# SeedWorld: Generative Simulation Automation for Compositional Robotic Skill Generalization via Seed Demonstrations


**Jialin Tang**[1,*], **Sen Cui**[2,*], **Zhirui Hu**[1], **Tianyuan Chen**[1], **Bangde Cao**[3], **Jingheng Ma**[2], **Changshui Zhang**[2], **Shiji Zhou**[1,4,†], **Xiao Zhang**[1,†]

[1]Beihang University, [2]Tsinghua University, [3]Beijing University of Posts and Telecommunication,
[4]Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing

[*]Equal contribution, [†]Corresponding author

jialintang@buaa.edu.cn, zhoushiji25@buaa.edu.cn, xiao.zh@buaa.edu.cn


## 1. Introduction

Simulation is an important paradigm for training robots on diverse tasks [2, 5, 15, 17, 21]. However, building useful simulation tasks still requires substantial human effort, such as teleoperation [1, 7]. Existing data generation paradigms face a dilemma: augmentation methods [9] based on static datasets are limited by the semantic boundaries of the original data, while unconstrained generative exploration [14, 16] often suffers from low efficiency. We propose **SeedWorld**, a multi-agent simulation framework that expands a single human demonstration into composable simulation tasks for robot skill learning (Fig. 1). SeedWorld coordinates task-proposal, task-construction, skill-learning, and feedback agents to generate related task variants, instantiate executable compositional tasks, train robot skills, and improve low-success tasks. Experiments show that SeedWorld generates diverse and physically plausible tasks from very few seeds, enabling compositional skill generalization across different scenes with minimal human input.

## 2. SeedWorld Framework

SeedWorld is a coordinated multi-agent framework that expands a small set of human seed demonstrations into compositional simulation tasks. Given a human seed demonstration $\tau_{\text{seed}}$, i.e., a teleoperated trajectory of one manipulation skill, SeedWorld extracts task information $\Omega_{\text{seed}} = \langle O_{\text{seed}}, C_{\text{seed}}, R_{\text{seed}}, L_{\text{seed}} \rangle$, where $O_{\text{seed}}$ denotes the scene configuration, $C_{\text{seed}}$ denotes physical properties, $R_{\text{seed}}$ denotes the robot embodiment, and $L_{\text{seed}}$ denotes the task goal. A coordinator agent routes this information among four specialized agents: a task proposal agent, a task construction agent, a skill learning agent, and a feedback agent. SeedWorld grows one demonstrated skill into many learnable task compositions and continuously improve difficult task skill learning through feedback enhancement.

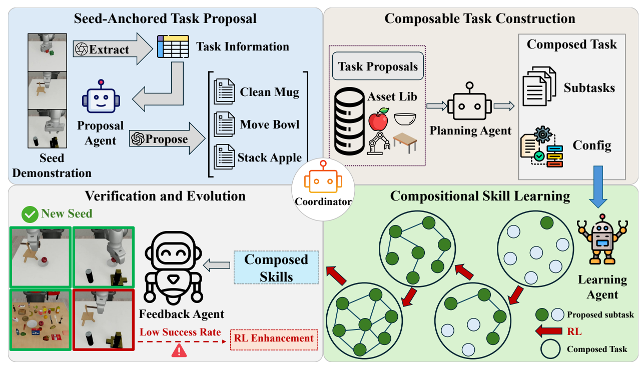



Figure 1. **Overview of SeedWorld.** A coordinator agent organizes four specialized agents. Starting from a seed demonstration, SeedWorld proposes related task variants, constructs executable and composable simulation tasks, trains robot skills, and evolves the task set through feedback enhancement from low-success cases.

