Совместное обучение высокоуровневого планировщика и низкоуровневых стратегий

Студент: Бухтуев Григорий
Руководитель: Ковалёв Алексей

Центр когнитивного моделирования МФТИ

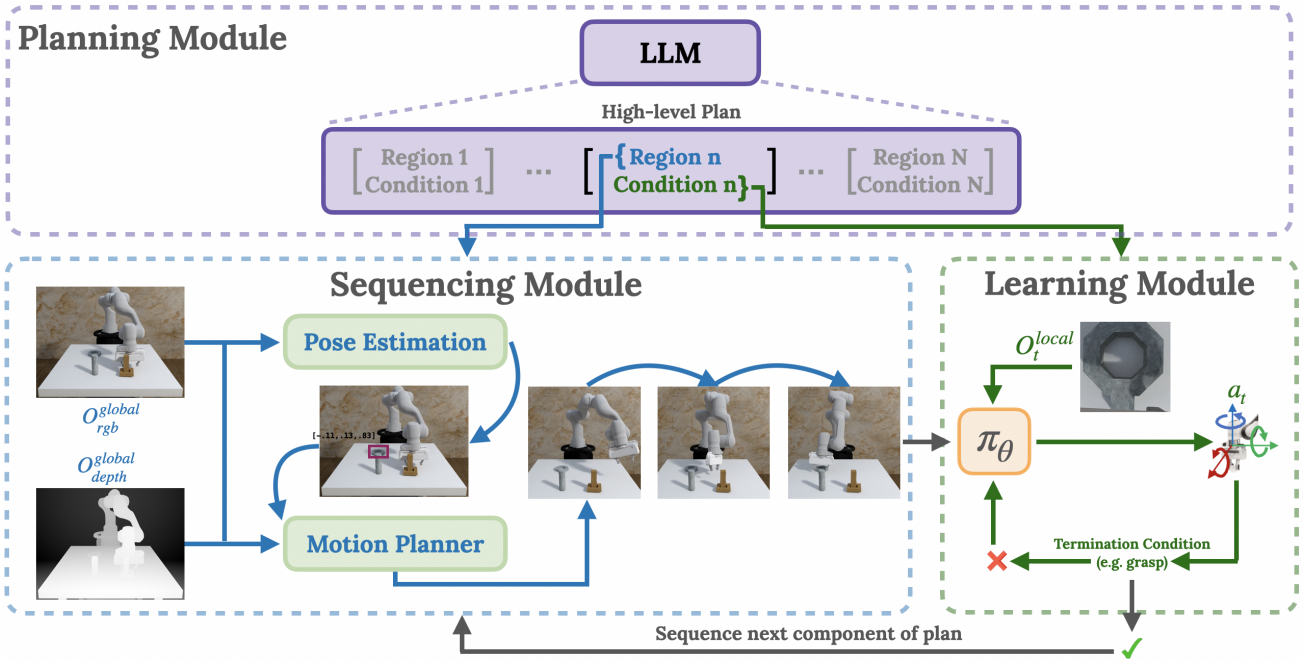
16 декабря – Обзор смежных работ





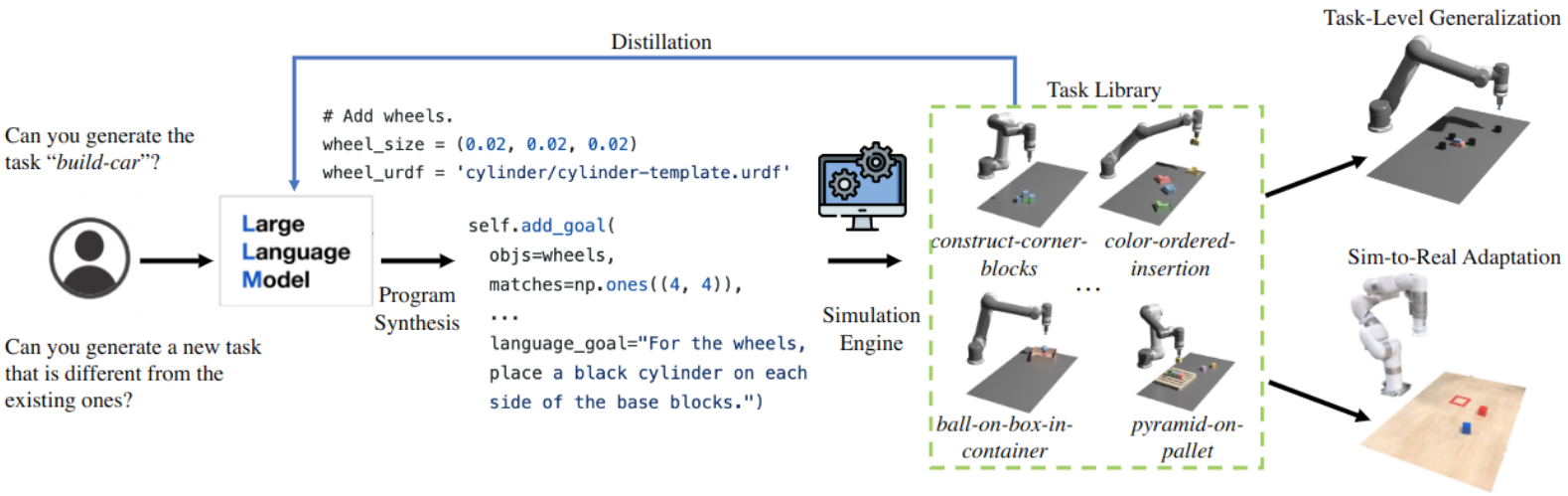
Постановка задачи

В последний год появились исследования, которые позволяют решать задачи с длительным горизонтом планирования в режиме онлайн. Это достигается за счет поочередного выполнения команд от высокоуровневого планировщика и мелких операций с применением низкоуровневых стратегий. Такой подход стал возможен благодаря использованию современных LLM, обладающих обширными знаниями об окружающей среде – PLAN-SEQ-LEARN



Постановка задачи

Параллельно с этим была разработана методика, позволяющая в реальном времени генерировать новые задачи для мобильных роботов в произвольной среде с помощью тех же LLM. Эта методика демонстрирует значительный прирост качества выполнения задач с коротким горизонтом планирования и хорошую обобщаемость на задачи, состоящие из нескольких последовательных действий – GenSim



