

*Глеб Карпов*¹

ДЕЛЕГИРОВАНИЕ ОБЯЗАТЕЛЬСТВ И АВТОРИТЕТНЫЕ ОТНОШЕНИЯ АГЕНТОВ В ЛОГИКЕ ДЕЙСТВИЙ²

Резюме: в статье, в рамках традиции логики действия, поставлена проблема делегирования указаний, порождающих обязательства агентов друг перед другом. Дан обзор существующих семантик логики действия и указаны их особенности, существенные для интерпретации взаимодействия агентов во времени, задающего общие условия делегирования. Построена семантическая модель, которая отражает связывающие агентов авторитетные отношения, рассматриваемые как следствия обязательств, порожденных отданными указаниями. Предложена интерпретация авторитетного отношения как возможности отдать указание и изложен вариант решения проблемы делегирования, данный через описание условий существования транзитивности авторитетного отношения на тройке агентов. Дан ряд замечаний относительно свойств структуры авторитетных отношений, зависящих от свойства агентов, связанных такими отношениями.

Ключевые слова: логика действий, делегирование, авторитетные отношения, NEXT-семантика.

Gleb Karpov

THE DELEGATION OF OBLIGATIONS AND THE AUTHORITY RELATIONS IN THE LOGIC OF ACTION

Abstract: The article sets up a problem of directives and obligations that are produced by them delegation within a framework of the logic of action. We give a review of the logic of action semantics and point out their features that are substantial for the interpretation of agentive interaction in time, that set the basic conditions for delegation. We built a semantic model of authority relations between agents that are regarded as the consequences of obligations generated by those agents directives. We present an interpretation of an authority relation as an ability

¹Карпов Глеб Викторович, кандидат философских наук, ассистент, кафедра логики, Санкт-Петербургский государственный университет.

Gleb Karpov, PhD, assistant professor, department of logic, Saint Petersburg State University.

²Исследование проведено в рамках проекта РГНФ 14-03-00650.

Автор благодарит анонимного рецензента за ряд замечаний, проработка которых позволила значительно улучшить текст статьи.

to provide a directive and thus formulate a candidate solution of the directives delegation problem via the description of the transitive authority relation on a set of three agents. Some comments on the properties of an authority structure that depend on the properties of agent under such an authority are given in the end.

Keywords: logic of action, delegation, authority relation, NEXT-semantics.

Для чего нужна логика действий?

Когда Хинтиikka объясняет для чего нужна первопорядковая логика, то он обращает внимание на то, что ее язык чрезвычайно удобен, чтобы передавать друг о друге сплетни. Даже если вы ничего не знаете о первопорядковой логике, то отсюда по крайности вам становится ясно, что она нужна хотя бы и для записи такого рода фактов: Берти любит тех и только тех людей, которые не любят самих себя; и, сверх того, — для записи и для сопоставления таких фактов друг с другом; наконец, вам становится ясно, что она нужна для ведения рассуждений, которые строятся на основе подобных фактов [Хинтиikka 2014: 20, 23]. Предположим, что нам нужно нечто большее, что умение строить рассуждения о свойствах людей и отношениях, которые их связывают, кажется нам недостаточным. Пусть требуется выяснить что-то об их поведении, рассказать о тех действиях, какие они совершают и порассуждать о том, что из таких действий получается.

К счастью, ученые уже давно позаботилась и об этом; о поступках мы можем рассуждать с помощью специально приспособленной для этих целей логической теории. Людей здесь принято называть *агентами*, а их поступки — *действиями*; сама теория носит название логики действия.³ Агенты совершают действия, побуждают разными средствами, в том числе с помощью слов, к совершению действий других агентов. В этой статье мы будем рассуждать о действиях двух видов: *прямых* — например, когда агент открывает окно, и *непрямых*, когда, например, один агент просит открыть окно другого агента. Эти последние будем называть *императивными действиями*, так как они реализуются, как правило, через произнесения агентом *императивного предложения*, например такого: “Послушайте, прикройте окно, пожалуйста”.

³Подробнее о логике действия, ее истории и современных направлениях (STIT-логике и динамической логике) — см. [Segeberg, Meyer, Kracht 2016].

Если с помощью этой логики мы захотим выявить логическую форму прямого действия агента α , приводящего к такому положению дел, что ϕ , то мы сможем воспользоваться для этого, например, такой формулой: $[\alpha]\phi$. Ее следует читать так: агент α делает так, что ϕ истинна. Тогда для записи непрямого действия, выраженного императивным предложением, можно прибегнуть к суперпозиции таких формул. В таком случае записи не прямых действий могут образовывать довольно длинные цепи. Вот, например, цепочка из трёх действий (двух не прямых и одного прямого): Ахиллес просит Агамемнона, чтобы тот приказал ахейским воинам заняться приготовлением еды в лагере. Формула логики действий, отражающая этот факт, выглядит так: “[Ахиллес] [Агамемнон] [ахейские воины] (приготовление еды в лагере началось)”.

Случай госпожи Октав, Франсуазы и судомойки

Теперь представим г-жу Октав, которая просит свою служанку Франсуазу приготовить на обед цыплёнок с овощами. Она говорит Франсуазе, возможно, отвечая той на вопрос, что готовить на обед: “На обед у нас будет цыплёнок с овощами”. Мы более чем уверены, что это высказывание г-жи Октав Франсуаза понимает как указание сделать так, чтобы к обеду на стол был подан цыплёнок и овощи. Франсуаза отправляется на кухню и, в свою очередь, отдаёт указание судомойке, чтобы та принималась за чистку и приготовление овощей (спаржи). Франсуаза делегирует передаёт только часть указания г-жи Октав — ту, что имеет отношение к овощам; приготовлением цыплёнка она решает заняться сама.

Нам бы хотелось знать, в связи с произошедшим, следующее: поступает ли в данном случае г-жа Октав так, как если бы она отдавала указание судомойке чистить и варить спаржу непосредственно? Будем ссылаться на этот вопрос как на вопрос о перепоручении или делегировании, а соответствующее явление называть собственно делегированием или перепоручением — указания, команды, просьбы и т. д.

