a warning to stop the process, and activates a counter action in the Engine-Intention to resist the robot arm.

The Joint Data and prediction data Trial 2 for the SPD-based model can be seen in Table 6.37, and Figure 6.22 depicts data pattern prediction of the experiment. Table 6.37 shows

Table 6.37 Joint Data and Prediction Data SPD-based Model Trial 2

No	Data	Prediction Data		
		5	6	7
1	0.02765			
2	0.02765			
3	0.02919			
4	0.2194			
5	0.39735	0.3974		
6	0.68421	0.57530	0.6842	
7	1.30548	0.75325	0.9711	1.30548
		0.93120	1.2579	1.92675
		1.1091	1.5448	2.54802
			1.8317	3.16929
				3.79056

Table 6.38 Prediction Error SPD-based Model Trial 2

Data	Prediction Cycles			Std D(σ)
5	0.00%			
6	10.89%	0.00%		0.08
7	55.22%	33.44%	0.00%	0.28

that the Robot Mind deciphers incoming information at sequence 4, with the real Joint Data being 0.2194. The first prediction constitutes the 5th sequence, 0.3974; the second prediction constitutes to the 6th sequence, 0.57530 and the third prediction for the 7th sequence is 0.75325. The first prediction produces a relatively high prediction error of 10.89%, which increases dramatically to 55.22% in the final prediction. However, the overall prediction

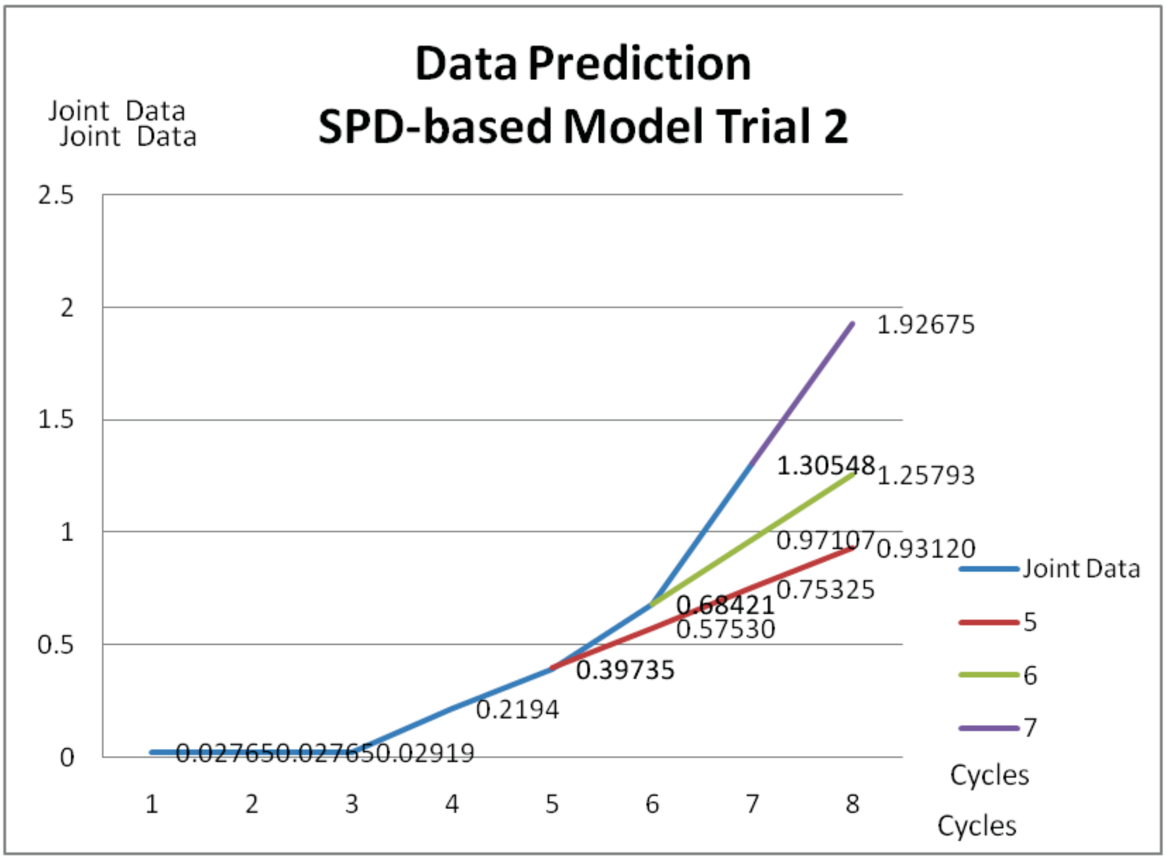

Fig. 6.22 Prediction Data SPD-based Model Trial 2

error in this sequence remains low, with a standard deviation of 0.08 (see Table 6.38 for prediction error during the SPD-based model Trial 2 experiment). When the robot processes the incoming data in the 6th sequence, the prediction error of the next sequence, 7th sequence, is high, 33.44% (Joint Data prediction is 0.9711). This process recommends that the robot should halt the experiment for Trial 2 as the predicted Joint Data will exceed the faulty joint value limit. Figure 6.22 depicts the prediction data pattern of the SPD-based model Trial 2 experiment. The Joint Data prediction inclines vertically with the final prediction data during the 7th sequence forcing the robot to stop. Table 6.40 and Table 6.41 show the pain activation and the Robot Mind recommendation during Trial 2. The initial state of the Robot Mind obtains consciousness in the unconstrained state and the awareness level is High Objective in Region 6 (CDV is 126).

Table 6.39 SPD Initial State Trial 2

No	Data Sensory	Internally	CDV	Region	Incoming Belief	Awareness	Consciousness State
1	0.02765		126	6	Current	High Objective	Unconstrained
2	0.02765		126	6	Current	High Objective	Unconstrained
3	0.02919		126	6	Current	High Objective	Unconstrained
4	0.2194		126	6	Current	High Objective	Unconstrained
5	0.39735		126	6	Current	High Objective	Unconstrained
		0.57530	126	6	Predict	High Objective	Unconstrained
		0.75325	126	6	Predict	High Objective	Unconstrained
		0.93120	126	6	Predict	High Objective	Unconstrained
		1.10915	126	6	Predict	High Objective	Unconstrained
6	0.68421		126	6	Current	High Objective	Unconstrained
		0.97107	126	6	Predict	High Objective	Unconstrained
		1.25793	126	6	Predict	High Objective	Unconstrained
		1.54479	126	6	Predict	High Objective	Unconstrained
		1.83165	126	6	Predict	High Objective	Unconstrained
7	1.30548		126	6	Current	High Objective	Unconstrained
		1.92675	126	6	Predict	High Objective	Unconstrained
		2.54802	126	6	Predict	High Objective	Unconstrained
		3.16929	126	6	Predict	High Objective	Unconstrained
		3.79056	126	6	Predict	High Objective	Unconstrained

It can be seen from Table 6.40 and Table 6.41 that the state of the Robot Mind during Trial 2 is similar to that in Trial 1, where no SPD recommendation passes to the Robot Mind during the 1st to 4th sequences. As a result, the Mind State remains unconstrained and no action is taken by the Engine-Intention. A similar situation occurs in the next two sequences and the SPD start to suggest pain activation in the 6th sequence, with the prediction for the 10th sequence (four interval) giving the kind of PP to be activated. The Robot Mind

Table 6.40 SPD Pain Activation Trial 2

SPD Recommendation - Region Mapping						
Current		**Prediction**				
Pain (Predefined)	**Pain Region**	**Pain (Predefined)**	**Pain Region**	**Danger Interval**	**Consciousness State**	**Warning**
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	Proprioceptive	3	4	None	Future Pain
None	1	N/A	N/A	N/A	None	None
N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Table 6.41 Robot Mind Recommendation Trial 2

Mind Recommendation - Goals			Action
Awareness	**Consciousness State**	**Warning**	**Engine-Intention**
	Unconstrained	No Danger	None
	None	None	None
	None	None	None
	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	Future Pain	Alert Future
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
High Subjective	Constrained	Stop Now	Arm Resist

then sends out a warning about future pain; however, a counter-physical action is sent to the Engine-Intention. The situation changes when the 7th Joint Data sequence is executed, in which all prediction cycles suggest that the predicted Joint Data is going to fall into the faulty joint regions. SPD recommends the Robot Mind to exercise constraint after the final prediction has been processed. The Robot Mind interprets this recommendation and constrains the Mind State, forces the Awareness to High Subjective type, sends out a warning to stop the process and activates a counter action in the Engine-Intention which is to resist the robot arm.

The Joint Data and prediction data for the SPD-based model Trial 3 experiment can be seen in Table 6.42, Table 6.43 and Figure 6.23 depicts the prediction data pattern of the experiment. During Trial 3, 10 sequences of Joint Data are obtained and the robot deciphers

Table 6.42 Joint Data and Prediction Data SPD-based Model Trial 3

No	**Data**	**Prediction Data**				
		6	**7**	**8**	**9**	**10**
1	0.02765					
2	0.02765					
3	0.02765					
4	0.06907					
5	0.29917					
6	0.52774	0.52774				
7	0.71642	0.75631	0.71642			
8	0.87902	0.98488	0.9051	0.87902		
9	1.0447	1.21345	1.09378	1.04162	1.0447	
10	1.41132	1.44202	1.28246	1.20422	1.21038	1.41132
			1.47114	1.20422	1.37606	1.77794
				1.52942	1.54174	2.14456
					1.70742	2.51118
						2.8778

the incoming data in the 6th Joint Data sequence. The first prediction cycle constitutes the 7th sequence, the second prediction constitutes the 8th sequence, the third constitutes the 9th sequence and the fourth constitutes the 10th sequence. All prediction values obtained during the 6th sequence are over-estimated predictions. The first three prediction cycles have an upward trend with the highest prediction error being for the 9th sequence, 16.87%. However, the error drops dramatically to 3.07%, with the predicted Joint Data being higher than the real Joint Data. In the 7th sequence, the first prediction error is 2.62% and increases to 12.89% in the final prediction. The 8th sequence has the highest prediction error, 20.71%, which is predicted for the 10th sequence and drops slightly to 20.09% during the 9th Joint Data

Table 6.43 Prediction Error SPD-based Model Trial 3

Data	Prediction Cycles 6	7	8	9	10	Std D(σ)
6	0.00%					
7	-3.99%	0.00%				0.03
8	-10.59%	2.61%	0.00%			0.06
9	-16.87%	4.91%	0.31%	0.00%		0.08
10	-3.07%	12.89%	20.71%	20.09%	0.00%	0.11
11						

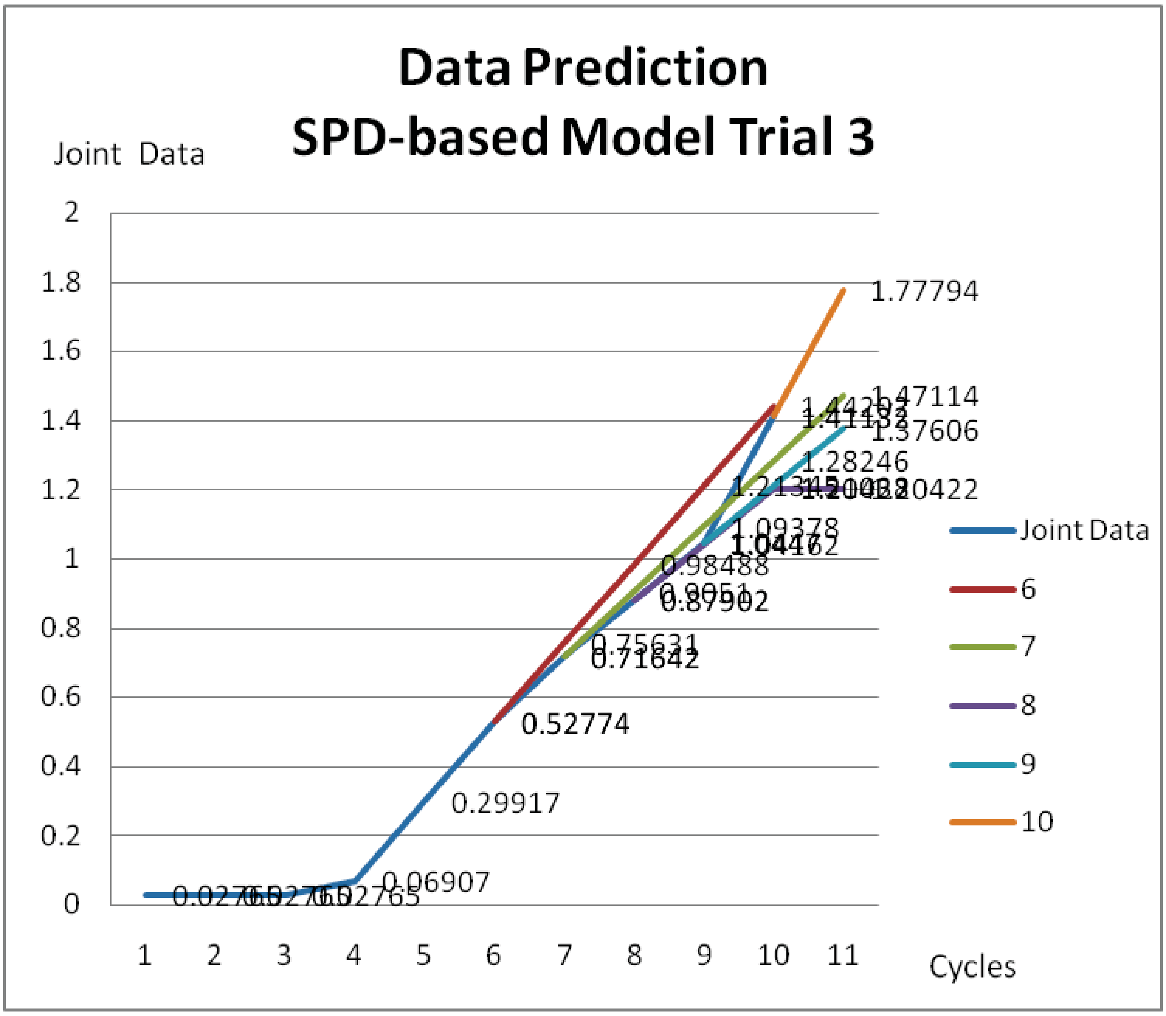

Fig. 6.23 Prediction Data SPD-based Model Trial 3

sequence. The pattern of the predicted Joint Data moves vertically and in the 10th sequence, the robot predicts that with the incoming next sequence, the Joint Data will fall into the faulty joint region.

Table 6.44 shows the initial state of the Robot Mind during the early stage of Trial 3. The prediction process commences in the 6th Joint Data sequence, and the states of the awareness is High Priority, the consciousness state is unconstrained and in Region 6 (CDV is 138). As can be seen from the table, all the elements of the Robot Mind remain the same. Pain activation also remains the same until the 9th Joint Data sequence is processed, where the final prediction suggests the kind of PP that should be invoked, the pain region is 3 and the pain is likely to occur in the 13th Joint Data sequence. However, the SPD suggests that the consciousness state should remain unconstrained with the warning signal of future pain. When the next sequence commences, the SPD predicts that the danger intervals consistently appear in all prediction cycles. As a result, the recommendation to be sent to the Robot Mind is that the consciousness state should be altered to the constrained state with the warning of pain evoked in the next sequence (see Table 6.45). The Goals are then modified, the robot awareness is forced to be in High Subjective condition under constraint, and the warning to be sent out is to stop straight away. The Engine-Intention executes a counter action by resisting the robot's arm, which is achieved by maximising the motor stiffness of the arm joint (see Table 6.46).

Table 6.47 shows the overall Joint Data being processed in Trial 4 along with the prediction data for each sequence of data, which lasts for seven sequences. Over four prediction cycles, three prediction data are taken into consideration. The first cycle of prediction, which constitutes the 5th sequence, predicts the Joint Data to be 0.42803 with prediction error 4.14%. The second cycle, which determines the predicted value of the 6th sequence, suggests a prediction error of 6.44%. The final prediction cycle generates a relatively high prediction error, 40.19% (predicted Joint Data is 0.82687 while the real Joint Data at predicted 7th sequence is 1.22878). The prediction error values for the next sequence of data decrease to 1.84% for the first prediction (under-estimated prediction) and 27.77% for the final prediction value, which drops slightly from the earlier prediction. In the 6th sequence of incoming Joint Data, 1.22878, the error value of the first prediction cycle increases to 31.45% (predicted Joint Data is 0.91431). In the last sequence of incoming data, the robot predicts that the Joint Data will be 1.76568 in the next sequence, which falls into the faulty joint region. As a result, the robot halts at the 7th Joint Data sequence (see Table 6.48 for prediction error Trial 4).

Figure 6.24 depicts the graphical-based analysis of the Joint Data and data prediction. It can be seen that from sequence 1 to sequence 3, the Joint Data remains the same, and in the 4th sequence, it starts to incline gradually as the prediction process takes place. In the

Table 6.44 SPD Initial State Trial 3

No	Data Sensory	Internally	CDV	Region	Incoming Belief	Awareness	Consciousness State
1	0.02765		138	6	Current	High Objective	Unconstrained
2	0.02765		138	6	Current	High Objective	Unconstrained
3	0.02765		138	6	Current	High Objective	Unconstrained
4	0.06907		138	6	Current	High Objective	Unconstrained
5	0.29917		138	6	Current	High Objective	Unconstrained
6	0.52774		138	6	Current	High Objective	Unconstrained
		0.75631	138	6	Prediction	High Objective	Unconstrained
		0.98488	138	6	Prediction	High Objective	Unconstrained
		1.21345	138	6	Prediction	High Objective	Unconstrained
		1.44202	138	6	Prediction	High Objective	Unconstrained
7	0.71642		138	6	Current	High Objective	Unconstrained
		0.9051	138	6	Prediction	High Objective	Unconstrained
		1.09378	138	6	Prediction	High Objective	Unconstrained
		1.28246	138	6	Prediction	High Objective	Unconstrained
		1.47114	138	6	Prediction	High Objective	Unconstrained
8	0.87902		138	6	Current	High Objective	Unconstrained
		1.04162	138	6	Prediction	High Objective	Unconstrained
		1.20422	138	6	Prediction	High Objective	Unconstrained
		1.20422	138	6	Prediction	High Objective	Unconstrained
		1.52942	138	6	Prediction	High Objective	Unconstrained
9	1.0447		138	6	Current	High Objective	Unconstrained
		1.21038	138	6	Prediction	High Objective	Unconstrained
		1.37606	138	6	Prediction	High Objective	Unconstrained
		1.54174	138	6	Prediction	High Objective	Unconstrained
		1.70742	138	6	Prediction	High Objective	Unconstrained
10	1.41132		138	6	Current	High Objective	Unconstrained
		1.77794	138	6	Prediction	High Objective	Unconstrained
		2.14456	138	6	Prediction	High Objective	Unconstrained
		2.51118	138	6	Prediction	High Objective	Unconstrained
		2.8778	138	6	Prediction	High Objective	Unconstrained

Table 6.45 SPD Pain Activation Trial 3

SPD Recommendation - Region Mapping						
Current		Prediction				
Pain (Predefined)	Pain Region	Pain (Predefined)	Pain Region	Danger Interval	Consciousness State	Warning
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	Proprioceptive	3	4	Unconstrained	Future Pain
None	1	N/A	N/A	N/A	None	None
N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Table 6.46 Robot Mind Recommendation Trial 3

Mind Recommendation - Goals			**Action**
Awareness	**Consciousness State**	**Warning**	**Engine-Intention**
	Unconstrained	No Danger	None
	None	None	None
	None	None	None
	None	None	None
	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	Future Pain	Alert Future
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
High Subjective	Constrained	Stop Now	Arm Resist

Table 6.47 Joint Data and Prediction Data SPD-based Model Trial 4

No	Data	Prediction Data			
		4	5	6	7
1	0.02612				
2	0.02765				
3	0.02919				
4	0.22861	0.22861			
5	0.46945	0.42803	0.46945		
6	0.69188	0.62745	0.71029	0.69188	
7	1.22878	0.82687	0.95113	0.91431	1.22878
		1.02629	1.19197	1.13674	1.76568
			1.43281	1.35917	2.30258
				1.5816	2.83948
				0.46945	3.37638

Table 6.48 Prediction Error SPD-based Model Trial 4

Data	Prediction Cycles 4	5	6	7	Std D(σ)
4	0.00%				
5	4.14%	0.00%			0.03
6	6.44%	-1.84%	0.00%		0.04
7	40.19%	27.77%	31.45%	0.00%	0.17

5th and 6th sequences, the prediction data almost converges to the real Joint Data, and in the 7th sequence, all prediction data are under-estimated. However, the prediction data for the 8th sequence is capable of preventing the robot from experiencing higher pain intensity and forcing the robot to execute a counter action by increasing joint stiffness of the motor of the arm joint.

