

$$b_{i,j} \frac{P_{i,j}^{n+1} - P_{i,j}^{n+1/2}}{\tau/2} = k_{i+1/2} \frac{P_{i+1,j}^{n+1/2} - P_{i,j}^{n+1/2}}{h^2} - k_{i-1/2,j} \frac{P_{i,j}^{n+1/2} - P_{i-1,j}^{n+1/2}}{h^2} +$$

$$+ k_{i,j+1/2} \frac{P_{i,j+1}^{n+1} - P_{i,j}^{n+1}}{h^2} - k_{i,j-1/2} \frac{P_{i,j}^{n+1} - P_{i,j-1}^{n+1}}{h^2} + f_{i,j}^{n+1/2};$$

при  $j = 0$   $P_{i,-1}^{n+1} = P_{i,1}^{n+1}$  ( $i = \overline{1, M-1}$ );

при  $i = N$   $P_{i,N}^{n+1} = \psi_i^{n+1}$  ( $i = \overline{1, M-1}$ ).

Поскольку на каждом полушаге задача оказывается фактически одномерной (неявной), то для ее решения можно использовать метод прогонки. Метод прогонки удобен тем, что требует относительно небольших объемов оперативной памяти и затрат времени на проведение расчетов.

Решив системы дважды, в результате получим решение на очередном шаге  $t = t_{n+1} t$ .

### Библиографический список:

1. Максимов М. М., Рыбицкая Л. П. Математическое моделирование процессов разработки нефтяных месторождений.-М.: Недра, 1976. 264 с.
2. Басниев К. С, Кочина И. Н., Максимов В. М. Подземная гидромеханика: Учебник для вузов.-М.: Недра, 1993. 416 с.

УДК 62-50:531.3

*Рамазанов Г.М.*

### РАЗРАБОТКА АЛГОРИТМА САМООБУЧЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПРИ НАЛИЧИИ В СРЕДЕ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ

*Ramazanov G.M.*

### DEVELOPMENT OF THE ALGORITHM OF SELF-TRAINING THE INTELLECTUAL SYSTEMS AT PRESENCE IN AMBIENCE OF THE CAUSAL RELATIONSHIPS

*Предложен и исследован алгоритм самообучения интеллектуальных систем в априори неопределенных проблемных средах при наличии в них причинно-следственных связей между происходящими событиями.*

**Ключевые слова:** интеллектуальная система, проблемная среда, алгоритм самообучения.

*It is offered and explored algorithm of self-training the intellectual systems in a priori undeclared problem-solving ambience at presence in them causal relationships between occurring events.*

**Key words:** intellectual system, problem-solving ambience, algorithm of self-training.

Одной из актуальных проблем современной науки является разработка интеллектуальных систем (ИС) способных автономно функционировать в априори неопределенных проблемных средах (ПС). Эффективное решение данной проблемы, прежде всего, связано с разработкой алгоритмов самообучения (АС) позволяющих ИС выявлять различные закономерности целенаправленного преобразования ситуаций ПС.

Заметим, что при проявлении в ПС независимых от действий ИС событий, связанных с появлением в ней соответствующих им сигналов среду следует отнести к классу динамических сред. Иными словами, в такой ПС происходят независимые от интеллектуальной системы преобразования допустимых ситуаций, влияющих на процесс ее адаптации к изменяющимся условиям функционирования. При этом перед ИС встает проблема в процессе формирования программ целесообразного поведения (ПЦП) в динамической ПС для выявления причинно-следственных связей между происходящими в ПС событиями.

В этом случае, в процессе самообучения ИС при наличии в ПС причинно-следственных связей, обусловленных выполнением в ней различных по своему содержанию событий, возникает необходимость исключения образования в формируемых ПЦП случайных элементарных актов поведения следующего вида:  $S_{\text{тек}} \& v_j \rightarrow S_{\text{цел}}^1$ , закрепляющихся в ней в результате неверного определения причин изменения ситуаций среды.

