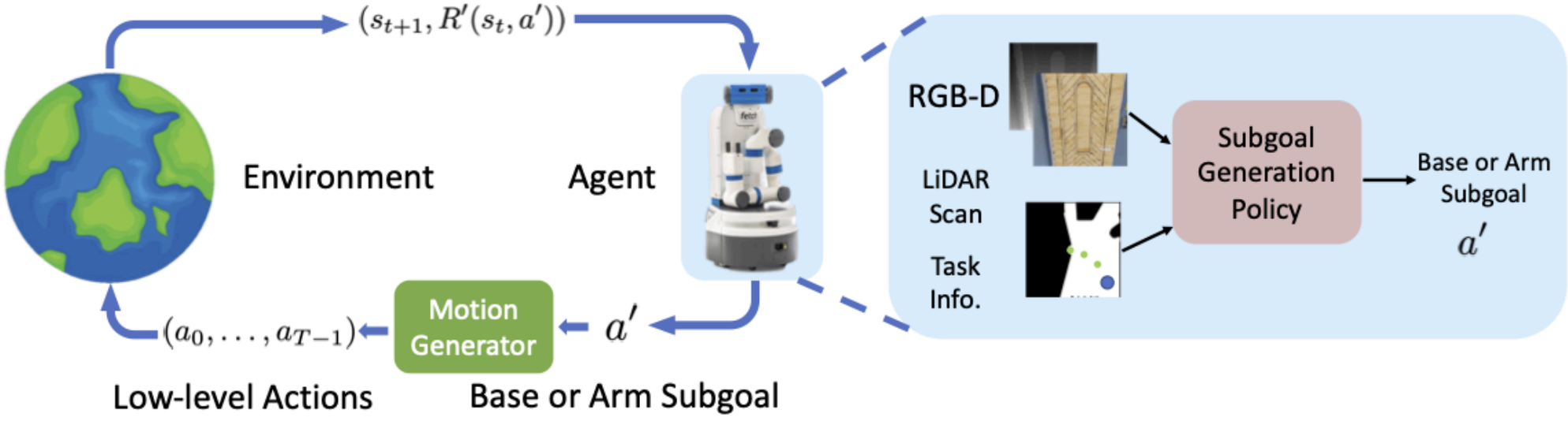
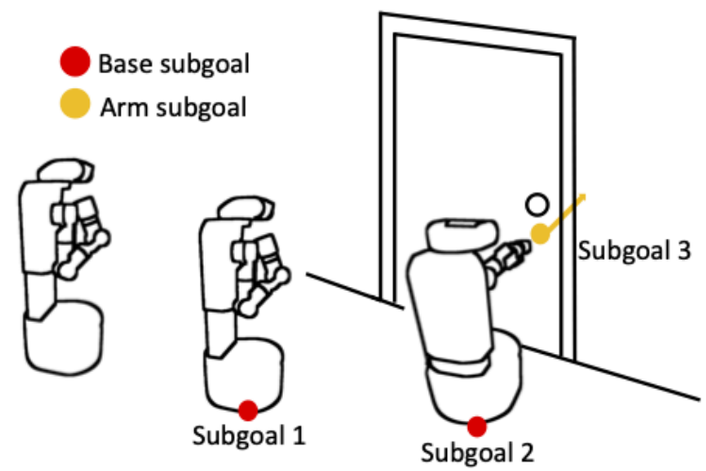
ReLMoGen

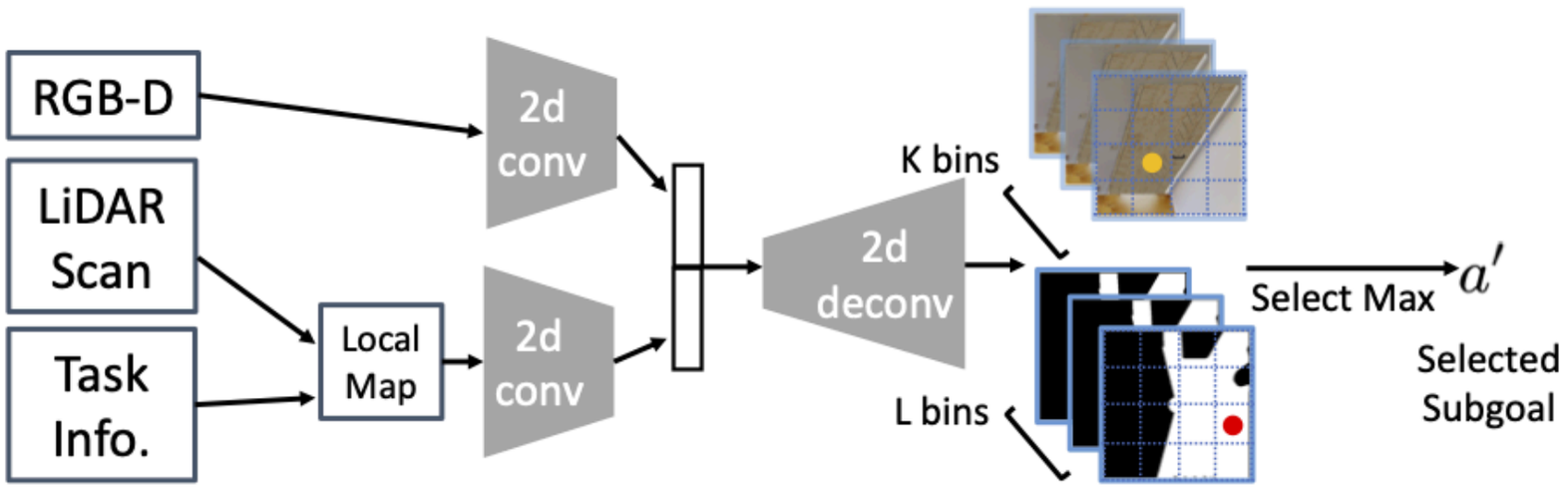


(a) Method overview.

