

Learning to Model the World with Language

Косса Николай Евгеньевич,
Студент группы БПМИ202

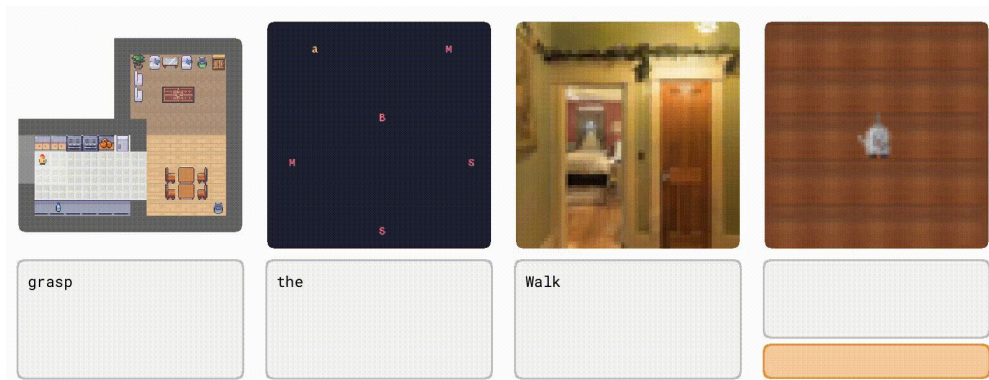
О чем эта статья?

- **Dynalang**
- **Основные идеи**
- **Существующие методы**
- **Постановка задачи**
- **Архитектура**
- **Эксперименты**
- **Результаты**

Dynalang

и какие задачи он решает

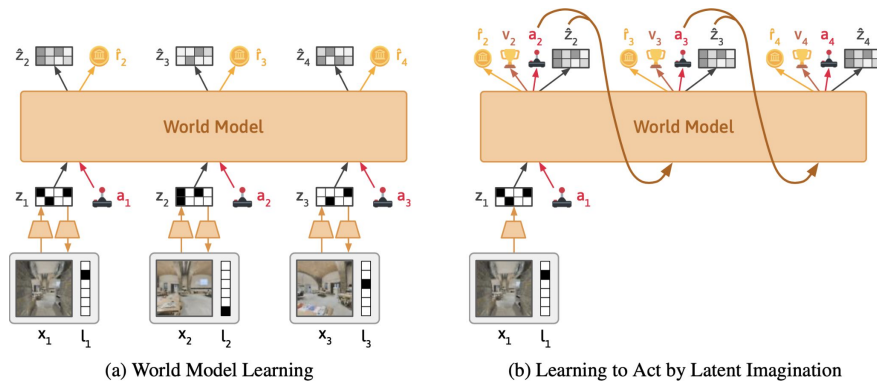
Dynalang представляет собой модель, направленную на взаимодействие с людьми и реальным миром, используя язык.



Основная задача, которую решает Dynalang — это ограничения нынешних моделей, которые могут следовать простым инструкциям, но испытывают трудности с различными способами использования языка, включая передачу знаний, предоставление ситуационной информации и координацию действий.

ОСНОВНЫЕ ИДЕИ

- Dynalang использует новый подход, предлагая агентам использовать язык для прогнозирования будущего.
- Прогнозируя будущие представления, Dynalang создает для агентов богатый обучающий сигнал для понимания языка и его связи с окружающей средой.



отличие от существующих методов

- Обучается использованию более широких типов языка, таких как описания того, как устроен мир.
- Вместо непосредственного изучения политик, обусловленных языком, Dynalang изучает модель мира.
- В отличие от архитектур, основанных на LLM, Dynalang является мультимодальным, основывая свои предсказания как на языке, так и на изображениях.

постановка задачи

- Среда при взаимодействии с агентом реагирует вознаграждением, флагом, указывающим, продолжается ли эпизод, и наблюдениями.
- Наблюдение состоит из изображения и языкового токена на каждом временном шаге.
- Цель агента состоит в том, чтобы выбрать действия, которые максимизируют ожидаемую сумму вознаграждений со скидкой.

