
VOYAGER: An Open-Ended Embodied Agent with Large Language Models

**Guanzhi Wang^{1 2}, Yuqi Xie³, Yunfan Jiang^{4*}, Ajay Mandlekar^{1*},
Chaowei Xiao^{1 5}, Yuke Zhu^{1 3}, Linxi Fan^{1†}, Anima Anandkumar^{1 2†}**

¹NVIDIA, ²Caltech, ³UT Austin, ⁴Stanford, ⁵ASU

*Equal contribution †Equal advising

<https://voyager.minedojo.org>

Interpreted by:

John Tan Chong Min

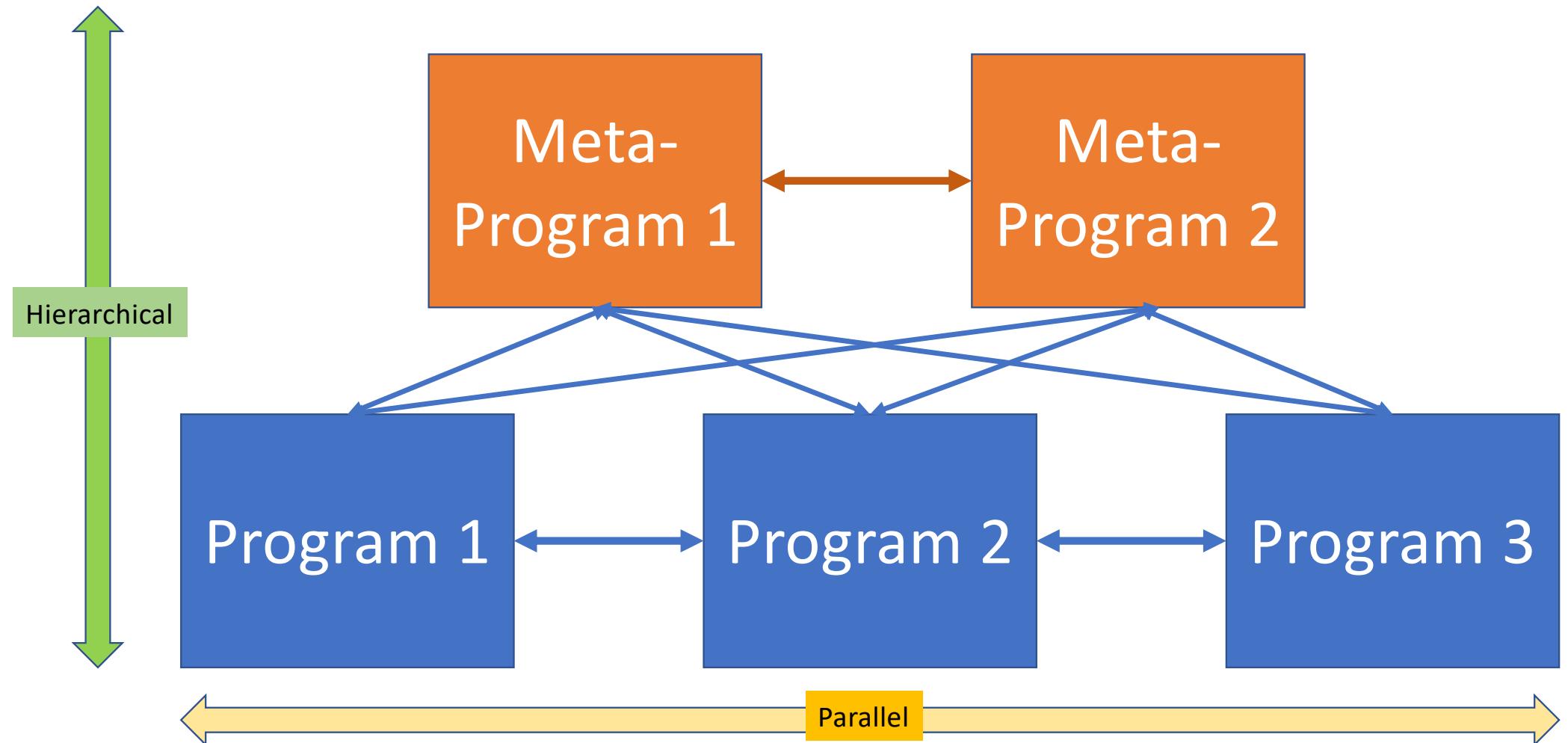
A New Way of doing RL

- Use LLM to decide actions (here given as skills via code) based on context
- Build upon a framework based on LLMOps which require just external database tuning without changing model weights