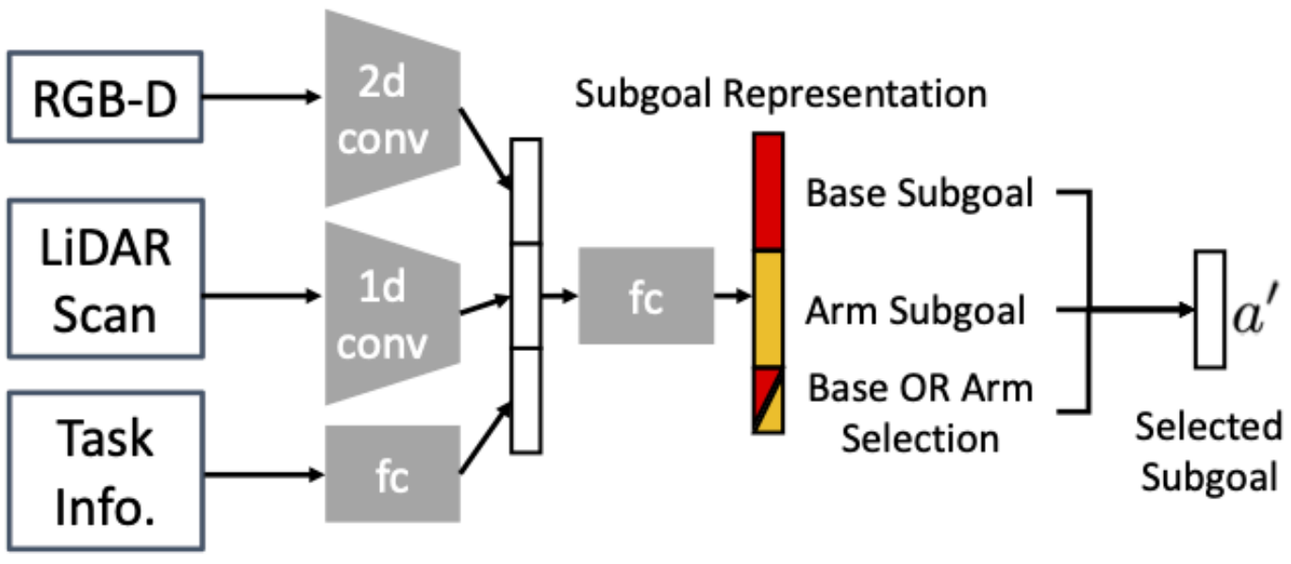


(b) A sequence of base and arm subgoals to push-open a door.

ReLMoGen



(a) Subgoal Generation Policy SGP-D



(b) Subgoal Generation Policy SGP-R

ReLMoGen

Algorithm 1: ReLMoGen Algorithm

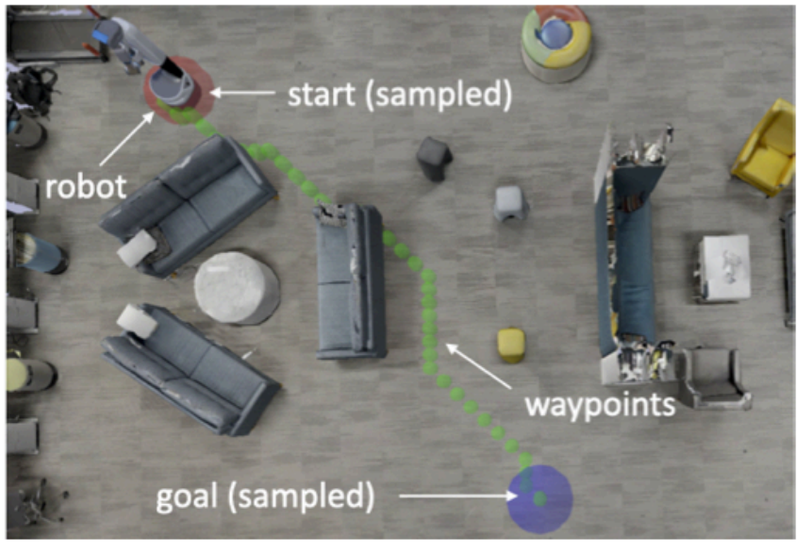
Input : env, MG, D
Output : π
Parameters: n_{iter}, n_{env_step}, n_{grad_step}

for $iter \leftarrow 1$ **to** n_{iter} **do**
  for $step \leftarrow 1$ **to** n_{env_step} **do**
    $a'_t \leftarrow \pi(o_t)$ // sample the next subgoal
    $\{a_t, a_{t+1}, ..., a_{t+T-1}\} \leftarrow \texttt{MG}(a'_t)$ // motion generator plans for T low-level actions; if the subgoal is infeasible, T = 0
    $r'_t = 0$
    for $i \leftarrow 0$ **to** $T-1$ **do**
      $o_{t+i+1}, r_{t+i+1} \leftarrow \texttt{env.step}(a_{t+i})$
      $r'_t \leftarrow r'_t + r_{t+i+1}$ // accumulate reward within a su
    end
    $\texttt{D} \leftarrow \texttt{D} \cup \{o_t, a'_t, r'_t, o_{t+T}\}$
  end
  for $step \leftarrow 1$ **to** n_{grad_step} **do**
    perform gradient updates for π with D as defined in [19] (policy gradient based) or [18
  end
end

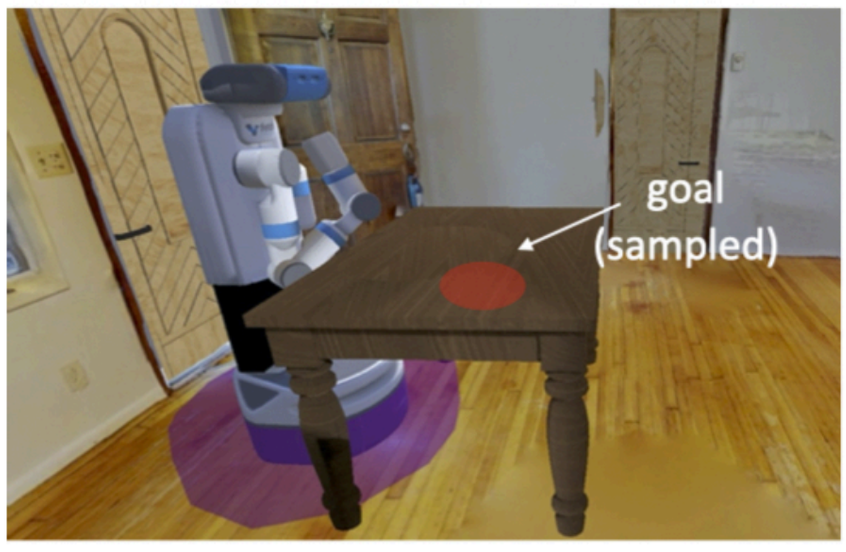
The motion planner for the base is a 2D Cartesian space RRT that searches for a collision-free path to the base subgoal location on the local map generated from the most recent LiDAR scan. The base subgoals are represented as the desired base 2D locations and orientations.

