

УДК 161/162 + 165 + 17

**Е. Н. Лисанюк
А. М. Павлова****ЛОГИЧЕСКИЕ АСПЕКТЫ МНОГООБРАЗИЯ АГЕНТОВ***

В статье предложена комплексная логическая классификация типов (презумпций) когнитивных агентов применительно к формальным теориям, моделирующим рациональное поведение, включая рассуждения. В классификации уточнены и систематизированы подходы к пониманию сути многообразия агентов, явным и неявным образом содержащиеся в логических формализмах, предназначенных для анализа знаний и убеждений агентов, их интеллектуальных способностей, а также особенностей рационального обоснования ими своих действий. Данная классификация может служить концептуальным основанием для решения проблемы многообразия агентов в логических теориях. Классификация типов когнитивных агентов включает три общих класса логических параметров — компетенций, по которым агенты могут различаться: эпистемические, целевые, или акционально-деятельностные, и дедуктивные. В статье мы отстаиваем положение о том, что эпистемические и целевые параметры выделения разных (презумпций) агентов выступают определенного рода надстройками над тем, каким образом трактуются дедуктивные параметры поведения рациональных агентов.

К л ю ч е в ы е с л о в а: логические компетенции, когнитивные презумпции, многообразие агентов, практические рассуждения, эпистемическая логика.

1. Введение

В чем состоит логическая проблема многообразия агентов. Логическая проблема многообразия агентов возникает при создании формализмов, претендующих на то, чтобы служить адекватными моделями действий, рассуждений и поведения рациональных агентов, и включает в себя корпус задач по уточнению и систематизации свойств, особенностей, профилей и тому подобных индивидов, групп индивидов, программных продуктов — субъектов анализируемых рассуждений, действий и поведения. Особенно остро эта проблема стоит, хотя и не всегда осознается, в двух ответвлениях современной философской, или модальной, логики — в эпистемической логике и в разнообразных логиках действий с модальными операторами, а также в других теориях, использующих эпистемические и акциональные операторы. Суть логической проблемы многообразия агентов проистекает из двух особенностей предмета их исследования — рационального поведения. То, насколько предлагаемые теории способны учесть и выразить каждую из этих особенностей, выступает ключевым фактором адекватности конструируемых в них формализмов. Первая особенность требует проводить границу между знаниями, убеждениями и действиями разных агентов, а вторая признает, что различия между агентами в аспекте их интеллектуальных способностей носят

* Статья написана при поддержке гранта РГНФ № 15-23-01002а.

необходимый характер¹. Таким образом, для решения логической проблемы многообразия агентов применительно к некоторой теории необходимо предусмотреть в ней две вещи: определить границу между агентами и описать корпус агентных способностей, или компетенций, наделив ими агентов в равной или неравной мере. Сделать и то и другое можно по-разному в зависимости от конкретных исследовательских задач, стоящих перед создаваемой формальной теорией.

Почему в логике возникает проблема многообразия агентов. Подобной проблемы в логических теориях не существовало до тех пор, пока по умолчанию подразумевалось, что правильность умозаключений и логические особенности их построения носят универсальный характер, и поэтому считалось, что в дедуктивном умозаключении всякий рациональный агент получит истинное заключение с необходимостью, если к истинным посылкам применит корректное правило вывода. Аналогичным образом обстоит дело в правдоподобных рассуждениях, несмотря на то, что у разных агентов заключения могут оказаться неодинаковыми [12, 370], тем не менее вероятность получения истинного заключения из идентичных посылок на основе идентичного правила вывода будет одинаковой в силу универсального алгоритма ее исчисления [13]. Поскольку ранее в логических теориях, как дедуктивных, так и недедуктивных, правильность умозаключения трактовалась как вообще не зависящая от того, какой именно рациональный агент берется его строить, при условии, что всякий придерживается необходимых правил, постольку от субъекта умозаключения можно было абстрагироваться.

Иначе обстоит дело, когда желают уточнить, каким образом можно построить правильное умозаключение в условиях неполной или не вполне достоверной информации, или избрать линию поведения на ее основе. Подобные ограничения могут носить внутренний или внешний характер. Внешние ограничения подразумевают, что то, как агент строит свои умозаключения, зависит от ситуации, в которой он находится, а внутренние ограничения предполагают, что на это влияют интеллектуальные особенности самого агента. Примером внешних ограничений могут служить законы морали и права, учитываемые или, наоборот, (сознательно) игнорируемые агентом при избрании линии поведения [7], а также недостоверность или неполнота информации о ситуации, в которой агент производит свои умозаключения [18, 20, 23, 25], или последствиях своих решений [3]².

Внутренние ограничения лежат в основе понятия ограниченной рациональности — так принято называть познавательную предпосылку, или презумпцию, определяющую интеллектуальную компетентность агентов в хранении информации

¹ Мы используем понятие интеллектуального агента, сформулированное в компьютерных науках и в особенности в программах искусственного интеллекта [32], откуда оно было отчасти заимствовано для логических исследований. Агент — это специальная абстракция для обозначения акторов определенного рода, выражающая субъектный источник рационального действия и поведения в функциональном смысле. Под агентами понимаются обладающие заданными характеристиками индивиды, группы индивидов, а также институты и программные продукты, наделенные способностями рассуждать, принимать решения и действовать.

² К внешним ограничениям относится и ограниченность доступа к информации, например, если агент вынужден действовать в ситуации жульничества со стороны других, сообщающих ему заведомо недостоверную информацию. Этот тип внешних ограничений мы частично учитываем в нашей классификации, но он в ней играет второстепенную роль.