архитектура

мультимодальное выравнивание

агент получает один видеокадр и один языковой токен на каждом временном шаге, и производит одно двигательное действие и один языковой токен на каждом временном шаге.

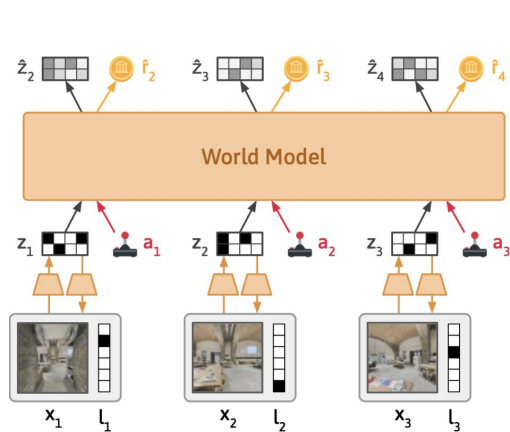
модель мира

служит множеству целей, обеспечивая богатый обучающий сигнал для закрепления языка в визуальном опыте.

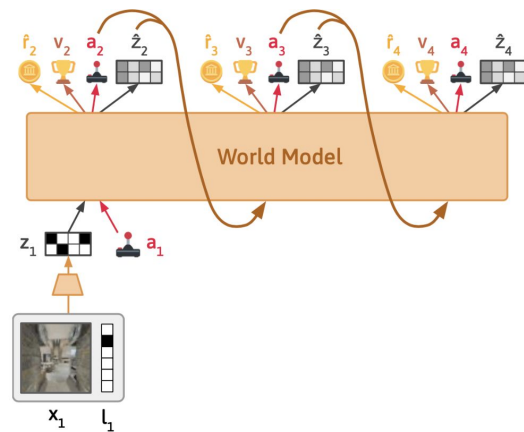
модель принятия решений

использует богатые мультимодальные представления, полученные с помощью модели мира, которые содержат сжатую информацию как о визуальных входных данных, так и о языке, чтобы предсказать распределение действий.

архитектура



(a) World Model Learning



(b) Learning to Act by Latent Imagination

эксперименты

HomrGrid

это многозадачная среда, где агенты получают спецификации задач на языке, а также языковые подсказки.

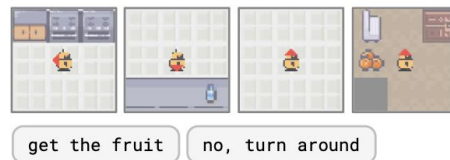
Виды подсказок:

- **Будущие наблюдения** о том, где находятся объекты в мире или куда они были перемещены.
- **Динамические описания** правильного действия достижения результата.
- **Исправления** сообщают агенту “нет, развернись”, когда расстояние до текущего объекта цели увеличивается.

