

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

Michał Lis

5 marca 2019

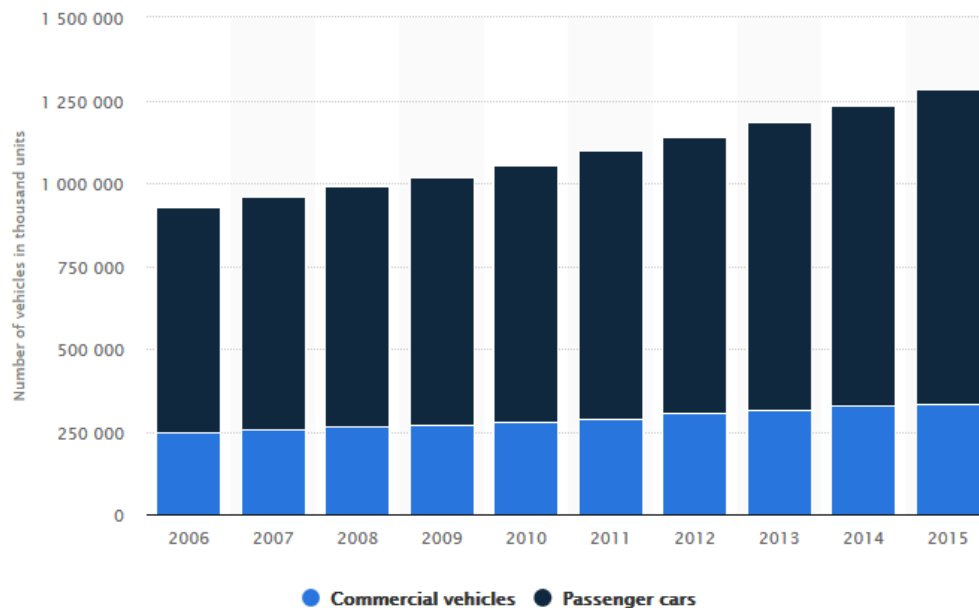
Spis treści

1	Wprowadzenie	5
2	Cel i zakres pracy	7
3	Siatka czasowa i przestrzenna	9
4	Makroskopowy model ruchu	11
4.1	Klasyfikacja modeli ruchu drogowego	11
4.2	Wstęp	11
4.3	Rozwój gęstości ruchu na drodze	12
4.4	Dyskretyzacja makroskopowego modelu ruchu	12
5	Model sieci dróg	15
5.1	Wstęp	15
5.2	Symbole i notacja obliczeniowa	15
5.3	Sieć dróg	15
5.4	Przykładowe skrzyżowanie	17
5.5	Macierz stanowa sieci dróg	17
5.6	Model przepływu - przykład	18
5.7	Model przepływu - ogólny wzór	19
5.8	Model przepływu z sygnalizacją świetlną - przykład	20
5.9	Model przepływu z sygnalizacją świetlną - ogólny wzór	21
5.10	Schematy skrzyżowań symulacyjnych	21
5.11	Przedstawienie problemu optymalizacyjnego	21
6	Przegląd metod optymalizacyjnych	25
6.1	Kategorie uczenia maszynowego	25
6.2	Uczenie ze wzmocnieniem	25
6.3	Programowanie dynamiczne	28

Rozdział 1

Wprowadzenie

Problem zatłoczonych ulic staje się coraz bardziej powszechny na całym świecie głównie ze względu na wzrost ekonomiczny. Według banku światowego w roku 1990 suma produktów krajowych brutto wszystkich państw świata wynosiła 22,6 biliona dolarów [1]. Najbardziej aktualne dane banku światowego dotyczą lat 2012-2017. W czasie ich trwania analogiczna wartość oscylowała wokół 75 bilionów dolarów. Nawet przy uwzględnieniu inflacji dolara, która od w okresie 1990 do 2015 wyniosła 80% światowy produkt brutto jest obecnie prawie dwukrotnie większy w porównaniu do roku 1990. Bezpośrednim skutkiem ogólnoświatowego wzrostu stopy życiowej jest zwiększenie ilości samochodów na drogach. Według danych firmy gromadzącej dane statystyczne Statista liczba zarejestrowanych pojazdów na świecie w roku 2006 wynosiła 947 tysięcy [2]. W 2015 roku po świecie jeździło już 1282 tysięcy pojazdów. Wzrost przez te 9 lat był niemalże liniowy - 39,4 tysięcy nowych samochodów rocznie wyznacza stopę wzrostu liczby pojazdów na poziomie 3,7%.



