

Optymalizacja systemu sygnalizacji świetlnej w oparciu o przepływowy model ruchu pojazdów.

Michał Lis

25 czerwca 2019

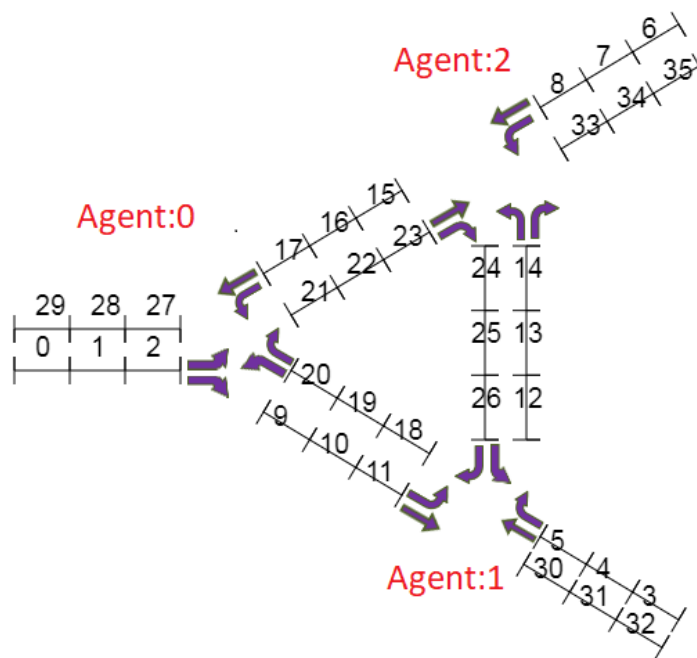
Spis treści

1	środowiska symulacyjne qwe	5
1.1	Srodowisko 4	5
1.2	Uczenie srodowiska 4	6
1.2.1	Podójście 1	6

Rozdział 1

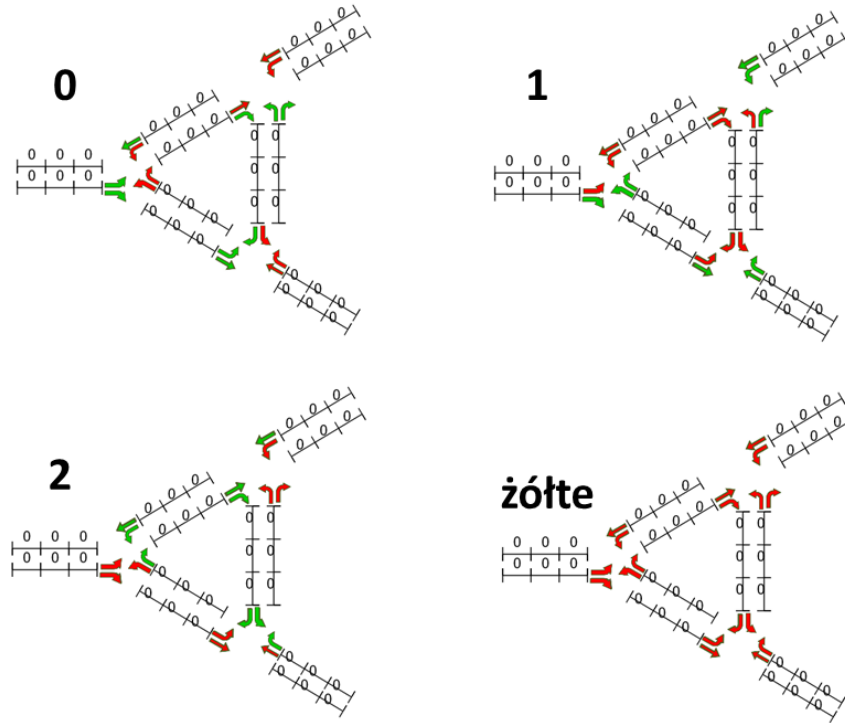
środowiska symulacyjne qwe

1.1 Środowisko 4



Rysunek 1.1: środowisko 4

środowisko posiada 12 jednokierunkowych dróg. Każda droga ma 3 odcinki co daje w sumie 36 odcinków (są numerowane od 0 co widać na rysunku 1.1). W sieci dróg znajdują się 3 skrzyżowania. Do każdego z nich jest przypisany agent, który odpowiada za sterowanie sygnalizacją świetlną. Pojazdy w jednym interwale czasowym pokonują jeden odcinek. Na skrzyżowaniach w przypadku zielonego światła przejeżdża maksymalnie 10 pojazdów w jedną stronę.



Rysunek 1.2: środowisko 4 - fazy świateł

Fazy świetlne: Każde skrzyżowanie posiada 4 fazy świetlne przedstawione powyżej. Fazy 0, 1 i 2 posiadają pewne zielone światła. Agent podejmuje decyzję o zmianie tych trzech faz. Zmiana faz świateł nie jest natychmiastowa i następuje dopiero po 2 interwałach czasowych fazy żółtych świateł. Agent może podjąć akcję a należącą do $[0,1,2]$ w przypadku gdy obecna faza f należy do $[0,1,2]$. W pozostałym przypadku agent jest zobowiązany do przekazania akcji 'yellow'.

1.2 Uczenie srodowiska 4

1.2.1 Podejście 1

Każdy z trzech agentów jako stan przyjmuje 10 elementową tablicę. 9 elementów to ilości pojazdów na odcinkach będących przed skrzyżowaniem przypisanym do agenta. Tablicę uzupełnia wartość obecnej fazy. Nagrody dla danego stanu są przyznawane jako suma pojazdów, które przejechały przez skrzyżowanie w trakcie najbliższych 4 interwałów czasowych. Początkowo przeprowadzana jest symulacja 100 epizodów z czego każdy trwa 90 interwałów czasowych. Do uczenia agent zapamiętuje jedynie te stany, których faza to 0, 1 lub 2. Nieistotne w procesie uczenia są zatem stany z fazą 'yellow' gdyż agent ma tylko 1 możliwą decyzję do podjęcia. Następnie sieć neuronowa przyjmująca na wejście stan - 10 elementowy wektor. Sieć na wyjściu zwraca 3-elementowy wektor określający przewidziane nagrody dla akcji podjętej w zadanym stanie. Podsumowując dla wybranego agenta:

- **Stanem** są ilości pojazdów przed skrzyżowaniem oraz aktualna faza świetlna
- **Nagroda** w chwili t jest suma pojazdów, które przejechały przez skrzyżowanie w trakcie najbliższych 4 interwałów czasowych czyli do momentu $t+4$.
- **Dane** są generowane poprzez przeprowadzenie 100 symulacji (każda ma 90 interwałów czasowych).
- **Sieć neuronowa** na podstawie wygenerowanych danych przewiduje najlepszą akcję dla obecnego stanu
- Końcowa symulacja zostaje przeprowadzona wedle przewidzianych przez sieć neuronową najlepszych akcji