

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 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	xv
List of Tables	xvii
1 Introduction	1
1.1 Overview of the Study Background	1
1.2 Current Issues	2
1.3 Description of Proposed Approach	2
1.4 Brief Description of Experiments	4
1.5 Contributions and Significance	4
1.6 Future Development	5
1.7 Structure of Thesis	5
2 Robot Planning and Robot Cognition	7
2.1 Motion Planning	7
2.1.1 Stimulus-based Planning	7
2.1.2 Reasoning-based Planning	9
2.2 Robot Cognition	13
2.2.1 Discussion on Theories of Mind	14
2.2.2 Self-Awareness	17
2.2.3 Empathy with the Experience of Pain	20
2.2.4 Robot Empathy	25
3 Perceptions, Artificial Pain and the Generation of Robot Empathy	29
3.1 Perceptions	30
3.1.1 Proprioception and Exteroception	31
3.2 Faulty Joint Setting Region and Artificial Pain	31
3.2.1 Proprioceptive Pain (PP)	32
3.2.2 Inflammatory Pain (IP)	33

3.2.3	Sensory Malfunction Pain (SMP)	33
3.3	Pain Level Assignment	35
3.4	Synthetic Pain Activation in Robots	36
3.4.1	Simplified Pain Detection (SPD)	37
3.4.2	Pain Matrix (PM)	39
3.5	Generation of Robot Empathy	42
3.5.1	Empathy Analysis	43
4	Adaptive Self-Awareness Framework for Robots	45
4.1	Overview of Adaptive Self-Awareness Framework for Robots	45
4.1.1	Consciousness Direction	45
4.1.2	Synthetic Pain Description	46
4.1.3	Robot Mind	47
4.1.4	Database	50
4.1.5	Atomic Actions	50
4.2	Reasoning Mechanism	51
4.2.1	Pattern Data Acquisition	51
4.2.2	Causal Reasoning	53
5	Integration and Implementation	57
5.1	Hardware Description	57
5.2	Experiment	57
5.2.1	Non-empathic Experiment	59
5.2.2	Empathic Experiment	60
5.3	Pre-defined Values	61
6	Results, Analysis and Discussion	65
6.1	Experiment Overview	65
6.2	Non-empathy based Experiments	67
6.2.1	SPD-based Model	67
6.2.2	Pain Matrix-based Model	127
6.3	Empathy-based Experiments	131
6.3.1	SPD Model	132
6.3.2	Pain Matrix Model	137
7	Conclusion and Future Work	145
7.1	Outcomes	145

7.1.1	Discussion Prompts	145
7.1.2	Framework Performance	146
7.1.3	Synthetic Pain Activation	146
7.1.4	Robot Empathy with Synthetic Pain	147
7.2	Future Works	149
7.2.1	Framework Development	149
7.2.2	Application Domain	150
References		155
Appendix A Terminology		169
Appendix B Documentation		171
B.1	Dimensions	171
B.2	Links	171
B.3	Joints and Motors	173
Appendix C Experiment Results Appendix		181
C.1	Non-Empathy Appendix	181
C.1.1	SPD-based Appendix	181
C.1.2	Pain Matrix-based Appendix	196

List of Figures

3.1	Synthetic Pain Activation PP and IP	34
3.2	Synthetic Pain Activation SMP	35
3.3	Pain Region Assignment	36
3.4	Pain Matrix Diagram	39
4.1	Adaptive Robot Self-Awareness Framework (ASAF)	46
4.2	Robot Awareness Region and CDV	47
4.3	Robot Mind Structure	55
4.4	Robot Mind Reasoning Process	56
5.1	NAO Humanoid Robot (Aldebaran, 2006)	58
5.2	Non Empathic Experiment	59
5.3	Initial Pose for Robot Experiments	61
5.4	Geometrical Transformation	64
6.1	Offline without Human Interaction Trial 1	68
6.2	Offline without Human Interaction Trial 2	69
6.3	Offline without Human Interaction Trial 3	70
6.4	Offline without Human Interaction Trial 4	74
6.5	Offline without Human Interaction Trial 5	74
6.6	Offline with Human Interaction Trial 1	79
6.7	Offline with Human Interaction Trial 2	80
6.8	Offline with Human Interaction Trial 3	81
6.9	Offline with Human Interaction Trial 4	83
6.10	Offline with Human Interaction Trial 5	85
6.11	Online without Human Interaction Trial 1	87
6.12	Online without Human Interaction Trial 2	88
6.13	Online without Human Interaction Trial 3	89
6.14	Online without Human Interaction Trial 4	90