### 2.1. Seed-Anchored Task Proposal

The task proposal agent is responsible for proposing task variants from a human seed demonstration. Given a seed $\tau_{\text{seed}}$, the agent analyzes the demonstrated task and extracts $\Omega_{\text{seed}} = \langle O_{\text{seed}}, C_{\text{seed}}, R_{\text{seed}}, L_{\text{seed}} \rangle$, where the four elements describe the scene configuration, physical properties, robot embodiment, and task goal. The agent then uses this information as constraints to produce a set of related but diverse task proposals $\mathcal{T} = \{\tau_1, \tau_2, \cdots, \tau_n\}$ by varying objects, goals, physical conditions, embodiments, and scene configurations. These proposals preserve the action logic and task goal demonstrated in the seed while exploring new task settings for later compositional construction.

### 2.2. Composable Task Construction

The task construction agent acts as a planning agent that turns coarse task proposals into executable compositional tasks. It selects compatible subtasks, orders them into a goal-conditioned sequence, and pairs each subtask with a configuration specifying the required objects, scene setup, initial state, goal condition, and success predicate. The

Table 1. **Quantitative assessment of task diversity.** Lower scores (↓) indicate higher diversity.

| Method | # Tasks | Self-BLEU ↓ | SBERT ↓ | ViT ↓ | CLIP ↓ |
|---|---|---|---|---|---|
| SemanticMimic | 100 | **0.252** | **0.163** | **0.184** | **0.752** |
| RoboGen [14] | 106 | 0.284 | 0.165 | 0.193 | 0.762 |
| Behavior-100 [12] | 100 | 0.299 | 0.210 | 0.389 | 0.833 |
| RLBench [6] | 106 | 0.317 | 0.200 | 0.375 | 0.864 |
| Meta-World [18] | 50 | 0.322 | 0.263 | 0.517 | 0.867 |
| ManiSkill2 [4] | 20 | 0.674 | 0.194 | 0.322 | 0.828 |
| GenSim [13] | 70 | 0.378 | 0.288 | 0.717 | 0.932 |

agent then builds a simulation instance for each subtask configuration and checks whether the full sequence is consistent with the task goal and physically feasible. Invalid plans are revised until they become valid training tasks.

### 2.3. Compositional Generalization in Skill Learning

The skill learning agent receives the valid compositional tasks constructed in the previous stage. Each task is organized as an ordered sequence of proposed subtasks, where each subtask is associated with a concrete configuration. Initialized with the seed-skill action prior, SeedWorld trains policies in simulation using reinforcement learning. This initialization reduces random exploration and helps the robot acquire subtask-level skills, which are then executed in the planned order to complete the full task composition.

### 2.4. Verification and Evolution

The feedback agent evaluates the learned subtask-level skills. It identifies low-success compositions and routes them back for targeted enhancement, using successful trajectories collected under the same subtask configurations as task-specific supervision. When a task composition is solved reliably, its successful trajectory sequence is promoted to a new seed demonstration. In this way, SeedWorld closes the loop from task proposal to skill learning: solved task compositions are reused as new demonstrations, enabling iterative task expansion and skill improvement.

### 2.5. Evaluation Metrics

We evaluate SeedWorld from two perspectives. **Task Diversity** measures whether the proposed tasks cover diverse language and visual variations. For task descriptions, we use Self-BLEU [10, 19] and Sentence-BERT similarity [19]; for rendered scene images, we use ViT [3] and CLIP [11] embedding similarity. **Skill Success** measures whether the learned policy can complete the generated tasks.

