

# One-Shot Learning from a Demonstration with Hierarchical Latent Language

Extended Abstract

Nathaniel Weir  
 Johns Hopkins University  
 Baltimore, United States of America  
 nweir@jhu.edu

Xingdi Yuan  
 Microsoft Research  
 Montreal, Canada  
 eric.yuan@microsoft.com

Marc-Alexandre Côté  
 Microsoft Research  
 Montreal, Canada

Matthew Hausknecht  
 Microsoft Research  
 Redmond, United States of America

Romain Laroche  
 Microsoft Research  
 Montreal, Canada

Ida Momennejad  
 Microsoft Research  
 New York, United States of America

Harm Van Seijen  
 Microsoft Research  
 Montreal, Canada

Benjamin Van Durme  
 Johns Hopkins University, Microsoft  
 Baltimore, United States of America

## ABSTRACT

Humans have the capability, aided by the expressive compositionality of their language, to learn quickly by demonstration. They are able to describe unseen task-performing procedures and generalize their execution to other contexts. This work introduces DescribeWorld, a Minecraft-like grid world environment designed to test this sort of generalization skill in grounded agents, where tasks are linguistically and procedurally composed of elementary concepts. The agent observes a single task demonstration, and is then asked to carry out the same task in a new map. To enable such a level of generalization, we propose a neural agent infused with hierarchical latent language—at the levels of task inference and subtask planning. Through a suite of generalization tests, we find agents that perform text-based inference are better equipped for the challenge under a random split of tasks.

## KEYWORDS

Demonstration Following; Compositional Generalization; Latent Language Policy; Embodied Language; Grounded Agents

### ACM Reference Format:

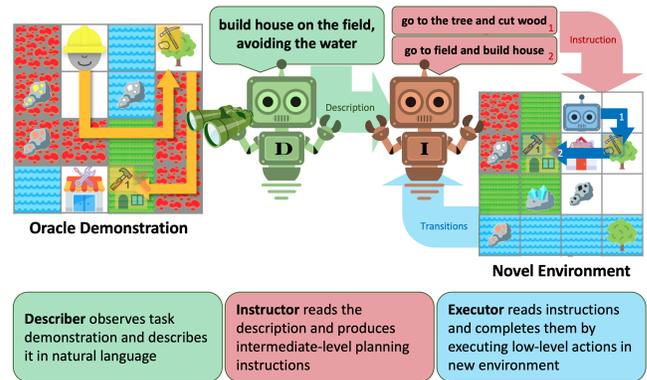
Nathaniel Weir, Xingdi Yuan, Marc-Alexandre Côté, Matthew Hausknecht, Romain Laroche, Ida Momennejad, Harm Van Seijen, and Benjamin Van Durme. 2023. One-Shot Learning from a Demonstration with Hierarchical Latent Language: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023)*, London, United Kingdom, May 29 – June 2, 2023, IFAAMAS, 3 pages.

This is an extended abstract of [8].

## 1 INTRODUCTION

Humans are highly capable of learning by example. If a child watches a school teacher draw a purple winged elephant then recite the alphabet backwards, they can replicate the sequence of activities at home with ease. This is in no small part due to the

*Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023)*, A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom. © 2023 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.



**Figure 1: Framework for learning from demonstration via latent language in DescribeWorld. The Describer module observes an oracle demonstration of an unseen task and describes it in text to the Instructor module, which infers necessary subtasks that the Executor module completes via low-level control actions.**

human ability to leverage the compositionality of language in order to comprehend new situations composed of familiar concepts [2]. Without this generalization, a human might overfit to the specifics of the classroom context.

In this work, we explore whether grounded artificial agents can similarly generalize from a demonstration: a single expert trajectory accomplishing a novel task (see Figure 1 for an overview). We construct DescribeWorld, a 2D grid environment containing a dataset of tasks involving collecting resources, building recipes, navigation, and interaction with objects and terrains.<sup>1</sup> Each task is parametrized by an end goal, the completion of which ends the session with a high reward. Test tasks are distinct from training tasks, but they are procedurally composed of the same subtasks and low-level actions. Figure 2 provides examples of such tasks.

