Hierarchical Reinforcement Learning for Workforce Engagement: Modeling Goal Formation, Task Acceptance, and Retention in On-Demand Platforms

Bjorn J. White
School of Computing, Clemson University, Clemson, SC, USA.
bjwhite@clemson.edu

Bjorn Fox
Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL, USA.
bfox@uab.edu

Ross Bdams
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
adams1976@unh.edu

Adisdya Dandey
Department of Computer Science, George Mason University, Fairfax, VA, USA.
adityap@gmu.edu

Abstract

On-demand labor platforms have become a central component of the modern gig economy, yet they face persistent challenges in sustaining worker engagement, balancing flexible task acceptance with platform profitability, and reducing high turnover rates. This paper proposes a hierarchical reinforcement learning framework that models workforce engagement as a multi-level decision process encompassing individual goal formation, task acceptance behavior, and long-term retention dynamics. At the lower level, agents representing workers learn personalized goal-setting strategies under uncertainty, drawing on cognitive heuristics and motivational constructs. The intermediate level addresses task acceptance by integrating real-time platform incentives, worker preferences, and opportunity costs derived from the external labor market. At the highest level, platform-level policies optimize retention through adaptive compensation schemes, feedback mechanisms, and algorithmic governance. We discuss the structural trade-offs inherent in aligning these three levels, including conflicts between short-term liquidity and long-term workforce stability, fairness in task allocation, and the emergent risks of algorithmic exploitation. The paper also examines deployment considerations for large-scale systems, such as computational scalability, data requirements for training hierarchical policies, and the robustness of learned strategies to shifts in market conditions. By situating the framework within the broader socio-technical infrastructure of gig platforms, we highlight implications for platform governance, regulatory oversight, and the design of sustainable digital labor markets. The analysis draws on interdisciplinary insights from reinforcement learning, behavioral economics, organizational psychology, and human-computer interaction to offer a comprehensive systems perspective on workforce engagement.

Keywords

hierarchical reinforcement learning, gig economy, workforce engagement, goal formation, task acceptance, retention, platform governance, algorithmic management, fairness, sustainability.

1. Introduction

The rapid expansion of on-demand labor platforms has fundamentally transformed how work is organized, shifting from long-term employment relationships to a marketplace of tasks mediated by algorithms. While these platforms offer flexibility and low barriers to entry, they also introduce profound challenges in engaging a distributed, transient workforce whose participation is voluntary and often intermittent. Platforms must simultaneously ensure that workers are motivated to set meaningful goals, accept tasks efficiently, and remain active over extended periods to maintain a reliable labor supply. Traditional approaches to workforce management, which rely on hierarchical supervision and fixed contracts, are ill-suited to the dynamic, self-directed nature of gig work. Instead, a new class of computational models is needed to capture the interplay between individual decision-making and platform-level incentives.

Reinforcement learning has emerged as a powerful paradigm for sequential decision-making under uncertainty, and its hierarchical variants offer a natural way to decompose complex coordination problems into manageable sub-tasks [1], [2]. Early work in hierarchical reinforcement learning demonstrated how temporal abstraction and subgoal decomposition can accelerate learning in structured environments [3], [4], [5]. These techniques have since been extended to multi-agent settings and real-world applications such as robotics, game playing, and resource allocation. However, the application of hierarchical reinforcement learning to human-centric socio-technical systems, particularly labor markets, remains underexplored. The present paper develops a conceptual framework that models workforce engagement on gig platforms as a hierarchy of learning and decision processes spanning three levels: individual goal formation, task acceptance, and platform-level retention policies.

The motivation for this hierarchical perspective stems from the observation that worker behavior on platforms is not governed by a single monolithic objective. Instead, workers form personal productivity goals that evolve over time, influenced by past earnings, social comparisons, and intrinsic motivational factors [6], [7]. These goals, in turn, shape the kinds of tasks they are willing to accept, given the platform's reward structure and the availability of alternative opportunities in the external labor market. Platform operators, meanwhile, design policies that affect both the immediate supply of labor and the long-term health of the workforce. A hierarchical reinforcement learning formulation allows each level to operate at a different temporal and decision-making granularity, enabling coordinated yet decentralized adaptation.

Beyond technical modeling, the framework raises critical questions about fairness, transparency, and sustainability. When platforms optimize engagement metrics through algorithms that learn from worker behavior, there is a risk that workers become trapped in suboptimal patterns or that certain groups are systematically disadvantaged. Balancing the pursuit of platform efficiency with the protection of worker autonomy and well-being requires careful consideration of structural trade-offs. In the following sections, we review related work, detail the proposed hierarchy, analyze trade-offs, and discuss deployment and governance implications.

