

УДК 81'322.2

Распознавание аргументативных связей в научно-популярных текстах

*Саломатина Н.В. (Институт математики им. С.Л. Соболева),
Кононенко И.С. (Институт систем информатики им. А.П. Ершова),
Сидорова Е.А. (Институт систем информатики им. А.П. Ершова),
Пименов И.С. (Новосибирский государственный университет)*

В статье представлено исследование эффективности использования признака принадлежности утверждений, участвующих в аргументации, к одному тематическому фрагменту текста. Работа проводилась с целью последующего применения этого признака в автоматическом распознавании аргументационных структур. Объектом исследования служили русскоязычные тексты научно-популярного жанра. Тематическая структура текста строилась на основе выявления сверхфразовых единств (фрагментов текста, объединенных одной темой) путем обнаружения кластеров слов и словосочетаний с помощью сканирующих статистик. Для верификации потенциально возможных связей, извлекаемых из тематической структуры, использовались тексты с ручной разметкой аргументационной структуры. Сопоставление связей, построенных «вручную» и потенциальных, определяемых из тематической структуры, проводилось автоматически. Полученные с макро-усреднением точность и полнота составили 48,6 % и 76,2 % соответственно.

***Ключевые слова:** извлечение аргументации, аргументативные связи, тематическая структура текста, профиль кластеризуемости, научно-популярные тексты.*

1. Введение

Область исследований, связанная с автоматическим извлечением аргументов, выделилась в самостоятельную из области тонального анализа текстов, поскольку стала важна не только эмоциональная оценка объектов, явлений и пр. – позитивная, негативная или нейтральная – и мнение о них, но и причины формирования определенных мнений и их обоснование [5]. В настоящее время автоматическое выявление аргументов, аргументационных стратегий в текстах разных жанров полезно, например, для понимания и ведения дебатов, принятия решений (за \ против) в рекомендательных системах, для обнаружения радикальных мнений,

вводящих в заблуждение текстов и пр. В приложениях, реализующих анализ такого типа, структура аргументации должна быть распознана автоматически.

Применение методов машинного обучения для извлечения аргументов и аргументационных структур проводится на базе размеченных специальным образом корпусов текстов. Работа по глубокой разметке корпуса, каковой является разметка аргументации, очень трудоемка, поэтому размеченных текстов часто недостаточно для качественного обучения систем распознавания. Информативность признаков, которые применяются в машинном обучении, в опубликованных работах обычно оценивается в совокупности, поэтому трудно понять, каков вклад того или иного признака в полученном результате. При разработке систем распознавания априорное знание о вкладе отдельных признаков позволит ускорить их отбор для дальнейшего использования.

Автоматическое распознавание структуры аргументов обычно проводится по следующей схеме:

- а) выделение утверждений (клауз) в исследуемом тексте;
- б) разделение утверждений на аргументативные и неаргументативные;
- в) выявление роли утверждений в структуре аргументации – главного тезиса, посылок и заключений;
- г) установление связности утверждений с использованием знаний о схемах рассуждений и \ или тематической структуры текста.

Каждый из этапов схемы требует проведения отдельного исследования и часто производится в предположении, что другие этапы некоторым образом реализованы.

Цель работы: оценить вклад тематического признака в решение задачи автоматического построения аргументативных связей между утверждениями на основе размеченного корпуса научно-популярных текстов.

Исследование выполнено при поддержке РФФИ в рамках проектов № 18-00-01376 (18-00-00889) и № 18-00-01376 (18-00-00760).

2. Обзор существующих методов

Автоматическое выделение утверждений в исследуемом тексте. Разбиение текста на элементарные единицы проводится с использованием простых синтаксических моделей или синтаксического анализатора [12].

Разделение утверждений на аргументативные и неаргументативные. Для этого, как правило, применяются методы машинного обучения, которое реализуется на размеченных вручную текстах определенного жанра. Показано, что качество поиска существенно зависит,

прежде всего, от выбора признаков, а также от сочетания методов машинного обучения с методами на основе экспертных правил. Например, на этапе выявления предложений, содержащих аргумент, использование шаблонов с supervised probabilistic sequence model повысило F-меру на 17% [7].

Выявление роли утверждений в структуре аргументации. В работе [8] посылки и заключения выявляются методом SVM с $F = 0.74$, на этапе предварительной классификации утверждений на аргументативные и не содержащие аргумента работает метод максимальной энтропии. Классификационные признаки наряду со структурными, лексическими, синтаксическими, включают модель ближайшего контекста, дискурсивных и риторических отношений.

