BEHAVIOR: Benchmark for Everyday Household Activities in Virtual, Interactive, and Ecological Environments

Sanjana Srivastava[*♠], Chengshu Li[*♠], Michael Lingelbach[*♡] Roberto Martín-Martín[*♠]
Fei Xia[♣], Kent Vainio[♠], Zheng Lian[♠], Cem Gokmen[♠], Shyamal Buch[♠]
C. Karen Liu[♠★], Silvio Savarese[♠★], Hyowon Gweon[◇★], Jiajun Wu[♠★], Li Fei-Fei[♠★]

Department of Computer Science[♠], Neurosciences IDP[♡], Electrical Engineering[♣], Psychology[◇]
Institute for Human-Centered AI (HAI)[★]
Stanford University

Abstract: We introduce BEHAVIOR, a benchmark for embodied AI with 100 activities in simulation, spanning a range of everyday household chores such as cleaning, maintenance, and food preparation. These activities are designed to be realistic, diverse and complex, aiming to reproduce the challenges that agents must face in the real world. Building such a benchmark poses three fundamental difficulties for each activity: definition (it can differ by time, place, or person), instantiation in a simulator, and evaluation. BEHAVIOR addresses these with three innovations. First, we propose a predicate logic-based description language for expressing an activity's initial and goal conditions, enabling generation of diverse instances for any activity. Second, we identify the simulator-agnostic features required by an underlying environment to support BEHAVIOR, and demonstrate in one such simulator. Third, we introduce a set of metrics to measure task progress and efficiency, absolute and relative to human demonstrators. We include 500 human demonstrations in virtual reality (VR) to serve as the human ground truth. Our experiments demonstrate that even state-of-the-art embodied AI solutions struggle with the level of realism, diversity, and complexity imposed by the activities in our benchmark. We make BEHAVIOR publicly available at behavior.stanford.edu to facilitate and calibrate the development of new embodied AI solutions.

Keywords: Embodied AI, Benchmarking, Household Activities

1 Introduction

Embodied AI refers to the study and development of artificial agents that can perceive, reason, and interact with the environment with the capabilities and limitations of a physical body. Recently, significant progress has been made in developing solutions to embodied AI problems such as (visual) navigation [1–5], interactive Q&A [6–10], instruction following [11–15], and manipulation [16–22].

To calibrate the progress, several lines of pioneering efforts have been made towards benchmarking embodied AI in simulated environments, including Rearrangement [23, 24], TDW Transport Challenge [25], VirtualHome [26], ALFRED [11], Interactive Gibson Benchmark [27], MetaWorld [28], and RLBench [29], among others [30–32]. These efforts are inspiring, but their activities represent only a fraction of challenges that humans face in their daily lives. To develop artificial agents that can eventually perform and assist with everyday activities with human-level robustness and flexibility, we need a comprehensive benchmark with activities that are more **realistic**, **diverse**, and **complex**.

But this is easier said than done. There are three major challenges that have prevented existing benchmarks from accommodating more realistic, diverse, and complex activities:

- Definition: Identifying and defining meaningful activities for benchmarking;
- Realization: Developing simulated environments that realistically support such activities;

[*] indicates equal contribution
correspondence to {sanjana2,chengshu,mjlbach,robertom}@stanford.edu

5th Conference on Robot Learning (CoRL 2021), London, UK.

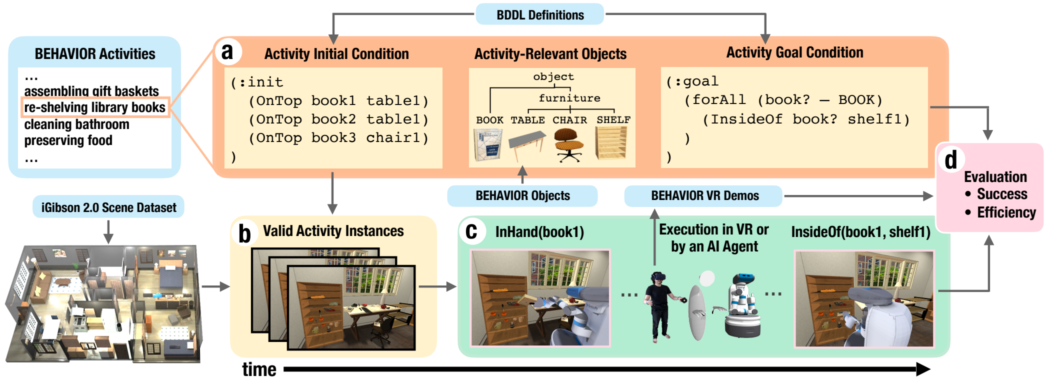

Figure 1: **Benchmarking Embodied AI with BEHAVIOR**: ⓐ We define 100 realistic household activities from the American Time Use Survey [34] and define them with a set of relevant objects, organized with WordNet [35], and logic-symbolic initial and goal conditions in BDDL (Sec. 4). ⓑ We provide an implementation of BEHAVIOR in iGibson 2.0 that generates potentially infinite diverse activity instances in realistic home scenes using the definition. ⓒ AI agents perform the activities in simulation through continuous physical interactions of an embodied avatar with the environment. Humans can perform the same activities in VR. BEHAVIOR includes a dataset of 500 successful VR demonstrations. ⓓ Changes in the scene are continuously mapped to their logic-symbolic equivalent representation in BDDL and checked against the goal condition; we provide intermediate success scores, metrics on agent's efficiency, and a human-centric metric relative to the demonstrations.

- Evaluation: Defining success and objective metrics for evaluating performance.

We propose **BEHAVIOR** (Fig. 1)–**B**enchmark for **E**veryday **H**ousehold **A**ctivities in **V**irtual, **I**nteractive, and ec**O**logical envi**R**onments, addressing the three key challenges with three technical innovations. First, we introduce BEHAVIOR Domain Definition Language (BDDL), a representation adapted from predicate logic that maps simulated states to semantic symbols. It allows us to define 100 activities as initial and goal conditions, and enables generation of potentially infinite initial states and solutions for achieving the goal states. Second, we facilitate its realization by listing environment-agnostic functional requirements for realistic simulation. With proper engineering, BEHAVIOR can be implemented in many existing environments; we discuss a fully functional instantiation in iGibson 2.0 [33] in this paper including the necessary object models (1217 models of 391 categories). Third, we provide a comprehensive set of metrics to evaluate agent performance in terms of success and efficiency. To make evaluation comparable across diverse activities, scenes, and instances, we propose a set of metrics relative to demonstrated human performance on each activity, and provide a large-scale dataset of 500 human demonstrations (1077.7 min) in virtual reality, which serve as ground truth for evaluation and may also facilitate developing imitation learning solutions.

