ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ / INFORMATICS AND INFORMATION PROCESSES

DOI: https://doi.org/10.60797/IRJ.2024.147.52

КОГНИТИВНЫЙ ПОДХОД К ГЕНЕРАЦИИ ПРОГРАММЫ ВИЗУАЛИЗАЦИИ ПО ТЕКСТУ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

Научная статья

Курбатов С.С.[1], Елистратова В.И.[2,] *
[1] ORCID : 0000-0003-4059-6702;
[1, 2] Научно-исследовательский центр электронной вычислительной техники, Москва, Российская Федерация

* Корреспондирующий автор (atrasword34512[at]gmail.com)

Аннотация

Разработан когнитивный подход для генерации программы визуализации по тексту на естественном языке. Для компьютерного воплощения подхода разработаны оригинальные алгоритмы, моделирующие целостное восприятие текста. Алгоритмы базируются на онтологии предметной области и ориентированы на комплексное решение на всех этапах обработки: лингвистическом, логическом, а также на этапе визуализации. При реализации алгоритмов акцент сделан на объяснимость действий системы и интерактивность визуализации. Инструментальные средства разработки: JavaScript, JSXGraph и MathJax. Рассмотрено взаимодействие этапов обработки на примере достаточно сложной задачи олимпиадного уровня. Обсуждаются результаты экспериментов и дается их оценка. Подчеркнута прикладная значимость подхода для целей образования.

Ключевые слова: когнитивная визуализация, онтология предметной области, обработка естественного языка, компьютерные образовательные ресурсы.

A COGNITIVE APPROACH TO GENERATING A VISUALIZATION PROGRAM FROM NATURAL LANGUAGE TEXT

Research article

Kurbatov S.C.[1], Yelistratova V.I.[2,] *
[1] ORCID : 0000-0003-4059-6702;
[1, 2] Research Centre of Electronic Computing, Moscow, Russian Federation

* Corresponding author (atrasword34512[at]gmail.com)

Abstract

A cognitive approach for generating a visualization programme from natural language text has been developed. For the computer implementation of the approach, original algorithms modelling holistic text perception are designed. The algorithms are based on the ontology of the subject domain and are oriented towards a comprehensive solution at all stages of processing: linguistic, logical, and visualization stages. When implementing the algorithms, the emphasis is on the explainability of system actions and interactivity of visualization. Development tools: JavaScript, JSXGraph and MathJax. The interaction of processing stages is discussed on the example of a rather complex problem of the Olympiad level. Experimental results are discussed, and their evaluation is given. The applied significance of the approach for educational purposes is emphasized.

Keywords: cognitive visualization, subject matter ontology, natural language processing, computer-based educational resources.

Введение

Когнитивная наука в высокой степени междисциплинарна. В психологии предметом когнитивных исследований являются способности к умственному восприятию и переработке внешней информации. В более общем контексте понятие «когнитивность» может быть интерпретировано в культурно-социальном смысле как возникновение знания и связанных с ним концепций. Типичные понятия, которыми оперировало исследование когнитивных процессов, – это память, внимание, восприятие, действие, принятие решений и воображение. Была разработана концепция когнитивных карт для организации пространственных знаний, которые использовались в дальнейшем как метафора для решения непространственных задач.

Из когнитивной науки выделилась важная область – когнитивная лингвистика, обладающая собственными методами исследования, базирующимися на материале языка. Далее парадигма когнитивности существенно расширялась, в частности для понимания того, каким образом на компьютере можно моделировать когнитивные процессы в конкретных предметных областях.

Именно в рамках этого направления, тесно связанного с проблематикой искусственного интеллекта (ИИ), выполнялось данное исследование. Целью данного исследования является разработка когнитивного подхода для генерации программы визуализации по тексту на естественном языке. Актуальность исследования определяется тем, что при обилии работ по обработке естественного языка, семантики предметных областей и визуализации, проблема интеграции результатов в перечисленных областях *с учетом когнитивного аспекта* далека от полноценного комплексного решения.