Rysunek 1.1: Liczba pojazdów na świecie

W Polsce wzrost ilości pojazdów w tych latach 2006 - 2015 był jeszcze większy [3]. W 2006 roku według GUS w Polsce było zarejestrowanych 13,4 miliona samochodów osobowych. W 2015 roku ich liczba wyniosła 20,7 miliona co oznacza 5 procentowy roczny wzrost. Najbardziej zatłoczonym polskim miastem jest Łódź. Zajmuje ona w rankingu stworzonym przez firmę TomTom bardzo wysokie 5 miejsce na świecie i 1 w Europie pod względem zatłoczenia dróg [4]. Oprócz Łodzi w pierwszej setce najbardziej zatłoczonych miast świata są także inne polskie miasta: Lublin(34), Kraków(48), Warszawa(50), Wrocław(63), Poznań(69), Bydgoszcz(83). Problem mocno dotyczy jednak także całej Europy. Spośród 100 najbardziej zatłoczonych miast świata aż 45 znajduje się w Europie. W 2008 roku Unia Europejska oszacowała iż koszty zatłoczenia dróg to 0,9% – 1,5% PKB unijnego [5]. Raport z 2017 roku może napawać optymizmem, gdyż przedstawione w nim wyliczenia określiły jedynie 0,77% straty całkowitego PKB wspólnoty [6]. Ten sam raport ocenia koszty zatorów komunikacyjnych w Polsce na poziomie 1,2% polskiego PKB. Problem w miastach jest o tyle trudny do rozwiązania, gdyż często nie ma miejsca na wybudowanie drogi o większej przepustowości. Rozwiązaniem może być wprowadzenie większej ilości sygnalizacji świetlnej. Istotną kwestią jest optymalizacja ustawień sygnalizacji świetlnej. W celu uzyskania danych potrzebnych do tego procesu rozstawiane są na drogach detektory ruchu.



Rysunek 1.2: System pętli do detekcji osi pojazdów zainstalowany na badawczym stanowisku pomiarowym na drodze DK81 będący w dyspozycji Katedry Metrologii i Elektroniki AGH.

Rozdział 2

Cel i zakres pracy

Celem pracy jest optymalizacja sygnalizacji świetlnej. Jako środowisko uczenia zostanie stworzony symulator ruchu drogowego. Symulacje ruchu będą w pełni zgodny z makroskopowym modelem ruchu. Sam makroskopowy model ruchu zostanie przedstawiony w rozdziale X. Jest to model ciągły. Pożądanym jest dyskretny model ruchu drogowego ze względu na łatwość implementacji komputerowej. Zostanie zatem przedstawiona w sekcji X.Y dyskretyzacja makroskopowego modelu ruchu. W rozdziale X zostanie zdefiniowany model sieci dróg. Na początku będzie on przedstawiony w sposób bardzo prosty, wręcz trywialny pomijając wiele aspektów. W każdej kolejnej sekcji model będzie stopniowo rozwijany. Sieci dróg zdefiniowane według końcowego modelu będą środowiskiem treningowym dla algorytmów uczenia maszynowego. Rozdział X opisuje uczenie ze wzmocnieniem - algorytm treningowy procesu optymalizacji sygnalizacji świetlnej. Rozdział X przedstawia cztery modele sieci dróg dla których został stworzony program symulacyjny. Rozdział X opisuje optymalizację sygnalizacji świetlnej dla wspomnianych sieci dróg.

Rozdział 3

Siatka czasowa i przestrzenna

Dyskretny charakter modelu przedstawianego w pracy obliguje do określenia siatki czasowej i przestrzennej. Dla par czasu i miejsc należących do tych dwóch siatek będą określane zmienne stanowe.

Siatka czasowa jest zdefiniowana jako skończony ciąg liczb naturalnych:

$$(0, 1, \dots, K). \quad (3.1)$$

Niech będzie ustalona droga e , która jest odcinkiem $[0, L_e]$. Droga zostaje podzielona na $L + 1$ odcinków o równej długości $\Delta x = \frac{L_e}{L+1}$. **Siatka przestrzenna** drogi to ciąg odcinków:

$$(b_l)_{l=0}^L = [l\Delta x, (l+1)\Delta x]$$



Rysunek 3.1: Siatka przestrzenna

Rozdział 4

Makroskopowy model ruchu

4.1 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 [7] - makroskopowy, mezoskopowy oraz mikroskopowy. Czasem [8] 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 [9]. Przykładowo pojazdy są zgrupowane w grupę podróżującą z pewnego punktu startowego do celu. Inne modele [10] 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 [11]. 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 [12] przedstawili pomysł przyrównania ruchu ulicznego na zatłoczonych drogach do przepływu wody w rzekach. Z tego powodu nie rozróżniamy w nim indywidualnie pojazdów, ani też nawet grupowo. Rozważamy natomiast gęstość ruchu w danym punkcie na drodze i czasie - czyli ilość pojazdów na danym odcinku drogi. Sposób w jaki poruszają się pojazdy jest wyliczany jedynie na podstawie gęstości ruchu. Jest to najmniej kosztowny obliczeniowo model. Właśnie w modelu makroskopowym zostało stworzone środowisko symulacyjne. Szczegóły modelu są przedstawione w następnym podrozdziale.