LLMOps

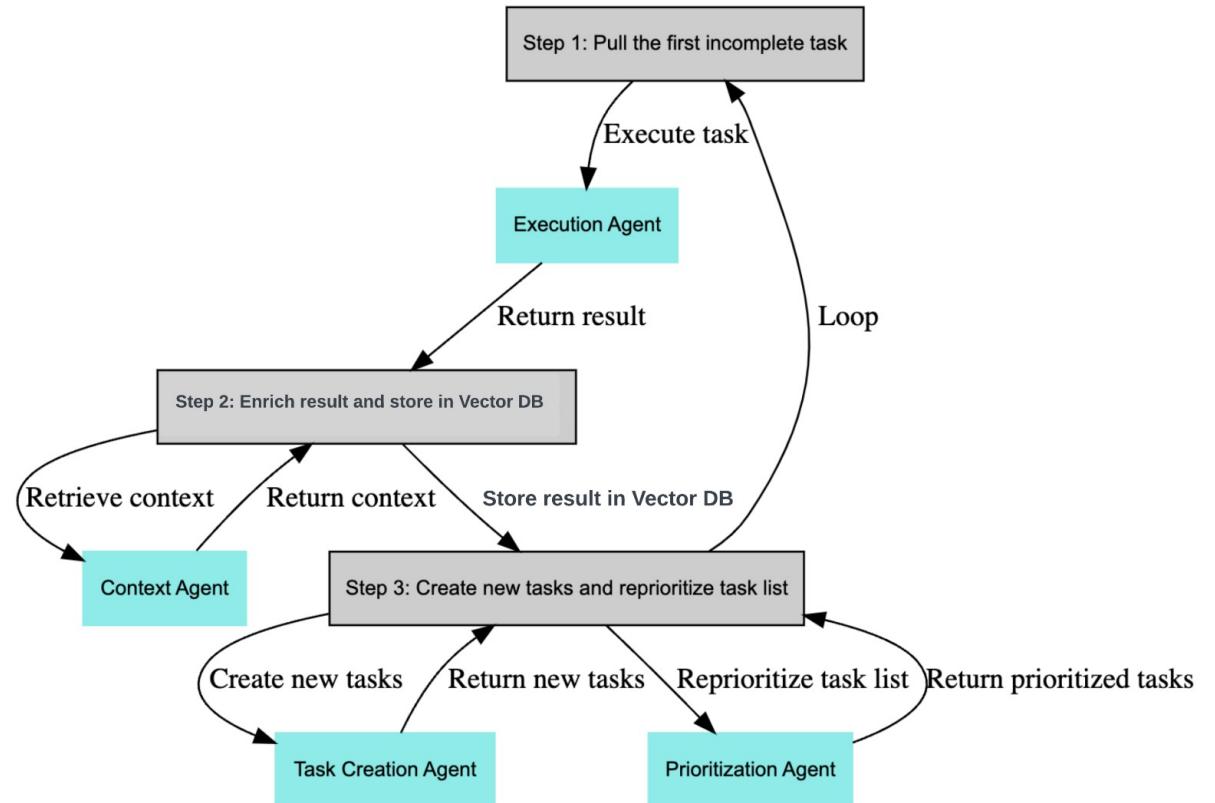
Integrating LLMs in a framework

Systems-level approach to LLMs



Example: BabyAGI / AutoGPT / AgentGPT

- Uses GPT4 to in a larger ecosystem to:
 - Create Tasks
 - Execute Tasks
 - Prioritize Tasks
- Uses memory to store and retrieve task/result pairs
 - Stored by vector embeddings



<https://github.com/yoheinakajima/babyagi>

Let's see it in action!

The screenshot shows the AgentGPT interface in Beta mode. At the top, it says "Assemble, configure, and deploy autonomous AI Agents in your browser." Below the header, there's a section titled "AgentGPT-3.5 (3 / 5 runs)" which lists the following tasks:

- ★ Embarking on a new goal: Create a comprehensive report of the Nike company
- Thinking... (Restart if this takes more than 30 seconds)
- ★ Added task: Gather financial statements and annual reports of Nike from the past few years
- ★ Added task: Conduct a SWOT analysis of the company
- ★ Added task: Research and analyze Nike's market share and competition
- ★ Added task: Write a comprehensive report summarizing all findings
- Thinking... (Restart if this takes more than 30 seconds)

At the bottom of this section is a "Pause Mode" toggle switch. To the right, there's a "Current tasks" sidebar with four items:

- Gather financial statements and annual reports of Nike from the past few years
- Conduct a SWOT analysis of the company
- Research and analyze Nike's market share and competition
- Write a comprehensive report summarizing all findings

Below the sidebar are "Custom task" and "Add" buttons. At the very bottom, there are fields for "Name" (set to "ResearchGPT") and "Goal" (set to "Create a comprehensive report of the Nike company").

