

## АДАПТИВНЫЙ РЕГУЛЯТОР НА НЕЙРОПОДОБНЫХ СТРУКТУРАХ

Леднев М. Ю.<sup>1</sup>, Бодякин В.И.<sup>2</sup>

(Учреждение Российской академии наук  
Институт проблем управления РАН, Москва)

*Рассматривается новый метод управления динамическими объектами. В качестве примера анализируется нейросемантический регулятор для управления априорно неизвестным объектом. Анализируются теоретические и практические результаты процесса обучения и функционирования регулятора.*

Ключевые слова: адаптивные регуляторы, адаптивное управление, нейроподобная структура.

### 1. Введение

Для управления все усложняющимися производственными процессами необходимы новые методы, основой которых могут стать нейросетевые регуляторы. Для настройки рассматриваемого в статье нейросетевого регулятора необходимо только задать классы состояний, определяемые, например, как: "нормальные" или "ненормальные" и, независимо от характера возмущений, такая система управления посредством внешних воздействий должна найти алгоритм (последовательность команд: ситуация – команда управления) по поддержанию управляемого объекта в "нормальном" режиме [2].

---

<sup>1</sup> Михаил Юрьевич Леднев, аспирант (miklesus@mail.ru).

<sup>2</sup> Владимир Ильич Бодякин, кандидат физико-математических наук, (body@ipu.ru).

Фактически, при минимуме информации ("нормальное - ненормальное" состояние) и при возможности исследования вариантов управляющих воздействий (этап обучения), регулятор способен вывести управляемый объект в область нормальных состояний его функционирования при любых допустимых начальных значениях, а также при случайных возмущениях внешней среды. В общем случае, нейросетевой подход – это "размен" сложной алгоритмической процедуры некоторой задачи на распределенную структуру простых преобразований, решающих эту же задачу.

В работе рассматривается особый класс классических нейросетевых систем – нейросемантические сети (НСС). В отличие от традиционных нейронных сетей, в НСС каждому нейроподобному элементу соответствует определенная семантическая единица исходного текстового потока<sup>1</sup>, что позволяет разработчику анализировать логику функционирования элементов системы в терминах языка предметной области. Второй существенной характеристикой НСС является ее ассоциативность.[1]

Построение теории нейросетевых адаптивных регуляторов весьма актуально как для технической сферы, так и для моделирования ими возможных механизмов в живых системах при поддержании ими гомеостаза в непрерывно изменяющихся средах.

## **2. Постановка задачи**

Рассмотрим объект, характеризующийся некоторой детерминированной функцией перехода внутренних состояний –  $f(X)$  и областью существования –  $\Delta S$ , по некоторому вектору состояний –  $X$ .

---

<sup>1</sup> Семантическая единица текста - это последовательность знаков, описывающая некоторый образ (процесс, объект), целостность и смысловую нагрузку которой, в свою очередь, однозначно определяют эксперты, понимающие данный текстовой материал.

В качестве примера таких объектов можно привести их физические аналоги: "обратный маятник" (по одной координате), "шарик на наклонной плоскости" и др., см. рис. 1 а) и б).

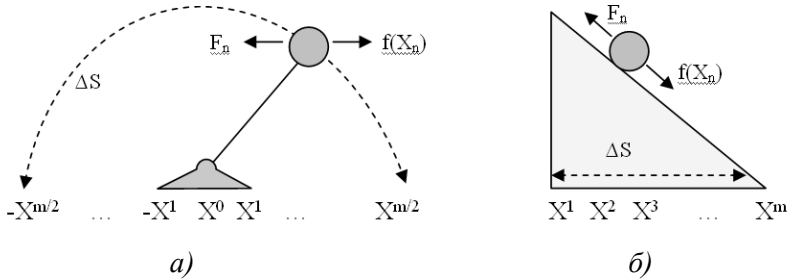


Рис. 1. Примеры динамических объектов исследования

Задача управления будет заключаться в выборе на  $n$ -ом шаге управляющего вектора воздействия  $F_n$  на объект, находящегося в состоянии  $X_n$ . Необходимо, чтобы в этом процессе НСС-регулятор обучился максимально долго ( $n > \max, n \rightarrow \infty$ ) удерживать объект в его области существования  $-\Delta S$ , за минимально возможное число попыток обучения  $-J$ .

