

УДК 004.896

**Модель коммуникации агентов на основе
концептуальных каркасов среды функционирования¹**
**Model of communications of agents on the basis of conceptual
templates of the environment of functioning**

А.А. Кулинич²
A.A. Kulinich

// Конгресс по интеллектуальным и информационным технологиям IS&IT' 15. Труды конгресса по интеллектуальным системам и информационным технологиям «AIS-IT' 15», 2-9 сентября. Россия, Краснодарский край, пос. Дивноморское. Научное издание в 3-х томах. – Таганрог.: Из-во ЮФУ, 2015 – Т.1. – стр. 140-147.

Представлена модель коммуникаций агентов (роботов) на основе структуризации их среды функционирования в виде качественного концептуального каркаса. В рамках этой модели формально определены: возможность и необходимость коммуникации, согласования убеждений агентов, а также условия для совместной работы для достижения агентами их целей. Многоагентная система, коммуникация агентов, концептуальный каркас, кооперация агентов

The model of communications of agents (robots) on the basis of structuration of their functioning environments in the form of a qualitative conceptual template is presented. Within the limits of this model are formally defined: possibility and necessity of communications, the coordination of belief of agents, and also teamwork conditions for achievement by agents of their purposes.

Multiagent system, communications of agents, a conceptual template, cooperation of agents

¹ Работа выполнена при частичной финансовой поддержке РФФИ (проект № №15-01-07900 и проект № 14-01-00422)

² 117997, Москва, ул. Профсоюзная 65, Институт проблем управления РАН, e-mail: kulinich@ipu.ru; alexkul@rambler.ru

Введение

Интерес к моделированию коллективного поведения агентов появился в 60-е годы прошлого столетия. Так в работах [1-3], на основе изучения коллективного поведения конечных автоматов были заложены теоретические основы кооперативного взаимодействия агентов.

Агент с простой реактивной архитектурой (стимул-реакция) реагирует некоторым действием на любое изменение среды, в которой он находится.

С помощью простых агентов реализуется, как правило, несложное командное поведение агентов – это образование стаи или роя агентов [4]. Исследуются разнообразные алгоритмы обхода стаей (роем) агентов препятствий, уклонение от столкновения [5] и т.д. Однако, агенты с реактивной архитектурой могут решать и более сложные задачи, если определены правила их поведения в среде и модели среды их функционирования. Так, например, в работе [6] представлена математическая модель командной работы агентов на основе модели поведения людей в малых социальных группах. В основу упомянутой модели положены работы социальных психологов Д. Хоманса [7] и Л. Фестингера [8]. В рамках этой модели, на основе критериев полезности агентов и взаимного диссонанса разработаны простые правила поведения агентов в среде, отражающие закономерности предметной области. Разработана имитационная модель игры в виртуальный футбол, позволяющая моделировать и исследовать командное и эгоистичное поведение агентов.

Для моделирования более сложного и интеллектуального поведения агентов в рамках теории многоагентных систем был предложен ряд теоретических концепций, определяемых как теории командной работы агентов. Это теория общих намерений [9] и теория общих планов [10]. Эти две теории командной работы ориентированы на *BDI (Belief-Desire-Intention)* [11] архитектуру агентов, которая считается основной при реализации интеллектуальных агентов. В рамках *BDI* архитектуры агентов описываются «ментальные» состояния агентов в терминах их убеждений, желаний (целей) и намерений. Задача теорий командной работы заключается в том, чтобы построить модели и методы согласования убеждений, желаний множества агентов, а также намерений для реализации их совместной работы. Естественно такое согласование требует многочисленных коммуникаций между агентами.

