УДК 681.5+519.7

Ф.Е. ГЕЧЕ, О.Ю. МУЛЕСА

АЛГЕБРАЇЧНІ ВЛАСТИВОСТІ ЯДЕР УЗАГАЛЬНЕНИХ НЕЙРОФУНКЦІЙ

Анотація. Розглянуто узагальнені нейронні елементи, встановлено умови реалізовності функцій алгебри логіки на цих елементах. Уведено поняття модифікованого ядра булевих функцій відносно системи характерів групи, на якій задаються функції алгебри логіки, наведено критерії належності цих функцій до класу узагальнених нейрофункцій. Досліджено алгебраїчну структуру ядер булевих нейрофункцій, на основі властивостей матриць толерантності отримано низку необхідних умов реалізовності булевих функцій одним узагальненим нейронним елементом. Одержані результати дають змогу розробити ефективні методи синтезу цілочислових узагальнених нейронних елементів з великим числом входів, які можна успішно застосовувати в задачах компресії і передавання інформації, а також у задачах розпізнавання дискретних сигналів.

Ключові слова: узагальнений нейронний елемент, ядро функції, спектр функції, характер групи, синтез, матриця толерантності.

ВСТУП

Останніми роками спостерігається активізація теоретичних і практичних розроблень у галузі нейрокомп'ютерів і зростання інтересу до нейроподібних структур, які знайшли широке застосування в задачах розпізнавання образів, компресії дискретних сигналів та зображень, прогнозуванні. У багатьох випадках ці задачі формулюють як моделі комбінаторної оптимізації [1, 2], а їхня поява пов'язана з потребою у розв'язанні проблем, що виникають у бізнесі, медицині, техніці.

Широке застосування нейромереж для розв'язування прикладних задач стане можливим, якщо будуть розроблені практично придатні методи синтезу нейронних елементів з різними функціями активації та синтезу логічних схем із них.

У задачах розпізнавання образів, ущільнення та передачі дискретних сигналів треба вміти здійснювати синтез надійних (цілочислових) нейронних елементів з великою кількістю входів. Класичні методи апроксимації та різні ітераційні методи синтезу нейронних елементів не можуть бути застосовані на практиці для знаходження векторів структур нейроелементів з метою реалізації дискретних функцій з великою кількістю аргументів.

Штучні нейронні мережі та нейроподібні структури ефективно використовують для класифікації і розпізнавання зображень [3], а також для покращення їхньої якості [4]. На основі зазначених мереж і структур розробляють інтелектуальні блоки різних систем для керування хімічними процесами [5], прогнозування економічних [6, 7] та біологічних [8] процесів. Як показують дослідження, нейромережеві методи успішно застосовують для ущільнення сигналів та зображень [9–11], у банківській сфері для оцінки кредитного ризику [12].

Вибираючи математичні моделі нейронних елементів для побудови нейроподібних структур, потрібно зважати на функціональну можливість цих елементів, яка має суттєве значення для оптимізації кількості елементів у відповідних логічних структурах. У роботі [13] розглянуто нейронні елементи з узагальненою пороговою функцією активації. Також за допомогою характеристичних векторів функцій алгебри логіки, побудованих з певних спектральних коефіцієнтів цих функцій у базисі Уолша–Адамара [14], отримано критерії реалізовності булевих функцій одним нейронним елементом з узагальненою пороговою функцією активації. Ці критерії встановлюють спектральні властивості булевих нейрофункцій і слугують основою різних ітераційних та апроксимаційних методів синтезу узагальнених нейронних елементів. Слід зазначити, що ці методи синтезу узагальнених нейронних елементів можна застосувати на практиці у тих випадках, коли кількість входів цих елементів є невеликим числом і не перевищує кількох десятків. Однак, з практичної точки зору, для успішного використання нейронних елементів у задачах розпізнавання дискретних сигналів та зображень, задачах кодування, ущільнення та передачі дискретних сигналів потрібно вміти здійснювати синтез (проводити навчання) нейронних елементів з великою кількістю входів (порядку кількох сотень, а, можливо, й кількох тисяч входів). У зв'язку з цим встановлення нових властивостей узагальнених булевих нейрофункцій і розроблення ефективних методів синтезу узагальнених нейронних елементів з великим числом входів та використання цих логічних елементів у різних нейроподібних структурах є актуальною і практично важливою задачею.

У цій роботі введено поняття модифікованого ядра булевих функцій. З використанням цього поняття отримано критерії реалізовності булевих функцій одним узагальненим нейронним елементом. Наведено ефективні необхідні умови перевірки належності функцій алгебри логіки до класу узагальнених нейрофункцій.

ПОБУДОВА ЯДЕР УЗАГАЛЬНЕНИХ НЕЙРОФУНКЦІЙ ТА ЇХНІ ОСНОВНІ ВЛАСТИВОСТІ

Нехай $H_2 = \{-1, 1\}$ — циклічна група 2-го порядку, $G_n = H_2 \otimes \ldots \otimes H_2$ — прямий добуток n циклічних груп H_2, $\chi(G_n)$ — група характерів [15] групи G_n над полем дійсних чисел R. На множині $R \setminus \{0\}$ визначимо функцію

$$\operatorname{Rsign} x = \begin{cases} 1, & \text{якщо } x > 0, \\ -1, & \text{якщо } x < 0. \end{cases} \qquad (1)$$

Нехай $Z_2 = \{0, 1\}$, $i \in \{0, 1, 2, \ldots, 2^n - 1\}$; $(i_1, \ldots, i_n)$ — двійковий код числа i, тобто $i = i_1 2^{n-1} + i_2 2^{n-2} + \ldots + i_n$, $i_j \in \{0, 1\}$. Значення характеру χ_i на елементі $\mathbf{g} = ((-1)^{\alpha_1}, \ldots, (-1)^{\alpha_n}) \in G_n$ $((\alpha_1, \ldots, \alpha_n) \in Z_2^n$ — n-й декартовий степінь $Z_2 = \{0, 1\})$ визначається так:

$$\chi_i(\mathbf{g}) = (-1)^{\alpha_1 i_1 + \alpha_2 i_2 + \ldots + \alpha_n i_n}. \qquad (2)$$

