

УДК 004.512

Л.О. Катеринич

## КЕРУВАННЯ СИНТЕЗОМ НЕЙРОННИХ МЕРЕЖ

Пропонується підхід до керування синтезом нейронних мереж. Розглядається керування синтезом нейронних мереж з використанням комутаційного елемента, а саме етапи проектування та керування мережею з такою архітектурою. Формалізовано опис архітектури мережі з комутаційним елементом.

### Вступ

Синтез – процес, як правило цілеспрямований, поєднання або об'єднання раніше розрізнених речей або понять в дещо якісно нове, ціле.

Керування – процес організації такої цілеспрямованої дії на об'єкт, внаслідок якого цей об'єкт переводиться у потрібний (цільовий) стан.

Керування синтезом нейронних мереж – процес певної структурної організації нейронних мереж у єдиний якісно новий об'єкт під впливом керуючого суб'єкта.

Розглядаючи керування, як цілеспрямований процес, необхідно також ввести поняття «суб'єкта». Суб'єкт задає цілі, які реалізуються керуванням, при синтезі нейронних мереж (НМ). Відповідні цілі виникають при взаємодії різноманітних потреб та об'єктом керування.

Якщо стан НМ задовольняє потребам суб'єкта при взаємодії, то ніякого керування не потрібно. Якщо стан не задовольняє, то суб'єкт має організувати такий вплив на НМ, яке привело її у новий стан, що задовольняє суб'єкта при вирішенні його задач. Елемент, на який відбувається вплив – це комутаційний елемент [1, 2]. Вплив виконується вибором певних дій, що призводять до активації комутаційного елемента.

### Архітектура нейронної мережі з комутаційним елементом

Стан НМ змінюється під впливом зовнішнього середовища, в якому він знаходиться. Позначимо  $Y$  – стан НМ,  $X$  – стан середовища. Тоді НМ можна представити, як перетворювач  $F_0$  стану середо-

вища в стан об'єкта:  $Y = F_0(X)$ , де  $F_0$  оператор зв'язку вхідних даних  $X$  і виходів  $Y$  нейронної мережі (рис.1).

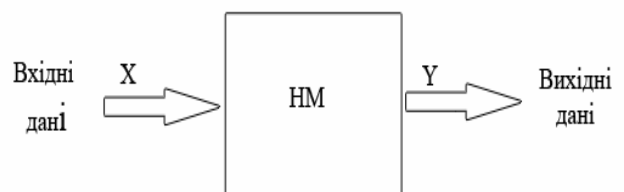


Рис. 1. Стан НМ (звичайний)

Оскільки роль керуючого елемента важлива для керування системою виділимо його з середовища. Позначимо  $D_x$  та  $D_y$  – структурований набір даних за допомогою яких він сприймає вхідні дані та дані НМ. Інформація  $\langle X', Y' \rangle$  утворює структурований простір вхідної і вихідної інформації, тобто той простір  $\langle X, Y \rangle$  інформації, яку комутаційний елемент (КЕ) здатен сприйняти й обробити у поточний момент часу.

В ідеальному випадку (хоча цього не може відбуватися практично ніколи) КЕ завжди для НМ формує цільові потреби  $Z^*$  (остаточний результат) які призведуть до досягнення мети у конкретному випадку.

Виконання цільових потреб  $Z^*$  задається рівністю  $Y = Z^*$ , а не виконання – нерівністю  $Y \neq Z^*$ .

В останньому випадку цілі не реалізуються. Необхідно вирішити наступні проблеми:

- все залишається без змін. Цілі не досягнуті;

- створити систему керування, яка реалізує цілі  $Z^*$  в НМ.

Для цього необхідно визначити, як можна впливати на НМ, тобто створити засіб керування. Такими засобами можуть бути додаткові входи НМ, що не завжди є оптимальним рішенням, або створити додатковий елемент мережі – комутаційний елемент [2, 3] (рис. 2), який буде керувати проміжними даними НМ.

