

**Engineering Sciences****Технические науки**

UDC 004.85 : 004.81

**Method of Computer-aided Instruction in Situation Control Systems**<sup>1</sup> Anatoliy O. Kargin<sup>2</sup> Mykola V. Krachkovsky

<sup>1</sup> Donetsk National University, Ukraine  
Ukraine, 83000, Donetsk, Teatralnyi av., 13  
Dr. (Engineering), Professor  
E-mail: a.kargin@donnu.edu.ua

<sup>2</sup> Donetsk National University, Ukraine  
Ukraine, 83000, Donetsk, Teatralnyi av., 13  
Lecturer  
E-mail: w01dnick@gmail.com

**Abstract.** The article considers the problem of computer-aided instruction in context-chain motivated situation control system of the complex technical system behavior. The conceptual and formal models of situation control with practical instruction are considered. Acquisition of new behavior knowledge is presented as structural changes in system memory in the form of situational agent set. Model and method of computer-aided instruction represent formalization, based on the nondistinct theories by physiologists and cognitive psychologists.

The formal instruction model describes situation and reaction formation and dependence on different parameters, effecting education, such as the reinforcement value, time between the stimulus, action and the reinforcement. The change of the contextual link between situational elements when using is formalized.

The examples and results of computer instruction experiments of the robot device “LEGO MINDSTORMS NXT”, equipped with ultrasonic distance, touch, light sensors.

**Keywords:** situational control; computer-aided instruction; reinforcement learning; cognitive psychology; physiology; context; motive.

**Введение.** Адаптация к изменяющимся условиям поведения сложных робототехнических комплексов (РТК) путём самообучения связана со снижением производственных затрат, а также устранением человека из вредной среды, когда непосредственное управление человеком или постройка алгоритмов и оборудования затруднены. Требуемое поведение для разных условий запрограммировать заранее затруднительно либо по причине отсутствия полной информации на предварительном этапе (функционирование в открытой среде), либо трудоёмко. Для управления такими комплексами применяются методы ситуационного управления [1–3]. Одним из направлений развития теории ситуационного управления является модель и подход контекстно-цепочного мотивированного ситуационного управления (КЦМСУ), представляющая собой модификацию системы ситуационного управления [4, 5] на основе данных когнитивной психологии [6, 7]. Модель заимствует из когнитивной психологии организацию памяти: в базе хранится не набор прототипов «ситуация-действие», характерный для классических систем ситуационного управления, а прототип последовательностей действий, связанных контекстом. На концептуальном уровне модель КЦМСУ базируется на односторонней зависимости правил.

$$P_i : \text{ЕСЛИ } \{cont_{j,i}, S \subset \hat{S}_i, M\} \text{ ТО } \{u_i, cont_{i,l}\}, \quad (1)$$

где  $M$  — мотив,  $S$  — текущая ситуация,  $\hat{S}$  — эталонная ситуация-прототип,  $u$  — управляющее воздействие,  $cont_{j,i}$  — контекстная связь между правилами  $\Pi_j$  и  $\Pi_i$ .

Организацию контекстно-цепочного мотивированного ситуационного управления можно представить в виде, показанном на рис. 1. Она включает следующие компоненты: сенсорную память  $SN = \{sn_i\}_{i=1}^{ns}$ , множество ситуационных агентов  $CA = \{CA_j\}_{j=1}^{na}$ , множество эффекторов  $U = \{u_k\}_{k=1}^{nu}$ , образующих эффекторную память, множество мотивов  $M = \{m_l\}_{l=1}^{nm}$ . Поведение системы определяется взаимодействием агентов с окружением: ситуаций, формирующей значения нечётких характеристик сенсорных элементов и мотивов.

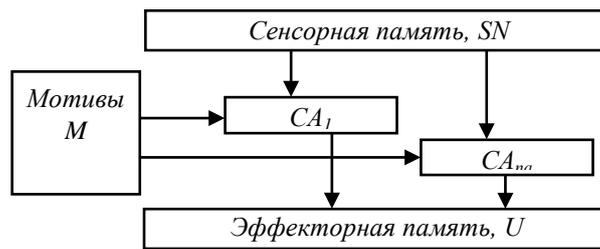


Рис. 1. Общая схема системы управления

Ситуационный агент представлен упорядоченной контекстом цепочкой ситуационных элементов  $\{ce_i\}_{i=0}^n$ . Структура ситуационного агента показана на рисунке 2. Модель ситуационного элемента есть (1).

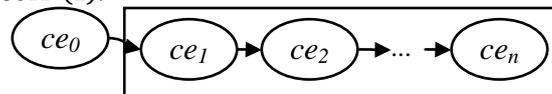


Рис. 2. Схематичное строение ситуационного агента

Знания о возможных вариантах поведения РТК при разных заданиях и состояниях (мотивах), в различных ситуациях представлены множеством ситуационных агентов СА. Для появления новой модели поведения, адаптированной к новой ситуации, требуется ввести в базу знаний новый СА<sub>j</sub> — цепочку связанных ситуационных элементов. Пополнение базы знаний путём обучения в ситуационных системах управления, как механизм адаптации к новым требованиям к поведению, повышает гибкость робототехнических комплексов, упрощает проектирование системы, что в конечном итоге, приводит к повышению их эффективности. Известные подходы к обучению, в основном, базируются на моделях искусственных нейронных сетей [8, 9], нейро-нечётких сетей [10], поведенческих сетей [11], развивающегося интеллекта [12]. В статье рассматривается обучение в КЦМСУ, как механизм формирования новых ситуационных агентов.

