



Faculty of Engineering and Information Technology

A Methodology for Operationalising the Robot Centric HRI Paradigm:

Enabling Robots to Leverage Sociocontextual Cues During Human-Robot Interaction

A thesis submitted for the degree of
Philosophiae Doctor (PhD)

Sonja Caraian

SUPERVISORS

Principal Supervisor

Dr. Nathan Kirchner

Alternate Supervisor

Dr. Alen Alempijevic

Senior Lecturers, School of Electrical, Mechanical and Mechatronic Systems

Center for Autonomous Systems, University of Technology Sydney

EXAMINERS

Prof. Dr. Vanessa Evers

Professor of Human Media Interaction

University of Twente, Enschede, Netherlands

Takayuki Kanda

Senior Research Scientist

ATR Intelligent Robotics and Communication Laboratories, Kyoto, Japan

October 2015

Certificate of Original Authorship

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 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.

Signed,

Production Note:
Signature removed prior to publication.

Sonja Caraian

Date: 14/10/15

Acknowledgments

The work presented in this thesis was funded by the Australian Research Council (ARC) through my having received an Australian Postgraduate Award (APA) which paid for my fees and scholarship; the University of Technology, Sydney (UTS); the Centre for Autonomous Systems (CAS); and RobotAssist.

Completing this thesis has been one of the biggest challenges I have faced, and it would not have been possible without the help, guidance, support and love of those around me. First and foremost, I wish to thank my principal supervisor, Dr. Nathan Kirchner. He has been a tremendous mentor for me, encouraging and enabling me to become both a better researcher and to grow as a person. His guidance, support, patience, immense knowledge and enthusiasm, and maddening attention to detail have been invaluable throughout this thesis work.

My sincere gratitude also goes to Dr. Alen Alempijevic, who, as my alternate supervisor, has been an inspiration and role model, and his technical know-how has been invaluable. I thank him for his encouraging and constructive feedback.

I also wish to thank Dr. Teresa Vidal-Calleja for her guidance and counsel, Dr. Brad Skinner and Dr. Gavin Paul for taking the time to listen to and encourage me, and Prof. Gamini Dissayanake, who, as the director of CAS, has enabled me and other researchers to learn and grow.

Getting through this thesis required more than academic support, and I would also like to thank all of my friends who supported and guided me in finding my way. My gratitude and appreciation for their friendship is endless.

Finally, none of this would have been possible without my family. Words cannot express how grateful I am to my parents and brother. This thesis stands as a testament to a lifetime of unconditional love and support.

Contents

List of Figures	vii
List of Tables	xi
1 Introduction	1
1.1 Leveraging Sociocontextual Cues to Increase the Effectiveness of HRI	3
1.1.1 The Increasing Presence of Social Robots in Society	3
1.1.2 Social Robots as Interaction Peers Leveraging Sociocontextual Cues	6
1.1.3 Effects of Robot Human-Likeness on Sociocontextual Cues	7
1.1.4 The Robot Centric HRI Paradigm and Robot Interactivity Through Sociocontextual Cues	9
1.2 Research Questions	12
1.2.1 Methodology for Robot Centric HRI Paradigm Operationalisation	12
1.2.2 Transferability of Sociocontextual Cues to HRI	12
1.2.3 Robot Interactivity Moderating Effectiveness	15
1.2.4 Summary	15
1.3 Principal Contributions	16
1.4 Outline of Thesis	19
2 Background and Aspects of the Transferability of an Exemplar Sociocontextual Cue to HRI	21
2.1 Bodily Sociocontextual Cues in HHI	22
2.1.1 The Role of Gaze Cues During Interactions	22
2.1.2 Dynamics of Gaze and the <i>Interaction Zone</i>	23