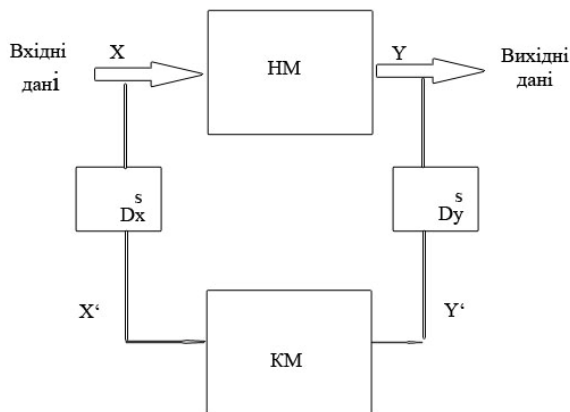


Рис. 2. Стан НМ (комутаційний елемент)

Розглянемо другий засіб. Нехай  $U$  – додатковий засіб керування. Тоді НМ можна представити як  $Y = F_0(X, U)$ . Виходячи з цього систему керування можна описати наступним чином:

$D_x$  – вхідні дані до системи (перевіряються експертом);

$D_y$  – вхідні дані до КЕ (вихідні данні НМ);

$X' = D_x(X)$  – структуровані вхідні дані експерта;

$Y' = D_y(Y)$  – структуровані вихідні дані до КЕ.

Структуровані дані передаються до КЕ, який перетворює їх на команди керування  $U$ . Ці команди мають бути оброблені КЕ, для того, щоб змінити стан керованого входу  $U'$  об'єкта.

Для функціонування КЕ необхідно поставити ціль  $Z^*$  керування, а також алгоритм керування  $\mu$  – вказівки, як досягти цілей, маючи інформацію про стан середовища, об'єкта і цілі (рис. 3):

$$U = \mu(X', Y', Z^*). \quad (1)$$

Керування пов'язане перш за все з цілями  $\{Z^*\}$ , які надходять з іншої НМ до КЕ. Ці цілі утворює суб'єкт, який і є споживачем з іншого боку виходів КЕ.

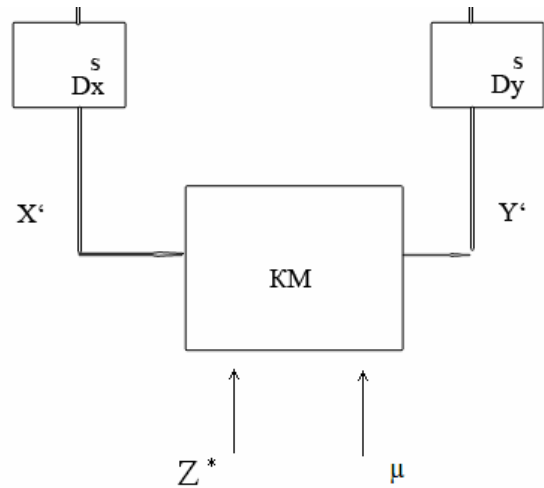


Рис. 3. Комутаційний елемент

### Етапи керування НМ з використанням комутаційного елемента

Керування складною системою можна описати такими етапами:

#### 1. Формування цілей керування.

На цьому етапі визначаються загальні цілі (множина цілей), які мають бути реалізовані у процесі керування з використанням КЕ.

КЕ сприймає вхідну інформацію, як скінчений набір параметрів  $S = (S_1, \dots, S_e)$ . Інакше кажучи КЕ завжди може сприйняти і обробити вхідну інформацію, тобто така подія завжди керована.

$$S(U) = (S_1(U), \dots, S_e(U)), \quad (2)$$

де  $U$  – керування комутаційним елементом.

Ведемо поняття простір ситуацій  $\{S\}$ , яке утворюється зазначеними параметрами  $S_i (i = 1, \dots, e)$ . Кожна точка цього простору визначає довільну конкретну ситуацію, яка виникла навколо КЕ (рис. 4). Отже простір  $\{S\}$  є простором сприйняття, тобто елементами множини є вхідні-

вихідні дані інших мереж або користувача системи.

Зміна ситуації (набір вхідних даних) навколо КЕ призводить до зміщення точки  $S$  вздовж траєкторії (лінії на рис. 4).