**Постановка задачи.** Управление рассматривается как многошаговый дискретный процесс обработки информации в моменты времени  $t, t + T, t + 2T, \dots, t + kT, \dots$ . Этапы обработки информации на одном шаге управления схематично показаны на рисунке 3: данные от датчиков на этапе предобработки фазифицируются и в сенсорной памяти представлены виде нечётких характеристик элементарных сенсоров; на основании близости текущей ситуации, представленной нечёткими характеристиками сенсорной памяти, и прототипов ситуации ситуационных агентов из прототипной памяти, формируются нечёткие характеристики состояния ситуационных элементов и прототипов реакций эффекторной памяти. На последнем этапе происходит дефазификация активированных прототипов реакции в непосредственную реакцию, которая подаётся на исполнительный механизм.

Научение системы новому поведению сводится к модификации прототипной памяти путём: 1) добавлению новых ситуационных элементов  $se_i$  и 2) установлению новых контекстных связей между существующими ситуационными элементами (ситуационные агенты). Отдельный ситуационный агент представляет некоторое отдельное законченное действие — фрагмент поведения.

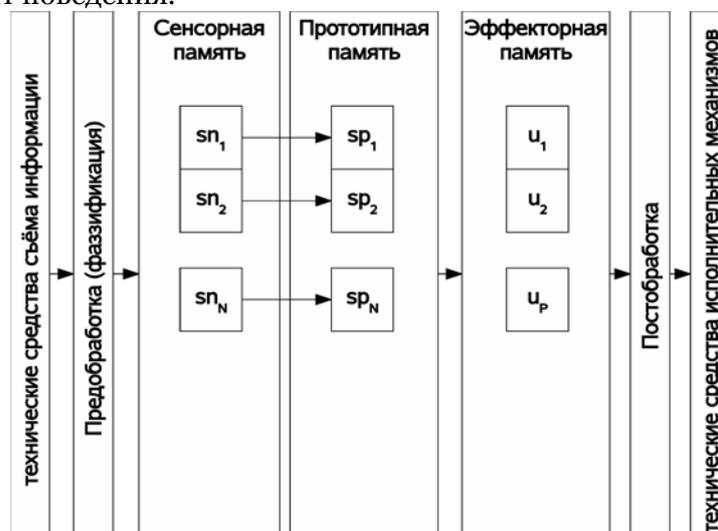


Рис. 3. Этапы шага управления

Обучение в вышеупомянутом смысле осуществляется на каждом шаге параллельно с управлением в оперативном режиме [3]. В качестве исходного метода обучения, который развивается применительно к рассматриваемому классу систем, применяется обучение с подкреплением [13].

В статье излагается модель и метод обучения с подкреплением для оперативного режима формирования новых ситуационных агентов из нескольких ситуационных элементов, связанных в контекстную цепочку. Модификация метода обучения с подкреплением базируется на обобщении теорий научения (Э. Торндайка, Б. Скиннера, И. Павлова [14]), изученных в когнитивной психологии.

**Формальная модель управления.** Формализм рассматриваемой системы базируется на моделях [4, 5], адаптированных к структуре системы, показанной на рис. 1. Сенсорный элемент  $sn_i$  формализован моделью элементарного свойства [3], и описывается нечёткой активностью:

$$A_i^{sn}(kT) = \left\{ x \left| \mu(x) = e^{-\frac{(x-\alpha)^2}{2\beta^2}}, x \in [-1, +1] \right. \right\}, \quad (2)$$

где  $\alpha$  – уровень активности  $\alpha \in [-1, +1]$ ,  $\beta$  – параметр, характеризующий актуальность информации,  $\beta \geq 0$ .

На этапе предобработки, показанной на рисунке 3, формируются значения  $\alpha, \beta$  в (2).

Модели их формирования для различных типов данных (сенсоров) могут различаться. Здесь примем:

$$\alpha_{A_i^{sn}(kT)} = F_1\left(I_{A_i^{sn}(kT)}\right), \beta_{A_i^{sn}(kT)} = F_2\left(\beta_{A_i^{sn}((k-1)T)}, \Delta I_{A_i^{sn}(kT)}\right), \quad (3)$$

где  $I_{\tilde{A}_i^{sn}(kT)}$  – показания сенсора  $sn_i$  в момент времени  $kT$ ,

$$\Delta I_{\tilde{A}_i^{sn}(kT)} = I_{\tilde{A}_i^{sn}(kT)} - I_{\tilde{A}_i^{sn}((k-1)T)}.$$

Сенсорная память на каждый момент времени  $kT$  моделирует ситуацию в виде (4).

$$S(kT) = \left\{ A_{\tilde{i}}^{sn}(kT), \Delta A_{\tilde{i}}^{sn}(kT) \right\}, \quad (4)$$

где  $A_{\tilde{i}}^{sn}(kT)$ ,  $\Delta A_{\tilde{i}}^{sn}(kT)$  – нечёткие множества-характеристики сенсора  $sn_i$  (активности и её изменения), в момент времени  $kT$ .

Ситуационный элемент  $ce$  контекстной цепочки (рис. 2), характеризуется:

- 1) нечётким прототипом ситуации –  $\hat{S}$ ;
- 2) нечётким прототипом управления –  $\hat{R}$ ;
- 3) контекстной связью –  $K$ ;
- 4) мотивированной связью –  $M$ .