4.2 Wstęp

Istotą makroskopowego modelu ruchu jest pojęcie gęstości ruchu. Jest to zmienna stanowa określona dla każdego punktu drogi w czasie. Formalnie gęstość można rozumieć jako czynnik definiujący dynamikę ruchu. Im większa gęstość tym mniejsza prędkość ruchu. W niektórych

artykułach gęstość ruchu [13] jest przedstawiona jako iloraz ilości pojazdów znajdujących się na pewnym odcinku i długości tego odcinka drogi. Nie są to jednak czysto matematyczne formalne definicje. W makroskopowym modelu nie rozróżniamy pojedynczych pojazdów, ani nawet grup, więc taka definicja gęstości ruchu może być odebrana jako nieściśła z ideą modelu.

4.3 Rozwój gęstości ruchu na drodze

Makroskopowy model ruchu jest oparty o równanie różniczkowe (4.2) wraz z warunkiem początkowym (4.1). Model makroskopowy traktuje ruch uliczny na drogach podobnie do przepływu wody w rzece[ref]. Gęstość ruchu można utożsamiać z polem powierzchni przekroju poprzecznego rzeki, co dla ustalonej szerokości rzeki - upraszcza się do wysokości wody w rzece. Istotną uwagą w tym miejscu jest zaznaczenie, iż rzeka zazwyczaj posiada pewien spadek, który zapewnia ruch cieczy ze źródła do ujścia. Ruch makroskopowy zdefiniowany przez równanie (4.2) z kolei odnosi się do rzeki która jest na całym swoim odcinku pozioma. W takim przypadku de facto nie ma zdefiniowanego zwrotu ruchu.

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

$$\begin{cases} p(x, 0) = p_0(x) \\ \frac{\partial p(x, t)}{\partial t} + \frac{\partial f(p(x, t))}{\partial x} = 0 \end{cases} \quad (4.1)$$

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

Równanie (4.1) zakłada istnienie pewnej z góry nałożonej początkowej gęstości drogi $p_0(x)$. Równanie (4.2) określa wedle założeń modelu makroskopowego [12] rozwój gęstości ruchu na drodze. Funkcja płynności ruchu f powinna być wklęsła [ref]. W przedstawionym w tej pracy modelu funkcja ma następującą definicję:

$$f(p) = \begin{cases} \lambda p & \text{dla } p \in [0, p^*] \\ \lambda \cdot (2p^* - p) & \text{dla } p \in (p^*, p^{max}] \end{cases} \quad (4.3)$$

$$(4.4)$$

Gdzie λ jest stałym parametrem funkcji trójkątnej oraz $p^* = \frac{1}{2}p^{max}$.

4.4 Dyskretyzacja makroskopowego modelu ruchu

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:

$$p_l^k = \int_{b_l} \frac{p(x, k)}{\Delta x} dx. \quad (4.5)$$

Na podstawie (4.2) można wywnioskować, że:

$$\int_{b_l} p(x, k+1) - p(x, k) dx + \int_k^{k+1} f(b_{l+1}, k) - f(b_l, k) dk = 0 \quad (4.6)$$

Upraszczając otrzymujemy:

$$\Delta x(p_l^{k+1} - p_l^k) + \int_k^{k+1} f(b_{l+1}, k) - f(b_l, k) dk = 0 = 0 \quad (4.7)$$

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:

$$\Delta x(p_l^{k+1} - p_l^k) + (f(b_{l+1}, k) - f(b_l, k)) = 0 \quad (4.8)$$

Rezultatem jest końcowy rekurencyjny wzór na gęstość ruchu:

$$p_l^{k+1} = p_l^k - \frac{1}{\Delta x}(f(b_{l+1}, k) - f(b_l, k)) \quad (4.9)$$

Rozdział 5

Model sieci dróg

5.1 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: ujęcie sygnalizacji świetlnej, brak kolizyjnych manewrów, makroskopowy przepływ ruchu, przepływ ruchu na skrzyżowaniu, struktura sieci dróg. Zestawienie w jednej sekcji wszystkich tych kwestii byłoby bardzo przytłaczające.

5.2 Symbole i notacja obliczeniowa

Obliczenia w tym rozdziale będą wykonywane głównie na macierzach. Zazwyczaj macierze będą wyrazem ciągu, którego indeksem jest czas. Przykładowo $P(k)$ odnosi się do wartości w chwili k . Notacja $P(k, l)$ odnosi się do l -tej kolumny macierzy $P(k)$.

$\overrightarrow{P(k)}$ jest następującą operacją macierzową:

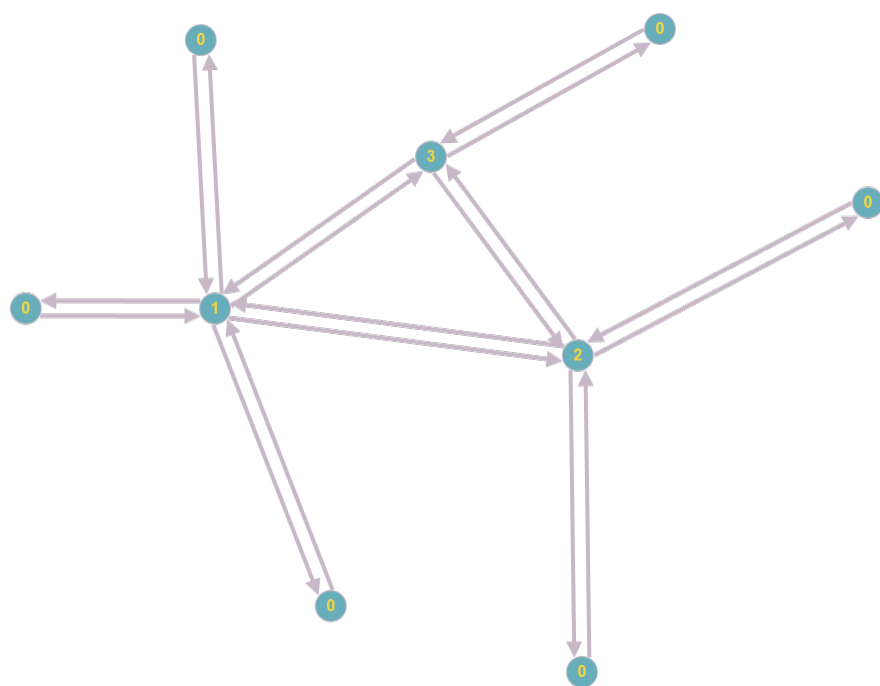
$$\overrightarrow{\begin{bmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}} = \begin{bmatrix} 0 & a_{11} & \dots & a_{1m-1} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n1} & \dots & a_{nm-1} \end{bmatrix} \quad (5.1)$$

Jak łatwo zauważyć przesuwana ona wszystkie kolumny o jeden indeks w prawo. Pierwsza kolumna jest zerowa.

5.3 Sieć dróg

Sieć dróg przedstawia uporządkowana dwójka $G = (V, E)$, gdzie:

- V to zbiór skrzyżowań
- E to zbiór dróg



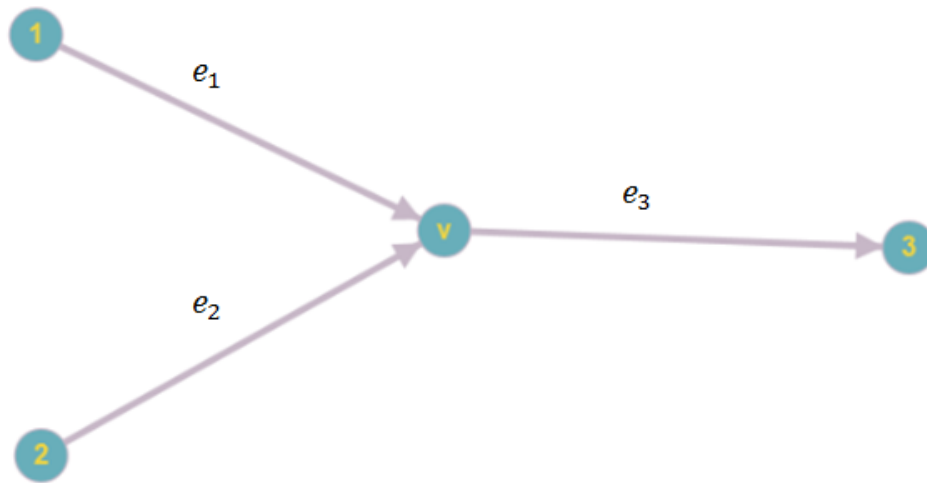
Skrzyżowania są oznaczone numerami: 1,2,3. Punkty z numerem 0 są jedynie początkami i końcami dróg.

Rysunek 5.1: Przykładowa sieć dróg

5.4 Przykładowe skrzyżowanie

Niech dane będzie skrzyżowanie dwóch dróg wlotowych e_1, e_2 i jednej wylotowej e_3 . Każda z dróg jest jednakowej długości. Struktura skrzyżowania jest przedstawiona na rysunku 5.2. Macierz B jest **macierzą przejścia** skrzyżowania v . Wartości 1 oznaczają możliwy przejazd na skrzyżowaniu z drogi odpowiadającej indeksowi kolumny do drogi zadanej przez indeks wiersza.

$$B = \begin{matrix} & \begin{matrix} e_1 & e_2 & e_3 \end{matrix} \\ \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{matrix} \quad (5.2)$$



Rysunek 5.2: Skrzyżowanie 1

5.5 Macierz stanowa sieci dróg

Strukturą przedstawiającą stan sieci dróg jest **macierz stanowa** $P(k)$. Składa się ona z zmiennych stanowych w danym momencie k . Każdy wiersz macierzy odpowiada jednej drodze e . Indeks kolumny określa konkretny odcinek tej drogi.

