Texto, Pizarra

Descripción generada automáticamente

**Licenciatura en Inteligencia Artificial y Robótica**

**Trabajo practico Nro. 3**

**Seminario de practica en Inteligencia Artificial y Robótica**

**Desarrollo de Agente de Toma de Decisiones**

**Autor: Donadío Gerónimo**

**Fecha de entrega: 27/10/2024**

El experimento tuvo como objetivo diseñar un agente de toma de decisiones que pudiera aprender de su experiencia en tres contextos distintos a lo largo de 100,000 iteraciones en total, 10,000 por cada contexto. El agente registraba en una base de datos las probabilidades, contexto, prioridad y resultados de cada acción y utilizaba OpenCL para acelerar el procesamiento de ajustes en las probabilidades. La metodología de aprendizaje del agente se basaba en un sistema de refuerzo y penalización que aumentaba la probabilidad de seleccionar acciones con resultados exitosos y disminuía la probabilidad de aquellas con fracasos. En el primer contexto el agente alternó inicialmente entre la acción uno y la acción tres, pero hacia el final mostró una preferencia clara por la acción tres debido a la mayor frecuencia de éxitos acumulados. Esto indica que el agente aprendió a priorizar esta acción gracias al refuerzo positivo. En el segundo contexto predominó la acción dos, la cual tuvo resultados mixtos de éxitos y fracasos; no obstante, los éxitos fueron lo suficientemente frecuentes como para que la probabilidad de esta acción se mantuviera alta, mostrando que la penalización por fracasos fue efectiva, pero no al punto de eliminar opciones con buenos resultados parciales. En el tercer contexto la acción cuatro fue la opción predominante a pesar de varios fracasos, y aunque logró algún éxito aislado en las iteraciones avanzadas este éxito ocasional permitió que la probabilidad de esta acción no disminuyera por completo, dando lugar a una exploración continua y controlada.

En general el agente demostró una capacidad de adaptación efectiva, lo cual refleja que el sistema de refuerzo y penalización funcionó según lo esperado promoviendo acciones exitosas y permitiendo una exploración moderada de alternativas en presencia de resultados neutros o fracasos ocasionales. Esto es ideal para sistemas de toma de decisiones donde el historial de éxitos debe guiar el comportamiento sin restringir opciones de mejora potencial. Como observaciones finales se sugiere que para futuros experimentos se podría fortalecer la penalización para fracasos repetidos y así fomentar una mayor diversidad en las elecciones en contextos de fracasos frecuentes. También se podría implementar un sistema de refuerzo dinámico donde los éxitos acumulados en una misma acción aumenten su probabilidad de manera exponencial, acelerando la selección de la mejor opción en cada contexto. Un enfoque de enfriamiento donde se reduzca gradualmente la exploración también podría mejorar la precisión en contextos con alta incertidumbre. Revisar las probabilidades finales de cada acción al término de las iteraciones proporcionaría un punto de referencia útil para futuras pruebas del agente permitiendo que retome el aprendizaje acumulado sin necesidad de reiniciar sus probabilidades.

En conclusión, el agente ha mostrado un buen balance entre exploración y explotación logrando una selección de acciones adaptativa y efectiva en cada contexto. Los resultados obtenidos reflejan una toma de decisiones precisa impulsada por el aprendizaje de su historial de resultados. Con estos ajustes el agente podría afinar aún más su capacidad de adaptación a situaciones complejas optimizando la precisión en la selección de acciones exitosas en cada contexto