The motion planner for the arm comprises a 3D Cartesian space RRT and a simple Cartesian space inverse kinematics (IK) based planner. The arm motion is made of two phases: 1) the motion from the initial configuration to the selected subgoal location, and 2) the pushing interaction starting from the subgoal location. For the first phase, the 3D RRT searches for a collision-free path to reach the subgoal location. If the first phase succeeds, as the second phase, the simple IK-based planner is queried to find a sequence of joint configurations to move the end effector in a straight line from the subgoal location along the specified pushing direction. Since the intent of the second phase is to interact with the environment, the path is not collision-free. The arm subgoals are thus represented as the desired end-effector 3D locations and parameterized pushing actions. We hypothesize that the pushing actions can be replaced by other types of parameterized actions (e.g. grasping and pulling). More details about algorithm description, network structure, training procedure and hyperparameters can be found on our website.

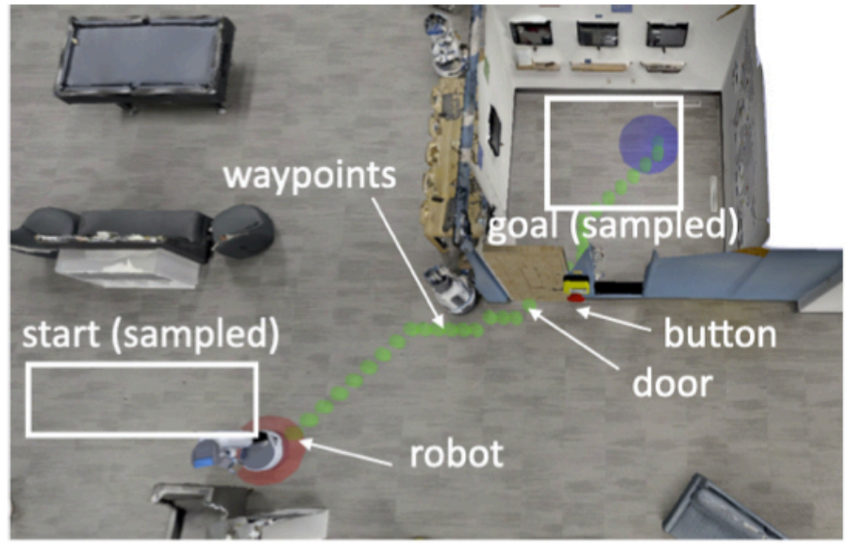
ReLMoGen



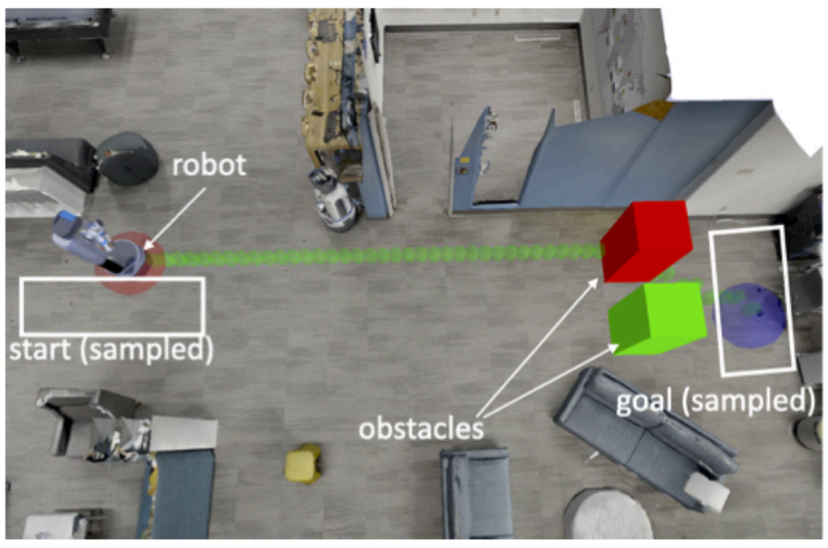
(a) `PointNav`



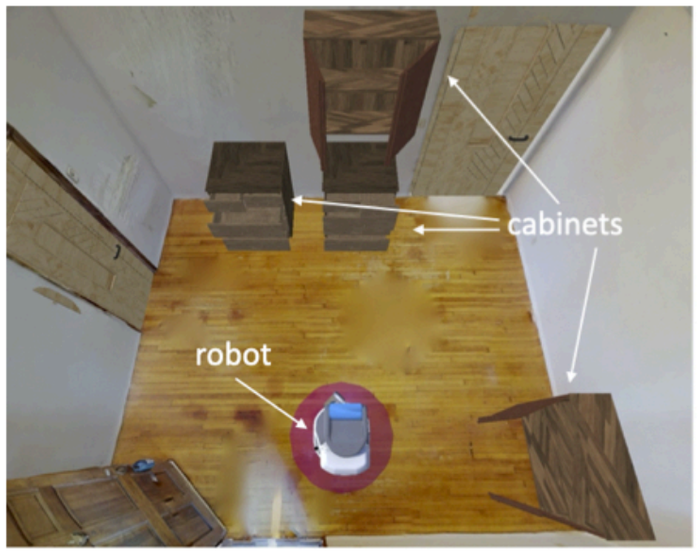
(b) `TabletopReachM`



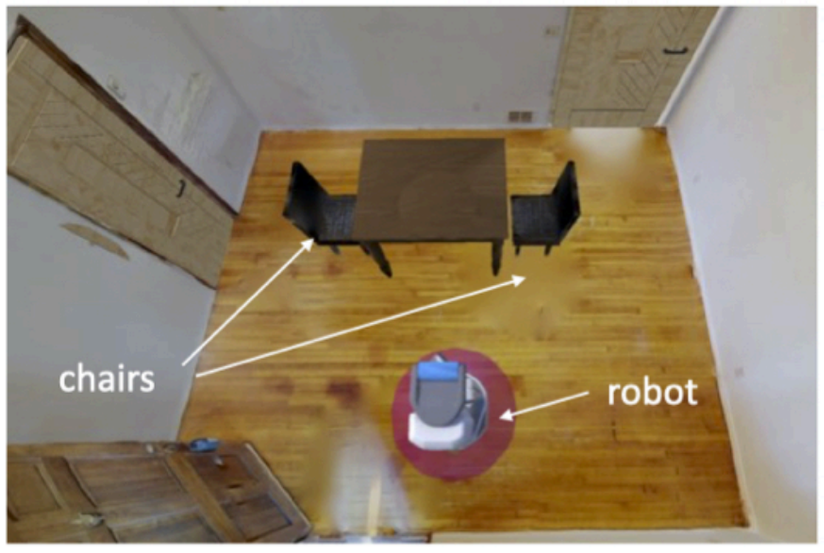
(c) `PushDoorNav`, `ButtonDoorNav`



(d) `InteractiveObstaclesNav`



(e) `ArrangeKitchenMM`



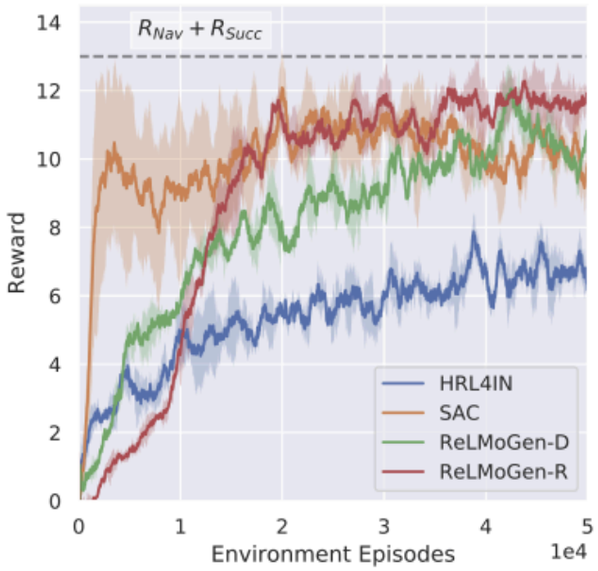
