

УДК 164.3

DOI: 10.17223/1998863X/64/2

**Г.В. Карпов**

## **ЛОГИКА ДЕЙСТВИЙ КАК ИНСТРУМЕНТ ДЛЯ АНАЛИЗА И РЕШЕНИЯ ФИЛОСОФСКИХ ПРОБЛЕМ**

Исследование выполнено за счет гранта Российского научного фонда  
(проект № 20-18-00158).

*Статья посвящена вопросу о том, допустимо ли сегодня современную логику рассматривать как инструмент, которым может эффективно пользоваться специалист в области философии. Мы исследуем методы, которые предлагает одна из современных логик – логика действия, и показываем, как ее арсенал может быть использован в области философии. В заключение мы высказываем мысль о том, что логика действий прочно войдет в набор исследовательских приемов специалиста-философа.*

*Ключевые слова: логика действий, философская проблема, агент, логическая семантика, уклонение*

### **Логика – это инструмент философского знания**

Заявления специалистов в области философии о том, что логика является инструментом, позволяющим формулировать новые вопросы относительно того или иного предмета, вносить ясность в обсуждаемые проблемы и даже решать их, являются общим местом как в вводных курсов по логике и философии и энциклопедических статей, так и специальных работ. Инструментальный характер логики в самом деле не вызывает сомнений, если речь идет, например, о традиционной логике как дисциплине и такой сфере профессиональной деятельности, как, например, юриспруденция. Однако в том случае, когда речь заходит об использовании специалистами-философами классических и неклассических логических теорий, например модальной логики, то их инструментальный, прикладной и эвристический характер уже не так очевиден.

В качестве примера, демонстрирующего саму возможность и удобство такого использования логических теорий, мы сошлемся на знаменитый онтологический аргумент Алвина Плантинга, в основе которого лежат аксиомы модальной логики S5 [1]. Широкая дискуссия, которая развернулась в среде специалистов, занятых философией религии, в связи с результатом Плантинга, говорит о том, что если подобное применение одной из логических теорий и не решает существующие философские проблемы, то, по крайней мере, рождает новые, что в данном случае не следует расценивать с практической точки зрения, т.е. в отрицательном ключе. Что же касается не столь провокационных примеров использования логических теорий, то здесь сами философы не столько демонстрируют весь потенциал названного инструмента, сколько, напротив, прячутся за общими фразами или, под видом дискуссий об использовании логического знания в области философии, говорят о дру-

гом. Так, например, Rod Girle в книге *«Modal Logics and Philosophy»* [2], вопреки обещающему названию, во второй части, которая посвящена приложениям логики к областям философии, дает обзор современных логик (классической, модальной, темпоральной, динамической, эпистемической, деонтической) точно так же, как это делают большинство стандартных логических учебников, где ни о каком специальном применении рассматриваемых логических теорий речи не идет. Подобным образом действует автор книги *«Modal Logic for Philosophers»* [3], которая также содержит лишь описание нескольких нормальных систем модальной логики и никаких инструкций о том, как работать с этими системами, т.е. как их применять. Это замечание справедливо и в отношении логики действия – направления, нацеленного, как об этом говорит Стэнфордская философская энциклопедия, на изучение философских аспектов действий формальными средствами [4]. Кроме того, если взглянуть на любую из логических статей последних лет (и статьи по логике действий здесь не являются исключением, см., например: [5, 6]) то можно обнаружить, что почти все эти тексты на три четверти состоят из описания теорий и на одну четверть из выяснения того, что можно созданными средствами утверждать содержательно о том или ином предмете. Все это вызывает рост разобщения специалистов в области логики и философии, отечественные философские журналы, публикующие статьи по логике и философии, все чаще принимают лишь те работы, где не допускается сколь бы то ни было громоздких формальных построений, а сами логики замыкаются в своих «чистых» логических исследованиях, считая, что все существенное о философских вопросах скажут в других, не сугубо логических журналах и без их участия.

Этим текстом мы хотим показать, что если не вся логика, то, по крайней мере, ее часть, а именно, – логика действий, была и остается отраслью философского знания, что она обладает потенциалом, который позволяет ей формулировать нетривиальные утверждения, имеющие отношение к исконным философским вопросам, и что перечень этих утверждений можно множить, если научиться правильно использовать те инструменты, которые она предлагает. Мы также надеемся, что эта статья послужит толчком для специалистов-логиков выступить в печати с аналогичными утверждениями об эффективности и удобстве использования их логик в области философии. Тем самым тезис о логике как инструменте философского знания получит как бы индуктивное обоснование, в то время как эта статья сегодня даст ему обоснование через пример.

### **Какие инструменты предлагает логика действий?**

Логика действий существует сразу в нескольких вариантах: как формализованная теория с креном в сторону семантических и собственно философских исследований [7–10]; как формальная теория с приоритетом логикоматематических штудий, в которых философские задачи, например в области приложения логики действий к вопросам этики, отходят на второй план [11–14]. На сегодняшний день логика действий располагает несколькими семантиками: классической и упрощенной семантиками ветвящегося времени, разработанными пионерами этого направления – Дж. Хорти и Н. Белнапом, – семантикой, основанной на модельных структурах Крипке, и семантикой для последующих состояний (так называемой NEXT-семантикой). Классическая

семантика является самой старой и самой мощной из всех: ее выразительных средств достаточно для того, чтобы интерпретировать описательные предложения, выражающие самые разнообразные смысловые оттенки совершаемых агентами действий. Упрощенная семантика используется преимущественно тогда, когда не возникает необходимости в интерпретации предложений, описывающих стратегические действия, когда решение о действии и последствие действия отставали бы друг от друга во времени. Семантика, в основе которой лежат структуры Крипке, служит средством теоретико-игровой интерпретации логики действия, что, в свою очередь, составляет часть программы по поиску оснований, которые бы позволили объединить ее с другими логиками, где задаются и решаются сходные вопросы – прежде всего с пропозициональной динамической логикой (см.: [15]). Семантика для последующих состояний используется специально для удобства интерпретации предложений, описывающих различные по своему виду стратегические действия агентов (см.: [16]).