(памяти), ее обновлении и пополнении, в способности строить выводы, производить вычисления и т. п. Термин «(когнитивная) презумпция» характеризует такую предпосылку обобщенно как «в первом лице», когда речь идет о представлениях агентов о самих себе, так и «в третьем лице», когда имеют в виду интеллектуальные компетенции других агентов [6]. Далее там, где это не рискует повлечь путаницы, мы будем подобные предпосылки называть когнитивными презумпциями рациональных агентов, имея в виду, что это название объединяет в одну группу представления агента о самом себе и о других агентах. В некоторых теориях эти представления необходимо разграничивать, например, когда речь идет об анализе исполнения норм различными агентами [5].

Сочетание внешних и внутренних ограничений позволяет по-разному конструировать различные типы агентов. Первые попытки ввести логические разграничения между агентами, устанавливающие зависимость между агентом и тем, как строятся умозаключения, появляются в середине XX в. [15]. В качестве этих разграничений могут быть поняты характеристики знаний и убеждений агентов, вытекающие из предложенных полноценных систем эпистемической логики, а также собственно логические аспекты различий между знаниями и убеждениями. Выяснилось, что содержательно многообразие агентов можно рассматривать двояко, как эпистемическое или эпистемологическое [33], хотя в логическом смысле такая граница между ними условна.

Эпистемическое многообразие агентов носит преимущественно семантический характер, оно указывает на содержательные различия в знаниях и убеждениях агентов и подразумевает, что агенты могут обладать нетождественными представлениями об истинности предложений. Например, агент α знает, что формула A истинная, потому что A — это теорема некоторой системы эпистемической логики, однако агент β может и не знать этого, если доказательство A не было получено или в другой системе A не является теоремой, либо β может считать A ложной, если существует система, в которой теоремой является $\neg A$.

Для того чтобы провести границу между разными в эпистемическом смысле знаниями агентов, в дальнейшем было предложено отличать виды группового знания агентов [29]. Эпистемологическое многообразие агентов говорит о том, что используемые ими способы получения и верификации знаний могут тоже различаться, если мы готовы допустить, что «существуют такие α , p и q , что α знает, что p , p логически влечет q (то есть $p \rightarrow q$ логически истинно), но α не знает, что q » [15, 231]. В логическом смысле эпистемологическое многообразие есть обобщение многообразия эпистемического. Эпистемологическое многообразие обычно трактуют в синтаксическом ключе, например, с помощью соответствующих аксиоматических базисов формальных систем, что равнозначно допущению о том, что агенты могут полагаться на разные определения дедуктивной выводимости [8, 288]. Эпистемологическая трактовка многообразия ограничивает абсолютный до этого характер эпистемического многообразия, когда от идеи о том, что «каждый эпистемический возможный мир логически возможен», отказываются в пользу идеи о существовании возможных миров, которые возможны эпистемически, но не являются при этом возможными логически. Такой подход позволяет миры,

возможные эпистемически, но не логически, отождествить с разными моделями того, каким образом познающий агент отличает эти миры один от другого на различных уровнях их анализа [15, 241].

В дальнейшем уточнение структуры интеллектуального агента и включение в нее наряду со знаниями и убеждениями также и интенций агента к совершению определенных действий, ценностей, желаний и т. п. открыло перспективы построения более богатых в выразительном смысле моделей поведения агентов на основе мультимодельных логик действий [28]. В теории игр и ее разнообразных практических приложениях выделение типов агентов сообразно их уточненным интеллектуальным компетенциям позволило получить важные результаты в области моделирования социального поведения, включая поведение экономическое [17], речекоммуникативное [31], а также принятие решений в условиях ограниченного доступа к информации [3].

Актуальность исследования, его цели и задачи. Вместе с тем, несмотря на многообразие предложенных логических систем, моделирующих поведение, рассуждения и взаимодействие агентов, имеется по меньшей мере два основания утверждать, что создание комплексной классификации типов агентов является актуальной задачей современной логики и смежных с ней областей исследований. Первое и, пожалуй, не столь серьезное для этого утверждения основание состоит в том, что во многих логических системах особенности изучаемых агентов принимаются по умолчанию и неявно, без каких-либо уточнений и обсуждений. Например, в большинстве популярных задач динамической эпистемической логики, таких как «чумазные дети», «профессор и бармен», «день рождения Шерил» [29], подразумевается, что действующие и рассуждающие агенты презумптивно одинаковы в смысле интеллектуальных способностей и доступа к обновлению собственных сведений. Однако известная древнеиндийская притча о трех слепых мудрецах, рассуждающих о слоне, подсказывает, что это упрощение ситуации, требующее подчас отдельного рассмотрения. Иными словами, многие стандартные формальные модели нормальных логических теорий не учитывают ограниченный характер способностей и возможностей агентов в реальной ситуации.

Второе основание более серьезное, и именно к его преодолению мы стремимся в нашем исследовании. Суть его заключается в том, что логическая проблема многообразия агентов открывает своеобразный «ящик Пандоры»: на место универсальных рациональных агентов, которыми можно пренебречь в силу их универсальности, приходит замечательное разнообразие, рискующее впасть в дефляционизм идеалов рациональности [10, 26; 2, 126]. Даже если допустить, что нам удалось провести границу между контингентными и необходимыми аспектами многообразия агентов, тем не менее неотъемлемым этапом логического моделирования рационального поведения становится то, что требуется всякий раз принимать решение о том, какие характеристики агентов следует брать во внимание при построении формализмов, а какими можно пренебречь. Свой вклад в решение этой проблемы мы видим в двух плоскостях: во-первых, мы формулируем базовый подход для конструирования локальных ответов на этот

серьезный вызов формальным теориям, и, во-вторых, мы привлекаем внимание академического сообщества к этой проблеме.

