

Nombres: Diego Martínez, José Gregorio Contreras Suárez
Profesor: Jesús Pérez
Asignatura: Sistemas Computacionales A2023

Lecture 7: DynaQ y DynaQ+

Comparación del desempeño de los Agentes DynaQ y DynaQ+ en el entorno de Blocks-v0

Experimento 1:

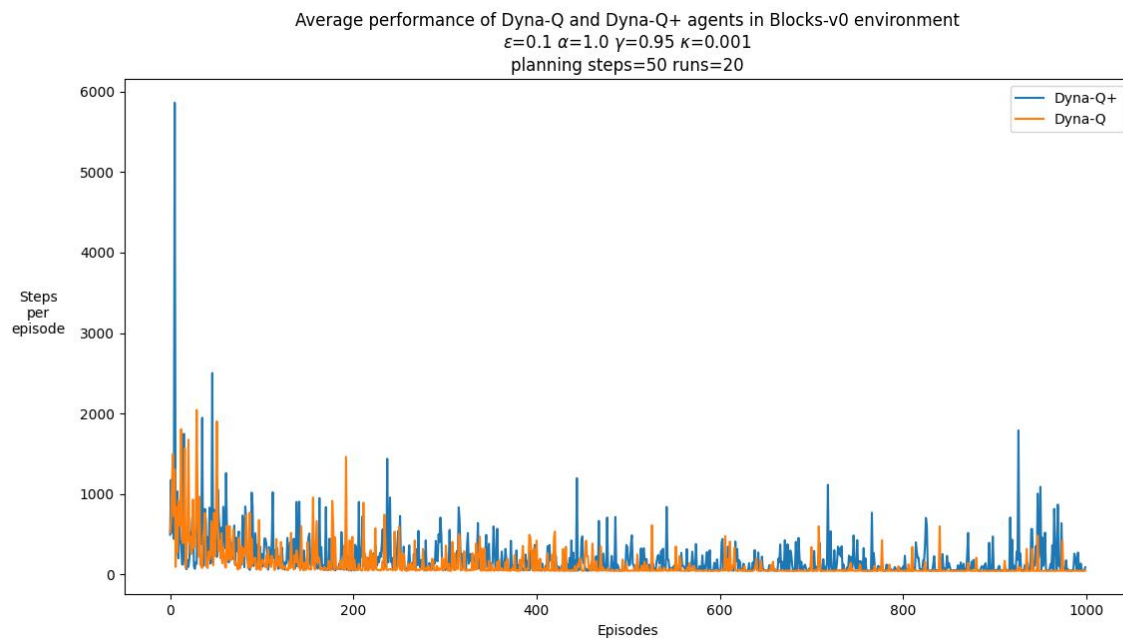


Figura 1

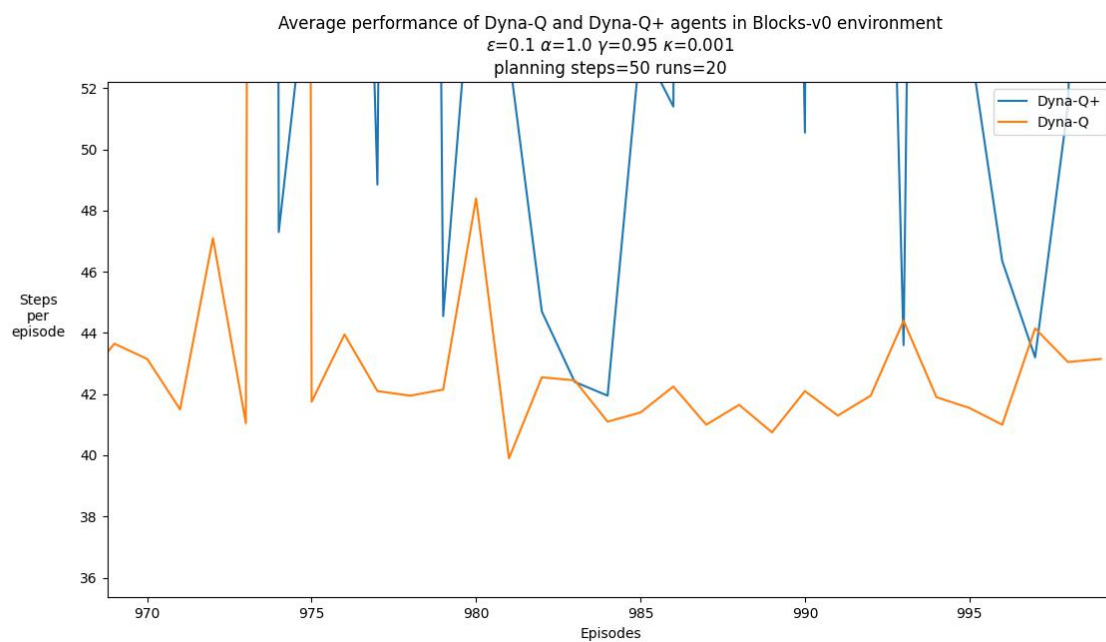


Figura 2

De la figura 1 se puede observar que el promedio de la cantidad de pasos por episodio es menos de 10000 y que tiende a un valor cercano a 40 como se muestra en la figura 2. Es evidente por los picos del gráfico, que el agente Dyna-Q Plus tuvo un mejor desempeño que el de Dyna-Q, alcanzando picos mucho más bajos desde un comienzo, y estabilizándose en menor cantidad de episodios.

Experimento 2:

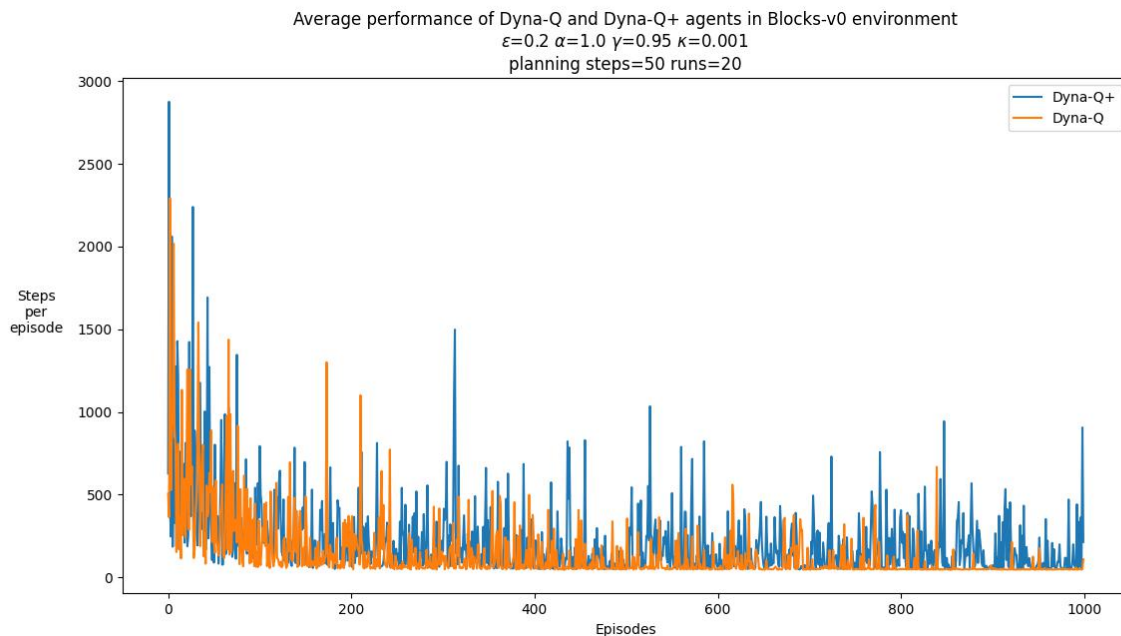


Figura 3

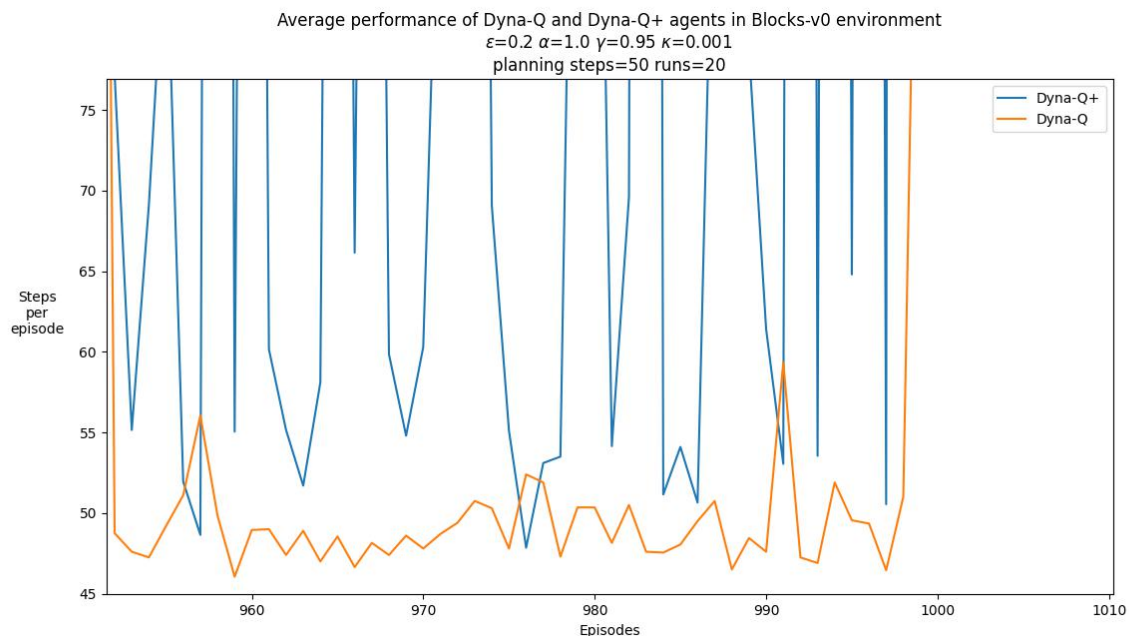


Figura 4

El gráfico de la figura 3 es muy similar a la de la figura 1, el cambio que más se aprecia a simple vista es que el promedio de pasos por episodios es menos de 3000, y como son valores promedio luego de varias corridas da a entender que el valor de ϵ influye directamente en la cantidad de pasos de un episodio, lo cual tiene sentido porque cómo el valor de ϵ

cambió a 0.2, el agente tenía más chance de explorar y por eso la cantidad de pasos del episodio es menor.

Experimento 3:

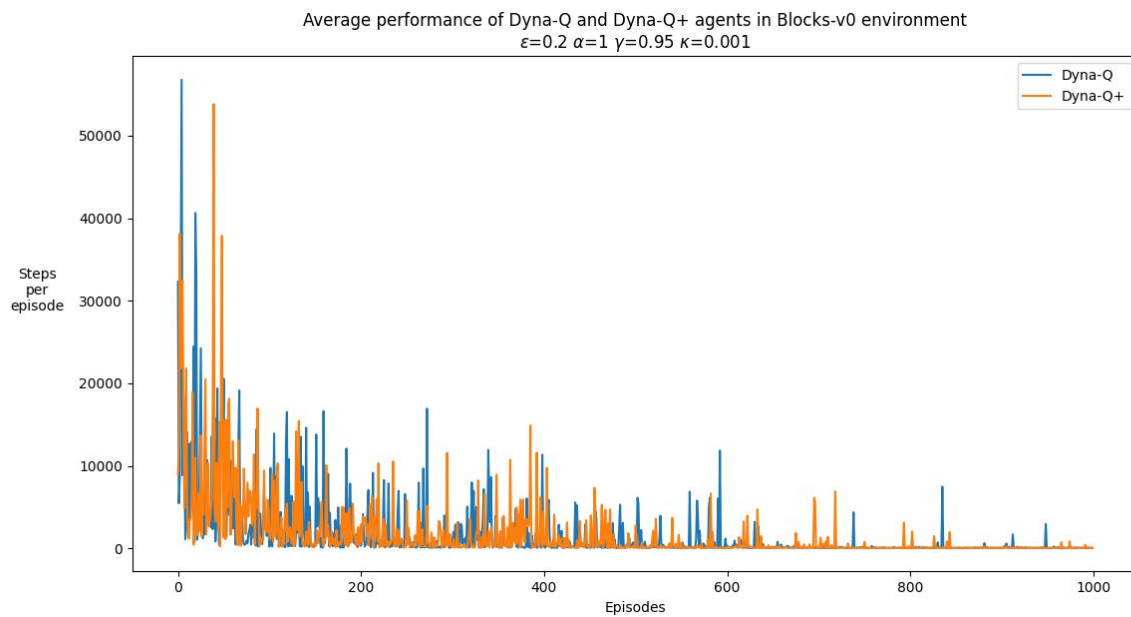


Figura 5

Experimento 4:

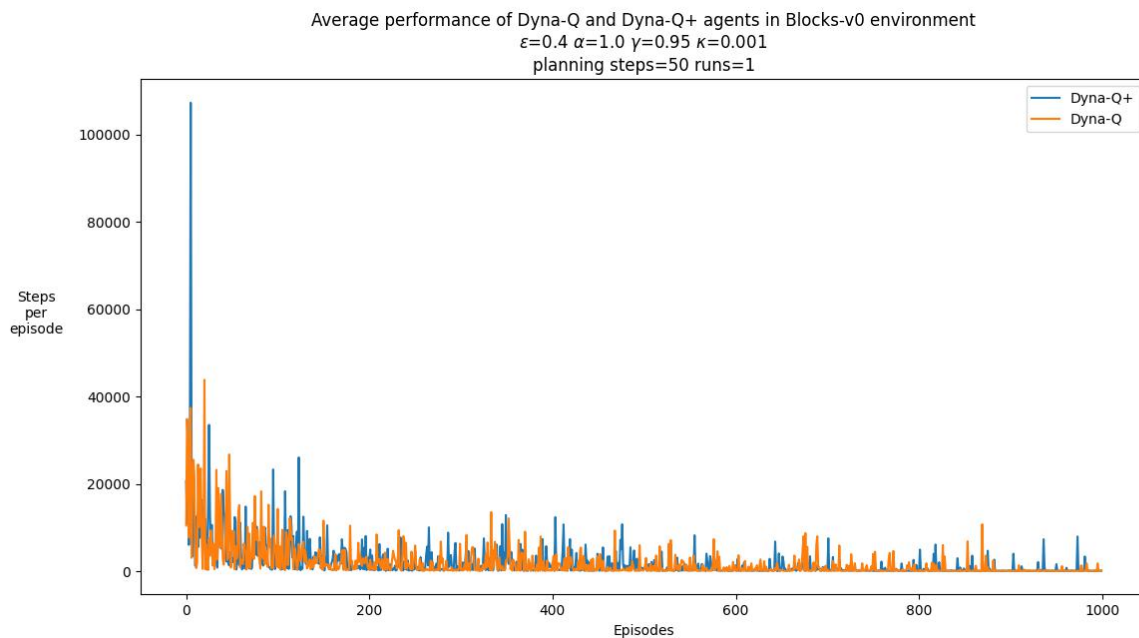


Figura 6

En estos dos experimentos no se promediaron los valores de la cantidad de pasos por episodio, porque solo se hizo una sola corrida.

Sin embargo, el análisis es similar a los experimentos anteriores, solo que por la aleatoriedad, al comienzo la cantidad de pasos por episodio puede ser gigante. En el caso de la figura 5 la

cantidad de pasos por episodio máxima fue mayor de 50000, y en el de la figura 6 fue mayor de 100000.

Lo que quiere decir que aunque no se hayan realizado varias corridas, mientras más grande sea el valor de ϵ , mayor chance tendrá para explorar, pero también menor para explotar, y por eso en la figura 6 la cantidad de pasos máxima fue mucho mayor que la de la figura 5.

No se realizaron experimentos variando el valor de α y γ , pero por los resultados obtenidos de los experimentos podemos decir que variando el valor de α y de γ tendremos un comportamiento similar, pero que puede ser más suave la transición a la convergencia o más lenta/rápida.

Cómo conclusión, el agente Dyna-Q Plus tiene mejor desempeño en el entorno de Blocks-v0 que el agente Dyna-Q.