


UNIVERSITY OF TECHNOLOGY SYDNEY

# **Planning and Sequential Decision Making for Human-Aware Robots**

by  
Tarek Taha



A thesis submitted in partial fulfilment for the  
degree of Doctor of Philosophy

in the  
Faculty of Engineering  
Mechatronics and Intelligent Systems

May 2012

## Declaration of Authorship

I, Tarek Taha, declare that this thesis titled, 'Planning and Sequential Decision Making for Human-Aware Robots' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Production Note:  
Signature removed prior to publication.

---

Date: 07/05/2012

*'The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge'*

Stephen W. Hawking

---

UNIVERSITY OF TECHNOLOGY SYDNEY

## *Abstract*

Faculty of Engineering  
Mechatronics and Intelligent Systems

Doctor of Philosophy

by Tarek Taha

This thesis explores the use of probabilistic techniques for enhancing the interaction between a human and a robotic assistant. The human in this context is regarded as an integral part of the system, providing a major contribution to the decision making process and is able to overwrite, re-evaluate and correct decisions made by the robot to fulfil her or his true intentions and ultimate goals and needs. Conversely, the robot is expected to behave as an intelligent collaborative agent that predicts human intentions and makes decisions by merging learned behaviours with the information it currently possesses. The work is motivated by the rapid increase of the application domains in which robotic systems operate, and the presence of humans in many of these domains. The proposed framework facilitates human-robot social integration by increasing the synergy between robot's capabilities and human needs, primarily during assistive navigational tasks.

The first part of the thesis sets the groundwork by developing a path-planning/re-planning strategy able to produce smooth feasible paths to address the issue of navigating a robotic wheelchair in cluttered indoor environments. This strategy integrates a global path-planner that operates as a mission controller, and a local reactive planner that navigates locally in an optimal manner while preventing collisions with static and dynamic obstacles in the local area. The proposed strategy also encapsulates social behaviour, such as navigating through preferred routes, in order to generate socially and behaviourally acceptable plans.

The work then focuses on predicting and responding to human interactions with a robotic agent by exploiting probabilistic techniques for sequential decision making and planning under uncertainty. Dynamic Bayesian networks and partially observable Markov decision processes are examined for estimating human intention in order to minimise the flow of information between the human and the robot during navigation tasks. A framework to capture human behaviour, motivated by the human action cycle as derived from the psychology domain is developed. This framework embeds a human-robot interaction layer, which defines variables and procedures to model interaction scenarios, and facilitates the transfer of information during human-robot collaborative tasks.

Experiments using a human-operated robotic wheelchair carrying out navigational daily routines are conducted to demonstrate the capacity of the proposed methodology to understand human intentions and comply with their long term plans. The results obtained are presented as the outcome of a set of trials conducted with actor users, or simulated experiments based on real scenarios.



## *Acknowledgements*

This work was made possible by many people other than myself, and I am extremely grateful for their encouragement and support.

Firstly, I would like to thank my supervisors, Dr Jaime Valls Miró and Professor Gamini Dis-sayanake, for their support, guidance and encouragement.

I would also like to thank all the bright minds who were around me at the Centre of Excellence for Autonomous Systems, for all the intelligent conversations, coffee breaks and research culture. Nathan Kirchner, Gavin Paul, Stephan Sehestedt and Weizhen Zhou, your presence around helped more than you can imagine.

Many thanks to Professor Raymond Apthorpe for his words of wisdom and support, and Christer for bearing with me during my intense times.

I would also like to acknowledge Dr Lisa Lines for her editorial intervention, which was restricted to Standards D and E of the Australian Standards for Editing Practice.

I would like to reserve my final thanks to my parents for their years of blessings and motivation.

# Contents

<b>Declaration of Authorship</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xiii</b>
<b>Abbreviations</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Research Problem and Motivation	1
1.2 Approach and Methodology	2
1.3 Contributions	4
1.4 Thesis Structure	5
1.5 Publications	6
<b>2 CSTEP for Navigation in Cluttered Environments</b>	<b>8</b>
2.1 Related Work	9
2.1.1 Configuration Space Construction	10
2.1.2 Global Path-planning	11
2.1.3 Local Reactive Planning	14
2.2 Navigation Strategy for Robotic Wheelchair in Indoor Cluttered Environments	16
2.3 Global Planning Approach	16
2.3.1 Generating the Search Space	19
2.3.2 Finding Paths in the Search Space	25
2.4 Reacting to Environmental Changes	25
2.5 Experimental Results	30
2.5.1 Experimental Platform	30
2.5.2 Effect of the Sampling Approach	31