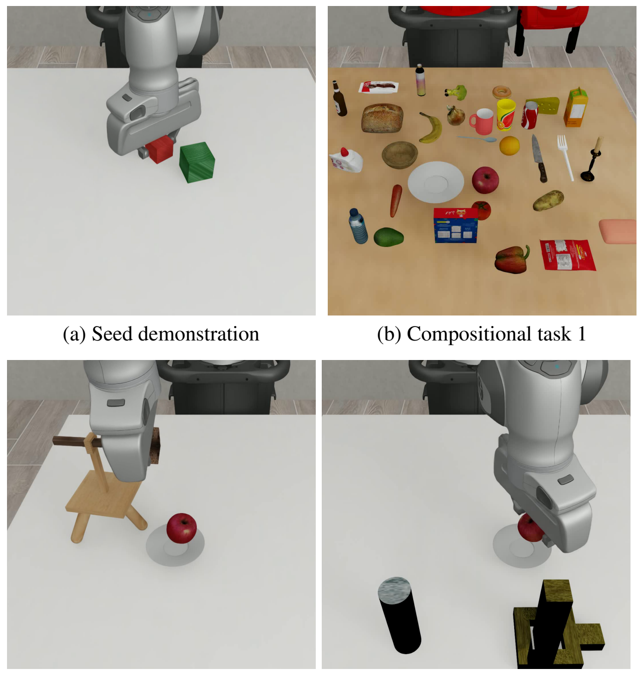
## 3. Results and Discussion

**Task Diversity Evaluation.** We first generate 100 task variations from different seed demonstrations and compare them with representative robot learning and simulation benchmarks. As shown in Table 1, SeedWorld achieves low task-description similarity and image similarity, indicating that the proposal agent produces diverse task variations. Figure 2 further illustrates how seed demonstrations can be expanded into multiple compositional tasks.

**Skill Learning Evaluation.** We evaluate whether the generated compositional tasks can support robot skill learning. Experiments are conducted in Robosuite [20], and policy training follows the default configuration of Robomimic [8]. As shown in Table 2, the learned policies can complete the generated tasks, suggesting that SeedWorld produces not only diverse but also executable and learnable tasks. For low-success tasks, feedback enhancement further improves performance by using successful trajectories collected from the corresponding task instances. Results show that SeedWorld can expand limited human demonstrations into useful compositional training tasks while preserving action guidance for skill learning.

Table 2. **Success rate of learned skills.** We compare the initial policy trained on generated tasks with two enhancement strategies: fine-tuning from the best checkpoint and retraining from scratch.

| Task | Success Rate | | |
|---|---|---|---|
| | Initial | Fine-tuning | Retraining |
| Apple Plate | 0.16 | **0.46** | 0.44 |
| Drink Cake Plate | 0.52 | **0.68** | 0.52 |
| Drink Cleanup | 0.20 | **0.42** | 0.36 |



(a) Seed demonstration
(b) Compositional task 1
(c) Compositional task 2
(d) Compositional task 3

Figure 2. **Task variation.** SeedWorld expands a seed demonstration into related but diverse compositional tasks.

## Acknowledgments and Disclosure of Funding

This work is supported by the National Science Foundation of China (62401327).

## References

[1] Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Joseph Dabis, Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, Jasmine Hsu, et al. Rt-1: Robotics transformer for real-world control at scale. *arXiv preprint arXiv:2212.06817*, 2022. 1

[2] Tao Chen, Jie Xu, and Pulkit Agrawal. A system for general in-hand object re-orientation. In *Conference on Robot Learning*, pages 297–307. PMLR, 2022. 1

[3] Alexey Dosovitskiy. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 2

[4] Jiayuan Gu, Fanbo Xiang, Xuanlin Li, Zhan Ling, Xiqiang Liu, Tongzhou Mu, Yihe Tang, Stone Tao, Xinyue Wei, Yunchao Yao, et al. Maniskill2: A unified benchmark for generalizable manipulation skills. *arXiv preprint arXiv:2302.04659*, 2023. 2

[5] Tuomas Haarnoja, Ben Moran, Guy Lever, Sandy H Huang, Dhruva Tirumala, Jan Humplik, Markus Wulfmeier, Saran Tunyasuvunakool, Noah Y Siegel, Roland Hafner, et al. Learning agile soccer skills for a bipedal robot with deep reinforcement learning. *Science Robotics*, 9(89):eadi8022, 2024. 1

[6] Stephen James, Zicong Ma, David Rovick Arrojo, and Andrew J Davison. Rlbench: The robot learning benchmark & learning environment. *IEEE Robotics and Automation Letters*, 5(2):3019–3026, 2020. 2

