

RU

Сочетаемость аргументов разных функциональных групп в научных текстах

Пименов И. С.

Аннотация. Цель исследования - определить сочетаемость аргументов разных функциональных групп в коллекции научных текстов. Научная новизна работы заключается в разработке функциональной классификации аргументационных схем и выявлении особенностей применения аргументов разных функциональных групп в коллекции русскоязычных научных текстов (такой анализ функциональной сочетаемости аргументов проведён впервые как для текстов научного жанра, так и для текстов на русском языке). На основе сопоставительного анализа семантики аргументов и функциональных особенностей их употребления разработана классификация аргументационных схем с разграничением четырёх способов доказательства (от авторитета, от практической значимости, через детализацию либо анализ причинно-следственных связей). Употребление аргументов четырёх групп исследовано на наборе из 1030 цепочек рассуждения, извлечённых из экспертно размеченных научных статей по лингвистике и компьютерным технологиям. Показано, что проанализированные статьи характеризуются активным сочетанием аргументов разных функциональных групп при их неравномерном позиционном расположении в отдельных цепочках в зависимости от расстановки акцентов при доказательстве. Работа включает следующие части: моделирование аргументации, функциональное сопоставление аргументационных схем, представление рассуждений через функциональные блоки, анализ сочетаемости таких аргументов.

EN

Compatibility of Arguments from Different Functional Groups in Scientific Texts

Pimenov I. S.

Abstract. The study aims to determine the compatibility of arguments from different functional groups in a collection of scientific texts. The study is novel in that it develops a functional classification of argumentation schemes and identifies the features of using arguments from different functional groups in the collection of Russian-language scientific texts (it is the first time that such an analysis of functional compatibility of arguments has been carried out both for texts of the scientific genre and for texts in Russian). Based on a comparative analysis of the semantics of arguments and the functional features of their use, a classification of argumentation schemes has been developed differentiating four methods of proof (from authority, from practical value, through elaboration or causal analysis). The use of arguments from four groups has been investigated using a set of 1030 reasoning sequences extracted from expertly annotated scientific papers on linguistics and computer technology. It has been shown that the analysed papers are characterised by an active combination of arguments from different functional groups with their uneven positional arrangement in some sequences, depending on the emphasis in the proof. The work includes the following parts: argumentation modelling, a functional comparison of argumentation schemes, presentation of reasoning through functional blocks, a compatibility analysis of such arguments.

Введение

Аргументационная структура отдельного текста обеспечивает его связность на уровне глубинной организации высказываемых идей в целостном рассуждении, направленном на формирование или изменение некоторой позиции адресата по рассматриваемому вопросу (Kononenko, Sidorova, Akhmadeeva, 2020). Вследствие этого анализ связей между приводимыми доводами, в частности выявление их аргументационной семантики, выступает особенно значимым для исследований аргументации, в том числе в аспекте её автоматической

обработки (Lawrence, Reed, 2019). Так, особенности построения рассуждений в отдельном тексте способны передавать его жанровую или тематическую специфику (Al-Khatib, Wachsmuth, Hagen et al., 2017), отражать степень его доступности для восприятия (Barbieri, Aggujaro, Molteni et al., 2015). Кроме того, анализ аргументационных связей важен для оценки убедительности выстраиваемых рассуждений (Zagorulko, Domanov, Sery et al., 2020). Соответственно, изучение сочетаемости аргументов разных функциональных групп, разграничиваемых согласно особенностям их употребления, позволяет выявить общие тенденции организации рассуждений в текстах отдельного жанра, а также предоставляет возможность сравнения текстов в аспекте их аргументационной схожести.

Актуальность исследования заключается в расширении методов формального представления и автоматической обработки аргументации (посредством анализа употребления аргументов разных типов в функциональном аспекте). В свою очередь, подобное расширение методов формального аргументационного анализа представляется значимым ввиду сравнительно недавнего становления автоматической обработки аргументации как самостоятельного направления: первая специализированная конференция была проведена в 2014 году (Lawrence, Reed, 2019, p. 776), тогда как осуществление подобных исследований на материале русскоязычных текстов ограничивается недостатком специализированных корпусов с формальным представлением аргументационных структур (Котельников, 2018).

В свою очередь, указанный анализ сочетаемости аргументов в коллекции научных текстов предполагает решение следующих задач:

- 1) определить общую специфику формального представления аргументационных структур для последующего анализа их образующих связей;
- 2) разграничить функциональные группы аргументов, представленных в текстах коллекции, путём сопоставительного анализа особенностей их реализации на уровне связанных утверждений (посылок и заключений);
- 3) построить представление цепочек рассуждения из анализируемых текстов на уровне выделенных функциональных аргументационных групп;
- 4) выявить типовые цепочки аргументов посредством частотного анализа, изучить распределение в них доводов разных типов с интерпретацией результатов на материале конкретных примеров.

Решение перечисленных задач осуществляется методами функциональной семантики и частотного анализа на основе общих методов моделирования аргументации.

