

УДК 162

DOI: 10.17223/1998863X/57/2

**Д.В. Зайцев, А.А. Беликов**

## **МОДЕЛИРУЯ АРГУМЕНТАЦИЮ: ОЦЕНКИ И РАССУЖДЕНИЯ<sup>1</sup>**

*Рассматривается проблема формального моделирования аргументации. Выявляются основные характеристики аргументативных рассуждений, играющие ключевую роль в оценке аргументации. Авторы предлагают формальную экспликацию отношения подтверждения между аргументами и тезисом, которая учитывает вариативность отношения следования, т.е. позволяет иметь дело как с дедуктивным, так и с правдоподобным следованием в одном формальном контексте.*

*Ключевые слова: аргументация, аргументативные рассуждения, оценка аргументации, формальные модели аргументации.*

### **Введение**

Несмотря на многовековую историю, изучение феномена аргументации все так же далеко от идеалов строгой научности, пример следования которым демонстрирует ее младшая сестра – логика, не менее тесно связанная с исследованием рассуждений. По сути дела, аргументация представляет собой сегодня, как, впрочем, и сотни лет назад, нечто наподобие сакрального текста – разновидности магического заговора или заклинания, вызывающего как минимум некоторые вербально оформленные действия (выражение согласия с обосновываемым тезисом), а как максимум – изменения в сознании рационального агента (перубеждение оппонента). В отличие от так называемых «малых фольклорных текстов» аргументация все-таки явно предполагает опору на рациональные средства убеждения, что в целом соответствует парадигме рациональной мультиагентности. Соответственно, изучение этого волшебного феномена протекает в широком спектре от сугубо практически ориентированных «цифровых самоучителей игре на гармошке без нот» до попыток истолкования – теоретического осмысления того, «что же стоит за аргументацией» и «почему она бывает убедительной». Чуть в стороне от этого многообразия идей располагаются работы в области формального моделирования аргументации (formal argumentation), инспирированные преимущественно потребностями искусственного интеллекта и computer science. При этом любой подход или направление напрямую или косвенно оказывается связанным с решением, можно сказать, «основного вопроса аргументации» – каковы критерии оценки аргументации? Для каких бы целей ни предпринимались попытки описать, систематизировать, моделировать аргументацию, каких бы философскометодологических позиций ни придерживались их авторы, обойти вопрос о том, в каких терминах и на каких основаниях оценивается аргументация, не удастся.

---

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФ в рамках научного проекта № 20-18-00158.

Эти небольшие вводные наблюдения приводят к очевидному, казалось бы, выводу: прежде чем предпринимать какие бы то ни было попытки «аргументативного теоретизирования», необходимо четко сформулировать эффективные критерии оценки аргументации. Однако все не так просто, существует целый ряд возможных параметров оценки аргументации, в разной степени поддающихся формализации и строгому учету. К примеру, субъективная, прагматическая эффективность аргументации, характеризующая личностные цели агента, в большинстве случаев остается неактуализированной, и мы как объективные исследователи аргументации можем только догадываться о том, какие цели ставил перед собой участник аргументативного процесса и оказались ли эти цели достигнуты. Кроме того, применительно к формальным моделям аргументации оказывается, что выбор инструментов моделирования уже накладывает определенные ограничения на возможности оценки аргументации. Таким образом, получается, что выбор исходной модели и параметров оценки аргументации тесно связаны.

Данная работа, с одной стороны, продолжает цикл наших исследований в области аргументации [1–4], а с другой – вызвана к жизни запуском амбициозного проекта «Формальная философия аргументации и комплексная методология поиска и отбора решений спора» (грант РНФ № 20-18-00158), участниками которого являются авторы. Мы оптимистично надеемся, что эта статья откроет новый цикл наших работ, посвященных формальному анализу аргументации. Принимая во внимание изложенное выше, в этой статье мы попытаемся предложить некоторый достаточно абстрактный подход к формализации аргументативных рассуждений и обсудить те возможности оценки аргументации, которые он предполагает. В соответствии с этими задачами и строится дальнейшее изложение. В первом разделе излагаются наиболее общие характеристики так называемой «аргументативной семантики». Второй раздел посвящен неформальному рассмотрению типов рассуждений, используемых в аргументации. Наконец, третий раздел содержит набросок формализации аргументативных рассуждений.

## **1. Формальная аргументация и аргументативная семантика**

Процесс аргументации характеризуется несколькими структурными компонентами, среди которых принципиальное значение имеет учет субъектов аргументации (рациональных когнитивных агентов). Активное развитие компьютерных наук и искусственного интеллекта оказывает значительное влияние и на развитие исследований аргументации, которая уже не трактуется как неформальная и сугубо практическая дисциплина, описывающая те или иные виды аргументов, уловок и полемических ситуаций. Наряду с человеком в категорию «субъект аргументации» теперь попадает и компьютер, что вполне предсказуемо породило новую волну интереса не только к точному и строгому изучению базовых понятий аргументации, но и к переосмыслению самой теории в целом.