З ортогональності характерів [15] випливає, що група характерів $X(G_n)$ утворює ортогональний базис векторного простору $V_R = \{\phi \mid \phi : G_n \to R\}$. Оскільки булева функція від n змінних в алфавіті $\{-1, 1\}$ задає однозначне відображення $f : G_n \to H_2$, то $f \in V_R$. Це означає, що довільну булеву функцію f однозначно можна записати у вигляді

$$f(\mathbf{g}) = s_0 \chi_0(\mathbf{g}) + s_1 \chi_1(\mathbf{g}) + \ldots + s_{2^n-1} \chi_{2^n-1}(\mathbf{g}). \qquad (3)$$

Вектор $\mathbf{s}_f = (s_0, s_1, \ldots, s_{2^n-1})$ називається спектром булевої функції f у системі характерів $\chi(G_n)$ (у системі базисних функцій Уолша–Адамара [14]).

З різних характерів $\chi(G_n)$, крім головного, побудуємо m-елементну множину $\chi = \{\chi_{i_1}, \ldots, \chi_{i_m}\}$. Відносно обраної системи характерів згідно з [16] розглянемо математичну модель нейронного елемента з узагальненою пороговою функцією активації (узагальненого нейронного елемента)

$$f(x_1(\mathbf{g}), \ldots, x_n(\mathbf{g})) = \operatorname{Rsign}\left(\sum_{j=1}^{m} \omega_j \chi_{i_j}(\mathbf{g}) + \omega_0 \right). \qquad (4)$$

Вектор $\mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0)$ називається вектором структури узагальненого нейронного елемента (УНЕ) відносно системи характерів $\chi \subset \chi(G_n)$ і $\mathbf{g} \in G_n$.

Нехай $\mathrm{w}(\mathbf{g}) = \omega_1 \chi_{i_1}(\mathbf{g}) + \ldots + \omega_m \chi_{i_m}(\mathbf{g}) + \omega_0$. Якщо $\mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0)$ є вектором структури УНЕ відносно системи характерів $\chi = \{\chi_{i_1}, \ldots \chi_{i_m}\}$ групи $\chi(G_n)$, що реалізує булеву функцію $f: G_n \to H_2$, то з (1) і (4) безпосередньо випливає, що

$$\forall \mathbf{g} \in G_n \quad \mathrm{w}(\mathbf{g}) \neq 0. \tag{5}$$

Далі будемо розглядати тільки такі нейронні елементи, вектори структури яких задовольняють умову (5). Множину всіх таких $m+1$-вимірних дійсних векторів, що задовольняють умову (5) відносно системи χ, позначимо $\mathrm{W}_{m+1}(\chi) = \mathrm{W}_{m+1}(\chi_{i_1}, \ldots \chi_{i_m})$.

Означення 1. Булева функція $f: G_n \to H_2$ називається узагальненою нейрофункцією відносно системи характерів $\chi = \{\chi_{i_1}, \ldots, \chi_{i_m}\} \subset \chi(G_n)$, якщо існує такий вектор $\mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0) \in \mathrm{W}_{m+1}(\chi)$, що має місце рівність (4).

Для введення поняття ядер узагальнених нейрофункцій і дослідження їхніх основних властивостей розглянемо булеві функції як в алфавіті $H_2 = \{-1, 1\}$, так і в алфавіті $Z_2 = \{0, 1\}$.

Нехай $f(x_1, \ldots, x_n)$ — булева функція в алфавіті $\{-1, 1\}$, тобто $f: G_n \to H_2$. Розглянемо задачу про реалізовність булевої функції $f(x_1, \ldots, x_n)$ одним УНЕ відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ в алфавіті $\{0, 1\}$. За допомогою перетворення $\mathbf{x}' = \frac{1}{2}(\mathbf{x} + 1)$ реалізуємо відображення $\{-1, 1\} \to \{0, 1\}$ і побудуємо систему $\chi' = \left\{ \chi'_{i_1} = \frac{1}{2}(\chi_{i_1} + 1), \chi'_{i_2} = \frac{1}{2}(\chi_{i_2} + 1), \ldots, \chi'_{i_m} = \frac{1}{2}(\chi_{i_m} + 1) \right\}$.

Нехай $f^{-1}(-1) = \{\mathbf{g} \in G_n \mid f(\mathbf{g}) = -1\}$ і $f^{-1}(1) = \{\mathbf{g} \in G_n \mid f(\mathbf{g}) = 1\}$. За допомогою системи χ' визначимо

$$f_\chi^{-1}(0) = \bigcup_{\mathbf{g} \in f^{-1}(-1)} \{(\chi'_{i_1}(\mathbf{g}), \ldots, \chi'_{i_m}(\mathbf{g}))\},$$

$$f_\chi^{-1}(1) = \bigcup_{\mathbf{g} \in f^{-1}(1)} \{(\chi'_{i_1}(\mathbf{g}), \ldots, \chi'_{i_m}(\mathbf{g}))\}.$$

Означення 2. Ядром булевої функції $f: G_n \to H_2$ відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ в алфавіті $\{0, 1\}$ називається множина $K(f_\chi)$, яка визначається так:

$$K(f_\chi) = \begin{cases} f_\chi^{-1}(1), & \text{якщо } |f_\chi^{-1}(1)| \leq |f_\chi^{-1}(0)|, \\ f_\chi^{-1}(0), & \text{якщо } |f_\chi^{-1}(1)| > |f_\chi^{-1}(0)| \end{cases}$$

за умови, що $f_\chi^{-1}(1) \cap f_\chi^{-1}(0) = \varnothing$, де $|f_\chi^{-1}(i)|$ — кількість елементів множини $f_\chi^{-1}(i)$ $(i \in \{0, 1\})$.

