

ГРАФЫ ГРАФЫ ГРАФЫ

Доступное описание графового способа представления знаний. Изложение преимуществ, недостатков, способов применения в настоящем и будущем.

ВВЕДЕНИЕ

Графы — не самая распространённой структуры данных, однако они являются мощным инструментом для представления сложных взаимосвязей.

В повседневных приложениях часто используются другие структуры данных, такие как массивы, списки, деревья и хэш-таблицы, поскольку они проще и быстрее для выполнения типичных операций. Однако графы незаменимы в задачах, где необходимо моделировать взаимосвязи между объектами, например, в социальных сетях, рекомендационных системах, картографических приложениях или биоинформатике.

I. ЧТО ТАКОЕ ГРАФ

Семантический граф — один из способов представления знаний и данных. Графы знаний состоят из двух основных элементов — вершин и рёбер. Как правило, вершины представляют какие-либо объекты или множества, а рёбра — отношения между ними.

II. ПРЕИМУЩЕСТВА ГРАФОВ

Удобство визуализации и восприятия

Одним из главных преимуществ графа знаний как способа представления информации является его одновременная понятность и человеку, и компьютеру. Это отличает его от машинного кода или традиционных энциклопедических текстов, которые в полной мере понятны лишь одной из сторон. Благодаря своей чёткой и упорядоченной структуре граф знаний эффективно решает эту проблему. Он легко интерпретируется компьютером, что позволяет машине понять, как и какие узлы связаны между собой, а визуальное представление графа удобно для восприятия человеком, поскольку явно и наглядно демонстрирует отношения между объектами.

Преимущества расширения графа

Краеугольной особенностью графов является простота их дополнения. Графы позволяют легко добавлять новые узлы и рёбра без необходимости изменения существующей структуры. Это означает, что вы можете расширять граф, вводя новые понятия и связи, не затрагивая уже существующие данные. Можно легко добавлять новые типы отношений между узлами. Например, если у вас есть узел «человек», вы можете добавить разные отношения, такие как «друг»,

«коллега», «родитель» и т.п. Это позволяет гибко моделировать различные аспекты реального мира.

Моделирование семантических знаний

Для начала определим понятие семантической информации, она описывает, что означают данные, а не просто как они представлены.

Семантическая информация всегда находится в контексте. Значение слов или объектов может меняться в зависимости от ситуации. Например, слово «класс» может означать учебную группу в одном контексте и категорию объектов в другом.

Графы являются идеальной структурой для работы с семантическими данными, что делает их особенно полезными в областях, где важно понимать значения и связи между различными объектами.

Возможность неоднократно использовать знания

Основным требованием при создании графов служит следование определённому стандарту формализации. Необходимо, чтобы все знания, представляемые в графах, имели одинаковую структуру и способы представления определённых отношений. Соблюдение таких стандартов гарантирует, что графы будут легко масштабируемы и совместимы с другими системами. Важно отметить, что общие стандарты позволяют без каких либо трудностей добавлять в одну систему, функционал другой, не теряя времени на пересоздание множество раз одних и тех же функций, под каждую систему отдельно.

Однако важно понимать, что для получения максимальной эффективности необходимо создание и строгое соблюдение определенного стандарта формализации.

Такой подход может значительно ускорить разработку любой системы, использующей графовую модель представления знаний.

Выявление неявных связей

Графы позволяют выявлять и использовать неявные связи между элементами, что особенно полезно в анализе данных. Эти неявные связи могут быть неочевидны при линейном анализе, но становятся явными при использовании графового подхода. Например, два пользователя соцсети могут не быть друзьями, но иметь общих знакомых, что может свидетельствовать о возможности взаимодействия между ними.

III. ИСПОЛЬЗОВАНИЕ ГРАФОВ В НАСТОЯЩЕЕ ВРЕМЯ

Поисковые системы

Google начал активно использовать графы с 2012 года, когда был запущен Google Knowledge Graph. Технология используется для улучшения результатов поиска. Граф знаний помогает связывать объекты, такие как люди, места и события, чтобы дать более точные и полезные результаты.

Графы активно используются не только в Google, но и в других крупных поисковых системах, таких как Яндекс, Microsoft Bing, Yahoo.

До внедрения графа знаний поисковые запросы обрабатывались на уровне ключевых слов, что могло давать неоднозначные или недостаточно релевантные результаты. Графы позволили системам «понимать» семантические связи между объектами. Например, если пользователь ищет «Данила Козловский», система понимает, что это актёр, и предлагает информацию не только о его фильмах, но и о людях и событиях, с которыми он связан.

Транспортные системы

На данный момент все современные системы навигации используют графы для представления транспортных маршрутов, так как графовая структура идеально подходит для моделирования транспортных сетей: автомобильных дорог и железнодорожных путей, а также для поиска кратчайшего маршрута и выявления пробок.

Обработка естественного языка

Использование графов в обработке естественного языка позволяет моделировать сложные связи и структуры. Их применение в семантическом анализе, синтаксическом разборе и извлечении информации способствует созданию современных интеллектуальных систем, которые могут более точно и эффективно взаимодействовать с человеческим языком. Также их применение повышает эффективность алгоритмов рекомендаций и понимания контекста запросов в поисковых системах. Это позволяет лучше интерпретировать намерения пользователей и давать более подходящие результаты.

IV. ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ГРАФОВ

Графовые базы данных

Графовые базы данных представляют собой системы управления данными, которые оптимизированы для хранения и обработки графовых структур. В последние годы интерес к данным системам возрос благодаря их способности эффективно работать со сложными взаимосвязями между данными.

Традиционные базы данных не всегда справляются с задачами, в которых важно учитывать связь между объектами, поскольку их структура не предназначена для быстрых и гибких запросов по отношениям. Такие графовые базы данных, как *Neo4j*, *Amazon Neptune* и *OrientDB* предоставляют более производительный способ хранения и использования связанный информации.

Разработка интеллектуальных систем

С ростом популярности сферы искусственного интеллекта возникает необходимость разрабатывать более сложные и универсальные интеллектуальные системы, причём важно создавать их со способностью глубокой обработки информации на уровне семантики, поддержки решения сложных задач и возможностью гибкого масштабирования. Технология OSTIS благодаря SC-коду — универсальному графовому языку представления знаний — позволяет одновременно и использовать преимущества графов, и устранить недостатки разработки интеллектуальных систем.

V. ВЫВОД

Бам бам бам мы стреляем по тупоголовым чиновникам

Гетман Данила Русланович, студент кафедры интеллектуальных информационных технологий БГУИР, danilahetman@gmail.com.

Титов Артём Вадимович, студент кафедры интеллектуальных информационных технологий БГУИР, ***почта***

Научный руководитель: Зотов Никита Владимирович, стажёр младшего научного сотрудника НИЛ 3.7, ассистент кафедры интеллектуальных информационных технологий, n.zotov@bsuir.by.