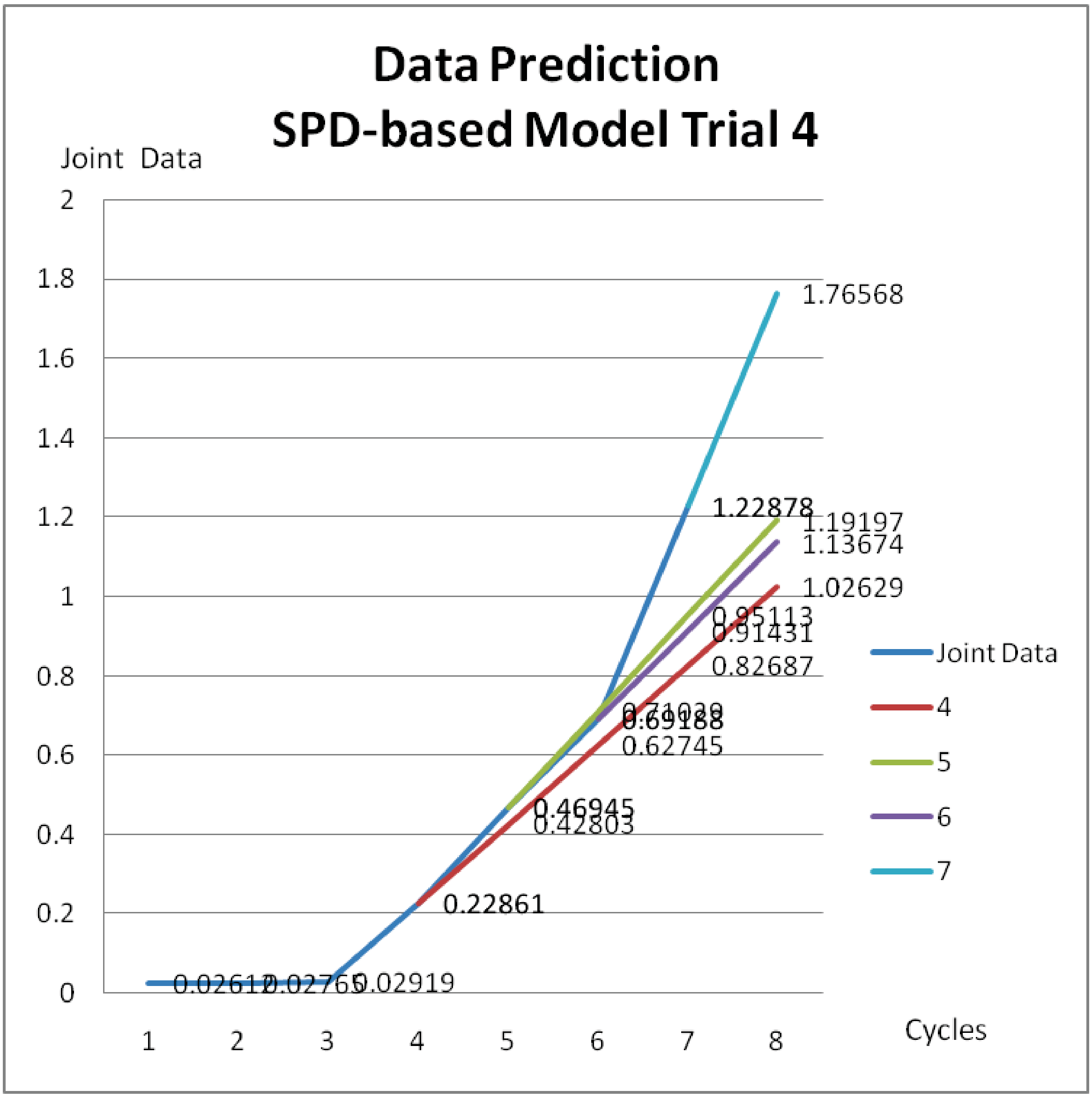

Fig. 6.24 Prediction Data SPD-based Model Trial 4

The initial state of the Robot Mind is shown in Table 6.49. The state of consciousness remains unconstrained from the 1st sequence to the 7th sequence. Similarly, the awareness state is Left Subjective, the activated Region is 3 and the CDV is 65. From the initial state, the

Table 6.49 SPD Initial State Trial 4

No	Data Sensory	Internally	CDV	Region	Incoming Belief	Awareness	Consciousness State
1	0.02612		65	3	Current	Left Subjective	Unconstrained
2	0.02765		65	3	Current	Left Subjective	Unconstrained
3	0.02919		65	3	Current	Left Subjective	Unconstrained
4	0.22861		65	3	Current	Left Subjective	Unconstrained
		0.42803	65	3	Predict	Left Subjective	Unconstrained
		0.62745	65	3	Predict	Left Subjective	Unconstrained
		0.82687	65	3	Predict	Left Subjective	Unconstrained
		1.02629	65	3	Predict	Left Subjective	Unconstrained
5	0.46945		65	3	Current	Left Subjective	Unconstrained
		0.71029	65	3	Predict	Left Subjective	Unconstrained
		0.95113	65	3	Predict	Left Subjective	Unconstrained
		1.19197	65	3	Predict	Left Subjective	Unconstrained
		1.43281	65	3	Predict	Left Subjective	Unconstrained
6	0.69188		65	3	Current	Left Subjective	Unconstrained
		0.91431	65	3	Predict	Left Subjective	Unconstrained
		1.13674	65	3	Predict	Left Subjective	Unconstrained
		1.35917	65	3	Predict	Left Subjective	Unconstrained
		1.5816	65	3	Predict	Left Subjective	Unconstrained
7	1.22878		65	3	Current	Left Subjective	Unconstrained
		1.76568	65	3	Predict	Left Subjective	Unconstrained
		2.30258	65	3	Predict	Left Subjective	Unconstrained
		2.83948	65	3	Predict	Left Subjective	Unconstrained
		3.37638	65	3	Predict	Left Subjective	Unconstrained

Robot Mind with SPD-based model sets the default values and deciphers the incoming data, which does not recommend that any changes should take place. Similarly, the Goals remain the same as the Robot Mind has not received a new recommendation. These conditions last until the 6th Joint Data sequence has been processed, particularly in the final prediction cycle. The SPD suggests that the kind of PP is generated and the pain region is 3 which will occur in the 4th interval. Since the interval is still far from taking place, the SPD recommends that the state of consciousness should continue to be unconstrained with a future pain warning. This recommendation is passed to the Goals and a warning is generated that this type PP will occur as future pain. The Engine-Intention then executes a counter response by sending out an alert about the future pain. When the 7th sequence is processed, the kind of IP is assessed, but there is no indication that the current Joint Data will fall into the faulty joint region. As a result, there is no recommendation to be passed to the Robot Mind's Goals. The kind of PP, however, is generated in all four prediction cycles when the pain prediction region is 3. Hence, the SPD Pain Activation mechanism generates recommendations that the state of the consciousness should be constrained with the next warning to indicate pain. The Goals then capture these recommendations, constraint the consciousness state and force the awareness to change to High Subjective with a warning to stop. At the same time, the Engine-Intention is activated which causes the counter action to be executed by increasing the robot joint stiffness (resistance on the robot's arm). Table 6.50 and Table 6.51 show the SPD-based pain activation and the Goals of the Robot Mind respectively.

In Trial 5, the experiment only lasts for 6 sequences, and the robot commences deciphering incoming data in the 4th Joint Data sequence. Table 6.52 shows the results obtained for the prediction data and Table 6.53 depicts the prediction error for each sequence and the deviation values. The robot starts prediction processes in the 4th Joint Data sequence; the first prediction cycle constitutes the 5th sequence and the final prediction is for the 6th sequence. The first prediction obtains an over-estimated prediction, resulting a higher value in the predicted Joint Data at 0.77318, while the real Joint Data is 0.74403 with a relatively low prediction error of 2.92%. The final prediction, however, has higher prediction error of 24.24%. In the 5th sequence, the prediction error of the first prediction cycle increases to 30.07%. The 6th sequence, which generates prediction data for the 7th sequence, obtains Joint Data of 1.76875. The Robot Mind then halts the experiment as the predicted Joint Data falls into the faulty joint region. The Joint Data remain the same in the 1st and 2nd sequences, and starts to incline in the 3rd sequence until the end of experiment Trial 5. The prediction data in the 6th sequence increases dramatically and peaks at 1.76875, which predicts that the 7th sequence will cause the robot to experience pain. This situation suggests that the Robot Mind should take preventive action by increasing its arm joint stiffness and discontinuing the

Table 6.50 SPD Pain Activation Trial 4

SPD Recommendation - Region Mapping						
Current		Prediction				
Pain (Predefined)	Pain Region	Pain (Predefined)	Pain Region	Danger Interval	Consciousness State	Warning
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	Proprioceptive	3	4	Unconstrained	Future Pain
None	1	N/A	N/A	N/A	None	None
N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

experiment (see Figure 6.25). Table 6.54 shows the initial state of the Robot Mind with the state of consciousness remaining unconstrained from the 1st sequence to the 6th sequence. Similarly, the awareness state is Low Objective; the activated Region is 5 and the CDV is 105.

Table 6.55 and Table 6.56 show the SPD-based pain activation and the Goals of the Robot Mind respectively. The anaylsis of Trial 5 is similar to that of Trial 4 as their internal states are uniform. The Robot Mind with SPD-based model sets the default values and deciphers the incoming data, which does not recommend that any change should be processed. Similarly, the Goals remain the same as the Robot Mind has not received a new recommendation. These conditions last until the 6th Joint Data sequence has been processed, particularly in the final prediction cycle. The SPD suggests that the kind of PP is generated; the pain region is 3 and the pain will occur in the 4th interval. Since the interval is still far from occurring, the SPD recommends that the state of consciousness should remain unconstrained with a future pain warning. This recommendation is passed to the Goals and the warning is generated that this kind of PP will occur in the future. The Engine-Intention then executes a counter response by sending out an alert about the future pain. When the 7th sequence is processed, the kind of IP is assessed, but there is no indication that the current Joint Data will fall into the faulty joint region. As a result, there is no recommendation to be passed into the Robot Mind's

Table 6.51 Robot Mind Recommendation Trial 4

Mind Recommendation - Goals			Action
Awareness	**Consciousness State**	**Warning**	**Engine-Intention**
	Unconstrained	No Danger	None
	None	None	None
	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	Future Pain	Alert Future
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
High Subjective	Constrained	Stop Now	Arm Resist

Table 6.52 Joint Data and Prediction Data SPD-based Model Trial 5

No	Data	Prediction Data		
		4	**5**	**6**
1	0.02765			
2	0.03839			
3	0.2915			
4	0.53234	0.53234		
5	0.74403	0.77318	0.74403	
6	1.25639	1.01402	0.95572	1.25639
		1.25486	1.16741	1.76875
		1.4957	1.3791	2.28111
			1.59079	2.79347
				3.30583

Table 6.53 Prediction Error SPD-based Model Trial 5

Data	**Prediction Cycles**			**Std D(σ)**
	4	**5**	**6**	
4	0.00%			
5	-2.92%	0.00%		0.02
6	24.24%	30.07%	0.00%	0.16

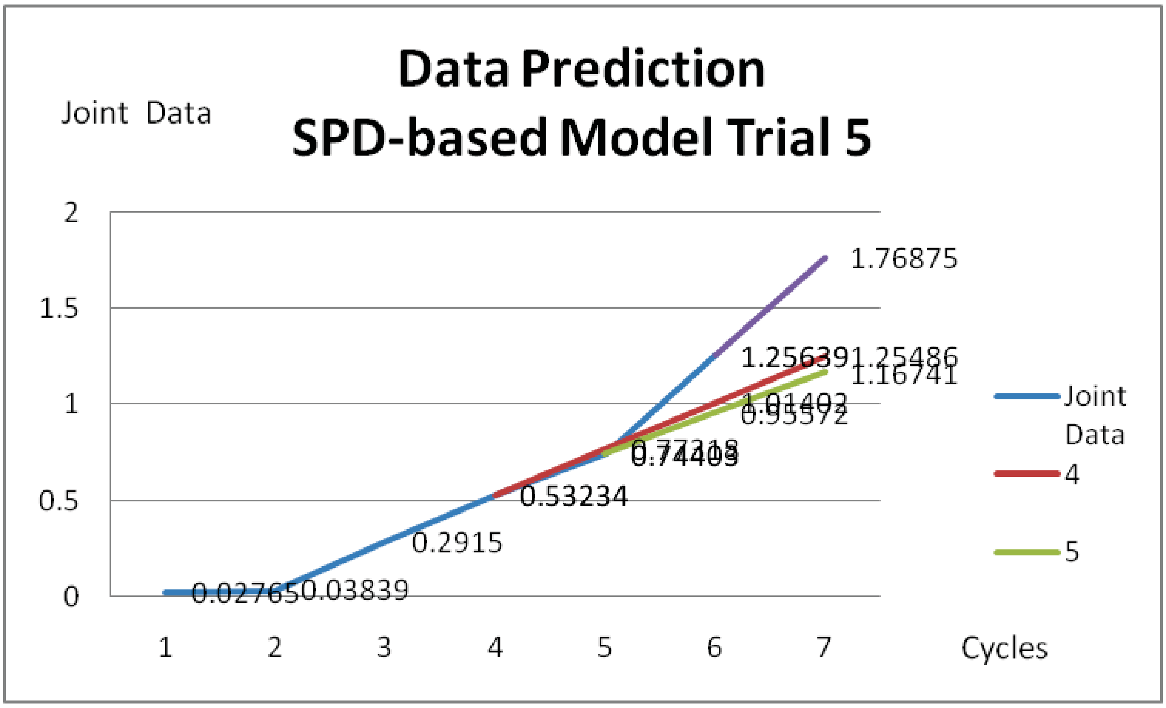

Fig. 6.25 Prediction Data SPD-based Model Trial 5

Table 6.54 SPD Initial State Trial 5

No	Data Sensory	Internally	CDV	Region	Incoming Belief	Awareness	Consciousness State
1	0.02765		105	5	Current	Low Objective	Unconstrained
2	0.03839		105	5	Current	Low Objective	Unconstrained
3	0.2915		105	5	Current	Low Objective	Unconstrained
4	0.53234		105	5	Current	Low Objective	Unconstrained
		0.77318	105	5	Predict	Low Objective	Unconstrained
		1.01402	105	5	Predict	Low Objective	Unconstrained
		1.25486	105	5	Predict	Low Objective	Unconstrained
		1.4957	105	5	Predict	Low Objective	Unconstrained
5	0.74403		105	5	Current	Low Objective	Unconstrained
		0.95572	105	5	Predict	Low Objective	Unconstrained
		1.16741	105	5	Predict	Low Objective	Unconstrained
		1.3791	105	5	Predict	Low Objective	Unconstrained
		1.59079	105	5	Predict	Low Objective	Unconstrained
6	1.25639		105	5	Current	Low Objective	Unconstrained
		1.76875	105	5	Predict	Low Objective	Unconstrained
		2.28111	105	5	Predict	Low Objective	Unconstrained
		2.79347	105	5	Predict	Low Objective	Unconstrained
		3.30583	105	5	Predict	Low Objective	Unconstrained

Goals (the kind of PP, however, is generated in all four prediction cycles). Hence, the SPD Pain Activation mechanism generates recommendations that the state of the consciousness is to be constrained and a warning that the pain will occur in the next sequence. The Goals then capture these recommendations, constrain the consciousness state and force the awareness to change to High Subjective with the warning to stop. At the same time, the Engine-Intention is activated which causes the counter action to be executed by increasing the robot joint stiffness (resistance on the robot's arm).

The results of the average SPD pain activation and the Robot Mind's recommendations are shown in Tables 6.57 and 6.58 respectively (the data is obtained from the second online experiment with physical interaction). It can be seen that the SPD activation system does not make any recommendation for four data cycles, hence, the Goals do not change the condition of the robot awareness. No pain is generated until the sixth data cycle is processed, which produces the kind of PP for which the warning recommendation is Future Pain. However, the Goals remain None, preventing the Intentions from making counter responses. As the next incoming data analysis predicts that the type of PP will persist throughout the prediction cycles, the Goals now constrain the state of the Robot Mind and recommend that the Intentions should take action by increasing the stiffness of the joint (Arm Resist).

Table 6.55 SPD Pain Activation Trial 5

SPD Recommendation - Region Mapping						
Current		Prediction				
Pain (Predefined)	Pain Region	Pain (Predefined)	Pain Region	Danger Interval	Consciousness State	Warning
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
None	1	N/A	N/A	N/A	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	None	1	None	None	None
N/A	N/A	Proprioceptive	3	4	Unconstrained	Future Pain
None	1	N/A	N/A	N/A	None	None
N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Table 6.56 Robot Mind Recommendation Trial 5

Mind Recommendation - Goals			Action
Awareness	Consciousness State	Warning	Engine-Intention
	Unconstrained	No Danger	None
	None	None	None
	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
None	None	Future Pain	Alert Future
None	None	None	None
None	None	None	None
None	None	None	None
None	None	None	None
High Subjective	Constrained	Stop Now	Arm Resist

Table 6.57 SPD Pain Activation - Average

No	Data Sensory	Internally	SPD Recommendation - Region Mapping						
			Current		Prediction				
			Pain	Pain Region	Pain	Pain Region	Interval	Mind State	Warning
1	0.02765								
2	0.02765								
3	0.02919								
4	0.2194								
5	0.39735		None	1	N/A	N/A	N/A	None	None
		0.57530	N/A	N/A	None	1	None	None	None
		0.75325	N/A	N/A	None	1	None	None	None
		0.93120	N/A	N/A	None	1	None	None	None
		1.10915	N/A	N/A	None	1	None	None	None
6	0.68421		None	1	N/A	N/A	N/A	None	None
		0.97107	N/A	N/A	None	1	None	None	None
		1.25793	N/A	N/A	None	1	None	None	None
		1.54479	N/A	N/A	None	1	None	None	None
		1.83165	N/A	N/A	Proprioceptive	3	4	None	Future Pain
7	1.30548		None	1	N/A	N/A	N/A	None	None
		1.92675	N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
		2.54802	N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
		3.16929	N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
		3.79056	N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Table 6.58 Robot Mind Recommendations

No	Data Sensory	Internally	Mind Recommendation - Goals			Action
			Awareness	Mind State	Warning	Engine-Intention
1	0.02765			Unconstrained	No Danger	None
2	0.02765			None	None	None
3	0.02919			None	None	None
4	0.2194			None	None	None
5	0.39735		None	None	None	None
		0.57530	None	None	None	None
		0.75325	None	None	None	None
		0.93120	None	None	None	None
		1.10915	None	None	None	None
6	0.68421		None	None	None	None
		0.97107	None	None	None	None
		1.25793	None	None	None	None
		1.54479	None	None	None	None
		1.83165	None	None	Future Pain	Alert Future
7	1.30548		None	None	None	None
		1.92675	None	None	None	None
		2.54802	None	None	None	None
		3.16929	None	None	None	None
		3.79056	High Subjective	Constrained	Stop Now	Arm Resist

6.2.2 Pain Matrix-based Model

The data collected during the upward and downward hand movement direction experiments are shown in Tables 6.59 and 6.60 respectively.

