М. М. Виньков, И.Б. Фоминых

Рассуждения о знаниях и проблема логического всеведения. Часть 2. Введение темпорального параметра в метаязык[1]

Аннотация. В статье проведен обзор логических систем, формализующих рассуждения о знаниях, в которых реализован подход к решению проблемы логического всеведения, в основе которого лежит введение темпорального параметра в метаязык. Данный подход к решению проблемы логического всеведения, в отличие от модального подхода, рассмотренного в первой части статьи, не связан с какими-либо ограничениями дедуктивных способностей интеллектуального агента, поведение которого описывают логические системы, реализующие этот подход. Кроме того, такой агент в ходе своих рассуждений способен оценивать имеющийся у него временной ресурс, что особенно важно для интеллектуальных систем жесткого реального времени.

Ключевые слова: эпистемическая логика, проблема логического всеведения, темпоральный параметр, рассуждения с ограниченным ресурсом, временной ресурс.

Введение

Рассмотренные в первой части данной статьи [2] способы решения проблемы логического всеведения относятся к модальному подходу [1] и образуют два направления. Одно из них связано с искусственным ослаблением дедуктивных возможностей агента, другое имеет в своей основе введение в логический язык специальных модальных операторов, трактуемых как необходимость применения умственных усилий для получения какого либо знания (выраженного в виде логических формул).

Хотя в рамках этих направлений проблема логического всеведения преодолевается в разной степени всеми рассмотренными способами, однако ни один из них не предоставляет возможности для ответа на важный с практической точки зрения вопрос: сможет ли агент, потратив какое-то время на проведение рассуждений, не просто решить стоящую перед ним задачу, но уложится при этом в определенные, подчас весьма жесткие сроки? Это особенно важно для интеллектуальных систем, решающих задачи в т.н. жестком реальном времени, когда выход за отведенные временные рамки чреват серьезными, быть может, катастрофическими последствиями [6]. Заметим, что без возможности ответа на данный вопрос решение проблемы логического всеведения в значительной степени теряет свою практическую ценность.

Во второй части данной статьи будет изложен иной, не связанный с модальной логикой, подход к решению не только проблемы логического всеведения как таковой, но и вопросов, подобных указанному выше. Его суть состоит во введении темпорального параметра в метаязык, используемый для описания операционной семантики логических систем, которые находятся в рамках рассматриваемого подхода.

1. Шаговая логика

Рассмотренная в первой части данной статьи динамическая эпистемическая логика дает возможность рассуждать о процессе рассуждения агента в целом, *после* его фактического осуществления. Для того, чтобы дать возможность на-

[1] Работа выполнена при финансовой поддержке РФФИ (проекты № 11-07-00042, № 12-07-00677)

блюдать за процессом рассуждения агента *во время* его осуществления, была предложена логическая система, получившая название «шаговая логика» [8], ставшая исторически первой реализаций более общей концепции, получившей название «активная логика» [3, 10, 11]. Сохраняя возможность рассуждать об агентах, как бы «глядя на них со стороны» (как это имеет место, например, в случае динамической эпистемической логики), шаговая логика в то же время позволяет самому агенту соотносить процесс своего рассуждения с событиями, происходящими во внешней среде в результате его деятельности или помимо нее. Систему шаговой логики образуют пары вида $<SL_n, SL^n>$, где SL^n - метатеория поведения агента, соответствующая его «внутренней» (далее собственной) теории SL_n (n – параметр, характеризующий уровень сложности теорий, о котором будет сказано ниже). Как модель дедукции, активная логика характеризуется языком, множеством дедуктивных правил, а также множеством «наблюдений». Если предположить, что агент рассуждает, находясь в статической среде, множество наблюдений может рассматриваться как часть исходной базы знаний дедуктивной системы, т.е. как множество утверждений, поддерживающих дедуктивный процесс, в результате которого порождаются новые знания. Однако, использование т.н. *функции наблюдения* позволяет моделировать динамическую среду, информация о которой поступает к агенту по мере происходящих в этой среде изменений.

Рассуждение во времени характеризуется выполнением циклов дедукции, называемых *шагами*. Так как в основе шаговой логики лежит дискретная модель времени, то эти шаги играют роль временного эталона – время измеряется в шагах. Знания агента ассоциируются с индексом шага, на котором они были впервые получены.

