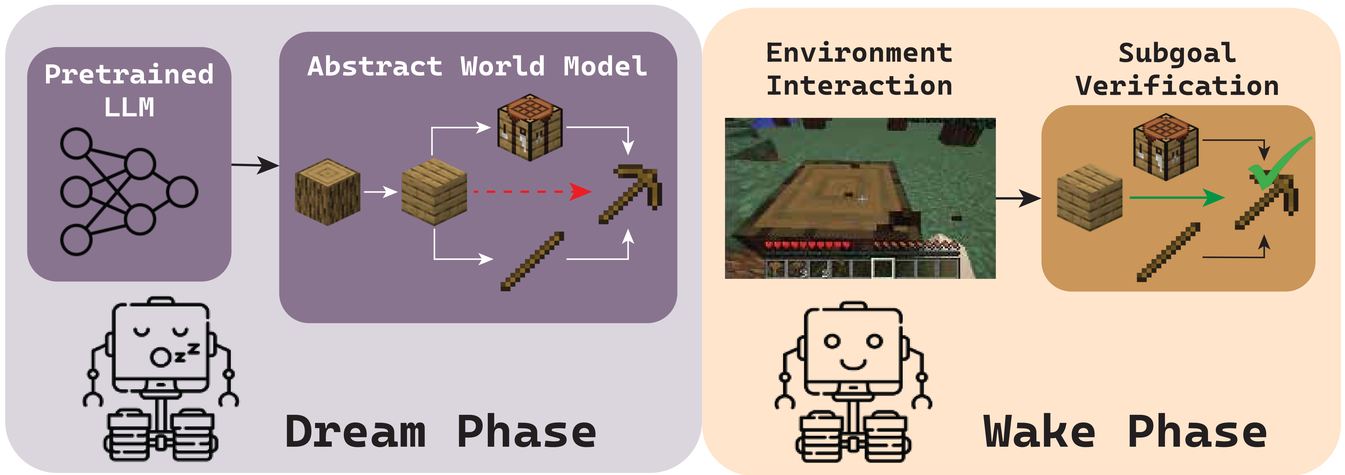
Do Embodied Agents Dream of Pixelated Sheep?

Embodied Decision Making using Language Guided World Modelling

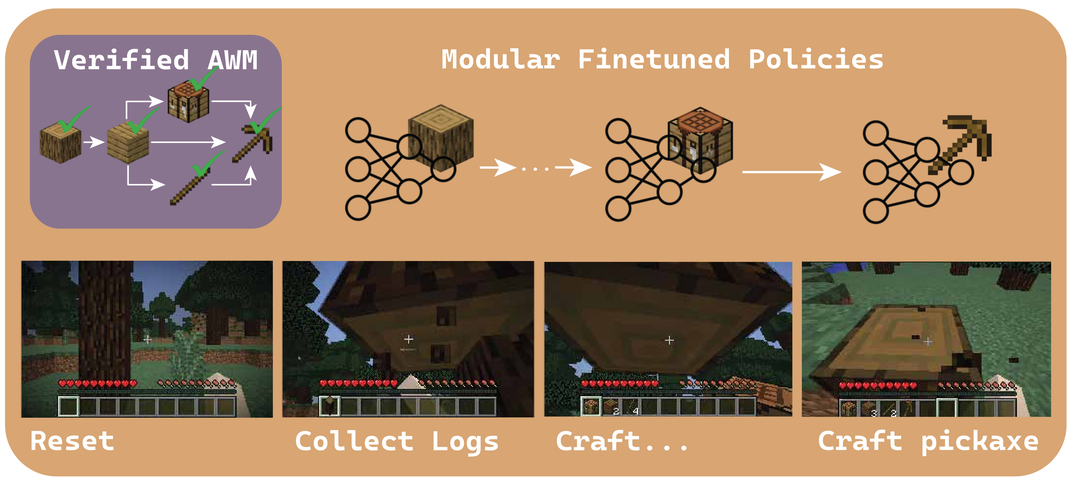
Kolby Nottingham[1] Prithviraj Ammanabrolu[2] Alane Suhr[2] Yejin Choi[3,2] Hannaneh Hajishirzi[3,2] Sameer Singh[1,2] Roy Fox[1]