Материалом исследования служит коллекция из 32 научных статей на русском языке, принадлежащих двум предметным областям (16 текстов по лингвистике, 16 – по информационным технологиям). Проанализированные статьи взяты из свободно доступной научной электронной библиотеки «КиберЛенинка» (<https://cyberleninka.ru>). Они были опубликованы в период с 2010 по 2020 год и характеризуются объёмом от 800 до 1500 слов.

Теоретической базой исследования служат, с одной стороны, работы Рут Амосси (Amossy, 2009), посвящённые дискурсивному подходу в изучении аргументации, а с другой – труды с представлением стандарта аргументационного моделирования Argument Interchange Format (Rahwan, Reed, 2009) и списка моделей рассуждения (аргументационных схем) Дугласа Уолтона (Walton, Reed, Macagno, 2008). В частности, дискурсивный подход Рут Амосси основан на трёх ключевых принципах: аргументационные связи между доводами проявляются на уровне целостных текстов; формальное устройство аргументационных структур доступно для реконструкции на основе выражающих их языковых средств; составляющими аргументационного воздействия признаются как абстрактные схемы убеждения, так и реализующие их языковые средства. В свою очередь, Argument Interchange Format предполагает выделение двух типов элементов в структурах аргументации: утверждений и аргументационных схем (абстрактных моделей рассуждения), реализация которых и задаёт связи между утверждениями. Для определения таких схем авторы формата предлагают использовать классификации моделей рассуждения, в частности рекомендуют компендиум Уолтона ввиду его применимости к текстам разных жанров.

Практическая значимость исследования заключается в упрощении разметки аргументационных структур путём двухэтапного определения семантики связей между доводами (посредством указания обобщённой функциональной группы с возможной детализацией модели рассуждения); в применимости классификации для определения аргументационной схожести текстов (например, в задачах кластеризации) с учётом функциональной близости отдельных аргументов; а также в выявлении научных статей с нетипичной организацией рассуждений. Кроме того, материалы исследования могут найти применение в процессе учебно-методической деятельности в области академического письма, компьютерной (прикладной) лингвистики, риторики, лингвистики текста.

Основная часть

Моделирование аргументационной структуры текста

Описание аргументационных отношений между доводами, реализуемыми в сегментах текста, а соответственно, и реконструкция их интегрирующей аргументационной структуры, предполагает представление таких связей посредством выбранной абстрактной модели. Таким актуальным стандартом для формального представления аргументации выступает Argument Interchange Format (AIF), разработанный Иядом Рахваном и Крисом Ридом (Rahwan, Reed, 2009).

AIF основан на моделировании аргументации посредством графов, или математических структур с двумя видами элементов: вершинами (отдельные объекты) и рёбрами (связи между вершинами). При этом в AIF разграничиваются два вида вершин: информационные (соответствуют пропозициональному содержанию отдельного

аргументативного утверждения) и вершины-схемы (указывают на типовую модель рассуждения, или аргументационную схему, посредством которой осуществляется переход между связанными аргументативными утверждениями). Одно ребро в аргументационном графе, отражающем аргументационную структуру текста, не может соединять две информационные вершины: любой переход от одного аргументационного утверждения к другому рассматривается как реализация некоей модели рассуждения. Таким образом, целостному аргументу в AIF соответствует совокупность аргументационной схемы и связываемых ею утверждений (одно из которых, заключение, обосновывается иными, именуемыми посылками). В свою очередь, аргументы способны пересекаться в аргументационной структуре полного текста: заключение одного может являться посылкой в другом.

Как следствие, реконструкция аргументационной структуры отдельного текста посредством AIF (его аргументационная разметка или аннотация) заключается в решении трёх частных задач: 1) выявлении аргументационных утверждений, или компонентов отдельных аргументов (посылок и заключений), соответствующих шагам в представленном рассуждении; 2) установлении связей между выделенными утверждениями; 3) определении модели рассуждения в основе каждой отмеченной связи. Выполнение третьей задачи подразумевает использование некоей классификации аргументационных схем, которая может быть адаптирована под анализируемый текст в зависимости от его жанровых и тематических особенностей. Классификацией, рекомендуемой авторами AIF в качестве образца, выступает компендиум аргументационных схем Дугласа Уолтона (Walton, Reed, Macagno, 2008), который в издании 2008 года содержит 60 моделей рассуждения, свойственных дискурсам в разных областях (например, повседневному общению, научной коммуникации, юридическим диспутам). На Рисунке 1 приведён пример фрагмента аргументационной структуры научного текста, построенный и визуализированный с помощью платформы ArgNetBankStudio, разработанной в Институте систем информатики имени А. П. Ершова Сибирского отделения Российской академии наук (Сидорова, Ахмадеева, Загоруйко и др., 2020).

