

Reinforcement Learning

1. Jelaskan cara kerja dari algoritma Q-Learning dan SARSA!

⇒ Baik Q-Learning maupun SARSA, keduanya menggunakan *q-table* untuk “belajar” dan melakukan pengambilan keputusan. *Q-table* ini berisi *Q-value* dimana semakin besar *Q-value* pada suatu pilihan keputusan, maka pilihan keputusan tersebut adalah yang terbaik.

⇒ Perbedaan signifikan antara Q-Learning dan SARSA adalah yaitu Q-Learning merupakan algoritma yang *off-policy* dalam artian dia *greedy* dan terkadang suka bereksplorasi. Sedangkan SARSA merupakan algoritma yang *on-policy* yang berarti selalu mengambil keputusan sesuai “kebijakan” yang ada dan konsisten demi menghindari resiko.

2. Bandingkan hasil dari kedua algoritma tersebut, bagaimana hasil perbandingannya? Jika ada perbedaan, jelaskan alasannya!

⇒ Berdasarkan hasil percobaan, didapatkan bahwa algoritma SARSA selalu memerlukan langkah yang lebih sedikit untuk mencapai 500 poin daripada algoritma Q-Learning. Hal tersebut tentunya masuk akal karena *environment* disini hanyalah 1 dimensi dan peraturannya cukup *straightforward* yaitu jika menyentuh titik nol akan mendapat -100 poin dan jika mencapai apel akan +100 poin, maka hal tersebut sangat menguntungkan bagi SARSA yang selalu mengikuti “kebijakan”, berbeda dengan algoritma Q-Learning yang *off-policy* dan terkadang suka bereksplorasi. Karena apel selalu berada di sebelah kanan, maka melakukan “eksplorasi” dan mencoba langkah ke kiri merupakan hal yang tidak berguna, oleh karena itu dalam kasus ini algoritma SARSA lebih unggul dibandingkan algoritma Q-Learning. Meskipun demikian, kedua algoritma tersebut sama-sama berhasil memenangkan permainan dalam 1x percobaan (tidak pernah mendapatkan -200 poin), tetapi tidak menutup kemungkinan bisa saja mengalami kekalahan tergantung *threshold* nilai epsilon untuk melakukan eksplorasinya.