Эти обстоятельства стимулировали развитие уже упоминавшейся формальной аргументации, подхода, в рамках которого значительно проще пытаться строить более или менее строгую модель аргументации. Ключевыми понятиями формальной аргументации стали понятия аргументативного фрейма и аргументативной семантики, предложенные Ф. Дунгом [5]. Именно

благодаря результатам Дунга мы имеем возможность использовать единый формализм для моделирования аргументативных рассуждений, что, безусловно, очень важно в контексте компьютерных наук. Еще один плюс подхода Дунга – удивительная простота формальных построений. Это, в свою очередь, позволяет расширить границы применимости данного подхода и распространить его на широкую область гуманитарных дисциплин: когнитивные науки, философию и пр.

Формализм Дунга может быть описан следующим образом. Базовым понятием является понятие аргументативной структуры, состоящей из множества аргументов (высказываний) и бинарного отношения «атаки» на множестве аргументов.

DF.1. Аргументативная структура (*argumentative framework*) – это пара (*Arg*, *attacks*), где *Arg* – множество аргументов, а *attacks* – бинарное отношение на множестве *Arg*.

Очевидной характеристикой любой аргументации является конфликт субъектов аргументации, состоящий в том, что аргументы одной стороны направлены на опровержение (или критику) аргументов другой, и наоборот. Именно эта простая интуиция выражена в использовании отношения «атаки» между аргументами. Уже из этих примитивных понятий мы можем вывести ряд интуитивно приемлемых свойств, которым должны удовлетворять аргументативные структуры, направленные на моделирование позиции абстрактного субъекта аргументации.

DF.2. Множество аргументов *S* называется неконфликтным (*conflict-free*), если не существует таких аргументов *A* и *B* из *S*, что *attacks*(*A*, *B*).

DF.3. Аргумент называется *приемлемым относительно множества S*, если для всякого аргумента *B* из *Arg* верно: если *B* атакует *A*, то *S* атакует *B*.

DF.4. Неконфликтное множество *S* называется *допустимым*, если всякий аргумент из *S* приемлем относительно *S*.

Наконец, мы подошли к самому главному. Структуры, построенные в соответствии с определениями 1–4, могут использоваться для моделирования конкретных полемических ситуаций. Таким образом, следующая задача заключается в том, чтобы установить, какими именно критериями мы должны руководствоваться, чтобы выявить максимально выигрышный набор аргументов в рамках той или иной ситуации. Подобные критерии и есть то, что Дунг называет «аргументативной семантикой».

Далее Дунг вводит понятие «расширения» (*E*, от *extension*) для построения аргументативной семантики. Под расширением понимается некоторое подмножество аргументов из базового множества *Arg*. Выделяется два основных типа расширений:

– предпочтительное расширение (*preferred extension*) – это максимальное (относительно теоретико-множественного включения) допустимое множество аргументов *E* из *Arg*,

– стабильное расширение (*stable extension*) – это множество аргументов *E* из *Arg* такое, что оно атакует каждый аргумент из множества *Arg* \ *E*.

Стоит отметить, что Дунг выделяет два дополнительных типа расширений – основательные (*grounded*) и полные (*complete*) (подробно см.: [5]).

Разработка различных по своей структуре и свойствам аргументативных семантик является сегодня одной из горячих тем в формальном моделирова-

нии аргументации. Следуя работе [6], аргументативные семантики можно классифицировать на три основных группы: семантики расширений (*extension semantics*), оценочные семантики (*weighting semantics*) и ранжированные семантики (*ranking semantics*). Различие между ними состоит в том, какого рода результат мы получаем после применения той или иной семантики. Руководствуясь, скажем, семантикой расширений, мы получаем некоторое подмножество аргументов, а если мы используем оценочную семантику, то результат состоит в некотором наборе оценок аргументов, которые позволяют нам сравнивать аргументы между собой по силе, основываясь лишь на некоторых числовых значениях этих аргументов. В свою очередь, ранжированная семантика дает на выходе структуры типа пред-порядка на множестве аргументов, тем самым фиксируя информацию об упорядочивании аргументов относительно того, насколько сильно они атакованы. Это, опять-таки, позволяет сравнивать аргументы по силе.

В работе [7] рассматриваются различные подходы к оценке аргументов в рамках ранжированных семантик. Так, оценка аргументации может быть основана на силе атакующих аргументов [8], их социальной поддержке [9], весе и длине атакующей цепи [10], а также учете всех атакующих и поддерживающих эти аргументы предшественников в атакующей цепи [11].

