Achieving Long-term Fairness in Sequential Decision Making

Yaowei Hu, Lu Zhang

University of Arkansas
{yaoweihu, lz006}@uark.edu

Abstract

In this paper, we propose a framework for achieving long-term fair sequential decision making. By conducting both the hard and soft interventions, we propose to take path-specific effects on the time-lagged causal graph as a quantitative tool for measuring long-term fairness. The problem of fair sequential decision making is then formulated as a constrained optimization problem with the utility as the objective and the long-term and short-term fairness as constraints. We show that such an optimization problem can be converted to a performative risk optimization. Finally, repeated risk minimization (RRM) is used for model training, and the convergence of RRM is theoretically analyzed. The empirical evaluation shows the effectiveness of the proposed algorithm on synthetic and semi-synthetic temporal datasets.

Introduction

Fair machine learning has received increasing attention in the past years, especially in decision making tasks such as hiring (Caton and Haas 2020), college admissions (Zhang, Wu, and Wu 2017a) and bank loans (Johnson, Foster, and Stine 2016). Many algorithms for achieving fair decision making have been proposed based on various fairness notions (e.g. demographic parity (Zemel et al. 2013), equalized odds (Hardt, Price, and Srebro 2016) and counterfactual fairness (Kusner et al. 2017)). At present, the majority of studies on fair machine learning focus on the static or one-shot classification setting. However, in practice, decision making systems are usually operating in a dynamic manner such that the classifier makes sequential decisions over a period of time. In many situations, each decision made by the classifier may change the underlying data population and further affect subsequent decisions. For example, suppose a person applies to a bank for a loan and the bank estimates the risk of default according to his/her credit score. Then, the bank's decision on the loan application (e.g., whether to grant the loan and the interest rate assigned) may in turn affect the default risk and change the person's credit score (e.g., the credit score will decrease if the loan is granted but he/she defaults on the loan) which will affect his/her next loan application. If the bank's decision leads to a long-term decrease in the credit score, then it imposes a negative long-term effect

on future decisions for this person. Therefore, fair decision making should concern not only the fairness of a single decision but more importantly, whether a decision model can impose fair long-term effects on different groups. This notion of fairness is referred to as long-term fairness in recent studies (Liu et al. 2018; Hu and Chen 2018; Ge et al. 2021).

The challenge of achieving long-term fairness comes in two folds. Firstly, different from static settings, decisions made by models may change users' behaviors, and/or affect their status such as reputation, qualification, etc., and impact subsequent decisions via feedback loops. Without knowing how the population would be reshaped by decisions, enforcing any fairness constraint may create negative feedback loops and eventually harm fairness in the long run. Recent research has demonstrated that existing fairness criteria cannot guarantee fairness and sometimes undermine fairness even if only one time step is taken into consideration (Liu et al. 2018; Kannan, Roth, and Ziani 2019; D'Amour et al. 2020; Creager et al. 2020). Secondly, due to the feedback loops, the deployment of the decision model will cause changes in the data distribution that is originally used for training. This can be viewed as a distribution shift problem as the distribution of the training data (i.e., distribution before the model deployment) is different from the distribution of the test data (i.e., distribution after the model deployment). Ignoring the distribution shift will critically affect the achievement of long-term fairness, as long-term fairness is affected by all decisions made by the model along the time.

In this paper, we propose a framework for achieving long-term fair sequential decision making by addressing both above challenges. We model the dynamics of the decision-making process by employing Pearl's Structural Causal Model (SCM) (Pearl 2009), in which the relations among user features and decisions and how those decisions affect the data distribution can be encoded in a probabilistic graphical model. Specifically, we leverage the time-lagged causal graph (Runge et al. 2019) to describe the causal relations over time, and adopt the soft intervention (Correa and Bareinboim 2020) for modeling the model deployment and inferring its impacts on the underlying population. Then, we measure long-term fairness as the path-specific effect on the time-lagged causal graph under both the hard intervention on the sensitive attribute and the soft intervention on the predicted decisions. A constrained optimization problem is

formulated to strike a trade-off between long-term fairness and model utility, as well as certain short-term fairness requirement that may be stipulated by law or regulations. On the other hand, we show that the constrained optimization problem can be converted to a performative risk optimization problem (Perdomo et al. 2020). Then, we employ the repeated risk minimization (RRM) training technique (Perdomo et al. 2020) for dealing with the distribution shift problem. The performative optimality and stability properties of the proposed method are theoretically and empirically evaluated which shows its effectiveness.

