

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

Michał Lis

30 sierpnia 2019

Spis treści

1	Wprowadzenie	5
2	Cel i zakres pracy	7
3	Makroskopowy model ruchu	9
3.1	Klasyfikacja modeli ruchu drogowego	9
3.2	Wstęp	9
3.3	Rozwój gęstości ruchu na drodze	10
3.4	CTM - Cell Transmission Model	10
3.4.1	Wprowadzenie	10
3.4.2	Przedstawienie modelu	10
3.4.3	Przepływ pojazdów na skrzyżowaniu	13
3.5	Gottlich dyskretyzacja - to jest bardzo źle	14
4	Model sieci dróg	15
4.1	Wstęp	15
4.2	Wektor stanu drogi	15
4.2.1	Przykład	15
4.3	Rozwój wektora stanu drogi i wprowadzenie macierzy systemu	16
4.3.1	Przykład bez zatoru	16
4.4	Wektor stanu sieci dróg	17
4.4.1	Przykład	17
4.5	Rozwój wektora stanu sieci dróg w oparciu o CTM	18
4.5.1	Przykład macierzy prawdopodobieństwa	19
4.6	Rozwój wektora stanu sieci dróg z przepływem CTM	19
4.6.1	Przykład	20
4.7	Wprowadzenie sygnalizacji świetlnej	22
4.7.1	Przykład macierzy sygnalizacji świetlnej	22
4.7.2	Macierz systemu uwzględniająca sygnalizację świetlną	23
4.7.3	Przykład rozwoju wektora stanu	23
4.8	Wprowadzenie źródeł ruchu	24
4.8.1	Przykład	25
4.9	Wprowadzenie modelu ruchu CTM	26
4.9.1	Przykład zatoru na pojedynczej drodze	27
4.9.2	Przykład zatoru na skrzyżowaniu	28

5	Środowiska symulacyjne i ich nauka	31
5.1	Środowisko 1	31
5.1.1	Sygnalizacje świetlne	32
5.1.2	Przestrzeń decyzyjna	32
5.1.3	Zatory drogowe	32
5.1.4	Uczenie	32
5.2	Środowisko 2	34
5.2.1	Sygnalizacje świetlne	34
5.2.2	Przestrzeń decyzyjna	34
5.2.3	Zatory drogowe	35
5.2.4	Uczenie	35
5.3	Środowisko 3	38
5.3.1	Sygnalizacje świetlne	38
5.3.2	Przestrzeń decyzyjna	39
5.3.3	Zatory drogowe	40
5.3.4	Uczenie	41
5.4	Środowisko 4 - kampus B Politechniki Łódzkiej	43
5.4.1	Sygnalizacje świetlne	43
5.4.2	Przestrzeń decyzyjna	45
5.4.3	Zatory drogowe	45
5.4.4	Uczenie	45
5.4.5	Modyfikacja środowiska	45
6	Przegląd metod optymalizacyjnych	47
6.1	Kategorie uczenia maszynowego	47
6.2	Uczenie ze wzmocnieniem	47
6.3	Metoda Monte Carlo On-Policy	50
6.4	Metoda Monte Carlo On-Policy z użyciem sieci neuronowej	51
6.4.1	Wariant z fazą żółtych świateł	52

Rozdział 1

Wprowadzenie

Problem zatłoczonych ulic staje się coraz bardziej powszechny na całym świecie. W ostatnich latach obserwowany jest ogromny wzrost ilości pojazdów na drogach. Według danych statystycznych firmy *Statista* liczba zarejestrowanych pojazdów na świecie w roku 2006 wynosiła 947 tysięcy [13]. W 2015 roku na świecie jeździło już 1282 tysięcy pojazdów. Wzrost przez te 9 lat był niemalże liniowy. Co roku rejestrowano około 40 tysięcy nowych samochodów rocznie, co wyznacza stopę wzrostu liczby pojazdów na poziomie 3,7% w skali roku.



Rysunek 1.1: Liczba pojazdów na świecie

W Polsce wzrost ilości pojazdów w latach 2006 - 2015 był jeszcze większy [12]. W 2006 roku według GUS w Polsce było zarejestrowanych 13,4 miliona samochodów osobowych. W 2015 roku ich liczba wynosiła już 20,7 miliona, co wyznacza 5 procentowy roczny wzrost ilości pojazdów. Najbardziej zatłoczonym polskim miastem jest Łódź. Według rankingu firmy *TomTom* Łódź zajmuje bardzo wysokie 5 miejsce na świecie i 1 w Europie pod względem zatłoczenia dróg [15]. Oprócz Łodzi w pierwszej setce najbardziej zatłoczonych miast świata

są inne polskie miasta: Lublin(34), Kraków(48), Warszawa(50), Wrocław(63), Poznań(69), Bydgoszcz(83). Zatory drogowe są jednak problem całej Europy. Spośród 100 najbardziej zatłoczonych miast świata aż 45 znajduje się w Europie. W 2008 roku Unia Europejska oszacowała, iż koszty zatłoczenia dróg kształtują się na poziomie 0,9% – 1,5% PKB unijnego [1]. Następny raport z 2017 roku może napawać optymizmem, gdyż przedstawione w nim wyliczenia określiły jedynie 0,77% straty całkowitego PKB wspólnoty [2]. Ten sam raport ocenia koszty zatorów komunikacyjnych w Polsce na poziomie 1,2% polskiego PKB. Problemy zatorów komunikacyjnych w miastach są o tyle trudniejsze do rozwiązania niż poza miastem, ponieważ na terenach zurbanizowanych często brakuje miejsca na wybudowanie dróg o większej przepustowości. Rozwiązaniem może być wprowadzenie większej ilości sygnalizacji świetlnej. Istotną kwestią jest optymalizacja ustawień sygnalizacji świetlnej. Praca jest poświęcona przede wszystkim temu aspektowi. Zostaną w niej także przedstawione matematyczne modele sieci drogowej oraz ich komputerowe symulacje.

Rozdział 2

Cel i zakres pracy

W rozdziale 4 przedstawiony zostanie model sieci dróg. Stworzony przez autora pracy został program symulujący ruch drogowy. Jest on napisany w języku Python (wersja 3.6). Oddzielna warstwa wizualizacyjna utworzona w typescriptowym frameworku Angular w wersji 6. Grafiki przedstawiające środowiska drogowe są narysowane przy pomocy tagów canvas. Kroki potrzebne do uruchomienia programu zostały przedstawione w TODO rozdział do napisania na sam koniec. Bezpośrednim celem pracy jest uzyskanie płynnego ruchu przy możliwie dużej ilości pojazdów wjeżdżających do rozważanego środowiska. Ostatnie środowisko testowe w pracy przedstawia sieć dróg wokół kampusu B Politechniki Łódzkiej i to głównie na optymalizacji tego środowiska będzie skupiona uwaga. Wszystkie poprzednie sieci dróg są mniej skomplikowane i służą głównie przy przedstawieniu modelu i procesu optymalizacji. Środowiska będą skonstruowane w pełni zgodne z makroskopowym modelem ruchu, który nie rozróżnia indywidualnych pojazdów, a traktuje je jako ilościowo. Sam makroskopowy model ruchu zostanie przedstawiony w rozdziale 3. Początkowy model będzie bardzo prosty z pominięciem większości aspektów. W każdej kolejnej sekcji model będzie stopniowo rozwijany.

Rozdział 3

Makroskopowy model ruchu

3.1 Klasyfikacja modeli ruchu drogowego

Modele ruchu drogowego mają na celu ukazanie rzeczywistego przepływu pojazdów w sposób czysto matematyczny. Ważnym kryterium doboru modelu jest przystępność jego implementacji informatycznej. Powszechnie klasyfikuje się 3 podejścia modelowe dla omawianego problemu [4] - makroskopowy, mezoskopowy oraz mikroskopowy. Czasem [10] wyróżnia się także czwarte podejście - submikroskopowe. Jest to podział ze względu na poziom modelu. Najniższy poziom i najbardziej dokładny model gwarantuje podejście mikroskopowe. Rozważa ono pojazdy indywidualnie w czasoprzestrzeni. Przyspieszenie pojazdu jest wyliczane na podstawie dynamiki (prędkości, przyspieszenia) i pozycji pojazdu bezpośrednio przed nim. Model mezoskopowy zapewnia indywidualne rozróżnienie pojazdów, jednak ich zachowanie jest wyliczane na danych zagregowanych [5]. Przykładowo pojazdy są podzielone na grupy podróżujące od pewnego punktu startowego do punktu końcowego. Inne modele [3] mezoskopowe wyliczają dynamikę ruchu na podstawie aktualnego zatłoczenia drogi. Poziom mezoskopowy jest obliczeniowo bardziej opłacalny od mikroskopowego. Wiele symulatorów stosujących model mezoskopowy oferuje symulację w czasie rzeczywistym dla sieci dróg całego miasta [16]. Ideą modelu makroskopowego jest traktowanie ruchu ulicznego identycznie jak ruchu cieczy lub gazów. Po raz pierwszy w roku 1956 M. J. Lighthill i G. B. Whitham [11] przedstawili pomysł przyrównania ruchu ulicznego na zatłoczonych drogach do przepływu wody w rzekach. Z tego powodu nie uznajemy w nim pojazdów jako niepodzielne jednostki. Pojazd można przyrównać do pewnej ilości wody w rzece. Jest to najmniej kosztowny obliczeniowo model. Właśnie w modelu makroskopowym zostało stworzone środowisko symulacyjne. Szczegóły modelu są przedstawione w następnym podrozdziale.

3.2 Wstęp

Jako zmienną stanową makroskopowego modelu ruchu zazwyczaj wybierana jest gęstość ruchu [4, 7]. Formalnie gęstość można rozumieć jako czynnik definiujący dynamikę ruchu. Im większa jest gęstość tym mniejsza prędkość ruchu. Gęstość ruchu [9] jest przedstawiona jako iloraz ilości pojazdów znajdujących się na pewnym odcinku i długości tego odcinka drogi. Prawidłowe jest także uznanie ilości pojazdów na pojedynczym odcinku drogi jako zmienną

stanową. Należy jednak pamiętać, że pojazdy w ruchu makroskopowym nie są nierozłączną jednostką. Pojazd w analogii do przepływu wody w rzece jest pewną jednostką wody. Wielokrotnie w symulacjach zdarza się, że na odcinku jest zatem niecałkowita liczba pojazdów.

3.3 Rozwój gęstości ruchu na drodze

Makroskopowy model ruchu jest oparty o równanie różniczkowe (3.2) wraz z warunkiem początkowym (3.1) [7]. W analogii do przepływu wody w rzece gęstość ruchu można utożsamiać z polem powierzchni przekroju poprzecznego rzeki, co dla stałej szerokości rzeki - upraszcza się do wysokości wody w rzece.

Dla ustalonej pojedynczej drogi zmianę gęstości ruchu definiuje następujący układ równań:

$$\begin{cases} p(x, 0) = p_0(x) & (3.1) \\ \frac{\partial p(x, t)}{\partial t} + \frac{\partial f(p(x, t))}{\partial x} = 0 & (3.2) \end{cases}$$

Gdzie $p(x, t)$ to gęstość ruchu w punkcie x i czasie t . Wartość funkcji gęstości należy do przedziału $[0, p^{max}]$.

Równanie (3.1) zakłada istnienie pewnej z góry nałożonej początkowej gęstości drogi $p_0(x)$.

Równanie (3.2) określa wedle założeń modelu makroskopowego [11] rozwój gęstości ruchu na drodze. Funkcja płynności ruchu f powinna być wklęsła [7].

3.4 CTM - Cell Transmission Model

3.4.1 Wprowadzenie

W tym rozdziale zostanie przedstawiony model CTM będący dyskretyzacją pierwotnego modelu makroskopowego przepływu [11]. Jego zmienną stanową jest ilość pojazdów. Model został opisany w artykule Carlosa Daganzo [6] w 1992. Ta pozycja jest podstawą wszelkich wyliczeń w tym rozdziale.

3.4.2 Przedstawienie modelu

Siatka czasowa i przestrzenna drogi

Dyskretny charakter modelu CTM obliguje do określenia siatki czasowej i przestrzennej drogi. Dla par czasu i miejsc należących do tych dwóch siatek będą określone zmienne stanu.

Siatka czasowa jest zdefiniowana jako skończony ciąg liczb naturalnych:

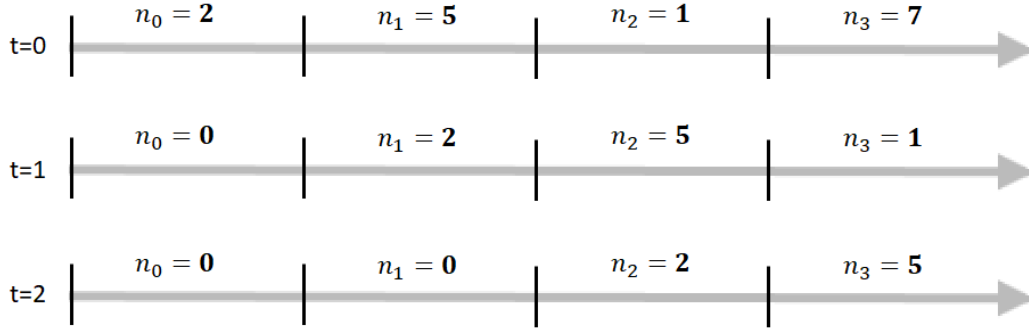
$$(0, 1, \dots, T). \quad (3.3)$$

Następnie zdefiniowana zostanie siatka przestrzenna drogi e . Ustalona droga e to odcinek $[0, L]$. Zostaje ona podzielona na L odcinków o równej długości. **Siatka przestrzenna** drogi to ciąg odcinków:

$$([0, 1), [1, 2), \dots, [L - 1, L])$$

Przepływ pojazdów na drodze

Niech dana będzie droga e i pewien odcinek j na tej drodze. Zmienna stanowa w modelu CMT jest oznaczana jako $x_j(t)$ i określa ona ilość pojazdów na odcinku j w chwili t . Przepływ pojazdów opiera się o założeniu, iż przy braku zatorów pojazdy pokonują w jednym interwale czasowym jeden odcinek drogi. Przykładowy przepływ jest następujący:



Rysunek 3.1: Przykładowy przepływ bez zatorów w modelu CTM

Łatwo zauważyć, że w tym przypadku

$$x_{j+1}(t+1) = x_j(t) \quad (3.4)$$

Wprowadzenie zatorów

W celu wprowadzenia pojęcia zatoru zostaną zdefiniowane następujące zmienne:

- $N_j(t)$ - maksymalna ilość pojazdów, które mogą znajdować się na odcinku j w chwili t .
- $Q_j(t)$ - maksymalna ilość pojazdów, które mogą napłynąć do odcinka j w momencie $t+1$.