Формально запись постановки задачи выглядит следующим образом:

$$(1) \quad (X_{n+1} = f(X_n) + F_n + Z_n)^J$$

Где:  $X_0, X_n = \{X^1, X^2, X^3, \dots, X^m\} \in \Delta S$ ;

$X, F, Z \in S$ ;

$X_n$  – параметр состояния объекта;

$X_0$  – начальное значение параметра объекта;

$m$  – число дискрет параметра  $X \in \Delta S$ ;

$S$  – область параметров состояния объекта и его управления;

$\Delta S$  – область параметров существования объекта;

$f(X)$  – функция перехода внутреннего состояния объекта;

$F_n$  – управление значением параметра объекта.  $F_n = \{F^1, F^2, \dots, F^q\}$ . При этом,  $\exists F_n : (\forall X_n + F_n) \in \Delta S$ ;  
 $Z_n$  – возмущение значений параметра объекта ( $Z_n < X_n$ );  
 $n$  – число шагов удовлетворительного управления в  $j$ -й попытке;  
 $J \rightarrow \min$  (число попыток обучения).

Мощность ( $M$ ) числа состояний  $\Delta S$  равна  $M_S = (N^2 + N) / 2$  и  $Q$ , соответственно, равна  $M_Q = ((N-1)^2 + N-1) / 2$ , что при принятом, в качестве примера,  $N=7$ ,  $M_S = 28$  и  $M_Q = 21$ .  $M_w = (N * N)^1 = N^2 = M_S + M_Q = 28 + 21 = 49$ .

Отметим, что  $M_S \sim M_Q$  ( $M_S = M_Q + N$ ). Эта оценка примерного равенства мощностей множеств  $\Delta S$  и  $Q$  будет нам удобна при оценке вероятностей случайного формирования траекторий.

### 3. Управление объектом

Из уравнения (1) следует, что процесс управления заключается в выборе из декартового пространства состояний объекта и управления ( $X * F$ ) выбрать все пары  $(X_n, F_n)$ , которые принадлежат  $\in Q$ .

Решить задачу выбора  $F_n$ , на текущее  $X_n$ , при априорно неизвестном объекте, можно только перебором (случайным, направленным, ...) в пространстве состояний ( $X * F$ ). Это пространство будет разделено на два множества подпространств: пары  $(X_n, F_n)$ , которых  $\in \Delta S$  и пары  $(X_n, F_n)$ , которых  $\notin \Delta S$ .

Область значений "ненормальных" состояний объекта  $Q$  ( $X_{n+1} = 1 + X_n + F_n + Z_n < 1$  и  $X_{n+1} = 1 + X_n + F_n + Z_n > 7$ ) совместно с областью "нормальных" состояний объекта  $\Delta S$  ( $1 \leq X_{n+1} = 1 + X_n + F_n + Z_n \leq 7$ ) представлены в таблице 1.

Таблица 1. Область значений состояний объекта

		$F_n$						
		$-a_1$	$-a_2$	$-a_3$	$-a_4$	$-a_5$	$-a_6$	$-a_7$
$X_n$	$a_1$	$a_1$	Q	Q	Q	Q	Q	Q
	$a_2$	$a_2$	$a_1$	Q	Q	Q	Q	Q
	$a_3$	$a_3$	$a_2$	$a_1$	Q	Q	Q	Q
	$a_4$	$a_4$	$a_3$	$a_2$	$a_1$	Q	Q	Q
	$a_5$	$a_5$	$a_4$	$a_3$	$a_2$	$a_1$	Q	Q
	$a_6$	$a_6$	$a_5$	$a_4$	$a_3$	$a_2$	$a_1$	Q
	$a_7$	$a_7$	$a_6$	$a_5$	$a_4$	$a_3$	$a_2$	$a_1$
		S						

