

УДК 004.896

МЕТОД ПОСТРОЕНИЯ СЕМИОТИЧЕСКОЙ СРЕДЫ ФУНКЦИОНИРОВАНИЯ ГРУППОЙ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ

А.А. Кулинич (*kulinich@ipu.ru*)Институт проблем управления РАН, Москва
НИЦ «Курчатовский институт», Москва

Аннотация. Предложен метод построения семиотической среды функционирования группы агентов, основанный на выделении и структуризации возможных классов состояний динамической системы «Группа роботов - Среда» и обозначения этих классов именами-символами. Предложенный метод позволяет группе агентов сформировать согласованное множество имен классов состояний среды функционирования, собственные понятийные системы, что позволяет реализовать их командное поведение для достижения общей цели.¹

Ключевые слова: Роботы, агенты, среда функционирования, имена-символы, семиотическая среда, кооперация.

Введение

Интеллектуального робота (агента) можно определить перечислением его следующих способностей [Wooldridge, 1994]: 1) *Автономность* – способность действовать самостоятельно; 2) *Реактивность* – способность реагировать на изменение состояний среды функционирования; 3) *Проактивность* – способность проявлять инициативу для достижения поставленных целей; 4) *Социальность* – способность взаимодействовать и договариваться с другими агентами для достижения общей цели.

Реализуются эти способности в рамках различных архитектур интеллектуальных агентов, одна из которых – это BDI-архитектура (*Belief-Desire-Intention*) [Rao, 1995]. BDI архитектура – это ментальная архитектура агента, в рамках которой определяются три составляющие: убеждения (*Belief*) – это знания агента о предметной области, желания

¹ Работа выполнена при поддержке грантов РФФИ, проекты: ОФИ_м № 16-29-04412 и № 15-01-07900 А.

(*Desire*) - это цели агентов и намерения (*Intention*) – это возможные действия агентов для достижения поставленных целей.

Вопросы командного поведения групп интеллектуальных агентов с BDI архитектурой исследовались в работе [Cohen, 1991], в которой вводятся абстрактные принципы (спецификации) организации совместной работы группы агентов на основе их общих намерений, которые выражаются через обязательства, принимаемые агентами при их совместной работе. В другой работе [Grosz, 1996] изложена теория общих планов, основными понятиями которой является понятие группового плана, для выполнения которого агенты должны прийти к соглашению о действиях, которые они будут выполнять, реализуя этот групповой план.

В обзорных работах [Городецкий, 2011] приводятся основные теоретические модели, языки и спецификации командного поведения автономных интеллектуальных агентов. Все эти исследования и разработки в основном направлены на решение конкретных задач группового поведения в рамках классических работ [Cohen, 1991] и [Grosz, 1996]. При этом системные вопросы формирования и функционирования групп автономных агентов и их самоорганизации, как правило, не рассматриваются. Такие вопросы обычно рассматриваются для больших команд агентов с простой реактивной архитектурой (рой, стая) [Карпов, 2016]. Однако, вопросы, например, социального поведения групп интеллектуальных агентов с BDI архитектурой мало изучены.

Сложность моделирования групповой работы интеллектуальных агентов с BDI-архитектурой заключается в том, что у каждого BDI-агента имеется индивидуальная модель мира (*убеждения, знания*), которая определяет его поведение (*намерения*) и может меняться по мере накопления агентом опыта. Элементы этой архитектуры (*убеждения, цели и действия*) выражаются не в числовом, а в символическом виде. По сути, формальное описание всех составляющих архитектуры агента – это его знания о предметной области задаются математическими структурами, элементами которых являются множества понятий, их признаки, значения признаков и множество отношений на всех этих множествах.

Формальное описание знаний агента, несмотря на то что, в настоящее время существует языки описания знаний агентов (агентов) при моделировании кооперативной работы больших групп неоднородных интеллектуальных агентов достаточно трудоемко.

В этих случаях, интерес представляют методы и алгоритмы самостоятельного построения агентами своих моделей знаний о предметной области, на основе опыта работы в этой области и некоторых базовых принципов их построения, представленных, например, в виде метаонтологии (самых общих знаний, правил, методов и т.д.).

