

Об одном подходе к распознаванию рациональности в коллективах агентов¹

Аннотация. Работа посвящена формализации поведения искусственного агента и группы агентов на основе структурированного представления агента и его действий. Предлагаемый подход позволяет осуществлять анализ и предсказание поведения отдельного агента, а также формировать классы агентов на основе детерминант набора действий. Аппарат может быть использован для анализа рациональности как отдельных агентов, так и много-агентных систем.

Ключевые слова: агент, многоагентные системы, ДСМ-метод, автоматическое порождение гипотез, рациональность.

Введение

Понятие «агента» рассматривается многими исследователями как базовое для Искусственного интеллекта (ИИ). Так, русский перевод фундаментального (по объему) труда Рассела и Норвига «Искусственный интеллект» [1] издан с шапкой «Книга об интеллектуальных агентах». Агент понимается здесь как воспринимающий среду с помощью датчиков и воздействующий на нее с помощью исполнительных механизмов. Это позволяет трактовать его чрезвычайно широко – от индивидуумов до компьютерных программ. При этом обязательной характеристикой агента является его рациональность, понимаемая как принятие наилучшего (с точки зрения тех или иных целевых критериев) решения. В работе [2], являющейся почти канонической для специалистов в области многоагентных систем, утверждается, что общепринятого определения «агента» не существует. Тем не менее, определение здесь все-таки дается: агент рассматривается как компьютерная система, находящаяся в некоторой среде и способная автономно действовать в среде своего существования для достижения поставленных целей.

Для авторов работы [3] ключевым является понятие «познающего агента», который обладает способностями к обучению, рассуждению, рефлексии для решения задач. Описанные характеристики представлены крайне формально, что затрудняет возможности их практической интерпретации и реализации. Наконец, наиболее детализированные требования к агенту представлены в [4]: агент должен быть активным и одновременно реактивным (воспринимающим среду), автономным и одновременно коммуникабельным и обязательно обладающим целенаправленным поведением. Всеми этими качествами обладает интеллектуальный робот [5], который, следовательно, и является реализацией «правильного» искусственного агента.

Однако интеллектуальный робот является заключительным звеном цепочки «Интеллектуальные системы – Когнитивные системы – Интеллектуальные роботы», построение которой может рассматриваться как «когнитивная магистраль» для развития искусственного интеллекта [5]. Соответственно, формализацию и изучение деятельности агентов следует начинать с наиболее простого уровня – действующего агента – в особенности, рационального (возможно, с функцией обучения).

¹ Работа выполнена при поддержке РФФИ (проект № 08-01-00797а) и программы фундаментальных исследований Президиума РАН № 14 (проект № 205)

1. Формализация действий агентов

Действующие агенты в сложных для автономной деятельности условиях объединяются в коллективы агентов на самых различных основаниях – от одинакового восприятия тех или иных аспектов окружающей среды до решения общих задач и достижения общих оптимизационных целей. При изучении таких коллективов напрашиваются естественные аналогии между социальными сообществами и общностями искусственных агентов, рассмотрением которых в Искусственном интеллекте занимается направление «Многоагентные системы». Так, в уже упоминавшейся работе [2] подчеркивается, что «...многоагентные системы и социальные науки могут многое подсказать друг другу». Многоагентные системы обеспечивают новые инструменты для анализа социумов, а социальные науки поставляют идеи и понятия для понимания и конструирования многоагентных систем. Тем не менее, достаточно сложно найти работы, явным образом демонстрирующие междисциплинарное взаимодействие этих столь различных научных областей – в [2] специально оговаривается дистанцирование от социальных исследований. Мы же, напротив, хотели бы попытаться распространить формальные средства, предложенные для анализа мнений в социологических опросах [6], на системы (коллективы) искусственных агентов.

Для описания таких коллективов в рамках направления, изучающего многоагентные системы, предлагается множество формальных описаний. Мы будем опираться на вариант описания многоагентной системы (МАС), предназначенный для характеристики ее действий [4]. Итак, многоагентная система $MAC = (Ag, ACT, F, L)$. Здесь $Ag = \{C_1, \dots, C_m\}$ – множество агентов, $ACT = \{p_1, \dots, p_n\}$ – множество действий агентов в МАС, $F: Ag \rightarrow 2^{ACT}$, $ACT_{C_i} = F_i$ – множество действий агента $C_i \in Ag$, L – подмножество расширенного множества ACT' , которое описывает действие всей МАС.

Будем считать, что отношение агента к действию может быть положительным (агент выполняет его) – (+1), отрицательным – (–1), противоречивым (агент может как выполнить, так и не выполнить действие) – (0), и неопределенным (неизвестно, выполняет ли агент действие) – τ . В этом случае $F: Ag \rightarrow 2^{ACT} \times \{1, -1, 0, \tau\}$.

Пусть t – логическая истина, f – логическая ложь, $J_v \Phi$ – оператор, $J_v \Phi = t$, если $v(\Phi) = v$, $J_v \Phi = f$, если $v(\Phi) \neq v$, где $v[\Phi]$ – функция оценки, а переменная v принимает значения из $\{1, -1, 0, \tau\}$ («фактическая истина», «фактическая ложь», «фактическое противоречие», «неопределенность», соответственно). Таким образом, $J_1 p$ означает, что действие p выполняется, $J_{-1} p$ – действие p не выполняется, $J_0 p$ – выполнение действия противоречиво, $J_\tau p$ – выполнение действия не определено. Тогда, по аналогии с [6], мы можем назвать *комплексным* действием ϕ_j j -го агента (X_j) максимальную конъюнкцию атомов $J_{v_i^{(j)}} p_i$,

$\phi_j = J_{v_1^{(j)}} p_1 \& \dots \& J_{v_n^{(j)}} p_n$, где $v_i^{(j)} \in \{\pm 1, 0, \tau\}$, $i = 1, \dots, n$; $j = 1, \dots, 4^n$ (« $\equiv$ » – предикат графического равенства формул). Множество членов этой конъюнкции, представляющее множество действий ACT_{X_j} агента X_j , обозначим $[\phi_j] = \{J_{v_1^{(j)}} p_1, \dots, J_{v_n^{(j)}} p_n\}$. Тогда множество возможных атомарных действий для агентов МАС есть $U^{(2)} = \{\psi \mid (\psi \equiv J_{v_i} p_i) \& (v_i \in \{1, -1, 0, \tau\}), i = 1, \dots, n\}$, число элементов этого множества $|U^{(2)}| = 4n$. Множество возможных комплексных действий $K = \{\phi_i \mid \phi_i = J_{v_1^{(i)}} p_1 \& \dots \& J_{v_n^{(i)}} p_n, v_j^{(i)}[p_j] = v_j^{(i)}, v_j^{(i)} \in \{1, -1, 0, \tau\}, j = 1, \dots, n, i = 1, \dots, 4^n\}$.

2. Аргументация

В последнее время в Искусственном интеллекте растет интерес к изучению аргументации [7] – как логических теорий аргументации, так и практических приложений, в частности, использования аргументации при принятии решений, анализе конфликтов, взаимодействии агентов в многоагентных системах. При этом очевидная связь аргументированного и рационального поведения упускается из виду – рациональное поведение повсеместно понимается (и не только в ИИ – к примеру, [8]) как оптимальное с точки зрения соотношения между целью действия и имеющимися знаниями, объективными возможностями и избранными средствами. Однако если вспомнить о классическом для социологии представлении М. Вебера о рациональном поведении и, соответственно, об аналогии между коллективами искусственных агентов и социумами, естест-

венным оказывается рассмотрение такого вида рациональности, как аргументированное принятие решений (выбор действий), не сводимое к дедуктивному рассуждению [6].

