

---

*Философская логика*  
*Philosophical Logic*

---

Г.В. КАРПОВ

**Императивы в stit-подходе<sup>1</sup>**

**Карпов Глеб Викторович**

Кафедра логики, Институт философии,  
Санкт-Петербургский государственный университет.  
199034, Российская Федерация, Санкт-Петербург, Менделеевская линия, д. 5.  
E-mail: glebsight@gmail.com

В статье обосновывается возможность рассматривать императив как предложение, описывающее побуждающую к действию стратегию, которая реализуется одним рациональным агентом в отношении другого. Отстаивается тезис, что такого рода стратегии, выступающие основанием практических умозаключений, могут быть описаны средствами модальных логик, в частности, средствами stit-логики. Дается изложение основ семантики stit-логики и предпринимается исследование, связанное с поиском модальной формулы, наилучшим образом отражающей множество свойств императива и соответствующего ему действия рационального агента. Особое внимание уделяется формулам, в которых различные модальные stit-операторы стоят в суперпозиции. На том основании, что обычно суперпозиция таких операторов лишает агента-адресата выбора, а агента-адресанта — незнания о том, будет ли успешным его побуждающее действие, вносится предложение использовать в определениях истинности различные точки соотнесения для различных по типу операторов, что позволяет сохранить основные свойства императива в отношении данных формул, теряемые в обычном случае. В заключении показывается, что для таких формул невозможно возникновение парадоксов, обычных для предлагаемых ранее императивных логик.

*Ключевые слова:* stit-логика, императив, логика императивов, логика действий

## 1. Введение

Возможность установить одно-однозначное соответствие между множеством описательных предложений и множеством императивных предложений давно была подмечена исследователями, предпринимавшими еще в конце 40-х гг. XX в. попытки построить императивную логику

---

<sup>1</sup>Работа выполнена в рамках проекта РГНФ «Логика норм и нормативные системы» № 12-03-00196.

по аналогии с классической логикой. Для одних это замечание служило основным методологическим принципом работы<sup>2</sup>; другие, напротив, видели в наличии такого соответствия затруднение, препятствовавшее возникновению новой области в логике — логики императивных предложений, или императивов<sup>3</sup>. Последующие события показали, что отчасти неправы были представители обоих лагерей, так как и тот, и другой подходы не справлялись с возложенной на них задачей. Аналогия между существующей пропозициональной логикой и возможной логикой императивов пала под натиском множества парадоксов, которые она породила. Полноценных самостоятельных теорий, описывающих логическую форму императивных предложений и логические законы, которым они подчиняются, с помощью понятий истинностных значений, или понятий, альтернативных им, так и не возникло, а те, что появились, были подвержены парадоксам в такой же степени, как и теории, основанные на аналогии с пропозициональной логикой. Уже в 1970-е гг. почти все попытки описать хоть какое-то множество законов, в соответствии с которыми императивные предложения используются в рассуждениях, сошли на нет.

Мы склонны считать, что подобное положение дел, тот факт, что сегодня в логике императивным предложениям уделяется мало внимания, требует исправления. Причин тому несколько: во-первых, императивы составляют значительный объем того речевого материала, с которым нам приходится сталкиваться повседневно, и несправедливо то, что большинство умозаключений, использующих этот материал, так называемые практические силлогизмы, до сих пор осуществляются интуитивно и регулируются скорее этическими и социальными нормами, чем логикой; во-вторых, любой исследователь сегодня, естественно, обладает несравненно большим арсеналом подходов и техник, чем те, кто предпринимал попытки создания императивной логики уже более полувека назад. Исследованием, отчасти, руководит любопытство и, в нашем случае, это любопытство связано с вопросом о том, можно ли современными средствами, прежде всего — средствами модальной логики, заново поставить и решить вопрос о возможности существования логики императивов.

---

<sup>2</sup>Самым ярким примером здесь, пожалуй, являются Ричард Хеар и Вальтер Дубислав. Подробнее об их методологической установке в отношении императивов см. [2].

<sup>3</sup>Среди наиболее непримиримых критиков данной позиции Альф Росс — автор самого знаменитого парадокса, демонстрирующего неприемлемость приложения законов пропозициональной логики к области императивов.

Этот вопрос, впрочем, помимо вышеизложенного, имеет еще одно обоснование. Среди современных модальных логик нам удалось найти такую логику, которая, как мы надеемся показать, предоставляет достаточное методологическое основание для решения поставленного вопроса, в отличие от аналогии с пропозициональной логикой или понятий «удовлетворенности» и «удовлетворительности»<sup>4</sup>. Речь идет о *stit*-логике или *stit*-подходе, том ее варианте, что представлен в статьях 1990-х годов ученых Питтсбургского университета Нуэля Белнапа и Майкла Перлофа<sup>5</sup>. Разработанный ими инструмент позволяет исследователю с достаточной точностью вести речь о результатах действий агентов, анализируя на временной модельной структуре истинностные условия пропозициональных высказываний, которые находятся в области действия специальных модальных операторов. Очевидное сходство основной идеи *stit*-подхода и классической трактовки императивов, которые с самого начала интереса к ним, проявленного философами языка, рассматривались именно как действия, облеченные в форму предложения, и послужило основанием предположения о том, что решение многих вопросов в отношении императивов может быть найдено, если императив удастся представить средствами *stit*-логики.

Результаты применения *stit*-подхода в решении ряда «наболевших» вопросов логики действий позволяют нам высоко оценивать его методологический потенциал. Брайн Челлас уже в 1992 году, то есть еще до того момента, как в широкое употребление вошли разнообразные *stit*-операторы, существенным образом расширившие выразительные возможности данного подхода, называет *stit*-логику «привлекательной и (многое) объясняющей теорией» [4, р. 485]. Действительно, Белнапом в [5] было, например, показано, что на модельных структурах *stit*-логики не выполняется *stit*-формула, аналогичная принципу **K** логики алетических модальностей; что существует фиксируемая формальными средствами разница между действием агента  $\alpha$  по осуществлению такого положения дел, что  $\neg P$ , простым отсутствием действия по осуществле-

<sup>4</sup> Авторство этих двух понятий, выступавших эрзацами по отношению к понятию «истинно», принадлежит А. Дж. Кенни, см. [3].

<sup>5</sup> *Stit*-логика (от “see to it that”, буквально “следить за тем (гарантировать), чтобы нечто было сделано”) — разновидность логики действий (наряду с более известным AGM-подходом и различными динамическими логиками, сформулированными на его основе: динамической эпистемической, динамической деонтической и др.), где язык классической пропозициональной логики дополнен *stit*-операторами. Формулы со *stit*-операторами отсылают к осуществлению агентом действия, результатом которого является положение дел, описываемое пропозициональной переменной.

нию  $P$ , как такового, и уклонением от осуществления  $P$ ; что выполнение аналогов таких правил, как *modus ponens*, возможно в *stit*-логике только при соблюдении определенной последовательности моментов, в которых агент делает выбор из представленного набора историй, и т.п.

Мы считаем, что этот потенциал *stit*-подхода можно применить и в области императивов. Это тем более верно, что само содержание части проектов императивной логики с самого начала, то есть с конца 40-х годов XX века, было связано с поисками, в условиях ожидаемого отсутствия параллелизма с пропозициональной логикой, таких правил, которые бы описывали и объясняли логические переходы от одного императива к другому. Несогласованность результатов, которые дает *stit*-подход, с правилами пропозициональной логики и с некоторыми теоремами модальной логики, аналогичная той «непохожести» поведения императивных формул, что предполагалась частью исследователей (например Россом и Кенни) интуитивно, но не получала должного формального представления, есть то общее основание, которое мы усматриваем в *stit*-формулах и в императивных предложениях.

Начнем наше исследование с краткого описания языка *stit*-логики, на котором, по нашему предположению, мы надеемся говорить об императивных действиях.

### 1.1. Язык *stit*-логики

Язык *stit*-логики состоит из бесконечного счетного множества пропозициональных переменных  $p_1, p_2, \dots$ , бесконечного счетного множества агентов  $\alpha_1, \alpha_2, \dots, \beta_1, \beta_2, \dots$ , истинностно-функциональных связок  $\neg$  и  $\wedge$ , операторов  $[_{astit}: \_]$  и  $[_{dstit}: \_]$ <sup>6</sup>. Правильно построенные выражения языка (формулы) задаются обычным образом. Выражения  $P, Q, \dots$ , используются в качестве метавариабельных для формул. Выражения  $[\alpha_{astit}: P]$  и  $[\alpha_{dstit}: P]$  суть формулы, когда  $\alpha$  обозначает агента, и  $P$  обозначает формулу. Истинностно-функциональные связки  $\vee$ ,  $\rightarrow$  и  $\leftrightarrow$  задаются как обычно.