6.15	Online without Human Interaction Trial 5	91
6.16	Online with Human Interaction Trial 1	92
6.17	Online with Human Interaction Trial 2	94
6.18	Online with Human Interaction Trial 3	95
6.19	Online with Human Interaction Trial 4	97
6.20	Online with Human Interaction Trial 5	97
6.21	Prediction Data SPD-based Model Trial 1	104
6.22	Prediction Data SPD-based Model Trial 2	108
6.23	Prediction Data SPD-based Model Trial 3	112
6.24	Prediction Data SPD-based Model Trial 4	118
6.25	Prediction Data SPD-based Model Trial 5	123
6.26	Observer Data	133
6.27	Region Mapping of Joint Data - Upward Experiment	138
6.28	Region Mapping of Joint Data - Downward Experiment	138

List of Tables

2.1	Hierarchical Model of Consciousness and Behaviour	16
2.2	Modalities of Somatosensory Systems (<i>Source: Byrne and Dafny, 1997</i>) . .	21
3.1	Artificial Pain for Robots	32
3.2	SPD Recommendation	38
3.3	Pain Matrix Functionality	41
4.1	Elements of the Database	50
5.1	Pre-Defined Values in the Database	62
5.2	Awareness State	62
5.3	Synthetic Pain Experiment	63
6.1	Experiment Overview	65
6.2	Offline Pre-Recorded without Physical Interaction Trial 1 to Trial 3	68
6.3	Offline Pre-Recorded without Physical Interaction Trial 4 and Trial 5	69
6.4	Offline Pre-Recorded with Physical Interaction Trial 1 to Trial 3	70
6.5	Offline Pre-Recorded with Physical Interaction Trial 4 and Trial 5	71
6.6	Online without Physical Interaction Trial 1 to Trial 3	71
6.7	Online without Physical Interaction Trial 4 and Trial 5	72
6.8	Online with Physical Interaction Trial 1 to Trial 3	72
6.9	Online with Physical Interaction Trial 4 and Trial 5	72
6.10	Offline without Physical Interaction - Interval Time	73
6.11	Prediction Error - Offline No Interaction	73
6.12	Interval Joint Data and Time Offline with Physical Interaction Trial 1 to Trial 3	76
6.13	Interval Joint Data and Time Offline with Physical Interaction Trial 4 and Trial 5	77
6.14	Prediction Error - Offline Physical Interaction Trial 1	78
6.15	Prediction Error - Offline Physical Interaction Trial 2	80

6.16 Prediction Error - Offline Physical Interaction Trial 3	81
6.17 Prediction Error - Offline Physical Interaction Trial 4	82
6.18 Prediction Error - Offline Physical Interaction Trial 5	84
6.19 Prediction Error - Online without Physical Interaction	86
6.20 Prediction Error - Online without Physical Interaction Trial 1	87
6.21 Prediction Error - Online without Physical Interaction Trial 2	88
6.22 Prediction Error - Online without Physical Interaction Trial 3	89
6.23 Prediction Error - Online without Physical Interaction Trial 4	90
6.24 Prediction Error - Online without Physical Interaction Trial 5	91
6.25 Prediction Error - Online with Physical Interaction Trial 1	92
6.26 Prediction Error - Online with Physical Interaction Trial 2	93
6.27 Prediction Error - Online with Physical Interaction Trial 3	94
6.28 Prediction Error - Online with Physical Interaction Trial 4	96
6.29 Prediction Error - Online with Physical Interaction Trial 5	96
6.30 State of Awareness	99
6.31 Internal States after Reasoning Process	100
6.32 Joint Data and Prediction Data SPD-based Model Trial 1	102
6.33 Prediction Error SPD-based Model Trial 1	103
6.34 SPD Initial State Trial 1	103
6.35 SPD Pain Activation Trial 1	105
6.36 Robot Mind Recommendation Trial 1	106
6.37 Joint Data and Prediction Data SPD-based Model Trial 2	107
6.38 Prediction Error SPD-based Model Trial 2	107
6.39 SPD Initial State Trial 2	109
6.40 SPD Pain Activation Trial 2	110
6.41 Robot Mind Recommendation Trial 2	110
6.42 Joint Data and Prediction Data SPD-based Model Trial 3	111
6.43 Prediction Error SPD-based Model Trial 3	112
6.44 SPD Initial State Trial 3	114
6.45 SPD Pain Activation Trial 3	115
6.46 Robot Mind Recommendation Trial 3	116
6.47 Joint Data and Prediction Data SPD-based Model Trial 4	117
6.48 Prediction Error SPD-based Model Trial 4	117
6.49 SPD Initial State Trial 4	119
6.50 SPD Pain Activation Trial 4	121
6.51 Robot Mind Recommendation Trial 4	122