Якщо $f_\chi^{-1}(1) \cap f_\chi^{-1}(0) \neq \varnothing$, тоді не існує ядра $K(f_\chi)$ і це означає, що функція f не реалізується одним УНЕ відносно системи χ.

Нехай Z_2^m — m-й декартовий степінь множини $Z_2 = \{0, 1\}$. Припустимо, що функція $f: G_n \to H_2$ має ядро $K(f_\chi)$ відносно системи $\chi = \{\chi_{i_1}, \ldots, \chi_{i_m}\}$, тобто $f_\chi^{-1}(1) \cap f_\chi^{-1}(0) = \varnothing$. Множина Z_2^m задовольняє одну з таких умов:

1) $Z_2^m = f_\chi^{-1}(1) \cup f_\chi^{-1}(0)$ — функція f_χ повністю визначена на множині Z_2^m;

2) $Z_2^m \neq f_\chi^{-1}(1) \cup f_\chi^{-1}(0)$ — функція f_χ частково визначена на множині Z_2^m.

У першому випадку ядро $K(f_\chi)$ визначається однозначно. Покладемо $K(f_\chi, \varnothing) = K(f_\chi)$.

У другому випадку введемо поняття розширеного ядра відносно системи характерів χ.

Нехай $K(f_\chi) = \{\mathbf{a}_1, \ldots, \mathbf{a}_q\}$ — ядро булевої функції f_χ відносно системи $\chi = \{\chi_{i_1}, \ldots, \chi_{i_m}\}$ і $f_\chi^{-1}(\cdot)$ — множина тих наборів із Z_2^m, на яких функція не визначена. Тоді під розширеним ядром функції f_χ відносно системи χ розуміють $K(f_\chi, A) = \{\mathbf{a}_1, \ldots, \mathbf{a}_q, \mathbf{a}_{q+1}, \ldots, \mathbf{a}_{q+s}\}$, де $\mathbf{a}_{q+1}, \ldots, \mathbf{a}_{q+s}$ — довільні різні елементи множини з $f_\chi^{-1}(\cdot)$, $q + s \leq 2^{m-1}$ і $A = \{\mathbf{a}_{q+1}, \ldots, \mathbf{a}_{q+s}\}$. Зазначимо, що множина A може бути порожньою.

Введемо поняття модифікованого ядра $K(f_\chi, M)$ булевої функції $f : G_n \to H_2$ відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ таким чином:

$$K(f_\chi, M) = \begin{cases} K(f_\chi), & \text{якщо } A = \varnothing, \\ K(f_\chi, A), & \text{якщо } A \neq \varnothing. \end{cases}$$

Означення 3. Булева функція $f : G_n \to H_2$ реалізується одним узагальненим нейронним елементом з вектором структури $\mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0^*)$ над полем дійсних чисел R відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ в алфавіті $\{0, 1\}$, якщо існує таке модифіковане ядро $K(f_\chi, M)$, що

$$\mathbf{a} \in K(f_\chi, M) \Leftrightarrow \begin{cases} \omega_1 \chi'_{i_1}(\mathbf{a}) + \ldots + \omega_m \chi'_{i_m}(\mathbf{a}) > \omega_0^*, & \text{якщо } K(f_\chi) = f_\chi^{-1}(1), \\ \omega_1 \chi'_{i_1}(\mathbf{a}) + \ldots + \omega_m \chi'_{i_m}(\mathbf{a}) < \omega_0^*, & \text{якщо } K(f_\chi) = f_\chi^{-1}(0). \end{cases}$$

Враховуючи перетворення $\chi' = \frac{1}{2}(\chi + 1)$ $(\{-1, 1\} \to \{0, 1\})$ або $\chi = 2\chi' - 1$ — перехід від $\{0, 1\} \to \{-1, 1\}$, можна встановити такий зв'язок між координатами векторів структур $\mathbf{w} = (\omega_1, \ldots, \omega_m; \omega_0) \in \mathrm{W}_{m+1}(\chi)$ (в алфавіті $\{-1, 1\}$) і $\mathbf{w}^* = (\omega_1^*, \ldots$ $\ldots, \omega_m^*; \omega_0^*) \in \mathrm{W}_{m+1}(\chi)$ (в алфавіті $\{0, 1\}$) узагальнених нейронних елементів, які реалізують одну й ту саму булеву функцію в різних алфавітах:

$$\omega_i^* = \omega_i \quad (i = 1, 2, \ldots, m) \text{ і } \omega_0^* = \frac{1}{2}\left(\sum_{i=1}^{m} \omega_i - \omega_0\right). \tag{6}$$

Нехай задано функцію $f : G_n \to H_2$, систему $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ і модифіковане ядро $K(f_\chi, M)$. Побудуємо для функції f відносно системи χ та модифікованого ядра $K(f_\chi, M)$ булеву функцію $f_\chi^* : Z_2^m \to Z_2$ таким чином:

$$\forall \mathbf{a} \in K(f_\chi, M) \;\; f_\chi^*(\mathbf{a}) = f_\chi(\mathbf{a}),$$

$$\forall \mathbf{a} \in Z_2^m \setminus K(f_\chi, M) \;\; f_\chi^*(\mathbf{a}) = \bar{f}_\chi(\mathbf{a}),$$

де риска над символом означає логічну операцію заперечення.

За визначенням функції f_χ^* і в силу $K(f_\chi) \subset K(f_\chi, M)$ маємо

$$f_\chi^{-1}(\alpha) \subset f_\chi^{*\,-1}(\alpha), \tag{7}$$

де $\alpha \in \{0, 1\}$.

