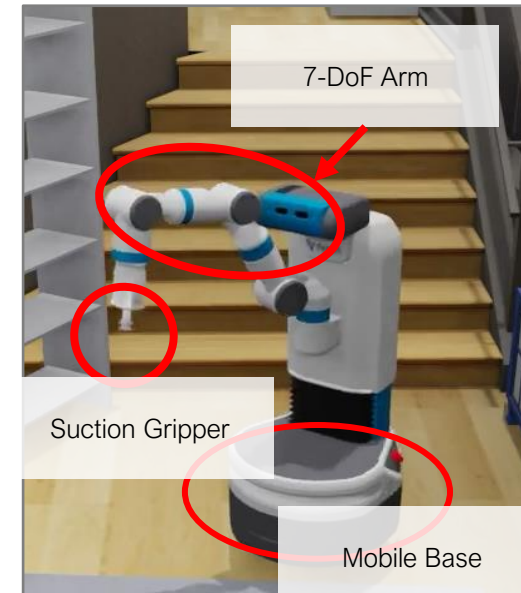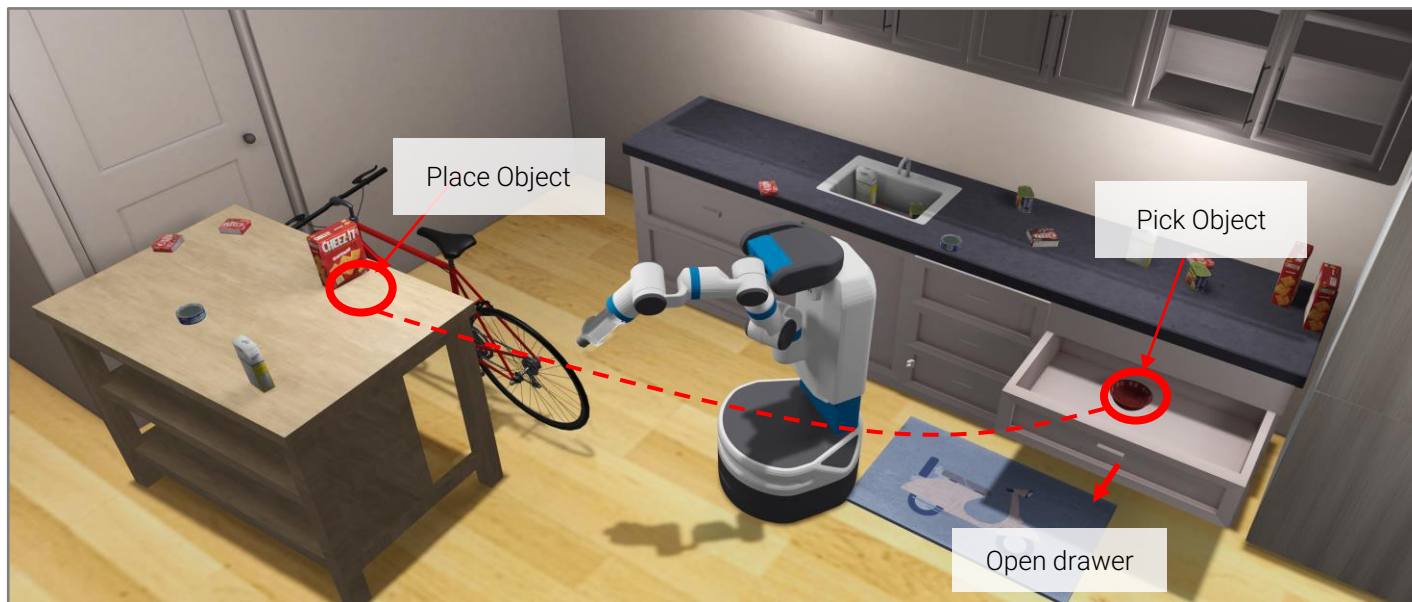# Reinforcement Learning via Auxiliary Task Distillation

Consider the task of using a robot to rearrange an object in the house
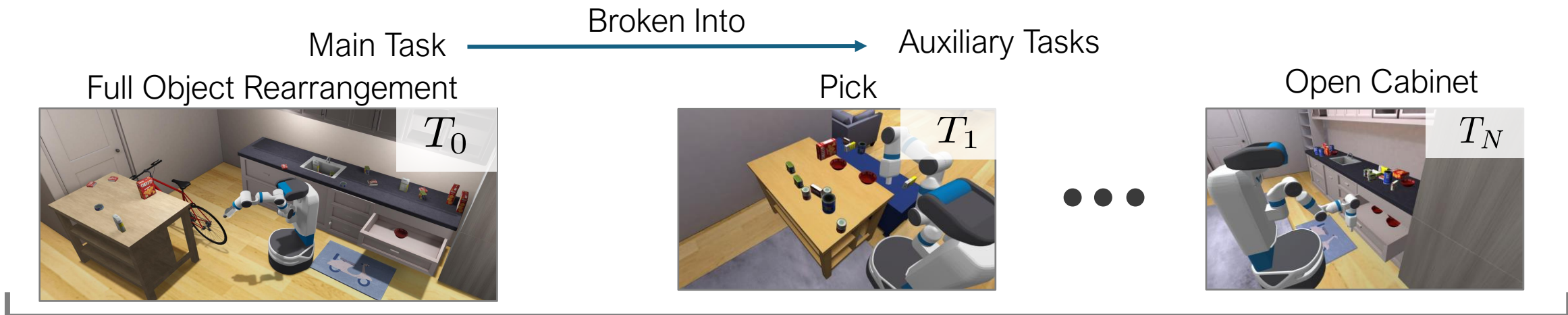
- Fetch-Robot with 10-DOF and a suction gripper
- Requires diverse skills like <u>Navigating,</u> <u>Opening</u> a cabinet, <u>Picking</u> up, and <u>Placing</u>

