

Дополненная логикой генерация

Дата: 2025-01-14 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2411.14012>

Рейтинг: 62

Адаптивность: 75

Ключевые выводы:

Исследование представляет концепцию Logic Augmented Generation (LAG) - новую парадигму, объединяющую семантические графы знаний (SKG) с большими языковыми моделями (LLM). Основная цель - преодолеть ограничения обоих подходов, используя LLM как реактивные непрерывные графы знаний (RCKG), способные генерировать потенциально бесконечные отношения и неявные знания по запросу, при этом используя SKG для обеспечения логической согласованности и фактических границ.

Объяснение метода:

Исследование представляет ценную концепцию интеграции структурированных знаний и генеративных возможностей LLM. Хотя полная реализация LAG технически сложна для обычных пользователей, основные принципы структурирования запросов и извлечения неявных знаний могут быть адаптированы для повседневного использования через продуманные промпты, что существенно улучшит качество взаимодействия с LLM.

Ключевые аспекты исследования: 1. **Logic Augmented Generation (LAG)** - новая парадигма, объединяющая преимущества семантических графов знаний (SKG) и языковых моделей (LLM), где LLM используются как реактивные непрерывные графы знаний (RCKG), а SKG обеспечивают логические границы и фактическую основу.

Reactive Continuous Knowledge Graphs (RCKG) - концепция использования LLM для динамического создания знаний по запросу, что позволяет генерировать потенциально бесконечные связи и неявные знания, адаптированные к конкретному контексту.

Трехэтапный процесс преобразования - от мультимодальных сигналов к естественному языку (супрамодалное преобразование), затем к структурированному графу знаний (амодальное преобразование) и, наконец, расширение графа неявными знаниями.

Интеграция явных и неявных знаний - LAG извлекает неявные (tacit) знания с помощью LLM и структурирует их в соответствии с логической моделью SKG, что

особенно ценно в задачах коллективного интеллекта, таких как медицинская диагностика.

Практическое применение в медицине и климатологии - демонстрация работы LAG на примере медицинской диагностики, где система связывает симптомы с возможными причинами на основе как явной информации, так и неявных знаний.

Дополнение: Для полной реализации LAG, как описано в исследовании, действительно требуется доступ к API и потенциально дообучение моделей. Однако ключевые концепции и подходы можно адаптировать для использования в стандартном чате без этих технических требований.

Вот концепции, которые можно применить в стандартном чате:

Структурированные промпты с фактическими границами: Пользователи могут включать в свои запросы структурированную фактическую информацию, тем самым ограничивая "пространство генерации" LLM. Например: "Основываясь на следующих фактах: [список фактов], сделай вывод о возможных причинах [проблемы]".

Явное запрашивание неявных связей: Можно попросить LLM выявить неявные связи между фактами: "Учитывая следующую информацию, какие неявные связи или причинно-следственные отношения могут существовать между этими фактами?"

Поэтапное структурирование знаний: Пользователи могут запрашивать структурированное представление информации в виде триплетов "субъект-предикат-объект" или JSON, что соответствует подходу RCKG: "Представь эту информацию в виде структурированных отношений: [информация]".

Интеграция с существующими знаниями: Можно попросить LLM объединить новую информацию с уже известными фактами: "Объедини эту новую информацию с ранее известными фактами и представь целостную картину".

Контекстное расширение знаний: Пользователи могут запрашивать расширение базовых фактов с учетом контекста: "Расширь эти базовые факты, добавив контекстуально релевантную информацию для [конкретной задачи]".

Результаты применения этих подходов: - Повышение точности и надежности ответов LLM за счет четких фактических границ - Более структурированное и логически последовательное представление информации - Выявление неявных связей и отношений, которые не очевидны из исходных данных - Лучшее понимание контекста и более релевантные ответы

Хотя такой подход и не будет обладать всей мощностью полной реализации LAG с использованием специализированных графов знаний и API, он позволит значительно улучшить качество взаимодействия с LLM в стандартном чате, применяя основные принципы исследования.

