

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Fútbol Robótico . . . . .	3
1.2. Estado del Arte . . . . .	3
1.3. Objetivos . . . . .	5
1.3.1. Objetivos Generales . . . . .	5
1.3.2. Objetivos Específicos . . . . .	5
1.4. Estructura de la Memoria . . . . .	6
<b>2. Marco Teórico</b>	<b>7</b>
2.1. Aprendizaje Reforzado . . . . .	7
2.1.1. Motivación . . . . .	8
2.1.2. Formulación del Aprendizaje Reforzado . . . . .	10
2.1.3. Objetivo del Aprendizaje Reforzado . . . . .	12
2.1.4. Taxonomía del Aprendizaje Reforzado . . . . .	14
2.1.5. Aprendizaje Reforzado en Robótica . . . . .	16
2.2. Algoritmos tradicionales de Aprendizaje Reforzado . . . . .	18
2.2.1. Aproximación Funcional . . . . .	18
2.2.2. Q-Learning . . . . .	20
2.2.3. Sarsa . . . . .	24
2.2.4. Actor-Crítico Lineal . . . . .	27
2.2.5. Aprendizaje Reforzado Descentralizado . . . . .	27
2.3. Algoritmos de Deep Reinforcement Learning . . . . .	30
2.3.1. Experience Replay . . . . .	31
2.3.2. Target Networks . . . . .	32
2.3.3. Deep Q-Learning . . . . .	32
2.3.4. Deep Deterministic Policy Gradients (DDPG) . . . . .	34
2.3.5. Delimitación en espacios de acciones continuos . . . . .	34
2.3.6. Observabilidad Parcial y Redes Recurrentes . . . . .	36
<b>3. Diseño e Implementación</b>	<b>38</b>
3.1. Librería de Aprendizaje Reforzado - UChileRL . . . . .	39
3.1.1. Motivación . . . . .	39
3.1.2. Herramientas utilizadas . . . . .	40
3.1.3. Arquitectura e Implementaciones . . . . .	41
3.2. Caso de estudio . . . . .	49
3.2.1. Introducción y Objetivo . . . . .	49
3.2.2. Herramientas Utilizadas . . . . .	50

3.2.3.	Ball Dribbling . . . . .	51
3.2.3.1.	Modelamiento . . . . .	52
3.2.3.2.	Recompensa y Rendimiento . . . . .	54
3.2.3.3.	Condiciones de entrenamiento . . . . .	56
3.2.4.	In-walk Kick . . . . .	57
3.2.4.1.	Modelamiento . . . . .	58
3.2.4.2.	Recompensa y rendimiento . . . . .	59
3.2.4.3.	Condiciones de entrenamiento . . . . .	60
3.2.5.	Navegación Visual sin Mapas . . . . .	61
3.2.5.1.	Modelamiento . . . . .	62
3.2.5.2.	Recompensa y Rendimiento . . . . .	64
3.2.5.3.	Condiciones de entrenamiento . . . . .	65
<b>4.</b>	<b>Análisis de Resultados</b>	<b>67</b>
4.1.	Ball Dribbling . . . . .	68
4.1.1.	Actor-Crítico Lineal . . . . .	69
4.1.2.	SARSA . . . . .	72
4.1.3.	DDPG . . . . .	77
4.1.4.	Proceso de Aprendizaje . . . . .	83
4.1.5.	Política Resultante . . . . .	85
4.1.6.	Análisis . . . . .	87
4.2.	In-walk Kick . . . . .	93
4.2.1.	Comparación de modelos . . . . .	94
4.2.2.	Comparación de funciones base . . . . .	95
4.2.3.	Análisis . . . . .	95
4.3.	Navegación visual . . . . .	99
4.3.1.	Comparación de modelos . . . . .	101
4.3.2.	Análisis . . . . .	105
<b>5.</b>	<b>Conclusiones</b>	<b>107</b>
5.1.	Trabajo Futuro . . . . .	109
	<b>Bibliografía</b>	<b>111</b>