В настоящей работе мы предложим комплексную классификацию типов агентов с ограниченной рациональностью. Это означает, что мы сфокусируемся на внутренних ограничениях интеллектуальных свойств, или компетенций, агентов и почти не будем касаться внешних ограничений. Мы также изучим представления об агентах, явным или неявным образом содержащиеся в некоторых логических формализмах, предложенных для анализа рационального поведения, и сделаем это при помощи трех специальных параметров, характеризующих интеллектуальные способности агентов. Назовем их условно эпистемическим, целевым и дедуктивным. Под эпистемическим параметром будем понимать эпистемический аспект многообразия агентов, для его эпистемологического аспекта во избежание путаницы используем название «дедуктивный», а целевой параметр будет указывать на отличные от знаний и убеждений акциональные аспекты агентов — интенции, желания и т. п. Поскольку отнести ту или иную интеллектуальную компетенцию агента к одному из этих трех параметров можно весьма условно, постольку мы, с одной стороны, будем пояснять эти вопросы, когда это важно. С другой стороны, хотя в предложенной классификации границы между тремя группами агентных компетенций не носят жесткого характера, сами обсуждаемые компетенции в целом и каждая из них в отдельности являются необходимыми для решения проблемы логического многообразия.

Для создания комплексной классификации нам потребуется различать два подхода к созданию логических моделей рационального поведения агентов — моноагентный и мультиагентный. К первому относятся теории, рассматривающие действия или рассуждения одного или более агентов, которые считаются идентичными, т. е. наделяются одинаковыми интеллектуальными способностями, обуславливающими необходимый характер знаний или действий агентов, а ко второму — теории, где исследуют взаимодействия двух и более агентов и подразумевается, что они неидентичны, т. е. наделены неодинаковыми способностями, как в [25]³.

В соответствии с этой целью работа имеет следующую структуру: в первом параграфе кратко изложена суть проблемы многообразия агентов, во втором мы обсудим, какие именно интеллектуальные компетенции агентов выражает каждый из трех параметров, в третьем параграфе мы рассматриваем дедуктивный параметр выделения различных типов агентов на примере диалоговой логики П. Лоренцена и К. Лоренца и сопоставляем его с эпистемическим — на примере классификации Р. Смаллиана. Подробное исследование целевых презумпций мы оставляем для другого исследования.

³ Отметим, что (не)идентичность агентов в смысле их интеллектуальных компетенций носит необходимый характер в силу того, что указывает не столько на эпистемический, сколько и прежде всего на эпистемологический аспект когнитивного многообразия.

2. Три группы интеллектуальных компетенций агентов

Рассмотрим подробнее три группы интеллектуальных компетенций агентов, выделенных выше.

I. Эпистемические. Под данным видом когнитивных презумпций мы будем понимать следующие разновидности:

а) собственно эпистемические, представляющие знания агентов о положении дел и о других агентах и их когнитивных презумпциях;

б) докстатические, т. е. убеждения агентов⁴;

в) способности агента, влияющие на способы получения, хранения и изменения информации, в частности знаний и убеждений.

Заметим, что разграничения агентов, накладываемые особенностями «а» и «б», представляют собой конфигурацию конкретной модели, на которой определены знания и убеждения агентов, тогда как подпункт «в» связан со свойствами фреймов.

II. Целевые, связанные с практическими рассуждениями. Данный вид когнитивных презумпций содержит цели, интенции (намерения) и желания агентов во взаимодействии и моделируется в рамках ВДИ-логики. Обстоятельное рассмотрение целевых компетенций мы отложим до дальнейшего исследования.

III. Дедуктивные. Под данным видом презумпций понимаются различные немодальные логические системы, которые лежат в основании рассуждений (вычислений) агентов. В зависимости от того какую систему тот или иной агент использует в качестве базовой, будут соответственно различаться и понимания им истины и логического следования. Здесь имеется в виду, что агенты не только могут иметь разные способности, знания и желания, но они также могут использовать разные трактовки логической истины и следования. В качестве примера можно предположить, что один агент может полагаться на интуиционистское понимание истинности и следования, а другой — на классическое. Подобные презумпции могут влиять на рассуждения агентов, приводя к тому, что одинаковое изменение баз знаний двух разных агентов (например, в результате публичного анонсирования предложения ϕ) может привести к различным результатам благодаря тому, что агенты используют различные логические правила и разные формальные модели для получения следствий [11].

Почему мы выделяем именно данные группы презумпций агентов? С формальной точки зрения эти группы презумпций можно объединять по-разному. Например, в первой группе презумпций рассматриваются свойства агентов, связанные с эпистемическими и докстатическими модальностями и — неявным образом — также и с модальностями, выражающими действия, поскольку речь идет об изменениях в знаниях и убеждениях агентов. Все три вида модальностей

⁴ Под знаниями мы понимаем относительно постоянный и непротиворечивый корпус известных агенту сведений. Важным свойством знания в стандартной эпистемической логике является то, что оно не может быть ложным. Убеждения — это подверженные пересмотру правдоподобные сведения, в истинность которых агент верит. Убеждение агента — то же самое, что ранее в эпистемической логике называли мнением [7], иногда их также называют верой агента в истинность тех или иных положений [4].

широко представлены в динамических эпистемических логиках, изучающих знания, убеждения и процедуры их модификации агентами.

Вторая группа презумпций также связана с модальными свойствами, но здесь уже речь идет только о модальностях цели, интенции (намерения) и им подобных, которые непосредственно влияют на действия агентов. Таким образом, вторая группа презумпций относится исключительно к поведению, практическим рассуждениям агентов и принципам их осуществления, которые у разных агентов могут не совпадать, так как желания и интенции по-разному регулируют как сами рассуждения, так и действия агента. Несмотря на то что эти две группы объединяет модальный характер изучающих их формализмов, виды используемых в них модальностей различны, так как предназначены для моделирования разных интеллектуальных компетенций и обладают неодинаковыми формальными свойствами.

