SPECIALIZATION AND EXCHANGE IN NEURAL MMO

Joseph Suarez and Phillip Isola
Massachusetts Institute of Technology
jsuarez@mit.edu

ABSTRACT

We present a simulated profession and exchange system for use in multi-agent intelligence research. Each of the eight implemented jobs produces items required by other professions. As a result, each profession must purchase items that they cannot produce themselves from other professions. These items are then used to produce increasingly high-quality goods for resale on a global market. Better and better goods enter the market as trade among professions creates a feedback loop. We integrate our profession and exchange system with Neural MMO, an existing multi-agent reinforcement learning platform capable of efficiently simulating populations of tens to 1000+ agents. We hope that our work will help support new research on *emergent specialization* — the ability to select and commit to a specific long-term strategy that fills a niche left by other learning agents. All of our code, including scripted baseline agents for each profession, will be free, open-source, and actively maintained.

1 INTRODUCTION

Specialization is a hallmark of human intelligence and achievement. Primitive societies featured distinct hunter and gatherer roles. Later, more specialized tradesman emerged, with different goods and services requiring different expertise. The modern world is so tremendously specialized that no single individual could produce just about any common household object from scratch. Milton Friedman famously described (Friedman & Friedman, 1981) this phenomenon as applied to a pencil: the wood, the saws used to cut the wood, the metal for those saws, the equipment to mine and process that metal, the graphite, the process for compacting it into a pencil, the rubber eraser and glue, etc. Thousands of people minimum are involved in these processes, and yet a pencil is both more effective and cheaper than older implements such as a quill and ink.

Societies of specialized individuals possess a great potential for invention and efficiency. Most or all modern multi-agent learning environments do not incentivize specialization in the way the real world does – we cannot expect specialization to simply emerge if, from an agent's perspective, there is no benefit to doing so. Our work introduces mechanics that enable specialization in simulated environments. We do not claim to accurately mimic human specialization structures – rather, our intent is to at least enable some form of learnable and quantifiable specialization for study.

2 RELATED WORK

Multiagent specialization is a common topic in early artificial life research (Sims, 1994). The problem has been studied more recently in reinforcement learning contexts. For example, (Trueba et al., 2012) explore the problem in the context of multi-robot self-organization while (Gasparrini et al., 2019) explore per-policy entropy regularization as a mechanism to enable convergence to different specializations. (Derek & Isola, 2021) even represent specializations implicitly in latent space. The precise history of algorithmic innovation in this area is outside the scope of the present work, as we do not build upon it. Rather, we aim to create better environments for the study of this problem. Excepting the below, previous works are limited to either physics simulations or simple one-off environments. Physics simulators are important in robotics and control, but this work bets against them for broader emergent behavior research on the basis of inefficiency and difficulty to develop. To our knowledge, there are presently no physics-based environments with support for higher-level mechanics such as progression and exchange (see Environment).

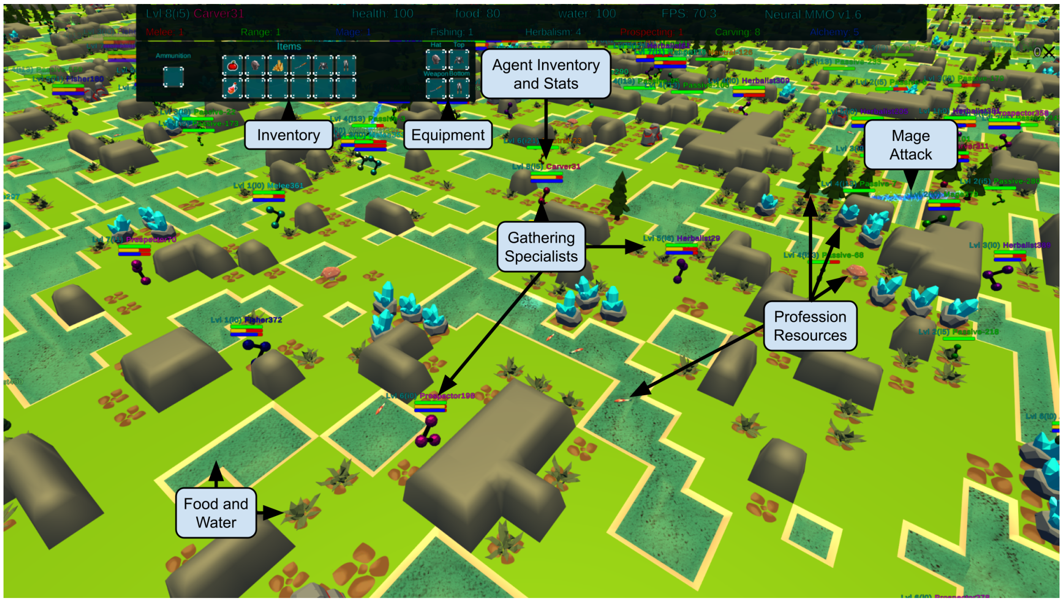

Figure 1: Annotated screenshot of a Neural MMO environment with profession and exchange systems. Around a hundred specialized agents are engaging in combat, gathering profession-specific resources, and exchanging goods on a global exchange.