В работе [11] при выявлении утверждений (клауз) четырех типов (главный тезис, тезис, посылка, неаргументативное) лучшим из методов (SVM, наивный Байес, деревья решений и случайный лес) был признан SVM. Классификаторы тренировались на структурных, лексических (n-граммы ($n = 1, 2, 3$), глаголы, наречия, модальные слова), синтаксических, контекстуальных признаках, дискурсивных маркерах (объем словаря – 55 единиц).

Установление аргументативной связности утверждений. Иногда решение этой задачи предполагает и возможность вывода содержания одного предложения из другого, как в работе, проведенной несколькими группами исследователей [2] на парах фрагментов (предложений). Показано, что наилучшим образом такого рода связность может быть установлена со средней точностью 0.8 при использовании лексических, синтаксических признаков, меток семантических ролей, статистики n-граммного покрытия.

В работе [3] тематически сходные утверждения определяются с помощью модели на основе скрытого распределения Дирихле (LDA). Каждое слово характеризуется вероятностью принадлежности к некоторой теме. Расстояние между утверждением и его предшественником рассчитывается как евклидово расстояние между оценками по темам. Если расстояние ниже установленного порога, утверждение связывается с предыдущим. Если порог превышен, тогда вычисляется расстояние между утверждением и всеми предыдущими утверждениями и, если самое близкое из них находится на определенном расстоянии, связь устанавливается. Если ни один из этих критериев не соблюдается, предложение считается несвязанным ни с каким другим, встретившимся ранее. Доля совпавших с ручной разметкой связей равна 76,5 %. Направление связи не указывается.

Используя байесовский классификатор, который тренируется на посылках (premises) и заключениях (conclusion), авторы работы [4] идентифицируют отдельные компоненты схемы аргументации (рассмотрены две схемы – Expert Opinion and Positive Consequences) в тексте.

А затем с помощью дискурсивных маркеров определяют отношения атаки \ поддержки между утверждениями. Оставшиеся не присоединенными элементы схемы с помощью результатов тематического анализа текста LDA объединяются в структуру.

Следует отметить, что тематический анализ при установлении связей разного типа применяется все чаще. В настоящей работе для определения фрагментов текста, объединенных одной темой, используется профиль кластеризуемости слов и словосочетаний [1], с помощью которого можно установить потенциальные связи между утверждениями.

3. Постановка задачи

Предположим, что текст с помощью некоторого метода (например, одного из указанных в п. 2) разделен на утверждения $P = \{p_i\}$ (i – номер утверждения в тексте $i = 1, \dots, M$; M – число утверждений в тексте), из которых извлечено подмножество аргументативных $\{p'_j\}$ ($j \leq M$). Метод, описанный ниже в п.4, позволяет разбить текст на фрагменты $B^k = \{p_i\}^k$, содержащие утверждения, относящиеся к одной теме (k – номер фрагмента текста, $k = 1, \dots, K$; K – число выделенных фрагментов). В данном исследовании нас интересует вклад тематического признака в эффективность распознавания аргументационной структуры. Будем считать, что множество $\{p'_j\}$ совпадает с множеством вручную выделенных аргументативных утверждений, а n_{re} – число связей, установленных вручную между утверждениями. Обозначим символами n_{rb} число потенциальных связей между утверждениями, которые выявляются по признаку принадлежности к одной теме. Пусть из них совпали с установленными вручную n_{ra} связей.

Задача состоит в получении оценок полноты (R) и точности (Pr) применения признака принадлежности утверждений (p'_{j1} и p'_{j2}) к одной теме B^k для установления связей между ними на основе вручную размеченных аргументативных утверждений и связей между ними: $R = n_{ra} / n_{re}$; $Pr = n_{ra} / n_{rb}$.

4. Построение профиля кластеризуемости

Построение тематической структуры текста в виде профиля кластеризуемости решается в предположении, что слова, неравномерно распределенные в тексте, являются тематически значимыми. Тогда задача построения тематической структуры может быть сведена к выявлению неравномерностей в позиционном распределении отдельных слов и словосочетаний в тексте согласно различным схемам расстановки точек на линии, если каждую точку рассматривать как место вхождения анализируемого слова (словосочетания) в

текст. Статистически значимые кластеры в этом случае успешно выявляются с помощью сканирующих статистик [9].

Пусть x_1, x_2, \dots, x_N – произвольный набор точек из единичного интервала $(0, 1]$. Требуется проверить гипотезу о равномерности (H_0) против альтернативы (H_1), связанной с тем или иным типом отклонения от равномерности. Для случая кластеризации эффективное решение основано на использовании сканирующей статистики $n(d)$, фиксирующей максимальное число точек n , попавших в интервал длины d при всевозможных расположениях этого интервала внутри единичного отрезка. Вычисление $n(d)$ ведется путем подсчета числа точек, попавших в окно ширины d , скользящее вдоль отрезка.