2.5.3	Comparison with Existing Randomised Planners . . . . .	33
2.5.4	Path Following . . . . .	34
2.5.5	Handling Dynamic Environments . . . . .	36
2.6	Summary . . . . .	38
<b>3</b>	<b>Socially Predictive Path-planning</b>	<b>40</b>
3.1	Intention Prediction . . . . .	41
3.2	Capturing Social Navigation Behaviour . . . . .	44
3.2.1	Preferred Route . . . . .	44
3.2.2	Places of Interest . . . . .	46
3.2.3	Time-based Behaviour . . . . .	47
3.3	Intention Predictive Navigation using DBNs . . . . .	47
3.3.1	DBN . . . . .	49
3.3.2	Global Wheelchair Social Predictive Navigation . . . . .	50
3.4	Intention Predictive Navigation using POMDPs . . . . .	55
3.4.1	POMDPs . . . . .	56
3.4.2	POMDP-based Wheelchair Predictive Navigation . . . . .	60
3.4.3	State Space . . . . .	60
3.4.4	Transition Model . . . . .	62
3.4.5	Observation Model . . . . .	62
3.4.6	Reward Function . . . . .	63
3.4.7	Interaction Strategies . . . . .	64
3.5	Summary . . . . .	69
<b>4</b>	<b>Modelling HRI using Human-Aware POMDPs</b>	<b>71</b>
4.1	Background on Human Psychology . . . . .	72
4.1.1	Human Action Cycle . . . . .	72
4.1.2	POMDPs in Assistive HRI Applications . . . . .	75
4.2	A Human-Aware POMDP Model for HRI . . . . .	77
4.2.1	HRI Variables . . . . .	78
4.3	Modelling and Solving Human-Aware POMDP . . . . .	81
4.3.1	ADDs . . . . .	82
4.3.2	Factoring Human-Aware POMDP with ADDs . . . . .	85
4.4	Selected Applications . . . . .	87
4.4.1	Wheelchair Navigation . . . . .	87
4.4.2	Intelligent Walker . . . . .	90
4.4.3	Robotic Tutors . . . . .	93
4.4.4	Wheelchair Experimental results . . . . .	95
4.4.5	Walker Experimental results . . . . .	99
4.5	Summary . . . . .	104
<b>5</b>	<b>Conclusion</b>	<b>105</b>
5.1	Summary . . . . .	105
5.2	Future Research Directions . . . . .	106
5.2.1	Enhancing Modelling . . . . .	106



5.2.2 Applications to Explore . . . . .	108
5.3 Conclusion . . . . .	109

Appendices

A ADDs . . . . .	112
A.1 ADD Formal Definition . . . . .	112
A.2 Basis in Boolean Algebra . . . . .	114
A.3 Properties of ADDs . . . . .	116
A.4 Matrix Multiplication . . . . .	116
A.4.1 Matrix multiplication in semi-rings . . . . .	117
A.4.2 Matrix multiplication in quasi-rings . . . . .	118

Bibliography . . . . .	120
------------------------	-----

# List of Figures

1.1	Interaction between a human, a robotic agent and their environment . . . . .	4
2.1	A robot represented by a configuration at point $r(x, y)$ in an environment with a single obstacle . . . . .	12
2.2	The configuration space constructed by sliding the robot around the obstacle (only translations using point $r(x, y)$ ) are considered . . . . .	12
2.3	The global navigation flowchart. . . . .	17
2.4	The navigation system architecture and the links between the underlying components. . . . .	18
2.5	Steps performed to generate the off-line search space, and how this search space is used to find paths on-line. . . . .	18
2.6	Largest robot dimension obstacle expansion method . . . . .	19
2.7	Narrow passage blocked as a result of largest robot dimension obstacle expansion . . . . .	20
2.8	The area of the robot is covered by circles of radius $R$ , the centres of these circles will be the points to be checked for collision . . . . .	21
2.9	The result of expanding obstacles in the map . . . . .	22
2.10	Bridge test performed in two locations. The bridge test favours samples in narrow areas and ignores samples in wide-open space. . . . .	23
2.11	Path generated using CSTEP. The green network represents the search tree expanded while searching for the optimal path. The path is depicted in the grey line segments, and the brown rectangles represent the robot's configurations on the path. . . . .	26
2.12	An initial path generated to goal. Red line segments represent connections between neighbouring nodes, the blue line is the generated path, the green lines represent the laser scan and the green triangles indicate the start and the end destinations. . . . .	26
2.13	The process of updating the search space from performing iterative closest point algorithm between the laser scans and the internal map. . . . .	27
2.14	A re-planning scenario. Changes in the environment are detected which trigger local modification of the search space and regeneration of the global path. . . . .	28
2.15	Illustration of how way-points are passed to the local planner. The red arrows in the figures denote way-points that are selected on the generated path from the unobstructed laser field of view denoted by the light-blue region. . . . .	29
2.16	Autonomous wheelchair platform . . . . .	31
2.17	Uniform sampling: the whole environment is equally mapped with a constant density of search nodes . . . . .	32