Related environments include the AI Economist (Zheng et al., 2020) and the recent (Lux-AI-Challenge). Neither of these environments include progression systems that enable experienced agents to collect or produce better goods. Different agents in LuxAI can collect different resources, but that is where support for specialization ends. AI Economist implements a basic resource and exchange system for the purpose of studying tax policy. It assumes agents are differentiated by their skills upon creation and has limited potential for emergent specialization.

This work required a simulation platform with support for adding new game systems to existing environments. The only reasonable choice was Neural MMO (Suarez et al., 2021), an open-source and computationally accessible research platform that simulates populations of agents in procedurally generated virtual worlds. Our work builds on this platform by adding new profession and exchange systems. Crucially, we do so in a way that is straightforward to integrate with modern reinforcement learning frameworks and maintains the efficiency of the base platform. We refer to our work as an "environment" for simplicity, but really, Neural MMO is a much larger platform capable of simulating a wide variety of environments. Our profession and exchange systems hook into the platform's configuration system to provide these features in all such environments.

3 Environment

Maps are procedurally generated with seven different types of resource. Two of these resources, food and water, are required by all agents to survive. The other five resources are gathered using the associated profession. Figure 1 shows a map with generated terrain and resources. Any agent can harvest any resource, but agents trained in the corresponding profession will obtain better items. Of these five resources, two are used to make consumable items that restore food and water or health. The other three are used as ammunition in the three available combat styles. Combat is used to fight both other agents and scripted non-player characters (NPCs) which drop gold, equipment, weapons, and tools upon defeat. Equipment is useful to all agents as a means of defense. Each weapon type corresponds to a specific combat skill. Each tool type corresponds to a specific gathering skill. Agents can trade consumables, ammunition, weapons, tools, and equipment on a global market using gold. Figures 2 and 3 show the relationships between professions and items. Below, we describe each of these systems. **All of the constants used below are configurable**, but we provide the defaults to give a sense of scale.

Profession	Produces	Requires
Melee	Equipment	Sword, Scrap
Range	Equipment	Bow, Shaving
Mage	Equipment	Wand, Shard
Fishing	Ration	Rod
Herbalism	Poultice	Gloves
Prospecting	Scrap	Pickaxe
Carving	Shaving	Chisel
Alchemy	Shard	Arcane Focus

Figure 2: Each profession produces some items and requires others. For example, herbalists require increasingly high quality gloves in order to produce increasingly high quality poultices, which agents can use to restore health.

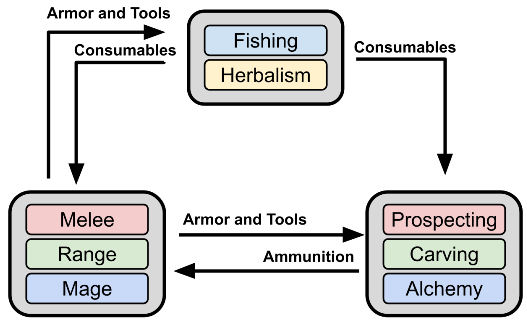

Figure 3: Each profession is directly or indirectly dependent upon goods from several other professions. This incentivizes exchange among agents: each profession sells excess products and buys items needed for survival and advancement.

Survival: Agents start with 100 food, water, and health. They lose 5 food and water per time step and begin losing health if either hits 0. Agents die at 0 health. Health slowly regenerates if agents have over 50 food and water. Agents restore water by moving adjacent to a water tile and restore food by stepping on a food tile. Food tiles decay once harvested and replenish slowly over time.

Combat: There are 3 combat profession: melee, range, and mage. These use rock-paper-scissors dominance: mage beats melee, melee beats range, and range beats mage. In this context, attacking a specialized agent with their weakness inflicts 1.5x damage.