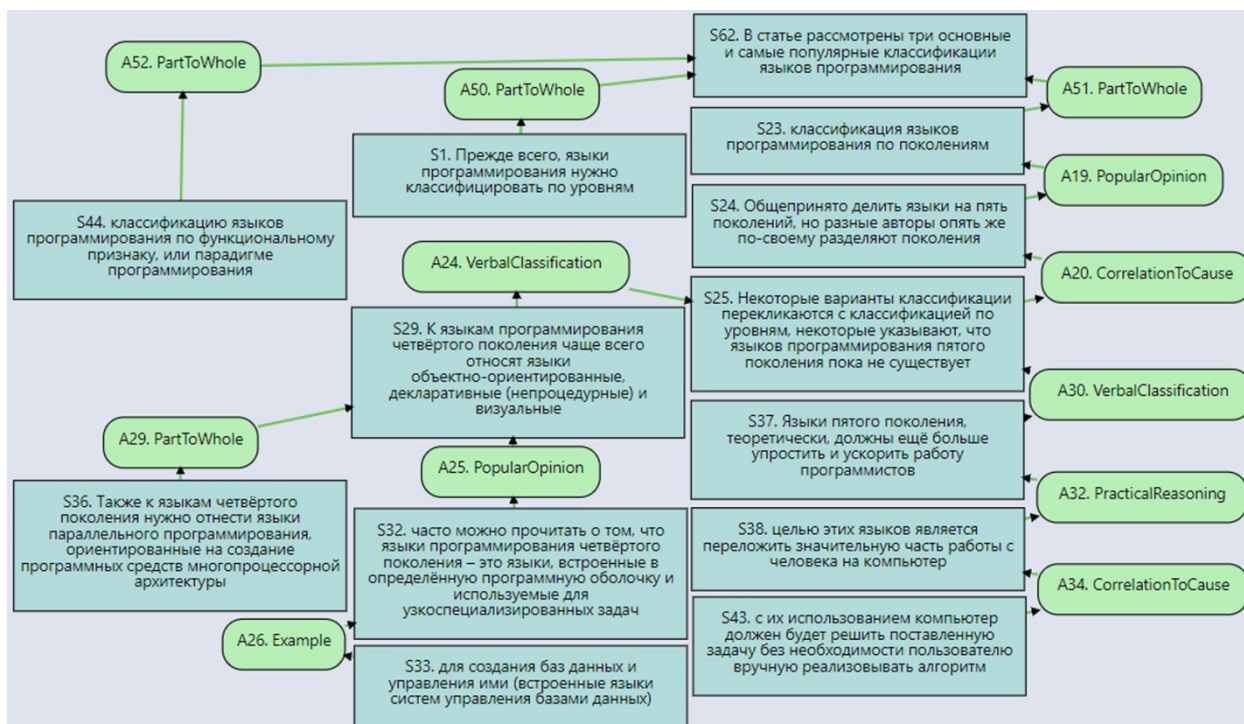


Рисунок 1. Пример фрагмента аргументационной структуры текста

Информационные вершины, соответствующие аргументативным утверждениям, представлены на рисунке в прямоугольных блоках и индексированы через букву S (номера определяются порядком их реализации в тексте). В свою очередь, модели рассуждения представлены в овальных блоках с индексами на А. Во фрагменте содержатся аргументы таких типов, как доказательство через пример (A26), через распространённое мнение (A19, A25), рассуждение от практической цели (A32), а также анализ корреляции через причину (A20, A34), целого через часть (A29, A50, A51, A52), класса явлений через сопоставление с иными элементами классификации (A24, A30). Текст, из аннотации которого извлечён фрагмент, представляет собой обзор трёх классификаций языков программирования, а его основной вывод (главный тезис, подводящий итог всего текста и раскрываемый на его протяжении) представлен на рисунке блоком S62 (единственным, не являющимся посылкой ни для какого другого утверждения). Наоборот, ряд утверждений вводится в структуру аргументации без их дальнейшего обоснования: S33, S36, S43 (S1 и S44 не отнесены к этому списку, поскольку уточняются посылками в полной версии разметки, приведённой сокращённо). От этих утверждений, являющихся листьями в терминах теории графов, можно построить полные цепочки рассуждения к главному тезису: S43 → S38 → S37 → S25 → S24 → S23 → S62, S33 → S32 → S29 → S25 → S24 → S23 → S62, S36 → S29 → S25 → S24 → S23 → S62.

Таким образом, моделирование аргументационной структуры текста посредством AIF позволяет наглядно представить организацию излагаемого в нём рассуждения с ясным разграничением как отдельных доводов (выражаемых языковыми средствами), так и реализуемых типовых аргументационных схем (раскрывающих формальную специфику связей между такими утверждениями). Однако практическое аннотирование текстов сопряжено с рядом сложностей, одной из которых выступает определение реализуемой модели рассуждения в каждом конкретном аргументе.

Функциональное сопоставление аргументационных схем

Представление аргументационной структуры текста посредством модели AIF предполагает семантическую интерпретацию связи между утверждениями в составе каждого аргумента. Такое определение реализуемой схемы рассуждения может осуществляться посредством как одноэтапного указания её детальной специфики (например, через выбор подходящей модели из компендиума Уолтона), так и через вспомогательное указание обобщённых функциональных особенностей перехода в качестве промежуточного шага. Первый вариант сопряжён с большими практическими сложностями при аргументационной разметке коллекции текстов, тогда как определение общих функциональных черт для впоследствии уточняемых моделей рассуждения позволяет соотносить их по степени схожести или различия.