To the best of our knowledge, this paper is the first to propose a causality-based long-fairness notion. The proposed learning framework is general such that it could incorporate different combinations of surrogate functions, utility loss functions, as well as causal paths regarding long-term fairness used to fit different applications. The experiment results show that the proposed method can achieve long-term fairness for multiple time steps, while the fairness performance deteriorates with time if no fairness constraint or static fairness constraints are used.

Related Work. Fair machine learning research in past years has been focused on static settings with one-shot decisions being made. To extend fair machine learning to dynamic settings, some efforts have been devoted to a compound decision-making process called pipeline (Bower et al. 2017; Dwork and Ilvento 2018). In pipelines, individuals may drop out at any stage and classification in subsequent stages depends on the remaining cohort of individuals. In a more general setting, decisions made in the past will affect the underlying population, and then affect future decisions (Zhang and Liu 2020). In this setting, a number of studies have demonstrated the inadequacy of static fairness approaches in various scenarios, including credit lending (Liu et al. 2018), college admission (D'Amour et al. 2020), labor market (Hu and Chen 2018). Creager et al. (2020) proposes to use causal directed acyclic graphs (DAGs) as a unifying framework to study fairness in dynamical systems, but does not propose an approach to achieve long-term fairness. On the other hand, some works (Jabbari et al. 2017; Zhang et al. 2020; Wen, Bastani, and Topcu 2021) study long-term fairness in the context of reinforcement learning. The most relevant work to this paper is (Hu et al. 2020) on fair multiple decision making, which also applies SCM and leverages soft interventions to model the deployment of decision models. However, (Hu et al. 2020) is still focused on the static fairness of each decision model separately other than the long-term fairness.

Preliminaries

Throughout this paper, variables and their values are denoted by uppercase and lowercase letters respectively, i.e., X and x. The sets of variables and their values are denoted by bold letters, i.e., $\mathbf{X}$ and $\mathbf{x}$.

Structural Causal Model

Our work is based on Pearl's structural causal models (Pearl 2009) which describe the causal mechanisms of a system

by a set of structural equations, i.e., $x = f_X(\mathbf{pa}_X, \mathbf{u}_X)$ for each X, where $\mathbf{pa}_X$ is a realization of a subset of endogenous variables, and $\mathbf{u}_X$ is a realization of a set of exogenous variables. Each causal model $\mathcal{M}$ is associated with a causal model graph $\mathcal{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ where $\mathbf{V}$ is a set of nodes and $\mathbf{E}$ is a set of directed edges for representing the direct causal relations. This paper assumes the Markovian model in which all exogenous variables are mutually independent.

Quantitatively measuring causal effects is facilitated with the (hard) intervention (Pearl 2009) which forces some variables to take certain values. Formally, the intervention that sets the value of X to x is denoted by $do(x)$. In a SCM, intervention $do(x)$ is defined as the substitution of equation $x = f_X(\mathbf{pa}_X, \mathbf{u}_X)$ with constant x. An intervention on a variable affects its descendants via causal paths. For an observed variable Y affected by X, its variant under intervention $do(x)$ is denoted by $Y(x)$. The distribution of $Y(x)$, also referred to as the post-intervention distribution of Y, is denoted by $P(Y(x))$. The soft intervention (Pearl 2009) extends the hard intervention such that it forces variable X to take functional relationship $g(\mathbf{z})$ in responding to some other variables $\mathbf{Z}$, which is denoted by σ in (Correa and Bareinboim 2020). The soft intervention substitutes equation $x = f_X(\mathbf{pa}_X, \mathbf{u}_X)$ with a new function $x = g(\mathbf{z})$. After performing the soft intervention, X will be associated with a new distribution determined by function g. In this paper, the function g is parameterized by θ, and we denote the new distribution by $P_\theta(x|\mathbf{z})$. We also denote the distribution of Y after performing the soft intervention by $P(Y(\theta))$.