Каждая из данных характеристик представляет собой множество нечётких характеристик [3] вида (2):

$$\hat{A}(kT) = \left\{ A_{\tilde{i}}^a(kT) \right\}, \quad (5)$$

В момент времени  $kT$  в результате обработки информации в прототипной памяти находятся нечёткие характеристики:

$$A_{\tilde{i},j,h}^{ce}(kT) = f_1 \left( A_{\tilde{i}}^M(kT), \rho(S(kT), \hat{S}_{i,j,h}), A_{\tilde{i},j,h}^{co}(kT) \right), \quad (6)$$

где  $\rho(S(kT), \hat{S}_{i,j,h})$  – нечёткая оценка близости текущей ситуации  $S_{kT}$  к прототипу  $\hat{S}_{i,j,h}$ ,  $A_{\tilde{i}}^M(kT)$  – нечёткая характеристика активности мотива,  $A_{\tilde{i},j,h}^{co}(kT)$  – нечёткая характеристика активности контекстной связи элемента  $ce_{i,j,h}$ .

Параметр  $\alpha_{\rho(S(kT), \hat{S})}$  нечёткой оценки близости  $\rho(S(kT), \hat{S}_{i,j,h})$  задаётся в виде

$$\alpha_{\rho(S(kT), \hat{S})} = \frac{\sum_{z=1}^N \left( \gamma_{\tilde{A}_z^{sn}(kT)} \cdot \gamma_{\tilde{A}_z^{sp}(kT)} \cdot \left( 1 - \left| \alpha_{\tilde{A}_z^{sn}(kT)} - \alpha_{\tilde{A}_z^{sp}(kT)} \right| \right) \cdot e^{-\beta_{\tilde{A}_z^{sp}(kT)}} \right)}{\sum_{z=1}^N e^{-\beta_{\tilde{A}_z^{sp}(kT)}}}, \quad (7)$$

где  $\alpha_{\tilde{A}_z^{sn}(kT)}$  – параметр  $\alpha$  нечёткой активности сенсора  $sn_z$  в текущей ситуации,

$\alpha_{\tilde{A}_z^{sp}(kT)}$  – параметр  $\alpha$  нечёткой активности признака  $sp_z$  в прототипе  $\hat{S}$ ,  $\beta_{\tilde{A}_z^{sp}(kT)}$  –

актуальность признака  $sp_z$  в прототипе  $\hat{S}$ .

Эффекторная память содержит множество  $U$  элементарных управляющих реакций – эффекторов. Модель эффектора включает модели управляющей и исполнительной компонент. Исполнительная включает непосредственные изменения значений сигналов по

нечёткой характеристике. Управляющая обрабатывает информацию о накоплении нечёткой активности эффектора от разных ситуационных агентов.

Нечёткая характеристика активности эффектора вычисляется на основании нечётких характеристик состояния активности всех ситуационных элементов и их прототипов реакции  $A_{i,j,h,x}^r(kT)$ .

$$A_x^u(kT) = \bigoplus_{ce_{i,j,h}} \left( A_{i,j,h}^{ce}(kT) \otimes A_{i,j,h,x}^r(kT) \right), x \in \overline{1, p}, \quad (8)$$

где  $\oplus$  – операция нечёткого накопления,  $\otimes$  – нечёткое произведение,  $ce_{i,j,h}$  –  $h$ -й ситуационный элемент  $j$ -го агента, относящегося к мотиву  $i$ .

**Концептуальная модель обучения.** Из 4-х концептуальных моделей обучения в контекстно-цепочной мотивированной ситуационной системе управления [15] в статье рассматривается модель обучения, сводящаяся к формированию нового ситуационного агента. Создание нового ситуационного агента рассматривается как многоэтапный процесс, на каждом этапе которого формируется отдельный ситуационный элемент путём нахождения характеристик: прототипов ситуации ( $\hat{S}$ ) и управления ( $\hat{R}$ ); нечётких характеристик мотива ( $M$ ) и контекстной связи ( $K$ ).

Первый шаг каждого этапа обучения начинается с формирования характеристик выделенного не специфицированного „пустого“ ситуационного элемента, который будет служить базой для образования нового элемента. «Пустой» элемент имеет потенциальные связи со всеми существующими компонентами: контекстные с агентами; информационные с сенсорами, сигнальные с управлением и связи с мотивами. Изначально этим связям присваиваются нейтральные значения нечётких характеристик. Структура пустого ситуационного элемента с его связями показана на рис. 4.

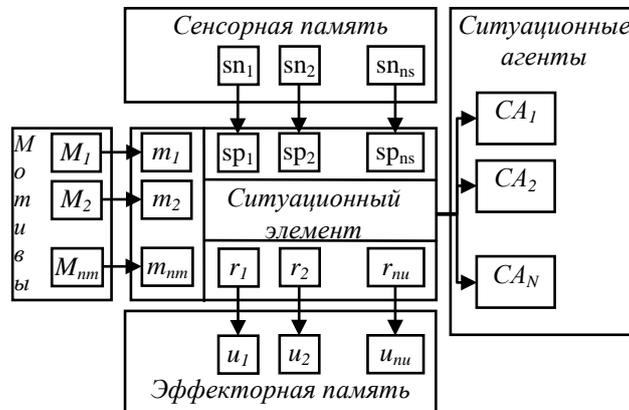


Рис. 4. Ситуационный элемент

На последующих  $i$ -х шагах обучение происходит на основе подкрепления: произошло изменение (падение) нечёткой характеристики активности мотива. Нечёткие характеристики ситуационного элемента в моменты времени  $kT$  находятся по модели обучения  $F$  на основании значений этих характеристик в предыдущий момент времени, векторов нечётких характеристик активности сенсоров ( $MS$ ), действий ( $MR$ ) и мотива ( $MM$ ).

$$\langle \hat{S}, \hat{R}, M, K \rangle_{kT} = F \left( \langle \hat{S}, \hat{R}, M, K \rangle_{(k-1)T}, MS, MR, MM \right), \quad (9)$$

где  $MS = \langle S((k-j)T) \rangle_{j=0}^d$ ,  $MR = \langle R((k-1)T) \rangle_{j=0}^d$ ,  $MM = \langle M((k-j)T) \rangle_{j=0}^d$ ,

$$S(kT) = \left\{ A_i^{sn}(kT) \right\}_{i=1}^{ns}, \quad R(kT) = \left\{ A_i^u(kT) \right\}_{i=1}^{nr}, \quad M(kT) = \left\{ A_i^m(kT) \right\}_{i=1}^{nm},$$

$d$  – глубина памяти.