<sup>6</sup>Оператор  $[_{astit}: \_]$  — так называемый “achievement stit” (от англ. to achieve — выполнять, добиваться), введенный Белнапом в 1988 году и широко используемый в статьях, посвященных *stit*-логике, середины 90-х годов;  $[_{dstit}: \_]$  или “deliberative stit” (от англ. to deliberate — обдумывать, размышлять) — оператор, предложенный в 1989 году Хорти. Широкое применение в *stit*-логике *dstit*-получил после публикации Хорти в соавторстве с Белнапом статьи [11]. Различие данных операторов будет определено ниже семантически. Мы также будем использовать имя *stit* в качестве родового по отношению к *astit* и *dstit*.

ОПРЕДЕЛЕНИЕ 1 (МОДЕЛЬНАЯ СТРУКТУРА  $\mathcal{S}$ ). Модельная структура  $\mathcal{S}$  stit-логики задается с помощью кортежа  $\langle \mathbf{T}, \leq, \mathbf{Instant}, \mathbf{Agent}, \mathbf{Choice} \rangle$ .

1.  $\mathbf{T}$  есть множество, элементами которого являются моменты времени  $m_1, m_2, \dots, w_1, w_2, \dots$ .
2.  $\leq$  есть отношение порядка, заданное на множестве  $\mathbf{T}$ , такое, что для любых элементов, принадлежащих  $\mathbf{T}$ , выполняются отношения рефлексивности, транзитивности и антисимметричности.
3. Максимальная цепь моментов  $\langle m_0, m_1, \dots \in \mathbf{T} \rangle$  называется историей  $h \in \mathbf{H}$ . То есть история  $h$  есть такая последовательность моментов времени, на которой выполняются все отношения, заданные  $\leq$ , что для нее всегда можно указать момент, относительно которого не существует более раннего момента, а также момент, относительно которого не существует более позднего момента. Запись  $m \in h$  означает, что момент  $m$  встречается в истории  $h$ . Так как множество  $\mathbf{T}$  является ветвящимся «наверх» деревом, что соответствует идеи «открытого» будущего, мыслимого как множество альтернативных друг другу историй, то очевидно, что один и тот же момент  $m$  может встречаться в разных историях. Будем обозначать множество историй, каждая из которых проходит через момент  $m$ , как  $\mathbf{H}_m$ .
4. **Instant** есть множество элементов из  $\mathbf{T}$ , полученное в результате «горизонтального» разбиения  $\mathbf{T}$ , такого, что каждый инстант  $i \in \mathbf{Instant}$  включает в себя множество моментов времени альтернативных друг другу. Например, если  $i = \{m_1, m_2\}$ , то в действительности с течением времени может быть реализован только один элемент из множества  $i : m_1$  либо  $m_2$ .
5. **Agents** есть непустое множество агентов, элементами которого являются  $\alpha_1, \alpha_2, \dots, \beta_1, \beta_2, \dots$ .
6. **Choice** есть функция, которая каждому  $\alpha \in \mathbf{Agents}$  и каждому  $m \in \mathbf{T}$  ставит в соответствие множество  $\mathbf{Choice}(\alpha, m)$ .  $\mathbf{Choice}(\alpha, m)$  является разбиением множества историй  $\mathbf{H}_{(m)}$ , то есть оно указывает на те альтернативные друг другу множества историй, которые доступны агенту  $\alpha$  в момент времени  $m$ .

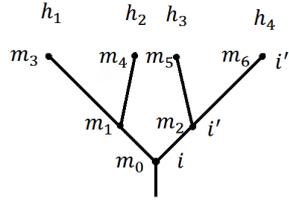


Рис. 1.

Пусть, например, модельная структура образована моментами времени  $m_0, m_1, \dots, m_6$ , отношением  $\leq$  и тремя инстантами  $i = \{m_0\}$ ,  $i' = \{m_1, m_2\}$  и  $i'' = \{m_3, m_4, m_5, m_6\}$ , такими, что вместе они определяют четыре истории  $h_1, h_2, h_3$  и  $h_4$ :  $h_1 = \{m_0, m_1, m_3\}$ ,  $h_2 = \{m_0, m_1, m_4\}$ ,  $h_3 = \{m_0, m_2, m_5\}$  и  $h_4 = \{m_0, m_2, m_6\}$ . Тогда функция **Choice** ставит в соответствие агенту  $\alpha$  в момент времени  $m_0$  множество **Choice**( $\alpha, m_0$ ), которое, в случае на рис. 1, может быть образовано набором  $K_1 = \{h_1, h_2\}$  или набором  $K_2 = \{h_3, h_4\}$ . Каждый набор  $K_i$  есть возможный выбор для  $\alpha$  в  $m_0$ . Очевидно, что, например, **Choice**( $\alpha, m_1$ ) =  $\{h_1, h_2\}$ , то есть возможным выбором для  $\alpha$  в  $m_1$  является каждая из историй  $m_1$  и  $h_2$ . Договоримся называть истории  $h_1$  и  $h_2$  неразделенными относительно некоторого момента времени  $m$ , если найдется такой  $m'$ , более поздний по отношению к  $m$ , что  $m' \in h_1$  и  $m' \in h_2$ . В нашем примере истории  $h_1$  и  $h_2$  не разделены относительно момента  $m_0$  (то же справедливо в случае с историями  $h_3$  и  $h_4$ ). Тогда верно, что между неразделенными историями невозможно сделать выбор: осуществляя выбор в момент времени  $m_0$ , агент  $\alpha$  располагает лишь двумя альтернативами, принадлежащими множеству **Choice**( $\alpha, m_1$ ), и, выбирая, например, первую из них —  $\{h_1, h_2\}$ , — он не может в этот момент времени предпочесть историю  $h_1$  истории  $h_2$  (и наоборот). Будем говорить, что в этом случае истории  $h_1$  и  $h_2$  эквивалентны относительно выбора агента  $\alpha$  в момент времени  $m_0$  ( $h_1 =_{m_0}^\alpha h_2$ ). Если  $h_1 =_{m_0}^\alpha h_2$  и  $m_1 \in h_1$ , а  $m_2 \in h_2$ , то будем называть моменты  $m_1$  и  $m_2$  эквивалентными относительно выбора агента  $\alpha$  в момент  $m_0$ .

## 1.2. Свойства императива

Императивное действие обладает рядом свойств, существенно отличающих его от любого другого предложения, и описательного в особенности. Следует обратить на эти свойства внимание, так как именно их нам предстоит выразить средствами языка stit-логики.

1. Мультиагентность (биагентность) искомой императивной формулы. Императив должен выражаться такой формулой, истинность которой поставлена в зависимость от совместных действий рациональных агентов — того, кто отдает некоторое приказание, и того, кто его выполняет. Мы будем требовать от формального языка, в котором выражается императив, фиксации этого отличия агентов друг от друга — за счет положения модального оператора в соответствующей формуле, или за счет ее вида.
2. Искомая формула должна сохранять незнание агента-адресанта о конечном результате его императивного действия, что, вероятно, может быть представлено через понятие эквивалентных историй или моментов.
3. Все действия, которые описывает искомая формула, должны удовлетворять критерию акциональности, в соответствии с которым действие считается таковым только тогда, когда агент имеет возможность и совершить его, и воздержаться от его совершения. Это в полной степени соответствует «критерию К», сформулированному в [1, с. 7]. В отношении императива это требование распространяется на всех рациональных агентов. Будем считать, что агент является рациональным агентом до тех пор, пока в отношении него продолжает действовать данный критерий.
4. С помощью искомой формулы должны выражаться такие свойства императивного действия, как осуществление и выполнение. Будем понимать под осуществлением императивного действия такое его свойство, которое приобретает им как следствие наличия факта произнесения соответствующего предложения а также, факта, свидетельствующего о том, что, например, приказ отдан, просьба высказана. Под выполнением будем понимать такое свойство императива, которое приобретает им в случае, если агент-адресат выполняет назначенное ему к выполнению агентом-адресантом: приказ в этом случае исполняют, просьба удовлетворяется. На модельной структуре нужно будет учитывать, по крайней мере, три возможных исхода для императива: его неосуществление, его осуществление и невыполнение, его осуществление и выполнение.
5. К очевидным свойствам императивного предложения, которые, вероятно, не составит труда выразить, отнесем:

- (а) Бикомпонентность — наличие в структуре императива описательной и модальной частей.
- (б) Возможность одномоментного или последовательного осуществления императивов разными агентами.
- (в) Возможность осуществления простых императивных действий, где содержание императива представлено формулой без бинарных пропозициональных связок, и осуществление сложных императивных действий, предполагающих наличие в содержании императива формул с бинарными пропозициональными связками.