Для выявления этих множеств в траекториях  $X_0 \rightarrow X_1, F_1 \rightarrow X_2, F_2 \rightarrow \dots \rightarrow X_{n+1} : X_{n+1} \notin \Delta S$  в  $j$ -ой траектории обучения, мы выбираем последнюю пару  $X_n, F_n$  и ставим запрет на  $F_n$ , при текущем состоянии объекта –  $X_n$ . Таким образом, значение  $F_n$  при значении  $X_n$  попадает под запрет управления регулятором. Целенаправленно перебрав все пространство состояний ( $X^*F$ , т.е. все пары  $X_n, F_n = \{X^1, X^2, X^3, \dots, X^m\} * \{F^1, F^2, \dots, F^q\}$ ) за  $J$  траекторий, которые назовем, числом попыток обучения, мы обучим нейросемантический регулятор устойчиво управлять данным объектом (см. рис. 2). Процесс обучения регулятора подробно рассматривается в [3].

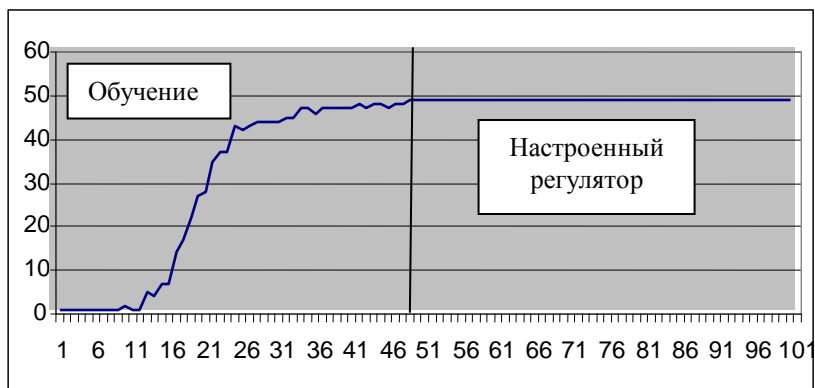


Рис. 2. График обучения нейросемантического регулятора

Для последующего теоретического исследования механизмов адаптации нейросемантического регулятора рассмотрим физическую модель, представленную на рис. 1(б).

Рассмотрим схему адаптивного регулятора на нейросемантической сети, см. рис. 3.

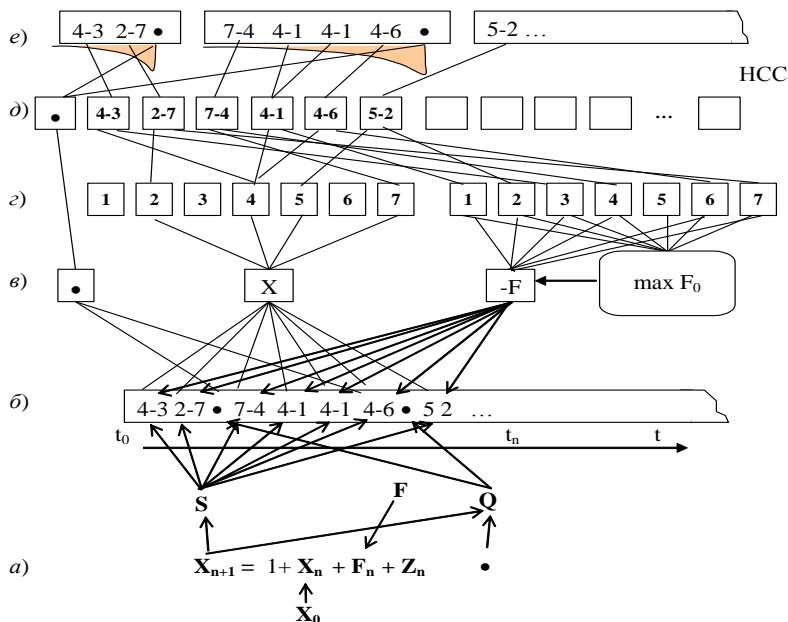


Рис. 3. Схема адаптивного нейросемантического регулятора

Это двухслойная НСС (слои д и е, см. рис. 3) с функционально упрощенными нейроподобными N-элементами [2]. Основное их отличие от классических нейронов [1] с алгебраической суммацией сигналов на входе заключается в векторной (пространственно-временной) обработке входных сигналов. Механизм функционирования N-элементов станет ясным из описания алгоритма настройки рассматриваемого адаптивного регулятора.[3]