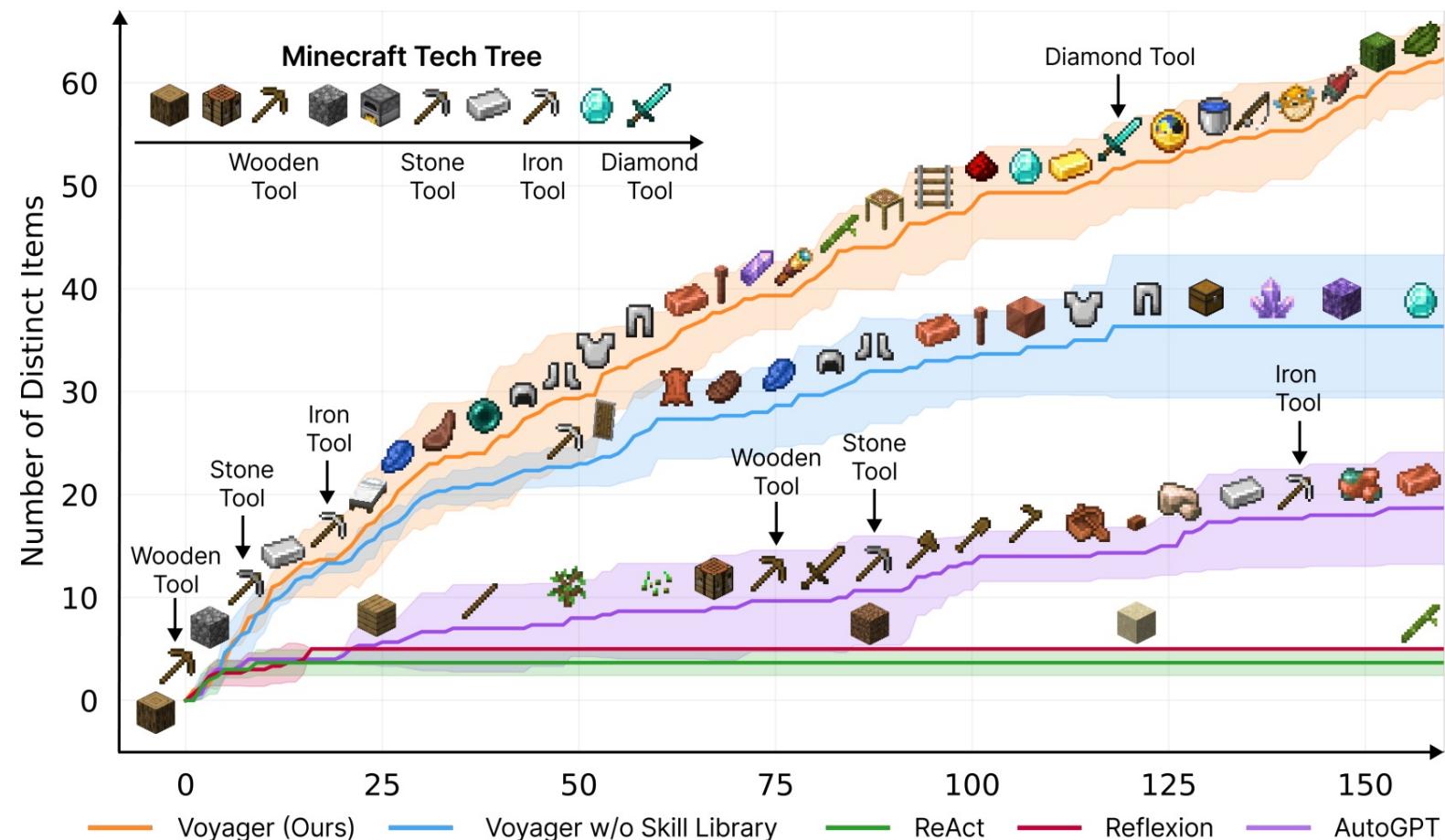
<https://agentgpt.reworkd.ai/>

VOYAGER: An Open-Ended Embodied Agent with Large Language Models

Better than AutoGPT...

Wait, what??!

Voyager has the best performance out of LLM-based methods



Components of Voyager

Teacher comes up with something manageable for student

Automatic Curriculum

Student learns (multiple tries)

Iterative Prompting Mechanism

Student remembers (only good stuff)