Для этой цели при разработке алгоритма самообучения необходимо использовать механизмы действия многократного совпадения двух условных сигналов во времени. Под условным сигналом в этом случае будем понимать любое наблюдаемое ИС изменение текущей ситуации ПС, например, изменение состояния конкретного объекта, появление в среде нового объекта, изменение значений отношений складывающихся между объектами среды и т.д.

Приведем один из возможных принципов реализации механизма закрепления элементарных актов поведения ИС в формируемой ПЦП. Введем параметр  $\delta$ , принимающий бинарные значения  $[0,1]$  и определяющий степень уверенности ИС в истинности сформированного ею отдельного элементарного акта поведения. Считаем, что если параметр  $\delta$  для выявленного акта поведения в процессе многократной его реализации принимает значение, большее единицы, то данное правило отражает закономерное преобразование ПС. В противном случае, при  $\delta=0$  принимается решение о том, что элементарный акт поведения носит случайный характер. Для ПЦП, состоящей из  $k_2$  элементарных актов поведения, значение степени уверенности ИС  $\delta^*$  для эффективного применения этой программы поведения может определяться согласно выражению  $\delta^* = \min \delta_w, w = \overline{1, k_2}$ , где  $\delta_w$  – степень уверенности в истинности  $w$ -го элементарного акта поведения, входящего в структуру сформированной программы поведения

Если у ИС сформировала несколько альтернативных ПЦП  $L_{i44} \in L$ , приводящих к достижению определенной цели, определяемой сигналом  $s_{i31}^*$ , то предпочтение отдается наращиванию и реализации программы поведения с наибольшим значением параметра  $\delta^*$ . При наличии двух ПЦП с одинаковыми значениями параметра  $\delta^*$  используется программа поведения, имеющая меньшую сложность реализации.

Для реализации рассмотренного механизма многократного закрепления элементарных актов поведения в формируемых ПЦП, необходимо выбрать количество  $k_3$ - совпадений условных сигналов во времени, при котором принимается решение об истинности закрепляемого акта поведения. Значение  $k_3$  можно определить эвристическим путем, опираясь на накопленный опыт поведения, или на основе результата опроса экспертов. Зная  $k_3$  можно вычислить шаг приращения  $\Delta\delta$  параметра  $\delta_w$ , как  $\Delta\delta = 1/k_3$ . Отсюда, удачная отработка действия элементарного акта поведения поощряется увеличением параметра  $\delta$  на  $\Delta\delta$  т.е.  $\Delta\delta = \delta_w + \Delta\delta$ , в противном случае следует штраф и  $\Delta\delta = \delta_w - \Delta\delta$ .

Используя описанный механизм, ИС, наблюдая за происходящими в ПС событиями, принимает следующее решение: если после восприятия условного сигнала  $S_{meki}$  через интервал времени  $t$  появляется сигнал  $S_{i+1mek}$  и данное совпадение проявляется многократно, т.е. для него  $\delta=1$ , то сигнал  $S_{meki}$  рассматривается как предвестник появления условного сигнала, на который интеллектуальная система реагирует определенным образом. Далее,

если сигнал  $S_{теки}$  появляется часто, например, периодически по истечении интервала времени  $t_2$ , то изменений в характере поведения ИС не происходит. Например, робот, ожидает самопроизвольного появления требуемого условного сигнала  $S_{i+1тек}$  и реагирует на него определенным образом согласно сформированной ПЦП, в которой закреплён этот сигнал.

В противном случае ИС принимает решение о необходимости достижения одного из условных сигналов  $S_{теки}$  или  $S_{i+1тек}$  путем отработки различных действий  $b_{j1} \in B$ . Последовательная цепочка действий, приводящих к появлению условных сигналов  $S_{теки}$  и  $S_{i+1тек}$  в среде, ИС формирует на основании алгоритма самообучения 1, вырабатывая на его основе соответствующие ПЦП.