Пример механизма изменения состояний объекта (X<sub>n</sub>) отображен на уровне а рисунка 3. На уровне б отображен протокол

изменения состояний объекта и управляющих воздействий регулятора ( $F_n$ ). Протокол привязан к дискретной временной оси ( $t_n$ ). Алфавит протокола состоит из  $N+2$  знаков (1, 2, 3, 4, 5, 6, 7, -, •). Грамматика формирования протокола – это последовательность попыток обучения, каждая из которых - это последовательность пар ( $X_n F_n$ ), завершающаяся знаком "•". Если же длина последовательности пар превышает некоторую длину (например, равную  $N^2 \sim 2 \cdot M_S$ ), то она рассматривается как положительный результат обучения, и при обучении "методом наказания" в памяти регулятора может не отображаться.

На уровне  $\epsilon$  отображены компоненты процесса: состояния объекта ( $X_n$ ), воздействий регулятора ( $F_n$ ), завершения  $i$ -й попытки (•), а также механизм блока управляющего воздействия (блок  $\max F_0$ , см. рис. 3). Работа блока заключается в выборе элемента из  $F$ -группы элементов уровня  $\epsilon$  с максимальным значением некоторой величины  $U(n)$ , которая количественно отражает величину меры "близости к наказанию", т.е. к "•", вычисляемую на  $N$ -элементах.

На уровне  $\epsilon$  отображены две группы элементов алфавита, отображающие значения состояния объекта ( $X_n$ ) и воздействия регулятора ( $F_n$ ). Эти элементы выполняют функцию простейшей ассоциативной коммутации различных пар ( $X_n F_n$ ) для  $N$ -элементов слоя  $\delta$  при реализации конкретных попыток обучения регулятора.

Уровень  $\delta$  – это первый слой  $N$ -элементов нейросемантической сети адаптивного регулятора.  $N$ -элементы этого слоя имеют два входа, для компоненты  $X_n$  состояния объекта (1-й вход) и для  $F_n$  ответного воздействия регулятора (2-й вход). Это слой ассоциативных  $N$ -элементов, отображающих простейшие "условные рефлексы" регулятора. В нем находятся только различные (уникальные) пары ( $X_n F_n$ ). Если в протоколе встречаются тождественные пары, например, "4-1... и 4-1", то они отображаются только одним  $N$ -элементом с информационным содержанием "4-1". Каждая же новая реализация пары, тождественная информационному содержанию  $N$ -элемента, только добавляет ссылку (связь) с  $N$ -элементом вышележащего слоя

(уровня  $e$ ), отображающим конкретную попытку обучения. Общее число  $N$ -элементов  $d$  слоя не превышает  $N^{m^2}$ .

$N$ -элементы уровня  $e$  имеют число входов, равное числу пар  $(X_n F_n)$  в каждой попытке обучения. *Завершающая пара* формируется либо при попадании  $X_n$  в множество состояний  $Q$  (сигнал - "•", это дополнительный вход, показанный на рис. 5 и приведен лишь для большей иллюстративности материала) либо, если число пар превышает заданное  $N^2$  – "положительная траектория". Новый  $N$ -элемент уровня  $e$  формируется из резервного  $N$ -элемента (самый правый свободный в слое). Сигналом к формированию нового  $N$ -элемента служит появление завершающей пары. Резервный  $N$ -элемент имеет практически бесконечное число входов.