Хотя абстрактный характер аргументативных фреймов Дунга ранее оценивался нами как преимущество, в некоторых контекстах подобные обобщения избыточны и, возможно, лишают нас другого взгляда на вещи. Один из возможных путей развития подхода Дунга состоит в уточнении структуры аргументов. Это позволяет использовать аппарат дедуктивных логических теорий, что уже полезно для анализа не только аргументативных рассуждений, но и чисто логических проблем, связанных, например, с экспликацией немонотонных рассуждений. По существу, в рамках данного подхода мы можем рассматривать множество *Arg* как замыкание некоторого формального языка (например, пропозиционального) относительно логического следования. Таким образом, множество *Arg* уже не будет состоять из каких-то абстрактных и не специфицированных объектов, которые мы называем «аргументами»; теперь под «аргументами» понимаются упорядоченные пары типа  $(\Gamma, A)$ , репрезентирующие утверждения о следовании между множеством формул  $\Gamma$  и формулой  $A$ . Этот подход послужил основой для возникновения целой области под названием «логическая аргументация» (*logic-based argumentation*).

Как уже было сказано, аргументативная семантика полезна для решения проблем на пересечении аргументации и логики. Хорошо известно, что начиная с 70-х гг. прошлого века в логике особый статус заняли исследования семантических моделей, допускающих ту или иную степень противоречивости информации. Среди разделов логики, касающихся этой проблематики, можно выделить, например, паранепротиворечивые и релевантные логики. Любопытно, что некоторые подходы были инспирированы в том числе и моделированием рассуждений компьютеров, искусственного интеллекта и пр. (см.: [12, 13]). В то же время они подвергались критике как подходы, не вполне адекватные реальной практике компьютерных рассуждений [14]. На сегодняшний день логическая аргументация дает возможность не только анализировать процесс аргументативных рассуждений, учитывая логическую

структуру аргументов, но и в некотором смысле ответить на критику упомянутых логических теорий. В последние годы опубликовано множество работ по моделированию аргументативных рассуждений, основанных на паранепротиворечивых [15–17], немонотонных [18], релевантных [19] и модальных логиках [20]. Важно, что выбор логического инструментария влияет не только на выразительные возможности построенных на его основе моделей аргументации, но и на возможности оценки аргументации.

## 2. Рассуждения в аргументации

В данном разделе мы постараемся наметить некоторые перспективы моделирования аргументации, обеспечивающего различные возможности для ее оценки.

Первый вопрос, касающийся аргументации, связан с трактовкой аргументативных рассуждений. Имеют ли они какую-то специфику, тождественны ли тому, что мы привыкли называть естественными рассуждениями, или для рассуждений в рамках аргументативного контекста оправданно говорить о каких-то специальных требованиях и критериях, что позволило бы характеризовать их, скорее, как «искусственные» (пусть и не в той степени, как, скажем, строгие математические рассуждения)? Этот вопрос в определенном смысле является риторическим, поскольку предполагает не столько ответ, сколько просто констатацию выбранной позиции. С нашей точки зрения, которая в этом отношении близка к идеям Х. Мерсье и Д. Сперберга [21], естественные рассуждения аргументативны, а аргументативные рассуждения естественны. Другими, менее метафоричными, словами, любое рассуждение, по сути дела, представляет аргументацию по отношению к некоторому внутреннему субъекту (суб-личности). В противном случае, если бы отсутствовала необходимость кого-то (в том числе, и себя) в чем-то убеждать, если бы истинность или приемлемость некоторого утверждения была бы очевидна и бесспорна, то и необходимость в конструировании рассуждения тоже бы не возникла. Раз так, то становится очевидным, что аргументативные рассуждения должны обладать «трудными» чертами естественных рассуждений: немонотонностью (модифицируемостью) и правдоподобностью. Остановимся на этих характеристиках рассуждений чуть подробнее.

В современной логике принят подход, согласно которому отношение между посылками и заключением в правильном рассуждении представляет собой логическое следование. Таким образом, из посылок логически следует заключение, если и только если не существует такой интерпретации нелогических параметров в их составе, при которой посылки оказались бы истинными, а заключение ложным. Этот критерий носит явный дедуктивный характер, логическое следование гарантирует, что в правильном рассуждении при истинности посылок заключение обязательно будет истинным. Применительно к естественным рассуждениям требование наличия логического следования между посылками и заключением является слишком сильным.

Во-первых, далеко не все примеры естественных рассуждений ему в полной мере соответствуют. Очень часто посылки не гарантируют истинности заключения, а лишь подтверждают его, делают более вероятным, правдоподобным. Скажем, схема рассуждения «Если А, то В. Имеет место А. Следовательно, имеет место В» не является дедуктивно корректной, но

достаточно часто используется в естественных рассуждениях. По данным исследований (см., напр.: [22]), до 60% респондентов считают такое рассуждение правильным. Более того, примерно 70% людей склонны считать правильные рассуждения (например, *Modus Tollens*) неправильными. Конечно, эти хорошо известные «когнитивные омрачения» можно считать следствием недостаточной логической грамотности, но даже такое возражение не снимает проблемы – как вести аргументацию с такими «неграмотными» оппонентами. Следовательно, адекватное моделирование аргументации требует возможности учета и подобных правдоподобных рассуждений.

