

# Remember, but also, Forget: Bridging Myopic and Perfect Recall Fairness with Past-Discounting

Ashwin Kumar

Washington University in St. Louis  
Saint Louis, Missouri, USA  
ashwinkumar@wustl.edu

William Yeoh

Washington University in St. Louis  
Saint Louis, Missouri, USA  
weoh@wustl.edu

## ABSTRACT

Dynamic resource allocation in multi-agent settings often requires balancing efficiency with fairness over time—a challenge inadequately addressed by conventional, myopic fairness measures. Motivated by behavioral insights that human judgments of fairness evolve with temporal distance, we introduce a novel framework for temporal fairness that incorporates past-discounting mechanisms. By applying a tunable discount factor to historical utilities, our approach interpolates between instantaneous and perfect-recall fairness, thereby capturing both immediate outcomes and long-term equity considerations. Beyond aligning more closely with human perceptions of fairness, this past-discounting method ensures that the augmented state space remains bounded, significantly improving computational tractability in sequential decision-making settings. We detail the formulation of discounted-recall fairness in both additive and averaged utility contexts, illustrate its benefits through practical examples, and discuss its implications for designing balanced, scalable resource allocation strategies.

## KEYWORDS

Resource Allocation, Fairness, Multi-Agent RL

## 1 INTRODUCTION

Dynamic resource allocation is central to many real-world applications—from matching passengers to taxis [3, 20] and distributing aid to the homeless [9], to allocating life-saving vaccines [1]. In these settings, decision-makers must balance diverse agent preferences against resource constraints to maximize cumulative utility. Traditional fairness approaches impose constraints at each decision step (e.g., maximin fairness) but ignore the temporal dimension by treating each allocation in isolation. Such myopic methods fail to account for the evolving nature of fairness when past decisions and future opportunities are involved.

Recent work in multi-agent learning addresses temporal fairness by evaluating cumulative or terminal utilities [1, 8, 25]. Yet, these approaches either assume perfect recall—giving equal weight to all past allocations—or assess each allocation independently. In contrast, insights from behavioral economics and moral psychology reveal that human perceptions of fairness evolve over time. For example, events in the distant past tend to be perceived more abstractly [23], and individuals naturally devalue outcomes that are further removed in time [7]. Empirical studies also indicate that forgiveness increases as the temporal distance from a transgression grows [13, 24].

Motivated by these findings, we propose incorporating **past discounting mechanisms** into dynamic resource allocation. By

discounting historical utilities, our approach offers a principled compromise between instantaneous and perfect-recall fairness. This method not only aligns more closely with observed human behavior but also ensures that the augmented state space remains bounded—a critical property for the convergence and scalability of reinforcement learning algorithms.

In this paper, we:

- Highlight the limitations of traditional fairness approaches in dynamic settings.
- Propose a framework that incorporates past discounting to balance short-term allocations with long-term fairness considerations.
- Provide theoretical insights demonstrating how past discounting bounds the state space, thereby improving computational tractability.

By integrating theoretical analysis with behavioral insights, we aim to provide a strong argument for using past discounts for fair resource allocation in dynamic environments.

## 2 RELATED WORK

A considerable body of research has explored fairness in resource allocation from both economic and algorithmic perspectives. Traditional fair division literature has focused on static fairness notions such as proportionality, envy-freeness, and maximin share guarantees [5, 17]. These approaches typically consider a one-shot allocation problem, enforcing fairness at each individual decision point. However, in dynamic or sequential settings—such as taxi matching, aid distribution, or vaccine allocation—the myopic application of static fairness criteria neglects the evolution of cumulative utilities over time.

Recent work in multi-agent reinforcement learning has begun to address fairness in dynamic settings. Several approaches enforce fairness by constraining per-step allocations or by evaluating the accumulated utilities at an intermediate or terminal stage [8, 25]. More recently, Alamdari et al. [1] introduce non-Markovian fairness frameworks that explicitly incorporate past data by augmenting the state-space with historical allocations, looking at concepts like long-term, anytime, and bounded fairness.

Beyond resource allocation, sequential fairness is also explored in multi-issue voting and online fair division. In these settings, fairness must account not only for a single decision but for a series of interdependent choices—ranging from perpetual voting schemes [12] to online food bank allocation protocols [2]. Such cross-disciplinary work further highlights that static fairness notions must be adapted to dynamic environments if they are to reflect long-term, evolving perceptions of fairness.