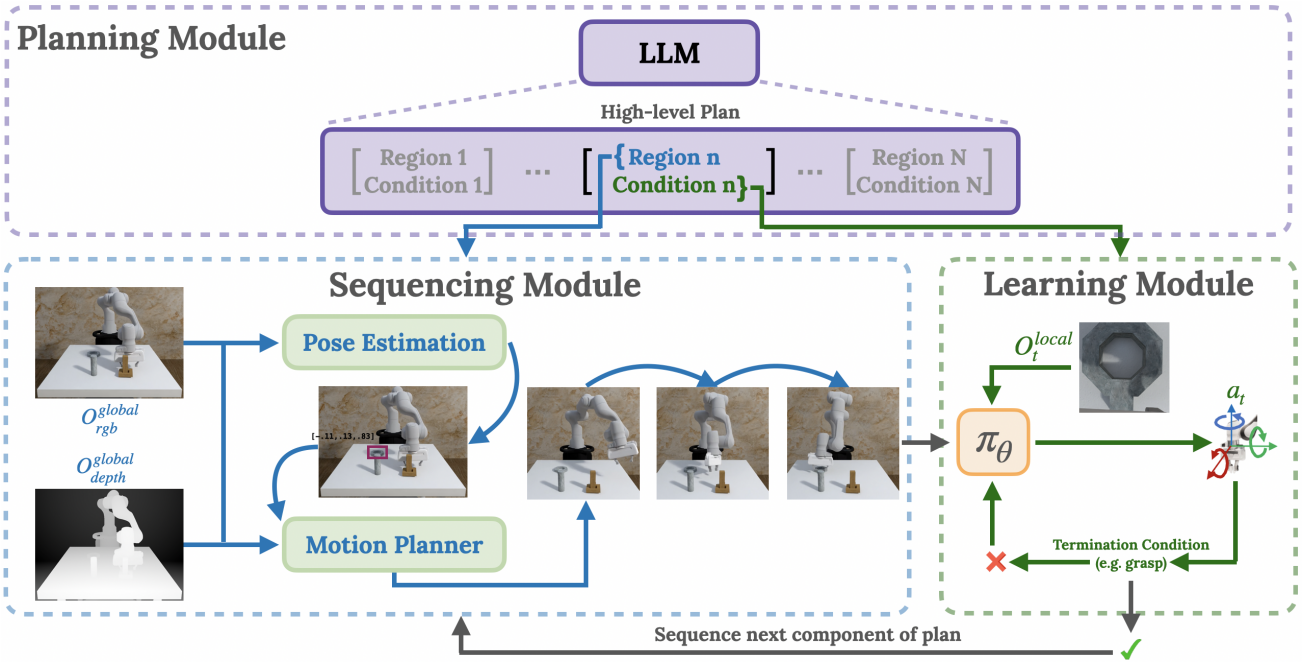
Цель работы

Целью данной работы является объединение этих двух подходов. Мы стремимся обучать низкоуровневые стратегии не только на заранее определенных задачах из заданной среды, но и на задачах, которые будут генерироваться в процессе обучения с помощью LLM.

Обзор статей

PLAN-SEQ-LEARN: (2024) Language Model Guided RL for Solving Long Horizon Robotics Tasks

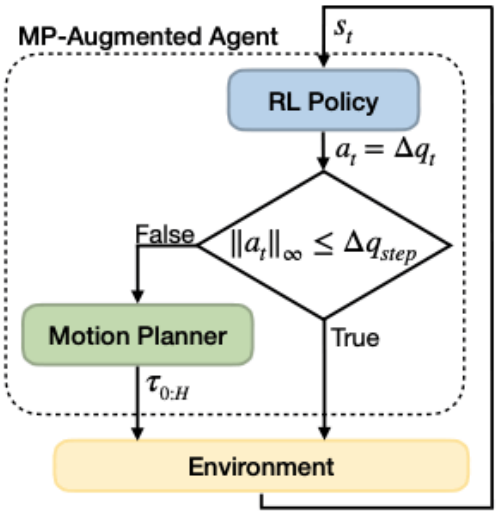
- Планирование (Plan): Генерация высокоуровневого плана с использованием LLM
