

# SAMOUCZĄCA SIĘ MASZYNA

Waldemar Wietrzykowski

Digital Intelligence Laboratory, grudzień 2016

**Streszczenie** *W niniejszej pracy przedstawiono matematyczne podstawy samouczącej się maszyny, w tym pokazano, dlaczego istnieje maszyna Turinga.*

**Abstract** *In this work presented mathematical basics of machine self-learning, in this shown, why there is a Turing machine.*

Rozważmy abstrakcyjną samouczącą się maszynę, która na etapie uczenia opisana jest funkcją wejściową  $g()$  działającą na strukturze opisanej funkcjami wyjściowymi  $h[i]()$ , zwanymi maszynami jednofunkcyjnymi, w sposób następujący:

$$i = g(X, U, D, Q) \\ Y = h[i](X)$$

gdzie:

$$X = (x_1, \dots, x_n), \quad U = (u), \quad D = (d), \\ Q = (q), \quad Y = (y, \delta, s)$$

w której funkcja wejściowa  $g()$  wyznacza dla sygnałów  $X, U, D, Q$  indeks takiej jednowartościowej funkcji wyjściowej  $h[i](X)$  która spełnia:

$$U.(u) = Y.(y) \\ D.(d) = Y.(\delta) \\ Q.(q) = Y.(s)$$

Ponadto każda z funkcji wyjściowej  $h[i](X)$  spełnia warunek jednoznaczności (ze względu na sygnał uczący  $U$ ) oznaczający, że dla każdego sygnału wejściowego  $X$  wartości funkcji wyjściowych  $Y.(y)$  dla wszystkich indeksów nie powtarzają się lub są puste (gotowe do przyjęcia wartości).

Oznacza to, że jeżeli dla jednego sygnału wejściowego  $X$  (obiektu) przyporządkowanych jest wiele wartości sygnałów uczących  $U.(u)$  (własności obiektu) to dla tego obiektu odpowiadające sygnały wyjściowe  $Y.(y)$

znajdują się w różnych funkcjach wyjściowych  $h[i](X)$  występujących pod różnymi indeksami.

Oznaczenia:

- $X$  - jest wektorem złożonym z wartości sygnałów wejściowych w chwili  $t$ , zwany sygnałem wejściowym, przeznaczonym do rozpoznania obiektu lub zdarzenia,
- $U$  - jest wektorem złożonym z wartości sygnałów uczących w chwili  $t$ , zwany sygnałem uczącym,
- $D$  - jest wektorem złożonym z wartości sygnałów dodatkowych w chwili  $t$ , zwany sygnałem dodatkowym,
- $Q$  - jest wektorem złożonym z wartości sygnałów wewnętrznych maszyny w chwili  $t$ , zwany sygnałem wewnętrznym,
- $Y$  - jest wektorem złożonym z zapamiętanych wartości sygnałów: uczącego, dodatkowego i wewnętrznego zwany sygnałem wyjściowym,
- $i$  - jest wartością indeksu wskazującą taką funkcję wyjściową  $f[i](X)$ , która sygnałom wejściowych  $X$  i wartościom sygnałów uczących  $U.(u)$  przyporządkowuje dokładnie jedną wartość sygnału wyjściowego  $Y.(y)$ .

Dodać należy, że wektory  $U, D, Q, Y$  składać się mogą z większej ilości wartości niż wskazano, lecz uczyniono to celowo dla możliwie najprostszego objaśnienia istoty uczenia się maszyny.

Na etapie eksploracji zgromadzonej przez maszynę wiedzy, gdzie nie występuje już

sygnał uczący  $U$ , maszynę tą opisuje się funkcjami:

$$\begin{aligned} i &= g(X, D, Q) \\ Y &= h[i](X) \end{aligned}$$

gdzie:

$$\begin{aligned} X &= (x_1, \dots, x_n), \quad D = (d), \quad Q = (q), \\ Y &= (y, \delta, s) \end{aligned}$$

w której funkcja wejściowa  $g()$  wyznacza dla sygnałów  $X$ ,  $D$ ,  $Q$  indeks takiej funkcji wyjściowej  $h[i]()$  która spełnia:

$$\begin{aligned} D.(d) &= Y.(\delta) \\ Q.(q) &= Y.(s) \end{aligned}$$

Ponadto każda z funkcji wyjściowej  $h[i](X)$  spełnia warunek jednoznaczności (ze względu na sygnał dodatkowy  $D$  lub sygnał wewnętrzny  $Q$ ) oznaczający, że dla każdego sygnału wejściowego  $X$  wartości funkcji wyjściowych  $Y.(\delta)$  lub  $Y.(s)$  dla wszystkich indeksów nie powtarzają się lub są puste (gotowe do przyjęcia wartości).

Oznacza to, że jeżeli dla jednego sygnału wejściowego  $X$  (obiektu) przyporządkowanych jest wiele wartości sygnałów dodatkowych  $U.(d)$  lub sygnałów wewnętrznych  $U.(q)$  to dla tego obiektu odpowiadające sygnały wyjściowe  $Y.(\delta)$  lub  $Y.(s)$  znajdują się w różnych funkcjach wyjściowych  $h[i](X)$  występujących pod różnymi indeksami

Z powyższego wynika, że na etapie eksploracji zgromadzonej wiedzy pomimo braku sygnału  $U$  wiedzę można jednoznacznie odtworzyć na podstawie sygnałów  $D$  lub  $Q$  jeżeli każda z funkcji wyjściowej  $h[i](X)$  spełnia warunek jednoznaczności ze względu na  $D$  lub  $Q$ . Odtwarzanie wiedzy na podstawie sygnałów  $D$  lub  $Q$  nosi nazwę odtwarzania nauczonej wiedzy poprzez kojarzenie.

Jeżeli w maszynie na etapie eksploracji zgromadzonej wiedzy sygnały  $D$  lub  $U$  nie są jednoznaczne lub w funkcji wejściowej  $g()$  całkowicie pominięto sygnały  $D$  i  $U$  to nie mamy żadnej gwarancji, aby funkcja wejściowa  $g()$  była jednoznaczna i wówczas funkcję  $g()$

musimy przedstawić jako funkcję wielowartościową (multifunkcję):

$$i \approx g(X)$$

Istnieje sposób na poradzenie sobie z wielowartościowością funkcji wejściowej  $g(X)$  poprzez zastosowanie automatycznego indeksu polegającego na tym, że zamiast wyliczać indeks funkcji wyjściowej  $h[i](X)$  za pomocą funkcji  $g(X)$  możemy pobierać go wprost z wartości  $Y_{t-1}.(i_t)$  ostatnio aktywowanego w chwili  $t-1$  wektora wyjściowego  $Y_{t-1}$ . Maszynę z takim automatem na etapie eksploracji zgromadzonej wiedzy możemy zapisać:

$$\begin{aligned} i_t &= Y_{t-1}.(i_t) \\ Y_t &= h[i_t](X) \end{aligned}$$

gdzie:

$$\begin{aligned} g(X) &= Y_{t-1}.(i_t), \quad Y_{t-1} = (y, \delta, s, i_t), \\ Y_t &= (y, \delta, s, i_{t+1}) \end{aligned}$$

przy czym warunek jednoznaczności jest spełniony, gdyż  $i_t = g(X) = Y_{t-1}.(i_t)$  a to oznacza, że w chwili  $t$  zostaje wyznaczony tylko jeden indeks  $i_t$  równy wartości  $(i_t)$  ostatnio aktywowanego w chwili  $t-1$  wektora wyjściowego  $Y_{t-1}$ . Jak z powyższego wynika indeks  $i_t$  nie jest zależny od sygnału wejściowego  $X$ , a więc dla całego zakresu  $X$  w chwili  $t$  indeks ten jest stały co oznacza, że cały zakres  $X$  w chwili  $t$  jest obsługiwany tylko przez jedną funkcję wyjściową  $Y_t = h[i_t](X)$  zwaną maszyną jednofunkcyjną.