Базой интеграции является онтология, концентрирующая лингвистические, логические и графические знания. Алгоритмы взаимодействия с онтологией имитируют когнитивные функции целостного анализа текста, логических

рассуждений и устранения когнитивного диссонанса при визуализации. Далее более подробно описываются принципы такой имитации и их реализация в рамках конкретной экспериментальной системы.

Трудности обработки естественного языка в широком контексте описаны в классических работах [1], [2], во многом не утративших своего значения и в наши дни. Трудности реализация ЕЯ-общения в действующей системе, при весьма жестких ограничениях на язык, описаны в [3]. Аналогичные трудности не преодолены и в последующих реально действующих разработках (ЕЯ-интерфейсы к базам данных, экспертным системам и т. п.). В формализованных онтологиях описываются как лингвистические знания, так и семантика предметных областей. Разработано достаточно много инструментов работы с онтологиями (RDF-схемы, варианты OWL и т.д.), обычно используемых в конкретных приложениях.

С когнитивных позиций иерархический подход в онтологиях критикуется, ему противопоставляется теория прототипов, обеспечивающая более гибкую (градуированную) степень принадлежности объекта к концептуальной категории. В когнитивном аспекте визуализация исследуется в педагогике [4]. Справедливо подчеркивается её отличие от собственно визуализации, связанное с направленностью именно на осмысление образного материала, его целостности, на развитие познавательных и критических способностей.

С появлением и бурным развитием (примерно с 2018 года) больших языковых моделей (LLM) фокус исследований обработки естественного языка сместился с традиционной парадигмы. LLM предоставляют возможности обобщения, прогноза и генерации текстов на естественном языке практически в любой области. LLM обучаются на гигантских объёмах текстовых данных и успешно имитируют понимание синтаксиса и семантики текстов. Вопрос о повсеместной применимости LLM требует исследований, поскольку опыт не только показал впечатляющие успехи, но и выявил целый ряд проблем.

Одной из существенных проблем является сложность объяснения, как именно модель машинного обучения пришла к тем или иным выводам. Развиваемое для решения этой проблемы направление «объяснимый искусственный интеллект (Explainable AI)» пока находится на стадии исследований и разработок [5]. Для когнитивного подхода вопрос обоснования принимаемых системой решений является исключительно важным.

Методы и принципы исследования

В соответствии с тем, что целью данного исследования является разработка когнитивного подхода для генерации программы визуализации по тексту на естественном языке, использовались методы обработки естественного языка (ЕЯ), организации онтологии предметной области и интерактивной визуализации. В основу исследования положены принципы *целостности, объяснимости* и *интерактивности*.

Согласно принципу целостности в результате лингвистической обработки текста должна быть выявлена максимально крупная целостная онтологическая структура. С этой структурой связан описывающий её ЕЯ-текст и программа визуализации. Последующий анализ вносит в эту структуру (и программу) уточнения для формирования визуализируемого объекта, в наибольшей степени соответствующего исходному ЕЯ-описанию. Объяснимость обеспечивается протоколом действий системы на каждом шаге принятия решения. В отличие от стандартного log-файла протокол интерактивен и включает объяснения в терминах предметной области с использованием онтологических знаний. Интерактивность позволяет проследить процесс перехода от ЕЯ-описания к визуализации на понятийно-содержательном уровне и в дальнейшем изменять результат визуализации для его корректировки. Это важно не только для пользователя, но и для разработчика, поскольку позволяет с высокой степенью наглядности выявить недостатки системы.

Онтологические структуры включают *комплексные* знания о естественном языке, знания о предметной области и о визуализации (в виде шаблонов программ). Именно комплексность, ориентация на целостный (сверхфразовый) анализ ЕЯ позволяет трактовать их как ***когнитивные***. Знания о логике предметной области могут быть как четко формализованы (в виде теорем), так и носить эвристический характер (правдоподобные рассуждения). Стиль работы с онтологией в рамках конкретной реализации описан в следующем разделе.