Gathering: There are 5 gathering professions: fishing, herbalism, prospecting, carving, and alchemy. The former two produce consumable items which restore food, water, and health. These act as supplies when exploring resource-sparse areas or as immediate healing in a pinch. The latter three professions produce ammunition that strengthens the attacks of corresponding combat styles.

Items: There are 17 types of items: 2 consumables, 3 armor pieces, 3 munitions, 3 weapons, 5 tools, and 1 currency (gold). Except for gold, each category contains items from levels 1 through 10, for a total of 161 unique items. Higher level consumables restore more food, water, and health. Higher level armor provides better defensive bonuses in combat. Higher level munitions and weapons increase damage in the respective combat style. Higher level tools enable gathering higher level items using the respective profession. Gold is inherently valuable as the only currency of exchange.

Progression: Each profession begins at level 1 and can be raised up to level 10. Experience points are awarded for using the skill. For example, fighting with melee will raise the melee skill. Higher level agents can use higher level items. This places a strong emphasis on exchange with other agents as a means for acquiring relevant items.

Equipment: Agents can equip a helmet, chestplate, platelegs, held item, and ammunition. The former three armor pieces reduce damage taken in combat. Equipping weapons and ammunition increases damage with the associated combat profession. Note that munitions are consumed upon use, so combat agents must constantly replenish their supply. Wielding a tool confers a significant bonus to defense, enabling gathering agents to flee from poorly equipped aggressors. Equipping armor requires that agents have at least one skill of the same level of the armor piece. Equipping a held item or ammunition requires at least the same level in the associated skill. Higher-level tools enable agents to harvest higher-level resources. Higher-level ammunition and weapons enable agents to inflict more damage in combat.

Exchange: Agents can buy and sell items on a global market using gold. Gold can be obtained from defeating scripted non-players and is inherently valuable because it is the sole currency of exchange. To sell an item, agents specify an item in their inventory and a price. To buy an item, agents purchase one of the current market offers.