Предлагаемая формализация наиболее естественным образом интерпретируется для интеллектуальных (в частности, интенциональных) агентов, если воспользоваться предложенной в [9] логикой аргументации. В самом деле, можно представить, что когнитивный рассуждающий агент действует (выбирает или не выбирает то или иное действие) на основании некоторых аргументов (представляющих собой, например, внутренние убеждения или мотивы). Тогда A – множество аргументов (аргументационная база) относительно принятия или непринятия некоторых утверждений (выполнения или невыполнения некоторых действий), например, составляющих действия $ACT = \{p_1, \dots, p_n\}$ многоагентной системы МАС. Определим для данного множества ACT функции $g^+(p_i)$ и $g^-(p_i)$ следующим образом: $g^\sigma: ACT \rightarrow 2^A$, где $\sigma \in \{+, -\}$.

Атомарная оценка для множества истинностных значений $\{1, -1, 0, \tau\}$ определяется следующим образом [9]:

$$v[p_i] = 1 \leftrightarrow g^+(p_i) \neq \emptyset, g^-(p_i) = \emptyset;$$

$$v[p_i] = -1 \leftrightarrow g^+(p_i) = \emptyset, g^-(p_i) \neq \emptyset;$$

$$v[p_i] = 0 \leftrightarrow g^+(p_i) \neq \emptyset, g^-(p_i) \neq \emptyset;$$

$$v[p_i] = \tau \leftrightarrow g^+(p_i) = g^-(p_i) = \emptyset;$$

Разумеется, каждый агент X_j имеет свой набор аргументационных функций $g_{X_j}^+, g_{X_j}^-, g_{X_j}^\sigma$, $\sigma \in \{+, -\}$. При этом предполагается, что агент действует рационально, т.е. $\forall X_i \forall p_j (g_{X_i}^+(p_j) \cap g_{X_i}^-(p_j) = \emptyset)$.

Отметим, что аргументационная база A рассматривается как общая для всех агентов – это может быть, к примеру, объединение аргументационных баз всех агентов. Однако если речь идет об описании агента (о чем будет сказано ниже), аргументационная база j -го ($j = 1, \dots, m$) агента A_j , наряду с его аргументационной функцией $\vec{G}_j = \langle g_{X_j}^+(p_1), \dots, g_{X_j}^+(p_n), g_{X_j}^-(p_1), \dots, g_{X_j}^-(p_n) \rangle$, может включаться в это описание.

3. ДСМ-метод и поведение агентов

Как уже говорилось, анализ и предсказание поведения агентов осуществляется средствами

ДСМ-метода автоматического порождения гипотез [10] в рамках модели, предложенной в [6] для анализа мнений респондентов в социологических опросах. Приведем здесь минимальные сведения, необходимые для понимания дальнейшего (подробное описание ДСМ-метода можно найти также в [11]). Условия применимости ДСМ-метода включают возможность алгебраической формализации сходства объектов и их свойств, на основе которого формируются гипотезы о причинах. Соответственно, необходимо структурированное описание объектов и свойств – в виде множеств в самом простом случае. Так, для анализа социального поведения (мнений, в том числе) индивидуума описывается тремя множествами характеристик, представляющими социальный характер субъекта, индивидуальные черты личности и биографические данные (выбор такого описания есть практическая реализация так называемого «постулата поведения»).

Пусть для описания агентов Ag из многоагентной системы МАС используется конечное множество дифференциальных признаков $U^{(1)} = \{d_1, \dots, d_r\}$ (более детально об описании агентов будет сказано ниже), множество возможных действий для агентов есть $U^{(2)} = \{\psi \mid (\psi \models J_{v_i} p_i) \& (v_i \in \{1, -1, 0, \tau\}), i=1, \dots, n\}$.

Массив начальных данных (БФ) содержит утверждения типа “высказывание «агент C_i выбирает действия $[f_i]$ » имеет истинностное значение $\langle v, 0 \rangle$ ”, что выражается элементарной формулой $J_{\langle v, 0 \rangle}(C_i \Rightarrow_1 [f_i])$ в ДСМ-языке [10, 11], $v \in \{1, -1, 0, \tau\}$. В результате применения правил индуктивного вывода (правил правдоподобного вывода 1 рода, п.п.в.-1 для обратного метода²) порождаются гипотезы вида $J_{\langle v, n \rangle}([\psi_i]_3 \Leftarrow C'_i)$, n – номер шага вычислений, выражающий степень правдоподобия истинностного значения, $n > 0$. Это выражение означает, что “высказывание «действия ψ_i есть следствие характеристик агента C'_i » имеет истинностное значение $\langle v, n \rangle$ ”. Как и выше, $J_{\langle v, n \rangle} \phi = t$, если $v[\phi] = \langle v, n \rangle$; $J_{\langle v, n \rangle} \phi = f$, если $v[\phi] \neq \langle v, n \rangle$, $v[\phi]$ есть функция оценки, $\langle v, n \rangle$ представляет

² Обратный ДСМ-метод формализует вывод «от следствия – к причине» и используется для анализа явлений, где свойства представлены подробнее самих объектов. Например, мнения респондентов, как правило, представлены подробнее самих респондентов.

«внутренние» истинностные значения фактов и гипотез, t , f – «внешние» истинностные значения двузначной логики. Порожденные детерминанты действий в дальнейшем используются для прогнозирования действий с помощью правил вывода по аналогии (п.п.в.-2), а также могут служить основанием для выделения классов агентов с детерминированными действиями. Здесь C_i , C'_i , $[\varphi_i]$, $[\psi_i]$ – константы, C_i , $C'_i \in 2^{U^{(1)}}$, $[\varphi_i]$, $[\psi_i] \in 2^{U^{(2)}}$, высказывания $J_{(v,0)}(C \Rightarrow_1 [\varphi])$ суть факты, $J_{(v,n)}([\psi] \Leftarrow C)$, $n > 0$ – гипотезы.

Приведем краткое описание процедур индуктивного анализа сходства мнений, а также вывода по аналогии и абдуктивного принятия гипотез в ДСМ-методе; детальное описание можно найти в [10, 11].

В [11] показано, что необходимыми составляющими предикатов, уточняющих и формализующих индуктивные методы Д.С. Милля и использующихся в качестве посылок при формулировке правил п.п.в.-1, являются экзистенциальные условия (ЭУ) существования не менее двух ($k \geq 2$) сходных позитивных (+) или негативных (–) примеров, и условия исчерпываемости (УИ), т.е. рассмотрение как можно более полного множества сходных примеров. При этом предикаты описывают эмпирическую зависимость (ЭЗ), представляющую причинное вынуждение соответствующего следствия.

