

МЕТОД РАЗРАБОТКИ СРЕДСТВ УПРАВЛЕНИЯ ОНТОЛОГИЧЕСКИМИ ДАННЫМИ В ТРЕХМЕРНОМ ВИРТУАЛЬНОМ ПРОСТРАНСТВЕ

Д. С. Раздьяконов,
НИУ ИТМО

Аннотация: Цель исследования – сокращение времени и трудоёмкости построения визуальных репрезентаций онтологических данных в трехмерном виртуальном пространстве для навигации и поиска по графам знаний. Сокращение времени достигается за счёт оптимизации пользовательского интерфейса с использованием модели интерактивной визуализации и применения контекстно-зависимых идентификаторов анонимных узлов.

Ключевые слова: визуализация, графические интерфейсы, онтологические данные, графы знаний, анонимные узлы, трехмерная графика.

Онтологии применяются в широком спектре технических сфер от областей, связанных с машинным обучением и моделированием бизнес-процессов, до искусственного интеллекта и поисковых систем. Для ускорения выполнения пользовательских задач применяются программные средства визуализации и визуального редактирования онтологических данных.

При использовании средств визуального редактирования онтологических данных (ОД) эффективность взаимодействия пользователя с интерфейсом приложения напрямую влияет на скорость выполнения целевых задач. Особенно остро проблема взаимодействия пользователя с графическим интерфейсом встаёт при работе с трехмерными визуализациями.

В рамках данной работы мы выделяем узкие места существующих средств управления онтологическими данными и формируем метод

оптимизации графического интерфейса приложения, включающий применение подхода к визуализации анонимных узлов онтологий [11]. Анонимные узлы являются узким местом большинства средств визуализации ОД, так как не имеют фиксированных идентификаторов, что делает невозможным осуществлять порционную визуализацию ОД.

Работы [1], [2], [3] выполняют подробный и многосторонний разбор средств и методик визуализации ОД, а в работах [4], [5], [6], [7] производится разбор методик разработки онтологий и даётся обзор средств, поддерживающих разработку. [8], [9], [10] посвящены разработки пользовательского интерфейса.

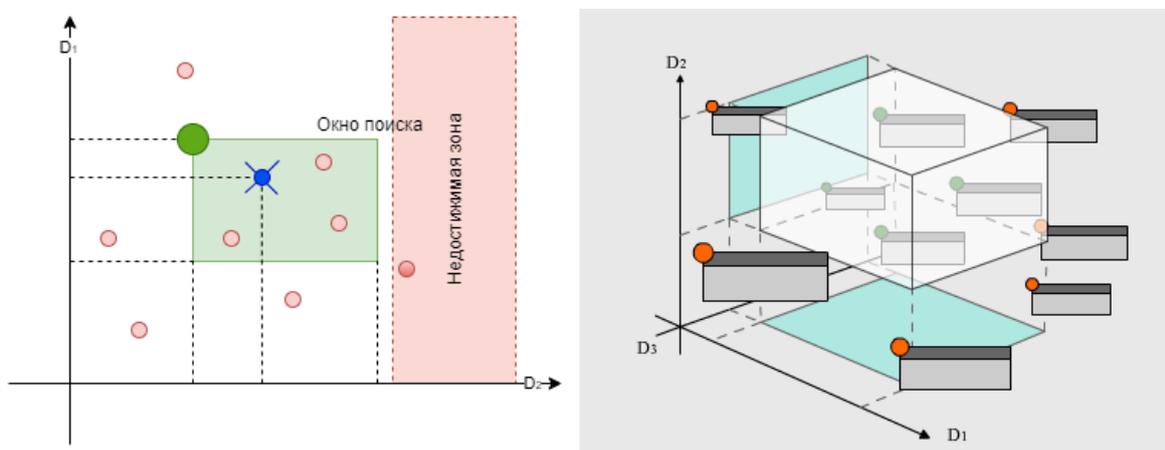


Рисунок 1 – 2D/3D пространства данных и окно поиска

Суть метода заключается в применение модели интерактивной визуализации для оптимизации пользовательского интерфейса. Основа модели – это онтологический граф. Онтологический граф – это конечный набор троек $\langle \text{subject, predicate, object} \rangle$. В рамках модели мы формируем пространство данных (рисунок 1), где каждому измерению соответствует предикат из онтологии. То есть измерение – это множество возможных объектов троек с соответствующим предикатом. Имея многомерное декартово пространство данных, мы выдаём каждому элементу визуализации координаты в этом пространстве.

Конечная визуализация содержит некоторое множество элементов визуализации, которые имеют свои координаты в пространстве данных, что даёт возможность определить параметры визуализации, то есть позицию и объём визуализации («Окно Поиска» – рисунок 1). В процессе выполнения целевой задачи пользователь изменяет состояние визуализации, то есть меняет положение окна поиска в пространстве данных. В рамках нашей работы мы ведём учёт действий пользователя и положений окна поиска в пространстве данных и составляем на их основе граф состояний визуализации, где каждая точка графа состояний имеет координату в пространстве данных. Вершинами графа в данном случае становятся визуализации, а ребра обозначают возможность перехода от одного состояния визуализации к другому, более того мы учитываем то, сколько действий совершил пользователь и на какое расстояние в пространстве данных он при этом продвинулся, это позволяет оптимизировать путь в графе (рисунок 2) состояний и исключить неэффективные пути.