Ответ на этот вопрос можно было бы попытаться найти прямо сейчас, руководствуясь в поиске интуицией или субъективным представлением о социальной норме, регулирующей поведение агентов из примера. Однако, традиция исследований, к которой принадлежит логика действий, является преимущественно формальной: здесь ученые стремились к созданию не только содержательной, но и формализованной

(аксиоматической) теории действия. Последнее стремление реализовано в полной мере в работах Минга Сюй [Xu 1998], Филиппа Бальбиани и др. [Balbiani, Herzig, Troquard 2008] и Яна Брурсена [Broersen 2009] и [Broersen 2011]. Направление логики действий, опирающееся на неформальные методы, пик популярности которого пришелся на середину 60-х годов прошлого века (работы Гектора Кастанеды [Castaneda 1960], Питера Гича [Geach 1963], Энтони Кенни [Kennedy 1966]) сегодня по всей видимости окончательно сошло на нет.

Формальная традиция, в самом общем смысле, предполагает, что истинность содержательного предложения можно трактовать как истинность формулы, соотносимой с данным предложением. Следовательно, вопрос о делегировании следует сначала представить в формализованном языке. И далее, — установить условия истинности полученных формул с помощью моделирования. Для этого нам потребуется логическая семантика, задающая интерпретацию формул, и тем самым растолковывающая их смысл. В логике действий используются или семантика ветвящегося времени (BTS, Branching Time Semantics) [Belnap 1991], [Belnap 1991a], или семантика, основанная на модельных структурах Крипке и аксиоме произведения (product axiom) — СКФ (Choice Kripke Frame) [Benthem, Pasuit 2014], или NEXT-семантика [Broersen 2011]. Таким образом, план нашего исследования, в целом, уже определен той традицией, к которой принадлежит исследование. Мы позволим себе предпослать неизбежному вопросу о формализации непрямого действия параграф, разъясняющий, без погружения в сугубо технические детали, семантические и философские основы проводимой работы, и более того — служащий введением в семантику логики действий вообще.

Обзор семантик логики действий. Семантика ветвящегося времени (BTS)

Самой распространённой семантикой, которая используется в логике действия, до сегодняшнего дня остаётся семантика ветвящегося времени (BTS). Для интерпретации формул в BTS прибегают к ветвящимся модельным структурам, подобным той, что представлена на рисунке 1.

Точки здесь обозначают моменты времени, а линии, связывающие моменты времени, — истории. История — это вариант развития событий

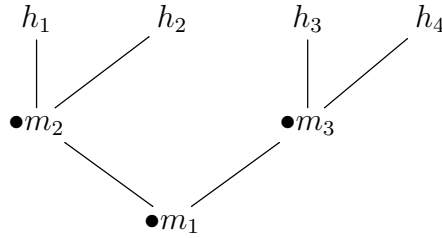


Рис. 1: Древовидная модельная структура BTS.

в моменте, для которого уже нет момента более раннего, чем он сам.⁴ Деревья, формируемые моментами времени и историями, ветвящимися вверх, иллюстрируют философскую основу BTS: единственность прошлого и открытость, потенциальную множественность, будущего.

Оригинальная идея, оправдывающая использование BTS в логике действий, заключалась в том, что действуя, агент, в некоторый момент времени, делает так, что в некотором последующем моменте или моментах (связанных с настоящим историями) истинной становится некоторая формула. Например, агент α , в момент времени t_1 может сделать так, что в момент времени t_2 (и во всех историях, проходящих через t_2 , т. е. в h_1 и в h_2) истинной будет формула ϕ . Если это верно, то говорят, что формула $[\alpha]\phi$ истинна в t_1 (точнее: в t_1/h_1 или в t_1/h_2).

Обычно эту идею сопровождают следующие замечания. Первое: если формула ϕ уже имеется в t_2 и t_3 , то мы не можем сказать, что $[\alpha]\phi$ истинна в t_1/h_1 или в t_1/h_2 . (Действительно, в чем тогда заключается действие агента α в t_1 , если во всех последующих моментах уже имеется ϕ ? Что бы α в t_1 не выбрал, все равно, всегда будет так, что ϕ . В таких случаях говорят, что α не может — не имеет возможности — совершить действие, результатом которого будет ϕ .) Второе замечание: действуя так, что ϕ истинно в t_2/h_1 или в t_2/h_2 агент α , вместе с тем, в момент t_1 не может сделать так, что ϕ будет истинно только в t_2/h_1 или только в t_2/h_2 . Это верно, так как обе истории h_1 и h_2 в моменте t_1 , там, где α делает выбор, как поступить, не разделены. Разделенными в t_1 являются классы историй $\{h_1, h_2\}$ и $\{h_3, h_4\}$. Ясно, что выбор агента в некотором моменте подразумевает разбиение множества историй, проходящих через этот момент, и определяется самой модельной структурой,

⁴На рис. 1 следовательно представлено четыре истории h_1, h_2, h_3 и h_4 , каждая из которых имеет в качестве начального момента момент t_1 .

самим устройством дерева. Далее мы будем пользоваться словами выбор агента и разбиение (историй в моменте) для обозначения одного и того же.

Отвлекаясь от этих замечаний общего, философского характера, скажем, что условия истинности в логике действий традиционно (начиная с середины 90-х годов прошлого века и до настоящего времени) задаются так: $m_1/h_1 \models [\alpha]\phi$ е. т. е. $m_1/h' \models \phi$, для всех историй h' , принадлежащих тому же разбиению, что и история h_1 .

Интересно обратить внимание на то, что при таком определении в деревьях, включающих последовательности моментов, просто не возникает необходимости! В такой последовательности нет необходимости, ведь всё ограничивается одним моментом и множеством историй, проходящих через него. Единственный случай, когда в действительности используется последовательность моментов, — белнаповское определение действия, задающееся через специальный оператор действия⁵. Деревья, следовательно, используются в логике действий больше для сохранения семантического единообразия определений и в значительной части случаев, там где не применяется оператор Белнапа (то есть в большинстве случаев), попросту не нужны. Другими словами, их применение — дань традиции, а не необходимость, продиктованная требованиями содержательного или формального характера той или иной проблемы.

Модельные структуры Крипке в семантике логики действий

Если мы сосредоточим наше внимание на единственном моменте и проходящих через него историях, то мы увидим, что такие истории есть не что иное, как возможные миры, разбитые действием агента на классы эквивалентности. Такой взгляд предполагает переход от ветвящихся структур к реляционным структурам. Если действуют несколько агентов, то, в соответствии с аксиомой произведения⁶, их действия должны

⁵Речь идет об операторе действия *achievement stit* и определениях Нуэля Белнапа, предложенных для него. В этом случае единственно и нужны деревья — как частично упорядоченные множества моментов времени. Подробнее об этом см. работу самого Белнапа [Belnap 1991a], ставшую классикой логики действия.

