УДК 164.03:811.1/8:004.838 Вестник СПбГУ. Философия и конфликтология. 2026. Т. 42. Вып. 1

# Разрешение споров при помощи абстрактных аргументационных структур и его автоматизация на примере научного рецензирования

*Е. Н. Лисанюк[1,2], Д. Е. Прокудин[3], И. Р. Баймуратов[4]*

[1] Институт философии РАН,
Российская Федерация, 109240, Москва, ул. Гончарная, 12
[2] Национальный исследовательский университет «Высшая школа экономики»,
Российская Федерация, 101000, Москва, ул. Мясницкая, 20
[3] Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9
[4] Ганноверский университет имени Лейбница,
Германия, 30167, Ганновер, Велфенгартен, 1

**Для цитирования:** *Лисанюк Е. Н., Прокудин Д. Е., Баймуратов И. Р.* Разрешение споров при помощи абстрактных аргументационных структур и его автоматизация на примере научного рецензирования // Вестник Санкт-Петербургского университета. Философия и конфликтология. 2026. Т. 42. Вып. 1. С. 118–130. https://doi.org/10.21638/spbu17.2026.110 EDN MZHSFO

В статье рассмотрено научное рецензирование как единичный смешанный аргументативный спор между авторами рукописи, направляемой для публикации, и рецензентами, оценивающими ее по определенным критериям. Для нахождения его решений авторы применяют формализм, основанный на абстрактных аргументационных структурах (ААС) с доверительной семантикой расширений. Предложенный метод нахождения решений споров в ходе научного рецензирования состоит из: 1) разметки рецензии; 2) конвертации разметки в JSON; 3) генерации ее OWL-представления; 4) выявления решения спора путем установления предпочтительного расширения, т. е. максимального подмножества приемлемых аргументов, выдвинутых авторами и рецензентами. В результате были разработаны схема разметки научных рецензий и алгоритм генерации их OWL-представлений, обеспечивающие формализацию аргументативных споров в ходе научного рецензирования с помощью ААС и нахождение их решений посредством автоматического логического вывода. Авторы валидировали предложенный метод с помощью разметки корпуса рецензий и применения к нему proof-of-concept реализации алгоритма. Результаты оценки предложенной схемы разметки на основе согласованности разметок, выполненных различными разметчиками, показали, что она обеспечивает однозначную разметку аргументов авторов, тогда как разметка аргументов рецензентов зависит от интерпретации разметчиков. Применение искусственного интеллекта ускоряет процесс научного рецензирования, но имеет ряд ограничений, включая воспроизводство предвзятостей, заложенных в обучающих данных. Предложенный метод представления и анализа аргументации вносит вклад в повышение качества научного рецензирования.

*Ключевые слова:* научное рецензирование, решение споров, аргументационная структура, разметка текста, представление знаний, OWL, автоматический логический вывод.

© Е. Н. Лисанюк, Д. Е. Прокудин, И. Р. Баймуратов, 2026

## Введение

Процесс рецензирования научных публикаций затруднен ростом объема подаваемых материалов [1], предубеждениями участников рецензирования [2], конфликтами интересов в академической среде и снижением качества рецензирования [3]. Инициативы по использованию искусственного интеллекта (ИИ) для преодоления этих и иных трудностей в научном рецензировании сопряжены с рисками ненадежности цифровых сервисов и воспроизводства предвзятостей в системах ИИ [4]. Наш вклад в преодоление подобных затруднений состоит в логическом и информационно-технологическом рассмотрении процесса рецензирования как аргументативного диалога, или спора определенного типа, между авторами рукописи, направляемой для публикации в научный журнал или в оргкомитет конференции, и рецензентами, оценивающими ее по критериям, предусмотренным содержанием рецензии, шаблон которой задает формат аргументирования.

Для оценки аргументации мы используем схему операций поиска и отбора решений споров (СПОР) [5], обеспечивающую логическое выведение результата — решения рецензентов принять рукопись, отклонить или отправить авторам на доработку, поддержанного наиболее сильными аргументами, с помощью абстрактных аргументационных структур (ААС) — одного из ведущих современных формальных методов анализа аргументации [6], рассматривая научное рецензирование как модифицируемую, или отменяемую (defeasible), аргументацию [7], на основе доверительной семантики [8], разработанной для немонотонной логики умолчаний (default logic) [9] и позволяющей выводить заключения в условиях неполной и неконсистентной информации.

Для автоматизации СПОРа мы воспользовались корпусом рецензий MDPI [10], разметили их согласно предложенной схеме аннотирования, что позволило автоматически конвертировать размеченные рецензии в текстовый формат JSON (JavaScript Object Notation), применяемый для хранения и передачи структурированных данных в интернете, сгенерировать репрезентации формализованных с помощью ААС споров на языке описания онтологий OWL (Web Ontology Language), и установить решения споров с помощью автоматического логического вывода. Во избежание наследования предвзятости предложенную схему разметки мы оценили на основе согласованности разметок, выполненных двумя различными разметчиками, а также подсчитали характеристики ААС, сгенерированные на основе разметок. Таким образом, предложенный нами метод автоматизации СПОРа на примере научного рецензирования состоит из следующих этапов: 1) разметка рецензии; 2) конвертация размеченной рецензии в формат JSON; 3) генерация OWL-репрезентации; 4) нахождение решения спора помощью автоматического логического вывода. Здесь мы излагаем формально-философскую часть этапа 4 и даем краткое описание этапов 1–3 без технической детализации, которая дана в [11].

