

В. Б. Мелехин, д-р техн. наук, проф., pashka1602@rambler.ru,

Дагестанский государственный технический университет, г. Махачкала,

М. В. Хачумов, канд. физ.-мат. наук, ст. науч. сотр., khmike@ribox.ru,

Федеральный исследовательский центр "Информатика и управление" Российской Академии наук,
Российский университет дружбы народов, Москва

Инструментальные средства управления целесообразным поведением самоорганизующихся автономных интеллектуальных агентов*

Сформулированы основные принципы построения знаковосигнального управления целесообразным поведением автономных интеллектуальных агентов в априори неописанных условиях проблемной среды. Уточнено понятие самоорганизующегося автономного интеллектуального агента как системы, способной к автоматическому целеполаганию при появлении в проблемной среде определенного вида условных и безусловных сигналов — знаков. Разработаны процедуры планирования целесообразного поведения автономных интеллектуальных агентов, предусматривающие имитацию отработки пробных действий в условиях неопределенности в процессе изучения закономерностей преобразования ситуаций проблемной среды, что позволяет избежать не связанных с достижением заданной цели изменений среды в процессе самообучения. Определены граничные оценки сложности предложенных процедур планирования целесообразного поведения, подтверждающие возможность их эффективной реализации на бортовом вычислителе автоматической системы управления целесообразной деятельностью автономных интеллектуальных агентов. Проведена имитация на персональном компьютере предложенных процедур планирования целенаправленного поведения, подтвердившая эффективность их использования для построения интеллектуальных решателей проблем автономных интеллектуальных агентов в целях наделяния их способностью адаптироваться к априори неописанным условиям функционирования. Структурированы основные виды связей между различными условными и безусловными сигналами — знаками проблемной среды, что позволяет автономным интеллектуальным агентам приспосабливаться к сложным априори неописанным, нестабильным условиям функционирования.

Ключевые слова: автономный интеллектуальный агент, целесообразное поведение, условия неопределенности, закономерности проблемной среды, самообучение, элементарный акт поведения

Введение

Актуальной проблемой создания автономных интеллектуальных агентов (АИА) различного назначения, оснащенных техническим зрением и развитой моторикой, которые способны эффективным образом целенаправленно функционировать в условиях неопределенности, является формализация мыслительных актов наглядно-действенного мышления живых систем [1—3] и построение на этой основе инструментальных средств управления целесообразным поведением. Наглядно-действенное мышление АИА в отличие, например, от

наглядно-образного мышления, к которому можно отнести процедуры вывода, обеспечивающие возможность планирования целенаправленного поведения на основе заданной модели проблемной среды (ПС) [4, 5], связано с отработкой в среде пробных действий и определением результативных действий в целях выявления закономерностей преобразования ее ситуаций [6].

Одним из подходов к решению данной проблемы является разработка эффективного знаковосигнального принципа управления целесообразным поведением, позволяющим самоорганизующимся АИА адаптироваться к априори неописанным условиям ПС [7, 8].

В этом случае роль различных по содержанию и структуре условных сигналов, появление которых в ПС влияет на поведение АИА

*Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 17-29-07003 офи_м, №18-07-00025 а.

в соответствии с его функциональным назначением и заданной целью, могут играть:

- определенные состояния отдельных объектов, находящихся в ПС;
- отношения, складывающиеся между различными объектами ПС;
- фрагменты ситуаций, описывающих текущие условия ПС и др.

Следует отметить, что наиболее важное практическое значение для самоорганизующихся АИА имеют знаки в виде фрагментов ситуаций, описывающих текущие условия ПС. Это обусловлено тем, что такие знаки позволяют определять сложные цели и подцели поведения.

Произвольный условный сигнал, наблюдаемый в ПС, становится знаком, вызывающим у АИА определенные поведенческие реакции, если в процессе самообучения была установлена его связь с достижением заданной на текущий момент времени цели поведения. Таким образом, ориентация АИА в процессе целесообразного поведения на условные сигналы — знаки — обеспечивает ему возможность избирательного восприятия и переработки поступающей из ПС информации.

Различные закономерности появления в ПС связанных во времени и пространстве между собой и заданной целью знаков АИА может выявить на основе алгоритмов самообучения, предложенных в работе [9]. Организовать самообучение АИА в априори неописанных условиях ПС также можно, используя роевые алгоритмы поведения живых организмов, формализуемые с помощью различного вида конечных автоматов [10], а также генетические алгоритмы [11]. Однако данные алгоритмы предусматривают в процессе организованного на их основе самообучения по методу проб и ошибок необходимость непосредственной отработки АИА пробных действий, что может привести к необратимым и негативным изменениям в ПС, не связанным с достижением заданной цели. Обойти данный недостаток управления целесообразным поведением АИА на основе известных алгоритмов планирования можно, разработав процедуры самообучения, которые не требуют непосредственной отработки пробных действий в ПС, заменив отработку действий имитацией их выполнения на формальном описании текущей ситуации ПС. В данной работе рассматривается вариант решения данной проблемы.

Постановка задачи

Пусть самоорганизующийся АИА в соответствии с функциональным назначением может обрабатывать множество действий $B = \{b_k\}$, $k = \overline{1, n_1}$. Под самоорганизацией следует понимать способность АИА к автоматическому целеполаганию при возникновении в ПС или в его внутренней среде определенного вида сигналов — знаков, связанных, например, с безопасностью деятельностью [12].

ПС можно охарактеризовать множеством условных $A = \{a_i\}$, $i = \overline{1, n_2}$, и безусловных $C = \{c_j\}$, $j = \overline{1, n_3}$, сигналов, которые появляются в ней либо в результате отработки АИА различных действий $b_k \in B$, либо в результате происходящих в ней независимых от АИА событий $D = \{d_\varphi\}$, $\varphi = \overline{1, n_4}$.