Эксперимент

Моделирующие когнитивные процессы алгоритмы исследовались главным образом в предметной области «планиметрическая геометрия». Именно в этой области удалось наиболее отчётливо продемонстрировать суть предлагаемого подхода. В качестве исходных текстов использовались ЕЯ-формулировки геометрических задач. Текст после лингвистической обработки транслировался в семантическое представление (ориентированный граф). Элементы графа имеют вид:

«тип_объекта_1 имя_объекта_1 отношение тип_объекта_2 имя_объекта_2», например «*точка* A является началом *отрезок* AB».

Лингвистическая обработка – помимо построения синтаксической структуры задачи — выделяет словоформы, непосредственно связанные с семантикой. Принципы обработки ЕЯ в области геометрии описаны в [6].

Рассмотрим взаимодействие лингвистики, логики и визуализации на примере достаточно сложной задачи. Она рассматривается как серьёзный успех в автоматическом решении геометрических задач олимпиадного уровня [7]. Фрагмент ЕЯ-описания задачи, достаточный для предварительного построения чертежа: «Пусть ABC – остроугольный треугольник, в котором AB > AC. Пусть okr – окружность, описанная около него, H – его ортоцентр. Пусть M – середина стороны BC. Пусть Q – точка на окружности okr такая, что ∠HQA = 90°».

Фрагмент насыщен семантическими словоформами: «треугольник», «окружность», «ортоцентр», «угол» – объекты, «>», «описанная», «на» – отношения, «остроугольный», «90°» – атрибуты. Наиболее крупная целостная онтологическая структура, извлекаемая по этим маркерам, имеет ЕЯ-описание, семантическое представление и заготовку программного текста для визуализации. ЕЯ-описание: «Окружность #O описана около треугольника #X#Y#Z» с ортоцентром #H». Символ # означает, что это шаблонные элементы, которые должны быть означены

реальными данными из текста (A, B и т.д.). Тем самым в когнитивном подходе *имитируется* возникновение в памяти человека аналогичной структуры и её адаптации для построения конкретного чертежа (визуализация).

Семантическое представление имеет представленный выше вид и по сути детально описывает процесс формирования чертежа (построить точку A, построить точку B, построить отрезок AB и т.д.). Программа визуализации формируется именно по семантическому представлению. С помощью менее крупных онтологических структур семантическое представление дополняется новыми элементами на основе уже означенными на предыдущем шаге (например, «точка A на окружности» и «точка B на окружности» могут формировать дополнение «хорда AB»). Так *имитируется* детализация и дополнение предварительного чертежа.

Вид чертежа на этих этапах отображен на рисунке 1.

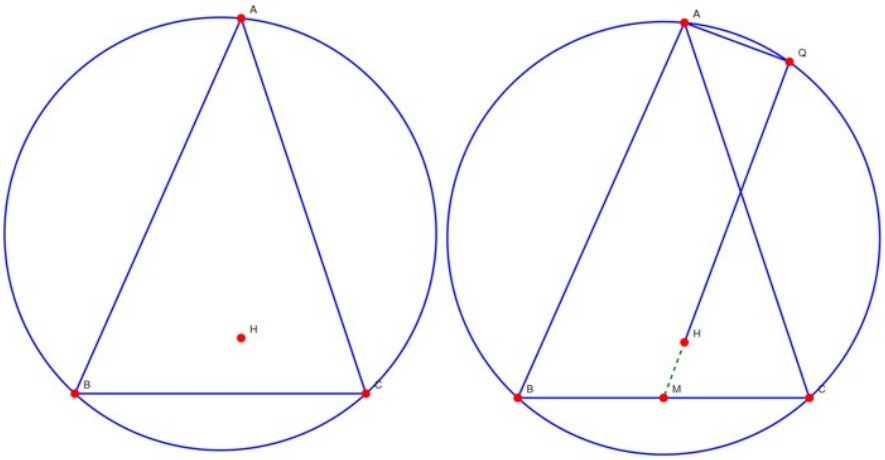

Рисунок 1 - Вид предварительного чертежа до и после вызова онтологической структуры
DOI: https://doi.org/10.60797/IRJ.2024.147.52.1