Fairness-aware Classification

The classification problem is to learn a functional mapping $f : \mathcal{X} \mapsto \mathcal{Y}$ from the labeled training data $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$ where $\mathbf{x}_i \in \mathcal{X}$, $y_i \in \mathcal{Y}$ and $\mathcal{Y} = \{-1, 1\}$, by minimizing the 0-1 loss function $\mathbb{E}_{\mathbf{X}, Y}[\mathbb{1}[f(\mathbf{X}) \neq Y]]$ where $\mathbb{1}[\cdot]$ is an indicator function. In general, f is made up of another function h set up in the real number domain, i.e., $h : \mathcal{X} \mapsto \mathbb{R}$ and $\mathbb{1}[f(\mathbf{X}) \neq Y] = \mathbb{1}[Yh(\mathbf{X}) \geq 0]$. Since directly minimizing the indicator is intractable, one can replace it with a smooth and differentiable surrogate function ϕ. Then, the loss function can can be reformulated as $\mathbb{E}_{\mathbf{X}, Y}[\phi(Yh(\mathbf{X}))]$. Similarly, one can also formulate fairness constraints as smoothed expressions using surrogate functions. As a result, fair classification problems can be formulated as constrained optimization problems (Wu, Zhang, and Wu 2019; Hu et al. 2020). We follow the notations used in (Wu, Zhang, and Wu 2019; Hu et al. 2020) in our formulations.

Formulating Long-term Fairness

We start by formally formulating the long-term fairness in sequential decision making. Assume we have access to a temporal dataset $\mathcal{D} = \{(S, \mathbf{X}^t, Y^t)\}_{t=1}^l$ where S is a time-invariant protected attribute, $\mathbf{X}^t$ is a set of time-dependent unprotected attributes and Y^t is a time-dependent class label. Note that this setting can be viewed as observing the data of a set of individuals at all time steps, or a more general situation where a population is subject to the decision cycles and the data is sampled at each time step. For ease

of discussion, we assume both class label and protected attribute are binary variables, i.e., $S = \{s^+, s^-\}$ with s^+ denoting the unprotected group and s^- denoting the protected group, and $Y = \{1, -1\}$ with 1 denoting the positive decision and -1 denoting the negative decision, but proposed concepts could be extended to multiple protected attributes and multiple/continuous labels situations. A predictive decision model $h_\theta(\cdot)$ parameterized by θ is trained on $\mathcal{D}$. Then, it is deployed to make predicted decisions $\hat{Y}^t$ from $(S, \mathbf{X}^t)$ repeatedly at each time, i.e., $\hat{Y}^t = 1$ if $h_\theta(s, \mathbf{x}^t) \geq 0$ and $\hat{Y}^t = -1$ otherwise, forming a sequential decision making process. Such sequential decision making process is common in practice. For example, a bank repeatedly makes lending decisions based on applicants' profile such as credit score, income, etc., and a predictive policing algorithm repeatedly makes decision about where to send police for patrolling based on the crime discovered in the neighborhood. The ultimate goal of long-term fair machine learning is to ensure that the model $h_\theta(\cdot)$ is fair in a long-term stage denoted by t^*. In this paper, we assume there is sufficient historical training data such that $l \geq t^*$.