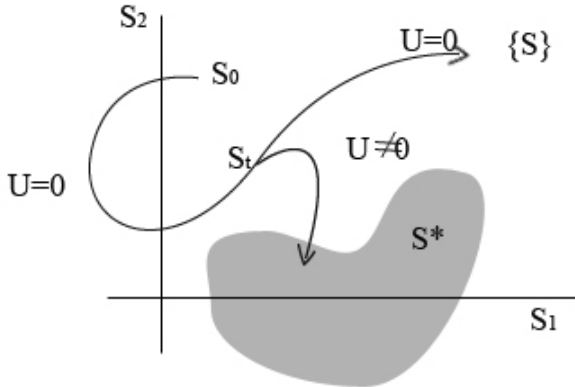


Рис. 4. Простір ситуацій

Але свої цілі користувач формує не в термінах множини  $S$ , а більш зрозумілишими термінами і поняттями, а саме симптомами. Опишемо ці цільові поняття вектором

$$Z = (z_1, \dots, z_k), \quad (3)$$

де кожен параметр  $z_i$  однозначно визначається ситуацією  $S$  (стан пацієнта), тобто

$$Z_i = \Psi_i(S) \quad (i = 1, \dots, k), \quad (4)$$

а функція  $\Psi_i(\bullet)$  визначає зв'язок стану середовища і цільового параметра  $z_i$ . У векторній формі цей зв'язок виражається у вигляді

$$Z = \Psi(S), \quad (5)$$

$\Psi(S) = (\Psi_1(S), \dots, \Psi_k(S))$  — деяка визначена вектор-функція.

Перетворення  $\Psi$  інформації у форму  $Z$  необхідно ще тому, що користувач, зазвичай, формує свої цілі в термінах, які зв'язані з термінами системи. Інколи може виникнути ситуація, коли  $Z = Y$  ( $k = m$ ).

Розглянемо  $k$ -мірний простір цілей  $\{Z\}$ , який утворюється точками (3). Цей простір зручний тим, що кожна координата є ціллю, виконання якої призведе до виконання поставленої задачі перед НМ з КЕ. Відповідну ціль сформулюємо у вигляді вектора-цілі

$$Z^* = (z_1^*, \dots, z_k^*),$$

де  $z_i^*$  -  $i$ -та потреба до стану середовища  $S$ , виражена за допомогою функції  $\Psi_i(S)$ . Цілі можуть бути різні, але форма запису має бути уніфікована.

Розглянемо уніфікований вигляд цілей  $Z^*$ . Для формалізації їх необхідно звести до однієї з наступних форм:

- $z_i^* : z_i = a_i$   $i$ -та цільова змінна  $z_i = \Psi_i(S)$  має дорівнювати заданій величині  $a_i$ ;
- $z_j^* : z_j \geq b_j$   $j$ -та цільова змінна не має бути менша заданого порогу  $b_j$ ;
- $z_v^* : z_v \rightarrow \min$   $v$ -та цільова змінна має бути мінімальна.

Якщо цілі не можна звести до цих форм, то не можна створити формальний опис системи керування НМ, а саме КЕ.

Розглянемо відображення цілей  $Z^*$  у просторі ситуацій  $\{S\}$ :

$$S^* : \begin{cases} \Psi_i(S) = a_i, (i = 1, \dots, s), \\ \Psi_j(S) \geq b_j, (j = s + 1, \dots, s + p), \\ \Psi_v(S) \rightarrow \min, (v = s + p + 1, \dots, s + p + l), \end{cases} \quad (6)$$

де  $s + p + l = k$ .

Точка або область  $S^*$  яка задовольняє цим потребам і є тим станом, що необхідно досягти НМ з КЕ. Відбудеться це чи ні залежить від можливостей впливу на КЕ, тобто від вигляду  $S = S(U)$ , де  $U \in R$ ,  $R$  — ресурси, що доступні.