Важно, что даже после предварительного построения система может обратиться к логике предметной области (теоремы) для деталей организации чертежа. В данном случае для проверки эмпирического (на чертеже!) нахождения точек Q, H, и M на одной прямой вызывается соответствующая теорема. Процесс её вызова базируется на анализе семантической структуры и относится скорее к решению задачи. Технически вызов теоремы достаточно сложен, но в идейном плане аналогичен вызову целостной структуры для визуализации. Результатом применения теоремы является включение в семантическую структуру элементов типа «точка на прямой». Когнитивные аспекты решения задачи здесь не обсуждаются.

В результате визуализации на экран выводится ЕЯ-текст задачи и чертёж. Интерактивность означает, что процесс построения чертежа можно проследить по шагам с соответствующими пояснениями. Кликом по элементам чертежа и по ЕЯ-тексту можно получить дополнительную информацию о данном элементе. Стиль пояснения: «Сторона AB треугольника, больше стороны AC по условию». Элементы чертежа можно перемещать (сохраняя по возможности условия задачи) для устранения недостатков визуализации (значения атрибутов объектов, отношения перпендикулярность/параллельность и т.д.).

Важность средств, позволяющих скрывать элементы чертежа и маркировать их, демонстрирует рисунок 2.

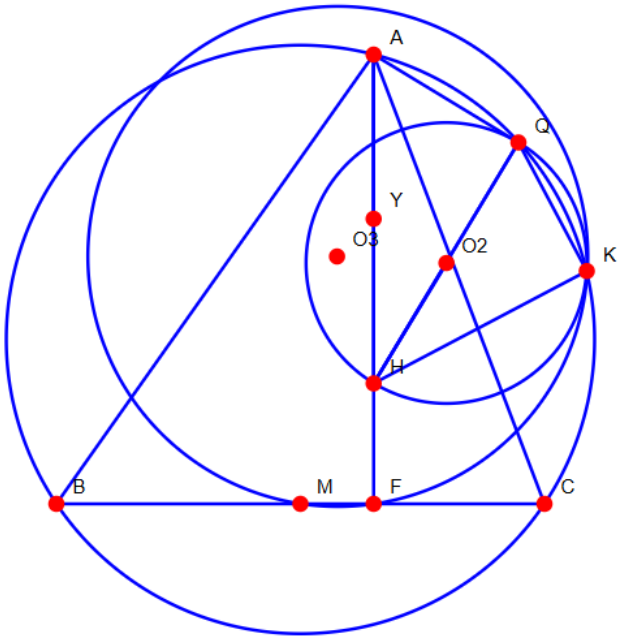

Рисунок 2 - Реальный чертеж, соответствующий полному тексту задачи
DOI: https://doi.org/10.60797/IRJ.2024.147.52.2

Обсуждение

Описанный в предыдущем разделе эксперимент даёт общее представление о выполнении алгоритмов в конкретной предметной области. Однако область применения подхода значительно шире, разумеется, при соответствующей модификации онтологии. Отдельные расширения были намечены для алгебраических задач ЕГЭ и генерации 3D-конструкций. Расширение для алгебры важно, поскольку в тексте геометрических задач возможны математические выражения (арифметические, алгебраические, логические, аналитические).

В текущей реализации семантический граф допускает тип «выражение» и обеспечивает древовидное представление выражений. Арность операторов выражения не ограничена, например, «определённый интеграл» обладает аргументами «нижний предел», «верхний предел», «подынтегральная функция» и «элемент интегрирования». Визуализация семантического представления выражения сводится к формированию линейной структуры с учетом типов операторов.

