

**Reinforcement Learning:
Un estudio comparativo de la performance de sus principales métodos**

María Liz Crespo, Marcelo Luis Errecalde, Cecilia Inés Montoya

Proyecto UNSL N° 338403¹
Departamento de Informática
Universidad Nacional de San Luis
Ejército de los Andes 950 - Local 106
5700 - San Luis - Argentina
E-mail: {mcrespo, merreca, cmontoya}@unsl.edu.ar
Fax: +54 652 30224

En los últimos años, el interés por el concepto de Reinforcement Learning (RL) se ha incrementado en forma considerable dentro de la comunidad de investigadores de Machine Learning e Inteligencia Artificial en general. El principal motivo fue el suceso que los métodos de RL tuvieron en la resolución de problemas, que no lograban atacar en forma satisfactoria enfoques tradicionales como Programación Dinámica y aprendizaje supervisado (por ejemplo Redes Neuronales).

RL ataca el problema de aprender a controlar agentes autónomos (como por ejemplo robots), mediante interacciones por prueba y error con un ambiente dinámico, el cual le provee señales de refuerzo por cada acción que realiza.

La principal virtud de RL es que permite atacar el *problema de la asignación de crédito temporal*, el cual consiste en asignar un apropiado crédito o censura a las acciones individuales cuando el efecto o recompensa de dichas acciones es demorado hasta que una serie de acciones se han realizado.

Los conceptos teóricos fundamentales de RL, como así también algunos de sus principales métodos son descriptos a modo de survey, dirigidos a aquellas personas que tienen interés en introducirse en este área.

Se presenta un análisis comparativo de los resultados obtenidos mediante métodos libres de modelo (Q-Learning) y métodos que integran aprendizaje y planificación (Dyna-Q y Prioritized Sweeping), tomando como referencia los valores obtenidos con los métodos clásicos de Programación Dinámica (Value Iteration). También se analiza el problema conocido como el *dilema de la exploración-explotación*, ya que en RL es el agente quien controla la distribución de los ejemplos de entrenamiento, eligiendo la secuencia de acciones a tomar. Estos métodos se aplicaron a problemas del mundo de los laberintos, típicamente usados en el área.

Palabras Claves: Inteligencia Artificial, Reinforcement Learning, Machine Learning, Programación Dinámica Asíncrona, Métodos Libres de Modelo, Arquitecturas para Aprendizaje y Planeamiento

¹ El grupo de investigación está soportado por la UNSL y la ANPCyT (Agencia Nacional para la promoción de la Ciencia y Tecnología).