- (Seq): Использование планирования движений для достижения целевых поз
- Обучение RL-политики для выполнения низкоуровневых действий



Обзор статей

MoPA-RL: (2020) Motion Planner Augmented Reinforcement Learning for Robot Manipulation in Obstructed Environments

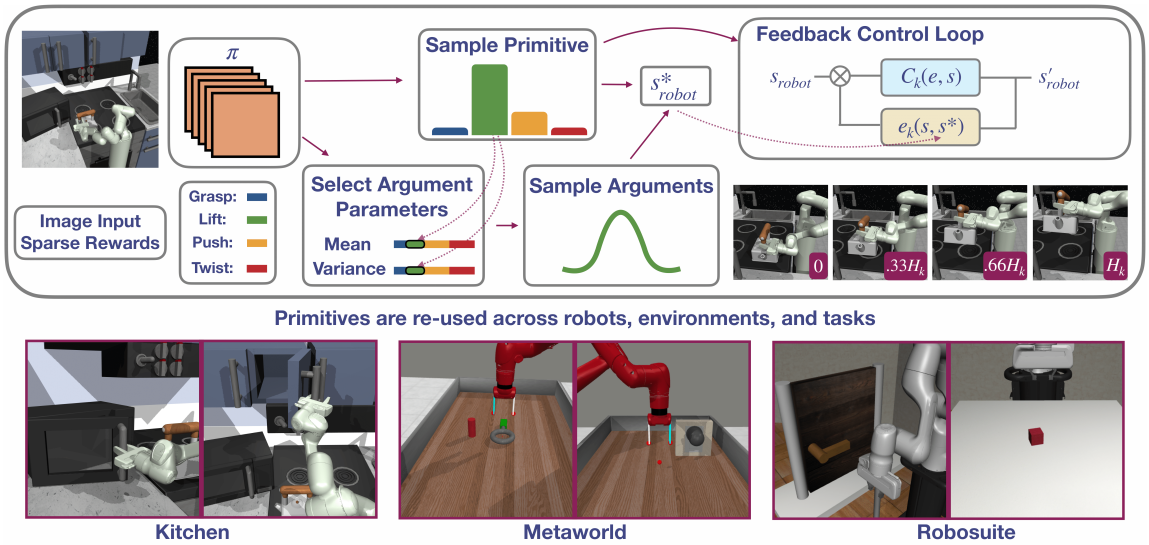
- Планирование: Агент использует планирование движений для нахождения безопасных путей к целям.
- Обучение: RL-агент обучается выполнять действия, используя как MP, так и прямое выполнение действий в зависимости от величины действия.