(f) `ArrangeChairMM`

ReLMoGen

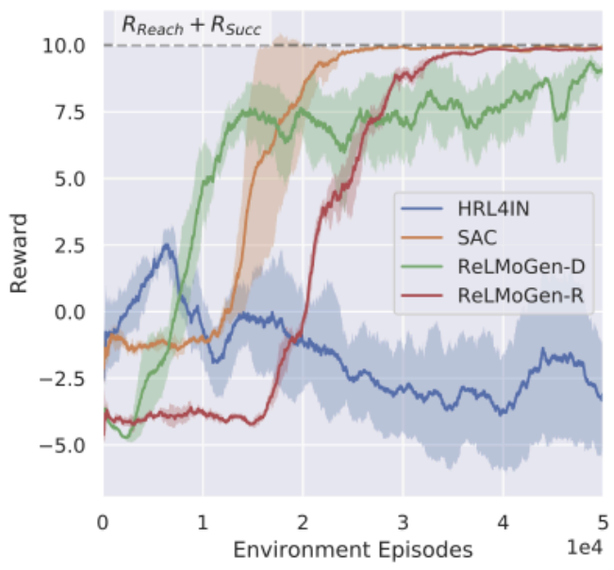
Task	PointNav		TabletopReachM	ArrangeKitchenMM		ArrangeChairMM	
Metric	SPL	SR	SR	# Closed 5°/5 cm	# Closed 10°/10 cm	# Closed 5 cm	# Closed 10 cm
ReLMoGen-D (ours)	0.57/0.02/0.58	0.68/0.01/0.68	0.95/0.02/0.96	**4.35/1.20/5.72**	**6.10/1.05/7.3**	**0.21/0.03/0.23**	**0.36/0.06/0.43**
ReLMoGen-R (ours)	**0.63/0.09/0.67**	**0.72/0.06/0.77**	**1.0/0.0/1.0**	3.43/0.61/3.94	4.91/0.51/5.25	0.06/0.10/0.17	0.11/0.20/0.34
HRL4IN [17]	0.27/0.01/0.28	0.33/0.01/0.35	0.09/0.07/0.19	3.0/0.23/3.3	4.67/0.20/4.95	0.0/0.0/0.0	0.0/0.0/0.0
SAC (joint vel.) [19, 10]	0.60/0.04/0.65	0.60/0.04/0.65	**1.0/0.0/1.0**	3.42/0.19/3.6	4.95/0.29/5.24	0.0/0.0/0.0	0.0/0.0/0.0
OAC (joint vel.) [41]	0.45/0.01/0.46	0.46/0.01/0.47	**1.0/0.0/1.0**	1.99/0.61/2.60	3.55/0.48/4.02	0.0/0.0/0.0	0.0/0.0/0.0

Task	PushDoorNav		ButtonDoorNav		InteractiveObstaclesNav	
Metric	SPL	SR	SPL	SR	SPL	SR
ReLMoGen-D (ours)	0.36/0.36/0.72	0.41/0.40/0.80	0.42/0.17/0.57	0.50/0.19/0.66	0.54/0.011/0.55	0.58/0.02/0.60
ReLMoGen-R (ours)	**0.80/0.02/0.83**	**0.97/0.02/0.99**	**0.51/0.15/0.61**	**0.73/0.21/0.87**	**0.76/0.01/0.87**	**0.79/0.11/0.91**
HRL4IN [17]	0.0/0.0/0.0	0.0/0.0/0.0	0.0/0.0/0.0	0.0/0.0/0.0	0.0/0.0/0.0	0.0/0.0/0.0
SAC (joint vel.) [19, 10]	0.0/0.0/0.0	0.0/0.0/0.0	0.00/0.01/0.01	0.01/0.01/0.01	0.50/0.36/0.84	0.51/0.37/0.87
OAC (joint vel.) [41]	0.0/0.0/0.0	0.0/0.0/0.0	0.00/0.00/0.01	0.01/0.00/0.01	0.00/0.00/0.01	0.01/0.01/0.01