## Научное рецензирование и искусственный интеллект

На преодоление вышеуказанных трудностей научного рецензирования нацелен ряд альтернативных систем рецензирования [12–14]. Для снижения трудоемкости первичного отбора поданных материалов используют ИИ, чтобы автомати-

зировать их распределение между рецензентами [15], прогнозировать результаты рецензирования [16; 17], улучшить редакционные стратегии [18], а также оценить согласованность статистических данных авторов [19]. Проверять соответствие рукописи формальным требованиями помогают Penelope.ai[1] и StatReviewer[2].

Использование подобных подходов влечет этические проблемы, связанные с рисками воспроизводства предвзятости и непрозрачностью в системах ИИ [4].

### *Представление аргументации*

К основным областям применения технологий ИИ к аргументации относят: извлечение аргументов (argument mining), анализ и представление аргументации (argument analysis and representation), поддержку аргументации и принятия решений (argument assistance and decision support), мультиагентные системы (multiagent systems). Мы рассматриваем методы формального анализа и представления аргументации в качестве инструмента для обеспечения объяснимости и непредвзятости в рецензировании, представляем рецензирование как аргументативный спор между авторами рукописи и ее рецензентами, решение которого возможно установить через формализацию и автоматизацию оценки их аргументации.

В настоящее время существует большое количество программного обеспечения, предназначенного для репрезентации и поддержки аргументации. Наиболее распространенным форматом репрезентации аргументации является формальная онтология AIF (Argument Interchange Format), представляющая аргументацию в виде направленных графов, называемых аргументационными сетями, где узлы представляют собой утверждения, или паттерны аргументации [20]. Созданы различные реализации AIF, в том числе и в формате RDF (Resource Description Framework), для которых была разработана репрезентация на языке OWL DL [21; 22]. Выразительные возможности AIF могут быть усилены с учетом схем аргументации и других ее аспектов [23]. AIF используется в OVA — онлайн-приложении для анализа аргументации [24].

В области юридической аргументации используется LKIF (Legal Knowledge Interchange Format), включающий онтологию для юридических баз знаний и формат обмена сообщениями и реализованный не в синтаксисе OWL DL, а как расширение SWRL с поддержкой отменяемых рассуждений [25], или в формате XML в системе Carneades [26]. Предложена и формализована на языке OWL DL модель представления законодательства с различными типами и атрибутами постановлений, вместе с доменными онтологиями используемая для структурной разметки юридических текстов, в которой можно логически выводить косвенные отношения между постановлениями, запрашиваемыми с помощью языка SPARQL [27]. Отдельное направление формальной репрезентации аргументации составляют ее автоматические решатели (argumentation solvers), реализуемые на языках C+++ [28] и Python [29], высокие вычислительные способности которых, однако, вряд ли применимы к решению нашей задачи из-за того, что такие решатели не подразумевают

---

[1] Penelope. ai. URL: https://www.penelope.ai/ (дата обращения: 02.04.2025).

[2] *A peerless review? Automating methodological and statistical review*. URL: http://blogs.biomedcentral.com/bmcblog/2016/05/23/peerless-review-automating-methodological-statistical-review (дата обращения: 02.04.2025).

формализации аргументации, выраженной в естественном языке, и абстрагируются от ее метаданных, важных в научном рецензировании.

Наиболее близким к нашей разработке является приложение PyArg, в котором на языке Python реализованы формализмы и алгоритмы на основе AAC и ASPIC+ [30]. PyArg обеспечивает импорт и экспорт аргументации в различных форматах: JSON, ASPARTIX (.APX), Trivial Graph (.TGF), и визуализирует схемы аргументации без обращения к стандартам семантического веба, таким как RDF и OWL.

### *Аргумент, аргументация и спор*

Под аргументом мы понимаем рассуждение или умозаключение, обосновывающее какое-либо утверждение, — его заключение путем выведения его из других утверждений, или посылок, а термином «аргументативный спор» называем диалог, где стороны — рациональные агенты, обнаружив расхождение во мнениях по поводу истинности каких-либо предложений, приводят аргументы для его разрешения или уточнения путем обоснования своей точки зрения или критики других точек зрения. Расхождение во мнениях — необходимое условие для начала спора. В его отсутствие — в случае, если рецензенты решают принять рукопись как есть, нет смысла спорить и приводить аргументы, в отличие от требующих обоснования решений отклонить ее или вернуть на доработку.