Так, предикат простого обратного положительного сходства $\tilde{M}_{a,n}^+(V, W, k)$ распознает локальное

сходство $(\bigcap_{i=1}^k C_i = C') \& (C' \neq \emptyset) \& ((\bigcap_{i=1}^k Q_i = Q') \& (Q' \neq \emptyset))$

на множестве (+)-примеров $J_{(1,n)}(C_i \Rightarrow_1 Q_i)$, $i = 1, \dots, k$, ($k \geq 2$), которое является основанием для правдоподобного вывода. Эмпирическая зависимость типа «сходство мнений субъектов в (+)-примерах влечет сходство самих субъектов и притом для всех рассматриваемых мнений (условие исчерпываемости мнений в (+)-примерах)» представлена подформулой $\forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \& \forall U (J_{(1,n)}(X \Rightarrow_1 U) \rightarrow U \subseteq Y) \& W \subseteq Y) \rightarrow (V \subseteq X \& V \neq \emptyset))$. В правилах же используются непараметрические предикаты простого обратного положительного сходства $\tilde{M}_{a,n}^+(V, W) \Rightarrow \exists k \tilde{M}_{a,n}^+(V, W, k)$ и симметрично формулируемый предикат простого обратного отрицательного сходства $\tilde{M}_{a,n}^-(V, W)$.

Случаи неопределенности в БФ $J_{(t,n)}(V \Rightarrow_1 W)$ уточняются с помощью п.п.в.-2 – выводов по аналогии, использующих в качестве посылок предикаты $\tilde{P}_n^\sigma(V, W) \Rightarrow \exists k \tilde{P}_n^\sigma(V, W, k)$, $\sigma = +, -$.

Подформула $\bigcap_{i=1}^k \exists X_i (J_{(1,n)}(Y_i \Leftarrow X_i) \& X_i \subseteq V \& (\bigcup_{i=1}^k Y_i = W))$ предиката $\tilde{P}_n^+(V, W, k)$ выражает

утверждение, что V содержит положительные причины $X_1, \dots, X_k$ (полученные обратным методом) наличия свойств (мнений, действий) $Y_1, \dots, Y_k$ соответственно, причем изучаемый эффект W полностью покрывается этими множествами свойств $(\bigcup_{i=1}^k Y_i = W, k - \text{параметр})$. При

этом также существенно выполнение условия «исчерпываемости» – рассмотрение всех релевантных свойств. Предикат также включает запрет на включение в рассматриваемый объект V отрицательных или противоречивых причин Z для любых подмножеств U свойств из W : $\forall U (((U \subseteq W) \& U \neq \emptyset) \rightarrow \neg \exists Z ((J_{(1,n)}(U \Leftarrow Z) \vee J_{(0,n)}(Z \Rightarrow_2 U) \& Z \subseteq V)))$. Предикат $\tilde{P}_n^-(V, W)$ определяется аналогично.

Завершающим этапом ДСМ-рассуждения является абдуктивное объяснение начального состояния БФ, т.е. принятие порожденных гипотез первого и второго рода на основании проверки так называемых аксиом каузальной полноты предметной области. Смысл АКП⁽⁺⁾ состоит в следующем: для каждого (+)-факта «объект C обладает множеством свойств Q » из начального состояния БФ, $J_{(1,0)}(C \Rightarrow_1 Q)$, существуют (+)-причины C' для фрагментов Q' , $J_{(1,n)}(Q' \Leftarrow C')$, такие, что $C' \subseteq C$ и свойства Q полностью покрываются фрагментами Q' , $\bigcup Q' = Q$. Аналогично формулируется АКП⁽⁻⁾.

Как видно из сказанного, анализ средствами ДСМ-метода предполагает определение сходства рассматриваемых объектов и их свойств. В нашем случае объектами анализа являются агенты, их свойствами – определенные выше варианты поведения $[\varphi_j]$, $[\varphi] \in 2^{U^{(2)}}$. При этом, как уже говорилось, описание самих агентов, может включать самые различные характеристики, в том числе – для рациональных агентов, – аргументационную базу и аргументационные функции агента, $Ag_j = \langle C_j, \langle A_j, g_{X_j}^+, g_{X_j}^- \rangle \rangle$ или

$Ag_j = \langle C_j, \langle A_j, \vec{G}_j \rangle \rangle$. Если аргументационная база общая для всех агентов, ее не нужно включать в описание $Ag_j = \langle C_j, \vec{G}_j \rangle$. Здесь C_j – характеристики агента, существенные для используемой модели МАС. Так, например, интенциональный агент может быть описан тремя множествами (ср. с постулатом поведения): $B \subseteq Bel$ – множество убеждений агента, $D \subseteq Des$ – множество его желаний, $I \subseteq Int$ – его намерения [4]. Сходство агентов с единой аргументационной базой A определяется как $Ag_i \sqcap Ag_j = \langle C_i \cap C_j, \vec{G}_i \cap \vec{G}_j \rangle$, где $C_i \cap C_j$ – сходство выбранных характеристик i -го и j -го агентов, $\vec{G}_i \cap \vec{G}_j$ – пересечение кортежей $\vec{G}_i = \langle g_{X_i}^+(p_1), \dots, g_{X_i}^+(p_n), g_{X_i}^-(p_1), \dots, g_{X_i}^-(p_n) \rangle$ и $\vec{G}_j = \langle g_{X_j}^+(p_1), \dots, g_{X_j}^+(p_n), g_{X_j}^-(p_1), \dots, g_{X_j}^-(p_n) \rangle$, определяемое покомпонентным пересечением $g_{X_i}^\sigma(p_l) \cap g_{X_j}^\sigma(p_l)$, где $\sigma \in \{+, -\}$, $l = 1, \dots, n$ (n – число возможных действий множества АСТ = $\{p_1, \dots, p_n\}$), $g_{X_i}^\sigma(p_l) \subseteq A$, $g_{X_j}^\sigma(p_l) \subseteq A$, (аргументационная база A – множество аргументов относительно принятия или непринятия действий из АСТ). Подчеркнем, однако, что выбор характеристик описания агентов остается за рамками нашей работы и целиком находится в компетенции создателей МАС.

Отвлечемся от конкретного представления агентов, будем считать, что задана операция сходства агентов $Ag_i \sqcap Ag_j$ и определено включение $Ag_i \subset Ag_j = \langle C_i \subset C_j, \vec{G}_i \subseteq \vec{G}_j \rangle$, где $\vec{G}_i \subseteq \vec{G}_j$

$\Rightarrow \bigwedge_{s=1}^n ((g_i^+(p_s) \subseteq g_j^+(p_s)) \& (g_i^-(p_s) \subseteq g_j^-(p_s)))$, $p_h \in \text{АСТ}$, $h = 1, \dots, n$. Исходная БФ содержит высказывания $J_{\langle v, 0 \rangle}(\langle C_j, \langle g_j^+, g_j^- \rangle \Rightarrow_1 [\phi_j] \rangle)$, $v \in \{1, -1, 0, \tau\}$, $j = 1, \dots, s$, s – число примеров. В результате анализа исходных фактов (агентов и их действий) средствами ДСМ-метода порождаются гипотезы вида $J_{\langle v, n \rangle}([\psi] \Leftarrow \langle C', \vec{G}' \rangle)$, где $v \in \{1, -1, 0, \tau\}$, C' – пересечение описаний агентов $C_{j_1}, \dots, C_{j_h}$, $\vec{G}'$ – результат пересечения $\vec{G}_{j_1}, \dots, \vec{G}_{j_h}$. Эти гипотезы представляют детерминанты, определяющие тот или иной выбор действий агента. А включение в представление знаний об агенте X_j ($j = 1, \dots, m$)

компонентов $\vec{G}_j = \langle g_{X_j}^+(p_1), \dots, g_{X_j}^+(p_n), g_{X_j}^-(p_1), \dots, g_{X_j}^-(p_n) \rangle$, $g_{X_j}^+(p_i) \subseteq A$, $g_{X_j}^-(p_i) \subseteq A$, $p_i \in \text{АСТ}$, $i=1, \dots, n$, позволяет явным образом выразить аргументацию агента при выполнении действий ϕ_j , т.е. рассмотреть рациональное поведение агента.

