

LAB LOG

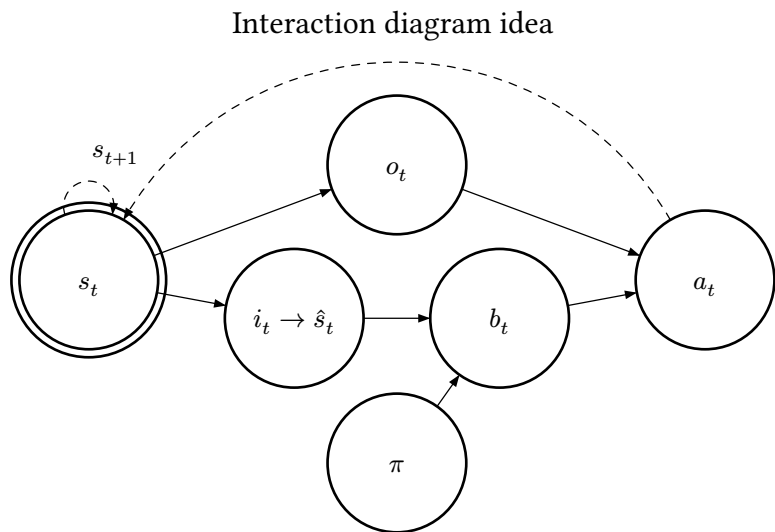
Noah Syrkis

June 27, 2025

1 | c2sim

2 | miiii

3 | aigs



- Policy π gets intel based $s_{\hat{t}}$ (not s_t itself)
- intel_fn map s_t to i_t . detel_fn map i_t to $s_{\hat{t}}$
- π map from $s_{\hat{t}}$ to b_t (could use MCTS also)

Figure 1: State s_t , intel i_t , behavior b_t (assigned to units by policy π weighing i_t), and action a_t (by b_t weighing observation o_t)

1.1 | `detel_fn(intel_fn(s))`

- ▶ Using gamma (jax native and easy fine tuning)
- ▶ As per Figure 2 we:
 1. We generate language intel i_t from state s_t
 2. Mask away some (maybe all) of state (s_{mt})
 3. Decode i_t and s_{mt} to get estimate $s_{\hat{t}}$
- ▶ See Appendix A for intel string templates
- ▶ Status: did `intel_fn` and doing `detel_fn`

```
1 Function IntelFunction( $s_t$ )
2   Generate mask for units not in sight
3   Generate  $i_t$  from  $s_t$  (could be lies)
4   Hide parts of  $s_t$  using mask to produce  $s_{mt}$ 
5   return  $i_t, s_{mt}$ 
6 end
7 Function DetelFunction( $i_t, s_{mt}$ )
8   Create prompt requesting indices to update
9   Use model to interpret  $i_t$  and  $s_{mt}$ 
10  Update  $s_{mt}$  with interpreted values
11  return updated state estimate  $\hat{s}_t$ 
12 end
13  $\hat{s}_t$ ) = DetelFunction(IntelFunction( $s_t$ ))
```

Figure 2: Pseudo code

Frequency spike in MLP layer around generalization

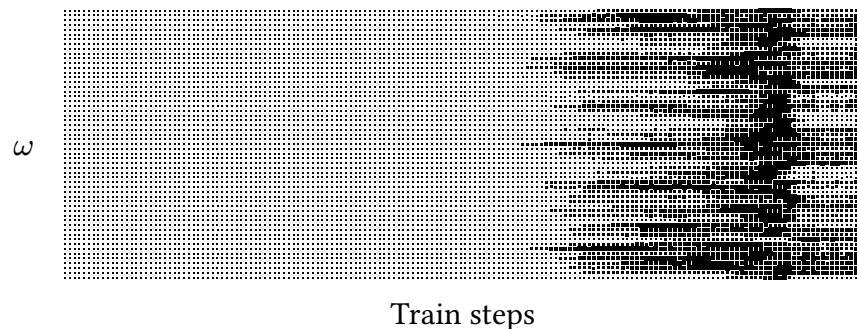


Figure 3: The spike in active frequencies during generalization indicate the presence of a non-generalizing and non-overfitting gradient component

- ▶ Grads have leaning and memory comps [1]
- ▶ Figure 3 Indicate a third, support-wheel comp
- ▶ Goal: publish in ICLR (better establish comp?)
- ▶ Now: chaning to better show spike across runs

3 | aigs

MCTS

- ▶ Connect 4 pettingZoo [2]
- ▶ Implement MCTS
- ▶ Tweak params and compete

DRL

- ▶ Get unity ml-agent to run
- ▶ pick game. Use PPO.
- ▶ play against

QD

- ▶ implement map elite
- ▶ generate dataset of levels
- ▶ Play lebel with drl bot

Index of Sources

- [1] J. Lee, B. G. Kang, K. Kim, and K. M. Lee, “Grokfast: Accelerated Grokking by Amplifying Slow Gradients,” no. arXiv:2405.20233. Jun. 2024.
- [2] J. Terry *et al.*, “Pettingzoo: Gym for Multi-Agent Reinforcement Learning,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 15032–15043, 2021.

A | Intel templates

```
> "Breaking news from the battlefield: Allied forces report enemy combatant spotted at {pos} with approximately {hp} health remaining."
> "Hey, did you hear? My cousin saw someone lurking around {pos} yesterday. They looked pretty beat up, maybe only {hp} health left. Be careful out there."
> "URGENT DISPATCH: Target acquired at coordinates {pos}. Visual assessment indicates {hp} vitality points. Proceed with caution."
> "Journal Entry, Day 47: Today I encountered a strange figure at {pos}. They appeared wounded, perhaps {hp} strength remaining.."
> "According to reliable sources, an individual was recently sighted at {pos} in poor condition, estimated at {hp} health. Local authorities knows."
> "Overheard at the tavern: 'I'm telling you, I saw them clear as day at {pos}! Could barely stand, maybe {hp} health at most. Something's not right.'"
> "Scout's Log: Entity detected at position {pos}. Current status: {hp} hit points. Monitoring situation closely."
> "My grandmother always said to watch out for strangers at {pos}. Well, I just saw one there, and they only had about {hp} health by the looks of it."
> "MEDICAL REPORT: Patient last seen at location {pos} with critical injuries. Estimated {hp} health remaining. Immediate assistance required."
> "Text message received: 'omg just saw someone at {pos}!! they look hurt bad, maybe like {hp} health?? should we call someone???'"
```