6.52	Joint Data and Prediction Data SPD-based Model Trial 5	122
6.53	Prediction Error SPD-based Model Trial 5	123
6.54	SPD Initial State Trial 5	124
6.55	SPD Pain Activation Trial 5	125
6.56	Robot Mind Recommendation Trial 5	125
6.57	SPD Pain Activation - Average	126
6.58	Robot Mind Recommendations	126
6.59	Upward Hand Movement Direction	127
6.60	Downward Hand Movement Direction	127
6.61	Upward Hand Movement Prediction	128
6.62	Belief State During Non-Empathy Experiment Using Pain Matrix Model . .	128
6.63	Pain Activation During Non-Empathy Experiment Using Pain Matrix Model	129
6.64	Pain Matrix Output During Non-Empathy Experiment	130
6.65	Goals - Intentions During Non-Empathy Experiment Using Pain Matrix Model	131
6.66	Faulty Joint Regions	131
6.67	Observer Data with SPD Model in Empathy Experiments	132
6.68	Belief State of the Observer in SPD Model	132
6.69	Observer and Mediator Data During Upward Experiment	133
6.70	Observer and Mediator Data During Downward Experiment	134
6.71	SPD Recommendations - Upward Experiment	134
6.72	SPD Recommendations - Downward Experiment	135
6.73	Goals and Intentions - Upward Experiment	136
6.74	Goals and Intentions - Downward Experiment	136
6.75	Observer Data with Pain Matrix Model	137
6.76	Belief State During Upward Experiment	139
6.77	Belief State During Downward Experiment	139
6.78	Belief State Recommendation During Upward Experiment	140
6.79	Belief State Recommendation During Downward Experiment	140
6.80	Pain Matrix Activation with Current Data - Upward Experiment	141
6.81	Pain Matrix Activation with Prediction Data - Upward Experiment	141
6.82	Goals and Intentions of Observer During Upward Experiment	142
6.83	Goals and Intentions of Observer During Downward Experiment	143
B.1	Body Dimensions	171
B.2	Link and Axis Definitions	171
B.3	Head Definition	172
B.4	Arm Definition	172

B.5	Leg Definition	173
B.6	Head Joints	174
B.7	Left Arm Joints	175
B.8	Right Arm Joints	175
B.9	Pelvis Joints	175
B.10	Left Leg Joints	176
B.11	Right Leg Joints	177
B.12	Motors and Speed Ratio	177
B.13	Head and Arms	178
B.14	Hands and Legs	178
B.15	Camera Resolution	178
B.16	Camera Position	179
B.17	Joint Sensor and Processor	179
B.18	Microphone and Loudspeaker	179
C.1	Experiment Overview-Appendix	181
C.2	Offline without Human Interaction Trial 1 with Prediction Data	182
C.3	Offline without Human Interaction Trial 2 with Prediction Data	183
C.4	Offline without Human Interaction Trial 3 with Prediction Data	184
C.5	Offline without Human Interaction Trial 4 with Prediction Data	185
C.6	Offline without Human Interaction Trial 5 with Prediction Data	186
C.7	Offline with Human Interaction Trial 1 with Prediction Data	187
C.8	Offline with Human Interaction Trial 2 with Prediction Data	188
C.9	Offline with Human Interaction Trial 3 with Prediction Data	189
C.10	Offline with Human Interaction Trial 4 with Prediction Data	190
C.11	Offline with Human Interaction Trial 4 with Prediction Data	191
C.12	Online without Human Interaction Trial 1 with Prediction Data	192
C.13	Online without Human Interaction Trial 2 with Prediction Data	192
C.14	Online without Human Interaction Trial 3 with Prediction Data	193
C.15	Online without Human Interaction Trial 4 with Prediction Data	193
C.16	Online without Human Interaction Trial 5 with Prediction Data	194
C.17	Online with Human Interaction Trial 1 with Prediction Data	194
C.18	Online with Human Interaction Trial 2 with Prediction Data	195
C.19	Online with Human Interaction Trial 3 with Prediction Data	196
C.20	Online with Human Interaction Trial 4 with Prediction Data	197
C.21	Online with Human Interaction Trial 5 with Prediction Data	197
C.22	Pain Matrix Without Human Interaction Appendix	198

C.23 Pain Matrix Without Human Interaction Incoming Belief Appendix	199
C.24 Pain Matrix Without Human Interaction SPD Recommendation	200
C.25 Pain Matrix Without Human Interaction SPD Goals	201