Текущая реализации алгоритмов выполнена на VBA (Excel), программа визуализации формируется на JavaScript, вызываемой из HTML-страницы. Инструментальные средства разработки: JavaScript, JSXGraph [8] и MathJax. Две последние библиотеки написаны на JavaScript: одна используется для построения чертежей, вторая для отображения сложных математических формул.

Анализ выражений, входящих в ЕЯ-описание, выполняется по упрощенной схеме (ограничение вложенности операций, предобработка для сложных случаев и т.п.). Тем не менее для корпуса из более 600 планиметрических задач эта схема оказалась успешной. Развитие алгоритмов анализа выражений и их интерпретации в стиле [9] предполагается в дальнейшем.

Элементы теории прототипов включены в онтологию. Не жесткая таксономия, но более гибкая классификация, включающая маркеры «более близок прототипу», «более знаком», «более интересен». Аналогичны маркеры для теорем, например, «сложная теорема» определяется длиной пути вывода до аксиоматики (все предшествующие теоремы). При этом когнитивным понятиям в онтологии придаётся строгий смысл. Например, «равносторонний/прямоугольный треугольник» *более интересны* с точки зрения решения задачи, поскольку они требуют *меньше* параметров для своего описания. Поэтому для выполнения дополнительных построений выбираются именно они (объяснимость). Аналогично, углы 60°, 90°, 45° и т.д. также *более интересны*, поскольку известны значения их тригонометрических функций и эти углы входят в ряд важных теорем.

Наглядность чертежа является хорошей основой для выдвижения важных для решения индуктивных предположений: параллельности или перпендикулярности линий, соотношений углов/отрезков и т.п. Интерактивные возможности позволяют выделить ключевые моменты чертежа (скрыть или выделить цветом отдельные элементы). Близкие вопросы, связанные с визуализацией и ориентированным на человека объяснением в диалоге, обсуждаются в работах [10], [11]. Недостатки нейросетей для задач на ЕЯ по математике отмечены в [12].

Полученные результаты позволили сделать вывод о перспективности развиваемого подхода, и не только в узких предметных областях. В подходе намечено моделирование «здравого смысла», что является важной проблемой в области ИИ. Ошибки нейронных сетей при ответе на вопросы, не вызывающие трудностей у человека, связаны именно с отсутствием в модели здравого смысла. Эти ошибки весьма редки, но принимаемые на основании рекомендаций нейросети решения могут быть весьма серьёзными.

На примере известной фразы «обезьяны не стали есть бананы, потому что *они* были сыты» обсудим различие нейросетевого и когнитивного подходов. Нейросеть на основе анализа огромного массива текстов выясняет по

признаку частотности, что «*они*» относится к обезьянам. Аналогично, если вместо «сыты» во фразе будет «гнилые», то «они» — это бананы (также по частоте). Человек по здравому смыслу определяет, что фраза описывает процесс: есть тот, *кто ест*, и то, *что едят*.

Человек быстро осознаёт, какие характеристики могут препятствовать этому процессу: «сыты, болели животы, боялись наказаний…» или «гнилые, отравленные, воткнуты гвозди, с шипами…». Поэтому не столько частотность, сколько наличие когнитивной структуры (здравого смысла) позволяет сделать правильный вывод, даже если вместо «обезьяны» и «бананы» будут стоять фиктивные имена (типа X и Y).

Заключение

В научном плане развиваемый подход позволит по-новому взглянуть на проблему разработки систем, моделирующих когнитивные способности человека. Комплексная, семантически-ориентированная интеграция лингвистических, логических и интерактивно-визуальных возможностей системы безусловно будет востребована в когнитивных исследованиях. В прикладном плане исследование открывает перспективу создания образовательных ресурсов, демонстрирующих учащимся тесную связь школьного курса и современных проблем информационной обработки.

Конфликт интересов

Не указан.