Для обозначения набора функционально близких моделей аргументации, представляемых через одну обобщённую модель на промежуточном этапе, мы применяем понятие функциональной аргументационной группы (ФАГ). Состав каждой группы определяется через сопоставительный анализ детализированных аргументационных схем (из компендиума Уолтона) по содержательным особенностям их реализации: явность или неявность убеждающего воздействия, ракурс представления сведений, сочетаемость с иными схемами в процессе рассуждения, позиционная вариативность в структуре полного текста (равномерное или неравномерное расположение в общей структуре, повторяемость схемы между последовательными или параллельными связями).

Материалом для сопоставительного анализа схем послужила коллекция из 1174 аргументов, извлечённых из 30 научных текстов с аргументационной разметкой, созданной тремя экспертами. Размеченные тексты характеризуются объёмом от 800 до 1500 слов, ориентированы на доказательство одной центральной идеи (общего вывода, обосновываемого на протяжении научной статьи). Представленные статьи относятся к двум предметным областям: лингвистике (преимущественно корпусной лингвистике и лексической семантике) и компьютерным технологиям (различные прикладные задачи, такие как информационная безопасность, обработка изображений и др.). Аргументационная разметка осуществлена на уровне детализированных схем рассуждения из компендиума Уолтона с помощью вспомогательных правил для определения главного тезиса текста, выявления связанных утверждений и установления аргументационной схемы в основе их связи. Более подробное представление размеченной коллекции доступно в работе (Pimenov, Salomatina, Timofeeva, 2022).

По итогам функционального анализа моделей рассуждения подготовлена классификация 14 наиболее частотных схем (встречающихся в 20% и более текстов коллекции), объединённых по четырём ФАГ. В Таблице 1 дана статистика употребления детализированных аргументационных схем с указанием для каждой её абсолютной частоты F_{abs} , доли от числа всех аргументов коллекции $F_{abs}(\%)$ и числа текстов с реализацией этой схемы F_{txt} .

Таблица 1. Частоты встречаемости детализированных аргументационных схем

Схема	Обозначение (аббревиатура)	F_{abs}	$F_{abs}(\%)$	F_{txt}
От примера	Example (E)	210	18	29
Через классификацию	Verbal Classification (VC)	184	16	26
От взаимосвязи к причине	Correlation to Cause (CtC)	149	13	26
От причины к следствию	Cause to Effect (CtE)	146	13	27
От части к целому	Part to Whole (PtW)	111	10	16
От практической цели	Practical Reasoning (PR)	72	6	22
От знака к означаемому	Sign (S)	58	5	14
От экспертного мнения	Expert Opinion (EO)	51	5	18
От применяемого метода	Applied Method (AM)	32	3	16
От положительных результатов	Positive Consequences (PC)	23	2	8
От негативных результатов	Negative Consequences (NC)	18	2	8
От распространённого мнения	Popular Opinion (PO)	15	1	7
По аналогии	Analogy (A)	13	1	10
С позиции знающего	Position to Know (PtK)	12	1	6
Суммарно	-	1094	96	-

Как показано в Таблице 1, 14 схем, отобранных для функциональной классификации, характеризуют практически полный набор аргументов коллекции (96%). В зависимости от специфики употребления отдельных схем различаются их соотношения абсолютной и текстовой частоты: например, рассуждения от экспертного мнения и практической цели реализуются в большем числе текстов, чем анализ целого через часть, однако уступают ему по абсолютной частоте (поскольку для такого анализа характерно интенсивное проявление через поочерёдное рассмотрение нескольких частей одной сущности). Схожим образом доказательства через аналогию встречаются в трети текстов, однако в каждом из них используются менее активно, чем аргументы от распространённого мнения или с позиции знающего, выявленные в каждой пятой статье.

Все перечисленные выше схемы могут быть отнесены по функциональной схожести к четырём различным классам, которые представлены в Таблице 2. Перевод англоязычных названий схем аргументации (выполнен автором статьи. – И. П.) приведён выше в Таблице 1.

Таблица 2. Состав функциональных аргументационных групп из детальных схем

ФАГ	Аббр.	Схемы аргументации
От авторитета (From Authority)	FA	Expert Opinion, Popular Opinion, Position to Know
От практической значимости (From Practical Value)	PV	Applied Method, Practical Reasoning, Positive/Negative Consequences
Анализ причинно-следственных связей (Causal Analysis)	CA	Cause to Effect, Correlation to Cause
Через детализацию (Through Elaboration)	TE	Analogy, Example, Part to Whole, Sign, Verbal Classification

Первая группа (аргументы от авторитета, FA) включает три схемы, доказательство посредством которых обращается вовне рассматриваемых реалий ко мнениям на их счёт. Прямое указание на авторитет эксплицирует убеждающее воздействие таких моделей. В рассмотренных научных текстах для рассуждений от авторитета характерно употребление компактных посылок, не обосновываемых иными доказательствами: указание на весомость источника делает менее актуальным воспроизведение частных шагов при достижении этим источником цитируемого вывода (а раскрытие приводимых рассуждений снижает потребность в акцентировании авторитетности источника). Тем не менее, на Рисунке 1 приведены два примера нетипичного употребления схемы Popular Opinion (с развитием посылки о распространённом мнении через её дополнительное раскрытие).