В этой работе предлагается метод построения семиотической среды функционирования автономными агентами. Суть метода заключается в том, что в качестве моделей знаний о предметной области используются концептуальные каркасы предметной области, грубо описывающие понятийную структуру предметной области, построение которых не вызывает трудностей [Кулинич, 2014].

1. Метаонтология интеллектуального агента

В программной когнитивной архитектуре реактивного агента его когнитивные функции: убеждения, целеполагание, предпочтения агента, реализуются разработчиком агента, как некоторые метазнания, в виде алгоритмов функционирования агента и, фактически, отражают представления разработчика о закономерностях предметной области.

Интеллектуальные агенты с BDI архитектурой по сравнению с реактивными агентами обладают большей свободой выбора поведения агента. Элементом архитектуры, обеспечивающим эту свободу, являются метазнания, представленные в виде метаонтологии, в которой определены самые общие знания о предметной области, модели, правила и др.

Основным элементом метаонтологии агента являются его знания о среде функционирования. Обычно это логическая модель представления знаний, основанная на логике предикатов первого порядка. В этой работе в качестве системной модели знаний о среде функционирования будет рассмотрена динамическая модель многоагентной системы [Каляев и др., 2009]. Здесь группа роботов $R_i, i=1, \dots, N$, каждый из которых характеризуется вектором состояний $R_i=(r_{i1}, \dots, r_{in})$ (вектором значений его характеристик) находится в среде E , которая также характеризуется вектором состояния, $E=(e_1, \dots, e_n)$. Пара векторов $S=\langle R, E \rangle$ характеризует состояние системы «группа роботов – среда» в пространстве состояний (SS), которое определяется как прямое произведение параметров всех агентов R_i , объектов и состояний среды (E), т.е., $SS=\prod_i R_i \times E$.

Пространство состояний SS может интерпретироваться как признаковое семантическое пространство, в котором объекты реального мира, различные ситуации представляются как понятия. В семантическом пространстве объекты (ситуации), по сути, денотаты, определяются именами и векторами значений признаков, определяющие их содержание (смысл) и могут интерпретироваться как знаки-символы в определении Г. Фреге [Бирюков, 1960]. Знак определяется тройкой: имя, смысл и значение знака. Имя – это символ (образ, икона) обозначающее объект реального мира, смысл определяет свойства этого объекта, а значение – это и есть сам объект – денотат. В семантических пространствах реальный объект

(денотат) имеет имя (символ) и представляется в признаковом пространстве как точка, координаты которой определяют значение признаков (смысл).

Определим множество $F=\{f_i\}$ как множество всех свойств (признаков) всех агентов и объектов, а Z_i как возможные значения этих свойств. Тогда среда функционирования определится как прямое произведение значений, т.е. $SF=\prod_i Z_i$, а агенты и объекта как точки в этом пространстве.

В рамках этой системной модели знаний метаонтология каждого агента включает: основные понятия предметной области $D=\{d^H\}$; свойства (признаки) этих понятий и их возможные значения (Z_i) - $F=\{f_i, Z_i\}$; тип шкалы значений признаков {числовые, качественные, номинальные, отношений}; типы отношений, заданных на множестве понятий предметной области и их признаков (это отношения «Род-Вид», («Класс-Подкласс»), «Причинно-следственные» и др.); закономерности предметной области, представленные как отношения «Род-Вид» множестве понятий предметной области, т.е. $R\subseteq D\times D$, и причинно-следственные отношения на множестве значений признаков, т.е. $W\subseteq\prod_i Z_i$.

Метаонтология определяет основные элементы модели знаний о предметной области каждого агента из которых агентами могут быть построены их индивидуальные модели предметной области, в зависимости от их опыта работы и индивидуальных возможностей.

Для реализации социальной способности интеллектуальных агентов (их коммуникации), каждый агент строит свою модель знаний о предметной области, опираясь на правила, определенные в метаонтологии. В этой работе эти правила задает качественный концептуальный каркас этой области [Кулинич, 2014], который, по сути, является идеализированной онтологией этой предметной области.

Формально качественный концептуальный каркас – это частично упорядоченное множество имен классов состояний агента (робота) в среде функционирования, $(\{d^H\}, \leq)$. При этом каждое имя (d^H) определяет подпространство $SS(d^H)\subseteq\prod_i Z_i$, определяющее этот класс состояний.