Выделение детерминант, помимо прочего, позволяет отсеять «лишние» характеристики агента, не влияющие на его поведение, и даёт возможность использовать более «компактного» агента, наделенного лишь необходимыми функциями.

4. Предсказание действий агентов

Рассмотрим множество комплексных действий K' агентов системы МАС, $K' \subseteq K$, $K' = \{\phi | \exists X \exists v J_{\langle v, 0 \rangle}(X \Rightarrow_1 [\phi]) \& (v \in \{1, -1, 0, \tau\})\}$ (K определено выше). Множество агентов Ag представлено как агентами Ag' , действия которых заранее определены, так и агентами Ag^* , действия которых необходимо предсказать, $Ag = Ag' \cup Ag^*$.

$Ag' = \{X | \exists \phi \exists v J_{\langle v, 0 \rangle}(X \Rightarrow_1 [\phi]) \& (\phi \in K') \& (v \in \{1, -1, 0\})\}$, $Ag^* = \{X | \exists \phi J_{\langle \tau, 0 \rangle}(X \Rightarrow_1 [\phi]) \& (\phi \in K')\}$.

Предсказание действий отдельного агента из Ag^* в рассматриваемой модели выполняется с помощью правил вывода по аналогии ДСМ-метода п.п.в.-2, использующих в качестве посылок предикаты $\check{\Pi}_n^\sigma(V, W)$, $\sigma \in \{+, -, 0, \tau\}$. Так, правило $(\check{\Pi})_n^+$ имеет вид

$$(\check{\Pi})_n^+ \frac{J_{\langle \tau, n \rangle}(V \Rightarrow_1 W), \check{\Pi}_n^+(V, W)}{J_{\langle 1, n+1 \rangle}(V \Rightarrow_1 W)}.$$

Аналогично определяются правила $(\check{\Pi})_n^\sigma$ ($\sigma \in \{-, 0, \tau\}$), $\check{\Pi}_n^\tau(V, W) = \neg(\check{\Pi}_n^+(V, W) \vee \check{\Pi}_n^-(V, W) \vee \check{\Pi}_n^0(V, W))$.

Схема предсказания принципиально не меняется и в случае, когда требуется рассмотреть ранее не встречавшиеся (комплексные) действия, т.е. $Ag^* = \{X | \exists \phi J_{\langle \tau, 0 \rangle}(X \Rightarrow_1 [\phi]) \& (\phi \in \tilde{K} = K \setminus K')\}$. Это объясняется особенностями вывода по аналогии в ДСМ-методе: действия ϕ из $\tilde{K} = K \setminus K'$ могут покрываться фрагментами ψ , для которых порождены детерминанты на основании примеров с действиями из K' . Иными словами,

описание агента C из $Ag^* (J_{(\tau, 0)}(C \Rightarrow_1[\varphi]))$ включает (например) положительные причины C'_i (полученные, напомним, на основании индуктивного вывода из имеющихся примеров), $C'_i \subset C$, $J_{(1, n)}([\psi_i] \Leftarrow C'_i)$, $i = 1, \dots, k$, причём $\bigcup_{i=1}^k [\psi_i] = [\varphi]$, но $(-)$ - и (0) -причины (для $[\psi_i]$)

не содержатся в C . Тогда порождается гипотеза $J_{(1, n+1)}(C \Rightarrow_1[\varphi])$. Аналогично порождаются гипотезы для типов истинностных значений $-1, 0, \tau$. Предсказанные действия (на n -м шаге ДСМ-рассуждений) представлены множеством $K_n^* = K_{1,n}^* \cup K_{-1,n}^* \cup K_{0,n}^* \cup K_{\tau,n}^*$, где $K_{v,n}^* = \{\varphi | \exists X J_{(v, n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1)\}$ – множество предсказанных мнений с оценкой v , $v \in \{\pm 1, 0\}$, $K_{\tau,n}^* = \{\varphi | \exists X J_{(\tau, n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1)\}$. Множество ранее не встречавшихся действий $\tilde{K}_n^* = K_n^* \cap \tilde{K}$, $\tilde{K}_{v,n}^* = K_{v,n}^* \cap \tilde{K}$, $v \in \{\pm 1, 0, \tau\}$. Если на m -м шаге достигается стабилизация процесса ДСМ-рассуждений [11], имеем $K^* = K_1^* \cup K_{-1}^* \cup K_0^* \cup K_{\tau}^*$, $K_v^* = \{\varphi | \exists X J_{(v, m)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (m \geq 1)\}$, $v \in \{\pm 1, 0, \tau\}$.

Отдельного подхода требует задача возможного предсказания поведения агента, если задано только его описание (без указания φ). Пусть задано множество действий агентов $K' = \{\varphi | \exists X \exists v J_{(v, 0)}(X \Rightarrow_1[\varphi]) \& (v \in \{1, -1, 0\})\}$, множество агентов представлено как $Ag = Ag' \cup \tilde{Ag}$ (множество агентов Ag' с известными действиями определено выше), $\tilde{Ag} = \{X | \forall \varphi J_{(\tau, 0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K)\}$. Итак, пусть $Ag' = \{C_1, \dots, C_q\}$, $K' = \{\varphi_1, \dots, \varphi_q\}$, $\tilde{Ag} = \{C_{q+1}, \dots, C_s\}$. В этом случае каждому агенту C_j ($j = q+1, \dots, s$) могут быть сопоставлены множества возможных действий $\tilde{K}^j = \tilde{K}_+^j \cup \tilde{K}_-^j \cup \tilde{K}_0^j$, $\tilde{K}_v^j = \{Y | \tilde{P}_n^v(C_j, Y) \& (Y = [\varphi]) \& (\varphi \in K)\}$, $v \in \{+1, -1, 0\}$, где $\tilde{P}_n^v$ – соответствующие предикаты для п.п.в.-2. Соответственно, $\tilde{K}_\tau^j = K \setminus \tilde{K}^j$. Все возможные (предсказанные)

действия агентов $K^* = \bigcup_{j=q+1}^s \tilde{K}^j$.

Разумеется, описанная процедура предсказания возможных действий агентов может рассматриваться лишь как теоретическая – ввиду высокой сложности вычислений. Можно пред-

ложить её упрощение, в особенности, если нас интересует, в первую очередь, рациональное поведение. Заметим, что поскольку для искусственного агента возможно внешнее назначение действий, естественно выбирать из возможных предсказанных действий те, которые можно считать рациональными (в том или ином смысле).