$BT + AC + I^1$  – классическая семантика логики действий. В ее основе лежит древовидная структура<sup>2</sup>, объединяющая множество моментов  $T = \{m_1, m_2, \dots, m_n\}$ , которые находятся в отношении частичного порядка  $<$ . В том случае, когда мы выделяем внутри  $T$  максимальное множество линейно упорядоченных элементов, мы говорим об *истории*  $h$ . Каждая такая история  $h$  – это ветка дерева  $T$ . В том случае, когда мы берем горизонтальный срез всех возможных историй из  $T$ , мы говорим об *альтернативе*  $i$  из множества альтернатив  $I$ , куда попадают по одному моменту из каждой истории, находящемуся как бы на одном временном уровне. Функция *Choice* для каждого момента  $m$  из  $T$  и каждого агента, совершающего выбор в данном моменте, разбивает множество историй  $H_{(m)}$ , проходящих через этот момент, на классы так, что, например, в один класс попадают множество историй, в каждой из которых следует ожидать последствия выбора агента в момент  $m$ , а во все прочие классы – множество историй, где не содержится ни одной, в которой эти последствия имеют место. Выражение «*Choice* ( $\alpha, m$ ), ( $h$ )» обычно используют для обозначения множества историй, эквивалентных истории  $h$ , которую агент  $\alpha$  выбирает в момент  $m$ . Структуры  $\langle T, <, I, \text{Choice}, \text{Agents} \rangle$  оказывается достаточно для того, чтобы интерпретировать разнообразные предложения произвольного искусственного языка, описывающие действия агентов  $\alpha, \beta, \gamma, \dots$  из множества *Agents*.

Наиболее существенным преимуществом классической семантики, ради которого исследователь терпит некоторые неудобства, связанные, как мы сейчас увидим, с громоздкостью определений, является уже отмечавшийся выше ее универсальный характер. Средств  $BT + AC + I$  хватает на то, чтобы формулировать и затем изучать осмысленные утверждения о самых разных типах действий: когда агент совершает некоторое действие намеренно или же случайно, по неосторожности; когда он или она обдуманно действуют или,

<sup>1</sup>  $BT$  – от *branching time* (ветвящееся время, англ.),  $AC$  – *agent choice* (выбор агента, англ.),  $I$  – *instants* (буквально – моменты времени, англ.; здесь мы передаем  $I$  как множество альтернатив, чтобы отличать его элементы, отдельные альтернативы  $i_1, i_2, \dots$ , от моментов времени  $m_1, m_2$  из множества  $T$ ).

<sup>2</sup> «Древовидный универсум», в терминологии Блинова и Петрова, авторов единственной известной нам монографии по логике действий на русском языке, вышедшей в 1991 г., т.е. в то самое время, когда это направление переживало свой первый расцвет (см.: [17]).

напротив, оказываются в ситуации, порожденной скорее не-действием, невмешательством в некоторые события, происходящие как бы сами собой. Для каждого такого действия  $BT + AC + I$  предлагает или специальный символ предметного языка (так называемый агентный оператор, например:  $astit$ ,  $bstit$ ,  $cstit$ ,  $dstit$ , ...,  $xstit$ ) или сочетания одного из таких символов с символами, выражающими операторы обыкновенной модальной логики (например, алетическими: «необходимо, что», «возможно, что», или деонтическими: «должно быть так, что...», «разрешено, что...», или временными: «будет так, что...», «всегда было так, что...»).

Рассмотрим подробнее несколько наиболее употребимых агентных операторов и их интерпретации в классической семантике. Традиционно интерпретация задается с помощью выделения внутри класса модельных структур класса моделей, т.е. таких модельных структур, в которых интерпретируемая формула истинна. Для  $cstit$ -оператора такая модель описывается следующим образом. Формула  $[\alpha \text{ cstit}: \varphi]$ <sup>1</sup> считается истинной в некотором мире  $m/h$  (который традиционно обозначается индексом, объединяющим символы для момента и истории), если и только если в каждом мире  $m/h'$ , где история  $h'$  принадлежит множеству историй  $\text{Choice}(\alpha, m)$ ,  $(h)$ , которые агент  $\alpha$  выбирает в момент  $m$ , истинной оказывается формула  $\varphi$ <sup>2</sup>. Если нам нужно описать такое действие, для которого характерны автоматизм (т.е. отсутствие сколь бы то ни было значимого временного зазора между принятием решения и собственно совершением действия), а, кроме того, отсутствие значимых альтернатив (ситуаций, в которых бы оказался агент, если бы он не совершил указанное действие), то для этого сложно подобрать кандидатуру, лучшую, чем  $cstit$ -оператор, и образованное с его помощью выражение предметного языка. Так действует, например, агент-водитель, который ведет машину по пустому и прямому шоссе без разветвлений и примыканий дорог. Едва ли его поведение можно рассматривать как такое, когда каждую секунду он принимает решение о том, как именно ему вести машину, и едва ли он сам рассматривает в качестве сколь бы то ни было значимой альтернативы своему действию ситуацию, когда он, вместе со своим ТС, съезжает с проезжей части и оказывается в канаве. Но что происходит с нашим агентом-водителем, когда вдруг справа от него появляется примыкание дороги – поворот?

Эту ситуацию отличает от предыдущей наличие для водителя некоторой существенной альтернативы: возможности сделать поворот. Из-за того, что этот поворот есть, водитель как бы становится другим по своему типу агентом, так как получает теперь возможность поступить по-другому. Вне зависимости от того, была ли эта возможность реализована, и даже вне зависимости от того, придавался ли он в связи с этим соответствующим размышлениям, тот факт,

<sup>1</sup> Или, как мы будем писать далее, следуя принятой сегодня нотации, –  $[\alpha]^c \varphi$ , где « $[\alpha]^c$ » есть  $cstit$ -оператор. Для всех прочих операторов действия применяется тот же принцип обозначения, когда первая буква их названия выносится в верхний индекс, а символ, обозначающий агента, совершающего действие того типа, на который указывает оператор, заключается в квадратные или в угловые скобки.

<sup>2</sup> Мы не даем здесь определения истинности формулы  $\varphi$ , неагентной формулы логики действий, т.е. той, которая не содержит в качестве своего элемента агентный оператор, так как оно не отличается от множества других аналогичных определений любой из модальных логик с реляционной семантикой. Аналогично мы поступаем со всеми прочими неагентными формулами. Никаких других формул в языке логики действий нет.