Rozwój zmiennej stanowej jest zadany poprzez następujące równanie:

$$x_j(t+1) = x_j(t) + y_j(t) - y_{j+1}(t) \quad (3.5)$$

Gdzie:

- $y_j(t)$ to ilość pojazdów napływających do odcinka j w chwili $t+1$.
- $y_{j+1}(t)$ to ilość pojazdów napływających do odcinka $j+1$ w chwili $t+1$. Jest to zatem także ilość pojazdów opuszczających odcinek j w interwale $[t, t+1)$.

Przepływ pojazdów $y_j(t)$ jest zdefiniowany jako minimum z następujących trzech wartości:

- $x_{j-1}(t)$ - ilość pojazdów na poprzedzającym odcinku (t.j. $j-1$)
- $Q_i(t)$ - maksymalna ilość pojazdów, które mogą napłynąć do odcinka i w chwili $t+1$.

- $N_i(t) - x_j(t)$ - ilość wolnych miejsc na odcinku j w chwili t

Co wyraża się wzorem:

$$y_j(t) = \text{Min}\{x_{j-1}(t), Q_i(t), N_i(t) - x_i(t)\} \quad (3.6)$$

Źródło i ujście ruchu

Powyższe wzory potrzebują ustaleń odnośnie źródła i ujścia ruchu. Bez nich wzór (3.6) nie ma sensu dla krańcowych odcinków $j = 0$ oraz $j = L$. Dla ułatwienia pogrubione zostały ustalone w tym rozdziale wartości zmiennych źródła i ujścia ruchu.

Źródło ruchu może być traktowane jako odcinek o indeksie $j = -1$. Jedyną wartością, która musi być ustalona dla źródła jest ilość pojazdów w nim, czyli $x_{j-1}(t)$. Przyjęte zostaje, że $\mathbf{x}_{j-1}(\mathbf{t}) = \infty$ - wtedy napływ pojazdów do odcinka $j = 0$ jest ograniczony z dołu już tylko przez $Q_0(t)$ oraz $N_0(t) - x_0(t)$.

Ujście ruchu z kolei może być traktowane jako odcinek $i = L + 1$. Istotne w kontekście wzoru (3.6) są następujące wartości:

- $Q_{L+1}(t)$ - ilość pojazdów, które mogą opuścić odcinek L w chwili t . Ustalono jest, że jest to dowolna ilość, wtedy $\mathbf{Q}_{L+1}(\mathbf{t}) = \infty$
- $N_{L+1}(t)$ - ilość pojazdów, które może pomieścić ujście. Ustalono jest, że $\mathbf{N}_j(\mathbf{t}) = \infty$, gdyż ujście może zawierać nieograniczoną ilość pojazdów.
- $x_{L+1}(t)$ - ilość pojazdów w ujściu. Służy jako licznik pojazdów, które opuściły układ.

Przykład przepływu na pojedynczej drodze

Zostają ustalone następujące parametry implementacji przykładowego modelu

$$L = 8$$

$$N_i(t) = 15 \text{ dla dowolnego } i, t$$

$$Q_i(t) = \begin{cases} 1 & \text{dla } i = 5, t \in \{0, \dots, 7\} \\ 5 & \text{dla } i = 5, t \in \{7, \dots, 20\} \\ 4 & \text{dla pozostałych przypadków} \end{cases} \quad (3.7)$$

$$(3.8)$$

$$(3.9)$$

Przeprowadzona zostanie symulacja przy warunku początkowym określającym liczebność pojazdów na poszczególnych odcinkach. Niech zatem $x_i(0) = 3$ dla $i \in \{0, \dots, 8\}$. Wtedy wyniki symulacji przedstawia poniższa tabela.

		Odcinek													
		-1	0	1	2	3	4	5	6	7	8	9	Q5		
Legenda Źródło ruchu Odcinek drogi Ujście ruchu	Czas	0	999	3	3	3	3	3	3	3	3	3	0	1	
		1	999	4	3	3	3	5	1	3	3	3	3	1	
		2	999	4	4	3	3	7	1	1	3	3	6	1	
		3	999	4	4	4	3	9	1	1	1	3	9	1	
		4	999	4	4	4	4	11	1	1	1	1	12	1	
		5	999	4	4	4	4	14	1	1	1	1	13	1	
		6	999	4	4	4	7	14	1	1	1	1	14	1	
		7	999	4	4	4	4	10	14	1	1	1	15	5	
		8	999	4	4	4	4	13	10	5	1	1	16	5	
		9	999	4	4	6	11	9	6	4	1	1	17	5	
		10	999	4	4	6	11	8	7	4	4	1	18	5	
		11	999	4	4	6	11	7	8	4	4	4	19	5	
		12	999	4	4	6	11	6	9	4	4	4	23	5	
		13	999	4	4	6	11	5	10	4	4	4	27	5	
		14	999	4	4	6	11	4	11	4	4	4	31	5	
		15	999	4	4	6	11	4	11	4	4	4	35	5	
		16	999	4	4	6	11	4	11	4	4	4	39	5	
		17	999	4	4	6	11	4	11	4	4	4	43	5	
		18	999	4	4	6	11	4	11	4	4	4	47	5	
		19	999	4	4	6	11	4	11	4	4	4	51	5	
		20	999	4	4	6	11	4	11	4	4	4	55	5	

Rysunek 3.2: Tabela ilości pojazdów

3.4.3 Przepływ pojazdów na skrzyżowaniu

Wprowadzona zostanie modyfikacja funkcji $y_j(t)$ określającej ilość pojazdów wjeżdżających na odcinek j w chwili $t+1$. Do tej pory dotyczyła ona jedynie ilości pojazdów które napływają do odcinka j z odcinka $j-1$. W przypadku sieci dróg większej niż 1 droga należy rozważyć możliwość wjazdu z innych odcinków niż $j-1$. Niech zatem:

$$y_{ij}(t) = \text{Min}\{p_{ij}x_i(t), Q_j(t), N_j(t) - n_j(t)\}$$

Określa ilość pojazdów, które wjeżdżają do odcinka i z odcinka j w chwili $t+1$. Ustalona z góry wartość p_{ij} określa jaką część pojazdów z odcinka i ma zamiar wjechać na odcinek j . Dla $i=j$ oraz niemożliwych manewrów oczywiście $a_{ij}=0$.

Niech Ω będzie zbiorem wszystkich odcinków w układzie. Wtedy suma pojazdów wyjeżdżających z odcinka i to:

$$y_{i*}(t) = \sum_{j \in \Omega/\{i\}} y_{ij}(t)$$

Rozwój zmiennej stanu jest wtedy zdefiniowany przez równanie:

$$x_j(t+1) = x_j(t) + y_{j*}(t) - y_{j*}$$

Przykład

Zostanie on przedstawiony w sekcji xx rozdziału yy przy wykorzystaniu macierzy systemowej.
 TODOOOOO

3.5 Gottlich dyskretyzacja - to jest bardzo źle

Początek drogi (k=0)

$$p_0^{t+1} = \frac{1}{4}(3p_0^t + p_1^t) - \frac{\Delta t}{2\Delta x}(f(p_1^t) + f(p_0^t) - 2\bar{\gamma}^t) \quad (3.10)$$

Centralny odcinek drogi (k=1,...,n-1)

$$p_k^{t+1} = \frac{1}{4}(p_{k-1}^t + 2p_k^t + p_{k+1}^t) - \frac{\Delta t}{2\Delta x}(f(p_{k+1}^t) - f(p_{k-1}^t)) \quad (3.11)$$

Końcowy odcinek drogi (k=n)

$$p_n^{t+1} = \frac{1}{4}(3p_n^t + p_{n-1}^t) - \frac{\Delta t}{2\Delta x}(2\hat{\gamma}^t - f(p_n^t) - f(p_{n-1}^t)) \quad (3.12)$$

$$\hat{\gamma}_i^t := f(p_{0,i}^t)$$

$$\bar{\gamma}_i^t := f(p_{n,i}^t)$$

Dla dróg i nie będących drogami źródłowymi:

$$\bar{\gamma}_i^t := \sum_{i \in \delta_v^{tn}} d_{ij} \hat{\gamma}_i(t)$$

Dla dróg źródłowych $\bar{\gamma}_i^t$ jest odgórnie ustalone i ustala jak dużo pojazdów wjeżdża do układu z rozważanej drogi.

STAREEEEEEEEEEE

Niech będzie ustalona droga e oraz jej siatka przestrzenna b_l . Celem jest przedstawienie wartości gęstości dla odcinków siatki przestrzennej w chwilach $k = 0, 1, \dots, K$. Gęstość w odcinku b_l i czasie k jest zdefiniowana jako:

$$p_l^k = \int_{b_l} \frac{p(x, k)}{\Delta x} dx. \quad (3.13)$$

Na podstawie (3.2) można wywnioskować, że:

$$\int_{b_l} p(x, k+1) - p(x, k) dx + \int_k^{k+1} f(b_{l+1}, k) - f(b_l, k) dk = 0 \quad (3.14)$$

Upraszczając otrzymujemy:

$$\Delta x(p_l^{k+1} - p_l^k) + \int_k^{k+1} f(b_{l+1}, k) - f(b_l, k) dk = 0 = 0 \quad (3.15)$$

Wartości gęstości zmieniają się w tylko w chwilach k . Wtedy wartości $f(b_{l+1}, k)$ i $f(b_l, k)$ są stałe na całym przedziale całkowania $[k, k+1)$. Otrzymujemy równanie:

$$\Delta x(p_l^{k+1} - p_l^k) + (f(b_{l+1}, k) - f(b_l, k)) = 0 \quad (3.16)$$

Rezultatem jest końcowy rekurencyjny wzór na gęstość ruchu:

$$p_l^{k+1} = p_l^k - \frac{1}{\Delta x}(f(b_{l+1}, k) - f(b_l, k)) \quad (3.17)$$

Rozdział 4

Model sieci dróg

4.1 Wstęp

Ze względu na dużą złożoność końcowego modelu zostanie przedstawiony najpierw bardzo prosty, podstawowy model. W każdej kolejnej sekcji dodawane będą zmiany przybliżające do ostatecznej postaci. Jest to podejście pozwalające na proste przedstawienie modelu, który zawiera bardzo wiele aspektów m.in: sygnalizacji świetlnej, wielu dróg, zatoru drogowego, źródła i ujścia ruchu. Zestawienie w jednej sekcji wszystkich tych kwestii byłoby bardzo przytłaczające.

4.2 Wektor stanu drogi

Wektor stanu jest strukturą przedstawiającą stan środowiska. Dla każdego odcinka drogi składa się on wartości zmiennych stanowych. Początkowo zmienną stanu jest identyfikowana jako ilość pojazdów na danym odcinku drogi.

4.2.1 Przykład

Niech będzie dana droga z wydzielonymi czterema odcinkami. Przykładowy wektor stanu takiego środowiska to

$$\mathbf{x}(t) = \begin{bmatrix} 2 \\ 4 \\ 3 \\ 0 \end{bmatrix} \quad (4.1)$$

Zawiera on w sobie następujące informacje dla chwili t :

- Są 2 pojazdy na zerowym odcinku
- Są 4 pojazdy na pierwszym odcinku
- Są 3 pojazdy na drugim odcinku
- Nie ma żadnego pojazdu na trzecim odcinku



Rysunek 4.1: Droga z ilością pojazdów na poszczególnych odcinkach

4.3 Rozwój wektora stanu drogi i wprowadzenie macierzy systemu

Formalnym wzorem definiującym rozwój wektora stanu jest:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) \quad (4.2)$$

Gdzie \mathbf{A} jest macierzą systemu. Wartość macierzy \mathbf{A} na przecięciu kolumny i i wiersza j określa jaka część pojazdów z odcinka i przejeżdża do odcinka j . W dalszych rozważaniach manewr przejazdu z odcinka i do j będzie oznaczany jako $[j, i]$.

