УДК 681.14

Ф.Е. Гече, В.М. Коцовський, С.А.Ковальов
Ужгородський нац. ун-т,
А.Є. Батюк
Національний університет "Львівська політехніка"

БУЛЬОВІ НЕЙРОФУНКЦІЇ І СИНТЕЗ РОЗПІЗНАЮЧОГО ПРИСТРОЮ У НЕЙРОБАЗИСІ

Анотація

Вивчаються достатні умови зображення множин бульових векторів матрицями толерантності. На основі цього доводяться теореми, що є достатніми умовами належності бульових функцій до класу нейрофункцій.

Resume

In this paper we have investigated sufficient conditions of representing sets of the Boolean vectors with help of the tolerancy matrixes.

Для вирішення різних задач з області біології, медицини, теорії інформації і в цілому ряді інших технічних наук успішно використовуються елементи нейротехнологій. Широкого застосування набувають нейромережі і нейрофункції, за допомогою яких досліджується багато різних природних явищ. Особливе значення ці технології мають при обробці та розпізнаванні дискретних сигналів і зображень.

Визначення належності бульових функцій до класу нейрофункцій на основі дослідження властивостей їх ядра[1] є особливо актуальною задачею. Тому важливим є вивчення зображення ядер бульових нейрофункцій матрицями толерантності, що дає можливість синтезу нейроелементів[1]-[3].

Нехай $Z_2 = \{0,1\}$ і Z_2^n–n-а декартова степінь множини Z_2 .Нехай $f\colon Z_2^n \to Z_2$ - бульова функція і

$f^{-1}(1) = \{x \in Z_2^n \mid f(x) = 1\}$, $f^{-1}(0) = \{x \in Z_2^n \mid f(x) = 0\}$

Кажуть, що НЕ з вектором структури $[w = (w_1, \ldots, w_n); \omega_0]$, (w " n-вимірний дійсний вектор, що називається ваговим, ω_0 - дійсне число – поріг) реалізує бульову функцію $f(x_1,..,x_n)$, якщо для кожного $x \in Z_2^n$ виконується умова $x \in f^{-1}(0) \Leftrightarrow x \cdot w^T < \omega_0$, де T і $\cdot$ відповідно символи транспонування матриць і матричного множення.

Бульова функція , яка реалізується хоча б на одному нейронному елементі, називається нейрофункцією.

Нехай $A = \{a_1,..,a_n\}$ довільна підмножина множини Z_2^n і $A' = Z_2^n \setminus A$. Із елементів множини A побудуємо матрицю $M(A)$ наступним чином : першим рядком матриці $M(A)$ буде вектор $a_1 = (a_{11},..,a_{1n})$ із A , другим рядком матриці $M(A)$ буде вектор $a_2 = (a_{21},..,a_{2n})$ і т.д. Через S_q позначимо симетричну групу і $A_\xi = (a_{\xi(1)},..,a_{\xi(q)})$, де $\xi(i)$ - дія підстановки $\xi \in S_q$ на i.

Нехай Ω_n - множина всіх n – вимірних дійсних векторів $w = (w_1,..,w_n)$ таких, що для різних $x_1, x_2 \in Z_2^n$, числа $x \cdot w_1^T , x \cdot w_2^T$ різні.

Нехай $c_1 > c_2 > .. > c_{2^n}$ розташовані в порядку спадання зважені суми $x \cdot w^T$ при фіксованому $w \in \Omega_n$ для всіх $x \in Z_2^n$ і $c_w = (c_1, c_2,.., c_{2^n})$.

В [1] показано , якщо $w \in \Omega_n$ і R_w матриця над Z_2 розмірності $2^n \times n$, яка задовольняє умову $R_w \cdot w^T = c_w^T$, то має наступне зображення , де () матриця толерантності і , де (риска над означає операцію інвертування).

Нехай $E_n = \bigcup\limits_{w \in \Omega_n} L_w$. Матрицю N побудовану із перших r рядків матриці толерантності $L \in E_n$ називають передматрицею толерантності і пишуть $N \triangleleft L$ або $N = L(r)$.