Wartość zmiennej stanowej początkowo jest ilością pojazdów na danym odcinku. Jako przykład zostanie przedstawiona początkowa macierz stanowa skrzyżowania z rysunku 5.2. Niech siatki przestrzenne dróg składają się z 4 odcinków. Początkowa macierz stanu $P(0)$ może być przedsta-

wiona jako:

$$P(0) = \begin{array}{cccc} b_0 & b_1 & b_2 & b_3 \\ \left[\begin{array}{cccc} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{array} \right] & \begin{array}{l} e_1 \\ e_2 \\ e_3 \end{array} \end{array} \quad (5.3)$$

Powyższa macierz przekazuje informację iż w chwili 0 pojazdy są na następujących drogach:

- e_1 - na odcinkach b_1, b_3
- e_2 - na odcinku b_3
- e_3 - na odcinku b_0

5.6 Model przepływu - przykład

Początkowy model przepływu pojazdów zakłada, iż wszystkie pojazdy w chwili $k + 1$ są o jeden odcinek dalej w swojej podróży niż w momencie k . Założone jest iż żadne nowe pojazdy nie pojawiają się w sieci dróg, a pojazdy będące w chwili k w ostatnim odcinku drogi e_3 opuszczają układ. Wtedy kolejne macierze stanowe dla przykładu 5.3 są następujące:

$$P(1) = \begin{array}{cccc} b_0 & b_1 & b_2 & b_3 \\ \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{array} \right] & \begin{array}{l} e_1 \\ e_2 \\ e_3 \end{array} \end{array} \quad (5.4)$$

$$P(2) = \begin{array}{cccc} b_0 & b_1 & b_2 & b_3 \\ \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \end{array} \right] & \begin{array}{l} e_1 \\ e_2 \\ e_3 \end{array} \end{array} \quad (5.5)$$

$$P(3) = \begin{array}{cccc} b_0 & b_1 & b_2 & b_3 \\ \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 1 \end{array} \right] & \begin{array}{l} e_1 \\ e_2 \\ e_3 \end{array} \end{array} \quad (5.6)$$

$$P(4) = \begin{array}{cccc} b_0 & b_1 & b_2 & b_3 \\ \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \end{array} \right] & \begin{array}{l} e_1 \\ e_2 \\ e_3 \end{array} \end{array} \quad (5.7)$$

5.7 Model przepływu - ogólny wzór

W poprzedniej sekcji został przedstawiony przykład wyliczenia macierzy stanowej dla kolejnych momentów w czasie. Następnym krokiem jest przedstawienie rozwiązania dla ogólnego przypadku.

Ogólny wzór przedstawiający rozwój macierzy stanowej to:

$$P(k+1) = \overrightarrow{P(k)} + S(k) \quad (5.8)$$

Macierz $S(k)$ nazywana będzie **macierzą źródła**. Dotyczy ona ruchu wpływającego do poszczególnych dróg w momencie $k+1$. Jest to macierz rzadka, istotna obliczeniowo jest jedynie pierwsza kolumna, a pozostałe wartości są zerowe by osiągnąć wymiar macierzy pozwalający na dodawanie z $\overrightarrow{P(k)}$. Pierwsza kolumna jest zdefiniowana w oparciu o ostatnią kolumnę macierzy $P(k)$ i jest równa:

$$S(k, 0) = B \cdot P(k, l) \quad (5.9)$$

Dla przykładu z poprzedniej sekcji, gdzie

$$P(0) = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Dla kolejnego momentu $k=1$ korzystając ze wzoru (5.8):

$$P(1) = \overrightarrow{P(0)} + S(0) \quad (5.10)$$

Należy wyliczyć macierz źródła, jej pierwsza kolumna to:

$$S(0, 0) = B \cdot P(0, l) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Zatem

$$S(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}$$

Oczywiście

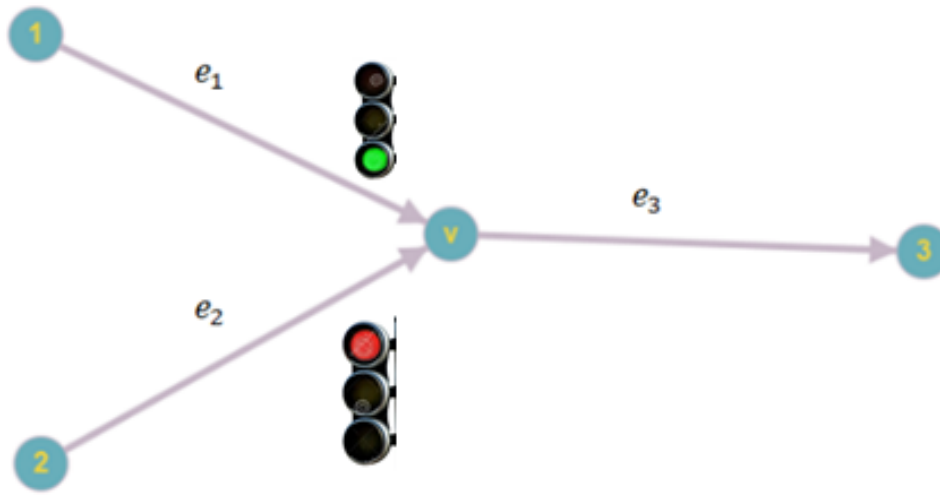
$$\overrightarrow{P(0)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Ostatecznie ze wzoru (5.10) wynika, że

$$P(1) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{bmatrix}$$

Warto zauważyć, że w chwili 0 na obydwu drogach wlotowych e_1, e_2 pojazdy znajdowały się na końcowym odcinku. Problemem okazuje się kolizyjność manewrów wjazdu na drogę e_3 z dróg odpowiednio e_1 i e_2 . Początkowo ta kwestia została ona pominięta. Następna sekcja przedstawia rozwiązanie tego problemu.