Алгоритм самообучения 1.

**Исходные данные:** ситуации  $S_{исх}$ ,  $S_{цел}$ , множество действий  $B$ ; заданная цель,

**Входные переменные:** воспринимаемые в ПС ситуации  $S_{итек} \in S$

**Выходные переменные:** сформированная ПЦП:  $S_{тек} \& b_j \rightarrow S^1_{теки} \& b_{j+1} \rightarrow \dots \rightarrow S_{теки+n} \& b_{jn} \rightarrow \dots \rightarrow$  заданная цель .

1. Начало.
2. Описать текущую ситуацию внешней среды  $S_{цел i}$ .
3. Сформировать множество действий  $b_i \in B$ , которые можно отработать в ПС.
4. Согласно равномерному закону вероятностей распределения выбрать действие  $b_i$  из множества  $B$ . Отработать выбранное действие.
5. Описать текущую ситуацию внешней среды  $S_{тек i+1}$ .
6. Проверить условие: « $S_{цел i} = S_{теки+1}$ », если условие выполняется, перейти к п. 8; в противном случае, перейти к п. 9.
7. Исключить проверяемое действие  $b_j$  из числа закрепляемых на текущем шаге самообучения, перейти к п. 5.
8. Проверить условие: «цель достигнута?» Если нет, то перейти к п.11, если да, перейти к п. 10.
9. Запомнить в формируемой ПЦП элементарный акт поведения:  
 $S_{тек i+1} \& b_i \rightarrow S_{цел i}$ . Принять на следующем шаге самообучения за  $S_{цел}$  ситуацию  $S_{тек i+1}$ , к пункту п. 4 .
10. Конец.

Для выполнения анализа и количественной оценки сложности алгоритма самообучения 1 в различных условиях ПС вводим следующие ограничения:

1) обязательным условием выработки ПЦП является наличие или появление в ПС необходимого целевого сигнала;

2) суммарная вероятность перехода ПС от состояния  $S_{теки}$  к другому состоянию такому, что на каждом этапе активного самообучения ИС равна единице при условии, что интеллектуальная система выполнит все действия из заданного множества  $B$  (частным случаем данного условия является то, что для каждого действия  $b_i$  из множества  $B$  в ПС существует переход от одного состояния к другому);

3) действия ИС из множества  $B_j$  равновероятны в условиях полной априорной неопределенности;

4) количество различных сигналов в ПС таково, что для хранения информации, связанной с формированием в среде любой по сложности УПП, достаточно оперативной памяти ИС;

5) из ограничений пп.4 вытекает, что для определения функциональной и емкостной сложности АС можно воспользоваться равномерными весовыми критериями, приняв за единицу функциональной сложности алгоритмов самообучения отработку ИС одного пробного действия.

**Аксиоматическое выражение 1.** Функциональная сложность алгоритма самообучения

1 выработки УПП в статических ПС определяется выражением  $\beta_1 = k_0 \sum_{i_{42}=1}^{\beta'} \sum_{i_{43}}^{\mathcal{E}_{i_{42}}} \beta_{i_{43}}$ , где  $\beta'$  - количество активных шагов самообучения ИС (каждый такой шаг завершается при достижении цели  $S_{цел}$ );  $\mathcal{E}_{i_{42}}$  - количество этапов на  $i_{42}$  шаге (каждый этап  $\mathcal{E}_{i_{42}}$  завершается при изменении текущего состояния ПС);  $\beta_{i_{43}}$  - количество пробных действий, обрабатываемых ИР на  $i_{43}$  этапе  $i_{42}$ -го шага.