2. Related Work

Research on gig platform dynamics has drawn from multiple disciplines. Economists have studied the labor supply decisions of workers, showing that financial incentives, transaction costs, and non-monetary benefits all play significant roles [7], [8]. Sociologists and organizational scholars have examined how algorithmic management reshapes worker autonomy and control, often leading to new forms of precarity [9], [11]. In the human-computer interaction community, studies have documented how feedback mechanisms, reputation systems, and goal-setting features influence worker motivation and satisfaction [12]. These empirical findings provide a rich foundation for computational modeling.

Reinforcement learning has been applied to several related problems in online labor markets. For example, researchers have used multi-armed bandit models to optimize task pricing and allocation [13], and inverse reinforcement learning to infer worker preferences from observed acceptance decisions. However, these approaches typically treat each decision in isolation and ignore the hierarchical structure of goals and retention. Hierarchical reinforcement learning, by contrast, has been successfully used in other domains to handle long-horizon planning and credit assignment across multiple timescales [3], [5]. The concept of options (temporally extended actions) is particularly relevant for modeling work shifts or sequences of tasks that contribute to a larger goal.

Behavioral economics and psychology provide essential insights into goal formation and motivation. Prospect theory explains how workers evaluate gains and losses relative to reference points, which can shift over time [14]. Self-determination theory distinguishes between intrinsic and extrinsic motivation, suggesting that autonomy, competence, and relatedness are key drivers of sustained engagement [15]. Goal-setting theory emphasizes that specific and challenging goals, when combined with feedback, improve performance [16]. These theories have been validated in laboratory and field settings, but their integration into algorithmic platform design is nascent. Min et al. [10] conducted a field experiment demonstrating that self-set goals can significantly increase gig workers' output and retention, providing direct evidence for the importance of goal formation in platform contexts. Their work underscores the potential of incorporating goal-setting mechanisms into reinforcement learning models.

3. Hierarchical Reinforcement Learning Framework for Workforce Engagement

We propose a three-level hierarchy for modeling workforce engagement. At the lowest level, each worker is represented as a reinforcement learning agent that learns to set personal productivity goals. These goals, such as earning a target amount of money or completing a certain number of tasks within a day, serve as subgoals that structure the worker’s subsequent decisions. The goal formation process is influenced by the worker’s internal state, including accumulated earnings, fatigue, self-efficacy beliefs, and prior experiences [17], [18]. Because goals may be revised daily or weekly, this level operates at a relatively fine temporal granularity.

The intermediate level concerns task acceptance. At this level, the worker observes a stream of task offers from the platform, each characterized by a wage, required effort, time commitment, and possibly a reputational signal. The worker decides whether to accept or reject each offer, given their current goal, remaining time, and alternative opportunities (e.g., other platforms or leisure). This decision is influenced by the worker’s valuation of the task, which may be nonlinear due to reference-dependent preferences and diminishing sensitivity [14]. The task acceptance policy can be learned using standard reinforcement learning or options-based methods, where a chosen option corresponds to committing to a sequence of tasks that collectively move the worker toward their goal.

At the highest level, the platform itself is modeled as a reinforcement learning agent that selects policies affecting the entire workforce. These policies include dynamic pricing, task allocation rules, bonus structures, feedback mechanisms, and communication strategies. The platform's objective may be to maximize long-term profit, market share, or social welfare, but it must account for the reactions of workers who are themselves learning and adapting. This creates a hierarchical co-evolutionary system, where each level's policy changes over time in response to the others.

The hierarchical framework offers several advantages. First, it allows the decomposition of a complex learning problem into subproblems with different time horizons, reducing the curse of dimensionality. Second, it provides a natural way to incorporate human cognitive biases and motivational factors into the agent models, making the simulations more realistic. Third, it enables the platform to reason about the long-term consequences of its policies on worker retention and well-being, rather than simply optimizing short-term metrics such as the number of completed tasks per hour.

4. Goal Formation and Task Acceptance Mechanisms

Goal formation in the context of gig work is not a static process but a dynamic one that adjusts to feedback. Workers often set income targets based on past earnings, social comparisons with peers, and external obligations. The hierarchical reinforcement learning model can capture this by endowing the goal-setting agent with a representation of its own historical performance and a learned model of how effort translates into reward. The agent selects a goal as an action at the start of each decision episode (e.g., a day) and then receives a reward signal based on whether the goal was achieved, as well as any surplus earnings or costs incurred. The reward function for goal setting can incorporate both intrinsic satisfaction from achieving the goal and the instrumental value of income.

Task acceptance at the intermediate level is modulated by the current goal. A worker with a high earnings target may be more willing to accept low-paying tasks that accumulate toward the goal than a worker who has already met their target and now values leisure more highly. This behavior is consistent with reference-dependent preferences: workers exhibit loss aversion relative to their goal, so they work harder to avoid falling short than to exceed the target [14]. Empirical studies have shown that such goal-gradient effects increase effort as workers approach a target, and the framework can replicate this by including a positive reward for goal proximity [10]. Moreover, the hierarchical setup allows the task acceptance policy to be learned using state representations that include the goal, elapsed time, and recent task outcomes, enabling the agent to adapt its strategy over multiple days.