Во-вторых, у отношения следования (и ассоциированного с ним синтаксического отношения выводимости) есть еще одно «лишнее» в контексте естественных рассуждений свойство – это монотонность следования. Монотонность означает возможность произвольным образом расширять множество посылок, не разрушая полученных ранее следствий:

$$\frac{\Gamma \models A}{\Gamma \cup \Delta \models A}.$$

Очевидно, что многие случаи естественных рассуждений демонстрируют нарушение этого условия. Применительно к аргументации немонотонность рассуждений проявляет себя в том, что под действием критики исходное обоснование тезиса вполне может быть разрушено благодаря новым (контр) аргументам.

Таким образом, необходимыми условиями моделирования рассуждений является возможность формальной экспликации немонотонности и правдоподобности перехода от посылок (аргументов) к заключению (тезису).

Важно отметить, что выполнение этих условий одновременно предоставляет дополнительные возможности для оценки аргументации. Так, можно предусмотреть различную качественную оценку дедуктивной и правдоподобной аргументации. Также в принципе можно отследить использование особых «немонотонных» правил, что будет свидетельствовать об эффективности критики и т.п.

Итак, возникает необходимость особым образом квалифицировать и представить формально те типы рассуждений, которые не соответствуют базовой логике (кстати, совершенно не обязательно классической), но принимаются и используются одним или несколькими участниками аргументации. Представляется, что добиться такого результата можно разными способами. Например, Дж. Поллок в своей компьютерной программе OSCAR проводит различие между «окончательными» (*conclusive*) рассуждениями и потенциально опровержимыми «рассуждениями при отсутствии доказательств в пользу противного» (*prima facie*) [23]. Для этого он использует так называемые показатели силы посылок и заключения, которые в случае дедуктивных рассуждений имеют бесконечно большое значение, а для правдоподобных следствий определяются по минимуму показателей посылок. Обобщая, можно сказать, что основные возможные стратегии распадаются на три типа: использование своеобразной «штрафной» (*penalty*) логики, когда рассужденческое действие разного типа имеет свою цену; постулирование методом грубой силы для каждого конкретного случая набора приемлемых типов рассуждений, релятивизированных относительно рационального агента и / или предметной области; разведение тем или иным способом двух отношений

типа следования – базового логического и правдоподобного. В данной работе мы попробуем представить последний вариант экспликации правдоподобных рассуждений.

Что касается немонотонности (модифицируемости рассуждений), здесь формальных экспликаций еще больше. Хороший обзор можно найти в Стэнфордской энциклопедии [24]. Ниже мы будем использовать вариант формализации «рассуждений по модулю» (см. напр.: [25, 26]). В основе этого подхода лежит очень прозрачная идея:  $A$  является следствием  $B$  по модулю множества допущений  $K$  тогда и только тогда, когда из  $K \cup B$  логически следует  $A$ . При фиксированном множестве  $K$  это отношение является вполне монотонным, модифицируемость появляется при варьировании  $K$ .

### 3. Аргументативное отношение подтверждения

Начнем с основных определений, проясняющих суть нашего подхода. Позднее вернемся к более детальному рассмотрению введенных понятий.

Итак, что означает, что некоторое утверждение  $A$  (пока для простоты изложения ограничим себя случаем однопосылочной аргументации) обосновывает (или поддерживает, подтверждает) тезис  $T$ ? С нашей точки зрения, формальная трактовка аргументативного рассуждения должна учитывать (1) возможность как дедуктивного, так и правдоподобного обоснования, (2) его немонотонность (опровержимость в свете новых аргументов) и (3) надежность использованных в обосновании аргументов (в нашем случае это отдельное высказывание  $A$ ). Как будет показано ниже, для формализации условия (3) нам потребуется ввести отношение «атаки» между высказываниями.

Таким образом,  $A$  подтверждает тезис  $T$  ( $A \Leftrightarrow T$ ), если и только если аргумент  $A$  является неотбитым (т.е. надежным –  $!A$ ) и при этом из  $A$  дедуктивно следует  $T$  ( $A \vDash T$ ) или из  $A$  правдоподобно следует  $T$  ( $A \Vdash T$ ). Аргумент является неотбитым, если на всякий атакующий его аргумент найдется контратакующий аргумент. Аргумент  $A$  атакует аргумент  $B$ , если  $A$  влечет (в широком, пока не уточненном смысле) отрицание  $B$ .