⁶Аксиома произведения (product axiom) утверждает, что если существует выбор агента α , такой, что ϕ , и если существует выбор агента β , такой, что ψ , то существует такой мир, где истинно, что $\phi \wedge \psi$: $(E[\alpha]\phi \wedge E[\beta]\psi) \rightarrow E(\phi \wedge \psi)$. Подробнее об аксиоме выбора и ее использовании в логике действий — см. [Benthem, Pacuit 2014].

быть независимы друг от друга. На рисунке 2 с помощью реляционной модельной структуры представлен выбор пары агентов α и β . Выбор, открытый для агента α , — это выбор между двумя классами эквивалентности $[h_1]_\alpha = \{h_1, h_2\}$ и $[h_3]_\alpha = \{h_3, h_4\}$; выбор, открытый для агента β , — выбор между классами $[h_1]_\beta = \{h_1, h_3\}$ и $[h_2]_\beta = \{h_2, h_4\}$.

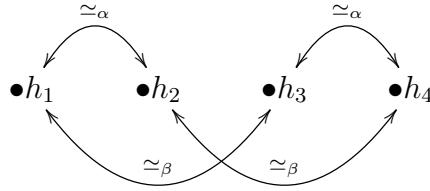


Рис. 2: Выбор пары агентов, представленный средствами реляционной семантики.

Читатель может убедиться самостоятельно, что при подобном разбиении множества историй-возможных миров, когда оба агента действуют независимо, истинным оказывается, например, такое утверждение: $[\alpha][\beta]\phi \leftrightarrow [\beta][\alpha]\phi$. Это означает, что допустима перестановка агентов, из-за чего теряются различия, к примеру, между командующим и тем, кто принимает команды.

Уже на основании этого можно заключить, что применение реляционной семантики в логике действий (традиционного инструмента исследования во многих обстоятельствах) затруднено, если мы собираемся использовать ее для моделирования непрямых действий агентов; ее доработка для решения вопроса о делегировании — тема отдельного исследования.

NEXT-семантика

Завершает наш обзор появившаяся относительно недавно (см. работы [Broersen 2011] и [Broersen 2009]) NEXT-семантика. В ее основе лежит упорядочивание моментов, отсутствие которого в реляционной семантике и дало указанный выше принцип, допускающий перестановку агентов, и останавливающий её дальнейшее использование для интерпретации непрямых действий. Вместе с тем, каждый момент рассматривается в NEXT-семантике как отдельная реляционная структура. Переход во

времени, то есть переход от одного момента к другому, от решения, принимаемого агентом, к его выполнению, связывает классы эквивалентности, принадлежащие разным моментам.

Пример NEXТ-структуры, включающей в себя те же элементы, что и структуры на рис. 1 и рис. 2, можно видеть на рис. 3.

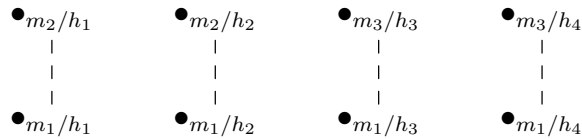


Рис. 3: NEXТ-структура, образованная моментами m_1 , m_2 и m_3 , и историями h_1 , h_2 , h_3 и h_4 .

Идея здесь, в общем, повторяет идеи реляционной и ветвящейся семантик: действуя в некотором мире, агент разбивает множество всех миров на классы эквивалентности. Отличие NEXТ-структур от ветвящихся и реляционных структур в том, что они позволяют учитывать время: действие, решение о котором агент принимает в некотором текущем моменте, реализуется только в последующем моменте.

Как мы увидим далее, нам достаточно того, что благодаря этой особенности NEXТ-структур, принцип, допускающий перестановку агентов, оказывается некорректен. Кроме этого, ее применение позволяет без труда истолковывать не прямые действия (например команды и просьбы) как простые описания будущего положения дел. Команду “Выполни ϕ !”, исходящую от α и предназначенную β , нетрудно рассматривать как описание будущей ситуации: “Будет так, что агент β сделает так, что ϕ ”, или как такое описание происходящего в настоящем: “Агент α делает так, что будет так, что агент β сделает так, что ϕ ”. Такой подход, помимо технического удобства, исключающего работу с какими бы то ни было дополнительными “императивными” операторами, отражает реальный способ употребления языка, и потому не является формалистской выдумкой. Так, например, предложения “Ваш сын отправится вместе с вами”, или “Ты оставишь альбом с марками дома” действительно используются именно как способы отдать указание или команду.

Определения, изложенные в разделе “Определения NEXТ-семантики”, позволят неподготовленному читателю сориентироваться в дальнейших рассуждениях. Те, кто знаком с NEXТ-семантикой, могут

продолжать чтение начиная с раздела “NEXT-модель для действий г-жи Октав, Франсуазы и судомойки”.

Определения NEXT-семантики

Формулы в NEXT-семантике интерпретируются с помощью модельных структур, упорядочивающих множества моментов времени по принципу раньше/позже. Более точно эту идею можно сформулировать в виде следующего определения:

модельная структура \mathcal{F} есть упорядоченная пятерка элементов $\langle M, H, R_X, R_\square, R_\circ, R_\Gamma \rangle$, где:

- $M = \{m_1, m_2, \dots\}$ есть бесконечное множество моментов времени;
- $H = \{h_1, h_2, \dots\}$ есть бесконечное множество историй, такое, что каждая история h представляет собой множество упорядоченных элементов из M ; множество пар $\{\langle m/h \rangle, \langle m'/h' \rangle\} \in M \times H$ обозначает множество миров — индексов, относительно которых осуществляется означивание формул.
- R_X есть отношение на множестве миров $M \times H$, такое что если $\langle m/h \rangle R_X \langle m'/h' \rangle$, то $h = h'$. Другими словами, для каждого момента времени m отношение R_X определяет последующий, более поздний, момент m' , принадлежащий той же истории, что и m . R_X антирефлексивно, транзитивно и антисимметрично.
- R_\square есть отношение на $M \times H$, такое что если $\langle m/h \rangle R_\square \langle m'/h' \rangle$, то $m = m'$. Так отношение R_\square для каждой истории h указывает такую историю h' , которая проходит через тот же момент времени, что и история h . R_\square рефлексивно, транзитивно и симметрично.
- R_\circ есть отношение на $M \times H$, такое, что если $\langle m/h \rangle R_\circ \langle m'/h' \rangle$, то $m = m'$, и каждая h' есть “идеальная” или “деонтически совершенная” история. R_\circ серийно (serial), т. е. для каждого мира $\langle m/h \rangle$ оно находит хотя бы один мир $\langle m'/h' \rangle$, такой, где h' — “идеальная” история.⁷

⁷Сейчас, приблизительно, скажем, что “идеальная” история — та, где события развиваются так, как они должны развиваться (вообще, или с точки зрения некоторого

- R_Γ есть отношение на $M \times H$, такое, что $R_\Gamma \subseteq R_X \circ R_\square$ и $R_\emptyset \subseteq R_\square \circ R_X$, где Γ — множество агентов $\{\alpha, \beta, \dots\}$. Мы также будем называть R_Γ отношением эффективности для агентов из Γ .