Вторая группа (аргументы от практической значимости, PV) образована схемами, при которых переход между утверждениями строится через сопоставление собственных особенностей представляемой сущности и внешних ориентиров предпочтительности или нежелательности актуализации таких свойств. Оценочный компонент, характеризующий применимость описываемых средств или методов, также усиливает убеждающее воздействие соответствующих схем. При этом для посылок, приводимых при реализации данных схем, характерно дальнейшее обоснование с раскрытием факторов, влияющих на применимость метода или средства. На Рисунке 2 приведён пример реализации схемы Practical Reasoning (наиболее частотной в этой ФАГ) с построением рассуждения от эксплицитно обозначенной задачи, чья актуальность описывается дополнительно.

Третья группа (анализ причинно-следственных связей, CA) содержит две частотные схемы, различающиеся направлением перехода. Использование каждой из них предполагает разграничение двух или более явлений, самодостаточных на уровне выражения (причина и следствие соответствуют отдельным событиям, корреляция проявляется между сопоставляемыми реалиями). Высокая общность значения этих схем обеспечивает их наибольшую позиционную гибкость при расположении в аргументационной структуре полного текста. Например, на Рисунке 1 указаны два случая реализации схемы Correlation To Cause: как при присоединении исходной посылки, не обосновываемой далее, так и в центре цепочки рассуждения.

Наконец, четвёртая группа (через детализацию, TE) объединяет схемы, основанные на самодостаточном представлении некоторого отдельного явления через освещение его сторон. Это представление может производиться посредством уточнения связей между частью и целым, частным и общим, означающим и означаемым, а также иллюстрации явления с помощью примера или аналогии. Как и в случае с предыдущей группой, для рассуждений по этим схемам нехарактерно эксплицитное убеждающее воздействие. При этом в текстах коллекции схемам именно этой группы наиболее свойственно активное параллельное повторение: так, раскрытие одной части или класса дополняется контрастом с иной частью или классом. На Рисунке 1 представлены примеры употребления схем Part To Whole, Verbal Classification и Example, где первые две реализуются именно при таком параллельном повторении (соответственно A50, A51, A52 и A24, A30).

Представление цепочек рассуждения через функциональные блоки

Построенная функциональная классификация схем может применяться и для анализа сочетаемости разных типов моделей рассуждения в текстах коллекции. В сравнении с характеристикой по детализированным схемам при анализе текстов по укрупнённым группам менее точно представляется специфика отдельных статей, что компенсируется учётом функциональной схожести разных схем, а соответственно, позволяет выявлять тенденции более общего плана в организации аргументов. В этой связи функциональная классификация особенно значима для автоматического определения схожести текстов по аргументационной структуре.

В представленной работе сочетаемость ФАГ анализируется посредством обработки полных цепочек рассуждений, элементы которых преобразованы в функциональные аргументационные блоки (ФАБ). Такое преобразование производится в соответствии со следующим набором шагов:

1. Исходные цепочки рассуждений извлекаются из размеченных текстов и соответствуют последовательности аргументационных схем в каждом полном пути (от начальной посылки, приводимой без дальнейшего обоснования, к главному тезису текста). Соответственно, в цепочках учитываются только модели рассуждения без содержания соединяемых ими утверждений.
2. Последовательные детализированные схемы в цепочке заменяются блоками (ФАБ) соответствующих им функциональных групп (ФАГ).

3. Редкие схемы, не включенные в классификацию, не учитываются в преобразованной цепочке. Также пропускается схема логического конфликта, встречающаяся в 2% аргументов коллекции, ввиду контрастной специфики её употребления (что отличает логический конфликт от иных схем, отражающих позитивное развитие рассуждения).

4. Одинаковые функциональные блоки (соответствующие одной ФАГ) в смежных позициях преобразованной цепочки объединяются в один ФАБ. Это производится для учёта непосредственно качественных, а не количественных особенностей организации рассуждений на уровне функциональных групп.

Процесс преобразования цепочек рассуждения в последовательность ФАБ удобно проиллюстрировать на примере фрагмента аргументационной структуры, приведённого на Рисунке 1. Так, в этом фрагменте представлены три полные цепочки рассуждения, принимающие следующий вид на уровне аргументационных схем: 1) CorrelationToCause → PracticalReasoning → VerbalClassification → CorrelationToCause → PopularOpinion → PartToWhole (от утверждения S43 к главному тезису S62), 2) Example → PopularOpinion → VerbalClassification → CorrelationToCause → PopularOpinion → PartToWhole (от S33), 3) PartToWhole → VerbalClassification → CorrelationToCause → PopularOpinion → PartToWhole (от S36). Одноэлементные цепочки от S1 и S44 не учитываются, так как в исходном тексте они развиваются дополнительными доводами (по аналогии с S23, однако на рисунке приведены компактно). Цепочка (1) преобразуется в последовательность CA → PV → TE → CA → FA → TE (включает 6 ФАБ, реализующих все 4 ФАГ из классификации), цепочка (2) – в TE → FA → TE → CA → FA → TE (6 ФАБ, 3 ФАГ), а цепочка (3) – в TE → CA → FA → TE (4 ФАБ, 3 ФАГ). Таким образом, и в (1), и в (2) по две ФАГ представлены несколькими ФАБ (соответственно два ТЕ и два СА, три ТЕ и два FA), которые чередуются с функционально отличающимися доводами, причём в первой цепочке реализуются все четыре выделенные ФАГ. В свою очередь, в цепочке (3) две смежные схемы из одной ФАГ (PartToWhole, VerbalClassification) объединяются в один ФАБ.