5.8 Model przepływu z sygnalizacją świetlną - przykład



Rysunek 5.3: Skrzyżowanie 1 z sygnalizacją świetlną w chwili 0

Wprowadzona zostaje sygnalizacja świetlna dla dwóch dróg wlotowych e_1, e_2 . Sygnalizacja świetlna będzie oparta o ciąg **macierzy sygnalizacji**. Określają one z której drogi pojazdy będą mogły wjechać na drogę e_3 . Dla chwili 0 przyjęte jest, że zgodnie z rysunkiem 5.3 pojazdy na ostatnim odcinku drogi e_2 będą musiały poczekać. Przykładowa początkowa macierz to:

$$M(0) = \begin{matrix} & \begin{matrix} e_1 & e_2 & e_3 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.11)$$

Przekazuje ona informację, że:

- Jest zielone światło na drodze e_1 w chwili 0
- Jest czerwone światło na drodze e_2 w chwili 0

Poniższy przykład obrazuje rozwój macierzy stanowej dla skrzyżowania z sygnalizacją świetlną.

$$P(0) = \begin{matrix} & \begin{matrix} b_0 & b_1 & b_2 & b_3 \end{matrix} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.12)$$

Zgodnie z rysunkiem 5.3 czerwone światło na drodze e_2 powoduje, że pojazdy nie opuszczają ostatniego odcinka tej drogi. Z kolei pojazdy ostatniego odcinka drogi e_1 wjeżdżają na drogę e_3 . Zatem:

$$P(1) = \begin{matrix} & \begin{matrix} b_0 & b_1 & b_2 & b_3 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.13)$$

Sygnalizacja świetlna nie zmienia się w chwili 1. Dochodzi do sytuacji gdy ruch kumuluje się na ostatnim odcinku drogi e_2 .

$$P(2) = \begin{matrix} & \begin{matrix} b_0 & b_1 & b_2 & b_3 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.14)$$

Niech w chwili 2 będzie zielone światło dla drogi e_2 , czyli

$$M(2) = \begin{matrix} & \begin{matrix} e_1 & e_2 & e_3 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.15)$$

Początkowo założone jest, że wszystkie pojazdy na ostatnim odcinku drogi przejeżdżają przez skrzyżowanie. Wtedy:

$$P(3) = \begin{matrix} & \begin{matrix} b_0 & b_1 & b_2 & b_3 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix} & \begin{matrix} e_1 \\ e_2 \\ e_3 \end{matrix} \end{matrix} \quad (5.16)$$

5.9 Model przepływu z sygnalizacją świetlną - ogólny wzór

5.10 Schematy skrzyżowań symulacyjnych

5.11 Przedstawienie problemu optymalizacyjnego

Następnym krokiem jest wyznaczenie ogólnego wzoru macierzy stanowej z uwzględnieniem sygnalizacji świetlnej.

Macierz źródła z uwzględnieniem sygnalizacji świetlnej jest przedstawiona wzorem:

$$S(k, 0) = M(k) \cdot P(k, l) \quad (5.17)$$

Oprócz pierwszej kolumny macierz S dalej posiada wartości zerowe.

Wprowadzony zostanie ciąg $W(k)$ pomocniczych **macierzy oczekujących pojazdów**. Będą to macierze rzadkie z wartościami innymi niż 0 tylko w ostatniej kolumnie. Macierze te będą składowały informację o pojazdach oczekujących na zielone światło w swojej ostatniej kolumnie. Jest ona zdefiniowana jako:

$$W(k, l) = P(k, l) - S(k, 0)$$

Wzór określający rozwój macierzy stanowej jest następujący:

$$P(k+1) = \overrightarrow{P(k)} + S(k) + W(k) \quad (5.18)$$

Dla wcześniej przedstawionego przykładu w tabeli są wypisane obliczenia wedle wzoru (5.18).

Tablica 5.1: Rozwój macierzy stanowej na skrzyżowaniu 1 z sygnalizacją świetlną

k	M(k)	P(k)	$\overrightarrow{P(k)}$	S(k,0)	W(k,1)
0	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & \mathbf{1} \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$
1	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & \mathbf{0} \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$
2	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & \mathbf{2} \\ 0 & 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
3		$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$			

Pogrubione wartości odnoszą się do pojazdów, które przejeżdżają w momencie k przez skrzyżowanie.

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 [14].

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 wiedzy 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.