Прежде чем перейти к формальным экспликациям этих определений необходимо сделать еще одно важное замечание. Мы будем опираться на некоторое базовое отношение (дедуктивного) следования ( $\vDash$ ). В качестве такого может быть выбрано и неклассическое отношение того или иного типа, но мы на данном этапе будем рассматривать базовое следование как классическое следование между формулами языка пропозициональной логики.

DF.1.  $A \vDash_C T \Leftrightarrow A, C \vDash T$  и  $A \not\vDash T$  и  $C \not\vDash T$ .

Таким образом, отношение следования задается с использованием формализма «рассуждений по модулю» с определенными модификациями. Первая часть определяющей части просто констатирует наличие базового отношения следования между множеством, включающим посылку и принимаемое фиксированное допущение  $C$ . Вторая и третья части приближают вводимое отношение следования к релевантному: из посылки и фиксированного допущения по отдельности тезис не следует, т.е., другими словами, в обосновании тезиса должны быть использованы все посылки и допущения.

DF.2.  $A \vDash_C T \Leftrightarrow A \vDash_T C$  и  $A \not\vDash_C T$ .

Предлагаемая трактовка отношения правдоподобного следования в определенном смысле выражает интуицию о том, что правдоподобные, но дедуктивно не корректные рассуждения являются приемлемыми для многих рациональных агентов, поскольку они похожи на определенные виды дедуктивных рассуждений. Так, например, поскольку очевидно  $A \supset B \vDash_A B$ , будет иметь место  $A \supset B \Vdash_B A$ . Аналогично обстоит дело с «отрицающим» модусом условно-категорического умозаключения:  $A \supset B \vDash_{\neg B} \neg A$ , следовательно,  $A \supset B \Vdash_{\neg A} \neg B$ .

Кроме того, в силу второго условия в определяющей части вводимое отношение правдоподобного следования не обобщает дедуктивное следование, а является альтернативой ему.

DF.3.  $A \Vdash T \Leftrightarrow \exists B(A \vDash_B T \text{ или } A \Vdash_B T)$ .

Зато отношение  $\Vdash$  достаточно прозрачно обобщает понятие следования, объединяя его дедуктивный и правдоподобный варианты.

Теперь можно ввести отношение атаки в полном соответствии с высказанной ранее интуицией.

DF.4.  $A > B \Leftrightarrow A \Vdash \text{not} B$ .

Мы намеренно используем метазнак отрицания *not*, оставляя тем самым возможность для различных формальных трактовок отрицания и соответствующего уточнения определения 4.

DF.5.  $!A \Leftrightarrow \text{Если } \exists \Psi(\Psi > A \text{ и } \Psi = C), \text{ то } \exists \Phi(\Phi > C)$

Здесь  $\Psi$  и  $\Phi$  – это переменные для аргументов, а само определение в точности выражает интуицию относительно неотбитого аргумента. Теперь в нашем распоряжении есть все для формулировки главного определения отношения подтверждения (обоснования).

DF.6.  $A \Rightarrow T \Leftrightarrow !A \text{ и } A \Vdash T$ .

Кратко остановимся на некоторых важных свойствах введенных отношений. Во-первых, уже отношение следования по модулю является немонотонным. В самом деле, пусть  $A \vDash_C T$ . По определению 1 это означает, что  $A, C \vDash T$ , но при этом  $A \not\equiv T$  и  $C \not\equiv T$ . Означает ли это, что для всякого  $\Psi$  верно, что  $\Psi, A \vDash_C T$ ? Естественно для всякого  $\Psi$  выполняется  $\Psi, A, C \vDash T$  и  $C \not\equiv T$ , но при этом вполне может найтись такая формула  $\Psi$ , что  $\Psi, A \vDash T$ , в частности это верно для  $\Psi := T$ . Все это заставляет следующим образом сформулировать условие ограниченной монотонности:

$$\frac{A \vDash_C B; \Psi, A \not\equiv B}{\Psi, A \vDash_C B}$$

Во-вторых, отношение следования по модулю над базовым (классическим) следованием не совпадает с ним. Некоторые стандартные дедуктивные постулаты (например, принцип введения дизъюнкции) для следования по модулю не проходят. В общем случае имеет место следующее утверждение: Если  $A \vDash B$ , то существует такая формула  $\Psi$ , что  $A \not\equiv_{\Psi} B$ .

В-третьих, отношение подтверждения (определение 2) также не является вполне монотонным относительно расширения множества допущений. Пусть имеет место  $A \Vdash_C B$ ; в частности, это означает, что  $A \vDash_B C$ . Но этого явно недостаточно, чтобы получить  $A \Vdash_{C \& D} B$ , необходимым условием для чего является  $A \vDash_B C \& D$ .