2.18	Non uniform sampling: density is increased around narrow passages, leaving open spaces with a lower sampling density . . . . .	32
2.19	Non uniform sampling: path generated using CSTEP method . . . . .	34
2.20	Path generated by MSL: this is an example of a path generated using MSL library with the RRTextExt planner . . . . .	35
2.21	Another path generated by MSL: this is an example of a path generated using MSL library with the RRTDual planner . . . . .	35
2.22	Mobile platform navigational architecture with linear controller . . . . .	36
2.23	Result of navigating a path generated by the planner, where the thick line represents the feedback position as estimated by the localiser, and the thin line represents the actual planned path . . . . .	37
2.24	A case study of the robot navigation strategy in the presence of static and dynamic obstacles. . . . .	38
3.1	Topological map structure of an office area is shown on top of the metric representations. Rectangles in this structure represent a possible destination node, circles represent topological nodes in the map and the line segments represent a viable physical path between two nodes. . . . .	45
3.2	A map with a topology of the traffic. The thickness of the green line indicates how often the user visited that path segment. Gray squares represent topology nodes in the map and red squares indicate destinations. The percentages show how often the user went to each destination. . . . .	46
3.3	A 3D top view of the map. The height of the green vertical bars and the percentages on them represent the popularity of the destinations. . . . .	47
3.4	A representation of the traffic in the map. The time of the day is divided into 4 intervals, each of which has a different traffic dynamics. The thickness of the line segments connecting any two nodes show how heavy the traffic is on that line segment. . . . .	48
3.5	A two slice DBN used for destination prediction. Loc is the current location of the robot, Joy is the joystick input and Dest is the destination goal to be predicted. . .	52
3.6	Two paths generated by the same path-planner discussed in Chapter 2, but with two different cost functions. In the first approach, only path optimality is taken into account, while in the other approach, the generated path takes into consideration the preferred routes in the map. . . . .	54
3.7	A navigation route to a predicted destination based on joystick inputs and a learned DBN. . . . .	55
3.8	The sequence of observations and the predicted destination shown at each topological state during Navigation. States $Sx$ correspond to the actual location on the map as shown previously in Figure 3.1. . . . .	56
3.9	The evolution of the destinations probabilities. The observations obtained at each topological state are used to predict the destinations. At each step, the destination with the highest probability is used to navigate to the next step. . . . .	57
3.10	Two time slices of general POMDP represented as a DBN . . . . .	58



3.11	The POMDP model generation architecture. The map topology together with the training data are used to determine the transition model. The training data is also used to determine the observation model of the POMDP. User's joystick calibration determines the uncertainty in the observations. . . . .	61
3.12	The POMDP driver assistance architecture. The user's input together with the current location generate an observation that helps in updating the belief in the destination. The appropriate action will be selected based on that belief, and the next state will then be determined and given to the navigator to drive the wheelchair to the next state. . . . .	63
3.13	The map topology used for intention recognition. Circles represent intersections and cannot be a destination while squares represent rooms or open spaces and can be considered a possible destination. The numbers inside the circles represent the possible wheelchair locations and are used to build the transition model. Gray shaded rectangles represent learned destinations. . . . .	64
3.14	The result of a wheelchair experiment showing the path (dashed line) and observations (arrows). The wheelchair starts in location 2 and drives the user successfully to location 26 by updating the belief at each step from the obtained observation. . . . .	67
3.15	Integrating belief into planning. . . . .	69
3.16	POMDP as a belief-based planner. . . . .	69
4.1	The Seven Stages of Action [1]. . . . .	74
4.2	Interaction strategy: the robot combines observations from the human partner and from the environment to generate the beliefs needed to select appropriate actions. . . . .	78
4.3	An extension to the seven stages of action theory to include a third gulf (Interaction gulf) acting as a communication bridge between the human and the robot. . . . .	79
4.4	Decision Diagram represented by a CPT . . . . .	83
4.5	ADD of the same decision diagram in Figure 4.4 . . . . .	83
4.6	Action network for (a) GetC and (b) TestC (reproduced from [2]) . . . . .	85
4.7	Reward function network (reproduced from [2]) . . . . .	85
4.8	Progression from (a) the classical POMDP structure, to (b) a two-slice DBN with the interaction layer/gulf. . . . .	87
4.9	The instrumented robotic wheelchair platform used in the experiments. . . . .	88
4.10	Two slice DBN where: $(Ta)$ is the variable set, $(Sa)$ is the user's satisfaction, $(In)$ is the user's intention, $(St)$ is the user's status, $(Joy)$ is the user's joystick observation, and $(Loc)$ is the location. . . . .	89
4.11	The instrumented walker platform. . . . .	92
4.12	Path traversed during a navigation experiment. Arrows represent joystick observations, dashed line represents the traversed path. . . . .	96
4.13	The values of the state variables during the navigation example shown in Figure 4.12. . . . .	96
4.14	Change of destination probabilities over time. $Sxx$ represents the topological states that the wheelchair had to pass through. . . . .	97
4.15	State of the system variables at each topological state. $Sat$ refers to the satisfaction variable, locations denoted by $Sxx$ represents the topological states. . . . .	98
4.16	Map with annotation. . . . .	100
4.17	Set of tasks logged from one user. Each colour represents a navigation task followed by the user. . . . .	101