Сочетаемость функциональных групп в цепочках рассуждений

Употребление моделей рассуждения разных функциональных групп проанализировано на наборе из 1030 цепочек аргументов, извлечённых из коллекции аргументационных аннотаций для 30 научных текстов. Цепочки моделей рассуждения, анализируемые в эксперименте, соответствуют полным путям от исходных посылок каждого текста (утверждений, приводимых без дальнейшего доказательства) к его главному тезису. Такие цепочки представляются в виде последовательностей аргументационных схем (между связанными утверждениями без учёта их содержания) и варьируются длиной от одной схемы (если одно из утверждений, напрямую поддерживающих главный тезис, приведено без дальнейшего обоснования) до двенадцати, а их средняя длина равна 5,43 схемам. Для анализа сочетаемости семантически близких и отдалённых моделей рассуждения эти цепочки преобразовываются в последовательности ФАБ.

Так, при объединении семантически близких схем из 1030 исходных цепочек получено 120 различных последовательностей ФАБ (длиной от 1 до 10 элементов со средним показателем в 4,6 блока). Уменьшение средней длины цепочек связано с объединением смежных функционально близких схем (из одной ФАГ) в общий блок. Только одна исходная цепочка выведена из рассмотрения ввиду преобразования в пустую последовательность (эта цепочка состоит из единственной редкой схемы «От исключительного случая», ExceptionalCase, с частотой в 0,3%, не учитываемой при функциональной группировке). В Таблице 3 приведены количественные характеристики для 120 цепочек ФАБ (Sequencies) с их группировкой по числу представленных ФАГ (Groups) и указанием количества исходных путей (Paths), реализующих соответствующее число уникальных групп.

Таблица 3. Распределение цепочек по числу функциональных групп

Groups	Sequencies	Paths
1	4	67
2	26	398
3	69	394
4	21	170

Как показывает Таблица 3, проанализированным научным статьям свойственно построение многоаспектных доказательств (с реализацией 2 и более групп), сочетающих аргументационные схемы разных ФАГ даже внутри последовательно развиваемых цепочек рассуждения (не только между параллельными цепочками). Так, в 93% изначальных полных цепочек (от исходной посылки до главного тезиса текста) употребляются аргументационные схемы нескольких ФАГ, тогда как все четыре группы реализуются в 17% путей рассуждения (каждом шестом). Приведённые характеристики уточняются в Таблицах 4 и 5, в первой из которых отражены частоты однородных цепочек из одной ФАГ (F_a – число путей, соответствующих этой цепочке, F_t – число текстов, содержащих такие пути).

Примечательной особенностью размеченной коллекции является выделение единственного пути рассуждения из 1030, содержащего лишь схемы аргументации через авторитетное мнение (без дополнения апелляций к авторитету доводами иных типов). Этот путь отмечен в обзорной статье о машинном переводе идиоматических выражений, в главном тезисе которой утверждается сохраняющаяся сложность перевода таких конструкций (доказываемая на протяжении текста через разбор возникающих частных затруднений). При этом главный тезис дополнительно поддерживается сторонним аргументом о предпочтительности человеческой проверки и редактирования любых автоматически переведённых текстов: данный довод основан

на цитировании авторитетного источника и не требует дополнительного раскрытия из-за выхода вонне поля зрения работы (обращённой к частной проблеме машинного перевода, а не всей задаче в целом). Вся цепочка из единственной авторитетной схемы имеет следующий вид (с учётом содержания утверждений): «*Не смотря на интенсивные исследования в сфере машинного перевода и перспективность данного направления, до сих пор существует ряд проблем и задач, стоящих перед разработчиками программного обеспечения, и одной из основных проблем является перевод идиоматических выражений*» ← {ExpertOpinion} ← «*Осуществление перевода в существующих системах требует вмешательства человека в качестве предредактора, интерредактора или постредактора* [Зубов, 2007, с. 254]».

Таблица 4. Частоты однородных цепочек ФАБ

Цепочка	F _a	F _t
CA (Causal Analysis)	30	11
TE (Through Elaboration)	22	4
PV (From Practical Value)	14	7
FA (From Authority)	1	1

Следует отметить, что схемы ФАГ "FromAuthority" встречаются в 29% исходных цепочек рассуждения (297 из 1030). Таким образом, почти полное отсутствие однородных цепочек рассуждений исключительно от авторитета обусловлено жанровыми особенностями: для научных статей нетипично приведение чужого мнения без его расширения авторским комментарием.

Кроме того, для однородных цепочек отмечается сравнительно низкая частота наиболее активно реализуемой из них: так, рассуждения на основе единственной ФАГ CausalAnalysis встречаются только в трети текстов. Эта тенденция указывает на предпочтение авторами многоаспектных доказательств (цепочек рассуждения с сочетанием нескольких ФАГ).