Наконец, и само отношение подтверждения не является монотонным в силу условия неотбитого аргумента. В обобщенном случае определения от-

ношений типа следования между множествами посылок, множествами гипотез и заключением-тезисом это нарушение немонотонности выглядело бы более привычно – как запрет перехода от  $\Gamma \Rightarrow B$  к  $\Gamma \cup \Delta \Rightarrow B$ . В нашем упрощенном варианте немонотонность отношения подтверждения запрещает переходить от  $!A$  к  $!(A \& D)$ . В свою очередь, такой переход требует получения  $C \succ A$  на основании допущения  $C \succ (A \& D)$ . Однако это невозможно, поскольку имеет место только выводимость в обратную сторону, что в принципе представляется вполне естественным и соответствующим нашим интуициям относительно атаки аргументов. В самом деле, критика одного конкретного аргумента означает критику конъюнкции аргументов, но не наоборот.

## Заключение

Резюмируя, можно констатировать, что предложенные формальные экспликации аргументативных рассуждений (критики и обоснования) позволяют различить разные типы немонотонности (модифицируемости) аргументативных рассуждений. Кроме того, явным образом разграничивается несколько типов возможных отношений между посылками (аргументами) и заключением (тезисом): дедуктивное «рассуждение по модулю», правдоподобное следование и (аргументативное) отношение подтверждения.

Все это позволяет уже на данной стадии исследования ввести целый спектр оценок, основанных на специфике использованных в обосновании и критике типов рассуждений. Было ли рассуждение немонотонным и в каком смысле? Основано ли обоснование только на дедуктивном следовании или в нем использованы правдоподобные рассуждения? Ответ на первый вопрос косвенно свидетельствует о степени состоятельности полемики. Отвечая на второй вопрос, мы получаем возможность оценить «строгость» рассуждений.

Наконец, отдельной, весьма интересной темой, заслуживающей самостоятельного исследования, является возможная формализация так называемых «рассуждений по модулю».

Таким образом, намеченные в статье направления дальнейшей работы открывают широкий спектр исследований аргументации с формальных позиций.

## Литература

1. *Беликов А.А.* Логика Данна–Белнапа, ее родственники и формальное моделирование аргументации // РАЦИО.ru. 2017. Т. 18 (1). С. 36–48.
2. *Зайцев Д.В.* Теории аргументации и их практические реализации // РАЦИО.ru. 2015. № 14. С. 4–15.
3. *Зайцев Д.В.* Логика и принятие решений // РАЦИО.ru. 2013. Т. 10, доп. вып. С. 21–22.
4. *Zaitsev D.V.* Who's the Tsar? Towards the True Science of Argument and Reasoning the 7th International Conference on Cognitive Science. 2016. Vol. 13, № 2. P. 231–232.
5. *Dung P.M.* On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games // Artificial Intelligence. 1995. Vol. 77 (2). P. 321–357.
6. *Amgoud A.* Replication study of semantics in argumentation // Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, main track (10–16 August 2019). Macao, 2019. P. 6260–6266. DOI: 10.24963/ijcai.2019/874.
7. *Bonzon E. et al.* A comparative study of ranking-based semantics for abstract argumentation // Thirtieth AAAI Conference on Artificial Intelligence (12–17 February 2016). Phoenix, 2016. P. 8.