Table 6.59 Upward Hand Movement Direction

No	Data	Prediction Data									
		4	5	6	7	8	9	10	11	12	13
1	-0.0153										
2	-0.03064										
3	-0.11501										
4	-0.26534										
5	-0.44328	-0.416									
6	-0.50464	-0.566	-0.621								
7	-0.50464	-0.716	-0.799	-0.566							
8	-0.71787	-0.867	-0.977	-0.627	-0.505						
9	-0.96331		-1.155	-0.689	-0.505	-0.931					
10	-1.13052			-0.750	-0.505	-1.144	-1.209				
11	-1.27164				-0.505	-1.358	-1.454	-1.298			
12	-1.4772					-1.571	-1.700	-1.465	-1.413		
13	-1.61526						-1.945	-1.632	-1.554	-1.683	
								-1.799	-1.695	-1.888	-1.753
									-1.836	-2.094	-1.891
										-2.299	-2.029
											-2.167

Table 6.60 Downward Hand Movement Direction

No	Data	Prediction Data					
		4	5	6	7	8	9
1	0.05527						
2	0.08901						
3	0.31758						
4	0.5937						
5	0.87749	0.86982					
6	1.14594	1.14594	1.16128				
7	1.38678	1.42206	1.44507	1.41439			
8	1.54478	1.69818	1.72886	1.68284	1.62762		
9	1.68284		2.01265	1.95129	1.86846	1.70278	
				2.21974	2.1093	1.86078	1.8209
					2.35014	2.01878	1.95896
						2.17678	2.09702
							2.23508

Pain Matrix-based Model Data Analysis

The two kinds of hand movement experiments have the same results and the Robot Mind produces data predictions at the 4th data cycle obtained from the joint position sensor of the robot. Because of this similarity, the discussion of the results is limited to one typical experiment, which is the Upward hand movement direction experiment.

We divided the data into prediction cycles, each of which contains of four time predictions. We obtained the highest error prediction occurring in the seventh data cycle with a standard deviation of about 0.5. The overall error prediction is shown in Table 6.61).

Table 6.61 Upward Hand Movement Prediction

No	Data	Prediction Data									
		4	5	6	7	8	9	10	11	12	13
1	-0.0153										
2	-0.03064										
3	-0.11501										
4	-0.26534	0.000									
5	-0.44328	0.020	0.000								
6	-0.50464	0.043	0.082	0.000							
7	-0.50464	0.150	0.208	0.043	0.000						
8	-0.71787	0.105	0.183	0.064	0.151	0.000					
9	-0.96331		0.136	0.194	0.324	0.023	0.000				
10	-1.13052			0.269	0.443	0.010	0.055	0.000			
11	-1.27164				0.542	0.061	0.129	0.018	0.000		
12	-1.4772					0.066	0.157	0.009	0.046	0.000	
13	-1.61526						0.233	0.012	0.043	0.048	0.000

Pain Matrix-based Model Activation

The state of the Belief, the pain activation, and the Goals and Intentions are shown in Table 6.62, Table 6.63, and Table 6.65 respectively. The state of the Belief of the robot is

Table 6.62 Belief State During Non-Empathy Experiment Using Pain Matrix Model

No	Data		CDV	Region	Incoming Belief	Awareness	Mind State	Extero ceptive	Proprio- ceptive
	Sensory	Internally							
1	-0.0153								
2	-0.0306								
3	-0.115				Current				
4 to 11	-0.2653								
		-0.41567							
		-0.566							
		-0.71633			Prediction		Un constrained		
		-0.86666							
12	-1.4772				Current	Lower Objective			
		-1.68276							
		-1.88832							
		-2.09388	119	5	Prediction			FALSE	TRUE
		-2.29944							
13	-1.61526				Current				
		-1.75332							
		-1.89138							
		-2.02944			Prediction				
		-2.1675							

comprised of several elements such as the data originating from the Sensory Mechanisms

Table 6.63 Pain Activation During Non-Empathy Experiment Using Pain Matrix Model

No	Activation		
	Booster	**Pain Init**	**Cons Modifier**
1	N/A	N/A	N/A
2	N/A	N/A	N/A
3	N/A	N/A	N/A
4 to 11	0.000856642	0.266197	FALSE
	0	0.41567	FALSE
	0	0.566	FALSE
	0	0.71633	FALSE
	0	0.86666	FALSE
12	0.000856642	1.4780566	FALSE
	0	1.68276	FALSE
	0	1.88832	FALSE
	0	2.09388	FALSE
	0	2.29944	FALSE
13	0.000856642	1.6161166	TRUE
	0	1.75332	FALSE
	0	1.89138	FALSE
	0	2.02944	FALSE
	0	2.1675	FALSE

(either exteroception or 'proprioception', the CDV, the values that correspond to the region of robot awareness, and the condition of the Robot Mind. During the experiment, which lasts for 13 cycles, the knowledge of the robot remains the same. The Consciousness Modifier remains inactive throughout out the experiment cycles. When the Incoming Belief data is Current, the Booster affects the functionality of the Pain Matrix, whereas when it is Prediction, the Booster is not allowed to affect the Pain Matrix.

From the 4th cycle to the 11th cycle, the Pain Matrix recommendations are uniform, and the kind of pain to be originated is the IP at pain level *None*. At the 12th cycle, the Current analysis does not recommend an increase in the level of detection, but the Prediction analysis estimates that the robot is suffering from PP and should take different preventive actions over the next four cycles of data. For the 13th and 14th cycles, the level of preventive actions is Low; for the 15th and 16th cycles, the level is Medium. It is also recommended that the state of Robot Mind should be altered to Constrained and that the Warning to be invoked should be Future Pain. The Robot Mind assesses this recommendation and concludes that there is no need to make any change to the state of the Robot Mind, and that this is the Goal to achieve at this stage.

In cycle 13, the Consciousness Modifier is set to TRUE allowing the Pain Matrix to influence the robot's awareness framework at the Standard level of influence. When the data is at Current level, the IP level increases to *Moderate* and the level of response action is Medium. The Prediction analysis suggests that the level of pain will increase to *Slight* as the predicted response actions increase through four of the subsequent data cycles (14th, 15th,

Table 6.64 Pain Matrix Output During Non-Empathy Experiment

No	**Current**						
	Consc Modifier	**Pain Dist**	**Kind of Pain**	**Pain level**	**empathised actions**	**self actions**	
1	N/A	N/A	N/A	N/A	N/A	N/A	
2	N/A	N/A	N/A	N/A	N/A	N/A	
3	N/A	N/A	N/A	N/A	N/A	N/A	
4 to 11	-	0.267053284	Inflammatory	None	N/A	None	
	-	0.41567	Proprioceptive	N/A	N/A	N/A	
	-	0.566	Proprioceptive	N/A	N/A	N/A	
	-	0.71633	Proprioceptive	N/A	N/A	N/A	
	-	0.86666	Proprioceptive	N/A	N/A	N/A	
12	-	1.478913284	Inflammatory	None	N/A	None	
	-	1.68276	Proprioceptive	N/A	N/A	N/A	
	-	1.88832	Proprioceptive	N/A	N/A	N/A	
	-	2.09388	Proprioceptive	N/A	N/A	N/A	
	-	2.29944	Proprioceptive	N/A	N/A	N/A	
13	Standard	1.616973284	Inflammatory	Moderate	N/A	Medium	
	-	1.75332	Proprioceptive	N/A	N/A	N/A	
	-	1.89138	Proprioceptive	N/A	N/A	N/A	
	-	2.02944	Proprioceptive	N/A	N/A	N/A	
	-	2.1675	Proprioceptive	N/A	N/A	N/A	
			Prediction			**Output**	
	pain level	**empathy action**	**self action**	**danger low**	**danger medium**	**Mind State**	**Warning**
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4 to 11	N/A	N/A	N/A	N/A	N/A	None	None
	None	-	-	-	-	None	None
	None	-	-	-	-	None	None
	None	-	-	-	-	None	None
	None	-	-	-	-	None	None
12	N/A	N/A	N/A	N/A	N/A	None	None
	None	-	Low	1	-	None	None
	None	-	Low	2	-	None	None
	None	-	Medium	-	3	Constrained	Future Pain
	None	-	Medium	-	4	Constrained	Future Pain
13	N/A	N/A	N/A	N/A	N/A	Constrained	Next Pain
	None	-	Low	1	-	None	None
	None	-	Low	2	-	None	None
	None	-	Low	3	-	None	None
	Slight	-	Medium	-	4	Constrained	Future Pain

Table 6.65 Goals - Intentions During Non-Empathy Experiment Using Pain Matrix Model

No	Mind Recommendation - Goals					Action Engine-Intention
	CDV	Region	Awareness	Mind State	Warning	
1	N/A	N/A	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A
4 to 11	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	None	None	None	None	No Danger	None
12	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A
	55	1	None	None	N/A	None
13	55	1	Left Subjective	Constrained	Next Pain	N/A
	None	None	N/A	N/A	N/A	N/A
	None	None	N/A	N/A	N/A	N/A
	None	None	N/A	N/A	N/A	N/A
	55	1	High Subject	Constrained	Future Pain	Resist Now

17th, and 18th). Hence, the Pain Matrix converges to a recommendation that the Mind State should be changed to Constrained and the Warning to Next Pain. The Robot Mind then sets the Goal to modify the CDV value (the new CDV is 55 and the region is 1) which will force the Awareness to change to Left Subjective Awareness and the Mind State to Constrained. Based on these Goals, the robot's Intentions function executes a counter reaction, which is Arm Resistance achieved by increasing the stiffness of the shoulder joint of the robot.

6.3 Empathy-based Experiments

Two phases of experiments are carried out in which the SPD model and the Pain Matrix model are implemented into the framework. The faulty joint regions throughout these experiments remain the same, as shown in Table 6.66.

Table 6.66 Faulty Joint Regions

Element		Region/Ordinate	Values
Faulty Joint Region	**1**	**Upper High**	-2.08313
	2	**Upper Medium**	-1.58313
	3	**Upper Low**	-1.38313
	4	**Lower Low**	1.385669
	5	**Lower Medium**	1.585669
	6	**Lower High**	2.085669

6.3.1 SPD Model

Two kinds of data are obtained during this phase of experimentation as shown in Table 6.67. As the Upward and Downward hand movement experiments share the same results,

Table 6.67 Observer Data with SPD Model in Empathy Experiments

No	Upward		Downward	
	Data	Prediction 4	Data	Prediction 4
1	1.214		0.242	
2	-0.291		1.186	
3	-0.777		1.805	
4	-1.165		2.105	
		-1.553		2.405
		-1.941		2.705
		-2.329		3.005
		-2.717		3.305

the discussion will present only one of the hand movement experiments, the Upward hand movement. However, for data that differs during the experiments, the discussion will present both results.

SPD-based Model Analysis

The Robot Mind starts the reasoning process, predicts the data and establishes Goals and Intentions about the future at the 4th cycle (see Figure 6.26). The internal state of the Robot Mind from which the Belief State is derived is shown in Table 6.68 below.
There are two kinds of raw data; Current data, which originates from the Sensory Mechanism,

Table 6.68 Belief State of the Observer in SPD Model

No	Data		CDV	Region	Incoming Belief	Awareness	Mind State
	Sensory	Internally					
1	1.214		2	1	Current	Upper Subjective	Unconstrained
2	-0.291		2	1	Current	Upper Subjective	Unconstrained
3	-0.777		2	1	Current	Upper Subjective	Unconstrained
4	-1.165		2	1	Current	Upper Subjective	Unconstrained
		-1.553	2	1	Prediction	Upper Subjective	Unconstrained
		-1.941	2	1	Prediction	Upper Subjective	Unconstrained
		-2.329	2	1	Prediction	Upper Subjective	Unconstrained
		-2.717	2	1	Prediction	Upper Subjective	Unconstrained

and Prediction data, which is generated internally. The CDV, in particular, can be generated

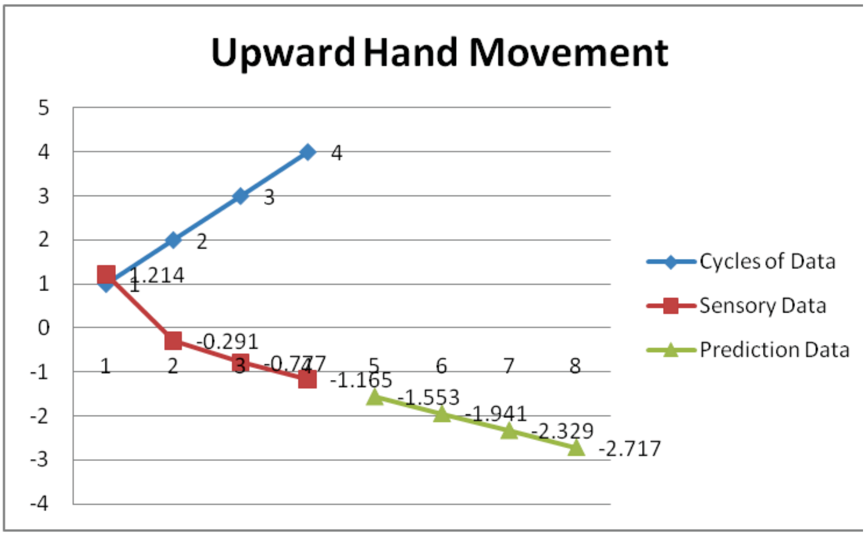

Fig. 6.26 Observer Data

either internally or externally. Manually predefining the CDV, however, is only available at the initiation stage, and during the process, it is governed by the Robot Mind. The CDV determines the active Region, the initial state of Awareness and the Mind State, while the Incoming Belief, which refers to whether the data classification is Current or Prediction, is obtained after the Robot Mind has reasoned about the incoming data from the Sensory Mechanism. The comparison of the Observer and Mediator data is shown Tables 6.69 and 6.70. It can be seen that the data cycles of the Observer and the Mediator differ significantly.

Table 6.69 Observer and Mediator Data During Upward Experiment

No	Mediator Data Cycle	Mediator Data	Observer Data Cycle	Observer Data	Observer Prediction
1	1	0.01845	1	1.214	
2	2643	-0.29449	2	-0.291	
3	3244	-0.77923	3	-0.777	
4	3695	-1.16273	4	-1.165	
5	4214	-1.5493			-1.553
6	4695	-1.94354			-1.941
7	5595	-2.08567			-2.329
8		-2.08567			-2.717

During the Upward experiment, for instance, the total number of data occurrences in the Mediator are 5595 cycles while in the Observer, they are only four cycles. A similar pattern also occurs in the Downward experiment, with fewer data occurrences in the Mediator (469 cycles), while those in the Observer remain the same at four cycles. Both Observer

Table 6.70 Observer and Mediator Data During Downward Experiment

No	Mediator		Observer		
	Data Cycle	Data	Data Cycle	Data	Prediction
1	75	0.23474	1	0.242	
2	227	1.18429	2	1.186	
3	379	1.80556	3	1.805	
4	415	1.93902	4	2.105	
5	437	2.01112			2.405
6	450	2.04026			2.705
7	469	2.08567			3.005
8		2.08567			3.305

experiments halt at the $4th$ cycle and the Observer follows up with the execution of empathic action. However, from the Mediator's perspective, the process continues and stops in a later data cycles. This shows that the Observer does not perform empathic reactions immediately as the robot requires additional time to execute the action; for example, approaching the scene of interaction. The Robot Mind of the Observer additionally performs more complex data analysis, which increases the execution time.

SPD-based Model Activation

The kinds of synthetic pain to be evoked by the SPD model are predefined prior to experiments and the kinds of pain to be modelled are the PP and IP, mentioned in Chapter 5.

The SPD Model recommendations are different during the Upward and Downward experiments, as shown in Table 6.71 and Table 6.72 respectively. During the Upward Hand

Table 6.71 SPD Recommendations - Upward Experiment

Data			SPD Recommendation - Region Mapping						
No	Sensory	Internally	Current		Prediction			Mind State	Warning
			Pain	Pain Region	Pain	Pain Region	Danger Interval		
1	1.214		N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	-0.291		N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	-0.777		N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	-1.165		None	1	N/A	N/A	N/A	None	None
		-1.553	N/A	N/A	Proprioceptive	2	1	Constrained	Next Pain
		-1.941	N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
		-2.329	N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
		-2.717	N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

Movement experiment, the first prediction cycle recommends that the robot should experience no pain at the joint value of -1.553, hence, the Mind State and the Warning recommendation do not produce any new amendment. However, the SPD suggests that the future incoming

Table 6.72 SPD Recommendations - Downward Experiment

Data			SPD Recommendation - Region Mapping						
			Current		Prediction				
No	Sensory	Internally	Pain	Pain Region	Pain	Pain Region	Danger Interval	Mind State	Warning
1	0.242		N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	1.186		N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	1.805		N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	2.105		Inflammatory	3	N/A	N/A	N/A	Constrained	In Pain
		2.405	N/A	N/A	Proprioceptive	3	1	Constrained	Next Pain
		2.705	N/A	N/A	Proprioceptive	3	2	Constrained	Next Pain
		3.005	N/A	N/A	Proprioceptive	3	3	Constrained	Next Pain
		3.305	N/A	N/A	Proprioceptive	3	4	Constrained	Next Pain

data should be predicted to generate the kind of PP which is persistent throughout four cycles of predictions. This suggests that the new recommendation for the Mind State will be Constrained and the kind of Warning to be sent out is Next Pain. In contrast, the 4th data cycle in the Downward experiment invokes the IP Pain Region 3. The SPD then recommends that the Mind State should be Constrained, while the kind of Warning is In Pain, which means that the robot has been suffering from the IP. This recommendation is halted until the Robot Mind has processed the prediction analysis. Unfortunately, the SPD predicts that the PP type will occur in the same region, *Region*3, and this pain will persist through all the prediction cycles.

Based on these recommendations, the Robot Mind sets up new Goals and directs the Intentions to generate the execution of empathic reactions by the Action Engine. In the first data cycle, the Robot Mind establishes the state as Unconstrained and as this occurs internally, there are no external actions to be executed. This internal state lasts until the 4th cycle of data occurs which allows the Robot Mind to reason and present four cycles of data prediction. The difference between the Upward and Downward experiments is that during the Upward experiment, there is no recommendation for the Goals for the Current data, while in the Downward experiment, the Robot Mind makes recommendations and sets up new Goals. In both cases, the Intentions execute empathic actions after the Prediction data analysis has confirmed the recommendation of the SPD model. The final Awareness of the robot is High Subjective, in which the Mind State is Constrained and the Warning to be sent out is Stop Now, which forces the Intentions to execute the counter action Arm Resist. As the counter action is in the form of an empathic action, the Observer alerts the human peer to stop and alerts the Mediator to increase its joint stiffness. The Mediator follows up by approaching the scene of the interaction to physically stop the interaction. The full range of data analyses from the Upward and Downward experiments are shown in Table 6.73 and Table 6.74 respectively.