Skill Library

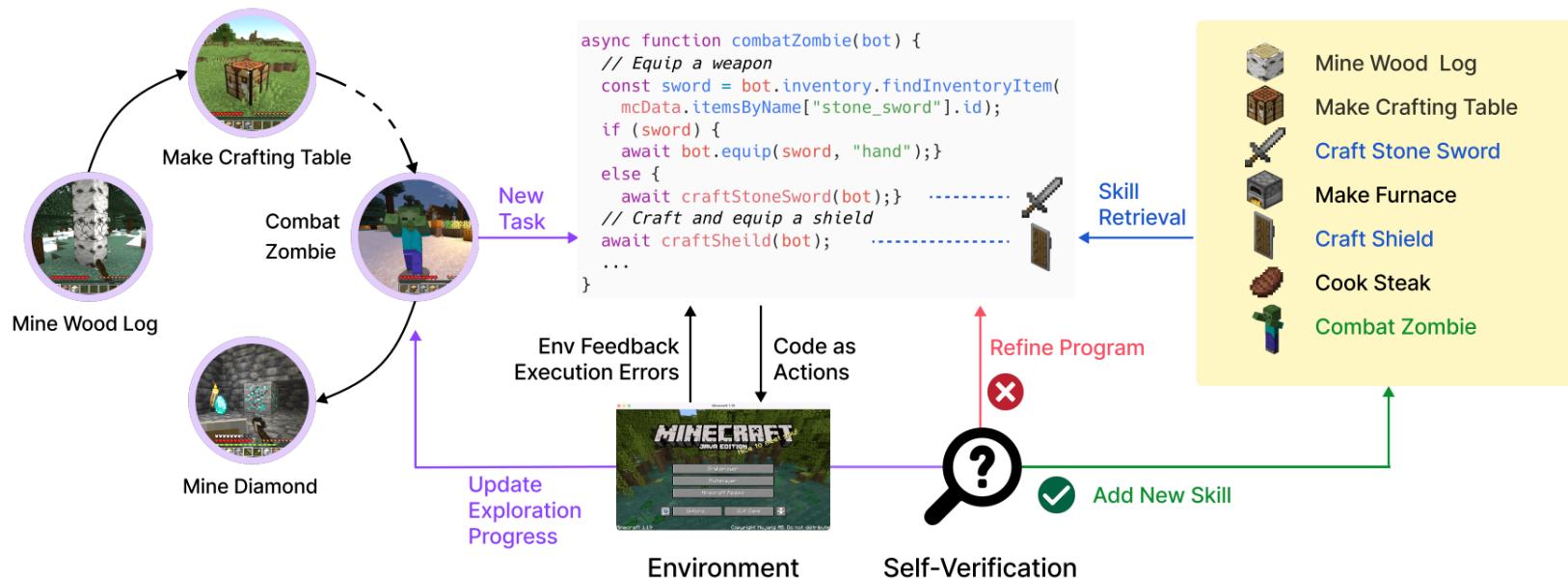


Figure 2: VOYAGER consists of three key components: an automatic curriculum for open-ended exploration, a skill library for increasingly complex behaviors, and an iterative prompting mechanism that uses code as action space.

Automatic Curriculum Generator

The teacher

Automatic Curriculum Generator

- Generated by GPT-4 based on the overarching goal of “discovering as many diverse things as possible”



Inventory (5/36): {'oak_planks': 3, 'stick': 4, 'crafting_table': 1, 'stone': 3, 'wooden_pickaxe': 1}

GPT-4

Reasoning: Since you have a wooden pickaxe and some stones, it would be beneficial to upgrade your pickaxe to a stone pickaxe for better efficiency.
Task: Craft 1 stone pickaxe.



Biome: river

Inventory (4/36): {'sugar_cane': 3, 'paper': 3, 'fishing_rod': 1, 'wooden_pickaxe': 1}

GPT-4

Reasoning: Since you have a fishing rod in your inventory and you are near a river biome, it would be a good opportunity to catch some fish for food and experience.
Task: Catch 1 fish.



Nearby entities: pig, cat, villager

Health: 12/20

Hunger: 0/20

GPT-4

Reasoning: Your hunger is at 0, which means you need to find food to replenish your hunger. Since there are pigs nearby, you can kill one to obtain raw porkchops.
Task: Kill 1 pig.



Inventory (6/36): {'furnace': 1, 'stone_pickaxe': 1, 'oak_planks': 7, 'crafting_table': 1, 'raw_iron': 4, 'coal': 1}

GPT-4

Reasoning: You have raw iron and coal, and you have a furnace. It's time to smelt the iron to make iron ingots, which can be used to craft better tools and armor.
Task: Smelt 4 raw iron.



Time: night

Nearby entities: zombie

Equipment: 'stone_sword', 'shield'

GPT-4

Reasoning: Since it's night and there's a zombie nearby, it's a good opportunity to try killing the zombie now that you have a stone sword and a shield equipped.
Task: Kill 1 zombie.

Prompt for Teacher: Curriculum Generator

- **Overall prompt encouraging diverse behaviors and imposing constraints**
 - “My ultimate goal is to discover as many diverse things as possible ... The next task should not be too hard since I may not have the necessary resources or have learned enough skills to complete it yet.”
- **The agent's current state:** Inventory, equipment, nearby blocks and entities, biome, time, health and hunger bars, and position
- **History:** Previously completed and failed tasks, reflecting the agent's current exploration progress and capabilities frontier
- **Memory retrieval:** GPT-3.5 to self-ask questions based on the agent's current state and exploration progress and self-answer questions with a wiki knowledge base to provide additional context to GPT-4.