Кажуть , що множина $A \subseteq Z_2^n$ допускає зображення матрицями толерантності із E_n , якщо існує такий елемент $\xi \in S_q (q = |A|)$ і матриця толерантності $L \in E_n$, що має місце одна із умов:

1. $M(A_\xi) \lhd L$, якщо $q \le 2^{n-1}$;

2. $M(A'_\xi) \lhd L$, якщо $q > 2^{n-1}$.

В [1] показано, що бульова функція $f(x_1,..,x_n)$ є нейрофункцією, тоді і тільки тоді, коли її ядро $K(f)$ допускає зображення матрицями толерантності із E_n.

Отже, задача по перевірці реалізованості бульової функції на одному нейроелементі зводиться до перевірки зображення ядра даної бульової функції матрицями толерантності [3].

Нехай $a = (\alpha_1,..,\alpha_j,..,\alpha_{j+s},..,\alpha_n)$ довільний вектор множини $Z_2^n, s \in \{0,1,..,n-j\}$ і $j \in \{2,..,n\}$. На множині координат $(\alpha_1,..,\alpha_n)$ вектора $a \in Z_2^n$ для фіксованих s та j визначимо функцію ε_j^k наступним чином

$$\forall k \in \{0,1,..,s\} \qquad \varepsilon_j^k(\alpha_i) = \begin{cases} \alpha_i, & \text{якщо } i \le j-1; \\ \alpha_i(j - r_k), & \text{якщо } i = j+k; \\ \alpha_i j, & \text{якщо } i > j+s; \end{cases} \tag{1}$$

де $r_0, r_1,.., r_s \in \{1,2,.., j-1\}$.

За означенням очевидно, що функція ε_j^k залежить від r_k, α_i, і j.

Через функції $\varepsilon_j^k(k \in \{0,1..,s\}$ при фіксованих $s \in \{0,..,n-j\}$ та задаємо відображення $\varepsilon_s^j : Z_2^n \to Z_j^n (2 \le j \le n)$ наступним чином

$$\varepsilon_j^s(a) = (\varepsilon_j^s(\alpha_1),..,\varepsilon_j^s(\alpha_{j-1}), \varepsilon_j^0(\alpha_j), \varepsilon_j^1(\alpha_{j+1}),.., \varepsilon_j^s(\alpha_{j+s}), \varepsilon_j^s(\alpha_{j+s+1}),.., \varepsilon_j^s(\alpha_n))$$

і визначимо функціонал на множині за формулою

$$\forall a \in Z_2^n, \qquad \upsilon_j^s(a) = \sum_{i \in I_s(j)} \varepsilon_j^s(\alpha_i) + \sum_{i=0}^{s} \varepsilon_j^i(\alpha_{j+i}) \tag{2}$$

де $I_s(j) = \{1,2,..,n\} \setminus \{j, j+1,.., j+s\}$. За допомогою функціонала υ_j^s для кожного $k \in \{0,1,..,s\}$ формуємо множину бульових векторів $F_{j+k}^{(r_k,s)}$ так: $F_{j+k}^{(r_k,s)} = \{a \in m(L_{j+k}^* 0...0) \mid \upsilon_j^s(a) \le j-1\}$,

де $m(L_{j+k}^* 0...0)$ — множина бульових векторів, що побудована з рядків матриці $(L_{j+k}^* 0...0)$, а $r_0,..., r_s \in \{1,.., j-1\}$.

Нехай $a \in Z_2^n, \sigma \in S_n$, визначимо $a^\sigma = (\alpha_{\sigma(1)},..., \alpha_{\sigma(n)})$, $A^\sigma = (a_{\sigma(1)},..., a_{\sigma(n)})$, де $\sigma(i)$ - дія підстановки σ на i, a_i–i-вий стовпець матриці A.

Теорема 1. *Якщо в множині $A \subset Z_2^n(|A| \le 2^{n-1})$, групі S_n відповідно існують такі елементи a, σ і такі цілі числа $r_0 \ge r_1 \ge ...r_s \ge 0 (r_0 \le j-1)$, що*

$$a^\sigma A^\sigma = m(L_j \underbrace{0...0}) \cup (\bigcup_{i=0}^{s} F_{j+i}^{(r_i,s)}), \tag{3}$$

тоді множина A допускає зображення матрицями толерантності з E_n.