В Таблице 5 указано распределение длин цепочек ФАБ для многоаспектных последовательностей (UT – число уникальных групп, Length – количество ФАБ в цепочке, N(Paths) – число таких цепочек и соответствующих исходных путей).

Таблица 5. Распределение длин многоаспектных цепочек рассуждения

UT	Length	N (Paths)	UT	Length	N (Paths)	UT	Length	N (Paths)
2	2	9 (249)	3	3	11 (83)	4	4	1 (1)
2	3	7 (69)	3	4	23 (156)	4	5	5 (60)
2	4	6 (61)	3	5	15 (90)	4	6	8 (77)
2	5	3 (16)	3	6	13 (37)	4	7	5 (30)
2	6	1 (3)	3	7	3 (6)	4	9	1 (1)
-	-	-	3	8	1 (8)	4	10	1 (1)
-	-	-	3	9	2 (13)	-	-	-
-	-	-	3	10	1 (1)	-	-	-

Другой особенностью научных текстов является активное чередование аргументов разных ФАГ при развитии цепочек рассуждений. Как показано в Таблице 5, даже при ограничении двумя группами возможно их чередование на протяжении пяти-шести блоков. Такое варьирование поддерживается для трёх разных групп: четыре отмеченные цепочки имеют вид: CA → PV → CA → PV → CA → PV, PV → CA → PV → CA → PV, PA → PV → PA → PV → PA, CA → PA → CA → PA → CA.

В частности, первая цепочка (тройное повторение CA → PV) в одном из текстовых выражений соответствует представлению общего решения для крупной задачи через описание последовательных этапов с их решениями и влиянием прошлого частного решения на новый этап (схемы из группы CausalAnalysis передают следственные связи, из группы FromPracticalValue – практические ориентиры при выборе решения). С учётом содержания утверждений (приведённого сокращённо) цепочка имеет вид «*На разных уровнях управления АСУ имеют неодинаковый характер и преследуют различные цели...*» → {CorrelationToCause} → «*Проблема возникает при разработке...*» → {PracticalReasoning} → «*Необходимо хранить пересекающуюся информацию...*» → {CauseToEffect} → «*Дублирование данных приводит к...*» → {PracticalReasoning} → «*Дублирование можно исключить при...*» → {CauseToEffect} → «*Возникает задача исключения...*» → {PracticalReasoning} → «*При проектировании корпоративных АСУ необходимо...*».

В свою очередь, цепочка из пяти блоков с чередованием CA → PA в одном из текстов характеризует вывод следствия из классификации системных причин для языкового явления, анализируемого через примеры с детальным представлением внутренних причинных связей («*Обоснование причины происходит с помощью...*» → {CauseToEffect} → «*Реагирующая реплика с reason, I can see several possible reasons why, акцентирует внимание...*» → {Example} → «*При анализе микротекстов обнаруживается, что функция reason не всегда одинакова...*» → {CorrelationToCause} → «*В реагирующих репликах она реализует несколько функций...*» → {VerbalClassification} → «*Риторическую, обращенную к общим фоновым знаниям...*» → {CauseToEffect} → «*Общим и для иницирующих, и для реагирующих реплик является...*»). Цепочка же из повторяющихся PA → PV появляется в тексте с описанием методов в бизнес-процессе на разных уровнях.

Характеризуя общее распределение числа реализуемых ФАГ и длины цепочки в ФАБ, следует подчеркнуть тенденцию к повторению схожих схем в многоаспектных рассуждениях (цепочках с тремя или четырьмя функциональными группами). Так, для цепочек с двумя ФАГ свойственно единоразовое применение каждой без их дублирования (249 цепочек из 398 с двумя группами, 63% случаев), а повторение одного блока столь же частотно, сколько и повторение обоих. Наоборот, среди цепочек ровно с тремя ФАГ преобладают последовательности с повторением любого одного блока (156 из 394, 40%), а единоразовое употребление каждой из сочетаемых групп отмечается в каждом пятом случае. Такая компактная реализация всех применяемых ФАГ ещё менее характерна для цепочек, реализующих все четыре выделенные группы: лишь в одной из 170 каждая ФАГ представлена ровно одним блоком, тогда как наиболее частые цепочки такого типа (6 блоков) содержат либо повторение одного ФАБ дважды, либо повторение двух по разу.

