

# MultiAgentBench: Комплексный анализ сотрудничества и конкуренции многоагентных систем

## Основные принципы и методология

**MultiAgentBench** — это комплексный бенчмарк для оценки систем на основе LLM, работающих в многоагентном режиме. Исследование фокусируется не только на выполнении задач, но и на качестве сотрудничества и конкуренции между агентами.

### Ключевые компоненты исследования:

#### 1. Разнообразие сценариев

- Включает шесть различных интерактивных сценариев: от исследовательских задач до кодирования, игр и переговоров
- Охватывает как кооперативные, так и конкурентные модели взаимодействия

#### 2. Топологии координации

- Исследует четыре структуры взаимодействия между агентами:
  - Звезда (центральный координатор)
  - Цепочка (последовательная передача информации)
  - Дерево (иерархическая структура)
  - Граф (все агенты могут напрямую взаимодействовать)
- Выявлено, что графовая структура показывает наилучшие результаты в исследовательских сценариях

#### 3. Стратегии планирования

- Когнитивное самоэволюционирующее планирование, при котором агенты сравнивают ожидаемые и фактические результаты
- Улучшает достижение ключевых этапов на 3%

#### 4. Системы оценки

- Новые метрики для измерения качества координации и коммуникации между агентами
- KPI на основе достижения конкретных этапов

# Практические примеры применения

## 1. Многоагентное исследование рынка

### Структура и роли:

Ты будешь действовать как система из 3 агентов с графовой структурой координации:

1. Аналитик данных - собирает и анализирует информацию о рынке
2. Маркетолог - интерпретирует данные с точки зрения потребительского поведения
3. Стратег - формулирует итоговые рекомендации

Метод координации:

- Используй графовую структуру взаимодействия, где каждый агент может напрямую общаться с любым другим
- Применяй когнитивное самоэволюционирующее планирование: в начале работы каждого агента формулируй ожидаемый результат, а после выполнения сравнивай его с фактическим
- Ограничь обсуждение 5-7 итерациями для оптимальной эффективности

Задача:

Исследуй рынок электросамокатов в городской среде и разработай стратегию вывода нового продукта.

## 2. Решение исследовательской задачи с многоагентным подходом

В этом сценарии несколько агентов с разной специализацией работают над общей проблемой, каждый применяя свои уникальные компетенции. Например, один агент может заниматься поиском литературы, другой — анализом данных, третий — синтезом результатов.

## Почему это работает

### 1. Разделение сложных задач

- Распределение когнитивной нагрузки позволяет каждому агенту сосредоточиться на своей области
- Преодоление ограничений контекста — каждый агент работает со своим сегментом информации

### 2. Оптимальная координация

- Графовая структура обеспечивает прямую коммуникацию, что особенно эффективно в исследовательских задачах

- Ограничение в 5-7 итераций предотвращает "зацикливание" и снижение эффективности при слишком длительных обсуждениях

### **3. Самокорректирующееся планирование**

- Когнитивное планирование позволяет агентам корректировать свою работу через сравнение ожидаемых и фактических результатов
- Это обеспечивает постепенное улучшение качества результатов

### **4. Специализация и синергия**

- Разные "личности" агентов обеспечивают разнообразие мышления и подходов
- Команды с разными личностными характеристиками могут работать эффективнее на определенных типах задач

## **Практическое применение**

Важно отметить, что этот многоролевой подход можно реализовать в стандартном чате через правильное структурирование промптов, без необходимости дополнительного API или дообучения моделей. Пользователь может определить:

- Роли агентов и их специализацию
- Структуру взаимодействия между ними
- Процесс планирования и самопроверки
- Количество итераций для обсуждения

Такой подход позволяет повысить качество решения сложных задач, получить более структурированные результаты и эффективнее использовать возможности языковых моделей даже в рамках стандартного интерфейса.