Посмотрим, какие средства выражения предлагает нам *stit*-подход для того, чтобы записать с помощью формул императивное действие. Все наши пробы мы будем осуществлять, придерживаясь достаточно простого метода: относительно формулы-кандидата, после ряда содержательных пояснений, будет предприниматься попытка построения модели на модельной структуре  $\mathcal{S}$  в соответствии с задаваемыми ниже определениями истинности для *stit*-формул. Условия истинности *stit*-формулы будут определяться исходя из значений ее пропозициональных (немодальных) подформул. Договоримся записывать эти значения непосредственно над линиями, которые обозначают истории. *Stit*-формулы, истинность которых является следствием данного истинностного набора пропозициональных формул, будем записывать над пунктирной чертой и над той веткой, где данная *stit*-формула принимает значение «истинно». Договоримся считать модель некоторой формулы построенной, если не существует противоречия между формулами, находящимися над пунктирной линией, и формулами, находящимися под ней. Также будем считать, что *stit*-формула не выражает закон *stit*-логики, если удастся подобрать такой истинностный набор ее пропозициональных переменных, который дает противоречие между «нижними» и «верхними» относительно пунктирной линии формулами. Данные договоренности лежат в основе всех доказательств приемлемости или неприемлемости некоторой *stit*-формулы, приводимых в настоящей статье.

## 2. *Astit*- и *astit-astit*

В данном разделе мы покажем, что для того, чтобы выразить императивное действие, возможностей *astit*- и *dstit*-операторов оказывается достаточно. Опишем условия истинности формулы с *astit*-оператором.

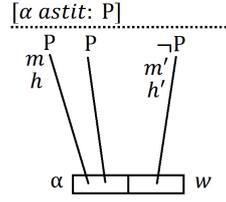


Рис. 2.

**ОПРЕДЕЛЕНИЕ 2.** Истинность  $[\alpha \text{ astit} : P]$  в модели  $\mathcal{M}$ . Пусть модель  $\mathcal{M}$  модельной структуры  $\mathcal{S}$  задается двойкой  $\langle \mathcal{S}, v \rangle$ , где  $v$  есть функция оценки, которая ставит в соответствие каждому выражению языка stit-логики пару  $m/h$ , указывающую момент и историю, где данное выражение становится истинным. Формула  $[\alpha \text{ astit} : P]$  истинна в модели  $\mathcal{M}$ , относительно  $m/h$ , если выполняются следующие два условия:

1. Найдется момент времени  $w$ , предшествующий  $m$ , в котором агент  $\alpha$  делает такой выбор, что любая из возможных историй  $h$ , заключенная в рамки данного выбора, с необходимостью повлечет истинность формулы  $P$  в  $m/h$ .
2. Найдется момент времени  $m'$ , являющийся частью истории  $h'$ , отсеченной выбором агента в моменте  $w$ , альтернативный для  $m$  и предшествующий  $w$ , относительно которого формула  $P$  является ложной.

Пункты 1 и 2 данного определения выполняются в модели, представленной на рис. 2, где, в отличие от рис. 1, момент  $w$  изображен так, что видно из каких множеств  $K \in \mathbf{Choice}(\alpha, w)$  он состоит. В дальнейшем все моменты, в которых агент делает выбор, мы будем представлять таким же образом. Место над пунктирной линией относительно  $m'/h'$  оставлено пустым, так как данная модель допускает здесь истинность нескольких astit-формул: можно с одинаковым успехом утверждать, что в ней относительно  $m'/h'$  истинными оказываются и формула  $[\alpha \text{ astit} : \neg P]$ , и формула  $\neg[\alpha \text{ astit} : P]$ . Истинность  $[\alpha \text{ astit} : \neg P]$  в  $m'/h'$  вытекает из определения 2 непосредственно. Истинность  $\neg[\alpha \text{ astit} : P]$  в  $m'/h'$  устанавливается на основании ложности  $[\alpha \text{ astit} : P]$  в  $m'/h'$  ввиду наличия над этой веткой  $\neg P$ . В [5, р. 795–797] показано, что формулы  $[\alpha \text{ astit} : \neg P]$  и  $\neg[\alpha \text{ astit} : P]$  обладают различным смыслом: первая говорит об осуществлении действия, описываемого пропозициональной формулой  $\neg P$ ; вторая — о воздержании

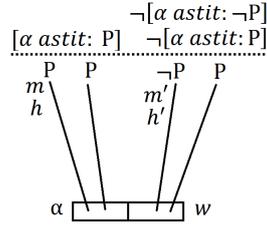


Рис. 3.

от осуществления действия  $P$ . Так как в модели на рис. 2 данные формулы неразличимы, мы оставляем это место незанятым; аналогично мы будем поступать во всех подобных ситуациях.

На рис. 3 представлена модель, позволяющая отличить формулы  $[\alpha \text{ astit} : \neg P]$  и  $\neg[\alpha \text{ astit} : P]$ : последняя истинна в  $m'/h'$ , в то время как первая, ввиду наличия в правой клетке выбора истории, оканчивающейся моментом, где истинна формула  $P$ , — нет. Действия, выражаемые формулой  $[\alpha \text{ astit} : P]$ , как бы растянуты во времени между моментом  $w$  выбора и моментом  $m$  осуществления выбранного. И то, и другое совершается одним и тем же агентом. Осуществление действия, такого, что формула  $P$  становится истинной, гарантировано выбором в предшествующем моменте. При этом агент не знает, какая именно из возможных историй, каждая из которых принадлежит  $\mathbf{Choice}(\alpha, w)$ , будет реализована. (Это незнание фиксируется в модельной структуре через понятия эквивалентных моментов и историй). Всегда найдется такая альтернативная история, в которой формула  $P$  будет ложной на том основании, что в предшествующем моменте агент решил или воздержаться от действия, или сделать так, что  $\neg P$ . Эта возможность «отрицательного» выбора существенна: ее отсутствие уничтожает саму идею действия на основании выбора: в том случае, если во всех историях, исходящих из  $w$ , имеет место или  $P$ , или  $\neg P$ , формула  $[\alpha \text{ astit} : P]$  не может быть объявлена истинной в какой бы то ни было истории и относительно любого момента (см. рис. 4). Агент в этом случае перестает быть рациональным агентом, так как его выбор и действие не приводит к изменению положения дел в мире. Очевидно, что формула  $[\alpha \text{ astit} : P]$  не может быть использована в чистом виде для записи императива. Наиболее существенным препятствием в этом отношении является неизбежность истинности формулы  $P$  относительно  $m/h$  в том случае, если истинна  $[\alpha \text{ astit} : P]$  относительно  $m/h$ . Так как  $[\alpha \text{ astit} : P]$  может быть либо истинной, либо ложной, то императивное действие,

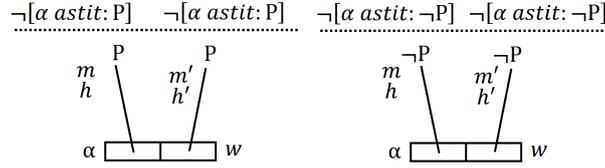


Рис. 4.

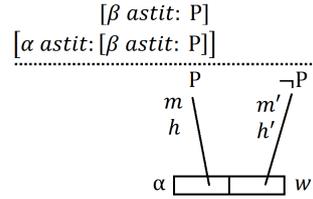


Рис. 5.

выражаемое с помощью  $[\alpha \text{ astit} : P]$ , может быть либо осуществленным и выполненным, либо неосуществленным и невыполненным. Ситуацию уклонения от исполнения приказа с помощью  $[\alpha \text{ astit} : P]$  выразить не удастся.

Нежелательное, в вопросе выражения в stit-подходе императивного действия, свойство одноагентности можно устранить соединяя astit-модальности в одной формуле так, что каждая такая модальность будет указывать на действие одного-единственного агента, и в зависимости от положения astit-модальности в формуле будет ясно, о каком типе агента идет речь — о том, кто отдает приказание, или о том, кто его принимает.

Рассмотрим формулу  $[\alpha \text{ astit} : [\beta \text{ astit} : P]]$  в качестве возможной кандидатуры на роль адекватного набору свойств императивного действия способа записи императива. Сформулируем для нее условия истинности относительно  $m/h$ . Если  $[\alpha \text{ astit} : [\beta \text{ astit} : P]]$  истинна в  $m/h$ , то должен существовать момент  $w$ , в котором агент  $\alpha$  делает выбор — принимает решение, силами агента  $\beta$  сделать так, чтобы имело место  $P$ . Тогда, в некотором моменте  $m$ , следующем за  $w$ , в соответствии с п. 1 определения 2, подформула  $[\beta \text{ astit} : P]$  формулы  $[\alpha \text{ astit} : [\beta \text{ astit} : P]]$  будет истинной. Так как  $[\beta \text{ astit} : P]$  истинна в  $m/h$ , то здесь же истинной оказывается и формула  $P$  (см. рис. 5). Выбор, сделанный агентом  $\alpha$  в момент  $w$ , с необходимостью влечет истинность формул  $[\beta \text{ astit} : P]$  и  $P$  в  $m/h$ . Агент  $\beta$  — предположительно адресат императивного действия, осуществляемого агентом  $\alpha$ , теряет

свое свойство агентности: момент  $t$  не может стать для него моментом выбора относительно того, осуществлять или нет  $P$ , ведь  $P$  уже истинно в  $t/h$ , и у  $\beta$ , следовательно, нет возможности выполнить требование п. 2 определения 2.

