

Table2: **Ablation study on expert memory components.** We report success rate as the evaluation metric. w/o memory denotes the complete removal of the expert memory module. w/ log-experts, w/ stone-experts, etc., represent cumulative inclusion of expert memories in a fixed order (e.g., w/ stone-experts includes both log and stone experts). A dash (-) indicates that a specific expert type is not applicable to the task. Results demonstrate that performance improves with the inclusion of relevant expert memories, supporting the modular and compositional design of our memory architecture.

Group	Task	w/o memory	w/ log-experts	w/ stone-experts	w/ stone	w/ redstone-experts	w/ gold-experts	w/ diamond-experts	MoEA
Wooden	composter	0.8298	0.8936	-	-	-	-	-	0.8936
	bowl	0.9362	0.9574	-	-	-	-	-	0.9574
	wooden_sword	0.9848	1.0000	-	-	-	-	-	1.0000
	chest	1.0000	1.0000	-	-	-	-	-	1.0000
Stone	stone_shovel	0.8649	0.8919	0.9459	-	-	-	-	0.9459
	stone	0.8413	0.8413	0.8730	-	-	-	-	0.8730
	furnace	0.9130	0.9275	0.9596	-	-	-	-	0.9596
	stone_axe	0.9474	0.9825	1.0000	-	-	-	-	1.0000
Iron	chain	0.3390	0.3729	0.4746	0.4576	-	-	-	0.4576
	shears	0.3281	0.3281	0.5781	0.5781	-	-	-	0.5781
Aram	shield	0.4242	0.4091	0.4394	0.5303	-	-	-	0.5303
	golden_helmet	0.0667	0.0667	0.0800	0.0800	-	0.1067	-	0.1067
Redstone	piston	0.2532	0.2658	0.2532	0.2785	0.3038	-	-	0.3038
	redstone_torch	0.2247	0.2472	0.2472	0.2697	0.2809	-	-	0.2809
Golden	golden_hoe	0.0290	0.0145	0.0435	0.0580	-	0.0580	-	0.0580
	golden_shovel	0.0548	0.0548	0.0548	0.0685	-	0.0882	-	0.0882
Diamond	diamond_pickaxe	0.0665	0.0679	0.0665	0.0780	-	-	0.0882	0.0882
	diamond_axe	0.0423	0.0423	0.0282	0.0563	-	-	0.0845	0.0845