Ниже рассматривается формализация процедуры обучения (9).

Обобщения известных теорий научения из физиологии и когнитивной психологии [14], таких как теорий Э. Л. Торндайка, К.Л. Халла, Э. Ч. Толмена, А. Бандуры и других, позволили сформировать и выделить следующие модели обучения, применимые к контекстно-цепочной мотивированной ситуационной системе управления:

1. Модель образования (изменения) контекстной связи между ситуационными агентами (агентами).

Это происходит в случае, когда ситуация, которая появилась как следствие выполнения функции  $СЭ_i$ , сопоставима с прототипом, активизирующим некоторый другой, например, агент  $СЭ_j$ . Многократное повторение такой последовательности с последующим подкреплением (ослабление мотива) приводит к усилению контекстной связи  $СЭ_i \rightarrow СЭ_j$  и в дальнейшем даже при значительном отклонении ситуации контекстная связь может обеспечить активацию ситуационного агента  $СЭ_j$ .

2. Модель образования прототипов нового ситуационного элемента.

Если в процессе реализации некоторого управления появилась ситуация, подходящая под прототип уже ситуационного элемента агента  $СA_k$ , поведение согласно которому привело к изменению мотива, то формируются на нечёткие характеристики нового ситуационного элемента. Многократное повторение такой последовательности с последующим подкреплением (ослабление мотива) приводит к образованию прототипов ситуации и управления.

На рис. 5 приведено схематическое объяснение приведённых моделей. На рис. 5.а показано изменение контекстной связи (сплошная стрелка) между ситуационными агентами  $СA_i$  и  $СA_j$ . При этом связь формируется односторонняя — в том же порядке, в котором происходит обработка информации. На рис. 5.б показано формирование (модификация) ситуационного агента  $СA_k$  путём включения в контекстную последовательность нового ситуационного элемента. На рис. 5.в приведён новый ситуационный агент  $СA_z$ , сформированный на базе одного элемента.

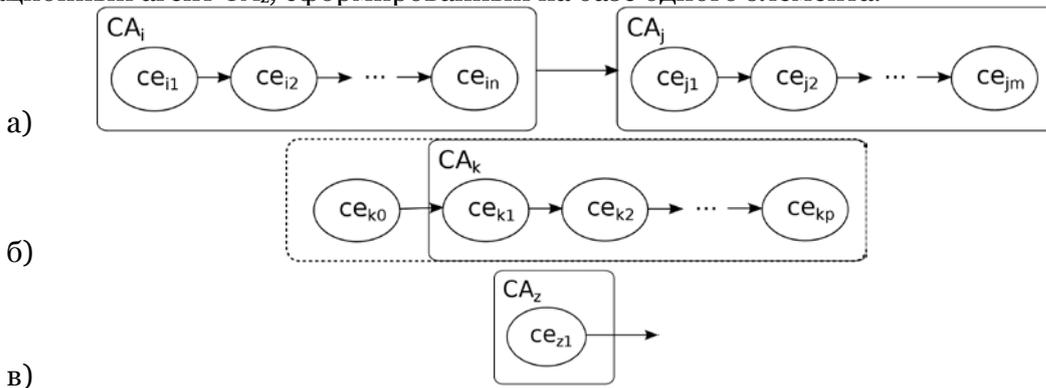


Рис. 5. Изменения структуры системы, вносимые обучением

В первом варианте механизм обучения формирует нечёткое множество контекстной связи  $CO_{i,n,j,1}$ , которая влияет на активность суммарного контекстного входа  $A_{i,1}^{co}((k-1)T)$ , входящую в модель управления (6) при расчёте активности ситуационного элемента  $ce_{j,1}$ .

Во втором варианте механизм обучения формирует нечёткие множества прототипов ситуации  $\hat{S}_{k,0}$ , и реакции  $\hat{R}_{k,0}$ . Также формируется контекстная связь  $CO_{k,0,k,1}$ .

В третьем варианте, как и во втором, происходит формирование нечётких множеств прототипа ситуации  $\hat{S}_{z,1}$  и реакции  $\hat{R}_{z,1}$ . Вместо контекстной связи формируется связь с мотивом. Данная связь используется при управлении для выбора подходящего агента, в случае наличия альтернатив, а также при дальнейшем обучении.

Значения всех сформированных нечётких множеств в каждом из вариантов 1–3 зависят от времени, прошедшего между реализацией управляющего воздействия и изменением активности мотивов.

**Формальная модель обучения.** Формирование прототипов ситуации и реакции нового ситуационного элемента сводится к определению актуальных признаков из всего множества признаков, представленных сенсорами в  $\{sp_i\}$ , и актуальных управлений из множества возможных  $\{u_x\}$ . Актуальность формально задаётся параметром  $\beta$  нечёткой характеристики (2): при значениях  $\beta$  близких к 0, актуальность информации максимальна, при  $\beta \geq 10 \div 12$  — актуальность оценивается числом равным нулю [16]. Поэтому, на каждом шаге обучения модифицируется значение  $\beta$  для всех признаков и управлений нового СЭ: для тех признаков, которые присутствовали на предыдущем шаге научения.