2.1.3	Mutual Gaze and Joint Attention	24
2.1.4	Characteristic Effects of Joint Attention	29
2.2	Feasibility of Robot-Issued Gaze Cues	31
2.2.1	Human-Likeness of Exemplar Humanoid Social Robots	31
2.2.2	Robot-Generated Joint Attention Cues	33
2.2.3	Summary	35
2.3	Considerations Surrounding Sensing Human-Issued Gaze Cues . .	36
2.3.1	Head Features Exploitable for Head Yaw Estimation . . .	37
2.3.2	Characteristics of Sensors Available for Head Yaw Estimation	38
2.3.3	Leveraging Multiple Imperfect Head Yaw Estimates	40
2.3.4	Summary	42
2.4	Conclusion	43
3	Methodology for Robot Centric HRI Paradigm	
	Operationalisation	44
3.1	Introduction	45
3.2	Methodology for Paradigm Operationalisation	46
3.2.1	Target Problem and Robot Goal Definition	46
3.2.2	Application Space Definition	47
3.2.3	Robot Centric HRI Paradigm Design	48
3.2.4	Implementation Design	55
3.3	Conclusion	57
4	<i>Elicit</i> – Exploring the Effects of Robot-Issued Cues During Real- World HRI	58
4.1	Introduction	59
4.2	Measures of Expected Human Gaze Behaviour in HRI Joint Attention Scenarios	60
4.3	Empirical Evaluation of the Effects of Joint Attention During HRI	64
4.3.1	Hypotheses	64
4.3.2	Participants	65
4.3.3	Setting and Setup	65
4.3.4	Experimental Conditions	66
4.3.5	Procedure	66
4.3.6	Measurement	69
4.3.7	Results	70

4.3.8	Discussion	74
4.4	Conclusions	76
5	<i>Read</i> – Enabling Robots to Decipher Human-Issued Cues	77
5.1	Investigating Human Gaze Behaviour in the HRI Space	78
5.1.1	Hypotheses	79
5.1.2	Participants	79
5.1.3	Setting and Setup	79
5.1.4	Procedure	80
5.1.5	Experimental Conditions	80
5.1.6	Measurement	81
5.1.7	Results	81
5.1.8	Discussion	82
5.1.9	Conclusions	84
5.2	Development of Head Yaw Estimation for the HRI Space	85
5.2.1	Existing Head Yaw Estimation Approaches	86
5.2.2	Developed Head Yaw Estimation Framework	88
5.2.3	Online Head Yaw Estimation Framework Evaluation	101
5.3	Conclusions	116
6	<i>Interactivity</i> – Demonstrating the Relationship Between Interactivity and Effectiveness	118
6.1	Introduction	119
6.2	Exploration of the Value of the <i>Elicit</i> and <i>Read</i> Branches	121
6.2.1	Hypotheses	121
6.2.2	Participants	122
6.2.3	Setting	123
6.2.4	Experimental Conditions	124
6.2.5	Procedure	124
6.2.6	Measurement	128
6.2.7	Results	129
6.2.8	Discussion	133
6.3	Conclusions	137
7	Generalising the Methodology and Reinforcing and Deepening Previous Findings	139
7.1	Introduction	140

7.2	Moderating a Lower Human-Likeness Social Robot’s Ability to Influence Through its Interactivity	141
7.2.1	Design of the Robot Centric HRI Paradigm	141
7.2.2	Empirical Explorations	144
7.2.3	Discussion	152
7.3	Conclusion	154
8	Conclusions	155
8.1	Specific Conclusions on Contributions	156
8.1.1	Methodology for operationalisation of the Robot Centric paradigm during real-world HRI	156
8.1.2	Demonstration that sociocontextual cues can be successfully leveraged during HRI via the Robot Centric HRI paradigm	156
8.1.3	Deepened understanding of the Robot Centric HRI paradigm	157
8.1.4	Demonstration of generalisability of the Robot Centric HRI paradigm and the devised methodology	159
8.2	Future Research	160
A	Publications and Other Outcomes	161
A.1	Directly Related Publications	162
A.2	Related Publications	162
A.3	Awards and Recognition	163
	References	164