Требуется разработать такое формальное описание обрабатываемых АИА действий, текущих ситуаций ПС S_t , $t = \overline{1, n_4}$, и инструментальных средств обработки информации и вывода решений, которые позволяют АИА в процессе самообучения имитировать отработку пробных действий $b_k \in B$ в ПС и на этой основе формировать условные программы целесообразного поведения, имеющие, например, следующую структуру:

$$a_i^1 \& b_k^1 \rightarrow a_i^2 \& b_k^2 \rightarrow \dots \rightarrow a_i^z \& b_k^z \rightarrow \dots \rightarrow a_i^n \& b_k^n \rightarrow c_j, \quad (1)$$

где a_i^1 — условный сигнал — знак, который АИА воспринимает в исходной ситуации ПС S_i ; $a_i^z \& b_k^z \rightarrow a_i^{z+1}$ — элементарный акт поведения, означающий, что если АИА наблюдает в ПС знак a_i^z , то обрабатываемое им действие b_k^z приводит к появлению в ней знака a_i^{z+1} ; c_j — безусловный сигнал — знак, вызывающий у АИА безусловные реакции $B_j^* \subseteq B$, представляющие собой заданную упорядоченную последовательность действий $b_k \in B$, отработка которых приводит к достижению текущей цели поведения c_j^* при восприятии в ПС этого знака.

Следует отметить, что после закрепления в условной программе поведения условные сигналы — знаки — $a_i \in A$ приобретают роль ориентиров или подцелей поведения, что позволяет АИА в процессе самообучения строить на их основе разветвленные программы целесообразной деятельности, охватывая различные области текущих условий функционирования.

Организация знаковосигнального принципа управления поведением АИА без непосредственной отработки в ПС пробных действий

Как отмечено выше, наглядно-действенное мышление служит для адаптации АИА к априори неописанным условиям функционирования путем выявления в процессе самообучения закономерностей их преобразования, связанных с достижением заданной на текущий момент времени цели c_j^* поведения. На этой основе АИА формирует условные программы целесообразного поведения, опираясь на формальное описание текущей ситуации среды S , которое строится им в виде графа видимости в пределах разрешающей способности технического зрения. Данный вид мышления позволяет АИА реализовать в априори неописанных условиях ПС следующие два вида целесообразной деятельности:

- *мотивационное, или самоорганизующееся, поведение*, которое характеризуется автоматическим целеполаганием при появлении либо в ПС, либо во внутренней среде АИА определенных по содержанию безусловных сигналов — знаков $c_j \in C$, приводящих к необходимости перепланирования текущей целесообразной деятельности.

Примером появления сигналов в ПС может быть безусловный сигнал $c_j \in C$, который связан с достижением соответствующей ему цели c_j^* , позволяющей АИА избежать столкновения с появившимся в ПС подвижным объектом. В другом случае (при появлении сигнала во внутренней среде АИА) на основе данных, поступающих из информационно-измерительной подсистемы, в интеллектуальном решателе задач автоматически формулируется цель поведения c_j^* , достижение которой позволяет АИА устранить аварийную ситуацию, возникшую в результате выхода из строя одной из его подсистем и т.п.;

- *целесообразное поведение*, в процессе реализации которого АИА изучает закономерности преобразования текущей ситуации ПС, связанные с достижением заданной ему извне цели поведения c_j^* .

В общем случае целесообразное поведение АИА, основанное на знаковосигнальном принципе управления, может быть организовано по следующей схеме:

1) в интеллектуальном решателе задач на основе данных, поступающих из информаци-

онно-измерительной подсистемы технического зрения АИА, формируется формальное описание текущей ситуации ПС S_i в виде нечеткой семантической сети [13], на основе которой затем осуществляется имитация отработки пробных действий в процессе самообучения;

2) на основе принятого критерия отбора АИА методом проб и ошибок выявляет полезные с точки зрения достижения заданной цели действия без непосредственной их отработки в ПС. На каждом шаге такого самообучения формируется элементарный акт поведения, например, следующего вида: $a_i^z \& b_k^z \rightarrow a_i^{z+1}$;

3) процесс самообучения продолжается до закрепления в формируемой условной программе элементарного акта поведения, который содержит заданный безусловный сигнал $c_j \in C$, т.е. имеет, например, следующую структуру: $a_j^* \& b_k^* \rightarrow c_j$. Появление в ПС безусловного сигнала — знака c_j — вызывает у АИА соответствующие ему безусловные ответные реакции B_j^* , приводящие к достижению соответствующей ему цели поведения c_j^* ;

4) формируемая таким образом условная программа запоминается и используется в дальнейшем в качестве накопленного опыта поведения в различных условиях ПС, аналогичных условиям, изученным в процессе самообучения.

Модель представления знаний, используемых АИА в процессе наглядно-действенного мышления, определяется следующими типовыми конструкциями:

- ♦ структурированным описанием в виде нечеткой семантической сети подситуаций $s_{\text{И}}^k$, определяющих условия, которые должны выполняться в ПС для успешной отработки действий $b_k \in B$, а также подситуации $s_{\text{Р}}^k$, получаемой в результате их выполнения. Типовые элементы представления знаний имеют следующий формат описания:

<Идентификатор действия b_k ,
"вход $s_{\text{И}}^k$ " & $b_k \rightarrow$ "выход $s_{\text{Р}}^k$ ">.