Свойство слоев  $N$ -элементов адаптивного регулятора таково, что в каждом  $N$ -элементе содержится ровно одна семантическая единица своего иерархического уровня ("условный рефлекс" как конкретная пара  $X_n F_n$  – слой  $d$ , или конкретная траектория – в слое  $e$ ). Семантическая единица представляет последовательность знаков в рамках алфавита объекта (1, 2, 3, 4, 5, 6, 7, –, •). Свойство соответствия семантической единицы в конкретном физическом отображении ( $N$ -элемент) чрезвычайно важно в адаптивных информационных системах, поэтому это свойство и подчеркивается в его названии – нейросемантический. Более подробно этот аспект рассматривается в [2].

#### 4. Заключение

Результаты экспериментов с адаптивным регулятором подтвердили ожидаемые аналитические расчеты. Так при мощности  $M_s \sim M_Q$ , вероятность ( $p$ ) завершения траектории, т.е. попадания пары  $(X_n F_n)$  в область  $Q$  в начале процесса обучения регулятора в зависимости от числа шагов траектории ( $n$ ) велика и имеет следующий числовой вид:  $n=2$ ,  $p=0.50$ ;  $n=4$ ,  $p=0.88$ ; ...  $n=8$ ,  $p=0.99$ ; Соответственно этим и объясняются короткие начальные траектории в экспериментах.[3]

На этапе завершения обучения (число попыток  $\sim M_Q$ ) длины траекторий должны несколько возрасть, но ненамного. Веро-



ятность ( $p$ ) того, что регулятор сможет построить траекторию длиной до  $N^*$  пар  $(X_n, F_n)$  при последней ненайденной паре  $(X_n, F_n)$  из области  $Q$  равна  $p \sim 0.007$ . Таким образом, при числе попыток  $MW$  ( $MW = N2 = MS + MQ$ ) вероятность нахождения всех пар  $(X_n, F_n)$  области  $Q$  и запоминания их регулятором в форме  $N$ -элементов уровня  $d$  порядка  $p \sim (1 - 0.00005)$ . Следовательно, за  $M_w$  обучающих траекторий адаптивный регулятор на нейросемантической сети способен обучиться и устойчиво поддерживать ( $p > 0.9999$ ) объект в области  $S$ , по сравнению с аналогичными регуляторами на классических нейронных сетях.

[4][5]

Эксперименты по синтезу алгоритмов управления более сложными динамическими объектами, по сравнению с представленными на рисунке 1а и 1б, подтверждают эффективность предложенного метода и доказывают возможность его применения без какого-либо дополнительного анализа объекта управления. Преимуществом метода является возможность быстро обучаться управлению объектами, о которых нет никакой априорной информации.

### **Литература**

1. БОДЯКИН В.И. *Концепция построения самообучающихся информационно-управляющих систем на базе нейросемантической парадигмы* // Управление развитием крупномасштабных систем (MLSD'2012): Труды шестой международной конференции. Том 2. – М.: Учреждение Российской академии наук Институт проблем управления им. В.А.Трапезникова РАН – 2012. – С. 289-298
2. БОДЯКИН В.И., *Куда идешь, человек? Основы эволюциологии. Информационный подход.* – М. СИНТЕГ, 1998. – 332с.
3. БОДЯКИН В.И., ВЛАДИМИРОВА С.С., ЧИСТЯКОВ А.А., Построение адаптивного регулятора на нейросемантической сети, //Устойчивость и колебания нелинейных систем управления, VIII Международный семинар, 74-79с, М., ИПУ РАН, 2004

4. УОСЕРМЕН Ф. *Нейрокомпьютерная техника* / М.: Мир. 1992. – 240 с.
5. SUTTON R.S., Barto A.G. "Reinforcement Learning: An Introduction [Электронный ресурс]. – Режим доступа: <http://www-anw.cs.umass.edu/~rich/book/the-book.html>

## ADAPTIVE CONTROLLER ON NEURAL STRUCTURES

**Mikhail Lednev**, Institute of Control Sciences of RAS, Moscow, graduate student (miklesus@mail.ru).

*Abstract: A new method of control of dynamic objects. As an example, analyzes the neuro-semantic controller for controlling a priori unknown object. The theoretical and practical results of the learning process and the operation of the regulator.*

**Keywords:** adaptive controllers, adaptive control, neural structure.