While these models capture the temporal nature of decision-making, they typically assume either perfect recall (where past allocations are fully aggregated) or evaluate each allocation in isolation without considering how agents discount historical outcomes. In contrast, our work proposes a third paradigm—incorporating past-discounting mechanisms—motivated by evidence from behavioral economics and moral psychology.

Behavioral studies indicate that human fairness judgments are sensitive to temporal distance. Construal level theory posits that events in the distant past are represented more abstractly and evoke less emotional intensity [23]. Empirical research on time discounting shows that individuals systematically devalue outcomes as they recede into the past [7]. These findings, along with studies on forgiveness in intergroup contexts [13, 24], imply that fairness assessments in real-world scenarios may benefit from discounting earlier allocations. This concept is reminiscent of the temporal discounting used in reinforcement learning [15, 22], yet its explicit integration into fairness metrics for resource allocation remains largely underexplored.

Our work is also related to recent advances in fair multi-agent learning mechanisms. For example, methods for ride-hailing applications incorporate forecasting to balance future utility against current fairness [10], and hierarchical frameworks have been proposed to mediate between efficiency and fairness in multi-agent settings [8]. However, these approaches do not systematically account for the decay in the perceived value of past allocations. Instead, they either assume that all past outcomes are equally important (perfect recall) or that each allocation is evaluated in isolation (myopic fairness).

In summary, while the literature offers robust methods for achieving either instantaneous fairness or cumulative fairness with perfect recall, a practical middle ground remains unexplored. Our work fills this gap by introducing a past discounting framework that integrates historical context into fairness evaluation while accounting for the human tendency to weigh recent outcomes more heavily, all while keeping the state space computationally tractable.

### 3 PRELIMINARIES

Social welfare functions provide a mathematical formulation to evaluate both fairness and efficiency in resource allocation. Given a utility vector

$$Z = (z_1, z_2, \dots, z_n) \quad (1)$$

representing the utilities received by  $n$  agents, many social welfare functions have been considered in the literature, including:

- **Utilitarian Welfare**, which maximizes the total utility without explicit fairness considerations [19].

$$W_U(Z) = \sum_{i=1}^n z_i, \quad (2)$$

- **Egalitarian Welfare**, which prioritizes the well-being of the worst-off agent. This is also known as Rawlsian or maximin fairness [18].

$$W_{MMF}(Z) = \min_i z_i, \quad (3)$$

- **Nash Welfare**, which balances fairness and efficiency. This measure is rooted in Nash’s bargaining solution [16] and has

been influential in fair division research [6].

$$W_N(Z) = \prod_{i=1}^n z_i, \quad (4)$$

- **Generalized Gini Welfare**, which is a family of functions that applies rank-based weights to the agent utilities, offering a flexible approach to balancing equity and efficiency [4, 14, 25].

Traditionally, these functions evaluate fairness at a single point in time, thereby ignoring the history of past allocations and expectations for future ones. In dynamic settings, several approaches extend these welfare functions by incorporating historical and predictive elements. For example, some works cast the allocation problem as a Multi-Agent Reinforcement Learning (MARL) task that optimizes fairness at an intermediate or terminal state [8, 11, 21, 25], while others employ Non-Markovian Decision Processes that explicitly account for the entire past trajectory of allocations [1].

In many formulations, the allocation at time  $t$ , denoted by  $\mathcal{A}^t$ , is defined as a mapping of resources to agents such that  $\mathcal{A}_i^t$  represents the resources allocated to agent  $i$ . The welfare corresponding to the post-allocation utility vector  $Z^t|\mathcal{A}^t$  is then given by  $W(Z^t|\mathcal{A}^t)$ , and the optimal allocation is defined as:

$$\mathcal{A}^{t*} = \underset{\mathcal{A}}{\operatorname{argmax}} W(Z^t|\mathcal{A}). \quad (5)$$

Here, we denote by  $u_i^{\mathcal{A}}$  the utility derived by agent  $i$  from allocation  $\mathcal{A}$ , and by  $u_i^t$  the utility actually received by agent  $i$  at time  $t$ .

There exist various methods to compute the utility vector  $Z^t|\mathcal{A}$ , and these choices influence the resulting allocation. We discuss some popular approaches below, motivating and building up to past-discounted fairness.