Отметим, что после закрепления в формируемой условной программе поведения действий $b_k \in B$ соответствующие им входные $s_{\text{И}}^k$ и выходные $s_{\text{Р}}^k$ подситуации становятся условными сигналами — знаками, связанными с достижением заданного на текущий момент времени безусловного сигнала — знака c_j ;

- ♦ типовыми структурами формального описания в виде нечеткой семантической сети безусловных знаков $c_j \in C$ и соответствующих им безусловных реакций B_j^* , представленных в виде элементарных актов поведения $c_j \& B_j^* \rightarrow c_j^*$. Безусловные реакции B_j^* отрабатываются АИА в ответ на появление в ПС соответствующего им безусловного сигнала. Следует отметить, что безусловные реакции АИА могут быть связаны и с необходимостью перепланирования текущей деятельности в нестабильных условиях ПС;
- ♦ структурированным описанием текущей ситуации ПС S_i в виде нечеткой семантической сети S_i^* .

Такое представление знаний позволяет реализовать эффективные процедуры планирования целесообразного поведения с имитацией отработки АИА различных пробных действий $b_k \in B$ в процессе самообучения. Основными операциями, выполняемыми в процессе вывода решений в этом случае, являются операции сравнения между собой нечетких семантических сетей на предмет определения нечеткого вложенного изоморфизма одной сети в другую и нечеткого равенства сравниваемых между собой сетей [13].

В общем виде процедуры планирования АИА целесообразного поведения имеют следующую структуру.

Входные переменные. Заданный безусловный сигнал — знак c_j ; текущая ситуация ПС S_i .

Промежуточные переменные. Воспринимаемые в ПС условные $a_i \in A$ и безусловные $c_j \in C$ сигналы — знаки; формируемые элементарные акты поведения $a_i^z \& b_k^z \rightarrow a_i^{z+1}$ и $a_i^n \& b_k^n \rightarrow c_j$.

Выходные переменные. Сформированная условная программа поведения:

$$\begin{aligned} s_{\text{И}}^k(1) \& b_k(1) \rightarrow s_{\text{Р}}^k \& b_k(2) \rightarrow s_{\text{И}}^k(2) \& b_k(3) \rightarrow \\ \rightarrow \dots \rightarrow s_{\text{И}}^k(j_1 - 1) \& b(j_1) \rightarrow \dots \\ \rightarrow s_{\text{Р}}^k(j_1) \& b_k(j_1 + 1) \rightarrow s_{\text{И}}^k(n) \& b_k(n) \rightarrow c_j, \end{aligned}$$

где j_1 — индекс, соответствующий номеру шага планирования поведения.

Метод.

1. *Начало.* Построить в решателе задач формальное описание:

— текущей ситуации ПС S_i в виде нечеткой семантической сети S_i^* ;

— целевой ситуации ПС S_i^{**} путем замены в нечеткой семантической сети S_i^* значений

отношений на значения биективно соответствующих им одноименных отношений в формальном описании заданного безусловного сигнала — знака c_j .

2. $j_1 = j_2 = 1$; приступить к планированию поведения, позволяющего изучить закономерности преобразования ситуации S_i в целях появления в ПС заданного на текущий момент времени безусловного сигнала — знака c_j .

3. Проверить условие: "Множество B включает действие b_k , у которого формальное представление входной подситуации $s_{\text{И}}^k$ нечетко изоморфно вложено в нечеткую семантическую сеть S_i^* , а для выходной подситуации $s_{\text{Р}}^k$ справедливым является одно из следующих условий:

а) семантическая сеть, определяющая условный сигнал c_j , нечетко изоморфно вложена в формальное описание подситуации $s_{\text{Р}}^k$;

б) расплывчатые семантические сети, соответствующие безусловному знаку c_j и подситуации $s_{\text{Р}}^k$, нечетко равны между собой".

Если условие выполняется, то перейти к п. 4; в противном случае перейти к п. 5.

4. Сохранить в качестве формируемой условной программы элементарный акт поведения $s_{\text{И}}^k(1) \& b_k(1) \rightarrow c_j$, отработка действия $b_k(1)$ которого в ситуации S_i , приводит к появлению в ней безусловного сигнала — знака c_j . *Конец.*

5. Определить на текущем j_1 шаге самообучения в условиях неопределенности случайным образом пробное действие $b_k(j_1)$ согласно равномерному закону распределения вероятностей его выбора во множестве действий B .

6. Проверить условие: "Вход $s_{\text{И}}^k(j_1)$ выбранного действия $b_k(j_1)$ является нечетко равным структурно эквивалентной ему части нечеткой семантической сети S_i^* ". Если условие выполняется, то перейти к п. 8; в противном случае исключить действие $b_k(j_1)$ из числа результативных действий на текущем шаге самообучения, перейти к п. 7.

7. Проверить условие: " $j_1 = n_1$ ". Если условие выполняется, то перейти к п. 17; в противном случае $j_1 := j_1 + 1$, перейти к п. 5.

8. Выполнить замену значений отношений в нечеткой семантической сети S_i^* на значения биективно соответствующих им одноименных отношений выхода $s_{\text{Р}}^k(j_1)$ пробного действия $b_k(j_1)$, сформировав, таким образом, нечеткую семантическую сеть S_i^T .

9. Проверить условие: "Условный сигнал c_j является вложено изоморфным в нечеткую семантическую сеть S_i^T ". Если условие выпол-

няется, то перейти к п. 10: в противном случае перейти в п. 11.

10. Сохранить в формируемой условной программе элементарный акт поведения $s_{\text{И}}^k(j_1) \& b_k(j_1) \rightarrow c_j$. *Конец.*

11. Рассчитать степени нечеткой близости $\rho_1(S_t^*, S_t^{**})$ и $\rho_2(S_t^T, S_t^{**})$ [13], соответственно, для пар нечетких семантических сетей $\langle S_t^*, S_t^{**} \rangle$ и $\langle S_t^T, S_t^{**} \rangle$.