**Предложение 1.** Максимальная функциональная сложность  $\beta_{1max}$  выработки УПП в статических ПС первой категории алгоритма 1 не превышает значения  $R(V_3) \sum_{v^{j^2}_{i35} \in R(v_3)'} (n_1 - \rho(v^{j^2}_{i35}) + 1)$ , где  $n_1$  - мощность множества  $B^*$  или количество разнообразных действий, обрабатываемых ИР в процессе самообучения;  $\rho(v^{j^2}_{i35})$  - локальная полустепень вершины  $\rho(v^{j^2}_{i35}) \in R(V_3)'$  для исходящих дуг;  $R(V_3)$  - максимальный ранг графа для исходной  $v_{i35}^0$  и целевой  $v_{i35}^k$  вершин, т.е. длина максимального пути между вершинами;  $R(V_3)'$  - множество всех вершин этого пути.

**Доказательство.**

1. Исходя из пп. 4-7 алгоритма самообучения 1 и при условии, что мощность множества В равна n, максимальное количество пробных действий, обрабатываемых на каждом этапе активного самообучения, не может быть более n.

2. Исходя из п.8 алгоритма самообучения 1, каждый этап самообучения завершается при появлении в ПС сигнала такого, что  $S_{цел} \neq S_{имме}$ , где  $S_{имме}$  - сигнал, характеризующий ПС до начала текущего шага изучения закономерностей ПС на  $i$  шаге. Следовательно, максимальное количество пробных действий, обрабатываемых ИС на каждом  $i$  шаге самообучения не больше величины  $n_1 - \rho^*(v_{i35}^{02}) + 1$ . Здесь единица учитывает действие, приводящее к завершению  $i$  этапа самообучения.

3. Согласно условию завершения каждого этапа самообучения  $\mathcal{E}_i$  после преобразования текущей ситуации ПС и из условия, что граф является ациклическим, вытекает, что максимальное количество этапов самообучения  $\mathcal{E}_i$  на каждом  $i$  шаге может быть не более  $R(V_3)$ .

4. Из условия завершения каждого шага самообучения при изменении текущей ситуации ПС и достижением цели следует, что количество таких шагов  $\beta'_1$  в ациклической ПС может быть не более  $R(V)$ .

5. Из пп. Доказательства 1-4 и аксиоматического выражения 1. следует  $\beta_{1max} \leq R(V_3) \sum_{v^{j^2}_{i35} \in R(v_3)'} (n_1 - \rho(v_{i35}) + 1)$ .

В случае, когда предвестниковый сигнал  $S_{мекі}$  предсказывает появление вредного или тормозного сигнала  $S_{i+1мек}$ , то для достижения стоящей перед ИС целью  $S_{рез}$  то вырабатывается ПЦП, приводящая к выключению сигнала  $S_{мекі}$  в ПС до появления связанного с ним условного сигнала  $S_{i+1мек}$ . Если по истечении заданного промежутка времени  $t_2$  ПЦП, приводящая к выключению сигнала  $S_{мекі}$ , полностью еще не сформирована, то по мере появления условного сигнала  $S_{i+1мек}$  ИС вырабатывает программу поведения,

связанную с его устранением, т.д., пока не будут выработаны обе части формируемой ПЦП, приводящие к выключению (устранению) в среде тормозных сигналов  $S_{meki}$  и  $S_{i+1mek}$ .

В случае неудачной попытки сформировать ПЦП, связанную с выключением (или включением) предвестникового сигнала  $S_{i+1mek}$ , ИС уточняет условный сигнал путем наблюдения за независимыми от нее событиями  $q_i \in Q$ , приводящими к включению предвестника при наличии в ПС необходимых для этого условий, определяющихся условным сигналом  $S_{i+2mek}$ . После формирования элементарного акта поведения вида:  $S_{meki} \& q_i \rightarrow S_{i+1mek}$  по результатам наблюдения за происходящими в среде событиями, формируется комбинированный условный сигнал  $S_{meki} \& q_i$ . Затем ИС определяет действие  $b_j \in B$ , приводящее к появлению в среде необходимого события  $q_i \in Q$  и формирует элементарный акт поведения следующего вида:  $S_{meki} \& b_j \rightarrow q_i$ . После этого формируется ПЦП, которая приводит к появлению в ПС события  $q_i \in Q$  из различных следующих друг за другом допустимых ситуациях среды.