Принципиальное отличие шаговой логики от других темпоральных эпистемических логик состоит в том, что «темпоральные аргументы введены в язык собственных теорий агентов» [8]. Таким образом, временной параметр связывается не только с каждым утверждением (формулой), которое эксплицитно знает агент, но и с дедуктивными правилами вывода. То, что узнал агент на шаге t (t-знания), используется для вывода новых знаний на шаге (t + 1). Дедуктивные правила вывода в активной логике имеют следующий вид:

$$\frac{t:\ \phi, \psi}{t+1:\ \phi \wedge \psi} \qquad \text{конъюнкция}$$

$$\frac{t:\ \phi \wedge \psi}{t+1:\ \phi} \qquad \text{отсоединение,}$$

$$\frac{t:\ \phi}{t+1:\ \phi} \qquad \text{наследование,}$$

$$\frac{t:\ \phi,\ \phi \rightarrow \psi}{t+1:\ \psi} \qquad \textit{modus ponens}\ ,$$

$$\frac{t:\ \varphi,\ sub\ (\ \lceil\phi\rceil, \lceil\varphi\rceil), [\phi]}{t+1:\ \neg K\ (t,\ \phi)} \qquad \text{самопознание,}$$

где ϕ - любая формула, не известная агенту i на шаге t, но являющаяся подформулой некоторой известной ему формулы φ, т.е. *осознаваемая* агентом, sub (.,.) – двухместный метапредикат, выражающий отношение «быть подформулой», $\lceil\phi\rceil$, $\lceil\varphi\rceil$ - константы, обозначающие имена формул ϕ и φ. $[\phi]$ – нотация, означающая, что формула ϕ отсутствует в текущих знаниях агента на шаге t. K(.,.) – двухместный метапредикат, используемый для выражения того факта, что агенту известна некоторая формула в некоторый момент времени.

$$\frac{t:\ \phi,\ \neg\phi}{t+1:\ contra\ (t,\ \lceil\phi\rceil,\ \lceil\neg\phi\rceil)} \quad \text{выявление «прямых» противоречий,}$$

где *contra(.,.)* – двухместный метапредикат, используемый для выражения того факта, что в текущих знаниях агента присутствуют формула ϕ и ее отрицание $\neg\phi$ в некоторый момент времени t.

Наблюдения могут осуществляться на любом шаге дедуктивного процесса. Результатом наблюдения является формула, выражающая некоторое утверждение и ассоциированная с соответствующим шагом. Для иллюстрации осуществляемого по шагам процесса рассуждений предположим, что агент изначально знает (на шаге t), что $\phi \rightarrow \psi$ и $\psi \rightarrow \chi$ и на шаге (t + 1) он наблюдает ϕ. В приведенной ниже таблице показано, какие *новые* знания появляются на каждом шаге при использовании агентом дедуктивных правил наследования и *modus ponens*.

$t: \quad \ldots \quad \phi \to \psi,\ \psi \to \chi \ldots$
$t+1: \quad \ldots \quad \phi \quad \ldots$
$t+2: \quad \ldots \quad \psi \quad \ldots$
$t+3: \quad \ldots \quad \chi \quad \ldots$

Различные уровни сложности теорий шаговой логики связаны с вовлечением в процесс рассуждения агентов трех различных механизмов:

- самопознания (способности агентов осознавать как то, что они на данный момент времени знают, так и то, чего они на данный момент времени *не* знают);
- отсчета времени;
- обнаружение и устранения противоречий в знаниях.

Простейшему уровню соответствует логика $<SL_0, SL^0>$, не содержащая ни одного из перечисленных механизмов, наиболее сложному – логика $<SL^7, SL_7>$, включающая все три механизма. Остальные логики перечислены ниже:

$<SL_1, SL^1>$: самопознание;

$<SL_2, SL^2>$: отсчет времени;

$<SL_3, SL^3>$: обнаружение и устранение противоречий;

$<SL_4, SL^4>$: самопознание и обнаружение и устранение противоречий;

$<SL_5, SL^5>$: самопознание и отсчет времени;

$<SL_6, SL^6>$: обнаружение и устранение противоречий и отсчет времени.

