

## УПРАВЛЕНИЕ ЭСКАЛАЦИЯМИ В LLM-АГЕНТНОЙ РАЗРАБОТКЕ: ОПТИМИЗАЦИЯ КАЧЕСТВА, СТОИМОСТИ И ВРЕМЕНИ

Басилаев Д. И.<sup>1</sup>

Научный руководитель – Ковальчук М. И.<sup>1</sup>

<sup>1</sup>Университет ИТМО

### Введение

LLM-агенты применяются для автоматизированной разработки программного обеспечения в итеративном цикле «генерация - тестирование - исправление» [1, 2]. Несмотря на рост качества моделей, такой процесс сопровождается значительными затратами: увеличивается число вызовов моделей, растёт время прогонов CI, в сложных случаях требуется участие человека [3]. Эффективность системы во многом определяется выбором следующего шага после получения результатов проверки.

### Основная часть

Процесс разработки рассматривается как последовательность решений в цикле «изменение - проверка - выбор действия», где состояние определяется результатами тестов, динамикой исправления ошибок и накопленными затратами. Возможные действия включают смену модели, подключение инструмента, откат или привлечение человека. Цель управления состоит в получении корректного результата при минимальных затратах и времени выполнения. В отличие от фиксированных сценариев и пороговых эвристик предлагается обучаемый выбор следующего шага на основе наблюдаемых сигналов качества и стоимости. Оценка проводится на тест-ориентированных наборах задач, включая SWE-bench [4], с использованием метрик доли решённых задач, суммарных затрат и времени выполнения.

### Выводы

Управление эскалациями существенно влияет на эффективность LLM-агентной разработки. Формализация процесса как последовательности решений позволяет перейти от фиксированных правил к адаптивному выбору действий. Ожидается, что обучаемый подход обеспечит снижение затрат и времени выполнения без ухудшения качества решения задач.

### Литература

1. Qian C., Liu W., Liu H., Chen N., Dang Y., Li J. et al. ChatDev: Communicative Agents for Software Development // arXiv preprint. – 2023.
2. Hong S., Zhuge M., Chen J., Zheng X., Cheng Y., Zhang C. et al. MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework // arXiv preprint. – 2023.
3. Bouzenia I. et al. RepairAgent: Autonomous Program Repair with LLM Agents // arXiv preprint. – 2025.
4. Jimenez C.E., Yang J., Wettig A., Yao S., Pei K., Press O., Narasimhan K. SWE-bench: Can Language Models Resolve Real-World GitHub Issues? // arXiv preprint. – 2023.