В зависимости от количества высказываний, составляющих пропозициональное содержание точек зрения сторон, споры делятся на единичные и множественные, а в зависимости от намерений сторон — играть роль аудитории, не выдвигая своей точки зрения, но сомневаясь в точках зрения других, отстаивать свою точку зрения, критиковать другие мнения или делать и то, и другое, — соответственно, на несмешанные и смешанные [31]. В данном исследовании мы рассматриваем аргументативные споры в научном рецензировании как единичные смешанные споры, где одна сторона при помощи аргументов защищает свою точку зрения по поводу истинности предложения $P$, выступая с положительной точкой зрения: $T+(P)$, а другая сторона с этим не согласна и критикует ее и аргументы, предъявленные в ее поддержку, т. е. выдвигает отрицательную точку зрения: $T–(P)$, но помимо этого не выдвигает противоположной положительной точки зрения, что $P$ неверно, $T+(\neg P)$, или аргументов в ее защиту, как в смешанном множественном споре. Единичный смешанный спор подразумевает, что в итоге утверждение об истинности $P$ будет либо защищено, либо отброшено, но не подразумевает, что будет защищено $\neg P$, что позволяет для моделирования такого спора использовать доверительную семантику для немонотонной логики, обладающую рядом свойств, обеспечивающих решение нашей задачи. Она позволяет выводить заключения в условиях недостаточности или неконсистентности информации в посылках, например, когда разные рецензенты одной рукописи выносят противоположные решения; допускает вывод компромиссного решения, образованного совместимыми между собой заключениями наиболее сильных аргументов спора независимо от того, кто их выдвинул; и, наконец, гарантирует, что множество решений спора будет непустым, а также уникальным, с учетом того, что всякий процесс рецензирования должен завершиться каким-либо решением по поводу рукописи.

### *Абстрактные аргументационные структуры*

Формальный аппарат СПОР основан на ААС, принимающих неструктурированный аргумент за атомарную единицу аргументативного спора, а за его наименьшую молекулярную единицу — пару, состоящую из аргумента и возражения против него. Абстрактные аргументационные структуры формально эксплицируют ключевую содержательную идею аргументативного рассуждения: убежден или нет рациональный агент в некоем положении, зависит от того, можно ли его «отстоять перед лицом атакующих аргументов» [6, с. 323].

Использующая ААС логика аргументации трактует процесс убеждения как обмен рассуждениями между рациональными агентами и «изучает критерии… определяющие степень, в которой разумно принимать логически выведенные заключения, хотя некоторые из этих заключений устанавливаются недедуктивным рассуждением. Иногда разумно принять выведенное из посылок заключение, даже если этот аргумент не является сильным, и его недостаточно, чтобы установить выведенное заключение с абсолютной уверенностью» [32, с. 225]. Для моделирования степеней подобной разумности в логике аргументации полагаются на вычислительные семантики, одной из которых, наряду с семантикой маркировок [33], является используемая нами семантика расширений (extensions) относительно теоретико-множественного включения [34], первоначально предложенная в логике умолчаний [9].

Отвлекаясь от внутреннего строения аргументов и от изучаемой в формальной логике демонстрации как способа подтверждения заключения посредством посылок, ААС представляют последнее производным от упорядочения множества аргументов на графе при помощи бинарного абстрактного отношения атаки, несимметричного и нетранзитивного, символизирующего критику или контраргументацию, принимаемую за базовое отношение между аргументами в споре.

**Определение 1.** Аргументационная структура $AF$ — это пара

$$AF = \langle AR, attacks \rangle,$$

где $AR$ — множество аргументов и $attacks \subseteq AR \times AR$.

Будем говорить, что аргумент $\alpha$ атакует аргумент $\beta$, или что $\beta$ атакован аргументом $\alpha$, если выполняется отношение $attacks(\alpha, \beta)$. Аналогично, мы говорим, что множество аргументов $S \subseteq AR$ атакует α, или что $\alpha$ атакован $S$, если некоторый аргумент из $S$ атакует α.

**Пример 1.** Пусть $attacks(\alpha, \beta)$. Если найдется аргумент $\gamma$, критикующий $\alpha$, т. е. выполняется $attacks(\gamma, \alpha)$, то аргумент $\beta$ возвращен в $AR$ в качестве защищенного, как на рис. 1, где отношение атаки обозначено сплошными линиями стрелок, а производное от них отношение поддержки — пунктирной линией. Говорить, что $\alpha$ атакует $\beta$ — то же самое, что сказать, что $\beta$ отклонен $\alpha$, в соответствии с абстрактной трактовкой $attacks(\alpha, \beta)$, отождествляющей атаку и отклонение, которой мы здесь придерживаемся. Выразительные возможности ААС можно усилить, обусловив способность $\alpha$ успешно атаковать и отклонить $\beta$ дополнительными требованиями для $\alpha$ быть сильнее $\beta$, например, по защищенности, о которой скажем ниже, способу демонстрации, модальности и т. д. Аналогичным образом могут быть уточнены свойства отношений атаки и поддержки.

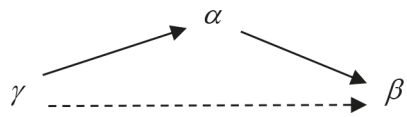


*Рис. 1.* Пример 1

Упорядочение множества аргументов $\alpha$, $\beta$ и $\gamma$ при помощи отношения атаки выражает расхождение во мнениях — заключениях аргументов $\alpha$ и $\beta$ между сторонами спора, когда для того, чтобы защитить свою точку зрения, агент — автор аргумента $\beta$, отклоненного аргументом $\alpha$, выступает с контраргументацией при помощи нового аргумента $\gamma$, отклоняющего $\alpha$. Можно понимать $\beta$ как выдвижение в споре аргумента вида «*P*, потому что *Q*», и тем самым — положительной точки зрения по поводу предложения *P*, *T+(P)*; а выдвижение аргумента $\alpha$ — как возражение вида «Но *Q* неверно и не поддерживает *P*», т. е. как отрицательную точку зрения по поводу предложения *P*, *T–(P)*; и, наконец, выдвижение нового аргумента $\gamma$ — как контраргумент вида «Однако *P*, потому что *R*», критикующего $\alpha$ и защищающего $\beta$.