A critical aspect of this level is the opportunity cost of time. In the gig economy, workers typically have multiple platforms available, as well as non-work activities. The task acceptance model must therefore incorporate the value of alternative actions, which may be learned through experience or inferred from market data. Hierarchical reinforcement learning can abstract these alternatives as options that the worker can switch to, such as "work on Platform A," "work on Platform B," or "take a break." The worker's policy at the intermediate level thus chooses among options, each of which contains a sub-policy for accepting specific tasks.

5. Retention Dynamics and Platform Governance

Retention is the ultimate concern for platform sustainability. High turnover rates impose significant costs on platforms in terms of recruitment, training, and loss of experienced workers who contribute to efficient task completion and quality control. The hierarchical framework models retention as a macro-level outcome that emerges from the interaction of goal setting and task acceptance across many workers over time. Workers who consistently fail to meet their goals may experience reduced self-efficacy and ultimately disengage, while those who exceed their goals and receive positive feedback are more likely to remain [15], [16].

Platform governance encompasses the set of policies that influence these dynamics. For instance, platforms can offer bonuses for achieving daily goals, providing extra motivation that shifts the goal-setting policy toward higher targets. They can also adjust the task pricing algorithm to make it easier to reach goals during periods of high labor demand. However, these interventions must be designed carefully to avoid unintended consequences. If bonuses are too generous, workers may set unrealistically high goals and become disappointed, or the platform may incur excessive costs. Conversely, if goals are set too low, workers may underperform and the platform may suffer from insufficient labor supply.

The hierarchical reinforcement learning perspective enables the platform to learn these policies in a principled manner. The platform agent treats the entire workforce as a partially observable environment, receiving aggregate statistics such as completion rates, average goal attainment, and retention rates. By learning a policy that selects governance actions (e.g., adjusting bonus parameters, sending motivational messages, or modifying task allocation rules), the platform can optimize long-term retention while accounting for the adaptive responses of individual workers. This resembles a multi-agent reinforcement learning problem with a hierarchical structure, where the platform acts at a higher level and workers act at lower levels.

Fairness considerations become paramount in such a system. Platform algorithms may inadvertently discriminate against certain groups if their goal-setting or task acceptance behaviors differ due to demographic or contextual factors, such as access to high-paying tasks or family responsibilities. The hierarchical framework can be extended to incorporate fairness constraints, for example by penalizing policies that lead to disparate outcomes across worker segments. This requires careful design of reward functions and the inclusion of diversity metrics as part of the platform's objective.

6. Structural Trade-offs and Fairness Implications

The hierarchical framework reveals several inherent trade-offs that platforms must navigate. One key trade-off is between short-term task completion and long-term worker well-being. Aggressive pricing strategies that maximize immediate task acceptance may cause workers to burn out or become dissatisfied, leading to higher churn rates. Conversely, overly generous compensation may reduce platform profitability and attract workers with low intrinsic motivation. The hierarchical reinforcement learning model allows platforms to explore the Pareto frontier of these objectives by tuning the relative weights of different rewards.

Another trade-off involves the granularity of goal setting. Permitting workers complete autonomy to set their own goals may increase engagement through greater self-determination, but also may result in goals that are too easy or too difficult, reducing the platform's ability to predict labor supply. The field experiment by Min et al. [10] demonstrates that self-set goals can be highly effective, but their success depends on the platform's ability to provide timely feedback and adjust incentives accordingly. In a hierarchical system, the platform can learn to dynamically offer goal suggestions or default targets that nudge workers toward productive goal ranges without eliminating autonomy.

Fairness arises not only along demographic lines but also across temporal contexts. Workers who join the platform during boom periods may develop higher goal aspirations that become unsustainable during downturns, leading to systematic dropouts. The reinforcement learning algorithm must be robust to such non-stationarities. Moreover, because workers learn from their own experiences, initial conditions can lock in persistent inequalities. A platform policy that favors workers who already have high acceptance rates may exacerbate disparities. Hierarchical models can incorporate counterfactual reasoning and inverse propensity weighting to reduce such biases, but these techniques require careful implementation in large-scale systems.

7. Deployment, Infrastructure, and Sustainability

Deploying a hierarchical reinforcement learning framework for workforce engagement on real gig platforms presents significant infrastructure challenges. The system must handle millions of workers and tasks, each with its own state and action space. Scalable implementations require distributed computing architectures, efficient approximation methods for policy learning (e.g., deep neural networks), and streaming data pipelines that continuously update worker models. Moreover, the platform must maintain low latency for task offers and real-time decision support.