Обзор статей

RAPS: (2021) Accelerating Robotic Reinforcement Learning via Parameterized Action Primitives

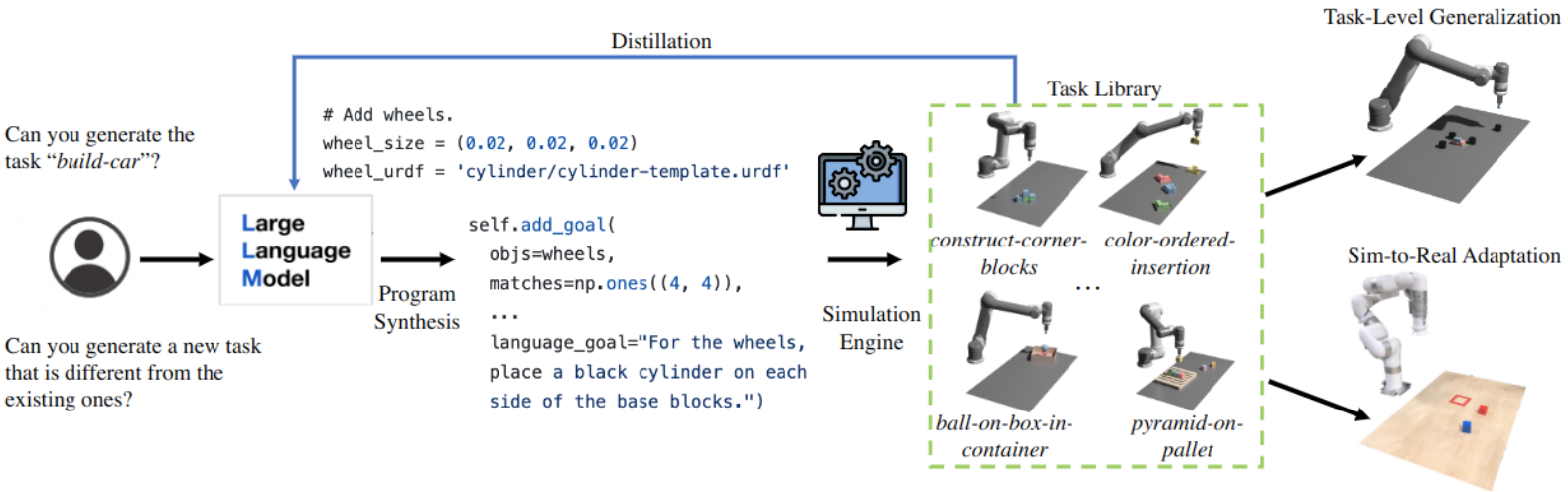
- Примитивы действий: Каждый примитив представляет собой заранее определенное действие, параметризованное аргументами, которые агент обучается выбирать
- RL-агент использует выбранные примитивы для выполнения действий, что позволяет ему сосредоточиться на том, что нужно сделать, а не на том, как это выполнить



Обзор статей

GenSim: (2024) Generating Robotic Simulation Tasks via Large Language Models

- Целевая генерация — когда LLM получает конкретную задачу и предлагает учебный план для ее решения.
- Исследовательская генерация — когда LLM использует предыдущие задачи для итеративного создания новых задач.



Текущий статус

Center for Cognitive Modeling

- Изучены различные статьи по совместному обучению высокоуровневого планировщика и низкоуровневых стратегий: [1], [2], [3]
- Успешно выполнена настройка окружения для запуска архитектур PLAN-SEQ-LEARN и GenSim на кластере
- Начато изменение архитектуры GenSim для ее применения в средах, на которых обучалась PLAN-SEQ-LEARN

Спасибо за внимание!

bukhtuev.ga@phystech.edu