12. Проверить условие: " $\rho_2(S_t^T, S_t^{**}) > \rho_1(S_t^*, S_t^{**})$ ". Если условие выполняется, то перейти к п. 13; в противном случае исключить действие $b_k(j_1)$ из числа результативных действий на текущем шаге самообучения, $j_1 := j_1 + 1$, перейти к п. 7.

13. Сохранить в формируемой условной программе элементарный акт поведения $s_{\text{И}}^k(j_1) \& b_k(j_1) \rightarrow s_{\text{Р}}^k(j_1)$; $j_2 := j_2 + 1$, где j_2 — число элементарных актов поведения, закрепленных на текущий момент времени в формируемой условной программе поведения.

14. Выбрать во множестве B действие $b_k(j_1 + 1)$, у которого вход $s_{\text{И}}^k(j_1 + 1)$ совпадает с выходом $s_{\text{Р}}^k(j_1)$ действия $b_k(j_1)$ последнего элементарного акта поведения, закрепленного в формируемой условной программе целесообразной деятельности.

15. Если такое действие $b_k(j_1 + 1)$ найдено, и оно отличается от действия, исключенного на предыдущем шаге планирования, то принять $j_1 := j_1 + 1$, перейти к п. 6; в противном случае перейти к п. 16.

16. Исключить действие $b_k(j_1)$ как не результативное для достижения заданной цели c_j на текущем шаге самообучения и соответствующий ему элементарный акт поведения из формируемой условной программы; $j_2 := j_2 - 1$.

17. Проверить условие: " $j_2 = 0$ ". Если условие выполняется, перейти к п. 18; в противном случае перейти к п. 14.

18. У АИА отсутствует возможность построения необходимой условной программы целесообразного поведения, приводящей к достижению заданного безусловного сигнала — знака c_j в текущей ситуации ПС S_t . *Конец.*

Для оценки сложности приведенных выше процедур планирования целесообразного поведения АИА введем следующие понятия:

- функциональная сложность β_1 процедур планирования, зависящая от общего числа действий $b_k \in B$, апробируемых АИА в процессе формирования условной программы целесообразного поведения;

- временная сложность β_2 процедур планирования, которая определяется числом элементарных операций сравнения между вершинами и ребрами нечетких семантических сетей при выводе решений в процессе формирования условной программы целесообразного поведения.

Утверждение 1. Функциональная сложность β_1 , процедур планирования целесообразного поведения АИА, в общем случае, определяется следующими граничными оценками:

$$m_1 \leq \beta_1 \leq n_2 n_1,$$

где m_1 — число элементарных актов поведения, закрепленных в сформированной условной программе целесообразного поведения; n_2 — число сигналов, характеризующих закономерности преобразования текущей ситуации проблемной среды S_t ; n_1 — число различных действий, которые АИА способен отрабатывать согласно своему назначению.

Доказательство. Справедливость сформулированного утверждения 1 вытекает из следующих соображений.

1. Согласно пп. 5–7 и пп. 14–17 процедур формирования условных программ целесообразного поведения вполне вероятно, что в лучшем случае на каждом j_1 шаге планирования первым выбирается результативное действие $b_k(j_1)$, а затем происходит его отработка. Следовательно, нижнее граничное значение оценки сложности β_1 , в общем случае, может быть не менее величины, равной m_1 .

Исключением является частный случай (см. пп. 3, 4 процедуры планирования), когда $m_1 = 1$, который соответствует условию, что заданная цель поведения c_j может быть достигнута в текущих условиях функционирования в результате отработки действия $b_k(1)$, выявленного АИА на первом шаге самообучения.

2. В худшем случае результативное действие $b_k(1)$ на каждом j_1 шаге планирования целесообразного поведения может быть выбрано в последнюю очередь. Отсюда на каждом шаге планирования целесообразного поведения апробируется не более n_1 действий $b_k(j_1) \in B$.

В этом случае для включения в ПС необходимого безусловного сигнала c_j в процессе выявления как результативных (см. пп. 5, 14 процедуры планирования), так и нерезультативных (см. п. 16 процедуры планирования)

элементарных актов поведения число выполняемых АИА шагов планирования не может превышать числа сигналов n_2 , характеризующих текущие условия ПС.

Следовательно, функциональная сложность β_1 процедуры планирования целесообразного поведения АИА в худшем случае может быть не более величины, равной $n_2 n_1$.

3. Из пп. 1, 2 проведенного доказательства с очевидностью следует справедливость утверждения 1.

Утверждение 2. Временная сложность β_2 предложенной процедуры планирования целесообразного поведения АИА определяется следующими граничными оценками:

$$2m_1 m_2^2 t_{\text{CP}} \leq \beta_2 \leq 2n_2 n_1 m_2^2 t_{\text{CP}},$$

где m_2 — число вершин в нечеткой семантической сети S_t^* ; t_{CP} — среднее время выполнения одной элементарной операции сравнения при сопоставлении между собой различных нечетких семантических сетей.

Доказательство. Справедливость утверждения 2, с учетом граничных оценок функциональной сложности β_1 , можно обосновать, проведя следующие рассуждения.

1. Согласно утверждению 1 в лучшем случае в процессе формирования условной программы целесообразного поведения совершаются m_1 шагов планирования. На каждом таком шаге выполняется одно сравнение нечетких семантических сетей между собой в процессе выбора результативного действия (см. п. 6 процедур планирования) и одно сравнение для проверки условия достижения заданного безусловного сигнала c_j (см. п. 9 процедур планирования). Учитывая, что в процессе вывода ре-

шений в каждом рассмотренном выше случае сравниваются между собой помеченные графы (см. пп. 11, 12 процедуры планирования), число сопоставлений между собой структурно эквивалентных элементов нечетких семантических сетей на каждом шаге планирования поведения может быть не менее величины, равной m_2^2 .