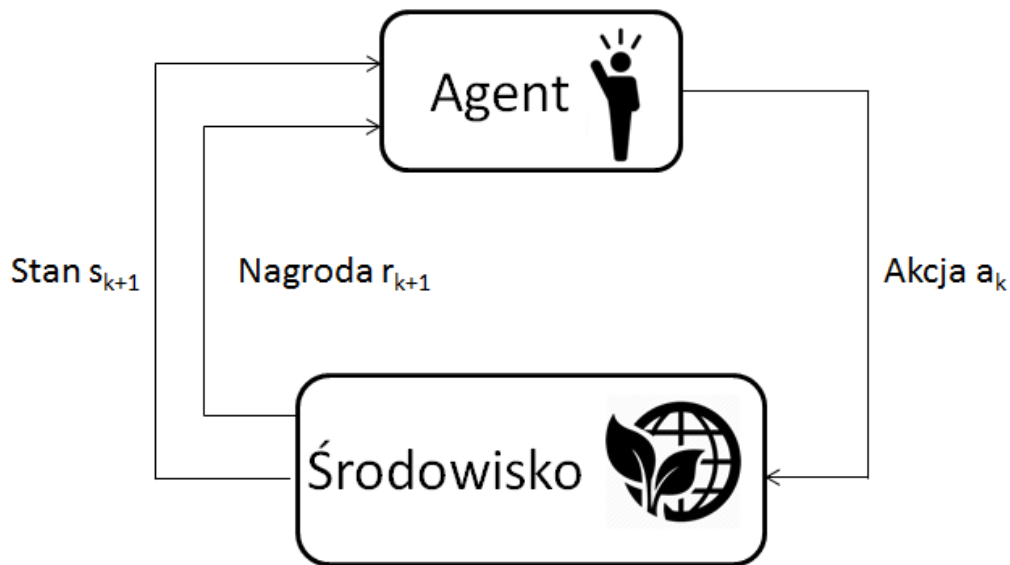
Ze wszystkich trzech kategorii uczenie ze wzmocnieniem odpowiada problemowi rozważanym w rozdziale Y. Szczegółowy opis uczenia ze wzmocnieniem zostanie przedstawiony w następnej sekcji.

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 wpływają na stan środowiska.

2. **Środowisko** jest przestrzenią posiadającą dynamiczny stan widoczny dla agenta. Chociaż to 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. 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 ma 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. $P_a(s, s') = Pr(s_{t+1} = s' | s_t = s, a_t = a)$ 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:

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

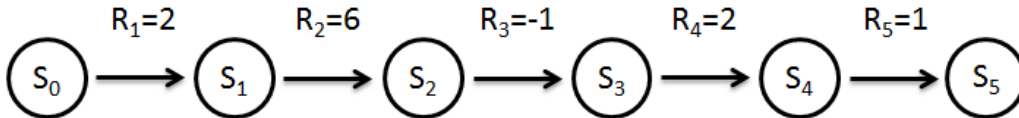
Chociaż strategii $\pi(s)$ nie ma we wzorze 6.1, to strategia wpływa na otrzymywane nagrody R_k w każdej chwili k .

$\gamma \in (0, 1]$ jest czynnikiem dyskontującym. Idea dyskontowania nagród zaczerpnięta jest z rachunku finansowego. Przykładowo wpływ 1000 złotych po upływie roku czasu jest z pewnością bardziej wartościowy niż za 20 lat. Innymi słowy - pieniądze są 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 , ponieważ $G_t = R_t + \gamma G_{t+1}$.
 $G_5 = R_5 = 1$.

$$\begin{aligned}
G_4 &= R_4 + \gamma G_5 = 2 + 0,9 \cdot 1 = 2,9 \\
G_3 &= R_3 + \gamma G_4 = -1 + 0,9 \cdot 2,9 = 1,61 \\
G_2 &= R_2 + \gamma G_3 = 6 + 0,9 \cdot 1,61 \approx 7,45 \\
G_1 &= R_1 + \gamma G_2 = 2 + 0,9 \cdot 7,45 \approx 8,70 \\
G &= R_0 + \gamma G_1 = 0 + 0,9 \cdot 8,70 \approx 7,83
\end{aligned}$$

6.3 Programowanie dynamiczne

Termin programowania dynamicznego odnosi się do algorytmów wyliczających optymalne strategie procesu decyzyjnego Markowa w przypadku posiadanej całkowitej wiedzy na temat modelu środowiska [15]. Środowisko nie musi być w pełni deterministyczne tzn. nie za każdym razem akcja przeprowadzana ze stanu s_k musi w efekcie doprowadzić do tego samego stanu s_{k+1} . Jednak w takim przypadku musi być znany rozkład prawdopodobieństwa przydzielania nowego stanu na podstawie poprzedniego i właśnie podjętej przez agenta akcji. Dodatkowo wymagana jest możliwość ustalenia dowolnego stanu w trakcie uczenia.

Początkowa strategia $\pi(s)$ jest dowolna, najczęściej losowa. Przedstawiony algorytm jest podzielony na 2 części. Część predykcji(prediction) oraz kontroli(control).

Proces predykcji ma za zadanie ustalenie wartości stanów na podstawie ustalonej strategii. Jej algorytm jest następujący:

1. Przyjmij daną z góry π jako strategię podejmowania akcji
2. Zainicjuj tablicę wartości stanów $V(s)$. Dla wszystkich możliwych stanów $s \in S$ przyjmij wartość 0.
3. $\Delta = 0$
4. Dla każdego $s \in S$:
 - (a) $v = V(s)$
 - (b) $V(s) = \sum_{s'} Pr(s'|s, \pi(s))[r + \gamma V(s')]$
 - (c) $\Delta = \max(\Delta, |v - V(s)|)$
5. Jeśli $\Delta < \theta$ to wróć do 3
6. Zwróć $V(s)$ jako tablicę wartości stanów $V_\pi(s)$ dla strategii π .