4.18 Observation samples collected to train the model. Each coloured segment represents the location from which sensor data was averaged to obtain a valid observation with minimal sensor noise. . . . . 102

4.19 The states of the system during a task performed by the walker user. . . . . 103

A.1 ADD representation of graphs and matrices.(Image reproduced from [3]) . . . . . 114



# List of Tables

2.1	Comparison of uniform and non-uniform <i>C-space</i> generation . . . . .	33
2.2	Comparison between our method and some of the most significant randomised planners available in the MSL . . . . .	34
3.1	List of Tasks recorded from the user's activities. Paths in this table are represented by the topological node number and the joystick observation obtain in that node. Up, Down, Right, Left and NoInput represent the joystick directional inputs. . . .	45
3.2	List of Tasks recorded from the user's activities . . . . .	47
3.3	POMDP Model Variables . . . . .	60
3.4	The result of an experiment on a real platform. The wheelchair starts in location 22 and tries to predict where the user is going to based on the joystick inputs (observations). The wheelchair in this case successfully takes the user's joystick inputs and decides on the correct actions that take the user to location 30. . . . .	66
3.5	The results of a navigation task with minimal interaction, the user gave only two inputs during navigation . . . . .	68
3.6	DBNs vs POMDPs . . . . .	70
4.1	Wheelchair Model Variables . . . . .	90
4.2	Walker Model Variables . . . . .	91
4.3	Robotic Tutor Variables . . . . .	95
4.4	Walker Model Variables . . . . .	99
4.5	Tasks Route Definition . . . . .	101
4.6	Model Accuracy . . . . .	102

# Abbreviations

<b>ADD</b>	<b>A</b> lgebraic <b>D</b> ecision <b>D</b> iagrams
<b>AI</b>	<b>A</b> rtificial <b>I</b> ntelligence
<b>AMCL</b>	<b>A</b> daptive <b>M</b> onte- <b>C</b> arlo <b>L</b> ocalisation
<b>BDD</b>	<b>B</b> inary <b>D</b> ecision <b>D</b> iagrams
<b>BN</b>	<b>B</b> ayesian <b>N</b> etwork
<b>CAS</b>	<b>C</b> entre of <b>A</b> utonomous <b>S</b> ystems
<b>CPT</b>	<b>C</b> onditional <b>P</b> robability <b>T</b> able
<b>CSTEP</b>	<b>C</b> ontrolled <b>S</b> ampling and <b>T</b> ime <b>E</b> fficient <b>P</b> lanner
<b>CVM</b>	<b>C</b> urvature <b>V</b> elocity <b>M</b> ethod
<b>DAC</b>	<b>D</b> igital to <b>A</b> nalog <b>C</b> onverter
<b>DBN</b>	<b>D</b> ynamic <b>B</b> ayesian <b>N</b> etwork
<b>DWA</b>	<b>D</b> ynamic <b>W</b> indow <b>O</b> bstacle <b>A</b> voidance
<b>EM</b>	<b>E</b> xpectation <b>M</b> aximisation
<b>GDWA</b>	<b>G</b> lobal <b>D</b> ynamic <b>W</b> indow <b>A</b> pproach
<b>HRI</b>	<b>H</b> uman <b>R</b> obot <b>I</b> nteraction
<b>ITE</b>	<b>I</b> f- <b>T</b> hen- <b>E</b> lse
<b>MDP</b>	<b>M</b> arkov <b>D</b> ecision <b>P</b> rocesses
<b>MSE</b>	<b>M</b> ean <b>S</b> quare <b>E</b> rror
<b>ND</b>	<b>N</b> earness <b>D</b> iagram
<b>POMDP</b>	<b>P</b> artially <b>O</b> bservable <b>M</b> arkov <b>D</b> ecision <b>P</b> rocesses
<b>PRM</b>	<b>P</b> robabilistic <b>R</b> oad <b>M</b> aps
<b>RRT</b>	<b>R</b> apidly- <b>E</b> xploring <b>R</b> andom <b>T</b> rees
<b>VFH</b>	<b>V</b> ector <b>F</b> ield <b>H</b> istogram

*Dedicated to my parents*