BEHAVIOR's 100 activities are realistic, diverse, and complex. They are often performed by humans in their homes (e.g., cleaning, packing or preparing food) and require long-horizon solutions for changing not only the position of multiple objects but also their internal states or texture (e.g., temperature, wetness or cleanliness levels). As we demonstrate by experimentally evaluating the performance of two state-of-the-art reinforcement learning algorithms (Section 7), these properties make BEHAVIOR activities extremely challenging for existing solutions. By presenting well-defined challenges beyond the capabilities of current solutions, BEHAVIOR can serve as a unifying benchmark that guides the development of embodied AI.

2 Related Work

Benchmarks and datasets have played a critical role in recent impressive advances in AI, particularly computer vision. Image [36–39] and video datasets [40–45] enable study and development of solutions for important research questions by providing both training data and fair comparison. These datasets, however, are passive observations that are not well-suited for development of embodied AI that must control and understand the consequences of their own actions.

Benchmarks for Embodied AI: Although real-world challenges [46–53] provide the ultimate testbed for embodied AI agents, benchmarks in simulated environments serve as useful alternatives with several advantages; simulation enables faster, safer learning, and supports more reproducible,

		Mobile manipulation								Static manipulation								Navigation	
		BEHAVIOR	AI2THOR Vis. Room Rearr.	TDW Transport	Rearrangement T5 (Habitat)	ManipulaTHOR ArmPointNav	Interactive Gibson Benchmark	VirtualHome	ALFRED	Rearrangement T2 (OCRTOC)	IKEA Furniture Assembly	RLBench	Metaworld	Robosuite	SoftGym	DeepMind Control Suite	OpenAIGym	Habitat 1.0	Gibson
Realism	Activity selections reflect human behavior	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
	Kinematics, dynamics	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Continuous extended states (e.g. temp., wetness)	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
	Changing flexible materials	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗
	Realistic action execution	✓	✗	✓	✗	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Scenes reconstructed from real homes	✓	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓
Diversity	# Activities	100	1	1	1	1	2	**549**	7	5	100	50	1	5	10	28	8	2	3
	Infinite scene-agnostic instantiation	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	N/A
	Object categories	391	118	50	YCB	150	5	**509**	84	12	73+	28	7	10	4	4	4	Matterport	N/A
	Object models	**1217**	118	112	YCB	150	152		84	101 + YCB	73+	28	80	10	4	4	4	N/A	N/A
	Scenes / Rooms	**15 / 100**	**- / 120**	**15 / 90-120**	55 static / -	- / 30	10 / -	7 / -	- / 120	1 / -	1 / -	1 / -	1 / -	1 / -	1 / -	1 / -	1 / -	Matterport + Gibson	572 static
Complexity	Activity length[2](steps)	**300-20000**	<100	100-1000	100-1000	<100	100-1000	<100	<100	100-1000	<1000	<100	<100	<100	<100	<100	<100	<100	100-1000
	Objs. per activity	**3-34**	5	7-9	2-5	2-3	10	1-24	2	5-10	1-2	1-2	1	1-3	1-3	1-3	1	0-1	N/A
	Benchmark focus: Task-Planning and/or Control	TP+C	TP	TP+C	TP+C	TP+C	C	TP	TP	TP+C	C	TP+C	C	C	C	C	C	C	C
	Diff. state changes required per activity (see A.2)	**2-8**	4	4	4	2	1-3	1-7	2-3	1	1-3	1-4	4	1	1-3	1-2	1-2	1	1
	# Human VR demos	**500**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

[1] Estimate of a near-optimal, e.g. human, execution of the activity given the platform's action space

Table 1: **Comparison of Embodied AI Benchmarks:** BEHAVIOR activities are exceptionally realistic due to their grounding in human population time use [34] and realistic simulation (sensing, actuation, changes in environment) in iGibson 2.0. The activity set is diverse in topic, objects used, scenes done in, and state changes required. The diversity is reinforced by the ability to generate infinite new instances scene-agnostically. BEHAVIOR activities are complex enough to reflect real-world housework: many decision steps and objects in each activity. This makes BEHAVIOR uniquely well-suited to benchmark task-planning and control, and it is the only one to include human VR demonstrations (see Table A.1 for more detail).

accessible, and fair evaluation. However, in order to serve as a meaningful proxy for real-world performance, simulation benchmarks need to achieve high levels of 1) **realism** (in the activities, the models, the sensing and actuation of the agent), 2) **diversity** (of scenes, objects and activities benchmarked), and 3) **complexity** (length, number of objects, required skills and state changes). Below we review existing benchmarks based on these three criteria (see Table 1 for a summary).

Benchmarks for *visual navigation* [54, 55] provide high levels of visual realism and diversity of scenes, but they often lack interactivity or diversity of activities. The Interactive Gibson Benchmark [27] trades off some visual realism for physically realistic object manipulation in order to benchmark interactive visual navigation. While benchmarks for *stationary manipulation* [56, 29, 28, 30, 57, 31, 32] fare well on physical realism, they commonly fall short on diversity (of scenes, objects, tasks) and complexity (often having simple activities that take a few seconds). Benchmarks for *instruction following* [11, 26] provide diversity of scenes, objects and possible changes of the environment, but with low levels of complexity; the horizon of the activities is shorter as the agents decide among a discrete set of predefined action primitives with full access to the state of the world.

Closer to BEHAVIOR, a recent group of benchmarks has focused on *rearrangement tasks* [23–25] in realistic simulation environments with diverse scenes. The initial Rearrangement position paper [23] poses critical questions such as how to define embodied AI tasks and measure solution quality. Importantly however, most household activities go far beyond the scope of rearrangement (see comparison in Fig. A.2). While such focus can inspire new solutions for solving rearrangement tasks, these solutions may not generalize to activities that require more than physical manipulation of object coordinates. Indeed, the majority of household activities involve other state changes (cooking, washing, etc.) (Fig. A.2, [34]). BEHAVIOR therefore incorporates 100 activities that humans actually spend time on at home [34] (Sec. 3). To express such diverse activites in a common language, we present a novel logic-symbolic representation that defines activities as initial and goal states, inspired by but distinct from the Planning Domain Definition Language (PDDL) [58] (see Sec. 4). These definitions yield in principle infinite instances per activity and accept any meaningful solution. We implement activity-independent metrics including a human-centric metric normalized to human performance; to facilitate comparison and development of new solutions, we also present a dataset of 500 successful VR demonstrations.

3 BEHAVIOR: Benchmarking Realistic, Diverse, Complex Activities

Building on the advances led by existing benchmarks, BEHAVIOR aims to reach new levels of realism, diversity, and complexity by using household activities as a domain for benchmarking AI. See Table 1 for comparisons between BEHAVIOR and existing benchmarks.