Так как, модель знаний любого агента является частью качественного концептуального каркаса предметной области, то и коммуникация между агентами будет осуществляться с помощью искусственных имен (d^H) классов состояний концептуального каркаса и будет понятна всем агентам. Концептуальный каркас предметной области в дальнейшем будем называть семиотической средой функционирования.

2. Алгоритм построения семиотической среды функционирования

Пусть названия признаков (имена) множества признаков среды функционирования $\{f_i\} \in F$ агентов являются элементами метаонтологии и заложены в память агентов и определены ресурсы, которые позволяют агенту изменять значения признаков среды функционирования. Под ресурсом агента понимаем возможность агента изменять значения своих свойств (признаков), свойств других агентов или объектов.

То есть, ресурс по (j -у свойству) признаку f_{ij} i -о агента это пара $(f_{ij}; Z_{ij}^R)$, где $Z_{ij}^R = \{z_{ie}, \dots, z_{iq}\} \subseteq Z_j$ – строго упорядоченное множество значений. Если каждый агент может изменить значения нескольких своих свойств, то тогда его ресурсы будут заданы вектором, например, для первого агента – это вектор: $(f_{11}; Z_{11}^R), \dots, (f_{1m}; Z_{1m}^R)$.

Считаем, каждый из N агентов знает о своих ресурсах:

$$\begin{aligned} &(f_{11}; Z_{11}^R), \dots, (f_{1m}; Z_{1m}^R); \\ &\dots\dots\dots \\ &(f_{n1}; Z_{n1}^R), \dots, (f_{nm}; Z_{nm}^R). \end{aligned} \tag{1}$$

Задача заключается в том, чтобы по имеющейся у агентов информации о ресурсах построить концептуальный каркас среды функционирования $(\{d^H\}, \leq)$, определить множество имен классов состояний $(\{d^H\})$, соответствующие им подпространства $SS(d^H)$, т.е. определить семиотическую среду функционирования и свою понятийную систему в терминах искусственных имен этой среды.

Эта задача решается алгоритмически и включает ряд шагов, основанных на обмене информацией через «доску объявлений».

Алгоритм включает следующие шаги, выполняемые каждым агентом:

1. Передача информации о собственных ресурсах на «доску объявлений» для определения пространства их возможного совместного функционирования;
2. Определение базового класса состояний концептуального каркаса среды функционирования;
3. Построение концептуального каркаса среды функционирования и символизация классов состояний (определение искусственных имен);
4. Определение агентами собственной понятийной системы в терминах искусственных имен семиотической среды функционирования.

На первом шаге алгоритма агенты помещают информацию о своих ресурсах на «доску объявлений». Для определения пространства их возможного совместного функционирования агенты, по имеющейся информации (1) о ресурсах каждый агент объединяет эти значения. Т.е.

$$(f_{\cup 1}; \bigcup_{i=1}^n Z_{i1}^R), \dots, (f_{\cup m}; \bigcup_{i=1}^n Z_{im}^R). \quad (2)$$

Определение 1. Средой возможного кооперативного взаимодействия агентов будем называть прямое произведение объединения их ресурсов,

т.е. $\times \bigcup_{j=1}^n Z_{ij}^R$.

На **втором шаге** каждый агент определяет базовый класс состояний концептуального каркаса среды. Для этого, агенты определяют границы (максимальное и минимальное значение) одноименных ресурсов и находят среднее значение для каждого ресурса, полученных от всех других агентов.

Далее мы будем считать, что существует отображение значений объединенных ресурсов, представленных в порядковых или номинальных шкалах, на отрезок числовой оси $[0,1]$, т.е. $\varphi: (\bigcup_{i=1}^n Z_{ij}^R) \rightarrow [0,1]$.