что поворот есть или был и что, следовательно, водитель, только если бы он захотел, мог бы повернуть, делает его тем, кого в логике действий принято называть dstit-агентом. Dstit-оператор, использующийся для описания подобных ситуаций, определяется так: формула  $[\alpha]^d\varphi$  истинна в некотором мире  $m/h$ , если и только если в каждом мире  $m/h'$  (где снова история  $h'$  принадлежит множеству историй  $\text{Choice}(\alpha, m)$ ,  $(h)$ , которые агент  $\alpha$  выбирает в момент  $m$ ) истинной оказывается формула  $\varphi$ , и, кроме того, если найдется история  $h''$ , которая принадлежит множеству историй  $H_{(m)}$ , проходящих через момент  $m$ , такая, что в мире  $m/h''$  будет истинной формула  $\sim\varphi$ . Как видим, наличие существенной альтернативы, которой мог бы воспользоваться, но не воспользовался агент, делает его или ее именно dstit-агентом.

Наконец, вообразим ситуацию, когда перед агентом-водителем возникает *разветвление* дорог, – так, что никакой возможности оставаться «автоматическим» cstit- или dstit-агентом больше нет. В данном случае приходится принимать решение, так как без него движение вперед невозможно. Такого рода действия наиболее адекватно описывает astit-оператор. Формула  $[\alpha]^a\varphi$  истинна в мире  $m/h$ , если и только если найдется такой момент  $w$ , предшествующий моменту  $m$ , что, во-первых, во всех мирах  $m'/h'$ , которые принадлежат множеству историй  $\text{Choice}(\alpha, w)$ ,  $(h)$ , где  $h'$  принадлежит множеству  $H_{(m)}$ , истинной оказывается формула  $\varphi$ , и, во-вторых, если найдется такой момент  $m''$  из множества альтернатив  $i_{(m)}$  по отношению к моменту  $m$ , когда в некотором мире  $m''/h''$  имеет место формула  $\sim\varphi$ . Отмеченную выше громоздкость определений (являющуюся, впрочем, лишь следствием стремления к максимально возможной для формализованной теории точности) несколько уравнивают изображения описанных моделей, аналогичные по стилистике знаменитым «stit-pictures» из статей Н. Белнапа и др. начала 90-х гг. прошлого века.

На рис. 1 во втором ряду даны изображения моделей для формул с astit-, dstit- и cstit-операторами; первый ряд рис. 1 содержит схематическое изображение аналогий, каждая из которых иллюстрирует отличительные особенности формул, использующих указанные операторы. Для каждой схемы первого ряда справедливо следующее: если агент-водитель был в точке 1, а оказался в точке 2, то он действовал как astit-, dstit- или cstit-агент соответственно. В случае *в* – продолжил движение прямо, при условии, что других альтернатив этому действию не было; в случае *б* – также продолжил движение прямо, но при условии, что он мог поступить и по-другому; в случае *а* – повернул, предварительно приняв решение о том, в какую сторону поворачивать.

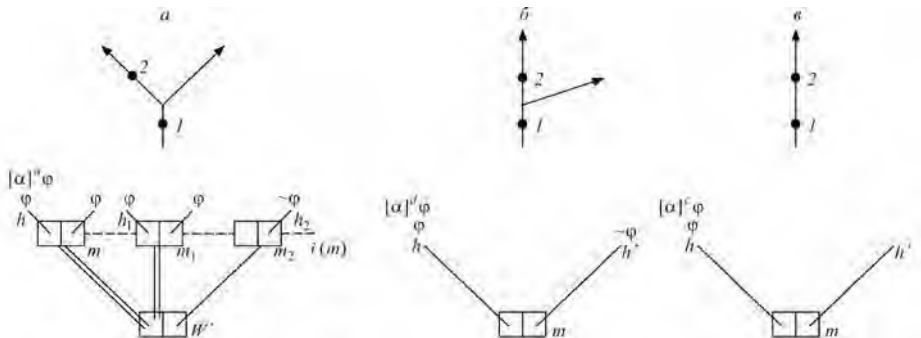


Рис. 1. Нижний ряд – семантические модели для формул с astit-, dstit- и cstit-операторами; верхний ряд – аналогия с действиями агента-водителя для каждого вида оператора действия

Таким образом, мы видим, что логика действий предлагает исследователю предметный язык и средства интерпретации его выражений, на котором можно описывать разнообразные ситуации совершения действий агентами, иногда трудно отличимые друг от друга – при первом рассмотрении или при использовании одного лишь естественного языка. Тем не менее разница, на которую мы можем ясно указать, воспользовавшись для этого средствами логики действий, может иметь решающее значение как для понимания природы действия вообще, так и для исследования отдельных кейсов, касающихся специфических действий, их этического и социального измерений.

### Степень агентности

Посмотрим на случаи *a* и *b* на рис. 1 внимательнее. В случае *b* выбор одной из альтернатив осуществляется как бы сам собой, так как агент-водитель может продолжать движение «прямо» или следуя некоторому заранее определенному порядку действий, или не следуя никакому решению вовсе. Напротив, совершить поворот, как это происходит в случае *a*, едва ли возможно без того, чтобы не было принято соответствующее решение. Мы хотим обратить внимание на то, что в случае *a* обе альтернативы предполагают одинаковые по сути действия, так как никакой возможности просто двигаться «прямо», поддавшись силе привычки или естественному ходу вещей, здесь нет. Равным образом, например, курильщик, который не закуривает очередную сигарету, совершает действие, которое можно назвать действием в гораздо большей степени, чем действие того агента, кто не закуривает, так как не курит вообще и не имеет к этому ни малейшей тяги. Это значит, что в связи с одним и тем же положением дел («агент  $\alpha/\beta$  делает так, что он не курит») курильщик ( $\alpha$ ) и некурильщик ( $\beta$ ) являются агентами в разной степени: агент  $\beta$  – в меньшей, так как ему это действие не стоит ни малейшего труда, агента  $\alpha$  – в большей, так как аналогичное действие стоит ему подчас многого. Мы предполагаем далее, что в отношении каждой пары вида  $\langle \alpha, \varphi \rangle$  ( $\alpha$  – агент, положение дел) можно говорить о степени агентности  $\alpha$  в связи с  $\varphi$ .