В данной работе вместо статистики $n(d)$ используется связанная с ней статистика $d(n)$, фиксирующая длину минимального интервала, содержащего ровно n точек ($2 \leq n \leq N$). Поскольку табулирование распределения этой статистики в широком диапазоне значений n и N представляется достаточно трудоемким, для оценки значимости отклонения вычисленной на конкретном тексте статистики $d(n)$ от значения, постулируемого гипотезой H_0 (равномерность), можно прибегнуть к имитационному моделированию.

Схема выявления позиционных аномалий (типа кластеризации) в распределении лексических единиц выглядит следующим образом (см. [1]).

1. В нормализованном тексте подсчитывается частота встречаемости каждой леммы.

2. Пусть x – произвольная лемма (или лемматизированное словосочетание), $F(x)$ – суммарное число ее (его) вхождений в текст, n – фиксированное число последовательных вхождений x в текст ($2 \leq n \leq F(x)$), $d(n)$ – длина минимального фрагмента текста, содержащего n вхождений x . Для дальнейшего анализа отбираем леммы x со значением $F(x) \geq F_b$, где F_b – пороговое значение частоты, зависящее от длины текста N в словоформах.

3. Для каждой леммы проводим перебор по всем допустимым значениям n ($F_b \leq n \leq F(x)$). Для фиксированного n :

а) вычисляем значение $d(n)$ в анализируемом тексте;

б) с помощью имитационного моделирования оцениваем распределение этой статистики при гипотезе H_0 . Для этого путем многократного перемешивания слов в исходном тексте формируем m его рандомизированных аналогов с равномерным распределением слова x по тексту (приемлемыми считаются значения $m \geq 100$). По полученной подборке вычисляем оценки минимального и среднего значения статистики $d(n)$ (соответственно, S_{\min} , S_{avr}), а также среднеквадратичное отклонение s .

4. Сравниваем наблюдаемое в исходном тексте значение статистики $d(n) = S_o$ с оценками, полученными в имитационном эксперименте. Считаем, что аномальное (неслучайное) отклонение от равномерности типа «кластеризация» имеет место, если выполняется условие: $(S_o \leq S_{\min})$ and $(S_o \leq S_{\text{avr}} - 3s)$.

Значимость выделенного кластера можно характеризовать безразмерной величиной $\delta(x)$, равной отношению среднего расстояния между вхождениями x к среднему внутрикластерному расстоянию между вхождениями x .

Понятие *профиля кластеризуемости* вводится для того, чтобы аккумулировать на одном графике информацию обо всех участках кластеризации разных лемм. Формально, профиль кластеризуемости – это ступенчатая функция, аргументом которой является порядковый номер предложения в тексте, а значение равно числу различных кластеров, включающих в себя данное предложение. При этом в отдельно рассматриваемом предложении могут не присутствовать одновременно все леммы, кластеризующиеся в данном участке текста. Пики профиля кластеризуемости обычно соответствуют отдельным темам текста, а провалы между ними – переходу от одной темы к другой. В профиле отражаются различные связи (часть\целое, общее\частное, ассоциации) между леммами, кластеризующимися в одном и том же участке текста.

Ниже приведен пример профиля кластеризуемости лемм научно-популярного текста «Люди, которые создают нормы» (см. Приложение), рассказывающего о том, как могут и должны формироваться языковые нормы ($\delta(x) = 2$).

n_s	n_c	lemma
1	1	РУССКИЙ ЯЗЫК;
7	3	НОРМА; ВАРИАТИВНОСТЬ; РУССКИЙ ЯЗЫК;
8	2	НОРМА; РУССКИЙ ЯЗЫК;
10	1	РУССКИЙ ЯЗЫК;
12	2	ЯЗЫК; РУССКИЙ ЯЗЫК;
17	1	ЯЗЫК;
18	2	ЯЗЫК; ИЗМЕНЕНИЕ;
19	1	ИЗМЕНЕНИЕ;
24	2	ЯЗЫК; ИЗМЕНЕНИЕ;
27	1	ИЗМЕНЕНИЕ;
30	2	ИЗМЕНЕНИЕ; ПРОСТРАНСТВО;
34	1	ПРОСТРАНСТВО;
37	3	ЯЗЫК; ЗАДАЧА; ПРОСТРАНСТВО;
38	4	ЯЗЫК; ЗАДАЧА; ПРОСТРАНСТВО; НОРМА;
40	3	ЗАДАЧА; ПРОСТРАНСТВО; НОРМА;
43	1	ЗАДАЧА;
46	3	ЗАДАЧА; ЭТАП; ТЕЛЕВИДЕНИЕ;
47	4	ЗАДАЧА; РЕЧЬ; ЭТАП; ТЕЛЕВИДЕНИЕ;
51	2	ЗАДАЧА; РЕЧЬ;
53	1	РЕЧЬ;

