

# Contents

<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Definition of the Problem . . . . .	1
1.2 Objectives . . . . .	2
1.2.1 General Objective . . . . .	2
1.2.2 Specific Objectives . . . . .	2
1.3 Hypothesis . . . . .	2
1.4 Contributions . . . . .	2
1.4.1 Bibliographic and technical production . . . . .	4
1.5 Structure of this Document . . . . .	4
<b>2 Background</b>	<b>5</b>
2.1 Single-Agent Reinforcement Learning . . . . .	5
2.2 Multi-Agent Reinforcement Learning . . . . .	7
2.3 Independent Learners . . . . .	7
<b>3 Decentralized Reinforcement Learning</b>	<b>9</b>
3.1 An Introduction to DRL . . . . .	9
3.2 Potential Advantages of DRL . . . . .	10
3.3 Challenges in DRL . . . . .	11
3.4 DRL Algorithms . . . . .	11
3.4.1 Independent DRL (DRL-Ind) . . . . .	12
3.4.2 Lenient DRL (DRL-Lenient) . . . . .	13
3.4.3 Cooperative Adaptive DRL (DRL-CA) . . . . .	17
3.5 Related Work . . . . .	20
3.6 Summary . . . . .	21
<b>4 Proposed Methodology for Modeling Decentralized Reinforcement Learning Systems</b>	<b>22</b>
4.1 Determining if the Problem is Decentralizable . . . . .	22
4.2 Identifying Common and Individual Goals . . . . .	23
4.3 Defining the Reward Functions . . . . .	24
4.4 Determining if the Problem is Fully Decentralizable . . . . .	25
4.5 Completing RL Single Modelings . . . . .	25

4.6	Summary	26
<b>5</b>	<b>Experimental Validation of Decentralized Reinforcement Learning of Robot Behaviors</b>	<b>28</b>
5.1	Experimental Validation	28
5.2	Three-Dimensional Mountain Car	29
5.2.1	Centralized Modelings	30
5.2.2	Proposed Decentralized Modelings	31
5.2.3	Performance Index	32
5.2.4	RL Algorithm and Optimized Parameters	32
5.2.5	Results and Analysis	33
5.3	Ball-pushing	35
5.3.1	Centralized Modeling	36
5.3.2	Decentralized Modeling	37
5.3.3	Performance index	38
5.3.4	RL algorithm and optimized parameters	38
5.3.5	Physical setup	39
5.3.6	Results and analysis	39
5.4	Ball-Dribbling	41
5.4.1	Proposed Decentralized Modeling	42
5.4.2	Centralized Modeling	42
5.4.3	Performance Indices	44
5.4.4	RL Algorithm and Optimized Parameters	44
5.4.5	Results and Analysis	45
5.5	SCARA Real-Time Trajectory Generation	47
5.5.1	Decentralized Modeling	47
5.5.2	Performance Index	49
5.5.3	RL Algorithm and Optimized Parameters	49
5.5.4	Results and Analysis	49
5.6	Discussion	50
5.7	Summary	53
<b>6</b>	<b>Accelerating Decentralized Reinforcement Learning</b>	<b>54</b>
6.1	DRL and Transfer Knowledge Overview	54
6.2	Literature Review	55
6.3	Proposed KT-based DRL	56
6.3.1	Control Sharing (CoSh)	57
6.3.2	Nearby Action Sharing (NeASh)	57
6.4	Experimental Validation	58
6.4.1	Inwalk Kicking	61
6.4.2	Ball-Dribbling	65
6.5	Discussion	67
6.6	Summary	69
<b>7</b>	<b>Conclusions and Future Direction</b>	<b>70</b>
<b>8</b>	<b>Bibliography</b>	<b>72</b>

