# End-to-End Learning to Follow Language Instructions with Compositional Policies

Vanya Cohen*, Geraud Nangue Tasse*, Nakul Gopalan, Steven James, Raymond Mooney, Benjamin Rosman

Our method leverages boolean compositional value functions (Nangue Tasse et al., 2020) and pretrained LLM representations to enable compositional generalization in a language-conditioned goal-reaching task

Compositional value functions and **large language model representations** enable **substantial sample efficiency gains** in an instruction following task.

METHODS

- Pretrain compositional value functions for each task attribute ("pick up the red object")
- Learn to compose value functions. Condition layout on the mission description ("pick up the red key")
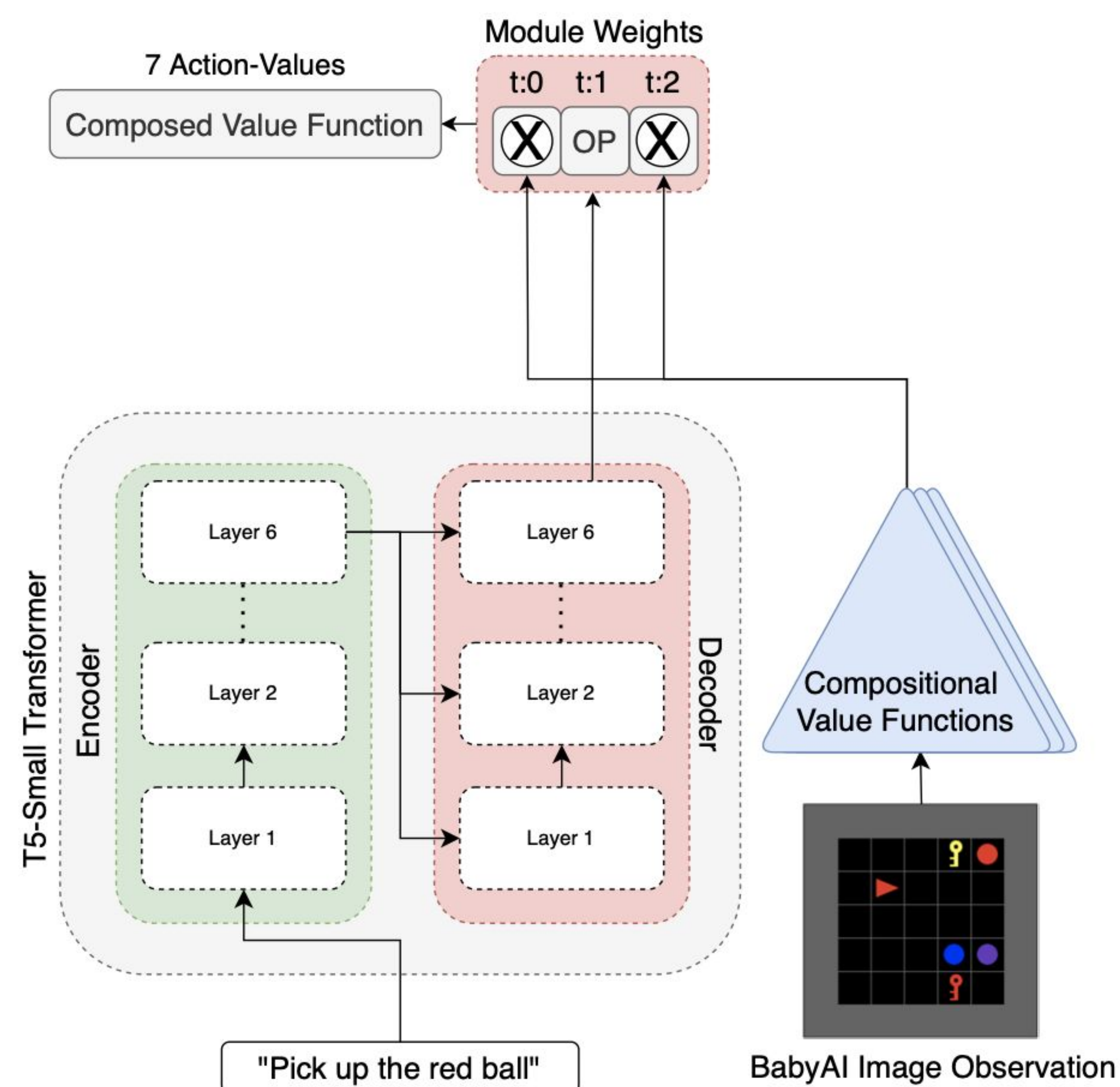
Neural Module Value Network (NMVN)
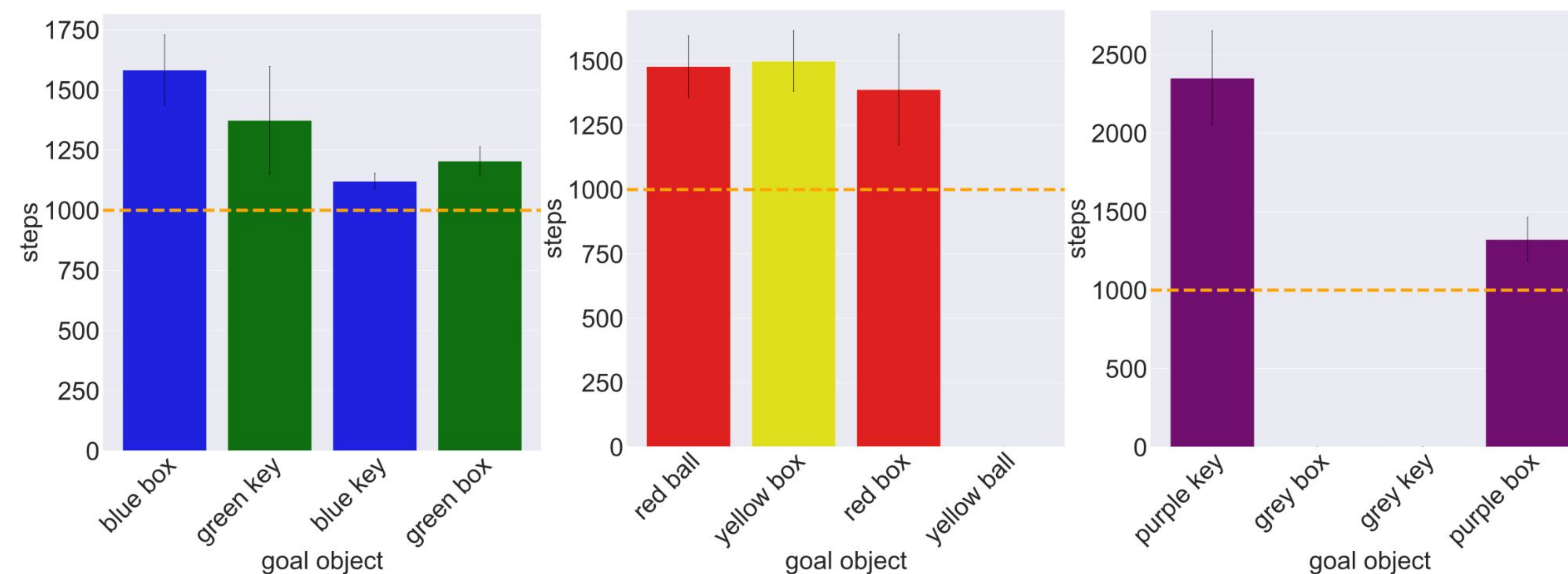
Task set for the attributes {box, key, green, blue}. The agent requires significantly fewer learning steps to learn each additional task. The agent requires fewer steps to learn the held-out combinations of attributes. For the two held-out attribute combinations, the agent requires at less than 250 environment steps to learn each task.

