

Curso de Sistemas Inteligentes

Práctica de Laboratorio No. 1

Aprendizaje por Refuerzo

Prof. Francisco Cruz

Ayudante Angel Ayala

Semestre 2 / 2018

1. Planteamiento del Problema

El enfoque computacional del aprendizaje por interacción con el entorno es el problema que aborda el método de aprendizaje por refuerzo. Este presenta un agente cognitivo que sensa una señal de entrada con la información sobre el estado del entorno, además de la posibilidad de ejecutar una determinada acción que influencia el estado siguiente, con el objetivo de obtener la mayor cantidad de señal de recompensa posible. Esto se puede representar mediante el diagrama de la Fig. 1

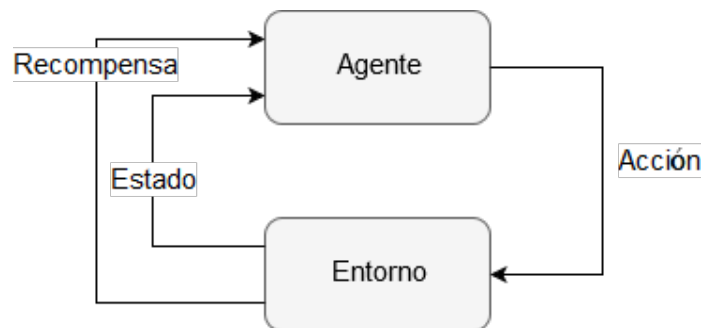


Figura 1: Diagrama de Aprendizaje por Refuerzo

1.1. Elementos de Aprendizaje por Refuerzo

Para un problema de aprendizaje por refuerzo existen tres elementos esenciales y uno opcional para la interacción entre el agente y el entorno:

1. **Política** (*policy*): define la acción que debe ejecutar el agente dado un determinado estado.
2. **Señal de recompensa** (*reward*): establece el objetivo del problema de aprendizaje por refuerzo, definiendo eventos positivos o negativos.
3. **Función de valor** (*value function*): valor Q que determina la calidad de la recompensa a través del tiempo.
4. Opcionalmente, **modelo del entorno** (*model of the environment*): permite inferir el estado y recompensa a futuro.

2. Actividades

Desarrollar y entrenar un agente de aprendizaje por refuerzo para un entorno presente en OpenAI Gym.

- (a) Utilizando Gym, implementar el entorno “CartPole-v0”.
- (b) Desarrollar un agente para el entorno antes mencionado.
- (c) Entrenar el agente durante 1000 episodios con $\alpha = 0,001$.
- (d) Grafique la cantidad de acciones y la posición del carro para cada episodio.

3. Informe

Basandose en el agente anterior, desarrollar un nuevo agente para:

- (a) Entrenar variando cantidad de episodios y el valor de α , en al menos 3 configuraciones diferentes.
- (b) Implementar la política ϵ -greedy y entrenar con valores de $\epsilon = [0, 1; 0, 3; 0, 5]$
- (c) Cambiar la política anterior por softmax y entrenar con valores de $\tau = [0, 1; 0, 3; 0, 5]$
- (d) Graficar los resultados obtenidos.

3.1. Estructura del Informe

El informe debe ser realizado en formato IEEE de doble columna con un máximo de 4 páginas. Incluir al menos las siguientes secciones:

- Título, autores y filiación.
- Resumen.
- Introducción.
- Fundamentos teóricos.
- Desarrollo y explicación de las actividades.
- Conclusiones.
- Referencias científicas en formato IEEE.

3.2. Entrega del Informe

Para la entrega considerar lo siguiente:

- Solo se aceptarán informes en el formato solicitado.
- La entrega del trabajo es individual.
- El informe debe ser enviado al correo angel4ayala@gmail.com con copia a francisco.cruz@ucentral.cl el día de la entrega hasta las 11:59pm.
- Se aceptarán informes de laboratorio atrasados sujetos a castigo de un punto menos por día (incluido sábado, domingo y feriados).
- No adjuntar el archivo o adjuntar el archivo incorrecto es responsabilidad del alumno.
- El archivo debe ser enviado en formato PDF usando como nombre de archivo el siguiente formato:

<Nombre><Apellido>LabSI<N° de lab><Semestre><Año>.pdf (1)

Por ejemplo, IsaacNewtonLabSI222018.pdf correspondería al laboratorio 2 del alumno Isaac Newton.

- El plagio será sancionado con la nota mínima, sin posibilidad de realizar el trabajo nuevamente.
- **Fecha de entrega:** Martes 25 de septiembre.