Таким образом, нижняя граничная оценка временной сложности β_2 процедуры планирования целесообразного поведения АИА не может быть менее величины, равной $2m_1 m_2^2 t_{\text{CP}}$.

2. Согласно утверждению 1 в худшем случае в процессе построения условной программы целесообразного поведения требуется апробировать $n_2 n_1$ действий. В процессе каждой такой апробации выполняется по два сравнения между собой нечетких семантических сетей. Следовательно, верхняя граничная оценка временной сложности β_2 предложенной процедуры планирования целесообразного поведения АИА определяется на основе следующего выражения: $2n_2 n_1 m_2^2 t_{\text{CP}}$.

3. Из пп. 1, 2 проведенного доказательства с очевидностью следует справедливость утверждения 2.

Полученные таким образом граничные оценки позволяют в первом приближении оценить требуемую для реализации предложенных процедур планирования целесообразного поведения или самообучения АИА производительность его бортовой вычислительной системы.

Рассмотрим ПС, в которой связь между обрабатываемыми АИА действиями $b_k \in B$ и появляющимися в ней в результате этого различными сигналами $a_i \in A$ и $c_j \in C$ определяется приведенными в таблице данными.

Согласно этим данным ПС характеризуется появлением в ней 15 условных сигналов

Закономерности появления в ПС различных сигналов в результате обрабатываемых АИА действий

Regularities of the appearance in the problem environment of various signals as a result of the actions performed by the autonomous intelligent agent

b_k	a_i														
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
b_1	a_3	a_7	a_5	a_2	a_3	a_6	a_9	a_{12}	a_{13}	a_{15}	a_7	a_8	c_j	a_{15}	a_9
b_2	a_8	a_1	a_3	a_2	a_5	a_7	a_8	a_{12}	a_{13}	a_{14}	a_{15}	a_{15}	a_9	a_7	a_6
b_3	a_1	a_3	a_2	a_6	a_6	a_{13}	a_6	a_{13}	a_8	a_{13}	a_{12}	a_{14}	a_7	a_9	a_{12}
b_4	a_3	c_j	a_1	a_3	a_6	a_{11}	a_6	a_{12}	a_{11}	a_{12}	a_{13}	a_{12}	a_{10}	a_{11}	a_{11}
b_5	a_4	a_4	a_4	a_4	a_4	a_8	a_8	a_{14}	a_6	c_j	a_9	c_j	a_9	a_{14}	a_{13}
b_6	a_5	a_3	a_5	a_5	a_{12}	a_{12}	a_{12}	a_{15}	a_{12}	c_j	a_8	a_{13}	a_7	c_j	c_j

$a_i \in A, i = \overline{1, 15}$, и одного безусловного сигнала c_j . Смена сигналов в ПС происходит в результате шести обрабатываемых АИА действий $b_k \in B, k = \overline{1, 6}$. Пересечение столбцов и строк таблицы определяет закономерности формирования АИА элементарных актов поведения. Например, пересечение седьмого столбца и четвертой строки обеспечивает АИА формирование следующего элементарного акта поведения $a_7 \& b_4 \rightarrow a_6$.

На основе приведенных в таблице данных на персональном компьютере проведено имитационное моделирование процесса самообучения АИА по предложенным процедурам планирования целесообразного поведения (при удалении из них пунктов планирования 3 и 4) как переход конечного автомата Q со случайными реакциями [14] от исходного состояния, характеризующегося условным сигналом $a_2 \in A$, к необходимому состоянию, которому соответствует безусловный сигнал c_j .

В результате проведенного эксперимента была получена следующая условная программа целесообразного поведения:

$$a_2 \& b_2 \rightarrow a_1 \& b_2 \rightarrow a_8 \& b_4 \rightarrow a_{12} \& b_5 \rightarrow c_j.$$

В процессе формирования данной программы АИА проверил целесообразность отработки в ПС 21 пробного действия. Таким образом, полученная при проведении эксперимента оценка сложности β_1 предложенной процедуры планирования целесообразного поведения АИА не превышает найденного аналитическим путем ее верхнего граничного значения, равного $n_2 n_1 = 96$.

После восстановления в процедуре планирования решающих правил пп. 3, 4 АИА в процессе имитационного моделирования сформировал условную программу целесообразного поведения следующего вида: $a_2 \& b_4 \rightarrow c_j$, апробировав в процессе самообучения отработку четырех действий, что значительно меньше, чем в первом случае.

Таким образом, найденные аналитическим путем граничные оценки сложности предложенной процедуры планирования целесообразного поведения, а также проведенные на персональном компьютере эксперименты подтверждают эффективность ее применения для создания решателя задач АИА, способного целенаправленно функционировать в априори неописанных условиях ПС.

Структура основных закономерностей преобразования ситуаций проблемной среды

По аналогии со структурой сигналов, закрепляемых в условных рефлексах живыми организмами [15], элементарные акты поведения, входящие в формируемые АИА условные программы целесообразного поведения и происходящие в ПС самопроизвольные изменения, могут иметь следующую структуру.

1. Элементарные акты поведения типа "простое звено", имеющие следующую структуру: $a_i \& b_k \rightarrow a_i'$ или $a_i \& b_k \rightarrow c_{ji}'$. Отдельные условные сигналы $a_i \in A$, входящие в структуру простого звена, как правило, определяются отдельными фрагментами $s_i \subseteq S_i$ текущей ситуации среды S_i . Закономерности ПС вида "простое звено" позволяют АИА формировать программы условного поведения в виде простой цепи (1).