Data requirements are substantial. The platform needs to collect detailed logs of worker actions, task outcomes, and goal-setting behavior over long time horizons. Privacy concerns must be addressed through differential privacy and data anonymization. Additionally, the hierarchical structure demands careful credit assignment: rewards for goal setting are only observed at the end of a period, while task acceptance rewards are immediate. Algorithms such as hierarchical Q-learning or the options framework can handle this, but they require careful tuning of discount factors and state abstraction.

Sustainability in the context of gig platforms goes beyond economic viability to include environmental and social dimensions. The computational resources consumed by large-scale reinforcement learning systems are non-negligible, and platforms should consider energy-efficient methods. Social sustainability requires that the platform's algorithms do not exploit worker vulnerabilities or contribute to a race to the bottom in terms of wages and conditions. The hierarchical framework can be designed to explicitly incorporate worker welfare as a primary objective, rather than a secondary constraint. This aligns with emerging regulatory frameworks that call for algorithmic accountability and fair treatment of gig workers [9].

Robustness is another critical concern. The learned policies must generalize to shifts in market conditions, such as changes in the demand for services or the entry of new competitors. Hierarchical models that learn representations of exogenous variables (e.g., unemployment rates, seasonal trends) can adapt faster. However, overfitting to historical patterns can lead to brittle policies. Regularization strategies and online learning with periodic policy updates are essential for maintaining performance over time.

8. Conclusion

This paper has introduced a hierarchical reinforcement learning framework for modeling workforce engagement on on-demand platforms, covering individual goal formation, task acceptance, and retention dynamics. By decomposing the complex interplay between worker motivation and platform governance into three interacting levels, the framework provides a systematic lens for analyzing structural trade-offs, fairness implications, and sustainability concerns. The inclusion of behavioral and psychological factors, informed by empirical studies such as the field experiment on self-set goals [10], enhances the realism of the model and its relevance to real-world platform design.

The framework also highlights the need for interdisciplinary collaboration in the development of platform algorithms. Technical advances in hierarchical reinforcement learning must be combined with insights from behavioral economics, organizational psychology, and fairness-aware machine learning to create platforms that are not only efficient but also equitable and resilient. Future work should focus on empirical validation of the hierarchical model using platform data, development of scalable learning algorithms that respect worker privacy, and the design of governance mechanisms that can be audited and regulated. As on-demand work continues to grow, ensuring that platforms serve both economic and humanistic goals will become increasingly urgent, and hierarchical reinforcement learning offers a promising path forward.

References

1. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

2. Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 4, 237–285.

3. Dietterich, T. G. (2000). Hierarchical reinforcement learning with the MAXQ value function decomposition. Journal of Artificial Intelligence Research, 13, 227–303.

4. Parr, R., & Russell, S. (1998). Reinforcement learning with hierarchies of machines. Advances in Neural Information Processing Systems, 10, 1043–1049.

5. Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artificial Intelligence, 112(1–2), 181–211.

6. Vamplew, P., Dazeley, R., Berry, A., Issa, R., & Dekker, E. (2021). Empirical evaluation of multi-objective reinforcement learning. Autonomous Agents and Multi-Agent Systems, 35(1), 1–39.

7. Chen, M. K., & Horton, J. J. (2016). Are online labor markets spot markets for tasks? A field experiment on the gig economy. Journal of Labor Economics, 34(S2), S197–S242.

8. Horton, J. J., & Chilton, L. B. (2010). The labor economics of paid crowdsourcing. In Proceedings of the 11th ACM Conference on Electronic Commerce (pp. 209–218).

9. Sundararajan, A. (2016). The sharing economy: The end of employment and the rise of crowd-based capitalism. MIT Press.

10. Min, X., Chi, W., Hu, X., & Ye, Q. (2024). Set a goal for yourself? A model and field experiment with gig workers. Production and Operations Management, 33(1), 205-224.

11. Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. Work, Employment and Society, 33(1), 56–75.

12. Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1603–1612).

13. Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. The Quarterly Journal of Economics, 115(3), 791–810.

14. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291.

15. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78.

16. Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. American Psychologist, 57(9), 705–717.

17. Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman.

18. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131.

19. Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. Management Science, 49(10), 1407–1424.

20. Huberman, B. A., Romero, D. M., & Wu, F. (2009). Crowdsourcing, attention and productivity. Journal of Information Science, 35(6), 758–765.

21. Mason, W., & Watts, D. J. (2010). Financial incentives and the "performance of crowds." ACM SIGKDD Explorations Newsletter, 11(2), 77–84.

22. Bernstein, A. (2012). Crowdsourcing and human computation. In Encyclopedia of Database Systems (pp. 1–5). Springer.