Для того чтобы определить исходы и затем отобрать из них решения споров, определим понятия приемлемого аргумента и расширений *AR*.

**Определение 2.** Аргумент $\alpha \in AR$ является приемлемым (acceptable) по отношению к множеству аргументов $S \subseteq AR$, если и только если для каждого аргумента $\beta \in AR$ выполняется, что если *attacks(β, α)*, то *attacks(S, β)*.

В примере 1 приемлемыми являются $\beta$ и $\gamma$, но не $\alpha$, т. е. всякий приемлемый аргумент либо не атакован, либо защищен.

**Определение 3.** Множество аргументов $S \subseteq AR$ называется бесконфликтным (conflict-free) $S^{cf}$, если в *S* не найдется аргументов $\alpha$ и $\beta$, таких что *attacks(α, β)*.

Не подразумевающее расхождения во мнениях $S^{cf}$ совместимо с любым подмножеством *AR* данного спора и выражает часть позиции агента спора, которую можно считать последовательной, но оно непригодно в качестве исхода и решения спора. Наименьшей степенью разумности, характеризующей аргументы, убедительные для рационального агента, обладает состоящее из приемлемых аргументов допустимое подмножество.

**Определение 4.** $S^{cf}$ является допустимым (admissible) $S^{adm}$, если и только если каждый аргумент в *S* является приемлемым по отношению к *S*.

Используемая нами СПОР задачу установления исходов спора сводит к поиску допустимых подмножеств, а задачу установления решения единичного смешанного спора — к нахождению максимального из них [5].

**Определение 5.** $S^{adm}$ является предпочтительным (preferable) расширением $S^{pref}$, если и только если оно максимальное (по отношению к теоретико-множественному включению) допустимое множество *AR*.

В примере 1 (рис. 1) имеется несколько допустимых подмножеств: *{β}*, *{γ}*, *{γ, β}*, а предпочтительным расширением является только *{γ, β}*. Предпочтительных расширений *AR* может быть более одного, поскольку *AR* не обязательно конечно: например, оно может содержать бесконечную цепочку аргументов $\alpha_0$, $\alpha_1$,…,$\alpha_{n-1}$, $\alpha_n$…, каждый из которых атакован последующим, т. е. для всякого $\alpha_{i+1}$ верно, что *attacks*($\alpha_{i+1}, \alpha_i$), либо в нем аргументы могут взаимно атаковать друг друга, как в примере 2 (рис. 2), изображающем новый раунд спора из примера 1.

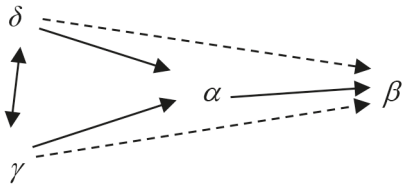


*Рис. 2.* Пример 2

Вместе с тем количество раундов рецензирования всегда конечно и должно привести к какому-либо решению по поводу рукописи, поэтому требуется, чтобы множество аргументов, составляющих решение спора, было непустым и единственным. Первое требование выполняется благодаря доверительной семантике, так как для того, чтобы в *AR* нашелся хотя бы один приемлемый аргумент, достаточно, чтобы имело место $S^{pref} \subseteq AR$, исключающее атакованные и незащищенные аргументы, как на примере 1, который можно понимать как завершение спора на втором раунде принятием рукописи после однократной доработки авторами. Требованию к авторам дать ответы на все без исключения замечания рецензентов, часто выступающему одним из правил рецензирования, соответствуют стабильное или полное расширения.

**Определение 6.** $S^{cf}$ является стабильным (stable) расширением $S^{stab}$, если и только если оно атакует всякий не принадлежащий ему аргумент, т. е. для всякого аргумента $\alpha \notin S$ верно, что $attacks(S,\alpha)$.

**Определение 7.** $S^{adm}$ является полным (complete) расширением $S^{compl}$, если и только если ему принадлежит всякий приемлемый относительно $S^{adm}$ аргумент.

Второе требование помогает выполнить скептическая семантика, устанавливающая прочное (grounded) расширение *AR*, которое является единственным для любого *AR* и позволяет выводить демонстративные заключения, но может оказаться пустым.

**Определение 8.** $S^{compl}$ является прочным расширением $S^{gr}$, если и только если оно минимальное (по отношению к теоретико-множественной операции включения) полное множество *AR*.

В силу конечности множества раундов рецензирования нас интересуют только конечные *AR*, а для них существует только одно полное расширение, которое является также прочным, предпочтительным и стабильным, как доказано в [6]. В примере 1 $\{\gamma, \beta\}$ является стабильным, прочным и полным расширением, а в примере 2 стабильных и прочных расширений два — $\{\beta,\gamma\}$ и $\{\beta,\delta\}$, но из-за циклической атаки между $\gamma$ и $\delta$ полное расширение пусто. Таким образом, поскольку рецензирование должно завершиться либо отклонением, либо принятием рукописи после доработки за конечное количество раундов, постольку для установления исходов такого спора необходимо найти все приемлемые аргументы, а для установления решения единичного смешанного спора, каковым мы считаем спор между авторами и рецензентами, достаточно найти его предпочтительное расширение, которое также будет полным и прочным. Это значит, что СПОР гарантирует вывод решения и его уникальность.