Realism in BEHAVIOR Activities: To effectively benchmark embodied AI agents in simulation, we need realistic activities that pose similar challenges to those in the real world. BEHAVIOR achieves this by using a data-driven approach to identify activities that approximate the true distribution of real household activities. To this end, we use the American Time Use Survey (ATUS, [34]): A survey from the U.S. Bureau of Labor Statistics on how Americans spend their time. BEHAVIOR activities come from, and are distributed similarly to, the full space of simulatable activities in ATUS (see Fig. A.2). The use of an independently curated source of real-world activities is a unique strength of BEHAVIOR as a benchmark that reflects natural behaviors of a large population.

BEHAVIOR also achieves realism by simulating these activities in reconstructions of real-world homes. We use iGibson 2.0, a simulation environment with realistic physics simulation from the Bullet [59] physics engine and high-quality virtual sensor signals (see Fig. A.7), which includes 15 ecological, fully interactive 3D models of real-world homes with furniture layouts that approximate their real counterparts. These scenes are further populated with object models created by professional artists from the new BEHAVIOR Object dataset, which includes 1217 models of 391 categories grounded in the WordNet [35] taxonomy. The dataset covers a data-driven selection of activity-related objects (see Fig. A.8). Figs. A.10 and A.9 illustrate examples of objects and taxonomic arrangement. The 100 BEHAVIOR activities, visualized in Fig. A.1, go beyond comparable benchmarks that evaluate a few hand-picked activities in less realistic setups (see Table 1 Realism). iGibson 2.0 also provides a wide variety of realistic simulated robots that have real-world counterparts, e.g. LoCoBot, Quadrotor, Fetch, the last of which we can use to fulfill the BEHAVIOR activities (see Sec. 5).

Diversity in BEHAVIOR Activities: Benchmarks with diverse activities demand generalizable solutions. In real-world homes, agents encounter a range of activities that differ in 1) the capabilities required for achieving them, 2) the environments in which they occur (e.g., scenes, objects), and 3) the initial states of a particular scene. BEHAVIOR presents extensive diversity in all these dimensions. We include 100 activities that require a wide variety of state changes (e.g., moving objects, soaking materials, cleaning surfaces, heating/freezing food) demanding a broad set of agent capabilities (see Fig A.2). To reflect the diversity in the ways humans encounter, understand, and accomplish these activities, we provide two example definitions per activity.

BDDL, our novel representation for activity definition, allows new valid instances to be sampled from each definition, providing potentially infinite number of instances per activity. The resulting instances vary over scene, object models, and configuration, supported by implementation in iGibson 2.0 and BEHAVIOR Object dataset. Related benchmarks focus on fewer tasks, mostly limited to kinematic state changes and with scene- or position-constant instantiation (see Table 1 Diversity).

Complexity in BEHAVIOR Activities: Beyond diversity across activities, BEHAVIOR also raises the complexity of the activities themselves by benchmarking full household activities that parallel the length (number of steps an agent needs), the number of objects involved, and the number of required capabilities of real-world chores (see Fig. A.3, comparison in Table 1 Complexity). Compared to activities in existing benchmarks, these activities are very long-horizon with some requiring several thousand steps (even for humans in VR; see Fig. A.12), involve more objects (avg. 10.5), and require a heterogeneous set of capabilities (range: 2 - 8) to change various environment states.

4 Defining Realistic, Diverse, and Complex Household Activities with BDDL

BEHAVIOR challenges embodied AI agents to achieve a diverse set of complex long-horizon household activities through physical interactions in a realistically simulated home environment.

Adopting the common formalism of partially-observable Markov decision processes (POMDP), each activity has a state space $\mathcal{S}$ (see more details in A.3.2).

We define an *activity* τ as two sets of states, $\tau = \{S_{\tau,0}, S_{\tau,g}\}$, where $S_{\tau,0}$ is a set of possible initial states and $S_{\tau,g}$ is a set of acceptable goal states. In an *activity instance*, the agent must change the world state from some concrete $s_0 \in S_{\tau,0}$ to any $s_g \in S_{\tau,g}$. However, describing activities in the physical state space generates scene- or pose-specific definitions (e.g., [23, 30, 29]) that are far more specific than how humans represent these activities, limiting the diversity and complexity of existing embodied AI benchmarks. To overcome this, we introduce *BEHAVIOR Domain Definition Language* (BDDL), a predicate logic-based language that establishes a symbolic state representation built on predefined, meaningful predicates grounded in simulated physical states; its variables and constants represent object categories from the BEHAVIOR object dataset. Each activity is defined in BDDL as an initial and goal condition parametrizing sets of possible initial states and satisfactory goal states $\bar{S}_{\tau,0}$ and $\bar{S}_{\tau,g}$. BDDL predicates create symbolic counterparts of the physical state, $\bar{S}$ (see Fig. 2).

BDDL overcomes limitations that hinder diversity through two mechanisms: first, an initial condition maps to infinite physical states in diverse scenes. Second, a goal condition detects all semantically satisfactory solutions. By contrast, other benchmarks support either infinite distinct instantiations but only in one scene per definition, because they sample from hard-coded regions; or instantiation in multiple scenes, but not infinitely because object poses are hard-coded on furniture objects in those scenes. BEHAVIOR is the only benchmark with both. BEHAVIOR also includes a systematic generation pipeline (Sec. A.3.3) enabling unlimited definitions per activity to formalize the subjectivity of household activities. We include 200 definitions and 100 instances in simulation (Sec. 5). BEHAVIOR is thus the only benchmark equipped to formalize unlimited human-defined versions of an activity and create practically infinite unique instantiations in any scene. Finally, BEHAVIOR has purely declarative definitions of initial and goal condition, whereas some benchmarks provide imperative plans for getting from initial to goal [26]. The declarative nature creates a true test of an agent's capability of task planning.

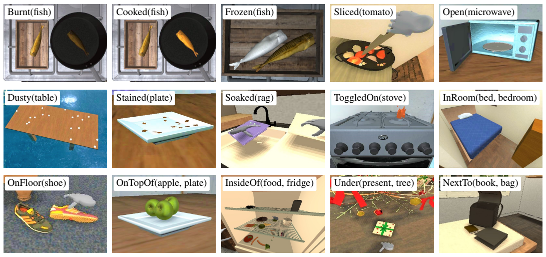

Figure 2: **Unary and Binary Predicates in BDDL:** We represent object states and relationships to other objects based on their kinematics, temperature, wetness level and other physical and functional properties, enabling a diverse and complex set of realistic activities

5 Instantiating BEHAVIOR in a Realistic Physics Simulator

While BEHAVIOR is not bounded to any specific simulation environment, there are a set of functional requirements to be able to simulate BEHAVIOR activities: 1) maintain an object-centric representation (object identities enriched with properties and states), 2) simulate physical forces and motion, and generate virtual sensor signals (images), 3) simulate additional properties per object (e.g. temperature, soak level, cleanliness level) necessary for BEHAVIOR activities, 4) implement functionality to **generate** valid instances based on the literals defining an activity's initial condition, e.g., instantiating an object `insideOf` another, and 5) implement functionality to **evaluate** the atomic formulae relevant to the goal condition, e.g. checking whether an object is `cooked` or `onTopOf` another.