4.3.1 Przykład bez zatoru

Dla przykładu przedstawionego w (4.2.1) zostanie przedstawiony rozwój wektora stanu. Niech zatem

$$\mathbf{x}(0) = \begin{bmatrix} 2 \\ 4 \\ 3 \\ 0 \end{bmatrix} \quad (4.3)$$

Macierzą systemu zapewniającą przejazd bez zatorów jest:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.4)$$

Co można zdefiniować także jako:

$$A[j, i] = \begin{cases} 1 & \text{dla manewrów } [j, i] = \{[1, 0], [2, 1], [3, 2]\} \\ 0 & \text{dla pozostałych manewrów} \end{cases} \quad (4.5)$$

$$(4.6)$$

Wedle wzoru (4.2) wyliczone zostają kolejne wartości wektora stanu.

$$\mathbf{x}(1) = \mathbf{A}\mathbf{x}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 4 \\ 3 \end{bmatrix} \quad (4.7)$$

$$\mathbf{x}(2) = \mathbf{Ax}(1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 4 \end{bmatrix} \quad (4.8)$$

$$\mathbf{x}(3) = \mathbf{Ax}(2) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \quad (4.9)$$

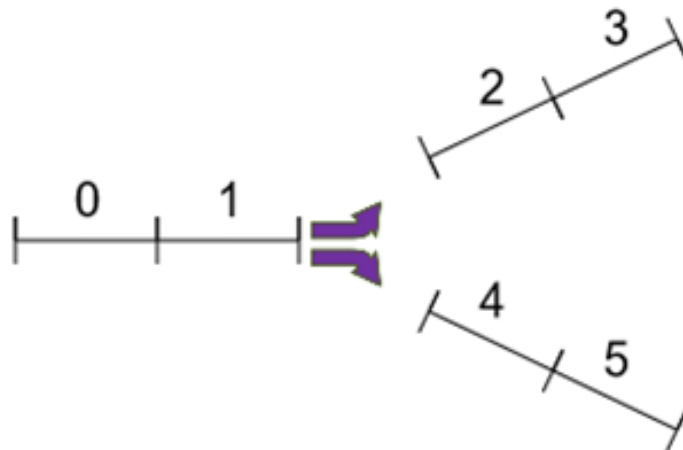
$$\mathbf{x}(4) = \mathbf{Ax}(3) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (4.10)$$

4.4 Wektor stanu sieci dróg

W rozdziale (4.2) przedstawiony został wektor stanu dla pojedynczej drogi. W tym rozdziale zostanie sformułowany wektor stanu dla bardziej ogólnego przypadku - sieci dróg. Sposób przedstawienia wartości stanów jednak jest bardzo podobny. Dla środowiska zawierającego drogi e_1, \dots, e_n wektor stanu jest utworzony z wartości stanu kolejnych odcinków tych dróg.

4.4.1 Przykład

Niech będzie dana sieć składająca się z trzech dróg. Dla każdej drogi zostaną wydzielone 2 odcinki. W sumie środowisko posiada 6 odcinków. Odcinki są ponumerowane od 0 co przedstawia poniższy rysunek.

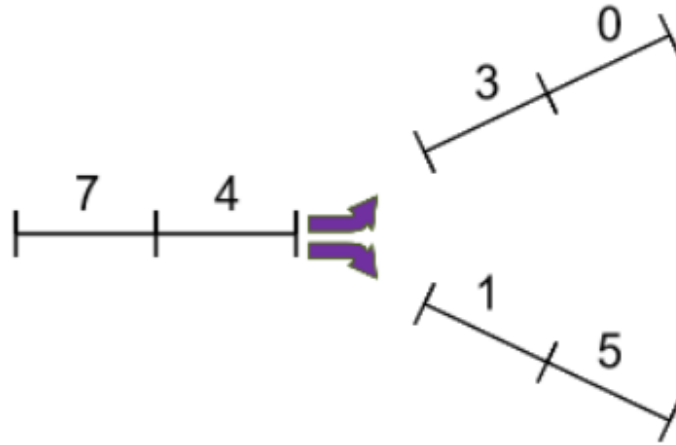


Rysunek 4.2: Numeracja odcinków środowiska

Niech przykładowym wektorem stanu będzie:

$$\mathbf{x}(t) = \begin{bmatrix} 7 \\ 4 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix} \quad (4.11)$$

Zawiera on w sobie informacje dotyczące ilości pojazdów na poszczególnych odcinkach w chwili t . Poniższy obraz przedstawia ilości pojazdów na poszczególnych odcinkach dla stanu zadanego przez wektor $\mathbf{x}(t)$



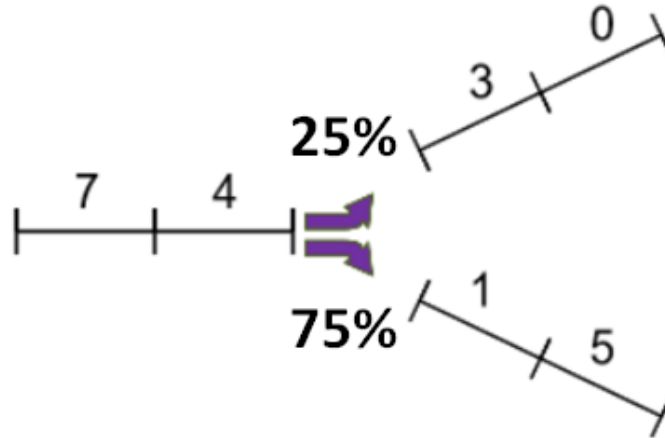
Rysunek 4.3: Sieć dróg z ilościami pojazdów na poszczególnych odcinkach

4.5 Rozwój wektora stanu sieci dróg w oparciu o CTM

Przedstawiony zostanie rozwój stanu środowiska (4.4.1). Jako model ruchu wykorzystany zostanie model CTM opisany w sekcji (3.4). Macierz systemu \mathbf{A} powinna uwzględnić przepływy pojazdów na skrzyżowaniach. Do tej pory wszystkie pojazdy poruszały się jedynie prosto po odcinkach jednej drogi i posiadały jedno ujęcie. Niech Ω_{out} będzie zbiorem wszystkich odcinków bezpośrednio przed jakimkolwiek ujęciem ruchu. W przypadku sieci dróg trzeba uwzględnić także przypadek skrzyżowania, gdzie pojazdy mogą obrać różne kierunki ruchu. W tym celu zdefiniowana zostanie macierz prawdopodobieństwa \mathbf{P} . Jest to niezmienna w czasie rzadka macierz. Wartości macierzy $P[j, i]$ są prawdopodobieństwem przejazdu z odcinka i do j . Dla niemożliwych manewrów $[j, i]$ wartości macierzy P są równe 0. Kolumny zadane przez $i \in \Omega_{out}$ mają zawsze wartości równe 0.

4.5.1 Przykład macierzy prawdopodobieństwa

Rozważmy poprzednią sieć dróg z dodatkowym założeniem, że 25 procent pojazdów ma zamiar dokonać manewru skrętu w lewo. Pozostałe 75 procent pojazdów z tego odcinka wybiera jazdę w prawo.



Rysunek 4.4: Sieć dróg z ilościami pojazdów na poszczególnych odcinkach

Wtedy macierz prawdopodobieństwa jest następująca:

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.12)$$

Dodatknie wartości macierzy P są następujące:

- $P[1, 0] = 1$ - manewr jazdy prosto z odcinka 0 do 1
- $P[3, 2] = 1$ - manewr jazdy prosto z odcinka 2 do 3
- $P[5, 4] = 1$ - manewr jazdy prosto z odcinka 4 do 5
- $P[2, 1] = 0.25$ - manewr jazdy w lewo z odcinka 1 do 2
- $P[4, 1] = 0.75$ - manewr jazdy w prawo z odcinka 1 do 4

4.6 Rozwój wektora stanu sieci dróg z przepływem CTM

Przedstawiona w tej sekcji macierz systemu \mathbf{A} powinna uwzględnić przepływy pojazdów na skrzyżowaniach oraz model ruchu CTM. Poniższa definicja macierzy A gwarantuje takie

zachowanie:

$$A[j, i](t) = \begin{cases} \frac{y_{ij}(t)}{x_i(t)} & \text{dla } i \neq j, i \notin \Omega_{out} \\ \frac{y_{ij}(t) - y_{i*}(t)}{x_i(t)} & \text{dla } i = j, i \notin \Omega_{out} \\ 0 & \text{dla } i \in \Omega_{out} \end{cases} \quad (4.13)$$

$$(4.14)$$

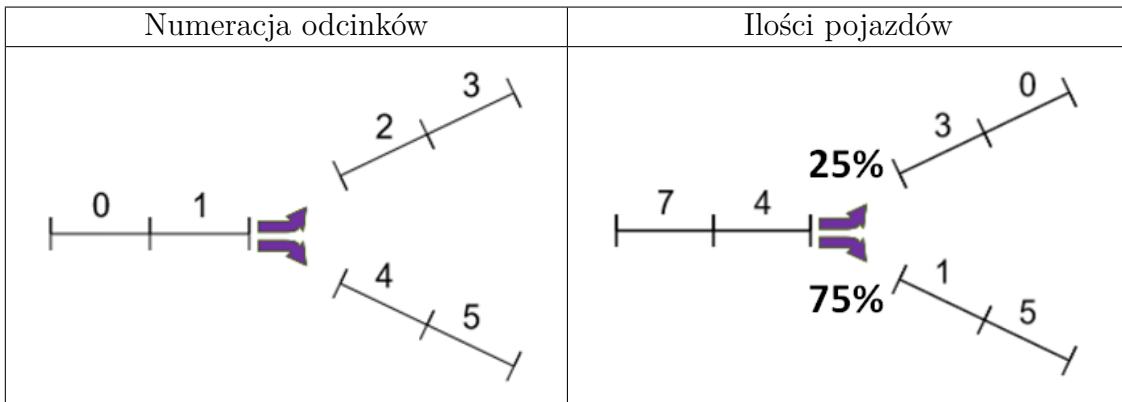
$$(4.15)$$

Powyższy wzór potrzebuje dokładnego omówienia. Jest on zgodny z początkową definicją macierzy systemu. Dla przypomnienia wartości macierzy systemu są określone jako część pojazdów będących na odcinku i dokonujących manewru $[j, i]$. Zgodność wszystkich 3 przypadków jest omówiona poniżej:

- (4.13) - Wartość $\frac{y_{ij}(t)}{x_i(t)}$ to stosunek pojazdów, które przejechały z odcinka i do odcinka j (t.j. $y_{ij}(t)$) do całkowitej liczby pojazdów na odcinku i (t.j. $x_i(t)$).
- (4.14) - Licznik $x_{ij}(t) - y_{i*}(t)$ określa ilość pojazdów, które pozostały na odcinku i .
- (4.15) - Wszystkie pojazdy z odcinka i bezpośrednio przed ujęciem przejeżdżają do ujęcia. Zatem nie wjeżdżają do żadnego innego odcinka j ,

4.6.1 Przykład

Rozważone zostanie środowisko (4.4.1) z założeniem, że 75 procent pojazdów ma zamiar skręcić w prawo. Pozostałe 25 procent wybiera skręt w lewo. Środowisko przedstawiają poniższe obrazki.



Powyższy sposób przepływu przez skrzyżowanie definiuje macierz prawdopodobieństwa \mathbf{P} określoną w przykładzie (4.12).

Ustalane zostaje, iż maksymalna ilość pojazdów będąca na odcinku to $N_j(t) = 7$ dla dowolnego i, t . Przepływ pojazdów w jednym interwale jest ograniczony przez $Q_j(t) = 4$. Wtedy następująca tabela przedstawia rozwój środowiska

t	Równanie stanu	Podgląd środowiska
0	$\mathbf{x}(0) = \begin{bmatrix} 7 \\ 4 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix}$	
1	$\mathbf{x}(1) = \mathbf{A}(0)\mathbf{x}(0) = \begin{bmatrix} \frac{4}{7} & 0 & 0 & 0 & 0 & 0 \\ \frac{3}{7} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}$	
2	$\mathbf{x}(2) = \mathbf{A}(1)\mathbf{x}(1) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 1 \\ 3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ \frac{3}{4} \\ 1 \\ 2\frac{3}{4} \\ 3 \end{bmatrix}$	
3	$\mathbf{x}(3) = \mathbf{A}(2)\mathbf{x}(2) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ \frac{3}{4} \\ 1 \\ 2\frac{3}{4} \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 6 \\ 0 \\ 0 \\ 2 \end{bmatrix}$	

4.7 Wprowadzenie sygnalizacji świetlnej

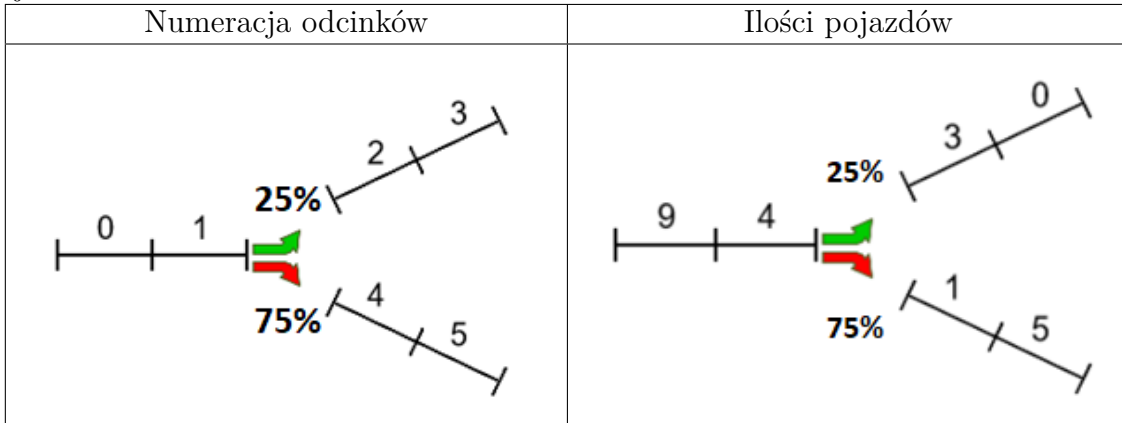
Kolejnym etapem rozwoju modelu jest wprowadzenie sygnalizacji świetlnej. Warto zauważyć, że do tej pory rozważane układy były pozbawione jakiegokolwiek sterowania, czego bezpośrednim skutkiem była niezmiennosc macierzy \mathbf{A} w czasie. Niech wektor $[i, j]$ będzie manewrem polegającym na bezpośrednim przejeździe z odcinka i na odcinek j . Zostanie teraz zdefiniowana macierz sygnalizacji świetlnej \mathbf{S} . Określa ona wykonalność dowolnego manewru.

$$S_{ij} = \begin{cases} 1 & \text{dla manewru zezwolonego przez zielone światło} & (4.16) \\ 1 & \text{dla manewru nie wymagającego sygnalizacji świetlnej} & (4.17) \\ 0 & \text{dla manewru wstrzymanego przez czerwone światło} & (4.18) \\ 0 & \text{dla niemożliwego manewru} & (4.19) \end{cases}$$

Sygnalizacja świetlna zwykle posiada pewną fazę, która określa dozwolone (4.16) oraz nie-dozwolone (4.18) manewry. Fazy świetlne będą oczywiście zmieniane w trakcie symulacji zatem macierz \mathbf{S} też jest zmienna w czasie. Wartości macierzy \mathbf{S} obejmujące przypadki niemożliwych manewrów (4.19) oraz nie wymagających sygnalizacji świetlnej (4.17) są z kolei niezmiennie dla ustalonego środowiska.