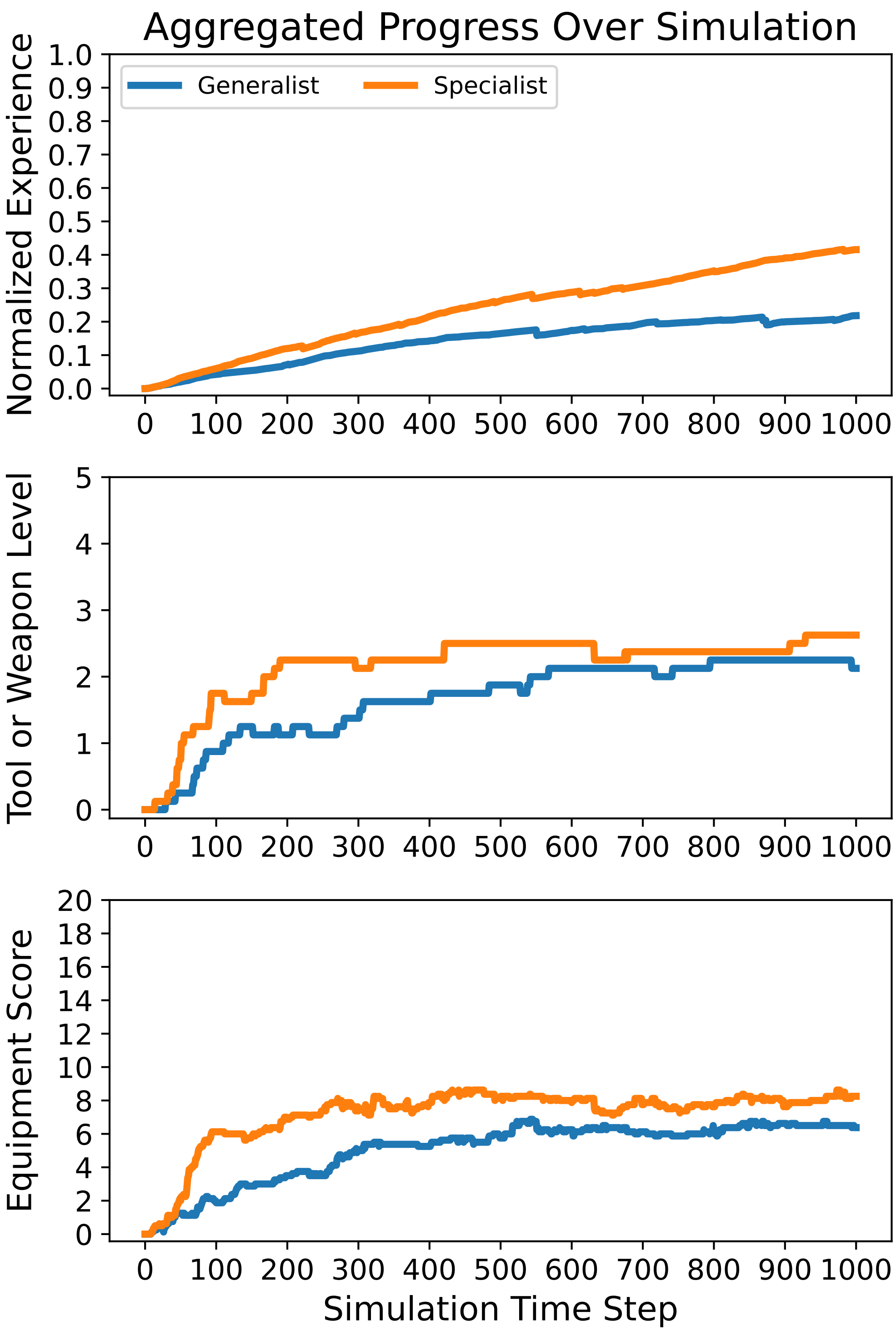

Figure 4: These metrics are aggregated over sets of scripted and generalist policies, respectively. Specialists gain experience points faster and aquire better tools, weapons, and equipment for their respective professions. See the appendix for replication details and per-profession statistics.

4 BASELINES

We provide 8 scripted policies. Agents spawn in the environment continuously and are uniformly sampled from these 8 (typically 10-30 of each are alive at any one time). Each policy corresponds to a single profession. Crucially, these define strategies for speciali*zing* rather than for speciali*sts* — agents begin with no inherent talents and may train any professions they wish. However, as per the equipment system description in the previous section, agents are only able to use better items according to their *highest* profession level. This strongly incentivizes specialization, as training a single skill will increase max level the fastest. At the same time, the dependency of each profession upon others for items makes it impractical for too many agents to share the same profession, as demand will greatly exceed supply. Thus our choice of baselines is reasonable, though likely not optimal.

The scripted policies share most of the same core logic. All agents process observations of themselves, nearby terrain, the market, and the inventory. In this stage, we extract higher-level attributes useful for scripting — lists of items owned that are not needed and can be sold, lists of equipment owned, lists of equipment that needs to be upgraded, lists of affordable items on the market, and so forth. Next, agents scan through these attributes and select actions. Neural MMO enables agents to take multiple actions simultaneously as long as they are in different categories. In particular, agents may move, attack, use an item, buy an item, and sell an item at the same time. We prioritize some actions over others in case of conflicts. For example, agents will use supplies (to regain food/health/water) before using armor (to equip it). All agents also share foraging, pathfinding, and exploration logic. There are two main distinctions among the 8 policies. First, the 3 combat policies fight using their respective attack type and the 5 gathering policies collect the resource specific to their profession. Second, each policy buys and sells different items according to its profession.

The purpose of the profession and exchange systems is to create a feedback loop in which agents specialize in a profession, acquire items using that profession, trade with other agents depending on their needs, and use their purchases to further enhance their profession, producing better items, and so on. These dynamics emerge even with our fairly simplistic scripted policies. Figure 4 shows that specialist agents gain experience in their respective professions and acquire tools and equipment over the course of simulation. These specialist policies also outperform ”generalist” policies in all three metrics. We implemented these by modifying the specialists to spend time training other skills: combat agents use all attack styles and gathering agents collect all profession-based resources. See the appendix for precise replication details and per-profession data.

We also perform a variety of ”simulation-first” logging that tracks the first time an event happens during the course of a simulation — for example, when new items enter the market, when more resource from the map have been concurrently depleted than ever before, when an agent inflicts more damage than ever before, when an agent reaches a new total item score (sum of equipment levels), and so on. We include the associated log file in the appendix.

5 DISCUSSION AND LIMITATIONS

It is important that we distinguish between the scope of the base Neural MMO platform and our profession and exchange system. The current public release of Neural MMO already includes resource, combat, and progression systems, as well as a (greatly simplified) equipment system. We added the (full) equipment system, professions, and exchange to this. In addition, we integrated the new equipment and professions with the existing progression system. As opposed to general game development, we had to design these systems from the ground up to be compatible with and efficient using Neural MMO's observation and action processing system, which is the main bottleneck in the platform's simulation speed. Examples of these low-level design requirements include allowing buy prices to fluctuate slightly when agents attempt to buy the same item (race conditions), structuring the market to only serialize observations for the best price of each item, and using a coarse-grained leveling system to avoid overly inflating the observation space.