Для построения алгоритма самообучения ИС для выявления в среде причинно-следственной взаимосвязи условных сигналов  $S_{meki} \in S$  и событий  $q_i \in Q$  используем следующие обозначения и операторы:

$t$  - отрезок времени, отведенный ИС под ожидание самопроизвольного появления в ПС необходимого события  $q_i \in Q$ ;

$t_m$  - отрезок времени, в течение которого ИС ожидает самопроизвольного появления в ПС события  $q_i \in Q$ ;

$\subset$  - обозначение условия нахождения условного сигнала в текущей ситуации ПС, например,  $S_{meki} \subset S$ ;

$s^{1t} = s^{2t}$  - равенство ситуаций  $s^{1t}, s^{2t} \in S$  между собой;

$S_{meki} \& S_{i+1mek}$  - обозначение одновременного присутствия в ПС сигналов  $S_{i+1mek}$  и  $S_{meki}$ .

С учетом принятых обозначений алгоритм самообучения ИС при наличии в ПС причинно-следственных связей представится следующим образом.

Алгоритм самообучения 2.

**Исходные условия:**  $S_{исх}, S_{цел}$ ;

**Входные переменные:**  $S_{itek} \in S, q_i \in Q$ ;

**Выходные переменные:** элементарные акты поведения следующего вида:  $S_{meki} \& b_{j1} \rightarrow S_{i+1mek}$ ;  $(S_{i+2mek} \& S_{i+1mek}) \& b_{j1} \rightarrow S_{meki}$ .

**1. Начало:**

2. Сформировать текущую ситуацию внешней среды  $S_{teki}$ .

3. Запустить таймер отсчета времени  $t$ , наблюдения за произвольными изменениями ситуаций ПС. С появлением в среде каждого нового события  $q_i \in Q$  сформировать текущую ситуацию  $S_{teki+1}$ .

4. Проверить условие: « $S_{teki} = S_{teki+1}$ »? Если да, то перейти к п.5; если нет, то перейти к п.6.

6. Сформировать сигналы:  $a_m = S_{teki+1} \setminus S_{teki}$  и  $a'_j = S_{teki} \setminus S_{teki+1}$ ;  $\Delta\delta = \delta_w + 1/k_3$ . Записать в память импликативное решающее правило

$$a'_j \& q_i \xrightarrow{T_m} a_m, \text{ где } T_m = t_1.$$

7. Принять за подцель функционирования в процессе самообучения условный сигнал  $q_i \xrightarrow{T_m} a_m$ .

8. Сформировать текущую ситуацию ПС  $S_{teki+2}$ .

9. Принять за подцель функционирования в процессе самообучения ИС появление в ПС события  $q_i$ .

10. Проверить условие: «событие  $q_i$  включено»? Если да, то перейти к п.16;если нет, то перейти к п.11.
11. Сформировать множество действий  $b_j \in B$ .
12. Выбрать действие  $b_j$  согласно равномерному закону распределения вероятностей их выполнения.
13. Отработать выбранное действие  $b_j^*$ .
14. Проверить условие: «событие  $q_i$  включено»? Если да, то перейти к п.11; если нет, то перейти к п.16.
15. Исключить действие  $b_j^*$  из множества  $B$  на текущем шаге самообучения.
16. После включения события  $q_i$  сформировать текущую ситуацию ПС  $S_{теки+4}$ .
17. Проверить условие: « $S_{цел} \subset S_{теки+4}$ »? Если да, то выполняются безусловные реакции и перейти к п.18;если нет, то перейти к п. 24.
18. Проверить условие: «событие  $q_i$  включено» в результате отработки действия  $b_j^*$ ? Если да, то перейти к п.19 ;если нет, то перейти к п.24.
19. Сформировать сигнал  $a'_j = S_{теки+4} \setminus S_{теки+3}$ ; Занести в память элементарный акт поведения:  $a'_j \& b_{j \rightarrow q_i}^*$ .
20. Проверить условие «необходимо появление заданного сигнала  $S_{цел}$ ». Если такая необходимость есть, то перейти к п.21; в противном случае закончить самообучение.
- 21.Сформировать текущую ситуацию ПС  $S_{теки+5}$ .
22. Провести имитацию отработки действия  $b_j^*$ . Проверить условие: «в среде произошло включение события  $q_i$ »? Если да, то перейти к п.23; если нет, то перейти к п.24.
23. Проверить условие:  $\delta > 0$ ? Если да, то перейти к п. ;если нет, то перейти к п. .
24. Используя алгоритм самообучения 1 сформировать ПЦП приводящую к включению в ПС события  $g_i$ ; перейти к п. 22.
25. Используя алгоритм самообучения 1 сформировать УПП, приводящую к включению в ПС сигнала  $S_{цел}$ .