Отличительной чертой третьей группы презумпций является то, что она охватывает формы и правила, в соответствии с которыми агент устанавливает необходимые связи между своими знаниями, убеждениями, желаниями и интенциями и строит умозаключения, где они выступают в качестве структурных элементов. С точки зрения конструирования формальных систем все три группы презумпций носят логический характер, но при этом дедуктивные презумпции характеризуют логику, используемую в качестве базовой в системе, к которой затем добавляется ее модальная часть, например эпистемическая или акциональная. Если не совпадают базовые немодальные части логических систем, применяемых агентами для построения своих умозаключений, то и общие формальные постулаты поиска выводов будут неодинаковыми. Это означает, что могут оказаться неидентичными и множества полученных таким путем предложений, выражающих знания или убеждения агентов, даже если исходные посылки были идентичны.

Рассмотрим более детально некоторые из вышеперечисленных агентных компетенций и для начала остановимся на подгруппе эпистемических презумпций, характеризующей познавательные способности, касающиеся приобретения знаний и убеждений. Даже если допустить, что данные способности носят универсальный характер, настаивать на том, что они представлены у различных агентов с одинаковой интенсивностью, было бы чрезмерным упрощением. Ф. Лью выделяет следующие параметры [25], определяющие конкретный тип агента с точки зрения *эпистемической логики и игр с несовершенной информацией* [23]:

а. *сила дедуктивных способностей* (inferential power), т. е. способность совершать необходимые шаги вывода. Под дедуктивными способностями мы подразумеваем умение строить корректные выводы, поэтому, говоря о том, что они могут различаться у разных агентов, мы подразумеваем, что эти способности тем выше, чем больше таких выводов агент может построить. В этом ключе можно выделить два количественных аспекта дедуктивных способностей: внешний, или количество выводов, которые агент способен осуществить; и внутренний, или количество шагов, за которые агент способен получать выводы. Что означают дедуктивные способности с точки зрения выражающих их логических формализмов, если отождествить последние с соответствующими модальными системами? В нормальных модальных системах дедуктивные способности агентов всегда будут

одинаковыми относительно данной системы и «идеальными» в смысле внешнего и внутреннего качественных аспектов в силу особенностей этих формализмов.

Наибольшее затруднение связано с K -аксиомой $K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$, из которой вытекает, что агент знает все тавтологии данной системы, а также все, что следует из множества его знаний. Суть логического всеведения заключается не только в том, что агент способен вывести все следствия из своих знаний, но также и в том, что он уже знает все следствия, не совершая никаких шагов в рассуждении, хотя с формальной точки зрения их количество может различаться, особенно относительно выведения следствий в разных системах. При формализации рассуждений и взаимодействия агентов с ограниченной рациональностью, отвечающих представлениям о людях или программах, логического всеведения необходимо избегать, с чем и связано стремление варьировать силы дедуктивных способностей в рамках эпистемических презумпций, как одного из параметров, характеризующих тип агентов [21]. Ограниченными силами дедуктивных способностей характеризуются агенты в диалоговой логике — в одном из игровых подходов к моделированию их взаимодействий, где агенты различаются в зависимости от того, какие ходы они имеют право совершить, например, способны ли они отвечать на любые атаки или только на последнюю из них [11];

в. интроспекция, т. е. способность выводить пропозиции по поводу своего собственного эпистемического состояния (своих знаний, убеждений и т. п.). В рамках эпистемической логики обычно говорят о позитивной и негативной интроспекциях, которые выражаются соответственно аксиомами $K\varphi \rightarrow KK\varphi$ и $\neg K\varphi \rightarrow K\neg K\varphi$. Согласно аксиоме позитивной интроспекции агент знает все, что он знает, это одна из особенностей относительно «логически всеведущего» агента, моделируемого посредством модальной системы $S4$. Если же агент не знает чего-либо, то он знает, что этого он не знает, согласно аксиоме негативной интроспекции. Эта компетенция негативной интроспекции характеризует относительно «логически всеведущего» агента, подразумеваемого в системе $S5$. Моделируя взаимодействия рациональных агентов, можно наделить их компетенциями, выраженными посредством этих двух постулатов интроспекции, только какой-либо одной из них или воздержаться от этого, так что агент вообще не будет обладать способностями интроспекции. С этой целью можно использовать более слабые аксиомы, позволяющие моделировать степень интроспекции, как это сделано у Смаллиана [14], о чем речь пойдет ниже. Ясно, что любое из данных решений необходимым образом влияет на результаты взаимодействий таких агентов;

с. способность к наблюдению. Под данным параметром мы подразумеваем способность агентов наблюдать события и состояния, выступающую важным фактором в связи с доступом агента к информации. Взаимодействие агентов в условиях ограничений, накладываемых на способность к наблюдению, моделируется в играх с *несовершенной информацией* и с *неполной информацией* и исследуется посредством теоретико-игровой семантики. В играх с несовершенной информацией агентов различают в зависимости от того, доступно ли им знание о своей позиции (раунде) в игре или нет, в то время как игры с неполной информацией характеризуются тем, что не все агенты обязательно должны знать структуру самой игры;

d. *память*. Данный параметр связан с количеством запоминаемой информации. При моделировании реальных агентов (т. е. агентов с ограниченной рациональностью) необходимо учитывать ограниченный характер объема хранимой информации, как в случае реальных людей, так и компьютерных вычислений. В рамках информационных технологий можно распространить данное ограничение на любую хранимую информацию;

e. *способность и тип пересмотра своих убеждений*. Данный параметр, пожалуй, является одним из наиболее обсуждаемых в современной литературе по доксистической и эпистемической логике. В русле этой компетенции выделяют агентов, обладающих такой способностью, и лишенных ее. Агентов первой группы далее разграничивают по доступным им типам пересмотра (ревизии) убеждений, которые, в свою очередь, связаны с тем, какие логические операции доступны агентам для такого пересмотра. Так, агенту может быть доступно только распространение, т. е. простое добавление нового убеждения к уже имеющимся, или расширение, когда добавление сопровождается проверкой на совместимость элементов вновь полученного множества убеждений. В последнем случае агенты могут отличаться друг от друга тем, каким образом определена доступная им операция взятия логических следствий, при помощи которой осуществляется эта проверка.