## 4 TEMPORAL FAIRNESS IN RESOURCE ALLOCATION

In dynamic resource allocation, fairness must be evaluated not only on the basis of the current decision but also by considering past allocations and future expectations. In this section, we present three paradigms for temporal fairness: *Instantaneous fairness*, *perfect-recall historical fairness*, and *discounted-recall historical fairness*. We define each approach, illustrate them with examples, and discuss their inherent limitations.

### 4.1 Instantaneous Fairness

Instantaneous fairness considers only the current allocation decision. Formally:

$$Z_i^t|\mathcal{A} = u_i^{\mathcal{A}}, \quad (6)$$

which implies that fairness is assessed solely on the utility  $u_i^{\mathcal{A}}$  derived from the current allocation. In this formulation, the welfare function  $W$  is optimized based solely on the immediate utilities, yielding an allocation that is deemed fair at that specific time step.

**Definition 1 (INSTANTANEOUS FAIRNESS).** *An allocation exhibits instantaneous fairness if it optimizes a welfare function solely based on the one-step utilities  $u_i^t$ .*

Although this approach often produces a solution that is optimal for that particular time step, it neglects the temporal dimension by ignoring both historical allocations and anticipated future resources. For example, consider:

**Example 1.** Two agents, Alice and Bob, compete for two indivisible items: a cake and a donut. Suppose

$$\text{Alice: } (u_{\text{cake}}, u_{\text{donut}}) = (0.2, 0.5), \quad \text{Bob: } (0.3, 0.5).$$

In a purely instantaneous allocation, the donut is assigned to the agent who slightly benefits from it more in that step (Alice). Over repeated interactions, however, Alice may receive a disproportionate number of donuts, leading to a cumulative imbalance.

Thus, while instantaneous fairness might maximize short-term efficiency, its disregard for temporal dynamics can result in significant long-term disparities.

## 4.2 Perfect-Recall Historical Fairness

To capture the temporal aspect of fairness, perfect-recall historical fairness incorporates all past allocations into the fairness evaluation. Instead of relying solely on the instantaneous utility, we define an adjusted utility vector  $Z^t$  that aggregates utilities over all previous steps:

$$Z_i^t = \sum_{\tau=0}^t u_i^\tau.$$

Consequently, the post-allocation utility becomes:

$$Z_i^t | \mathcal{A} = \sum_{\tau=0}^{t-1} u_i^\tau + u_i^{\mathcal{A}},$$

$$Z_i^t | \mathcal{A} = Z_i^{t-1} + u_i^{\mathcal{A}}.$$

In some cases, averaging these utilities over time may be preferable:

$$Z_i^t | \mathcal{A} = \frac{Z_i^{t-1} \cdot (t-1) + u_i^{\mathcal{A}}}{t}.$$

**Definition 2 (PERFECT-RECALL FAIRNESS).** An allocation exhibits perfect-recall fairness if it optimizes a welfare function  $W$  over the cumulative (or averaged) utility vector  $(Z_1^t, \dots, Z_n^t)$ , where  $Z_i^t$  captures all past allocations.

This approach is especially relevant in domains such as long-term healthcare or education funding, where addressing historical disparities is crucial. However, perfect recall may overcompensate past imbalances. For instance:

**Example 2.** Suppose Alice is the sole participant for 10 steps and accumulates a high utility. If Bob joins at step 11, perfect-recall fairness might allocate many future resources to Bob to “catch him up.” This could be viewed as unfair to Alice, as her early contributions—made in Bob’s absence—should not overly penalize her in future allocations.

## 4.3 Discounted-Recall Historical Fairness

To balance the extremes of instantaneous and perfect-recall fairness, we propose *discounted-recall historical fairness*. This paradigm introduces a temporal decay factor  $\gamma_p \in [0, 1]$  that gradually diminishes the influence of older allocations. The intuition is that while historical context is important, its influence should naturally

decay over time. Such a decay mechanism is inspired by behavioral research, which indicates that humans discount temporally distant events. Furthermore, incorporating past discounts aligns how we consider past utilities with how future rewards are treated in sequential decision-making (SDM), where temporal decay is crucial for ensuring convergence of returns and for tractable computation.