Causality-based Long-term Fairness

We develop the long-term fairness notion by leveraging Pearl's SCM. First, we assume a time-lagged causal graph $\mathcal{G}$ for describing the causal relationship among variables over time. In recent years, structure learning algorithms have been proposed for constructing time-lagged causal graphs from data, including both constrained-based approaches (Runge 2018a,b, 2020) and continuous optimization-based approaches (Pamfil et al. 2020; Löwe et al. 2020) which can be leveraged to learn the time-lagged causal graph from data. Figure 1 shows a typical example of the time-lagged causal graph in our settings: the edge from S to $\mathbf{X}^0$ represents the bias in the distribution of $\mathbf{X}$ due to historical reasons; the edges from S and $\mathbf{X}^t$ to Y^t represent that S and $\mathbf{X}^t$ are used as the input to compute $\hat{Y}^t$; and the edges from $\mathbf{X}^t$ and Y^t to $\mathbf{X}^{t+1}$ represent how the distribution of $\mathbf{X}$ would be reshaped via feedback after each decision.

Next, we formulate long-term fairness as path-specific effects that are transmitted in the time-lagged causal graph along certain paths. The path-specific effects reflect how the intervention affects each variable on the path in a topological order and hence are appropriate for capturing dynamics in sequential decision making. Similar to the indirect discrimination in static fair machine learning (Zhang, Wu, and Wu 2017b; Nabi and Shpitser 2018), we can also justify the use of the path-specific effect by the need to distinguish discriminatory effects from explainable effects. We consider discriminatory effects as those which are due to biased decisions made by the decision making system in the past and will continue to influence future decisions. Correspondingly, we consider explainable effects as those which are attributed to external factors and cannot be eliminated within the decision making system. To this end, we categorize unprotected attributes $\mathbf{X}$ into two disjoint subsets: irrelevant attributes $\mathbf{X}_i$ and relevant attributes $\mathbf{X}_r$. We define irrelevant attributes as those which are justifiable in deci-

sion making, and meanwhile evolved autonomously or/and altered by external factors only. We define the rest of attributes as relevant attributes, which could be unjustifiable in decision making or reshaped by the decision over time. Then, we define long-term fairness as the causal effect where the influence of the hard intervention on S is transmitted in the causal graph by passing through relevant attributes only. Note that the influence of the soft intervention on Y is still transmitted through all causal paths.

Finally, we propose to adopt soft interventions as a key technique for modeling decision model deployment and inferring its impacts on the underlying population. We treat the deployment of the decision model at each time step as to perform a soft intervention on the decision variable. More specifically, we force the structural equation associated with Y^t in the causal model to be replaced by the decision model $h_\theta(\cdot)$ that outputs $\hat{Y}^t$. Thus, the change to underlying population could be inferred as the post-intervention distribution after performing the soft intervention. Meanwhile, to quantify fairness as causal effects of the protected attribute on the decision, we perform hard intervention on the protected attribute in order to answer the "what if" question, i.e., "what would the decision be if we intervene the gender of applications to female?" As a result, we perform both hard intervention and soft intervention simultaneously for measuring long-term fairness as causal effects.

Symbolically, denote by π the set of causal paths from S to $\hat{Y}^{t^*}$ through relevant attributes $\mathbf{X}_r^1, \cdots, \mathbf{X}_r^{t^*}$ and $\hat{Y}^1, \cdots, \hat{Y}^{t^*-1}$ but not through irrelevant attributes $\mathbf{X}_i^0, \cdots, \mathbf{X}_i^{t^*}$. Meanwhile, as we conduct path-specific hard intervention on S and soft interventions on Y to deploy decision model $h_\theta(\cdot)$, we denote the post-intervention distribution of $\hat{Y}^t$ by $\hat{Y}^t(s_\pi, \theta)$ which explicitly shows that the soft intervention depends on parameters θ. Then, we can readily propose the quantitative notion for long-term fairness.