4.7.1 Przykład macierzy sygnalizacji świetlnej

Rozważone będzie znane z poprzednich przykładów skrzyżowanie - tym razem mające sygnalizację świetlną z aktywną fazą zezwalającą na manewr $[1,2]$. Ta sama faza zabrania przejazdu z 1 odcinka do 4.



Poszczególne manewry są następujące:

- (4.16) - Manewrem zezwolonym przez zielone światło dla fazy 0 jest $[1, 2]$.
- (4.17) - Prawidłowym manewrem bez sygnalizacji świetlnej są manewry $[0, 1]$, $[2, 3]$, $[4, 5]$.
- (4.18) - Dla fazy świetlnej 0 manewrem zatrzymanym przez czerwone światło jest $[1, 4]$.
- (4.19) - Pozostałe manewry są niemożliwe. Przykładem jest $[0, 2]$, gdyż nie ma możliwości bezpośredniego przejazdu z odcinka 0 do 2.

Macierz sygnalizacji świetlnej dla tego przykładu jest następująca:

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

4.7.2 Macierz systemu uwzględniająca sygnalizację świetlną

Mając zdefiniowaną zarówno macierz prawdopodobieństwa przejazdów \mathbf{P} jak i macierz sygnalizacji świetlnej \mathbf{S} można ustalić macierz systemu \mathbf{A} uwzględniającą już sygnalizację świetlną. Niech Q będzie zbiorem odcinków, które znajdują się bezpośrednio przed ujściem ruchu. Wtedy macierz systemu zdefiniowana jest jako:

$$A[i, j] = \begin{cases} 0 & \text{dla } S[i, j] = 0 \vee i \in Q \\ P[i, j] & \text{dla } S[i, j] = 1 \\ 1 - \delta(i) & \text{dla } i = j \wedge i \notin Q \end{cases} \quad \begin{matrix} (4.20) \\ (4.21) \\ (4.22) \end{matrix}$$

Równanie systemu jest następujące:

$$\mathbf{x}(t+1) = \mathbf{A}(t)\mathbf{x}(t) \quad (4.23)$$

Macierz systemu jest teraz zmienna w czasie, zatem w miejsce \mathbf{A} dotychczasowego równania pojawiło się $\mathbf{A}(t)$.

4.7.3 Przykład rozwoju wektora stanu

Niech będzie dany układ z przykładu (4.7.1) z zielonym światłem dla lewoskrętu. W chwili $t=2$ zostanie zmieniona faza świetlna. Od tego momentu obydwa manewry na skrzyżowaniu są dozwolone. W chwili $t=3$ wartości stanu nie są całkowite, co nie jest sprzeczne z modelem matematycznym. Rozwój wektora stanu został przedstawiony w poniższej tabeli.

t	Równanie stanu	Podgląd środowiska
0	$\mathbf{x}(0) = \begin{bmatrix} 9 \\ 4 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix}$	
1	$\mathbf{x}(1) = \mathbf{A}(0)\mathbf{x}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 9 \\ 4 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 3 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	
2	$\mathbf{x}(2) = \mathbf{A}(1)\mathbf{x}(1) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 12 \\ 3 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 9 \\ 3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	
3	$\mathbf{x}(3) = \mathbf{A}(2)\mathbf{x}(2) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 9 \\ 3 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2\frac{1}{4} \\ 3 \\ 7\frac{3}{4} \\ 0 \end{bmatrix}$	

4.8 Wprowadzenie źródeł ruchu

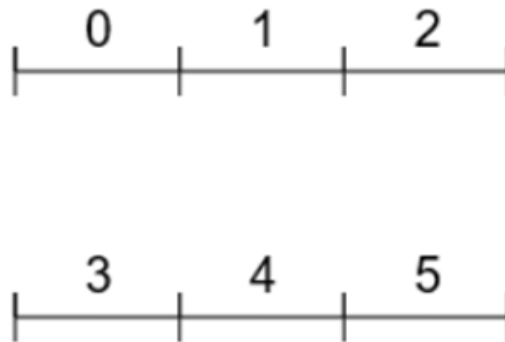
Wszystkie poprzednie przykłady układów ruchu drogowego szybko kończyły się stanem w którym nie było już żadnych pojazdów na drogach. W tej sekcji zostanie przedstawiony sposób napływania nowych pojazdów do układu. Wprowadzony zostanie wektor źródła $\mathbf{u}(t)$

definiujący pojazdy pojawiające się w układzie. Jest on zmienny w czasie, a jego wartości określają ile pojazdów pojawia się w następnej chwili w poszczególnych odcinkach układu. Równanie systemu uwzględniające źródła ruchu to:

$$\mathbf{x}(t) = \mathbf{A}(t-1)\mathbf{x}(t-1) + \mathbf{u}(t-1) \quad (4.24)$$

4.8.1 Przykład

Rozważony zostanie prosty przykład, który nie uwzględnia wcześniej wprowadzonych pojęć sygnalizacji świetlnej oraz zatoru. Przedstawione zostanie środowisko składające się z dwóch dróg, które nie są ze sobą w żaden sposób połączone.



Rysunek 4.5: Numeracja odcinków środowiska

Pojazdy wzdłuż dróg kończąc swój bieg po przejechaniu przez odcinki 5 i 2 co przedstawia macierz systemu A :

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Pragnąc by do odcinków 0 i 3 napływało odpowiednio po 4, 8, 20 oraz 3, 7, 5 pojazdów należy zdefiniować ciąg wektorów $\mathbf{u}(t)$ w sposób przedstawiony w poniższej tabeli.

t	Równanie stanu	Podgląd środowiska	$\mathbf{u}(t)$
0	$\mathbf{x}(0) = \begin{bmatrix} 0 \\ 8 \\ 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}$		$\begin{bmatrix} 4 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}$
1	$\mathbf{x}(1) = \mathbf{Ax}(0) + \mathbf{u}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ 1 \\ 3 \\ 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 8 \\ 3 \\ 3 \\ 3 \end{bmatrix}$		$\begin{bmatrix} 8 \\ 0 \\ 0 \\ 7 \\ 0 \\ 0 \end{bmatrix}$
2	$\mathbf{x}(2) = \mathbf{Ax}(1) + \mathbf{u}(1) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 8 \\ 3 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 8 \\ 0 \\ 0 \\ 7 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 0 \\ 7 \\ 3 \\ 3 \end{bmatrix}$		$\begin{bmatrix} 20 \\ 0 \\ 0 \\ 5 \\ 0 \\ 0 \end{bmatrix}$
3	$\mathbf{x}(3) = \mathbf{Ax}(2) + \mathbf{u}(2) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \\ 0 \\ 7 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 20 \\ 0 \\ 0 \\ 5 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 20 \\ 8 \\ 4 \\ 5 \\ 7 \\ 3 \end{bmatrix}$		

4.9 Wprowadzenie modelu ruchu CTM

Dotychczasowe przedstawienie macierzy systemu \mathbf{A} nie zawiera w sobie jeszcze pojęcia zatoru drogowego. W jednym interwale czasowym może przejechać przez odcinek astronomiczna wręcz liczba pojazdów. Wprowadzony do modelu systemu zostanie zatem sposób przepływu pojazdów przedstawiony w sekcji (3.4). Niech dane zatem środowisko z macierzą prawdopodobieństwa P .

Dziedziną funkcji zatoru f jest zbiór wszystkich manewrów $[i, j]$, takich że $i \neq j$. Funkcja zatoru jest jednak zależna także od wartości stanowych wektora \mathbf{x} , a przede wszystkim od ilości pojazdów na odcinkach i oraz j czyli $x[i], x[j]$. Wartości funkcji f należą do przedziału $[0, 1]$. Macierz stanowa \mathbf{A} z uwzględnieniem zatorów dla układu z n odcinkami przedstawia się następująco:

$$\mathbf{A} = \begin{bmatrix} 1 - \delta(0) & f(1, 0) & \dots & f(n, 0) \\ f(0, 1) & 1 - \delta(1) & \dots & f(n, 1) \\ f(0, 2) & f(1, 2) & \dots & f(n, 2) \\ \dots & \dots & \dots & \dots \\ f(0, n) & f(1, n) & \dots & 1 - \delta(n) \end{bmatrix} \quad (4.25)$$

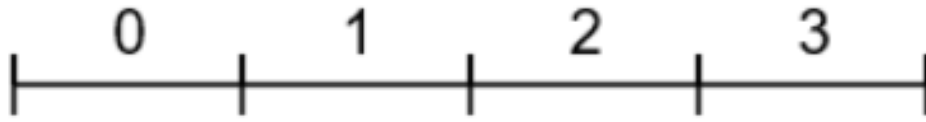
Gdzie δ to suma wszystkich wartości danej kolumny z wyłączeniem tych na głównej przekątnej. W przypadku odcinków zakończonych ujściem, aby wyzerować wartości δ przyjmuje wartość 1. Formalnie δ to:

$$\delta(i) = \begin{cases} \sum_{j \in \{0, \dots, n\}, j \neq i} f(i, j) & \text{dla odcinka } i \text{ nie zakończonego ujściem} \\ 1 & \text{dla odcinka } i \text{ zakończonego ujściem} \end{cases} \quad (4.26)$$

$$(4.27)$$

4.9.1 Przykład zatoru na pojedynczej drodze

Rozważmy poniższe środowisko składające się z drogi posiadającej 4 odcinki.



Rysunek 4.6: Droga z numeracją odcinków

Wykorzystany zostanie model przepływu CTM przedstawiony w sekcji 3.4. Niech maksymalny przepływ to $Q_i(t) = 10$ dla dowolnej chwili t . Należy skonstruować pewną funkcję zatoru. Niech jej celem będzie, aby przez jeden interwał czasowy maksymalnie 10 pojazdów przejeżdżało przez odcinek drogi. Następująca funkcja gwarantuje takie zachowanie:

$$f(i, j) = \begin{cases} \mathbf{P}[i, j] & \text{dla } j = i + 1 \wedge x[i] < 10 - \text{przejazd bez zatoru} \\ \frac{10}{x[i]} & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zatorem} \\ 0 & \text{dla pozostałych przypadków - niemożliwe manewry} \end{cases} \quad (4.28)$$

$$(4.29)$$

$$(4.30)$$

Macierze stanu zostały wyliczone na podstawie wzoru 4.25. W tym przypadku jest to:

$$A = \begin{bmatrix} 1 - \delta(0) & f(1, 0) & f(2, 0) & f(3, 0) \\ f(0, 1) & 1 - \delta(1) & f(2, 1) & f(3, 1) \\ f(0, 2) & f(1, 2) & 1 - \delta(2) & f(3, 2) \\ f(0, 3) & f(1, 3) & f(2, 3) & 1 - \delta(3) \end{bmatrix} \quad (4.31)$$

Funkcja δ dla tego środowiska to:

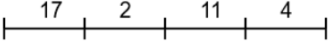
$$\delta(i) = \begin{cases} f(0, 1) + f(0, 2) + f(0, 3) & \text{dla } i = 0 \\ f(1, 0) + f(1, 2) + f(1, 3) & \text{dla } i = 1 \\ f(2, 0) + f(2, 1) + f(2, 3) & \text{dla } i = 2 \\ 1 & \text{dla } i = 3 \end{cases} \quad (4.32)$$

$$(4.33)$$

$$(4.34)$$

$$(4.35)$$

Rozważmy następujący stan początkowy środowiska:

t	Równanie stanu	Podgląd środowiska
0	$\mathbf{x}(0) = \begin{bmatrix} 17 \\ 2 \\ 11 \\ 4 \end{bmatrix}$	

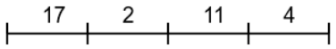
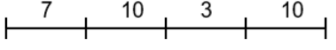
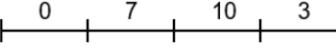
W celu wyliczenia $\mathbf{x}(1)$ należy obliczyć $\mathbf{A}(0)$, wedle wzoru (4.31). Wartości zerowej kolumny macierzy $\mathbf{A}(0)$ są wyliczane w następującej kolejności:

- $A[0, 1] = f(0, 1)$ to przypadek manewru z zatorem (4.29). Zatem $f(0, 1) = \frac{10}{17}$, gdyż $x[0] = 17$
- $A[0, 2] = f(0, 2) = 0$ $f(0, 3) = 0$ to przypadki niemożliwych manewrów (4.30).
- $A[0, 0] = 1 - \delta(0) = 1 - f(0, 1) - f(0, 2) - f(0, 3) = 1 - \frac{10}{17} - 0 - 0 = \frac{7}{17}$ pozostaje na odcinku 0.