Під ядром функції f_χ^* будемо розуміти модифіковане ядро функції f_χ, тобто

$$K(f_\chi^*) = K(f_\chi, M). \tag{8}$$

Нехай E_m — множина матриць толерантності [17] і $K(f_\chi^*) = \{\mathbf{a}_1, \ldots, \mathbf{a}_t\}$ — ядро булевої функції $f: G_n \to H_2$ ($f \neq \text{const}$) відносно системи характерів $\chi = \{\chi_{i_1}, \ldots, \chi_{i_m}\} \subset \chi(G_n)$. З елементів ядра $K(f_\chi^*)$ побудуємо матрицю $K_\xi(f_\chi^*)$ таким чином: першим рядком матриці $K_\xi(f_\chi^*)$ буде вектор $\mathbf{a}_{\xi(1)} = (\alpha_{\xi(1)1}, \ldots, \alpha_{\xi(1)m})$ з $K(f_\chi^*)$, другим рядком матриці буде вектор $\mathbf{a}_{\xi(2)} = (\alpha_{\xi(2)1}, \ldots, \alpha_{\xi(2)m})$, останнім рядком $K_\xi(f_\chi^*)$ буде $\mathbf{a}_{\xi(t)} = (\alpha_{\xi(t)1}, \ldots, \alpha_{\xi(t)m})$, де $\xi(i)$ — дія підстановки $\xi \in S_t$ (S_t — симетрична група степеня t) на i. Позначимо перші r рядків матриці $L \in E_m$ як $L(r)$, і якщо $r < 2^{m-1}$, то матрицю $L(r)$ будемо називати передматрицею матриці толерантності L. Введемо поняття зображення ядра $K(f_\chi^*)$ матрицями толерантності з E_m таким чином: ядро $K(f_\chi^*)$ допускає зображення матрицями толерантності з E_m, якщо існує такий елемент $\xi \in S_t$ і така матриця $L \in E_m$, що $K_\xi(f_\chi^*) = L(t)$.

Нехай Ω_m — множина всіх таких m-вимірних дійсних векторів, що $(\mathbf{x}_1, \mathbf{w}) \neq (\mathbf{x}_2, \mathbf{w})$, $(\mathbf{x}_1 \neq \mathbf{x}_2)$ і $\mathbf{x}_1, \mathbf{x}_2 \in Z_2^m$.

Визначимо опуклу лінійну оболонку $\text{conv}\, K(f_\chi^*)$ елементів ядра $K(f_\chi^*)$ так:

$$\text{conv}\, K(f_\chi^*) =$$
$$= \left\{ \mathbf{x} \in [0,1]^m \mid \mathbf{x} = \sum_{i=1}^{t} \lambda_i \mathbf{a}_i, \sum_{i=1}^{t} \lambda_i = 1, \lambda_1 \geq 0, \ldots, \lambda_t \geq 0; \mathbf{a}_1, \ldots, \mathbf{a}_t \in K(f_\chi^*) \right\}.$$

Аналогічно визначимо $\text{conv}\, K(f_\chi^*)^*$ для $K(f_\chi^*)^* = Z_2^m \setminus K(f_\chi^*)$.

Теорема 1. Булева функція $f: G_n \to H_2$ ($f \neq \text{const}$) реалізується одним узагальненим нейронним елементом відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots$ $\ldots, \chi_{i_m}\} \subset X(G_n)$ тоді й тільки тоді, коли існує таке модифіковане ядро $K(f_\chi, M)$, що $\text{conv}\, K(f_\chi^*) \cap \text{conv}\, K(f_\chi^*)^* = \varnothing$.

Необхідність доведемо від супротивного. Припустимо, що $\text{conv}\, K(f_\chi^*) \cap$ $\cap\, \text{conv}\, K(f_\chi^*)^* \neq \varnothing$, $K(f_\chi) = f_\chi^{-1}(1)$ і функція f_χ^* реалізується одним узагальненим нейронним елементом відносно системи χ. Тоді, як відомо [17], існує такий елемент ξ симетричної групи S_t (t — кількість елементів ядра $K(f_\chi^*)$) і матриця толерантності $L = L_\mathbf{w} \in E_m$, що $K_\xi(f_\chi^*) = L_\mathbf{w}(t)$.

З рівності $K_\xi(f_\chi^*) = L_\mathbf{w}(t)$ випливає, що для всіх $\mathbf{a} \in K(f_\chi^*)$ і для всіх $\mathbf{b} \in K(f_\chi^*)^*$

$$(\mathbf{a}, \mathbf{w}) > (\mathbf{b}, \mathbf{w}). \tag{9}$$

Нехай $\mathbf{d} \in \text{conv}\, K(f_\chi^*) \cap \text{conv}\, K(f_\chi^*)^*$. Тоді

$$\mathbf{d} = \sum_{i=1}^{t} \lambda_i \mathbf{a}_i, \ \sum_{i=1}^{t} \lambda_i = 1, \ \lambda_1, \ldots \lambda_t \geq 0, \ \mathbf{a}_1, \ldots, \mathbf{a}_t \in K(f_\chi^*), \tag{10}$$

$$\mathbf{d} = \sum_{i=1}^{k} \lambda'_i \mathbf{b}_i, \ \sum_{i=1}^{k} \lambda'_i = 1, \ \lambda'_1, \ldots \lambda'_k \geq 0, \ \mathbf{b}_1, \ldots, \mathbf{b}_k \in K(f_\chi^*)^*. \tag{11}$$