While BEHAVIOR activities are not tailored to a specific embodiment, we propose two concrete bodies to fulfill the activities offering different action spaces (see Fig. 1): a *bimanual humanoid* avatar (24 degrees of freedom, DoF), and a *Fetch robot* (12/13 DoF), both capable of navigating, grasping and interacting with the hand(s). Humans in VR embody the bimanual humanoid.

Because it models a real-world robot, agents trained with the Fetch embodiment could be directly tested with a real-world version of the hardware (see discussion on sim2real in Sec. A.8). Both agents receive sensor signals from the on-board virtual sensors, and perform actions at 30 Hz.

We provide a fully functional implementation of BEHAVIOR using iGibson 2.0, a new version of the open-source simulation environment iGibson that fulfills the requirements above. iGibson 2.0 provides an object-centric representation with additional properties, support for sources of heat and water, dust and stain particles, and changes in object appearance based on extended states. We also implement the two embodiments mentioned above and a set of action primitives inspired by [25, 55, 60, 24] to facilitate solution prototyping and task-planning research. The primitives are executing sequences of low-level actions resulting from a motion planning process (bilateral RRT* [61]) to `navigateTo`, `grasp`, `placeOnTop`, `placeInside`, `open`, and `close` the object given as argument. Further details can be found in Sec. A.4 and in [33]. Our implementation of BEHAVIOR in iGibson 2.0 goes beyond the capabilities of existing benchmarks and amplifies realism, diversity, and complexity (see Table 1).

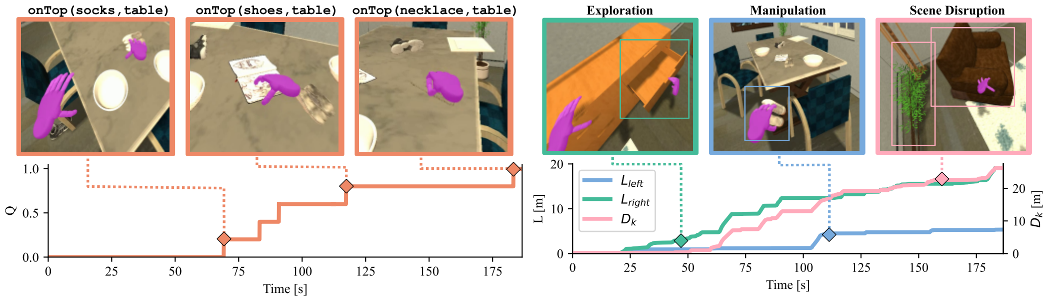

Figure 3: **Evaluation of human performance in `collect_misplaced_items`:** (*Left*) success score, Q; (*Right*) efficiency metrics: kinematic disarrangement, (D_k, dotted), hand interaction displacement (L_{right}, green, and L_{left}, blue); frames at the top depict significant events detected by the metrics; the success score detects the completion of activity-relevant steps; exploration, manipulation and scene disruption events are captured by the efficiency metrics that provide complementary information about the performance of the agent.

6 Evaluation Metrics: Success, Efficiency and Human-Centric Metric

BEHAVIOR provides evaluation metrics to quantify the performance of an embodied AI solution. Extending prior metrics suggested for Rearrangement [23], we propose a primary metric based on success and several secondary metrics for characterizing efficiency.

Primary Metric – Success Score Q: The main goal of an embodied AI agent in BEHAVIOR is to perform an activity successfully (i.e., all logical expressions in the goal condition are met). A binary definition of success, however, only signals the end of a successful execution and cannot assess interim progress. To provide more guidance to agents and enable comparisons of partial solutions, we propose **success score** as the primary metric, defined as the **maximum fraction of satisfied goal literals in a ground solution to the goal condition** at each step. More formally:

Given an activity τ with goal state set $\bar{S}_{\tau,g}$, its goal condition can be flattened to a set C of conjunctions C_i of ground literals l_{j_i}.

For any $C_i \in C$, if all $l_{j_i} \in C_i$ are true then the goal condition is satisfied (see A.3.2 for definitions and technical details on flattening), i.e. for some current environment state s, we have $\bigvee\limits_{C_i} \bigwedge\limits_{l_{j_i}} l_{j_i} = \texttt{True} \implies s \in \bar{S}_{\tau,g}$. We compute the fraction of literals l_{j_i} that are `True` for each C_i, and define the overall success score by taking the maximum: $Q = \max\limits_{C} \frac{|\{l_{j_i}|l_{j_i}=\texttt{True}\}|}{|C_i|}$, where $|\cdot|$ is set cardinality.

An activity is complete when all literals in *at least one* C_i of its goal condition are satisfied, achieving $Q = 1$ (100%). Fig. 3, left, depicts time evolution of Q during an activity execution. Q extends the fraction of objects in acceptable poses proposed as metric in [23], generalized to any type of activity.

Secondary Metrics – Efficiency: Beyond success, efficiency is critical to evaluation; a successful solution in real-world tasks may be ineffective if it takes too long or causes scene disruption. We propose six secondary metrics that complement the primary metric (see Fig. 3, right, for examples):

- *Simulated time*, T_{sim}: Accumulated time in simulation during execution as the number of simulated steps times the average simulated time per step. T_{sim} is independent of the computer used.
- *Kinematic disarrangement*, D_K: Displacement caused by the agent in the environment. This can be *accumulated* over time, or *differential*, i.e. computed between two time steps, e.g. initial, final.
- *Logical disarrangement*, D_L: Amount of changes caused by the agent in the logical state of the environment. This can be *accumulated* over time or *differential* between two time steps.
- *Distance navigated*, L_{body}: Accumulated distance traveled by the agent's base body. This metric evaluates the efficiency of the agent in navigating the environment.
- *Displacement of hands*, L_{left} and L_{right}: Accumulated displacement of each of the agent's hands while in contact with another object for manipulation (i.e., grasping, pushing, etc). This metric evaluates the efficiency of the agent in its interaction with the environment.

These efficiency metrics above can be quantified in absolute units (e.g., distance, time) for scene- and activity-specific comparisons (**general efficiency**). To enable fair comparisons cross diverse activities in BEHAVIOR, we also propose normalization relative to human performance (**human-centric**

efficiency); given a human demonstration for an activity instance in VR, each secondary metric can be expressed as a *fraction of the maximum human performance* on that metric.