Analogiczne operacje zostają przeprowadzone dla kolejnych kolumn co daje rezultat w postaci:

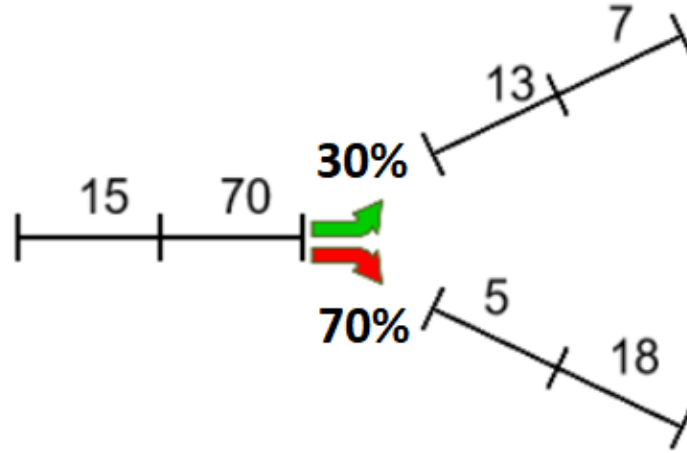
$$\mathbf{A}(0) = \begin{bmatrix} \frac{7}{17} & 0 & 0 & 0 \\ \frac{10}{17} & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{11} & 0 \\ 0 & 0 & \frac{10}{11} & 0 \end{bmatrix}$$

Dalszy rozwój wektora stanu jest przedstawiony w tabeli poniżej.

t	Równanie stanu	Podgląd środowiska
0	$\mathbf{x}(0) = \begin{bmatrix} 17 \\ 2 \\ 11 \\ 4 \end{bmatrix}$	
1	$\mathbf{x}(1) = \mathbf{A}(0)\mathbf{x}(0) = \begin{bmatrix} \frac{7}{17} & 0 & 0 & 0 \\ \frac{10}{17} & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{11} & 0 \\ 0 & 0 & \frac{10}{11} & 0 \end{bmatrix} \begin{bmatrix} 17 \\ 2 \\ 11 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \\ 3 \\ 10 \end{bmatrix}$	
2	$\mathbf{x}(2) = \mathbf{A}(1)\mathbf{x}(1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 7 \\ 10 \\ 3 \\ 10 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 10 \\ 3 \end{bmatrix}$	

4.9.2 Przykład zatoru na skrzyżowaniu

Niech funkcją zatoru będzie funkcja przedstawiona w (4.28-4.30) Rozważmy poniższą sytuację na drogach w chwili $t = 0$:



- Zatory są na odcinkach 0,1,2 i 5 (odpowiednio ilości pojazdów to 15,70,13,18).
- Obecna faza świetlna to 0.

Najistotniejszym punktem macierzy sygnalizacji świetlnej jest $S[1,2] = 1$, co odpowiada zielonemu światłu dla lewoskrętu. Cała macierz \mathbf{S} to:

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (4.36)$$

Macierz systemu to zatem:

$$\mathbf{A} = \begin{bmatrix} \frac{5}{15} & 0 & 0 & 0 & 0 & 0 \\ \frac{10}{15} & \frac{60}{70} & 0 & 0 & 0 & 0 \\ 0 & \frac{10}{70} & \frac{3}{13} & 0 & 0 & 0 \\ 0 & 0 & \frac{13}{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.37)$$

Dla przykładu szczegółowo zostaną policzone wartości zerowej kolumny macierzy \mathbf{A} , która dotyczy pojazdów wyjeżdżających z odcinka 0. Pozostałe wartości są liczone analogicznie.

- Korzystając ze wzoru 5.11 ustalone zostaje $f(0,1) = \frac{10}{15}$. Jest to przypadek zatoru i jedynie 10 pojazdów z 15 przemieszcza się do następnego odcinka.
- Wartości $f(0,2), (0,3), (0,4), (0,5)$ są równe 0, gdyż są to przypadki 4.19 nieistniejących manewrów.
- $A[0,0] = 1 - \delta(0) = 1 - f(0,1) - f(0,2) - f(0,3) - f(0,4) - f(0,5) = 1 - \frac{10}{15} - 0 - 0 - 0 - 0 = \frac{5}{15}$

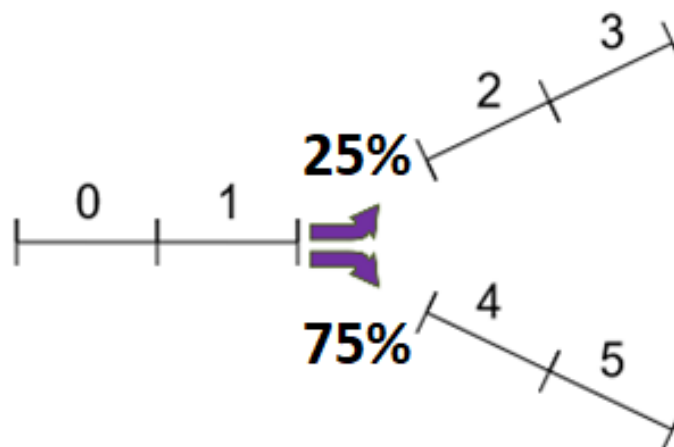
Rozdział 5

Środowiska symulacyjne i ich nauka

W tym rozdziale przedstawione zostaną środowiska symulacyjne. Głównym ich elementem jest sieć dróg. Podstawowe informacje na temat działania modelu sieci dróg i ruchu obowiązującego na nich zostały przedstawione w rozdziale 4. Każde środowisko posiada dokładną specyfikację możliwych faz świetlnych. Do każdego ze skrzyżowań jest przyporządkowany agent, który ma za zadanie podejmować akcje prowadzące do dobrej przepustowości ruchu. Dokładnie zostaną przedstawione przestrzenie decyzyjne i bezpośrednie konsekwencje każdej z akcji. Każde środowisko będzie posiadało swój problem optymalizacyjny i przedstawiony zostanie sposób uczenia, który prowadzi do rozwiązania tego problemu.

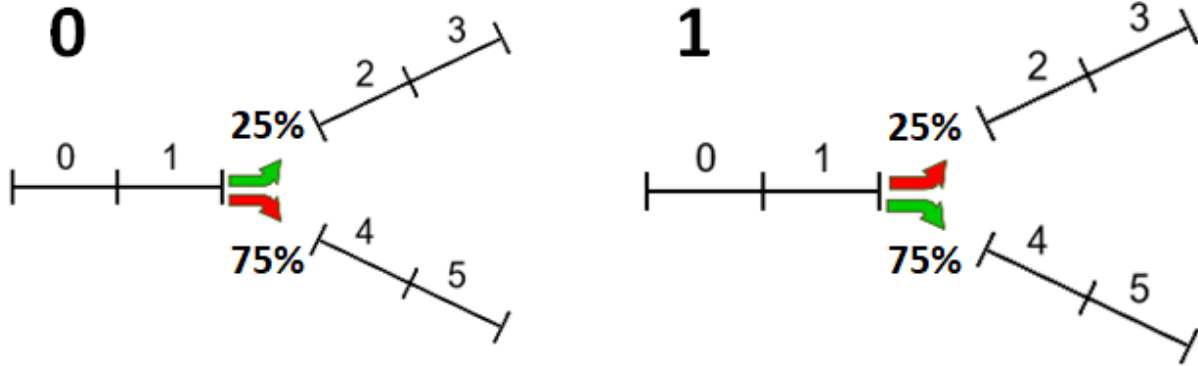
5.1 Środowisko 1

Elementy sieci dróg pierwszego środowiska były niejednokrotnie przedstawiane w rozdziale (4), jednak dopiero teraz nastąpi przedstawienie pierwszego pełnego środowiska symulacyjnego. Na model sieci składają się 3 drogi - każda ma po 2 odcinki. Jest jedno skrzyżowanie na którym 75 procent pojazdów skręca w prawo, a pozostałe obierają kierunek w lewo.



Rysunek 5.1: środowisko 11

5.1.1 Sygnalizacje świetlne



Rysunek 5.2: środowisko 11 - fazy świateł

Skrzyżowanie posiada 2 fazy świetlne przedstawione powyżej. Faza 0 umożliwia skręt w lewo (manewr [1,2]) z kolei faza 1 umożliwia skręt w prawo (manewr [1,4]).

5.1.2 Przestrzeń decyzyjna

Agent ma dwie możliwe akcje do podjęcia - 0 oraz 1. Obydwie powodują aktywację odpowiadającej im fazy świetlnej.

5.1.3 Zatory drogowe

Wykorzystana w modelu zostanie funkcja zatoru przedstawiona w rozdziale (4.9.1), czyli:

$$f(i, j) = \begin{cases} P[i, j] & \text{dla } j = i + 1 \wedge x[i] < 10 - \text{przejazd bez zatoru} & (5.1) \\ \frac{10}{x[i]} & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zatorem} & (5.2) \\ 0 & \text{dla pozostałych przypadków - manewr [i,j]} & (5.3) \end{cases}$$

5.1.4 Uczenie

Ogólny zarys

W trakcie symulacji treningowych do drogi wjeżdża 9 pojazdów w czasie jednego interwału. Po jednej iteracji algorytmu uczenia następuje symulacja w pełni zachłanna.

Cel nauki

Celem nauki jest uzyskanie jak największej liczby pojazdów, które opuszczą rozwidlenie dróg.

Pożądane zachowanie

Pożądanym zachowaniem jest utrzymywanie fazy 1. Jest ona w każdym przypadku lepsza lub równa fazie 0. Spowodowane jest to tym, że więcej pojazdów skręca w prawo. W przypadku gdy przed rozwidleniem drogi jest co najmniej 40 pojazdów poprawną akcją jest także wybór fazy 0. Nie powinno jednak dojść do aż takiego zatoru.

Sygnal wejściowy

Sygnal wejściowy to ilości pojazdów przed skrzyżowaniem.

Przykład sygnału wejściowego

Sygnal wejściowy $\mathbf{x} = [4, 7]$ zawiera w sobie informację, iż na odcinku zerowym są 4 pojazdy oraz na pierwszym - jest ich 7.

Sposób uczenia

Zastosowana została metoda uczenia przedstawiona w (6.4). Następujące stałe określające model uczenia to:

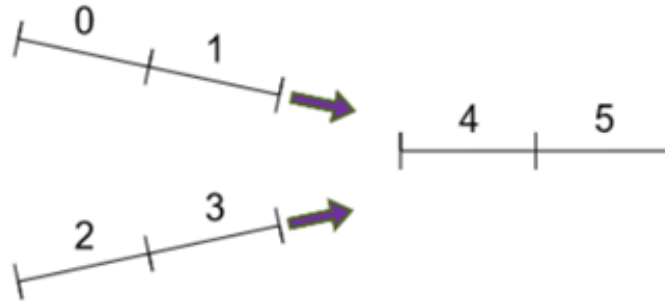
- $\gamma = 0.9$ - stopa dyskontowa
- $\alpha = 0.01$ - stała szybkości uczenia (ang. learning rate)
- Generowanych jest jednorazowo 50 epizodów. Każdy ma 200 interwałów czasowych.
- Wygenerowane zbiory danych są dzielone w proporcjach 80 i 20 procent dla odpowiednio treningu i walidacji.
- Warunkiem stopu jest 60 sekund działania algorytmu.
- Sieć neuronowa posiada 1 ukrytą warstwę z 10 neuronami.
- Za uczenie odpowiada sieć neuronowa biblioteki Keras. Posiada ukrytą warstwę z 10 neuronami i funkcją aktywacji relu. Optymalizacja jest przeprowadzana przez metodę Adam z funkcją straty błędu średniokwadratowego i parametrem *learning_rate* = 0,01.
- Epizod zachłanny trwa 200 interwałów czasowych.

Wyniki uczenia

Za każdym razem już po trzeciej iteracji algorytmu osiągnęte jest pożądane zachowanie agenta. Dla każdego stanu agent wybiera akcję 1. Nie tworzą się zatory i wszystkie pojazdy opuszczają układ (oczywiście z wyjątkiem tych, które pojawiły się w dwóch ostatnich interwałach czasowych).

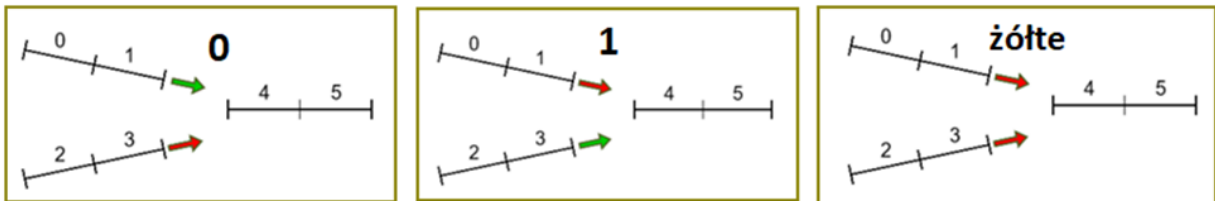
5.2 Środowisko 2

Na model sieci składają się 3 drogi - każda ma po 2 odcinki. Istnieje jedno skrzyżowanie z dwiema drogami wlotowymi i jedną wylotową.



Rysunek 5.3: środowisko 2

5.2.1 Sygnalizacje świetlne

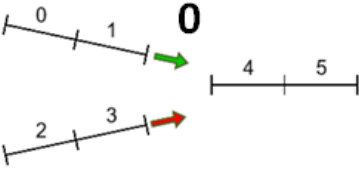
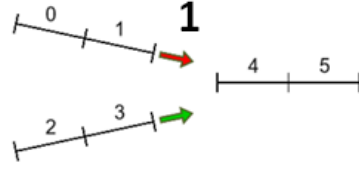
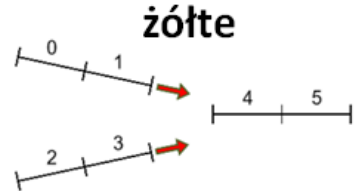


Rysunek 5.4: środowisko 2 - fazy świateł

Skrzyżowanie posiada 3 fazy świetlne przedstawione powyżej. Faza 0 umożliwia przejazd przez skrzyżowanie z górnej drogi (manewr [1,4]) z kolei faza 1 umożliwia wjazd z drogi dolnej (manewr [3,4]). Faza "żółte" blokuje przejazd przez skrzyżowanie.

