РЕФЛЕКСИЯ И УСТОЙЧИВОСТЬ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ В МНОГОАГЕНТНЫХ СИСТЕМАХ

Д.А. Новиков (novikov@ipu.ru)
Институт проблем управления РАН, Москва

Аннотация. Рассматривается влияние глубины рефлексии интеллектуальных агентов на устойчивость результатов их взаимодействия. Показано, что в случае, когда рефлексивные отображения агентов не стационарны, их коллективное поведение может быть не устойчиво (по параметрам модели).

Ключевые слова: многоагентные системы, коллективное поведение, рефлексия.

Введение

В многоагентных системах (МАС) [5] «степень интеллектуальности» агента традиционно оценивается разнообразием той информации, на основе которой он «принимает решения». С одной стороны, автономные агенты принимают решения самостоятельно; с другой стороны, взаимодействуя с другими агентами, они вынуждены прогнозировать поведение последних. Другими словами, высокоинтеллектуальный агент должен быть наделен рефлексией, причем, чем выше глубина рефлексии, тем, наверное, более сложное поведение демонстрирует агент. Однако, оказывается, что увеличение глубины рефлексии агентов может входить в противоречие с устойчивостью их коллективного поведения [4].

В настоящей работе рассматриваются свойства рефлексивных отображений агентов: показывается, что в модели принятия агентами решений на основании иерархии своих представлений (о существенных параметрах, представлениях других агентов об этих параметрах, представлениях о представлениях и т.д.) действия, выбираемые фантомными агентами различных уровней, определяются в общем случае системой нелинейных итерированных функций [2]. Исследование поведения МАС позволяет сделать вывод о возможной неустойчивости информационного равновесия [3] по параметрам модели и о нецелесообразности использования агентами высоких рангов рефлексии.

1. Информационное равновесие

Предположим, что имеется множество $N = \{1, 2, ..., n\}$ агентов, и их информированность описывается *информационной структурой* $I = (I_1, I_2, ..., I_n)$, где $I_i = (\theta_i, \theta_{ij}, \theta_{ijk}, \ldots)$, $i, j, k \in N$, – древовидная структура информированности i-го агента; $\theta_i \in \Omega$ – его представления о неопределенном параметре (так называемом «*состоянии природы*», то есть условиях, в которых функционируют агенты); $\theta_{ij} \in \Omega$ – его представления о представлениях j-го агента; $\theta_{ijk} \in \Omega$ – представления i-го агента о том, что j-ый агент «думает» о представлениях k-го агента и т.д. в общем случае до бесконечности [3]. Если задана структура информированности I, то тем самым задана и структура информированности каждого из агентов (как реальных, так и *фантомных* – то есть существующих в сознании других реальных и фантомных агентов). Выбор активным τ-агентом, где τ – некоторая последовательность индексов из множества N, своего действия x_τ в рамках гипотезы рационального поведения (стремления к максимизации своей целевой функции) определяется его структурой информированности I_τ, поэтому, имея эту структуру, можно смоделировать его «рассуждения» и определить его действие. Выбирая свое действие, агент «моделирует» действия других агентов (осуществляет *рефлексию*).

Обозначим Σ_+ – множество всевозможных конечных последовательностей индексов из N, Σ – объединение Σ_+ с пустой последовательностью, $|\sigma|$ – количество индексов в последовательности σ (для пустой последовательности принимается равным нулю).

Коллективное поведение агентов в МАС будем описывать в рамках игровой постановки, считая прогнозируемым состоянием многоагентной системы равновесие игры агентов.

Рефлексивной игрой называется игра «в нормальной форме», задаваемая кортежем $\{N, (A_i)_{i \in N}, (f_i(\cdot))_{i \in N}, I\}$, где N – множество игроков (агентов), A_i – множество допустимых действий i-го игрока, $f_i(\cdot)$: $\Omega \times A' \to \Re^1$ – его целевая функция, $A' = \prod_{i \in N} A_i$, $i \in N$, I – структура информированности. Определим равновесие этой игры. Набор действий x_τ^*, $\tau \in \Sigma_+$, называется *информационным равновесием* [3], если выполнены следующие условия:

1. структура информированности I имеет конечную сложность ν, то есть, дерево I содержит конечный набор попарно различных поддеревьев;
2. $\forall\ \lambda, \mu \in \Sigma_+ \quad I_\lambda = I_\mu \Rightarrow x_\lambda^* = x_\mu^*$;
3. $\forall\ i \in N, \forall\ \sigma \in \Sigma \quad x^*_{\sigma i} \in \operatorname{Arg}\max\limits_{y_i \in A_i} f_i(\theta_{\sigma i}, x^*_{\sigma i 1}, \ldots, x^*_{\sigma i, i-1}, y_i, x^*_{\sigma i, i+1}, \ldots, x^*_{\sigma i n})$.