Кроме того, даже в том случае, если бы нам удалось сохранить агентность  $\beta$ , мы, тем не менее, основывались бы на еще одном неявном, но существенном допущении, которое никак не отражено на модельной структуре — последовательном осуществлении действий агентами. В рамках классического *stit*-подхода агенты действуют одномоментно, причем так, что выбор каждого из них не оказывает никакого влияния на выбор другого. Это свойство зафиксировано в определении, получившем название «нечто происходит» (*something happened conditions*), подробнее о нем, см., например, [4, р. 494]. Очевидно, что в случае императивного действия от этого определения нужно будет отказаться.

Интересен в связи с этим предположением результат, полученный Челласом в 1992 году (см. [4, р. 503–505]). Им обосновывается истинность формулы  $[\alpha \text{ astit} : [\beta \text{ astit} : P]] \rightarrow [\alpha \text{ astit} : P]$ , которая является записанным средствами языка *stit*-логики принципом права «*Qui facit per alium facit per se*» («Кто действует посредством другого, тот действует сам»), или просто принципом **Q**. Истинность этой формулы говорит о том, что указание на выполнение  $P$ , отдаваемое агентом  $\alpha$  агенту  $\beta$ , рассматривается на модельной структуре так, как если бы оно выполнялось самим агентом  $\alpha$ , что неприемлемо в отношении императивного действия, где не только должна быть сохранена агентность  $\beta$ , но и должны быть отделены друг от друга две агентности, два способа действия: осуществление императивного действия агентом-адресантом и выполнение назначенного агентом-адресатом.

Заметим также, что обоснование истинности принципа **Q**, предложенное Челласом, осуществляется через демонстрацию того факта, что для антецедента формулы, выражающей принцип, не удастся найти модель, т.е. он оказывается всегда ложен.

Любопытное, на наш взгляд, решение в отношении способа записи императивного действия с помощью суперпозиции *astit*-операторов предложили в 1993 году Белнап и Перлоф (см. [8, р. 33]). Оно касается тех ситуаций, когда один агент не просто просит другого агента выполнить некоторое действие, но и указывает на те последствия, которые произойдут в случае невыполнения указанного. Пусть  $\alpha$  говорит  $\beta$ : «Или ты выполнишь  $P$ , или, если откажешься, то наступит крайне нежелательное  $Q$ ». В *stit*-подходе этому императивному дей-

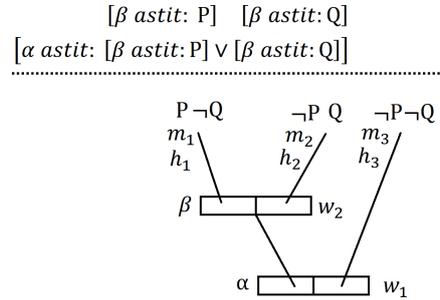


Рис. 6.

ствию может быть поставлена в соответствие следующая формула:  $[\alpha \text{ astit} : [\beta \text{ astit} : P] \vee [\beta \text{ astit} : Q]]$ . Попытка построить модель для этой формулы дает следующий результат: см. рис. 6. Здесь агенту  $\beta$  удастся отчасти сохранить свою агентность: ведь момент  $w_2$  является одновременно и моментом, в котором формула  $[\beta \text{ astit} : P] \vee [\beta \text{ astit} : Q]$  становится истинной в результате выбора, осуществленного агентом  $\alpha$  в  $w_1$ , и моментом, в котором  $\beta$  может осуществить выбор между  $P$  и  $Q$ , не нарушая при этом условий истинности всей дизъюнктивной подформулы рассматриваемой формулы.

Однако условия истинности формулы с модальным *astit*-оператором предписывают, в соответствии с п. 2 определения 2, необходимость наличия в отношении такой формулы истории и момента, где она могла бы быть ложной. И если применительно к формуле  $[\alpha \text{ astit} : [\beta \text{ astit} : P] \vee [\beta \text{ astit} : Q]]$  такой момент есть, то его нет в отношении  $[\beta \text{ astit} : P] \vee [\beta \text{ astit} : Q]$ , ведь вне зависимости от того, какую из историй выберет  $\beta$  в момент  $w_2$ , ни в одной из них формула  $[\beta \text{ astit} : P] \vee [\beta \text{ astit} : Q]$  не становится ложной. Именно поэтому мы и говорим лишь о частичном сохранении агентности в отношении  $\beta$ .

Эту «инвалидную» агентность можно было бы попытаться устранить через введение еще одного момента выбора для  $\beta$ , когда он сначала выбирал бы между возможностью принять альтернативу, предлагаемую ему агентом  $\alpha$ , и возможностью ее отклонить, а затем, в случае принятия альтернативы, делал бы выбор «внутри» нее (см. рис. 7). Такую же схему, но без указанных значений модальных и немодальных формул, мы можем найти и у Белнапа и Перлофа [8, р. 34, fig. 2]. Она действительно описывает ситуацию, когда  $\beta$  получает возможность отказать от предлагаемого ему выбора: ни  $[\beta \text{ astit} : P]$ , ни  $[\beta \text{ astit} : Q]$  не являются истинными в  $m_1/h_1$  и в  $m_2/h_2$  соответственно в отноше-



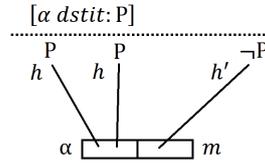


Рис. 8.

тернативных историй, проходящих через  $m$  и принадлежащих выбору  $\alpha$ .

2. Существует история  $h'$ , такая, что в ней, относительно  $m$ ,  $P$  ложна (см. рис. 8).

В данном случае в отличие от модели для *astit*-формулы истинность определяется уже не относительно пары моментов  $w$  и  $m$ , а только лишь относительно одного момента  $m$ . В одном и том же моменте  $m$  формула может быть и истинной, и ложной, в зависимости от того, относительно какой истории, проходящей через  $m$ , она оценивается. Заметим, что принцип работы *astit*-оператора фактически делает разговор об историях необязательным, так как каждая история заканчивается уникальным моментом и, следовательно, указав, что истинность некоторой формулы определяется относительно некоторого уникального момента и не сделав никакой отсылки к истории, мы не сказали бы ничего, что допускало бы различную истинностную оценку формулы. В остальном условия истинности для *dstit* повторяют условия истинности *astit*.

Что происходит, когда агент совершает действие  $[\alpha \text{ dstit} : P]$ ? Формула с *dstit*-оператором устроена так, что момент выбора и момент осуществления задуманного совпадают. Хорошо видно, что результат действия агента  $\alpha$  — истинность  $P$  — обнаруживает себя ровно в тот самый момент, когда  $\alpha$  решает выполнить  $P$ . Именно благодаря этому истинность  $[\alpha \text{ dstit} : P]$ , установленная относительно  $m/h$ , не является необходимой истиной, в отличие от истинности  $[\alpha \text{ astit} : P]$  (ведь если бы это было так, то не выполнялся бы п. 2 определения 3)<sup>7</sup>. Сравнение условий истинности *astit*- и *dstit*-формулы представлено на рис. 9. Квадратом обведен момент, в котором *astit*-формула принимает значение «истина», прямоугольником — момент и история, в которой *dstit*-формула принимает значение «истина». Покажем на примере некоторых логи-

<sup>7</sup>Подробнее об отличии *astit* от *dstit* в этом аспекте см. [11, p. 593].

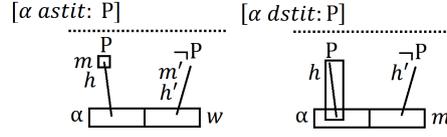


Рис. 9.

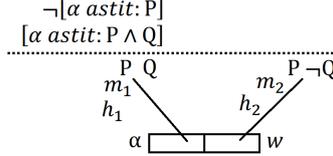


Рис. 10.

ческих принципов, как работают astit- и dstit-операторы в сравнении друг с другом. Их действие идентично, например, в случае принципа **M**:  $[\alpha stit : P \wedge Q] \rightarrow [\alpha stit : P] \wedge [\alpha stit : Q]$ , так как оба они позволяют построить такие модели, где будут истинны формулы  $[\alpha stit : P \wedge Q]$  и  $\neg[\alpha stit : P]$ , или  $[\alpha stit : P \wedge Q]$  и  $\neg[\alpha stit : Q]$ , что означает, что принцип **M** не подтверждается ни в отношении astit-, ни в отношении dstit-оператора.