Od sygnału wejściowego  $X$  zależy jedynie wybór jednego z wektorów wyjściowych  $Y_t$  należących do tylko jednej funkcji wyjściowej (maszyny jednofunkcyjnej)  $h[i_t](X)$  o indeksie  $i_t$ .

Sygnały wyjściowe  $Y_t$  są więc tak przydzielone to funkcji wyjściowych  $h[i_t](X)$ , aby następny sygnał wyjściowy  $Y_{t+1}$  należał albo do tej samej funkcji wyjściowej albo do innej tylko jednej określonej funkcji wyjściowej.

Można zauważyć, że powyższy zapis maszyny z automatycznym indeksem można przedstawić również następująco:

$$i_t, X \rightarrow y, \delta, s, i_{t+1}$$

Zapis ten przedstawia rozkaz maszyny Turinga znajdującej się w stanie ( $i_t$ ) który, jeżeli symbol odczytany z taśmy maszyny jest równy  $X$ , zapisuje na taśmie symbol ( $y$ ) i przesuwa głowicę w kierunku ( $\delta$ ) oraz przechodzi do stanu następnego ( $i_{t+1}$ ).

Jak z powyższego wynika, maszyna Turinga jest jedynie przypadkiem samouczącej się maszyny (z automatycznym indeksem) będącej na etapie eksploracji zgromadzonej wiedzy i dlatego do swojej pracy wymaga zgromadzonej już wiedzy, której sama się nie uczy lecz dostarcza ją programista maszyny za pomocą zaprojektowanej przez niego listy rozkazów.

Samouczącą się maszynę z automatycznym indeksem na etapie uczenia można przedstawić następująco:

$$i_t = g(X, U, D, Q) = Y_{t+1} \cdot (i_t)$$

$$Y_t = h[i_t](X)$$

gdzie:

$$Y_{t+1} = (y, \delta, s, i_t), \quad Y_t = (y, \delta, s, i_{t+1})$$

w której funkcja wejściowa  $g()$  wyznacza dla sygnałów  $X, U, D, Q$  indeks takiej funkcji wyjściowej  $h[i_t]()$  która spełnia:

$$U.(u) = Y.(y)$$

$$D.(d) = Y.(\delta)$$

$$Q.(q) = Y.(s)$$

Ponadto każda z funkcji wyjściowej  $h[i_t](X)$  spełnia warunek jednoznaczności ze względu na sygnał uczący  $U$  oznaczający, że dla

każdego sygnału wejściowego  $X$  wartości funkcji wyjściowych  $Y.(y)$  dla wszystkich indeksów nie powtarzają się lub są puste (gotowe do przyjęcia wartości) i dodatkowo są tak przydzielone to tych funkcji, aby następny sygnał wyjściowy  $Y_{t+1}$  należał albo do tej samej funkcji wyjściowej albo do innej tylko jednej określonej funkcji wyjściowej (czyli do tej samej lub innej tylko jednej określonej maszyny jednofunkcyjnej).

Oznacza to, że jeżeli dla jednego sygnału wejściowego  $X$  (obiektu) przyporządkowanych jest wiele wartości sygnałów uczących  $U.(u)$  (własności obiektu) to dla tego obiektu odpowiadające sygnały wyjściowe  $Y.(y)$  znajdują się w różnych funkcjach wyjściowych  $h[i_t](X)$  pod takimi indeksami, aby w kolejnej chwili  $t+1$  wyznaczony był ten sam lub tylko jeden inny określony indeks.

Od sygnału wejściowego  $X$  zależy jedynie wybór jednego z wektorów wyjściowych  $Y_t$  należących do tylko jednej funkcji wyjściowej (maszyny jednofunkcyjnej)  $h[i_t](X)$  o indeksie  $i_t$ .

Reasumując w samouczącej się maszynie z automatycznym indeksem w kolejnej chwili czasowej  $t+1$  dla wszystkich  $X$  zostaje wybrana tylko ta sama lub tylko jedna inna maszyna jednofunkcyjna, tym samym każdemu sygnałowi  $X$  przyporządkowany jest wektor wyjściowy  $Y_{t+1}$  należący do tej samej lub tylko jednej innej maszyny jednofunkcyjnej.