

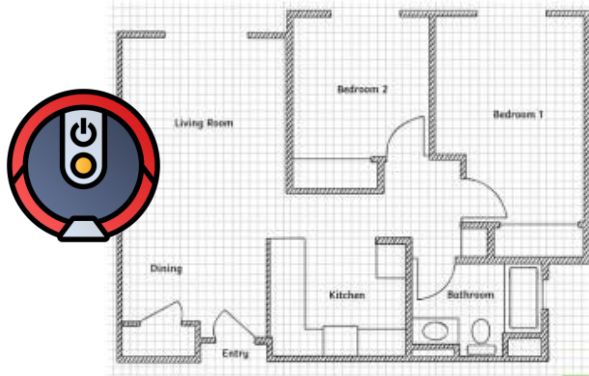


Elementos y proceso del aprendizaje por refuerzo



**DATA SCIENCE
ONLINE**

Elementos del aprendizaje por refuerzo



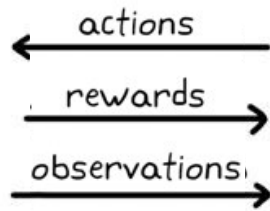
- **Agente (Agent):** Perrito y Robot
- **Entorno/Ambiente (Environment):** El ser humano y la zona de entrenamiento; la casa y sus condiciones de suciedad.
- **Observaciones (Observations) y Estado (State):** Las posturas y respuestas del dueño/entrenador del perrito; las diferentes combinaciones de posición (x,y), número de veces que se ha pasado por esa posición, y el volumen ocupado de la bolsa de vaciado.
- **Recompensa (Reward):** La galletita, los puntos positivos o negativos ganados por el robot
- **Acción (Action):** comportamientos del perro; avanzar, girar, aspirar, en el caso del robot
- **Política (Policy):** El criterio, reglas, programa o modelo que usa el agente para decidir cuál es la mejor acción dado un estado

Proceso

environment



agent



1. Para empezar se “observa” el **entorno**: Nos da las **observaciones** y el **agente** puede saber el estado en el que está
2. El **agente** decide cómo actuar, en general, en función de ese **estado** y de acuerdo a la **política** que tenga en ese momento
3. Actúa
4. El **entorno** le “devuelve” el **estado** (o las observaciones necesarias para definirlo) resultante y la **recompensa** (positiva o negativa)
5. El **agente** aprende de la experiencia: actualiza la **política** o estrategia.
6. Se itera hasta conseguir una estrategia o **política** óptima

Consideraciones y Aplicaciones

- El reinforcement learning se suele considerar como un tercer tipo de aprendizaje (ni supervisado, ni no-supervisado)
- La secuencia de acciones influye en la recompensa final. El “tiempo” sí es importante.
- Ser “glotón” (greedy) no siempre es lo deseado.
- Aplicado a todo tipo de bots: mecánicos, digitales (e.j: ChatGPT)
- ... pero también a problemas como los del principio, sobre todo en sistemas de recomendación
- Aunque para la investigación se emplean principalmente videojuegos