Тогда, среднее значение каждого из объединенных ресурсов агенты определяют по очевидной формуле:

$$z_{ij} = (\max(\varphi(\bigcup_{i=1}^n Z_{ij}^R) - \min(\varphi(\bigcup_{i=1}^n Z_{ij}^R)))/2) \quad (3)$$

Границы интервалов значений признаков базового класса каждый агент определяет как положительное и отрицательное отклонение от их средних значений: $\Delta i = z_{ij} \pm \varepsilon$, где $\varepsilon = (10-15\%) z_{ij}$, а сам базовый класс как прямое произведение, полученных интервалов, т.е.

$$SS(d_0) = \times i \Delta i. \quad (4)$$

Отметим, что поскольку все агенты работают по одному и тому же алгоритму определения базового класса состояний, то базовый класс состояний у всех агентов будет одинаков.

В качестве иллюстрации определения базового класса, на рисунке 1 показан базовый класс состояний среды функционирования для двух признаков $X \in [0,1]$ и $Y \in [0,1]$.

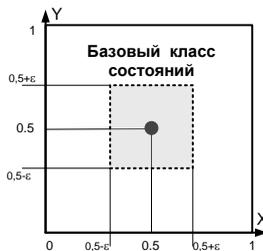


Рис. 1. Базовый класс состояний среды функционирования

Следующий шаг алгоритма – **построение и символизация концептуального каркаса среды функционирования**. Базовый класс состояний – это «идеальная точка» для всех агентов, относительно которой будет строиться концептуальный каркас, и символизироваться среда функционирования.

Общий принцип символизации среды функционирования заключается в качественном кодировании значений признаков, в виде множества символов, например, $Z_i^O = \{0, 1, 2\}$. Качественные значения обозначают интервалы значений признаков и определяются по правилу (5).

$$Z_i^O = \begin{cases} 0, & \text{если } z_i \in [z_{ie}^b; z_{iq}^b] \\ 1, & \text{если } \max(Z_{ij}^{R\cup}) \geq z_i > z_{iq}^b; \\ 2, & \text{если } \min(Z_{ij}^{R\cup}) \leq z_i < z_{ie}^b. \end{cases} \quad (5)$$

Т.е., значения признака $z_i \in Z_i$ попадающие в интервал значений признака базового понятия кодируются цифрой «0», превышающее значение базового интервала, цифрой «1», а если значение признака меньше базового значения, то, соответственно, «2».

Построение концептуального каркаса заключается в определении имен классов состояний среды функционирования и подпространств среды функционирования и соответствующих этим именам.

Определение 2. Именем класса состояний качественной среды функционирования будем называть конкатенацию подмножеств прямого произведения качественных значений всех признаков среды функционирования, т.е. d^{name} , где $name = K(\{*\})$, K – операция конкатенации, $\{*\} \in \times_i Z_i^O$.

Например, есть два признака, качественные значения которых соответственно, $Z_1^O = \{0, 1, 2\}$ и $Z_2^O = \{0, 1, 2\}$. Их прямое произведение ($Z_1^O \times Z_2^O$) равно: $\{\{0,0\}; \{0,1\}; \{0,2\}; \{1,0\}; \{1,1\}; \{1,2\}; \{2,0\}; \{2,1\}; \{2,2\}\}$, а конкатенация подмножеств этого произведения равна: $\{\langle 00 \rangle; \langle 01 \rangle; \langle 02 \rangle; \langle 10 \rangle; \langle 11 \rangle; \langle 12 \rangle; \langle 20 \rangle; \langle 21 \rangle; \langle 22 \rangle\}$. Имена классов состояний, соответственно, равны: $\{d^{00}; d^{01}; d^{02}; d^{10}; d^{11}; d^{12}; d^{20}; d^{21}; d^{22}\}$.

Элементы прямого произведения качественных значений Z_i^O признаков среды функционирования однозначно определяют подпространства, определяющие обобщенные классы состояний среды функционирования. Действительно, подмножество $\{0,0\}$ согласно правилу (3) определяет базовые интервалы по первому $[z_{1e}^b; z_{1q}^b]$ и по второму признаку $[z_{2e}^b; z_{2q}^b]$, а их произведение, $SS(d^{00}) = [z_{1e}^b; z_{1q}^b] \times [z_{2e}^b; z_{2q}^b]$ дает нам базовый класс состояний. Для случаев, когда качественное значение одного из признаков принимает 1, при определении подпространства в качестве интервала

берем интервал $[z_{ie}^b; \max(\bigcup_{i=1}^n Z_{ij}^R)]$, а если 2, т.е. значение меньше базового интервала, то $[\min(\bigcup_{i=1}^n Z_{ij}^R); z_{iq}^b]$.