These details are not relevant to an understanding of the main objective and contribution of this work, but is important to establish their existence. There are a total of 17 unique item types and, including item level, 161 unique items in the environment. To our knowledge, there exist no other research environments with an equipment and exchange system of this complexity. From our experience with this project, we believe that the most likely reason for this is the technical complexity of encapsulating these systems in standard observation and action systems.

There are other possible formulations of items and exchange, including far larger level spreads, fully procedurally generated item stats, peer-to-peer trading, local markets, and bartering. We chose a coarse item level system and global exchange because they are substantially simpler to integrate into decision making processes involving neural networks. Some parameterizations are prohibitively inefficient for use in modern deep learning. Most others are prohibitively complicated, requiring significant additions to the observation and action spaces. For example, bartering would require a trade request action, an incoming trade observation, a trade reply action, and a trade accept action, with each trade requiring a sequence of actions over 3 time steps from both parties to complete. Local markets would require dynamically adding a block of observations whenever a market is within range. These systems are certainly possible to implement, but they are not clearly better and, in the presence of uncertainty, we have chosen the implementation that makes our work the easiest for others to use and build upon.

The main limitation of this work is a lack of pretrained baselines. Thus far, we have modified the baseline networks provided by the official Neural MMO repository to run on our version of the environment. We have only run a couple of small training experiments so far, and we expect that some modifications to the network architecture will be required in order to fully take advantage of the new actions. We plan on doing this in the next few months and open-sourcing this code as well. We anticipate that better models will surpass the scripted baselines and likely learn qualitatively different behaviors. There are many possible intuitive strategies in the environment that were simply too difficult to script.

Our scripted baselines demonstrate one possible set of strategies in our Neural MMO derivative environments, but the strategy space as a whole is relatively unexplored. We expect that it is possible to substantially outperform the baseline by leveraging strategies we cannot currently anticipate. It is also possible that some of these strategies will do so in an ”uninteresting” manner by exploiting design flaws in the profession and exchange systems. This is a risk in all new environments, and we will make adjustments if needed. The maintainers of Neural MMO have agreed to include our systems in the next version of the platform, and the profession/exchange system is currently in public beta (though we cannot link it here during double-blind review). We will work with them to ensure it is up to standard with the rest of the platform.

6 Conclusion

We presented a profession and exchange system in Neural MMO that enables specialization to different roles, incentivizes trade among those roles, and creates a feedback loop by which specialized agents can produce better goods than comparable generalist agents. Our systems will be included in Neural MMO's next update and are currently in public beta. We commit to supporting this release and hope that it will be useful to researchers studying emergent-behavior in open-ended simulations.

References

Kenneth Derek and Phillip Isola. Adaptable agent populations via a generative model of policies. *CoRR*, abs/2107.07506, 2021. URL `https://arxiv.org/abs/2107.07506`.

Milton Friedman and Rose Friedman. *Free to choose: A personal statement*. Penguin Australia, 1981.

Marco Jerome Gasparrini, Ricard Solé, and Martí Sánchez-Fibla. Individual specialization in multi-task environments with multiagent reinforcement learners, 2019.

Lux-AI-Challenge. Lux-ai-challenge/lux-design-2021: Home to the design and engine of the @lux-ai-challenge season 1, hosted on @kaggle. URL `https://github.com/Lux-AI-Challenge/Lux-Design-2021`.

Karl Sims. Evolving virtual creatures. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pp. 15–22, 1994.

Joseph Suarez, Yilun Du, Clare Zhu, Igor Mordatch, and Phillip Isola. The neural mmo platform for massively multiagent research. 33, 2021. URL http://arxiv.org/abs/2110.07594.

Pedro Trueba, Abraham Prieto, Francisco Bellas, Pilar Caamaño, and Richard J. Duro. Self-organization and specialization in multiagent systems through open-ended natural evolution. In Cecilia Di Chio, Alexandros Agapitos, Stefano Cagnoni, Carlos Cotta, Francisco Fernández de Vega, Gianni A. Di Caro, Rolf Drechsler, Anikó Ekárt, Anna I. Esparcia-Alcázar, Muddassar Farooq, William B. Langdon, Juan J. Merelo-Guervós, Mike Preuss, Hendrik Richter, Sara Silva, Anabela Simões, Giovanni Squillero, Ernesto Tarantino, Andrea G. B. Tettamanzi, Julian Togelius, Neil Urquhart, A. Şima Uyar, and Georgios N. Yannakakis (eds.), *Applications of Evolutionary Computation*, pp. 93–102, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-29178-4.