Table 6.73 Goals and Intentions - Upward Experiment

No	Data		Mind Recommendation - Goals			Action
	Sensory	Internally	Awareness	Mind State	Warning	Engine-Intention
1	1.214		N/A	Unconstrained	No Danger	None
2	-0.291		N/A	None	None	None
3	-0.777		N/A	None	None	None
4	-1.165		N/A	N/A	N/A	N/A
		-1.553	N/A	N/A	N/A	N/A
		-1.941	N/A	N/A	N/A	N/A
		-2.329	N/A	N/A	N/A	N/A
		-2.717	N/A	N/A	N/A	N/A
			High Subjective	Constrained	Stop Now	Arm Resist

Table 6.74 Goals and Intentions - Downward Experiment

No	Data		Mind Recommendation - Goals			Action
	Sensory	Internally	Awareness	Mind State	Warning	Engine-Intention
1	0.242		N/A	Unconstrained	No Danger	None
2	1.186		N/A	None	None	None
3	1.805		N/A	None	None	None
4	2.105		None	Constrained	Stop Now	None
		2.405	N/A	N/A	N/A	N/A
		2.705	N/A	N/A	N/A	N/A
		3.005	N/A	N/A	N/A	N/A
		3.305	N/A	N/A	N/A	N/A
			High Subjective	Constrained	Stop Now	Arm Resist

6.3.2 Pain Matrix Model

Unlike the SPD Model, the Pain Matrix Model performs complex analyses in evoking the synthetic pain, so for the purpose of clarity, the following discussions are grouped according to the raw data analyses of the Joint Data and the pain activation analyses.

Raw Data Analyses

The results of the Observer data during the Upward and Downward experiments are shown in Table 6.75 below. The Robot Mind processes the incoming data at the 4th cycle and

Table 6.75 Observer Data with Pain Matrix Model

No	Upward				Downward			
	Data	Prediction 4	Prediction 5	Standard Deviation	Data	Prediction 4	Prediction 5	Standard Deviation
1	0.901				-0.015			
2	-1.057				1.137			
3	-2.047				2.376			
4	-2.084	-2.084		0	2.504	2.504		0
5	-2.1	-2.121	-2.1	0.01485	2.507	2.632	2.507	0.088388
		-2.158	-2.116			2.760	2.510	
		-2.195	-2.132			2.888	2.513	
		-2.232	-2.148			3.016	2.516	
			-2.164				2.519	

halts at the next data cycle, and the Current and Prediction data evoke the generation of synthetic pain. The results of the prediction are measurable only on the first cycle of the Prediction data, as the Robot Mind still allows the execution of the next Joint Data, with a standard deviation of 0.01485 for the Upward experiment and 0.088388 for the Downward experiment, both of which are relatively low.

The overall data mapping of the Current and Prediction data in relation to the Faulty Joint Region is shown in Figure 6.27 for the Upward experiment, and Figure 6.28 for the Downward experiment. It can be seen from the Upward experiment figure, 6.27, that the first two cycles of data are lower than the limit of the Lower Stage of the Faulty Joint Region. In data cycle 3, the Robot Mind determines that the Current and Prediction data have exceeded the Medium Stage. The critical analysis occurs at the 4th cycle when the value of the Current data almost falls into the Upper High Stage and the Prediction produces a consistent recommendation for the high possibility of pain invocation. A similar situation occurs in the Downward experiment; however, the Current data in the 3rd cycle has violated the High Stage of the Faulty Joint Region which causes the robot to experience a specific kind of synthetic pain. As a result of the noisy data suggested by the Robot Mind, the process

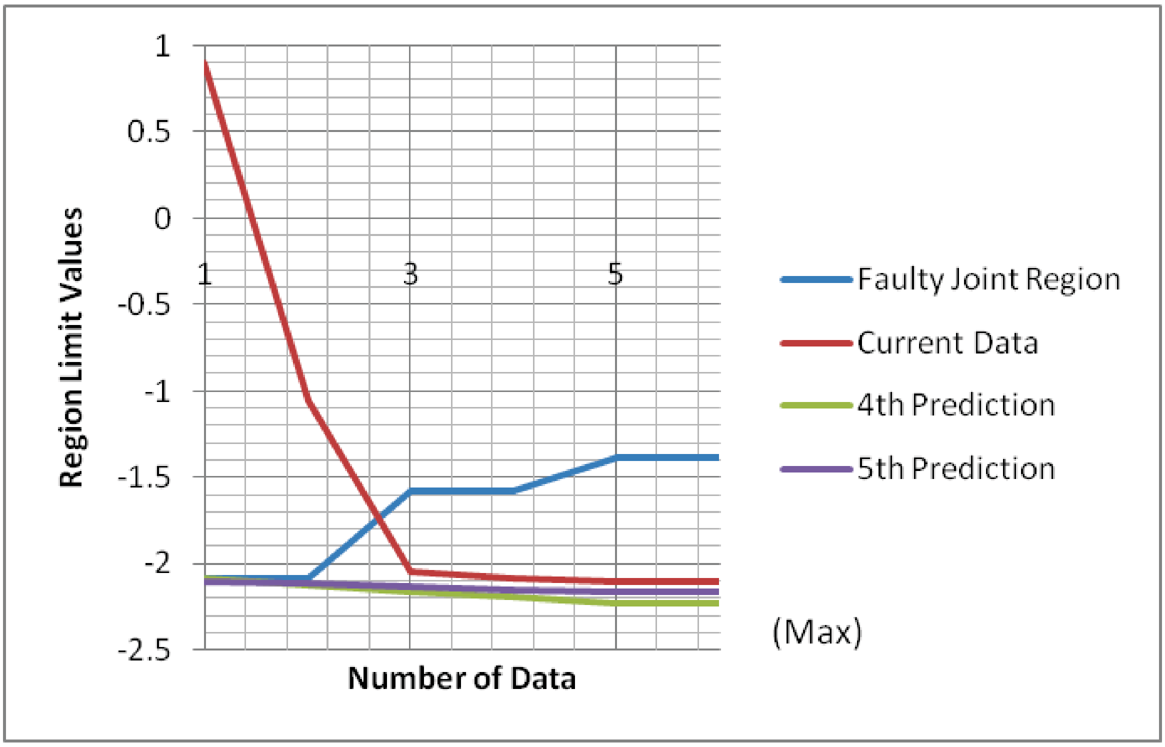

Fig. 6.27 Region Mapping of Joint Data - Upward Experiment

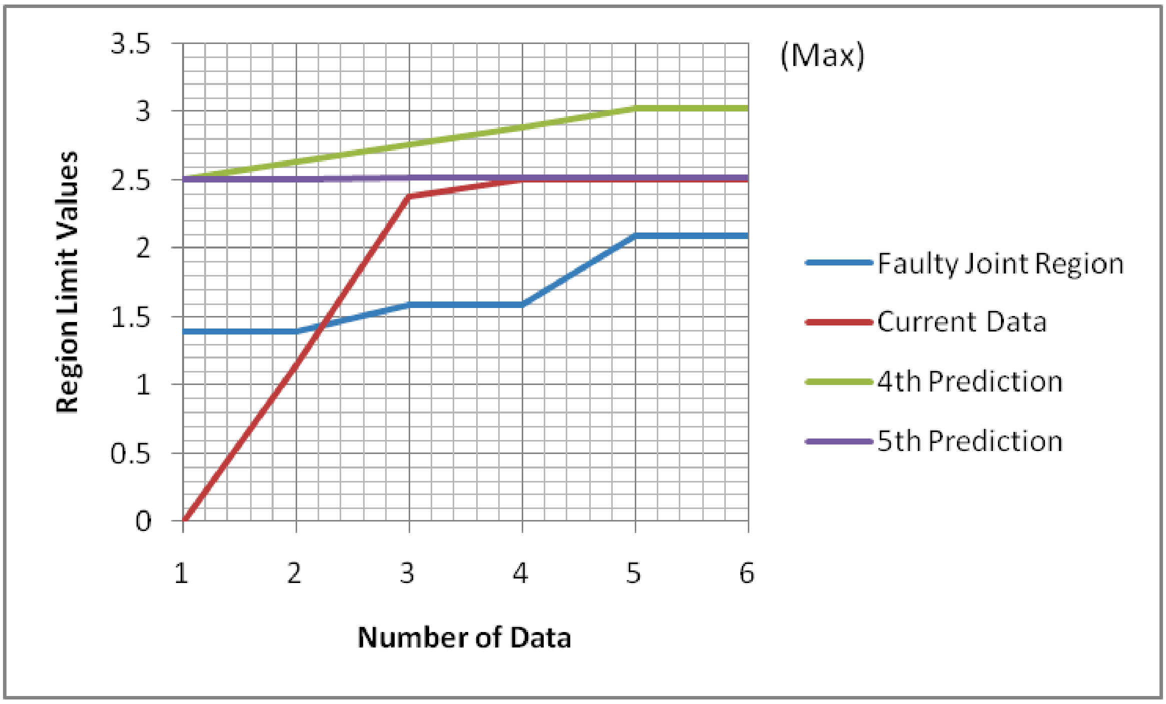

Fig. 6.28 Region Mapping of Joint Data - Downward Experiment

continues to assess the incoming data. The Robot Mind eventually reaches the conclusion that the incoming data is valid, supported by the Prediction data, which consistently shows that the robot is experiencing a specific kind of synthetic pain which worsens if the process is not forced to stop.

Pain Activation Analyses

The Belief State of the robot for the Upward experiment is summarised in Table 6.76. It

Table 6.76 Belief State During Upward Experiment

No	Data		CDV	Region	Incoming Belief	Awareness	Mind State	Extero-ceptive	Proprio-ceptive
	Sensory	Internally							
1	0.901		95	4	Current	Low Objective	Unconstrained	TRUE	FALSE
2	-1.057		95	4	Current	Low Objective	Unconstrained	TRUE	FALSE
3	-2.047		95	4	Current	Low Objective	Unconstrained	TRUE	FALSE
4	-2.084		95	4	Current	Low Objective	Unconstrained	TRUE	FALSE
		-2.121	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.158	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.195	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.232	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
5	-2.1		95	4	Current	Low Objective	Unconstrained	TRUE	FALSE
		-2.116	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.132	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.148	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE
		-2.164	95	4	Prediction	Low Objective	Unconstrained	TRUE	FALSE

Table 6.77 Belief State During Downward Experiment

No	Data		CDV	Region	Incoming Belief	Awareness	Mind State	Extero-ceptive	Proprio-ceptive
	Sensory	Internally							
1	-0.015		119	5	Current	Right Objective	Unconstrained	TRUE	FALSE
2	1.137		119	5	Current	Right Objective	Unconstrained	TRUE	FALSE
3	2.376		119	5	Current	Right Objective	Unconstrained	TRUE	FALSE
4	2.504		119	5	Current	Right Objective	Unconstrained	TRUE	FALSE
		2.632	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		2.760	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		2.888	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		3.016	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
5	2.507		119	5	Current	Right Objective	Unconstrained	TRUE	FALSE
		2.510	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		2.513	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		2.516	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE
		2.519	119	5	Prediction	Right Objective	Unconstrained	TRUE	FALSE

can be seen that the Exteroceptive values are TRUE and Proprioceptive values are FALSE throughout the experiments, which means that the active robot is the Observer and its main tasks are to observe and react empathically reactions towards the other robot (the Mediator).

Information is extracted from the Belief State and utilised to derive a set of recommendations for the Pain Matrix model (see Table 6.78 for the Upward experiment and Table 6.79 for the Downward experiment). Both tables show that during the Current data analyses, the

Table 6.78 Belief State Recommendation During Upward Experiment

No	Data		Belief Recommendations		
	Sensory	Internally	Booster	Pain Init	Consciousness Modifier
1	0.901	N/A	N/A	N/A	N/A
2	-1.057	N/A	N/A	N/A	N/A
3	-2.047	N/A	N/A	N/A	N/A
4	-2.084	N/A	0.000787	2.08434	TRUE
		-2.121	0	2.121	FALSE
		-2.158	0	2.158	FALSE
		-2.195	0	2.195	FALSE
		-2.232	0	2.232	FALSE
5	-2.1	N/A	0.000343	2.10034	TRUE
		-2.116	0	2.116	FALSE
		-2.132	0	2.132	FALSE
		-2.148	0	2.148	FALSE
		-2.164	0	2.164	FALSE

Table 6.79 Belief State Recommendation During Downward Experiment

No	Data		Belief Recommendations		
	Sensory	Internally	Booster	Pain Init	Consciousness Modifier
1	-0.015	N/A	N/A	N/A	N/A
2	1.137	N/A	N/A	N/A	N/A
3	2.376	N/A	N/A	N/A	N/A
4	2.504	N/A	0.00078697	2.504	TRUE
		2.632	0	2.632	FALSE
		2.760	0	2.760	FALSE
		2.888	0	2.888	FALSE
		3.016	0	3.016	FALSE
5	2.507	N/A	0.000343067	2.100343067	TRUE
		2.510	0	2.116	FALSE
		2.513	0	2.132	FALSE
		2.516	0	2.148	FALSE
		2.519	0	2.164	FALSE

Pain Matrix has full authority to alter the Consciousness State of the robot throughout the process, while in the Prediction data, the proposal made by the Pain Matrix is not allowed (achieved by setting the value of the Consciousness Modifier to FALSE). Furthermore, the Booster has no effect on the Prediction data because the value of the Booster is set to 0.

For the Downward experiment, the Pain Matrix Activation has similar results to the Upward experiment. During the Current data, it can be seen that as the result of Belief

Table 6.80 Pain Matrix Activation with Current Data - Upward Experiment

No	Consciousness Modifier	Pain Distribution	Current Kind of Pain	Pain Level	Empathic Actions	Self Actions
1	N/A	N/A	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A
4	Critical	0	Inflammatory	Severe	High	None
		2.121	N/A	N/A	N/A	N/A
		2.158	N/A	N/A	N/A	N/A
		2.195	N/A	N/A	N/A	N/A
		2.232	N/A	N/A	N/A	N/A
5	Critical	0	Inflammatory	Severe	High	None
		2.116	N/A	N/A	N/A	N/A
		2.132	N/A	N/A	N/A	N/A
		2.148	N/A	N/A	N/A	N/A
		2.164	N/A	N/A	N/A	N/A

Table 6.81 Pain Matrix Activation with Prediction Data - Upward Experiment

No	Prediction Kind of Pain	Pain Level	Empathic Actions	Self Actions	Danger Low	Danger Medium	Mind State	Warning
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	N/A	N/A	N/A	N/A	N/A	N/A	Unconstrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	1	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	2	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	3	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	4	Constrained	No Danger
5	N/A	N/A	N/A	N/A	N/A	N/A	Constrained	Next Pain
	Proprioceptive	Slight	Medium	N/A	N/A	1	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	2	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	3	Constrained	No Danger
	Proprioceptive	Slight	Medium	N/A	N/A	4	Constrained	Future Pain

recommendations, the Consciousness Modifier is set to the Critical level, which invokes the kind of IP at a pain level of *Severe* with the Empathic Actions recommendation at High. This situation occurs in the 4th cycle; however, the Pain Matrix assesses that the data is noisy and suggests that no change is required for the Mind State regardless of the fact that the Prediction data consistently predicts that the Mind State should be altered (Prediction invokes the kind of IP at the level of *Slight*). When the 5th cycle is processed, the Current data processing makes the same recommendation as in the previous cycle. The Pain Matrix therefore recommends that the Mind State should be changed to Constrained, while the Warning is at Next Pain. Based on these recommendations, the Robot Mind defines a set of Goals to be achieved, and maps the Goals into Intentions for further follow-up. Table 6.82 and Table 6.83 respectively show the Goals and Intentions of the Robot Mind throughout the experiments. The tables show that at the initial stage, the Goals focus on the Awareness

Table 6.82 Goals and Intentions of Observer During Upward Experiment

No	Mind Recommendation - Goals					Intentions	
	CDV	Region	Awareness	Mind State	Warning	Action	Empathy Delay
1	95	4	Low Objective	Unconstrained	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
5	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	24	1	Upper Subject	Constrained	Next Pain	Right Arm Resist	690.75

of the robot by setting the value of CDV to 95 during the Upward experiment and 119 during the Downward experiment. By doing this, the Robot Mind indirectly modifies the Awareness Region while the Mind State is configured separately (Unconstrained for both experiments). During this stage, there are no proposals for actions to be passed to the Action Execution Engine module by the Intentions. This condition prevails until the final cycle occurs, when the Goals are to alter the Awareness and the Mind State to the maximum level of the subjective element, Upper Subjective Awareness, and to constrain the Robot Mind. This situation modifies the Intentions and the Action Execution Engine modules and prompts them to send out an alert with a message of Right Arm Resist. This message is designed to alert the Mediator to increase the joint stiffness and the human peer to halt the interaction. The Observer then approaches the scene to stop the interaction. The time delay of executing

Table 6.83 Goals and Intentions of Observer During Downward Experiment

No	Mind Recommendation - Goals					Intentions	
	CDV	Region	Awareness	Mind State	Warning	Action	Empathy Delay
1	119	5	Right Objective	Unconstrained	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
	None	None	None	None	None	None	N/A
5	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	24	1	Upper Subject	Constrained	Next Pain	Right Arm Resist	894.23

empathic actions is 690.75 during the Upward experiment and 894.23 during the Downward experiment.

Chapter 7

Conclusion and Future Work

This chapter summarises the overall presentation of the thesis and highlights the findings obtained from the experiments as a proof-of-concept of the proposal of evolving robot empathy through the creation of artificial pain in an Adaptive Self-Awareness Framework (ASAF). The discussion also reveals possible future works that could be undertaken in different implementation domains.

7.1 Outcomes

This section is divided into two subsections. The first subsection briefly presents important discussion prompts and the second summarises the research findings obtained from the experiments.

7.1.1 Discussion Prompts

Two major concepts have been raised in this thesis which stem from the biological mechanism of humans. They are:

1. Self-Awareness and Consciousness
2. Pain and Empathy

From the first concept, we derive the Adaptive Robot Self-Awareness Framework (ASAF) and utilise it as a framework for the robot mind. Robot consciousness is directed towards the ability to focus a robot's attention on two elements of self, namely Subjective and Objective, which refer to the Awareness Region. The robot's awareness elements refer to the body parts as the Subjective element and the robot's goals as the Objective elements. Directing attention is achievable through modification of the Consciousness Direction Value (CDV) which results in the selection of a specific Awareness Region.

The second concept inspires the development of artificial pain, which consists of several classes of synthetic pain. By implementing pain detection and activation mechanisms on the ASAF, robot empathy can be realised. Changing the robot's attention to focus on one specific part of the awareness element, such as the arm joint, allows the robot to develop counter reactions to pain activation. The terms Beliefs, Desires or Goals and Intentions are used to communicate the internal state of the framework. The kinds of counter reactions produced depend on the robot's Beliefs, which inform whether the synthetic pain originates from the robot's own sensory mechanism or from another robot. The Goals are modified based on the reasoning recommendation which uses a causal reasoning approach to analyse the current Joint Data and predict the future data in the form of a sequential data pattern. In any situation where synthetic pain is invoked, the Intentions generate actions that may involve physical actions, which activate the Action Execution Engine or involve only changes in the state of the Robot Mind.