Теория общих намерений основана на двух типах базовых понятий. Первый тип понятий – это ментальные понятия: события (отражают состояние внешнего мира); убеждения (отражают достоверные знания агента о внешнем мире); цель агента; взаимные убеждения (отражающие общие достоверные знания множества агентов о внешнем мире). Понятия

этого уровня ответственны за отражение действительности в терминах убеждений агента. Второй тип понятий – это понятия для описания поведения агентов: индивидуальные и общие обязательства (это обязательства по достижению индивидуальной или общей цели коллектива агентов); соглашения (отражают условия отказа агентом от своих обязательств); индивидуальные и общие намерения (отражают особенности коллективного взаимодействия агентов). К недостаткам теории общих намерений относят трудности при построении плана работы команды агентов в динамических ситуациях [12].

В теории общих планов [10] основными базовыми понятиями являются понятия группового плана и индивидуальных ментальных понятий агентов (намерений, убеждений и т.д.). Для выполнения группового плана агенты должны прийти к соглашению о действиях, которые они будут выполнять, реализуя групповой план. Таким образом, теория общих планов описывает множество взаимосвязанных намерений и убеждений агентов, работающих в команде.

В теориях общих намерений и общих планов, часто агенты не могут построить общий план, выполняя только частичные планы, многократно пересматривая их. Такой перманентный пересмотр планов возникает в динамических ситуациях, а также при условии противодействия агентов препятствующих командной работе группы агентов [12].

Важность коммуникации между агентами при их коллективном решении сложных задач, привело к появлению множества подходов (протоколов коммуникации), основанных на различных принципах.

В этих подходах важным элементом, для обеспечения взаимопонимания между агентами являются модели их знаний о предметной области - онтологии. Если у агентов одинаковые знания о предметной области, то в этом случае они будут понимать друг друга.

Одинаковые онтологии предметной области у разных агентов позволяет им общаться с помощью знаков, которые, по сути, являются именами понятий их понятийной системы. Знак в онтологии представлен тройкой: имя, смысл, представление (значение) знака. Если у агентов одинаковые онтологии, то при коммуникации достаточно передать агенту только имя знака, которое однозначно определит его смысл и представление.

Важно отметить, что в многочисленных статьях и описаниях языков коммуникации агентов явным образом онтология как элемент архитектуры агента всегда присутствует, однако, конкретных примеров работы этих моделей знаний в процессах коммуникации практически нет.

Интересны вопросы коммуникации агентов (роботов), которые функционируют в условиях неопределенности среды или когда о среде есть неполные сведения.

1. Качественная среда функционирования агентов в условиях неопределенности

Обычно, предполагается, что агенты, выполняя некоторую совместную работу, могут общаться не только между собой (искусственными агентами), и с людьми (естественными агентами). В этом случае, онтологии искусственных агентов и людей должны совпадать. Однако, в случаях, когда агенты решают задачу автономно, без общения с людьми, язык их коммуникации может быть значительно упрощен, оставаясь при этом достаточно выразительным для решения поставленной задачи.

Упростить коммуникации между агентами можно, если сделать некоторые допущения о среде функционирования агентов. Дадим определение среды функционирования агентов. Считаем, что среда функционирования включает множество объектов B (в том числе и агентов A) и пусть известны множества свойств $F = \{f_i\}$ каждого из объектов этой среды.

Пусть для каждого свойства и каждого объекта известны множество возможных значений, $Z = \{Z_i\}$. Считаем, что множество значений свойства объекта – это упорядоченное множество, т.е. $Z_i = \{z_{i1}, \dots, z_{iq}\}$, $z_{iq+1} \succ z_{iq}$, $q = 0 \dots n-1$;

Среду функционирования агентов определим как гиперкуб, полученный прямым произведением значений всех упорядоченных множеств значений всех свойств объектов, $SF = \times_i Z_i$. При таком определении среды функционирования агентов, каждый ее объект (в том числе и агенты) могут быть представлены в виде точек с координатами – значениями их свойств (z_{1e}, \dots, z_{nq}) , т.е. $(z_{1e}, \dots, z_{nq}) \in \times_i Z_i$.

Состоянием среды функционирования будем называть вектор значений всех свойств объектов среды в некоторый момент времени t . Считаем, что агенты могут изменять значения некоторых своих свойств и значения свойств объектов. Тогда состояние среды будет меняться в случаях таких действий агентов.

Среду функционирования агентов можно интерпретировать как семантическое пространство, которое считается признаковой моделью знаний. В работе [13] было показано, что в семантическом пространстве можно выделить множество подпространств $\{SS(d^H)\}$, $SS(d^H) \subset FS$, удовлетворяющее свойствам рефлексивности, антисимметричности и транзитивности. Они образуют решетку $KK = (\{SS(d^H)\}, \cap, \cup)$, структурирующую это семантическое пространство, которая называется качественным концептуальным каркасом онтологии и отражает идеализированную структуру знаний о предметной области [13]. Все

подпространства $SS(d^H)$ в данном случае интерпретируются как классы состояний среды функционирования и характеризуются тройкой: d^H – имя класса; $SS(d^H)$ – содержание класса и $V(d^H)$ – объем класса – это множество объектов среды, значения свойств которых попадает в подпространство $SS(d^H)$. Мощность множества подпространств, включенных в концептуальный каркас, определится как 3^N , где N – общее количество всех свойств объектов среды функционирования.

Пусть известны некоторые закономерности в предметной области (среды функционирования) и они могут быть заданы в виде множества правил (причинно-следственных отношений), связывающих значения свойств разных объектов этой предметной области. Эти закономерности могут быть представлены в виде отображения: $W: \times_i Z_i \rightarrow \times_i Z_i$.

В работе [13] было показано, что известные закономерности порождают концептуальный каркас $KK^W = (\{SS(d^H)\}^W, \cap, \cup)$, множество классов которого $\{SS(d^H)\}^W$ есть подмножество классов концептуального каркаса $KK = (\{SS(d^H)\}, \cap, \cup)$, т.е. $\{SS(d^H)\}^W \subseteq \{SS(d^H)\}$.

Согласно [13] разным видам отображений W_q соответствуют разные концептуальные каркасы KK^{Wq} , т.е. $W_q \Leftrightarrow KK^{Wq}$.

Таким образом, среда функционирования определяется кортежем: $\langle A \cup B, F, Z, W, KK^W \rangle$, B – множество объектов и A – агентов, F – множество свойств объектов и агентов, $Z = \{Z_i\}$ – множества множеств значений каждого свойства, W – закономерности среды функционирования и ее концептуальный каркас – KK^W .

В рамках представленной модели среды функционирования представим элементы BDI архитектуры агентов, т.е. их убеждения (знания о предметной области), желания (цели) и намерения (действия).

Убеждения агентов – это его знания о среде функционирования. В случае полных знаний о среде, его убеждения выражаются кортежем $\langle A \cup B, F, Z, W, KK^W \rangle$. Однако, знания могут быть неполными и тогда убеждения каждого i -го агента выражаются так: $\langle A_i \cup B_i, F_i, Z_i, W_i, KK_i^W \rangle$, где $A_i \subseteq A$, $B_i \subseteq B$, $F_i \subseteq F$, $Z_i \subseteq Z$, $W_i \subseteq W$, $KK_i^W \subseteq KK^W$, $i = 1, \dots, n$.

Цели агентов выражаются некоторым целевым состоянием среды функционирования $G_i = (g_{i1}, \dots, g_{in})$, $g_{in} \in Z_i$. По сути, цель каждого агента – это точка в пространстве FS и она может быть охарактеризована классом состояний среды функционирования, к которому эта цель принадлежит. В этом случае цель определяется как знак, т.е. тройкой: d^{G_i} – имя класса цели; $SS(d^{G_i})$ – его содержание и $V(d^{G_i}) = (g_{i1}, \dots, g_{in})$ – объем класса цели.