8. *Besnard P., Hunter A.* A logic-based theory of deductive arguments // *Artificial Intelligence*. 2001. Vol. 128 (1-2). P. 203–235.
9. *Leite J., Martins J.* Social abstract argumentation // *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (19–22 July 2011)*. Barcelona, 2011. P. 2287–2292.
10. *Amgoud L., Ben-Naim J.* Ranking-based semantics for argumentation frameworks // *Proceedings of the 7th International Conference on Scalable Uncertainty Management (16–18 September 2013)*. Washington, 2013. P. 134–147.
11. *Cayrol C., Lagasque-Schiex M.* Graduality in argumentation // *Journal of Artificial Intelligence Research*. 2005. Vol. 23. P. 245–297.
12. *Belnap N.* A useful four-valued logic // *Modern Uses of Multiple-Valued Logic* / ed. by J.M. Dunn, G. Epstein. Boston : Reidel Publishing Company, 1977. P. 7–37.
13. *Belnap N.* How a computer should think // *Contemporary Aspects of Philosophy* / ed. by G. Rule. Stocksfield : Oriel Press, 1977. P. 30–56.
14. *Dubois D.* On Ignorance and Contradiction Considered as Truth-Values // *Logic Journal of the IGPL*. Vol. 16. (2). P. 195–216. DOI: 10.1093/jigpal/jzn003.
15. *Amgoud L., Ben-Naim J.* On Argumentation-based Paraconsistent Logics // *Computational Models of Rationality*. London : College Publications, 2016. P. 377–391.
16. *Ben-Naim J.* Argumentation-Based Paraconsistent Logics // *21st International Conference on Conceptual Structures (27–30 July 2014)*. Iasi, 2014. P. 19–24.
17. *Arieli O.* A sequent-based representation of logical argumentation // *International Workshop on Computational Logic in Multi-Agent Systems (16–18 September 2013)*. Corunna, 2013. P. 69–85.
18. *Borg A., Strasser C.* Relevance in structured argumentation // *Proceedings of the 27th International Joint Conference on Artificial Intelligence (13–19 July 2018)*. Stockholm, 2018. P. 1753–1759. DOI: 10.24963/ijcai.2018/242.
19. *Borg A., Arieli O., Strasser C.* Hypersequent-Based Argumentation: An Instantiation in the Relevance Logic RM // *International Workshop on Theories and Applications of Formal Argumentation (19–20 August 2017)*. Melbourne, 2017. P. 17–34.
20. *Borg A., Arieli O.* Hypersequential Argumentation Frameworks: an Instantiation in the Modal Logic S5 // *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (10–15 July 2018)*. Stockholm, 2018. P. 1097–1104.
21. *Mercier H., Sperber D.* Why do humans reason? Arguments for an argumentative theory // *Behavioral and Brain Sciences*. 2011. Vol. 34. P. 57–74.
22. *Evans J., Over D.* Rationality and reasoning. Psychology Press, 2013. 192 p.
23. *Pollock J.* Oscar: An architecture for generally intelligent agents // *Frontiers in Artificial Intelligence and Applications*. 2008. Vol. 171 (1). P. 275–286.
24. *Strasser C., Antonelli G.* Non-monotonic Logic // *The Stanford Encyclopedia of Philosophy* (Summer 2019 ed.) / E.N. Zalta (ed.) URL: <https://plato.stanford.edu/entries/logic-nonmonotonic/> (accessed: 01.05.2020).
25. *Gärdenfors P., Makinson D.* Nonmonotonic inference based on expectations // *Artificial Intelligence*. 1994. Vol. 65 (2). P. 197–245.
26. *Makinson D.* Bridges between classical and nonmonotonic logic // *Logic Journal of IGPL*. 2003. Vol. 11 (1). P. 69–96.

**Dmitry V. Zaitsev**, Saint Petersburg State University (Saint Petersburg, Russian Federation); Lomonosov Moscow State University (MSU) (Moscow, Russian Federation).

E-mail: [zaitsev@philos.msu.ru](mailto:zaitsev@philos.msu.ru)

**Alexander A. Belikov**, Saint Petersburg State University (Saint Petersburg, Russian Federation); Lomonosov Moscow State University (MSU) (Moscow, Russian Federation).

E-mail: [belikov@philos.msu.ru](mailto:belikov@philos.msu.ru)

*Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya – Tomsk State University Journal of Philosophy, Sociology and Political Science*. 2020. 57. pp. 13–24.

DOI: 10.17223/1998863X/57/2

#### **MODELLING ARGUMENTATION: VALUATIONS AND REASONING**

**Keywords:** argumentation; argumentative reasoning; evaluation of arguments; formal models of argumentation.

This article discusses the formal modelling of argumentation. The fundamental characteristics of argumentative reasoning, which play the key role in evaluating arguments, are examined. The authors focus on the distinctive features of natural reasoning. Real practice of reasoning is frequently accom-

panied by usage of deductively incorrect arguments as justifications of certain statements that arise within polemical situations. Another peculiarity of natural reasoning the authors analyse consists in its nonmonotonicity, i.e., extending the set of premises of a certain valid argument affects its validity. Thus, the possibility of constructing a simple and intuitively plausible formal theory that allows grasping the aforementioned phenomena is the main aim of the article. The authors suggest a formal explication of the evidence relation between an argument and a thesis. This explication can take into account the variability of the entailment relation, i.e., allows dealing with both deductive and inductive arguments within the same framework. The suggested explication of evidence relation is substantially grounded on the notion of reasoning by modulo, which, in turn, is a suitable tool for analysing deductively incorrect arguments. The suggested explication of argumentative reasoning (critique and refutation) allows distinguishing between different kinds of nonmonotonicity (modifiability) of argumentative reasoning and, furthermore, drawing the distinction between several types of relations between premises (arguments) and conclusion (thesis): deductive reasoning by modulo, inductive consequence, and (argumentative) evidence relation.

