



UNIVERSIDAD ALFONSO X EL SABIO

Q-LEARNING PARA PRINCIPIANTES

Germán Llorente y Carlos Puigserver

December 26, 2023

1 Reflexiones sobre el Aprendizaje por Refuerzo

1. **Exploración y Explotación:** El aprendizaje por refuerzo aborda el dilema de exploración y explotación, permitiendo que un agente explore diferentes acciones para descubrir cuáles son más beneficiosas.
2. **Adaptabilidad al Entorno:** El aprendizaje por refuerzo permite a un agente adaptarse a entornos cambiantes, ajustando sus estrategias para maximizar las recompensas.
3. **Optimización Continua:** A diferencia de los condicionales `if...else`, el aprendizaje por refuerzo optimiza continuamente el comportamiento del agente basándose en la retroalimentación recibida.
4. **Manejo de Complejidad:** En entornos complejos, el aprendizaje por refuerzo puede ser más efectivo, ya que no requiere una programación manual de todas las posibles situaciones.
5. **Generalización:** El aprendizaje por refuerzo tiende a generalizar el conocimiento, desarrollando estrategias que pueden aplicarse a situaciones similares.
6. **Eficiencia en Ambientes Dinámicos:** En entornos dinámicos, el aprendizaje por refuerzo puede ser más eficiente que depender de reglas estáticas.

Diferencias con Soluciones más Simples, como Condicionales `if...else`:

1. **Manual vs. Automático:** Las soluciones basadas en estamentos condicionales `if...else` requieren programación manual de reglas, mientras que el aprendizaje por refuerzo automatiza el ajuste de estrategias. Aunque las manuales son muy útiles para códigos simples, puede ser muy tedioso programarlas para códigos complejos. Además, el peso programacional es mayor.

2. **Flexibilidad:** Las reglas `if...else` son estáticas, mientras que el aprendizaje por refuerzo es más flexible y puede adaptarse a cambios sin necesidad de reprogramación. No solo eso, sino que el segundo es mucho más directo que el primero, que requeriría de reprogramación en función del problema.
3. **Capacidad de Aprendizaje:** El aprendizaje por refuerzo tiene la capacidad de aprender de manera autónoma y mejorar con el tiempo. Por otro lado, tal y como ya se ha explicado, los condicionales son estáticos.
4. **Manejo de Incertidumbre:** El aprendizaje por refuerzo puede lidiar mejor con la incertidumbre explorando y aprendiendo en lugar de depender completamente de reglas predefinidas.
5. **Generalización:** Mientras que el aprendizaje por refuerzo tiende a generalizar el conocimiento, las soluciones `if...else` pueden requerir reglas específicas para cada situación.
6. **Escalabilidad:** El aprendizaje por refuerzo puede ser más escalable en entornos complejos ya que no requiere la especificación manual de todas las condiciones posibles.

2 Decisiones y Desafíos en el Aprendizaje por Refuerzo en Frozen Lake

2.1 Decisiones Tomadas:

1. **Inicialización del Q-table:** Se inicializó el Q-table con ceros para representar los valores iniciales de las acciones en cada estado. Luego ya estos valores se cambiarían con el desarrollo del proceso.
2. **Visualización:** Se utilizó Matplotlib para visualizar el rendimiento del agente a lo largo de los episodios.

2.2 Desafíos Enfrentados:

1. **Determinación de Hiperparámetros:** La elección adecuada de hiperparámetros es crucial y puede requerir experimentación.

2. **Manejo de Ambientes Resbaladizos:** Adaptar el enfoque para ambientes resbaladizos puede requerir ajustes adicionales. Cabe destacar que mientras el lago helado nos daba una eficacia perfecta, la del lago helado resbaladizo era menor a 30 por ciento. Esto se debe a que la complejidad es mayor, y es posible que haya habido algún problema en el desarrollo.
3. **Aprendizaje Continuo:** Implementar el aprendizaje continuo y actualizar el Q-table a lo largo de múltiples episodios puede requerir consideraciones especiales.
4. **Visualización Dinámica:** La visualización dinámica del agente puede ser desafiante debido a restricciones de entornos de ejecución.
5. **Manejo de Recompensas y Éxito:** Definir adecuadamente recompensas y el criterio de éxito son aspectos críticos para la efectividad del aprendizaje.

En resumen, el desarrollo implica equilibrar la teoría del aprendizaje por refuerzo con la implementación práctica, ajustando parámetros y abordando desafíos específicos del entorno seleccionado. En este proyecto, hemos aplicado estos conceptos al ejemplo del lago congelado, consiguiendo una precisión perfecta. Sin embargo, el lago congelado ha sido más tedioso de realizar, y no se ha conseguido un resultado tan óptimo.