Формально активная логика определяется следующим образом. Пусть $\mathcal{L}$ –стандартный язык первого порядка, $\mathcal{W} \subseteq \mathcal{L}$ – множество формул, $\mathbb{N}$ – множество натуральных чисел. Далее, исходя из тематики данной статьи, основное внимание будет уделено механизмам самопознания и отсчета времени в ущерб механизму обнаружения и устранения противоречий, не имеющему к проблеме логического всеведения непосредственного отношения.

Определение 1. Функция наблюдения есть отображение OBS: $\mathbb{N} \Rightarrow \mathcal{P}(\mathcal{W})$, где $\mathcal{P}(\mathcal{W})$ – множество всех подмножеств $\mathcal{W}$ и для любого $t \in \mathbb{N}$ $OBS(t)$ есть конечное множество.

Определение 2. История есть конечный кортеж, компонентами которого являются пары подмножеств множества $\mathcal{W}$. $\mathcal{H}$ – есть множество историй.

Определение 3. Функция вывода есть отображение $INF : \mathcal{H} \Rightarrow \mathcal{P}(\mathcal{W})$, где для любого $h \in \mathcal{H}$ $INF(h)$ есть конечное множество.

Содержательно история представляет собой хронологическую последовательность пар множеств выведенных теорем/наблюдений агента, циклически порождавшихся (по шагам рассуждений) вплоть до некоторого момента времени. Функция вывода расширяет историю, являющуюся ее аргументом, реализуя новый шаг рассуждения и порождая множество новых формул (= теорем), выведенных агентом к моменту времени завершения соответствующего шага рассуждений. Следующие два определения формализуют эти соображения в терминах шаговой логики SL_n .

Определение 4. SL_n –теория для языка $\mathcal{L}$ есть тройка $<\mathcal{L}, OBS, INF>$, где $\mathcal{L}$ есть язык первого порядка, OBS – функция наблюдения и INF – функция вывода. Будем использовать нотацию SL_n (OBS, INF) для обозначения такой теории ($\mathcal{L}$ имплицируется определениями OBS и INF).

Определение 5. Пусть trm_0 – пустое множество. Для всех t > 0 множеством теорем t-го шага trm_i (t-теорем) назовем INF $(<<trm_0, OBS(0)>, <trm_1, OBS(1)>, \ldots, <trm_{t-1}, OBS(t-1)>>$. Мы пишем $SL_n(OBS, INF) \vdash_t \varphi$ имея в виду, что φ является t-теоремой теории $SL_n(OBS, INF)$.

Определение 6. Для данной теории SL_n (OBS, INF) соответствующей метатеорией (обозначаемой как $SL^n(OBS, INF)$) является теория первого порядка, включающая числа, имена формул из $SL_n(OBS, INF)$, двухместный метапредикатный символ K и такая, что

$$SL^n(OBS, INF) \vdash K(t, ¦\varphi¦) \text{ если и только если } SL_n(OBS, INF) \vdash_t \varphi .$$

Таким образом, в $SL^n(OBS, INF)$ формула $K(t, ¦\varphi¦)$ выражает тот факт, что φ является t-теоремой теории $SL_n(OBS, INF)$. Понятие полноты для мета-теории вводится следующим образом.

Определение 7. Говорят, что мета-теория SL^n аналитически полна, если для каждого целого положительного t и каждой константы $¦\varphi¦$, именующей некоторую формулу φ из соответствующей собственной теории агента $SL_n(OBS, INF)$, имеет место $SL^n(OBS, INF) \vdash K(t, ¦\varphi¦)$ или $SL^n(OBS, INF) \vdash \neg K(t, ¦\varphi¦)$.

Пусть $\mathcal{L}' \supseteq \mathcal{L}$ есть язык первого порядка, отличающийся от $\mathcal{L}$ тем, что в его алфавите содержатся метапредикатные символы – двухместный K и одноместный Now.

Определение 8. Шаговая интерпретация для $\mathcal{L}'$ есть последовательность M = <M_0, M_1, …, M_t, …>, где

1. Каждая M_t есть обычная интерпретация первого порядка для $\mathcal{L}'$;

2. $M_t \models Now(t)$.