Агентам может быть доступно не только расширение, но и удаление убеждений, которое также может быть реализовано как простое теоретико-множественное вычитание или как сокращение, т. е. удаление с проверкой оставшегося множества на совместимость, как в случае с расширением. В случае когда агенты полагают, что новая информация может оказаться ложной, но при этом все же доверяют источнику, требуются так называемые мягкие способы обновления информации (soft information) — консервативное и радикальное. Данные типы обновления информации задаются на моделях, в которых помимо миров, отношения достижимости и означивания также представлено отношение порядка, когда агент считает то или иное положение дел x «как минимум столь же правдоподобным, как» и положение дел y . Такие миры можно назвать наилучшими или наиболее релевантными для агента. Суть радикального обновления (radical upgrade) заключается в том, что после поступления информации о том, что ϕ , все миры, в которых имеет место ϕ , становятся более правдоподобными, или лучше, чем все миры, где имеется не- ϕ , при этом внутри этих двух классов отношение порядка остается прежним.

Ван Бентем [19, 139] сравнивает данный вид обновления с социальной революцией. Консервативное обновление (conservative upgrade) предполагает, что только наиболее правдоподобные миры, где имеет место ϕ , оказываются выше в данной иерархии, чем миры, где нет ϕ . В остальном же старое отношение порядка сохраняется. Проблема пересмотра убеждений агента обсуждается не только в рамках эпистемико-доксистических теорий, но также и с теоретико-игровой точки зрения, например в [22]. Различные виды ревизии убеждений подробно рассмотрены в работе [30]. Для моделирования обновления убеждений концепция АГМ предлагает специальный «оператор ревизии» [1]. Один из примеров алгоритма пересмотра множества убеждений, представляющий собой проверку на совместимость с новой информацией (из достоверного источника), был

предложен в качестве интерпретации средневекового диспута с предписаниями (de obligationibus) в работе [24].

Когда мы говорим о том, что многообразие агентов возникает из варьирования вышеуказанных параметров, то имеем в виду, что относительно них существует так называемый тип совершенных агентов, к примеру, агенты, обладающие идеальной памятью, которые помнят все предыдущие состояния и действия в системе (агенты с *Perfect Recall*). Другим примером идеализации могут служить агенты, обладающие логическим всеведением.

3. Дедуктивные компетенции агентов

При построении формальных моделей поведения и взаимодействия агентов их свойства можно фиксировать на синтаксическом и семантическом уровнях теории. В зависимости от того, каким образом в логических теориях определены синтаксические и семантические постулаты, будут различаться не только сами теории, но и типы рациональных агентов, подразумеваемые в них. Эта идея не нова. Выше мы говорили о постулатах эпистемической и особенностях динамической логики, которые задают тот или иной тип агента, и обнаружили, что дедуктивные свойства знаний и мнений агентов, моделируемые при помощи нормальных модальных систем, могут носить контринтуитивный и даже спорный характер, например, когда речь идет о логическом всеведении, пересмотре убеждений или интроспективных компетенциях. Несмотря на то что средствами разных систем можно конструировать различных агентов с неидентичными компетенциями по силе и их качественному набору, относительно самого агента подобные системы являются конститутивными, потому что они не только определенным образом описывают его эпистемическое состояние, т. е. его знания о себе и о мире, но и задают определенный тип структуры семантической модели, служащий для означивания предложений, выражающих эти знания и мнения, и, следовательно, оказывают решающее влияние на умозаключения агентов. Это обстоятельство подчеркивает предложенная Р. Смаллианом классификация, где в зависимости от дедуктивных компетенций выделяется четыре типа агентов, различающихся «степенью самосознания» [14].

Агент типа 1 обладает наименьшим уровнем самосознания и имеет следующие характеристики:

а) Он знает некоторые, а возможно, все логические тавтологии (теоремы), или верит в них;

б) множество его убеждений логически замкнуто, т. е. для любых предложений X и Y , если он верит в X и верит в $X \rightarrow Y$, тогда он верит в Y .

Агент типа 2 знает, что его множество знаний и убеждений замкнуто относительно *modus ponens*, поэтому к условиям «а» и «б» добавляется следующее условие:

с) он верит в $(Vp \wedge B(p \rightarrow q)) \rightarrow Bq$ (т. е. выполняется формула $B((Vp \wedge B(p \rightarrow q)) \rightarrow Bq)$). В связи с этим утверждается, что данный тип агента обладает более высоким уровнем самосознания, чем агент типа 1. Очевидно, что любой агент, обладающий

компетенцией позитивной интроспекции, может получить данную формулу в качестве следствия из соответствующего постулата. Однако эксплицитное добавление аксиомы позитивной интроспекции к компетенциям агента автоматически превращает его в агента типа 3.

Агент типа 3, или нормальный агент (у Р. Смаллиана «Нормальный мыслитель»), характеризуется тем, что для него выполняется условие $Bp \rightarrow BBp$, т. е. если агент верит в p , то он верит, что верит в p , что соответствует нормальной модальной логике $S4$.

Агент типа 4 — это Нормальный мыслитель с наиболее развитым самосознанием, как называет его Р. Смаллиан. В отличие от Нормального мыслителя, агента типа 3, агент типа 4 верит в то, что относительно его знаний выполняется постулат позитивной интроспекции и для него имеет место в качестве аксиомы формула вида $B(Bp \rightarrow BBp)$.