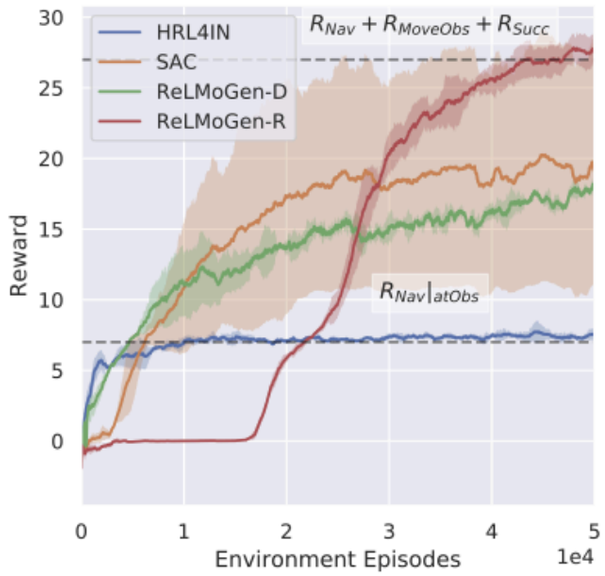
ReLMoGen



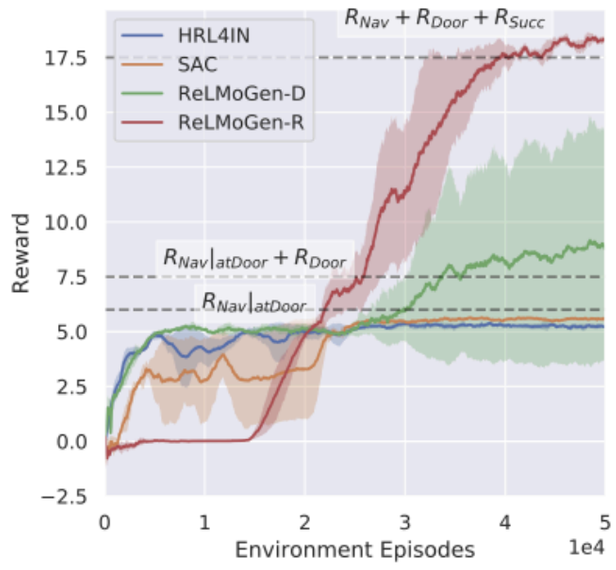
(a) PointNav



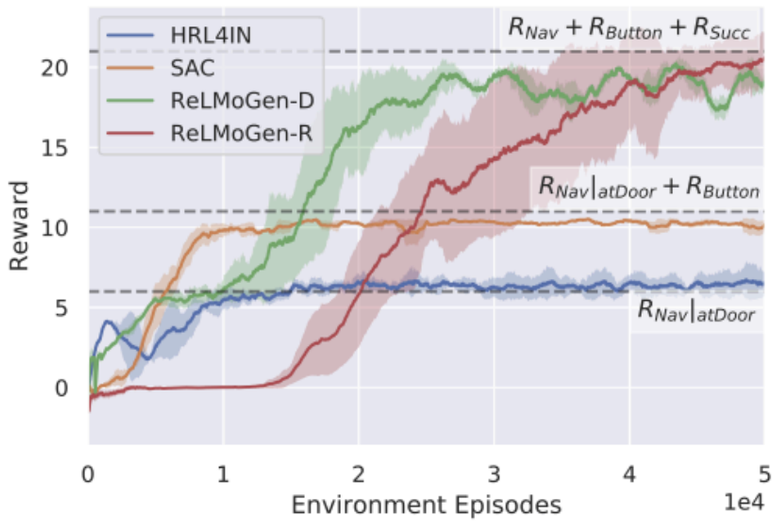
(b) TabletopReachM



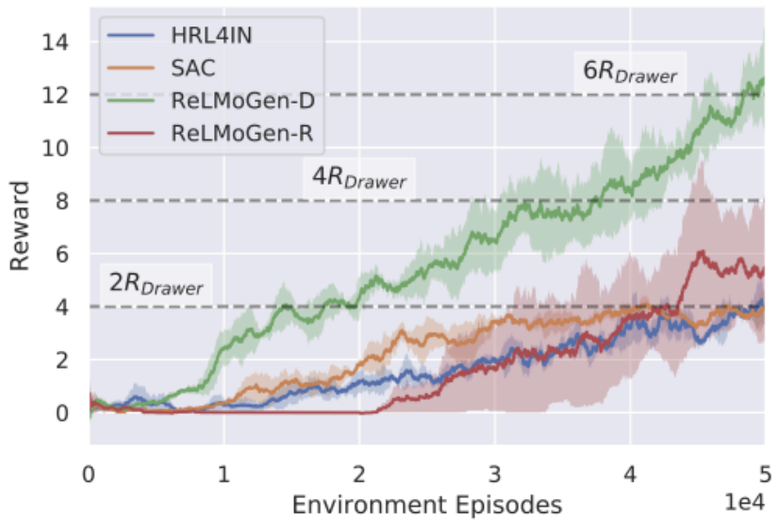
(c) InteractiveObstaclesNav



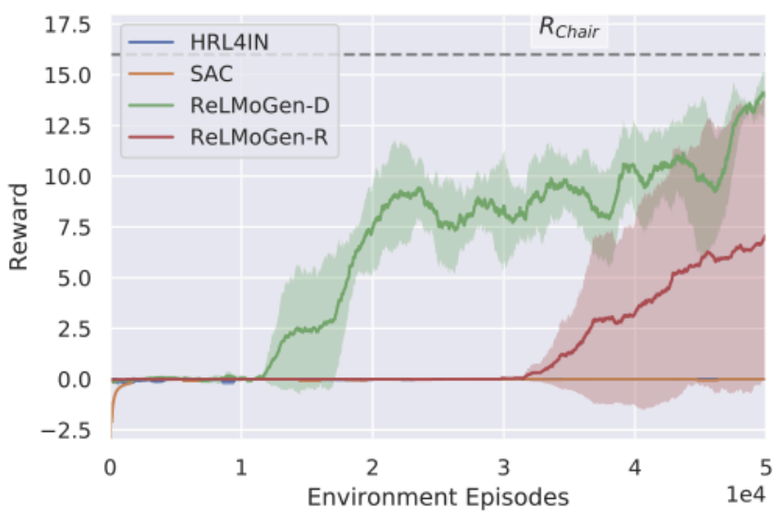
(d) PushDoorNav



(e) ButtonDoorNav

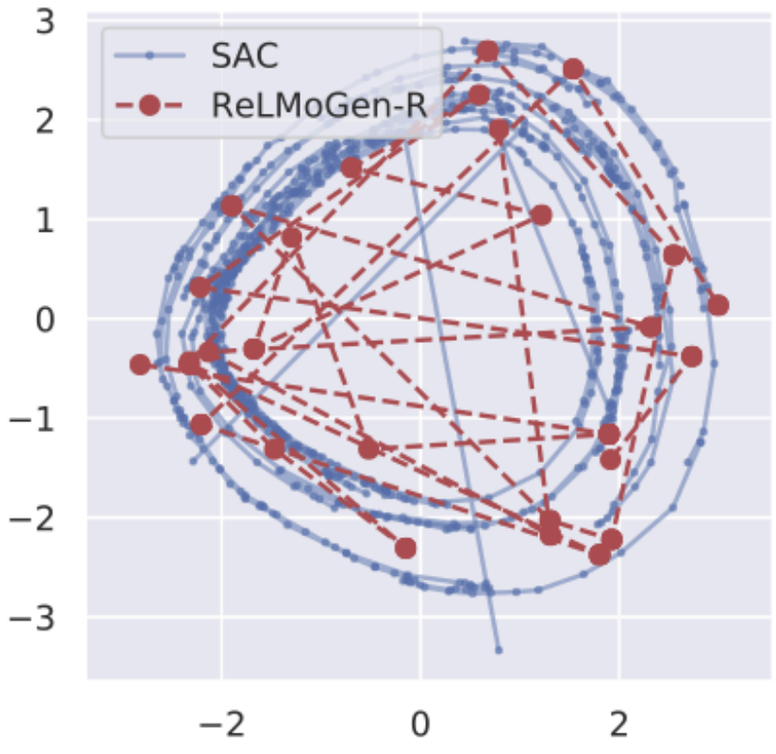