### *Формализация научного рецензирования с помощью абстрактных аргументационных структур*

Мы разработали алгоритм генерации представлений AAC на языке OWL DL, получающий в качестве входных данных множества аргументов и отношения атаки [11]. Алгоритм состоит из трех шагов.

Шаг 1. Сначала формируются базовые сущности, включая наборы аргументов, сами аргументы, отношения ‘attacks’ и обратное отношение ‘isAttackedBy’, необходимые для формирования допустимых подмножеств. Здесь и далее для ясности пользовательские сущности, определенные на языке OWL, указаны в единичных кавычках. Шаг 1 имеет линейную сложность, так как он только единожды проходит по множествам аргументов и по парам атак.

Шаг 2 «закрывает» аргументы относительно отношения ‘attacks’ и ‘isAttackedBy’ для обеспечения логического вывода в рамках предположения об открытости мира (open world assumption). Этот шаг также имеет линейную сложность в зависимости от количества аргументов.

Шаг 3. На этом этапе для каждого входного множества аргументов формируются бесконфликтные подмножества и подмножества допустимых аргументов и с помощью автоматического логического вывода классифицируются аргументы по ним.

Поскольку формирование бесконфликтного подмножества для некоторого множества аргументов при открытом мире требует построения списка других множеств аргументов, данный шаг имеет квадратичную сложность относительно количества множеств аргументов.

Для проверки алгоритма был разработан proof-of-concept. Он был реализован на языке Python с использованием библиотеки Owlready2 для работы с языком OWL и механизма логического вывода Pellet [35]. Входные данные для этого proof-of-concept представляются в формате JSON.

Мы рассматриваем научное рецензирование как единичный смешанный аргументативный спор между авторами рукописи, направляемой для публикации в научный журнал или для выступления в оргкомитет конференции, и рецензентами, оценивающими ее по определенным критериям, как-то: новизна и оригинальность, обоснованность и актуальность отстаиваемых в ней положений, релевантность источников, соответствие правилам оформления, комплектность и т. п. Подобные критерии можно понимать как шаблон для аргументативных рассуждений авторов рукописей и их научных рецензентов, регламентирующий количество, содержание и строение аргументов.

Посредством направления своей рукописи авторы выступают с аргументативным рассуждением, в котором заявляют, что она достойна быть принятой, потому что отвечает всем предъявляемым требованиям, т. е. выдвигают положительное мнение $T+(P)$ по поводу истинности предложения

($P$) [Нашу] рукопись $X$ стоит принять [для публикации в журнале $Z$].

а также аргументы в его защиту, вроде

[Наша] рукопись $X$ отвечает всем требованиям $Z$.

Посредством оценивания рукописи рецензенты в своем решении могут согласиться с таким мнением авторов, и тогда их рукопись принимают как есть, либо возразить против него и отклонить рукопись. В этих случаях диалог авторов и рецензентов завершается на первом раунде. Рецензенты могут возразить авторам и путем указания на необходимость доработки рукописи ввиду неполного соответствия требованиям, и тогда диалог продолжается, по меньшей мере, еще один раунд, в котором авторы отвечают рецензентам новым аргументативным рассуждением, сообщая о произведенных в рукописи исправлениях, а рецензенты в своем аргументативном рассуждении оценивают его по тому же шаблону. Мы рассматриваем подобные аргументативные диалоги, включающие не менее двух раундов, необходимых в случаях условного принятия рукописи и ее возвращения авторам для доработки, и трактуем возражения, замечания и рекомендации рецензентов как аргументативные рассуждения, содержащие отрицательную точку зрения по поводу истинности предложения ($P$): $T–(P)$, а также критику аргументов авторов.

$T–(P)$ Рукопись $X$ отвечает не всем требованиям $Z$ и нуждается в доработке, согласно требованиям $z_1$, $z_2$…, и т. д.

Это позволяет представить научное рецензирование при помощи ААС как единичный смешанный спор, в котором рецензенты нацелены критиковать мнение и аргументы авторов, но не выдвигают противоположного мнения о предложении $P$ и, соответственно, не намереваются его защищать посредством аргументов в качестве положительной точки зрения.

Таким образом, предлагаемый алгоритм разрешения споров при научном рецензировании состоит из следующих этапов:

1) разметка рецензии с помощью ААС;
2) конвертация размеченной рецензии в JSON формат;
3) генерация OWL репрезентации полученного ААС;
4) классификация аргументов по допустимым подмножествам для нахождения предпочтительного расширения.