5.2.2 Przestrzeń decyzyjna

Przestrzeń decyzyjną oraz konsekwencje akcji podjętych przez agenta przedstawia poniższa tabela:

Obecna faza	Możliwe akcje	Konsekwencje akcji
	0	Podtrzymanie fazy 0
	1	2 interwały fazy żółtych świateł, a następnie faza 1
	0	2 interwały fazy żółtych świateł, a następnie faza 0
	1	Podtrzymanie fazy 1
	żółte	Podtrzymanie fazy żółtych świateł. Jeśli żółta faza trwa dłużej niż 2 interwały czasowe – włączenie oczekiwanej fazy 0 lub 1.

5.2.3 Zatory drogowe

Wykorzystana w modelu zostanie funkcja zatoru przedstawiona w rozdziale (4.9.1), czyli:

$$f(i, j) = \begin{cases} P[i, j] & \text{dla } j = i + 1 \wedge x[i] < 10 - \text{przejazd bez zatoru} \\ \frac{10}{x[i]} & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zatorem} \\ 0 & \text{dla pozostałych przypadków - manewr [i,j]} \end{cases} \quad (5.4)$$

$$(5.5)$$

$$(5.6)$$

Zezwala ona na jednoczesny wjazd maksymalnie 10 pojazdów na dowolny odcinek drogi.

5.2.4 Uczenie

Ogólny zarys

Początkowo do każdej z dróg wjeżdżają po 2 pojazdy w jednym interwale czasowym. Na koniec każdej iteracji nauki przeprowadzane są symulacje w pełni zachłanne. Agent ma za zadanie wykonywać akcje, które w trakcie symulacji pozwolą na opuszczenie skrzyżowania przez co najmniej 98 procent pojazdów. W przypadku spełnienia tego warunku następuje zwiększenie ilości wjeżdżających do układu pojazdów. Ten proces jest powtarzany, aż do

momentu gdy przez 10 iteracji uczenia nie uzyskany zostanie warunek 98 procent pojazdów opuszczających skrzyżowanie. Moment stopu spodziewany jest na chwilę, gdy na każdą z dróg będzie wjeżdżało około 5 pojazdów. Nie ma możliwości, by przejeżdżało więcej niż 10 pojazdów, a tyle właśnie pojawia się w układzie. Muszą zatem wtedy pojawiać się zatory.

Cel nauki

Bezpośrednim celem nauki jest uzyskanie jak największej ilości pojazdów wjeżdżających do sieci. Aby to się stało agenci muszą podejmować decyzje, które pozwalają na opuszczenie sieci jak największej ilości pojazdów.

Spodziewany wynik

Moment stopu spodziewany jest na chwilę, gdy na każdą z dróg będzie wjeżdżało około 5 pojazdów (w sumie 10 pojazdów). Zgodnie z opisem modelu ruchu - maksymalnie 10 pojazdów może przejechać w jednej chwili przez skrzyżowanie, zatem muszą się tworzyć zatory przy większej ilości pojazdów.

Pożądane zachowanie agenta

Pożądane zachowanie agenta w tym przypadku nie jest jednoznaczne do określenia. Z pewnością jeśli na drodze na której jest zielone światło jest co najmniej 10 pojazdów, to warto podtrzymać fazę świetlną. Zmiana sygnalizacji powoduje 2 interwały karencji, podczas których żaden pojazd nie przejedzie i może to być czynnikiem decydującym o pozostaniu przy obecnej fazie. Jaka powinna być akcja w przypadku gdy zielone światło jest dla drogi z mniejszą niż 10 ilością pojazdów? Wtedy agent powinien odnaleźć złoty środek i jeśli na drugiej drodze jest dużo pojazdów - może zmienić światło.

Sygnał wejściowy

Wektor stanu posiada 4 wartości. Pierwsze dwie z nich to ilości pojazdów przed skrzyżowaniem. Kolejne określają aktualną fazę świetlną.

Przykłady sygnałów wejściowych

- Sygnał wejściowy $\mathbf{x} = [8, 2, 0, 1]$ zawiera w sobie informację, iż na odcinku zerowym jest 8 pojazdów oraz na pierwszym - jest ich 2. Aktywna jest faza 1.
- Sygnał wejściowy $\mathbf{x} = [7, 5, 1, 0]$ zawiera w sobie informację, iż na odcinku zerowym jest 7 pojazdów oraz na pierwszym - jest ich 5. Aktywna jest faza 0.

Sposób uczenia

Zastosowana została metoda uczenia przedstawiona w (6.4).

Następujące stałe oraz ustalenia określają szczegóły uczenia:

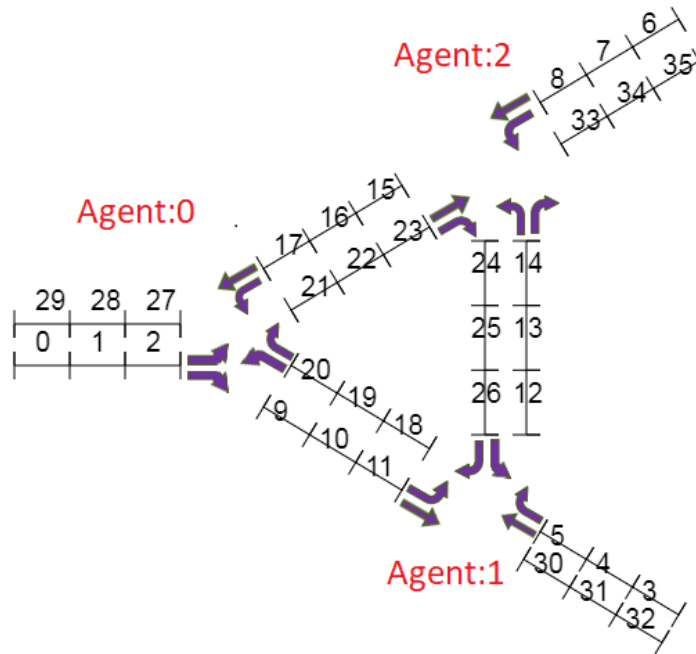
- $\gamma = 0.9$ - stopa dyskontowa

- Parametr ϵ jest początkowo równy 1 (w pełni losowe epizody). Z każdą iteracją jest zmniejszany o 0,01 aż do wartości $\epsilon = 0,2$.
- Generowanych jest jednorazowo 10 epizodów. Każdy z nich trwa 1000 interwałów czasowych.
- Wygenerowane zbiory danych są dzielone w proporcjach 80 i 20 procent dla odpowiednio treningu i walidacji.
- Za uczenie odpowiada sieć neuronowa biblioteki Keras. Posiada ona 3 ukryte warstwy (10,14,10 neuronów) z funkcjami aktywacji relu. Optymalizacja jest przeprowadzana przez metodę Adam z funkcją straty błędu średniokwadratowego i parametrem $learning_rate = 0,01$.

Wyniki uczenia

Algorytm szybko uzyskuje kolejne warunki 98 procent pojazdów przejeżdżających przez skrzyżowanie. Dopiero przy 4,97 wjeżdżających pojazdach na drogę widoczny jest zastój, który powoduje warunek stopu. Najlepszy wynik to średnio 9,48 pojazdów przejeżdżających w jednej chwili przez skrzyżowanie. Jest to dobry wynik zważając na to, iż potencjalnie najwyższy możliwy wynik do osiągnięcia jest w granicy 9,94. Z 9940 pojazdów 450 nie opuściło skrzyżowania. Zatem przy niemalże 5 pojazdach napływających do układu powstają korki, co jest naturalnym zjawiskiem - wynikającym z ograniczeń ruchu, a nie nieoptymalnej sygnalizacji świetlnej.

5.3 Środowisko 3

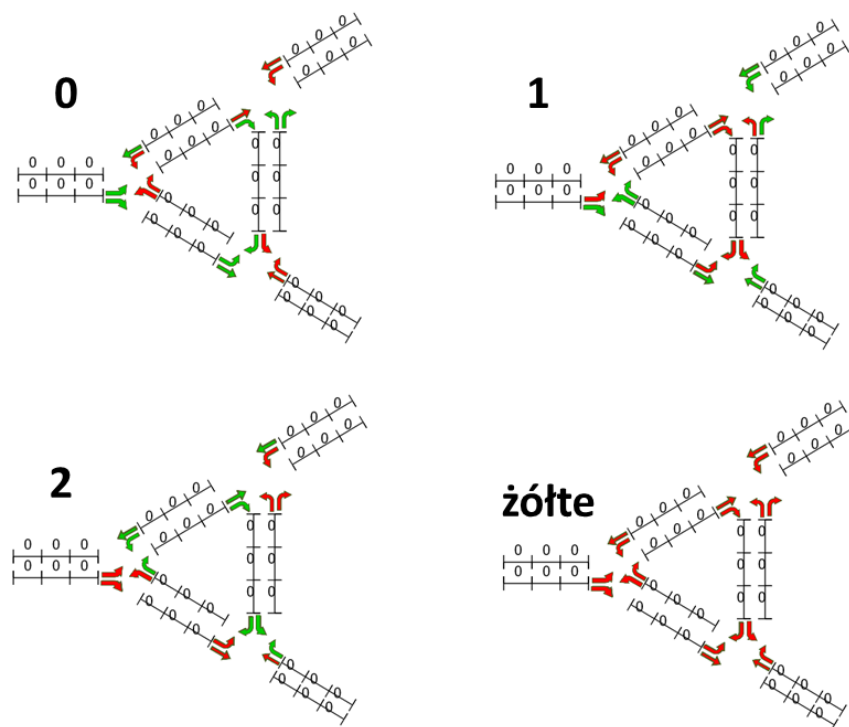


Rysunek 5.5: środowisko 4

Środowisko posiada 12 jednokierunkowych dróg. Każda droga ma 3 odcinki co daje w sumie 36 odcinków (są numerowane od 0 co widać na rysunku 5.3). W sieci dróg znajdują się 3 skrzyżowania. Do każdego z nich jest przypisany agent, który odpowiada za sterowanie sygnalizacją świetlną.

5.3.1 Sygnalizacje świetlne

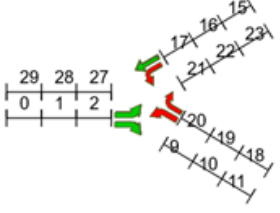
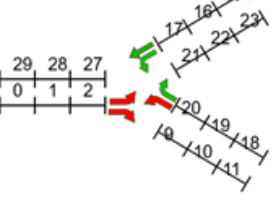
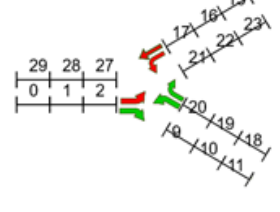
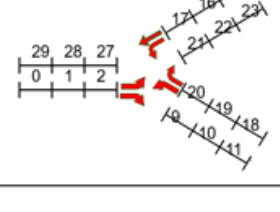
Każde skrzyżowanie posiada 4 fazy świetlne przedstawione poniżej. Fazy 0,1 i 2 są fazami, które posiadają pewne zielone światła. Zmiana pomiędzy tymi fazami nie jest natychmiastowa i następuje dopiero po 2 interwałach fazy żółtych światel.



Rysunek 5.6: środowisko 4 - fazy świateł

5.3.2 Przestrzeń decyzyjna

Przedstawiona zostanie przestrzeń decyzyjna jedynie dla agenta 0, jednak dla pozostałych dwóch przestrzenie decyzyjne są analogiczne. Schemat zmiany faz jest identyczny, jedynie manewry danych faz są inne.

Obecna faza	Możliwe akcje	Konsekwencje akcji
	0	Podtrzymanie fazy 0
	1	2 interwały fazy żółtych świateł, a następnie faza 1
	2	2 interwały fazy żółtych świateł, a następnie faza 2
	0	2 interwały fazy żółtych świateł, a następnie faza 0
	1	Podtrzymanie fazy 1
	2	2 interwały fazy żółtych świateł, a następnie faza 2
	0	2 interwały fazy żółtych świateł, a następnie faza 0
	1	2 interwały fazy żółtych świateł, a następnie faza 1
	2	Podtrzymanie fazy 2
	żółte	Podtrzymanie fazy żółtych świateł. Jeśli żółta faza trwa dłużej niż 2 interwały czasowe – włączenie oczekiwanej fazy 0 lub 1.

Rysunek 5.7: środowisko 4 - fazy świateł

5.3.3 Zatory drogowe

Wykorzystana w modelu zostanie funkcja zatoru przedstawiona w rozdziale (4.9.1), czyli:

$$f(i, j) = \begin{cases} P[i, j] & \text{dla } j = i + 1 \wedge x[i] < 10 - \text{przejazd bez zatoru} \\ 10 & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zatorem} \\ 0 & \text{dla pozostałych przypadków - manewr [i,j]} \end{cases} \quad (5.7)$$

$$f(i, j) = \begin{cases} \frac{10}{x[i]} & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zatorem} \end{cases} \quad (5.8)$$

$$f(i, j) = \begin{cases} 0 & \text{dla pozostałych przypadków - manewr [i,j]} \end{cases} \quad (5.9)$$

Zezwala ona na jednoczesny wjazd 10 pojazdów na dowolny odcinek drogi.

5.3.4 Uczenie

Sygnal wejściowy

Sieć neuronowa przypisana do każdego z agentów przyjmuje jako sygnał wejściowy 6 elementową tablicę. Pierwsze 3 elementy to ilości pojazdów na odcinkach będących przed skrzyżowaniem przypisanym do agenta. Pozostałe 3 elementy mają wartości z zakresu $\{0, 1\}$. Określają one aktualną fazę świetlną.