(f) ArrangeKitchenMM

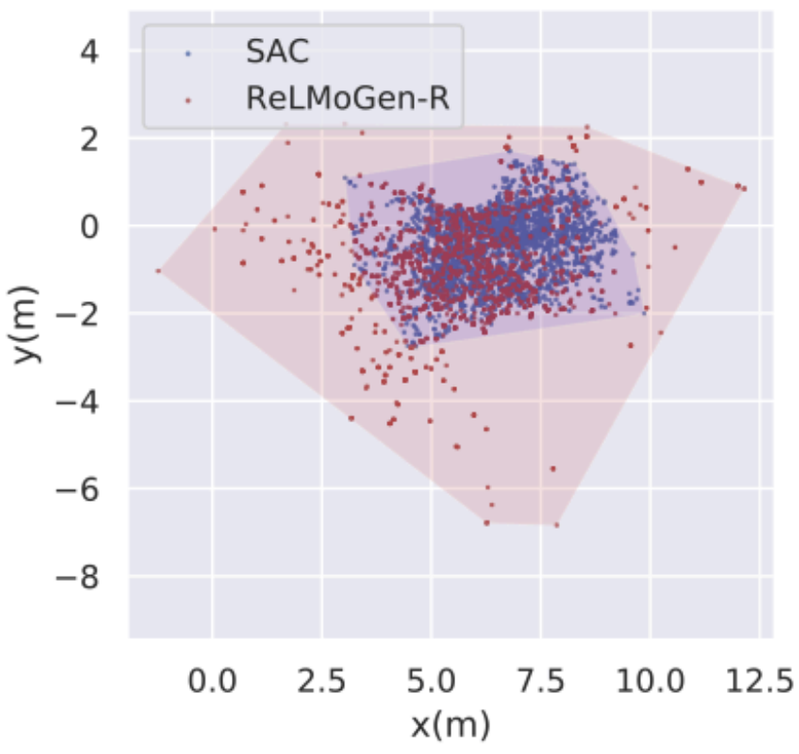


(g) ArrangeChairMM

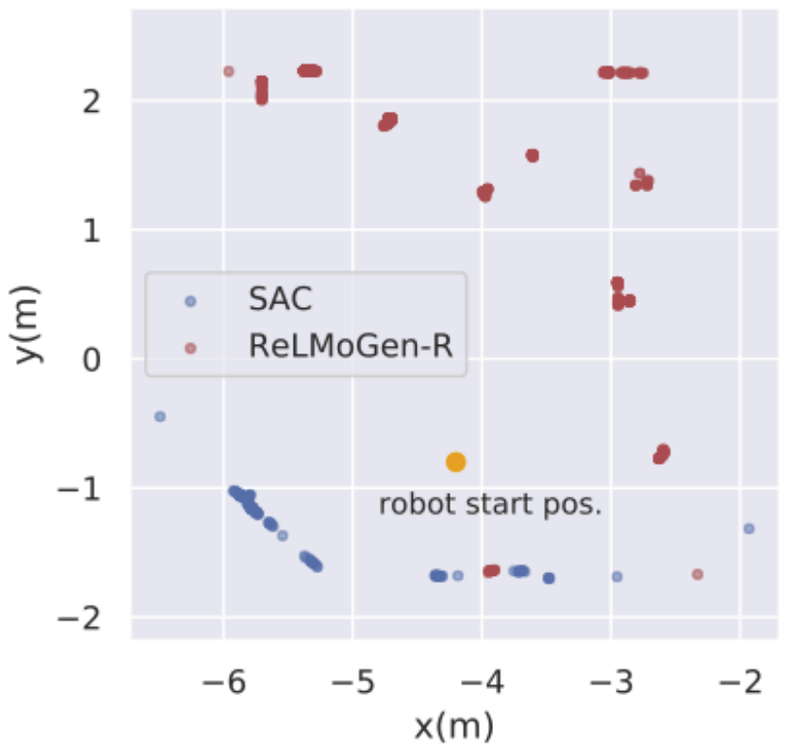
ReLMoGen



(a) Latent State Space



(b) Cartesian Space



(c) Interaction Heatmap

ReLMoGen

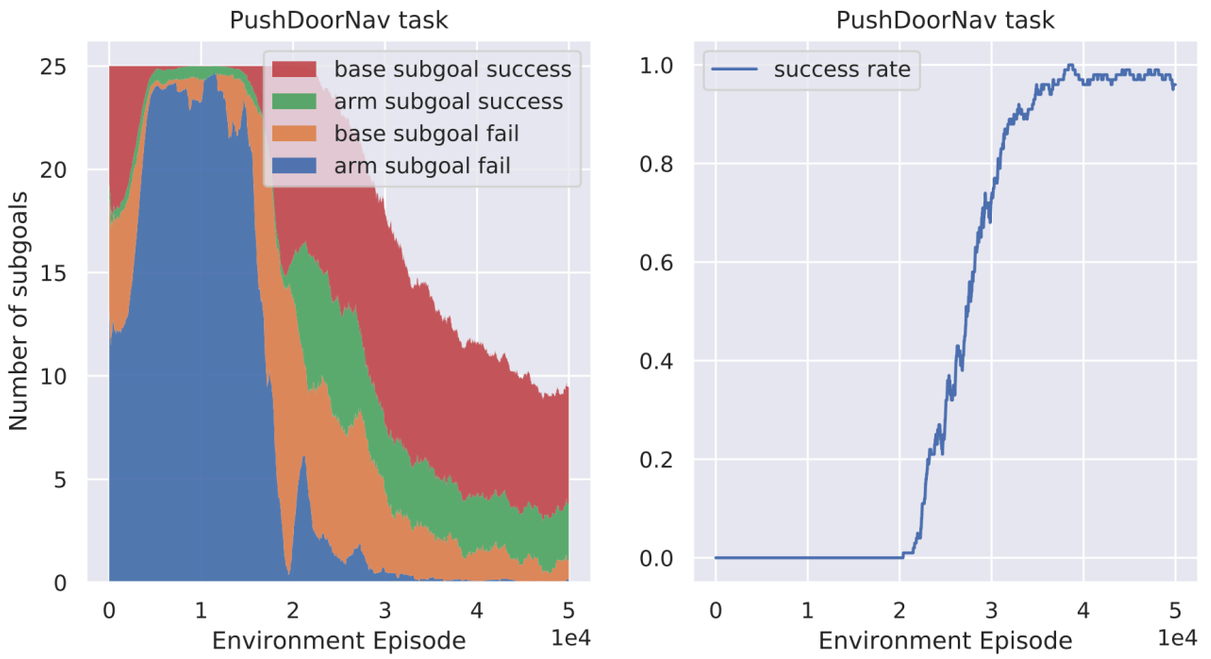
Base MP	Arm MP	Success rate
RRT-Connect	**RRT-Connect**	**0.99**