7.1.2 Framework Performance

Using the Adaptive Robot Self-Awareness Framework (ASAF), we are able to demonstrate that a robot is capable of developing accurate pain acknowledgements and appropriate responses. The performance of the framework is measured during collaborative tasks which involve one humanoid robot and one human peer for non-empathy based experiments and two humanoid robots and one human peer for empathy-based experiments. The framework utilises causal reasoning and is capable of analysing and producing reliable robot decisions which embrace past, current and the future considerations. The experiments show that the robot becomes aware of its body part by demonstrating the ability to foresee its future state from the current state. The skills of predicting the consequences of its body behaviour and making proper counter responses in a timely fashion are used to prevent the robot from experiencing synthetic pain. In addition, the ASAF successfully disregards noisy data from the robot's sensory mechanism by proposing amendment data sequences derived from previous data sequence prediction. This allows the robot to robustly generate accurate alternatives to its decision based on self-awareness.

7.1.3 Synthetic Pain Activation

Synthetic pain is comprised of several kinds of pain and the activation for each kind of synthetic pain is embedded in a group of faulty joint settings. If the body part moves into these designated faulty joint regions, the pain detection module captures this occurrence and activates associated kinds of synthetic pain.

The proof-of-concept implements synthetic pain by directing the robot's focus of attention to the subjective elements of the robot's awareness which are accessible in the Awareness Region, and by embodying the consciousness feature through robot body part motions which are integrated with the ASAF.

The experiments show that the robot is able to become aware of its body parts by demonstrating the ability to foresee its future state from its current state. The skills of predicting the consequences of its body behaviour and classifying these consequences into appropriate synthetic pain categories, and at the same time making proper counter-responses in a timely fashion, prevent the robot from experiencing other kinds of pain, such as Inflammatory Pain at a higher level which could lead to significant hardware damage to the robot. Also, the innovative ASAF successfully disregards noisy sensor data from the robot's physical body by proposing amended data sequences derived from previous data sequence prediction. This allows the robot to generate robustly accurate alternatives to its decisions based on self-awareness.

Two kinds of synthetic pain activation mechanisms, Simplified Pain Detection - SPD and Pain Matrix models, were successfully implemented and were relatively accurate in detecting faulty joint visitation occurrences and invoking appropriate kinds of synthetic pain. The experiment results also indicate that the SPD model is more straightforward in terms of synthetic pain generation than the Pain Matrix model, which requires more recommendation data from the robot's Beliefs. However, the Pain Matrix model offers a great deal of functionality such as the empathy response, which is directly acknowledgeable from the Sensory Mechanism. Our findings also suggest that when the Robot Mind explores its High Priority Subjective Awareness region, the state of the Robot Mind does not have to be at the Constrained level. In other words, even though the Robot Mind focuses attention on a specified element of the body, such as the arm, this does not mean that the robot is in a state of pain which requires an increase in the stiffness of the arm joint motor. This finding demonstrates that the ASAF is capable of exploring any of its consciousness regions, including focusing attention on its body parts even when synthetic pain is not being invoked.

7.1.4 Robot Empathy with Synthetic Pain

Our concept of empathy with pain is implemented in our proof-of-concept experiments in which we project the internal state of the human peer into the internal state of the Observer. As the sharing task focuses on the shoulder rotation of the human peer, the Mediator is introduced as a medium that maps the human shoulder region onto the shoulder region of the robot (which shares a uniform physical design with the Observer). Having realised this projection process, the Observer is able to infer what the other robot is experiencing, which allows the robot to eventually produce an accurate description of pain and generate appropriate empathy responses. Data transformation, which processes captured information through robot vision, determines the quality of the projection. The causal reasoning through sequential pattern prediction enables the robot's decision making to embrace past, current and future considerations. This ability allows the robot to build its expectations of the other object's internal state. The Observer's focus of attention switches as the reasoning process predicts the synthetic pain level experienced in the human shoulder.

The experiments show that the Observer is able to project the internal state of the Mediator through the embodiment feature of consciousness, which is expressed by the arm motions captured through the robot's exteroceptive sensor. This projection generally takes place accurately when both the Observer and the Mediator, share uniform internal states, particularly during the upward trend motion in the SPD model. As the Observer's mind focuses on the faulty joint region, the computation time increases and as a result, introduces data analysis discrepancies through the generation of a false alarm. However, the Observer and the Mediator converge to the same result, where the final synthetic pain region is Region 3 and the pain category is Category 2.1. During the downward trend, the final projection produces a significant difference when the Observer mis-projects the Mediator's internal state, as the pain region falls into Region 6 whereas in fact it is still in Region 4.

There are three main causes for this false projection: the limited amount of data to be used in the sequence data prediction process decreases the quality of reasoning, the hardware discrepancies of the arm joint motor motion areas, and the interactions with the human peer which are not uniform throughout the experiments, resulting in a variable speed in hand movement. The Pain Matrix model-based experiments identify that computational time increases, as the model requires more data analyses from various elements of the Pain Matrix, and this contributes to delays in performing empathy reactions. In other words, the Pain Matrix model offers more functionality for the detection and activation of synthetic pain, but its complex computation process causes an increase in the time response. Overall, the SPD model allows the kind of Proprioceptive Pain to stop the flow of the interaction, while in the Pain Matrix model, Proprioceptive Pain will not trigger the activation of the Action Execution Engine; rather, it helps the reasoning process by providing this kind of pain to the Robot Mind so that the Robot Mind can optimise its critical analysis concerning future decision making.

Projection thus takes place accurately when both robots, the Observer and the Mediator, share a unified internal state. As the Observer's mind starts to visit the faulty joint region, the computation time increases and as a result, data analysis discrepancies are introduced through the generation of a false alarm. This false projection has three main causes: (1) the limited amount of data to be used in the sequence data prediction process decreases the quality of the reasoning; (2) there are hardware discrepancies of the arm joint motor motion areas; (3) there is variable speed in the hand movements of the human peer. The experiment also shows that robot awareness may revisit any of its consciousness regions under the Unconstrained condition unless the mind switches to the Constrained condition. The research and experiments in this thesis demonstrate the strong potential of the Adaptive Self-Awareness Framework to be implemented in assistive robot applications where an empathic response generated by synthetic pain is desirable.

7.2 Future Works

Building on this implementation and proof-of-concept work, future research can be divided into two major development projects: those related to the framework and those related to the application domain.

7.2.1 Framework Development

Several possible improvements can be made to the ASAF and its integration into a robot platform, and they are listed below:

1. Integration of various sensors using more sophisticated data integration.
2. Development of a reasoning mechanism that utilises different learning approaches.
3. Exploration of the objective awareness region of the framework to allow the robot to perform a range of various tasks.
4. By manually designing a predetermined artificial dictionary in the framework, an artificial pain dictionary can be captured via human facial expression, which will increase the robustness of the framework.
5. Implementation in various robot joints which will increase the usability of the robot framework.
6. As a proof-of-concept and future adoption for wider research groups, conducting an evaluation of the awareness framework along with the synthetic pain categories in a human-centred experiment.

7.2.2 Application Domain

Having aware that the abilities of understanding pain and reacting with it carry a tremendous advantageous, its realisation could bring a wide open potential applications in human-robot interaction. Its potential applications could be identified as the following.

1. Search and rescue application area. When robots work under a human supervision,the task is less challenging compared to robots work as human peers (proximity). The use of robots in searching and rescuing victims of disasters, and recovering efforts is known as the Urban Search And Rescue (USAR), and this research area is the most prominent research profile of HRI in the United States (Casper and Murphy, 2003). In an emergency search and rescue situation, robot peers should be able to assess the current situation of the surrounding, including victims of a disaster that the robots manage to rescue. The rescue robots should be able to detect and understand in advance the kinds of pain that the human victims are experiencing, then to consider what are the most appropriate reactions. Embedding the proposed self-awareness concept into the robot's framework allows the rescue robots to understand the pain of the victims and develop the most suitable reactions or treatments to be delivered.

2. Military and police purposes. Working as a peer between robots and humans requires a very sophisticated skills such as understanding the human peer when suffers from pain as a result of a gun fire in a joint-patrol task. Accurate assessment of the human peer condition is crucial in taking appropriate counter actions. The ASAF framework could potentially be used to develop the robot's understanding of pain. In addition to that, delivering social cues skills could be equipped into the robot framework such as showing empathy towards the wounded peer.

3. Educational robotics. Another potential application is in educational robotics in which robots function as an assistance in classroom activities or laboratory experiments. The notion of robots in providing educational services has been an active and rapid growing research area in HRI and there have been several studies reported to conduct research works related to the notion (Ahlgren, 2002; Fournier and Riopel, 2007; Jeong-Hye et al., 2009; Kai-Yi et al., 2011; Karna-Lin et al., 2006; Stansbury, 2010). A recent study by Serholt et al. (2014) investigates how empathic technologies could be integrated into the robot tutor capabilities. The study further develops a scenario for an empathic robot tutor which transforms high-level action specification into a concrete set of words and behaviours for the robot to perform (Hall et al., 2016). The proposed ASAF framework would be able to handle the empathic embodiment through the initiation of the empathic actions on the lower level.

4. Entertainment. This area covers different applications of robots, such as storyteller robots (Duc-Minh et al., 2003; Miletitch et al., 2011), dance partner robots (Buondonno et al., 2015; Granados and Kosuge, 2015; Liu et al., 2010), and pet robots (Bharatharaj et al., 2015; Kubota et al., 2001; Toshimitsu et al., 2008). If all these robot types are embedded with the ability to perform empathic reactions through the understanding of pain, the acceptance of robots as partners will increase. As a result, robots may play more roles in human social life.

5. Health care industry. Large amount of studies in the field of robot applications for health care have been conducted and their applications can be classified into two major groups, which is health care robots for children and elder people (Manti et al., 2016; Sharkey and Sharkey, 2011), and assistive care robots for people with disabilities (Khosla et al., 2015; Magnenat-Thalmann and Zhang, 2014; Ranasinghe et al., 2014). These two major groups are in needs to have robots' ability to understand pain and react empathetically towards it. A possible major application of the framework and the synthetic pain proposed in this thesis is in assistive care robots, particularly robots which are used for therapy purposes. We present a brief overview of this potential future application below.

Background

Studies in the field of assistive robotics have growned rapidly for three main reasons: ageing populations, disability and factors related to independence (Mann, 2005). One area in which assistive robotics has attracted attention in the last decade is that of rehabilitation robotics. Until now, studies in rehabilitation robotics have focused on *orthopaedic* shoulder rehabilitation, which concerns the prevention or correction of injuries or disorders of the skeletal system and associated muscles, joints, and ligaments around the human shoulder. Sicuri et al. (2014) classify rehabilitative robots into two categories, each of which is comprised of the one or more of the groups described below:

1. Control Strategy is divided into:

    - Passive, which means that the robot controls the motion of the patient's arm.
    - Active Unassisted, which means that the patient performs actions without the robot's help.
    - Active Assisted which means that the patient performs actions with the assistance of the robot (assistance is voluntarily but inadequate or limited).

- Resistive, which requires the patient to exercise pressure against a force generated by the robot.

2. Mechanical Characteristics are divided into:

   - Exoskeletons, which are typically designed to align their mechanical joints to human limbs so that an articular decoupling and a good coverage of the whole arm can be achieved.
   - End Effectors, which are designed to connect to the patient at a single point, which restricts the patient-machine interaction at the end-effector level.
   - Cable-driven, which is designed to use cables with an attached end-effector which is held by a fixed frame support. This mechanism requires the patient's forearm to be fixed into a splint and the stimuli are sent through the upper limb by pulling on the cables.

Possible Application

Given the current limited technology in rehabilitative robots, the current focus of studies is only on the hardware design of the robot. During the therapy process, humans may suffer from pain as a result of the motion of the shoulder in a specific area. By detecting the pain in advance and generating preventive reactions, the robot would greatly reduce the possibility of the patient experiencing pain during therapy exercises.

Implementation of the Proposal

The following are crucial points related to the implementation of the proposal:

- Customisation is required, in this case, in relation to the human shoulder area. This stage is categorised as the teaching phase. A possible approach during this stage would be to model the patient's facial expressions when the shoulder motion provokes pain. These sets of faulty joint region data are stored in the database of the robot framework - ASAF.

- An improvement in the therapy is measured by the change in the faulty joint area which is automatically monitored by the Robot Mind.

- The application is limited to shoulder rehabilitation, however the framework is reconfigurable, which will possibly lead to the next step of improvement in the application of the ASAF.

Significance of the Outcomes

Our proposal makes the following significant contributions through the implementation of the Adaptive Self-Awareness Framework into robot applications for rehabilitation purposes.

- It develops the Mind Infrastructure for robots used in rehabilitation, which is of great assistance to healthcare providers
- It promotes a new approach to robot therapy in the way that robots interact with patients during the rehabilitation process
- It emphasises the need to incorporate the awareness concept in robots used for rehabilitation so that when assistance is delivered, these robots can develop a sense of awareness of the patient's pain.

References

Ackerman, E. (2013). Hexapod figures out how to walk after you chop its leg off. Technical report, IEEE Spectrum.

Agha-Mohammad, A.-A., Ure, N. K., How, J. P., and Vian, J. (2014). Health-aware stochastic planning for persistent package delivery missions using quadrotors. In *Proceedings of IEEE International Conference on Intelligent Robot Systems (IROS)*.

Agrawal, R. and Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the Eleventh International Conference on Data Engineering*, pages 3–14. IEEE.

Ahlgren, D. J. (2002). Meeting educational objectives and outcomes through robotics education. In *Automation Congress, Proceedings of the 5th Biannual World*, volume 14, pages 395–404.

Alami, R., Fleury, S., Herrb, M., Ingrand, F., and Robert, F. (1998). Multi-robot cooperation in the MARTHA project. *Robotics and Automation Magazine, IEEE*, 5(1):36–47.

Aldebaran (2006). Aldebaran documentation, https://www.aldebaran.com/en.

Anshar, M. and Williams, M.-A. (2007). Extended evolutionary fast learn-to-walk approach for four-legged robots. *Journal of Bionic Engineering*, 4(4):255–263.

Anshar, M. and Williams, M.-A. (2015). Evolving synthetic pain into adaptive self-awareness framework for robots. In *Annual International Conference on Biologically Inspired Cognitive Architectures Sixth Annual Meeting of the BICA Society*. BICA Society.

Avenanti, A., Paluello, I. M., Bufalari, I., and Aglioti, S. M. (2006). Stimulus-driven modulation of motor-evoked potentials during observation of others' pain. *NeuroImage*, 32(1):316–324.

Banissy, M. J., Kanai, R., Walsh, V., and Rees, G. (2012). Inter-individual differences in empathy are reflected in human brain structure. *NeuroImage*, 62(3):2034–2039.

Barandiaran, X. E., Di Paolo, E., and Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5):367–386.

Barrett-Lennard, G. T. (1986). The relationship inventory now: Issues and advances in theory, method, and use. In Pinsof, L. S. G. W. M., editor, *The Psychotherapeutic Process: A Research Handbook*, Guilford Clinical Psychology and Psychotherapy Series., pages 439–476, New York. Guilford Press.

Bharatharaj, J., Huang, L., and Al-Jumaily, A. (2015). Bio-inspired therapeutic pet robots: Review and future direction. In *10th International Conference on Information, Communications and Signal Processing (ICICS)*, pages 1–5.

Birlo, M. and Tapus, A. (2011). The testing crucial role of robot self-awareness in human-robot interaction. pages 115–116.

Blum, A. L. and Furst, M. L. (1997). Fast planning through planning graph analysis. *Artificial intelligence*, 90(1):281–300.

Bonet, B. and Geffner, H. (1999). Planning as heuristic search: New results in: Recent advances in ai planning. *Proc. 5th European Conference on Planning, Lecture Notes in Artificial Intelligence*, 1809:359–371.

Bonet, B. and Geffner, H. (2001). Planning as heuristic search. *Artificial Intelligence*, 129(1-2):5–33.

Bongard, J., Zykov, V., and Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science*, 314(5802):1118–1121.

Bordini, R. H. and Hübner, J. F. (2006). BDI agent programming in AgentSpeak using Jason. In *Computational Logic in Multi-Agent Systems*, Lecture Notes in Artificial Intelligence 1193, pages 143–164. Springer.

Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press.

Bratman, M. E., Israel, D., and Pollack, M. E. (1988). Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355.

Braubach, L., Pokahr, A., Moldt, D., and Lamersdorf, W. (2005). Goal representation for BDI agent systems. In *Programming Multi-Agent Systems*, Lecture Notes in Computer Science 3346, pages 44–65. Springer.

Brooks, R. (1991). Intelligence without reason. In *International Joint Conferences on Artificial Intelligence, IJCAI 91*, pages 569–595. Morgan Kaufmann.

Buondonno, G., Patota, F., Wang, H., Luca, A. D., and Kosuge, K. (2015). A model predictive control approach for the partner ballroom dance robot. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 774–780.

Busetta, P., Rönnquist, R., Hodgson, A., and Lucas, A. (1999). JACK intelligent agents-components for intelligent agents in Java. *AgentLink News Letter*, 2(1):2–5.

Byrne, J. H. and Dafny, N. (1997). Neuroscience online: An electronic textbook for the neurosciences. http://nba.uth.tmc.edu/neuroscience/.

Casper, J. and Murphy, R. R. (2003). Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 33(3):367–385.

Cervero, F. and Merskey, H. (1996). What is a noxious stimulus? In *Pain Forum*, volume 5, pages 157–161. Elsevier.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–219.

Chen, J. (2011). History of pain theories. *Neuroscience Bulletin*, 27(5):343–350.

Cossentino, M., Chella, A., Lodato, C., Lopes, S., Ribino, P., and Seidita, V. (2012). A notation for modeling Jason-like BDI agents. In *Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), 2012*, pages 12–19. IEEE.

Crick, F. (1994). *The Astonishing Hypothesis: The Scientific Search for the Soul*. Scribners, NY.

Cuff, B. M., Brown, S. J., Taylor, L., and Howat, D. J. (2014). Empathy: A review of the concept. *Emotion Review*, page 1754073914558466.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1):113–126.

De Santis, A., Siciliano, B., De Luca, A., and Bicchi, A. (2008). An atlas of physical human–robot interaction. *Mechanism and Machine Theory*, 43(3):253–270.

De Vignemont, F. and Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, 10(10):435–441.

Decety, J. and Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2):71–100.

Descartes, R. and Olscamp, P. J. (2001). *Discourse on Method, Optics, Geometry, and Meteorology*. Indianapolis, IN : Hackett Pub.

Dictionary.com Online Dictionary (2015). *Self-Awareness*. Dictionary.com.

Duc-Minh, P., Trong-Nhan, D.-N., Phuc-Thinh, N.-V., and Minh-Triet, T. (2003). Smart Teddy Bear a vision-based story teller. In *International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 257–262.

Edelkamp, S. (2000). Planning with pattern databases. In Cesta, I. A., editor, *Sixth European Conference on Planning (ECP'01)*, Recent Advances in AI Planning, pages 13–24. AAAI Press.

Edelkamp, S. (2002). Symbolic pattern databases in heuristic search planning. In *AIPS*, pages 274–283.

Edelkamp, S. (2014). Planning with pattern databases. In Cesta, I. A., editor, *Sixth European Conference on Planning (ECP'01)*, Recent Advances in AI Planning, pages 13–24. AAAI Press.

Ehrenfeld, S. and Butz, M. V. (2012). Autonomous failure detection and multimodal sensor fusion in a modular arm model. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2186–2191.

Elliott Fahlman, S. (1974). A planning system for robot construction tasks. *Artificial Intelligence*, 5(1):1–49.