Определение 9. Шаговой моделью для теории SL_n (OBS, INF) является шаговая интерпретация M = <M_0, M_1, …, M_t, …> , для любого t удовлетворяющая следующим условиям:

1. $M_t \models K(t, ¦\varphi¦)$ если и только если SL_n (OBS, INF) $\vdash_t \varphi$;

2. SL_n (OBS, INF) $\vdash_t \varphi$ если $M_t \models \varphi$.

Первое условие удостоверяет, что существует историческая запись об t-теореме, второе – что t-теорема на самом деле верна.

Определение 10. Формула φ t-истинна в шаговой модели M = <M_0, M_1, …, M_t, …> (пишется $M \models_t \varphi$) если $M_t \models \varphi$.

Определение 11. Функция наблюдения OBS конечна, если существует такое i, что для всех $j > i$ $OBS(j) = \varnothing$.

Теорема 2. Каждая t-теорема теории SL_n (OBS, INF) t-истинна в каждой шаговой модели M теории SL_n (OBS, INF).

2. Рассуждения, спланированные во времени и отвечающие семантике минимальных моделей

В данном разделе будет рассмотрена свободная от логического всеведения логика TRL (timed reasoning logic), введенная в [7], семантика которой схожа с семантикой минимальных моделей, хорошо известной в логическом программировании [9]. Общей особенностью подходов, воплощенных в системах TRL и активной логики является трактовка рассуждения как процесса. Этот процесс представляет собой упорядоченную во времени последовательность дедуктивных циклов, каждый из которых состоит в выполнении программы, представляющей собою множество правил вида Если <условие> то <действие>. Информация, полученная в результате наблюдения за внешней средой, также как и любые априорные знания запоминается в рабочей памяти. Правила сопоставляются с текущим содержимым рабочей памяти и некоторое их подмножество «срабатывает», что приводит к изменению содержимого рабочей памяти агента и/или к воздействию на внешнюю среду. В общем случае условия правил могут сопоставляться с содержимым рабочей памяти более чем одним способом. При этом порождаются различные *экземпляры* одного и того же правила. Следуя стандартной терминологии систем, основанных на правилах, множество экземпляров правил называется *конфликтным*, а процесс, в результате которого получается подмножество конфликтного множества, «срабатывающее» в ходе выполнения данного дедуктивного цикла, называется *разрешением конфликтов*. Агенты могут использовать различные стратегии для разрешения конфликтов (в системе TRL [7]). Однако, во всех случаях под мерой времени (эталоном) неявно подразумевается продолжительность дедуктивного цикла. Каждому выполнению дедуктивного цикла соответствует один «тик» виртуальных внутренних часов. Время рассматривается как бесконечная последовательность <0, 1, 2,… > из множества натуральных чисел N. В системе TRL, для обеспечения возможности рассуждать о конкретных агентах нам потребуется непустое множество $Ag = \{1, \dots i, \dots\}$, каждый элемент которого обозначает конкретного агента. Процесс рассуждения агента мыслится как упорядоченная во времени последовательность его локальных ментальных состояний. Каждое ментальное состояние агента $m^i{}_t$ индексируется элементом индексного множества $I = Ag \times \mathrm{N}$, которое является множеством пар вида (i,t), где i – агент, а t – момент времени завершения дедуктивного цикла, в результате которого было получено это состояние. Оно представляет собой конечное множество формул некоторого языка (пропозиционального, 1-го порядка, модального и т.д.) $\{\varphi_1, \dots \varphi_n\}$. Различные агенты могут иметь разный язык, который у них может меняться от одного ментального состояния к другому, благодаря приобретению новых имен понятий и т.д. Чтобы иметь возможность моделировать эти изменения, языки агентов также индексируются элементами индексного множества I: $\mathrm{L}^i{}_t$ - язык, на котором агент i «мыслил» при выполнении дедуктивного цикла, окончившегося в момент времени t.

Синтаксис языка агента $\mathrm{L}^i{}_t$ определяется обычным образом. Например, если $\mathrm{L}^i{}_0$ является языком пропозициональной логики с пропозициональными переменными $p_1, \dots, p_n$, то правильно построенная формула φ языка $L^i{}_0$ определится как

$\varphi = p_i \mid \neg\, \varphi \mid \varphi \rightarrow \varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi$

Если i – агент, t – момент времени и φ - формула языка $L^i{}_t$, то (i,t): φ есть *помеченная формула* TRL.