RRT-Connect	Lazy PRM	1.0 (+**0.01**)
Lazy PRM	RRT-Connect	0.99 (+**0.0**)
Lazy PRM	Lazy PRM	1.0 (+**0.01**)

(a) PushDoorNav Task

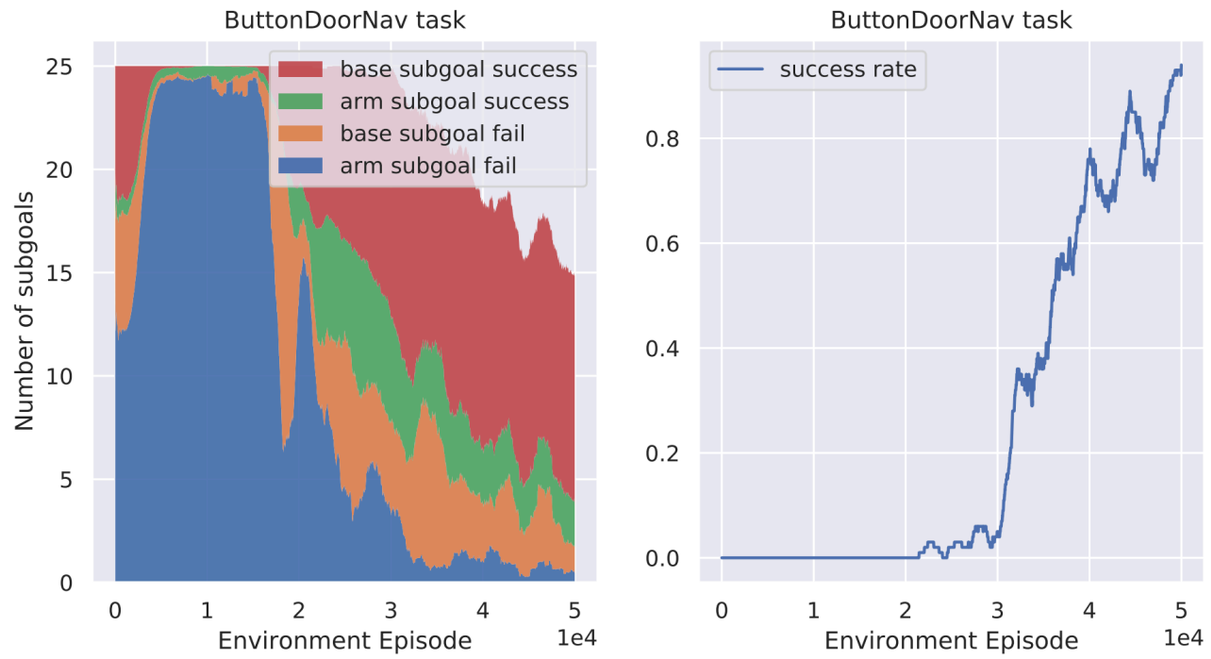
Base MP	Arm MP	# Closed (10°/10 cm)
RRT-Connect	**RRT-Connect**	**5.25**
RRT-Connect	Lazy PRM	5.0 (−**0.25**)
Lazy PRM	RRT-Connect	5.18 (−**0.07**)
Lazy PRM	Lazy PRM	5.09 (−**0.16**)