Can long-horizon robot control be learnt end-to-end without using demonstrations or a curriculum?
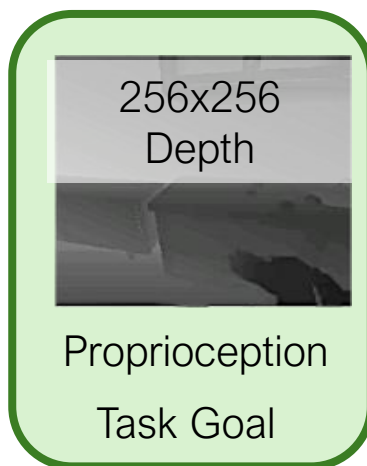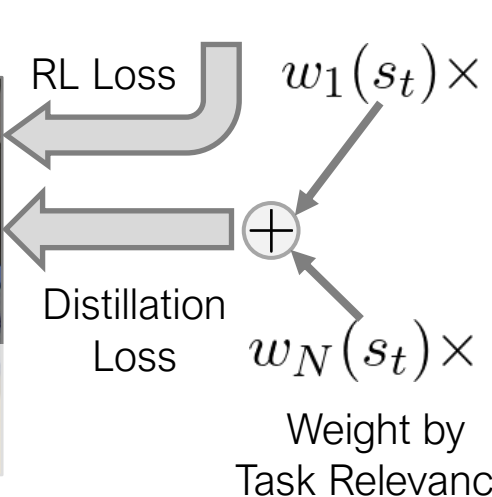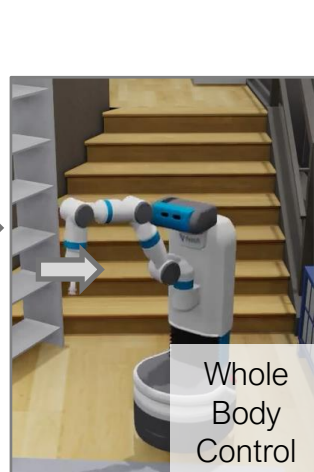
Yes, by using Auxiliary Tasks!

- Auxiliary tasks carry <u>relevant behaviors</u> which are <u>easier to learn</u> and transferred to the main task
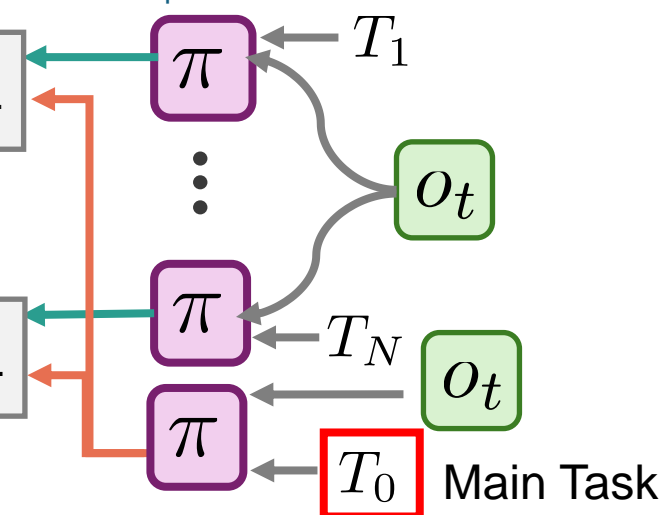- They are learnt simultaneously along with the main task

Broken Into

Main Task → Auxiliary Tasks

Full Object Rearrangement T_0

Pick T_1

$\bullet\bullet\bullet$

Open Cabinet T_N

2. Collect Rollouts

256x256 Depth

Proprioception

Task Goal

π

3. Multi-Task RL loss across tasks

RL Loss

Distillation Loss

Whole Body Control

$w_1(s_t) \times$

$w_N(s_t) \times$

Weight by Task Relevance

$+$

KL

KL

4. Compute Distillation Loss

$\pi \leftarrow T_1$

$\pi \leftarrow T_N$

$\pi \leftarrow T_0$ Main Task

o_t

o_t

Results

Outperforms a variety of end-to-end and hierarchical baselines by 2.3x

Easy: Episodes in which the object is placed in an open receptacle
Hard: Object is placed inside a closed receptacle

- M3 (+24%) ➔ Hierarchical RL with STRIPS planner with Navigate, Pick and Place skills
- Mono (+73%) ➔ end to end RL which directly maps observations to actions
- GALA (+24%): Scaling end to end RL with kinematic simulation (2B samples: x4 more than Aux-Distill)
- ST (+25%)➔ Transformer architecture for rearrangement using demonstrations