TRL позволяет моделировать различные стратегии разрешения конфликтов при применении агентами правил вывода. Ниже применительно к TRL будет рассмотрена стратегия, состоящая в том, что в рамках одного дедуктивного цикла применяются все экземпляры правил, попавшие в конфликтное множество. Эта же стратегия лежит и в основе всех вариантов упомянутой выше шаговой логики и ее обобщения активной логики. Синтаксис правил вывода TRL, отвечающих данной стратегии, задается следующим образом:

$$\frac{(i_1, t)\colon \varphi_1, \ldots, (i_n, t)\colon \varphi_n}{(i, (t+1))\colon\ \psi}\ .$$

Здесь t – универсально квантифицированная переменная, принимающая значения на множестве N, а $i_1,\ldots, i_n, i$ – не обязательно различные константы, соответствующие именам агентов.

Пусть R – множество правил вывода TRL. Помеченная формула (i, t): φ является *выводимой* из множества помеченных формул Γ с использованием R ($\Gamma \vdash_R (i,t)\colon \varphi$) если существует последовательность помеченных формул (i_1, t_1): $\varphi_1,\ldots, (i_n, t_n)$: φ_n такая, что:

1) каждая помеченная формула в последовательности либо принадлежит Γ, либо является непосредственным следствием Γ по одному из правил вывода из R;

2) последняя из помеченных формул в последовательности (i_n, t_n): φ_n совпадает с (i, t): φ.

Каждое локальное ментальное состояние агента формируется в результате применения некоторого множества правил к предыдущему состоянию, а также в результате его непосредственного наблюдения за внешней средой. Это моделируется при помощи множества функций вывода inf_i (по одной для каждого агента) и функции наблюдения *obs,* общей для всех.

Определение 12. *Пусть Ag есть множество агентов и* $\{ L^i{}_t : i\in Ag, t \in N \}$ *есть множество языков агентов.* TRL*-модель M это кортеж* $< N, obs, \{ inf_i : i\in Ag \}, \{ m^i{}_t : i\in Ag, t\in N \}>$, *где obj есть функция, отображающая каждую пару* (i,t) *в конечное множество формул из* $L^i{}_t$, inf_i *есть функция, отображающая конечные множества формул из* $L^i{}_t$ *в конечные множества формул из* $L^i{}_{t+1}$ *, и каждое* $m^i{}_t$ *есть конечное множество формул из* $L^i{}_t$, *такое, что* $m^i{}_{t+1} = inf_i\,(m^i{}_t) \cup obs\,(i, (t+1))$.

Любое множество правил вывода в TRL естественным образом можно разбить на два подмножества. Первое подмножество называется *внутренними правилами*. Это правила, в которых упоминается ровно один агент. Например, таким правилом является правило, относящееся ко всем агентам, использующим modus ponens:

$$\frac{(i, t)\colon \varphi,\ (i, t)\colon \varphi \rightarrow \psi}{(i, (t+1))\colon\ \psi}\ .$$

Оно означает, что если в любой момент времени t агент i верит, что верны формулы φ и $\varphi \rightarrow \psi$, то в следующий момент времени он будет верить, что верна формула ψ. Внутренние правила соответствуют функциям inf_i . Множество таких правил будем обозначать R_{inf}.

Второе подмножество множества правил вывода TRL назовем правилами коммуникации. Они имеют форму

$$\frac{(i, t)\colon \varphi}{(j, (t+1))\colon\ \psi}$$

и соответствуют функции *obs*. Множество таких правил будем обозначать R_{obs} .

Мы говорим, что модель *M конформна* множеству правил R если

1) для каждого правила из множества R_{inf} вида

$$\frac{(i, t)\colon \varphi_1, \ldots, (i, t)\colon \varphi_n}{(i, (t+1))\colon\ \psi}\ .$$

inf_i в M удовлетворяет свойству: из $\varphi_1,\ldots, \varphi_n \in m^i{}_t$ следует $\psi \in inf_i\,(m^i{}_t)$;

2) для каждого правила из множества R_{obs} вида

$$\frac{(i, t)\colon \varphi}{(j, (t+1))\colon\ \psi}$$

obs в M удовлетворяет свойству: из $\varphi \in m^i{}_t$ следует $\psi \in obs(j, (t+1))$.