Таким образом, отмеченные тенденции характеризуют следующие особенности цепочек рассуждения в проанализированных научных текстах. В первую очередь, авторам научных текстов свойственно построение многоаспектных доказательств с активным чередованием аргументов разных ФАГ в последовательном развитии доводов. Во-вторых, однородные цепочки рассуждения (с реализацией одной ФАГ) являются достаточно редкими, причём среди них почти не встречаются последовательности исключительно из апелляций к авторитету: хотя отсылки к экспертным мнениям (в частности, указания на цитируемую литературу) содержатся в 30% проанализированных цепочек, они неизменно дополняются авторскими комментариями. В-третьих, высокая вариативность аргументов разных ФАГ сопряжена с их неравномерным употреблением внутри цепочек рассуждения: для последовательностей более чем с двумя ФАГ более характерно повторение довода одного типа, чем равномерное употребление всех. В-четвёртых, выявленная вариативность цепочек доказательства на уровне ФАБ свидетельствует о применимости тех как для аргументационного аннотирования текстов (на начальном этапе, перед детализацией схем рассуждения в соответствии с углублёнными семантическими классификациями по типу компендиума Уолтона), так и для отдельных прикладных задач обработки текстов (например, для их кластеризации на основе аргументационной схожести). Кроме того, достаточная информативность аргументационного анализа на уровне функциональных блоков может снимать потребность в углублении к уровню детализированно размеченных схем для отдельных задач.

Заключение

В работе представлена функциональная классификация аргументационных схем (доводы от авторитета, от практической значимости, через анализ причинно-следственных связей, через детализацию), выделенных через многоаспектное сопоставление особенностей реализации типовых моделей рассуждения. Экспериментальное применение классификации к набору из 1030 цепочек рассуждения из 30 научных текстов позволило прийти к следующим выводам.

Во-первых, проанализированные научные тексты характеризуются преобладанием многоаспектных доказательств с сочетанием аргументов нескольких функциональных групп. Во-вторых, отмечается неравномерное употребление доводов разных групп (их позиционное расположение) в отдельных цепочках в зависимости от расстановки акцентов при рассуждении. В-третьих, исследованные научные тексты выделяются практическим отсутствием доказательств исключительно через аргументы от авторитета: такие доводы сопровождаются рассуждениями через иные функциональные группы.

В перспективе дальнейших исследований представленная классификация аргументационных схем и выявленная специфика сочетаемости функциональных групп могут уточняться и дополняться в работах, посвящённых процедуре аргументационной разметки текстов либо автоматической обработке структур рассуждения в прикладных задачах аргументационного анализа. Планируется развитие данного исследования в рамках компьютерной лингвистики путём как составления подробной методики аргументационной разметки (на основе функциональной классификации аргументов), так и проведения экспериментов по кластеризации научных статей с целью автоматического выявления работ со схожей организацией рассуждений, по оценке сложности аргументации в научных текстах (на основе функционального разнообразия доводов).

Источники | References

1. Котельников Е. В. Извлечение аргументации из текстов и проблема отсутствия русскоязычных текстовых корпусов // *Advanced Science*. 2018. № 3 (11).
2. Сидорова Е. А., Ахмадеева И. Р., Загоруйко Ю. А., Серый А. С., Шестаков В. К. Платформа для исследования аргументации в научно-популярном дискурсе // *Онтология проектирования*. 2020. Т. 10. № 4 (38).
3. Al-Khatib K., Wachsmuth H., Hagen M., Stein B. Patterns of Argumentation Strategies across Topics // *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (Copenhagen, September 7-11, 2017)*. Copenhagen, 2017.
4. Amossy R. Argumentation in Discourse. A Socio-Discursive Approach to Arguments // *Informal Logic*. 2009. Vol. 29 (3).
5. Barbieri E., Aggujaro S., Molteni F., Luzzatti C. Does Argument Structure Complexity Affect Reading? A Case Study of an Italian Agrammatic Patient with Deep Dyslexia // *Applied Psycholinguistics*. 2015. Vol. 36. Iss. 3.

6. Kononenko I., Sidorova E., Akhmadeeva I. Comparative Analysis of Rhetorical and Argumentative Structures in the Study of Popular Science Discourse // International Conference on Computational Linguistics and Intellectual Technologies “Dialogue”. Moscow, 2020.
7. Lawrence J., Reed C. Argument Mining: A Survey // Computational Linguistics. 2019. Vol. 45. No. 4.
8. Pimenov I., Salomatina N., Timofeeva M. The Quantitative Evaluation of the Pathos to Ethos Ratio in Scientific Texts // Proceedings of the 2022 IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM). Altai, 2022.
9. Rahwan I., Reed C. The Argument Interchange Format // Argumentation in Artificial Intelligence / ed. by I. Rahwan, G. Simari. Dordrecht - Heidelberg - L. - N. Y.: Springer, 2009.
10. Walton D., Reed C., Macagno F. Argumentation Schemes. N. Y.: Cambridge University Press, 2008.
11. Zagorulko Yu. A., Domanov O. A., Sery A. S., Sidorova E. A., Borovikova O. I. Analysis of the Persuasiveness of Argumentation in Popular Science Texts // Artificial Intelligence, Proceedings of the 18th Russian Conference RCAI. Moscow, 2020.

Информация об авторах | Author information



Пименов Иван Сергеевич¹

¹ Новосибирский национальный исследовательский государственный университет



Pimenov Ivan Sergeevich¹

¹ Novosibirsk State University

¹ pimenov.1330@yandex.ru

Информация о статье | About this article

Дата поступления рукописи (received): 06.10.2022; опубликовано (published): 30.11.2022.

Ключевые слова (keywords): автоматический анализ аргументации; моделирование аргументации; модели рассуждения; научные тексты; automatic argumentation analysis; argumentation modelling; reasoning models; scientific texts.