Stephan Zheng, Alexander Trott, Sunil Srinivasa, Nikhil Naik, Melvin Gruesbeck, David C. Parkes, and Richard Socher. The ai economist: Improving equality and productivity with ai-driven tax policies, 2020.

A Appendix

A.1 Specialist vs. Generalist Simulation Details

We ran 20 total simulations of 1000 steps each – 10 for both specialist and generalist policies. Each simulation is around 10 minutes of real-time play and contains around 100 agents (the precise number is variable and depends on how long agents survive).

We collect experience, tool level, and equipment score data for each profession in each simulation. In each simulation, we then take the max profession level across agents at each time step. Finally, we average across parallel simulations. The quantities plotted below are therefore the average maximum values obtained at each point in simulation. We chose these metrics to indicate how well, on average, the best agent in each profession performs. Figure 4 further averages over all professions, but the per-profession breakdown for both generalists and specialists is shown in Figures 5 and 6. This reveals that, despite higher initial performance, fisher and hunter levels actually drop off later in simulation for specialist policies. The generation patterns for the resources corresponding to fishing and hunting are slightly harder to gather than those in other professions, and our scripted exploration algorithm is not particularly good. In seeking to also gather other resources, generalist policies effectively explore more. A better exploration policy would likely eliminate this anomaly. Even with this oversight, specialists still surpass generalists in aggregate on all three metrics.

A.2 Event Logs

Our logging records ”interesting” events that occur for the first time during any given simulation. They provide another way of interpreting agent progress throughout the course of a simulation. We will be incorporating them into a user-friendly dashboard in the full release of our profession and exchange systems. Below are the first, middle, and final sections of logs for one environment simulation.