2. Акты поведения, включающие суммарные комплексы безусловных вспомогательных сигналов — знаков $\langle c_j(1) \wedge c_2(2) \wedge \dots \wedge c_j(l) \rangle$, имеющие следующую структуру:

$$(c_j(1) \wedge c_2(2) \wedge \dots \wedge c_j(d)) \& b_k \rightarrow c_j,$$

где c_j — основной безусловный сигнал — знак, который появляется в ПС в результате отработки АИА действия b_k только при одновременном присутствии в ней вспомогательных безусловных сигналов — знаков $c_j(j_3)$, $j_3 = 1, 2, \dots, d$; $\wedge$ — символ, означающий необходимость одновременного присутствия в ПС связанных им сигналов — знаков.

При наличии такого вида закономерностей преобразования ситуаций ПС для достижения основного безусловного сигнала — знака c_j — при отсутствии в среде вспомогательных безусловных сигналов — знаков $c_j(1), c_j(2), \dots, c_j(d)$ — АИА формирует разветвленную условную программу поведения, состоящую из l_{j_3} , $j_3 = \overline{1, d}$, простых цепей. Оработка каждой такой j_3 простой цепи АИА приводит к появлению в ПС соответствующего ей вспомогательного безусловного сигнала — знака $c_j(j_3)$. Таким образом, при одновременном наличии в ПС вспомогательных условных сигналов — знаков, входящих в комплекс $\langle c_j(1) \wedge c_2(2) \wedge \dots \wedge c_j(l) \rangle$, — обрабатываемое АИА действие b_k приводит к появлению в среде основного безусловного сигнала — знака c_j .

3. Элементарные акты поведения, связанные с устранением возникающих в нестабильной ПС тормозных сигналов — знаков a_i^T , которые имеют следующую структуру: $(a_i \wedge a_i^T) \& b_k \rightarrow \overline{a_i^*}$, $(a_i \wedge \overline{a_i^T} a_i^*) \& b_k \rightarrow a_i^*$, $(a_i \wedge a_i^T) \& b_k \rightarrow \overline{c_j}$ и $(a_i \wedge \overline{a_i^T}) \& b_k \rightarrow c_j$.

Здесь элементарный акт поведения $(a_i \wedge a_i^T) \& b_k \rightarrow \overline{a_i^*}$ означает, что ранее закрепленный в условной программе элементарный акт поведения $(a_i) \& b_k \rightarrow a_i^*$ перестает быть результативным после появления в ПС тормозного сигнала — знака a_i^T . После выключения тормозного сигнала a_i^T в текущей ситуации ПС (обозначается как $(a_i \wedge \overline{a_i^T} a_i^*) \& b_k \rightarrow a_i^*$) элементарный акт поведения $(a_i) \& b_k \rightarrow a_i^*$ вновь становится результативным. Аналогичным образом интерпретируется и элементарный акт поведения $(a_i \wedge a_i^T) \& b_k \rightarrow \overline{c_j}$.

Структура появляющихся в ПС тормозных сигналов a_i^T определяется АИА следующим образом:

$$a_i^T = S_t^{**} / S_t^*,$$

где S_t^{**} — нечеткая семантическая сеть, описывающая текущую ситуацию ПС S_t^T , в которой элементарный акт поведения $(a_i) \& b_k \rightarrow a_i^*$ потерял свою результативность; S_t^* — нечеткая семантическая сеть, соответствующая ситуации ПС S_t , в которой элементарный акт поведения $(a_i) \& b_k \rightarrow a_i^*$ был сформирован; / — теоретико-множественная операция, обозначающая разность нечетких семантических сетей S_t^{**} и S_t^* , которая выполняется с учетом пометок структурно эквивалентных в них вершин и ребер.

Другими словами, тормозной сигнал a_i^T определяется фрагментом нечеткой семантической сети S_t^{**} , $a_i^T \subseteq S_t^{**}$, который состоит из вершин и ребер, имеющих пометки, отличающиеся от пометок структурно эквивалентных им вершин и ребер в сети S_t^* .

Таким образом, при появлении в ПС тормозного сигнала a_i^T и наличии условной программы поведения, обеспечивающей достижение в ситуации S_t заданного безусловного сигнала c_j , АИА формирует новую условную программу поведения, отработка которой позволяет включить в среде сигнал a_i^T в результате устранения различий между ситуациями S_t и S_t^T .

Данная программа целесообразного поведения строится в дополнение к ранее сформированной результативной условной программе, содержащей элементарный акт поведения $(a_i) \& b_k \rightarrow a_i^*$ и заданный безусловный сигнал — знак c_j .

4. Самопроизвольные изменения условий функционирования АИА вида $a_i^{\Pi T} \rightarrow d_\phi \rightarrow a_i^*$ и $a_i^{\Pi T} \rightarrow d_\phi \rightarrow c_j$, отражающие причинно-следственные связи между различными сигналами — знаками нестабильной ПС с устойчиво повторяющимися в ней событиями $d_\phi \in D$, где a_i^* — условный сигнал — знак, закрепленный в ранее сформированной условной программе целесообразного поведения. Элементарный акт поведения, например, $a_i^{\Pi T} \rightarrow d_\phi \rightarrow c_j$ означает следующее. После появления пускового условного сигнала — знака a_i^{Π} — по истечению промежутка времени T в ПС происходит событие d_ϕ , которое, в свою очередь, включает в среде условный сигнал — знак a_i^* . Аналогичным образом интерпретируется и событие $a_i^{\Pi T} \rightarrow d_\phi \rightarrow c_j$.

Данного вида закономерности преобразований ситуаций ПС определяются АИА в процессе пассивного наблюдения за произвольными устойчиво повторяющимися в среде событиями $d_\phi \in D$ следующим образом. Если АИА неоднократно наблюдает, что после появления в ПС сигнала a_i^{Π} происходит событие d_ϕ , которое приводит к появлению в ней необходимого условного a_i^* или заданного ему безусловного сигнала c_j , то данная закономерность запоминается и используется в процессе дальнейшего целесообразного поведения. Если же у АИА возникает необходимость в появлении пускового сигнала a_i^{Π} в изменившихся условиях функционирования, то он формирует условную программу поведения, позволяющую обеспечить его включение в текущих условиях ПС.