Доведення. Визначимо $n-$ вимірний вектор $w=(\omega_1,..,\omega_n)$ наступним чином $\omega_1=...=\omega_{j-1}=-1,\ \omega_j=r_0-j,\ \omega_{j+1}=r_1-j,...,\omega_{j+s}=r_s-j,\ \omega_{j+s+1}=...=\omega_n=-j$. За побудовою вектора w маємо

$$\min\{x\cdot w^T \mid x\in m(L_j0...0)\}=1-j,$$

$$\forall k\in\{0,1,...,s\}\ \min\{x\cdot w^T \mid x\in F_{j+k}^{(r_k,k)}\}=j-1>\max\{x\cdot w^T \mid x\in m(L_{j+k}^{*}0...0)\setminus F_{j+k}^{(r_k,k)}\}=-j,$$

$$\forall t\in\{s+1,...,n\}\ \max\{x\cdot w^T \mid x\in m(L_t0...0)\}=-j\cdot$$

Тоді з (3) безпосередньо випливає

$$\forall t\in a^{\sigma}A^{\sigma},\forall y\in Z_2^n\notin a^{\sigma}A^{\sigma}\quad x\cdot w^T>y\cdot w^T \tag{4}$$

Отже, як показано в [4], існує такий вектор $v\in\Omega_n$, який також задовольняє умову (4). Це означає, що існує такий елемент $\xi\in S_q(q=|A|)$, що $a^{\sigma}(A_{\xi})^{\sigma}=L_v(q)$. Тоді з елементів множини А можна побудувати передматрицю толерантності $L_1(q)$ $(L_1=aL_v^{\sigma^{-1}})$ і теорема доведена. З теореми 1 безпосередньо випливає:

Теорема 2. *Бульова функція* $f(x_1,..,x_n)$ *з ядром* $K(f)$ *є нейрофункцією, якщо в* $K(f)$ *і групі* S_n *відповідно існують такі елементи* g,σ *і такі цілі числа* $r_0\ge r_1\ge...r_s\ge 0(r_0\le j-1)$, *що*

$$g^{\sigma}K(f)^{\sigma}=m(L_j\underbrace{0...0})\cup(\bigcup_{i=0}^{s}F_{j+i}^{(r_i,s)}).$$

Розглянемо узагальнення функцій ε_j^k і функціоналу υ_j^s. Нехай $a=(\alpha_1,...,\alpha_j,...,\alpha_{j+s},...,\alpha_n)\in Z_2^n,\ s\in\{0,1,...,n-j\}\ (j\ge 2)$. Для фіксованих $j\in\{2,3,...,n\}$ i $t\in\{1,2,j-1\}$ будуємо множину векторів $u_j(t)$:

$$u_j(t)=\{(u_1,...,u_n)\mid u_1+...+u_t=j-1,\ u_1,...,u_t\in\{1,2,...,j-t\}\}$$

і визначимо $U_j=\bigcup_{t=1}^{j-1}u_j(t)$. Нехай $u=(u_1,...,u_{l_u})\in U_j(l_u$ – розмірність вектора u) і $\forall k\in\{0,1,..,s\}$

$$\varepsilon_j^{(u,k)}(\alpha_i)=\begin{cases}\alpha_i, & i\le u_1,\\ \alpha_i 2^{r-1}, & \sum\limits_{p=1}^{r-1}u_p<i<\sum\limits_{p=1}^{r}u_p,\\ \alpha_i(\sum\limits_{p=1}^{l_u}u_p2^{p-1}-r_k+1), & i=j+k,\\ \alpha_i(\sum\limits_{p=1}^{l_u}u_p2^{p-1}+1), & j>j+s.\end{cases}$$

Для фіксованих $j\in\{2,3,..,n\},\ s\in\{0,1,...,n-j\}$ i $u=(u_1,...,u_n)\in U_j$ задаємо відображення $\varepsilon_j^{(u,s)}:Z_2^n\to Z_k^n$ $(k=\sum\limits_{p=1}^{l_u}u_p2^{p-1})$ так:

$$\varepsilon_j^{(u,s)}(a)=(\varepsilon_j^{(u,s)}(\alpha_1),...,\varepsilon_j^{(u,s)}(\alpha_{j-1}),\varepsilon_j^{(u,0)}(\alpha_j),\varepsilon_j^{(u,l)}(\alpha_{j+1}),...,$$
$$\varepsilon_j^{(u,s)}(\alpha_{j+s}),\varepsilon_j^{(u,s)}(\alpha_{j+s+1}),...,\varepsilon_j^{(u,s)}(\alpha_n))$$