**4.3.1 Discounted Recall with Additive Utilities.** In the additive setting, past utilities are discounted and then combined with the current utility:

$$Z_i^t | \mathcal{A} = \gamma_p Z_i^{t-1} + u_i^{\mathcal{A}}. \quad (7)$$

Here,  $\gamma_p$  governs the balance between immediate and historical considerations, with  $\gamma_p = 0$  reducing to instantaneous fairness and  $\gamma_p = 1$  recovering perfect-recall fairness.

**4.3.2 Discounted Recall with Averaged Utilities.** For averaged utilities, both the accumulated utility and the effective time denominator are discounted. Let  $d_t$  denote the past-discounted denominator at time  $t$ . Then:

$$Z_i^t | \mathcal{A} = \frac{\gamma_p Z_i^{t-1} \cdot d_{t-1} + u_i^{\mathcal{A}}}{\gamma_p d_{t-1} + 1}.$$

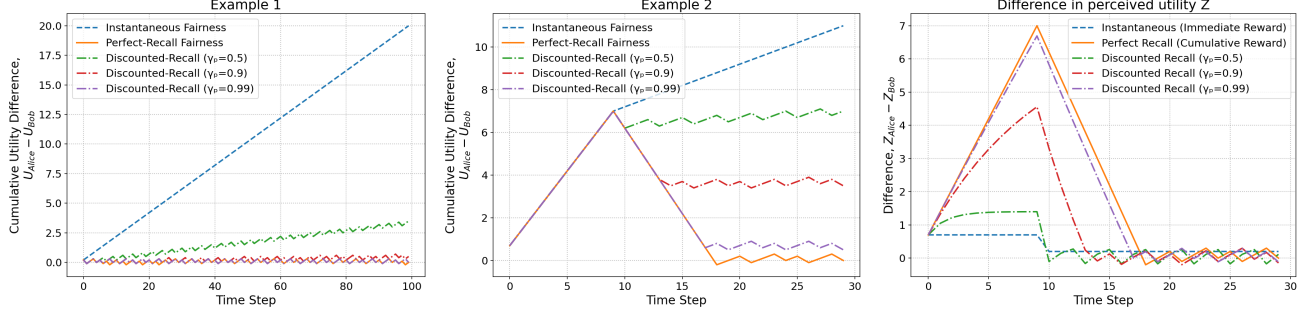
In both formulations, the computation of  $Z^t$  depends only on the previous state—specifically, the augmented state comprising  $Z^{t-1}$  (and  $d_{t-1}$  in the averaged case). This Markovian structure not only simplifies the computation but also provides a smooth interpolation between instantaneous and perfect-recall fairness, while ensuring that the state space is augmented in a tractable manner.

## 4.4 Comparison between the Different Paradigms

Figure 1 illustrates the evolution of the cumulative utility difference,  $\sum U_{\text{Alice}} - \sum U_{\text{Bob}}$ , under the three fairness approaches: Instantaneous, perfect-recall, and discounted-recall fairness, as discussed in Examples 1 and 2. The allocations are made with the MMF objective, using additive aggregation for perfect-recall and discounted-recall. In Figure 1(left), where both agents participate from the start, instantaneous fairness accumulates short-term differences leading to long-term unfairness, while perfect-recall fairness compensates by accounting for all past allocations. Discounted-recall fairness offers a tunable middle ground, with the discount factor  $\gamma_p$  controlling how quickly past utilities decay in importance.

In Figure 1(center), only Alice is active initially and Bob joins later. All approaches perform similarly in the initial phase. Instantaneous fairness keeps accumulating the imbalance regardless of the history, while both perfect-recall and discounted-recall mechanisms move towards equalizing the past imbalance. Perfect recall only starts allocating resources to Alice again after equalizing the total utility.  $\gamma_p$  serves as a tuning knob which lets us control how strongly we want the distant past to affect current allocations.