Рис.1. Профиль кластеризуемости текста «Люди, которые создают нормы»

Ось абсцисс с номерами предложений направлена вниз, а ось ординат (число кластеров) – по горизонтали: слева направо. Для экономии места ось ординат представлена в нелинейном масштабе: указаны номера (n_s) лишь тех предложений, на которых происходит изменение значений профиля, т.е. добавляются новые кластеры или исчезают старые (n_c – число кластеризующихся лемм). Наибольшее значение профиля равно 4 и зафиксировано в предложениях под номерами 38 и 47. Профиль на Рис.1 демонстрирует наличие нескольких явно выраженных тем в тексте, например, в диапазоне (1–10), (34–43), (43–53). Темы в диапазоне (10–34) являются менее выраженными и могут быть присоединены к смежным.

Рассматриваются возможные взаимные расположения кластеров – вложенность, пересечение, разнесенность, относительно которых сформирована гипотеза: связанные аргументативные утверждения могут находиться в тех фрагментах текста, которые входят в кластеры одной темы или в пересекающиеся по лексическому составу кластеры смежных тем. Вероятность существования связи аргументативных утверждений, принадлежащих фрагментам несмежных тем, существует, но она существенно ниже и ее установление требует введения дополнительных правил.

5. Результаты эксперимента

Материалом для эксперимента по оценке эффективности признака принадлежности \ не принадлежности связанных аргументативных утверждений к одной теме текста послужили 100 размеченных вручную текстов, длиной не менее 500 слов.

Для описания структур аргументации принят стандарт, фиксированный форматом AIF (Argument Interchange Format) [10]. Согласно данному формату аргументы представляются ориентированным графом, в котором выделяют два типа вершин: информационные вершины (вершины-утверждения) и вершины-схемы (вершины-аргументы). Вершинам-утверждениям сопоставляются послылки, заключения, а вершинам-схемам – типовые схемы (модели) рассуждений. Для разметки предлагались к использованию несколько десятков схем аргументации из компендиума Уолтона [13].

Всего 36 разных схем аргументации были употреблены экспертами при разметке текстов коллекции более 2-х раз. Наиболее частотные из них (с частотой встречаемости $F > 10$) представлены в таблице 1 ниже. Как можно видеть, примеры Example_Inference в научно-популярных текстах являются самым распространенным видом аргументации, их доля от всех реализованных в размеченной коллекции составляет ~ 17 %. Отстает по частоте, но является весьма характерным для научно-популярного дискурса обращение к аргументации, основанной на строгом (неоспоримом) и нестрогом причинно-следственном выводе

(ModusPonens_Inference и CauseToEffect_Inference). Стилистически значимое обращение к мнению экспертов в обсуждаемой области (ExpertOpinion_Inference) является сильным аргументом для убеждения аудитории как научно-популярного, так и научного текста. Наконец, еще одним значимым вариантом аргументации с помощью каузальных отношений является абдуктивный вывод, позволяющий осуществлять принятие объяснительных правдоподобных гипотез, как в случае аргументации по схеме Sign_Inference. Доля пяти редких схем (CausalSlipperySlope_Inference, DirectAdHominem_Inference, Gradualism_Inference, InconsistentCommitment_Inference, PropertyNotExistant_Conflict) с $F = 2$ составила в сумме 0,3 % от всех использованных в разметке коллекции.

Таблица 1. Частоты применения схем аргументации в научно-популярных текстах

№ п/п	Название схемы	F	№ п/п	Название схемы	F
1	Example_Inference	1008	12	PopularOpinion_Inference	70
2	CauseToEffect_Inference	686	13	VerbalClassification_Inference	50
3	ModusPonens_Inference	632	14	ExceptionalCase_Inference	38
4	ExpertOpinion_Inference	406	15	Logical_Conflict	30
5	Sign_Inference	308	16	NegativeConsequences_Inference	28
6	CorrelationToCause_Inference	228	17	PositiveConsequences_Inference	26
7	EstablishedRule_Inference	100	18	Bias_Inference	16
8	PracticalReasoning_Inference	94	19	PopularPractice_Inference	14
9	EvidenceToHypothesis_Inference	82	20	Waste_Inference	14
10	PositionToKnow_Inference	80	21	Commitment_Inference	14
11	Analogy_Inference	78			