<sup>1</sup>Examples and code found at [describeworld.github.io](https://github.com/describeworld).

	Training Tasks	Evaluation Tasks
Hidden Subtask	<ul style="list-style-type: none"> <li>erect wood shrine</li> <li>erect iron shrine</li> <li>place wood flooring</li> </ul>	<ul style="list-style-type: none"> <li>place iron flooring</li> </ul>
Hidden Use Case	<ul style="list-style-type: none"> <li>erect wood shrine on iron flooring</li> <li>erect wood shrine on field</li> <li>build diamond house</li> </ul>	<ul style="list-style-type: none"> <li>place iron flooring on field</li> <li>build diamond house on field</li> </ul>
Hidden Terrain Destination	<ul style="list-style-type: none"> <li>place wood flooring on field</li> <li>place iron flooring covering all fields</li> <li>place wood flooring, don't walk on water</li> </ul>	<ul style="list-style-type: none"> <li>place wood flooring on water</li> <li>place iron flooring covering all waters</li> </ul>
Length Generalization	<ul style="list-style-type: none"> <li>build diamond house on road</li> <li>place iron flooring covering all fields</li> </ul>	<ul style="list-style-type: none"> <li>place iron flooring covering all fields and build diamond house on road</li> </ul>

Figure 2: Systematic generalization splits with examples.

## 2 HIERARCHICAL LATENT LANGUAGE POLICY

In addition to DescribeWorld, we devise a novel three-level Hierarchical Latent Language Policy (**HLLP**) agent, depicted Figure 1, that represents both high-level tasks (“*build a house on field*”) and subtask plans (“*cut wood*”) in natural language (NL). The agent uses text representations at two levels of abstraction: identifying top-level tasks (via a *describer* module), and identifying a sequence of intermediate-level subtasks (via *instructor*). We train the agent via imitation learning on NL text tied to oracle actions. The **describer** module observes oracle demonstrations and describes them in text. The **instructor** and **executor** modules work in tandem to generate intermediate-level text instructions and choose low-level actions.

## 3 EXPERIMENTS

For our novel testing scenario of **demonstration following**, where the agent must replicate a demonstrated task in a new random map, we contrast approaches that leverage latent language policies versus those that instead use continuous representations. Specifically, we contrast the HLLP agent with a fully nonverbal baseline (**NV Baseline**) and a Latent Language Description Only (**LLD**) baseline (i.e., HLLP without the instructor).

DescribeWorld tests demonstration following agents for multiple forms of generalization. The **random task split**: is a pointwise random 70/30 split of the tasks grouped by their end goal. This is nontrivially difficult due to complex subtask dependencies. The **systematic generalization splits** each require a particular form of rule-based systematic generalization (Figure 2). For example, we hold specific crafting recipes out of training and require agents to compose them at test time from learnt recipes (e.g. train on *erect iron shrine* and *place wood flooring* and test on *place iron flooring*). Such systematic discrepancies between train and test cause catastrophic performance drops for most models.

Our main evaluation metric is the binary completion of the demonstrated task. We find that modeling agent policy as latent natural language improves the ability to generalize to demonstrations of unseen tasks. Table 1 shows task completion rate on the (upper) random task split and (lower) generalization challenge splits.

Table 1: Completion rates for demonstration following

Completion (%)	NV Baseline	LLD	HLLP
<b>Random task splits</b>			
<b>Overall</b>	25.2 ± 7.0	65.1 ± 3.2	68.4 ± 2.2
Navigation	45.6 ± 2.6	40.5 ± 1.3	46.5 ± 2.9
Crafting	44.4 ± 13.7	79.6 ± 3.2	85.5 ± 1.7
Craft then Nav	45.4 ± 14.3	89.4 ± 1.8	95.1 ± 1.4
Build on Terrain	9.1 ± 2.7	54.4 ± 4.1	63.0 ± 3.4
Cover Terrain	5.4 ± 2.9	61.2 ± 4.0	37.9 ± 1.7
Clear Items	11.6 ± 5.6	39.3 ± 0.6	27.0 ± 6.3
<b>Systematic generalization splits</b>			
<b>Hidden Subtask</b>	2.5 ± 1.4	1.3 ± 0.4	0.4 ± 0.3
<b>Hidden Use Case</b>	0.3 ± 0.5	5.1 ± 1.5	5.9 ± 3.3
<b>Hidden Terrain Destn</b>	1.6 ± 0.9	4.6 ± 0.5	3.7 ± 0.7
<b>Length Generalization</b>	6.0 ± 2.1	62.6 ± 3.8	57.9 ± 9.0