В модели, представленной на рис. 10, формула  $[\alpha astit : P \wedge Q]$  истинна относительно  $m_1/h_1$ . Вместе с тем, относительно  $m_1/h_1$  ложной оказывается формула  $[\alpha astit : P]$ , так как формула  $P$  имеет место и в  $m_1/h_1$ , и в  $m_2/h_2$ . Набор значений для формул  $P$  и  $Q$  в  $m_2/h_2$ , наличие которого выполняет требования п. 2 определения 2 для формулы  $[\alpha astit : P \wedge Q]$ , не удовлетворяет требованию того же пункта того же определения в отношении формулы  $[\alpha astit : P]$ . Модель, демонстрирующая невыполнение принципа **M** в отношении dstit-формул, строится и обосновывается аналогично (см. рис. 11). В отношении принципа **SMP** (stit modus ponens)  $[\alpha stit : P \rightarrow Q] \wedge [\alpha stit : P] \rightarrow [\alpha stit : Q]$  astit- и

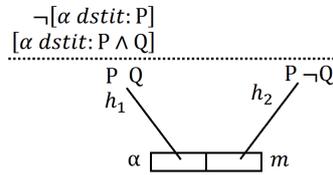


Рис. 11.

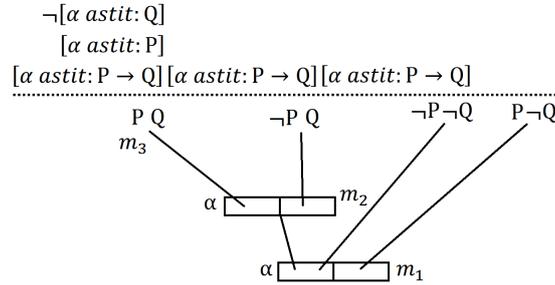


Рис. 12.

dstit-оператор ведут себя по-разному, исходя из чего можно ясно различить содержательные следствия тех формальных свойств, которыми они обладают.

Покажем некорректность принципа **SMP** в отношении astit-. Пусть формула  $P \rightarrow Q$  истинна в  $m_3$  на основании того выбора, который делает агент  $\alpha$  в  $m_1$  (см. рис. 12).

Тогда, множество историй, проходящих через  $m_1$ , разделяется на группы:  $K_1 = \{h_1, h_2, h_3\}$  и  $K_2 = \{h_4\}$ , где в выборе  $K_1$  оказываются все истории, в которых  $P \rightarrow Q$  истинна (и таких историй, очевидно, три), а в  $K_2$  — все истории, в которых  $P \rightarrow Q$  ложна. После того, как  $\alpha$  сделал выбор в  $m_1$ , он выбирает между  $P$  и  $\neg P$ . По условию, он выбирает  $P$ , и это действие он осуществляет в моменте  $m_2 > m_1$ . Таким образом, в  $m_3$   $\alpha$  получает, во-первых,  $P$ , на основании выбора, осуществленного в  $m_2$ , и, во-вторых, как мы уже сказали, —  $P \rightarrow Q$ , на основании выбора, осуществленного им в  $m_1$ . Легко убедиться в том, что для  $P$  и  $P \rightarrow Q$  в отношении  $m_3$  выполняются оба пункта определения 2. Однако формула  $Q$  не является истинной в  $m_3$ , так как в  $m_2$   $\alpha$  не может выбрать  $\neg Q$ , что означает, что в отношении  $Q$  он теряет свое свойство агентности, и  $Q$  наступает как бы само по себе. Агент  $\alpha$  не реализует  $Q$  непосредственно, на основании сознательного выбора, предшествующего моменту осуществления действия. Фактически  $\alpha$  лишается какого бы то ни было выбора в отношении  $Q$ , и происходит это из-за того, что между выбором в  $m_1$  и реализацией выбранного в  $m_3$  существует временной разрыв, в котором помещается действие по выбору  $P$  в  $m_2$ , что, в конце концов, и лишает  $\alpha$  возможности выбрать  $Q$ .

Действия, описываемые dstit-формулой, реализуются мгновенно, благодаря чему за агентом  $\alpha$  сохраняется возможность выбрать  $Q$  самостоятельно, а не иметь ее выполненной в результате логической необходимости.

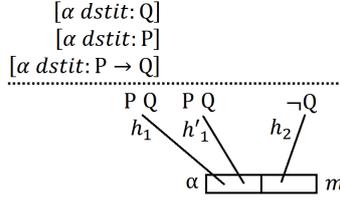


Рис. 13.

Белнап в [11, р. 599] предлагает следующую модель, на которой выполняются формулы  $[\alpha \text{ dstit} : P \rightarrow Q]$ ,  $[\alpha \text{ dstit} : P]$  и  $[\alpha \text{ dstit} : Q]$  (см. рис. 13). Пусть в  $m/h_1$  истинны  $[\alpha \text{ dstit} : P \rightarrow Q]$  и  $[\alpha \text{ dstit} : P]$ . Тогда, в соответствии с п. 1 определения 3, должна существовать (альтернативная по отношению к  $h_1$ ) история  $h'_1$ , принадлежащая множеству  $\mathbf{Choice}(\alpha, m)$ , такая, что в ней истинными оказываются формулы  $P \rightarrow Q$  и  $P$ . Следовательно, на основании правила modus ponens классической логики, в  $h'_1$  имеет место также и  $Q$ . Это, в свою очередь, должно означать, что  $Q$  имеет место также и во всех прочих историях, принадлежащих  $\mathbf{Choice}(\alpha, m)$  и, в частности, в истории  $h_1$ . Значит, в  $h_1$  истинна  $[\alpha \text{ dstit} : Q]$ , так как относительно этой формулы выполняется требование п. 1 определения 3. Требование п. 2 определения 3 также выполняется для  $[\alpha \text{ dstit} : Q]$ , так как в истории  $h_2$  присутствует формула  $\neg Q$ .

Заметим, что Белнап ничего не говорит о выполнении п. 2 определения 3 в отношении формулы  $[\alpha \text{ dstit} : P]$ , истинной в  $m/h_1$ . Если предположить, что п. 2 данного определения выполняет формула  $\neg P$ , истинная в  $m/h_2$ , то это нарушит условия истинности для  $[\alpha \text{ dstit} : P \rightarrow Q]$  в  $m/h_1$ . Модель, предложенная Белнапом, является неполной. Однако эту неполноту можно устранить, если например допустить, что множество  $\mathbf{Choice}(\alpha, m)$  образовано не парой элементов  $K_1 = \{h_1, h'_1\}$  и  $K_2 = \{h_2\}$ , а тройкой, в которой наряду с  $K_1$  и  $K_2$  присутствует также и  $K_3 = \{h_3\}$ , где  $h_3$  есть такая история, в которой истинной оказывается формула  $\neg P$ . Данное предположение позволяет сохранить условия истинности в модели относительно  $m/h_1$  и для  $[\alpha \text{ dstit} : P]$ , и для  $[\alpha \text{ dstit} : P \rightarrow Q]$ .

#### 4. Dstit-astit- и astit-dstit-

Astit- и dstit- операторы выражают разные действия разных агентов. Формула  $[\alpha \text{ astit} : P]$  указывает на то, что агент  $\alpha$  совершает на самом деле не одно, а два действия: сначала он делает выбор в моменте

$w$ , а затем реализует выбранную историю в моменте  $m$ , поступая так, что формула, описывающая выбранное им положение дел, становится истинной. Напротив, формула  $[\alpha \text{ dstit} : P]$  говорит о том, что агент  $\alpha$  совершает некоторое действие мгновенно, и в моменте  $m$  фиксируется только результат реализации выбранной им истории. Это различие имеет следствия как чисто формальные, накладывающие существенные ограничения на выразительные возможности языка, так и содержательные, позволяющие нам сделать одно важное предположение, непосредственно касающееся способа представления императива в stit-подходе.

Технически введение dstit- оборачивается потерей какого бы то ни было динамического аспекта в формальном языке. Его призрак еще был замечен, когда мы имели дело с astit-оператором: хотя в языке stit-логики и нельзя было описать само действие как таковое, выразительных возможностей хватало на то, чтобы говорить о действии-выборе и действии-реализации выбранной истории. Динамика обнаруживала себя в переходе агента от момента-выбора к моменту-осуществлению выбранного. В dstit-формулах такого перехода нет, а свое значение подформула astit-формулы, описывающая действие агента, получает мгновенно. Мы уже видели, как это свойство dstit-оператора сказывается на таких важных логических принципах, как принципы **M** и **SMP**. Однако не теряет ли dstit-формула, выражающая принцип **SMP**, слишком много для того, чтобы считаться формулой, описывающей действие? Одномomentное осуществление некоторого множества посылок, полное отсутствие возможности описывать взаимодействие агентов свидетельствуют о том, что «чистый» dstit- едва ли пригоден для записи императива.