Future Observations



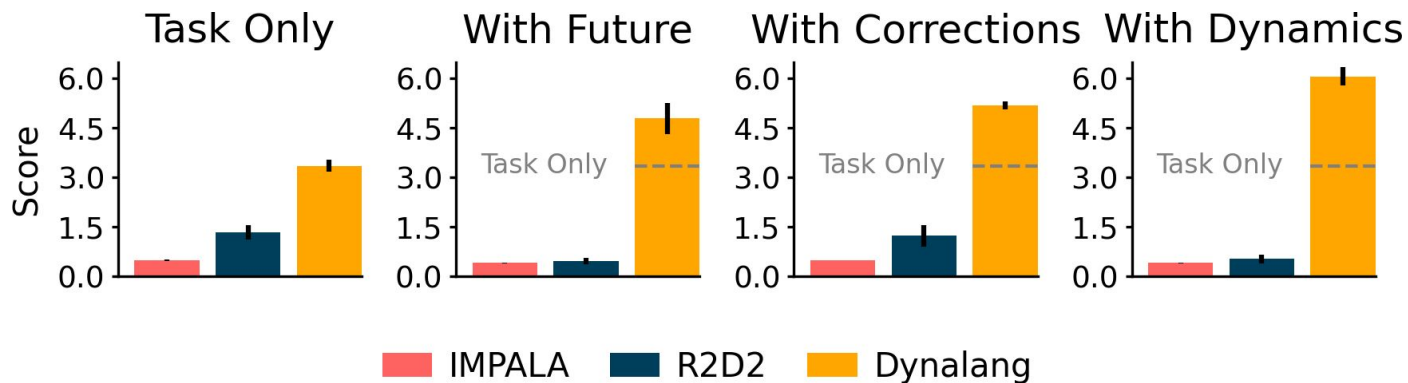
Corrections



Dynamics



ЭКСПЕРИМЕНТЫ



Качество для HomeGrid. Dynalang учится использовать все типы языковых подсказок для достижения более высоких, превосходя IMPALA и R2D2, где мы видим снижение производительности при включении языковых подсказок.

ЭКСПЕРИМЕНТЫ

Messenger

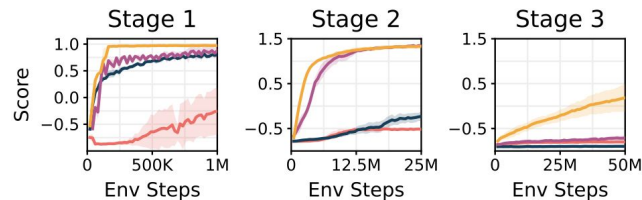
Агент должен получить сообщение от одного из объектов в среде и доставить его другому объекту, избегая при этом врагов.

Habitat

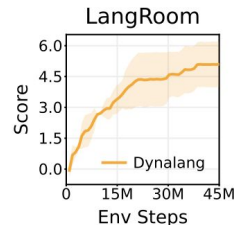
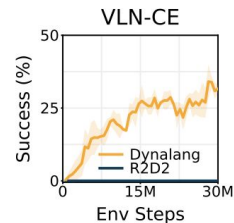
Агенты должны перемещаться по 3D-панорамам, снятым в реальных домах, следуя инструкциям на естественном языке, которые указывают, куда агенту следует идти.

LangRoom

Среда содержит комнату с объектами с фиксированным расположением, но рандомизированными цветами. Модель должна передвигаться по миру, чтобы отвечать на вопросы о нем.



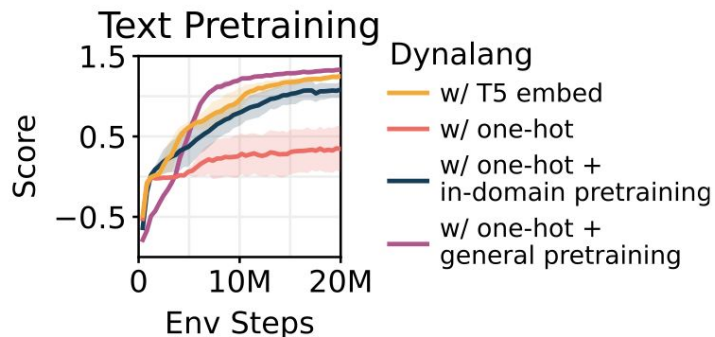
— Dynalang (ours) — EMMA — R2D2 — IMPALA



эксперименты

Text-only предобучение

Dynalang может быть предварительно обучен на одномодальных данных путем обнуления других входных данных о модальности и действиях. Это позволяет агентам RL извлекать выгоду из крупномасштабных автономных данных в рамках единой архитектуры.



Предобучение на Tiny Stories превосходит стандартное обучение, так как позволяет модели изучать динамику текста в автономном режиме, а не во время взаимодействия с окружающей средой.

результаты

Ограничения

- Предложенная архитектура может показать плохие результаты в задачах с длинным горизонтом планирования.
- Качество генерации уступает современным языковым моделям.

Заключение

- Dynalang сопоставляет язык и визуальные данные, действиям и вознаграждениям посредством прогнозирования будущего.
- Dynalang учится действовать на основе различных языковых данных в широком спектре задач, часто превосходя существующие методы, которые испытывают трудности с повышенной сложностью языка.
- Возможность предобучения на видео и тексте предполагает, что Dynalang можно масштабировать до больших наборов данных.

красивые гифки