Обратим внимание на то, что с точки зрения гипотезы ограниченной рациональности все четыре агента Смаллиана, начиная с агента типа 1, — это идеализированные типы и «страдают» логическим всеведением в той или иной степени, ведь даже для агента типа 1 выполняется постулат $(Bp \wedge B(p \rightarrow q)) \rightarrow Bq$. В соответствии с теоремой о дедукции мы можем переписать его в виде K -аксиомы: $B(p \rightarrow q) \rightarrow (Bp \rightarrow Bq)$, что означает, что агенты верят во все следствия (или знают их — в случае эпистемической логики), выводимые из их убеждений (или знаний). Для моделирования рассуждений реальных агентов данное утверждение является слишком сильным и несовместимо с гипотезой ограниченной рациональности.

Классификация Смаллиана показывает влияние эпистемических презумпций, таких как различный уровень интроспекции, или уровень самосознания — в терминах Смаллиана, не только на выводимость тех ли иных формул в рамках системы, но и на дедуктивные способности агентов, что хорошо иллюстрирует условный характер предлагаемой классификации. Тем не менее при характеристике выделенных нами типов презумпций можно говорить о том, что эпистемические и целевые параметры соответствуют модальному фрагменту, добавляемому к существующим пропозициональным или первопорядковым фрагментам.

Классификация агентов Смаллиана указывает еще на одно обстоятельство, характеризующее агентные презумпции и существенное для нашего исследования. Так как при их рассмотрении посредством эпистемических логик и логик действия наибольший интерес для исследователей обычно представляет именно модальный фрагмент, то базовой логике, к которой добавляются те или иные модальные аксиомы, редко уделяется внимание, какого она заслуживает. Обычно ограничиваются указанием на то, что подразумевается некоторая классическая логика. Тем не менее именно логика, подразумеваемая агентами в качестве базовой, будет представлять собой комплекс его дедуктивных презумпций. Так, в зависимости от того, какой тип следования агент использует, т. е. какими постулатами выражено определение следования, будут различаться те выводы, к которым он приходит. Нетрудно заметить, что в классификации Смаллиана в этом аспекте ключевую роль играют условия «а» и «б», характеризующие базовые дедуктивные

компетенции агента таким образом, что все агенты обладают ими в равной степени. Это касается совокупности немодальных тавтологий, принимаемых агентом, и правила отделения (*modus ponens*), позволяющего осуществлять первичное упорядочивание знаний. Стало быть, в силу «b», агент типа 1, равно как и агенты типов 2–4, наделен некоторыми способностями проверять свои знания, носящие композиционный характер — об этом говорит условие «a» [9, 47]. Следовательно, дедуктивные компетенции агентов Смаллиана идентичны, и различаются они лишь в связи со своими эпистемическими компетенциями.

Рассмотрим теперь пример формальной теории, в которой агенты обладают неодинаковыми дедуктивными компетенциями. Напомним, что агентные системы могут быть моно- или мультиагентными и могут предусматривать как одного-единственного агента, так и множество агентов, однако с точки зрения свойств агентов система будет мультиагентной только в том случае, если агенты наделены различными компетенциями. Дедуктивные презумпции агентов могут быть выражены по-разному, например посредством нетождественных представлений о следовании. В подобном случае встает вопрос о том, каким образом моделировать взаимодействие между такими агентами.

Одним из возможных способов представления агентного взаимодействия агентов логическими средствами является игра, в которой правила формулируются на нескольких уровнях, например, можно выделить уровни логических и структурных правил. Тогда логические правила, определяющие, что значит атаковать и защищать некую формулу, будут также задавать формальный язык игры и выступать правилами манипулирования атомарными и молекулярными формулами в ней, а структурные правила будут уточнять порядок и типы ходов, доступные игрокам. Ясно, что на класс выводимых в данной системе формул влияют и те и другие правила, хотя и делают это по-разному. В качестве примера возьмем диалоговую логику П. Лоренцена и К. Лоренца [26, 27]. В данной системе имеются следующие логические (или частичные) правила:

Связка	Атака	Защита
$X-!-A \wedge B$	$Y-?-L$	$X-!-A$
	$Y-?-R$	$X-!-B$
$X-!-A \vee B$	$Y-?-v$	$X-!-A$
		$X-!-B$
$X-!-A \rightarrow B$	$Y-!-A$	$X-!-B$
$X-!-\sim A$	$Y-!-A$	—
$X-!-\forall xA(x)$	$Y-?-\forall x/n$	$X-!-A[n/x]$
$X-!-\exists xA(x)$	$Y-?-\exists x$	$X-!-A[n/x]$

где X и Y — игроки (переменные для P — проponent или O — оппонент), причем $X \neq Y$, $?$ и $!$ — знаки для обозначения характера действия игроков: $?$ — атака в форме запроса, $!$ — утверждение некоторой формулы (в форме как атаки, так

и защиты). Задача проponentа заключается в том, чтобы отстаивать рассматриваемую формулу, а опponentа — ее оспаривать. Логические правила выполняют конститутивную функцию, и игра не может начаться, если какой-либо из игроков не придерживается их, поэтому на уровне логических правил между агентами не может быть различий при игровом моделировании их взаимодействий.

Структурные правила игры таковы:

1. Опponent может атаковать каждую формулу, выдвинутую проponentом, лишь один раз.
2. Проponent может атаковать формулу, выдвинутую опponentом, произвольное количество раз.
3. Игроки ходят поочередно.
4. Игроки могут атаковать формулы, выдвинутые на любом этапе диалога.
5. Проponent может утверждать атомарную формулу, если и только если она уже была выдвинута опponentом.
6. Правила защиты:
 - 6.1. Игроки могут защищать только последнюю атакованную и еще не защищенную формулу.
 - 6.2. Игроки могут защищать любую еще не защищенную формулу.