Предположим, что между возможными действиями агентов из множества АСТ = $\{p_1, \dots, p_n\}$ существуют зависимости (к примеру, выбор одного из действий должен сопровождаться отказом от другого действия). Эти зависимости могут быть логически выражены в виде непротиворечивого множества $\Sigma = \{\psi_1, \dots, \psi_r\}$ формул логики JA_4 [6] – варианта 4-х-значной логики аргументации [9], специально созданного для анализа рациональности мнений в социологических опросах. Непротиворечивость Σ устанавливается методом аналитических таблиц для логики JA_4 , также предложенным в [6]. Аналитическая таблица для непротиворечивого множества Σ является незамкнутой. Предполагается при этом, что $\psi_1 \& \dots \& \psi_r$ не является тавтологией, что также устанавливается методом аналитических таблиц (аналитическая таблица с корнем $f(\psi_1 \& \dots \& \psi_r)$ также является незамкнутой, t – логическая истина, f – логическая ложь).

Пусть $consis(\Sigma \cup \{\varphi\})$ – метапредикат непротиворечивости множества формул $(\Sigma \cup \{\varphi\})$. Тогда множество всех φ (комплексных действий агентов из определенного выше множества K), не противоречащих Σ , обозначим $K^+ = \{\varphi | consis(\Sigma \cup \{\varphi\}) \& (\varphi \in K)\}$; множество действий φ , противоречащих Σ , обозначим $\Delta = \{\varphi | \neg consis(\Sigma \cup \{\varphi\}) \& (\varphi \in K)\}$. Соответственно, действия из K^+ могут рассматриваться как рациональные (не противоречащие существующим зависимостям между элементами АСТ, мы будем называть их непротиворечивыми).

Тогда предсказание (точнее, выбор) возможного поведения агентов $C_{q+1}, \dots, C_s$ из $\tilde{Ag}$ может осуществляться на основе постепенного снижения рациональности поведения.

1). Рассмотрим множество имеющихся в системе агентов непротиворечивых действий, $K' \cap K^+$.

Будем считать, что $\tilde{Ag} = \{X | \exists \varphi J_{(\tau, 0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K' \cap K^+)\}$. Тогда доопределение поведения

осуществляется стандартной процедурой вывода по аналогии в ДСМ-методе. Если для всех C_j ($j = q+1, \dots, s$) удалось породить гипотезы вида $J_{(v,n)}(C_j \Rightarrow_1 [\varphi_j])$, $\varphi \in K' \cap K^+$, $v \in \{+1, -1, 0\}$, процедура заканчивается.

2). Агентов, поведение которых не удалось предсказать на предыдущем шаге, рассмотрим

как множество $\tilde{Ag}_1 = \{X | \exists \varphi J_{(\tau, 0)}(X \Rightarrow_1 [\varphi]) \& (\varphi \in K' \setminus (K' \cap K^+))\}$, т.е. нас интересуют непротиворечивые комплексные действия, не встречавшиеся ранее. Дальнейшая процедура – как описано выше.

3). Рассмотрим далее $\tilde{Ag}_2 = \{X | \exists \varphi J_{(\tau, 0)}(X \Rightarrow_1 [\varphi]) \& (\varphi \in K' \setminus (K' \cap K^+))\}$, т.е. доопределяются существующие противоречивые действия.

4). Наконец, $\tilde{Ag}_3 = \{X | \exists \varphi J_{(\tau, 0)}(X \Rightarrow_1 [\varphi]) \& (\varphi \in K' \setminus (K' \cap K^+))\}$, т.е. рассматриваются несуществующие противоречивые действия.

В наихудшем случае потребуются провести все четыре этапа доопределения, причем обязательно в конце концов всё будет доопределено. Это обстоятельство может быть связано с

тем, что в описание агента C из $\tilde{Ag}$ включаются, предположим, порожденные положительные причины C'_i , $C'_i \subset C$, $J_{(1,n)}([\psi_i] \Leftarrow C'_i)$, $i = 1, \dots, k$, одновременно не включаются $(-)$ - и (0) -причины для подмножеств φ , но $\bigcup_{i=1}^k [\psi_i] \neq [\varphi]$,

т.е. не все возможные действия из АСТ определены для C (напомним, что φ – максимальная конъюнкция логики JA_4). Поскольку, однако, требование определенности по отношению к каждому возможному действию не всегда актуально для агентов, может быть предложен ослабленный предикат для п.п.в.-2, в котором условие $\bigcup_{i=1}^k [\psi_i] = [\varphi]$ может быть заменено на

$$\bigcup_{i=1}^k [\psi_i] \subseteq [\varphi].$$

Однако, согласие на рассмотрение объединения $\bigcup_{i=1}^k [\psi_i]$, лишь частично покрывающего $[\varphi]$, порождает при реализации такого предиката для предсказания поведения агентов из $\tilde{Ag}$

(имеющих описание, но не имеющих действий) переборные трудности – разумеется, наибольший интерес имеют максимально возможные объединения. Далее, хотелось бы в этом случае избежать громоздкого сопоставления агенту последовательно каждого из возможных 4^n комплексных действий. Можно предложить следующую реализацию доопределения заранее не заданных φ , причем с учетом их рациональности. Рассмотрим процедуру для положительного доопределения (отрицательное осуществляется аналогично, противоречивое предполагает одновременное вхождение $(+)$ - и $(-)$ - гипотез). Пусть для агента C существует l гипотез таких, что C'_i , $C'_i \subset C$, $J_{(1,n)}([\psi_i] \Leftarrow C'_i)$, $i = 1, \dots, l$, но $\forall \varphi (\bigcup_{i=1}^l [\psi_i] \neq [\varphi] \& \varphi \in K)$ (будем счи-

тать, что отрицательные и противоречивые причины для соответствующих наборов действий в описание агента не входят). Упорядочим конъюнкции $\psi_1, \dots, \psi_l$ по сложности, в случае совпадения сложности выполняется лексикографическое упорядочение (по номерам p_k , $k = 1, \dots, n$, входящих атомов $J_{v_k^{(i)}} p_k$). Методом аналитических таблиц будем последовательно проверять $consis(\Sigma \cup \{\psi_1\})$, $\dots$, $consis(\Sigma \cup \{\psi_l\})$, найдем все непротиворечивые $\psi_{i_1}, \dots, \psi_{i_h}$. Открытые ветви аналитических таблиц с корнями $\Sigma \cup \{\psi_{i_1}\}$, $\dots$, $\Sigma \cup \{\psi_{i_h}\}$ достроим до максимальных конъюнкций логики JA_4 с помощью закона исключенного 5-го $((J_1 p \vee J_{-1} p \vee J_0 p \vee J_{0\tau} p) \leftrightarrow 1)$. Полученные выражения и будут представлять собой множество возможных рациональных комплексных действий агента C . Поскольку для искусственного агента возможно внешнее назначение предпочтительных действий, можно остановиться, дойдя до первого непротиворечивого ψ_{i_1} – и далее по указанной процедуре.