Both agents that leverage a predicted task description (HLLP and LLD) outperform the nonverbal baseline on the random unseen task split. Results are mixed on the suite of generalization challenge splits, where model performance is close to 0, though latent language contributes a couple of points particularly for generalization to tasks with longer trajectory length. Full results on other testing scenarios (e.g., **description following** and **instruction following** are found in [8]).

## 4 DISCUSSION AND CONCLUSION

DescribeWorld is a challenging environment that tests agents at learning new tasks, composed of familiar concepts, given just single demonstrations. Our suite of high-level tasks requires an agent to identify task concepts and their procedural roles in unseen combinations. This builds upon existing work that tests for compositional generalization in language processing models [4, 5, 7], but alters the input modality such that agents must respond to nonverbal demonstrations of novel tasks. We argue that agents *using language latently* for planning are better suited for this type of challenge.

We propose a *hierarchical latent language policy* agent, which makes decisions on the basis of text at multiple levels of abstraction. Our results suggest that language serves as an effective means to identify and act upon these novel combinations, thus providing an expressive, generalization-promoting representation for one-shot demonstration following agents. Intermediate-level planning on the basis of LM decoding provides incremental improvements upon nonverbal baselines on a random task split, suggesting improved generalization to other maps and unseen tasks under a random split. However, we find that instruction-level latent language does not solve the challenging DescribeWorld generalization splits.

We observe that models can accomplish systematically novel tasks provided the correct decision is made at a higher level of abstraction, exemplifying how hierarchical latent language provides a mechanism for isolating the level of policy abstraction in which a generalization might occur. Indeed, such challenges remain largely unsolved, though recent approaches have suggested incremental progress in specific cases [1, 3, 6]. We hope that our benchmark adds to this discourse, and that future work considers DescribeWorld.

## REFERENCES

- [1] Jacob Andreas. 2020. Good-Enough Compositional Data Augmentation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 7556–7566. <https://doi.org/10.18653/v1/2020.acl-main.676>
- [2] Noam Chomsky. 1957. *Syntactic Structures*. De Gruyter Mouton.
- [3] Henry Conklin, Bailin Wang, Kenny Smith, and Ivan Titov. 2021. Meta-Learning to Compositionally Generalize. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Online, 3322–3335. <https://doi.org/10.18653/v1/2021.acl-long.258>
- [4] Najoung Kim and Tal Linzen. 2020. COGS: A Compositional Generalization Challenge Based on Semantic Interpretation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Online, 9087–9105. <https://doi.org/10.18653/v1/2020.emnlp-main.731>
- [5] Brenden Lake and Marco Baroni. 2018. Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks. In *International conference on machine learning*. PMLR, 2873–2882.
- [6] Linlu Qiu, Hexiang Hu, Bowen Zhang, Peter Shaw, and Fei Sha. 2021. Systematic Generalization on gSCAN: What is Nearly Solved and What is Next?. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 2180–2188. <https://aclanthology.org/2021.emnlp-main.166>
- [7] Laura Ruis, Jacob Andreas, Marco Baroni, Diane Bouchacourt, and Brenden M Lake. 2020. A Benchmark for Systematic Generalization in Grounded Language Understanding. *Advances in Neural Information Processing Systems* 33 (2020).
- [8] Nathaniel Weir, Xingdi Yuan, Marc-Alexandre Côté, Matthew Hausknecht, Romain Laroché, Ida Momennejad, Harm Van Seijen, and Benjamin Van Durme. 2022. One-Shot Learning from a Demonstration with Hierarchical Latent Language. *arXiv preprint arXiv:2203.04806* (2022).