

Поиск и отбор решений спора и его автоматизация на примере научного рецензирования

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

Институт философии РАН · Санкт-Петербургский государственный университет ·
Университет Лейбница в Ганновере

e lenalisanyuk@yandex.ru · d.prokudin@spbu.ru · baimuratov.i@gmail.com

Аннотация. Мы автоматизируем поиск и отбор решений спора на примере диалога между авторами и рецензентами рукописи, направляемой для публикации в научный журнал или для выступления в оргкомитет конференции. Мы рассматриваем научное рецензирование как разновидность аргументативного спора и формализуем его с помощью абстрактных аргументационных фреймворков с доверительной семантикой расширений. Разработанный нами общий метод генерации представлений абстрактных аргументационных фреймворков на языке OWL DL обеспечивает разрешение споров с помощью автоматического логического вывода и применяется для генерации OWL-представлений размеченных научных рецензий. DOI: 10.52119/LPHS.2024.15.66.007.

Ключевые слова: логика аргументации, решение споров, аргументационный фреймворк, OWL, автоматический логический вывод, научное рецензирование.

Мы автоматизировали работу алгоритма поиска и отбора решений аргументативного спора — ПОРС [1] и применили его к научному рецензированию, представив его в качестве такого спора между авторами рукописи, направляемой для публикации в научный журнал или для выступления в оргкомитет конференции, и рецензентами, оценивающими ее по критериям принятия рукописей. Предложенный нами метод разрешения споров в ходе научного рецензирования состоит из трех этапов: формализации рецензии в виде абстрактного аргументационного фреймворка, генерации его представления на языке OWL DL и выявления решения спора с помощью автоматического логического вывода. Мы валидировали предложенный метод, реализовав proof of concept и применив его к корпусу научных рецензий, и подсчитали характеристики сгенерированных OWL-представлений [2].

Стандартный характер формата научного рецензирования позволил избежать обращения к содержательному картированию [3; 4], техникам извлечения аргументов (argument mining) вроде [5] и оценивания одиночных аргументов в связи с их внутренней структурой, как предложено в [6] или [7].

Посредством направления своей рукописи авторы выступают с аргументативным рассуждением, выдвигая положительное мнение $T + (P)$ по поводу истинности предложения (пропозиции) P .

(P) *Наша рукопись достойна быть принятой для...*,

а также аргументы в защиту этой точки зрения, утверждая, что рукопись отвечает всем предъявляемым требованиям. Диалог авторов и рецензентов завершается на первом раунде, если последние отклоняют рукопись вместе с мнением авторов, либо принимают ее как есть, соглашаясь с ним. Мы рассмотрели случаи, когда диалог между ними состоит из двух и более раундов, в ходе которых рецензенты возражают авторам, аргументируя это необходимостью доработки рукописи ввиду неполного соответствия требованиям, затем авторы отвечают новым аргументативным рассуждением, сообщая о произведенных в рукописи исправлениях, после чего рецензенты в своем аргументативном рассуждении оценивают рукопись по тому же шаблону. Мы представили возражения рекомендации рецензентов как аргументативные

рассуждения, обосновывающие отрицательную точку зрения по поводу истинности предложения (P): $T - (P)$ и включающие критику аргументов авторов. Результат такого спора, относящегося к единичным смешанным спорам [8], подразумевает, что P будет защищено и рукопись примут после доработки, либо P будет отклонено вместе с рукописью, но не подразумевает защиты, либо отклонения мнений о $\neg P$, которые в нем не были выдвинуты. В отличие от этого, в множественном смешанном споре хотя бы одна из сторон защищает свою точку зрения в качестве положительной и одновременно в качестве отрицательной критикует противоположную точку зрения, так что в итоге будет защищено мнение об истинности предложения A и мнение об истинности $\neg A$ будет отклонено, либо, наоборот, будет защищено $\neg A$, а A отклонено.

Алгоритм ПОРС основан на аргументационных фреймворках логики аргументации и содержательной идее о том, что убежденность рационального агента в истинности какого-либо предложения A или ее отсутствие зависит от того, «насколько успешно аргумент, поддерживающий это положение, может быть защищен от контраргументов» [9, р. 323]. В соответствии с ней в логике аргументации за базовое отношение аргументации принимают критику одного аргумента другим и демонстрацию как поддержку заключения посылками считают производным от нее. Аргументативный спор — это абстракция для обозначения диалога, в котором стороны — рациональные агенты расходятся во мнениях по поводу предложения A и приводят аргументы в поддержку своей точки зрения или для отклонения противоположных точек зрения. Атомом аргументативного спора является аргумент — одно или несколько предложений, упорядоченных определенным образом для поддержки какого-либо мнения, выступающего его заключением. Формальной репрезентацией аргументативного спора выступает множество аргументов, упорядоченное на графе при помощи абстрактного отношения атаки, символизирующего критику:

$$F = \langle Arg, attack \rangle.$$

Для некоторого подмножества аргументов $S \subseteq F$, такого, что $\{\alpha, \beta, \dots\} \in S$, где аргумент α критикует аргумент β , будем говорить, что α отклоняет β : $attack[\alpha, \beta]$, разве что в подмножестве $S \subseteq B$ найдется аргумент γ , такой, что γ отклоняет α : $attack[\gamma, \alpha]$ и тем самым возвращает β в F в качестве защищенного, как в примере на рис. 1, где отношения атаки обозначены сплошными линиями стрелок, а производное от них отношение поддержки — пунктирной линией.