Рассмотрим пример аргументативной разметки текста (Рис. 2). Текст полностью приведен в приложении. Номера предложений в тексте соответствуют указанным в профиле кластеризуемости (Рис. 1). Синим в тексте и в профиле выделен фрагмент, для которого на рисунке 2 приводится структура аргументации. В прямоугольники помещены аргументативные утверждения (в тексте пронумерованы с префиксом S и выделены квадратными скобками), в овалах указаны названия схем аргументации. Утверждение на сером фоне в тексте не присутствует явно, но вводится экспертом, который делает разметку, для устранения лакун в аргументации. Утверждения S23 и S24, а также S24 и S32 связаны, т.к. принадлежат одной теме. Утверждения S32 и S37, S32 и S27, S23 и S25, S24 и S25, S25 и S27 принадлежат смежным пересекающимся темам. Проблемы с установлением связи

возникают с позиционно разнесенными утверждениями (S37 и S2). Но в приводимом примере кластер, который формируется во фрагменте текста, содержащем утверждение S2, образован словосочетанием РУССКИЙ ЯЗЫК, а тема, к которой принадлежит утверждение S37 содержит слово ЯЗЫК. Если принять во внимание, что эти лексические единицы связаны отношением общее-частное, то можно предположить, что и утверждения связаны.