Правила выигрыша: окончанием диалога (игры) является состояние, когда один из игроков, делающий ход, не имеет возможности осуществить то или иное действие, т. е. для него нет ни одного возможного хода. В подобной ситуации выигрывает его противник.

В аспекте структурных правил нетрудно заметить два отличия игроков: у них неодинаковые роли, правила защиты формул от атак могут быть сформулированы по-разному. Различие ролей опponentа и проponentа, в соответствии с которым им доступны разные ходы, тесно связано с дедуктивными компетенциями агентов и носит конститутивный характер. В самом деле, опponent выступает в роли не только своего рода фальсификатора, но и несет «ответственность» за упорядочивание (пересмотр) защищенных формул, так как только те формулы, которые утверждает опponent, могут быть в дальнейшем использованы проponentом. Соответственно опponent является верификатором, поскольку на него возложена обязанность вносить формулы на обсуждение. Для защиты же своего тезиса он может атаковать выдвигаемые опponentом формулы неоднократно. Идентичность ролей агентов сделает невозможной игру, поэтому различие ролей агентов в диалоговой логике является конститутивным.

Формулировка правил защиты также влияет на дедуктивные компетенции агентов, однако если в аспекте атак и внесения формул на обсуждение роли агентов закреплены как неодинаковые, то в аспекте защит подразумевается, что в одной и той же игре агенты, напротив, одинаково руководствуются одной и той же формулировкой такого правила. В зависимости от правил защит, которые определены в системе, реализуется одна из концепций того, что считать истинной формулой. Так, если в системе имеется правило 6.1, то речь идет об интуиционистском варианте диалоговой логики, а правило 6.2 характеризует классическое определение отношения следования. Доказательство соответствия между данными типами

диалогов и интуиционистским и классическим следованием представлено в [16] с использованием секвенциального исчисления **ГКср**, и оно состоит в том, чтобы показать, что секвенциальный вывод для рассматриваемой формулы будет соответствовать выигрышной стратегии игрока в данной игре диалога. Посредством модификаций структурных правил в диалоговой логике можно задавать и более слабые системы, чем интуиционистская, например, минимальную логику. Минимальная логика в системе естественного вывода отличается от интуиционистской отсутствием правила $\perp \rightarrow D$, где D — произвольная формула, т. е. правила «из лжи следует все, что угодно». В генценовском секвенциальном исчислении это будет соответствовать уточнению в сукцеденте в том случае, если сукцедент верхней секвенции пуст, т. е. когда верхняя секвенция имеет вид $\Gamma \rightarrow \emptyset$, а нижняя $\Gamma \rightarrow D$, где Γ — множество формул, а D — произвольная формула. Таким образом, при помощи соответствующей формулировки структурного правила для защит, подразумевающего отсутствие обязательства ответить на все атаки, можно различать интуиционистский диалог от диалога для минимальной логики и тем самым различать между тем, что понимается под дедуктивной выводимостью в том и в другом случае. В качестве секвенциального исчисления для доказательства соответствия между такими диалогами и логиками можно использовать модифицированный вариант исчисления G_3 Клини.

Итак, структурные правила взаимодействия в диалоговых логиках могут рассматриваться как дедуктивные презумпции агентов в конкретной логической системе в двух ракурсах: как метаправила, определяющие тип отношения следования, и как соответствующие роли в игре. В первом ракурсе агенты, подразумеваемые в одной игре, не отличаются друг от друга, но могут быть неодинаковыми в разных играх. Во втором ракурсе нетождественность агентов носит конститутивный характер.

4. Заключение

К вопросу многообразия агентов можно подойти с двух различных позиций. Во-первых, можно считать, что многообразие агентов — это многообразие логических систем, в рамках которых формально задается или подразумевается тот или иной тип агента. В целом подобные системы могут быть как мультиагентными (допускающими более одного агента), так и моноагентными. Во-вторых, можно попытаться представить многообразие агентов, которые вступают друг с другом во взаимодействие в рамках одной системы. Данное понимание проблемы многообразия агентов кажется нам наиболее интересным для дальнейшего рассмотрения.

Источники многообразия, или агентные презумпции, можно условно разделить на три группы составляющих их компетенций, две из которых были более подробно описаны в данной статье:

1. Дедуктивные презумпции, связанные с тем, каким образом в базовой (не-модальной) системе задано определение логического следования.

2. Эпистемические презумпции — знания, убеждения, включающие компетенции агентов по доступу к информации и ее хранению, а также способности модифицировать множество своих убеждений.

3. Целевые презумпции, или интенции и предпочтения касательно действий и поведения агентов: как само содержание интенций и предпочтений, влияющее на поведение агента, так и тип интенций агента.

В русле этих трех групп презумпций, своеобразных источников многообразия агентов, важная граница пролегает между дедуктивными компетенциями агентов, с одной стороны, и эпистемическими и целевыми компетенциями — с другой. Многообразие в силу различий агентов в смысле эпистемических и целевых компетенций совместимо как с различиями агентов в части дедуктивных компетенций, так и с отсутствием таких различий между ними. Однако многообразие агентов с точки зрения дедуктивных компетенций влечет их многообразие в эпистемическом смысле, это мы продемонстрировали на примере классификации агентов Р. Смаллиана. Вместе с тем имеются определенные особенности формальных теорий, относительно которых невозможно многообразие агентов в дедуктивных компетенциях, — это касается определения формального языка и способов построения формул в ней, что было показано на примере диалоговой логики, где многообразие агентов в аспекте дедуктивных компетенций задается структурными правилами, но не логическими.