Содержательные следствия сравнения astit- и dstit- убеждают нас в том, что последний оператор, конечно же, не является полностью бесполезным формальным инструментом, не годящимся для анализа императивов. По крайней мере в одном отношении dstit- мог бы сослужить нам хорошую службу. Дело в том, что использование dstit- в соединении с astit- могло бы позволить сохранить агентность исполнителя. Действительно, одним из свойств astit- является то, что пропозициональные формулы, входящие в его область действия, становятся истинными во всех историях и моментах, следующих за моментом выбора для агента из формулы astit-. И это же свойство, как мы видели, распространяется и на модальные подформулы astit-формул. Напротив, dstit-формула работает так, что ее подформула, в соответствии с определением 3, не должна быть необходимо истинной: dstit-агент в том самом моменте, где

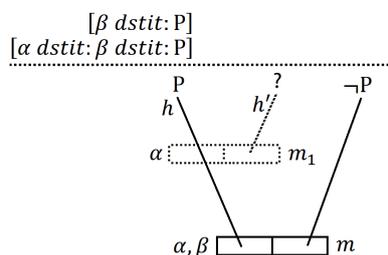


Рис. 14.

dstit-формула объявляется истинной, обладает возможностью выбрать отрицание подформулы dstit-формулы, не нарушая при этом условий истинности для dstit-. И именно этого свойства, в момент  $m$ , лишен astit-агент.

Данное различие Белнап предлагает фиксировать следующими формулами:  $[\alpha \text{ astit} : P] \rightarrow \Box P$  и  $[\alpha \text{ dstit} : P] \rightarrow \neg \Box P$ . (см. [11, р. 593]). Таким образом, вероятность использования dstit-оператора для выражения императивного действия сопряжена с положительным моментом сохранения агентности для возможных dstit-подформул astit-формул. В этом свойстве dstit-подформула в большей степени походит на императивное предложение, чем astit-формула, в отношении которой также должны выполняться некоторые постулаты, утверждающие свободу выбора агента. Негативное следствие в отношении dstit- заключается в том, что формулы с этим оператором не описывают взаимодействие агентов, так как для любой формулы, где в области действия dstit-оператора оказывается другой stit-оператор, построение модели чрезвычайно затруднено. Связано это с тем, что некоторые из определений модельной структуры  $\mathcal{S}$  перестают выполняться.

Так, например, если бы существовала модель для формулы  $[\alpha \text{ dstit} : [\beta \text{ dstit} : P]]$ , в которой она была бы истинна относительно  $m/h$ , то это означало бы, что, относительно той же пары момент/история, истинной должна быть формула  $[\beta \text{ dstit} : P]$  (см. рис. 14). Это в свою очередь влечет следствие, согласно которому истинная относительно  $m/h$  формула  $[\beta \text{ dstit} : P]$  дает истинную формулу  $P$ , — и снова в том же самом моменте, и в той же самой истории. Истинность  $P$  в  $m/h$  говорит нам о том, что агент  $\beta$  поступил в  $m$  так, что его действия гарантировали истинность  $P$ . Однако моменты действия агентов  $\alpha$  и  $\beta$  не могут совпадать. Эти моменты должны быть разнесены во времени, если формула  $[\alpha \text{ dstit} : [\beta \text{ dstit} : P]]$  действи-

тельно выражает императивное действие — например, приказ, который сначала отдается, и только после этого выполняется. Следовательно, мы вынуждены предположить наличие внутри истории  $h$  еще одного момента  $m_1$ , такого, что  $m < m_1$  и  $\mathbf{Choice}(\alpha, m_1) = \{h, h'\}$ , где  $h'$  означает альтернативную для  $\beta$  историю, выбрав которую он реализует свое право *dstit*-агента, заданное п. 2 определения 2. Но в таком случае история  $h$  перестает быть историей исходя из п. 3 определения 1, ведь она больше не является максимальной цепью. Предположив наличие модели у формулы  $[\alpha \textit{dstit} : [\beta \textit{dstit} : P]]$ , мы столкнулись с ограничениями на введение подобных формул, основанных на определениях модельной структуры<sup>8</sup>.

Аналогично складывается ситуация и в отношении формулы  $[\alpha \textit{dstit} : [\beta \textit{astit} : P]]$ , где структура формулы, требующая наличие пары моментов, также оказывается несовместимой с п. 3 определения 1. Но совершенно по-другому ведет себя формула  $[\alpha \textit{astit} : [\beta \textit{dstit} : P]]$ .

Здесь необходимо еще раз напомнить о той роли, которую играют истории в определении истинности *astit*- и *dstit*-формул. Выше мы уже упоминали о том, что отсылка к истории в случае определения условий истинности *astit*-формул является излишней, а потому и необязательной — различий в том случае, если мы не сошлемся на конкретную историю, в которой истинна *astit*-формула, не возникает. Подтверждение этому можно найти в статье [11, р. 593]: “Несмотря на то, что индекс, относительно которого оценивается *astit*-формула, содержит момент и историю, история упоминается в нем только для удобства и в целях соблюдения семантического единообразия. Истории бесполезны в семантических определениях (для *astit*-)”<sup>9</sup>. Это действительно так, однако противоречия могут иметь место там, где мы ссылаемся на историю, относительно которой определяется истинность такой *astit*-формулы, которая имеет в качестве подформулы *dstit*-. И ситуация с формулой  $[\alpha \textit{astit} : [\beta \textit{dstit} : P]]$  как раз одна из таких.

<sup>8</sup>Заметим, что определение истории, с нашей точки зрения, само по себе, вне зависимости от стремления использовать *stit*-подход для выражения императивов, нуждается в корректировке: ведь под историей понимается последовательность моментов во времени, однако в отношении моделей *dstit*-формул все ограничивается одним-единственным моментом  $m$ . Это существенное ограничение, затрудняющее построение модели для *dstit*-формул, имеющих отношение к нескольким агентам, вероятно, нужно будет снимать и в том случае, если мы хотим работать со сложными формулами, пусть и не имеющими отношения к нескольким агентам, но все же начинающимися с *dstit*-оператора, или имеющими *dstit*- где-то «в середине».

<sup>9</sup>“Для *astit*-” добавлено в конце предложения для ясности (Пер. мой — Г.К.).

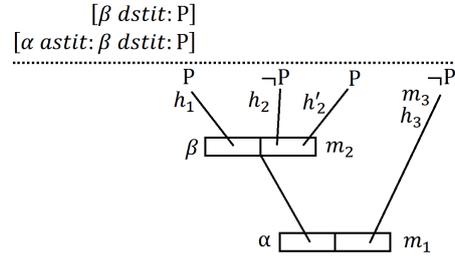


Рис. 15.

Осуществляя до настоящего времени привычную ссылку на момент и историю, предположим, что эта формула истинна относительно  $m_2/h_1$ , см. рис. 15. Тогда для агента  $\alpha$  должен найтись такой момент  $m_1$ , предшествующий  $m_2$ , когда он принял решение о том, что формула  $[\beta \text{ dstit} : P]$  истинна в  $m_2/h_1$  (в соответствии с п. 1 определения 2). Решение действовать так, чтобы формула  $[\beta \text{ dstit} : P]$  была истинной в  $m_2$ , принятое  $\alpha$  в  $m_1$ , условно может быть интерпретировано как решение об осуществлении императивного действия: агент  $\alpha$  делает так, что агент  $\beta$  оказывается в таких условиях, где он должен выбирать между осуществлением  $P$  и отказом от осуществления  $P$ . При этом формула  $P$  отождествляется с желаемым исходом для  $\alpha$ , с тем положением дел, реализации которого он ожидает от  $\beta$ . Благодаря действию агента  $\alpha$ , осуществленному в моменте  $m_1$ , перед агентом  $\beta$  в моменте  $m_2$  открывается выбор между осуществлением  $P$  и воздержанием от осуществления  $P$ . Заметим, что само наличие этого выбора еще не делает формулу  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  ложной. Напротив, альтернатива  $\neg P$  в  $m_2/h_2$  не только не противоречит истинности  $[\beta \text{ dstit} : P]$  в  $m_2/h_1$  (а значит и истинности  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  там же) — она необходима и для того, чтобы  $[\beta \text{ dstit} : P]$  была истинной в  $m_2/h_1$ , так как через это реализуется п. 2 определения 3, и для того, чтобы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  была истинной в  $m_2/h_1$ , так как это говорит о соблюдении требований п. 2 определения 2. Если бы такой альтернативы у  $\beta$  не было, если бы не было истории  $h_2$ , или если бы в  $h_2$  имела место формула  $P$ , то  $[\beta \text{ dstit} : P]$ , как результат предпринятого ранее действия агента  $\alpha$ , не могла бы быть истинной в  $m_2/h_1$ , так как в этом случае агент  $\beta$  не принимал бы решения о том, как ему поступить, а агент  $\alpha$  отдавал бы «вырожденное» указание, типа «Подчиняйся закону всемирного тяготения!».