і визначимо функціонал $\upsilon_j^{(u,s)}$ на множині Z_2^n за допомогою наступної формули

$$\forall a \in Z_2^n,\ \upsilon_j^{(u,s)}(a) = \sum_{i \in I_s(j)} \varepsilon_j^{(u,s)}(\alpha_i) + \sum_{i=0}^{s} \varepsilon_j^{(u,i)}(\alpha_{j+i}), \text{де } I_s(j) = \{1,2,\ldots,n\} \setminus \{j, j+1, \ldots, j+s\}.$$

Через функціонал $\upsilon_j^{(u,s)}$ задаємо множину бульових векторів $F_{j+k}^{(u,r_k,s)}$:

$$F_{j+k}^{(u,r_k,s)} = \left\{ a \in m(L_{j+k}^{*}0\ldots0) \mid \upsilon_j^{(u,s)}(a) \le \sum_{p=1}^{l_u} u_p 2^{p-1} \right\},$$

де $r_0, r_1, \ldots, r_s \in \left\{0,1,\ldots,\sum_{p=1}^{l_u} u_p 2^{p-1}\right\}$, $k \in \{1,2,\ldots,s\}$

Теорема 3. *Якщо в множині* $A \subset Z_2^n$ $(|A| \le 2^{n-1})$, *групі* S_n, *множині векторів* U_j *відповідно існують такі елементи* a, σ, u *і такі цілі числа* $r_0 \ge r_1 \ge \ldots \ge r_s \ge 0$ $(r_0 \le \sum_{p=1}^{l_u} u_p 2^{p-1})$, *що*

$$a^{\sigma} A^{\sigma} = m(L_j \underbrace{0\ldots0}) \cup (\bigcup_{i=0}^{s} F_{j+s}^{(u,r_i,s)}), \qquad (5)$$

тоді множина A допускає зображення матрицями толерантності з E_n.

Доведення. Дано, що відносно елементів $a \in A, \sigma \in S_n,\ i\ u = (u_1, \ldots, u_{l_u})$ справджується рівність (5). Покажемо, що можна побудувати вектор $w = (\omega_1, \ldots, \omega_n)$, який задовольняє умову

$$\forall x \in a^{\sigma} A^{\sigma},\ \forall y \in Z_2^n \setminus a^{\sigma} A^{\sigma} \quad . \qquad x \cdot w^T > y \cdot w^T \qquad (6)$$

Визначимо координати ω_i вектора w наступним чином:

$$\omega_1 = \ldots = \omega_{u_1} = -1,\ \omega_{u_1+1} = \ldots = \omega_{u_1+u_2} = -2,\ \omega_{u_1+u_2+1} = \ldots = \omega_{u_1+u_2+u_3} == -2^2,\ \ldots,$$

$$\omega_{u_1+u_2+\ldots+u_{l_u-1}+1} = \omega_{u_1+u_2+\ldots+u_{l_u}} = -2^{l_u-1},\ \omega_j = r_0 - (\sum_{p=1}^{l_u} u_p 2^{p-1} + 1),\ \omega_{j+1} = r_1 - (\sum_{p=1}^{l_u} u_p 2^{p-1} + 1), \ldots,$$

$$\omega_{j+s} = r_s - (\sum_{p=1}^{l_u} u_p 2^{p-1} + 1),\ \omega_{j+s+1} = \ldots = \omega_n = -\sum_{p=1}^{l_u} u_p 2^{p-1} - 1.$$