For this purpose, we present the BEHAVIOR Dataset of Human Demonstrations with 500 successful demonstrations of BEHAVIOR activities in VR (1077.7 min). Humans are immersed in iGibson 2.0, controlling the same embodiment used by the AI agents (details in Sec. A.6). The dataset includes a complete record of human actions including manipulation, navigation, and gaze tracking data (Fig. A.12, Fig. A.14, and Fig. A.16), supporting analysis and subactivity segmentation (Fig. A.11). Sec. A.6.2 presents a comprehensive analysis of these data; we quantify human performance in BEHAVIOR efficiency metrics (see Fig. A.12), and Fig. A.13 provides a further decomposition of room occupancy and hand usage across each BEHAVIOR activity. To our knowledge, this is the largest available dataset of human behavior in VR; these data can facilitate development of new solutions for embodied AI (e.g., imitation learning) and also support studies of human cognition, planning, and motor control in ecological environments.

7 Evaluating Reinforcement Learning in BEHAVIOR

In this section, we aim to experimentally demonstrate the challenges imposed by BEHAVIOR's realism, diversity, and complexity on state-of-the-art embodied AI solutions. BEHAVIOR is a benchmark for all kinds of embodied AI methods. Here, we evaluate two reinforcement learning (RL) algorithms that have excelled on simpler embodied AI tasks [62, 63, 21, 64–68]: Soft-Actor Critic (SAC [16]) and Proximal-Policy Optimization (PPO [17]). We use SAC to train policies in the original low-level continuous action space of the agent, and PPO for experiments using our implemented action primitives (for details on the agents, see Sec. 5). Due to limited computational resources, we evaluate on the 12 most simple activities (by distinct state changes involved) using the bimanual humanoid embodiment.

Reward is given by our staggered success score Q. We use as input a subset of the realistic agent's observations, RGB, depth and proprioception (excluding LiDAR, segmentation, etc.). Sec. A.7 includes more experimental details.

Results in the original activities: The first row of Table 2 shows the results of SAC (mean Q at the end of training for 3 seeds) on the original 12 activities with the standard setup: realistic robot actions and onboard sensing. Even for these "simpler" activities, BEHAVIOR is too great a challenge: the training agents do not fulfill any predicate in the goal condition ($Q = 0$). In the following, we will analyze how each dimension of difficulty (realism, diversity, complexity) contributes to these results.

		bringingInWood	collectMisplacedItems	movingBoxesToStorage	organizingFileCabinet	throwingAwayLeftovers	puttingDishesAway	puttingleftoversAway	re-shelvingLibraryBooks	layingTileFloors	settingUpCandles	pickingUpTrash	storingFood
	Q^{ca}	0	0	0	0	0	0	0	0	0	0	0	0
complexity	Q^{ap}	0.13	0	0	0	0	0	0	0	0	0	0	0
	$SR^{ca}@T-1\,\mathrm{s}$	1	1	1	0	0	0	0	0	1	1	0.97	1
	$SR^{ca}@T-2\,\mathrm{s}$	1	0.07	1	0	0	0	0	0	1	1	0.01	0
	$SR^{ca}@T-3\,\mathrm{s}$	1	0.21	1	0	0	0	0	0	1	0.01	0	0
	$SR^{ca}@T-10\,\mathrm{s}$	0	0	0	0	0	0	0	0	0	0	0	0
realism	$Q^{ca}_{FullObs}$	0	0	0	0	0	0	0	0	0	0	0	0
	$Q^{ap}_{FullObs}$	0.20	0.02	0.49	0	0	0	0	0.13	0	0.09	0	0
	Q^{ap}_{NoPhys}	0.92	0.47	0.73	0	0.32	0.55	0.44	0.04	0	0.27	0	0.32
	$Q^{ap}_{NoPhys,FullObs}$	1.0	0.95	0.83	0	0.56	0.94	0.55	0.56	0	0.5	0.67	1.0

Table 2: **Evaluation of state-of-the-art RL algorithms on BEHAVIOR** *Fully realistic, diverse and complex (row 1):* SAC [16] for visuomotor continuous actions (superindex ca) performs poorly in all activities; *Complexity analysis (rows 2-6):* reducing complexity (horizon) with temporally extended action primitives (superindex ap and gray cells, trained with PPO [17]) or by starting few seconds away from a goal state, lead to some non-zero success rate (SR). *Realism analysis (rows 7-10):* Only by reducing realism in observations and physics, and complexity through action primitives, RL achieves significant success in a handful of the activities.

Effect of complexity (activity length): First, we evaluate the impact of the activity complexity (time length) on performance. We begin with an RL algorithm using our implemented action primitives based on motion planning. These temporally extended actions effectively shorten the horizon and length of the activity. The results of training with PPO are depicted in the second row of Table 2. Even in these simpler conditions, agents fail in all but one activity (`bringingInWood`, $Q = 0.13$). In a second oracle-driven experiment, we take a successful human demonstration for each activity from the BEHAVIOR Dataset and set a state a few seconds before its successful execution at T as the activity initial state. We train agents with SAC: rows 3 to 6 of Table 2 show the mean success rate (SR, full accomplishment of the activity) in 100 evaluation episodes for the final policy resulting from training with three different random seeds (Q starts here close to 1 and is less informative). Even when starting 1 s away from a goal state, most agents fail. A few do better, but their performance decreases quickly as we start further away from the successful execution, being zero for all activities at $10\,\mathrm{s}$. This indicates that the long horizon of BEHAVIOR activities is in fact a paramount challenge for RL. We hypothesize that Embodied AI solutions with a hierarchical structure such as hierarchical-RL or task-and-motion-planning (TAMP) may help to overcome the challenges of high complexity (length) of the BEHAVIOR activities [69–72].

Effect of realism (in sensing and actuation): In a third experiment, we evaluate how much the realism in actuation and sensing affects the performance of embodied AI solutions. We train agents with continuous motion control (SAC), and motion primitives (PPO) assuming full-observability of the state, with results in Tables 2 (rows 7-8, subindex $FullObs$). Even with full observability, the complexity dominates policies in the original action space and they fail entirely. For policies selecting among action primitives, there is partial success in only five activities, indicating that perception is part of the difficulty in BEHAVIOR. To evaluate the effect of realistic actuation, we train an agent using action primitives that execute without physics simulation, achieving their expected outcome (e.g. grasp an object, or place it somewhere). Tables 2 (row 9-10, subindex $noPhys$) shows the results, also in combination with unrealistic full-observability. We observe that without the difficulties of realistic physics and actuation, the learning agents achieve an important part of most activities, accomplishing consistently two of them ($Q = 1$) when full-observability of the state is also granted. This indicates that the generation of the correct actuation is a critical challenge for embodied AI solutions, even when they infer the right next step at the task-planning level, supporting the importance of benchmarks with realistically action execution over predefined action outcomes.