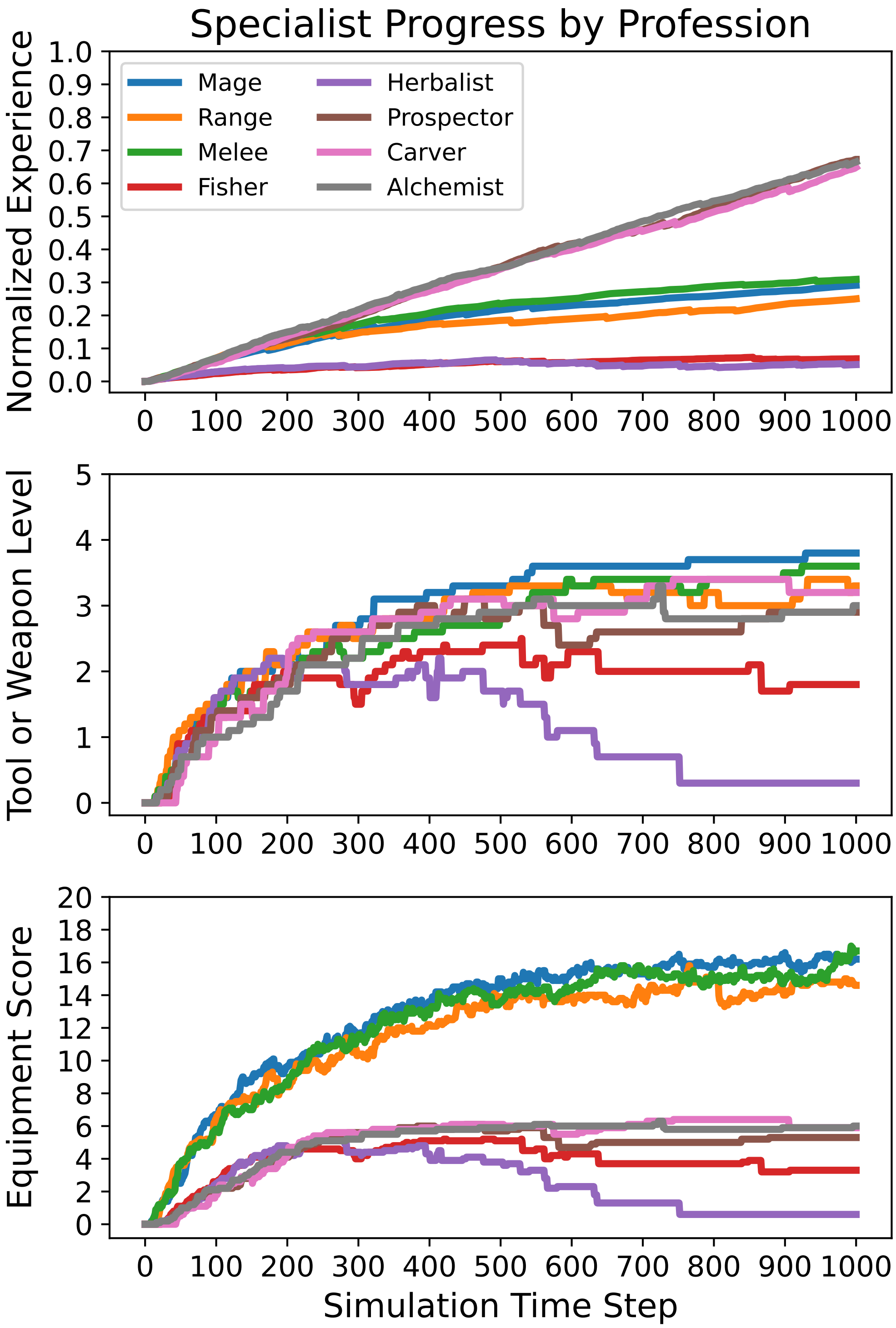

Figure 5: Per-profession specialist performance metrics. All but two professions improve across all metrics over the course of simulation. See above for possible explanations.

