Анализ Книги Роджера Фишера и Уилльяма Юри. Путь к согласию, или Переговоры без поражения

**Постановка задачи**

Цель эксперимента — проверить, как разные методы поиска могут находить информацию в книге и отвечать на заранее сформулированные вопросы.

Для проверки заранее были найдены ответы в книге, чтобы сравнивать качество выдачи разных алгоритмов.

**Вопросы для тестирования**

1. Кто автор книги?
2. В чем основная цель книги?
3. Какие четыре принципа переговоров выделяют авторы?
4. Что такое BATNA?
5. Чем отличается позиционный торг от переговоров на основе интересов?
6. Как бороться с нечестными тактиками в переговорах?
7. Что значит разделять людей и проблему?

**Методы поиска, использованные в эксперименте**

1. **MEMOS**

Векторный поиск с «операционной памятью» (Memory OS). MEMOS сам управляет хранением и поиском фрагментов текста (чанков), поддерживает семантический поиск и возвращает не только сырой контент, но и связные ответы. Хорош для задач, где нужен именно «агент», а не просто поиск.

1. **BM25**

Классический алгоритм информационного поиска, основанный на подсчете частоты встречаемости слов и их важности в документе. Сильный на точных совпадениях ключевых слов, но плохо справляется с переформулировками и синонимами.

1. **FAISS (с OpenAI embeddings)**

Векторный поиск на базе библиотеки FAISS от Facebook. Тексты и запросы переводятся в эмбеддинги (числовые векторы) с помощью модели OpenAI text-embedding-3-small. Работает лучше, чем BM25, так как учитывает семантическое сходство, а не только слова.

1. **SentenceTransformers**

Локальная модель эмбеддингов (all-MiniLM-L6-v2), которая также кодирует тексты и запросы в векторы. Это альтернатива FAISS с OpenAI, но полностью оффлайн и без затрат на API. Качество обычно чуть ниже, чем у свежих моделей OpenAI, но зато бесплатно и быстро.

1. **Combined (MEMOS + FAISS)**

Гибридный метод. Сначала ответы формируются MEMOS и FAISS, затем используется LLM (gpt-4o-mini) для реранкинга — сравнения и выбора лучшего финального ответа. Такой подход сочетает объяснительную силу MEMOS и точность FAISS.

**Ход работы**

* Загружался текст книги, разбитый на фрагменты (чанки).
* Для каждого метода строился индекс или модель.
* Задавались заранее сформулированные вопросы.
* Каждый метод возвращал свои ответы.
* Ответы сравнивались с эталонными по косинусной близости эмбеддингов.
* Для каждого вопроса фиксировались результаты в таблице results.csv и визуализировались на графике results.png.

**Артефакты**

* results.csv — таблица с вопросами, эталонными ответами, ответами всех методов, оценкой близости и указанием победителя.
* results.png — график, показывающий сравнение методов по всем вопросам.

**Основные выводы**

* MEMOS формирует более связные объяснения, но иногда уходит в сторону.
* FAISS выдает более точные по фактам ответы, но менее развернутые.
* BM25 ограничен и часто возвращает не тот контекст.
* SentenceTransformers работает стабильно, но качество ниже, чем у FAISS.
* Лучшие результаты показал комбинированный метод MEMOS + FAISS, где реранкер выбирал наиболее точный и полный ответ.

**Гипотезы для улучшения**

* Проверить гибридный поиск BM25 + FAISS, чтобы добавить точность к семантическому поиску.
* Добавить ещё один уровень проверки качества ответа (например, LLM, ориентированный только на факты).
* Протестировать подход на книгах из других доменов (продажи, дейтинг), чтобы проверить универсальность метода.