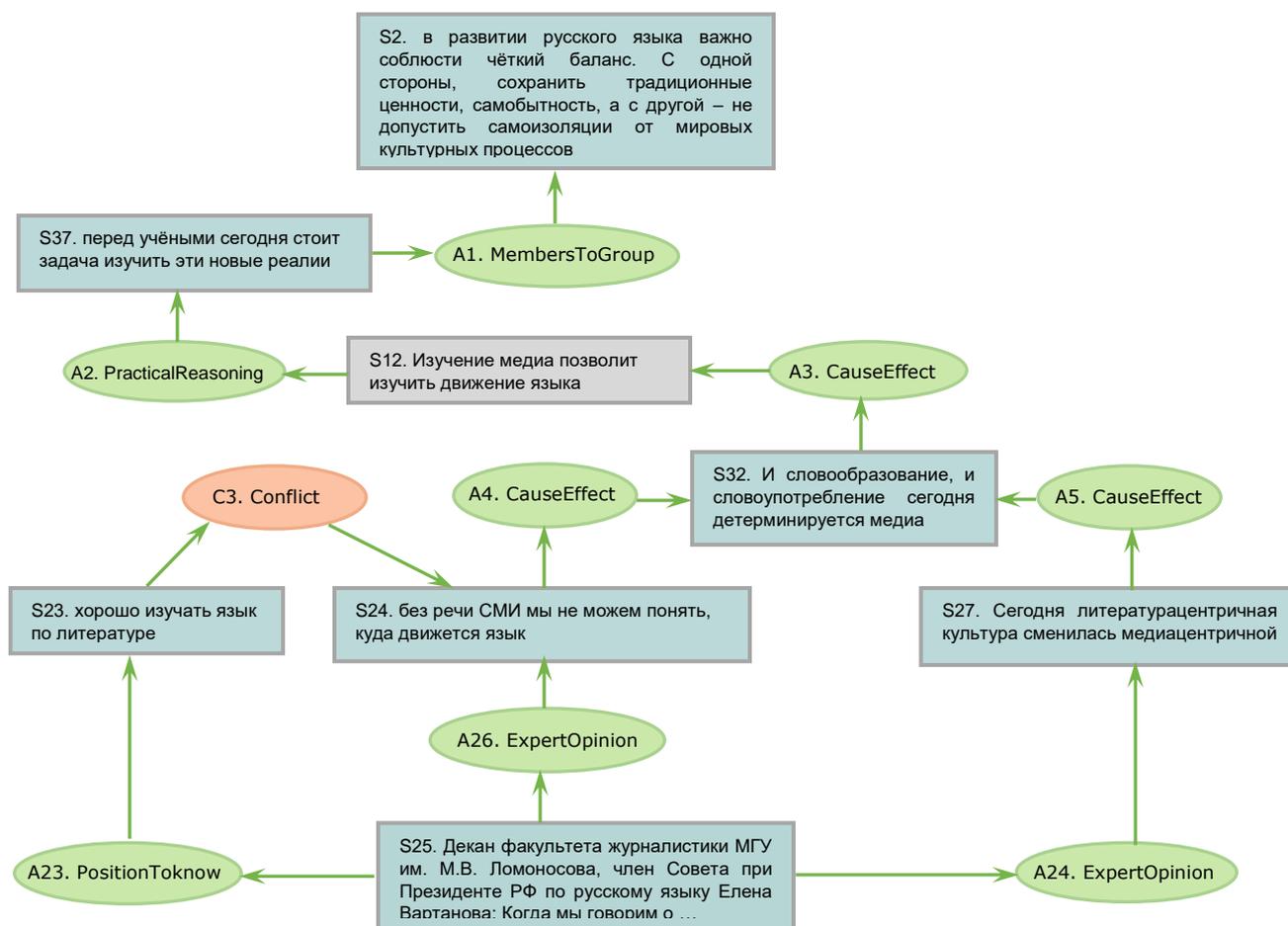


Рис. 2. Фрагмент аргументационной структуры текста «Люди, которые создают нормы»

Если схема аргументации реализуется в одном предложении (доля таких случаев от всех зафиксированных составляет около 15%), то нет необходимости использовать признак принадлежности к одной теме или смежным пересекающимся темам. В других случаях проверка тематического признака может быть полезна. В частности, для всех рассмотренных схем вычисленная с макро-усреднением полнота (R) установления связей между утверждениями составила 76,2%, а точность (Pr) – 48,6%. Явно непредставленные в тексте утверждения не учитывались при подсчете полноты и точности.

Соответствующая разным типам схем рассуждений (представлены 5 самых частотных) полнота установления связей указана в таблице 2. Распознавание этих схем могло бы

послужить основой для автоматического восстановления цепочек аргументов с целью получения максимально полной структуры аргументации.

Таблица 2. Полнота выявления связанных утверждений для разных схем рассуждений

№ п/п	Название схемы	R
1	Example _Inference	78%
2	CauseToEffect _Inference	86%
3	ModusPonens _Inference	77%
4	ExpertOpinion _Inference	72%
5	Sign _Inference	82%

Ошибки в установлении связей случаются из-за наличия в аргументационной структуре большого числа утверждений, явно непредставленных в тексте. Также связи не определяются, если утверждения находятся вне зоны кластеризации. Преимущественно такая ситуация характерна для утверждений, встречающихся в самом начале текста. Трудность в установлении связи представляют собой и позиционно разнесенные утверждения, принадлежащие разным темам. Пересечение множества терминов, характеризующих такие темы, часто отсутствует.

Сравнение точности и полноты распознавания связей, полученных в других работах, возможно только по результатам экспериментов, проведенных на материале английского языка. Но условия экспериментов слишком различаются, сравнение вряд ли можно считать правомерным. Самыми близкими являются результаты работы [3]. Полнота и точность (76,5 % и 72,2 % соответственно) вычислена на очень ограниченном наборе – всего 26 утверждений. В нашей работе получена сравнимая оценка по полноте (76,2 %) и уступающая по точности (48,6 %).

Возможности улучшения метода связаны с самыми разными аспектами автоматической обработки текста. К ним относятся: разрешение анафоры; отбор сочетаний слов определенного типа; изменение параметра, отвечающего за плотность кластера (его уменьшение приводит к образованию более слабых кластеров, но позволяет выявить недостающие связи); учет общих слов и словосочетаний в разных темах; учет конфигурации уже построенной части аргументационной структуры.

Существуют и неустранимые ограничения, чаще всего связанные с недостаточным объемом анализируемого текста, когда либо кластеры не обнаруживаются, либо весь текст попадает в один кластер или одну тему.

6. Заключение

В статье приведены результаты исследования признака принадлежности утверждений к одной теме для установления связи между ними. Тематическая структура строилась автоматически на основе позиционного распределения слов и словосочетаний, выявляемого с помощью сканирующих статистик. Для верификации связей, извлекаемых из тематической структуры, использована коллекция размеченных текстов. Рассматривалась возможность автоматического установления связи между утверждениями, выделенными экспертами в ходе разметки коллекции. Полученные оценки полноты сравнимы с известными, точность несколько уступает. Однако потенциал метода нельзя считать исчерпанным. Исследования по его применению в обнаружении связи между утверждениями будут продолжены.