По данным когнитивной психологии [14] процесс научения происходит с различной эффективностью, которая определяется такими параметрами, как время между предъявлением стимула, совершённой реакцией и полученным подкреплением. Предлагается эту зависимость представить в виде (10).

$$I_m(t) = \frac{t}{m} \cdot e^{1 - \frac{t}{m}}, t \geq 0, \quad (10)$$

где  $t$  — время от предъявления стимула до подкрепления, для  $t < 0$  можно считать значение равным 0;

$m$  — параметр, задающий значение оптимального времени.

Параметр времени обучения (10) для реакции должен быть меньше, чем параметр для ситуации, так как реакция выполняется с некоторой задержкой после предъявления стимула.

Вторым фактором, влияющим на эффективность научения, является величина подкрепления — явились ли последствия действия полезными для объекта. Формализация представлена ниже (11).

$$\alpha_{\tilde{Q}_j^M(kT)} = o\left(\alpha_{\tilde{A}^M((k-j)T)} - \alpha_{\tilde{A}^M(kT)}\right), \beta_{\tilde{Q}_j^M(kT)} = \max\left(\beta_{\tilde{A}^M((k-j)T)}, \beta_{\tilde{A}^M(kT)}\right), \quad (11)$$

$$o(x) = \frac{x}{2} \cdot \left|\frac{x}{2}\right|^\gamma, -1 < \gamma \leq 0,$$

где  $j$  — число тактов времени, за которое оценивается подкрепление,  $\gamma$  — параметр, влияющий на эффективность обучения при малых изменениях мотива.

Прототип формально представлен нечёткими характеристиками признаков сенсорной памяти или управлений эффекторной памяти. При формировании прототипа на момент времени  $kT$  рассматривается последовательность за время  $(k-j)T$  описаний ситуаций и формирующихся моделей прототипов (потенциальных прототипов)  $\tilde{A}'_i((k-j)T)$

$$\alpha_{\tilde{A}'_i((k-j)T)} = (1-q)\alpha_{\tilde{A}'_i^{sp}((k-1)T)} + q \cdot \alpha_{\tilde{A}'_i^{sn}((k-j)T)}, \quad (12)$$

$$\beta_{\tilde{A}'_i((k-j)T)} = (1 - q)\beta_{\tilde{A}'_i^{sp}((k-1)T)} + q \cdot \beta_2,$$

где

$$\beta_2 = \frac{4\beta_{\tilde{A}'_i^{sn}((k-j)T)}}{I_{m_S}(jT) \left( \alpha_{\tilde{Q}_j^{\bar{M}}(kT)} + 1 \right) \left( \alpha_{\tilde{A}'_i^{\Delta sn}((k-j)T)} + 1 \right)} + \frac{\beta_{\tilde{Q}_j^{\bar{M}}(kT)} + \beta_{\tilde{A}'_i^{\Delta sn}((k-j)T)}}{I_{m_S}(jT)};$$

$q = \varphi \cdot e^{-\beta_2}$ ;  $\varphi$  – параметр скорости обучения;

$\tilde{A}'_i^{\Delta sn}((k-j)T)$  – нечёткая характеристика скорости изменения сенсора  $sn_i$  в момент времени  $(k-j)T$ ;

$\tilde{Q}_j^{\bar{M}}(kT)$  – нечёткая характеристика подкрепления.

Из набора потенциальных прототипов выбирается один из них, который обладает максимальной актуальностью, что представлено.

$$\tilde{A}'_i^{sp}(kT) = \tilde{A}'_i((k-g)T) \Big|_{g \in [0, d]}, \beta_{\tilde{A}'_i((k-g)T)} = \min_{j=0, d} \beta_{\tilde{A}'_i((k-j)T)}, \quad (13)$$

где  $\tilde{A}'_i((k-j)T)$  – расчётная нечёткая характеристика элемента  $sp_i$ , модифицированная с учётом влияния эффективности обучения и величины подкрепления относительно момента времени;  $d$  – глубина сенсорной памяти.

При формировании контекстной связи (14) между ситуационными элементами  $ce_h$  и  $ce_g$  величина нечёткой характеристики этой связи модифицируется в соответствии с полученным подкреплением (11).

$$\alpha_{\tilde{A}'_{g,h}^{co}(kT)} = (1 - \varphi)\alpha_{\tilde{A}'_{g,h}^{co}((k-1)T)} + \varphi \cdot \alpha_{\tilde{Q}_j^{\bar{M}}(kT)}. \quad (14)$$

**Результаты.** Компьютерный эксперимент проводился на комплексе, состоящем из робота LEGO MINDSTORMS NXT, который удалённо управляется компьютером при помощи Bluetooth. Робот представлен на рисунке 6, имеет два независимо управляемых колеса. Третье колесо пассивное и обеспечивает устойчивость. Из датчиков робота использованы датчик соприкосновения и расстояния, оба направлены вперёд. Целью эксперимента было обучение робота поведению, которое позволяет избежать столкновения с препятствием.