Отношение R_Γ упорядочивает определенным образом множество последующих моментов для агента или группы агентов Γ , так что его элементами выступают множества миров $\langle m/h \rangle$, каждое из которых ассоциируется с тем или иным возможным выбором для Γ . Пусть имеет место древовидная структура, представленная на рисунке 1. Ей соответствует NEXТ-структура на рисунке 3: они аналогичны, и отличаются лишь тем, что во втором случае явным образом, эксплицитно, указаны все истории, проходящие через каждый момент. Пусть в момент времени m агент α совершает некоторое действие, т. е. делает выбор или — разбивает множество последующих состояний на непересекающиеся классы, классы эквивалентности. В этом случае, по определению отношения эффективности, $\langle m_1/h_1 \rangle R_X \langle m_2/h_1 \rangle$ и $\langle m_2/h_1 \rangle R_\square \langle m_2/h_1 \rangle$, и значит, $\langle m_1/h_1 \rangle R_\alpha \langle m_2/h_1 \rangle$; далее, так как снова $\langle m_1/h_1 \rangle R_X \langle m_2/h_1 \rangle$ и $\langle m_2/h_1 \rangle R_\square \langle m_2/h_2 \rangle$, то $\langle m_1/h_1 \rangle R_\alpha \langle m_2/h_2 \rangle$. Аналогично, так как $\langle m_1/h_2 \rangle R_X \langle m_2/h_2 \rangle$ и $\langle m_2/h_2 \rangle R_\square \langle m_2/h_2 \rangle$, то, $\langle m_1/h_2 \rangle R_\alpha \langle m_2/h_2 \rangle$; наконец, так как $\langle m_1/h_2 \rangle R_X \langle m_2/h_2 \rangle$ и $\langle m_2/h_2 \rangle R_\square \langle m_2/h_1 \rangle$, то $\langle m_1/h_2 \rangle R_\alpha \langle m_2/h_1 \rangle$. Таким образом, классом эквивалентности для агента α и мира $\langle m_1/h_1 \rangle$, на модельной структуре из рис. 3, является множество $\{\langle m_2/h_1 \rangle, \langle m_2/h_2 \rangle\}$; классом эквивалентности для α и $\langle m_1/h_2 \rangle$ выступает оно же. Предлагаем читателю определить класс эквивалентности для α и миров $\langle m_1/h_3 \rangle$ и $\langle m_1/h_4 \rangle$ на данной модельной структуре.

Для описания прямых и непрямых действий нам понадобится язык $\mathcal{L}_{NEXТ}$, такой, что для множества пропозициональных переменных Φ и множества агентов Γ правильно построенная формула ϕ задается следующим образом:

$$\phi := p \mid \neg\phi \mid \phi \wedge \phi \mid X\phi \mid \square\phi \mid [\alpha]\phi \mid \bigcirc\phi$$

где $p \in \Phi$, и $\alpha \in \Gamma$.

Модель \mathcal{M} образуется на основе модельной структуры \mathcal{F} , как обычно, присоединением к ней функции означивания v , которая ставит в соответ-

агента). Например, если я одалживаю книгу у своего приятеля, то я, как правило, должен эту книгу в течении некоторого времени вернуть. “Идеальной” историей, идеальным развитием событий, будет то, когда я эту книгу действительно возвращаю. Подробнее об “идеальных” историях см. [Horty 2001: 34 и далее].

ствие атомарным пропозициям языка $\mathcal{L}_{\mathcal{N}\varepsilon\chi\tau}$ множество миров из $M \times H$, в которых эти пропозиции истинны.

Условия истинности формул языка $\mathcal{L}_{\mathcal{N}\varepsilon\chi\tau}$ определяются следующими правилами:

- $\mathcal{M}, \langle m/h \rangle \models p$ е. т. е. $\langle m/h \rangle \in v(p)$;
- $\mathcal{M}, \langle m/h \rangle \models \neg\phi$ е. т. е. $\mathcal{M}, \langle m/h \rangle \not\models \phi$;
- $\mathcal{M}, \langle m/h \rangle \models \phi \wedge \psi$ е. т. е. $\mathcal{M}, \langle m/h \rangle \models \phi$ и $\mathcal{M}, \langle m/h \rangle \models \psi$;
- $\mathcal{M}, \langle m/h \rangle \models X\phi$ е. т. е. существует такой мир $\langle m'/h' \rangle$, что $\langle m/h \rangle R_X \langle m'/h' \rangle$ и $\mathcal{M}, \langle m'/h' \rangle \models \phi$;
- $\mathcal{M}, \langle m/h \rangle \models \Box\phi$ е. т. е. для каждого индекса $\langle m'/h' \rangle$, такого, что $\langle m/h \rangle R_{\Box} \langle m'/h' \rangle$ имеет место $\mathcal{M}, \langle m'/h' \rangle \models \phi$;
- $\mathcal{M}, \langle m/h \rangle \models [\Gamma]\phi$ е. т. е. для всех индексов $\langle m'/h' \rangle$, таких, что $\langle m/h \rangle R_{\Gamma} \langle m'/h' \rangle$ имеет место $\mathcal{M}, \langle m'/h' \rangle \models \phi$;
- $\mathcal{M}, \langle m/h \rangle \models \bigcirc\phi$ е. т. е. $\mathcal{M}, \langle m'/h' \rangle \models \phi$ для каждого мира $\langle m'/h' \rangle$, такого, что $\langle m/h \rangle R_{\bigcirc} \langle m'/h' \rangle$.

Вместе язык $\mathcal{L}_{\mathcal{N}\varepsilon\chi\tau}$ и модельная структура \mathcal{F} предоставляют достаточно средств для того, чтобы формализовать и интерпретировать прямые и не прямые действия агентов из примера в разделе “Случай госпожи Октав, Франсуазы и судомойки”.