Список литературы

1. Гусев В.Д., Мирошниченко Л.А., Саломатина Н.В. Тематический анализ и квазиреферирование текста с использованием сканирующих статистик // [Электронный ресурс]. URL: <http://www.dialog-21.ru/media/2374/gusevvd.pdf> (дата обращения: 27.08.2020).
2. Dagan I., Glickman O., and Magnini B. The PASCAL recognising textual entailment challenge // Machine Learning Challenges, Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Textual Entailment, ed. Quinonero-Candela J. and others, Springer, 2006. P. 177–190.
3. Lawrence J., Reed C. and others. Mining arguments from 19th century philosophical texts using topic based modelling. // Proc. of the First Workshop on Argumentation Mining, Baltimore, Maryland, June. Association for Computational Linguistics (ACL), 2014. P. 79–87.
4. Lawrence J., Reed C. Combining Argument Mining Techniques // Proc. of the 2nd Workshop on Argumentation Mining (Denver), 2015. Vol. 1. P. 127–136.
5. Lawrence J., Reed C. Argument Mining: A Survey // Computational Linguistics, 2019. Vol. 45. N 4. P.765–818.
6. Lippi M., Torrony P. Argumentation Mining: State of the Art and Emerging Trends // ACM Transactions on Internet Technology, 2016. Vol. 16. Article 10.
7. Madnani N., Heilman M., Tetreault J., Chodorow M. Identifying high-level organizational elements in argumentative discourse // Proc. of the 2012 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Montreal), 2012. P. 20–28.
8. Mochales Palau R., Moens M-F. Argumentation mining: The detection, classification and structure of arguments in text // Proc. of the 12-th Int. Conf. on Artificial Intelligence and Law (Barcelona). ACM, New York, 2009. P. 98–109.

9. Naus J. The distribution of the size of the maximum cluster of points on a line // Journal of the American Statistical Association. Electronic materials, 1965. Vol. 60. P. 532–538.
10. Rahwan I., Reed C. The argument interchange format // Argumentation in artificial intelligence, ed. Rahwan I. and Simari G. Springer, 2009. P. 383–402.
11. Stab C., Gurevych I. Identifying argumentative discourse structures in persuasive essays // Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing (Doha), 2014. P. 46–56.
12. UDpipe. URL: <http://github.com/ufal/udpipe> (дата обращения: 25.09.2020).
13. Walton D., Reed C., Macagno F. // Argumentation schemes Fundamentals of critical argumentation. New York: Cambridge University Press, 2008. 443 p.

Приложение

Текст статьи «Люди, которые создают нормы» (Авторы: М. Осадчий, Е. Вартанова, Т. Горяева.)

1. 19 мая состоялась встреча Президента Владимира Путина с представителями Совета по межнациональным отношениям и Совета по русскому языку.
2. Президент в ходе общения отметил, что [S2. в развитии русского языка важно соблюсти чёткий баланс.
3. С одной стороны, сохранить традиционные ценности, самобытность, а с другой – не допустить самоизоляции от мировых культурных процессов.]
4. На пресс-конференции, приуроченной ко дню русского языка 6 июня, эксперты обсудили, по каким законам развивается язык, роль СМИ в его изменении и как выработать ответственное отношение у носителей к существующим трансформациям.
5. Проректор по науке Государственного института русского языка им. А.С. Пушкина Михаил Осадчий:
6. Речевая коммуникация – это явление очень динамичное.
7. Мы всегда имеем дело с некоей вариативностью: вариативностью норм, вариативностью реализации тех или иных компетенций и так далее.
8. Сегодня не утихают споры по поводу существования нормы как таковой.
9. Когда мы говорим о нормах в языке, мы в первую очередь имеем в виду словарь.
10. Словарь составлен такими же людьми, причём, как правило, по результатам опросов.
11. То есть, идёт наблюдение за узусом, реальной речевой коммуникацией, на основе которой уже закрепляются те или иные правила.
12. Эти тренды в некоторой степени входят в противоречие с идеей государственного нормирования языка.
13. С одной стороны, русский язык – это язык государства, и тут он совершенно точно должен нормироваться и кодифицироваться.
14. С другой стороны, русский – это язык общества, инструмент межличностного и межнационального взаимодействия.