Definition 1 (Long-term Fairness). *The long-term fairness of a decision model $h_\theta(\cdot)$ is measured by $P(\hat{Y}^{t*}(s_\pi^+, \theta)) - P(\hat{Y}^{t*}(s_\pi^-, \theta))$ where π is a set of paths from S to $\hat{Y}^{t*}$ passing through $\mathbf{X}_r^1, \hat{Y}^1, \cdots, \mathbf{X}_r^{t*-1}, \hat{Y}^{t*-1}, \mathbf{X}_r^{t*}$, s_π represents the path-specific hard intervention and θ represents the soft intervention through all paths.*

Loss Function and Short-term Fairness

In addition to long-term fairness, a desired fair decision model should also satisfy two other requirements. Firstly, it is a natural desire for a predictive decision model to maximize the institution utility, e.g., the loan granting model of a bank certainly wants to maximize the expected return from loans. Secondly, the decision model should also satisfy certain short-term fairness requirement at each time step to enforce local equality, which may be stipulated by law or regulations. For example, the Equal Credit Opportunity Act, 1974, prohibits lending decisions from being influenced by race, age, religion, etc. Similar to the direct discrimination in static fair machine learning, we consider a subset of relevant attributes $\tilde{\mathbf{X}}_r \subset \mathbf{X}_r$ which are unprotected but cannot be justifiably used in the decision making either directly or
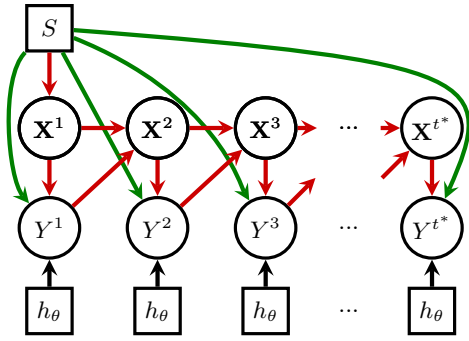
Figure 1: A time-lagged causal graph for sequential decision making. Long-term fairness is captured by paths in red, and short-term fairness is captured by paths in green.

indirectly, referred to as the redlining attributes (Zhang, Wu, and Wu 2017b). Then, we measure the short-term fairness by the causal effect of S on $\hat{Y}^t$ along paths that pass through $\tilde{\mathbf{X}}_r$, i.e., $S \to \tilde{\mathbf{X}}_r \to \hat{Y}^t$, as well as the direct edge $S \to \hat{Y}^t$ at each time step t.

We note that trade-off may exist between fairness and utility, as well as between long-term and short-term fairness. The long-term fairness focuses on remedying past discrimination existed in the system, but has no constraint on the biases in the decision at each time step. The short-term fairness, on the other hand, cares about fairness in the decision making process at each time step, but pays no attention in correcting past discrimination in the population. One should combine long-term and short-term fairness to force the decision model to take into consideration both factors and to remove discrimination in the system gradually with time. Therefore, we similarly propose quantitative notions for short-term fairness and institution utility as follows.

Definition 2 (Short-term Fairness). *The short-term fairness of a decision model $h_\theta(\cdot)$ at time t is measured by the causal effect transmitted through paths involved in time t, i.e., $P(\hat{Y}^t(s_{\pi^t}^+, \theta)) - P(\hat{Y}^t(s_{\pi^t}^-, \theta))$, where $\pi^t = \{S \to \tilde{\mathbf{X}}_r \to \hat{Y}^t, S \to \hat{Y}^t\}$ with redlining attributes $\tilde{\mathbf{X}}_r$, s_π is the path-specific hard intervention and θ represents the soft intervention.*

Definition 3 (Institution Utility). *The institution utility of decision model $h_\theta(\cdot)$ is measured by the aggregate loss given by $\sum_{t=1}^{t^*} \mathbb{E}[\mathcal{L}(Y^t, \hat{Y}^t)]$ where $\mathcal{L}(\cdot)$ is the loss function.*

Learning Fair Decision Models

After formulating related notions, we are ready to formulate the fair sequential decision making problem given a time-lagged causal graph. To ease the representation, in following discussions we consider the simplified causal graph shown in Figure 1 where only relevant attributes with no redlining attributes exist. In this case, the long-term fairness is captured by paths from S to $\hat{Y}^{t^*}$ through $\mathbf{X}^1, \hat{Y}^1, \cdots, \mathbf{X}^{t^*}$ as shown in red, and the short-term fairness is captured by the direct edge $S \to \hat{Y}^t$ at each time t as shown in green. How-

ever, all our discussions can be applied to our general formulation that includes both relevant and irrelevant features.