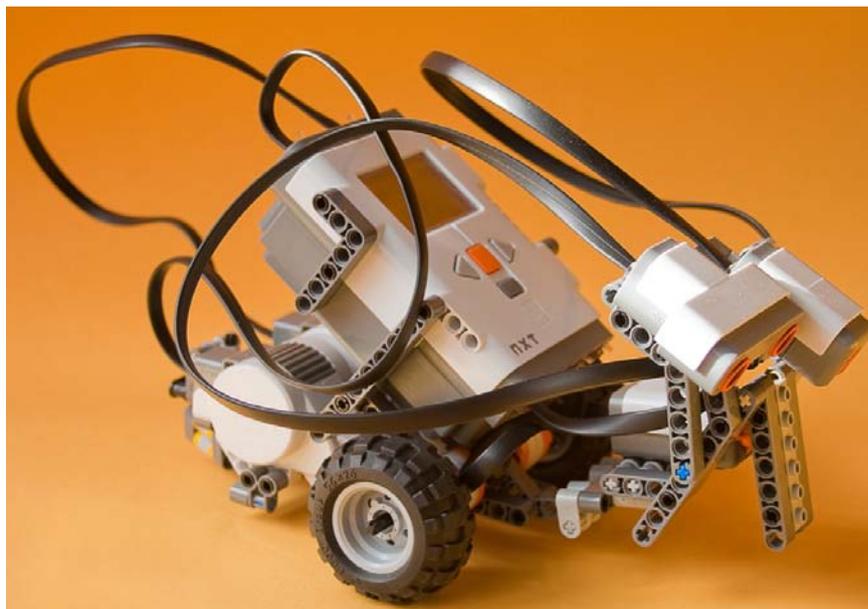


Рис. 6. Внешний вид робота LEGO MINDSTORMS NXT

В качестве примера обучения рассмотрим ситуацию, когда в прототипной памяти хранятся знания о стереотипах поведения: 1) при столкновении с препятствием отъехать назад, чтобы не повредить робота; 2) при низком заряде повернуться в сторону источника света и приблизиться к нему; 3) при высоком заряде уйти из освещённого места — в виде множества СА ( $1 - CA_1, 2 - CA_2-CA_5, 1 - CA_6-CA_8,$ ). Однако этих знаний недостаточно для того, чтобы избежать столкновения с препятствием, как показано на рис. Обучение избеганию столкновения сводится к формированию ситуационного агента, который управляет поворотом робота при приближении к препятствию, как показано на рис. 7.

Информация от двух датчиков гранулирована так, что по показаниям датчиков формируются нечёткие характеристики 28 элементарных сенсоров: 2 сенсора для датчика соприкосновения ( $snt_0, snt_1$ ), 20 для датчика расстояния для разного уровня детализации (от 2 до 6 сенсоров на область детектирования датчика:  $snd_{i,j}, i = \overline{1,5}, j = \overline{0,i}$ ), а также по 3 сенсора на каждое колесо ( $snr_{i,j}, i = \overline{0,1}, j = \overline{-1,1}$ ). Подробнее датчики описаны в [17].

В качестве мотива выбран мотив самосохранения робота, который основан на  $snt_1$ : если сработал датчик, то есть угроза столкновения. В случае столкновения робота с препятствием, обучение избеганию будет происходить в два этапа: формирование знания о столкновении и собственно обучение избеганию. В случае обучения сложному поведению эти этапы будут повторяться.

Обучение было выполнено по следующей схеме. Робот движется по прямой к стенке со средней скоростью. При столкновении со стенкой возрастает активность мотива самосохранения. Ситуация соответствует прототипу первоначально созданного элемента  $СЭ_1$ , выдаёт управление в соответствии с прототипом реакции — робот останавливается и отъезжает от стенки. До обучения траектория движения робота показана на рис. 7.а и представляет собой горизонтальную линию. Вертикальной линией показана стена.

В момент активизации мотива самосохранения происходит обучение упреждающей ситуации — формируется прототип нового элемента  $СЭ_2$ , описывающий малое расстояние до препятствия и движение вперёд. В процессе компьютерного эксперимента на основании нескольких столкновений данный прототип закрепляется и начинает активизировать как мотив самосохранения, так и состояние  $СЭ_2$ . Поскольку связи с управлениями в  $СЭ_2$  на

этот момент ещё не сформировались (нет прототипа управления) активное состояние  $CЭ_2$  случайным образом формирует активность какого-то управления, которое и реализуется. Это продолжается до тех пор, пока не появится случай, когда выработается управление поворот вправо, что при приближении робота к стене, позволит избежать столкновения, траектория показана на рис. 7.б. Мотив при этом деактивируется, поскольку изменится ситуация, соответствующая прототипу ситуации в  $CЭ_2$ . Падение активности мотива  $\bar{M}$  соответствует подкреплению. В этом случае происходит закрепление второго прототипа  $CЭ_2$ : формируется прототип реакции. На рис. 7.в показана траектория, выработанная вновь сформированным ситуационным элементом после многократного научения.

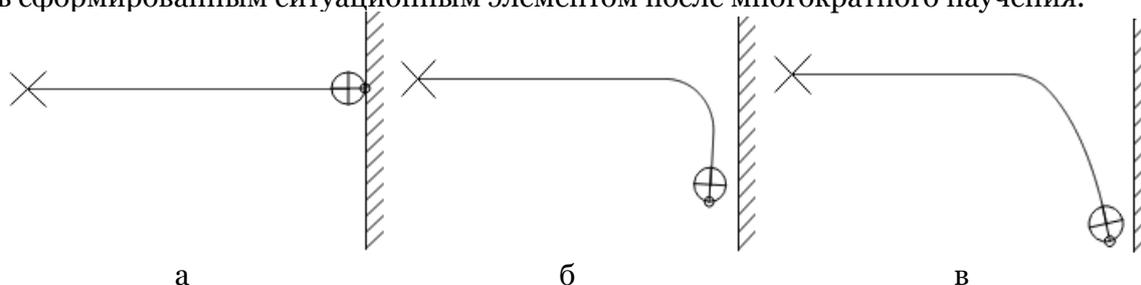


Рис. 7. Поведение робота при столкновении со стенкой: до обучения (а), эталонная реакция (б) и после обучения (в)

На следующем этапе аналогичным методом формируется третий ситуационный элемент  $CЭ_3$  агента, для случая, когда робот движется с высокой скоростью и не успевает повернуть. Данная ситуация представлена на рисунке 8.а.

