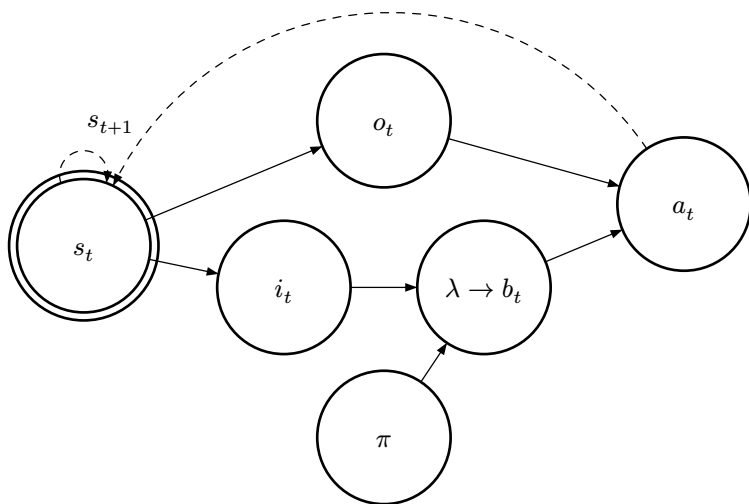# Lab Log

Noah Syrkis

June 26, 2025

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)

▸ Policy / plan π gets intel *about* s_t (not s_t itself)

▸ `intel_fn` that map s_t to i_t (see Appendix A).

▸ Next: use intel (not state) during plan eval

▸ For now λ is identity, but could do MCTS

▶ Grads have leaning and memory comps [1]

▶ Figure 2 Indicate a third, support-wheel comp

▶ Goal: publish in ICLR (better establish comp?)
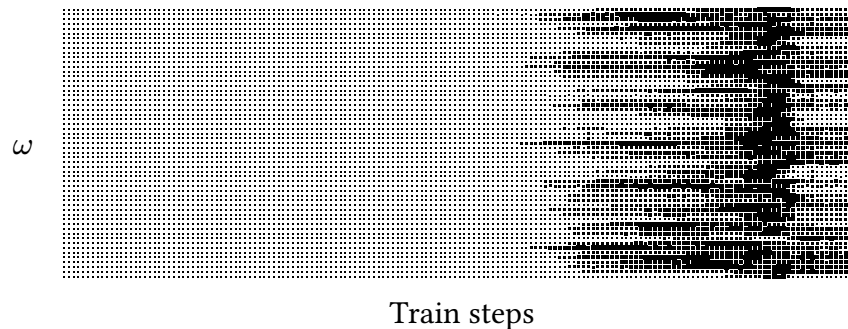
Frequency spike in MLP layer around generalization

ω

Train steps

Figure 2: The spike in active frequencies during generalization indicate the presence of a non-generalizing and non-overfitting gradient component

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.

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???'"