Будем считать, что степень агентности коррелирует с усилием, которое необходимо приложить агенту для реализации некоторого действия. Одни действия совершаются сами собой и как бы в силу привычки и не требуют сколь бы то ни было заметного усилия для своего осуществления. Для совершения других действий агентам приходится уговаривать себя, или подбадривать себя, или, например, обещать себе награду. Во всех этих случаях они оказываются в положении тех, кто смотрит на себя со стороны и кто действует исходя из этой перспективы – теперь сами агенты выступают для самих же себя объектом приложения волевого усилия, объектом собственных действий. Мы будем считать, что синтаксически степень агентности одного агента в отношении одного и того же положения дел можно выразить количеством агентных операторов. Пусть выражение  $\langle [\alpha]\varphi \rangle$  отвечает некоторой базовой степени агентности, а выражение  $\langle [\alpha][\alpha]\varphi \rangle$  (которое следует читать так: агент  $\alpha$  делает так, что агент  $\alpha$  делает так, что имеет место  $\varphi$ ) указывает на то, что агенту  $\alpha$  для выполнения действия, такого, что  $\varphi$ , приходится прикладывать дополнительное усилие.

Попытаемся перевести эту интуицию о степени агентности в знание, воспользовавшись для этого предметным языком логики действий, интерпре-

тированным в ее классической семантике. Сложность, с которой мы при этом сразу же сталкиваемся, сводится к следующему. В рамках семантики  $BT + AC + I$  действует принцип (\*)  $[\alpha]\phi \leftrightarrow [\alpha][\alpha]\phi$ , причем он справедлив вне зависимости от того, о каком типе агентного оператора идет речь: *astit*, *cstit* или *dstit*. Принцип (\*) говорит о том, что формулы « $[\alpha]\phi$ » и « $[\alpha][\alpha]\phi$ » семантически неотличимы друг от друга, следовательно, не могут выражать разные степени агентности. Таким образом, для того чтобы показать, что две данные формулы, по крайности, суть не одно и то же и что, более того, каждая из них может ассоциироваться с действием определенной степени, нам нужно тем или иным образом поставить принцип (\*) под сомнение. Это кажется невозможным, так как любое правильно построенное доказательство неуниверсальности (\*) свидетельствовало бы, на первый взгляд, о противоречивости базовых определений классической семантики. Однако в действительности подобное положение дел говорит в большей степени не столько о противоречивости семантики, сколько о наличии среди ее базовых определений некоторого скрытого компонента, умолчания, которое само по себе и является причиной противоречия. В самом деле, случаи обнаружения таких умолчаний в  $BT + AC + I$  уже встречались ранее. Так, например, в отношении принципа SMP, аналога *modus ponens* для агентных предложений, было установлено, что его выполнимость зависит от того, в какой порядок на древовидной структуре выстраиваются предложения, выражающие посылки, что делает SMP не универсально выполнимым в тех случаях, когда его посылки выражены с помощью предложений с *astit*-операторами (см.: [7. Р. 799–800]). Умолчание, напоминающее это, мы надеемся обнаружить и в отношении (\*).

Принцип (\*) образован соединением принципов 4 ( $[\alpha]^a\phi \rightarrow [\alpha]^a[\alpha]^a\phi$ ) и T ( $[\alpha]^a[\alpha]^a\phi \rightarrow [\alpha]^a\phi$ ). Обоснования, предложенные для 4 и T (см.: [8. Р. 501–502]), опираются на *witness identity lemma*, лемму о тождестве момента выбора (WIL). В соответствии с ней, если верно, что  $\phi \rightarrow \psi$ , и если формулы  $[\alpha]^a\phi$  и  $[\alpha]^a\psi$  оказываются истинными в одном и том же мире, то не может быть так, что момент выбора для  $[\alpha]^a\phi$  предшествует моменту выбора для  $[\alpha]^a\psi$  (см.: [8. Р. 498]). Доказательство принципа T использует WIL явным образом, что оправдано, так как в самом деле, во-первых,  $[\alpha]^a\phi$  влечет  $\phi$  и, во-вторых, и антецедент, и консеквент T выполняются в одном и том же мире. Другими словами, оба условия, гарантирующие саму возможность использования WIL в построении доказательства для T, наличествуют. Но в отношении доказательства принципа 4, в связи с возможностью использования WIL, все вкладывается иначе. Хотя и антецедент, и консеквент принципа 4 выполняются в одном и том же мире, неверно, что  $\phi$  влечет  $[\alpha]^a\phi$ , – т.е. одно из условий, гарантирующее возможность применения WIL, не выполняется. Что, в свою очередь, говорит о том, что мы можем исходить из допущения, согласно которому моменты выбора для антецедента и консеквента 4 – суть *разные моменты*. Таким образом, использование WIL при построении доказательств для 4 непосредственно, и, косвенно, – при доказательстве (\*) как раз и является тем умолчанием, которое мы надеялись обнаружить. Рассмотрение модельных структур с разными моментами выбора для агентных формул, задействованных в 4, может дать нам модель, где этот принцип не будет выполняться.

В самом деле, если допустить, что момент выбора  $w$  для  $[\alpha]^a\phi$  предшествует моменту выбора  $w'$  для  $[\alpha]^a[\alpha]^a\phi$  (рис. 2), то придется признать, что во всех мирах, следующих за  $w$  и входящих в множество  $\text{Choice}(\alpha, w)(h_1)$ , имеет место  $\phi$ . Так как между мирами внутри множества  $\text{Choice}(\alpha, w)(h_1)$  устанавливается отношение эквивалентности, то, следовательно, в каждом мире, принадлежащем данному множеству, имеет место не только  $\phi$ , но и  $[\alpha]^a\phi$ , потому что формула  $[\alpha]^a\phi$  истинна в мире, обозначенном  $m_1/h_1$  из  $\text{Choice}(\alpha, w')(h_1)$ . Допуская противоположное тому, что мы хотим показать, будем считать, что в мире, который мы обозначили индексом  $m_1/h_1$ , истинна  $[\alpha]^a[\alpha]^a\phi$ . Тогда среди миров, принадлежащих множеству  $\text{Choice}(\alpha, w')(h_5)$ , должен найтись хотя бы один такой, в котором истинной окажется формула  $\sim[\alpha]^a\phi$ . Однако так как множество  $\text{Choice}(\alpha, w')(h_5)$  принадлежит множеству  $\text{Choice}(\alpha, w)(h_1)$ , то здесь всюду, как мы только что установили, во всех мирах должна быть истинна формула  $[\alpha]^a\phi$ . Следовательно, неверно, что формула  $[\alpha]^a[\alpha]^a\phi$  истинна в  $m_1/h_1$ , наряду с формулой  $[\alpha]^a\phi$ .

