

ОНТОЛОГИИ И ЯЗЫКИ ДЛЯ ПРЕДСТАВЛЕНИЯ МАТЕМАТИЧЕСКИХ ЗНАНИЙ В SEMANTIC WEB

Стаселько И.Д., Позняков Т.Д., Протасов А.П.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

*Анисимов В.Я. – канд. физ.-мат. наук, доцент
Теслюк В.Н. – канд. мат. наук, доцент*

Предоставлена информация об использовании математических знаний совместно с технологиями Web 2.0 и Semantic Web.

Ключевые слова: математические знания, семантические технологии, искусственный интеллект, семантические сети, Web 2.0.

Обзор состояния математики в сети Интернет должен учитывать веб-сайты, которые находятся в повседневном употреблении: реферативные службы, такие как Math Zentralblatt и MathSciNet, arXiv – предварительный сервер печати, библиотеки формализованного и проверенного алгоритмом математического содержания, такие как Mizar Mathematical Library (MML). Эти сайты облегчили доступ к математическому материалу. Однако они предлагают ограниченный уровень взаимодействия и не облегчают пользование, и средства автоматического восстановления, использования и адаптивного представления знания через автоматизированных агентов ограничены.

Относительно сети Интернет в целом эта проблема и проблемы Семантической паутины (Semantic Web) были решены приложениями Web 2.0. В данной работе рассматриваются проблемы, связанные с получением и обработкой математических знаний в сети Интернет, и предложения по их преодолению с использованием новой комбинации Web 2.0 и Semantic Web. После того, как недавно математики освоили технологию Web 2.0, возросло число математиков, которые начали использовать данную технологию для того, чтобы совместно развивать новые идеи, а также использовать как новый поток для публикации полученных знаний. Типичные представители указанных технологий – блог исследований и энциклопедия Wiki. Такие блоги отражают традиционную математическую практику публикации кратких обзоров ранее изданного материала.

Исследователи также считают блоги полезными для того, чтобы собрать раннюю связь между предварительными результатами, тогда как традиционные государственные публикации, предназначенные для экспертной оценки, представляют зрелые идеи, скрывая альтернативы, которые когда-то рассмотрели и затем от них отказались, – такой метод работы сильно препятствует сотрудничеству между малочисленными исследовательскими группами. Успешное сотрудничество среди математиков, не знающих друг друга прежде, началось в блогах и имело результатом последующие совместные статьи. По сравнению с блогами исследования форум MathOverflow, где пользователи могут отправить свои проблемы и давать решения проблем других пользователей, предлагает больше помощи с меньшими рисками. Основой репутации этого форума является то, что он действует как качественное моделирование традиционных научных процессов публикации и экспертной оценки. Для развития идей, появившихся из обсуждений блога, или для создания постоянных, коротких, связанных описаний тем Wiki считают наиболее подходящей.

Также следует упомянуть сервис Tricky, который является хранилищем общих математических методов Wiki, работающий на Web 2.0. Wiki, которая собирает существующие математические знания для образовательных и ознакомительных целей, известна более широко. PlanetMath является математической энциклопедией и насчитывает более, чем 8000 записей на время данных исследований. Wikipedia, имеющая 15 миллионов статей на более чем 250 языках, также затрагивает математику. Предназначаясь для широкой аудитории, она опускает большинство формальных доказательств, но включает чистое математическое знание в более широкий контекст, а также включает, например, историю математики, биографию математиков и информацию об областях ее применения. Отсутствие доказательств частично компенсировано благодаря технически подобной ProofWiki или PlanetMath, содержащими более 2500 доказательств.

Наконец, Connexions, управляемый более традиционной системой работы с контентом, является открытым веб-хранилищем, специализированным на программном обеспечении учебного курса. Connexions продвигает более 17000 небольших модулей курса многократного использования его стандарту свободного содержания так, чтобы не только автор, но и другие пользователи могли свободно объединить их в сборники, такие как примечания к определенному курсу, приблизительно

4000 от математики и статистики и приблизительно 6000 от науки и техники. Разработчики установленных баз знаний также начинают использовать технологию Web 2.0 – например, недавно разработанный прототипный макет Wiki frontend для Mizar Mathematical Library (MML) – это крупная библиотека формализованного и проверенного машиной математического контента. Wiki намеревается осуществить общую поддержку и улучшение MML.

Сайты на Web 2.0 требуют крупных инвестиций рабочей силы для компилирования сборников знаний. Повторного использования знаний с применением машинного интеллекта, например из других сетевых коллекций знаний, не происходит. Различные базы знаний технически отделены друг от друга из-за форматов документа, которые подходят для одиночного представления данных, такие как XHTML с LATEX формулами.

Единственным способом присоединиться к другим базам знаний является гиперссылка. Методы доказательств, собранные в Tricky, не могут быть автоматически применены к проблемам разработки в блоге исследования, поскольку ни в одном из них недостаточно формул. С другой стороны, сообщество Polymath не имеет достаточно автоматизированных инструментов проверки, но оно полагается на краудсорсинговый принцип. Также плохо поддерживается поиск математической информации, являющийся предпосылкой для поиска знаний к применению и повторному использованию.

Например, Википедия указывает теорему Пифагора и регистрирует ее в категории «Статьи с доказательством» и «Математические теоремы». Представление формул в формате LATEX не поддерживает поиск функциональных структур.

Несмотря на то, что Википедия не может искать формулы, поиск эквивалентных выражений не привел бы к теореме, если они явно не указаны в статье. Если смотреть со стороны классификации, это в любом случае непонятно для машины (хотя, очень вероятно, понятно для человека), содержит ли статья доказательство конкретной теоремы и правильно ли оно.

Следовательно, сотрудники Polymath должны были искать предыдущие публикации и опровержения с доказательством и ключевым словом. Формализованные хранилища, такие как MML, специализированные поисковые системы, также поддерживают внутреннее повторное использование знаний, формулируя новые математические понятия для уже существующих и доказывая новые теоремы, которые уже были доказаны. Также находятся в дефиците математические сайты на Web

с автоматизированными услугами по рассуждению и вычислению. Интерактивное вычисление доступно в математических системах электронного обучения, таких как ActiveMath или MathDox, где авторы документа прежде, чем издать в цифровых библиотеках общего назначения и в среде сообщества, достаточно сформулировали основную математику в отдельных инструментах редактирования. Мэшапы, которые являлись движущей силой развития Web 2.0, сейчас едва используются для математических задач.

В начале 2000-х, когда XML все больше использовался для математики, особенно для формул, которые были первыми стандартными блоками Семантической паутины, такие как RDFS, приблизились к стандартизации. Эти события вызвали интерес к появившемуся междисциплинарному сообществу математического управления знаниями, состоящему из программистов, опытных математиков и цифровых исследователей библиотек, цель которых состоит в том, чтобы «развивать новые и лучшие способы управления математическими знаниями, используя сложные программные средства» [1]. Или, более конкретно, «служить математикам, ученым и инженерам, которые производят и используют математическое знание; педагогам и студентам, которые преподают и изучают математику; издателям, которые предлагают математические учебники и распространяют новые математические результаты; и библиотекарям и математикам, которые каталогизируют и организуют математическое знание» [1].

Список использованных источников:

1. Marchiori, M. *The Mathematical Semantic Web* / M. Marchiori // *Mathematical Knowledge Management : Second International Conference, Bertinoro, Italy, February 16–18, 2003* / Springer-Verlag. – Berlin, Heidelberg, 2003. – P. 216–223.