Effect of diversity (in activity instance and objects): Another cause of the poor performance of robot learning solutions in the 12 BEHAVIOR activities may be the high diversity in multiple dimensions, such as scenes, objects, and initial states. This diversity forces embodied AI solutions to generalize to all possible conditions. In a second experiment, we evaluate the effect of BEHAVIOR's diversity on performance. To present diversity across activities while alleviating their complexity, we train RL agents to complete five single-literal activities involving only one or two objects. Note that these activities are not part of BEHAVIOR. We evaluate training with RL (SAC) for each activity under diverse instantiations: initialization of the activity (object poses) and object instances. The results are shown in Table 3, where we report Q. First, we train without any diversity as baseline to understand the ground complexity of the single-literal activities. All agents achieve success. Then, we evaluate how well the RL policies train for a diverse set of instances of the activities, first changing objects' initial pose, then changing the object. Performance in all activities decreases rapidly, especially in `sliced` and `stained`. These experiments indicate that the diversity in BEHAVIOR goes beyond what current RL algorithms can handle even in simplified activities, and poses a challenge for generalization in embodied AI.

Diversity in...						
object poses	object instances	`ontop`	`sliced`	`soaked`	`stained`	`cooked`
✗	✗	1	0.15	1	1	1
✓	✗	0.825	0	0.935	0.28	0.66
✓	✓	0.46	0	0.925	0.11	0.265

Table 3: **Evaluation of the effect of BEHAVIOR's diversity:** Results of training agents with SAC [16] in single-predicate activities of increasing diversity; Even in these simple activities, performance degrades quickly indicating that current state-of-the-art cannot cope with the dimensions of diversity spanned in BEHAVIOR

8 Conclusion and Future Work

We presented BEHAVIOR, a novel benchmark for embodied AI solutions of household activities. BEHAVIOR presents 100 realistic, diverse and complex activities with a new logic-symbolic representation, a fully functional simulation-based implementation, and a set of human-centric metrics based on the performance of humans on the same activities in VR. The activities push the state-of-the-art in benchmarking adding new types of state changes that the agent needs to be able to cause, such as cleaning surfaces or changing object temperatures. Our experiments with two state-of-the-art RL baselines shed light on the challenges presented by BEHAVIOR's level of realism, diversity and complexity. BEHAVIOR will be open-source and free to use; we hope it facilitates participation and fair access to research tools, and paves the way towards a new generation of embodied AI.

Acknowledgments

We would like to thank Bokui Shen, Xi Jia Zhou, Jim Fan, Manasi Sharma, and Sidhart Krishnan for comments, ideas, and support in data collection. This work is in part supported by Toyota Research Institute (TRI), ARMY MURI grant W911NF-15-1-0479, Samsung, Amazon, and Stanford Institute for Human-Centered AI (SUHAI). S. S. and C. L. are supported by SUHAI Award #202521. S. S. is also supported by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP). R. M-M. and S. B. are supported by SAIL TRI Center – Award # S-2018-28-Savarese-Robot-Learn. S. B. is also supported by a National Defense Science and Engineering Graduate (NDSEG) fellowship, SAIL TRI Center – Award # S-2018-27-Niebles, and SAIL TRI Center – Award # TRI Code 44. S. S. and S. B. are supported by a Department of Navy award (N00014-16-1-2127) issued by the Office of Naval Research (ONR). F. X. is supported by the Qualcomm Innovation Fellowship and Stanford Graduate Fellowship. This article solely reflects the opinions and conclusions of its authors and not any other entity.

References

[1] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi. Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 3357–3364. IEEE, 2017.

[2] N. Hirose, F. Xia, R. Martín-Martín, A. Sadeghian, and S. Savarese. Deep visual mpc-policy learning for navigation. *IEEE Robotics and Automation Letters*, 4(4):3184–3191, 2019.

[3] E. Wijmans, A. Kadian, A. Morcos, S. Lee, I. Essa, D. Parikh, M. Savva, and D. Batra. Dd-ppo: Learning near-perfect pointgoal navigators from 2.5 billion frames. *arXiv preprint arXiv:1911.00357*, 2019.

[4] S. Gupta, J. Davidson, S. Levine, R. Sukthankar, and J. Malik. Cognitive mapping and planning for visual navigation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2616–2625, 2017.

[5] S. Bansal, V. Tolani, S. Gupta, J. Malik, and C. Tomlin. Combining optimal control and learning for visual navigation in novel environments. In *Conference on Robot Learning*, pages 420–429. PMLR, 2020.

[6] A. Das, S. Datta, G. Gkioxari, S. Lee, D. Parikh, and D. Batra. Embodied question answering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–10, 2018.

[7] L. Yu, X. Chen, G. Gkioxari, M. Bansal, T. L. Berg, and D. Batra. Multi-target embodied question answering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6309–6318, 2019.

[8] A. Das, G. Gkioxari, S. Lee, D. Parikh, and D. Batra. Neural modular control for embodied question answering. In *Conference on Robot Learning*, pages 53–62. PMLR, 2018.

[9] A. Mousavian, A. Toshev, M. Fišer, J. Košecká, A. Wahid, and J. Davidson. Visual representations for semantic target driven navigation. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8846–8852. IEEE, 2019.

[10] K. Marino, M. Rastegari, A. Farhadi, and R. Mottaghi. Ok-vqa: A visual question answering benchmark requiring external knowledge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3195–3204, 2019.

[11] M. Shridhar, J. Thomason, D. Gordon, Y. Bisk, W. Han, R. Mottaghi, L. Zettlemoyer, and D. Fox. Alfred: A benchmark for interpreting grounded instructions for everyday tasks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10740–10749, 2020.

[12] J. Fu, A. Korattikara, S. Levine, and S. Guadarrama. From language to goals: Inverse reinforcement learning for vision-based instruction following. In *International Conference on Learning Representations*, 2018.

[13] P. Anderson, Q. Wu, D. Teney, J. Bruce, M. Johnson, N. Sünderhauf, I. Reid, S. Gould, and A. van den Hengel. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[14] X. Wang, Q. Huang, A. Celikyilmaz, J. Gao, D. Shen, Y.-F. Wang, W. Y. Wang, and L. Zhang. Reinforced cross-modal matching and self-supervised imitation learning for vision-language navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6629–6638, 2019.

[15] D. Fried, R. Hu, V. Cirik, A. Rohrbach, J. Andreas, L.-P. Morency, T. Berg-Kirkpatrick, K. Saenko, D. Klein, and T. Darrell. Speaker-follower models for vision-and-language navigation. *arXiv preprint arXiv:1806.02724*, 2018.

[16] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International Conference on Machine Learning*, pages 1861–1870. PMLR, 2018.

[17] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[18] N. Watters, L. Matthey, M. Bosnjak, C. P. Burgess, and A. Lerchner. Cobra: Data-efficient model-based rl through unsupervised object discovery and curiosity-driven exploration. *arXiv preprint arXiv:1905.09275*, 2019.