1: University of California Irvine 2: Allen Institute for AI 3: University of Washington

DECKARD Training



DECKARD Inference



Motivation

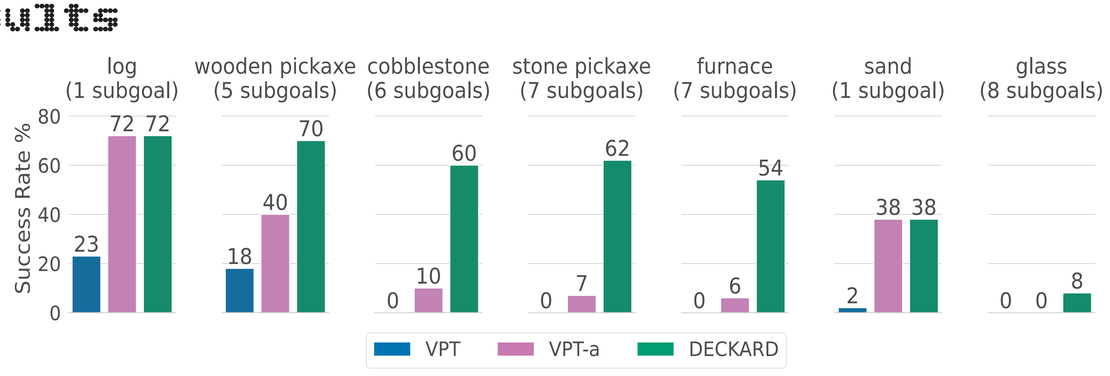
- Rather than learning tabula-rasa, RL agents can leverage existing knowledge.
- LLMs for decision making should have a way to verify output and correct errors.

Method

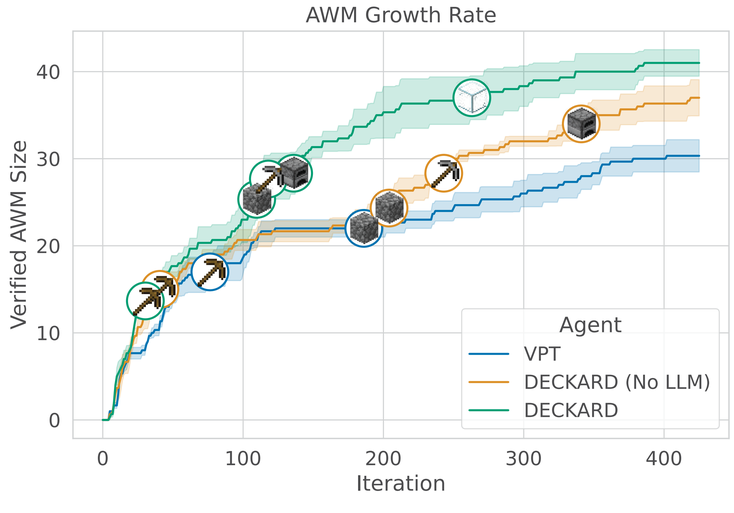
- DECKARD operates in two phases

1. Dream Phase
   - Sample the next subgoal from an LLM
2. Wake Phase
   - Agent explores to reach subgoal
   - Subgoal is verified or corrected

Results



- DECKARD outperforms pretrained agents (VPT[1]) and RL finetuned agents (VPT-a) on crafting tasks.



- DECKARD discovers new recipes faster than baseline agents and ablations.
- LLM guidance is key to exploring the state space.

deckardagent.github.io

knotting@uci.edu

1. Baker, B., Akkaya, I., Zhokov, P., Huizinga, J., Tang, J., Ecoffet, A., Houghton, B., Sampedro, R. and Clune, J., 2022. Video pretraining (vpt): Learning to act by watching unlabeled online videos. Advances in Neural Information Processing Systems, 35, pp.24639-24654.