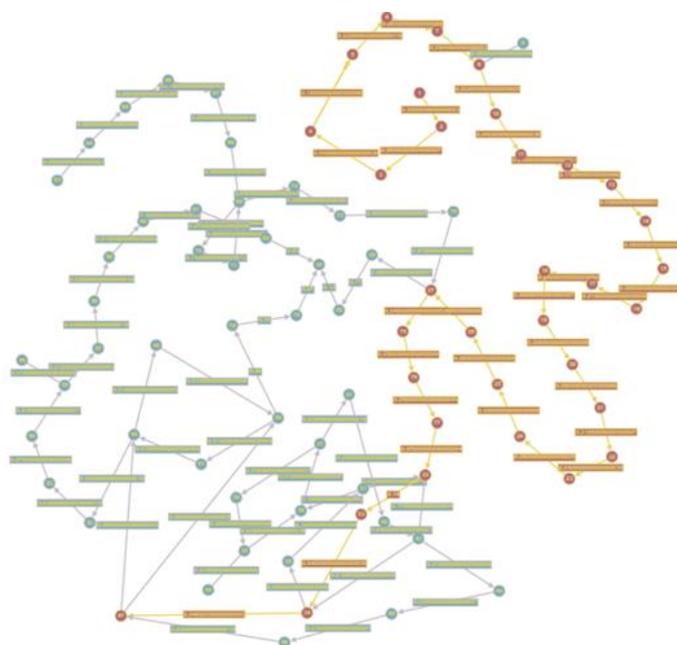


Рисунок 2 – Путь в пространстве состояний

Вторая часть метода описывает применение подхода к визуализации анонимных узлов в рамках порционной визуализации. Подробнее о подходе

можно прочитать в статье «Approach to blank node processing in incremental data visualization by the example of ontodia» [12]. При порционной визуализации структур, содержащих анонимные узлы, нет возможности частично визуализировать подобную структуру. Структура визуализируется или целиком, или не визуализируется вовсе.

Распространенные средства визуализации, чтобы иметь возможность работы с подобными структурами на уровне хранения ОД, заменяют анонимные идентификаторы стандартными и работают в обычном режиме, однако такая возможность есть не всегда, особенно при работе с большими данными. В таких случаях пользователям при использовании средств визуализации приходится пользоваться обходными путями, то есть: использовать именные метки, визуализировать различные части структур через сторонние предикаты, выполнять ручные SPARQL-запросы.

С точки зрения модели интерактивной визуализации это выливается в длинные неоптимальные пути на графе состояний. Для сокращения времени и трудоёмкости построения визуальных репрезентаций, содержащих анонимные узлы, в рамках метода мы применяем контекстно-зависимые идентификаторы для анонимных узлов.

Суть подхода заключается в том, чтобы анализировать результаты запросов к базе знаний (БЗ) и, в случае появления анонимных узлов, запрашивать весь контекст анонимного узла, где контекст анонимного узла – это подграф основного графа, содержащий целевой анонимный узел и транзитивно все анонимные узлы, заключенные между обычными узлами, окружающими данный подграф, и включающий их. После того как контекст анонимного узла собран, производится канонизация контекстного подграфа и сжатие по префиксному дереву.

На последнем шаге, сформированный сжатый граф используется для формирования контекстно зависимых идентификаторов (рисунок 3), которые представляют собой пары из индекса в графе контекста и самого сжатого контекста. Данные идентификаторы, при использовании в визуализации,

обрабатываются и используются для порционной визуализации ранее не визуализированных частей структур, содержащих анонимные узлы.