Emerson, E. A. and Srinivasan, J. (1988). Branching time temporal logic. In *Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency*, pages 123–172. Springer.

Erol, K., Hendler, J., and Nau, D. S. (1994). HTN planning: Complexity and expressivity. 94:1123–1128.

Erol, K., Hendler, J., and Nau, D. S. (1996). Complexity results for HTN planning. *Annals of Mathematics and Artificial Intelligence*, 18(1):69–93.

Fagiolini, A., Valenti, G., Pallottino, L., Dini, G., and Bicchi, A. (2007). Decentralized intrusion detection for secure cooperative multi-agent systems. In *46th IEEE Conference on Decision and Control*, pages 1553–1558.

Fikes, R. E. and Nilsson, N. J. (1972). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2(3):189–208.

Finger, S. (1994). Pain. In *Origins of Neuroscience: A History of Explorations into Brain Function*, pages 148–164. Oxford University Press.

Firby, R. J. (1987). An investigation into reactive planning in complex domains. In *Proceedings of the Sixth National Conference on Artificial Intelligence*, volume 1, pages 202–206. Seattle, WA.

Fitzpatrick, P. (2003). Perception and perspective in robotics. Technical report, DTIC Document.

Fournier, F. and Riopel, M. (2007). Innovative technologies for learning in science laboratories. In *Advanced Learning Technologies. ICALT 2007. Seventh IEEE International Conference on*, pages 497–498.

Franklin, S. and Graesser, A. (1997). Is it an agent, or just a program? A taxonomy for autonomous agents. In Müller, J., Wooldridge, M., and Jennings, N., editors, *Intelligent Agents III Agent Theories, Architectures, and Languages*, Lecture Notes in Computer Science 1193, pages 21–35. Springer Berlin Heidelberg.

Gamez, D. (2008). Progress in machine consciousness. *Consciousness and Cognition*, 17(3):887–910.

Garcia-Larrea, L. and Peyron, R. (2013). Pain matrices and neuropathic pain matrices: A review. *PAIN®*, 154, Supplement 1(0):S29–S43.

Georgeff, M. and Pell (1989). Decision-making in an embedded reasoning system. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 972–978. Morgan Kaufmann.

Georgeff, M., Pell, B., Pollack, M., Tambe, M., and Wooldridge, M. (1999). The belief-desire-intention model of agency. In *Intelligent Agents V: Agents Theories, Architectures, and Languages*, pages 1–10. Springer.

Gold, K. and Scassellati, B. (2009). Using probabilistic reasoning over time to self-recognize. *Robotics and Autonomous Systems*, 57(4):384–392.

Goldie, P. (1999). How we think of others' emotions. *Mind & Language*, 14(4):394–423.

Golombek, R., Wrede, S., Hanheide, M., and Heckmann, M. (2010). Learning a probabilistic self-awareness model for robotic systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2745–2750.

Gottifredi, S., Tucat, M., Corbatta, D., García, A. J., and Simari, G. R. (2008). A BDI architecture for high level robot deliberation. In *XIV Congreso Argentino de Ciencias de la Computación, 6-8 October*. Red de Universidades con Carreras en Informática (RedUNCI), Argentina.

Gottifredi, S., Tucat, M., Corbatta, D., García, A. J., and Simari, G. R. (2010). A BDI architecture for high level robot deliberation. *Inteligencia Artificial*, 46:74–83.

Goubert, L., Craig, K. D., Vervoort, T., Morley, S., Sullivan, M. J. L., Williams, A. C. d. C., Cano, A., and Crombez, G. (2005). Facing others in pain: The effects of empathy. *Pain*, 118(3):285–288.

Granados, D. F. P. and Kosuge, K. (2015). Design of a male-type dance partner robot for leading a physical human-robot interaction. In *IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 1234–1240.

Haikonen, P. O. (2012). Consciousness and robot sentience. In *Series on Machine Consciousness*, volume 2. World Scientific.

Hall, L., Hume, C., Tazzyman, S., Deshmukh, A., Janarthanam, S., Hastie, H., Aylett, R., Castellano, G., Papadopoulos, F., Jones, A., Corrigan, L. J., Paiva, A., Oliveira, P. A., Ribeiro, T., Barendregt, W., Serholt, S., and Kappas, A. (2016). Map reading with an empathic robot tutor. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 567–567.

Hanna, T. (1991). Somatics. *Journal of Behavioural Optometry*, 2(2):31–35.

Harold Merskey, D. and Bogduk, N. (1994). *Classification of Chronic Pain Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms*. 2nd edition. International Association for the Study of Pain Press, Seattle.

Hart, J. W. and Scassellati, B. (2011). A robotic model of the ecological self. In *11th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 682–688. IEEE.

Hashimoto, M., Kawashima, H., Nakagami, T., and Oba, F. (2001). Sensor fault detection and identification in dead-reckoning system of mobile robot: Interacting multiple model approach. In *Proceedings of 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 1321–1326.

Helmert, M. (2006). The fast downward planning system. *J. Artif. Intell. Res.(JAIR)*, 26:191–246.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33(3):307–316.

Hsu, M. C., Schubiner, H., Stracks, J. S., and Clauw, D. J. (2010). Sustained pain reduction through affective self-awareness in fibromyalgia: A randomized controlled trial. *Journal of General Internal Medicine*, 25(10):1064–1070.

Hummel, B., Kammel, S., Dang, T., Duchow, C., and Stiller, C. (2006). Vision-based path-planning in unstructured environments. In *Intelligent Vehicles Symposium, 2006 IEEE*, pages 176–181.

Hyun, D. J., Seok, S., Lee, J., and Kim, S. (2014). High speed trot-running: Implementation of a hierarchical controller using proprioceptive impedance control on the MIT Cheetah. *The International Journal of Robotics Research*, 33(11):1417–1445.

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, 24(3):771–779.

Jackson, P. L., Rainville, P., and Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain*, 125(1–2):5–9.

Jaerock, K. and Yoonsuck, C. (2008). Internal state predictability as an evolutionary precursor of self-awareness and agency. In *7th IEEE International Conference on Development and Learning (ICDL)*, pages 109–114.

Jain, D., Mosenlechner, L., and Beetz, M. (2009). Equipping robot control programs with first-order probabilistic reasoning capabilities. In *IEEE International Conference on Robotics and Automation, ICRA '09*, pages 3626–3631.

Jeong-Hye, H., Dong-Ho, K., and Jong-Won, K. (2009). Physical learning activities with a teaching assistant robot in elementary school music class. In *Fifth International Joint Conference on INC, IMS and IDC (NCM '09)*, pages 1406–1410.

Jolliffe, D. and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence*, 29(4):589–611.

Jorgensen, M. W., Ostergaard, E. H., and Lund, H. H. (2004). Modular atron: Modules for a self-reconfigurable robot. In *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, volume 2, pages 2068–2073.

Jung-Min, Y. (2003). Fault-tolerant gait generation for locked joint failures. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2237–2242.

Kai-Yi, C., Chin-Hsien, W., and Zeng-Wei, H. (2011). A humanoid robot as a teaching assistant for primary education. In *Fifth International Conference on Genetic and Evolutionary Computing (ICGEC)*, pages 21–24.

Kaminka, G. and Tambe, M. (1998). Social comparison for failure detection and recovery. In Singh, M., Rao, A., and Wooldridge, M., editors, *Intelligent Agents IV Agent Theories, Architectures, and Languages*, Lecture Notes in Computer Science 1365, pages 127–141. Springer Berlin Heidelberg.

Kappenman, E. S. and Luck, S. J. (2011). ERP components: The ups and downs of brainwave recordings. Oxford University Press.

Karna-Lin, E., Pihlainen-Bednarik, K., Sutinen, E., and Virnes, M. (2006). Can robots teach? preliminary results on educational robotics in special education. In *Sixth International Conference on Advanced Learning Technologies*, pages 319–321.

Kautz, H. and Selman, B. (1992). Planning as satisfiability. In Neumann, B., editor, *Proceedings of the 10th European Conference on Artificial Intelligence*, pages 359–363. John Wiley & Sons.

Kautz, H. and Selman, B. (1999). Unifying sat-based and graph-based planning. In *IJCAI*, volume 99, pages 318–325.

Kautz, H. and Selman, B. (2006). Satplan04: Planning as satisfiability. *Working Notes on the Fifth International Planning Competition (IPC-2006)*, pages 45–46.

Khosla, R., Nguyen, K., and Chu, M. T. (2015). Service personalisation of assistive robot for autism care. In *41st Annual Conference of the IEEE Industrial Electronics Society (IECON)*, pages 002088–002093.

Kitamura, T. (1998). Animal-like behavior design of small robots by consciousness-based architecture. *Advanced Robotics*, 12(3):289–307.

Kitamura, T., Tahara, T., and Asami, K.-I. (2000). How can a robot have consciousness? *Advanced Robotics*, 14(4):263–275.

Koditschek, D. E. (1992). Task encoding: Toward a scientific paradigm for robot planning and control. *Robotics and Autonomous Systems*, 9(1):5–39.

Koehler, J., Nebel, B., Hoffmann, J., and Dimopoulos, Y. (1997). *Extending planning graphs to an ADL subset*, pages 273–285. Springer Berlin Heidelberg, Berlin, Heidelberg.

Koos, S., Cully, A., and Mouret, J.-B. (2013). Fast damage recovery in robotics with the t-resilience algorithm. Cornell University Library.

Kowalski, R. and Sadri, F. (1996). An agent architecture that unifies rationality with reactivity. In *Proceedings of Logic and Databases, Italy*. Citeseer.

Krings, T., Buchbinder, B., Butler, W., Chiappa, K., Jiang, H., Cosgrove, G., and Rosen, B. (1997). Functional magnetic resonance imaging and transcranial magnetic stimulation complementary approaches in the evaluation of cortical motor function. *Neurology*, 48(5):1406–1416.

Kubota, N., Kojima, F., and Fukuda, T. (2001). Self-consciousness and emotion for a pet robot with structured intelligence. In *Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569)*, pages 2786–2791 vol.5.

Laird, P. (1993). Identifying and using patterns in sequential data. In *Proceedings of the 4th International Workshop on Algorithmic Learning Theory (ALT '93')*, pages 1–18. Springer-Verlag.

Lamm, C., Decety, J., and Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, 54(3):2492–2502.

Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., and Paiva, A. (2011). Social robots in learning environments: A case study of an empathic chess companion. In *Proceedings of the International Workshop on Personalization Approaches in Learning Environments*, volume 732, pages 8–12.

Lewis, M. (1991). Ways of knowing: Objective self-awareness or consciousness. *Developmental Review*, 11(3):231–243.

Liu, Z., Chenglong, F., Chen, K., Koike, Y., Hirata, Y., and Kosuge, K. (2010). A new control algorithm for a passive type dance partner robot. In *IEEE Conference on Robotics, Automation and Mechatronics*, pages 486–491.

Loggia, M. L., Mogil, J. S., and Catherine Bushnell, M. (2008). Empathy hurts: Compassion for another increases both sensory and affective components of pain perception. *Pain*, 136(1–2):168–176.

Long, D. and Fox, M. (1999). Efficient implementation of the plan graph in stan. *J. Artif. Intell. Res.(JAIR)*, 10:87–115.

Magnenat-Thalmann, N. and Zhang, Z. (2014). Assistive social robots for people with special needs. In *International Conference on Contemporary Computing and Informatics (IC3I)*, pages 1374–1380.

Mann, W. C. (2005). *Smart Technology for Aging, Disability, and Independence: The State of the Science*. John Wiley & Sons.

Manti, M., Pratesi, A., Falotico, E., Cianchetti, M., and Laschi, C. (2016). Soft assistive robot for personal care of elderly people. In *6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 833–838.

Marier, J. S., Rabbath, C. A., and Lechevin, N. (2013). Health-aware coverage control with application to a team of small UAVs. *IEEE Transactions on Control Systems Technology*, 21(5):1719–1730.

McDermott, D. (1992). Robot planning. *AI Magazine*, 13(2):55.

Mehrabian, A. (1996). Manual for the balanced emotional empathy scale (BEES). Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA, USA 93940.

Mehrabian, A. and Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40(4):525–543.

Melzack, R. and Wall, P. D. (1996). Pain mechanisms: A new theory. *Pain Forum*, 5(1):3–11.

Meng, J., Jackson, T., Chen, H., Hu, L., Yang, Z., Su, Y., and Huang, X. (2013). Pain perception in the self and observation of others: An ERP investigation. *NeuroImage*, 72(0):164–173.

Merriam-Webster Online Dictionary (2015). *Self-Awareness*. Merriam Webster.

Mesulam, M.-M. (1998). From sensation to cognition. *Brain*, 121(6):1013–1052.

Michel, P., Gold, K., and Scassellati, B. (2004). Motion-based robotic self-recognition. In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, volume 3, pages 2763–2768.

Miletitch, R., Sabouret, N., and Ochs, M. (2011). From socio-emotional scenarios to expressive virtual narrators. In *IEEE Workshop on Affective Computational Intelligence (WACI)*, pages 1–8.

Moayedi, M. and Davis, K. D. (2013). Theories of pain: From specificity to gate control. *Journal of Neurophysiology*, 109(1):5–12.

Morgenstern, L. and Stein, L. A. (1988). Why things go wrong: A formal theory of causal reasoning. In *Association for the Advancement of Artificial Intelligence*, pages 518–523.

Moseley, G. L. and Arntz, A. (2007). The context of a noxious stimulus affects the pain it evokes. *Pain*, 133(1):64–71.

Novianto, R. and Williams, M. A. (2009). The role of attention in robot self-awareness. In *The 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, pages 1047–1053.

Oxford Online Dictionary (2015). *Self-Awareness*. Oxford Dictionaries.

Padgham, L. and Lambrix, P. (2005). Formalisations of capabilities for BDI-agents. *Autonomous Agents and Multi-Agent Systems*, 10(3):249–271.

Paredis, C. J. J. and Khosla, P. K. (1995). Global trajectory planning for fault tolerant manipulators. In *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots'*, volume 2, pages 428–434 vol.2.

Pereira, A., Leite, I., Mascarenhas, S., Martinho, C., and Paiva, A. (2011). Using empathy to improve human-robot relationships. In Lamers, M. and Verbeek, F., editors, *Human-Robot Personal Relationships*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 59, pages 130–138. Springer Berlin Heidelberg.

Perl, E. R. (2007). Ideas about Pain, A historical view. *Nature Reviews Neuroscience*, 8(1):71–80.

Preston, S. D. and De Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(01):1–20.

Pütten, A. M. R.-V. D., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., Maderwald, S., Brand, M., and Krämer, N. C. (2013). Neural correlates of empathy towards robots. In *Proceedings of the 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 215–216.

Pütten, A. M. R.-V. D., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., Maderwald, S., Brand, M., and Krämer, N. C. (2014). Investigations on empathy towards humans and robots using functional magnetic resonance imaging (fMRI). In *Computers in Human Behaviour*, volume 33, pages 201–212.

Ralph, S. K. and Pai, D. K. (1997). Fault tolerant locomotion for walking robots. In *Proceedings of the 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA'97)*, pages 130–137.

Ranasinghe, R., Dantanarayana, L., Tran, A., Lie, S., Behrens, M., and Liu, L. (2014). Smart hoist: An assistive robot to aid carers. In *13th International Conference on Control Automation Robotics & Vision (ICARCV)*, pages 1285–1291.

Rao, A. S. and Georgeff, M. P. (1991). Modeling rational agents within a BDI-architecture. *Knowledge Representation*, 91:473–484.

Regan, D. (2000). *Human Perception of Objects: Early Visual Processing of Spatial Form Defined by Luminance, Color, Texture, Motion, and Binocular Disparity*. Sinauer Associates, Sunderland, MA.

Reggia, J. A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, 44:112–131.

Reisberg, D. (2013). *The Oxford Handbook of Cognitive Psychology*. Oxford University Press.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131–141.

Sacerdoti, E. D. (1975). *The Nonlinear Nature of Plans*, pages 206–214. Morgan Kaufmann Publishers Inc.

Salter, T., Michaud, F., Letourneau, D., Lee, D. C., and Werry, I. P. (2007). Using proprioceptive sensors for categorizing human-robot interactions. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*, pages 105–112.

Scassellati, B. (2002). Theory of mind for a humanoid robot. *Autonomous Robots*, 12(1):13–24.

Scheutz, M. and Kramer, J. (2007). Reflection and reasoning mechanisms for failure detection and recovery in a distributed robotic architecture for complex robots. In *IEEE International Conference on Robotics and Automation*, pages 3699–3704.

Schwind, C. (1999). A theory for causal reasoning. In Burgard, W., Cremers, A., and Cristaller, T., editors, *KI-99: Advances in Artificial Intelligence*, Lecture Notes in Computer Science 1701, pages 285–288. Springer Berlin Heidelberg.

Serholt, S., Barendregt, W., Leite, I., Hastie, H., Jones, A., Paiva, A., Vasalou, A., and Castellano, G. (2014). Teachers' views on the use of empathic robotic tutors in the classroom. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 955–960.

Sharkey, A. and Sharkey, N. (2011). Children, the elderly, and interactive robots. *IEEE Robotics & Automation Magazine*, 18(1):32–38.

Shear, J. (1995). *Explaining Consciousness: The Hard Problem*. MIT Press.

Sicuri, C., Porcellini, G., and Merolla, G. (2014). Robotics in shoulder rehabilitation. *Muscles, Ligaments and Tendons Journal*, 4(2):207.

Siegel, M. (2001). Smart sensors and small robots. In *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference, 2001 (IMTC)*, volume 1, pages 303–308.

Sievers, S., Ortlieb, M., and Helmert, M. (2010). Efficient implementation of pattern database heuristics for classical planning. In *SOCS*.

Silva, F., Alexandre Castilho, M., and Allan Künzle, L. (2000). *Petriplan: A New Algorithm for Plan Generation (Preliminary Report)*, pages 86–95. Springer Berlin Heidelberg, Berlin, Heidelberg.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661):1157–1162.

Singh, D., Sardina, S., and Padgham, L. (2010). Extending BDI plan selection to incorporate learning from experience. *Robotics and Autonomous Systems*, 58(9):1067–1075.

Singh, D., Sardina, S., Padgham, L., and James, G. (2011). Integrating learning into a BDI agent for environments with changing dynamics. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Volume Three*, pages 2525–2530. AAAI Press.

Soika, M. (1997). A sensor failure detection framework for autonomous mobile robots. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '97)*, volume 3, pages 1735–1740.

Spong, M. W., Hutchinson, S., and Vidyasagar, M. (2006). *Robot Dynamics and Control*. John & Wiley Sons, US, 1st edition.

Stahl, B. C., McBride, N., Wakunuma, K., and Flick, C. (2014). The empathic care robot: A prototype of responsible research and innovation. *Technological Forecasting and Social Change*, 84:74–85.

Stansbury, R. S. (2010). Enhancing student learning in artificial intelligence using robotics. In *Proceedings of the IEEE SoutheastCon 2010 (SoutheastCon)*, pages 168–171.

Steen, E. and Haugli, L. (2001). From pain to self-awareness — a qualitative analysis of the significance of group participation for persons with chronic musculoskeletal pain. *Patient Education and Counseling*, 42(1):35–46.

Stein, L. A. and Morgenstern, L. (1994). Motivated action theory: A formal theory of causal reasoning. *Artificial Intelligence*, 71(1):1–42.

Stueber, K. (2014). Empathy. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Standford University.