Рецензия

Гибадуллин Р.Ф., Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ, Казань, Российская Федерация
DOI: https://doi.org/10.60797/IRJ.2024.147.52.3

Conflict of Interest

None declared.

Review

Gibadullin R.F., Kazan National Research Technical University named after A.N. Tupolev – KAI, Kazan, Russian Federation
DOI: https://doi.org/10.60797/IRJ.2024.147.52.3

Список литературы / References

1. Мельчук И.А. Русский язык в модели «смысл — текст» / И.А. Мельчук. — Москва: Языки славянской культуры, 1995. — 682 с.

2. Апресян Ю.Д. Теоретические проблемы русского синтаксиса: взаимодействие грамматики и словаря / Ю.Д. Апресян, И.М. Богуславский, Л.Л. Иомдин [и др.]. — Москва: Языки славянской культуры, 2010. — 408 с.

3. Попов Э.В. Общение с компьютером на естественном языке / Э.В. Попов. — Springer, 2011. — 308 с.

4. Богатырева Ю.И. Методические аспекты активизации обучения с использованием цифровых образовательных ресурсов / Ю.И. Богатырева, А.П. Углова // Международный научно-исследовательский журнал. — 2024. — № 6(144). — URL: https://research-journal.org/en/archive/6-144-2024-june/10.60797/IRJ.2024.144.145 (дата обращения: 09.07.2024).

5. Корешкова Т. Объяснимый искусственный интеллект / Т. Корешкова. — URL: https://rdc.grfc.ru/2020/12/explainable-ai/ (дата обращения: 09.07.2024).

6. Курбатов С. Лингвистический процессор интегрированной системы решения планиметрических задач / С. Курбатов // Международный журнал когнитивной информатики и естественного интеллекта. — 2021. — № 15(4). — 14 с.

7. Тринх Т. Решение олимпиадной геометрии без демонстрации человека / Т. Тринх, Ю. Ву, Ц. Ле [и др.] // Nature. — 2024. — № 625. — С. 476–482.

8. JSXGraph. — URL: https://jsxgraph.uni-bayreuth.de/share/ (accessed: 09.07.2024).

9. Крюгер Р. Автоматическое построение диаграмм для олимпиадных задач по геометрии / Р. Крюгер, Дж.М. Хан, Д. Селсам // Материалы CADE 28 — 28-ой международной конференции по автоматической дедукции. — 2021. — 22 с.

10. Фонт Л. Автоматизация генерации доказательств по геометрии в средней школе с помощью Пролога в образовательном контексте / Л. Фонт, С. Сир, П. Ричард [и др.] // Материалы 8-го международного семинара по компонентам доказательства теорем для образовательного программного обеспечения, ThEdu@CADE. — Натал, 2019. — 16 с.

11. Фонт Л. Интеллект в QED-Tutrix: Баланс взаимодействия между естественным интеллектом пользователя и искусственным интеллектом программного обеспечения-репетитора / Л. Фонт, М. Ганьон, Н. Ледюк [и др.] // Математическое образование в эпоху искусственного интеллекта. — 2022. — Т. 17. — С. 45–76.

12. Коббе К. Обучение верификаторов решению словесных задач по математике / К. Коббе, В. Косараджу, М. Бавариан [и др.]. — URL: https://arxiv.org/abs/2110.14168 (дата обращения: 09.07.2024).

Список литературы на английском языке / References in English

1. Mel'chuk I.A. Russkij jazyk v modeli «smysl — tekst» [Russian language in the meaning-text perspective] / I.A. Mel'chuk. — Moscow: Languages of Slavic culture, 1995. — 682 p. [in Russian]

2. Apresjan Ju.D. Teoreticheskie problemy russkogo sintaksisa: vzaimodejstvie grammatiki i slovarja [Theoretical problem of Russian syntax; interaction of grammar and vocabulary] / Ju.D. Apresjan, I.M. Boguslavskij, L.L. Iomdin [et al.]. — Moscow: Languages of Slavic culture, 2010. — 408 p. [in Russian]