На рис. 2 показана модель, где истинными в одном и том же мире оказываются формулы  $[\alpha]^a\phi$  и  $\sim[\alpha]^a[\alpha]^a\phi$ , что опровергает принцип 4 и, следовательно, принцип (\*).

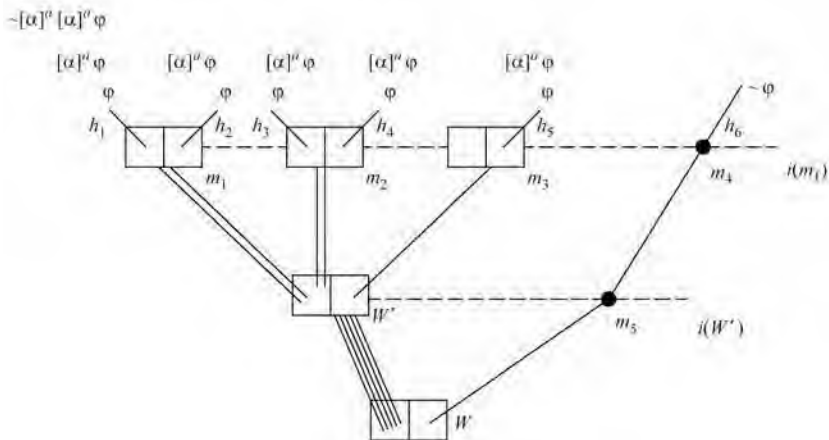


Рис. 2. Семантическая модель, где совместимы по истинности формулы  $[\alpha]^a\phi$  и  $\sim[\alpha]^a[\alpha]^a\phi$

Таким образом, формулы  $\langle[\alpha]^a[\alpha]^a\phi\rangle$  и  $\langle[\alpha]^a\phi\rangle$  могут выражать разные действия, которые не сводятся друг к другу, хотя и оканчиваются одним и тем же результатом, одним и тем же положением дел « $\phi$ », реализованным агентом  $\alpha$ . Мы предлагаем понимать формулу  $\langle[\alpha]^a\phi\rangle$  как выражение такой степени агентности, когда действие, хотя и не совершается как бы само собой, как это имеет место в случае с действиями, обозначенными формулами  $\langle[\alpha]^c\phi\rangle$  или  $\langle[\alpha]^d\phi\rangle$ , но требует некоторого минимального волевого усилия, минимального участия в самом себе как в деятеле. Так действуют заядлый курильщик, когда закуривает очередную сигарету, водитель, выбирающий поворот тогда, когда последствия этого выбора ему или ей в общем-то безразличны (ведь обе дороги, например, ведут в одно и то же место), любой другой агент, поставленный перед выбором, который не является существенно важным. Формула  $\langle[\alpha]^a[\alpha]^a\phi\rangle$ , напротив, указывает на то, что агенту, который принимает решения о каком-то действии, требуется приложение уси-

лия – как для того, чтобы принять это решение, так и для того, чтобы затем реализовать задуманное. Выше мы предложили понимать такое усилие как подбадривания, уговоры или обещания, адресованные агентом самому себе. Если трудно приступить к выполнению некоторого действия, трудно реализовать некоторое положение дел, такое, что «ф», то приходится сначала прикладывать усилия, чтобы приняться за дело – делать так, что делаешь так, что «ф». Следует ожидать, что награда, как и наказание, если того заслуживают действия агентов, в двух указанных случаях, отличающихся степенями агентности, будут разными. Учет степеней агентности открывает новое измерение в логике и философии действия, связанное с дифференциацией заслуг, ответственности агента за реализацию одного и того же действия в зависимости от приложенных усилий.

### Агенты и агентные домены

Заметим, во-первых, что для всякого агента  $\alpha$  следует, строго говоря, указывать множество предложений  $D_\alpha$ , т.е. его, как мы будем здесь говорить, *агентный домен*. Если предложение предметного языка логики действий входит в  $D_\alpha$ , то это означает, что  $\alpha$  может сделать так, что данное предложение истинно. Множество предложений из  $D_\alpha$  есть такое множество описаний положений дел, относительно которых  $\alpha$  и является агентом. Таким образом, сказать, что  $\alpha$  – агент, значит сказать недостаточно для того, чтобы ясно определить, кто такой  $\alpha$ ; напротив, сказать, что  $\alpha$  – агент в связи с положением дел, таким, что предложение «ф» истинно – значит определить агентный домен  $\alpha$  и тем самым сказать об  $\alpha$  достаточно для того, чтобы мыслить его как агента более или менее ясно.

Во-вторых, будем считать, что в обычных, базовых ситуациях для агента  $\alpha$  верно, что  $D_\alpha = \{\phi, \psi, \dots, \sim\phi, \sim\psi, \dots\}$ , т.е. что  $\alpha$  является агентом в связи с таким положением дел, что оказывается истинным множество неагентных формул, т.е. формул без операторов « $[\alpha]$ », « $[\beta]$ » и т.д. Однако ясно, что, например, множеству  $D_\alpha$  могут принадлежать и формулы с агентными операторами, например, такие: « $[\beta]\phi$ » или даже « $[\alpha]\phi$ ». В первом случае это означает, что агент  $\alpha$  может сделать так, что агент  $\beta$  делает так, что имеет место такое положение дел, что формула «ф» истинна. В этом случае  $\alpha$  может отдать, например, указание  $\beta$ , чтобы тот выполнил действие, результат которого мог бы быть описан предложением «ф». (Эта структура, на которую не раз обращали внимание специалисты в области логики действий, в деталях исследована в работе [18]). Во втором случае, если  $D_\alpha$  принадлежит формула « $[\alpha]\phi$ », то это значит, что  $\alpha$  может отдавать указание сам себе, т.е. реализовывать некоторое действие с приложением большего усилия, чем это имеет место в том случае, когда  $D_\alpha$  принадлежит исключительно формула «ф».