Так как $\psi_{i_1}, \dots, \psi_{i_h}$ также упорядочены по сложности, возможна модификация указанной процедуры. Будем последовательно рассматривать конъюнкции ψ_{i_1} , $\psi_{i_1} \& \psi_{i_2}$, $\dots$, $\psi_{i_1} \& \dots \& \psi_{i_h}$, следя при этом, чтобы $J_{v\mu} p_i$ и $J_{\mu\tau} p_i$, где $v \neq \mu$, $v, \mu \in \{\pm 1, 0, \tau\}$ не входили в конъюнкцию одновременно (в противном случае возникает противоречие, которое не слишком информативно при описании действий агентов). Как только это условие нарушается, отбрасываем добав-

ленное ψ_{i_p} и переходим к следующему. Новое построение начинаем с отброшенного ψ_{i_p} и повторяем всю процедуру. Таким образом мы получим несколько фрагментов максимальных конъюнкций из K , причем все они будут согласованы с Σ . Достроенные, как было описано выше, до максимальных конъюнкций из K^+ , они будут представлять возможные рациональные действия агента C .

Отметим, что абдуктивное объяснение и принятие полученных гипотез [11], являющееся завершением полной процедуры ДСМ-рассуждений, осуществляется для коллективов агентов стандартным образом. Заметим только, что для искусственных агентов невыполнение критерия достаточного основания принятия гипотез – аксиом каузальной полноты, описанных выше, – означает необходимость расширения имеющейся МАС. С одной стороны, это не всегда возможно, с другой же стороны, напротив, такое расширение может осуществляться искусственным подбором агентов с необходимым набором действий – например, для повышения степени каузальной полноты [11] (отношения числа агентов, действия которых объяснены, к общему числу агентов в системе).

5. Поведение коллектива агентов

В соответствии с правилами вывода по аналогии, для агента X , выполняющего действия $[\varphi]$, выполняется условие $\forall X(\tilde{P}_n^+(X, [\varphi]) \rightarrow J_{\langle 1, n+1 \rangle}(X \Rightarrow_1 [\varphi]))$, где $n \geq 0$. Следствием этого утверждения (теорема обратимости ДСМ-рассуждений [11]) является $\forall X \forall V (J_{\langle 1, n+1 \rangle}(X \Rightarrow_1 [\varphi]) \rightarrow (J_{\langle -1, n \rangle}([\varphi] \Leftarrow V) \rightarrow \neg (V \subset X)))$. Аналогично формулируется условие для $\tilde{P}_n^-(X, Y)$.

Пусть, далее, $Ag = \{C_1, \dots, C_m\}$ – множество всех агентов. Для решения задачи о поведении коллектива агентов рассмотрим каузальное покрытие φ , удовлетворяющее следующим условиям:

(a) $J_{\langle 1, n \rangle}([\psi_i] \Leftarrow Det_i)$, $i = 1, \dots, l$; $l \leq m$; где $n = \bigvee_{i=0}^n J_{\langle 1, i \rangle}$; здесь Det_i – детерминанты, полученные применением п.п.в.-1 для обратного метода, причем, как было сказано выше, Det_i , в зависимости от описания, могут включать дифференциальные признаки, аргументационные функции и аргументационную базу агента;

(b) $Ag = \{C_1, \dots, C_m\}$;

(c) $J_{\langle v, 0 \rangle}(C_j \Rightarrow_1 [\varphi_j])$, $j = 1, \dots, m$, где C_j – информативное описание агента X_j (индивидуальная константа) и существуют $J_{\langle 1, 0 \rangle}(C_i \Rightarrow_1 [\varphi_i])$, $1 \leq i \leq m$;

(d) $Det_j \subset C_j$, $j = 1, \dots, l$; $l \leq m$;

(e) $[\psi_1] \cup \dots \cup [\psi_l] = [\varphi]$, $2 \leq l \leq m$.

Если имеет место (a) – (e), тогда истинно

$J_{\langle 1, n+1 \rangle}(Ag \Rightarrow_1 [\varphi])$ при условии $\tilde{P}_n^+(X_j, [\psi_j])$ для всех X_j , $j = 1, \dots, m$. Таким образом, мы получаем возможность прогнозирования коллективного поведения агентов.

Расширение языка ДСМ-метода посредством введения термина $\{X_1, \dots, X_m\}$ (множества агентов) и формул $\{X_1, \dots, X_m\} \Rightarrow_1 Y$, где Y – переменная для мнений агентов и, соответственно, формул $J_{\bar{v}}(\{C_1, \dots, C_m\} \Rightarrow_1 [\varphi])$ ($\bar{v} = \langle v, n \rangle$, $v \in \{1, -1, 0, \tau\}$) создает возможность представления фактов в базе фактов (БФ) для коллективного поведения агентов. Таким образом, возникает задача формирования БФ о коллективном поведении агентов с использованием БФ об индивидуальных мнениях агентов. Решение её должно быть основано на результатах решения задачи о каузальном покрытии мнения φ . С этой целью для $Ag = \{C_1, \dots, C_m\}$ определяются функции $g_{Ag}^+(p_j) = g_{C_1}^+(p_j) \cap \dots \cap g_{C_m}^+(p_j)$ и $g_{Ag}^-(p_j) = g_{C_1}^-(p_j) \cap \dots \cap g_{C_m}^-(p_j)$, $p_j \in ACT$, $j = 1, \dots, n$.

Посредством $G_{Ag}^+ = \{g_{Ag}^+(p_1), \dots, g_{Ag}^+(p_n)\}$ и $G_{Ag}^- = \{g_{Ag}^-(p_1), \dots, g_{Ag}^-(p_n)\}$ аналогично [9] определяются функции оценки для формул $\{C_1, \dots, C_m\} \Rightarrow_1 [\varphi]$.

Однако, корректное описание коллективного поведения агентов средствами ДСМ-метода возможно, если сформулировать абдуктивное принятие порожденных гипотез ($J_{\langle v, n+1 \rangle}(Ag \Rightarrow_1 [\varphi])$, $v \in \{1, -1, 0\}$ или $J_{\langle \tau, n+1 \rangle}(Ag \Rightarrow_1 [\varphi])$) посредством объяснения начального состояния базы фактов [10, 11].

Рассмотрим еще один принцип оценивания высказываний о коллективе агентов посредством расширения ДСМ-метода. Пусть для $p_k \in ACT$ имеют место условия

(*) $\forall X_i \forall X_j \forall p_k (g_{X_i}^+(p_k) \cap g_{X_j}^-(p_k) = \emptyset)$. Это оз-

начает согласованность действий агентов и может рассматриваться как условие сильной рациональности.

(**) $\forall X_i \forall p_k (g_{X_i}^+(p_k) \cap g_{X_i}^-(p_k) = \emptyset)$. Это усло-

вие рациональности отдельного агента.

Тогда, если $\neg(*)$ или $\neg(**)$, то $J_{\langle 0, 0 \rangle}(Ag \Rightarrow_1 [\varphi])$ истинно, где $Ag = \{C_1, \dots, C_m\}$.

6. Рациональность коллектива агентов

Будем, как и прежде, считать, что действия агентов рациональны, если они не противоречат зависимостям Σ между возможными действиями. Рассмотрим для коллектива агентов такие характеристики рациональности, как степень непротиворечивости и степень противоречивости действий. Эти величины задаются, соответственно, функциями $\eta(K', K^+) = |K' \cap K^+|/|K'|$ и $\xi(K', \Delta) = |K' \cap \Delta|/|K'|$ [12], $\eta(K', K^+) + \xi(K', \Delta) = 1$ ³. Соответственно, можно определить степень v -непротиворечивости опроса $\eta^{(v)}(K'_v, K^+_v) = |K'_v \cap K^+_v|/|K'_v|$, $v \in \{+1, -1, 0, \tau\}$.