[19] Y. Wu, W. Yan, T. Kurutach, L. Pinto, and P. Abbeel. Learning to manipulate deformable objects without demonstrations. *arXiv preprint arXiv:1910.13439*, 2019.

[20] A. Billard and D. Kragic. Trends and challenges in robot manipulation. *Science*, 364(6446), 2019.

[21] O. M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, et al. Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, 39(1):3–20, 2020.

[22] F. Xia, C. Li, R. Martín-Martín, O. Litany, A. Toshev, and S. Savarese. ReLMoGen: Leveraging motion generation in reinforcement learning for mobile manipulation. In *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

[23] D. Batra, A. X. Chang, S. Chernova, A. J. Davison, J. Deng, V. Koltun, S. Levine, J. Malik, I. Mordatch, R. Mottaghi, M. Savva, and H. Su. Rearrangement: A challenge for embodied ai, 2020.

[24] L. Weihs, M. Deitke, A. Kembhavi, and R. Mottaghi. Visual room rearrangement. *arXiv preprint arXiv:2103.16544*, 2021.

[25] C. Gan, S. Zhou, J. Schwartz, S. Alter, A. Bhandwaldar, D. Gutfreund, D. L. Yamins, J. J. DiCarlo, J. McDermott, A. Torralba, et al. The threedworld transport challenge: A visually guided task-and-motion planning benchmark for physically realistic embodied ai. *arXiv preprint arXiv:2103.14025*, 2021.

[26] X. Puig et al. Virtualhome: Simulating household activities via programs. In *IEEE CVPR*, 2018.

[27] F. Xia, W. B. Shen, C. Li, P. Kasimbeg, M. E. Tchapmi, A. Toshev, R. Martín-Martín, and S. Savarese. Interactive gibson benchmark: A benchmark for interactive navigation in cluttered environments. *IEEE Robotics and Automation Letters*, 5(2):713–720, 2020.

[28] T. Yu, D. Quillen, Z. He, R. Julian, K. Hausman, C. Finn, and S. Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on Robot Learning*, pages 1094–1100. PMLR, 2020.

[29] S. James, Z. Ma, D. Rovick Arrojo, and A. J. Davison. Rlbench: The robot learning benchmark & learning environment. *IEEE Robotics and Automation Letters*, 2020.

[30] Y. Zhu, J. Wong, A. Mandlekar, and R. Martín-Martín. robosuite: A modular simulation framework and benchmark for robot learning. *arXiv preprint arXiv:2009.12293*, 2020.

[31] Y. Tassa, Y. Doron, A. Muldal, T. Erez, Y. Li, D. d. L. Casas, D. Budden, A. Abdolmaleki, J. Merel, A. Lefrancq, et al. Deepmind control suite. *arXiv preprint arXiv:1801.00690*, 2018.

[32] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

[33] C. Li, F. Xia, R. Martín-Martín, M. Lingelbach, S. Srivastava, B. Shen, K. E. Vainio, C. Gokmen, G. Dharan, T. Jain, A. Kurenkov, K. Liu, H. Gweon, J. Wu, L. Fei-Fei, and S. Savarese. igibson 2.0: Object-centric simulation for robot learning of everyday household tasks. In *5th Annual Conference on Robot Learning*, 2021. URL https://openreview.net/forum?id=2uGN5jNJROR.

[34] U.S. Bureau of Labor Statistics. American Time Use Survey. https://www.bls.gov/tus/, 2019.

[35] G. A. Miller. WordNet: a lexical database. *Communications of the ACM*, 38(11):39–41, 1995.

[36] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[37] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[38] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010.

[39] R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, S. Chen, Y. Kalantidis, L.-J. Li, D. A. Shamma, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International journal of computer vision*, 123(1):32–73, 2017.

[40] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361. IEEE, 2012.

[41] R. Goyal, S. Ebrahimi Kahou, V. Michalski, J. Materzynska, S. Westphal, H. Kim, V. Haenel, I. Fruend, P. Yianilos, M. Mueller-Freitag, et al. The" something something" video database for learning and evaluating visual common sense. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5842–5850, 2017.

[42] G. A. Sigurdsson, A. Gupta, C. Schmid, A. Farhadi, and K. Alahari. Actor and observer: Joint modeling of first and third-person videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7396–7404, 2018.

[43] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox. Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes. *arXiv preprint arXiv:1711.00199*, 2017.

[44] R. Martín-Martín, M. Patel, H. Rezatofighi, A. Shenoi, J. Gwak, E. Frankel, A. Sadeghian, and S. Savarese. Jrdb: A dataset and benchmark of egocentric robot visual perception of humans in built environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[45] F. Caba Heilbron, V. Escorcia, B. Ghanem, and J. Carlos Niebles. Activitynet: A large-scale video benchmark for human activity understanding. In *Proceedings of the ieee conference on computer vision and pattern recognition*, pages 961–970, 2015.

[46] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, E. Osawa, and H. Matsubara. Robocup: A challenge problem for ai. *AI magazine*, 18(1):73–73, 1997.

[47] T. Wisspeintner, T. Van Der Zant, L. Iocchi, and S. Schiffer. Robocup@home: Scientific competition and benchmarking for domestic service robots. *Interaction Studies*, 10(3):392–426, 2009.

[48] L. Iocchi, D. Holz, J. Ruiz-del Solar, K. Sugiura, and T. Van Der Zant. Robocup@ home: Analysis and results of evolving competitions for domestic and service robots. *Artificial Intelligence*, 229:258–281, 2015.

[49] M. Buehler, K. Iagnemma, and S. Singh. *The DARPA urban challenge: autonomous vehicles in city traffic*, volume 56. springer, 2009.

[50] E. Krotkov, D. Hackett, L. Jackel, M. Perschbacher, J. Pippine, J. Strauss, G. Pratt, and C. Orlowski. The darpa robotics challenge finals: Results and perspectives. *Journal of Field Robotics*, 34(2):229–240, 2017.

[51] N. Correll, K. E. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. M. Romano, and P. R. Wurman. Analysis and observations from the first amazon picking challenge. *IEEE Transactions on Automation Science and Engineering*, 15(1):172–188, 2016.

[52] C. Eppner, S. Höfer, R. Jonschkowski, R. Martín-Martín, A. Sieverling, V. Wall, and O. Brock. Lessons from the amazon picking challenge: four aspects of building robotic systems. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 4831–4835, 2017.

[53] M. A. Roa, M. Dogar, C. Vivas, A. Morales, N. Correll, M. Gorner, J. Rosell, S. Foix, R. Memmesheimer, F. Ferro, et al. Mobile manipulation hackathon: Moving into real world applications. *IEEE Robotics & Automation Magazine*, pages 2–14, 2021.