Przykład sygnału wejściowego

Sygnał wejściowy $x = [5, 17, 2, 0, 0, 1]$ przedstawia informację, iż przed skrzyżowaniem są kolejki odpowiednio 5, 17 i 2 pojazdów. Ostatnie 3 elementy oznaczają, że aktualna faza świetlna to 2.

Uczenie

Ogólny zarys

Początkowo do każdej z dróg wjeżdża po jednym pojeździe. Na koniec każdej iteracji nauki przeprowadzane są symulacje w pełni zachłanne. Agent ma za zadanie wybierać akcje, które pozwolą na opuszczenie sieci dróg przez co najmniej 95 procent pojazdów. Jeśli ten warunek zostanie spełniony następuje zwiększenie ilości wjeżdżających pojazdów o 20 procent dotychczasowej wartości. W przypadku gdy ten warunek nie zostanie spełniony przez 15 kolejnych symulacji następuje koniec nauki.

Cel nauki i spodziewany wynik

Bezpośrednim celem nauki jest uzyskanie jak największej ilości pojazdów wjeżdżających do sieci. Aby to się stało agenci muszą podejmować decyzje, które pozwalają na opuszczenie sieci jak największej ilości pojazdów. Moment stopu spodziewany jest na chwilę, gdy na każdą z dróg będzie wjeżdżało niecałe 10 pojazdów. Zgodnie z opisem modelu ruchu - maksymalnie 10 pojazdów może wjechać w jednej chwili na dany odcinek, zatem muszą się tworzyć zatory przy większej ilości pojazdów.

Sposób uczenia

Zastosowana została metoda uczenia przedstawiona w (6.4).

Następujące stałe oraz ustalenia określają szczegóły uczenia:

- $\gamma = 0.9$ - stopa dyskontowa
- Parametr zachłanności ϵ jest początkowo równy 1 (w pełni losowe epizody). Z każdą iteracją jest zmniejszany o 0,01 aż do wartości $\epsilon = 0,2$.
- Generowanych jest jednorazowo 20 epizodów. Każdy z nich trwa 300 interwałów czasowych.

- Wygenerowane zbiory danych są dzielone w proporcjach 80 i 20 procent dla odpowiednio treningu i walidacji.
- Za uczenie odpowiada sieć neuronowa biblioteki Keras. Posiada ona 3 ukryte warstwy (10,14,10 neuronów) z funkcjami aktywacji relu. Optymalizacja jest przeprowadzana przez metodę Adam z funkcją straty błędu średniokwadratowego i parametrem $learning_rate = 0,01$.
- Epizod zachłanny trwa 1000 interwałów czasowych.

Wyniki uczenia

Moment stopu jest osiągnięty przy 8.9 pojazdów napływających do sieci ruchu. Najwyższy osiągnięty wyniki 20 epizodów z takim napływem ruchu to 90 procent. Żadna z akcji nie jest zaniechana - agenci co najmniej kilkadziesiąt razy podejmują każdą z dostępnych akcji w zachłannym epizodzie.

5.4 Środowisko 4 - kampus B Politechniki Łódzkiej



Rysunek 5.8: Środowisko 5 - Skrzyżowania wokół kampusu B Politechniki Łódzkiej. Mapy Open Street Map

Struktura środowiska bazuje na rzeczywistej sieci dróg znajdującej się w Łodzi. Sieć wyznaczają ulice Wróblewskiego, Wólczajska, Radwańska i aleja Politechniki Łódzkiej. Środowisko posiada 40 odcinków w tym 4 odcinki źródłowe (0,8,30,39) i 4 odcinki wylotowe (1,8,13,19). Drogi krzyżują się w 4 miejscach. Na dowolnym skrzyżowaniu kierowcy mogą wykonać manewry skrętu w lewo, w prawo oraz jazdy prosto.

5.4.1 Sygnalizacje świetlne

Każde skrzyżowanie posiada 5 faz świetlnych. Fazy 0,1,2 i 3 są fazami, które posiadają pewne zielone światła. Zmiana pomiędzy tymi fazami nie jest natychmiastowa i następuje dopiero po 2 interwałach fazy żółtych światel. Skrzyżowanie ulic Wróblewskiego, Wólczajskiej i Czerwonej posiada inne od pozostałych trzech skrzyżowań fazy świetlne. Wynika to

z kolizyjności lewoskrętów z ulic Wróblewskiego i Czerwonej. Poniższa tabela przedstawia możliwe fazy świetlne.

Faza	Wólczańska/ Wróblewskiego/ Czerwona	Pozostałe 3 skrzyżowania
0		
1		
2		
3		
Żółte		

5.4.2 Przestrzeń decyzyjna

Analogicznie do poprzednich przestrzeni decyzyjnych z fazą żółtych świateł - przykładowo (5.3.2).

5.4.3 Zatory drogowe

Wykorzystana w modelu zostanie funkcja zatoru przedstawiona w rozdziale (4.9.1), czyli:

$$f(i, j) = \begin{cases} \mathbf{P}[i, j] & \text{dla } j = i + 1 \wedge x[i] < 10 - \text{przejazd bez zatoru} \\ \frac{10}{x[i]} & \text{dla } j = i + 1 \wedge x[i] \geq 10 - \text{przejazd z zaturem} \\ 0 & \text{dla pozostałych przypadków - manewr [i,j]} \end{cases}$$

(5.10)

(5.11)

(5.12)

Zezwala ona na jednoczesny wjazd maksymalnie 10 pojazdów na dowolny odcinek drogi.

5.4.4 Uczenie

Ogólny zarys

Do każdej z dróg źródłowych wjeżdża pewna ilość pojazdów co interwał. Początkowo ilość ta jest zadana przez zmienną losową o rozkładzie lognormalnym o parametrach $\delta = 1, \mu = 1$. Na koniec każdej iteracji nauki przeprowadzane są symulacje w pełni zachłanne. Agent ma za zadanie wykonywać akcje, które w trakcie symulacji pozwolą na opuszczenie skrzyżowania przez co najmniej 95 procent pojazdów. W przypadku spełnienia tego warunku następuje zwiększenie parametru μ o 20 procent dotychczasowej wartości. Konsekwencją zwiększenia μ jest większa ilość wjeżdżających do układu pojazdów. Ten proces jest powtarzany, aż do momentu gdy przez 15 kolejnych iteracji uczenia nie uzyskany zostanie warunek 95 procent pojazdów opuszczających skrzyżowanie.

Cel nauki

Bezpośrednim celem nauki jest uzyskanie jak największej wartości parametru μ . Aby to się stało agenci muszą podejmować decyzje, które pozwalają na opuszczenie sieci możliwie jak największej ilości pojazdów.

Wyniki uczenia

Moment stopu przypadł na $\mu = 7,43$ pojazdu wjeżdżającego na drogę. Okazuje się, że proces uczenia doprowadził do zaniechania przez agentów manewrów lewoskrętów (akcje 0 i 1). Jest to z pewnością niepożądane zachowanie, jednak można było się go spodziewać, gdyż nagrody zostały wymodelowane tak, by gratyfikować jedynie przepływ - nieistotne z jakiego manewru.

5.4.5 Modyfikacja środowiska

Zostaje nałożona dodatkowa restrykcja związana z przestrzenią akcji. Spośród 100 kolejno podjętych akcji powinna znaleźć się każda możliwa akcja.

Wyniki uczenia zmodyfikowanego środowiska

Wyniki uczenia są identyczne do wariantu bez restrykcji wystąpienia wszystkich faz. Wyeliminowany został jednak niepożądany efekt zaniechania lewoskrętów.

Rozdział 6

Przegląd metod optymalizacyjnych

6.1 Kategorie uczenia maszynowego

Uczenie maszynowe to dziedzina wchodząca w skład nauk zajmujących się sztuczną inteligencją. Samo uczenie w najprostszym kształcie może być rozumiane jako dobór parametrów na podstawie dostępnych danych. Uczenie maszynowe jest powszechnie dzielone na 3 kategorie nauki [8].

1. Nadzorowane

Dane używane do uczenia nazywane są zbiorem treningowym. Każdy pojedynczy element zbioru treningowego ma informacje wejściowe oraz pewną pożądaną wartość wyjściową. W trakcie uczenia algorytm dopasowuje swoje parametry tak aby na podstawie danych wejściowych mógł przewidzieć wartość wyjściową. Przykładami uczenia nadzorowanego jest np. rozpoznawanie tekstu pisanego, detekcja obiektów na zdjęciach.

2. Nienadzorowane

Uczenie nienadzorowane różni się od nadzorowanego tym, że nie są znane pożądane wartości wyjściowe. Celem nauki jest na podstawie podobieństw poszczególnych elementów zgrupowanie ich do odpowiednich klas. Przykładem uczenia nienadzorowanego może być np. klasyfikacja gatunków drzew na podstawie danych o ich wysokościach i szerokości korony drzew.

3. Wzmocnione

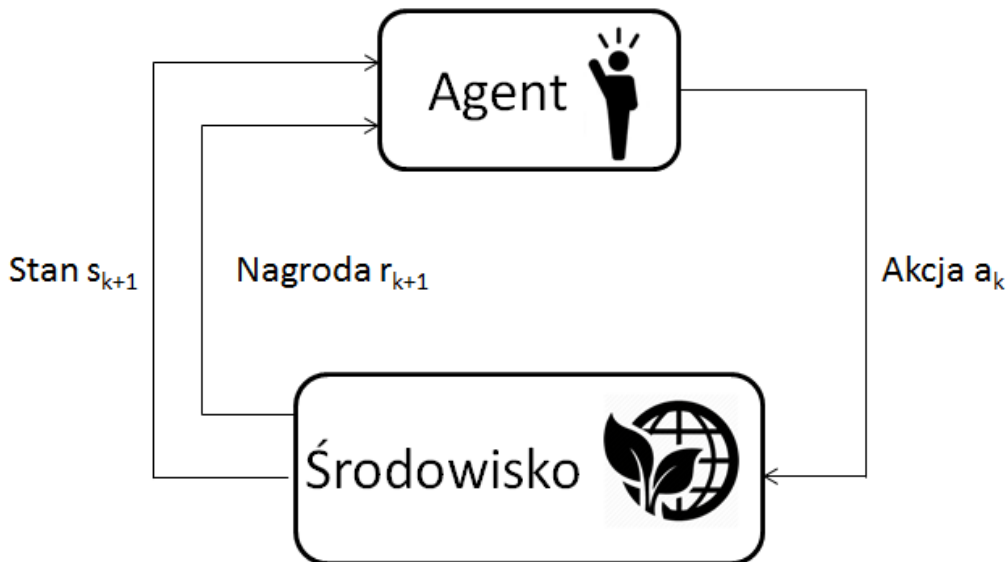
W środowisku uczenia ze wzmocnieniem istnieje agent, który jest odpowiedzialny za podejmowanie pewnych decyzji. Każda decyzja ma wpływ na stan środowiska, które zwraca agentowi nagrodę. Celem uczenia ze wzmocnieniem jest ustalenie strategii maksymalizującej skumulowaną wartość nagród.

Ze wszystkich trzech kategorii uczenie ze wzmocnieniem odpowiada problemowi rozważanym w rozdziale Y. Szczegółowy opis uczenia ze wzmocnieniem zostanie przedstawiony w następnej sekcji.

6.2 Uczenie ze wzmocnieniem

Schemat uczenia ze wzmocnieniem składa się z następujących elementów:

1. **Agent** jest odpowiedzialny za podejmowanie pewnych decyzji. Ma on wiedzę na temat obecnego stanu środowiska i otrzymuje w każdym kroku czasowym sygnał nagrody. Jego decyzje niebezpośrednio wpływają na stan środowiska.
2. **Środowisko** jest przestrzenią posiadającą dynamiczny stan widoczny dla agenta. Choć agent podejmuje akcje, to środowisko ma zdefiniowany model zmiany stanu. Model zmiany stanu może być stochastyczny oraz niewidoczny dla agenta. Oznacza to, że dwie te same akcje podjęte w tym samym stanie nie zawsze przyniosą identyczny następny stan. Innymi słowy agent nie może być stuprocentowo pewny rezultatów swoich akcji. Środowiska rozważane w rozdziale x są jednak w pełni deterministyczne. Środowisko jest także nadawcą sygnału nagrody.
3. **Strategia** definiuje sposób doboru akcji przez agenta w danej chwili. Jest to funkcja, która przyjmuje stan środowiska i zwraca akcję, która powinna być przeprowadzona.
4. **Sygnał nagrody** definiuje cel problemu uczenia ze wzmocnieniem. W każdym kroku czasowym środowisko wysyła agentowi liczbę rzeczywistą, która jest nazywana nagrodą (reward). Wartości nagród są czynnikiem wpływającym na zmianę strategii, gdyż zadaniem agenta jest maksymalizacja nagród. Wartość nagród zatem definiuje, które zdarzenia są dobre, a które złe dla agenta. Biologicznym odpowiednikiem dodatniej nagrody jest przyjemność, a ujemnej - ból [?].
5. **Funkcja wartości** zwraca wartość stanu czyli oczekiwaną sumę nagród jakie agent osiągnie w przyszłości będąc aktualnie w tym stanie.



Rysunek 6.1: Interakcje pomiędzy agentem a środowiskiem.

Algorytmy uczenia ze wzmocnieniem zazwyczaj stosuje się do rozwiązywania problemu procesu decyzyjnego Markowa. Sam **proces decyzyjny Markowa** jest zdefiniowany jako uporządkowana czwórka (S, A, P_a, R_a) , gdzie:

1. S to zbiór stanów
2. A to zbiór akcji. Notacją A_s oznaczane są możliwe akcje dla stanu s .
3. $P_a(s, s') = \text{Pr}(s_{t+1} = s' | s_t = s, a_t = a)$ to prawdopodobieństwo, że akcja a wykonana w stanie s w chwili t doprowadzi do stanu s' w chwili $t + 1$.
4. $R_a(s, s')$ to oczekiwana nagroda otrzymana w wyniku akcji podjętej w stanie s prowadzącej do stanu s' .

