

А.А. Кулинич¹

СТАЙНЫЕ АЛГОРИТМЫ ФОРМИРОВАНИЯ И ФУНКЦИОНИРОВАНИЯ КОМАНД АГЕНТОВ*

В работе проанализирован классический стайный алгоритм агентов, выявлены его недостатки при формировании общего плана решения задачи в динамических ситуациях. Предложены стайные алгоритмы на основе критериев возможности самостоятельного достижения агентом цели и их взаимной полезности. На основе этих критериев рассмотрены стайные алгоритмы «ленивых» и «эгоистичных» агентов.

Роботы, агенты, стайный алгоритм, взаимная полезность.

A.A. Kulinich

GREGARIOUS ALGORITHMS OF FORMATION AND FUNCTIONING OF AGENTS COMMANDS

In work the classical gregarious algorithm of robots is analyzed, its lacks at the general plan formation of the problem decision in dynamic situations are revealed . Gregarious algorithms on the basis of criteria of possibility of independent achievement by the agent of the purpose and mutual utility of agents are offered. On the basis of the same criteria gregarious algorithms of "lazy" and "selfish" agents are considered.

Robots, agents, gregarious algorithm, mutual usefulness.

Введение

В настоящее время вопросы формирования и функционирования команд искусственных агентов (роботов) исследуется в рамках в теории

¹ 117997, Москва, ул. Профсоюзная 65, Институт проблем управления РАН,
e-mail: kulinich@ipu.ru; alexkul@rambler.ru

* Работа выполнена при финансовой поддержке РФФ (проект № 16-11-00018), ОВИ_м (проект №16-29-04412).

мультиагентных систем и искусственного интеллекта. Здесь исследуются команды интеллектуальных роботов, обладающих свойствами: 1) *Автономности* – способность действовать самостоятельно; 2) *Реактивности* – способность реагировать на изменение состояний среды функционирования; 3) *Проактивности* – способность проявлять инициативу для достижения поставленных целей; 4) *Социальности* – способность взаимодействовать с другими агентами для достижения цели.

Для реализации всех свойств (1-4) интеллектуальных агентов исследовались различные абстрактные архитектуры интеллектуальных агентов. Одна из таких архитектур – BDI (*Belief-Desire-Intention*) архитектура, в рамках которой определены три составляющие агента: убеждения (*Belief*) - характеризуют знания агента, желания (*Desire*) - отражают цели агентов и намерения (*Intention*) - возможные действия агентов для достижения целей. Каждая составляющая у агента индивидуальна, может меняться по мере накопления агентом опыта работы в некоторой среде и характеризует его ментальное состояние.

Формальные модели агентов в терминах BDI архитектуры (по сути в терминах ментальных состояний агентов) рассмотрены в работах [1, 2]. В рамках BDI архитектуры предложены формальные модели реактивного, обучающегося, целенаправленного и др. абстрактных агентов, в которых показана связь ментальных состояний агента и его поведения.

Задача создания команды на множестве агентов с BDI архитектурой заключается в том, чтобы организовать между агентами обмен информацией об их ментальных состояниях, создающий мотивации для их совместной работы для достижения как индивидуальных, так и общей цели. Фактически речь идет о самоорганизации агентов в динамическом процессе, основанном на обменах агентами информацией об их ментальных состояниях, для поддержки структуры взаимодействия команды агентов и достижения общей цели без внешнего вмешательства.

Вопросы командного поведения агентов с BDI архитектурой исследовались в работах [3, 4]. Несмотря на разработки моделей представления знаний, языков, протоколов коммуникации BDI агентов [5] решение задач образования и функционирования команд с большим количеством агентов в терминах этой архитектуры оказывается нецелесообразным. Поэтому, в настоящее время есть интерес к подходам формирования команд в рамках теории группового управления.

Теоретические вопросы формирования команд роботов с системных позиций были рассмотрены в работе [6]. Здесь группа роботов, каждый из которых характеризуется вектором значений его свойств R , помещены в среду E , которая также характеризуется вектором состояния E . Пара векторов $S_c = \langle R, E \rangle$ характеризует состояние системы «роботы – среда».