NEXT-модель для действий г-жи Октав, Франсуазы и судомойки

Ситуация с приготовлением обеда образована четырьмя действиями и четырьмя результатами этих действий. Прямые действия: Франсуазы — приготовление цыпленка — обозначим как $[\Phi]\phi$; судомойки — приготовление овощей — обозначим как $[C]\psi$. Способ записи не прямых действий требует комментария. По предположению, не прямое действие (в частности, указание и его разновидности) порождает обязательство. Так, не прямое действие г-жи Октав порождает обязательство Франсуазы сделать так, чтобы на обед были поданы цыпленок и овощи. Это можно

записать так: $[\text{ОКТ}] \circ [\Phi](\phi \wedge \psi)$. Вот не прямое действие Франсуазы, перепоручающей судомойке приготовление овощей: $[\Phi] \circ [C]\psi$.

Для того, чтобы построить модель для формул, соответствующих указанным прямым и непрямым действиям и их результатам, нам потребуется такая модельная структура, которая будет учитывать моменты принятия решения агентами ОКТ, Φ и C и моменты, в которых будет (или не будет) реализовываться принятое решение. Модельная структура, учитывающая, для всех агентов, наличие моментов принятия решения и моментов, в которых реализуется принятое решение, и, кроме того, — принимающая в расчет, что у каждого агента из примера должен быть хотя бы один выбор (поступить так, как требуется или воздержаться от действия), представлена на рис. 4.

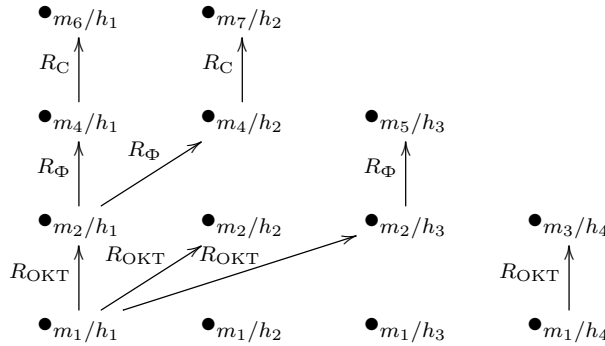


Рис. 4: Модельная структура для агентов ОКТ, Φ и C .

Момент m_1 — момент принятия решения γ -жой Октав о своем непрямом действии. По определению $R_{\text{ОКТ}} \subseteq R_X \circ R_{\square}$. Тогда действие γ -жи Октав в m_1 разбивает последующие моменты на два класса: в один из них входят пары $\langle \text{момент/история} \rangle \langle m_2/h_1 \rangle, \langle m_2/h_2 \rangle$ и $\langle m_2/h_3 \rangle$; в другой — момент/история $\langle m_3/h_4 \rangle$.⁸ Эти два класса — те, в которых соответственно реализуется и не реализуется не прямое действие γ -жи Октав.

⁸Происходит это на основании следующего: так как $\langle m_1/h_1 \rangle R_X \langle m_2/h_1 \rangle$, и $\langle m_2/h_1 \rangle R_{\square} \langle m_2/h_1 \rangle$, то $\langle m_1/h_1 \rangle R_{\text{ОКТ}} \langle m_2/h_1 \rangle$; так как $\langle m_1/h_1 \rangle R_X \langle m_2/h_1 \rangle$, и $\langle m_2/h_1 \rangle R_{\square} \langle m_2/h_2 \rangle$, то $\langle m_1/h_1 \rangle R_{\text{ОКТ}} \langle m_2/h_2 \rangle$; так как $\langle m_1/h_1 \rangle R_X \langle m_2/h_1 \rangle$, и $\langle m_2/h_1 \rangle R_{\square} \langle m_2/h_3 \rangle$, то $\langle m_1/h_1 \rangle R_{\text{ОКТ}} \langle m_2/h_3 \rangle$; наконец, так как $\langle m_1/h_4 \rangle R_X \langle m_3/h_4 \rangle$, и $\langle m_3/h_4 \rangle R_{\square} \langle m_3/h_4 \rangle$, то $\langle m_1/h_4 \rangle R_{\text{ОКТ}} \langle m_3/h_4 \rangle$. Все остальные действия разбивают последующие моменты сходным образом, в соответствии с определением R_{Γ} и тем, как

Кроме того, первый из них служит еще и отправной точкой для принятия решения Франсуазой. Именно здесь Франсуаза, получившая указание г-жи Октав, и отныне связанная обязательством сделать так, что “ ϕ и ψ ”, принимает решение выполнить одну часть указания самостоятельно, а другую поручить судомойке.

Классы эквивалентности для Франсуазы в m_2 : один, соответствующий решению готовить цыпленка самостоятельно и поручить судомойке приготовление овощей, — класс, куда входят $\langle m_4/h_1 \rangle$ и $\langle m_4/h_2 \rangle$; другой, связанный с возможными действиями иного рода, которые мы предполагаем на том основании, что Франсуаза — агент, что означает, что у нее всегда есть выбор, как поступить, — класс, куда входит $\langle \text{момент/история} \rangle \langle m_5/h_3 \rangle$. Класс $\{\langle m_1/h_1 \rangle, \langle m_1/h_2 \rangle\}$, куда нас приводит решение Франсуазы поручить судомойке приготовление овощей, является тем классом, где теперь уже судомойка делает выбор: приготовить овощи ($\{\langle m_6/h_1 \rangle\}$) или воздержаться от приготовления овощей ($\{\langle m_7/h_2 \rangle\}$).

Пусть “идеальной” историей будет история h_1 . Нетрудно увидеть, что на данной модельной структуре, результатами непрямых действий г-жи Октав и Франсуазы, и прямых действий Франсуазы и судомойки, в соответствии с указанными классами эквивалентности, образованными разбиениями последующих моментов функциями $R_{\text{ОКТ}}$, $R_{\text{Ф}}$ и $R_{\text{С}}$, являются: $\phi \wedge \psi$, истинная в $\langle m_4/h_1 \rangle$ и $\langle m_4/h_2 \rangle$ (результат непрямого действия г-жи Октав); ϕ , истинная там же (результат прямого действия Франсуазы); ψ , истинная в $\langle m_6/h_1 \rangle$ (результат прямого действия судомойки). Модель, таким образом, оказывается непротиворечива.

Интересно отметить, что даже в таком виде, как сейчас, модель позволяет сделать некоторые выводы о делегировании распоряжения, отданного г-жой Октав. В самом деле, на вопрос о том, является ли это распоряжение, в то же самое время, таким, как если бы оно было отдано судомойке непосредственно, хотя бы и исключительно в отношении приготовления овощей, может быть дан отрицательный ответ. Нет, не является, ведь формула $[\text{ОКТ}] \circ [\text{С}]\psi$, выражающая данное распоряжение, истинная в том же самом $\langle \text{момента/истории} \rangle$, что и изначальное указание г-жи Октав, не дает ψ в $\langle m_6/h_1 \rangle$, как это происходит в первом случае. т. е. при передаче указания судомойке Франсуазой.