Анализ практической применимости: 1. **Logic Augmented Generation (LAG):** -

Прямая применимость: Средняя. Обычные пользователи не смогут самостоятельно реализовать полную архитектуру LAG, но могут использовать принцип комбинирования фактических знаний с генеративными возможностями LLM. - Концептуальная ценность: Высокая. Понимание важности баланса между фактической точностью и генеративной гибкостью помогает пользователям формулировать более эффективные запросы к LLM. - Потенциал для адаптации: Высокий. Принцип "ограничения" генеративных возможностей LLM фактической информацией может быть применен через промпты, включающие структурированные знания.

Reactive Continuous Knowledge Graphs (RCKG): Прямая применимость: Низкая для реализации, средняя для концепции. Пользователи не смогут создать полноценные RCKG, но могут использовать подход запроса к LLM для получения структурированной информации. Концептуальная ценность: Высокая. Понимание LLM как источника динамически генерируемых знаний помогает пользователям лучше использовать их возможности. Потенциал для адаптации: Средний. Пользователи могут запрашивать у LLM генерацию структурированных данных (например, в формате JSON или трехчастных утверждений).

Трехэтапный процесс преобразования:

Прямая применимость: Низкая. Технический процесс, требующий специализированных знаний. Концептуальная ценность: Средняя. Понимание того, как происходит преобразование информации, может помочь в построении более эффективных запросов. Потенциал для адаптации: Низкий. Слишком технический для адаптации обычными пользователями.

Интеграция явных и неявных знаний:

Прямая применимость: Средняя. Пользователи могут структурировать запросы, включая как фактическую информацию, так и просьбу о выводе неявных связей. Концептуальная ценность: Высокая. Понимание различия между явными и неявными знаниями помогает формулировать более эффективные запросы. Потенциал для адаптации: Высокий. Пользователи могут явно запрашивать LLM о выявлении неявных связей между фактами.

Практическое применение:

Прямая применимость: Средняя. Пример медицинской диагностики демонстрирует конкретный способ использования комбинации структурированных и неструктурированных данных. Концептуальная ценность: Высокая. Показывает, как LLM могут дополнять фактические данные неявными знаниями. Потенциал для адаптации: Высокий. Подход может быть адаптирован для различных областей, где требуется сочетание фактов и выводов.

Prompt:

Использование концепции LAG в промптах для GPT ## Что такое LAG

Logic Augmented Generation (LAG) объединяет семантические графы знаний (SKG) с языковыми моделями (LLM), позволяя:

Использовать структурированные знания для логической согласованности
Извлекать неявные (tacit) знания с помощью LLM
Создавать причинно-следственные связи, не указанные явно в данных ## Пример промпта с применением LAG

[=====] Я хочу, чтобы ты действовал как медицинский эксперт, используя подход Logic Augmented Generation. Вот начальная информация о пациенте:

ИСХОДНЫЕ ДАННЫЕ: - Мужчина, 45 лет - Симптомы: лихорадка, головная боль, мышечные боли - Недавно вернулся из деловой поездки в Юго-Восточную Азию

СЕМАНТИЧЕСКИЙ ГРАФ: пациент -- имеет симптом --> лихорадка пациент -- имеет симптом --> головная боль пациент -- имеет симптом --> мышечные боли пациент -- совершил --> деловая поездка в Юго-Восточную Азию

Пожалуйста: 1. Расширь этот семантический граф, добавляя возможные причинно-следственные связи между симптомами и поездкой 2. Предложи 3 возможных диагноза, основываясь на расширенном графе 3. Для каждого диагноза укажи, какие неявные знания ты использовал для его формирования

При ответе сначала визуализируй расширенный граф знаний, затем опиши логику рассуждений. [=====]

Как работает этот промпт

Структурирование входных данных: Промпт содержит как неструктурированное описание, так и структурированный семантический граф, задающий логические ограничения.

Запрос на расширение графа: Мы просим GPT действовать как RCKG (реактивный непрерывный граф знаний), генерируя дополнительные связи между существующими узлами.

Извлечение неявных знаний: GPT должен использовать свои знания о географии, эпидемиологии и медицине для выявления потенциальных связей между поездкой и симптомами.

Логическое обоснование: Требование объяснить рассуждения заставляет модель следовать логическим ограничениям графа.

Этот подход помогает получить более структурированные, логически согласованные и фактически обоснованные ответы, сочетая преимущества графов знаний (логическая структура) и языковых моделей (извлечение неявных знаний).