Для оценки предложенного метода разрешения споров при научном рецензировании мы предложили схему разметки рецензий, которая обеспечивает их автоматическую конвертацию в JSON формат, используемый для генерации OWL-представлений ААС [11], разметили корпус рецензий, реализовали proof-of-concept, сгенерировали OWL-представления размеченных рецензий и нашли предпочтительное расширение для каждого из них. Для валидации предложенной схемы разметки и метода в целом был выбран открытый корпус рецензий от MDPI [10], содержащий 123 статьи с рецензиями, доступные по состоянию на 16 июня 2022 г. Для оценки предложенной схемы разметки каждая рецензия размечалась дважды, согласованность между различными разметками одних и тех же рецензий была подсчитана с помощью коэффициента Криппендорфа. Анализ полученных результатов позволяет сделать вывод, что разметки имеют высокую корреляцию для количества множеств аргументов, количества аргументов автора, количества аргументов рецензентов и количества допустимых аргументов автора. При этом количество допустимых аргументов рецензентов имеют невысокий уровень корреляции, а средняя длина цепочек атак между аргументами имеет отрицательную

корреляцию [36]. Таким образом, можно сделать вывод, что предложенная схема обеспечивает согласованную разметку аргументов авторов, тогда как разметка аргументов рецензентов зависит от интерпретации разметчиков.

## Заключение

Мы рассмотрели научное рецензирование как единичный смешанный аргументативный спор между авторами рукописи, направляемой для публикации в научный журнал или в оргкомитет конференции, и рецензентами, оценивающими ее по определенным критериям, и предложили формализовать научное рецензирование с помощью AAC и доверительной семантики расширений. Предложенная схема обеспечивает автоматическую конвертацию размеченной рецензии в формат JSON, который используется для генерации представления AAC на языке OWL, обеспечивающее решение споров с помощью автоматического логического вывода. Разработанный нами метод решения споров при научном рецензировании состоит из следующих этапов: 1) разметка рецензии; 2) конвертация размеченной рецензии в формат JSON; 3) генерация OWL-представления и 4) нахождение решений спора. Мы разработали алгоритм генерации JSON файлов на основе размеченных рецензий, формализовали их с помощью AAC и представили на языке OWL. Для оценки предложенного метода решения споров при научном рецензировании мы реализовали proof-of-concept, разметили корпус рецензий, сгенерировали OWL-репрезентации размеченных рецензии и нашли их предпочтительные расширения. Для валидации предложенной схемы разметки каждая рецензия размечалась двумя разметчиками. Результаты показали, что предложенная схема обеспечивает согласованную разметку аргументов автора, тогда как разметка аргументов рецензентов зависит от интерпретации разметчиков.

### Литература/References


1. Huisman, J. and Smits, J. (2017), Duration and quality of the peer review process: The author's perspective, *Scientometrics*, vol. 113, no. 1, pp. 633–650. https://doi.org/10.1007/s11192-017-2310-5

2. Lee, C. J., Sugimoto, C. R., Zhang, G. and Cronin, B. (2013), Bias in peer review, *Journal of the American society for information science and technology*, vol. 64, no. 1, pp. 2–17. https://doi.org/10.1002/asi.22784

3. D'Andrea, R. and O'Dwyer, J. P. (2017), Can editors save peer review from peer reviewers?, *PloS one*, vol. 12 (10), e0186111. https://doi.org/10.1371/journal.pone.0186111

4. Checco, A., Bracciale, L., Loreti, P., Pinfield, S. and Bianchi, G. (2021), Ai-assisted peer review, *Humanities and social sciences communications*, vol. 8, no. 8, art. no. 25. https://doi.org/10.1057/s41599-020-00703-8

5. Lisanyuk, E. (2022), Determining argumentative dispute resolution reveals deep disagreement over harassment issue (a case-study discussion in the Russian parliament), *Studia humana*, vol. 11, no. 3–4, pp. 30–45. https://doi.org/10.2478/sh-2022-0014

6. Dung, P. M. (1995), On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games, *Artificial intelligence*, vol. 77, pp. 321–357.

7. Pollock, J. L. (1987), Defeasible reasoning, *Cognitive science*, no. 11, pp. 481–518.

8. Prakken, H. and Vreeswijk, G. (2000), Credulous and sceptical argument games for preferred semantics, *JELIA '00: Proceedings of the European workshop on logics in artificial intelligence*, ACM Digital Libraries, pp. 239–253.

9. Reiter, R. (1980), A logic for default reasoning, *Artifical intelligence*, no. 13, pp. 81–132.

10. Miłkowski, M. and Jasieński, K. (2022), MDPI open peer review corpus, *RepOD*, V1. https://doi.org/10.18150/D5L2EK

11. Baimuratov, I., Lisanyuk, E. and Prokudin, D. (2023), Dispute resolution with OWL DL and reasoning, *Proceedings of the 36th international workshop on description logics (DL 2023), Rhodes, Greece, September 2-4, 2023*, CEUR Workshop Proceedings, vol. 3515, available at: https://ceur-ws.org/Vol-3515/paper-4.pdf (accessed: 02.04.2025).

12. Ross-Hellauer, T. (2017), What is open peer review? a systematic review, *F1000Research*, vol. 6, 588. https://doi.org/10.12688/f1000research.11369.2

13. Soergel, D., Saunders, A. and McCallum, A. (2013), Open scholarship and peer review: a time for experimentation, *Peer reviewing and publishing models: A workshop in conjunction with the 30th international conference on machine learning (ICML 2013), June 20, 2013*, Atlanta, Georgia, USA: JMLR: W&CP, vol. 28.