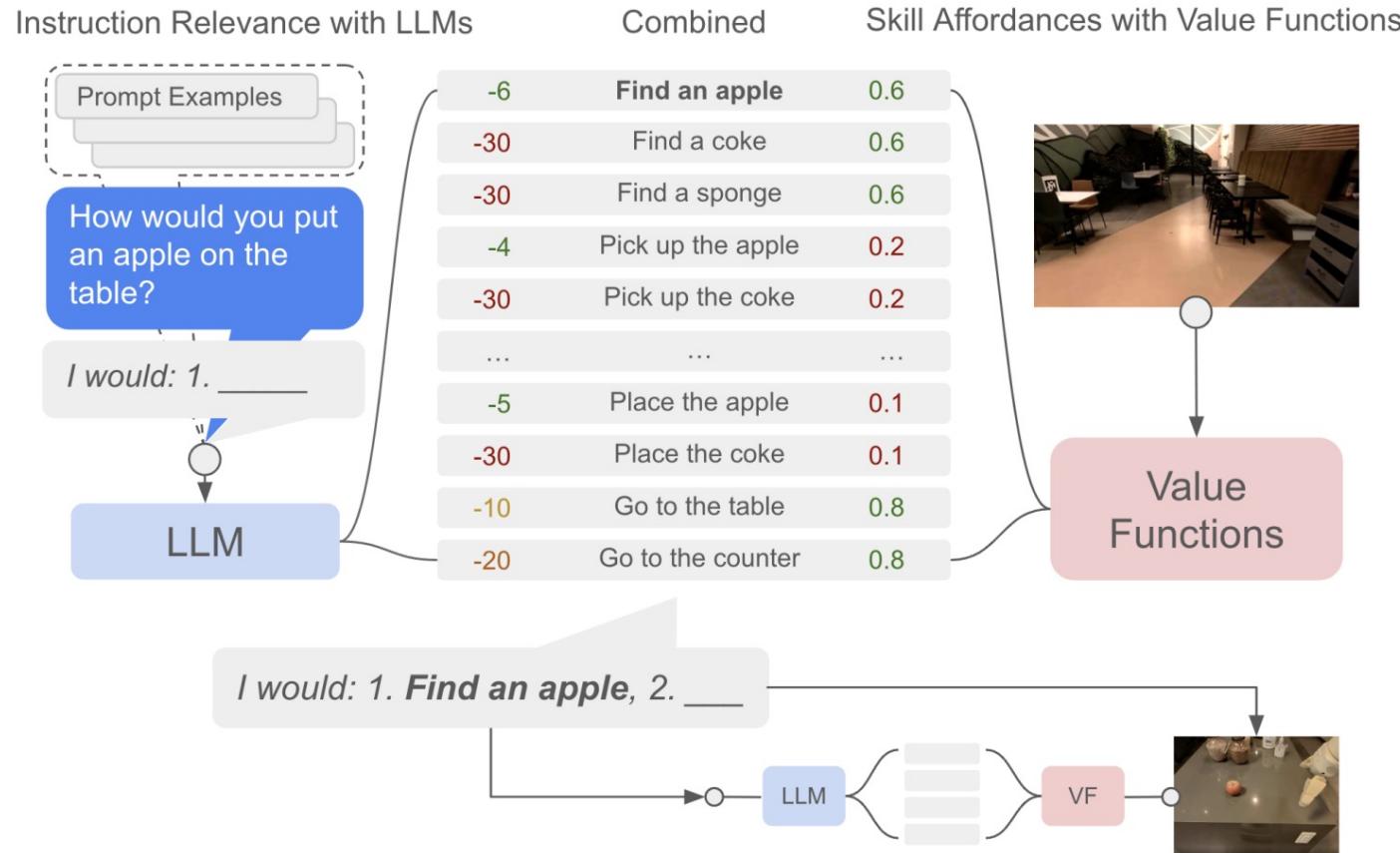
Iterative Prompting

Now that teacher has come up with a curriculum, student needs to learn from it

Problems with GPT generation

- Lacks grounding with the environment
- May not know what are the functions available to use in the javascript code for Minecraft
- Code may not lead to completion of task

Previously: Value Function Grounding



- This is a **slow way of learning** as value functions take a long time to converge
- Encoding the environment feedback via text may be more natural

Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. Ahn et al. 2022

Iterative Prompting : Give the agent environment and code feedback!

- **Environment Feedback**
 - `bot.chat()` inside control primitive APIs to generate environment feedback and prompt GPT-4 to use this function as well during code generation
- **Code Feedback**
 - Execution errors Generated by program interpreter that reveals any invalid operations or syntax errors in programs
- **Self-verification for checking task success**
 - Another GPT-4 agent for self-verification. Takes in current state and task, and outputs if program achieves the task.
 - Gives critique (constructive feedback) if unable to achieve.
 - **(Potentially limited) GPT-4 is still unable to simulate multiple time steps well without using the environment engine**

Actor: RL Loop now via past-history prompting!