В-третьих, мы хотим обратить внимание на то, что об агентах можно судить не только по тому, что они могут сделать, но и по тому, что они сделать не могут. Например, в большинстве ситуаций в отношении агента  $\alpha$  и положения дел, такого что «у  $\alpha$  не болит голова», верно, что указанное предложение входит в  $D_\alpha$ . Это так, ведь для того чтобы достичь этой ситуации в случае, когда в действительности имеет место обратное, в распоряжении  $\alpha$  достаточно средств. Однако утверждение о том, что формула « $\sim\phi$ » («у  $\alpha$  болит голова») принадлежит  $D_\alpha$ , также, как и формула «ф», вероятно, все же

неверно: в самом деле, за исключением случаев, когда некто намеренно вызывает у себя головную боль, как правило, никто не станет рассматривать себя как агента в связи с положением дел «у меня болит голова». Таким образом, вероятно, существуют положения дел, которые входят, вместе со своими отрицаниями, в агентные домены, и существуют положения дел и действия, которые входят в агентные домены или сами по себе, или с отрицаниями, но не то и другое одновременно.

Вернемся в связи со сказанным об агентных доменах к нашему примеру с агентом-водителем на рис. 1, а. Пусть формула  $\varphi$  означает «агент  $\alpha$  повернул налево». В этом случае агентный домен  $D_\alpha(\varphi)$  представлен множеством  $\{\varphi, \sim\varphi, \sim[\alpha]^a\varphi, \sim[\alpha]^a\sim\varphi\}$ . Наличие в нем формулы  $\varphi$  говорит о том, что найдется такая модель, в которой будет справедлива формула  $[\alpha]^a\varphi$ ; формула  $\sim\varphi$  из  $D_\alpha(\varphi)$  говорит о том, что допустима модель, где истинна формула  $\sim[\alpha]^a\varphi$ , описывающая ситуацию, когда  $\alpha$  делает так, что он не поворачивает налево; формула  $\sim[\alpha]^a\varphi$  говорит о том, что может быть построена модель для  $[\alpha]^a\sim[\alpha]^a\varphi$  ( $\alpha$  уклоняется от того, чтобы повернуть налево); а формула  $\sim[\alpha]^a\sim\varphi$  — о том, что может быть построена модель для  $[\alpha]^a\sim[\alpha]^a\sim\varphi$  ( $\alpha$  делает так, что он не делает так, что он не поворачивает налево, или:  $\alpha$  уклоняется от того, чтобы не повернуть налево). Иначе этот агентный домен можно представить в виде матрицы возможностей, открытых для нашего агента-водителя  $\alpha$ , в основе которой лежит следующее соображение: он может повернуть налево или не повернуть налево; он может сделать то и другое «специально» или «просто так» (таблица).

Агентный домен в виде матрицы возможностей

| «Просто»                | «Специально»                          |             |
|-------------------------|---------------------------------------|-------------|
| $[\alpha]^a\varphi$     | $[\alpha]^a\sim[\alpha]^a\sim\varphi$ | Повернул    |
| $[\alpha]^a\sim\varphi$ | $[\alpha]^a\sim[\alpha]^a\varphi$     | Не повернул |

Наш  $\alpha$  «просто» поворачивает налево, когда путь, который он выбирает, является одним из возможных путей, ведущих к точке назначения. Это не обязательное действие, в своем роде игра. Поворот налево, выполненный «специально», означает, что, например, иного пути нет или он есть, но чрезмерно долог. Здесь существенным оказывается момент обязательности действия, связанный с конечной целью всего путешествия. Аналогично, «просто» не повернуть означает, например, позволить себе задуматься и проехать поворот тогда, когда это можно сделать без боязни не достигнуть цели путешествия в срок. Не повернуть «специально» — уйти от неприятностей, которые, возможно, ждут за поворотом. Результат действий, описываемый парой формул « $[\alpha]^a\varphi$ » и « $[\alpha]^a\sim[\alpha]^a\sim\varphi$ », и парой формул « $\sim[\alpha]^a\varphi$ » и « $[\alpha]^a\sim[\alpha]^a\varphi$ », одинаков. Однако смысл действий, выраженных парами указанных формул, различен. Для того чтобы сделать смысловые отличия более выпуклыми, мы можем сказать, что  $\alpha$  (водитель и, вдобавок, бывший курильщик), «просто» поворачивая налево, выбирает один из возможных путей, в конечной точке которых бассейн или оздоровительный центр; поворачивая налево «специально»,  $\alpha$  уклоняется от всех возможных дорог, которые хотя и ведут в бассейн или оздоровительный центр, но в то же время проходят мимо торговых центров с табачными лавками.

Возможный агентный домен  $D_a(\varphi)$  приобретает пугающие черты, если в содержании « $\varphi$ » мыслить предложение « $\alpha$  совершает убийство». Вероятно, это тот случай, когда невозможно ни «просто» убить, ни «просто» не убить, если под убийством понимать умышленное причинение смерти. Таким образом, здесь общий вид агентного домена сокращается до множества, состоящего из двух элементов:  $\{\sim[\alpha]^a\varphi, \sim[\alpha]^a\sim\varphi\}$ , т.е. до его «специальной» части. Интересно, что формулу « $[\alpha]^a\sim[\alpha]^a\sim\varphi$ », выражающую «специальное» «положительное» действие, аналогичное специальному повороту налево, в этом случае следует читать так:  $\alpha$ , агент-убийца уклоняется от того, чтобы не убивать. Отсутствие действия убийства по умолчанию принимается, таким образом, за некую норму, или императив. Совершить убийство – нарушить этот императив. Чудовищем выглядит тот, кто «специально» не убивает, т.е. всякую минуту воздерживается от того, чтобы убить. Кажется, что здесь нет места «нормальным людям», и каждый является либо убийцей, либо тем, кто готов пойти на убийство во всякое время.