устроена модельная структура. Далее мы будем указывать классы эквивалентности без подробного описания работы функции $R_X \circ R_{\square}$.

Вместе с тем, если считать, что судомойка получает указание варить овощи как будто от г-жи Октав именно в тот момент, когда она слышит его из уст Франсуазы, т. е. если считать, что формула $[ОКТ] \circ [C]\psi$ истинна не в $\langle m_1/h_1 \rangle$, а, например, в $\langle m_2/h_1 \rangle$, то тогда условия истинности этой формулы совпадают с условиями истинности формулы $[Ф] \circ [C]\psi$, выражающей переданное Франсуазой судомойке указание г-жи Октав. Совпадение условий истинности двух формул, некоторым образом, свидетельствует в пользу того, что они могут рассматриваться как одно и то же действие. В этом случае перепоручение указания первого агента третьему агенту через второго, действительно означает то, как если бы оно было отдано непосредственно от первого агента третьему.

Авторитетные отношения агентов

Удовлетворительное решение вопроса о делегировании упирается в отслеживание исследователем и запись в языке-объекте авторства осуществляемых действий в связи с тем, чье авторство считается агентом, отдавшим некоторое указание, предпочтительным⁹. До настоящего момента мы считали, что, например, в формуле $[\alpha] \circ [\beta]\phi$, агент, обязываемый агентом α выполнить ϕ , — это агент β , и никто другой. Однако, такое личное обязательство есть скорее частный случай обязательства вообще, чем ситуация, отражающая большинство обязательств-следствий непрямых действий. В этом нет ничего необычного, ведь значительная часть обязательств подразумевает возможность их переадресации, делегирования; более того, переадресация часто рассматривается как само собой разумеющееся действие. Например, когда некто получает указание вымыть автомобиль, то это вовсе не означает, что он непременно должен сделать это самостоятельно. (А в случае отсутствия моек самообслуживания и наличия запрета на мытье машин “во дворах”, получивший такое указание и вовсе лишен возможности его самостоятельного исполнения.) Тем не менее, очевидно, его способность исполнить указание не ставится под сомнение, а значит, она обеспечивается исключительно переадресацией, и получивший такое указание осведомлен об этом в той же степени, как и тот, кто его выдал.

⁹В ситуации, когда агент α делает так, что ϕ , мы будем считать агента α автором ϕ . Тогда можно представить себе действие, не имеющее точного указания авторства: $\exists x$ такой, что $[x]\phi$. И, самое главное, — тогда такие действия могут быть целью не прямых действий других агентов: “Кто-нибудь, закройте окно!”

Будем говорить, что, в зависимости от того, кто должен исполнять указание, обязательства, порождаемые непрямыми действиями агентов, можно поделить на *личные* или *сугубо-личные* и *безличные* или *условно-личные* обязательства; вторые, в отличие от первых, допускают переадресацию обязательного к исполнению действия другому агенту. Сугубо-личное обязательство порождается, например, императивным предложением “Слушайся старших”. Условно-личное обязательство возникает, например, тогда, когда некто произносит: “Я напечатаю эти фотографии на следующей неделе”.

Это заставляет нас по-новому смотреть на понятие агента. Ясно, что некто признается агентом не вообще, а в связи с некоторым набором формул. В этот набор входят все те формулы, истинность которых он может гарантировать. С учетом прямых и не прямых действий, совершаемых агентом, это могут быть формулы вида p , ϕ , $\neg\phi$ и так далее, или же формулы вида $\bigcirc\phi$. Теперь, так как мы выделили среди обязательств сугубо-личные и условно-личные обязательства, мы можем высказаться о таком наборе формул более определенно, и сказать, что среди множества принадлежащих ему формул вида $\bigcirc\phi$, существуют формулы вида $\bigcirc[\alpha]\phi$, отражающие сугубо-личное обязательство агента α , которое возникает как результат непрямого действия, и формулы вида $\bigcirc[x]\phi$, где x есть переменная, значения которой пробегают по множеству Γ , — отражающие условно-личные обязательства некоторого агента x .

Будем называть множество формул D_α , истинность которых может гарантировать агент α , областью определения данного агента. Если $\phi \in D_\alpha$, то существует такой мир $\langle m/h \rangle$, что $\langle m/h \rangle \models [\alpha]\phi$. Возвращаясь к нашему примеру, скажем, что: $\psi \in D_C$, $\phi \in D_\Phi$, $\bigcirc[C]\psi \in D_\Phi$, а $\bigcirc[\Phi]\phi \wedge \psi \in D_{\text{ОКТ}}$. (Очевидно, что, например, $\bigcirc[\text{ОКТ}]\phi \notin D_\Phi$.)

Учитывая деления обязательств на сугубо-личные и условно-личные, мы можем уточнить область определения агента ОКТ: пусть $\bigcirc[x]\phi \in D_{\text{ОКТ}}$ — тогда, где-то Октава все равно, кто именно будет заниматься приготовлением обеда: Франсуаза или судомойка, или кто-то еще из помощниц Франсуазы — других слуг. Обязательства, которые она адресует Франсуазе — условно-личное.

Пусть состоялось не прямое действие, такое, что $[\alpha] \bigcirc [\beta]\phi$. Будем считать, что это говорит в пользу того, что между агентами α и β установлено авторитетное отношение. Авторитетное отношение будем понимать как возможность осуществить не прямое действие. Тогда, верно, что $[\alpha] \bigcirc [\beta]\phi \rightarrow \langle \alpha \rangle \bigcirc [\beta]\phi$, где $\langle \alpha \rangle \bigcirc [\beta]\phi$ есть синтаксическое выражение

авторитетного отношения между α и β по ϕ . Семантически же на авторитетное отношение можно сослаться, если описать область определения агента. Если $\bigcirc[\beta]\phi \in D_\alpha$, то, снова, α находится в авторитетном отношении по ϕ к β .

Сказанное выше относится к непрямым действиям, которые порождают порождающим сугубо-личные обязательства. Авторитетное отношение, существующее между агентами, в ситуации непрямого действия, порождающего условно-личное обязательство, можно представить так. Пусть агент β получает условно-личное обязательство от α . Тогда, он вправе выполнить требуемое самостоятельно или перепоручить его выполнение некоторому другому агенту x . Таким образом, агенты α и β , и агенты β и x связаны авторитетным отношением по ϕ : $\langle \alpha \rangle \bigcirc [\beta]\phi$ и $\langle \beta \rangle [x]\phi$ (где x — агент, такой, что $\bigcirc[x]\phi \in D_\beta$).