### References

1. Belikov, A.A. (2017) Logika Danna-Belnapa, ee rodstvenniki i formal'noe modelirovanie argumentatsii [Dunn-Belnap's logic, its relatives and formal modeling of argumentation]. *RATsIO.ru*. 18(1). pp. 36–48.
2. Zaytsev, D.V. (2015) Teorii argumentatsii i ikh prakticheskie realizatsii [Theories of argumentation and their practical implementation]. *RATsIO.ru*. 14. pp. 4–15.
3. Zaytsev, D.V. (2013) Logika i prinyatie resheniy [Logic and decision making]. *RATsIO.ru*. 10. pp. 21–22.
4. Zaitsev, D.V. (2016) Who's the Tsar? Towards the True Science of Argument and Reasoning. *The Seventh International Conference on Cognitive Science*. 13(2). pp. 231–232.
5. Dung, P.M. (1995) On the acceptability of arguments and its fundamental role in nonmonotonic rea-soning, logic programming and n-person games. *Artificial Intelligence*. 77(2). pp. 321–357. DOI: 10.1016/0004-3702(94)00041-X
6. Amgoud, A. (2019) Replication study of semantics in argumentation. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. Main Track. August 10–16, 2019. Macao. pp. 6260–6266. DOI: 10.24963/ijcai.2019/874
7. Bonzon, E. et al. (2016) A comparative study of ranking-based semantics for abstract argumentation. *Thirtieth AAAI Conference on Artificial Intelligence*. Phoenix, February 12–17, 2016. p. 8.
8. Besnard, P. & Hunter, A. (2001) A logic-based theory of deductive arguments. *Artificial Intelligence*. 128 (1–2). pp. 203–235. DOI: 10.1016/S0004-3702(01)00071-6
9. Leite, J. & Martins, J. (2011) Social abstract argumentation. *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*. Barcelona, July 19–22, 2011. pp. 2287–2292.
10. Amgoud, L. & Ben-Naim, J. (2013) Ranking-based semantics for argumentation frameworks. *Proceedings of the 7th International Conference on Scalable Uncertainty Management*. Washington, September 16–18, 2013. pp. 134–147.
11. Cayrol, C. & Lagasque-Schiex, M. (2005) Graduality in argumentation. *Journal of Artificial Intelligence Research*. 23. pp. 245–297. DOI: 10.1613/jair.1411
12. Belnap, N. (1977) A useful four-valued logic. In: Dunn, J.M. & Epstein, G. (eds) *Modern Uses of Multiple-Valued Logic*. Boston: Reidel Publishing Company. pp. 7–37.
13. Belnap, N. (1977) How a computer should think. In: Rule, G. (ed.) *Contemporary Aspects of Philosophy*. Stockfield: Oriel Press. pp. 30–56.
14. Dubois, D. (2008) On Ignorance and Contradiction Considered as TruthValues. *Logic Journal of the IGPL*. 16(2). pp. 195–216. DOI: 10.1093/jigpal/jzn003
15. Amgoud, L. & Ben-Naim, J. (2016) On Argumentation-based Paraconsistent Logics. In: Beierle, Ch., Brewka, G. & Thimm, M. (eds) *Computational Models of Rationality*. London: College Publications. pp. 377–391.
16. Ben-Naim, J. (2014) Argumentation-Based Paraconsistent Logics. *21st International Conference on Conceptual Structures*. Iasi, July 27–30, 2014. pp. 19–24.
17. Arieli, O. (2013) A sequent-based representation of logical argumentation. *International Workshop on Computational Logic in Multi-Agent Systems*. Corunna, September 16–18, 2013. pp. 69–85.
18. Borg, A. & Strasser, C. (2018) Relevance in structured argumentation. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. Stockholm, July 13–19, 2018. pp. 1753–1759. DOI: 10.24963/ijcai.2018/242

19. Borg, A., Arieli, O. & Strasser, C. (2017) Hypersequent-Based Argumentation: An Instantiation in the Relevance Logic RM. *International Workshop on Theories and Applications of Formal Argumentation*. Melbourne, August 19–20, 2017. pp. 17–34.
20. Borg, A. & Arieli, O. (2018) Hypersequential Argumentation Frameworks: An Instantiation in the Modal Logic S5. *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. Stockholm, July 10–15, 2018. pp. 1097–1104.
21. Mercier, H. & Sperber, D. (2011) Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*. 34. pp. 57–74. DOI: 10.1017/S0140525X10000968
22. Evans, J. & Over, D. (2013) *Rationality and Reasoning*. Psychology Press.
23. Pollock, J. (2008) Oscar: An architecture for generally intelligent agents. *Frontiers in Artificial Intelligence and Applications*. 171.
24. Strasser, C. & Antonelli, G. (2019) Non-monotonic Logic. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*. [Online] Available from: <https://plato.stanford.edu/entries/logic-nonmonotonic/> (Accessed: 1st May 2020).
25. Gärdenfors, P. & Makinson, D. (1994) Nonmonotonic inference based on expectations. *Artificial Intelligence*. 65(2). pp. 197–245. DOI: 10.1016/0004-3702(94)90017-5
26. Makinson, D. (2003) Bridges between classical and nonmonotonic logic. *Logic Journal of IGPL*. 11(1). pp. 69–96. DOI: 10.1093/jigpal/11.1.69