List of Figures

1.1	Ageing populations, such as that in Australia, are raising aged-care costs and necessitating ageing-in-place	4
1.2	Sydney Central Business District (CBD) station entries and exits by time of day and day type, showing high volume morning and afternoon peak periods	5
1.3	Paro, a socially interactive and assistive robot	5
1.4	With increasing human-likeness (HL), people prescribe robots a greater number of human characteristics	7
1.5	Non-verbal cues play an important role in communication	8
1.6	A contemporary Robot Centric Human-Robot Interaction (HRI) paradigm, which proposes robots as interaction peers with increased agency	10
2.1	Hall’s proxemic zones [57] depicted along with Kendon’s transactional segment [76], the overlap of which is a person’s likely <i>interaction zone</i>	25
2.2	Three people engaged in joint attention (JA)	27
2.3	Young children engaged in joint attention	28
2.4	The three-step joint attention sequence to increase object desirability	28
2.5	Visual search patterns during Human-Human Interaction (HHI) decision-making scenarios.	30
2.6	Two leading social robotics research platforms, illustrating their humanoid shape	32
2.7	The humanoid shape and capabilities of the RobotAssist platform make it likely that gaze cues issued by the robot will have similar effects and characteristics to gaze cues issued during HHI	33

2.8	A robot displaying two steps of the three step sequence to increase object desirability	34
2.9	The RobotAssist platform executing the three-step JA sequence	34
2.10	The RobotAssist platform executing joint attention left (JAL) and joint attention right (JAR) cues	35
2.11	Standard deviation of plane-fitting residuals at different distances of the plane to the sensor	38
2.12	Typical operation spaces of coarser and finer facial feature head yaw estimation (HYE) methods	39
2.13	Point cloud of a planar surface at different distances from the sensor, projected on the YZ plane	40
3.1	Process flow of the methodology for operationalisation of the Robot Centric HRI paradigm	47
3.2	Process flow of design of the <i>Read</i> and <i>Elicit</i> branches of the paradigm	49
3.3	Detail of the process flow of the ‘Design of the Robot Centric HRI Paradigm’ stage of the methodology	54
4.1	Object (OB) and Not Object (!OB) locations in a scene	61
4.2	Joint attention (JA) and Not joint attention (!JA) objects in a scene	62
4.3	Chosen (CH) and Not Chosen (!CH) objects in a scene	63
4.4	Experiment setup	67
4.5	Experiment scenario	67
4.6	Experimental procedure	69
4.7	The Robot Operating System (ROS) coding image and Graphical User Interface	70
4.8	Participants’ attendance to the OB and !OB locations during the experiment	72
4.9	Participants’ attendance the the JA and !JA objects during the experiment	72
4.10	Participants’ attendance of CH and !CH objects during the experiment	73
5.1	To ensure cues are witnessed, a robot needs an understanding of when people are looking at it	78
5.2	The experiment setup	80
5.3	The experiment scenario	81

5.4	Participants Not Looking (top row) and Looking (bottom row) at the RobotAssist platform during the experiment	82
5.5	Looking patterns of participants at the robot over the course of their interactions	83
5.6	The number of participants looking at the robot and those still interacting over time	83
5.7	Process and information flow of the Fanelli et al. HYE method . .	87
5.8	Process and information flow of the HYE framework	89
5.9	The segmented head point cloud	91
5.10	Illustration of the visible plane and Face Plane Yaw Estimation FPYE and FPYE' methods on a head point cloud viewed from the top down	94
5.11	Point clouds of two different people with 10 slice spans of the Head-to-Shoulder Signature (HSS) illustrated	95
5.12	Setup and procedure of Gaussian process (GP) model training data acquisition	99
5.13	Raw readings of the HYE methods from the training data	103
5.14	Raw error results of the HYE methods from the training data . .	104
5.15	Experiment setting and setup	106
5.16	Experiment Part A individual HYE method raw error results . . .	108
5.17	Conceptual depiction of the individual HYE methods' raw errors .	108
5.18	Experiment Part A fusion error results	109
5.19	Conceptual depiction of the fusion errors	110
5.20	Experiment Part B individual HYE method raw error results . . .	113
5.21	Experiment Part B fusion error results	114
5.22	Participant paths during the two parts of the experiment	115
6.1	Setting and setup of the final experiment	124
6.2	Flowchart of the experimental procedure	125
6.3	Diagrammatic representation of the possible experiment outcomes	128
6.4	Results of the influence of the JA cue on choice	130
6.5	Participants' two predominant patterns of contribution of the head and eyes to the horizontal re-orientation of gaze towards the boxes	133
6.6	Approximate ground truth and HYE framework estimates when participants looked towards the JA and !JA Boxes with different relative contributions of head and eyes to gaze re-orientation . . .	136