14. Tennant, J. P., Dugan, J. M., Graziotin, D., Jacques, D. C., Waldner, F., Mietchen, D., Elkhatib, Y., Collister, L. B., Pikas, C. K., Crick, T., Masuzzo P., Caravaggi, A., Berg, D. R., Niemeyer, K. E., Ross-Hellauer, T., Mannheimer, S., Rigling, L., Katz, D. S., Tzovaras, B. G., Pacheco-Mendoza, J., Fatima, N., Poblet, M., Isaakidis, M., Irawan, D. E., Renaut, S., Madan, Ch. R., Matthias, L., Kjær, J. N., O'Donnell, D. P., Neylon, c., Kearns, S., Selvaraju, M. and Colomb, J. (2017), A multidisciplinary perspective on emergent and future innovations in peer review, *F1000Research*, vol. 6, 1151. https://doi.org/10.12688/f1000research.12037.3

15. Price, S. and Flach, P. A. (2017), Computational support for academic peer review: A perspective from artificial intelligence, *Communications of the ACM*, vol. 60, no. 3, pp. 70–79. https://doi.org/10.1145/2979672

16. Ghosal, T., Verma, R., Ekbal, A. and Bhattacharyya, P. (2019), DeepSentiPeer: Harnessing sentiment in review texts to recommend peer review decisions, in: Korhonen, A., Traum, D. and Marquez, L. (eds), *Proceedings of the 57th annual meeting of the association for computational linguistics*, Florence, Italy: Association for Computational Linguistics, pp. 1120–1130. https://doi.org/10.18653/v1/P19-1106

17. Bharti, P. K., Ranjan, S., Ghosal, T., Agrawal, M. and Ekbal, A. (2021), Peerassist: Leveraging on paper-review interactions to predict peer review decisions, in: Ke, H.-R., Lee, C. S. and Sugiyama, K. (eds), *Towards open and trustworthy digital societies*, Cham: Springer International Publishing, pp. 421–435.

18. Mrowinski, M., Fronczak, P., Fronczak, A., Ausloos, M. and Nedic, O. (2017), Artificial intelligence in peer review: How can evolutionary computation support journal editors?, *PLOS one*, no. 12 (9), e0184711. https://doi.org/10.1371/journal.pone.0184711

19. Nuijten, M. B. and Polanin, J. R. (2020), "statcheck": Automatically detect statistical reporting inconsistencies to increase reproducibility of meta-analyses, *Research synthesis methods*, vol. 11, no. 5, pp. 574–579. https://doi.org/10.1002/jrsm.1408

20. Chesnevar, C., Mcginnis, M. S., Rahwan, I., Reed, C., Simari, G., South, M., Vreeswijk, G. and Willmott, S. (2006), Towards an argument interchange format, *The knowledge engineering review*, vol. 21, no. 4, pp. 293–316. https://doi.org/10.1017/S0269888906001044

21. Rahwan, I. and Reed, C. (2009), The argument interchange format, in: Simari, G. and Rahwan, I. (eds), *Argumentation in artificial intelligence*, Boston: Springer, pp. 383–402. https://doi.org/10.1007/978-0-387-98197-0_19

22. Lawrence, J., Bex, F., Reed, C. and Snaith, M. (2012), Aifdb: Infrastructure for the argument web, *Computational models of argument*, IOS Press, pp. 515–516.

23. Zagorulko, Y., Garanina, N., Sery, A. and Domanov, O. (2019), Ontology-based approach to organizing the support for the analysis of argumentation in popular science discourse, *Artificial intelligence: 17th Russian conference, RCAI 2019, Ulyanovsk, Russia, October 21–25, 2019, Proceedings 17, CCIS*, vol. 1093, Springer, pp. 348–362. https://doi.org/10.1007/978-3-030-30763-9_29

24. Janier, M., Lawrence, J. and Reed, C. (2014), Ova+: An argument analysis interface, *Frontiers in artificial intelligence and applications*, vol. 266, pp. 463–464. https://doi.org/10.3233/978-1-61499-436-7-463

25. Boer, A., Winkels, R. and Vitali, F. (2008), MetaLex XML and the Legal Knowledge Interchange Format, *Computable models of the law: Languages, dialogues, games, ontologies*, Berlin — Heidelberg: Springer-Verlag, pp. 21–41. https://doi.org/10.1007/978-3-540-85569-9_2

26. Gordon, T. F. (2013), Introducing the Carneades web application, *Proceedings of the fourteenth international conference on artificial intelligence and law (ICAIL '13),* New York: Association for Computing Machinery, pp. 243–244. https://doi.org/10.1145/2514601.2514637

27. Francesconi, E. (2016), Semantic model for legal resources: Annotation and reasoning over normative provisions, *Semantic web*, vol. 7, no. 3, pp. 255–265. https://doi.org/10.3233/SW-140150

28. Een, N., Sörensson, N. (2003), Temporal induction by incremental sat solving, *Electronic notes in theoretical computer science*, no. 89(4), pp. 543–560.

29. Dvorak, W., Gaggl, S. A., Rapberger, A., Wallner J. P. and Woltran, S. (2020), The aspartix system suite, *Computational models of argument*, IOS Press, pp. 461–462.