Sudeikat, J., Braubach, L., Pokahr, A., Lamersdorf, W., and Renz, W. (2007). Validation of BDI agents. In Bordini, R., Dastani, M., Dix, J., and Seghrouchni, A., editors, *Programming Multi-Agent Systems*, Lecture Notes in Computer Science 4411, pages 185–200. Springer Berlin Heidelberg.

Sussman, G. J. (1973). A computational model of skill acquisition. Technical Report Note AI TR-297, Massachusetts Institute of Technology.

Takeno, J. (2012). *Creation of a Conscious Robot: Mirror Image Cognition and Self-Awareness*. CRC Press.

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

Tashani, O. A. and Johnson, M. I. (2010). Avicenna's concept of pain. *The Libyan Journal of Medicine*, 5:10.3402/ljm.v5i0.5253.

Tate, A. (1977). *Generating project networks*, pages 888–893. Morgan Kaufmann.

Thangarajah, J., Harland, J., Morley, D., and Yorke-Smith, N. (2011). Operational behaviour for executing, suspending, and aborting goals in BDI agent systems. In Omicini, A., Sardina, S., and Vasconcelos, W., editors, *Declarative Agent Languages and Technologies VIII*, Lecture Notes in Computer Science 6619, pages 1–21. Springer Berlin Heidelberg.

Toshimitsu, H., Hiroki, O., Kazuya, I., Joji, M., Hisashi, O., Yoshihito, K., and Tomomi, H. (2008). Robot therapy as for recreation for elderly people with dementia - game recreation using a pet-type robot. In *The 17th IEEE International Symposium on Robot and Human Interactive Communication(RO-MAN)*, pages 174–179.

Tosunoglu, S. (1995). Fault-tolerant control of mechanical systems. In *Proceedings of the 1995 IEEE IECON 21st International Conference on Industrial Electronics, Control, and Instrumentation*, volume 1, pages 127–132.

Vidal, V. (2004). A lookahead strategy for heuristic search planning. In *ICAPS*, pages 150–160.

Vidal, V. (2011). Yahsp2: Keep it simple, stupid. *The*, pages 83–90.

Vidal, V. and Geffner, H. (2006). Branching and pruning: An optimal temporal pocl planner based on constraint programming. *Artificial Intelligence*, 170(3):298–335.

Watanabe, Y. and Yuta, S. (1990). Position estimation of mobile robots with internal and external sensors using uncertainty evolution technique. In *Proceedings of the 1990 IEEE International Conference on Robotics and Automation*, volume 3, pages 2011–2016.

Weld, D. S. (1999). Recent advances in ai planning. *AI magazine*, 20(2):93.

Woolf, C. J. (2010). What is this thing called pain? *The Journal of Clinical Investigation*, 120(11):3742.

Woolf, C. J. and Ma, Q. (2007). Nociceptors—noxious stimulus detectors. *Neuron*, 55(3):353–364.

Yan, H., Ang, MarceloH, J., and Poo, A. (2014). A survey on perception methods for human–robot interaction in social robots. *International Journal of Social Robotics*, 6(1):85–119.

Yi, Z., Jun, Z., Weiwei, L., Ying, H., and Jianwei, Z. (2012). Memory-based hierarchical task and motion planning. In *International Conference on Information and Automation (ICIA), 2012*, pages 412–417.

Zagal, J. C. and Lipson, H. (2009). Self-reflection in evolutionary robotics: Resilient adaptation with a minimum of physical exploration. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, pages 2179–2188. ACM.

Zhao, X. and Son, Y. J. (2008). Extended BDI framework for modelling human decision-making in complex automated manufacturing systems. *International Journal of Modelling and Simulation*, 28(3):347–356.

Zhu, S. (2014). Pain expression recognition based on pLSA model. *The Scientific World Journal, url:http://www.hindawi.com/journals/tswj/2014/736106/*, 2014.

Ziemke, T., Jirenhed, D.-A., and Hesslow, G. (2005). Internal simulation of perception: A minimal neuro-robotic model. *Neurocomputing*, 68:85–104.

Appendix A

Terminology

Qualia : The qualities of the world and body sensations which are obtained through perception (noun plural of quale).
Epiphenomenalism : The view that physical events cause mental events in the brain, although these mental events have no effect upon physical events.
Artificial functional pain: The process of emulating the functional elements of pain.
Synthetic: Not natural or man-made.
Consciousness: Signifies the cognitive focus, or the focus of attention. It should not be understood to mean human consciousness
Cognitive: Involving conscious mental activity such as thinking, learning, remembering, and understanding.
Pain matrix: A collection of elements in the human brain that accommodate the cycle of pain, from detection, recognition and acknowledgement to actions to counter the pain.
Proprioceptive: Perception which monitors and gathers information internally.
Exteroceptive: Perception which monitors and captures information from the external world.
Empathy: The ability to acknowledge, understand and share the internal state of others.
Stimulus: A thing or event that evokes a specific functional reaction.
Fault Tolerance: An area of robot control which focus on the continuation of system functionality continue in the event of hardware failure.
Noxious stimuli: Actual or potential damage to tissue which is liable to cause pain.
Nociceptor: Sensory perceptors which are responsible for capturing noxious stimuli.
Fibromyalgia: A widespread pain and tenderness in the human body which is sometimes accompanied by fatigue, cognitive disturbance and emotional distress.
Somatics: The field of study concerning the human body (soma) from self-perspective (the first person view).
Static: A situation that lacks of action or is unchanged.

Projection: The fusion of an observer's internal state and the object of the observer's perception, which occurs internally in the observer.

Appendix B

Documentation

This appendix presents the specifications of the NAO H25 robot, obtained from the official website of NAO Aldebaran, http://doc.aldebaran.com/1-14/family/nao_h25/. It provides descriptions of the hardware related to the experiments in Chapter 5 and Chapter 6.

B.1 Dimensions

Table B.1 Body Dimensions

Robot Version	Height (mm)	Depth (mm)	Width (mm)
V3.2	573.2	290	273.3
V3.3	573	311	275

B.2 Links

The robot links contain several elements such as the head, arms and legs.

Table B.2 Link and Axis Definitions

Element	Position
Torso	- 126.50 Z (mm) from HeadYaw
X Axis	Positive towards NAO's Front
Y Axis	From Right to Lelft
Z Axis	Vertical

Table B.3 Head Definition

From ...	To ...	V3.2. similar to V3.3 X (mm)	Y (mm)	Z (mm)
Torso	HeadYaw	0	0	126.5
HeadYaw	HeadPitch	0	0	0
	Main length (mm)			
	NeckOffsetZ	126.5		

Table B.4 Arm Definition

From ...	To ...	X (mm) V3.2	X (mm) V3.3	Y (mm) V3.2	Y (mm) V3.3	Z (mm) V3.2	Z (mm) V3.3
Torso	LShoulderPitch	0	0	98	98	100	100
LShoulderPitch	LShoulderRoll	0	0	0	0	0	0
LShoulderRoll	LElbowYaw	90	105	0	15	0	0
LElbowYaw	LElbowRoll	0	0	0	0	0	0
LElbowRoll	LWristYaw	50.55	55.95	0	0	0	0
	Main length (mm)	**V3.2**	**V3.3**				
	ShoulderOffsetY	98	98				
	ElbowOffsetY	-	15				
	UpperArmLength	90	105				
	LowerArmLength	50.55	55.95				
	ShoulderOffsetZ	100	100				
	HandOffsetX	58	57.75				
	HandOffsetZ	15.9	12.31				

Table B.5 Leg Definition

From ...	To ...	X (mm) V3.2	X (mm) V3.3	Y (mm) V3.2	Y (mm) V3.3	Z (mm) V3.2	Z (mm) V3.3
Torso	LHipYawPitch	0	0	50	50	-85	-85
LHipYawPitch	LHipRoll	0	0	0	0	0	0
LHipRoll	LHipPitch	0	0	0	0	0	0
LHipPitch	LKneePitch	0	0	0	0	-100	-100
LKneePitch	LAnklePitch	0	0	0	0	-102.9	-102.9
LAnklePitch	LAnkleRoll	0	0	0	0	0	0

	Main length (mm) V3.2	Main length (mm) V3.3
HipOffsetZ	85	85
HipOffsetY	50	50
ThighLength	100	100
TibiaLength	102.9	102.9
FootHeight	45.11	45.19

B.3 Joints and Motors

The robot joints contain several elements such as the head joints, right arm and left arm joints, pelvis joints, left leg and right leg joints. Each of these joints is driven by one motor.

Table B.6 Head Joints

Motion range (V3.2 similar to V3.3)			
Joint name	**Motion**	**Range** *degrees*	*radians*
HeadYaw	Head joint twist (Z)	-119.5 to 119.5	-2.0857 to 2.0857
HeadPitch	Head joint front and back (Y)	-38.5 to 29.5	-0.6720 to 0.5149

Anti collision limitation					
HeadYaw V3.2 similar to V3.3		**HeadPitch Min V3.2 similar to V3.3**		**HeadPitch Max V3.2 similar to V3.3**	
degrees	*radians*	*degrees*	*radians*	*degrees*	*radians*
-119.52	-2.086017	-25.73	-0.4491	18.91	0.330041
-87.49	-1.526988	-18.91	-0.33	11.46	0.200015
-62.45	-1.089958	-24.64	-0.43	17.19	0.300022
-51.74	-0.903033	-27.5	-0.48	18.91	0.330041
-43.32	-0.756077	-31.4	-0.548	21.2	0.37001
-27.85	-0.486074	-38.5	-0.672	24.18	0.422021
0	0	-38.5	-0.672	29.51	0.515047
27.85	0.486074	-38.5	-0.672	24.18	0.422021
43.32	0.756077	-31.4	-0.548	21.2	0.37001
51.74	0.903033	-27.5	-0.48	18.91	0.330041
62.45	1.089958	-24.64	-0.43	17.19	0.300022
87.49	1.526988	-18.91	-0.33	11.46	0.200015
119.52	2.086017	-25.73	-0.4491	18.91	0.330041

Table B.7 Left Arm Joints

Joint name	Motion	Range (degrees) V3.2	Range (degrees) V3.3	Range (radians) V3.2	Range (radians) V3.3
LShoulderPitch	Left shoulder joint front and back (Y)	-119.5 to 119.5		-2.0857 to 2.0857	
LShoulderRoll	Left shoulder joint right and left (Z)	0.5 to 94.5	-18 to 76	0.0087 to 1.6494	-0.3142 to 1.3265
LElbowYaw	Left shoulder joint twist (X)	-119.5 to 119.5		-2.0857 to 2.0857	
LElbowRoll	Left elbow joint (Z)	-89.5 to -0.5	-88.5 to -2	-1.5621 to -0.0087	-1.5446 to -0.0349
LWristYaw	Left wrist joint (X)	-104.5 to 104.5		-1.8238 to 1.8238	
LHand	Left hand	Open and Close			

Table B.8 Right Arm Joints

Joint name	Motion	Range (degrees) V3.2	Range (degrees) V3.3	Range (radians) V3.2	Range (radians) V3.3
RShoulderPitch	Right shoulder joint front and back (Y)	-119.5 to 119.5		-2.0857 to 2.0857	
RShoulderRoll	Right shoulder joint right and left (Z)	-94.5 to -0.5	-76 to 18	-1.6494 to -0.0087	-1.3265 to 0.3142
RElbowYaw	Right shoulder joint twist (X)	-119.5 to 119.5		-2.0857 to 2.0857	
RElbowRoll	Right elbow joint (Z)	0.5 to 89.5 89.5	2 to 88.5	0.0087 to 1.5621	0.0349 to 1.5446
RWristYaw	Right wrist joint (X)	-104.5 to 104.5		-1.8238 to 1.8238	
RHand	Right hand	Open and Close			

Table B.9 Pelvis Joints

Joint name	Motion (Y-Z 45°)	Range V3.2 Similar to V3.3 degrees	Range V3.2 Similar to V3.3 radians
LHipYawPitch	Left hip joint twist	-65.62 to 42.44	-1.145303 to 0.740810
RHipYawPitch	Right hip joint twist	-65.62 to 42.44	-1.145303 to 0.740810

*LHipYawPitch and RHipYawPitch are physically just one motor so they cannot be controlled independently.

Table B.10 Left Leg Joints

Motion range (V3.2 similar to V3.3)			
Joint name	**Motion**	**Range** **degrees**	**radians**
LHipRoll	Left hip joint right and left (X)	-21.74 to 45.29	-0.379472 to 0.790477
LHipPitch	Left hip joint front and back (Y)	-88.00 to 27.73	-1.535889 to 0.48409
LKneePitch	Left knee joint (Y)	-5.29 to 121.04	-0.092346 to 2.112528
LAnklePitch	Left ankle joint front and back (Y)	-68.15 to 52.86	-1.189516 to 0.922747
LAnkleRoll	Left ankle joint right and left (X)	-22.79 to 44.06	-0.397880 to 0.769001

Anti collision limitation V3.2 similar to V3.3					
LAnklePitch		**LAnkleRoll Min**		**LAnkleRoll Max**	
degrees	***radians***	***degrees***	***radians***	***degrees***	***radians***
-68.15	-1.189442	-2.86	-0.049916	4.3	0.075049
-48.13	-0.840027	-10.31	-0.179943	9.74	0.169995
-40.11	-0.700051	-22.8	-0.397935	12.61	0.220086
-25.78	-0.449946	-22.8	-0.397935	44.06	0.768992
5.73	0.100007	-22.8	-0.397935	44.06	0.768992
20.05	0.349938	-22.8	-0.397935	31.54	0.550477
52.87	0.922755	0	0	2.86	0.049916

Table B.11 Right Leg Joints

Motion range (V3.2 similar to V3.3)			
Joint name	Motion	Range degrees	Range radians
RHipRoll	Right hip joint right and left (X)	-45.29 to 21.74	-0.790477 to 0.379472
RHipPitch	Right hip joint front and back (Y)	-88.00 to 27.73	-1.535889 to 0.484090
RKneePitch	Right knee joint (Y)	-5.90 to 121.47	-0.103083 to 2.120198
RAnklePitch	Right ankle joint front and back (Y)	-67.97 to 53.4	-1.186448 to 0.932056
RAnkleRoll	Right ankle joint right and left (X)	-44.06 to 22.8	-0.768992 to 0.397935

Anti collision limitation V3.2 similar to V3.3					
RAnklePitch		**RAnkleRoll Min**		**RAnkleRoll Max**	
degrees	*radians*	*degrees*	*radians*	*degrees*	*radians*
-68.15	-1.189442	-4.3	-0.075049	2.86	0.049916
-48.13	-0.840027	-9.74	-0.169995	10.31	0.179943
-40.11	-0.700051	-12.61	-0.220086	22.8	0.397935
-25.78	-0.449946	-44.06	-0.768992	22.8	0.397935
5.73	0.100007	-44.06	-0.768992	22.8	0.397935
20.05	0.349938	-31.54	-0.550477	22.8	0.397935
52.87	0.922755	-2.86	-0.049916	0	0

Table B.12 Motors and Speed Ratio

Motors	Motor Type 1	Motor Type 2	Motor Type 3	Motor Type 4
Model	RE-Max 24	RE-Max 17	A-max12	GM20
No load speed	8 000 rpm	11 900 rpm	12 300 rpm	13 206 rpm
Stall torque	59.5 mNm	15,1 mNm	1.52 mNm	0.08 mNm
Nominal Torque	12.3 mNm	3.4 mNm	0.931 mNm	0.08 mNm
Speed Reduction ratio				
Type A	201.3	150.27	800	372
Type B	130.85	173.22		

Table B.13 Head and Arms

	Head			
Joints	**V3.2**		**V3.3**	
	Motor	**Reduction ratio**	**Motor**	**Reduction ratio**
HeadYaw	Type 2	Type A	Type 2	Type A
HeadPitch	Type 2	Type B	Type 2	Type B
		Arms		
ShoulderPitch	Type 2	Type A	Type 2	Type A
ShoulderRoll	Type 2	Type B	Type 2	Type B
ElbowYaw	Type 2	Type A	Type 2	Type A
ElbowRoll	Type 2	Type B	Type 2	Type B

Table B.14 Hands and Legs

	Hands			
Joints	**V3.2**		**V3.3**	
	Motor	**Reduction ratio**	**Motor**	**Reduction ratio**
WristYaw	Type 3	Type A	Type 2	Type C
Hand	Type 4	Type A	Type 2	Type D
		Legs		
HipYawPitch	Type 1	Type A	Type 1	Type A
HipRoll	Type 1	Type A	Type 1	Type A
HipPitch	Type 1	Type B	Type 1	Type B
KneePitch	Type 1	Type B	Type 1	Type B
AnklePitch	Type 1	Type B	Type 1	Type B
AnkleRoll	Type 1	Type A	Type 1	Type A

Table B.15 Camera Resolution

Elements	**Specifications**	**Resolution**	**Output Format**
Sensor Model	OV7670		
Camera output	VGA@30fps (YUV422 color space)	640 x 480	
Field of view	58°DFOV (47.8°HFOV, 36.8°VFOV)		YUV422
Focus range	30cm infinity	at 30 frame/s	
Focus type	Fixed focus		

Table B.16 Camera Position

Head Camera V3.2 similar to V3.3						
Camera name	**X(m)**	**Y(m)**	**Z(m)**	**WX(rd) [deg]***	**WY(rd) [deg]***	**WZ(rd) [deg]***
CameraTop	0.0539	0	0.0679	0	0	0
CameraBottom	0.0488	0	0.02381	0	0.6981 [40.0]	0

Table B.17 Joint Sensor and Processor

Elements	Details
Joint Position Sensor	12 bit 4096/turn = 0.1 precicion
Processor	x86 AMD GEODE 500MHz CPU 256 MB SDRAM / 2 GB flash memory

Table B.18 Microphone and Loudspeaker

Microphone						
Micro name	**X(m)**		**Y(m)**		**Z(m)**	
	V3.2	**V3.3**	**V3.2**	**V3.3**	**V3.2**	**V3.3**
MicroFront	0.041	0.0489	0	0	0.0915	0.076
MicroRear	-0.0577	-0.046	0	0	0.0693	0.0814
MicroLeft	-0.0195	-0.0195	0.0606	0.0606	0.0331	0.0331
MicroRight	-0.0195	-0.0195	-0.0606	-0.0606	0.0331	0.0331
Electrical Bandpass : [300Hz - 8kHz]						
Loudspeaker						
Left	0.0038		0.0453		0.0526	
Right	0.0038		-0.0453		0.0526	

Appendix C

Experiment Results Appendix

This appendix presents the additional results of experiments which cover only the Non-Empathy based Excperiment, particularly the SPD-based (see Table C.1 for the overall experiment result classification).

Table C.1 Experiment Overview-Appendix

Experiments	**Pain Activation**	**Scenario**			
		Offline		**Online**	
		No Shared Task	**Shared Task Hand Push Direction**	**No Shared Task**	**Shared Task Hand Push Direction**
Non-empathy	SPD-based	✓	Horizontal	✓	Horizontal
	Pain Matrix-based	x	x	✓	Vertical
Empathy	SPD-based	x	x	x	Vertical
	Pain Matrix-based	x	x	x	Vertical

C.1 Non-Empathy Appendix

Two sub experiments: SPD-based and Pain Matrix-based, each of which contains the online and offline experiments.