(b) ArrangeKitchenMM Task

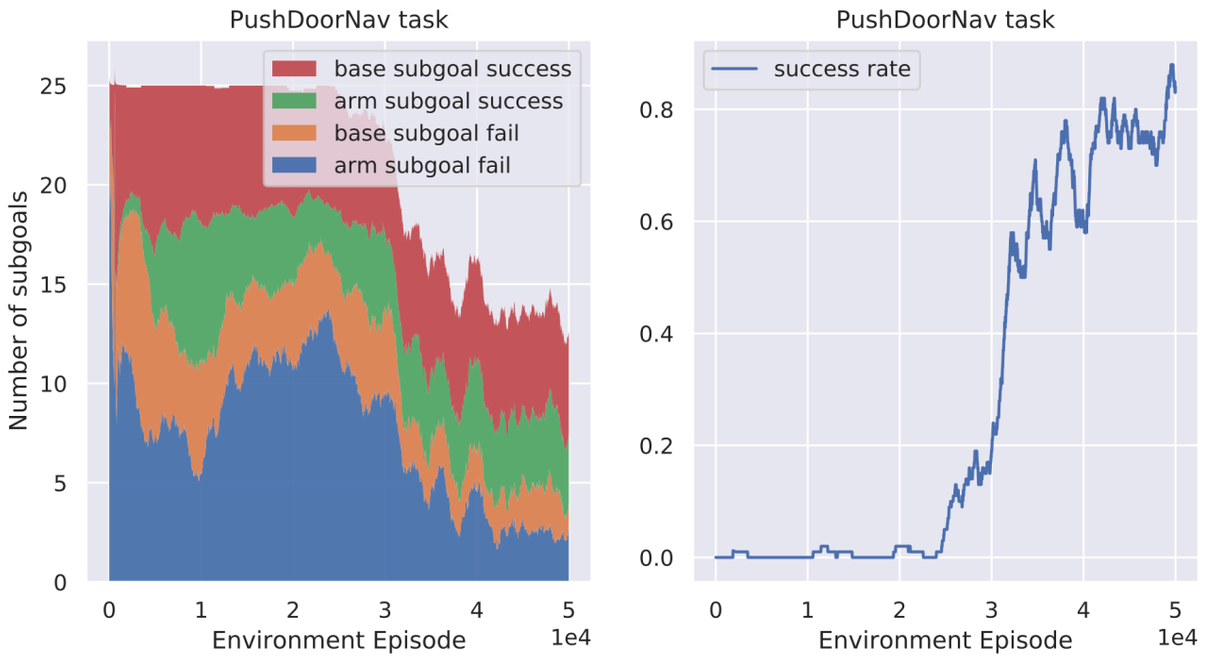
ReLMoGen



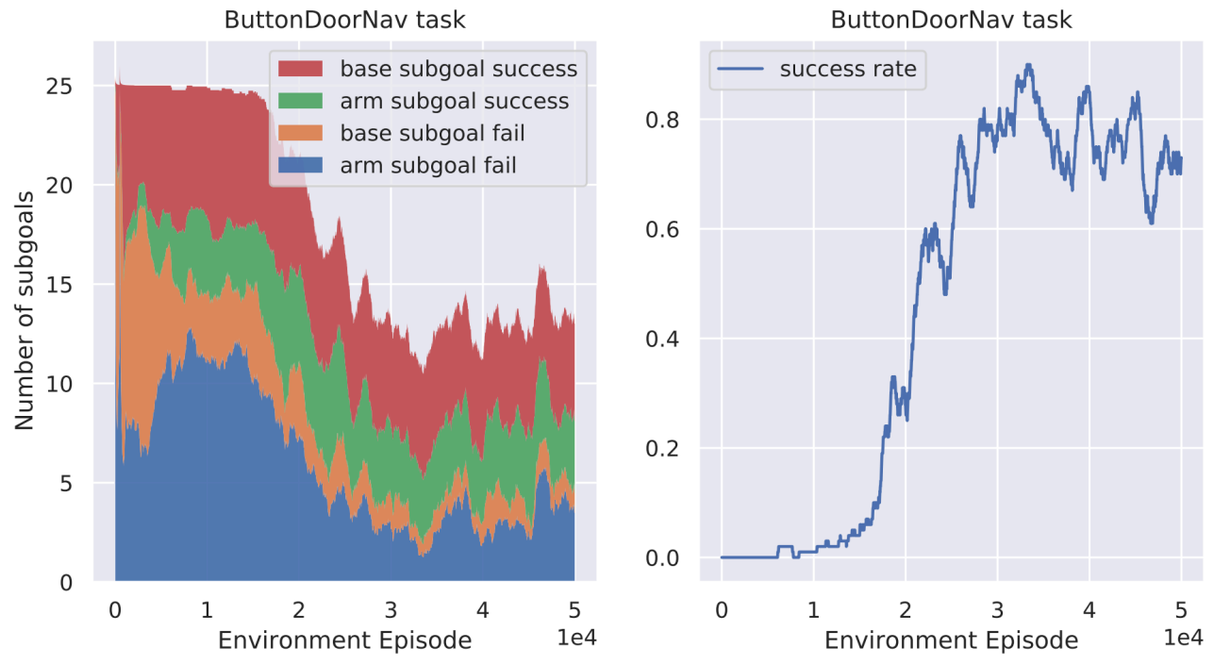
(a) ReLMoGen-R on `PushDoorNav`



(b) ReLMoGen-R on `ButtonDoorNav`



(c) ReLMoGen-D on `PushDoorNav`



(d) ReLMoGen-D on `ButtonDoorNav`

Thanks!