The goal is to learn a functional mapping $h_\theta : (\mathbf{X}^t, S) \mapsto Y^t$ parameterized with θ, i.e., $\hat{Y}^t = h_\theta(\mathbf{X}^t, S)$. Based on the discussions above, we formulate a constrained optimization problem which minimizes the loss while subject to long-term fairness and short-term fairness constrains simultaneously. The thresholds τ_l and τ_t control the strictness of constraints.

Problem Formulation 1. *The problem of fair sequential decision making is formulated as the constrained optimization:*

$$\arg\min_\theta \sum_{t=1}^{t^*} \mathbb{E}\left[\mathcal{L}(Y^t, \hat{Y}^t)\right]$$

$$s.t. \ P\left(\hat{Y}^{t^*}(s_\pi^+, \theta)=1\right) - P\left(\hat{Y}^{t^*}(s_\pi^-, \theta)=1\right) \le \tau_l,$$

$$P\left(\hat{Y}^t(s_{\pi^t}^+, \theta)=1\right) - P\left(\hat{Y}^t(s_{\pi^t}^-, \theta)=1\right) \le \tau_t, \ t=1,\cdots,t^*$$

where τ_l and τ_t are thresholds for long-term fairness and short-term fairness constraints, respectively.

Formulating as Performative Risk Optimization

Solving the optimization problem in Problem Formulation 1 is not trivial. According to the path-specific effect inference (Avin, Shpitser, and Pearl 2005) and the definition of soft intervention (Correa and Bareinboim 2020), post-intervention probability $P(\hat{Y}^{t^*}(s_\pi^+, \theta) = 1)$ is given by

$$\sum_{\mathbf{X}^1, Y^1, \cdots, \mathbf{X}^{t^*}} \left\{ P(\mathbf{x}^1 | s^+) P_\theta(y^1 | \mathbf{x}^1, s^-) \cdots \right.$$
$$\left. \cdots P(\mathbf{x}^{t^*} | \mathbf{x}^{t^*-1}, y^{t^*-1}) P_\theta(Y^{t^*}=1 | \mathbf{x}^{t^*}, s^-) \right\}, \quad (1)$$

where $P_\theta(y|\mathbf{x}, s^-)$ is a probabilistic function determined by $h_\theta(\cdot)$. As a result, $P(\hat{Y}^{t^*}(s_\pi^+, \theta) = 1)$ is a complex nonlinear function of θ, making Problem Formulation 1 difficult to solve. In the following, we show how Problem Formulation 1 is converted to a performative risk optimization problem and then propose an optimization algorithm by leveraging repeated risk minimization.

Following the notation of convex optimization of classification, we denote by ϕ a convex surrogate function. Then, we can formulate the loss function as

$$\mathcal{L}(Y^t, \hat{Y}^t) = \mathbb{1}\left[Y^t h_\theta(\mathbf{X}^t, S) < 0\right] = \phi\left(Y^t h_\theta(\mathbf{X}^t, S)\right).$$

We can also apply the surrogate function to the fairness constraints. For any t, we have

$$P\left(\hat{Y}^t(s_\pi^+, \theta)=1\right) = \sum_{\mathbf{X}^t} P\left(\mathbf{x}^t(s_\pi^+, \theta)\right) P_\theta(Y^t=1|\mathbf{x}, s^-).$$