Пусть, как было определено выше, $Ag = Ag' \cup Ag^*$, $Ag' = \{X|\exists \varphi \exists v_{(v,0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K') \& (v \in \{1, -1, 0\})\}$. Если при этом $Ag^* = \{X|\exists \varphi J_{(\tau,0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K')\}$, степень непротиворечивости при предсказании действий агентов из Ag^* не изменится, поскольку не изменится K' и, соответственно, $K' \cap K^+$. Рассмотрим еще одну характеристику рациональности – степень непротиворечивости действий коллектива агентов как единого сообщества [13]. Пусть $Ag^+ = \{X|\exists \varphi \exists v_{(v,0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K' \cap K^+) \& (v \in \{1, -1, 0\})\}$ – множество агентов с непротиворечивыми действиями. Тогда степень непротиворечивости коллектива агентов $\eta_{Ag}(Ag, Ag^+) = |Ag \cap Ag^+|/|Ag|$, причем $|Ag| \neq |K'|$ (разные агенты могут иметь одинаковый выбор действий), $|Ag^+| \neq |K^+|$, соответственно, $\eta(K', K^+) \neq \eta_{Ag}(Ag, Ag^+)$. Пусть n – номер шага стабилизации ДСМ-рассуждений (новые предсказанные действия не появляются)⁴, $Ag^{*+} = \{X|\exists n \exists \varphi \exists v_{(v,n)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K' \cap K^+) \& (v \in \{1, -1, 0\}) \& (n \geq 1)\}$. Поскольку в общем случае $|Ag^{*+}| \neq |Ag^+|$, при предсказании действий для агентов рациональность коллектива агентов может измениться, $\eta_{Ag}(Ag, Ag^{*+}) \neq \eta_{Ag}(Ag, Ag^+)$.

Рассмотрим случай $Ag^* = \{X|\exists \varphi J_{(\tau,0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in \tilde{K})\}$, где $\tilde{K} = K \setminus K'$. На n -м шаге ДСМ-рассуждений множество предсказанных действий с оценкой v , $v \in \{\pm 1, 0\}$, есть

$$K_{v,n}^* = \{\varphi|\exists X J_{(v,n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1)\},$$

$$K_{\tau,n}^* = \{\varphi|\exists X J_{(\tau,n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1)\}.$$

Тогда $\tilde{K}_{v,n}^* = K_{v,n}^* \cap \tilde{K}$ – множество новых (ранее не встречавшихся) предсказанных действий с оценкой v , $v \in \{\pm 1, 0, \tau\}$. Множество предсказанных действий есть $K_n^* = K_{1,n}^* \cup K_{-1,n}^* \cup K_{0,n}^* \cup K_{\tau,n}^*$, а множество новых предсказанных действий – $\tilde{K}_n^* = K_n^* \cap \tilde{K}$.

Пусть n – номер шага стабилизации. Тогда $K_v^* = \{\varphi|\exists n \exists X J_{(v,n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (v \in \{\pm 1, 0\}) \& (n \geq 1)\}$ – множество предсказанных действий с оценкой v , $K_\tau^* = \{\varphi|\exists n \exists X (J_{(\tau,n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1))\}$. Соответственно, множество предсказанных действий $K^* = \{\varphi|\exists n \exists X (((J_{(v,n)}(X \Rightarrow_1[\varphi]) \& (v \in \{\pm 1, 0\})) \vee J_{(\tau,n)}(X \Rightarrow_1[\varphi]) \& (X \in Ag^*) \& (n \geq 1)))\}$, множество новых предсказанных действий $\tilde{K}^* = K^* \cap \tilde{K}$. Таким образом, после предсказания мы получим множество действий $K'^* = K^* \cup K' = \tilde{K}^* \cup K'$, причем $\tilde{K}^* \cap K' = \emptyset$, откуда видно, что $K^* \cup K' = K'$ ($K^* \subseteq K'$), только если $\tilde{K}^* = \emptyset$, т.е. не удастся доопределить действия ни для одного агента из Ag^* . В этом случае, разумеется, степень непротиворечивости не меняется.

Для новых предсказанных действий $\eta_N(\tilde{K}^*, K^+) = |\tilde{K}^* \cap K^+|/|\tilde{K}^*|$, для всех предсказанных действий $\eta_P(K^*, K^+) = |K^* \cap K^+|/|K^*|$. После выполнения цикла ДСМ-рассуждений степень непротиворечивости действий будет определяться выражением $\eta(K'^*, K^+) = |K'^* \cap K^+|/|K'^*|$. Сравним $\eta(K', K^+) = |K' \cap K^+|/|K'|$ и $\eta(K'^*, K^+) = |K'^* \cap K^+|/|K'^*| = |(\tilde{K}^* \cup K') \cap K^+|/|(\tilde{K}^* \cup K')|$. Можно показать, что степень $\eta(K'^*, K^+) \geq \eta(K', K^+)$, если для новых предсказанных действий доля непротиворечивых не меньше доли непротиворечивых действий в исходной БФ, $|\tilde{K}^* \cap K^+|/|\tilde{K}^* \cup K^+| \geq |K' \cap K^+|/|K' \cup K^+|$.

Предсказание рациональности возможных действий для случая, когда известны только описания

агентов, $\tilde{Ag} = \{X|\forall \varphi J_{(\tau,0)}(X \Rightarrow_1[\varphi]) \& (\varphi \in K)\}$, соответствует избранной процедуре предсказания действий, т.е. выполнению последовательно шагов 1). – 4)., описанных выше. Так, для 1)., когда рассматриваются только существующие непротиворечивые действия, степень рациональности (непротиворечивости) не меняется. Для 2). – предсказания непротиворечивых ранее не встречавшихся действий – она увеличивается. Соответственно, для случаев 3). и 4).

³ Отметим, что введение двух функций связано с вычислительной сложностью процедуры построения множеств K' и Δ , которая зависит от структуры множества Σ .

⁴ Здесь и далее при рассмотрении рациональности предсказанных действий будем считать, что для полученного стабилизированного множества гипотез выполнены аксиомы каузальной полноты.

Степень рациональности может только уменьшаться.