Определение 13. TRL*-модель M, конформная множеству правил вывода* TRL *R, является* *минимальной моделью для множества помеченных формул* Γ *если для любых i, t и* φ, $\varphi \in m^i{}_t$ *если и только если*

1) *существует правило в* R_{inf} *вида*

$$\frac{(i, t)\colon \varphi_1, \ldots, (i, t)\colon \varphi_n}{(i, (t+1))\colon \varphi}$$

и $\varphi_1, \ldots, \varphi_n \in m^i{}_{t\text{-}1}$;

2) *или* $\varphi \in obs(i, t)$ *и* $(i, t)\colon \varphi \in \Gamma$ *или существует правило в* R_{obs} *вида*

$$\frac{(j, t)\colon \psi}{(i, (t+1))\colon \varphi}$$

и $\psi \in m^i{}_{t\text{-}1}$.

В работе [6] доказываются теоремы о полноте и семантической непротиворечивости логики TRL, а также ее алгоритмической разрешимости.

Теорема 3. *Для любого множества правил вывода* TRL *R, любого конечного множества помеченных формул* Γ *и любой помеченной формулы* φ, $\Gamma \vdash_R \varphi$ *если и только если* $\Gamma \vDash \mathcal{R}\varphi$, *где* $\mathcal{R}$ *– множество всех* TRL*-моделей, конформных R.*

Теорема 4. *Для любого множества правил вывода* TRL *R, любого конечного множества помеченных формул* Γ *и любой помеченной формулы* φ *разрешимы вопросы, имеет ли место* $\Gamma \vdash_R \varphi$ *или* $\Gamma \vDash \mathcal{R}\varphi$, *где* $\mathcal{R}$ *– множество всех* TRL*-моделей, конформных R.*

3. Время как внешняя сущность

Общим недостатком и активной логики, и TRL является трактовка времени, в некотором смысле, как внутренней сущности этих систем, течение которого определяется структурой правил вывода, используемых для получения новых формул из уже имеющихся. Однако такая его трактовка, как будет показано ниже, предполагает допущение, по своей нереалистичности сравнимое с допущением, составляющим проблему логического всеведения. Во всех случаях под мерой времени (эталоном) неявно подразумевается продолжительность дедуктивного цикла. Каждому выполнению дедуктивного цикла соответствует один «тик» виртуальных внутренних часов. При этом (также неявно) делается допущение, что длительность выполнения не меняется от цикла к циклу или, что изменения столь малы, что ими можно пренебречь. В реальности на длительность выполнения дедуктивного цикла оказывают влияние изменения, происходящие в составе и структуре знаний агента вследствие проводимых им рассуждений и наблюдений за внешней средой. Кроме того, на длительность дедуктивных циклов могут влиять случайные факторы, такие, как сбои в электропитании, в работе других технических систем и т.д. Также и «мыслительные способности», в данном случае, длительность одних и тех же вычислительных циклов у разных агентов могут быть различными. По существу, допущение о неизменной длительности дедуктивных циклов сродни логическому всеведению и, так же как и последнее, оно нередко вступает в противоречие с реальной действительностью.