Для подмножества $\{0,1\}$ подпространство, определяющее класс состояний с именем d^{01} равно $SS(d^{01})=[z_{ie}^b; z_{iq}^b] \times [z_{ie}^b; \max(\bigcup_{i=1}^n Z_{ij}^R)]$.

Таким образом, искусственные имена классов состояний определяют подпространства классов состояний среды функционирования и образуют частично упорядоченное множество, т.е. концептуальный каркас среды функционирования.

Определение 3. Качественным концептуальным каркасом среды функционирования агентов будем называть частично упорядоченное множество имен классов состояний среды функционирования ($\{d^H\}, \leq$), однозначно определяющие подпространства этой среды $SS(d^H)$.

Таким образом, агенты, обмениваясь информацией о своих ресурсах, могут построить концептуальный каркас среды функционирования, опираясь на общие правила его структуризации и символизации.

Следующий шаг – построение агентами собственной **понятийной системы** о среде функционирования.

Определение 4. Понятийной системой агента будем называть частично упорядоченное множество имен классов состояний полученное прямым произведением качественных доменов всех его ресурсов, т.е. ($\{d_i^H\}, \leq$), где $d_i^H \Leftrightarrow SS(d^H)$, $SS(d^H) \subseteq \times_j Z_{ij}^R$.

3. Пример

Пусть агенты *Агент 1* и *Агент 2* знают, что они находятся на плоскости $\{X, Y\}$ и могут изменить свое положение, изменив свои координаты и состояние среды функционирования.

У агента 1 есть ресурс, позволяющий ему перемещаться только по координате X, (горизонтально): $X_1 = \{x_{1i}\}$, $x_{1i} \in [0,1]$, $Y_1 = y_{1j}$, $y_{1j} = \text{const} = 0,5$. У агента 2 есть ресурс, позволяющий ему перемещаться только по координате Y, (вертикально): $Y_2 = \{y_{2j}\}$, $y_{2j} \in [0,1]$, $X_2 = x_{2i}$, $x_{2i} = \text{const} = 0$ (Рис. 2).

Объединенные ресурсы агентов: $Y_1 \cup Y_2 = [0,1]$ и $X_1 \cup X_2 = [0,1]$, т.е ресурсы агентов это вектор: $(f_{Y}; [0,1], f_{X}; [0,1])$, а среда возможного совместного функционирования агентов это произведение: $[0,1] \times [0,1]$.

Агенты определили базовый класс состояний по формулам 3 и 4, показанный на рис. 1 и символизировали среду функционирования согласно правилу (5). Пример символизации приведен в определении 2 и показан на рис. 2.

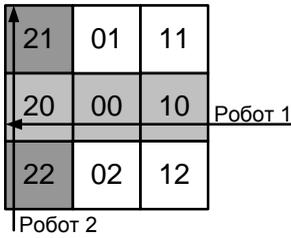


Рис. 2. Понятийные системы агента 1 и агента 2

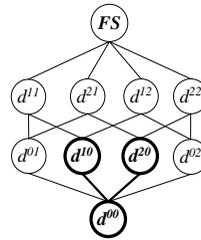


Рис. 3. Диаграмма Хассе понятийной системы агента 1

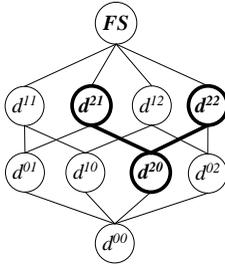


Рис. 4. Диаграмма Хассе понятийной системы агента 1

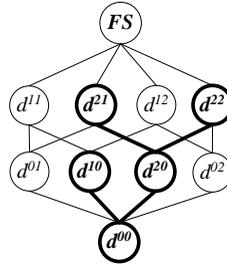


Рис. 5. Диаграмма Хассе области совместного функционирования агентов 1 и 2

Далее агенты определяют собственные понятийные системы. Агент 1 по координате Y может изменять значения этого параметра в пределах базового класса, т.е. $Z_Y^{Q1} = \{0\}$, а по координате X он может изменять значение этого параметра больше и меньше значения базового класса, т.е. качественные значения его ресурсов равны: $Z_X^{Q1} = \{0,1,2\}$. Тогда знания этого агента о классах среды функционирования определим: $Z_X^{Q1} \times Z_Y^{Q1} = \{0,1,2\} \times \{0\} = \{00, 10, 20\}$, или в именах классов состояний $\{d^{00}, d^{10}, d^{20}\}$.