Работа с агентными доменами, как и учет степени агентного действия, позволяет нам обращать внимание на различие в возможных смыслах тех действий, которые оканчиваются одинаковым результатом, видеть в таких действиях как бы второе дно, или, напротив, удовлетворяться тем, что этого дна в том или ином действии нет.

### Заключительные замечания

Мы надеемся, нам удалось показать, что использование логической теории, логики действий в области философии и само является философией, как процесс, и дает философию (формальную философию, философию аналитического толка, в нашем случае, но все же, как мы надеемся, именно философию) как свой результат. В подтверждение этому мы отдаем на суд читателя два сюжета собственного изобретения – о степени агентности и об агентных доменах, претендующие на то, чтобы также служить подтверждением главному тезису статьи.

До сих пор в рамках исследуемого направления можно было изучать действия, формулируя о них предложения на предметном языке с использованием разных операторов, каковых в распоряжении исследователя к 10-м годам текущего столетия набралось чуть ли не на целый алфавит: *astit*-, *bstit*-, *cstit*-, *dstit*-, ..., *xstit*-. Однако даже в прикладном аспекте этих исследований о действиях в-большей-или-в-меньшей-степени речи не заходило. Обратившись к примеру борьбы с дурными привычками, мы впервые указали как на принципиальную возможность фиксации этого отличия имеющимися средствами, так и на то новое в понимании природы действия, что удалось отсюда вывести. В частности, мы предложили интерпретировать суперпозицию агентных операторов вида « $[\alpha][\alpha]$ » как логическую форму усилия, которое прикладывает агент, совершая некоторое действие.

Во втором сюжете, который касается доменов агентности, мы показали, как использование базовых определений логики действий в совокупности с почти механистическими методами, вкладывает в руки исследователя средство, позволяющее различать смыслы таких действий, которые тем не менее одинаковы по своему результату.

Наконец, и понятие степени агентности, и понятие агентного домена вместе позволили нам обратить внимание на феномен действий агентов, направленных на самих себя. Мы полагаем, что в этой сфере открыть логических форм, логических законов и сформулировать нетривиальных философских утверждений можно не меньше, чем в традиционной сфере «неавтороферентных» действий, которые являлись основным предметом исследований направления до сих пор.

Для многих сочетание «логика действий» продолжает звучать так же, как сочетание «логика музыки» или «логика истории». Какая у музыки, истории или действия может быть логика? От этих, в чем-то профанных, вопросов все же не следует отмахиваться, ведь над всякой областью философии постоянно висит опасность превращения ее в работу по описанию правил выдуманных, не существующих в реальности игр, игр, в которые никто, кроме самих философов, не играет. В итоге они рискуют обнаружить себя занятыми производством текстов, содержащих высказывания, составленные из пустых понятий<sup>1</sup>. Эта статья дает некоторые основания для того, чтобы предложить переименовать нашу область и назвать ее не логикой действий, а логикой выбора. В том, что у выбора есть логика, и в том, что рассуждения о структуре выбора (чего бы он ни касался: желаемого предмета, будущего вообще, или того, как поступать или думать) иногда могут привести того, кто их начал, к обобщенным наблюдениям о природе выбора вообще, времени вообще, поступка вообще, стоящим, кроме того, как бы в преддверии философии, мы не сомневаемся. Это еще одно подтверждение нашему тезису: рассуждение о структуре феномена (впрочем, наверно, все же – развитое до уровня теории, хотя бы формализованной) с необходимостью дает следствия в область, где исследуются наиболее общие законы возникновения и существования всего многообразия предметов и явлений, каковой и является собственно философия.

Подобно тому, как в XVIII в. логика обыкновенно служила введением в философию вообще (по крайней мере, начиная с Христиана Вольфа, и, по крайней мере, – в философию Просвещения), можно надеяться, что и сегодня логика займется тем же, благо инструментов и потенциала, как это видно на примере логики действия, достаточно. Дело, таким образом, за энтузиастами, которых, если только идея, которую мы здесь высказали, хоть немного жизнеспособна, поставит рано или поздно сама корпорация, высшая школа.

### Литература

1. *Plantinga A.* The Nature of Necessity. Oxford : Oxford University Press, 1974. 255 p.
2. *Girle R.* Modal Logics and Philosophy, 2009. 240 p.
3. *Garson J.* Modal Logic for Philosophers. Cambridge University Press, 2006. 506 p.
4. *Segerberg K., Meyer J.-J., Kracht M.* The Logic of Action // The Stanford Encyclopedia of Philosophy / ed. by E.N. Zalta. URL: <https://plato.stanford.edu/archives/sum2020/entries/logic-action> (accessed: 15.05.2020).
5. *Lorini E., Sartor G.* A STIT Logic for Reasoning About Social Influence // *Studia Logica*. 2016. 104. P. 773–812.
6. *Broersen J.* Deontic epistemic stit logic distinguishing modes of mens rea // *Journal of Applied Logic*. 2011. Vol. 9. P. 137–152.
7. *Belnap N.* Backwards and forwards in the modal logic of agency // *Philosophy and Phenomenological Research*. 1991. Vol. 51, № 4. P. 777–807.

<sup>1</sup> Подробнее об этом см. в [19].

8. Chellas B.F. Time and Modality in the Logic of Agency // *Studia Logica: An International Journal for Symbolic Logic*. 1992. Vol. 51, № 3/4, Logic of Action. P. 485–517.
9. Horty J.F., Belnap N. The Deliberative Stit: A Study of Action, Omission, Ability, and Obligation // *Journal of Philosophical Logic*. Dec., 1995. Vol. 24, № 6. P. 583–644.
10. Ciuni R., Lorini E. Comparing Semantics for Temporal Stit Logic // *Logique Et Analyse*. 2018. Issue 243. P. 299–339.
11. Xu M. Axioms for Deliberative “Stit” // *Journal of Philosophical Logic*. Oct., 1998. Vol. 27, № 5. P. 505–552.
12. Payette G. Decidability of an Xstit Logic // *Studia Logica*. 2014. 102. P. 577–607.
13. Xu M. Decidability of Stit Theory with a Single Agent and Refref Equivalence // *Studia Logica*. 1994. 53. P. 259–298.
14. Balbiani P., Herzig A., Troquard N. Alternative Axiomatics And Complexity Of Deliberative Stit Theories // *Journal Of Philosophical Logic*. August 2008. Vol. 37, № 4. P. 387–406.
15. Benthem J. van, Pacuit E. Connecting Logics of Choice and Change // *Nuel Belnap on Indeterminism and Free Action* / ed. by T. Müller. Springer, 2014. P. 291–314.
16. Broersen J. A complete STIT logic for knowledge and action, and some of its applications / eds. M. Baldoni, T.C. Son, M.B. van Riemsdijk, M. Winikoff // *Declarative Agent Languages and Technologies VI*. DALT 2008. Lecture Notes in Computer Science, vol 5397. Springer, Berlin, Heidelberg. P. 47–59.
17. Блинов А.Л., Петров В.В. Элементы логики действий. М. : Наука, 1991. 232 с.
18. Logie R.D. A Study of Agent Influence in Nested Agent Interactions. PhD thesis. The Open University, 2010. 270 p. URL: <http://oro.open.ac.uk/62023/> (accessed: 20.06.2020).
19. Dennett D. Higher-order truths about chmess. *Topoi*, 2006. P. 39–41.