Будем рассматривать регулярные структуры информированности [3], для задания которых введем вспомогательное понятие *регулярного конечного дерева* (РКД), которое определим рекуррентно. Пусть в игре участвуют n агентов. Если (в простейшем случае) все агенты одинаково информированы, то структура информированности имеет сложность n и единичную глубину. Будем представлять эту ситуацию в виде дерева, состоящего из корневой вершины, n ребер и n висячих вершин. Далее РКД может «расти» следующим образом: к каждой висячей вершине τi, $\tau \in \Sigma$, присоединяется ровно $(n - 1)$ ребро, при этом возникает $(n - 1)$ висячая вершина τij, $j = 1, \ldots, i - 1, i + 1, \ldots, n$. Построенное РКД будем интерпретировать так: если имеется висячая вершина τi, $\tau \in \Sigma$, то τi-агент одинаково информирован с τ-агентом (если τ – пустая последовательность, то τi-агент является реальным, и его субъективные представления совпадают с объективными).

Максимальная глубина k_i РКД i-го реального агента является рангом его рефлексии. Известно, что конечная регулярная информационная структура однозначно (с учетом аксиомы автоинформированности – $\forall\ i \in N\ \forall\ \tau, \sigma \in \Sigma\ \theta_{\tau i i \sigma} = \theta_{\tau i \sigma}$) задается перечислением своих висячих вершин [3].

Частным случаем рефлексивной игры является классическая игра в нормальной форме [7], в которой значение состояния природы θ является *общим знанием* среди участников игры, то есть:

1) о нем известно всем агентам;
2) всем агентам известно 1;
3) всем агентам известно 2 и т.д. до бесконечности.

В условиях общего знания структура информированности тривиальна (представления всех агентов одинаковы), и информационное равновесие превращается в равновесие Нэша. Качественно, общее знание необходимо для того, чтобы каждый из агентов мог промоделировать поведение других агентов. Реализуемость равновесия Нэша подразумевает возможность агентов априори и независимо рассчитать равновесие Нэша, и в одношаговой игре сразу выбрать равновесные по Нэшу действия.

Таким образом, общее знание (единичная глубина рефлексии) представляет собой вырожденную ситуацию тривиальной взаимной информированности агентов. В общем случае, рассмотрение нетривиальной взаимной информированности агентов дает возможность, во-первых (с нормативной точки зрения), расширить множество исходов их игры, что, в свою очередь, увеличивает возможности по управлению МАС [3]. Во вторых (с дескриптивной точки зрения), многие ситуации, которые не могут быть интерпретированы как «обычные» равновесия Нэша в условиях общего знания, являются информационными равновесиями.

2. Рефлексивные отображения

Обозначим множество параметрических (параметр – вектор $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n) \in \Omega^n$) равновесий Нэша

$$E_N(\boldsymbol{\theta}) = \{\{x_i(\boldsymbol{\theta})\}_{i \in N} \in A' \mid \forall\ i \in N, \forall\ y_i \in A_i$$

$$f_i(\theta_i, x_1(\boldsymbol{\theta}), \ldots, x_n(\boldsymbol{\theta})) \geq f_i(\theta_i, x_1(\boldsymbol{\theta}), \ldots, x_{i-1}(\boldsymbol{\theta}), y_i, x_{i+1}(\boldsymbol{\theta}), \ldots, x_n(\boldsymbol{\theta}))\}, \qquad (1)$$

а объединение этих множеств по всевозможным представлениям о значении состоянии природы обозначим $E_N = \bigcup_{(\theta_1, \theta_2, \ldots, \theta_n) \in \Omega^n} E_N(\theta_1, \theta_2, \ldots, \theta_n)$.

