

Contents

1	Introduction	1
1.1	Motivation and Definition of the Problem	1
1.2	Objectives	3
1.2.1	General Objective	3
1.2.2	Specific Objectives	3
1.3	Hypotheses	4
1.4	Contribution	4
1.5	Outline	5
2	State of the Art	7
2.1	Reinforcement Learning	7
2.2	Interactive Machine Learning	9
2.2.1	Human feedback in actions domains	9
2.2.2	Human feedback in evaluative domains	11
2.3	The Correspondence Problem	13
3	An Interactive Framework for Learning Continuous Actions Policies based on Corrective Feedback	14
3.1	Introduction	14
3.2	COACH: CORrective Advice Communicated by Humans	15
3.2.1	General Aspects	15
3.2.2	General Framework for Problems of high frequency and complexity	20
3.2.3	COACH with Takagi Sugeno Fuzzy Systems	23
3.2.4	COACH Extension to Multi-Dimensional Action Problems	26
3.3	Experimental Validation and Analysis	26
3.3.1	Mountain Car Problem	27
3.3.2	Cart Pole Problem	29
3.3.3	Ball-Dribbling with Humanoid Robots	31
3.3.4	Learning to balance on a bike	36
3.3.5	Ablation Study	39
3.4	Discussion	41
4	Learning in MDPs with Human Corrective Feedback and Reinforcement Learning	44
4.1	Introduction	44
4.2	Episode based Exploration and Evaluation Policy Search Background	45

4.2.1	Cross-Entropy Method	46
4.2.2	Episode-Based PI ²	47
4.2.3	Episode-Based PI ² with Covariance Matrix Adaptation	47
4.3	Policy Search Guided with Human Corrective Feedback	49
4.3.1	Learning Sequentially with COACH and Policy Search	49
4.3.2	Learning Simultaneously with COACH and Policy Search	50
4.4	Experiments and Results	52
4.4.1	Learning with a Simulated Human Teacher	52
4.4.2	Learning with Real Human Feedback	56
4.4.3	Comparison between tele-operation and learning	61
4.5	Discussion	62
5	Policy Search with Human Corrective Feedback for Motor Primitives	
	Learning	64
5.1	Introduction	64
5.2	Related Work	66
5.3	Corrective Advice for shaping movement primitives	67
5.3.1	Case of policies defined in the Cartesian space	68
5.3.2	Case of policies defined in the joint space	71
5.4	Corrective Advice and Policy Search for shaping movement primitives	72
5.5	Experiments and Results	75
5.5.1	Learning movement primitives with corrective advice	75
5.5.2	Learning with Simultaneous Corrective Advice and Policy Search	83
5.6	Discussion	94
6	Conclusions and Future Work	96
	Bibliography	99