## 26.Конец.

Определим функциональную сложность алгоритма самообучения 2 ИС при наличии в ПС причинно-следственных связей между условными сигналами и событиями. Считаем, что значение функциональной сложности алгоритма самообучения для наиболее трудоемких условий ПС лежит в следующих граничных пределах  $2n_{35} \leq \beta_1 \leq n_{35}^2 n_1$  [1].

**Предложение 2.** Функциональная сложность  $\beta$  выработки ПЦП по алгоритму самообучения 2 при наличии в ПС причинно-следственных связей имеет следующие граничные значения  $6n_{35} \leq \beta \leq n_1(3n_{35}^2 + 1)$ ,

где  $n_{35}$ - количество сигналов, характеризующих среду;  $n_1$  - количество действий, обрабатываемых ИС.

**Доказательство.** Количество пробных действий  $\beta$ , обрабатываемых ИС по предложенному алгоритму самообучения в процессе формирования ПЦП при наличии в ПС причинно-следственных связей между условными сигналами, определяется согласно выражению  $\beta = \beta' + 3\beta_1$ , где  $\beta'$ - количество пробных операций, выполняемых ИС для включения событий  $q_i^* \in Q$ , связанных с появлением предвестников  $a_m \in A$ ;  $3\beta_1$  - количество пробных действий, обрабатываемых в процессе формирования ПЦП вида простой цепи.

2. Значение  $3\beta_1$  для наиболее трудоемких условий ПС лежит в пределах  $3\beta_1 \in [6n_{35}, 3n_{35}^2 n_1]$ , что вытекает из граничных значений функциональной сложности использованного алгоритма самообучения 1 для формирования ПЦП, приводящей в среде требуемого события  $q_i^* \in Q$ .

3. При формировании имплицативного решающего правила  $(S_{итек} \& S_{теки+1}) \& b_{j1} \rightarrow q_i$ , ИС при выполнении соответствующих условий в лучшем случае может отработать одно действие, если эффективное действие из множества  $B_1$  выбирается первым (это вполне вероятно, т.к. выбор действий выполняется согласно равномерному закону распределения

вероятностей их отработки). В худшем случае, ИС может выполнить все действия из множества  $B_1$  при условии, что нужное действие выбрано последним,  $|B_1| = n_1$ .

4. Из пп.1-3 доказательства следует, что для наиболее трудоемких условий ПС значение  $\beta$  лежит в следующих пределах:

$$6n_{35} \leq \beta \leq n_1(3n_{35}^2 + 1).$$

Таким образом, предложенные алгоритмы самообучения имеет полиномиальную сложность, а, следовательно, за ограниченное количество шагов обучения обеспечивает формирование ИС программы целесообразного поведения, позволяющей достигать требуемых целей в априори неописанных условиях ПС.

**Библиографический список:**

1. Берштейн Л.С., Мелехин В.Б. Планирование поведения интеллектуального робота. -М.: Энергоатомиздат, 1994. 240 с.