Similar to (Hu et al. 2020), we estimate $P_\theta(Y^t = 1|\mathbf{x}, s^-)$ by first treating it as $\mathbb{1}[h_\theta(\mathbf{x}^t, s^-) \ge 0]$ and then replacing the indicator function by $\phi(\cdot)$:

$$P\left(\hat{Y}^t(s_\pi^+, \theta)=1\right) = \sum_{\mathbf{X}^t} P\left(\mathbf{x}^t(s_\pi^+, \theta)\right) \phi\left(-h_\theta\left(\mathbf{x}^t, s^-\right)\right)$$
$$= \mathbb{E}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_\pi^+, \theta))}\left[\phi\left(-h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right].$$

Similarly, we have

$$-P\left(\hat{Y}^t(s_\pi^-,\theta)\!=\!1\right) = P\left(\hat{Y}^t(s_\pi^-,\theta)\!=\!0\right) - 1$$
$$= \mathbb{E}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_\pi^-,\theta))}\left[\phi\left(h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right] - 1.$$

Then, we define utility loss $l_u(\theta)$, long-term fairness loss $l_l(\theta)$, and short-term fairness loss $l_s(\theta)$ as follows.

$$l_u(\theta) = \sum_{t=1}^{t^*} \mathbb{E}_{S,\mathbf{X}^t,Y^t \sim P(S,\mathbf{X}^t,Y^t)}\left[\phi\left(Y^t h_\theta(\mathbf{X}^t, S)\right)\right],$$

$$l_l(\theta) = \frac{1}{2}\left\{\mathbb{E}_{\mathbf{X}^{t^*} \sim P(\mathbf{X}^{t^*}(s_\pi^+,\theta))}\left[\phi\left(-h_\theta\left(\mathbf{X}^{t^*}, s^-\right)\right)\right]\right.$$
$$\left. + \mathbb{E}_{\mathbf{X}^{t^*} \sim P(\mathbf{X}^{t^*}(s_\pi^-,\theta))}\left[\phi\left(h_\theta\left(\mathbf{X}^{t^*}, s^-\right)\right)\right] - 1 - \tau_l\right\},$$

$$l_s(\theta) = \frac{1}{t^*}\sum_{t=1}^{t^*}\left\{\mathbb{E}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_{\pi^t}^+,\theta))}\left[\phi\left(-h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right]\right.$$
$$\left. + \mathbb{E}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_{\pi^t}^-,\theta))}\left[\phi\left(h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right] - 1 - \tau_t\right\}.$$

By adding the long-term and short-term fairness losses as regularization terms into the objective function, we obtain an unconstrained optimization problem as given in Problem Formulation 2. The general formulation of the performative risk optimization can be given by $\arg\min_\theta \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\theta)} l(\mathbf{Z};\theta)$ where $\mathbf{Z}$ represents the set of all attributes and outcome (Perdomo et al. 2020). Thus, Problem Formulation 2 can be considered as a performative risk optimization problem as all terms in the objective function are represented as expectations of the loss function over the distributions that depend on the loss function parameters. Compare with Problem Formulation 1, Problem Formulation 2 relaxes the fairness constraints and certain amount of violations to the constraints are allowed. However, Problem Formulation 2 can be solve more efficiently by leveraging the repeated risk minimization technique as shown in the next subsection.

Problem Formulation 2. *The problem of fair sequential decision making is reformulated as the performative risk optimization:*

$$\arg\min_\theta l(\theta) = \lambda_u l_u(\theta) + \lambda_l l_l(\theta) + \lambda_s l_s(\theta) \quad (2)$$

where λ_u, λ_l and λ_s are weight parameters and satisfy $\lambda_u + \lambda_l + \lambda_s = 1$.

The Algorithm of Repeated Risk Minimization

Repeated risk minimization (RRM) is an iterative algorithmic heuristic for solving the performative risk optimization problem. The procedure of the RRM is to start from an initial model and repeatedly find a model that minimizes the loss function on the distribution resulting from the previous model, which can symbolically represented as the update

rule $\theta_{i+1} = \arg\min_\theta \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\theta_i)} l(\mathbf{Z};\theta)$ (Perdomo et al. 2020). The RRM converges if the model that minimizes the loss remains unchanged from the previous model, i.e., $\theta_{i+1} = \theta_i$.