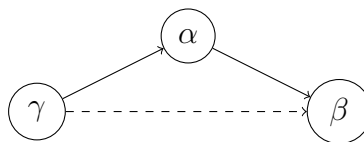


Рис. 1: Пример

Такое упорядочение подмножества аргументов α , β и γ при помощи отношения атаки символизирует расхождение во мнениях — заключениях аргументов α и β между сторонами спора, когда для того, чтобы защитить свою точку зрения, автор аргумента β , отклоненного аргументом α , выступает с контраргументацией при помощи нового аргумента γ , отклоняющего α .

Аргумент называют приемлемым на данном множестве аргументов, если в случае, когда он атакован, в нем найдется аргумент, отклоняющий атаковавший его аргумент. В ПОРС исходными спора выступают подмножества приемлемых аргументов, из которых отбирают решения в зависимости от типа спора. На рис. 1 приемлемыми являются β и γ , но не α , т. к. всякий приемлемый аргумент либо не атакован, либо защищен.

Наименьшей степенью разумности, характеризующей аргументы, которыми можно убедить рационального агента, обладает состоящее из приемлемых аргументов допустимое подмножество аргументов. На рис. 1 имеется несколько таких подмножеств: $\{\beta\}$, $\{\gamma\}$, $\{\gamma, \beta\}$. Для рационального агента наиболее убедительным будет, очевидно, максимальное допустимое подмножество $\{\gamma, \beta\}$, составляющее предпочтительное расширение *Arg* на основе доверительной семантики расширений — одной из семантик, предложенной в немонотонной логике [10; 11]. В дальнейшем мы планируем обратиться к моделированию решений множественных смешанных споров на основе скептической семантики.

Таким образом, мы рассмотрели научное рецензирование как единичный смешанный аргументативный спор между авторами и рецензентами рукописи, направляемой для публикации в научный журнал или для выступления в оргкомитет конференции, и формализовали научное рецензирование с помощью абстрактных аргументационных фреймворков и доверительной семантики расширений. Мы разработали общий метод генерации представлений абстрактных аргументационных фреймворков на языке OWL DL, обеспечивающих разрешение споров с помощью автоматического логического вывода, и применили его для генерации OWL-представлений размеченных научных рецензий.

Литература

1. Лисанюк Е. Н. Поиск и отбор решений спора *Формальная философия аргументации* / под ред. Е. Н. Лисанюк. СПб.: Алетейя, 2022, с. 166–194.
2. Baimuratov I., Lisanyuk E., Prokudin D. 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, 2023.
3. Лисанюк Е. Н., Прокудин Д. Е. Разработка классификации и каталога программного обеспечения для моделирования аргументации и делиберативных рассуждений. *International Journal of Open Information Technologies*, 2022, т. 10, № 11, с. 11–25.
4. Загорюлько Ю. А. и др. Программный комплекс для моделирования и анализа аргументации в научно-популярных текстах *Argnetbank studio*. Свидетельство о регистрации программы для ЭВМ RU 2020665092, 20.11.2020. Заявка № 2020663982 от 09.11.2020.
5. Habernal I., Gurevych I. Argumentation Mining in User-Generated Web Discourse. *Computational Linguistics* 43.1, 2017, pp. 125–179.
6. Walton D., Reed Ch., Macagno F. *Argumentation schemes*. Cambridge University Press, 2008.
7. Prakken H. An abstract framework for argumentation with structured arguments. *Argument and Computation* 2.1, 2011, p. 93–124.
8. Еемеерен Ф. ван, Гроотендорст Р. *Систематическая теория аргументации. Прагматодialeктический подход*. М.: Канон+, 2021.
9. Dung P. M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence* 77, 1995, pp. 321–357.
10. Prakken H., Vreeswijk G. Logics for defeasible argumentation. *Handbook of philosophical logic. 2nd edition* / ed. by D. Gabbay, F. Guenther. Vol. 4. Dordrecht: Kluwer, 2022, p. 219–318.

11. Pollock J. L. *Cognitive Carpentry: A Blueprint for How to Build a Person*. Cambridge, MA: MIT Press, 1995.

Финансирование. Исследование поддержано РФФ, проект № 20-18-00158 «Формальная философия аргументации и комплексная методология поиска и отбора решений спора», реализуемый в Санкт-Петербургском государственном университете.