Розглянемо взаємодію цільової зони  $S^*$  і траєкторії зміни середовища  $S(t)$ . Якщо траєкторія руху  $S(t)$  проходить через зону то керування не потрібно. В цьому випадку  $S(t)$  з часом буде належати  $S^*$ . Але така ситуація мало ймовірна, тому

слід впливати на середовище  $S_t = S(U, t)$  і змінювати так, щоб  $S(U, t) \in S^*$ , де  $t$  — це рух середовища.

## 2. Визначення об'єкта керування.

Інколи виникає ситуація коли зв'язок об'єкта і середовища настільки сильний, що їх відокремити неможливо. Коли предметна область чітка, то така задача не виникає. Це буває у таких випадках, коли предметна область автономна. Однак у реальному житті предметні області настільки зв'язані, що визначити чітку межу майже не можливо. Ця обставина є поштовхом для виділення процесу визначення меж предметної області та/або об'єкта в окремий етап.

У середовищі, що має формальний опис виділення об'єкта не є проблемою. Обмежуючи певні фактори середовища ми отримуємо певний екземпляр об'єкта і перевіряємо на моделі досяжність поставлених цілей. Також можливо оцінити чисельно ступінь некерованості цього варіанта об'єкта. Повторюючи цю процедуру для інших варіантів об'єктів буде отримано той екземпляр об'єкта на якому може бути досягнута максимальна керованість процесом.

Маючи формальний опис середовища це не було б проблемою. Оскільки існування такого опису майже неможливо звертаються до експертів. Для цього необхідно експертно синтезувати декілька варіантів об'єкта, а потім за допомогою експерта оцінити їх за критерієм і обрати найкращий.

Таким чином, виникає потреба виділити процес визначення об'єкта (КЕ) в самостійний етап керування. Задача зводиться до того, щоб для заданої множини цілей  $\{Z^*\}$  і ресурсів  $R$  визначити такий КЕ, який за критеріями досяжності цілей буде найкращий.

**3. Структурний синтез моделі НМ з КЕ.** Під структурою будемо розглядати залежність  $F$  стану виходів об'єкта  $Y$  (НМ) від його входів — не керованого  $X$  та керованого  $U$ .

$$Y = F(X, U).$$

У загальному випадку залежність  $F$  визначається деяким алгоритмом побудови НМ (алгоритм зворотнього розповсюдження помилки) [4], який вказує, як маючи інформацію про входи  $X$  і  $U$ , визначити вихід  $Y$ . Вид цього алгоритму з точністю до його параметрів визначає структуру  $F$ . Умовно будемо вважати, що модель  $F$  складається з структури і параметрів:  $F = \langle S_t, C \rangle$ , де  $S_t$  структура моделі  $F$ ,  $C = (C_1, \dots, C_k)$  — її параметри. Таким чином, ціллю третього етапу є визначення структури  $S_t$  з КЕ.

Наприклад, категорії лінійності, статичності, детермінованості, дискретності — структурні категорії. Лінійна статична неперервно детермінована структура однозначно визначає наступний вид для  $F$ :  $y = c_1 x_1 + \dots + c_n x_n$ , причому на стадії структурного синтезу конкретні значення  $c_1, \dots, c_n$  поки не суттєві. Важлива лише залежність  $F$  від цих параметрів і входів до КЕ.

Структурний синтез моделі можна поділити на такі етапи:

- 1) визначення входів і виходів НМ та КЕ;
- 2) експертне ранжування входів і виходів об'єкта;
- 3) декомпозиція моделі;
- 4) вибір структурних елементів моделі.

**4. Параметричний синтез моделі КЕ.** Цей етап зв'язаний з визначенням параметрів  $C = (c_1, \dots, c_k)$  моделі  $Y = F(X, U, C)$ , де обрана на попередньому етапі структура  $S_t$  відображена у модельному операторі  $F$ .

Для визначення параметрів  $C$  моделі необхідно мати інформацію про вхід  $X$ ,  $U$  та вихід  $Y$  НМ та КЕ.

