

УДК 004.896
ББК 32.813

МОДЕЛЬ КОМАНДНОГО ПОВЕДЕНИЯ АГЕНТОВ (РОБОТОВ): КОГНИТИВНЫЙ ПОДХОД¹

Кулинич А. А.²

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Предложена математическая модель кооперации агентов в динамических ситуациях, основанная на критериях полезности агентов и когнитивного диссонанса их отношений. Модель основана на работах социальных психологов, исследовавших особенности поведения людей в малых социальных группах и объясняющих принципы их функционирования и устойчивости. На основе предложенной модели разработана имитационная модель игры агентов (роботов) в виртуальный футбол. Разработанная имитационная модель позволяет моделировать различные аспекты командного и эгоистичного поведения агентов.

Ключевые слова: кооперация агентов, критерии полезности, когнитивный диссонанс, когнитивная карта, правило поведения агента, командное поведение, эгоистичное поведение.

1. Введение

Исследование кооперации агентов (автономных роботов) возможно с использованием абстрактных математических моделей их взаимодействия в виде многоагентной системы. Многоагентная система состоит из множества интеллектуальных, общающихся между собой агентов (роботов), которые могут объединять свои ресурсы для решения сложных задач.

¹ Работа выполнена при финансовой поддержке РФФИ (проект №14-01-00817).

² Александр Алексеевич Кулинич, кандидат технических наук, старший научный сотрудник (kulinich@ipu.ru).

В рамках теории многоагентных систем был предложен ряд теоретических концепций, определяемых как теории командной работы агентов. Это теория общих намерений [17] и теория общих планов [20]. Эти две теории командной работы ориентированы на BDI (Belief-Desire-Intention) [22] – 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 (конъюнкция знаний агентов)). Понятия этого уровня ответственны за отражение действительности в терминах убеждений (достоверных знаний) 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 [3, 4], когда план приходится пересматривать в динамическом контексте ситуации.

В теории общих планов [20] 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 агенты часто 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 соперников – агентов, препятствующих командной работе группы агентов [3, 4].

Для многих случаев моделирования командной работы оказывается достаточным использование агентов с простой реактивной архитектурой (стимул-реакция). Агент с реактивной архитектурой реагирует некоторым действием на любое изменение среды, в которой он находится. Для реализации реактивной архитектуры необходимо перечислить возможные состояния среды и всевозможные реакции на эти состояния. Интерес к моделированию коллективного поведения простых реактивных агентов появился в 60-е годы прошлого столетия. Так, в работах [1, 12, 16] на основе изучения коллективного поведения конечных автоматов были заложены теоретические основы коллективного взаимодействия агентов. С помощью реактивных агентов реализуется, как правило, несложное командное поведение агентов – это образование стаи или роя агентов, которое реализуется с помощью роевых и эволюционных алгоритмов [5]. Исследуются разнообразные алгоритмы обхода стаей (роем) агентов препятствий, уклонение от столкновения [9] и т.д.

Агенты с реактивной архитектурой могут решать и более сложные задачи, если определены правила их поведения в среде и модели среды их функционирования. Например, если предположить, что агенты рациональны и ведут себя в соответствии с некоторой функцией полезности, пытаясь ее максимизировать, то задачи коммуникации агентов для командной работы могут быть значительно сокращены.

Рассмотрим, например, алгоритмы образования коалиций, основанные на методах решения кооперативных игр. Задача образования коалиции в рамках этой теории решается при условии

рациональности агентов и полной информированности всех агентов о целях, ресурсах и стратегиях других агентов, а также при условии их рационального поведения и возможности дележа выигрыша, полученного коалицией агентов. Коалиционную игру представляют форме характеристической функции [7]. Игра задается множеством агентов A и характеристической функцией $\gamma(\cdot)$ на всех возможных подмножествах агентов – коалициях. Выигрыш коалиции определяется как сумма индивидуальных выигрышей всех участников коалиции. Считается, что агенты коалиции могут делиться своими выигрышами, рационально распределяя его между другими участниками коалиции. Дележ $x = (x_1, x_2, \dots, x_n)$ выигрыша коалиции $\gamma(K)$ должен удовлетворять определенным условиям, одно из которых – это условие индивидуальной рациональности агента. Это условие означает, что выигрыш любого агента коалиции x_i должен превышать его выигрыш $\gamma(\{i\})$, если бы этот агент действовал самостоятельно, т.е. $x_i > \gamma(\{i\})$, $i \in K$. Понятие решения в коалиционной игре связано с устойчивостью дележа. Дележ выигрыша должен лишать участников коалиции мотиваций покидать коалицию, обеспечивая ее устойчивость. Для этого он должен доминировать дележи выигрыша других коалиций. Существуют методы решения кооперативных игр (С-ядро, МН-решения и др.) [7]. Однако при обеспечении полной информированности агентов эти методы имеют экспоненциальную сложность вычислений и коммуникаций между агентами относительно их числа, что ограничивает возможность их практического применения [11].

Еще одна группа методов – это методы, основанные на моделировании поведения агентов при образовании коалиций на основе теорий социального выбора и социальной зависимости субъектов [18, 19, 21]. Эти методы позволяют оперативно решать вопросы образования открытых динамических коалиций, состав которых может меняться в процессе решения задачи, а агенты могут свободно входить в состав коалиции или выходить из нее [18, 19, 21].

Здесь задано множество агентов $A = \{a_i\}$, для каждого из которых определены: $ACT(a_i)$ – множество возможных действий агента a_i ; $RES(a_i)$ – множество ресурсов агента a_i ; $PL(a_i)$ – план действий агента a_i для достижения цели o_j . План действий агента реализуется им для достижения цели путем манипулирования имеющимися ресурсами. План можно представить парой $PL(a_i) = (act_i^j, res_i^j)$, где $act_i^j \in ACT(a_i)$, $res_i^j \in RES(a_i)$.

Если у агента a_i для реализации плана достижения собственной цели нет ресурса $res_i^j \notin RES(a_i)$, но этот ресурс есть у агента a_q , $res_i^j \in RES(a_q)$, то считается, что агент a_i может обратиться к агенту a_q с просьбой передать недостающий ресурс. Если ресурс будет выделен, то агент a_i становится зависимым от агента a_q . Эта зависимость становится устойчивой, если агент a_i в свою очередь поделится каким-либо своим ресурсом с агентом a_q . Зависимости между агентами могут возникать по отношению ресурсов и действий. К недостаткам этого подхода относят отсутствие стимула у агентов для кооперации. Это связано с отсутствием общей цели или отсутствием знаний у агентов о зависимости между их целями [13].

Использование в модели кооперации агентов известных закономерностей предметной области позволяет снизить сложность процесса коммуникаций между ними [11,19], повысить их автономность и возможности командной работы в динамических ситуациях.

В этой работе предложена модель кооперации агентов, основанная на моделях социального поведения, изложенных в работах социальных психологов Д. Хоманса [15] и Л. Фестингера [14]. В этих работах предложены: модель функционирования малых социальных групп (коалицию можно рассматривать как малую социальную группу) на основе обмена полезностями и концепция когнитивного диссонанса, позволяющая оценить субъективные ощущения агентов о справедливости в коалиции и моделировать ее устойчивость.

2. Базовая модель

Рассмотрим динамическую систему и множество агентов A , которые пытаются изменить состояние системы на выгодное для каждого из них состояние. В качестве математической модели системы выбрана модель представления экспертных знаний о процессах в динамической системе в виде качественной когнитивной карты [6]. В когнитивной карте определено множество факторов ситуации $F = \{f_i\}$, упорядоченное множество лингвистических значений факторов $L_i = \{l_{ij}\}$, $l_{ij} \prec l_{ij+1}$, $i = 1, \dots, n, j = 1, \dots, m$, и причинно-следственные отношения между факторами – W .

Динамика изменения факторов ситуации в когнитивной карте задается системой конечно-разностных логико-лингвистических уравнений

$$(1) \quad Y(t+1) = W^{\circ} Y(t),$$

где W – система правил «Если, то», заданная на множестве всех возможных значений факторов ситуации $W: \times_i L_i \rightarrow \times_i L_i$; $Y(t)$ – состояние ситуации – это вектор лингвистических значений (l_{1j}, \dots, l_{nb}) , т.е. $Y(t) = (l_{1j}, \dots, l_{nb})$, $l_{ij} \in L_i, \forall i$, – правило вывода.

Логико-лингвистические модели [10] позволяют описывать динамические ситуации, параметры которых могут быть измерены экспертным способом и представлены в качественных порядковых шкалах. Динамика моделируется с помощью системы продукционных правил, отражающих закономерности предметной области. Применение такого рода моделей значительно расширяют класс моделируемых ситуаций.

Каждый агент характеризуется следующим кортежем:

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

где

1) $g_q = (l_{1j}, \dots, l_{nb})$ – вектор целевых факторов и их лингвистических значений для агента q ;

2) $r_q = (l_{1j}, \dots, l_{nb})$ – стратегия управления агента q , где $r_q \in U_q$, $U_q = \times_i L_{qi}^r$, $L_{qi}^r \subseteq L_i$ – ресурсы агента q . Считается, что если агент q применяет стратегию r_q для достижения своей цели

g_q , то прогноз развития ситуации на n шагах моделирования определится из соотношения (1) при условии, что $Y_q(0) = r_q$, т.е.

$$3) Y_q(1) = W^{\circ} r_q, Y_q(2) = W^{\circ} Y_q(1), \dots, Y_q(n) = W^{\circ} Y_q(n-1).$$

4) $e_q: (u(g_q), (v(r_q))) \rightarrow R^+$ – эффективность достижения агентом целевой ситуации, где $u(g_q)$ – полезность целевой ситуации, $u: g_q \rightarrow R^+$; $v(r_q)$ – затраты агента для достижения целевой ситуации, $v: r_q \rightarrow R^+$;

5) $\mu_q(Y_q(n), g_q)$ – возможность достижения агентом q целевой ситуации за счет собственных ресурсов в условиях противодействия агентов противников.

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

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

При этом прогноз развития ситуации $Y_q(n)$ для каждого агента строится с учетом противодействия противников [6]. По сути, этот показатель определяет потенциальную «силу» каждого агента коалиции без поддержки потенциальных сторонников при условии противодействия противников.

2.1. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ ОБРАЗОВАНИЯ КОАЛИЦИИ

Очевидно, что в качестве потенциальных сторонников в конфликтной ситуации необходимо выбирать агентов, цели которых близки. Сторонниками считается множество агентов с близкими целями, т.е. $K \subseteq A$, $\rho(g_i, g_q) \leq \varepsilon$, $\forall i, q \in K$, ε – критерий близости целей, определяется экспертом. При этом остальные агенты AK считаются противниками коалиции. Близость целей агентов – это необходимые условия образования коалиции, не позволяющие судить об ее устойчивости. В настоящей работе на основе теории социального поведения субъектов [15] анализируется взаимная полезность агентов при объединении их ресурсов для достижения общей цели. Считается, что дисбаланс во взаимной полезности агентов в коалиции создает у них ощу-

щения несправедливости, которые увеличивают их мотивации выхода из коалиции, угрожающих ее устойчивости. Для оценки ощущений агентов о несправедливости в коалиции используется теория когнитивного диссонанса [14].

2.2. КРИТЕРИЙ ВЗАИМНОЙ ПОЛЕЗНОСТИ АГЕНТОВ В КОАЛИЦИИ

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

$$(2) \quad P(i, q) = 1 - \frac{\mu(Y_{i+q}, g_i)}{\mu(Y_i, g_i)},$$

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.

Полезность игрока q для игрока i максимальна, т.е. $P(i, q) = 1$, когда при объединении их стратегий игрок i достигнет своей цели g_i . В этом случае $\mu(Y_{i+q}, g_i) = \rho(Y_{i+q}, g_i) = 0$. Если при объединении стратегий расстояние до цели не изменилось, $\mu(Y_{i+q}, g_i) = \mu(Y_i, g_i)$, то полезность игрока нулевая, $P(i, q) = 0$. И, наконец, если при объединении стратегий расстояние до цели увеличилось, $\mu(Y_{i+q}, g_i) > \mu(Y_i, g_i)$, то игрок q вреден для игрока i .

Социальное поведение в социальной группе – это обмен ценностями между агентами (ресурсами, полезностями и т.д.). Целостность группы достигается при балансе «сил» между обменами при условии, что каждый агент получает при этом максимальную выгоду [15].

По критерию взаимной полезности коалиция K устойчива, если агенты коалиции одинаково полезны друг для друга

$P(i, q) = P(q, i)$, $\forall i, q \in K$, и имеют равную эффективность участия в коалиции, т.е. $e_i = e_q$.

2.3. КРИТЕРИЙ КОГНИТИВНОГО ДИССОНАНСА

Наличие дисбаланса взаимной полезности или эффективности агентов, включенных в коалицию, приводит к возникновению скрытых (латентных) конфликтов в коалиции, которые характеризуют ощущения каждого агента о несправедливости в коалиции, и определяются уровнем его когнитивного диссонанса. Когнитивный диссонанс – это противоречие в системе знаний человека, порождающее у него неприятные переживания и побуждающее его к действиям, направленным на устранение этого противоречия [14]. Когнитивный диссонанс вызывает у субъекта (агента) стремление уменьшить его или препятствовать его росту. Проявление этого стремления у агентов приводит к изменению его поведения [14], которое приводит к их выходу из коалиции, т.е. к неустойчивости коалиции.

Степень когнитивного диссонанса по взаимной полезности между агентами i, q определяется из соотношения

$$(3) \quad D^W(q, i) = \frac{P(i, q) - P(q, i)}{P(i, q) + P(q, i)}, \quad D^W(q, i) = D^W(i, q), \quad \forall i, q, i \neq q.$$

Степень когнитивного диссонанса по эффективности между агентами i, q определится из соотношения

$$(4) \quad D^W(i, q) = \frac{e_i - e_q}{e_i + e_q}, \quad D^E(i, q) = -D^E(q, i), \quad \forall i, q, i \neq q.$$

По критерию когнитивного диссонанса устойчивая коалиция определится следующим образом: коалиция K устойчива, если для всех агентов $i, q \in K$ их когнитивные диссонансы по полезности и эффективности близки к нулю: $D^W(i, q) \approx 0$, $D^E(i, q) \approx 0, \forall i, q$.

3. Пример

В качестве примера рассмотрим описание в терминах предложенной модели кооперативного взаимодействия агентов, иг-

рающих в футбол. Моделирование игры в футбол считается «полигоном» для демонстрации возможностей разнообразных методов искусственного интеллекта, определяющих поведение искусственных агентов – роботов. С 1997 года проходят международные соревнования роботов (RoboCup), играющих в футбол. На деле это соревнование научных и инженерных школ, разрабатывающих роботов, методы и алгоритмы их поведения в сложной динамической игре. Возникающая при этом конкуренция между разработчиками роботов не способствует опубликованию методов и алгоритмов, реализующих поведение роботов-футболистов. В работе [8] изложены основные методы и подходы к моделированию игры в виртуальный футбол.

В настоящее время популярность набирает виртуальный футбол с футбольными командами, состоящих из виртуальных игроков (программных агентов) [2]. В этой игре характеристики игрока и характеристики его командного поведения индивидуально настраиваются пользователем (футбольным менеджером) для создания виртуальной футбольной команды. Созданные таким образом многочисленные виртуальные футбольные команды участвуют в соревнованиях (чемпионатах) в интернете для определения лучшей команды и, следовательно, лучших настроек виртуальной футбольной команды. Несмотря на зрелищность 3D-визуализаций трансляций футбольных матчей виртуальных команд, в этом случае модель виртуального футболиста неизвестна и возможностей для экспериментов по моделированию его поведения на основе предложенных выше идей практически нет.

Поэтому далее опишем игру в футбол в терминах рассмотренной выше базовой модели кооперативного взаимодействия агентов (игроков) на основе взаимной полезности и когнитивного диссонанса.

Игру в футбол будем считать динамической системой. Параметры этой системы суть положение игроков (агентов) двух противоборствующих команд $A = \{a_i\}$ и $B = \{b_i\}$ на футбольном поле. Состояние футбольной игры как динамической системы будем характеризовать вектором положения всех

игроков – агентов и мяча на плоскости в последовательные моменты времени:

$$(5) \quad Y(t) = (x_{ai}(t); y_{ai}(t)), \dots, (x_{bi}(t); y_{bi}(t), \text{Ball}), i = 1, \dots, N,$$

где N – число агентов обеих команд; $t = 1, \dots, n$; $\text{Ball} = (x_m(t); y_m(t))$ – координаты мяча.

Координаты ворот противников заданы интервалами: ворота команды A – $G_A = (x_a; [y_{a1}, y_{a2}])$ и, соответственно, ворота команды B – $G_B = (x_b; [y_{b1}, y_{b2}])$.

Цели игроков команды A – забить мяч в ворота команды B , т.е. $\forall a_i, g_{ai} = (x_m(t); y_m(t)) \mid (x_m(t); y_m(t)) \in (x_b; [y_{b1}, y_{b2}])$, и, соответственно, игроков команды B – забить мяч в ворота команды A , т.е. $\forall b_i, g_{bi} = (x_m(t); y_m(t)) \mid (x_m(t); y_m(t)) \in (x_a; [y_{a1}, y_{a2}])$. То есть игроки каждой команды изменяют координаты мяча до тех пор, пока они не совпадут с координатами ворот противника.

Стратегии игроков (агентов) команд $A(B)$ – это их действия на каждом шаге дискретного времени. Далее будем описывать поведения игроков команды A , предполагая что поведение игроков команды B описывается аналогично. Итак, действия игроков команд следующие:

- $r_1(a_i, t)$ – движение к мячу;
- $r_2(a_i, t)$ – движение с мячом к воротам;
- $r_3(a_i, t)$ – удар по воротам;
- $r_4(a_i, t)$ – перемещение в позицию паса;
- $r_5(a_i, a_j, t)$ – пас партнеру по команде.

Любое из названных действий игроков изменяет состояние ситуации $Y(t)$.

Эффективность игрока – это соотношение его ожиданий вознаграждения за игру $u(a_i)$ к результативности его игры $v(r_q)$. Результативность игрока выражается интегральной оценкой его результативных действий. Под результативными действиями понимаются следующие действия: $r_3(a_i)$ – удар по воротам и $r_5(a_i, a_j, t)$ – пас партнеру по команде.

Возможность достижения цели игроками команд $A(B)$ определяется ресурсами агента. В данном случае под ресурсами агента понимаются его индивидуальные характеристики и особенности его позиция на поле.

Будем рассматривать следующие характеристики позиции игрока:

- $d(a_i, G_B, t) = \rho((x_{ai}(t); y_{ai}(t)), (x_{a(b)}; [y_{a1}, y_{a2}]))$ – близость агента к воротам противника;
- $d(a_i, b_j, t) = \rho((x_{ai}(t); y_{ai}(t)), (x_{bj}(t); y_{bj}(t)))$ – близость противников, которые могут помещать игроку.

Индивидуальные характеристики игрока – это его индивидуальная скорость $V(a_i)$ и его действия $r_1(a_i, t), \dots, r_5(a_i, a_j, t)$.

3.1. КООПЕРАЦИЯ АГЕНТОВ НА ОСНОВЕ КРИТЕРИЯ ИХ ПОЛЕЗНОСТИ

В работе [6] было показано, что полезными агентами для кооперации являются агенты, имеющие большую возможность достижения цели. Определим правило вычисления полезности агента, выразив ее через составляющие их возможности достижения цели, и определим их действия.

Агент a_i полезен своим партнерам по критерию близости к воротам противника (K_G^B), если $d(a_i, G_B) < d(a_q, G_B)$, $i \neq q$, $q = 1, \dots, N/2$. То есть агент полезен, если он находится ближе всех своих партнеров к воротам противника. В этом случае значение этого критерия равно единице, т.е. $K_G^B = 1$.

Агент a_i полезен своим партнерам по критерию близости противников (K^B), если $d(a_i, b_j) > d(a_q, b_j)$, $i \neq q$, $q = 1, \dots, N/2$, $j = N/2 + 1, \dots, N$. То есть агент полезен, если он находится дальше всех своих партнеров от игроков противника. В этом случае значение этого критерия равно единице, т.е. $K^B = 1$.

Агент a_i полезен своим партнерам по критерию скорости (K^V), если $V(a_i) > V(a_q)$, $i \neq q$, $q = 1, \dots, N/2$. То есть агент полезен, если его скорость больше скоростей остальных партнеров. В этом случае значение этого критерия равно единице, т.е. $K^V = 1$.

Интегральная характеристика полезности агента определяется значением функции: $P(a_i) = F(K_G^B, K^B, K^V)$.

В качестве одного из вариантов такой функции рассмотрим линейную свертку с коэффициентами предпочтений полезностей положения игрока – α, β, γ , полученных с помощью процедуры парных сравнений. В этом случае,

$$P(a_i) = \alpha K_G^B + \beta K^B + \gamma K^V, \alpha + \beta + \gamma = 1.$$

Тогда $P(a_i) = 1$ при $K_G^B = K^B = K^V = 1$, и $P(a_i) = 0$ при $K_G^B = K^B = K^V = 0$.

3.2. ВЫБОР СТРАТЕГИИ АГЕНТА НА ОСНОВЕ АНАЛИЗА ПОЛЕЗНОСТИ АГЕНТОВ

Построить план взаимодействия игроков одной команды для достижения цели в динамической ситуации и в случае противодействия игроков (агентов) противника оказывается затруднительным. Поэтому в данном случае для игроков (агентов) формулируются правила поведения агентов, учитывающие закономерности игры в футбол.

Правила поведения игроков предписывают им действия, которые способствуют командной работе для достижения цели. В этой работе рассмотрены два типа правил:

- командной игры;
- эгоистичного поведения.

Правило командной игры будем рассматривать для двух случаев: агент не владеет мячом; агент владеет мячом. Для агента, не владеющего мячом, определено следующее множество правил $W_b = \{w_{b1}, w_{b2}\}$:

- правило w_{b1} : если агент находится ближе к мячу, чем другие агенты команды, то он выбирает действие $r_1(a_i, t)$ (движение к мячу);
- правило w_{b2} : если есть другой агент, который находится ближе к мячу, то агент выбирает действие $r_4(a_i, t)$ (перемещение в позицию паса).

Для агента, владеющего мячом, выбор действия основан на анализе полезности партнеров по команде. В этом случае также определено множество правил $W_{nb} = \{w_{nb3}, w_{nb4}, w_{nb5}\}$:

- если полезность игрока, владеющего мячом, выше, чем полезность остальных партнеров, т.е. $P(a_i) > P(a_q)$, $i \neq q$, $q = 1, \dots, N/2$, то агент выбирает действие в зависимости от расстояния до ворот противника.

- правило w_{nb3} : при большом расстоянии выбирается действие $r_2(a_i, t)$ (движение с мячом к воротам);

- правило w_{nb4} : при близком расстоянии выбирается действие $r_3(a_i, t)$ (удар по воротам).
- правило w_{nb5} : если полезность игрока, владеющего мячом, ниже, чем полезность одного из партнеров, то выбирается действие $r_5(a_i, a_j, t)$ – пас партнеру по команде (a_j), имеющему большую полезность.

Правила эгоистичного поведения агентов будут рассмотрены в следующем разделе.

3.3. УСТОЙЧИВОСТЬ КООПЕРАЦИИ АГЕНТОВ

Ранее при принятии агентом решения учитывалась только полезность агентов для достижения цели на одном такте. Для анализа устойчивости кооперации агентов необходим анализ взаимной полезности агентов и анализ когнитивного диссонанса их эффективности на протяжении игры.

В базовой модели взаимная полезность игрока определялась (2) как соотношение возможностей достижения цели при объединении стратегий двух игроков и собственными силами. Считается, что объединение стратегий должно увеличивать (по крайней мере не уменьшать) возможность достижения цели каждым игроком. В рассматриваемом примере будем считать, что передача паса партнеру, основанная на анализе его полезности, увеличивает возможность достижения этим партнером цели, т.е. дает ему возможность забить гол.

Поэтому взаимная полезность агентов в игре в футбол определяется суммарным числом пасов между парой агентов:

$$P(a_i, a_j) = \frac{\sum_{t=1, \dots, n} r_5(a_i, a_j, t)}{\sum_{t=1, \dots, n} r_5(a_i, a_j, t) + \sum_{t=1, \dots, n} r_5(a_j, a_i, t)}, \quad P(a_j, a_i) = 1 - P(a_i, a_j).$$

Здесь $\sum_{t=1, \dots, n} r_5(a_i, a_j, t)$ и $\sum_{t=1, \dots, n} r_5(a_j, a_i, t)$ – число пасов агента a_i агенту a_j и, соответственно, a_j агенту a_i за промежуток времени игры $t = 1, \dots, n$.

Агенты с неравной взаимной полезностью характеризуются когнитивным диссонансом, вычисляемым по формуле (3), и могут изменять свое поведение. Изменение поведение проявляется

в изменении правил выбора агента на основе полезности агента партнера.

При анализе когнитивного диссонанса по эффективности агентов считается, что агенты получают вознаграждение пропорционально их результативности (забитые голы). Эффективность агента определится из соотношения

$$E(a_i) = \frac{\sum_{t=1, \dots, n} r_3(a_i, t)}{\sum_{\substack{t=1, \dots, n \\ i=1, \dots, N/2}} r_3(a_i, t)},$$

где в числителе число голов, забитых игроком a_i за период времени $t = 1, \dots, n$; в знаменателе общее число голов, забитых игроками команды.

Агенты с неравной эффективностью характеризуются когнитивным диссонансом по эффективности, вычисляемой по формуле (4), и могут изменять свое поведение, изменения правила выбора агента на основе полезности агента партнера.

Правило эгоистического поведения может применяться при наличии когнитивного диссонанса при выполнении агентами командной работы.

Множество правил эгоистического поведения агента $W_e = \{w_{e1}, w_{e2}\}$ следующее:

- правило w_{e1} : если игрок, владеющий мячом, выбирает действие $r_5(a_i, a_j, t)$ – пас партнеру по команде (a_j), имеющему большую полезность, и если когнитивный диссонанс этого игрока (a_i) по отношению к партнеру (a_j) близок нулю, $D^E(a_i, a_j) \approx 0$, то пас передается, т.е. выполняется действие $r_5(a_i, a_j, t)$;

- правило w_{e2} : если игрок, владеющий мячом, выбирает действие $r_5(a_i, a_j, t)$ – пас партнеру по команде (a_j), имеющему большую полезность, и если когнитивный диссонанс этого игрока (a_i) по отношению к партнеру (a_j) больше нуля, $D^E(a_i, a_j) > \varepsilon$, то пас не передается, выполняются действия: $r_2(a_i, t)$ – движение с мячом к воротам и (или) $r_3(a_i, t)$ – удар по воротам. Здесь ε определяется экспертом как область нечувствительности к небольшим значениям когнитивного диссонанса.

Иными словами, при эгоистическом поведении игрока он не передает пас наиболее полезному игроку, если этот игрок имеет большую эффективность, т.е. имеет лучшую результативность и получает большее вознаграждение. Эгоистический игрок действует самостоятельно, пытаясь повысить свою результативность (забить гол) и, соответственно, получить вознаграждение.

3.4. МОДЕЛИРОВАНИЕ ДИНАМИКИ

В базовой модели состояние управляемой множеством агентов динамической системы изменяется под действием стратегий всех агентов. При этом считается, что все агенты применяют свои стратегии одновременно на первом дискретном шаге моделирования, прогнозируя развитие ситуации на n шагов. Изменение состояния системы описывается системой правил W , представленных в уравнении (1).

При описании игры в футбол агенты изменяют состояние системы на каждом дискретном шаге моделирования, используя для этого продукционные правила из множеств правил W_b , W_{nb} , W_e , описанных выше. Каждое продукционное правило из этих множеств состоит из двух частей:

- часть «Если...» включает условия срабатывания правила. Условия срабатывания правила зависят от вектора состояния футбольной игры $Y(t)$ (5), а также от вычисляемых по этому вектору индивидуальных характеристик полезности, когнитивного диссонанса каждого игрока;

- часть «То...», включающая действие, которое будет выполнено при выполнении условий из части «Если». Под действиями здесь понимаются действия $r_1((a_i), t), \dots, r_5(a_i, a_j, t)$, которые изменяют вектор состояния футбольной игры $Y(t)$ (5) на новое состояние $Y(t+1)$.

Множество всех правил $W = W_b \cup W_{nb} \cup W_e \cup W_{pl}$ назовем базой правил, где

- W_b – множество правил поведения игрока, не владеющего мячом;
- W_{nb} – множество правил поведения игроков, владеющих мячом;

- W_e – множество правил эгоистичного поведения;
- W_{pl} – множество правил вычисления индивидуальных характеристик игроков по вектору состояния игры $Y(t)$.

Тогда уравнение динамики игры в футбол будет описано уравнением (1) базовой модели.

3.5. ЭКСПЕРИМЕНТ И ОБСУЖДЕНИЕ

Предложенная модель кооперации агентов была исследована экспериментально. Для этого в среде моделирования многоагентных систем VisualBots for Excel [23] была разработана программа, моделирующая игру в футбол двух команд, состоящих из четырех игроков. Агенты в каждой команде незначительно отличались по скорости. Так, агенты 1 и 3 команды A и агенты 1 и 4 команды B имеют скорость равную 1, агенты 2 обеих команд имеют скорость 0,9, а агент 4 и агент 3 команд соответственно A и B имеют скорость 0,7.

Считается, что все агенты получают вознаграждение пропорционально количеству забитых голов, независимо от выигрыша или проигрыша команды.

Перед экспериментальным исследованием ставились две задачи. Первая задача заключается в проверке правдоподобности теоретической модели кооперативного взаимодействия на основе их полезности и когнитивного диссонанса их взаимодействия. Вторая задача заключалась в проверке механизмов взаимодействия агентов при решении кооперативной задачи.

В качестве механизма взаимодействия агентов была рассмотрена модель с общей памятью. Суть этой модели заключается в том, что все агенты сохраняют координаты своего положения на поле в общей памяти и не общаются между собой непосредственно.

Каждый агент может считать информацию об агентах своей команды или агентах из команды противника из общей памяти и на основе ее анализа (полезности игрока (агента)) принимать действия, описанные выше в теоретической модели.

На рис. 1 показано окно системы моделирования, в котором игроки разных команд обозначены как фигуры насекомых:

«Пчелки» и «Божьи коровки». Ломаная линия в этом окне показывает траекторию мяча (мяч обозначен точкой) в процессе его распасовки игроками команды.

На рис. 2–5 показаны результаты экспериментов. Здесь по оси абсцисс показаны номера игроков команд, а по оси ординат – количество пасов партнерам (рис. 2 и рис. 4) и забитых голов (рис. 3 и рис. 5) в двух сериях игр: без учета и с учетом диссонанса игроков.

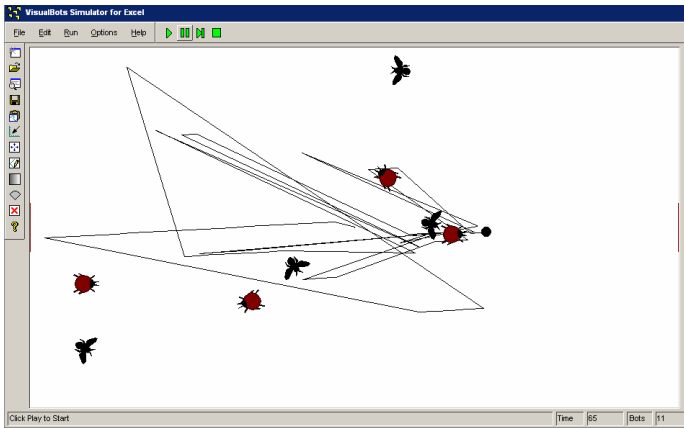


Рис. 1. Окно системы моделирования

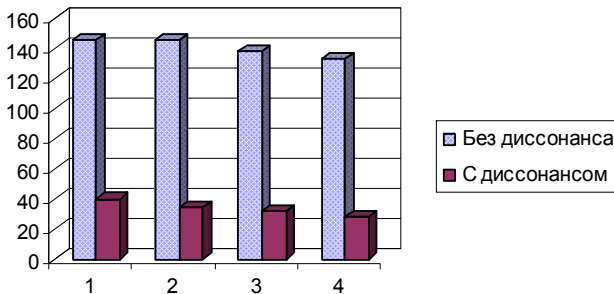


Рис. 2. Пасы игроков команды А

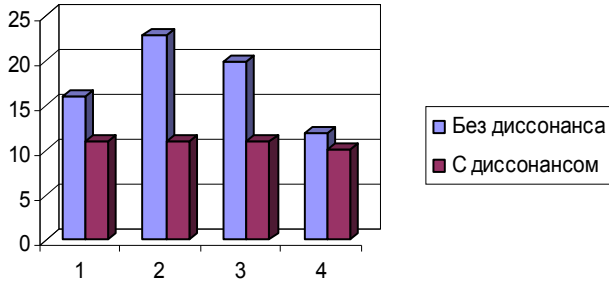


Рис. 3. Голы игроков команды А

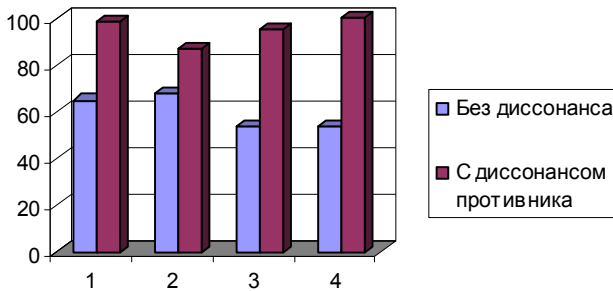


Рис. 4. Пасы игроков команды В

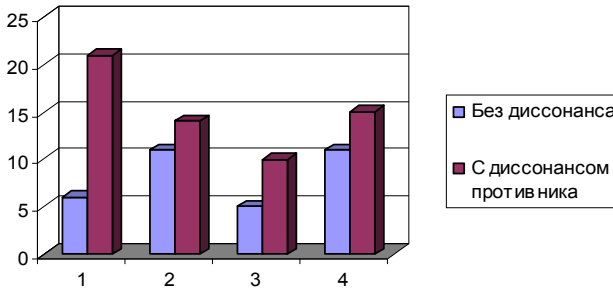


Рис. 5. Голы игроков команды В

Имитационное моделирование поведения агентов на основе анализа полезности агентов (без диссонансов) показало, что агенты с меньшими скоростями характеризуются меньшей результативностью по забитым голам, но не по количеству пасов партнерам (игрок №4 команды А, игрок №3 ко-

манды *B*). Это может быть интерпретировано как самоорганизация агентов в команде. То есть агенты с меньшей скоростью выполняют задачи защитников, а агенты с большей скоростью играют в нападении.

При моделировании поведения агентов с учетом диссонансов диссонансы учитывались только для игроков команды *A*. То есть все игроки команды *A* при передаче паса партнеру учитывают не только его полезность, но и диссонанс с партнером по эффективности. Если эффективность партнера выше, чем у игрока, передающего пас, то пас не передается. Этот игрок, из-за зависти или эгоистических соображений, пытается самостоятельно забить гол и получить вознаграждение.

На рис. 2 и рис. 3 показано, что результативность распасовки игроков команды *A* при учете диссонансов агентов резко снизилась. Снизился и диссонанс по эффективности (забитым голам) – все игроки забивают одинаковое количество голов и, следовательно, получают равное вознаграждение. Игроки же команды *B*, играющие с «эгоистичной» командой *A*, повысили свою результативность (рис. 4, рис. 5).

4. Заключение

Предложена математическая модель командной работы агентов на основе критериев полезности агентов и их когнитивных диссонансов. Модель основана на работах социальных психологов, исследовавших особенности поведения людей в малых социальных группах и объясняющих принципы их функционирования, устойчивости, сплоченности и т.д. На основе предложенной модели разработана имитационная модель игры агентов (роботов) в виртуальный футбол. Разработанная имитационная модель позволяет моделировать различные аспекты командного поведения агентов.

Имитационное моделирование командного поведения агентов на основе предложенных математических моделей выявило, что индивидуальные отличия агентов по скорости перемещения создают условия самоорганизации в команде. Так, в командах

появляются агенты, выполняющие разные задачи: «нападающие» и «защитники» – это агенты с большей и меньшей скоростями соответственно.

Моделирование эгоистичного поведения агентов одной из команд выявило факт уменьшения когнитивного диссонанса этих агентов (все агенты забывают равное число голов) и увеличения индивидуальной игры, что выражается в сокращении числа пасов в эгоистичной команде.

Предложенная модель может служить алгоритмической основой для реализации поведения автономных роботов в задачах кооперации в сложных динамических ситуациях и в условиях противодействия противников.

Литература

1. ВАРШАВСКИЙ В.И., ПОСПЕЛОВ Д.А. *Оркестр играет без дирижера*. – М: Наука 1984. – 208 с.
2. *Виртуальная футбольная лига*. – [Электронный ресурс]. – URL: <http://www.virtualsoccer.ru/> (дата обращения: 16.07.2014).
3. ГОРОДЕЦКИЙ В.И. *Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор (Часть 1)* // Искусственный интеллект и принятие решений. – 2011. – №2. – С. 19–30.
4. ГОРОДЕЦКИЙ В.И. *Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор (Часть 2)* // Искусственный интеллект и принятие решений. – 2011. – №3. – С. 34–47.
5. КАРПОВ В.Э. *Коллективное поведение роботов. Желаемое и действительное* // Современная мехатроника. Сб. научн. трудов Всероссийской научной школы (г. Орехово-Зуево, 22–23 сентября 2011). – Орехово-Зуево, 2011. – С. 35–51.
6. КУЛИНИЧ А.А. *Модель поддержки принятия решений для образования коалиций в условиях неопределенности* // Искусственный интеллект и принятие решений. – 2012. – №2. – С. 95–106.
7. ОУЭН Г. *Теория игр*. – М.: Мир, 1971. – 230 с.

8. ОХОЦИМСКИЙ Д.Е., ПАВЛОВСКИЙ В.Е., ПЛАХОВ А.Г., ТУГАНОВ А.Н., ПАВЛОВСКИЙ В.В. *Моделирование игры роботов-футболистов в пакете «Виртуальный футбол»* // Мехатроника. – 2002. – №1. – С. 2–5.
9. ПАВЛОВСКИЙ В.Е., КИРИКОВА Е.П. *Моделирование управляемого адаптивного поведения гомогенной группы роботов* // Искусственный интеллект. – 2002. – №4. – С. 596–605.
10. ПОСПЕЛОВ Д.А. *Логико-лингвистические модели в системах управления*. – М.: «Энергоиздат», 1981. – 232 с.
11. СМИРНОВ А.В., ШЕРЕМЕТОВ Л.Б. *Модели формирования коалиций кооперативных агентов: состояние и перспективы исследований* // Искусственный интеллект и принятие решений. – 2011. – №1. – С. 36–48.
12. СТЕФАНЮК В.Л., ЦЕТЛИН М.Л. *О регулировке мощности в коллективе радиостанций* // Проблемы передачи информации. – 1967. – Т. 3, №4. – С. 59–67.
13. ТАРАСОВ В.Б. *От многоагентных систем к интеллектуальным организациям*. – М.: Эдиториал УРСС, 2002. – 352 с.
14. ФЕСТИНГЕР Л. *Теория когнитивного диссонанса*. – СПб.: Ювента, 1999. – С. 15–52.
15. ХОМАНС ДЖ. *Социальное поведение как обмен. Современная зарубежная социальная психология*. – М.: Изд-во Московского университета, 1984. – С. 82–91.
16. ЦЕТЛИН М.Л. *Исследования по теории автоматов и моделированию биологических систем*. – М.: Наука, 1969. – 316 с.
17. COHEN P., LEVESQUE H.J. *Teamwork* // Nous. Special Issue on Cognitive Science and Artificial Intelligence. – 1991. – No. 25(4). – P. 487–512.
18. CONTE R., EDMONDS B., MOSS S. and etc. *Sociology and Social Theory in Agent Based Social Simulation* // A Symposium: Computational and Mathematical Organization Theory. – 2001. – Vol. 7, №3. – P. 183–205.
19. GENESERETH M., GINSBERG M. AND ROSENCHIN J. *Cooperation without Communication* // Proc. Annual Conf. Assoc. Artificial Intelligence, Philadelphia, 1986. – P. 51–57.

20. GROSZ B., KRAUS S. *Collaborative Plans for Complex Group Actions* // Artificial Intelligence. – 1996. – №86. – P. 269–358.
21. SICHMAN J., CONTE R., CASTELFRANCHI C. et al. *A social reasoning mechanism based on dependence networks* // Proc. 11th European Conference on Artificial Intelligence (ECAI), Amsterdam, Netherlands, 1994 – P. 188–192.
22. 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.
23. *Visual Bots. Visual Programming for Agent-Based Simulation.* – [Электронный ресурс]. – URL: <http://www.visualbots.com/downloads.htm> (дата обращения: 16.07.2014).

MODEL OF AGENTS (ROBOTS) COMMAND BEHAVIOUR: COGNITIVE APPROACH

Alexander Kulinich, Institute of Control Sciences of RAS, Moscow, Cand.Sc. (kulinich@ipu.ru)

We suggest a mathematical model of agents' cooperation in dynamics, which employs agents' utility functions and cognitive dissonance of their relations. The model is based on the theories of social psychologists investigating behavior of people in small social groups and explaining principles of their functioning and stability. We illustrate the suggested model with a simulation of a virtual soccer game of agents (robots). The developed simulation allows modeling diverse aspects of agents' cooperation and selfish behavior.

Keywords: cooperation of agents, criteria of utility, cognitive dissonance, a rule of the agent behavior, command behavior, selfish behavior.

Статья представлена к публикации членом редакционной коллегии П.Ю. Чеботарёвым

*Поступила в редакцию 17.07.2014.
Опубликована 30.09.2014.*