Нехай $\omega_{\min} = \min\{(\mathbf{a}_i, \mathbf{w}) \mid \mathbf{a}_i \in K(f_\chi^*)\}$ і $\omega'_{\max} = \max\{(\mathbf{b}_j, \mathbf{w}) \mid \mathbf{b}_j \in K(f_\chi^*)^*\}$. На основі (9)–(11) маємо:

$$(\mathbf{d}, \mathbf{w}) = \sum_{i=1}^{t} \lambda_i (\mathbf{a}_i, \mathbf{w}) \geq \left(\sum_{i=1}^{t} \lambda_i \right) \omega_{\min} > \omega'_{\max} = \left(\sum_{j=1}^{k} \lambda'_j \right) \omega'_{\max} \geq \sum_{j=1}^{k} \lambda'_j (\mathbf{b}_j, \mathbf{w}) = (\mathbf{d}, \mathbf{w}).$$

З отриманої нерівності $(\mathbf{d}, \mathbf{w}) > (\mathbf{d}, \mathbf{w})$ випливає, що у разі $\operatorname{conv} K(f_\chi^*) \cap$ $\cap \operatorname{conv} K(f_\chi^*)^* \neq \varnothing$ і $K(f_\chi) = f_\chi^{-1}(1)$ не існує такого модифікованого ядра функції f_χ, для якого функція f_χ^* реалізується одним узагальненим нейронним елементом відносно системи χ. Тоді з урахуванням (7) функція f_χ не реалізується одним узагальненим нейронним елементом відносно χ. Отримана суперечність доводить необхідність теореми для умови $K(f_\chi) = f_\chi^{-1}(1)$. Якщо $K(f_\chi) = f_\chi^{-1}(0)$, тоді маємо $(\mathbf{d}, \mathbf{w}) < (\mathbf{d}, \mathbf{w})$ і аналогічно попередньому випадку одержуємо суперечність, що доводить необхідність.

Достатність. Дано, що існує таке модифіковане ядро $K(f_\chi, M)$, для якого має місце рівність $\operatorname{conv} K(f_\chi^*) \cap \operatorname{conv} K(f_\chi^*)^* = \varnothing$. Покажемо, що функція f_χ^* реалізується одним узагальненим нейронним елементом відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$. За допомогою опуклих оболонок $\operatorname{conv} K(f_\chi^*)$ і $\operatorname{conv} K(f_\chi^*)^*$ побудуємо множину

$$D = \{\mathbf{d} = \mathbf{a} - \mathbf{b} \mid \mathbf{a} \in \operatorname{conv} K(f_\chi^*),\ \mathbf{b} \in \operatorname{conv} K(f_\chi^*)^*\},$$

яка, очевидно, є опуклою і не містить нульового вектора $\mathbf{0}$, оскільки $\operatorname{conv} K(f_\chi^*) \cap \cap \operatorname{conv} K(f_\chi^*)^* = \varnothing$. Опуклі лінійні оболонки $\operatorname{conv} K(f_\chi^*)$ і $\operatorname{conv} K(f_\chi^*)^*$ є компактними [18], отже, множина D також компактна і тому замкнена. Тоді на основі теореми про відокремлення [19] можна стверджувати, що для D у m-вимірному евклідовому просторі R^m існує така гіперплощина $\pi = \{\mathbf{x} \in R^m \mid (\mathbf{p}, \mathbf{x}) = p_0\}$ $(\mathbf{p} \neq \mathbf{0})$, $p_0 \in R$, яка задовольняє умови

$$p_0 = (\mathbf{p}, \mathbf{0}) = 0 \tag{12}$$

і для всіх $\mathbf{d} \in D$ маємо

$$(\mathbf{p}, \mathbf{d}) > p_0. \tag{13}$$

Враховуючи, що $\mathbf{d} = \mathbf{a} - \mathbf{b}$ $(\mathbf{a} \in \operatorname{conv} K(f_\chi^*),\ \mathbf{b} \in \operatorname{conv} K(f_\chi^*)^*)$, з останньої нерівності отримуємо

$$(\mathbf{p}, \mathbf{a}) > (\mathbf{p}, \mathbf{b}) \tag{14}$$

для будь-якого $\mathbf{a} \in \operatorname{conv} K(f_\chi^*)$ і для будь-якого $\mathbf{b} \in \operatorname{conv} K(f_\chi^*)^*$. Отже, нерівність (14) має місце для всіх $\mathbf{a} \in K(f_\chi^*)$ і для всіх $\mathbf{b} \in K(f_\chi^*)^*$. Тоді, як показано в [20], існує такий вектор $\mathbf{w} \in \Omega_m$, який задовольняє (14). Це означає, що з елементів ядра $K(f_\chi^*)$ можна побудувати таку матрицю $K_\xi(f_\chi^*)$, що $K_\xi(f_\chi^*) = L_{\mathbf{w}}(t)$ $(L_{\mathbf{w}} \in E_m)$. Отже, функція f_χ^* або $\bar{f}_\chi^*$ (залежно від того, чи $K(f_\chi) = f_\chi^{-1}(1)$ чи $K(f_\chi) = f_\chi^{-1}(0)$) реалізується одним узагальненим нейронним елементом відносно системи χ. Функції f_χ^* і $\bar{f}_\chi^*$ одночасно реалізуються або не реалізуються одним УНЕ відносно системи χ. Тоді за означенням 3 і співвідношенням (7) можна стверджувати, що функція f_χ реалізується одним узагальненим нейронним елементом відносно системи $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots$ $\ldots, \chi_{i_m}\} \subset X(G_n)$. Достатність доведено. ■

Зауваження 1. Якщо функція $f = \text{const}$, то вона завжди реалізується одним УНЕ відносно системи $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$. Дійсно, згідно з (4) функція $f = 1$ реалізується одним УНЕ відносно системи χ з вектором структури $\mathbf{w} = (0, \ldots, 0; \omega_0)$, якщо $\omega_0 > 0$, а функція $f = 0$ — відносно системи χ з вектором структури $\mathbf{w} = (0, \ldots, 0; \omega_0)$, коли $\omega_0 < 0$.