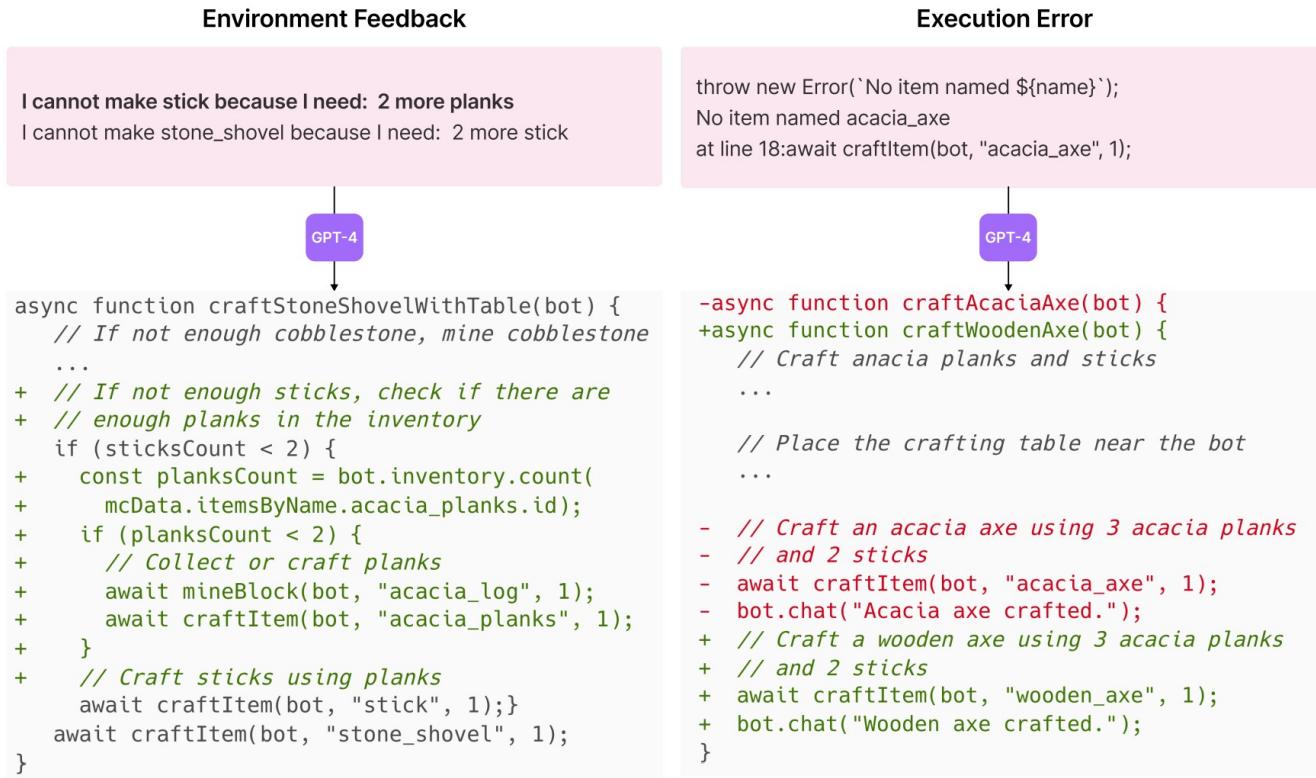


Figure 5: **Left: Environment feedback.** GPT-4 realizes it needs 2 more planks before crafting sticks. **Right: Execution error.** GPT-4 realizes it should craft a wooden axe instead of an acacia axe since there is no acacia axe in Minecraft. We only display the partial prompt for brevity. The full prompt structure for code generation is in Appendix, Sec. A.4.

- Mitigates lack of grounding with environment
- Interpreter solves to a limited extent GPT not knowing which functions are available

Critic: Using GPT4 to verify if task is completed



Does GPT4 already have the answers?

- Prompt: “Generate me the Minecraft Java code to craft a diamond tool using the Mineflayer API”

```
// Determine the recipe
const recipe = bot.recipesFor(Item.toNotch('diamond_pickaxe'))[0];

if (!recipe) {
    console.log("I don't know how to craft that!");
    return;
}
```

Does GPT4 already have the answers (CoT)?

- Prompt: “How do we generate a diamond tool in Minecraft?”
- Followed by “Generate me the Minecraft Java code to craft a diamond tool using the Mineflayer API”

```
if (recipe) {  
    bot.chat(`Crafting ${itemName}`);  
    bot.craft(recipe, 1, craftingTable, (err) => {  
        if (err) bot.chat(err.message);  
        else bot.chat(`Crafted ${itemName}`);  
    });  
} else {  
    bot.chat(`No recipe for ${itemName}`);  
}
```

Skill Generation

Student can learn better using past successful memories

Skill as Code

- Represent each skill with **executable code** that scaffolds temporally extended actions for completing a specific task proposed by the automatic curriculum



GPT Prompt for Skill Generation

- **Overall prompt to encourage generic skills:** “Your function will be reused for building more complex functions. Therefore, you should make it generic and reusable.”
- **Tool Description:** Control primitive APIs, and relevant skills retrieved from the skill library
- **History:** The generated code from the last round, environment feedback, execution errors, and critique, based on which GPT-4 can self-improve
- **Agent information:** The agent’s current state, including inventory, equipment, nearby blocks and entities, biome, time, health and hunger bars, and position
- **Chain-of-thought prompting** to do reasoning before code generation.

Skill Library (aka Memory)

Program Generated by GPT-4

```
async function combatZombie(bot) {  
    // Equip a weapon  
    const sword =  
    bot.inventory.findInventoryItem(  
        mcData.itemsByName[  
            "stone_sword"  
        ].id  
    );  
    if (sword) {  
        await bot.equip(sword, "hand");  
    } else {  
        await craftStoneSword(bot);  
    }  
    ...  
    // Craft and equip a shield  
    ...  
    // Recover hunger  
    ...  
    // Look for and combat a zombie  
    ...  
}
```

Program Description

```
async function combatZombie(bot) {  
    // The function is about  
    equipping a stone sword to combat  
    a zombie. If a stone sword is not  
    found, it will craft one.  
    Additionally, it crafts and equips  
    a shield for added protection.  
    Afterwards, it proceeds to cook  
    sticks in order to restore hunger.  
    Once hunger is replenished, it  
    actively searches for a zombie and  
    engages in combat with it.  
}
```

GPT-3.5

Embedding

Key

Add

Skill Library



Task: Craft Iron Pickaxe

How to craft an iron pickaxe in Minecraft?