3. Popov Je.V. Obshhenie s komp'juterom na estestvennom jazyke [Talking with computers in natural language] / Je.V. Popov. — Springer, 2011. — 308 p. [in Russian]

4. Bogatyreva Ju.I. Metodicheskie aspekty aktivizacii obuchenija s ispol'zovaniem cifrovyh obrazovatel'nyh resursov [Methodological aspects of activating learning using digital educational resources] / Ju.I. Bogatyreva, A.P. Uglova // Mezhdunarodnyj nauchno-issledovatel'skij zhurnal [International Research Journal]. — 2024. — № 6(144). — URL: https://research-journal.org/en/archive/6-144-2024-june/10.60797/IRJ.2024.144.145 (accessed: 09.07.2024). [in Russian]

5. Koreshkova T. Ob'jasnimyj iskusstvennyj intellekt [Explicable artificial Intelligence] / T. Koreshkova. — URL: https://rdc.grfc.ru/2020/12/explainable-ai/ (accessed: 09.07.2024). [in Russian]

6. Kurbatov S. Lingvisticheskij processor integrirovannoj sistemy reshenija planimetricheskih zadach [Linguistic processor of the integrated system for solving planimetric problems] / S. Kurbatov // Mezhdunarodnyj zhurnal kognitivnoj informatiki i estestvennogo intellekta [International Journal of Cognitive Informatics and Natural Intelligence]. — 2021. — № 15(4). — 14 p. [in Russian]

7. Trinh T. Reshenie olimpiadnoj geometrii bez demonstracii cheloveka [Solving olympiad geometry without human demonstrations] / T. Trinh, Ju. Vu, Q. Le [et al.] // Nature. — 2024. — № 625. — P. 476–482. [in Russian]

8. JSXGraph. — URL: https://jsxgraph.uni-bayreuth.de/share/ (accessed: 09.07.2024).

9. Krjuger R. Avtomaticheskoe postroenie diagramm dlja olimpiadnyh zadach po geometrii [Automatically Building Diagrams for Olympiad Geometry Problems] / R. Krjuger, Dzh.M. Han, D. Selsam // Materialy CADE 28 — 28-oj mezhdunarodnoj konferencii po avtomaticheskoj dedukcii [Automated Deduction — CADE 28: 28th International Conference on Automated Deduction]. — 2021. — 22 p. [in Russian]

10. Font L. Avtomatizacija generacii dokazatel'stv po geometrii v srednej shkole s pomoshh'ju Prologa v obrazovatel'nom kontekste [Automating the Generation of High School Geometry Proofs using Prolog in an Educational Context] / L. Font, S. Sir, P. Richard [et al.] // Materialy 8-go mezhdunarodnogo seminara po komponentam dokazatel'stva teorem dlja obrazovatel'nogo programmnogo obespechenija, ThEdu@CADE [Proceedings 8th International Workshop on Theorem Proving Components for Educational Software, ThEdu@CADE]. — Natal, 2019. — 16 p. [in Russian]

11. Font L. Intellekt v QED-Tutrix: Balans vzaimodejstvija mezhdu estestvennym intellektom pol'zovatelja i iskusstvennym intellektom programmnogo obespechenija-repetitora [Intelligence in QED-Tutrix: Balancing the interactions between the natural intelligence of the user and the artificial intelligence of the tutor software] / L. Font, M. Gan'on, N. Ledjuk [et al.] // Matematicheskoe obrazovanie v jepohu iskusstvennogo intellekta [Mathematics Education in the Digital Era]. — 2022. — Vol. 17. — P. 45–76. [in Russian]

12. Kobbe K. Obuchenie verifikatorov resheniju slovesnyh zadach po matematike [Training Verifiers to Solve Math Word Problems] / K. Kobbe, V. Kosaradzhu, M. Bavarian [et al.]. — URL: https://arxiv.org/abs/2110.14168 (accessed: 09.07.2024). [in Russian]