Резюмируя вышеизложенное, следует отметить, что необходимость в разработке процедур планирования целесообразного поведения АИА в условиях неопределенности обусловлено тем, что априори описать подробным образом модель реальной проблемной среды, позволяющей организовать целенаправленную деятельность на основе известных принципов построения интеллектуальных решателей задач [16—19], практически не представляется возможным. Следовательно, автономные интеллектуальные

системы различного назначения следует наде-
лать возможностью пополнять свои знания пу-
тем изучения закономерностей ПС в процессе
активного с ней взаимодействия.

Заключение

1. Проведенные исследования показали, что
предложенные процедуры планирования целе-
сообразного поведения АИА в априори неопи-
санных условиях ПС, согласно утверждениям
1 и 2, обладают полиномиальной сложностью
второго порядка, что обеспечивает возмож-
ность их эффективной реализации на борто-
вой ЭВМ различных по назначению автоном-
ных интеллектуальных систем, обладающей,
как правило, ограниченными вычислительны-
ми ресурсами.

2. Разработанные процедуры планирования
целесообразного поведения и инструментальные
средства, позволяющие определять связь между
различными наблюдаемыми в ПС сигналами,
наделяют АИА способностью изучать законо-
мерности преобразования ситуаций сложной
априори неописанной, нестабильной ПС и на
этой основе организовывать эффективным обра-
зом целенаправленную деятельность в реальных
ПС. Это, в свою очередь, позволяет создавать
автономные интегральные роботы наземного
и подводного назначения, а также автономные
беспилотные летательные аппараты, оснащен-
ные манипулятором, которые могут выполнять
достаточно сложные задания в реальных ПС.

3. Дальнейшее развитие полученных в работе
результатов может быть связано с разработкой
алгоритмов самообучения, обладающих, в от-
личие от известных алгоритмов [9], инструмен-
тальными средствами, обеспечивающими ими-
тацию отработки пробных действий в процессе
изучения закономерностей ПС, характеризую-
щихся различными сложными формами взаи-
мосвязи между собой условных и безусловных
сигналов. Целесообразно также рассмотреть
возможность организации на основе предло-
женных процедур планирования инструмен-
тальной деятельности АИА в сложных ПС.

4. Формализация мыслительных актов на-
глядно-действенного мышления живых систем
на основе знаков сигнального принципа управ-
ления целесообразным поведением является
одним из перспективных направлений созда-
ния АИА различного назначения, способных

эффективным образом действовать и решать
поставленные перед ними задачи в нестабиль-
ных, априори неописанных условиях функ-
ционирования.

Список литературы

1. Поддьяков Н. Н. Мышление дошкольника. М.: Педа-
гогика, 1977. 277 с.
2. Анохин П. К. Системные механизмы высшей нервной
деятельности. М.: Наука, 1979. 454 с.
3. Поспелов Д. А. Моделирование рассуждений. Опыт
анализа мыслительных актов. М.: Радио и связь, 1989. 184 с.
4. Kober J., Peters J. Learning Motor Skills: From Algo-
rithms to Robot Experiments. Cham: Springer, 2014. 201 p.
5. Kelly A. Mobile Robotics: Mathematics, Models, and
Methods. Cambridge: Cambridge University Press, 2013. 808 p.
6. Мелехин В. Б. Модель представления и получения
новых знаний автономным интеллектуальным роботом на
основе логики условно-зависимых предикатов // Известия
РАН. Теория и системы управления. 2019. № 5. С. 87–107.
7. Берштейн Л. С., Мелехин В. Б. Организация знако-
сигнального принципа управления целесообразным поведе-
нием интегрального робота // Автоматика и телемеханика.
1991. № 12. С. 118–127.
8. Шингаров Г. Х. Условные рефлексы и проблема знака
и значения. М.: Наука, 1986. 200 с.
9. Берштейн Л. С., Мелехин В. Б. Планирование поведе-
ния интеллектуального робота. М.: Энегтоатомиздат, 1994. 240 с.
10. Кулинич А. А., Карпов В. Э., Карпова И. П. Социальные
сообщества роботов. М.: URSS, ООО "ЛЕНАНД", 2019. 352 с.
11. Саймон Д. Алгоритмы эволюционной оптимизации.
М.: ДМК Пресс, 2020. 940 с.
12. Берштейн Л. С., Мелехин В. Б. Планирование поли-
фазного поведения самоорганизующихся интеллектуальных
систем // Известия РАН. Теория и системы управления.
2000. № 5. С. 62–65.
13. Мелехин В. Б., Хачумов М. В. Нечеткие семанти-
ческие сети как адаптивная модель представления знаний
автономных интеллектуальных систем // Искусственный
интеллект и принятие решений. 2020. № 3. С. 61–72.
14. Растринин Л. А., Риппа К. К. Автоматная теория слу-
чайного поиска. Рига: Зинатне, 1973. 340 с.
15. Брайнес С. Н., Напалков А. Н., Свечинский В. Б.
Нейрокибернетика. М.: Госмедиздат, 1962. 172 с.
16. Норвиг П. Искусственный интеллект. Современный
подход. Руководство. М.: Вильямс. 2015. 1408.
17. Kilani Y., Bsoul M., Alsarhan A., Al-Khasawneh A. A.
Survey of the Satisfiability-Problems Solving Algorithms // Intern.
J. Advanced Intelligence Paradigms. 2013. Vol.5, N.3. P. 233–256.
18. La Valle S. M. Planning Algorithms. Cambridge: Cam-
bridge University Press, 2006. 844 p.