-
1. *Альчуррон К. Э.* Отделение и отменяемость в деонтической логике / пер. с англ. Е. Н. Лисанюк, А. М. Павловой // «Нормативные системы» и другие работы по философии права и логике норм / К. Э. Альчуррон, Е. В. Булыгин, П. Герденфорс, Д. Макинсон ; под ред. Е. Н. Лисанюк. СПб., 2013. С. 358–370.
 2. *Драгаллина-Черная Е. Г.* Неформальные заметки о логической форме. СПб., 2015. С. 202.
 3. *Канеман Д., Словик П., Тверски А.* Принятие решений в неопределенности: Правила и предубеждения. Харьков, 2005. С. 632.
 4. *Караваев Э. Ф.* Вера и знание: Наследие Канта в философии и логике сегодня // Вестн. Санкт-Петербург. ун-та. Серия 17 : Философия. Конфликтология. Культурология. Религиоведение. 2014. № 3. С. 5–15.
 5. *Кислов А. Г.* Динамическая логика и деонтические операторы в «строгом смысле» // Философия науки. 2012. Т. 52, № 3. С. 65–80.
 6. *Лисанюк Е. Н.* Когнитивные характеристики агентов аргументации // Вестн. СПбГУ. Серия 6. Вып. 1. 2013. С. 13–21.
 7. *Лисанюк Е. Н.* Лояльный агент и отменяемость в деонтической логике // Изв. Урал. федер. ун-та. Серия 3 : Общественные науки. 2014. Вып. 1(125). С. 32–44.
 8. *Микиртумов И. Б.* Теория смысла и интенциональная логика. СПб., 2006. С. 288.
 9. *Микиртумов И. Б.* Композициональность и ее прагматика // Эпистемология и философия науки. 2013. Т. 36, № 2. С. 42–58.
 10. *Нийншилүто И.* Истина. Абсолютная или относительная? / пер. с англ. Е. Н. Лисанюк, А. А. Савченко // Вестн. СПбГУ. Серия 17. 2013. Вып. 4. С. 26.
 11. *Павлова А. М.* Истинность в диалоговой логике и теоретико-игровой семантике (GTS) // Логические исследования. 2015. Т. 21, № 2. С. 107–133.
 12. *Пойа Д.* Математика и правдоподобные рассуждения / пер. с англ. И. А. Вайнштейна ; под ред. С. А. Яновской. М., 1957. С. 370.

13. *Светлов В. А.* Современные индуктивные концепции (логико-методологический анализ). Л., 1988.
14. *Смаллиан Р.* Вовечи неразрешимое. Путь к Геделю через занимательные загадки / пер. В. В. Целищева. М., 2013. С. 303.
15. *Хинтиikka Я.* В защиту невозможных возможных миров / пер. А. Л. Никифорова // Хинтиikka Я. Логико-эпистемологические исследования / сост., общ. ред. В. Н. Садовского, В. А. Смирнова М., 1980. С. 231.
16. *Alama J., Knoks A., Uckelman S. L.* Dialogues games for classical logic (short paper) // M. Giese, & R. Kuznets (Eds.), *Tableaux 2011: Workshops, tutorials and short papers*. Bern : Universiteit Bern. 2011. P. 82–86.
17. *Aumann Robert J.* Agreeing to Disagree. *The Annals of Statistics*. 1976. Vol. 4, № 6. P. 1236.
18. *van Benthem J.* Logical Construction Games // *Acta Philosophica Fennica*. 2006. 78. P. 123–138.
19. *van Benthem J.* *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2010.
20. *van Benthem J., Liu F.* Diversity of Logical Agents in Games // *Philosophia Scientiæ*. 2004. 8(2). P. 165–181.
21. *Egre P.* Propositional Attitudes and Epistemic Paradoxes. PhD thesis, Universite Paris 1 et IHPST, 2004.
22. *Hintikka J.* *The Principles of Mathematics Revisited*. Cambridge University Press, Cambridge, 1996.
23. *Hötte T.* A Model for Epistemic Games. Master thesis, 2004.
24. *Lagerlund H., Olsson E.* Disputation and Change of Belief // *Burley's Theory of Obligations as a Theory of Belief Revision*, in Yrjönsuuri, 2001. P. 35–62.
25. *Liu F.* Diversity of Agents and their Interaction // *Journal of Logic, Language and Information*. 2008. 18(1). P. 23–53.
26. *Lorenzen P.* *Constructive philosophy* // Transl. by Pavlowic. Amherst : University of Massachusetts Press, 1987.
27. *Lorenzen P., Lorenz K.* *Dialogische Logik*. Darmstadt : Wissenschaftliche Buchgesellschaft, 1978.
28. *Meyer J.-J., Veltman F.* Intelligent agents and common-sense reasoning // Blackburn P., Benthem J. van, Wolter F. (eds). *Handbook of Modal Logic*. Elsevier. 2007. P. 991–1029.
29. *Pacuit E.* Dynamic Epistemic Logic I: Modeling Knowledge and Belief // *Philosophy Compass*. 2013. Vol. 8, № 9. P. 798–814.
30. *Pacuit E.* Dynamic Epistemic Logic II: Logics of Information Change // *Philosophy Compass*. 2013. Vol. 8, № 9. P. 815–833.
31. *Schumann A.* Reflexive games and non-Archimedean probabilities. *P-Adic Numbers, Ultrametric Analysis and Applications*. 2014(1). Vol. 6. P. 66–79.
32. *Shoham Y.* 1993. Agent-oriented programming. *Artificial intelligence* 60: 51–92.
33. *Stich St.* Reflective equilibrium, analytic epistemology and the problem of cognitive diversity // *Synthese*. 1988. 74(3). P. 391–413.

Рукопись поступила в редакцию 15 октября 2015 г.