Агент 2 по координате X может находиться в состояниях, когда его значения меньше значения X базового класса, $Z_X^{Q2} = \{2\}$, а по координате Y ресурсы позволяют ему переходить в классы $\{0,1,2\}$, т.е. $Z_Y^{Q2} = \{0,1,2\}$. Тогда его знания определяться так: $Z_X^{Q2} \times Z_Y^{Q2} = \{2\} \times \{0,1,2\} = \{20, 21, 22\}$, или в именах классов состояний $\{d^{20}, d^{21}, d^{22}\}$.

Символизованные понятийные системы агента (агента) 1 и 2 показаны на рис. 2. Диаграммы Хассе понятийных систем агента 1 и 2 показаны на рис. 3 и рис. 4. Диаграмма Хассе области возможного совместного функционирования показана на рис. 5.

Заключение

В работе предложен метод построения качественной семиотической среды функционирования группой агентов, основанный на выделении и структуризации возможных классов состояний динамической системы «Группа роботов - Среда». Классы состояний обозначены именами-символами и образуют частично упорядоченное множество имен, названное семиотической средой. Предложен алгоритм, позволяющий группе агентов сформировать и согласовать множество искусственных имен классов состояний среды и построить собственные понятийные системы - знания об этой среде в терминах этих имен. Предложенные метод и алгоритм являются частью исследований [Кулинич, 2016], направленных на моделирование социального поведения групп агентов с BDI-архитектурой, функционирующих в семиотической среде.

Список литературы

- [Wooldridge, 1994] M.Wooldridge and N.R.Jennings. Agent Theories, Architectures, and Languages: A Survey. In: Intelligent Agents. ECAI-94 Workshop on Agent Theories, Architecture and Languages. Amsterdam, The Netherlands, August 8-9, 1994, Proceedings. Springer Verlag: 3-39, 1994.
- [Rao, 1995] Rao A.S., Georgeff M.P. BDI Agents: From Theory to Practice // Proc. First International Conference on Multi-Agent Systems (ed. V.Lesser). – AAAI Press/The MIT Press, 1995. – P. 312–319.
- [Cohen, 1991] Cohen P., Levesque H.J. Teamwork. *Nous*, 25(4), (1991) Special Issue on Cognitive Science and Artificial Intelligence, P. 487-512.
- [Grosz, 1996] Grosz B., Kraus S. Collaborative Plans for Complex Group Actions // *Artificial Intelligence*. – 1996. – №86. – P. 269–358.
- [Городецкий, 2011] Городецкий В.И. Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор (Часть 1, Часть 2) // *Искусственный интеллект и принятие решений*. – 2011. – №2, №3 – С. 19–30, С. 34–47.
- [Карпов, 2016] Карпов В. Э. Модели социального поведения в групповой робототехнике / Управление большими системами. Выпуск 59. М.: ИПУ РАН, 2016. С.165-232.
- [Кулинич, 2014] Кулинич А.А. Концептуальные каркасы онтологий слабо структурированных предметных областей // *Искусственный интеллект и принятие решений*. 2014. - № 4. - С. 31-41
- [Каляев и др., 2009] Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов / М.: Физматлит, 2009.-279 с.
- [Бирюков, 1960] Бирюков Б.В. Теория смысла Готлоба Фреге / В кн.: Применение логики в науке и технике. – М.: Изд-во АН СССР, 1960. – С. 502–555.
- [Кулинич, 2016] Кулинич А.А. Модель командной работы агентов с BDI архитектурой. 15-я национальная конференция по искусственному интеллекту

КИИ-2016 (3-7 октября 2016 г., г. Смоленск, Россия): Труды конференции. В 3-х томах. Т.2. – Смоленск: Универсум, 2016. - с. 344-352.