Эта ситуация отражает ровно тот случай, что описан в разделе “Случай госпожи Октав, Франсуазы и судомойки”: г-жа Октав отдает указание Франсуазе; Франсуаза (часть этого указания) перепоручает судомойке. Тогда между г-жой Октав и Франсуазой существует авторитетное отношение по ϕ и по ψ , а между Франсуазой и судомойкой — авторитетное отношение по ψ . Это можно представить в виде дерева авторитетных отношений, связывающего области определений г-жи Октав, Франсуазы и судомойки. Если D_α расположена непосредственно над стрелкой, слева от которой стоит ϕ , а D_β — непосредственно под этой стрелкой, то агент α находится в авторитетном отношении к агенту β по ϕ . Другими словами, α может отдать β указание на выполнение ϕ .

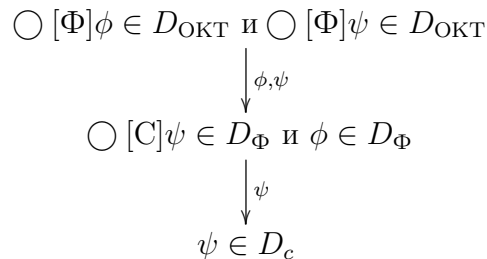


Рис. 5: Авторитетное отношение по ϕ и ψ между ОКТ, Ф и С.

Вопрос о делегировании указания, определивший содержание и направление исследования (означает ли эта рассматриваемая нами ситуа-

ция то, что г-жа Октав как будто бы отдает указание судомойке чистить и варить овощи непосредственно?) теперь может быть поставлен нами в несколько иной форме. Следует отказаться от этого “как будто” в его формулировке (напоминающего кантовское “как если бы”) и спросить о следующем: верно ли, что рассматриваемая нами ситуация означает то, что у г-жи Октав есть возможность отдать указание судомойке непосредственно? Другими словами, верно ли, что если между г-жой Октав и Франсуазой существует авторитетное отношение по ψ (в тот момент, когда г-жа Октав осуществляет свое не прямое действие), и если между Франсуазой и судомойкой существует авторитетное отношение по ψ (в тот момент, когда Франсуаза осуществляет свое не прямое действие), то между г-жой Октав и судомойкой также существует авторитетное отношение по ψ (в один из моментов, когда осуществляют свои действия г-жа Октав и Франсуаза)?

Этот вопрос означает, что мы должны убедиться в истинности следующей формулы $((\langle \text{ОКТ} \rangle \circ [\Phi]\psi) \wedge (\langle \Phi \rangle \circ [C]\psi)) \rightarrow (\langle \text{ОКТ} \rangle \circ [C]\psi)$ в модели, представленной на рис. 4. Предположим, что предполагаемого авторитетного отношения нет ни в одном из моментов. Тогда, по определению авторитетного отношения, г-жа Октав, во всех этих моментах, не является агентом в отношении обязательства судомойки сделать так, что ψ . Это означает, что г-жа Октав полностью лишена возможности отдать команду судомойке выполнить ψ . В частности, г-жа Октав лишена этой возможности в тот момент, когда соответствующее указание отдает судомойке Франсуаза, т. е. в момент m_2 . Отсутствие возможности у г-жи Октав в момент m_2 обязать судомойку сделать так, чтобы имело место ψ , представим в виде формулы $\neg \langle \text{ОКТ} \rangle \circ [C]\psi$, истинной в m_2 и во всех историях, проходящих через m_2 (т. е. в парах $\langle m_2/h_1 \rangle$, $\langle m_2/h_2 \rangle$ и $\langle m_2/h_3 \rangle$). Это равносильно $[\text{ОКТ}]\neg \circ [C]\psi$ (так как верно, что $\langle \alpha \rangle \phi \equiv \neg [\alpha]\neg \phi$) истинной там же.

Вне зависимости от того, как пойдет разбиение пар $\langle \text{момент} / \text{история} \rangle$, следующих за моментом m_2 (пар $\langle m_4/h_1 \rangle$, $\langle m_4/h_2 \rangle$ и $\langle m_5/h_2 \rangle$), истинность формулы $\neg \circ [C]\psi$ в любой из этих пар дает противоречие с формулой $\circ [C]\psi$, истинной в $\langle m_4/h_1 \rangle$ по условию (как результат действия Франсуазы — $[\Phi]\circ [C]\psi$ — в m_2). В самом деле, если допустить, что в одной из пар $\langle m_4/h_1 \rangle$, $\langle m_4/h_2 \rangle$ или $\langle m_5/h_2 \rangle$ имеет место $\neg \circ [C]\psi$, то следует признать, что существует такая “идеальная” история, где неверно, что истинна формула $[C]\psi$. Вместе с тем, по условию, во всех “идеаль-

ных” историях, т. к. $\bigcirc[C]\psi$ истинна в $\langle m_4/h_1 \rangle$, имеет место $[C]\psi$ (ведь, снова, по условию, судомойка не делегирует указание дальше).

Таким образом, ясно, что не прямое действие г-жи Октав и не прямое действие Франсуазы, вместе, дают основание утверждать, что имеет место $\langle \text{ОКТ} \rangle \bigcirc [C]\psi$; или, другими словами — что между агентами ОКТ и С по ψ существует авторитетное отношение.

Заключительные замечания к делегированию указаний

В предыдущем разделе мы определили понятие авторитетного отношения на синтаксическом, семантическом и содержательном уровне. Синтаксически авторитетное отношение между агентами α и β по ϕ представимо в виде формулы $\langle \alpha \rangle \bigcirc [\beta]\phi$. Семантически это означает, что $\bigcirc[\beta]\phi \in D_\alpha$. Содержательно — что агент α обладает возможностью отдать указание (или совершить какое-то другое не прямое действие) β на выполнение ϕ .

Мы установили, что $[\text{ОКТ}] \bigcirc [\Phi]\psi$ и $[\Phi] \bigcirc [C]\psi$ вместе влекут $\langle \text{ОКТ} \rangle \bigcirc [C]\psi$. Это значит, что указание на выполнение ψ , отданное г-жой Октав Франсуазе, и указание на выполнение ψ , отданное Франсуазой судомойке, позволяют заключить, что между г-жой Октав и судомойкой существует авторитетное отношение по ψ ; т. е. — что у г-жи Октав существует возможность отдать судомойке указание непосредственно на выполнение ψ .