Нехай $\mathbf{a}, \mathbf{b}$ — довільні елементи ядра $K(f_\chi)$ $(\mathbf{a} \neq \mathbf{b})$ булевої функції $f: G_n \to H_2$ відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ і $O(\mathbf{a}, \mathbf{b})$ — множина таких орт-векторів $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_s}$, що $\mathbf{a} \oplus \mathbf{b} = \mathbf{e}_{i_1} + \mathbf{e}_{i_2} + \ldots + \mathbf{e}_{i_s}$, де $\oplus$ — покоординатна сума векторів за модулем 2, $i_r \neq i_k$, якщо $r \neq k$. Позначимо $H(\mathbf{a}, \mathbf{b})$ підгрупу групи Z_2^m (Z_2^m утворює групу відносно операції $\oplus$), яка породжується елементами $O(\mathbf{a}, \mathbf{b})$, тобто $H(\mathbf{a}, \mathbf{b}) = \langle \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_s} \mid \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_s} \in O(\mathbf{a}, \mathbf{b}) \rangle$.

Нехай $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$, $\mathbf{b} = (\beta_1, \ldots, \beta_m) \in Z_2^m$. Покоординатну кон'юнкцію векторів $\mathbf{a}$ та $\mathbf{b}$ позначимо $\mathbf{a} \,\&\, \mathbf{b} = (\alpha_1 \& \beta_1, \ldots, \alpha_m \& \beta_m)$ і $H(\mathbf{a} \,\&\, \mathbf{b})$ позначимо суміжний клас групи Z_2^m за підгрупою $H(\mathbf{a}, \mathbf{b})$, що визначається елементом $\mathbf{a} \,\&\, \mathbf{b}$, тобто $H(\mathbf{a} \,\&\, \mathbf{b}) = \mathbf{a} \,\&\, \mathbf{b} \oplus H(\mathbf{a}, \mathbf{b})$.

Задамо метрику $\rho(\mathbf{a}, \mathbf{b})$ на Z_2^m таким чином:

$$\rho(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{m} (\alpha_i \oplus \beta_i),$$

де $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$, $\mathbf{b} = (\beta_1, \ldots, \beta_m) \in Z_2^m$.

Теорема 2. Якщо булева функція $f: G_n \to H_2$ ($f \neq \text{const}$) реалізується одним узагальненим нейронним елементом відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ з модифікованим ядром $K(f_\chi, M)$, тоді для будь-яких двох різних елементів $\mathbf{a}, \mathbf{b}$ з $K(f_\chi^*)$, для яких $|H(\mathbf{a} \,\&\, \mathbf{b}) \cap K(f_\chi^*)^*| \geq 2$, і для будь-яких двох різних елементів $\mathbf{g}, \mathbf{h}$ з $H(\mathbf{a} \,\&\, \mathbf{b}) \cap K(f_\chi^*)^*$ справджується нерівність $\rho(\mathbf{g}, \mathbf{h}) < \rho(\mathbf{a}, \mathbf{b})$.

Доведення. Нехай $\mathbf{a} = (\alpha_1, \ldots, \alpha_{\mathrm{m}})$, $\mathbf{b} = (\beta_1, \ldots, \beta_m)$ — довільні різні елементи з $K(f_\chi^*)$ $(\mathbf{a} \neq \mathbf{b})$, $\mathbf{g} = (\gamma_1, \ldots, \gamma_m)$, $\mathbf{h} = (\delta_1, \ldots, \delta_m)$ — довільні різні елементи з $H(\mathbf{a} \,\&\, \mathbf{b}) \cap K(f_\chi^*)^*$ і $\rho = \rho(\mathbf{a}, \mathbf{b})$. Не обмежуючи загальності міркувань, будемо вважати, що перші ρ координати векторів $\mathbf{a}$ і $\mathbf{b}$ різні, а інші — однакові, тобто $\alpha_i \neq \beta_i$ для $i = 1, 2, \ldots, \rho$ і $\alpha_i = \beta_i$, $i = \rho + 1, \ldots, m$. З теореми 1 та з того, що функція f реалізується одним УНЕ відносно системи χ, випливає

$$\operatorname{conv} K(f_\chi^*) \cap \operatorname{conv} K(f_\chi^*)^* = \varnothing.$$

Отже,

$$\lambda_1 \mathbf{a} + (1 - \lambda_1) \mathbf{b} \neq \lambda_2 \mathbf{g} + (1 - \lambda_2) \mathbf{h} \tag{15}$$

для всіх $\lambda_1, \lambda_2 \in [0, 1]$.

Враховуючи, що точки $\mathbf{a}, \mathbf{b}$ $(\mathbf{a} \neq \mathbf{b})$, $\mathbf{g}, \mathbf{h}$ $(\mathbf{g} \neq \mathbf{h})$ є кутовими точками відповідних множин $\operatorname{conv} K(f_\chi^*)$, $\operatorname{conv} K(f_\chi^*)^*$ і $K(f_\chi^*) \cap K(f_\chi^*)^* = \varnothing$, нерівність (15) можна замінити нерівністю

$$\lambda_1 (\mathbf{a} - \mathbf{b}) + \mathbf{b} \neq \lambda_2 (\mathbf{g} - \mathbf{h}) + \mathbf{h} \tag{16}$$

за умови, що

$$\lambda_1 \in (0,1) \text{ та } \lambda_2 \in (0,1). \tag{17}$$

З (17) випливає, що існує таке число $r \in \{1, \ldots, \rho\}$, для якого має місце нерівність

$$\lambda_1(\alpha_r - \beta_r) + \beta_r \neq \lambda_2(\gamma_r - \delta_r) + \delta_r. \tag{18}$$

Покажемо, що коли має місце (16), (17) і $\alpha_r \neq \beta_r$, то $\gamma_r = \delta_r$. Це означає, що $\rho(\mathbf{g}, \mathbf{h}) < \rho(\mathbf{a}, \mathbf{b})$.