****

C.1.1 SPD-based Appendix

Table C.2 Offline without Human Interaction Trial 1 with Prediction Data

No	Elbow Data	Time	Prediction Data						
			1st	2nd	3rd	4th	5th	6th	7th
1	0.00873	221.72							
2	0.00873	222.24							
3	0.00873	222.75							
4	0.00873	223.28	0.00873						
5	0.00873	223.84	0.00873	0.00873					
6	0.00873	224.40	0.00873	0.00873	0.00873				
7	0.00873	224.96	0.00873	0.00873	0.00873	0.00873			
8	0.00873	225.51	0.00873	0.00873	0.00873	0.00873	0.00873		
9	0.00873	226.07		0.00873	0.00873	0.00873	0.00873	0.00873	
10	0.00873	226.63			0.00873	0.00873	0.00873	0.00873	0.00873
11	0.00873	227.19				0.00873	0.00873	0.00873	0.00873
12	0.00873	227.76					0.00873	0.00873	0.00873
13	0.00873	228.32						0.00873	0.00873
14	0.00873	228.88							0.00873
15	0.00873	229.44							
16	0.00873	230.01							
17	0.00873	230.58							
18	0.00873	231.14							
19	0.00873	231.71							
20	0.00873	232.28							
21	0.00873	232.85							

Table C.3 Offline without Human Interaction Trial 2 with Prediction Data

No	Elbow Data	Time	Prediction Data						
			1st	2nd	3rd	4th	5th	6th	7th
1	0.00873	316.95							
2	0.00873	317.47							
3	0.00873	317.98							
4	0.00873	318.52	0.00873						
5	0.00873	319.07	0.00873	0.00873					
6	0.00873	319.63	0.00873	0.00873	0.00873				
7	0.00873	320.19	0.00873	0.00873	0.00873	0.00873			
8	0.00873	320.75	0.00873	0.00873	0.00873	0.00873	0.00873		
9	0.00873	321.31		0.00873	0.00873	0.00873	0.00873	0.00873	
10	0.00873	321.86			0.00873	0.00873	0.00873	0.00873	0.00873
11	0.00873	322.42				0.00873	0.00873	0.00873	0.00873
12	0.00873	322.98					0.00873	0.00873	0.00873
13	0.00873	323.54						0.00873	0.00873
14	0.00873	324.11							0.00873
15	0.00873	324.67							
16	0.00873	325.24							
17	0.00873	325.80							
18	0.00873	326.37							
19	0.00873	326.95							
20	0.00873	327.52							
21	0.00873	328.10							

Table C.4 Offline without Human Interaction Trial 3 with Prediction Data

No	Elbow Data	Time	Prediction Data						
			1st	2nd	3rd	4th	5th	6th	7th
1	0.00873	385.60							
2	0.00873	386.12							
3	0.00873	386.64							
4	0.00873	387.16	0.00873						
5	0.00873	387.72	0.00873	0.00873					
6	0.00873	388.30	0.00873	0.00873	0.00873				
7	0.00873	388.85	0.00873	0.00873	0.00873	0.00873			
8	0.00873	389.41	0.00873	0.00873	0.00873	0.00873	0.00873		
9	0.00873	389.97		0.00873	0.00873	0.00873	0.00873	0.00873	
10	0.00873	390.54			0.00873	0.00873	0.00873	0.00873	0.00873
11	0.00873	391.10				0.00873	0.00873	0.00873	0.00873
12	0.00873	391.66					0.00873	0.00873	0.00873
13	0.00873	392.22						0.00873	0.00873
14	0.00873	392.78							0.00873
15	0.00873	393.35							
16	0.00873	393.92							
17	0.00873	394.49							
18	0.00873	395.05							
19	0.00873	395.62							
20	0.00873	396.19							
21	0.00873	396.76							

Table C.5 Offline without Human Interaction Trial 4 with Prediction Data

No	Elbow Data	Time	Prediction Data						
			1st	2nd	3rd	4th	5th	6th	7th
1	0.00873	449.60							
2	0.00873	450.12							
3	0.00873	450.64							
4	0.00873	451.16	0.00873						
5	0.00873	451.72	0.00873	0.00873					
6	0.00873	452.28	0.00873	0.00873	0.00873				
7	0.00873	452.84	0.00873	0.00873	0.00873	0.00873			
8	0.00873	453.40	0.00873	0.00873	0.00873	0.00873	0.00873		
9	0.00873	453.95		0.00873	0.00873	0.00873	0.00873	0.00873	
10	0.00873	454.51			0.00873	0.00873	0.00873	0.00873	0.00873
11	0.00873	455.08				0.00873	0.00873	0.00873	0.00873
12	0.00873	455.65					0.00873	0.00873	0.00873
13	0.00873	456.21						0.00873	0.00873
14	0.00873	456.78							0.00873
15	0.00873	457.35							
16	0.00873	457.91							
17	0.00873	458.48							
18	0.00873	459.04							
19	0.00873	459.60							
20	0.00873	460.18							
21	0.00873	460.74							

Table C.6 Offline without Human Interaction Trial 5 with Prediction Data

No	Elbow Data	Time	Prediction Data						
			1st	2nd	3rd	4th	5th	6th	7th
1	0.00873	514.15							
2	0.00873	514.67							
3	0.00873	515.18							
4	0.00873	515.71	0.00873						
5	0.00873	516.28	0.00873	0.00873					
6	0.00873	516.84	0.00873	0.00873	0.00873				
7	0.00873	517.39	0.00873	0.00873	0.00873	0.00873			
8	0.00873	517.96	0.00873	0.00873	0.00873	0.00873	0.00873		
9	0.00873	518.51		0.00873	0.00873	0.00873	0.00873	0.00873	
10	0.00873	519.07			0.00873	0.00873	0.00873	0.00873	0.00873
11	0.00873	519.64				0.00873	0.00873	0.00873	0.00873
12	0.00873	520.21					0.00873	0.00873	0.00873
13	0.00873	520.77						0.00873	0.00873
14	0.00873	521.33							0.00873
15	0.00873	521.89							
16	0.00873	522.46							
17	0.00873	523.02							
18	0.00873	523.60							
19	0.00873	524.16							
20	0.00873	524.73							
21	0.00873	525.30							

Table C.7 Offline with Human Interaction Trial 1 with Prediction Data

No	Elbow Data	Time	Prediction Data					
			1st	2nd	3rd	4th	5th	6th
1	0.00873	583.66						
2	0.02765	584.17						
3	0.14884	584.69						
4	0.34519	585.22	0.34519					
5	0.57836	585.76	0.54154	0.57836				
6	0.78238	586.31	0.73789	0.81153	0.78238			
7	1.02782	586.85	0.93424	1.04470	0.98640	1.02782		
8	1.30701	587.40	1.13059	1.27787	1.19042	1.27326	1.30701	
9	1.51870	587.94		1.51104	1.39444	1.51870	1.58620	1.51870
10	1.56207	588.48			1.59846	1.76414	1.86539	1.73039
11	1.56207	589.03						
12	1.56207	589.56						
13	1.56207	590.08						
14	1.56207	590.63						
15	1.56207	591.18						
16	1.56207	591.74						
17	1.56207	592.30						
18	1.56207	592.85						
19	1.56207	593.40						
20	1.56207	593.95						

Table C.8 Offline with Human Interaction Trial 2 with Prediction Data

No	Elbow Data	Time	Prediction Data					
			1st	2nd	3rd	4th	5th	6th
1	0.02919	644.54						
2	0.02919	645.05						
3	0.03072	645.56						
4	0.03072	646.09						
5	0.13043	646.62						
6	0.45871	647.14	0.45871					
7	0.73023	647.66	0.78699	0.73023				
8	0.94959	648.19	1.11527	1.00175	0.94959			
9	1.14441	648.72	1.44355	1.27327	1.16895	1.14441		
10	1.37297	649.25	1.77183	1.54479	1.38831	1.33923	1.37297	
11	1.53251	649.77		1.81631	1.60767	1.53405	1.60153	1.53251
12	1.56207	650.30			1.82703	1.72887	1.83009	1.69205
13	1.56207	650.83				1.92369	2.05865	1.85159
14	1.56207	651.38					2.28721	2.01113
15	1.56207	651.93					amend	2.17067
16	1.56207	652.48						
17	1.56207	653.04						
18	1.56207	653.59						
19	1.56207	654.14						
20	1.56207	654.70						

Table C.9 Offline with Human Interaction Trial 3 with Prediction Data

No	Elbow Data	Time	Prediction Data					
			1st	2nd	3rd	4th	5th	6th
1	0.02919	726.37						
2	0.02919	726.89						
3	0.04606	727.40						
4	0.22861	727.93						
5	0.40348	728.46	0.40348					
6	0.60444	728.98	0.57835	0.60444				
7	0.88669	729.50	0.75322	0.80540	0.88669			
8	1.08765	730.03	0.92809	1.00636	1.16894	1.08765		
9	1.25485	730.55	1.10296	1.20732	1.45119	1.28861	1.25485	
10	1.42359	731.08		1.40828	1.73344	1.48957	1.42205	1.42359
11	1.56012	731.60			2.01569	1.69053	1.58925	1.59233
12	1.56207	732.14				1.89149	1.75645	1.76107
13	1.56207	732.66					1.92365	1.92981
14	1.56207	733.19						2.09855
15	1.56207	733.74						
16	1.56207	734.29						
17	1.56207	734.85						
18	1.56207	735.40						
19	1.56207	735.96						
20	1.56207	736.51						

Table C.10 Offline with Human Interaction Trial 4 with Prediction Data

No	Elbow Data	Time	Prediction Data					
			1st	2nd	3rd	4th	5th	6th
1	0.02919	773.30						
2	0.02765	773.82						
3	0.02919	774.34						
4	0.02919	774.86						
5	0.07214	775.38	0.07214					
6	0.36974	775.92	0.11509	0.36974				
7	0.74096	776.44	0.15804	0.66734	0.74096			
8	1.06924	776.96	0.20099	0.96494	1.11218	1.06924		
9	1.50643	777.49	0.24394	1.26254	1.48340	1.39752	1.50643	
10	1.56207	778.02		1.56014	1.85462	1.72580	1.94362	
11	1.56207	778.54			2.22584	2.05408	2.38081	
12	1.56207	779.07				2.38236	2.81800	
13	1.56207	779.62					3.25519	
14	1.56207	780.17						
15	1.56207	780.72						
16	1.56207	781.28						
17	1.56207	781.83						
18	1.56207	782.38						
19	1.56207	782.93						
20	1.56207	783.49						

Table C.11 Offline with Human Interaction Trial 4 with Prediction Data

No	Elbow Data	Time	Prediction Data										
			1st	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**11th**	**12th**
1	0.02919	823.77											
2	0.02765	824.29											
3	0.03072	824.81											
4	0.14884	825.33	0.14884										
5	0.28997	825.88	0.26696	0.28997									
6	0.44183	826.42	0.38508	0.43110	0.44183								
7	0.59370	826.96	0.50320	0.57223	0.59369	0.59370							
8	0.73176	827.51	0.62132	0.71336	0.74555	0.74557	0.73176						
9	0.87289	828.06		0.85449	0.89741	0.89744	0.86982	0.87289					
10	1.01402	828.60			1.04927	1.04931	1.00788	1.01402	1.01402				
11	1.13980	829.15				1.20118	1.14594	1.15515	1.15515	1.13980			
12	1.27480	829.69					1.28400	1.29628	1.29628	1.26558	1.27480		
13	1.39445	830.24						1.43741	1.43741	1.39136	1.40980	1.39445	
14	1.50643	830.79							1.57854	1.51714	1.54480	1.51410	1.50643
15	1.56207	831.34								1.64292	1.67980	1.63375	1.61841
16	1.56207	831.89									1.81480	1.75340	1.73039
17	1.56207	832.42										1.87305	1.84237
18	1.56207	832.94											1.95435
19	1.56207	833.50											
20	1.56207	834.06											

Table C.12 Online without Human Interaction Trial 1 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd	3rd
1	0.02765	367.27			
2	0.02765	367.79			
3	0.02765	368.30			
4	0.02765	368.83	0.02765		
5	0.02765	369.38	0.02765	0.02765	
6	0.02765	369.94	0.02765	0.02765	0.02765
			0.02765	0.02765	0.02765
			0.02765	0.02765	0.02765
				0.02765	0.02765
					0.02765

Table C.13 Online without Human Interaction Trial 2 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st
1	0.02765	551.84	
2	0.02765	552.36	
3	0.02612	552.87	
4	0.02765	553.4	
5	0.02765	553.93	
6	0.02765	554.46	0.02765
			0.02765
			0.02765
			0.02765
			0.02765

Table C.14 Online without Human Interaction Trial 3 with Prediction Data

No	Elbow Data	Time	Prediction Data			
			1st	2nd	3rd	4th
1	0.02765	793.73				
2	0.02765	794.25				
3	0.02765	794.76				
4	0.02765	795.28	0.02765			
5	0.02612	795.84	0.02765	0.02612		
6	0.02765	796.38	0.02765	0.02765	0.02765	
7	0.02612	796.94	0.02765	0.02918	0.02918	0.02612
			0.02765	0.03071	0.03071	0.02765
				0.03224	0.03224	0.02918
					0.03377	0.03071
						0.03224

Table C.15 Online without Human Interaction Trial 4 with Prediction Data

No	Elbow Data	Time	Prediction Data		
			1st	2nd	3rd
1	0.0277	971.52			
2	0.0277	972.04			
3	0.0277	972.55			
4	0.0277	973.08	0.02765		
5	0.0277	973.64	0.02765	0.02765	
6	0.0277	974.2	0.02765	0.02765	0.02765
			0.02765	0.02765	0.02765
			0.02765	0.02765	0.02765
				0.02765	0.02765
					0.02765

Table C.16 Online without Human Interaction Trial 5 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd
1	0.02612	354.21		
2	0.02612	354.73		
3	0.02765	355.24		
4	0.02765	355.77		
5	0.02612	356.3		
6	0.02612	356.82	0.02612	
7	0.02612	357.38	0.02612	0.02612
			0.02612	0.02612
			0.02612	0.02612
			0.02612	0.02612
				0.02612

Table C.17 Online with Human Interaction Trial 1 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd	3rd	4th	5th
1	0.22247	38.88					
2	0.26696	39.40					
3	0.37127	39.91					
4	0.49246	40.44	0.49246				
5	0.63205	41.00	0.61365	0.63205			
6	0.78852	41.54	0.73484	0.77164	0.78852		
7	0.95572	42.10	0.85603	0.91123	0.94499	0.95572	
8	1.32695	44.37	0.97722	1.05082	1.10146	1.12292	1.32695
				1.19041	1.25793	1.29012	1.69818
					1.41440	1.45732	2.06941
						1.62452	2.44064
							2.81187

Table C.18 Online with Human Interaction Trial 2 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd	3rd
1	0.02765	776.46			
2	0.02765	776.98			
3	0.02919	777.49			
4	0.21940	778.01			
5	0.39735	778.54	0.39735		
6	0.68421	779.11	0.57530	0.68421	
7	1.30548	781.36	0.75325	0.97107	1.30548
			0.93120	1.25793	1.92675
			1.10915	1.54479	2.54802
				1.83165	3.16929
					3.79056

Table C.19 Online with Human Interaction Trial 3 with Prediction Data

No	Elbow Data	Time	Prediction Data				
			1st	2nd	3rd	4th	5th
1	0.02765	267.12					
2	0.02765	267.64					
3	0.02765	268.15					
4	0.06907	268.68					
5	0.29917	269.20					
6	0.52774	269.73	0.52774				
7	0.71642	270.28	0.75631	0.71642			
8	0.87902	270.84	0.98488	0.90510	0.87902		
9	1.04470	271.39	1.21345	1.09378	1.04162	1.04470	
10	1.41132	273.68	1.44202	1.28246	1.20422	1.21038	1.41132
				1.47114	1.36682	1.37606	1.77794
					1.52942	1.54174	2.14456
						1.70742	2.51118
							2.87780

****

C.1.2 Pain Matrix-based Appendix

Table C.20 Online with Human Interaction Trial 4 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd	3rd	4th
1	0.0261	550.05				
2	0.0277	550.57				
3	0.0292	551.08				
4	0.2286	551.61	0.22861			
5	0.4695	552.16	0.42803	0.46945		
6	0.6919	552.71	0.62745	0.71029	0.69188	
7	1.2288	554.99	0.82687	0.95113	0.91431	1.22878
			1.02629	1.19197	1.13674	1.76568
				1.43281	1.35917	2.30258
					1.58160	2.83948
						3.37638

Table C.21 Online with Human Interaction Trial 5 with Prediction Data

No	Elbow Data	Time	Prediction Data 1st	2nd	3rd
1	0.02765	855.20			
2	0.03839	856.34			
3	0.29150	856.86			
4	0.53234	857.39	0.53234		
5	0.74403	857.94	0.77318	0.74403	
6	1.25639	860.23	1.01402	0.95572	1.25639
			1.25486	1.16741	1.76875
			1.49570	1.37910	2.28111
				1.59079	2.79347
					3.30583

Table C.22 Pain Matrix Without Human Interaction Appendix

No	Data	Prediction Data						
		4	**5**	**-**	**-**	**24**	**25**	**26**
1	-0.0153							
2	-0.0153							
3	-0.0153							
4	-0.0153	-0.0153						
5	-0.0153	-0.0153	-0.0153					
6		-0.0153	-0.0153					
7		-0.0153	-0.0153					
8			-0.0153					
-								
-								
24						-0.0153		
25						-0.0153	-0.0153	
26						-0.0153	-0.0153	-0.0153
						-0.0153	-0.0153	-0.0153
							-0.0153	-0.0153
								-0.0153

Table C.23 Pain Matrix Without Human Interaction Incoming Belief Appendix

No	Data		CDV	Region	Incoming Belief	Awareness	Consciousness State	Exteroceptive	Proprioceptive
	Sensory	Internally							
1	-0.0153		33	5	Current	Lower Objective	Unconstrained	FALSE	TRUE
2	-0.0153		33	5	Current	Lower Objective	Unconstrained	FALSE	TRUE
3	-0.0153		33	5	Current	Lower Objective	Unconstrained	FALSE	TRUE
4	-0.0153		33	5	Current	Lower Objective	Unconstrained	FALSE	TRUE
		-0.0153	33	5	Prediction	Lower Objective	Unconstrained	FALSE	TRUE
		-0.0153	33	5	Prediction	Lower Objective	Unconstrained	FALSE	TRUE
		-0.0153	33	5	Prediction	Lower Objective	Unconstrained	FALSE	TRUE
		-0.0153	33	5	Prediction	Lower Objective	Unconstrained	FALSE	TRUE

Table C.24 Pain Matrix Without Human Interaction SPD Recommendation

Pain Matrix Recommendation								
	Activation				Current			
Booster	Pain Init	Cons Modifier	Consc Modifier	Pain Dist	Kind of Pain	Pain level	empathised actions	self actions
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	-	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	-	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	-	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	-	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	-	N/A	N/A	N/A	N/A	N/A

Pain Matrix Recommendation									
	Activation					Prediction			
Booster	Pain Init	Cons Modifier	pain level	empathy action	self action	danger low	danger medium	Warning	Consc State
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0.0010293	0.016329273	N/A	N/A	-	-	-	-	N/A	N/A
0.0010293	0.016329273	N/A	N/A	-	-	-	-	N/A	N/A
0.0010293	0.016329273	N/A	N/A	-	-	-	-	N/A	N/A
0.0010293	0.016329273	N/A	N/A	-	-	-	-	N/A	N/A

Table C.25 Pain Matrix Without Human Interaction SPD Goals

Activation			Mind Recommendation - Goals				
Booster	**Pain Init**	**Cons Modifier**	**CDV**	**Region**	**Awareness**	**Consciousness State**	**Warning**
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A
0.001029	0.016329	N/A	N/A	N/A	N/A	N/A	N/A