Parametr $\theta \geq 0$ definiuje w kroku 5. moment stopu.

Wartość $Pr(s', s, \pi(s))$ to prawdopodobieństwo, że akcja $\pi(s)$ podjęta w stanie s doprowadzi do stanu s' . Z kolei r jest właśnie otrzymaną nagrodą.

Algorytm **kontroli** ma za zadanie odnaleźć bardziej optymalną strategię niż dotychczas. Jako argument przyjmuje on wyliczoną właśnie tablicę wartości stanów $V_\pi(s)$. Wprowadzona zostaje macierz $Q_\pi(s, a)$. Jest ona zdefiniowana następująco:

$$\begin{aligned}
Q_\pi(s, a) &= E[R_{t+1} + \gamma v_\pi(S_{t+1}) | S_t = s, A_t = a] \\
&= \sum_{s'} Pr(s'|s, a)[r + \gamma V_{\pi}(s')]
\end{aligned} \tag{6.3}$$

Macierz przedstawia wartość akcji a podjętej w stanie s . Algorytm kontroli jest następujący:

1. Przyjmij wyliczoną przez algorytm predykcji macierz wartości stanów $V_\pi(s)$
2. Zainicjuj tablicę wartości stanów $V(s)$. Dla wszystkich możliwych stanów $s \in S$ przyjmij wartość 0.
3. $policy_stable = false$
4. Dla każdego $s \in S$:
 - (a) $old_action = \pi(s)$
 - (b) $\pi(s) = \operatorname{argmax}_{a \in A_s} \sum_{s'} Pr(s'|s, a)[r + \gamma V(s')]$
 - (c) Jeśli $old_action \neq \pi(s)$ to $policy_stable = true$
5. Jeśli $policy_stable = true$ to zwróć strategię $\pi(s)$.

Bibliografia

- [1] World Bank, “World development indicators,” <https://databank.worldbank.org/data/reports.aspx?source=2&country=&series=NY.GDP.MKTP.CD&period=>, 2019 (accessed March 3, 2019).
- [2] 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).
- [3] polskawliczbach.pl, “Samochody osobowe w polsce w latach 2003-2016 (źródło: Gus),” <http://www.polskawliczbach.pl/#transport-i-komunikacja>, 2019 (accessed March 3, 2019).
- [4] 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).
- [5] “Plan działania na rzecz wdrażania inteligentnych systemów transportowych w europie,” *Komisja Wspólnot Europejskich, KOM*, vol. 886, 2008.
- [6] “Study on urban mobility – assessing and improving the accessibility of urban areas,” *Komisja Wspólnot Europejskich, KOM*, vol. 886, 2017.
- [7] S. Boubaker, F. Rehim, and A. Kalboussi, “Comparative analysis of microscopic models of road traffic data,” in *Logistics (LOGISTIQUA), 2011 4th International Conference on*. IEEE, 2011, pp. 474–478.
- [8] P. Kumar, R. Merzouki, B. Conrard, V. Coelen, and B. O. Bouamama, “Multilevel modeling of the traffic dynamic,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 3, pp. 1066–1082, 2014.
- [9] W. Burghout, H. N. Koutsopoulos, and I. Andreasson, “A discrete-event mesoscopic traffic simulation model for hybrid traffic simulation,” in *Intelligent Transportation Systems Conference, 2006. ITSC’06. IEEE*. IEEE, 2006, pp. 1102–1107.
- [10] M. Ben-Akiva, M. Bierlaire, D. Burton, H. N. Koutsopoulos, and R. Mishalani, “Network state estimation and prediction for real-time traffic management,” *Networks and spatial economics*, vol. 1, no. 3-4, pp. 293–318, 2001.

- [11] V. A. Vu and G. 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*. IEEE Press, 2017, pp. 127–135.
- [12] M. J. Lighthill and G. B. Whitham, “On kinematic waves ii. a theory of traffic flow on long crowded roads,” *Proc. R. Soc. Lond. A*, vol. 229, no. 1178, pp. 317–345, 1955.
- [13] D. Helbing, A. Hennecke, V. Shvetsov, and M. Treiber, “Master: macroscopic traffic simulation based on a gas-kinetic, non-local traffic model,” *Transportation Research Part B: Methodological*, vol. 35, no. 2, pp. 183–211, 2001.
- [14] Z. Guan, L. Bian, T. Shang, and J. Liu, “When machine learning meets security issues: A survey,” in *2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*. IEEE, 2018, pp. 158–165.
- [15] R. S. Sutton, A. G. Barto, and R. J. Williams, “Reinforcement learning is direct adaptive optimal control,” *IEEE Control Systems Magazine*, vol. 12, no. 2, pp. 19–22, 1992.