Finally, Figure 1(right) shows the inner workings of each fairness approach by plotting the perceived differences between Alice and Bob in terms of  $Z$ . Instantaneous fairness seems to always keep a low difference between the two agents locally, even as the total utility for Alice keeps rising. Perfect-recall keeps an exact track of resources, considering the distribution unfair even after many steps of allocating to Bob only. The decay of the line, particularly



**Figure 1: Comparison of cumulative utility differences under different fairness paradigms. (Left) Cumulative utility difference,  $\sum U_{\text{Alice}} - \sum U_{\text{Bob}}$ , over time for Example 1, where both agents participate from the start. (Center) Cumulative utility difference,  $\sum U_{\text{Alice}} - \sum U_{\text{Bob}}$ , over time for Example 2, where only Alice is active initially and Bob joins later. (Right) Difference in perceived utility between Alice and Bob for all three methods for Example 2. This plot shows the effect of  $\gamma_p$  on the perceived values, demonstrating how it changes the speed at which we forget past decisions, interpolating between perfect-recall and instantaneous fairness.**

visible for  $\gamma_p = 0.9$  shows how discounted-recall slowly forgets past decisions, with values close to 1 emulating longer memory.

#### 4.5 Practical Benefits of Past-Discounting

A key practical advantage of our past-discounted approach is that it bounds the cumulative utility over time. In typical non-discounted fairness frameworks (e.g., [1, 25]), the cumulative utility is computed as  $Z_i^t = \sum_{\tau=0}^t u_i^\tau$ , which grows linearly with the time horizon (i.e.,  $Z_i^t \leq (t+1)u_{\max}$  when  $u_i^\tau \in [0, u_{\max}]$ ). As a consequence, when the fairness state is augmented with these cumulative utilities, the state space expands unboundedly with time, severely hampering the scalability and learnability of the problem.

In contrast, by updating the cumulative utility with a past-discount factor  $\gamma_p \in [0, 1)$ :

$$Z_i^t = \gamma_p Z_i^{t-1} + u_i^t, \quad Z_i^0 = u_i^0,$$

we ensure that:

$$Z_i^t \leq \frac{u_{\max}}{1 - \gamma_p}.$$

**Proof Sketch:** We prove the bound by induction. For  $t = 0$ ,  $Z_i^0 \leq u_{\max} \leq \frac{u_{\max}}{1 - \gamma_p}$ . Assume  $Z_i^{t-1} \leq \frac{u_{\max}}{1 - \gamma_p}$ . Then,

$$Z_i^t = \gamma_p Z_i^{t-1} + u_i^t \leq \gamma_p \frac{u_{\max}}{1 - \gamma_p} + u_{\max} = \frac{u_{\max}}{1 - \gamma_p}.$$

Thus, the bound holds for all  $t$ .  $\square$

This boundedness ensures that, upon discretization, the number of distinct cumulative utility states per agent is fixed, only dependent on  $\gamma_p$  and independent of  $t$ , in stark contrast to the non-discounted approach, where the number of states grows approximately linearly in  $t$  per agent (and thus the joint augmented state space grows exponentially with both  $t$  and the number of agents). A bounded augmented state space is critical for the practical application of learning methods—particularly reinforcement learning—as it significantly improves convergence and reduces the sample complexity of the learning problem. Thus, past-discounting

not only provides a principled balance between short- and long-term fairness but also renders the underlying sequential decision problem computationally tractable.

## 5 CONCLUSION

In this work, we introduced a framework for incorporating past-discounted historical utilities into dynamic resource allocation—a strategy inspired by behavioral economics and moral psychology, which show that humans naturally discount the impact of distant past events [7, 23]. By applying a discount factor  $\gamma_p$  to past utilities, our method enables decision-makers to reason over accumulated utilities while effectively balancing short-term and long-term fairness considerations.

Crucially, the past-discounting approach ensures that the augmented state space remains bounded. In contrast to traditional non-discounted fairness methods—where the state space grows linearly with the time horizon and thus becomes computationally intractable—our framework yields a joint state space whose size is independent of time. This boundedness not only improves the sample complexity and convergence of reinforcement learning algorithms in multi-agent settings, but also provides a principled mechanism to manage the trade-off between immediate outcomes and historical context.

Looking ahead, future work will focus on extensive empirical evaluations to quantify the performance gains of past-discounting over non-discounted fairness in real-world allocation scenarios. We also plan to conduct human studies to assess how well our metrics align with people’s perceptions of fairness. These investigations will validate our theoretical insights and inform the design of practical, human-centered resource allocation systems.

Overall, we believe that past-discounted historical fairness offers a more expressive and versatile framework—one that deserves greater attention for achieving equitable and efficient outcomes in complex, dynamic environments.