В работе [4] предлагается на примере модифицированного варианта логики рассуждения, спланированного во времени, названного TRL*, подход, в котором время трактуется как внешняя сущность, не связанная со скоростью выполнения дедуктивных циклов. Такая модификация позволяет отказаться от нереалистичного допущения, связанного с внутренним временем и может быть выполнена также и для систем активной логики. В предлагаемой логической системе, как и в других логических системах из данного класса, время рассматривается как бесконечная последовательность натуральных чисел из множества N. Будем обозначать ее *Gck* (глобальные часы). Однако в данном случае учитывается, что основное назначение такого рода логических систем состоит в моделировании поведения многоагентной системы в различных условиях (= прогонах). Поэтому каждому такому прогону ставятся в соответствие т.н. часы прогона модели *Ck*, отражающие его специфику (принцип грануляции времени). Часы прогона модели – это конечная или бесконечная строго возрастающая подпоследовательность глобальных часов, члены которой интерпретируются как моменты времени (на глобальных часах) завершения дедуктивных циклов, например <3, 5, 7, 10,... > . Множество всех таких моментов времени будем обозначать *Ck**. Каждый «тик» часов прогона модели, как и «тик» рассмотренных выше виртуальных внутренних часов, соответствует одному выполнению конкретного дедуктивного цикла. При этом, порядковый номер этого цикла совпадает не с моментом времени его завершения (как это имеет место в активной логике или TRL), а только с порядковым номером этого момента времени на часах прогона модели. Данное обстоятельство дает возможность, меняя часы прогона модели, имитировать различные условия работы многоагентной системы и лучше отражать, например, такие особенности, как увеличение длительности дедуктивных циклов агента по мере увеличения количества известной ему информации. Кроме того, различным агентам можно назначать различные локальные часы, моделируя таким образом, например, их различную «сообразительность» (быстродействие) или то, что они вводятся в действие в различные моменты времени. В дальнейшем, однако, для простоты рассматривается случай, когда всем агентам назначены одни и те же часы прогона модели. Нам понадобятся также две функции - *clock* (.) и *rank* (.). Первая отображает множество N во множество *Ck**. Терм *clock* (*n*) интерпретируется как момент времени, имеющий порядковый номер *n* на часах прогона модели. Вторая функция является обратной по отношению к первой, т.е. для всех $t \in Ck^*$ $rank\ (t) = clock^{-1}\ (t)$, ее значением является порядковый номер момента времени *t* на часах прогона модели *Ck*. Тогда, если время завершения предыдущего дедуктивного цикла было t, следующий дедуктивный цикл завершится в момент времени *clock* (*rank*(*t*) +1). Теперь каждое ментальное состояние агента $m^i{}_t$ индексируется элементом индексного множества $I = Ag \times Ck^*$, которое является множеством пар вида (*i*,*t*), где *i* – агент, а *t* – момент времени завершения дедуктивного цикла с номером *rank* (*t*), в результате которого было получено это состояние. Определение TRL*-модели будет выглядеть так:

Определение 14. *Пусть Ag есть множество агентов, Ck есть часы прогона модели и* $\{ L^i{}_t : i\in Ag, t \in Ck^*\}$ *есть множество языков агентов.* TRL*-*модель M* это кортеж* $< Ck$, obs, $\{ inf_i : i\in Ag \}$, $\{ m^i{}_t : i\in Ag, t \in Ck^*\}>$, *где obj есть функция, отображающая каждую пару* (*i*,*t*) *в конечное множество формул из* $L^i{}_t$, inf_i *есть функция, отображающая конечные множества формул из* $L^i{}_t$ *в конечные множества формул из* $L^i{}_{\mathrm{clock(rank(t)+1)}}$, *и каждое* $m^i{}_t$ *есть конечное множество формул из* $L^i{}_t$, *такое, что* $m^i{}_{\mathrm{clock(rank(t)+1)}} = inf_i\,(m^i{}_t) \cup obs\,(i, (clock\,(rank(t) +1)))$.

Соответственно, правила вывода в TRL* примут следующий вид:

$$\frac{(i_1, t)\colon \varphi_1,\ldots, (i_n, t)\colon \varphi_n}{(i, (clock\ (rank\ (t) +1)))\colon\ \psi},$$

И далее во всех введенных выше понятиях для TRL, в TRL* следует (*t* +1) заменить на (*clock* (*rank*(*t*) +1)). При этом нетрудно доказать, что теоремы 3 и 4 сохраняют свою силу.

Возможность применения рассмотренной выше концепции часов прогона модели не зависит от используемой стратегии разрешения конфликтов, поэтому данная концепция может быть использована и для аналогичной модификации других известных вариантов TRL точно так же, как и для любого известного варианта Активной Логики. Результатом во всех случаях будет трактовка времени, совершенно отличная от той, которая сейчас имеет место во всех логических системах данного класса. Время перестает быть их внутренней сущностью. Течение времени не связывается со структурой множества правил вывода. Задача моделирования, о которой идет речь в данной статье, получение логических результатов, интерпретируемых как: агент *i* способен вывести формулу φ, не выходя за временную границу *t,* не может быть сведена к определению глубины вывода данной формулы. Это дает возможность более адекватно моделировать поведение многоагентных систем логическими средствами.