GPT-3.5

To craft an iron pickaxe, you
need to 3 iron ingots and 2
sticks. Once you have gathered
the materials,

Environment Feedback

Embedding

Query

Skill Library

Retrieve

Top-5 Relevant Skills



- **Storage:** Skill stored with summarized program description text embedding as key, program code as value

- **Retrieval:** Task converted into description taking into account environment feedback, use embedding to find closest match to skills

Results

Let's analyze why the results are better

Baselines

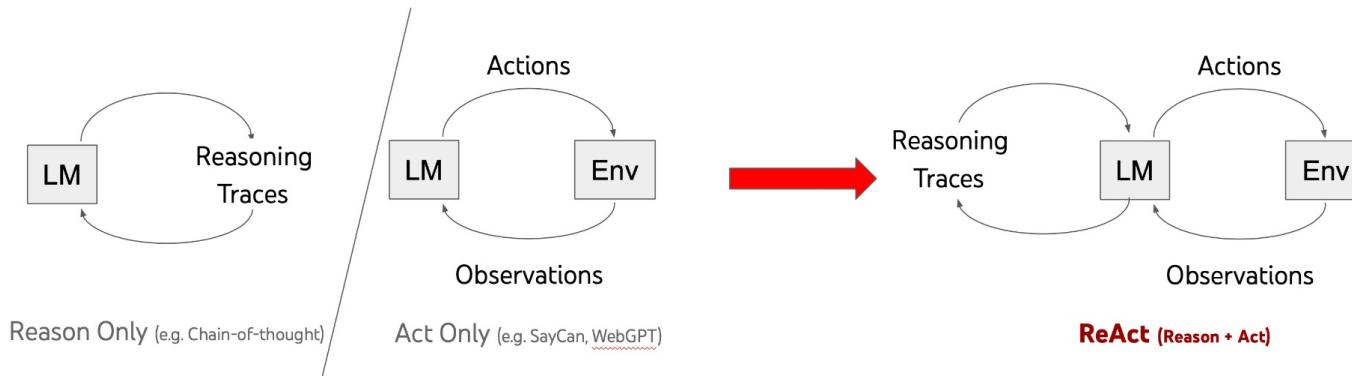


Image from: <https://react-lm.github.io/>

- **ReAct:** Chain of thought prompting plus environment grounding.
 - Thought, Action, Observation
- **Reflexion:** ReAct plus “Reflect on your answer”
- **AutoGPT:** Sub-task decomposition in a ReAct-like framework

Tech Tree Mastery

Table 1: Tech tree mastery. Fractions indicate the number of successful trials out of three total runs. 0/3 means the method fails to unlock a level of the tech tree within the maximal prompting iterations (160). Numbers are prompting iterations averaged over three trials. The fewer the iterations, the more efficient the method.

Method	Wooden Tool	Stone Tool	Iron Tool	Diamond Tool
ReAct [29]	N/A (0/3)	N/A (0/3)	N/A (0/3)	N/A (0/3)
Reflexion [30]	N/A (0/3)	N/A (0/3)	N/A (0/3)	N/A (0/3)
AutoGPT [28]	92 ± 72 (3/3)	94 ± 72 (3/3)	135 ± 103 (3/3)	N/A (0/3)
VOYAGER w/o Skill Library	7 ± 2 (3/3)	9 ± 4 (3/3)	29 ± 11 (3/3)	N/A (0/3)
VOYAGER (Ours)	6 ± 2 (3/3)	11 ± 2 (3/3)	21 ± 7 (3/3)	102 (1/3)