## REFERENCES

- [1] Parand Alizadeh Alamdari, Toryn Q. Klassen, Elliot Creager, and Sheila A. McIlraith. 2024. Remembering to be fair: On non-markovian fairness in sequential decision making. In *Proceedings of the 2024 International Conference on Machine Learning (ICML)*. 906–920.
- [2] Martin Aleksandrov, Haris Aziz, Serge Gaspers, and Toby Walsh. 2015. Online fair division: Analysing a food bank problem. *arXiv preprint arXiv:1502.07571* (2015).
- [3] Javier Alonso-Mora, Samitha Samaranayake, Alex Wallar, Emilio Frazzoli, and Daniela Rus. 2017. On-demand High-capacity Ride-sharing via Dynamic Trip-vehicle Assignment. *Proceedings of the National Academy of Sciences* 114 (2017), 462–467.
- [4] Anthony B Atkinson et al. 1970. On the measurement of inequality. *Journal of Economic Theory* 2, 3 (1970), 244–263.
- [5] Eric Budish. 2011. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103.
- [6] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D Procaccia, Nisarg Shah, and Junxing Wang. 2019. The unreasonable fairness of maximum Nash welfare. *ACM Transactions on Economics and Computation* 7, 3 (2019), 1–32.
- [7] Shane Frederick, George Loewenstein, and Ted O’donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* 40, 2 (2002), 351–401.
- [8] Jiechuan Jiang and Zongqing Lu. 2019. Learning fairness in multi-agent systems. In *Proceedings of the Conference on Neural Information Processing Systems*.
- [9] Amanda R. Kube, Sanmay Das, and Patrick J. Fowler. 2019. Allocating interventions based on predicted outcomes: A case study on homelessness services. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 622–629.
- [10] Ashwin Kumar, Yevgeniy Vorobeychik, and William Yeoh. 2023. Using simple incentives to improve two-sided fairness in ridesharing systems. In *Proceedings of the International Conference on Automated Planning and Scheduling*. 227–235.
- [11] Ashwin Kumar and William Yeoh. 2025. DECAF: Learning to be fair in multi-agent resource allocation. *arXiv preprint arXiv:2502.04281* (2025).
- [12] Martin Lackner. 2020. Perpetual voting: Fairness in long-term decision making. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 2103–2110.
- [13] Mengyao Li, Bernhard Leidner, Nebojša Petrović, and Nedim Prelic. 2021. Close or distant past? The role of temporal distance in responses to intergroup violence from victim and perpetrator perspectives. *Personality and Social Psychology Bulletin* 47, 4 (2021), 657–672.
- [14] Max O Lorenz. 1905. Methods of measuring the concentration of wealth. *Publications of the American Statistical Association* 9, 70 (1905), 209–219.
- [15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602* (2013).
- [16] John F Nash et al. 1950. The bargaining problem. *Econometrica* 18, 2 (1950), 155–162.
- [17] Ariel D Procaccia and Junxing Wang. 2014. Fair enough: Guaranteeing approximate maximin shares. In *Proceedings of the ACM Conference on Economics and Computation*. 675–692.
- [18] John Rawls. 1971. A Theory of Justice. *Cambridge (Mass.)* (1971).
- [19] Amartya Sen. 2017. *Collective Choice and Social Welfare: Expanded Edition*. Penguin UK.
- [20] Sanket Shah, Meghna Lowalekar, and Pradeep Varakantham. 2020. Neural approximate dynamic programming for on-demand ride-pooling. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 507–515.
- [21] Umer Siddique, Paul Weng, and Matthieu Zimmer. 2020. Learning fair policies in multi-objective (deep) reinforcement learning with average and discounted rewards. In *Proceedings of the International Conference on Machine Learning*. 8905–8915.
- [22] Richard S. Sutton and Andrew G. Barto. 1998. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA.
- [23] Yaacov Trope and Nira Liberman. 2010. Construal-level theory of psychological distance. *Psychological Review* 117, 2 (2010), 440.
- [24] Michael JA Wohl and April L McGrath. 2007. The perception of time heals all wounds: Temporal distance affects willingness to forgive following an interpersonal transgression. *Personality and Social Psychology Bulletin* 33, 7 (2007), 1023–1035.
- [25] Matthieu Zimmer, Claire Glanois, Umer Siddique, and Paul Weng. 2021. Learning fair policies in decentralized cooperative multi-agent reinforcement learning. In *Proceedings of the International Conference on Machine Learning*. 12967–12978.