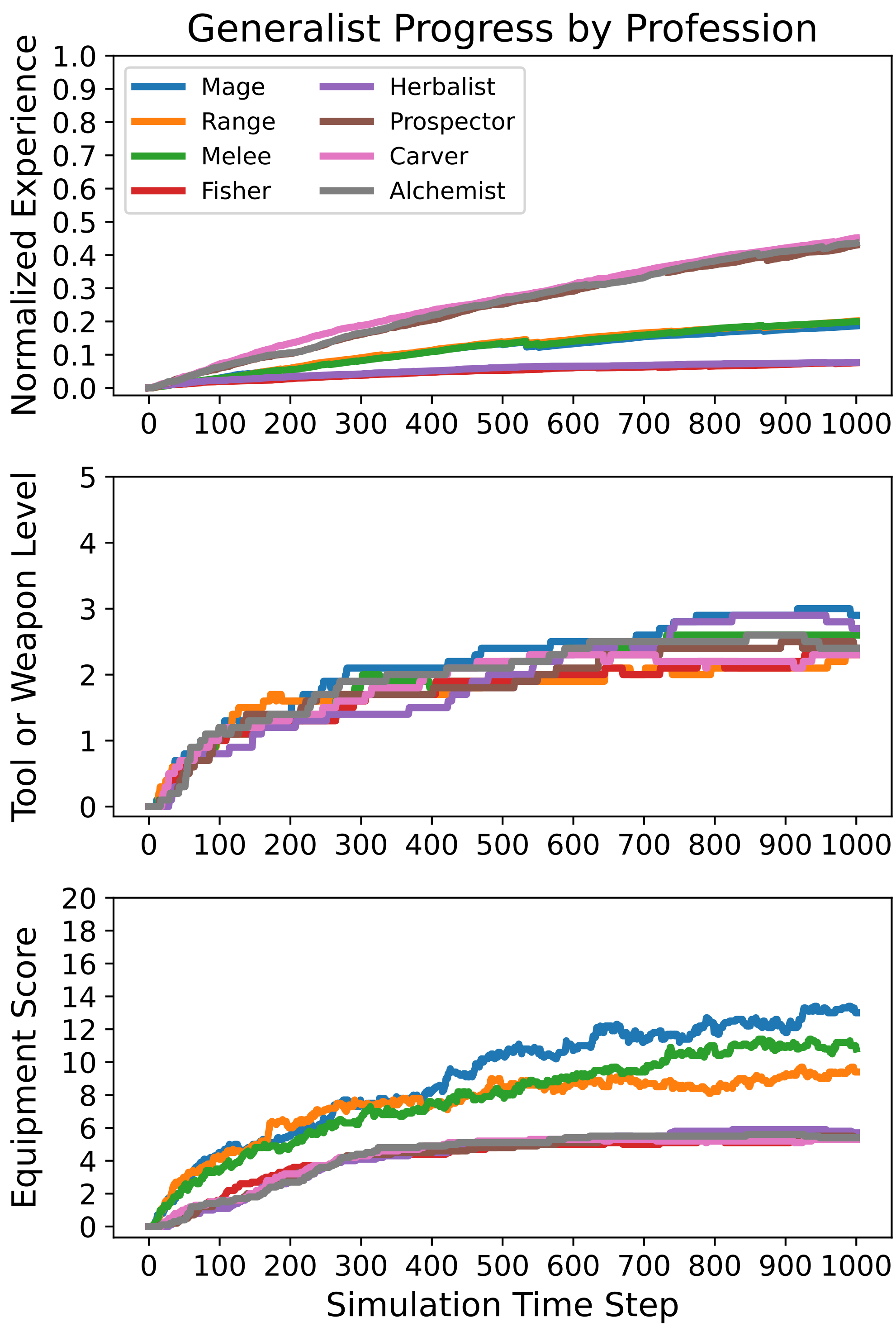

Figure 6: Per-profession generalist performance metrics. All but two professions under-perform the corresponding specialists over the course of simulation. See above for possible explanations.

```
Beginning
INFO:RESOURCE: Depleted 0 resource tiles
INFO:RESOURCE: Depleted 1 resource tiles
INFO:RESOURCE: Depleted 3 resource tiles
INFO:COMBAT: Inflicted 25 Mage damage (lvl 1 i0 vs lvl 1 i1)
INFO:PROGRESSION: Reached level 1 Mage
INFO:RESOURCE: Depleted 5 resource tiles
INFO:PROFESSION: Gathered level 1 Ration (level 1 Fishing)
INFO:PROGRESSION: Reached level 1 Fishing
INFO:COMBAT: Inflicted 30 Range damage (lvl 1 i0 vs lvl 1 i1)
INFO:PROGRESSION: Reached level 1 Range
INFO:RESOURCE: Depleted 8 resource tiles
INFO:COMBAT: Inflicted 25 Melee damage (lvl 1 i0 vs lvl 1 i1)
INFO:PROGRESSION: Reached level 1 Melee
INFO:EXCHANGE: Total wealth 16 gold
INFO:RESOURCE: Depleted 12 resource tiles
INFO:PROFESSION: Gathered level 1 Poultice (level 1 Herbalism)
INFO:PROGRESSION: Reached level 1 Herbalism
INFO:RESOURCE: Depleted 16 resource tiles
INFO:EQUIPMENT: Equipped level 1 Hat
INFO:EQUIPMENT: Item level 1

Middle
INFO:PROGRESSION: Reached level 5 Carving
INFO:EQUIPMENT: Item level 13
INFO:EXCHANGE: Offered level 3 Bottom for 4 gold
INFO:EXCHANGE: Bought level 3 Chisel for 4 gold
INFO:EQUIPMENT: Equipped level 3 Chisel
INFO:PROFESSION: Gathered level 3 Ration (level 3 Fishing)
INFO:PROGRESSION: Reached level 5 Alchemy
INFO:EXCHANGE: Offered level 4 Wand for 5 gold
INFO:EXCHANGE: Offered level 3 Pickaxe for 4 gold
INFO:EXCHANGE: Bought level 4 Wand for 5 gold
INFO:EQUIPMENT: Equipped level 4 Wand
INFO:EQUIPMENT: Item level 15
INFO:EQUIPMENT: Mage attack 70
INFO:EXCHANGE: Bought level 3 Pickaxe for 4 gold
INFO:EQUIPMENT: Equipped level 3 Pickaxe
INFO:COMBAT: Inflicted 142 Mage damage (lvl 4 i15 vs lvl 4 i1)
INFO:RESOURCE: Depleted 1167 resource tiles

End
INFO:EQUIPMENT: Equipped level 5 Top
INFO:EXCHANGE: Offered level 4 Bottom for 5 gold
INFO:EXCHANGE: Bought level 5 Bow for 6 gold
INFO:EXCHANGE: Transaction of 6 gold (level 5 Bow)
INFO:RESOURCE: Depleted 1519 resource tiles
INFO:EQUIPMENT: Equipped level 5 Bow
INFO:EXCHANGE: Bought level 4 Bottom for 5 gold
INFO:RESOURCE: Depleted 1524 resource tiles
INFO:EQUIPMENT: Item level 22
INFO:EQUIPMENT: Range attack 95
INFO:EXCHANGE: Bought level 4 Sword for 5 gold
INFO:PROGRESSION: Reached level 7 Carving
INFO:PROGRESSION: Reached level 7 Alchemy
INFO:COMBAT: Inflicted 142 Melee damage (lvl 6 i15 vs lvl 2 i0)
INFO:EXCHANGE: Offered level 5 Hat for 6 gold
INFO:COMBAT: Inflicted 165 Melee damage (lvl 5 i15 vs lvl 1 i0)
INFO:COMBAT: Inflicted 150 Mage damage (lvl 5 i14 vs lvl 1 i1)
```