Task set for the attributes {box, ball, yellow, red}. The agent requires no additional steps to learn the held-out "pick up the yellow ball" task but roughly the same number of steps to learn the "pick up the red box" task.

Task set for the attributes {key, box, purple, grey}. In this sequence the middle two tasks now share color attributes. The agent requires no additional steps to learn the held-out "pick up the grey box" and "pick up the grey key" tasks and requires fewer steps to learn the "pick up the purple box" task.
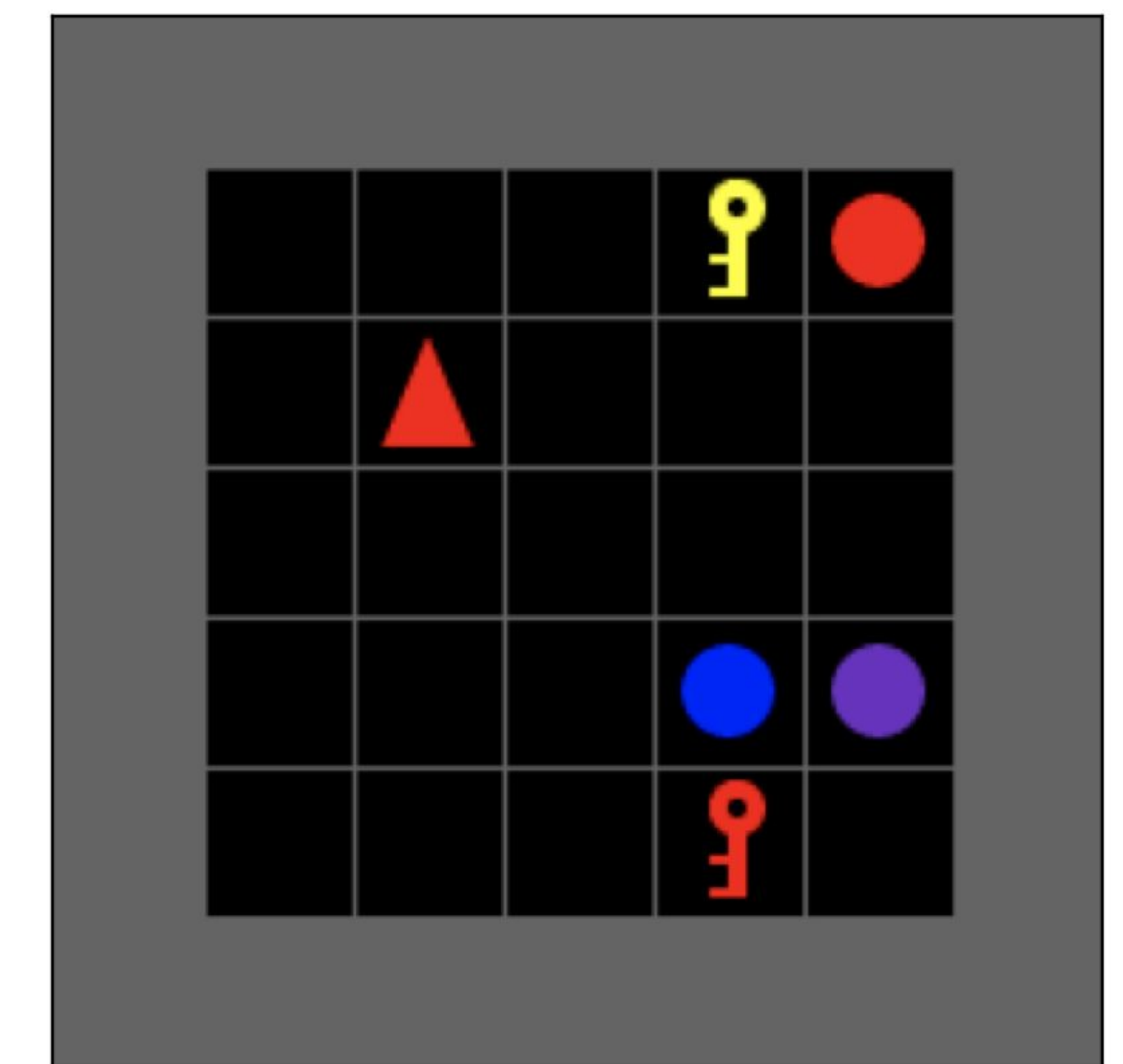
Our BabyAI domain consists of goal-reaching tasks in a gridworld. The agent must navigate to the correct object and pick it up while avoiding four distractor objects. Objects have type and color attributes.

RESULTS

- Evaluate the number of steps to learn tasks in series with shared attributes.
- Learning held-out tasks requires substantially fewer or no additional learning steps
- Compositional generalization from combining compositional value function and language representations

Scan the QR code to read the full paper