У залежності від того, як отримана інформація, розглянемо два підходи:

1. *Ідентифікація* параметрів моделі зв'язана з оцінкою чисельних значень шуканих параметрів у стандартному режимі функціонування НМ, тобто використання КЕ елемента не потрібна. Початковою інформацією для ідентифікації є структура  $S_t$  і спостереження за поведінкою входів  $X(t)$  і виходів  $Y(t)$ . Пара

$J(t) = \langle X(t), Y(t) \rangle$  отримана без використання КЕ є основною інформацією при ідентифікації.

У деяких випадках не всі входи  $X$  (не керовані) та  $U$  (керовані) змінюються в процесі нормального функціонування НМ. Не змінюються ті параметри з  $U$ , на які не впливають зовнішні фактори. Для з'ясування залежності виходів об'єкта  $Y$  від параметрів такого типу необхідно примусово їх змінювати, тобто експериментувати з ними. Оскільки будь-який експеримент порушує роботу НМ у нормальному режимі, то слід при мінімальному впливі на об'єкт отримувати максимальну інформацію про вплив цих параметрів на НМ. У цьому випадку слід звернутися до планування експерименту.

2. *Планування експериментів.* У процесі синтезується спеціальний підхід, який дозволяє при заданих обмеженнях з максимальною ефективністю визначити параметри  $S$  моделі об'єкта керування.

При використанні НМ мережі цей підхід має на увазі, що набір станів керованого входу НМ  $U_1, \dots, U_n$  належать заданій допустимій області експерименту, в результаті чого визначається його вихід  $Y_1, \dots, Y_n$ , тобто  $Y_i = F_0(U_i)$  ( $i = 1, \dots, n$ ). Отримані  $n$  пар  $\langle U_i, Y_i \rangle$  ( $i = 1, \dots, n$ ) є вихідними даними для визначення необхідних параметрів моделі НМ.

**5. Синтез та реалізація керування КЕ.** Синтез керування пов'язаний з прийняттям рішення про те, яке має бути керування  $U$ , таке, що при ситуації  $S$  досягнути поставленої цілі  $Z^*$  в об'єкті. Цей розв'язок спирається на існуючу модель об'єкта  $F$ , задану цілю  $Z^*$ , отриману інформацію про стан середовища  $X$  і об'єкта  $Y$ , а також на виділені ресурси  $R$  керування, які є обмеженнями на керування  $U$  у зв'язку зі специфікою об'єкта і можливостями середовища керування. Синтез керування зводиться до розв'язку варіаційної задачі, яка отримується з  $Y = F_0(X, U)$  шляхом відповідних перетворень, необхідних для преодоління багатокритеріальності.

Отримане керування має бути оптимальне з точки зору цілей керування і мати вигляд  $U^* = U^*(t)$ .

Реалізація керування зв'язана з процесом обробки об'єктом програми, отриманої на попередньому етапі. Такий процес для не активних об'єктів розв'язується методами теорії спостережуваних систем [5], обробляючи дану програму. Значно складніше реалізація керування активної системи. Ці труднощі мають бути вирішені на стадії синтезу моделі об'єкта керування, враховуючи його активність. Тоді обробка програми буде, як і при керуванні, пасивним об'єктом.

Якщо керування реалізоване, а його ціль не досягнута, необхідно повернутися на один з попередніх етапів. Але навіть коли цілі досягнуті, інколи необхідно повертатися до попередніх етапів, оскільки відбуваються зміни стану середовища  $X$  або цілей керування  $Z^*$ .

**6. Адаптація** (або корекція) всієї системи керування (всіх етапів) розглядається, як більш глибокий аналіз процесу керування складної системи.

**Параметрична адаптація** зв'язана з корекцією та налаштуванням параметрів  $S$  моделі. Така адаптація має місце при будь-якій зміні характеристик керуючого об'єкта, а саме КЕ. Адаптація дозволяє налаштувати модель на кожному кроці керування. Вихідною інформацією для цього виду адаптації є неспівпадання вихідної/вхідної інформації та необхідною вхідною для КЕ. Оскільки процес керування нейронною мережею з КЕ не завжди здатен представити необхідну інформацію для найбільш точного остаточного результату, необхідно вводити додаткові елементи керування, а саме поповнювати вхідною інформацією початкову структуру вхідних даних НМ, що, в свою чергу, призведе до формування більш точних вхідних даних для КЕ.