30. Borg, A. and Odekerken, D. (2022), PyArg for solving and explaining argumentation in Python: Demonstration, *Proceedings of (COMMA-22), Frontiers in artificial intelligence and applications*, vol. 353, pp. 349–350. https://doi.org/10.3233/FAIA220167

31. Van Eemeren, F. and Grootendorst, R. (2004), *A systematic theory of argumentation: the pragma-dialectical approach*, Cambridge, Cambridge University Press.

32. Prakken, H. and Vreeswijk, G. (2002), Logics for defeasible argumentation, in: Gabbay, D. and Guenthner, F. (eds), *Handbook of philosophical logic*, vol. 4, Dordrecht: Kluwer, pp. 219–318. https://doi.org/10.1007/978-94-017-0456-4_3

33. Caminada, M. W. A. and Gabbay, D. M. (2009), A logical account of formal argumentation, *Studia logica*, vol. 93, no. 2–3, art. no. 109. https://doi.org/10.1007/s11225-009-9218-x

34. Antonelli, G. A. (1999), A directly cautious theory of defeasible consequence for default logic via the notion of general extension, *Artificial intelligence*, no. 109 (1), pp. 71–109.

35. Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A. and Katz, Y. (2007), Pellet: A practical owl-dl reasoner, *Journal of web semantics*, vol. 5, no. 2, pp. 51–53. https://doi.org/10.1016/j.websem.2007.03.004

36. Baimuratov, I., Karpovich, A., Lisanyuk, E. and Prokudin, D. (2025), Argument identification for neuro-symbolic dispute resolution in scientific peer review, *Proceedings of the 24th ACM/IEEE joint conference on digital libraries (JCDL '24)*, New York: Association for Computing Machinery, art. 6. https://doi.org/10.1145/3677389.3702506

Статья поступила в редакцию 11 марта 2025 г.
рекомендована к печати 10 декабря 2025 г.

Контактная информация:

*Лисанюк Елена Николаевна* — д-р филос. наук, доц., вед. науч. сотр.; https://orcid.org/0000-0003-0135-4583, elisanyuk@hse.ru
*Прокудин Дмитрий Евгеньевич* — д-р филос. наук, доц.; https://orcid.org/0000-0002-9464-8371, d.prokudin@spbu.ru
*Баймуратов Ильдар Раисович* — канд. техн. наук; https://orcid.org/0000-0002-6573-131X, baimuratov.i@gmail.com

## Dispute resolution based on abstract argumentation frameworks and its automatization in peer-review

*E. N. Lisanyuk[1,2], D. E. Prokudin[3], I. R. Baimuratov[4]*

[1] Institute of Philosophy of the Russian Academy of Sciences,
12/1, ul. Goncharnaya, Moscow, 109240, Russian Federation
[2] HSE University,
20, ul. Myasnitskaya, Moscow, 101000, Russian Federation
[3] St. Petersburg State University,
7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation
[4] Leibniz University Hannover,
1, Welfengarten, Hannover, 30167, Germany

**For citation:** Lisanyuk E. N., Prokudin D. E., Baimuratov I. R. Dispute resolution based on abstract argumentation frameworks and its automatization in peer-review. *Vestnik of Saint Petersburg University. Philosophy and Conflict Studies*, 2026, vol. 42, issue 1, pp. 118–130.
https://doi.org/10.21638/spbu17.2026.110 EDN MZHSFO (In Russian)

The process of reviewing scientific publications is hampered by the constant increase in the volume of submissions and the bias of reviewers. Artificial intelligence can speed up the scientific review process, but it is also capable of reproducing the biases embedded in the training data. We propose a formal method for representing and analyzing arguments, which contributes to reducing the duration and different kinds of bias in scientific reviewing. We consider

scientific peer review as a single mixed argumentative dispute between authors of a manuscript submitted for publication and reviewers evaluating it according to certain criteria, and apply a formalism based on abstract argumentation frameworks with credulous extension semantics, which allows to find resolutions of the disputes. Our method for dispute resolution consists of 1) marking up a review, 2) converting the markup to JSON, 3) generating its OWL representation, and 3) identifying the dispute solution by finding its preferred extension, i.e. the maximum subset of acceptable arguments put forward by authors and reviewers. We developed a markup scheme for scientific reviews and an algorithm for generating their OWL representations, which provide a formalization of argumentation disputes during scientific peer review with the help of abstract argumentation frameworks and allow us to find solutions to the disputes by means of the automated logical inference. We validated the proposed method by annotating a corpus of peer reviews and applying the proof-of-concept algorithm to it. We evaluated the proposed annotation scheme with respect to the inter annotator agreement, which showed that the proposed scheme provides unambiguous markup of author arguments, whereas the markup of reviewers' arguments depends on the interpretation of the annotators.

*Keywords:* scientific peer review, dispute resolution, argumentation framework, text markup, knowledge representation, OWL, automatic logical inference.

Received: March 11, 2025
Accepted: December 10, 2025

Authors' information:

*Elena N. Lisanyuk* — Dr. Sci. in Philosophy, Associate Professor, Leading Researcher; https://orcid.org/0000-0003-0135-4583, elisanyuk@hse.ru
*Dmitry E. Prokudin* — Dr. Sci. in Philosophy, Associate Professor; https://orcid.org/0000-0002-9464-8371, d.prokudin@spbu.ru
*Ildar R. Baimuratov* — PhD in Technology; https://orcid.org/0000-0002-6573-131X, baimuratov.i@gmail.com