7.1	Directional indicators are likely to be more congruent with the lower-HL of Tillie, making them likely to be interpretable by interacting humans	143
7.2	The social robot utilised to generalise the findings of this work . .	145
7.3	Setting for the Part 1 study	147
7.4	Influencing people towards the left	148
7.5	The Part 2 study setting and setup	150
7.6	Influence reducing the extent of people cutting the corner in Part 2 of the study	152

List of Tables

4.1	JA experiment participant trial distribution	71
5.1	5-fold cross validation root mean square (RMS) model errors for different input feature vectors	99
5.2	Experiment Part A mean accuracies of the HYE methods	111
5.3	Experiment Part B mean accuracies of the HYE methods	114

List of Acronyms

Acronym	
ANOVA	Analysis of Variance
C	Compliant
CAS	Centre for Autonomous Systems
D	Dimension
DOF	Degree-of-freedom
FEIT	Faculty of Engineering and Information Technology
FPYE	Face Plane Yaw Estimation
GP	Gaussian process
GUI	Graphical User Interface
H	Hypothesis
HHI	Human-Human Interaction
HL	Human-Likeness
HRI	Human-Robot Interaction
HSS	Head-to-Shoulder Signature
HYE	Head Yaw Estimation
JA	Joint Attention
L	Looking
!L	Not Looking
NC	Non-Compliant
PCA	Principal Component Analysis
RGB	Red Green Blue
RGB-D	Red Green Blue Depth
ROS	Robot Operating System
RMS	Root Mean Squared
RQ	Research Question
UTS	University of Technology, Sydney

Abstract

A Methodology for Operationalising the Robot Centric HRI Paradigm:

Enabling Robots to Leverage Sociocontextual Cues During Human-Robot Interaction

Sonja Caraian

October 2015

The presence of social robots in society is increasing rapidly as their reach expands into more roles which are useful in our everyday lives. Many of these new roles require them to embody capabilities which were typically not accounted for in traditional Human-Robot Interaction (HRI) paradigms, for example increased agency and the ability to lead interactions and resolve ambiguity in situations of naïvety. The ability of such robots to leverage sociocontextual cues (i.e. non-verbal cues dependent on the social-interaction space and contextual-task space in order to be interpreted) is an important aspect of achieving these goals effectively and in a socially sensitive manner.

This thesis presents a methodology which can be drawn on to successfully operationalise a contemporary paradigm of HRI – Kirchner & Alempijevic’s Robot Centric HRI paradigm – which frames the interaction between humans and robots as a loop, incorporating additional feedback mechanisms to enable robots to leverage sociocontextual cues. Given the complexities of human behaviour and the dynamics of interaction, this is a non-trivial task. The Robot Centric HRI paradigm and methodology were therefore developed, explored and verified through a series of real-world HRI studies ($n_{total} = 435 = 16 + 24 + 26 + 96 + 189 + 84$).

Firstly, by drawing on the methodology, it is demonstrated that sociocontextual cues can be successfully leveraged to increase the effectiveness of HRI in both directions of communication between humans and robots via the paradigm. Specifically, cues issued by social robots are shown to be recognisable to people, who generally respond to them in line with human-issued cues. Further, enabling robots to read interaction partners' cues *in situ* is shown to be highly valuable to HRI, for example by enabling robots to intentionally and effectively issue cues. In light of the finding that people will display HHI-predicted sociocontextual cues such as gaze around robots, a novel head yaw estimation framework which showed promise for the HRI space was developed and evaluated. This enables robots to read human-issued gaze cues and mutual attention *in situ*.

Next, it is illustrated that a robot's effectiveness at achieving its goal(s) can be increased by adding to its ability to moderate the cues it issues based on information read from humans (i.e. increased interactivity).

Finally, the above findings are shown to generalise to other sociocontextual cues, social robots and application spaces, demonstrating that the developed methodology can be drawn on to successfully operationalise the Robot Centric HRI paradigm, enabling robots to leverage sociocontextual cues to more effectively achieve their goal(s) and meet the requirements of their expanding roles.