Instrumental Means for Managing the Rational Behavior of Self-Organizing Autonomous Intelligent Agents

V. B. Melekhin, pashka1602@rambler.ru,

Dagestan State Technical University, Makhachkala, 367015, Russian Federation,

M. V. Khachumov, khmike@ribox.ru,

Federal Research Center "Informatics and Control" Russian Academy of Sciences,

Moscow, 117312, Russian Federation,

People's Friendship University of Russia, 117198, Moscow, Russian Federation

Corresponding author: **Melekhin Vladimir B.**, D. Sc., Professor,
Makhachkala, 367015, Russian Federation, e-mail: pashka1602@rambler.ru

Accepted on January 12, 2021

Abstract

We formulate the basic principles of constructing a sign-signal control for the expedient behavior of autonomous intelligent agents in a priori undescribed conditions of a problematic environment. We clarify the concept of a self-organizing autonomous intelligent agent as a system capable of automatic goal-setting when a certain type of conditional and unconditional signal — signs appears in a problem environment. The procedures for planning the expedient behavior of autonomous intelligent agents have been developed, that imitate trial actions under uncertainty in the process of studying the regularities of transforming situations in a problem environment, which allows avoiding environmental changes in the process of self-learning that are not related to the achievement of a given goal. Boundary estimates of the proposed procedures complexity for planning expedient behavior are determined, confirming the possibility of their effective implementation on the on-board computer of the automatic control system for the expedient activity of autonomous intelligent agents. We carry out an imitation on a personal computer of the proposed procedures for planning purposeful behavior, confirming the effectiveness of their use to build intelligent problem solvers for autonomous intelligent agents in order to endow them with the ability to adapt to a priori undescribed operating conditions. The main types of connections between various conditional and unconditional signal — signs of a problem environment are structured, which allows autonomous intelligent agents to adapt to complex a priori undescribed and unstable conditions of functioning.

Keywords: autonomous intelligent agent, purposeful behavior, conditions of uncertainty, patterns of problem environment, self-learning, elementary act of behavior

Acknowledgements: The study was carried out with partial financial support from the Russian Foundation for Basic Research within the framework of scientific projects no 17-29-07003 ofi_m; no 18-07-00025 a.

For citation:

Melekhin V. B., Khachumov M. V. Instrumental Means for Managing the Rational Behavior of Self-Organizing Autonomous Intelligent Agents, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2021, vol. 22, no. 4, pp. 171–180.

DOI: 10.17587/mau.22.171-180

References

1. Podd'yakov N. N. Preschooler thinking, Moscow, Pedagogika, 1977, 277 p. (in Russian).
2. Anohin P. K. Systemic mechanisms of higher nervous activity, Moscow, Nauka, 1979, 454 p. (in Russian).
3. Pospelov D. A. Modeling reasoning. Experience in the analysis of mental acts, Moscow, Radio i svyaz', 1989, 184 p. (in Russian).
4. Kober J., Peters J. Learning Motor Skills: From Algorithms to Robot Experiments, Cham, Springer, 2014, 201 p.
5. Kelly A. Mobile Robotics: Mathematics, Models, and Methods, Cambridge, Cambridge University Press, 2013, 808 p.
6. Melekhin V. B. Model of Representation and Acquisition of New Knowledge by an Autonomous Intelligent Robot Based on the Logic of Conditionally Dependent Predicates, *Izvestiya Rossijskoj akademii Nauk. Teoriya i Sistemy Upravleniya*, 2019, no 5, pp. 87–107.
7. Bershtejn L. S., Melekhin V. B. Organization of a sign-signal principle for controlling the expedient behavior of an integral robot, *Avtomatika i Telemekhanika*, 1991, no 12, pp. 118–127 (in Russian).
8. Shingarov G. H. Conditioned reflexes and the problem of sign and meaning, Moscow, Nauka, 1986, 200 p. (in Russian).
9. Bershtejn L. S., Melekhin V. B. Planning the behavior of an intelligent robot, Moscow, Enegoatomizdat, 1994, 240 p. (in Russian).
10. Kulinich A. A., Karpov V. E., Karpova I. P. Social communities of robots, Moscow, URSS, LENAND, 2019, 352 p. (in Russian).
11. Sajmon D. Algorithms of evolutionary optimization, Moscow, DMC Press, 2020, 940 p. (in Russian).
12. Bershtejn L. S., Melekhin V. B. Planirovanie polifaznogo povedeniya samoorganizuyushchihsysa intellektual'nyh system, *Izvestiya Rossijskoj akademii Nauk. Teoriya i Sistemy Upravleniya*, 2000, no 5, pp. 62–65 (in Russian).
13. Melekhin V. B., Hachumov M. V. Fuzzy semantic networks as an adaptive model for representing knowledge of autonomous intelligent systems, *Iskusstvennyj Intellekt i Prinyatie Reshenij*, 2020, no 3, pp. 61–72 (in Russian).
14. Rastrigin L. A., Rippa K. K. Automata theory of random search, Riga, Zinatne, 1973, 340 p.
15. Brajnes S. N., Napalkov A. N., Svechinskij V. B. Neurocybernetics, Moscow, Gosmedizdat, 1962, 172 p. (in Russian).
16. Norvig P. Artificial intelligence. Modern approach, Leadership, Moscow, Vil'yams. 2015, 1408 p. (in Russian).
17. Kilani Y., Bsoul M., Alsarhan A., Al-Khasawneh A. A. Survey of the Satisfiability-Problems Solving Algorithms, *Intern. J. Advanced Intelligence Paradigms*, 2013, no. 3, vol. 5, pp. 233–256.
18. La Valle S. M. Planning Algorithms, Cambridge, Cambridge University Press, 2006, 844 p.