Будем считать, что если $\langle \alpha \rangle \bigcirc [\beta]\phi$ и $\langle \beta \rangle \bigcirc [\gamma]\phi$, то $\langle \alpha \rangle \bigcirc [\gamma]\phi$, т. е. что, в таком случае, между α , β и γ установлено транзитивное авторитетное отношение.

Отметим, что авторитетные отношения между г-жой Октав, Франсуазой и судомойкой определены нами не исходя из общих представлений о господах и их слугах, а исключительно на основе тех действий, которые были предприняты этими тремя женщинами, и о которых шла речь в примере. Авторитетное отношение между агентами требует не только указания той формулы, по которой оно устанавливается, но и того момента времени, когда оно действует. Интересно, что авторитетное отношение по ψ между ОКТ и С возникает в момент передачи указания Франсуазой “дальше” и, исходя из предложенной семантической модели, не существует как таковое всегда, по умолчанию.

В определении агента сказано: α — агент в отношении ϕ , только если ϕ истинна во всех мирах, связанных с данным миром отношением R_α . Теперь к этому можно добавить: быть агентом также означает находиться в авторитетном отношении по некоторой формуле к некоторому другому агенту.

Тогда можно выделить типы агентов, в зависимости от того, каким образом тот или иной агент гарантирует истинность формул из своей области определения. Пусть существуют агенты, которые обладают возможностью реализовывать только прямые действия (тип 0); пусть существуют те, кто способен гарантировать истинность формул из своей области определения через реализацию как прямых, так и непрямых действий (тип 1); наконец, пусть существуют такие агенты, кто гарантирует истинность формул из своей области определения только благодаря тому, что они связаны авторитетными отношениями с другими агентами — т. е. только благодаря своей возможности осуществлять не прямые действия (тип 2).

Если мы представим агента α , принадлежащего типу 2, то мы должны будем признать, что его существование как агента обеспечивает некий другой агент, например β (который сам принадлежит типу 0, 1 или 2). Если β принадлежит типу 2, то все, что сказано в отношении α , теперь можно повторить и в отношении β , и некоторого агента γ , такого, что $\bigcirc[\gamma]\phi \in D_\beta$, и так до бесконечности. Этого, однако, можно избежать, если предположить, что цепочку из агентов типа 2 замыкает агент, принадлежащий типу 0 — тот кто не может перепоручить выполнение некоторого действия кому-то еще. Интересно заметить, что в подобных случаях, агенты типа 2, составляющие цепочку, остаются агентами лишь благодаря тому, что с ними связан авторитетным отношением агент типа 0 — агент-исполнитель. (Так г-жа Октав остается агентом в отношении ψ — может гарантировать ψ — только потому, что в цепочке ОКТ-Ф-С находится агент С, принадлежащий типу 0 в связи с положением дел ψ .)

Действительно, пусть α — агент по отношению к ϕ и $\neg[\alpha]\phi$. Тогда должен найтись агент β , такой, что $\bigcirc[\beta]\phi \in D_\alpha$. Пусть β , снова, как и α — такой агент, что $\neg[\beta]\phi$. Тогда должен найтись агент γ , такой, что $\bigcirc[\gamma]\phi \in D_\beta$, и так далее *ad infinitum*. Если β не является таким агентом для которого должен найтись агент γ , находящийся в авторитетном отношении к β по ϕ , то должно быть так, что ϕ истинна во всех мирах, достижимых из данного мира, через отношение R_β . Это, помимо прочего, является и необходимым условием для того, чтобы агентом продолжал

оставаться, кроме самого β , еще и α . Ведь если неверно, что $[\beta]\phi$, то β не агент в связи с ϕ . А если это так, то нет такого агента, который был бы подчинен авторитету агента α . Тогда и α теперь (в силу того, что, по условию, $\neg[\alpha]\phi$) перестает, в свою очередь, быть агентом в связи с ϕ .

Таким образом “агентность” по некоторой формуле ϕ , в случае, если та обеспечивается исключительно авторитетным отношением по ϕ , упирается в существование агента-самостоятельного-исполнителя ϕ , агента, который может сделать так, что ϕ будет иметь место как результат его прямого действия. При отсутствии, среди агентов, связанных авторитетным отношением по ϕ , агента-исполнителя в самом “низу”, эти авторитетные отношения перестают существовать, а агенты, не обладающие возможностью выполнить ϕ и лишь назначающие ϕ к исполнению другим, перестают быть агентами в отношении ϕ .

Литература

- Хинтиikka 2014 — *Хинтиikka, Я.* О Гёделе. КАНОН+, 2014.
- Balbani, Herzig, Troquard 2008 — *Balbani, P., Herzig, A., Troquard, N.* Alternative axiomatics and complexity of deliberative stit theories // *Journal of Philosophical Logic*, 37(4):387–406, 2008.
- Belnap 1991 — *Belnap, N.* Before refraining: Concepts for agency // *Erkenntnis*, 34(2):137–169, 1991.
- Belnap 1991a — *Belnap, N.* Backwards and forwards in the modal logic of agency // *Philosophy and phenomenological research*, pages 777–807, 1991.
- Bentham, Pacuit 2014 — *Bentham, J., Pacuit, E.* Connecting logics of choice and change // *Nuel Belnap on Indeterminism and Free Action*, pages 291–314. Springer, 2014.
- Broersen 2009 — *Broersen, N.* A complete stit logic for knowledge and action, and some of its applications // *Declarative Agent Languages and Technologies VI*, pages 47–59. Springer, 2009.
- Broersen 2011 — *Broersen, N.* Making a start with the stit logic analysis of intentional action // *Journal of philosophical logic*, 40(4):499–530, 2011.
- Castaneda 1960 — *Castaneda, H.* Imperative reasonings // *Philosophy and Phenomenological Research*, 21(1):21–49, Sep. 1960.
- Geach 1963 — *Geach, P.* Imperative inference // *Analysis*, 23 (Supplement 1):30–42., Jan. 1963.

- Horty 2001 — *Horty, J.* Agency and deontic logic. Oxford University Press
Oxford, 2001.
- Kenny 1966 — *Kenny, A.* Practical inference // *Analysis*, 26(3):65–75,
Jan. 1966.
- Seegerberg, Meyer, Kracht 2016 — Seegerberg, K., Meyer, J., Kracht, M.
The logic of action // Edward N. Zalta, editor, *The Stanford
Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford
University, winter 2016 edition, 2016.
- Xu 1998 — *Xu, M.* Axioms for deliberative stit // *Journal of Philosophical
Logic*, 27(5):505–552, 1998.