Розглянемо два можливі випадки.

1. Нехай $\alpha_r = 1$. Тоді $\beta_r = 0$ і з (18) маємо $\lambda_1 \neq \lambda_2(\gamma_r - \delta_r) + \delta_r$. Звідси

$$\gamma_r - \delta_r \neq \frac{1}{\lambda_2}(\lambda_1 - \delta_r),\ \gamma_r, \delta_r \in Z_2. \tag{19}$$

Ліва частина нерівності (19) набуває значення з множини $\{-1, 0, 1\}$, оскільки $\gamma_r, \delta_r \in Z_2$.

Права частина нерівності (19) з урахуванням (17) не може дорівнювати 0 для жодних значень λ_1, λ_2. Отже, нерівність (19) має місце для довільних $\lambda_1, \lambda_2 \in (0,1)$ тільки у тому разі, коли $\gamma_r - \delta_r = 0$. Отже, $\rho(\mathbf{g}, \mathbf{h}) < \rho(\mathbf{a}, \mathbf{b})$.

2. Нехай $\alpha_r = 0$. Тоді $\beta_r = 1$ і з (18) випливає, що

$$-\lambda_1 + 1 \neq \lambda_2(\gamma_r - \delta_r) + \delta_r,$$

або

$$(\gamma_r - \delta_r) \neq \frac{1}{\lambda_2}(1 - \lambda_1 - \delta_r).$$

Як і в попередньому випадку, остання нерівність має місце для всіх $\lambda_1, \lambda_2 \in (0,1)$ тільки тоді, коли $\gamma_r - \delta_r = 0$. Отже, $\rho(\mathbf{g}, \mathbf{h}) < \rho(\mathbf{a}, \mathbf{b})$. ■

Нехай $K(f_\chi^*) = \{\mathbf{a}_1, \ldots, \mathbf{a}_t\}$ — ядро булевої функції $f: G_n \to H_2$ ($f \neq \mathrm{const}$) відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ з модифікованим ядром $K(f_\chi, M)$ і $K(f_\chi^*)_i = \{\mathbf{a}_i \oplus \mathbf{a}_1, \ldots, \mathbf{a}_i \oplus \mathbf{a}_t\}$ — зведене ядро [17] функції f_χ^* відносно елемента $\mathbf{a}_i \in K(f_\chi^*)$.

Множину всіх зведених ядер булевої функції f_χ^* позначимо $T(f_\chi^*) = \{K(f_\chi^*)_i = \mathbf{a}_i K(f_\chi^*) \mid i = 1, 2, \ldots, t\}$.

Кажуть, що вектор $\mathbf{a} = (\alpha_1, \ldots, \alpha_m) \in Z_2^m$ передує вектору $\mathbf{b} = (\beta_1, \ldots, \beta_m) \in Z_2^m$ $\mathbf{a} \prec \mathbf{b}$, якщо $\alpha_i \leq \beta_i$ $(i = 1, 2, \ldots, m)$. Позначимо $N_\mathbf{a}$ множину всіх таких векторів із Z_2^m, які передують вектору $\mathbf{a}$.

З означення ядра $K(f_\chi^*)$ та з теореми 3 і теореми 4 з [16] безпосередньо маємо такі твердження.

Теорема 3. Якщо булева функція $f: G_n \to H_2$ реалізується одним узагальненим нейронним елементом відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ з модифікованим ядром $K(f_\chi, M)$, тоді ядро $K(f_\chi^*)$ функції f_χ^* задовольняє умову

$$\mathbf{a} = (\alpha_1, \ldots, \alpha_m) \in K(f_\chi^*) \Rightarrow \overline{\mathbf{a}} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_m) \notin K(f_\chi^*),$$

де $\overline{\alpha}_i$ — інвертоване значення α_i.

Теорема 4. Якщо булева функція $f: G_n \to H_2$ реалізується одним узагальненим нейронним елементом відносно системи характерів $\chi = \{\chi_{i_1}, \chi_{i_2}, \ldots, \chi_{i_m}\} \subset X(G_n)$ з модифікованим ядром $K(f_\chi, M)$, тоді множина зведених ядер $T(f_\chi^*)$ містить такий елемент $K(f_\chi^*)_i$, що

$$\forall \mathbf{a} \in K(f_\chi^*)_i \Rightarrow N_\mathbf{a} \subset K(f_\chi^*)_i.$$

ВИСНОВКИ

У цій роботі досліджено властивості узагальнених нейронних елементів і умови реалізовності функцій алгебри логіки за допомогою цих елементів. Введено поняття модифікованого ядра булевих функцій відносно системи характерів групи, на якій задано ці функції. Встановлено критерії реалізовності булевих функцій одним узагальненим нейронним елементом відносно заданої системи характерів, а також ефективні необхідні умови перевірки реалізовності функцій алгебри логіки одним узагальненим нейронним елементом.

На основі критеріїв реалізовності булевих функцій одним узагальненим нейронним елементом можна отримати ряд достатніх умов, які можуть бути успішно використані для синтезу цілочислових узагальнених нейронних елементів з великою кількістю входів.