З побудови вектора випливає $\min\{x, w^T \mid x \in m(L_j 0\ldots0)\} = \sum_{p=1}^{l_u} u_p 2^{p-1}$,

$$\forall k \in \{0,1,\ldots,s\}\ \min\{x \cdot w^T \mid x \in F_{j+k}^{(u,r_k,k)}\} = -\sum_{p=1}^{l_u} u_p 2^{p-1} >$$

$$-\sum_{p=1}^{l_u} u_p 2^{p-1} - 1 = \max\{x \cdot w^T \mid x \in m(L_{j+k}^{*}0\ldots0) \setminus F_{j+k}^{(u,r_k,k)}\}$$

$$\forall t \in \{s+1,\ldots,n\}\ \max\{x \cdot w^T \mid x \in m(L_t^{*}0\ldots0)\} = -\sum_{p=1}^{l_u} u_p 2^{p-1} - 1.$$

Тоді з (5) безпосередньо випливає (6) і аналогічно тому, як в теоремі 1 можна показати існування такої матриці толерантності $L_1 \in E_n$, перші q рядки $(q = |A|)$ якої можуть бути побудовані з елементів множини А, тобто існує такий елемент $\xi \in S_q$, що $A_\xi = L_1(q)$. Отже, множина А допускає зображення матрицею толерантності $L_1 \in E_n$ і теорема доведена. З теореми 3 безпосередньо маємо:

ТТеорема 4. *Бульова функція* $f(x_1,..,x_n)$ *з ядром* $K(f)$ *є нейрофункцією, якщо в* $K(f)$ *, групі* S_n *, множині векторів* U_j *відповідно існують такі елементи* g,σ,u *і такі цілі числа*

$$r_0 \ge r_1 \ge ... \ge r_s \ge 0 \ (r_0 \le \sum_{p=1}^{l_u} u_p 2^{p-1}), \text{ що } g^\sigma K(f)^\sigma = m(L_j \underbrace{0...0}) \cup (\bigcup_{i=0}^{s} F_{j+s}^{(u,r_i,s)}).$$

Теорема 5. Бульова функція $f(x_1,..,x_n)$ з ядром $K(f)$ є нейрофункцією, якщо в $K(f)$ і групі S_n відповідно існують такі елементи g,σ і такі цілі числа $q_0 \ge q_1 \ge ...q_t > 0(q_0 \le j)$, що

$$g^\sigma K(f)^\sigma = m(L_j \underbrace{0...0}) \cup (\bigcup_{i=0}^{s} F_{j+i}^{(r_i,s)}) \text{ або } g^\sigma K(f)^\sigma = L_j \cup L_j^*(q_0) \cup L_{j+1}^*(q_1) \cup ... \cup L_{j+t}^*(q_t), \ .$$

Доведення теореми 5 випливає із теореми 1 і теореми 4.7[2].

Нехай $K_1, K_2,..., K_t$ навчаюча вибірка для класів об'єктів $K_1', K_2',..., K_t'$. Класи $K_i', K_j' (i \ne j)$ в тому числі і підмножини $K_i \subset K_i'$ і $K_j \subset K_j'$ можуть перетинатися.

Розглянемо задачу синтезу логічного пристрою з нейронних елементів (НЕ), який випадковий об'єкт v із $\bigcup_{i=1}^{t} K_i'$ віднесе до одного із класів об'єктів K_i', якщо елементи класів $K_1', K_2',..., K_t'$ закодовані бульовими векторами розмірності n, тобто:

$$K_1 = \{(\alpha_{11}^{(1)},...,\alpha_{1n}^{(1)}),(\alpha_{21}^{(1)},...,\alpha_{2n}^{(1)}),...,(\alpha_{k_11}^{(1)},...,\alpha_{k_1n}^{(1)})\},$$

....,

$$K_t = \{(\alpha_{11}^{(t)},...,\alpha_{1n}^{(t)}),(\alpha_{21}^{(t)},...,\alpha_{2n}^{(t)}),...,(\alpha_{k_t1}^{(t)},...,\alpha_{k_tn}^{(t)})\}.$$

Для вирішення поставленої задачі виберемо наступну конфігурацію трьохшарової нейромережі, де на вхід наступного шару попадають вихідні значення з попереднього:

1. Перший шар з t груп нейроелементів, де кожна група складається з певної кількості елементів, у кожного з яких є n входів, на які поступає код $(x_1,...,x_n)$ випадкового об'єкту v із $\bigcup_{i=1}^{t} K_i'$.

2. Другий шар з t нейроелементів з векторами структури $[(1,1,...,1);1]$, які відповідно формують значення $f_1,..., f_t$, бульових нейрофункцій.