15. И в этом плане многообразие языка, не выходящее за рамки приличия, должно поддерживаться, развиваться и изучаться.
16. Необходимо поддерживать проекты, связанные со сбором, анализом и популяризацией различной литературы на русском языке, принимая во внимание, что литература является высшей формой функционирования языка.
17. Если носитель, изучая литературу различных эпох, видит, как развивается язык, он по-другому воспринимает саму идею изменчивости языка.
18. Сегодня я часто сталкиваюсь в школе с тем, что учителя однозначно негативно относятся к любым изменениям в языке, будь то произношение, грамматика или общие нормы коммуникации.
19. Но нужно приучать и преподавателей, и учеников к тому, что эти изменения неизбежны.
20. Для этого необходимо вырабатывать некое осмысленное и ответственное отношение к изменениям.
21. В менеджменте есть такой важный элемент любого управленческого процесса – менеджмент изменений.
22. Мы, носители языка, тоже являемся своего рода менеджерами своего языка (кто нам может приказывать, как говорить).
23. Но навыками управления изменениями надо овладевать уже сейчас, чтобы понимать, какие модификации мы принимаем, а какие, мотивированно и осознанно, отвергаем.
24. В любом случае цель государства и образования – формирования у своего населения элитарного отношения к языку и элитарных навыков пользования этим языком.
25. [S25. Декан факультета журналистики МГУ им. М.В. Ломоносова, член Совета при Президенте РФ по русскому языку Елена Варганова:
26. Когда мы говорим о распространении новых норм, всех, конечно, волнует речь журналистов.]
27. Здесь я принимаю на себя все удары.
28. Конечно, журналисты могут говорить неправильно, но я бы хотела выступить в защиту медиаречи.
29. Думаю, что [S23. хорошо изучать язык по литературе.]
30. Так было, особенно в XIX веке, когда литература была главным культурным, интеллектуальным и даже академическим пространством.
31. Но сегодня, в эпоху, когда случилась цифровая революция, когда интернет и маленькие гаджеты проникли в самое интимное пространство человеческого бытия, [S24. без речи СМИ мы не можем понять, куда движется язык.]
32. [S32. И словообразование, и словоупотребление сегодня детерминируется медиа.]
33. При этом, даже если что-то в этом движении нам не нравится, присмотреться к изменениям, которые приживаются в языке после использования журналистами, кажется целесообразным.
34. [S27. Сегодня литературацентричная культура сменилась медиацентричной.]
35. Если раньше в центре был текст, то сейчас его место заняла картинка.

36. Но картинка не обедняет речь, она просто приводит звучащую речь, которая становится важным фактором влияния на речевую среду.
37. Поэтому [S37. перед учёными сегодня стоит задача изучить эти новые реалии], а перед журналистами стоит задача повышения квалификации в области русского языка.
38. К тому, что должно детерминировать норму – правила, разговорный язык или, может быть, язык журналиста – может быть разное отношение.
39. Но очевидно одно: если норма не будет учитывать движение языка, даже в самых просторечных его проявлениях, тогда она потеряет свою легитимность для значительного числа людей.
40. Тут важно прислушаться и к таким новым участникам медиaprостранства, как блогеры.
41. Закон о блогерах, принятый около года назад, может быть несовершенен, но одно он показывает точно: эти люди становятся ориентиром для многих, а значит, должны нести определённую ответственность.
42. С другой стороны, понятно, что без нормы мы можем оказаться в пространстве социального непонимания, потому что речевая коммуникация является основой социальной коммуникации, и здесь нужны правила игры.
43. Сегодня учёные начинают говорить о лингвоэтике как о важнейшем критерии оценки речи в межнациональном государстве, когда словоупотребление может стать причиной конфликтов.
44. На совместном заседании Советов Владимир Путин признал, что важно, чтобы государственные служащие тоже хорошо владели языком и периодически проходили курсы повышения квалификации наравне с работниками СМИ.
45. Директор Российского государственного архива литературы и искусства, член Совета при Президенте РФ по русскому языку Татьяна Горяева: Мы выступаем как фиксаторы нормы.
46. Интересно, что и радиовещание, и телевидение сегодня вернулись к своему первоначальному этапу – к прямому эфиру.
47. Известно, что на пионерском этапе развития телевидения в основном передавались прямые трансляции, и естественно, речь там звучала живая и не всегда корректная.
48. Потом наступила эпоха звукозаписи.
49. Тому поспособствовали разные обстоятельства, в том числе, развитие технических средств записи.
50. Тогда телевидение и радио перешли на этап фиксирования.
51. И уже на наших глазах расцвело непосредственное вещание, когда мы слышали живую речь.
52. И задача здесь – сохранить баланс между непосредственной речью журналиста и соблюдением им норм литературного языка.
53. Что касается цензуры и проверки не только содержания, но и формы подачи информации, сегодня скорее речь идёт о самоцензуре.

54. Ведь внешние институты – Главлиты, Главреперткомы и т.д. – сегодня не функционируют, после знаменитого постановления 1969 года о передаче власти Главлита фактически в руки редакторов СМИ, журналов и газет.

55. С этого момента ответственность за содержание полностью ложилась на плечи главного редактора.

56. Сейчас есть определённые факторы, которые напоминают эти времена.