Предположим, что на нижнем уровне $\{\theta_{\tau ij}\}_{j \in N}$ конечной регулярной структуры информированности имеет место субъективное общее знание [3] фантомных агентов. Тогда с точки зрения τi-агента возможными являются равновесия их игры из множества $E_N(\{\theta_{\tau ij}\}_{j \in N})$. Введем *множество наилучших ответов* i-го агента на выбор оппонентами действий из множества X_{-i} при множестве Ω возможных состояний природы:

$$\mathrm{BR}_i(\Omega, X_{-i}) = \bigcup_{x_{-i} \in X_{-i},\ \theta \in \Omega} \operatorname{Arg} \max_{x_i \in A_i} f_i(\theta, x_i, x_{-i}),\ i \in N,$$

где $x_{-i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, а также следующие величины и множества: $E_N = \bigcup_{\boldsymbol{\theta} \in \Omega^n} E_N(\boldsymbol{\theta})$,

$X_i^0 = \mathrm{Proj}_i\, E_N$, $X_{-i}^k = \prod_{j \neq i} X_i^k$, $i \in N$, $k = 0, 1, 2, \ldots$, где

$$X_i^k = \mathrm{BR}_i(\Omega, X_{-i}^{k-1}),\ k = 1, 2, \ldots,\ i \in N. \qquad (2)$$

Отображение $\mathrm{BR}_i(\cdot, \cdot)$: $\Omega \times A_{-i} \to A_i$ называется *рефлексивным отображением* i-го агента, $i \in N$. В [3] доказано, что $X_i^k \subseteq X_i^{k+1}$, $k = 0, 1, \ldots$, $i \in N$, то есть с ростом ранга рефлексии множества возможных наилучших ответов агентов не сужаются.

3. Проблема глубины рефлексии агентов

Таким образом, информационное равновесие может быть вычислено следующим образом. Если на нижнем уровне конечной регулярной структуры информированности имеет место субъективное общее знание, то исходом игры соответствующих фантомных агентов будет параметрическое равновесие Нэша (1). Обозначим это равновесие g, $g \in A'$. Тогда агенты следующего (более высокого) уровня выберут действия, являющиеся в рамках их информированности наилучшими ответами на обстановку, соответствующую этому равновесию. Аналогичным образом поступят агенты следующего уровня и т.д., вплоть до реальных агентов. Поясним описанную конструкцию на примере двух агентов. Если на нижнем уровне РКД имеется равновесие g, то с точки зрения, например, первого – реального – агента он должен выбрать действие $x_1 = \mathrm{BR}_1(\theta_1, \mathrm{BR}_2(\theta_{12}, \ldots \mathrm{BR}_i(\theta_{1\tau i}, g)))$ ($i = 1$ или 2 в зависимости от четности глубины РКД). В общем же случае действия реальных и фантомных агентов будут описываться системой итерированных отображений (2), начальной точкой для которых будет параметрическое равновесие Нэша, сложившееся на нижнем уровне РКД.

Рассуждения о свойствах рефлексивных отображений оказываются существенными при рассмотрении задачи о *максимальном целесообразном субъективном ранге рефлексии*, в рамках которой для каждого реального агента требуется определить минимальный ранг рефлексии, при котором он охватывает все многообразие своих возможных выигрышей в рефлексивной игре (при различных вариантах своей структуры информированности). Данная задача является математической формулировкой вопроса о том, какова «оптимальная» глубина рефлексии.

Интуитивно кажется, что чем выше ранг рефлексии агента, тем для него «лучше» (и тем более разнообразное поведение может демонстрировать такой агент), и идеалом является бесконечная глубина рефлексии. Однако на самом деле ответ не столь очевиден. Во-первых, существуют информационные ограничения, которые делают бессмысленными большие значения рангов рефлексии. Во-вторых, существуют многочисленные примеры (см. [3]), свидетельствующие, что увеличение ранга рефлексии агента не приводит к увеличению его выигрыша. Помимо этих двух качественных доводов, можно привести несколько формальных.

Рефлексивное отображение i-го агента называется *стационарным*, если $X_i^k = X_i^{k+1}$, $k = 0, 1, \ldots$. В [3] доказано, что, если рефлексивные отображения агентов стационарны, то максимальный целесообразный ранг рефлексии равен двум и множество действий i-го агента, которые могут быть реализованы как компоненты информационного равновесия, составляет X_i^0, $i \in N$. При этом множество информационных равновесий составляет $E = \prod_{i \in N} X_i^0$. Таким образом, если рефлексивные отображения стационарны, то увеличивать ранг рефлексии, свыше второго, не имеет смысла (условия стационарности рефлексивных отображений для ряда частных случаев описаны в [1]).