Map Exploration

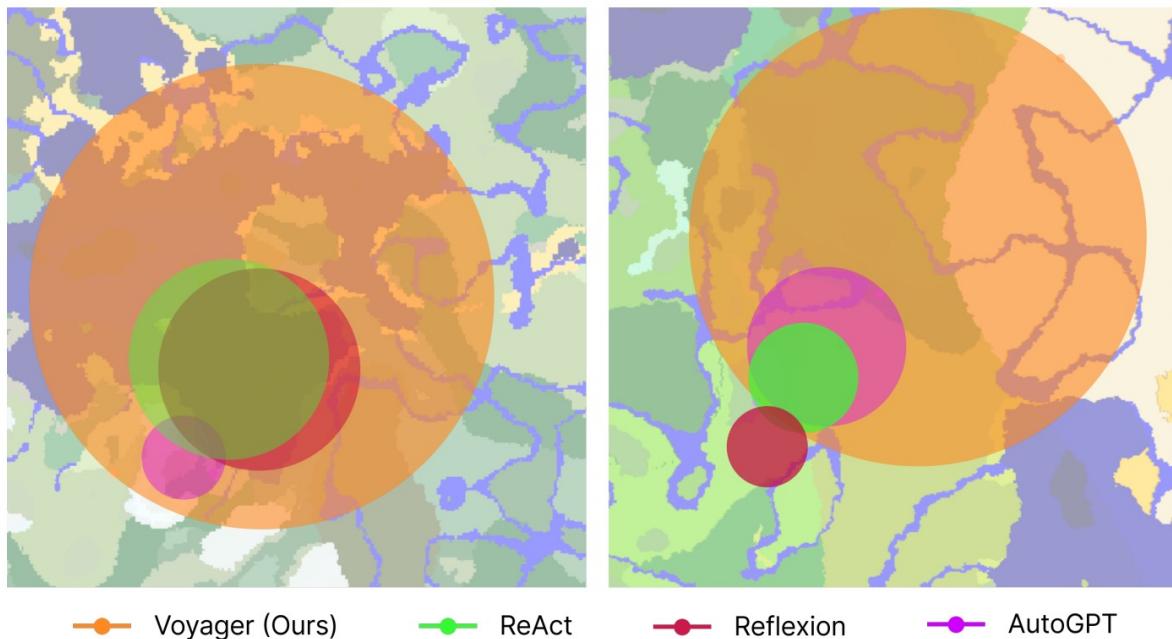


Figure 7: Map coverage: bird's eye views of Minecraft maps. VOYAGER is able to traverse $2.3\times$ longer distances compared to baselines while crossing diverse terrains.

Zero-shot generalization to unseen tasks

- Skill library (aka memory) makes a huge difference here

Table 2: Zero-shot generalization to unseen tasks. Fractions indicate the number of successful trials out of three total attempts. 0/3 means the method fails to solve the task within the maximal prompting iterations (50). Numbers are prompting iterations averaged over three trials. The fewer the iterations, the more efficient the method.

Method	Diamond Pickaxe	Golden Sword	Lava Bucket	Compass
ReAct [29]	N/A (0/3)	N/A (0/3)	N/A (0/3)	N/A (0/3)
Reflexion [30]	N/A (0/3)	N/A (0/3)	N/A (0/3)	N/A (0/3)
AutoGPT [28]	N/A (0/3)	N/A (0/3)	N/A (0/3)	N/A (0/3)
AutoGPT [28] w/ Our Skill Library	39 (1/3)	30 (1/3)	N/A (0/3)	30 (2/3)
VOYAGER w/o Skill Library	36 (2/3)	30 ± 9 (3/3)	27 ± 9 (3/3)	26 ± 3 (3/3)
VOYAGER (Ours)	19 ± 3 (3/3)	18 ± 7 (3/3)	21 ± 5 (3/3)	18 ± 2 (3/3)

Conclusion

- GPT learns very fast: it is very good at zero-shot description-based learning, and few-shot example-based learning
- Imbuing GPT prompts with in-game context can help ground it better
 - Having a feedback loop to the environment via code can help
- Having external memory in the form of code-blocks of skills can help improve performance

Critique (my view)

- Prompts are too hand-tuned (look at Appendix). It cannot generalize beyond Minecraft. Perhaps we can have retrieval-augmented generation so that the prompts can be made more generic and augmented with domain-specific knowledge
- Agent does not start from scratch here (MineFlayer API has some basic functions like movement). We can perhaps also learn the basic level directly via an embedding method like Decision Transformers.
- (In paper) Can combine with visual models to augment more information to GPT-4
 - My view: No need to have visual models as it is too high-dimensional. Need to better improve vision-to-text models like BLIP for text input.

Potential Improvements (my view)

- Learning from expert trajectories
 - Humans also learn from each other, especially the best performing one
 - Helps a lot to have model answer (refer to synthetic examples in AlphaTensor / HyperTree Proof Search)
- Prompting from broad to specific actions might help better with hierarchical planning
- Critique can be given the environment (something like world model) to evaluate code

Questions to Ponder

- Curriculum Generation: How does GPT know what are the tasks to create given the inventory, nearby entities and equipment?
- Is there a better way to represent skills other than through code? Should skills be represented as a sequence of primitive skills (Ghost in the Minecraft paper)?
- Can all action spaces be mapped into code? What are the pros and cons of denoting actions as code?
- For the skill library (aka memory), is there a need to also store negative examples (aka code that did not work)?
- Should memory be purely just skills alone? Should there be memory at more levels?

Voyager Pseudocode

Pseudocode 1: VOYAGER algorithm.

```
def voyager(
    environment, # environment that uses code as action space
    curriculum_agent, # curriculum agent for proposing the next task
    action_agent, # action agent for code generation
    critic_agent, # critic agent for self-verification
    skill_manager, # skill manager for adding new skills and skill
    retrieval
):
    agent_state = environment.reset()
    while True:
        exploration_progress = (
            curriculum_agent.get_exploration_progress(
                curriculum_agent.get_completed_tasks(),
                curriculum_agent.get_failed_tasks(),
            )
        )
        task = curriculum_agent.propose_next_task(
            agent_state, exploration_progress
        )
        code = None
        environment_feedback = None
        execution_errors = None
        critique = None
        success = False
```

Voyager Pseudocode

```
# try at most 4 rounds before moving on to the next task
for i in range(4):
    skills = skill_manager.retrieve_skills(
        task, environment_feedback
    )
    code = action_agent.generate_code(
        task,
        code,
        environment_feedback,
        execution_errors,
        critique,
        skills,
    )
    (
        agent_state,
        environment_feedback,
        execution_errors,
    ) = environment.step(code)
    success, critique = critic_agent.check_task_success(
        task, agent_state
    )
    if success:
        break
if success:
    skill_manager.add_skill(code)
    curriculum_agent.add_completed_task(task)
else:
    curriculum_agent.add_failed_task(task)
```