Таким образом, с помощью формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , истинной в  $m_2/h_1$ , выражается императивное действие, в котором  $\alpha$  отдает некоторое недифференцированное указание, а  $\beta$  его выполняет. То, что  $\alpha$  отдает указание на выполнение именно  $P$ , следует, естественно, из самой рассмотренной формулы; то, что  $\beta$  реализует  $P$ , следует из того, что истинность формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  задается относительно  $m_2/h_1$ , т.е. тех истории и момента, где формула  $P$  оказывается истинной.

Но зададимся вопросом о том, может ли формула  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  оставаться истинной в том случае, когда агент  $\beta$  выбирает  $\neg P$ ? Утвердительный ответ на этот вопрос будет означать, что данная формула выражает осуществленное императивное действие агента  $\alpha$ , связанное с побуждением агента  $\beta$  к реализации некоторого действия  $P$ , запрошенного агентом  $\alpha$ , исход которого неизвестен и определяется либо отсылкой к истории, где формула  $P$  истинна, либо отсылкой к истории, где формула  $\neg P$  истинна. Отрицательный ответ на вопрос покажет, что на самом деле под видом  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  нам приходится иметь дело с не до конца корректным представлением императива, так как в этом случае, несмотря на то, что модельная структура сохраняет за агентом  $\beta$  возможность выбрать  $\neg P$  в  $m/h_2$  и тем самым отказаться от выполнения того действия, к которому принуждает его агент  $\alpha$ , — эта возможность, в случае ее реализации, делает *astit*-формулу агента  $\alpha$  ложной, что равносильно неосуществлению императивного действия. И здесь упомянутая выше необязательная отсылка к истории в случае с *astit*-формулой играет решающую роль, так как мы действительно получаем этот отрицательный ответ, когда оцениваем истинность формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  относительно момента  $u$  истории.

В самом деле, если формула  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  претендует на то, чтобы служить корректным выражением императивного действия, то она, даже в том случае, если агент  $\alpha$  предпочтет воздержаться от осуществления действия  $P$ , не должна становиться ложной ни в одной из возможных историй, примыкающих к моменту  $m_2$ , ведь в таком случае это означало бы то, что приказ вообще не был осуществлен — формулу  $\neg[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  следует понимать именно так. Если же предположить, что, несмотря на выбор истории  $h_2$  агентом  $\beta$  в моменте  $m_2$ , формула  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  остается истинной относительно  $m_2/h_2$ , то это будет противоречить п. 1 определения 2. Действительно,  $\neg P$  в  $m_2/h_2$  дает, в модели на рис. 15, формулу  $\neg[\beta \text{ dstit} : P]$  в  $m_2/h_2$ , которая, в свою очередь, влечет формулу  $\neg[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  в  $m_2/h_2$ , что

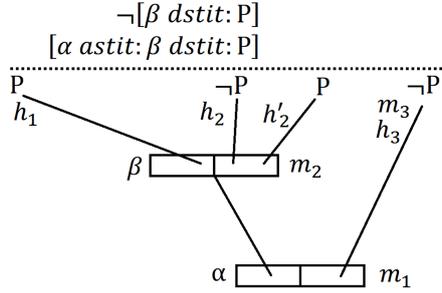


Рис. 16.

противоречит начальному допущению о выборе агента  $\beta$  и сохранении при этом значения *astit*-формулы, выражающей приказ (см. рис. 16). Однако если задавать истинность формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  только на основании ее соотнесения с моментом  $m_2$ , то противоречие между ее значением и выбором агентом  $\beta$  формулы  $\neg P$  в  $m_2/h_2$  тем самым удастся снять.

Предположим, что  $\alpha$  осуществляет императивное действие, выраженное формулой  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , истинной относительно момента  $m_2$ . Тогда *dstit*-подформула  $[\beta \text{ dstit} : P]$  *astit*-формулы в соответствии с п. 1 определения 2 также истинна в  $m_2$ , при том, что это значение в моменте  $m_2$  данная формула будет сохранять вне зависимости от того, как поступит  $\beta$  — реализует он  $P$  или нет. Если  $\beta$  выполнит указание  $\alpha$ , то тем самым формула  $[\beta \text{ dstit} : P]$  как *dstit*-формула станет истинной в  $m_2$  и относительно истории  $h_1$ ; если  $\beta$  не выполнит указание  $\alpha$ , то формула  $[\beta \text{ dstit} : P]$  как *dstit*-формула станет ложной относительно момента  $m_2$  и истории  $h_2$ , но не относительно  $m_2$ .

Таким образом, формула  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , истинность которой задается относительно момента  $m$ , не выражает «обреченный на успешность» императив, как это происходит в случае с формулой  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , истинность которой задается относительно момента  $m$  и истории  $h$ , — она представляет собой, на наш взгляд, корректное отображение в формальном языке ситуации осуществляемого указания, которое может быть исполнено или не исполнено.

Это подтверждается в том числе и следующими соображениями. Как мы указывали ранее, следуя Белнапу, верно, что если имеет место *astit*-формула, то любая ее подформула необходимо истинна:  $[\alpha \text{ astit} : P] \rightarrow \Box P$ . Тогда  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  влечет  $\Box[\beta \text{ dstit} : P]$ . Вместе с тем также верно, что *dstit*-оператор не дает в качестве своего следствия

необходимую истинность своих подформул:  $[\alpha \text{ dstit} : P] \rightarrow \neg \Box P$ . Тогда  $\Box[\beta \text{ dstit} : P]$  влечет формулу  $\Box \neg \Box P$ , которая равносильна формуле  $\Diamond P$ . Таким образом,  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  гипотетически дает  $\Diamond P$ , что означает, что осуществление указания к исполнению  $P$  предполагает лишь возможность такого положения дел, что  $P$ , и даже отсутствие  $P$ , после того как указание было выполнено, не отрицает  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , т.е. не отрицает того факта, что это указание было осуществлено<sup>10</sup>.

Еще раз отметим, что истинность формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , как и любой другой *astit*-формулы, имеющей в качестве подформулы *dstit*-формулу, устанавливается только относительно момента  $m$ , в данном случае — относительно  $m_2$ . Будем считать ошибкой всякую попытку определить истинность этой формулы в отношении момента  $m_2$  и некоторой истории  $h_1$  или  $h_2$  на том основании, что данные истории являются эквивалентными для агента  $\alpha$  в моменте  $m_1$ :  $h_1 \equiv_{m_1}^{\alpha} h_2$ . (Это же замечание справедливо и применительно к другим *astit*-формулам, где в качестве подформул фигурируют *dstit*-формулы). Так как в отношении эквивалентных историй  $h_1$  и  $h_2$  агент  $\alpha$  не может сделать выбор, то он, начиная с этого момента, перестает нести ответственность за истинность подформулы *astit*-формулы, выражающей императивное действие, которая, будучи истинной, выражает осуществленный, но неизвестно, выполненный или нет, приказ.

Напротив, определение истинности *dstit*-подформулы осуществляется посредством соотнесения ее с моментом и историей, что сообразуется, во-первых, с определением 3, и, во-вторых, с тем, что истории  $h_1$  и  $h_2$  не являются эквивалентными для агента  $\beta$  в моменте  $m_2$ :  $h_1 \not\equiv_{m_2}^{\beta} h_2$ . То, что подформула  $[\beta \text{ dstit} : P]$  истинна в  $m_2/h_1$ , свидетельствует о том, что  $\beta$  выполнил приказ  $\alpha$ ; истинность  $\neg[\beta \text{ dstit} : P]$  в  $m_2/h_2$  (ложность  $[\beta \text{ dstit} : P]$  в  $m_2/h_2$ ) говорит о том, что  $\beta$  отказался от выполнения приказа. Определение истинности для  $[\beta \text{ dstit} : P]$  на основании ее соотнесения только с моментом  $m_2$  является следствием п. 1 определения 2, когда  $[\beta \text{ dstit} : P]$  рассматривается как подформула *astit*-

<sup>10</sup>То, что промежуточным звеном этого рассуждения выступает формула  $\Box[\beta \text{ dstit} : P]$ , видится нам естественным, несмотря на то что у классиков *stit*-подхода можно встретить такие формулы, как  $\Diamond[\alpha \text{ dstit} : P]$  [10, р. 25] и  $\neg \Box[\alpha \text{ dstit} : P]$  [11, р. 593], претендующие на то, что с их помощью выражается существенное логическое свойство *dstit*-оператора. Формула  $\Box[\beta \text{ dstit} : P]$  указывает на то, что агент  $\beta$  поставлен перед необходимостью принять решение о  $P$ . После получения указания он вынужден принимать решение: действовать или воздержаться от действия. Какое-то решение должно быть принято, — и это единственная необходимость, о которой говорит формула  $\Box[\beta \text{ dstit} : P]$ .