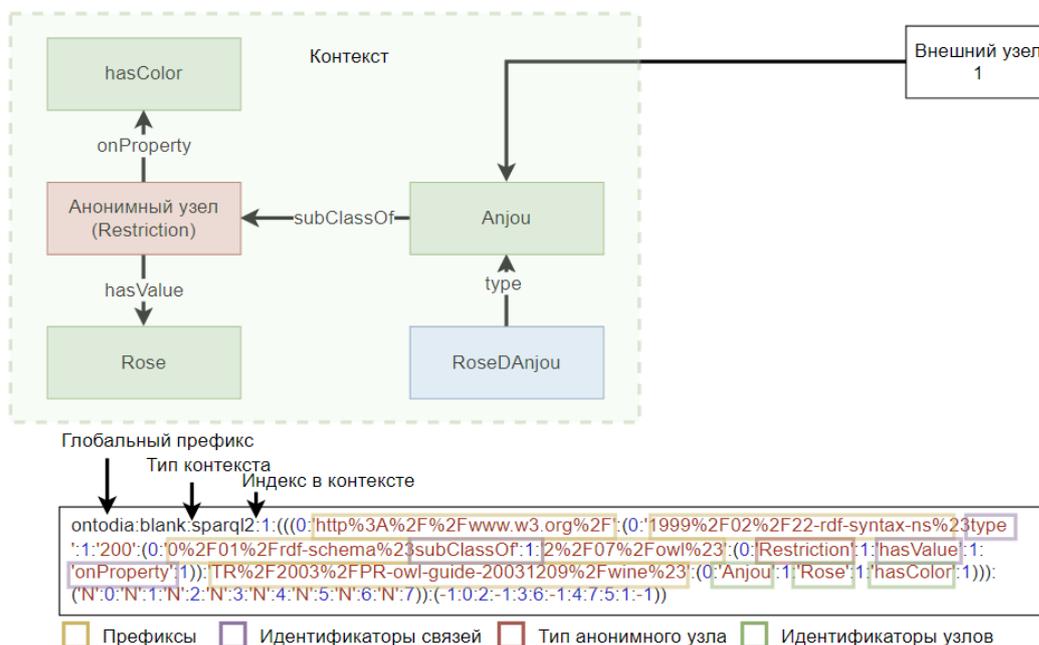


Рисунок 3 – Граф контекста и контекстно зависимый идентификатор

В случае повторного получения похожих структур, благодаря алгоритму канонизации графа, генерируется эквивалентный идентификатор, что позволяет однозначно установить была ли целевая структура уже визуализирована или нет, что невозможно при прямой визуализации подобных структур.



Рисунок 4 – Граф контекста и контекстно зависимый идентификатор

На финальном этапе работы были проведены эксперименты на испытательном стенде. В качестве целевого инструмента для модификации с

использованием предложенного метода был выбран инструмент Ontodia3d [13], а в качестве онтологии для тестирования была выбрана бенчмарк онтология вин (<http://www.w3.org/TR/owl-guide/wine.rdf>).

Эксперименты показали, что средняя трудоемкость взаимодействия пользователя с графическим интерфейсом сократилась на 21% (рисунок 4), а использование контекстно-зависимых идентификаторов позволило увеличить объём доступного для визуализации пространства с 31% до 100% (процентное соотношение 31/100 верно только для целевой онтологии вин). Уменьшение трудоёмкости и увеличение охвата пространства позволило ускорить построения визуальных репрезентаций онтологических данных пользователями, что и являлось целью проделанной работы.

Список литературы

[1] Katifori A. et al. Ontology visualization methods—a survey //ACM Computing Surveys (CSUR). – 2007. – Т. 39. – №. 4. – С. 10-es.

[2] Dudáš M. et al. Ontology visualization methods and tools: a survey of the state of the art //The Knowledge Engineering Review. – 2018. – Т. 33. – С. e10.

[3] Sivakumar R., Arivoli P. V. Ontology visualization PROTÉGÉ tools—a review //International Journal of Advanced Information Technology (IJAIT) Vol. – 2011. – Т. 1.

[4] Gokhale P., Deokattey S., Bhanumurthy K. Ontology development methods //DESIDOC Journal of Library & Information Technology. – 2011. – Т. 31. – №. 2.

[5] Jones D., Bench-Capon T., Visser P. Methodologies for ontology development. – 1998.

[6] Aminu E. F. et al. A review on ontology development methodologies for developing ontological knowledge representation systems for various domains. – 2020.

[7] Breitman K. K., Casanova M. A., Truszkowski W. Methods for ontology development //Semantic Web: Concepts, Technologies and Applications. – 2007. – С. 155-173.

[8] Fischer H., Yigitbas E., Sauer S. Integrating human-centered and model-driven methods in Agile UI development //Proceedings of 15th IFIP TC. 13 International Conference on Human-Computer Interaction (INTERACT), S. – 2015. – С. 215-221.

[9] Nassiet D. et al. Paprika: Rapid UI development of scientific dataset editors for high performance computing //SDL 2011: Integrating System and Software Modeling: 15th International SDL Forum Toulouse, France, July 5-7, 2011. Revised Papers 15. – Springer Berlin Heidelberg, 2012. – С. 69-78.

[10] Масленников О. П. и др. Разработка системы интерактивного визуального анализа многомерных данных //Научная визуализация. – 2014. – Т. 6. – №. 4. – С. 30-49.

[11] Раздьяконов Д. С. и др. ПОДХОД К ОБРАБОТКЕ ПУСТЫХ УЗЛОВ ПРИ ПОРЦИОННОЙ ВИЗУАЛИЗАЦИИ ДАННЫХ НА ПРИМЕРЕ ИНСТРУМЕНТА ONTODIA //Программирование. – 2020. – №. 6. – С. 16-29.

[12] Razd'yakonov D. S. et al. Approach to Blank Node Processing in Incremental Data Visualization by the Example of Ontodia //Programming and Computer Software. – 2020. – Т. 46. – С. 384-396.

[13] Daniil R. et al. A new tool for linked data visualization and exploration in 3D/VR space //The Semantic Web: ESWC 2019 Satellite Events: ESWC 2019 Satellite Events, Portorož, Slovenia, June 2–6, 2019, Revised Selected Papers 16. – Springer International Publishing, 2019. – С. 167-171.

© Д.С. Раздьяконов 2023