**Структурна адаптація** має місце при неспівпаданні структури моделі та структури НМ. За різних обставин структура НМ може змінитися, що, в свою чергу, буде вимагати зміни структури моделі,

а саме до механізмів структурної адаптації. Перехід від однієї структури до іншої виконується процедурою вибору альтернативної структури моделі. Різниця може полягати як у кількості входів/виходів до моделі, варіантами декомпозиції або структурою елементів моделі. Слід зазначити, що альтернативні моделі при ідентифікації параметрів теж використовують методи параметричної адаптації.

**Адаптація об'єкта** необхідна в тому випадку, коли структурна адаптація не ефективна, тобто деякі цілі керування не досягаються. Ця адаптація зв'язана зі зміною організації потоків інформації в НМ. Додавання нової інформації призводить до підвищення його керованості, але вимагає більше затрат при параметричній та структурній адаптації. Вибір найкращого варіанта розподілення та переміщення потоків інформації у НМ в процесі керування і є основою процесу адаптації на цьому етапі.

**Адаптація цілей керування.** В робочому режимі системи де використовується НМ може виникати ситуація, коли виникають нові цілі, або їх уточнення. В цьому випадку необхідно визначити нову множину цілей  $\{Z^*\}$ , адаптовану під поточні потреби. Це є основою цього етапу.

Розглянемо задачу адаптації об'єкта  $F^0$ , як задачу керування. Ціль  $Z^*$  адаптації визначає вимоги до критеріїв. Ці потреби можуть бути записні наступним чином:

1. Критерій-нерівність  
 $H(U) = (h_1(U), \dots, h_p(U)) \geq 0$ .
2. Критерій-рівність  
 $G(U) = (g_1(U), \dots, g_s(U)) = 0$
3. Критерії, що мінімізують  
 $Q(U) = (q_1(U), \dots, q_l(U)) \rightarrow \min$ ,

де  $h_i = M_x h_i'(Y) = M_x h_i'(F^0(X, U))$

$(i = 1, \dots, p)$ ;

$g_j = M_x g_j'(Y) = M_x g_j'(F^0(X, U))$

$(j = 1, \dots, s)$ ;

$g_k(U) = M_x q_k'(Y) = M_x q_k'(Y) = M_x q_k'(F^0(X, U)); \quad k = (1, \dots, l)$ ;

$M_x$  – оператор усереднення по  $X$ :

$M(\bullet) = \int (\bullet) \rho(X) dX$ , де  $\rho(X)$  – щільність розподілу стану  $X$  середовища.

Ціль адаптації це розв'язок наступної задачі:

$$Q(U) \rightarrow \min_{U \in S} \Rightarrow U^*, \quad (7)$$

$$\text{де } S : \begin{cases} H(U) \geq 0 \\ G(U) = 0. \end{cases}$$

Такі задачі при заданій моделі  $F$  об'єкта  $F^0$  і розподілу  $\rho(X)$  називають задачами стохастичного програмування [6]. Проблема розв'язку поставлених задач очевидна. Але задачі адаптації ускладнюються ще й тим фактором, що модель об'єкта  $F$  відсутня. Також відсутні інформація про розподіли  $\rho(X)$ . Більш того, ці фактори мають тенденцію змінюватися у часі.

Отже, маючи алгоритм  $\varphi$  розв'язку задачі (7) з обмеженнями вигляду (4) у вигляді

$$U^* = \varphi(F, \rho(X)), \quad (8)$$

не можна вважати задачу адаптації розв'язаною. Необхідно додатково ідентифікувати об'єкт, що змінюється  $F^0$  і статистичні властивості середовища, щоб отримати

$$W = V(F_t, \rho_t(X)), \quad (9)$$

де  $F_t$  – модель об'єкта, а  $\rho_t$  – щільність розподілу  $X$ . Ці обставини змушують відмовитися від пошуку алгоритму  $\varphi$  розв'язку поставленої задачі адаптації, оскільки для складного об'єкта адаптації залежності  $F_t$  і  $\rho_t(X)$  достатньо складні.

