Politicas_rl

Política

Una política es la regla que un agente sigue para **decidir qué acción tomar** en cada estado. Es clave para **maximizar recompensas a largo plazo**. Es como la estrategia del jugador en un videojuego.

El problema del bandido de múltiples brazos

Imagina una máquina con varios botones (brazos), cada uno con distinta probabilidad de darte una recompensa. Para saber cuál es el mejor, necesitas **probar todos suficientes veces**, pero si **exploras demasiado, pierdes tiempo para explotar el mejor**. Ahí nace…

Dilema Exploración - Explotación

Un clásico: ¿probar nuevas opciones (explorar) o seguir con la que ya conoces que funciona (explotar)?

Explorar mucho = Menos recompensas inmediatas.

Explotar rápido = Riesgo de perder mejores opciones.

Valores Q

Son las "expectativas de recompensa" por tomar una acción en un estado. Pero **saber el valor Q no basta**: el agente debe convertir esos valores en acciones reales mediante una **regla de elección**.

Reglas de elección

Son fórmulas o algoritmos que **transforman valores Q en decisiones**. Te ayudan a decidir entre múltiples opciones.

Regla de Maximización (Greedy)

Selecciona **siempre la acción con mayor Q**.

Problemas:

-
- Las personas **no eligen siempre lo mejor** (variabilidad empírica).
-
- Se puede **quedar atrapado en un máximo local**, sin explorar mejores opciones.
-

Regla épsilon-codiciosa

Solución para el dilema exploración/explotación:

-
- Con **probabilidad (1 -)** elige la mejor acción.
-
- Con **probabilidad** , elige aleatoriamente.
-

Ejemplo:

-
- = 0.1: Explora el 10% del tiempo.
-
- = 0.01: Explora menos, pero puede encontrar mejores resultados a largo plazo.
-

Funciones de respuesta probabilísticas

Modelos de psicología que **explican elecciones aleatorias**, especialmente cuando las diferencias entre opciones no son tan claras. Aquí entran:

Funciones psicométricas

Muestran cómo, cuando las diferencias entre estímulos (o Qs) son pequeñas, las elecciones se vuelven menos consistentes. Se derivan dos modelos:
