

МОДЕЛЬ КООПЕРАЦИИ АГЕНТОВ (РОБОТОВ)¹

А.А. Кулинич (kulinich@ipu.ru)
Институт проблем управления РАН, Москва

Исследована модель кооперации агентов (роботов) в динамических ситуациях, основанная на критериях полезности агента и когнитивного диссонанса их отношений.

Введение

Исследование вопросов кооперации агентов (роботов) возможно с использованием абстрактных математических моделей абстрактных многоагентных систем. Многоагентная система состоит из множества интеллектуальных, общающихся между собой агентов (роботов), которые могут объединяться в коалицию для решения сложных задач, решение которых невозможно отдельными агентами.

Интересны две группы методов образования коалиций (коалиционных алгоритмов). Первая группа – это алгоритмы образования коалиций, основанные на методах решения кооперативных игр. Задача решается при условии полной информированности всех агентов о целях, ресурсах и стратегиях других агентов, а также, при условии возможности дележа выигрыша, полученного коалицией агентов. Методы решения кооперативных игр (С-ядро, МН-решения и др.) [Оуэн, 1971] имеют экспоненциальную сложность вычислений и коммуникаций между агентами относительно числа агентов, что ограничивает возможность их практического применения [Смирнов и др., 2011].

Вторая группа методов – это методы, основанные на моделировании поведения агентов при образовании коалиций на основе теорий социальной зависимости субъектов [Conte, 2001]. Эти методы позволяют 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 [Conte, 2001].

Далее рассматривается модель, основанная на моделях социального поведения, изложенных в работах социальных психологов Д. Хоманса [Хоманс, 1984] и Л. Фестингера [Фестингер, 1999]. В этих работах

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

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.

1. Базовая модель

Рассмотрим динамическую систему и множество агентов N , которые пытаются изменить состояние системы на выгодное для каждого из них состояние. В качестве математической системы выбрана модель представления экспертных знаний о процессах в динамической системе в виде когнитивной карты [Кулинич, 2012]. В когнитивной карте определено множество факторов ситуации $A=\{a_{ij}\}$, упорядоченное множество лингвистических значений факторов $L_i=\{l_{ij}\}$, $l_{ij}<l_{ij+1}$, $i=1\dots n, j=1\dots m$, причинно-следственные отношения между значениями факторов ситуации $W\subseteq\times_i L_i$. Динамика изменения факторов ситуации в когнитивной карте, задается системой конечно-разностных уравнений:

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

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. $Y(t)=(l_{1j}, \dots, l_{nb})$, $l_{ij} \in L_{ij} \quad \forall i$.

Каждый агент характеризуется 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.

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

3. $e_q:(u(g_q), v(r_q)) \rightarrow R^+$ - 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, $v: r_q \rightarrow R^+$;

4. $\mu_q(Y_q(n), g_q)$ - 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 $(\times L_i)$ определена метрика $\rho(a, b)$, $a, b \in \times L_i$. Тогда возможность достижения цели определяется как близость прогнозной ситуации $Y_q(n)$ и его целевой ситуацией g_q :

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

По сути, этот показатель определяет «силу» каждого агента коалиции без поддержки сторонников, при условии противодействия противников.

1.1. Необходимые и достаточные условия образования коалиции

Сторонниками по коалиции считаются агенты с близкими целями, т.е. $K \subseteq N$, $\rho(g_i, g_q) \leq \varepsilon$, $\forall i, q \in K$, а агенты $M \setminus K$ – это противники. Для оценки устойчивости коалиции, в этой работе, на основе теории социального поведения субъектов [Хоманс, 1999] анализируется взаимная полезность агентов при объединении их ресурсов для достижения общей цели. Считается, что дисбаланс во взаимной полезности агентов в коалиции создает у них ощущения несправедливости, которые увеличивают их мотивации выхода из коалиции, угрожающих ее устойчивости. Для оценки ощущений агентов о несправедливости в коалиции используется теория когнитивного диссонанса [Фестингер, 1999].

1.2. 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 . Т.е. 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:

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

где $\mu(Y_{i+q}, g_i)$, $\mu(Y_q, g_q)$, $\mu(Y_i, g_i) \in [0, 1]$ – возможности достижения агентами q и i своих целей при объединении их стратегий, и, соответственно, собственными силами в условиях противодействия противников.

Социальное поведение в социальной группе – это обмен ценностями между агентами (ресурсами, полезностями и т.д.). Целостность группы достигается при балансе «сил» между обменами [Хоманс, 1984]. Т.е. 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. $e_i=e_q$.

1.3. Критерий когнитивного диссонанса

Наличие дисбаланса взаимной полезности или эффективности агентов, включенных в коалицию, приводит к возникновению латентных конфликтов в коалиции, которые характеризуют ощущения агентов о несправедливости в коалиции, и определяются уровнем 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 [Фестингер, 1999]. Когнитивный диссонанс вызывает у субъекта (агента) стремление уменьшить его или препятствовать его росту. Проявление этого стремления у агентов приводит к изменению его поведения, 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:

$$D^W(q,i) = \frac{P(i,q) - P(q,i)}{P(i,q) + P(q,i)}, D^W(q,i) = -D^W(i,q), \forall i,q, i \neq q. \quad (2)$$

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

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

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

2. Пример

В качестве примера рассмотрим описание в терминах предложенной модели кооперативного взаимодействия агентов, играющих в футбол.

Игру в футбол будем считать динамической системой. Параметры этой системы суть положение игроков на футбольном поле двух противоборствующих команд $A=\{a_i\}$ и $B=\{b_i\}$. Состояние футбольной игры динамической системы будем характеризовать вектором положения всех игроков – агентов на плоскости в последовательные моменты времени: $(x_{ai}(t); y_{ai}(t)), \dots, (x_{bi}(t); y_{bi}(t))$, $i \in N$, где N – число агентов обеих

команд, $t=1, \dots, n$. Кроме этого, заданы координаты мяча $Ball=(x_m(t); y_m(t))$. Координаты ворот противников, заданные интервалами: ворота команды $A - G_A=(x_a; [y_{a1}, y_{a2}])$ и, соответственно, команды $B - G_B=(x_b; [y_{b1}, y_{b2}])$.

Цели игроков команды A – забить мяч в ворота команды B , т.е. $\forall a_i, g_{ai}=(x_m(t); y_m(t))|(x_m(t); y_m(t)) \in (x_b; [y_{b1}, y_{b2}])$, и соответственно, игроков команды B забить мяч в ворота команды A , т.е. $\forall b_i, g_{bi}=(x_m(t); y_m(t))|(x_m(t); y_m(t)) \in (x_a; [y_{a1}, y_{a2}])$.

Стратегии игроков (агентов) – это их следующие действия на каждом шаге дискретного времени: $r1(a_i(b_i), t)$ – движение к мячу; $r2(a_i(b_i), t)$ – движение с мячом к воротам; $r3(a_i(b_i), t)$ – удар по воротам; $r4(a_i(b_i), t)$ – перемещение в позицию паса; $r5(a_i, a_j(b_i, b_j), t)$ – пас партнеру по команде.

Эффективность игрока – это соотношение его ожиданий вознаграждения за игру $u(a_i(b_i))$ к результативности его игры, которая выражается интегральной оценкой его результативных действий: удар по воротам и пас партнеру по команде, т.е. $\forall (r3(a_i(b_i)), r5(a_i, a_j(b_i, b_j), t))$.

Возможность достижения цели игроками команд $A(B)$ определяется его ресурсами: близостью агента к воротам противника, т.е. $d(a(b)_i, G_{B(A)}, t) = \rho((x_{a(b)_i}(t); y_{a(b)_i}(t)), (x_{a(b)_1}; [y_{a(b)_1}, y_{a(b)_2}]))$; близостью противников $d(a_i, b_j, t) = \rho((x_{ai}(t); y_{ai}(t)), (x_{aj}(t); y_{aj}(t)))$; индивидуальной скоростью агента – $V(a(b)_i)$ и его действиями $r1(a_i(b_i), t), \dots, r5(a_i, a_j(b_i, b_j), t)$.

2.1. Кооперация агентов на основе критерия полезности

В работе [Кулинич, 2012] 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 $d(a_i, G_B) < d(a_q, G_B)$, $i \neq q$, $q=1, \dots, N/2$. Т.е. агент полезен, если он находится ближе всех своих партнеров к воротам противника, ($K_G^B=1$).

Агент a_i полезен своим партнерам по критерию близости противников (K^B), если $d(a_i, b_j) > d(a_q, b_j)$, $i \neq q$, $q=1, \dots, N/2$, $j=N/2+1, \dots, N$. Т.е. агент полезен, если он находится дальше всех своих партнеров от игроков противника, ($K^B=1$).

Агент a_i полезен своим партнерам по критерию скорости (K^V), если $V(a_i) > V(a_q)$, $i \neq q$, $q=1, \dots, N/2$. Т.е. агент полезен, если его скорость больше чем у всех партнеров, ($K^V=1$).

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 - α , β , γ , полученных с помощью процедуры парных сравнений. В этом случае,

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

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

2.2. Выбор стратегии агента на основе анализа его полезности

Выбор действий агентов будем рассматривать для двух случаев: агент не владеет мячом; агент владеет мячом. Агент, не владеющий мячом, выбирает действия: $r1(a_i(b_i), t)$ - движение к мячу; $r4(a_i(b_i), t)$ - перемещение в позицию паса. Эти действия не кооперативные и их выбор не основан на анализе полезности агента. Для агента, владеющего мячом, выбор действия основан на анализе полезности партнеров по команде. При условии, что полезность игрока, владеющего мячом выше, чем полезность остальных партнеров, т.е. $P(a(b)_i) > P(a(b)_q)$, $i \neq q$, $q = 1, \dots, N/2$, то агент выбирает действия: $r2(a_i(b_i), t)$ - движение с мячом к воротам; $r3(a_i(b_i), t)$ - удар по воротам.

Если полезность игрока, владеющего мячом ниже, чем полезность одного из партнеров, то выбирается действие: $r5(a_i, a_j(b_i, b_j), t)$ - пас партнеру по команде ($a_j(b_j)$), имеющему большую полезность.

2.3. Устойчивость кооперации агентов

Для анализа устойчивости кооперации агентов необходим анализ взаимной полезности агентов и анализ когнитивного диссонанса их эффективности на протяжении игры.

Взаимная полезность агентов в игре определяется суммарным числом пасов между парой агентов:

$$P(a_i, a_j) = \frac{\sum_{t=1, \dots, n} r5(a_i, a_j, t)}{\sum_{t=1, \dots, n} r5(a_i, a_j, t) + \sum_{t=1, \dots, n} r5(a_j, a_i, t)}, \quad P(a_j, a_i) = 1 - P(a_i, a_j).$$

Здесь $\sum_{t=1, \dots, n} r5(a_i, a_j, t)$ и $\sum_{t=1, \dots, n} r5(a_j, a_i, t)$ 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 (2) и могут изменять свое поведение. Изменение поведения 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:

$$E(a_i) = \frac{\sum_{t=1, \dots, n} r3(a_i, t)}{\sum_{i=1, \dots, N/2} r3(a_i, t)}$$

Агенты с неравной эффективностью характеризуются когнитивным диссонансом по эффективности, вычисляемой по формуле (3) и могут изменять свое поведение, изменения правила выбора агента на основе полезности агента партнера.

2.4. Эксперимент и обсуждение

Предложенная модель кооперации агентов была исследована экспериментально. Для этого в среде моделирования многоагентных систем VisualBots for Excel была разработана программа, моделирующая игру в футбол двух команд, состоящих из четырех игроков. Агенты в каждой команде незначительно отличались по скорости. Считается, что все агенты получают вознаграждение пропорционально количеству забитых голов, независимо от выигрыша или проигрыша команды.

На рисунках 1-4 показаны результаты экспериментов. Здесь по оси абсцисс показаны номера игроков команд, а по оси ординат количество пасов партнерам (рис. 1 и рис. 3) и забитых голов (рис. 2 и рис. 4) в двух сериях игр: без учета и с учетом диссонанса игроков. 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 *A*, игрок №3 команды *B*). Это может быть интерпретировано как самоорганизация агентов в команде. Т.е. 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*. Т.е. все игроки команды *A*, при передаче паса партнеру учитывают не только его полезность, но и эффективность. 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. На рис.1 и рис.2 показано, что результативность распасовки игроков команды *A*, при учете диссонансов агентов, резко снизилась. Снижился 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.3, Рис.4).

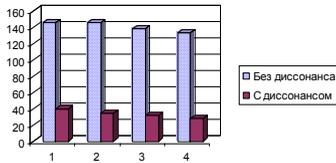


Рис.1. Пасы игроков команды А.

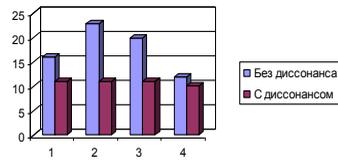


Рис.2. Голы игроков команды А.

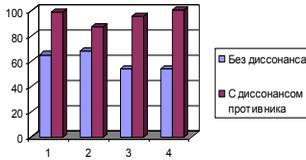


Рис.3. Пасы игроков команды В.

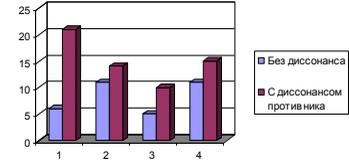


Рис.4. Голы игроков команды В.

Заключение

Имитационное моделирование поведения агентов на основе предложенных математических моделей их кооперации на основе критериев полезности агентов и их когнитивных диссонансов показали правдоподобный результат. Такие модели могут служить алгоритмической основой для реализации поведения автономных роботов в задачах кооперации.

Список литературы

- [Оуэн, 1971] Оуэн Г. Теория игр. М.: Мир, 1971.
- [Смирнов, 2011] Смирнов А.В., Шереметов Л.Б. Модели формирования коалиций кооперативных агентов: состояние и перспективы исследований // Искусственный интеллект и принятие решений. № 1, 2011, стр. 36-48.
- [Conte, 2001] Conte R., Edmonds B., Moss S., Sawyer K., *Sociology and Social Theory in Agent Based Social Simulation: A Symposium. Computational and Mathematical Organization Theory*, 2001, v.7, n.3, pp 183-205.
- [Хоманс, 1984] Хоманс Дж. Социальное поведение как обмен. // Современная зарубежная социальная психология. М., 1984. С. 83-91.
- [Фестингер, 1999] Фестингер Л. Теория когнитивного диссонанса. СПб.: Ювента, 1999.
- [Кулинич, 2012] А.А. Модель поддержки принятия решений для образования коалиций в условиях неопределенности// Искусственный интеллект и принятие решений. № 2, 2012, стр. 95-106.