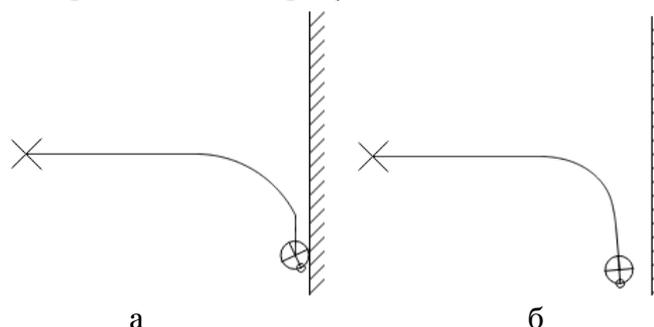


Рис. 8. Траектория движения при высокой начальной скорости

Обучение в данном случае было проведено аналогично ранее описанному, только на других сенсорах.

**Заключение.** Рассмотренный механизм обучения в КЦМСУ, основанный на теориях научения, исследованным в когнитивной психологии, формализован путём описания в виде нечётких конструкций таких когнитивных понятий как прототип ситуации и реакции, мотив и контекстная связь. Машинное обучение сведено к формированию в оперативном режиме новых структур в базе знаний (прототипной памяти) по методу, заимствованному из теорий научения когнитивных психологов и физиологов Э. Торндайка, Б. Скиннера, И. Павлова. Показано, что формализация указанного механизма научения возможна на базе специализированной модели ситуационного управления — контекстно-цепочной мотивированной ситуационной системы управления. Новая модель и подход ситуационного управления с оперативным машинным обучением проверены на серии экспериментов с реальным робототехническим комплексом. При дальнейшем функционировании системы может также происходить обучение, влияние которого отражается не на структуре системы, а на значении прототипов существующих элементов и контекстных связей между ситуационными элементами.

**Примечания:**

1. Поспелов Д. А. Ситуационное управление: Теория и практика [текст] / Д. А. Поспелов. М.: Наука. – Гл. ред. физ.-мат. Лит., 1986. 288 с.
2. Мелихов А.Н. Ситуационные советующие системы с нечеткой логикой [текст] / Мелихов А.Н., Берштейн Л.Е., Коровин С.Д. М.: Наука, 1990.
3. Каргин А. А. Введение в интеллектуальные машины. Книга 1. Интеллектуальные регуляторы [текст] / А. А. Каргин. Донецк: Норд-Пресс, ДонНУ, 2010. 526 с.
4. Каргин А.А. Модели динамических ситуационных интеллектуальных машин / А.А. Каргин, Т.Г. Петренко // Искусственный интеллект. 1999. №2. С. 41-47.
5. Каргин А.А. Модель организации памяти в системах, построенных по технологии ситуационного логогена / А.А. Каргин, Т.Г. Петренко. Наукові праці Донецького національного технічного університету. Серія «Обчислювальна техніка та автоматизація». Випуск 12(118): Донецьк-ДонНТУ, 2007, С.120-124.
6. Солсо Р. Когнитивная психология [текст] / Р. Солсо. СПб.: Питер, 2002. 592 с.
7. Андерсон Дж. Р. Когнитивная психология [текст] / Дж. Р. Андерсон. СПб.: Питер, 2002. 496 с.
8. Терехов В.А. Нейросетевые системы управления [текст] / В.А. Терехов, Д.В. Ефимов, И.Ю. Тюкин. М.: Высш. шк., 2002.
9. Tan A.-H. Intelligence through interaction: towards a unified theory for learning [текст] / A.-H. Tan, G.A. Carpenter, S. Grossberg – Advances in neural networks. N. 1., 2007. P. 1094–1103.
10. Lin C.-T. Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems / Chin-Teng Lin, C. S. George Lee – Upper Saddle River, NJ: Prentice Hall – 1996. – 797p.
11. Maes P. Learning to Coordinate Behaviors [текст] / P. Maes, P. Brooks – AAAI Press/MIT Press – Proceedings of the Eighth National Conference on Artificial Intelligence, 1990, P.796-802.
12. Meng Y. Bio-Inspired Self-Organizing Robotic Systems [текст] / Yan Meng, Yaochu Jin. Springer-Verlag Berlin Heidelberg. 2011. 273p. ISBN 978-3-642-20759-4.
13. Саттон Р. Обучение с подкреплением [текст] / Р. Саттон, Э. Барто. СПб.: Бином, 2011, 399 с.
14. Хегенхан Б. Теории научения [текст] / Б. Хегенхан, М. Олсон. ; Перевод на русский язык ЗАО Издательский дом «Питер». – [6-е издание]. СПб. : Питер, 2004. 474 с. : ил. (Серия «Мастера психологии»). – ISBN 5-94723-033-X.
15. Каргин А. А. Модели обучения системы мотивированного контекстного ситуационного управления [текст] / А. А. Каргин, Крачковский Н. В. Херсон: Вісник ХНТУ. 2012.-№1(44). С.257–260.
16. Каргин А. А. Модель ситуационного управления роботом, учитывающая актуальность сенсорной информации [Текст] / А.А. Каргин // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту: Матеріали наук. міжн. конференції. Херсон: ХНТУ, 2013. С. 155–157.
17. Каргин А. А. Об одной модели ситуационного управления подвижным роботом [текст] / А. А. Каргин, Н.В. Крачковский. Харків: Науково-технічний журнал «Інформаційно-керуючі системи на залізничному транспорті». 2011. №4(89). С. 12-17