Действия агентов. В приведенном выше определении среды функционирования считается, что агенты способны изменить ее состояние. Напомним, что изменение состояния среды выражается в изменении значений свойств объектов. Активность агентов возможна,

благодаря наличию у них некоторого ресурса $U_i^R=(u_{i1}^R, \dots, u_{in}^R)$, $u_{in}^R \in Z_i$. Изменение состояния среды, i -м агентом осуществляется с учетом его знаний о среде W_i и может быть выражена системой конечно-разностных уравнений: $Z(t+1)=W_i \circ (Z(t) \cup U_i(t))$, где \circ - правило вывода, $Z(t+1)$, $Z(t)=(z_1, \dots, z_q)$ – состояние ситуации – это векторы значений признаков в моменты времени t , $U_i(t)$ – вектор управляющих воздействий в моменты времени t .

В среде функционирования FS ресурсы агента, также могут быть представлены в виде класса состояний среды функционирования и задаваться тройкой: d^{U_i} – имя класса ресурсов; $SS(d^{U_i})$ – его содержание и $V(d^{U_i})=(u_{i1}, \dots, u_{in})$ – объем класса ресурсов.

Формально поиск действий для достижения заданного вектора цели $G_i=(g_{i1}, \dots, g_{in})$ сводится к решению уравнения $G_i=W_i \circ U_i$ относительно управляющих воздействий U_i , т.е. к решению обратной задачи. Решение запишется в следующем виде: $U^*=G_i \circ W_i^{-1}$, где U^* – новые значения свойств среды, позволяющие достичь цели G_i , \circ – процедура обратного вывода. Решение обратной задачи – это множество решений $U^*=\{U_i\}$, где U_i – векторы значений свойств среды, позволяющие достичь цели.

Опишем возможный вариант коммуникации агентов для организации совместной работы в терминах среды их функционирования.

2. Модель коммуникации агентов

Для коммуникации агентов нужно определить язык коммуникации агентов. Обычно язык определяется как знаковая система, предназначенная для передачи информации о предметной области. Естественной возможной знаковой системой может служить онтология предметной области, обычное представление которой – это иерархия понятий, представленных тройкой, определяющей знак: имя понятия, содержание и объем. Идеальный вариант для коммуникации возможен, когда у всех агентов есть одинаковая онтология, которая полностью описывает предметную область на естественном языке.

В условиях неопределенности в качестве онтологии предметной области может использоваться концептуальный каркас предметной области. Концептуальный каркас является знаковой системой (все понятия определены тройкой: имя, содержание и объем), в которой имена понятий (знаков) заменены искусственными именами.

В условиях неопределенности концептуальный каркас онтологии предметной области может быть использован для организации коммуникации между агентами с помощью искусственных имен.

Для организации коммуникации сделаем следующее допущение. Будем считать, что полные или частичные знания о свойствах среды каждого агента заданные вектором F_i , позиция каждого свойства

фиксирована у разных агентов, а недостающие сведения кодируются, например, нулем.

Напомним, что убеждения агентов формально в рамках среды функционирования определяется кортежем: $\langle A_i \cup B_i, F_i, Z_i, W_i, KK^W_i \rangle$. Допустим, имеем двух агентов i и j с убеждениями, соответственно, $\langle A_i \cup B_i, F_i, Z_i, W_i, KK^W_i \rangle$ и $\langle A_j \cup B_j, F_j, Z_j, W_j, KK^W_j \rangle$. Рассмотрим процесс коммуникации двух этих агентов в среде функционирования FS .

Коммуникация между агентами возможна, если в концептуальных каркасах этих агентов существуют общие элементы, т.е. если $KK^W_i \cap KK^W_j = \emptyset$.

Если принципиальная возможность коммуникации имеется, то дальнейшее общение агентов связано с переговорами о совместной работе. Необходимость в совместной работе у агента i , имеющего знания о предметной области W_i , появляется в случае, если при решении им обратной задачи для цели G_i во множество возможных решений U^* для ее достижения не попадает вектор, имеющихся у него ресурсов U_i^R , т.е. если $U_i^R \notin U_i^*$.

В этом случае, агент i передает агенту j информацию об имеющихся ресурсах у него ресурсах U_i^R и об имени класса состояний, к которому принадлежит его цель G_i .

Агент j получив эту информацию, решает обратную задачу на ее основе и основе собственных знаний о закономерностях предметной области W_j . Если решение, в рамках знаний агента j найдено, т.е. если $U_i^R \in U_j^*$, то этот агент инициирует переговоры о согласовании (изменении) знаний W_i агента i . Для этого он передает ему собственные знания о закономерностях предметной области W_j . В этом случае, агент j может достичь цели G_i самостоятельно, если он не занят достижением собственной цели G_j .

Введем еще одну характеристику агента – это возможность достижения и цели за счет собственных ресурсов. Считаем, что в среде функционирования, которую мы определили, как пространство SF определена метрика $\rho(\cdot)$. Возможность достижения цели μ_i i -ым агентом определяется как расстояние между его целью G_i прогнозом достижения цели $Z(t+n)$ в случае применения собственных ресурсов U_i^R , т.е. $\mu_i(U_i^R) = \rho(G_i, Z(t+n))$.

Условием совместной работы агентов i и j будем считать следующее условие: если объединение ресурсов этих агентов увеличивает возможность достижения цели каждого из них, т.е. если $\mu_i(U_i^R + U_j^R) > \mu_i(U_i^R)$ и если $\mu_j(U_i^R + U_j^R) > \mu_j(U_j^R)$, то агенты могут сотрудничать.

Таким образом, для начала совместной работы агенты обмениваются именами классов собственных ресурсов и именами классов своих целей.

Если условие совместной работы выполняются, то агенты начинают переговоры и вырабатывают план совместных действий.

В этой статье представлена модель коммуникаций агентов (роботов) на основе структуризации их среды функционирования в виде качественного концептуального каркаса. В рамках этой модели формально определены: возможность и необходимость коммуникации, согласования убеждений агентов, а также условия для совместной работы для достижения агентами их целей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Варшавский В.И., Поспелов Д.А. Оркестр играет без дирижера. – М: Наука 1984. – 208 с.
2. Стефанюк В.Л., Цетлин М.Л. О регулировке мощности в коллективе радиостанций // Проблемы передачи информации. – 1967. – Т. 3, №4. – С. 59–67.
3. Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. – М.: Наука, 1969. – 316 с.
4. Карпов В.Э. Процедура голосования в однородных коллективах роботов //XIV национальная конференция по искусственному интеллекту с международным участием КИИ-2014 (24-27 октября 2014 г., Казань, Россия): Труды конференции. Т.2. – Казань: Изд-во РИЦ «Школа», 2014, -341 с. 159-167.
5. Павловский В.Е., Кирикова Е.П. Моделирование управляемого адаптивного поведения гомогенной группы роботов // Искусственный интеллект. – 2002. – №4. – С. 596–605.
6. Кулинич А. А. Модель командного поведения агентов (роботов): когнитивный подход / Управление большими системами. Выпуск 51. М.: ИПУ РАН, 2014. С.174-196.
7. Хоманс ДЖ. Социальное поведение как обмен. Современная зарубежная социальная психология. – М.: Издательство Московского университета, 1984. – С. 82–91.
8. Фестингер Л. Теория когнитивного диссонанса. – СПб.: Ювента, 1999. – С. 15-52.
9. Cohen P., Levesque H.J. Teamwork. *Nous*, 25(4), (1991) Special Issue on Cognitive Science and Artificial Intelligence, P. 487-512.
10. Grosz B., Kraus S. Collaborative Plans for Complex Group Actions // *Artificial Intelligence*. – 1996. – №86. – P. 269–358.
11. 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.
12. 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.

13. Кулинич А.А. Концептуальные каркасы онтологий слабо структурированных предметных областей // Искусственный интеллект и принятие решений. 2014. - № 4. - С. 31-41.

Автор: Кулинич Александр Алексеевич, Институт проблем управления Российской академии наук, кандидат технических наук, старший научный сотрудник, e-mail: kulinich@ipu.ru.

Author: Kulinich Alexander Alexseevich, Institute of Control Sciences, Russian Academy of Sciences, Ph.d., senior scientist, e-mail: kulinich@ipu.ru.