1. Третій шар складається з одного нейроелементу і формує на виході значення $f_1 + f_2 2^1 + ... + f_t 2^{t-1}$.

Причому, нейроелементи першого і другого шару мають порогову функцію активації і якщо на вхід подати бульовий набір $\bar{a} = (\alpha_1,...,\alpha_n) \in K_i'$, то значення вихідного сигналу і-го блоку буде 1, тобто $f_i = 1$.

Максимальну підмножину $p(aA)$ множини aA [2], що задовільняє умову

Максимальну підмножину $p(aA)$ множини aA [2], що задовільняє умову

$$p(aA)_\xi^\sigma = (L_j 0_j ... 0_j) \subset (\bigcirc_{i=0}^{n-j} (L_{j+1}^*(q) \underbrace{0...0}_{n-(j+i)}))$$

назвемо p - множиною A відносно $\bar{a}$ з індексом j.

Будь-яку множину A_s [2] бульових наборів можна записати через p - множини так:

$$A_s = a_1^s p(a_1^s A_s) \cup a_2^s p(a_2^s A_s) \cup \ldots \cup a_{r_s}^s p(a_{r_s}^s A_s) \quad (7)$$

Рівність (7) задає p -розклад множини A_s відносно точок розкладу $a_1^s, \ldots, a_{r_s}^s$.

Для кожної p -множини $a_k^s p(a_k^s A_s)$ в [2] розроблено ефективний алгоритм знаходження такого вектора $w_{ks} = (w_1, \ldots, w_{j_k}, w_{j_{k+1}}, \ldots, w_n)$, що задовільнятиме умову $\forall x \in a_k^s p(a_k^s A_s), \forall y \in Z_2^n \setminus a_k^s p(a_k^s A_s)^{\sigma_k}$ $x \cdot w_{ks}^T > y \cdot w_{ks}^T$. Вектор w_{ks}^T буде ваговим вектором НЕ, що реалізує характеристичну функцію множини $a_k^s p(a_k^s A_s)$.

Алгоритм синтезу розпізнаючого пристрою

Крок 1. Нехай $K_1, K_2, \ldots, K_t$ навчаюча вибірка, $s = 1$, переходимо до кроку 2.

Крок 2. Побудуємо множину $A_s = K_s \cup \left(Z_2^n \setminus \bigcup_{i=1,\, i \neq s}^{t} K_i \right)$

Знаходимо p - розклад множини K_i в A_i і для кожної p – множини знаходимо ваговий вектор w_{ks} і поріг τ_{ks} відповідного НЕ. Система векторів $\{[w_{1s}; \tau_{1s}], \ldots, [w_{rs}; \tau_{rs}]\}$ задає вектори структури нейроелементів 1-го шару блока s .

Крок 3. Якщо $s < t$, то $s = s + 1$ і переходимо до кроку 2, а в протилежному випадку синтез мережі завершено.

Клас K_j , до якого відноситься об'єкт $v \in \bigcup_{i=1}^{t} K_i'$ визначається по значенню виразу $f_1 + f_2 2^1 + \ldots + f_t 2^{t-1}$.

Розроблений алгоритм розпізнаючого пристрою в нейробазисі працює у реальному часі і може бути застосовано для розпізнавання будь-яких класів об'єктів, закодованих бульовими векторами.

Список використаних джерел:

1. *Н.Н.Айзенберг, А.А. Бовди, Э.Й.Герго, Ф.Э.Гече. Некоторые алгебраические аспекты пороговой логики/ /Кибернетика.-К.,1980.-№2.-с.26-30.*
2. *Параллельная обработка информации//Проблемно-ориентированые средства обработки информации в 5-ти т./Под ред. Б.Н.Малиновского. – К.:Наук. думка,1990 – т.5.–502с.*
3. *Гече Ф.Э., Поливко В.П., Роботишин В.А. Реализация функций алгебры и логики на ПЭ.//Кибернетика –К.,1983,–№6.–с.62-67.*
4. *Яджима С., Ибараки Т. Нижняя оценка числа пороговых функций.//Кибернетический сборник, новая серия, вып. 6. М.:Мир, 1969 – с.71-81.*