формулы и описывает ситуацию необходимости сделать выбор для  $\beta$ . Определение истинности для  $[\beta \text{ dstit} : P]$  на основании ее соотнесения с моментом  $m_2$  и историями  $h_1$  или  $h_2$  является следствием п. 1 определения 3, когда  $[\beta \text{ dstit} : P]$  рассматривается как формула, описывающая действие, осуществленное  $\beta$ , или отсутствие действия со стороны  $\beta$ .

## 5. Заключительные замечания

На основании вышеизложенного мы считаем формулу  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$  наиболее корректным способом записи императивного действия в stit-подходе. С ее помощью в модели  $\mathcal{M}$  удастся отразить такие свойства императива, как биагентность, которая предполагает выполнение аналога «критерия К» в отношении всех агентов, участвующих в императивном действии, незнание агента-адресата о конечном результате осуществляемого им императива и свойства «быть осуществленным» и «быть выполненным».

Stit-подход, примененный к императивам, позволяет заново ставить и решать вопросы, возникающие всякий раз, когда логическое исследование императивов начинает вестись с оглядкой на аналогию с классической логикой или логикой алетических модальностей. Источником парадокса Россса, о котором мы упоминали во введении, является, вероятно, убеждение, что в отношении императивов как модальных высказываний должно выполняться правило аналогичное правилу **DR1** алетической модальной логики:  $\vdash P \rightarrow Q$  влечет  $\vdash \Box P \rightarrow \Box Q$ . В отношении stit-формул, в том числе тех, что выражают императивы, это означает справедливость следующего утверждения: если  $P \rightarrow Q$  является законом классической логики высказываний, то  $[\alpha \text{ stit} : P] \rightarrow [\alpha \text{ stit} : Q]$  является законом stit-логики. Однако нетрудно построить модель, опровергающую истинность  $[\alpha \text{ stit} : P] \rightarrow [\alpha \text{ stit} : Q]$  при том что  $P \rightarrow Q$  в данной модели будет истинна. Заметим также, что доказательство самого правила **DR1** основывается на принципе **K**, аналога которому нет в stit-логике, так как неверно, что  $[\alpha \text{ stit} : P \rightarrow Q]$  влечет  $[\alpha \text{ stit} : P]$  и  $[\alpha \text{ stit} : Q]$ . В свете сказанного нет необходимости демонстрировать неприемлемость формулы  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]] \rightarrow [\alpha \text{ astit} : [\beta \text{ dstit} : P \vee Q]]$ , выражающей в stit-подходе парадокс Россса и являющейся, конечно же, правилом введения дизъюнкции классической логики высказываний, снабженным stit-операторами. Мы предлагаем интересующемуся читателю самостоятельно построить модель, в которой был бы истинен антецедент данной формулы и ложен ее консеквент. Формула  $[\alpha \text{ astit} : [\beta \text{ dstit} : P]]$ , вместе с условиями ее истинно-

сти в модели  $\mathcal{M}$ , выражает описательную гипотезу о функционировании императивного действия. Настоящим исследованием обосновывается отсутствие в этой гипотезе внутренних противоречий. Stit-подход, предоставивший методологические, инструментальные принципы для изучения императивов, тем самым позволил заложить основание для содержательной stit-теории императивных действий. Дальнейшее исследование в этом направлении, очевидно, должно ставить перед собой задачу превращения этой содержательной теории в теорию формализованную, подобно тому, как это произошло в случае с «чистой» stit-логикой, когда сначала в 1998, а затем в 2008 году (см. [13] и [12]) были получены аксиоматические системы для dstit-предложений. Мы надеемся, что подобный путь удастся пройти и stit-логике для императивов.

### Литература

- [1] *Блинов А.Л., Петров В.В.* Элементы логики действий. М.: Наука, 1991. 232 с.
- [2] *Карпов Г.В.* Правила вывода в императивной логике // Логико-философские штудии. 2013. Вып. 11. № 2. С. 5–18.
- [3] *Kenny A. J.* Practical Inference // Analysis. 1966. Vol. 26. № 3. P. 65–75.
- [4] *Chellas B. F.* Time and Modality in the Logic of Agency // Studia Logica. 1992. Vol. 51. № 3/4. P. 485–517.
- [5] *Belnap N.* Backwards and Forwards in the Modal Logic of Agency // Philosophy and Phenomenological Research. 1991. Vol. 51. № 4. P. 777–807.
- [6] *Belnap N.* Before Refraining: Concepts for Agency // Erkenntnis. 1991. Vol. 34. № 2. P. 137–169.
- [7] *Belnap N., Perloff M.* The Way of the Agent // Studia Logica. 1992. Vol. 51. № 3/4. P. 463–484.
- [8] *Belnap N., Perloff M.* In the Realm of Agents // Annals of Mathematics and Artificial Intelligence 9. 1993. P. 25–48.
- [9] *Horty J. F.* An alternative stit operator. Maryland: Manuscript, 1989.
- [10] *Horty J. F.* Agency and Deontic Logic. Oxford, 2001. 208 p.
- [11] *Horty J. F., Belnap N.* The Deliberative Stit: A Study of Action, Omission, Ability, and Obligation // Journal of Philosophical Logic. 1995. Vol. 24. № 6. P. 583–644.
- [12] *Balbani P., Herzig A., Troquard N.* Alternative axiomatics and complexity of deliberative stit theories // Journal of Philosophical Logic. 2008. Vol. 37. № 4. P. 387–406.
- [13] *Xu M.* Axioms for Deliberative Stit // Journal of Philosophical Logic. 1998. Vol. 27. № 5. P. 505–552.

G.V. KARPOV

## Stit-logic for Imperatives

**Karpov Gleb Victorovich**

Department of Logic, Institute of Philosophy,  
Saint-Petersburg State University.

5 Mendeleevskaya Liniya, St. Petersburg, 199034, Russian Federation.

E-mail: glebsight@gmail.com

In this paper the possibility of interpreting imperatives as sentences that are used by rational agents to impel or motivate other rational agents to act in a desired way is investigated by the author. We claim that such impelling strategies, that ground practical reasoning, can be pictured formally by means of stit-logic. We introduce the basic semantical ideas of stit-logic and discuss the most acceptable way of formal representation of imperative and its corresponding impelling effect. Special attention is paid to embedded stit-formulas. We demonstrate that such formulas cannot serve as an appropriate way of imperative formalization since an agent-addressee lacks the possibility of choice, and the agent-sender becomes “omnipotent” about the future course of agent-addressee’s action. We show that this unwanted situation can be eliminated, and basic imperative properties can still be expressed with embedded stit-formulas, if different kinds of indexes are used in semantic definitions for different kind of stit-operators. Finally we put an assertion that using such formulas, evaluated with respect to different kind of indexes, leaves no room for usual paradoxes of imperative logic.

*Keywords:* stit-logic, imperative, logic of imperatives, logic of action

### References

- [1] Blinov, A.L., Petrov, V.V. *Elementy logiki deistvii*. M.: Nauka, 1991. 232 pp. (In Russian)
- [2] Karpov, V.G. “Pravila vyvoda v imperativnoi logike”, *Logiko-filosofskie shtudii*, 2013, vol.11. no 2, pp. 5–18. (In Russian)
- [3] Kenny, A. J. “Practical Inference”, *Analysis*, 1966, vol. 26, no 3, pp. 65–75.
- [4] Chellas, B. F. “Time and Modality in the Logic of Agency”, *Studia Logica*, 1992, vol. 51, no 3/4, pp. 485–517.
- [5] Belnap, N. “Backwards and Forwards in the Modal Logic of Agency”, *Philosophy and Phenomenological Research*, 1991, vol. 51, no 4, pp. 777–807.
- [6] Belnap, N. “Before Refraining: Concepts for Agency”, *Erkenntnis*, 1991, vol. 34, no 2, pp. 137–169.
- [7] Belnap, N., Perloff, M. “The Way of the Agent”, *Studia Logica*, 1992, vol. 51, no 3/4, pp. 463–484.
- [8] Belnap, N., Perloff, M. “In the Realm of Agents”, *Annals of Mathematics and Artificial Intelligence* 9, 1993, pp. 25–48.

- [9] Horty, J. F. *An alternative stit operator*. Maryland: Manuscript, 1989.
- [10] Horty, J. F. *Agency and Deontic Logic*. Oxford, 2001. 208 pp.
- [11] Horty, J. F., Belnap, N. “The Deliberative Stit: A Study of Action, Omission, Ability, and Obligation”, *Journal of Philosophical Logic*, 1995, vol. 24, no 6, pp. 583–644.
- [12] Balbiani, P., Herzig, A., Troquard, N. “Alternative axiomatics and complexity of deliberative stit theories”, *Journal of Philosophical Logic*, 2008, vol. 37, no 4, pp. 387–406.
- [13] Xu, M. “Axioms for Deliberative Stit”, *Journal of Philosophical Logic*, 1998, vol. 27, no 5, pp. 505–552.