Каждый робот может выполнять действия A , изменяющие состояние системы S_c . Действия роботов направлены на то, чтобы изменить текущее состояние системы S_c на желаемое состояние S_c^f . Ставится задача определения совместных действий роботов, позволяющих достичь поставленной цели S_c^f . Эта задача определяется как задача группового управления (управления группой роботов), решение которой – это векторы действий всех роботов $A_c(t)$, обеспечивающие экстремум некоторого функционала, характеризующего качество управления.

Среди методов группового управления интерес представляет стайное и коллективное децентрализованное управление группой роботов. Как первый, так и второй вид управления предполагает командное поведение роботов для достижения общей цели. Однако, при стайном управлении роботы не обмениваются между собой информацией, а строят свою стратегию достижения цели на основе анализа текущего состояния среды функционирования. В стайном управлении выделяется класс роевых подходов к управлению группой роботов. Роевые алгоритмы, основанные на локальном взаимодействии множества однородных роботов (роя роботов), обеспечивают их скоординированное движение к цели, обход препятствий и т.д. [7]. Стайные алгоритмы предполагают работу функционально разнородных роботов на основе их общих знаний о состоянии среды функционирования и индивидуальных правил поведения агентов в стае.

В живой природе существует множество примеров стайного поведения насекомых и животных [8]. Считается, что такое поведение является оптимальным. Поэтому, в качестве одной из возможностей повышения качества командной работы агентов является моделирование поведения стайных насекомых в алгоритмах командной работы искусственных агентов. В этой работе, на основе исследований социальных психологов предложены модели стайного поведения агентов в процессах образования и функционирования команд агентов.

1. Стайное поведение агентов (постановка задачи)

Общая постановка стайной работы команды роботов следующая. Рассматривается множество агентов (роботов) $A = \{a_i\}$, обладающих свойствами (параметрами) $F = \{f_i\}$. Для каждого свойства каждого из объектов (агентов) этой среды определено упорядоченное множество их возможных значений, $Z = \{Z_i\}$, где $Z_i = \{z_{i1}, \dots, z_{iq}\}$, $z_{iq+1} \succ z_{iq}$, $q = 0 \dots n-1$.

Среда функционирования агентов определяется как прямые произведение множеств значений всех свойств агентов, $SF = \times_i Z_i$.

Вектор значений всех агентов $Y(t) = (z_{1j}, \dots, z_{nb})$, $z_{ij} \in Z_i, \forall i$, определяет состояние среды функционирования.

Динамика изменения состояния среды функционирования происходит в случаях изменения агентами своих свойств и представляется как отображение:

$$W: Y(t) \rightarrow Y(t+1), \quad (2)$$

где W – система правил, заданных на множестве всех возможных состояний среды $W: \times_i Z_i \rightarrow \times_i Z_i$; $Y(t), Y(t+1)$ – состояния среды в последовательные моменты времени.

Отметим, что отображение W определяет закономерности среды функционирования, на основе которых агенты принимают решения о своих действиях по достижению цели.

Каждый агент характеризуется следующим кортежем:

$$\langle g_q, r_q, \mu_q(Y_q, g_q) \rangle,$$

где

1) $g_q = (z_{1j}, \dots, z_{nb})$ – вектор целевых значений для агента q ;

2) $r_q = (z_{1j}, \dots, z_{nb})$ – стратегия достижения цели агента q , где $r_q \in U_q$, $U_q = \times_i Z_i^r, Z_i^r \subseteq Z_i$ – ресурсы агента q .

Считается, что агент q применяет стратегию r_q для достижения своей цели g_q , предполагая, что другие агенты никаких действий не совершали. Прогноз изменения состояния среды на n шагах определится из соотношения (1) при условии, что $Y_q(0) = r_q$, т.е. $Y_q(1) = W^{\circ} r_q, Y_q(2) = W^{\circ} Y_q(1), \dots, Y_q(n) = W^{\circ} Y_q(n-1)$.

3) $\mu_q(Y_q(n), g_q)$ – возможность достижения агентом q целевого состояния за счет собственных ресурсов в условиях противодействия других агентов.

При определении возможности достижения целевого состояния считается, что в пространстве состояний $(\times_i Z_i)$ определена метрика $\rho(a, b), a, b \in \times_i Z_i$. Тогда возможность достижения цели агентом определяется как близость прогнозной $Y_q(n)$ и его целевой ситуации g_q :

$$\mu_q(Y_q(n), g_q) = \rho(Y_q(n), g_q).$$

По сути, этот показатель определяет потенциальную «силу» каждого агента команды без поддержки других членов команды при условии противодействия противников [9].

В такой общей постановке задачи необходимым условием командной работы агентов будем считать критерий близости целей агентов. В команду K в стайном алгоритме включаются агенты a_i , цели которых близки, т.е. $\forall a_i \in K, K \subseteq A, \rho(g_i, g_q) \leq \varepsilon, \forall i, q \in K, \varepsilon$ – критерий близости целей. При этом остальные агенты $A \setminus K$ считаются противниками.

1.2. Итерационный стайный алгоритм. В работе [6] приводится итерационный алгоритм согласования действий всех роботов (агентов), который предполагает наличие между ними канала обмена информацией. Согласно этому алгоритму агенты упорядочены, например, по номерам и выбирают свои действия последовательно в порядке возрастания номеров. Причем первый, выбрав действие, сообщает о нем второму, который выбирает свое действие, учитывая действие первого, и сообщает об этом третьему агенту и т.д. Такая последовательная передача информации о действиях от роботов с меньшими номерами всем последующим роботам позволяет согласовать их действия некоторым оптимальным образом.

В этом алгоритме, до начала решения некоторой задачи группой роботов на основе последовательного обмена информацией о своих действиях формируется общий план последовательных действий агентов достижения общей цели. Далее, считается, что реализация этого плана позволяет оптимальным образом решить поставленную задачу.

Такой алгоритм не применим в динамичных средах функционирования. Если есть агенты–противники, препятствующие достижению цели команды агентов, то команда будет вынуждена, при каждой смене состояния среды функционирования строить новый план оптимального взаимодействия агентов. Построение общего плана может потребовать много времени, что в динамических ситуациях неприемлемо.

Далее будут рассмотрены стайные алгоритмы, в которых применяется широкоэвентательный канал обмена информацией (доска объявлений), не предполагающий скоординированного диалога агентов.

2. Стайный алгоритм на основе критериев возможности достижения цели и взаимной полезности агентов

Этот алгоритм основан на анализе критериях возможности самостоятельного достижения цели каждым агентом и их взаимной полезности. Формально определим взаимную полезность агентов [9].

Агенты q и i называются взаимно полезными, если объединение их стратегий $r_i \oplus r_q$ увеличивает возможность достижения собственных целей (g_q, g_i) агентами q и i в условиях противодействия их противников. То есть если $\mu(Y_{i+q}, g_i) < \mu(Y_i, g_i)$, $\mu(Y_{i+q}, g_q) < \mu(Y_q, g_q)$. Степень полезности $P(i, q)$ агента q для агента i определяется из соотношения [9]:

$$P(i, q) = 1 - \frac{\mu(Y_{i+q}, g_i)}{\mu(Y_i, g_i)} \quad (4)$$

где $\mu(Y_{i+q}, g_i)$, – возможности достижения агентами q и i своих целей при объединении их стратегий; $\mu(Y_q, g_q)$, $\mu(Y_i, g_i)$ – возможности достижения агентами q и i своих целей собственными силами в условиях противодействия противников.

Степень полезности следующего агента, для команды из двух агентов q и i будем определять из соотношения:

$$P((i, q), j) = 1 - \frac{\mu(Y_{((i+q)+j)}, g_i)}{\mu(Y_{((i+q))}, g_i)} \quad (5)$$

Таким образом, для определения полезности вновь добавляемого агента к существующей команде, команда считается единым агентом.

Считается, что каждый агент или команда на очередном такте работы помещает на доску объявлений информацию: о действиях r_{iq} (стратегию достижения цели) и о возможности достижения цели самостоятельно - $\mu_q(Y_q(n), g_q)$. Эти сведения доступны всем агентам.

Эту информацию каждый агент анализирует автономно по одинаковым алгоритмам. Поэтому результаты вычислений у всех агентов будут одинаковыми, что исключает необходимость их согласования.

Алгоритм функционирования агентов следующий.

1. Каждый агент упорядочивает всех агентов (или команду агентов) в порядке возрастания их возможностей достижения цели.
2. Если в перечне агентов есть агент (команда), способные достичь цели самостоятельно, то агент (команда) и будут решать задачу достижения цели, т.е. $\exists i | \mu(Y_i, g_i) < \epsilon$. Остальные агенты в этом процессе не участвуют.
3. Если такого агента нет, то в качестве члена команды выбирается агент с большей возможностью достижения цели, для которого по формулам (4,5) подбирается агент с наибольшей взаимной полезностью. Эта команда будет работать на следующем такте работы стайного алгоритма. Остальные члены команды бездействуют.

Особенность стайного алгоритма заключается в том в том, что на каждом шаге работы алгоритма изменяются возможности достижения цели и, соответственно, взаимные полезности агентов.

4. Шаги алгоритма 1,2 и 3 повторяются до достижения цели, т.е. $\rho(Y^*(t+n), g) < \epsilon$, для новых значений возможности достижения цели агента и их взаимной полезности.

Таким образом, на каждом шаге этого алгоритма состав команды может меняться по количеству агентов или составу. Но важно, то, что на каждом шаге команда всегда будет содержать агента (команду) с наибольшей возможностью достижения цели и самого взаимно полезного

агента. Считается, что такой подбор команды позволит повысить качество командной работы агентов и приблизить их поведение к оптимальному.

Далее рассмотрим стайные алгоритмы, в которых алгоритмы командной работы стаи искусственных агентов (роботов) строятся на основе моделирования поведения людей в социуме.

2.1 Стайный алгоритм «ленивых» агентов. Этот алгоритм будем формулировать в терминах ранее введенной постановки задачи. Общее описание алгоритма следующее.

1. Каждый агент упорядочивает всех потенциальных членов команды по возможности достижения ими цели.
2. Если есть агент, способный достичь цели самостоятельно, то он достигает ее. Остальные агенты бездействуют.
3. Если такого агента нет, (здесь начинаются отличия от ранее описанного алгоритма), то команда принимает решение, чтобы лучший агент попытался самостоятельно достичь целевого состояния.
4. Остальные агенты, наблюдая за состоянием среды функционирования, ждут, когда этот агент исчерпает свои возможности достижения цели. Т.е. множество агентов находятся в состоянии «ленивого» ожидания результатов работы лучшего агента.
5. Если, наблюдаемый агент исчерпал свои возможности, то из множества «ленивых» агентов выбирается агент с большей возможностью достижения цели, который пытается ее достичь.
6. Множество «ленивых» агентов ждет, когда и этот агент исчерпает свои возможности и назначает нового исполнителя.
7. Процесс заканчивается, когда цель будет достигнута, т.е. если $\rho(Y^*(t+n), g) < \varepsilon$.

Этот алгоритм похож на итерационный алгоритм [Каляев и др., 2009], но его отличие заключается в том, что выбор очередного агента для включения его в команду происходит в процессе решения агентами задачи, а не до его начала.

2.2. Стайный алгоритм «эгоистичных» агентов. Этот алгоритм предполагает, что устойчивая команда многократно выполняет некоторую работу. Например, в социальных системах, в бригаде рабочих при многократном выполнении проектов и т.д. между членами коллектива могут сложиться отношения (симпатии, неприязни, лидерства и т.д.), которые могут сказаться на работе коллектива. Отношения между агентами, возникающие в коллективе связаны с их индивидуальными отличиями и благами, которые могут получать агенты с большими возможностями. Например, в ранее рассмотренных алгоритмах, основанных на возможности достижения цели и взаимной полезности,

основную командную работу будут выполнять только агенты с лучшими возможностями. Остальные будут «отдыхать». Если лучшие агенты за свою работу будут получать большие блага, то остальные («ленивые») агенты могут стремиться получать эти блага, перехватывая работу у лучших агентов, даже в том случае если качество работы всей команды ухудшится. Таких агентов будем называть эгоистичными.

В алгоритме «эгоистичных» агентов каждый агент кроме характеристики возможности достижения цели самостоятельно имеет еще и характеристику его работы O_i в предыдущих циклах работы, величина которой пропорциональна объему выполненной им работы и, соответственно, благам, которые агент за эту работу получил. Уровень эгоистичности агентов может быть определен из соотношения:

$$D_{ij} = \frac{|O_i - O_j|}{O_i + O_j}$$

В этом соотношении эгоистичность агентов $D_{ij}=0$, если они выполнили равную по объему работу, и соответственно, равна 1, если один из них никакой работы не выполнял – «отдыхал» («ленился»).

В алгоритме «эгоистичных» агентов считается, что сведения о выполненной работе каждого агента является общедоступными и, следовательно, агенты могут быть упорядочены по этой характеристике.

Рассмотрим алгоритм «эгоистичных» агентов.

1. Каждый агент упорядочивает всех потенциальных членов команды в два упорядоченных множества: по возможности достижения ими цели и по уровню эгоистичности агента относительно агента с максимальным объемом, выполненной работы.
2. Если есть агент, способный достичь цели самостоятельно, и его эгоистичность максимальна (или эгоистичности всех агентов равны), то он достигает ее. Остальные агенты бездействуют.
3. Если такого агента нет, то команда принимает решение, чтобы агент с максимальным уровнем эгоистичности попытался самостоятельно достичь целевого состояния.
4. Процесс заканчивается, когда цель будет достигнута, т.е. если $\rho(Y^*(t+n), g) < \varepsilon$.

Этот алгоритм позволяет равномерно распределить работу между всеми членами команды и обеспечить им одинаковые вознаграждения за работу в команде.

Заключение

В работе проанализирован классический стайный алгоритм роботов, выявлены его недостатки при формировании общего плана решения задачи в динамических ситуациях. Предложены стайные алгоритмы на основе критериев возможности самостоятельного достижения агентом цели и взаимной полезности агентов. На основе этих же критериев рассмотрены стайные алгоритмы «ленивых» и «эгоистичных» агентов.

Дальнейшие имитационные исследования формальных стайных алгоритмов позволят предложить их для реализации в командах роботов.

Библиографический список

1. Wooldridge M. J. The Logical Modeling of Computational Multi-Agent Systems. PhD thesis. _Manchester. _1992.. 153 p.
2. 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.
3. Cohen P., Levesque H.J. Teamwork. *Nous*, 25(4), (1991) Special Issue on Cognitive Science and Artificial Intelligence, P. 487-512.
4. Grosz B., Kraus S. Collaborative Plans for Complex Group Actions // *Artificial Intelligence*. – 1996. – №86. – P. 269–358.
5. Городецкий В.И. Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор. // *Искусственный интеллект и принятие решений*. – 2011. – №2, №3 – С. 19–30, С. 34–47.
6. Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов / М.: ФИЗМАТЛИТ, 2009.-279 с.
7. Павловский В.Е., Павловский В.В. Математическая модель двумерной гомогенной стаи роботов// *Искусственный интеллект и принятие решений*. М.: URSS. 2015. №4. - с. 62-71.
8. Карпов В. Э. Модели социального поведения в групповой робототехнике / *Управление большими системами*. Выпуск 59. М.: ИПУ РАН, 2016. С.165-232.
9. Кулинич А. А. Модель командного поведения агентов (роботов): когнитивный подход / *Управление большими системами*. Выпуск 51. М.: ИПУ РАН, 2014. С.174-196.

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.