[54] F. Xia, A. R. Zamir, Z. He, A. Sax, J. Malik, and S. Savarese. Gibson env: Real-world perception for embodied agents. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[55] Manolis Savva*, Abhishek Kadian*, Oleksandr Maksymets*, Y. Zhao, E. Wijmans, B. Jain, J. Straub, J. Liu, V. Koltun, J. Malik, D. Parikh, and D. Batra. Habitat: A Platform for Embodied AI Research. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

[56] Y. Lee, E. S. Hu, Z. Yang, A. Yin, and J. J. Lim. Ikea furniture assembly environment for long-horizon complex manipulation tasks. *arXiv preprint arXiv:1911.07246*, 2019.

[57] X. Lin, Y. Wang, J. Olkin, and D. Held. Softgym: Benchmarking deep reinforcement learning for deformable object manipulation. In *Conference on Robot Learning*, 2020.

[58] D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, and D. Wilkins. Pddl - the planning domain definition language. Technical report, Technical Report 1165, Yale Computer Science, 1998.(CVC Report 98-003), 1998.

[59] E. Coumans and Y. Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. *Technical Report*, 2016.

[60] E. Kolve et al. AI2-THOR: An interactive 3d environment for visual ai. *arXiv preprint arXiv:1712.05474*, 2017.

[61] M. Jordan and A. Perez. Optimal bidirectional rapidly-exploring random trees. Technical Report MIT-CSAIL-TR-2013-021, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, August 2013. URL http://dspace.mit.edu/bitstream/handle/1721.1/79884/MIT-CSAIL-TR-2013-021.pdf.

[62] R. Martín-Martín, M. A. Lee, R. Gardner, S. Savarese, J. Bohg, and A. Garg. Variable impedance control in end-effector space: An action space for reinforcement learning in contact-rich tasks. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1010–1017. IEEE, 2019.

[63] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.

[64] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine. Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *arXiv preprint arXiv:1709.10087*, 2017.

[65] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4(26), 2019.

[66] I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, et al. Solving rubik's cube with a robot hand. *arXiv preprint arXiv:1910.07113*, 2019.

[67] T. Haarnoja, S. Ha, A. Zhou, J. Tan, G. Tucker, and S. Levine. Learning to walk via deep reinforcement learning. *arXiv preprint arXiv:1812.11103*, 2018.

[68] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Pérez. Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[69] A. G. Barto and S. Mahadevan. Recent advances in hierarchical reinforcement learning. *Discrete event dynamic systems*, 13(1):41–77, 2003.

[70] A. S. Vezhnevets, S. Osindero, T. Schaul, N. Heess, M. Jaderberg, D. Silver, and K. Kavukcuoglu. Feudal networks for hierarchical reinforcement learning. In *International Conference on Machine Learning*, pages 3540–3549. PMLR, 2017.

[71] C. Li, F. Xia, R. Martín-Martín, and S. Savarese. Hrl4in: Hierarchical reinforcement learning for interactive navigation with mobile manipulators. In *Conference on Robot Learning*, pages 603–616. PMLR, 2020.

[72] C. R. Garrett, T. Lozano-Pérez, and L. P. Kaelbling. Ffrob: An efficient heuristic for task and motion planning. In *Algorithmic Foundations of Robotics XI*, pages 179–195. Springer, 2015.

[73] Y. Gil. Description logics and planning. *AI Magazine*, 26(2):73–84.

[74] A. Olivares-Alarcos, D. Besler, A. Khamis, P. Goncalves, M. K. Habib, , J. Bermejo-Alonso, M. Barreto, M. Diab, J. Rosell, J. Quintas, J. Olszewska, H. Nakawala, E. Pignaton, A. Gyrard, S. Borgo, G. Alenya, M. Beetz, and H. Li. A review and comparison of ontology-based approaches to robot autonomy. *The Knowledge Engineering Review*, 34(0):1–29, 2019.

[75] A. Aho and J. Ullman. *Foundations of Computer Science*. W. H. Freeman, 1992.

[76] Upwork Global Inc. Upwork. https://www.upwork.com/, 2021. Accessed: 2021-06-16.

[77] wikiHow, Inc. wikihow. https://www.wikihow.com, 2021. Accessed: 2021-06-16.

[78] Google Alphabet. Blockly. https://developers.google.com/blockly/, 2021. Accessed: 2021-06-16.

[79] B. Shen, F. Xia, C. Li, R. Martín-Martín, L. Fan, G. Wang, S. Buch, C. D'Arpino, S. Srivastava, L. P. Tchapmi, M. E. Tchapmi, K. Vainio, L. Fei-Fei, and S. Savarese. iGibson, a Simulation Environment for Interactive Tasks in Large Realistic Scenes, 2020.

[80] F. Xiang, Y. Qin, K. Mo, Y. Xia, H. Zhu, F. Liu, M. Liu, H. Jiang, Y. Yuan, H. Wang, et al. SAPIEN: A simulated part-based interactive environment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11097–11107, 2020.

[81] HTC Corporation. HTC Vive Pro Eye. https://www.vive.com/us/product/vive-pro-eye/, 2021. Accessed: 2021-06-16.

[82] R. Kothari, Z. Yang, C. Kanan, R. Bailey, J. B. Pelz, and G. J. Diaz. Gaze-in-wild: A dataset for studying eye and head coordination in everyday activities. *Scientific reports*, 10(1):1–18, 2020.

[83] J. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. *International journal of computer vision*, 1(4):333–356, 1988.

[84] D. H. Ballard. Animate vision. *Artificial intelligence*, 48(1):57–86, 1991.

[85] J. Bohg, K. Hausman, B. Sankaran, O. Brock, D. Kragic, S. Schaal, and G. S. Sukhatme. Interactive perception: Leveraging action in perception and perception in action. *IEEE Transactions on Robotics*, 33(6):1273–1291, 2017.

[86] A. Sipatchin, S. Wahl, and K. Rifai. Accuracy and precision of the htc vive pro eye tracking in head-restrained and head-free conditions. *Investigative Ophthalmology & Visual Science*, 61(7): 5071–5071, 2020.

[87] S. Guadarrama, A. Korattikara, O. Ramirez, P. Castro, E. Holly, S. Fishman, K. Wang, E. Gonina, N. Wu, E. Kokiopoulou, L. Sbaiz, J. Smith, G. Bartók, J. Berent, C. Harris, V. Vanhoucke, and E. Brevdo. TF-Agents: A library for reinforcement learning in tensorflow. https://github.com/tensorflow/agents, 2018. URL https://github.com/tensorflow/agents. [Online; accessed 25-June-2019].

[88] K. Kang, S. Belkhale, G. Kahn, P. Abbeel, and S. Levine. Generalization through simulation: Integrating simulated and real data into deep reinforcement learning for vision-based autonomous flight. In *2019 international conference on robotics and automation (ICRA)*, pages 6008–6014. IEEE, 2019.