**Gleb V. Karpov**, St. Petersburg State University (St. Petersburg, Russian Federation).

E-mail: [glebsight@gmail.com](mailto:glebsight@gmail.com)

*Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya – Tomsk State University Journal of Philosophy, Sociology and Political Science*. 2021. 64. pp. 16–29.

DOI: 10.17223/1998863X/64/2

## THE LOGIC OF ACTION AS A TOOL FOR ANALYZING AND SOLVING PHILOSOPHICAL PROBLEMS

**Keywords:** logic of action; philosophical question; agent; logical semantics; refraining

The study is supported by the Russian Science Foundation, Project No. 20-18-00158.

The article investigates the question about the possibility of treating contemporary logic with its methods of analysis and its discipline-specific subjects as a tool that can be used efficiently and comfortably by specialists in philosophy. From this point of view, I firstly investigate the methods of the logic of action – one of the branch of contemporary logic – and then demonstrate how the arsenal of the logic of action can be applied in the field of philosophy today. Firstly, I demonstrate the richness of the logic of action methods by expressing with the means of classical semantics of the notion of the agentive degree or the degree of agency. Secondly, by using the set-theoretical approach, I point out the new way of development for the logic of action that lies in the traditional area of philosophical knowledge. Here I argue that it is possible to regard as a new subject for the logic of action the selfward actions and the structures of so-called agentive domains, or the sets of state of affairs that can be put into existence via agentive actions. In the conclusion, I claim that the logic of action, or, following my proposal, the logic of choice with its methods will be well adopted into the set of investigation practices of those who are professionals in the realm of philosophy.

## References

1. Plantinga, A. (1974) *The Nature of Necessity*. Oxford: Oxford University Press.
2. Gärle, R. (2009) *Modal Logics and Philosophy*. Routledge.
3. Garson, J. (2006) *Modal Logic for Philosophers*. Cambridge University Press.
4. Segerberg, K., Meyer, J.-J. & Kracht, M. (2020) The Logic of Action. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*. [Online] Available from: <https://plato.stanford.edu/archives/sum2020/entries/logic-action> (Accessed: 15th May 2020).
5. Lorini, E. & Sartor, G. (2016) A STIT Logic for Reasoning About Social Influence. *Studia Logica*. 104. pp. 773–812. DOI: 10.1007/s11225-015-9636-x

6. Broersen, J. (2011) Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of Applied Logic*. 9. pp. 137–152. DOI: 10.1016/j.jal.2010.06.002
7. Belnap, N. (1991) Backwards and forwards in the modal logic of agency. *Philosophy and Phenomenological Research*. 51(4). pp. 777–807.
8. Chellas, B.F. (1992) Time and Modality in the Logic of Agency. *Studia Logica: An International Journal for Symbolic Logic*. 51(3/4). pp. 485–517.
9. Horty, J.F. & Belnap, N. (1995) The Deliberative Stit: A Study of Action, Omission, Ability, and Obligation. *Journal of Philosophical Logic*. 24(6). pp. 583–644. DOI: 10.1007/BF01306968
10. Ciuni, R. & Lorini, E. (2018) Comparing Semantics for Temporal Stit Logic. *Logique Et Analyse*. 243. pp. 299–339. DOI: 10.2143/LEA.243.0.3285131
11. Xu, M. (1998) Axioms for Deliberative “Stit”. *Journal of Philosophical Logic*. 27(5). pp. 505–552. DOI: 10.1023/A:1004274131669
12. Payette, G. (2014) Decidability of an Xstit Logic. *Studia Logica*. 102. pp. 577–607. DOI: 10.1007/s11225-013-9492-5
13. Xu, M. (1994) Decidability of Stit Theory with a Single Agent and Reflex Equivalence. *Studia Logica*. 53. pp. 259–298. DOI: 10.2307/2275842
14. Balbiani, P., Herzig, A. & Troquard, N. (2008) Alternative Axiomatics and Complexity of Deliberative Stit Theories. *Journal of Philosophical Logic*. 37(4). pp. 387–406. DOI: 10.1007/s10992-007-9078-7
15. Benthem, J. van & Pacuit, E. (2014) Connecting Logics of Choice and Change. In: Müller, T. (ed.) *Nuel Belnap on Indeterminism and Free Action*. Springer. pp. 291–314.
16. Broersen, J. (2008) A complete STIT logic for knowledge and action, and some of its applications. In: Baldoni, M., Son, T.C., van Riemsdijk, M.B. & Winikoff, M. (eds) *Declarative Agent Languages and Technologies VI. DALT 2008. Lecture Notes in Computer Science*. Vol 5397. Berlin, Heidelberg: Springer. pp. 47–59.
17. Blinov, A.L. & Petrov, V.V. (1991) *Elementy logiki deystviy* [Elements of the logic of actions]. Moscow: Nauka.
18. Logie, R.D. (2010) *A Study of Agent Influence in Nested Agent Interactions*. PhD thesis. The Open University. [Online] Available from: <http://oro.open.ac.uk/62023/> (Accessed: 20th June 2020).
19. Dennett, D. (2006) Higher-order truths about chess. *Topoi*. 25(1). pp. 39–41. DOI: 10.1007/s11245-006-0005-2