

# Contents

Abstract	iii
Contents	xiii
1 Introduction	1
1.1 Motivation and objectives . . . . .	1
1.2 Structure of the Thesis . . . . .	2
1.3 Publications . . . . .	3
I State of the art	7
2 Dynamic Programming and Reinforcement Learning	9
2.1 Introduction . . . . .	9
2.2 Background. . . . .	11
2.3 Dynamic Programming . . . . .	15
2.4 Reinforcement Learning. . . . .	18
2.5 Summary . . . . .	21

3	Approximate Dynamic programming	23
3.1	Introduction	23
3.2	Function approximation for reinforcement learning	25
3.3	Policy iteration and Value iteration algorithms in continuous spaces	27
3.4	Batch reinforcement learning	30
3.5	Deep Reinforcement Learning	32
3.6	Linear programming approach to ADP	32
3.7	Summary	33
4	Policy Search	35
4.1	Introduction	35
4.2	Model-free policy search	37
4.3	Model-based policy search	42
4.4	Learning Upper-Level Policy using Importance Sampling-based Policy Search Method	43
4.5	Summary	50
II	Contributions	53
5	Fitted Q-function Control Methodology based on Takagi-Sugeno Systems	55
5.1	Introduction	56
5.2	Preliminaries and problem Statement	57
5.3	Methodology Proposal	62
5.4	TS models and guaranteed cost controllers	63
5.5	Improved TS guaranteed-cost controllers	66
5.6	Q-function parametrisation for Takagi-Sugeno systems	68
5.7	A generic-optimisation approach	70
5.8	Case study: inverted pendulum	72
5.9	Conclusions	82
5.10	Appendix	83

6 A Linear Programming Methodology for Approximate Dynamic Programming	87
6.1 Introduction . . . . .	88
6.2 Preliminares . . . . .	90
6.3 ADP via linear programming: problem setup . . . . .	96
6.4 Regressor choice/regularisation methodology. . . . .	101
6.5 Computation of a value function upper bound. . . . .	105
6.6 Discussion. . . . .	109
6.7 Examples and experimental results . . . . .	111
6.8 Conclusions. . . . .	126
6.9 Appendix . . . . .	128
7 An Incremental Approximate Dynamic Programming methodology	131
7.1 Introduction . . . . .	132
7.2 Problem statement . . . . .	132
7.3 Problem setup . . . . .	132
7.4 Simulation Example . . . . .	134
7.5 Conclusions. . . . .	141
8 Conclusions and future work	143
8.1 Conclusions. . . . .	143
8.2 Future work . . . . .	146
Bibliography	147