Заключение

Рассмотренный выше подход к решению проблемы логического всеведения обеспечивает его в полном объеме и не связан с ограничениями дедуктивных способностей интеллектуального агента, чье поведение этот подход позволяет моделировать. При этом такой интеллектуальный агент способен оценивать на любом шаге своих рассуждений величину оставшегося в его распоряжении временного ресурса и менять свое поведение в зависимости от результатов этой оценки. Данное обстоятельство особенно ценно для обеспечения устойчивости к непредвиденным ситуациям интеллектуальных многоагентных систем жесткого реального времени, как было показано в работе [5]. Вне рамок данного обзора осталась задача построения паранепротиворечивой декларативной семантики для систем активной логики, которая в полном объеме до сих пор не решена, однако эта проблема не связана напрямую с проблемой логического всеведения, являющейся предметом данной статьи.

Литература

1. В.Н. Вагин, Е.Ю. Головина, А.А. Загорянская, М.В. Фомина. Достоверный и правдоподобный вывод в интеллектуальных системах. Монография. Под ред. В.Н. Вагина и Д.А. Поспелова. М.: Физматлит, 2004, 704 с.
2. М.М. Виньков, И.Б. Фоминых. Рассуждения о знаниях и проблема логического всеведения. Часть I: Модальный подход. // Искусственный интеллект и принятие решений. М.: УРСС, №4, 2011, - С. 3-13.
3. Виньков М.М. Активная логика с точки зрения фундированной семантики логических программ с приоритетами. // Труды IX-й национальной конференции по искусственному интеллекту с международным участием конференции КИИ’2004. Т1, М.: Физматлит, 2004.
4. Виньков М.М. Время, как внешняя сущность при моделировании рассуждений рационального агента с ограниченными ресурсами //Труды XI-й национальной конференции по искусственному интеллекту с международным участием КИИ-2008. — М.: Физматлит, 2008.
5. Виньков М.М., Фоминых И.Б. Повышение устойчивости к аномалиям интеллектуального агента с ограниченным временным ресурсом: метакогнитивный подход. //Труды XII-й национальной конференции по искусственному интеллекту с международным участием КИИ-2010. — М.: Физматлит, 2010.
6. Емельянов В.В. Многоагентная модель децентрализованного управления потоком производственных ресурсов// Труды Международной конференции <Интеллектуальное управление: новые интеллектуальные технологии в задачах управления> (ICIT'99, Переславль-Залесский, 6-9 декабря, 1999). - М.: Наука. Физматлит, 1999. - С. 121-126.
7. Alechina N., Logan B., and Whitsey M. A complete and decidable logic for resource-bounded agents. In Proc. Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2004),
8. Elgot-Drapkin J. Step Logic: Reasoning situated in time. PhD thesis. Department of computer science, University of Maryland, Colledge-Park, Maryland, 1988.
9. Konolige K. A Deduction Model of Belief. Morgan Kaufman, San Francisco, Calif., 1986.
10. Perlis D., Purang K., Purushothaman D., Andersen C., Traum D. Modeling time and meta-reasoning in dialog via active logic //Working Notes of AAAI Fall Symposium on Psychological Models of Communication. — 2005.
11. Purang K., Systems that detect and repair their own mistakes. PhD thesis, Department of Computer Science, University of Maryland, College Park, Maryland, 2001.

Виньков Михаил Михайлович. Заведующий отделом Российского НИИ информационных технологий и систем автоматизированного проектирования. Окончил МАИ в 1977 году. Кандидат технических наук, доцент. Автор 47 научных работ. Область научных интересов: динамические экспертные системы, индуктивное обучение, неклассические логики, формализация рассуждений. E-mail:vinkovmm@mail.ru

Фоминых Игорь Борисович. Заместитель директора по научной работе Российского НИИ информационных технологий и систем автоматизированного проектирования. Окончил МГТУ им. Н.Э.Баумана в 1963 году и МГУ им. М.В.Ломоносова в 1969 году. Д.т.н., профессор. Автор 110 печатных работ, в том числе шести монографий. Область научных интересов: динамические интеллектуальные системы, нейроинформатика, неклассические логики, информационные проблемы мозга. E-mail: igborfomin@mail.ru