СПИСОК ЛІТЕРАТУРИ

1. Hulianytskyi L.F., Riasna I.I. Formalization and classification of combinatorial optimization problems. *Optimization Methods and Applications*. Butenko S., Pardalos P.M., Shylo V. (Eds.). Springer International Publishing AG, 2017. P. 239–250.
2. Гуляницький Л.Ф., Мулеса О.Ю. Прикладні методи комбінаторної оптимізації: навч.посібник. Київ: Видавничо-поліграфічний центр «Київський університет», 2016. 133 с.
3. Azarbad M., Hakimi S., Ebrahimzadeh A. Automatic recognition of digital communication signal. *International Journal of Energy, Information and Communications*. 2012. Vol. 3, Iss. 4. P. 21–33.
4. Ізонін І.В., Ткаченко Р.О., Пелешко Д.Д., Батюк Д.А. Нейромережевий метод зміни роздільної здатності зображень. *Системи обробки інформації*. 2015. Вип. 9(134). С. 30–34.
5. Amato F., González-Hernández J. L., Havel J. Artifical neural networks combined with experimental desing: a “soft” approach for chemical kinetics. *Talanta*. 2012. Vol. 93. P. 72–78.
6. Гече Ф., Мулеса О., Гече С., Вашкеба М. Розробка методу синтезу прогнозуючої схеми на основі базових прогнозуючих моделей. *Технологічний аудит та резерви виробництва*. 2015. № 3/2(23). С. 36–41.
7. Kuchansky A., Biloshchytskyi A. Selective pattern matching method for time-series forecasting. *Eastern-European Journal of Enterprise Technologies*. 2015. Vol. 6, N 4(78). P. 13–18.
8. Dey P., Lamba A., Kumary S., Marwaha N. Application of an artifical neural network in the prognosis of chronic myeloid leukemia. *Analytical and Quantitative Cytology and Histology*. 2011. Vol 33, N 6. P. 335–339.
9. Liu A.S., Zhu Q. Automatic modulation classification based on the combination of clustering and neural network. *The Journal of China Universities of Ports and Telecommunication*. 2011. Vol. 18, N 4. P. 13–19.
10. Pathok A., Wadhwani A.K. Data compression of ECG signals using error back propagation (EBP) algorithm. *International Journal of Engineering and Advence Technology (IJEAT)*. 2012. Vol. 1, Iss. 4. P. 2249–8958.
11. Bodyansky Ye., Grimm P., Mashtalir S., Vinarski V. Fast training of neural networks for image compression. *Advances in Data Mining*. *Applications and Theoretical Aspects. Lecture Notes in Computer Science*. 2010. Vol. 6171. P. 165–173.
12. Шовгун Н.В. Анализ эффективности нечетких нейронных сетей в задаче оценки кредитного риска. *Information Technologies & Knowledge.* 2013. Vol. 7, N 3. P. 286–293.
13. Грицик В.В., Гече Ф.Е. Реалізація бульових та багатозначних логічних функцій на нейронних елементах. *Доповіді Національної академії наук України.* 2004. № 5. С. 65–68.
14. Голубов Б.И., Ефимов А.В., Скворцов В.А. Ряды и преобразования Уолша. Теория и применения. Москва: Наука, 1987. 343 с.
15. Кертис Ч., Райнер И. Теория представлений конечных групп и ассоциативных алгебр. Москва: Наука, 1969. 667 с.
16. Geche F., Mulesa O., Buchok V. Synthesis of generalized neural elements by means of the tolerance matrices. *Eastern-European Journal of Enterprise Technologies*. 2017. Vol. 4, N 4(88). P. 50–62.

17. Айзенберг Н.Н., Бовди А.А., Герго Э.Й., Гече Ф.Э. Некоторые алгебраические аспекти пороговой логики. *Кибернетика*. 1980. № 2. С. 26–30.
18. Лейхтвейс К. Выпуклые множества. Москва: Наука,1985. 335 с.
19. Карманов В.Г. Математическое программирование. Москва: Наука, 1986. 285 с.
20. Яджима С., Ибараки Т. Нижняя оценка числа пороговых функций. *Кибернитический сборник: новая серия*. Москва: Мир, 1969. Вып. 6. С. 72–81.

Надійшла до редакції 02.02.2018

Ф.Э. Гече, О.Ю. Мулеса

АЛГЕБРАИЧЕСКИЕ СВОЙСТВА ЯДЕР ОБОБЩЕННЫХ НЕЙРОФУНКЦИЙ

Аннотация. Рассмотрены обобщенные нейронные элементы, определены условия реализуемости функций алгебры логики на таких элементах. Введено понятие модифицированного ядра булевых функций относительно системы характеров группы, на которой задаются функции алгебры логики. Приведены критерии принадлежности этих функций к классу обобщенных нейрофункций. Исследована алгебраическая структура ядер булевых нейрофункций. На основе свойств матриц толерантности установлен ряд необходимых условий реализуемости булевых функций одним обобщенным нейронным элементом. Полученные в работе результаты позволяют разработать эффективные методы синтеза целочисленных обобщенных нейронных элементов с большим числом входов, которые могут быть успешно применены в задачах компрессии и передачи информации, а также в задачах распознавания дискретных сигналов.

Ключевые слова: обобщенный нейронный элемент, ядро функции, спектр функции, характер группы, синтез, матрица толерантности.

F. Geche, O. Mulesa

THE ALGEBRAIC PROPERTIES OF CORES OF GENERALIZED NEUROFUNCTIONS

Abstract. In this paper, we consider generalized neural elements and study the conditions for the implementation of the functions of the algebra of logic with such elements. The concept of a modified core of Boolean functions is introduced in relation to the systems of groups' character where the functions of the algebra of logic are defined. The criteria for belonging these functions to a class of generalized neural functions are given. The algebraic structure of the core of Boolean neurofunctions is investigated and number of necessary conditions for the implementation of Boolean functions by one generalized neural element are established based on the properties of tolerance matrices. The received results allow elaborating efficient methods of synthesis of integer generalized neural elements with many inputs, which can be used in problems of information compression and transmission and discrete signal recognition.

Keywords: generalized neural element, core of function, function's spectrum, group character, synthesis, tolerance matrix.

Гече Федір Елемирович,
доктор техн. наук, професор, завідувач кафедри Державного вищого навчального закладу «Ужгородський національний університет», e-mail: fgeche@hotmail.com.

Мулеса Оксана Юріївна,
кандидат техн. наук, доцент, доцент кафедри Державного вищого навчального закладу «Ужгородський національний університет», e-mail: Oksana.mulesa@uzhnu.edu.ua.