**References:**

1. Pospelov D. A. Situatsionnoe upravlenie: Teoriya i praktika / D.A. Pospelov. M.: Nauka. Gl. red. fiz.-mat. Lit., 1986. 288 s.
2. Melikhov A.N. Situatsionnye sovetuyushchie sistemy s nechetkoi logikoi / Melikhov A.N., Bershtein L.E., Korovin S.D. M.: Nauka, 1990.
3. Kargin A. A. Vvedenie v intellektual'nye mashiny. Kniga 1. Intellektual'nye regulatory / A.A. Kargin. Donetsk: Nord-Press, DonNU, 2010. 526 s.
4. Kargin A.A. Modeli dinamicheskikh situatsionnykh intellektual'nykh mashin / A.A. Kargin, T.G. Petrenko // Iskusstvennyi intellekt. 1999. №2. S. 41-47.
5. Kargin A.A. Model' organizatsii pamyati v sistemakh, postroennykh po tekhnologii situatsionnogo logogena / A.A. Kargin, T.G. Petrenko. – Naukovi pratsi Donets'kogo natsional'nogo tekhnichnogo univ. Seriya "Obchislyval'na tekhnika ta avtomatizatsiya". Vipusk 12(118): Donets'k-DonNTU, 2007, S.120-124.
6. Solso R. Kognitivnaya psikhologiya / R. Solso. SPb.: Piter, 2002. 592 s.
7. Anderson Dzh. R. Kognitivnaya psikhologiya / Dzh. R. Anderson. SPb.: Piter, 2002. 496 s.
8. Terekhov V.A. Neirosetevye sistemy upravleniya / V.A. Terekhov, D.V. Efimov, I.Yu. Tyukin. M.: Vyssh. shk., 2002.
9. Tan A.-H. Intelligence through interaction: towards a unified theory for learning / A.-H.Tan, G.A. Carpenter, S. Grossberg. Advances in neural networks. N. 1. 2007. P. 1094–1103.
10. Lin C.-T. Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems / Chin-Teng Lin, C. S. George Lee – Upper Saddle River, NJ: Prentice Hall/ 1996. 797 p.
11. Maes R. Learning to Coordinate Behaviors / R. Maes, P. Brooks – AAAI Press/MIT Press – Proceedings of the Eighth National Conference on Artificial Intelligence, 1990, P.796-802.

12. Meng Y. Bio-Inspired Self-Organizing Robotic Systems / Yan Meng, Yaochu Jin – Springer-Verlag Berlin Heidelberg / 2011. 273 p. – ISBN 978-3-642-20759-4.
13. Sattou R. Obuchenie s podkrepleniem / R. Sattou, E. Barto. SPb.: Binom, 2011. 399 s.
14. Khegenkhan B. Teorii naucheniya / B. Khegenkhan, M. Olson. ; Perevod na russkii yazyk ZAO Izdatel'skii dom «Piter». – [6-e izdanie]. SPb. : Piter, 2004. 474 s. : il. – (Seriya «Mastera psikhologii»). – ISBN 5-94723-033-X.
15. Kargin A.A. Modeli obucheniya sistemy motivirovannogo kontekstnogo situatsionnogo upravleniya / A.A. Kargin, N.V. Krachkovskii. Kherson: Visnik KhNTU. 2012.-№1(44).-S. 257–260
16. Kargin A.A. Model' situatsionnogo upravleniya robotom, uchityvayushchaya aktual'nost' sensornoi informatsii / A.A. Kargin // Intelektual'ni sistemi priinyattya rishen' i problemi obchislyval'nogo intelektu: Materiali nauk. mizhn. konferentsii. Kherson: KhNTU, 2013. S. 155–157.
17. Kargin A.A. Ob odnoi modeli situatsionnogo upravleniya podvizhnym robotom / A.A. Kargin, N.V. Krachkovskii. Kharkiv: Naukovo-tekhnichnii zhurnal «Informatsiino-keruyuchi sistemi na zaliznichnomu transporti». 2011. №4(89).-S. 12-17.

УДК 004.85 : 004.81

### Метод машинного обучения в системах ситуационного управления

<sup>1</sup> Анатолий А. Каргин

<sup>2</sup> Николай В. Крачковский

<sup>1</sup> Донецкий национальный университет, Украина  
Украина, 83000, г. Донецк, пр. Театральный, 13  
Доктор технических наук, профессор  
E-mail: a.kargin@donnu.edu.ua

<sup>2</sup> Донецкий национальный университет, Украина  
Украина, 83000, г. Донецк, пр. Театральный, 13  
Старший преподаватель  
E-mail: w01dnick@gmail.com

**Аннотация.** Рассматривается задача машинного обучения в системах контекстно-цепочного мотивированного ситуационного управления поведением сложного технического комплекса. Рассмотрены концептуальная и формальная модели ситуационного управления с оперативным обучением. Приобретение знаний о новом поведении представлено структурными изменениями в памяти системы в виде множества ситуационных агентов. Модель и метод машинного обучения представляют формализацию на базе нечётких конструкций теорий научения физиологов и когнитивных психологов.

Формальная модель обучения описывает формирование прототипов ситуации и реакции и зависимость от различных параметров, влияющих на обучение, таких как величина подкрепления, время между стимулом, действием и подкреплением. Дается формализация контекстной связи между ситуационными элементами при её использовании.

Приведены примеры и результаты компьютерных экспериментов обучения робота типа «LEGO MINDSTORMS NXT», оснащённого сенсорами: ультразвуковой расстояния, соприкосновения, освещённости.

**Ключевые слова:** ситуационное управление; машинное обучение; обучение с подкреплением; когнитивная психология; физиология; контекст; мотив.