Проблема анализа и предсказания рациональности может быть рассмотрена для так называемых обобщенных агентов $B_1, \dots, B_s$, каждый из которых представляет собой коллектив агентов (МАС) с соответствующими множествами мнений $K'_1, \dots, K'_s$ и множествами зависимостей $\Sigma_1, \dots, \Sigma_s$. Возникает вопрос: действия каких обобщенных агентов могут быть согласованы друг с другом? Определим максимальные и минимальные n -арные ($2 \leq n \leq s$) сходства обобщенных агентов, соответственно, как

$\bigcup_{i \leq s} \Sigma_i \left(\bigcup_{j=1}^k \Sigma_{i_j}, i_j = 1, \dots, s, k = 2, \dots, s \right)$ при условии непротиворечивости таких объединений ($\text{consis}(\bigcup_{i \leq s} \Sigma_i)$), всего возможно $2^s - s$ таких

сходств) и $\bigcap_{i \leq s} \Sigma_i \left(\bigcap_{j=1}^k \Sigma_{i_j}, i_j = 1, \dots, s, k = 2, \dots, s \right)$, $\text{consis}(\bigcap_{i \leq s} \Sigma_i)$. Заметим, не все n -арные ($2 \leq n \leq s$) объединения непротиворечивы, т.е. в действительности максимальных сходств может быть меньше, чем $2^s - s$. При этом и минимальных сходств не больше $2^s - s$, так как пересечение непротиворечивых множеств также непротиворечиво, но среди пересечений могут встречаться пустые. Множества сходных обобщенных агентов

$B_r^k = \{B_{i_1}, \dots, B_{i_k}\}, k = 2, \dots, s, r = 1, \dots, C_s^k$, для максимальных и минимальных сходств могут быть рассмотрены по отдельности (заметим, что как для максимальных, так и для минимальных сходств некоторые $B_r^k = \emptyset$). Соответственно, для каждого такого множества определяются степени непротиворечивости η_r^k на основании $\Sigma_r^k = \Sigma_{i_1} \cup \dots \cup \Sigma_{i_k}$ для максимальных или $\Sigma_{i_1} \cap \dots \cap \Sigma_{i_k}$ для минимальных

сходств, соответствующих K_r^{+k} и $K_r'^k = K_{i_1}' \cup \dots \cup K_{i_k}'$, – что позволяет задать на них упорядочение. Отдельного рассмотрения требует вопрос, сохраняется ли порядок на новых степенях непротиворечивости η_r^{k*} при предсказании действий в

рамках каждого максимально или минимально сходного множества обобщенных агентов B_r^k .

Заметим здесь, что для анализа рациональности мнений в [12] предложены также критерии ρ – степень близости к «идеальному» мнению – и λ – степень согласованности (однородности) мнений одной социальной общности. Соответственно, рациональности нескольких социальных общностей соответствует частичный порядок на $\langle \eta, \rho, \lambda \rangle$, который может меняться в случае предсказания мнений и порождения новых $\langle \eta^*, \rho^*, \lambda^* \rangle$. Разумеется, нечто подобное может быть рассмотрено как для коллективов обобщенных агентов, так и для предсказанных действий таких агентов. Для этого необходимо определить понятие «идеального» действия агента (например, им может быть рациональное в стандартном смысле действие, позволяющее достичь цели оптимальным способом – в соответствии с тем или иным критерием оптимальности). Далее, согласованность действий агентов также требует специального определения в случае систем искусственных, а также обобщенных агентов. Более того, такое же упорядочение может быть осуществлено для каждого построенного пространства толерантности – множества (максимально или минимально) сходных коллективов агентов, а также для предсказанных значений всех трех критериев рациональности.

Заключение

Рассмотренные средства – ДСМ-метод, семантика логики аргументации и критерии рациональности действий – позволяют решать задачи анализа и предсказания действий агентов и рациональности (непротиворечивости) этих действий, а также выбирать решения, увеличивающие степень рациональности как одного агента, так и системы агентов. Решение этих задач опирается на предварительное структурированное описание агента. Дальнейшее развитие предложенных средств позволяет также приблизиться к решению задачи принятия решения множеством агентов, а также рассмотреть возможность согласованных действий коллективов агентов, обладающих единой непротиворечивой системой возможных действий.

Работа в этом направлении может рассматриваться как движение вдоль «когнитивной магистрали» «Интеллектуальные системы – Когнитивные системы – Интеллектуальные роботы». Из изложенного выше следует, что указанная «магистраль» может быть продлена до пункта «Коллективы роботов». Выделение этой магистрали есть следствие понимания фундаментальной цели Искусственного интеллекта как имитации и усиления способностей, характеризующих феноменологию естественного интеллекта [5]. При этом широта исследований в этой области заставляет обращаться к возможностям и аналогиям самых различных направлений – как в случае взаимодействия в области описания и изучения социальных систем и систем искусственных агентов.

Литература

1. Рассел С., Норвиг П. Искусственный интеллект. Современный подход. – Москва. Санкт-Петербург. Киев. Изд. 2-е. 2007. (1407 с.)
2. Wooldridge M.J. An Introduction to Multiagent Systems. – John Wiley&Sons, Chichester, United Kingdom, 2009 (2-nd edition).
3. Anshakov O.M., Gergely T. Cognitive reasoning. A formal approach. Springer, 2010.
4. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: УРСС, 2002.
5. Финн В.К. К структурной когнитологии: феноменология сознания с точки зрения искусственного интеллекта // Вопросы философии, № 1, 2009, с. 88 – 103.
6. Финн В.К., Михеенкова М.А. О логических средствах концептуализации анализа мнений // В сб.: Многозначные логики и их применения, т. 2: Логики в системах искусственного интеллекта (под ред. проф. В.К. Финна). М.: Издательство ЛКИ, 2008, с. 152–199.
7. Bench-Capon T.J.M., Dunne Paul E. Argumentation in Artificial Intelligence // Artificial Intelligence, V. 171, issue 10, 2007, pp. 619 – 641.
8. Девятко И.Ф. Социологические теории деятельности и практической рациональности. – М.: АВАНТИ ПЛЮС, 2003.
9. Финн В.К. Стандартные и нестандартные логики аргументации // В сб.: Многозначные логики и их применения, т. 2: Логики в системах искусственного интеллекта (под ред. проф. В.К. Финна). М.: Издательство ЛКИ, 2008, с. 59–91.
10. Финн В.К. Синтез познавательных процедур и проблема индукции // В кн. Автоматическое порождение гипотез в интеллектуальных системах. Под ред. проф. В.К. Финна. М.: Книжный дом «Либроком», 2009, с. 92 – 158.
11. Финн В.К. Индуктивные методы Д.С. Милля в системах искусственного интеллекта. Часть I. // Искусственный интеллект и принятие решений, 2010, наст №..
12. Михеенкова М.А., Финн В.К. Интеллектуальный анализ данных для проблем когнитивной социологии // XI Национальная конференция с международным участием “Искусственный интеллект-2008”, Дубна, Сентябрь 29 – Октябрь 2, 2008, Труды конференции в 3 томах, т.2, с. 61 – 69.
13. Михеенкова М.А., Финн В.К. О количественных характеристиках в качественном анализе социологических данных // Третья международная конференция “Системный анализ и информационные технологии” САИТ-2009, Звенигород, 14-18 сентября 2009 г., Труды конференции, с. 191 – 196.

Михеенкова Мария Анатольевна. Старший научный сотрудник Всероссийского института научной и технической информации РАН (ВИНИТИ РАН). Окончила Московский физико-технический институт (МФТИ) в 1981 году. Кандидат технических наук. Имеет 59 печатных работ. Область научных интересов: искусственный интеллект, логика и методология социальных наук, интеллектуальный анализ данных, качественный анализ. E-mail: mmikh@viniti.ru.

Финн Виктор Константинович. Заведующий сектором Всероссийского института научной и технической информации РАН (ВИНИТИ РАН). В 1957 году окончил Философский факультет, в 1966 году – Механико-математический факультет Московского государственного университета им. М.В. Ломоносова. Доктор технических наук, профессор. Заслуженный деятель науки РФ. Имеет 194 печатные работы. Область научных интересов: искусственный интеллект, математическая логика, интеллектуальные системы, машинное обучение, теория познания, логика и методология науки. E-mail: finn@viniti.ru