Problemem procesu decyzyjnego Markowa jest odnalezienie optymalnej strategii. Strategia określona jest jako funkcja $\pi(s)$ przyjmująca jako argument stan, a zwracająca podejmowaną akcję. Celem optymalizacji jest odnalezienie strategii maksymalizującej wartość:

$$G = \sum_{k=0}^K \gamma^k R_k \quad (6.1)$$

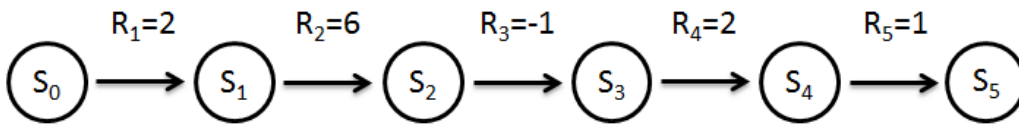
Chociaż we wzorze (6.1) nie ma strategii $\pi(s)$, to wpływa ona na otrzymywane nagrody R_k w każdej chwili k .

$\gamma \in (0, 1]$ jest czynnikiem dyskontującym. Idea dyskontowania nagród zaczerpnięta jest z rachunku finansowego. Przykładowo wpływ na konto kapitału 1000 złotych po upływie roku czasu jest z pewnością bardziej wartościowy niż za 20 lat. Pieniądze w matematyce finansowej są prawie zawsze liczone w czasie i tak samo należy postępować z nagrodami. Im wartość γ jest bliższa 0 tym bardziej istotne są początkowe nagrody. Dla $\gamma = 1$ wszystkie nagrody są równie istotne - bez względu na czas ich otrzymania.

Analogicznie do (6.1) jest ustalona funkcja wartości stanu. Jako wartość stanu s określone jest:

$$G_t = \sum_{k=t}^K \gamma^k R_k = R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + \dots + \gamma^K R_K \quad (6.2)$$

Zostanie przedstawiony teraz przykład obliczeniowy. Agent podejmuje decyzje na których podstawie otrzymuje ciąg nagród $R_0 = 0, R_1 = 2, R_2 = 6, R_3 = -1, R_4 = 2, R_5 = 1$. Czynniki dyskontujący γ jest równy 0,9. Jaka jest wartość G oraz G_1, G_2, G_3, G_4, G_5 ?



Rysunek 6.2: Ciąg nagród i stanów.

Najłatwiej obliczenia rozpocząć od G_5 i skończyć na G_1 .

$$G_5 = R_5. \quad (6.3)$$

Dla pozostałych t :

$$G_t = R_t + \gamma G_{t+1}. \quad (6.4)$$

$$G_5 = R_5 = 1.$$

$$G_4 = R_4 + \gamma G_5 = 2 + 0,9 \cdot 1 = 2,9.$$

$$G_3 = R_3 + \gamma G_4 = -1 + 0,9 \cdot 2,9 = 1,61.$$

$$G_2 = R_2 + \gamma G_3 = 6 + 0,9 \cdot 1,61 \approx 7,45.$$

$$G_1 = R_1 + \gamma G_2 = 2 + 0,9 \cdot 7,45 \approx 8,70.$$

$$G = R_0 + \gamma G_1 = 0 + 0,9 \cdot 8,70 \approx 7,83.$$

6.3 Metoda Monte Carlo On-Policy

Metody Monte Carlo są szeroką klasą algorytmów, których wyniki oparte są o losowe próbkowanie. Nie potrzebują one żadnej wiedzy na temat środowiska - akceptowalne są zarówno środowiska deterministyczne jak i stochastyczne. Algorytmy Monte Carlo uczenia ze wzmocnieniem są dzielone na dwie kategorie - On-Policy oraz Off-policy. W pracy zostanie przedstawiona metoda on-policy, która różni się od off-policy jedynie sposobem doboru momentu eksploracji. Metoda on-policy zakłada, iż ustalana jest pewna liczba ϵ bliska zeru. Określa ona prawdopodobieństwo kroku eksploracji, czyli wyboru losowej akcji. Pozostałe akcje są wybierane w sposób zachłanny, czyli podejmowana jest najbardziej wartościowa dostępna akcja. Algorytm przedstawiony jest poniższym pseudokodem.

1. Zainicjuj słowniki:

- (a) $Q(s, a)$ - Wartość określa opłacalność wyboru akcji a w stanie s
- (b) $Returns(s, a)$ - Wartość słownika to tablica wartości G wyliczonych na podstawie wzoru (6.2).
- (c) $\pi(s)$ - Określa jaką akcja ma zostać podjęta dla stanu s . Początkowo wszystkie akcje są wybrane losowo.

2. Zasymuluj pełny epizod wedle strategii π

3. Zoptymalizuj strategię π - dla każdej pary (s, a) - stanu i akcji, które pojawiły się w epizodzie

- (a) Wylicz G rekurencyjnie wedle wzoru (6.4) t.j. $G_t = R_t + \gamma G_{t+1}$.
- (b) Do tablicy $Returns(s, a)$ dodaj wartość G
- (c) $Q(s, a) = avg(Returns(s, a))$
- (d) $\pi(s) = a^*$. Gdzie a^* to taka akcja, że $Q(s, a^*) \geq Q(s, a)$ dla każdego $a \in A$. Jeśli został jednak wylosowany krok eksploracji, wtedy $\pi(s)$ jest losowo wybraną akcją $a \in A$.

Uwagi

Rozsądnym jest iterowanie w kroku 3a) w kolejności odwrotnej do chronologicznej, gdyż ułatwia wykorzystanie wzoru (6.4)

Warto zauważyć następujące zalety metody Monte Carlo:

1. Brak potrzeby wiedzy na temat modelu środowiska - t.j. rozkładu prawdopodobieństwa konsekwencji akcji.
2. Nie wymaga często nierealnego założenia, iż stan środowiska może zostać sztywno ustalony przez programistę. Wymagana jest jedynie możliwość przeprowadzania pełnych epizodów, co jest nieporównywalnie słabszym założeniem.
3. Adekwatne dane - czasu treningowy jest poświęcony dla stanów, które są często odwiedzane. Z kolei te, które nie pojawiają się prawie wcale - nie marnują dużo czasu w procesie nauki.

Jak się okazało dla bardziej złożonych środowisk (3 i 4 dodaj refa) - z dużą przestrzanią stanu i akcji powyższy algorytm jest niewystarczający. Należy zauważyć, że bardzo podobne stany nie mają ze sobą w tym algorytmie żadnej korelacji. Gdy przestrzeń jest bardzo duża, ograniczenia czasowe mogą nie pozwolić na eksplorację każdego stanu. Jako przykład problemu niech stanem środowiska będzie kolejka przed skrzyżowaniem. Wtedy dwa bardzo podobne stany [3, 22, 72] oraz [3, 22, 73] są od siebie różne i symulacje muszą prześledzić każdy z tych stanów. Zakładając, że wartości każdej zmiennej są ze zbioru $\{0, 1, \dots, 100\}$ liczność przestrzeni stanu jest równa aż 1 000 000. Lepiej skalującym rozwiązaniem jest zastosowanie sieci neuronowych zamiast tablic. Zostanie ono przedstawione w następnej sekcji.

6.4 Metoda Monte Carlo On-Policy z użyciem sieci neuronowej

Podstawową różnicą względem algorytmu Monte Carlo opartego na tablicach jest inny wzór rozwoju wartości. Formuła jest następująca:

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha[r + \gamma \max_{a' \in A'} Q(s', a')] \quad (6.5)$$

- s, a to para stanu i akcji, które pojawiły się w trakcie symulacji
- α to stała szybkości uczenia (ang. learning rate)
- r to nagroda przydzielona podczas symulacji dla akcji a podjętej w stanie s
- $\gamma \in [0, 1]$ to stopa dyskontowa
- s' to stan do którego doprowadziła akcja a podjęta w stanie s podczas symulacji
- A' to przestrzeń możliwych akcji do podjęcia w stanie s'

Nie jest ten wzór jednak używany do składowania wartości Q w tablicach, a do generowania danych do sieci neuronowej. Całkowity algorytm jest następujący:

1. Dla każdego agenta utwórz sieć neuronową. Będzie ona na wejście przyjmować stany środowiska. Wyjściowym sygnałem jest opłacalność akcji możliwych do podjęcia w stanie danym na wejściu - czyli Q .
2. Zasymuluj pewną ilość epizodów wedle losowej strategii. Akcje są wybierane w ϵ -greedy sposób. Zapisuj dane wygenerowane podczas tych epizodów:
 - (a) Stany środowiska
 - (b) Podjęte akcje
 - (c) Przydzielone nagrody
3. Na podstawie zapisanych danych utwórz następujące zbiory wykorzystane w sieci neuronowej:
 - (a) Zbiór sygnałów wejściowych (x-batch). Pojedynczym elementem tego zbioru jest tablica przedstawiająca stan środowiska.
 - (b) Zbiór oczekiwanych sygnałów wyjściowych (y-batch), czyli opłacalności akcji do podjęcia w stanie zadanym na wejściu. Dla elementu odpowiadającego akcji, która została podjęta w symulacji wylicz wartość wedle równania (6.5). Dla elementów odpowiadających pozostałym akcją, które nie zostały podjęte - sieć neuronowa przewiduje nagrodę.
4. Podziel powyższe zbiory na segment treningowy i walidacyjny.
5. Trenuj sieć neuronową do momentu, gdy wyniki dla zbioru walidacyjnego się polepszają.
6. Jeśli spełniony jest warunek stopu - przerwij algorytm. W przeciwnym razie przejdź do 2.

6.4.1 Wariant z fazą żółtych świateł

W momencie fazy żółtych świateł agent podejmuje zawsze akcję "żółte". Nie potrzebne jest zatem wyliczanie wartości Q dla tej wymuszonej akcji. Aby nie obciążać sieci neuronowej do aproksymacji nieistotnych wartości zostaną podjęte następujące kroki. Jeśli dana jest wiedza na temat długości fazy żółtych świateł należy zastąpić wzór (6.5) poniższymi formułami. Niech faza żółtych świateł trwa k interwałów czasowych, wtedy dla akcji a utrzymania obecnej fazy:

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha[r + \gamma \max_{a' \in A'} Q(s', a')], \quad (6.6)$$

dla akcji a zmiany obecnej fazy:

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha[\gamma^k r + \gamma^{k+1} \max_{a' \in A'} Q(s', a')]. \quad (6.7)$$

Bibliografia

- [1] Plan działania na rzecz wdrażania inteligentnych systemów transportowych w europie. *Komisja Wspólnot Europejskich, KOM*, 886, 2008.
- [2] Study on urban mobility – assessing and improving the accessibility of urban areas. *Komisja Wspólnot Europejskich, KOM*, 886, 2017.
- [3] Moshe Ben-Akiva, Michel Bierlaire, Didier Burton, Haris N Koutsopoulos, and Rabi Mishalani. Network state estimation and prediction for real-time traffic management. *Networks and spatial economics*, 1(3-4):293–318, 2001.
- [4] Samia Boubaker, Ferid Rehim, and Adel Kalboussi. Comparative analysis of microscopic models of road traffic data. In *Logistics (LOGISTIQUA), 2011 4th International Conference on*, pages 474–478. IEEE, 2011.
- [5] Wilco Burghout, Haris N Koutsopoulos, and Ingmar Andreasson. A discrete-event mesoscopic traffic simulation model for hybrid traffic simulation. In *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE*, pages 1102–1107. IEEE, 2006.
- [6] Carlos F Daganzo. The cell transmission model, part ii: network traffic. *Transportation Research Part B: Methodological*, 29(2):79–93, 1995.
- [7] Simone Göttlich, Michael Herty, and Ute Ziegler. Modeling and optimizing traffic light settings in road networks. *Computers & operations research*, 55:36–51, 2015.
- [8] Zhenyu Guan, Liangxu Bian, Tao Shang, and Jianwei Liu. When machine learning meets security issues: A survey. In *2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, pages 158–165. IEEE, 2018.
- [9] Dirk Helbing, Ansgar Hennecke, Vladimir Shvetsov, and Martin Treiber. Master: macroscopic traffic simulation based on a gas-kinetic, non-local traffic model. *Transportation Research Part B: Methodological*, 35(2):183–211, 2001.
- [10] Pushpendra Kumar, Rochdi Merzouki, Blaise Conrard, Vincent Coelen, and Belkacem Ould Bouamama. Multilevel modeling of the traffic dynamic. *IEEE Transactions on Intelligent Transportation Systems*, 15(3):1066–1082, 2014.
- [11] Michael James Lighthill and Gerald Beresford Whitham. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A*, 229(1178):317–345, 1955.

- [12] polskawliczbach.pl. Samochody osobowe w polsce w latach 2003-2016 (źródło: Gus). <http://www.polskawliczbach.pl/#transport-i-komunikacja>, 2019 (accessed March 3, 2019).
- [13] Statista. Number of passenger cars and commercial vehicles in use worldwide from 2006 to 2015 in (1,000 units). <https://www.statista.com/statistics/281134/number-of-vehicles-in-use-worldwide/>, 2019 (accessed March 3, 2019).
- [14] Richard S Sutton, Andrew G Barto, and Ronald J Williams. Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems Magazine*, 12(2):19–22, 1992.
- [15] TomTom. Tomtom traffic index measuring congestion worldwide). https://www.tomtom.com/en_gb/trafficindex/list?citySize=ALL&continent=ALL&country=ALL, 2019 (accessed March 3, 2019).
- [16] Vinh An Vu and Gary Tan. High-performance mesoscopic traffic simulation with gpu for large scale networks. In *Proceedings of the 21st International Symposium on Distributed Simulation and Real Time Applications*, pages 127–135. IEEE Press, 2017.