[7] Ajay Mandlekar, Yuke Zhu, Animesh Garg, Jonathan Booher, Max Spero, Albert Tung, Julian Gao, John Emmons, Anchit Gupta, Emre Orbay, et al. Roboturk: A crowdsourcing platform for robotic skill learning through imitation. In *Conference on Robot Learning*, pages 879–893. PMLR, 2018. 1

[8] Ajay Mandlekar, Danfei Xu, Josiah Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, and Roberto Martín-Martín. What matters in learning from offline human demonstrations for robot manipulation. In *arXiv preprint arXiv:2108.03298*, 2021. 2

[9] Ajay Mandlekar, Soroush Nasiriany, Bowen Wen, Iretiayo Akinola, Yashraj Narang, Linxi Fan, Yuke Zhu, and Dieter Fox. Mimicgen: A data generation system for scalable robot learning using human demonstrations. *arXiv preprint arXiv:2310.17596*, 2023. 1

[10] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318, 2002. 2

[11] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PmLR, 2021. 2

[12] Sanjana Srivastava, Chengshu Li, Michael Lingelbach, Roberto Martín-Martín, Fei Xia, Kent Elliott Vainio, Zheng Lian, Cem Gokmen, Shyamal Buch, Karen Liu, et al. Behavior: Benchmark for everyday household activities in virtual, interactive, and ecological environments. In *Conference on robot learning*, pages 477–490. PMLR, 2022. 2

[13] Lirui Wang, Yiyang Ling, Zhecheng Yuan, Mohit Shridhar, Chen Bao, Yuzhe Qin, Bailin Wang, Huazhe Xu, and Xiaolong Wang. Gensim: Generating robotic simulation tasks via large language models. *arXiv preprint arXiv:2310.01361*, 2023. 2

[14] Yufei Wang, Zhou Xian, Feng Chen, Tsun-Hsuan Wang, Yian Wang, Katerina Fragkiadaki, Zackory Erickson, David Held, and Chuang Gan. Robogen: Towards unleashing infinite data for automated robot learning via generative simulation. *arXiv preprint arXiv:2311.01455*, 2023. 1, 2

[15] Thomas Weng, Sujay Man Bajracharya, Yufei Wang, Khush Agrawal, and David Held. Fabricflownet: Bimanual cloth manipulation with a flow-based policy. In *Conference on Robot Learning*, pages 192–202. PMLR, 2022. 1

[16] Zhou Xian, Theophile Gervet, Zhenjia Xu, Yi-Ling Qiao, Tsun-Hsuan Wang, and Yian Wang. Towards generalist robots: A promising paradigm via generative simulation. *arXiv preprint arXiv:2305.10455*, 2023. 1

[17] Zhenjia Xu, Zhou Xian, Xingyu Lin, Cheng Chi, Zhiao Huang, Chuang Gan, and Shuran Song. Roboninja: Learning an adaptive cutting policy for multi-material objects. *arXiv preprint arXiv:2302.11553*, 2023. 1

[18] Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on robot learning*, pages 1094–1100. PMLR, 2020. 2

[19] Yaoming Zhu, Sidi Lu, Lei Zheng, Jiaxian Guo, Weinan Zhang, Jun Wang, and Yong Yu. Texygen: A benchmarking platform for text generation models. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 1097–1100, 2018. 2

[20] Yuke Zhu, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Abhishek Joshi, Kevin Lin, Soroush Nasiriany, and Yifeng Zhu. robosuite: A modular simulation framework and benchmark for robot learning. In *arXiv preprint arXiv:2009.12293*, 2020. 2

[21] Ziwen Zhuang, Zipeng Fu, Jianren Wang, Christopher Atkeson, Soeren Schwertfeger, Chelsea Finn, and Hang Zhao. Robot parkour learning. *arXiv preprint arXiv:2309.05665*, 2023. 1