Кроме того, во многих случаях увеличение ранга рефлексии приводит к росту неопределенности – неустойчивости коллективного поведения агентов, что может быть объяснено нелинейностью нестационарных рефлексивных отображений. Приведем иллюстративный пример из области «экономических» многоагентных моделей.

4. Пример

Пусть имеются два агента, выбирающих действия из единичного отрезка и имеющих следующие целевые функции (экономической интерпретацией данной модели является «дуополия Курно» [2]): $f_1(\theta, x_1, x_2) = 4\,\theta\, x_1\, x_2\,(1 - x_2) - x_1^2 / 2$, $f_2(\theta, x_1, x_2) = x_1\, x_2 - x_2^2 / 2$, а состояние природы принимает значения из множества $\Omega = (1/4; 1]$.

Предположим, что на нижнем уровне конечного РКД, имеющего глубину m_0, имеет место общее знание с некоторым значением $\theta_0 \in \Omega$ состоянии природы. Вычисляем равновесие Нэша игры фантомных агентов: $x_1(\theta_0) = x_2(\theta_0) = 1 - 1 / (4\,\theta_0)$ и находим наилучшие ответы первого и второго агентов на действия оппонентов:

$$\mathrm{BR}_1(\theta, x_2) = 4\,\theta\, x_2\,(1 - x_2),\ \mathrm{BR}_2(\theta, x_1) = x_1,\ \theta \in \Omega.$$

Получаем, что наилучшие ответы $\tau 1$-агентов, $\tau \in \Sigma$, $|\tau| \le m_0$, удовлетворяют логистическому отображению

$$x_1^m = 4\,\theta\, x_1^{m-1}\,(1 - x_1^{m-1}),\quad m = 1, 2, \ldots, [m_0/2], \tag{3}$$

с начальной точкой $x_1^0 = 1 - 1 / (4\,\theta_0)$ (здесь за $[\cdot]$ обозначена целая часть).

Анализируя (3), получаем, что в зависимости от информированности θ_τ τ-агентов (отметим, что эта информированность для всех агентов с первого по $(m_0 - 2)$-й уровень включительно считается одинаковой, т.е. $\theta_\tau \equiv \theta$ для некоторого $\theta \in \Omega$ при $|\tau| \le m_0 - 2$, и в случае различной информированности агентов может наблюдаться еще более сложное поведение) возможны следующие варианты асимптотически (при $m_0 \to \infty$) устойчивых и слабо зависящих от начальной точки стратегий первого реального агента: выбор единственного действия; периодическое поведение; хаотическое или периодическое поведение. Содержательные интерпретации и недостатки подобной неопределенности коллективного поведения агентов в МАС очевидны.

Заключение

Наряду с информационными ограничениями, нелинейность нестационарных рефлексивных отображений является одной из причин ограниченности ранга рефлексии – увеличение рангов рефлексии агентов приводит к росту неопределенности результатов их взаимодействия и неустойчивости коллективного поведения по параметрам МАС (в частности, взаимным представлениям агентов). Поэтому при разработке и анализе интеллектуальных многоагентных систем, в которых агенты обладают способностью к рефлексии, необходимо тщательное исследование стационарности рефлексивных отображений агентов. Если эти отображения не стационарны, то при выборе глубины рефлексии агентов следует иметь в виду, что ее увеличение может приводить к росту неопределенности исходов взаимодействия агентов.

Список литературы

1 *Казанцев С.Б., Новиков Д.А., Чхартишвили А.Г.* Об условиях стационарности линейных рефлексивных отображений // Системы управления и информационные технологии. – 2004. – № 5. – С. 20 – 22.
2 *Малинецкий Г.Г., Потапов А.Б.* Современные проблемы нелинейной динамики. – М.: Едиториал УРСС, 2002. – 360 с.
3 *Новиков Д.А., Чхартишвили А.Г.* Рефлексивные игры. – М.: Синтег, 2003. – 160 с.
4 *Опойцев В.И.* Равновесие и устойчивость в моделях коллективного поведения. – М.: Наука, 1977. – 248 с.
5 *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 353 с.
6 *Mas-Collel A., Whinston M.D., Green J.R.* Microeconomic theory. – N.Y.: Oxford Univ. Press, 1995. – 981 p.
7 *Myerson R.B.* Game theory: analysis of conflict. – London: Harvard Univ. Press, 1991. – 568 p.