Тому для розв'язку задачі (7) необхідно звертатися до алгоритмів адаптації, використовуючи лише значення функцій

$$h_i'(\bullet), \quad g_j'(\bullet) \quad \text{і}$$

$$q_k'(\bullet) \quad (i = 1, \dots, p; j = 1, \dots, s; k = 1, \dots, l)$$

у визначені моменти часу. В загальному випадку має такий вигляд:

$$U_{N+1} = \psi(\overrightarrow{U_N}, \overrightarrow{H_N}, \overrightarrow{G_N}, \overrightarrow{Q_N}), \quad (10)$$

де  $\psi$  — алгоритм рекурентної адаптації.

$$\overrightarrow{U_N} = (U_N, U_{N-1}, \dots, U_{N-W});$$

$$\overrightarrow{H_N} = (H'_N, H'_{N-1}, \dots, H'_{N-W});$$

$$\overrightarrow{G_N} = (G'_N, G'_{N-1}, \dots, G'_{N-W});$$

$$\overrightarrow{Q_N} = (Q'_N, Q'_{N-1}, \dots, Q'_{N-W});$$

$$H' = (h'_1, \dots, h'_p);$$

$$G' = (g'_1, \dots, g'_s);$$

$$Q' = (q'_1, \dots, q'_l)$$

де  $U'_N, H'_N, G'_N, Q'_N$  — значення параметрів, що адаптуються,  $N$  — момент часу,  $W$  — глибина пам'яті алгоритму  $\psi$ .

### Висновки

Запропонований підхід до керування синтезом нейронних мереж з використанням комутаційного елемента дозволяє прискорити роботу систем з відповідною архітектурою мережі. Кожен з етапів вимагає ретельного аналізу з боку розробника, оскільки деякі етапи проектування вимагають не лише чіткого формального опису, а й чітко сформовану інформацію від експерта. Наприклад, простір цілей формується майже виключно або під контролем експерта предметної області для якої розробляється відповідна НМ. Таким чином, можна виділити наступну перевагу: запропонована модель враховує різні чинники, що можуть впливати на стан нейронної мережі та комутаційного елемента. Зокрема, при формуванні цілей, враховується не лише загальна ціль для нейронної мережі, а й проміжні – для комутаційного елемента.

Отже, штучні нейронні мережі можна вважати універсальним потужним засобом моделювання нелінійних об'єктів. Синтез оптимальних алгоритмів керування багатозв'язними нелінійними об'єктами є оптимальним кроком розв'язання поставлених задач.

1. Катеринич Л., Проватар А. Синтез нейронных сетей на основе информационных гранул // International Book Series "Information Science and Computing: Advanced Research in Artificial Intelligence". – Sofia, 2008. – V.1. – P. 179–182.
2. Катеринич Л., Проватар А. Диагностирование на нейронных сетях в системе Голеопат // XIII-th International Conference: Knowledge Dialogue Solution. – Sofia, 2007. – V.1. – P. 64–68.
3. Bargiela, Andrzej and Pedrycz, Witold. Granular Computing: An introduction. – Kluwer Academic Publishers, 2003. – 5 p.
4. Хайкин С. Нейронные сети: полный курс, 2-е изд. – М.: ООО «И.Д. Вильямс», 2006. – 220 с.
5. Ньютон Дж. К., Гулд Л.А., Кайзер Дж. Ф. Теория линейных следящих систем. Аналитические методы расчета. – М.; 1981. – 29 с.
6. Юдин Д.Б. Математические методы управления в условиях неполной информации. – М.: Сов. радио, 1974. – 400 с.

Отримано 05.12.2008

### Про автора:

Катеринич Лариса Олександрівна,  
асистентка факультету кібернетики  
Київського національного університету  
імені Тараса Шевченка,  
e-mail: [katerinich@rambler.ru](mailto:katerinich@rambler.ru)

### Місце роботи автора:

Факультет кібернетики  
Київського національного університету  
імені Тараса Шевченка.