To implement the RRM algorithm in our context with three different loss terms, we sample different distributions at each iteration. For computing $l_u(\theta)$, the data distribution does not change with the deployment of new models, and we always use the original dataset $\mathcal{D}$ to compute $l_u(\theta)$. For computing $l_l(\theta)$, the data distribution follows the post-intervention distribution $P(\mathbf{X}^{t^*}(s_\pi^+,\theta))$ (resp. $P(\mathbf{X}^{t^*}(s_\pi^-,\theta))$). Thus, we sample the data according to the inference formula that is similar to Eq. (1) where a smooth probabilistic function $P_\theta(y|\mathbf{x},s)$ is used. Specifically, we first sample $\mathbf{X}^1$ according to the distribution $P(\mathbf{X}^1|s^+)$ (resp. $P(\mathbf{X}^1|s^-)$), and sample the decision for each sample according to $P_\theta(Y^1|\mathbf{x}^1,s^-)$. Then, we sample $\mathbf{X}^2$ according to the distribution $P(\mathbf{X}^2|\mathbf{X}^1,Y^1)$ upon the samples obtained at the first time step. We repeat this process until time t^* to obtain samples for $\mathbf{X}^{t^*}$ for computing $l_l(\theta)$. For computing $l_s(\theta)$, we similarly sample the distributions $P(\mathbf{X}^t(s_{\pi^t}^+,\theta))$ and $P(\mathbf{X}^t(s_{\pi^t}^-,\theta))$ for each time step t. The procedure of our algorithm starts from an initial model h_{θ_0} directly trained on $\mathcal{D}$, and repeatedly train the model on the re-sample data at each iteration, until the model converges to performative stability. The pseudocode of this procedure is described in Algorithm 1.

Algorithm 1: Repeated Risk Minimization

Input : Dataset $\mathcal{D} = \{(S,\mathbf{X}^t,Y^t)\}_{t=1}^l$, time -lagged causal graph $\mathcal{G}$, convergence threshold δ
Output: The stable model h_θ

1 Train a classifier on $\mathcal{D}$ according to Eq. (2) without the soft intervention to obtain the initial parameter θ_0;
2 $i \leftarrow 0$;
3 **repeat**
4   Sampled the post-intervention distributions $P\left(\mathbf{X}^{t^*}(s_\pi^+,\theta_i)\right)$ and $P\left(\mathbf{X}^{t^*}(s_\pi^-,\theta_i)\right)$;
5   Sampled the post-intervention distributions $P\left(\mathbf{X}^t(s_\pi^+,\theta_i)\right)$ and $P\left(\mathbf{X}^t(s_\pi^-,\theta_i)\right)$ for each t;
6   Minimize $l(\theta)$ according to Eq. (2) to obtain θ_{i+1};
7     $\triangle = \|\theta_{i+1} - \theta_i\|_2$;
8     $i \to i + 1$;
9 **until** $\triangle < \delta$;
10 $\theta \leftarrow \theta_i$;
11 **return** h_θ;

Convergence Analysis of RRM

We now conduct performative stability analysis for our algorithm. The convergence of the RRM algorithm depends on the smoothness and convexity of the loss function, as well as the sensitivity of the distribution to the parameters (Perdomo et al. 2020). Specifically, given a general RRM formulation $\theta_{i+1} = \arg\min_\theta \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\theta_i)} l(\mathbf{Z};\theta)$, if loss function $l(\cdot)$ is β-jointly smooth and γ-strongly convex, and distribution $\mathcal{D}(\theta)$ is ε-sensitive, then the RRM converges to a stable point if

$\varepsilon < \frac{\beta}{\gamma}$. We similarly analyze these factors for our problem and then give the theoretical convergence result.

Lemma 1. *If the surrogated loss function $(\phi \circ h)(\cdot)$ is γ-strongly convex, then $f(\cdot)$ is γ-strongly convex.*

Lemma 1 can be directly proven according to the sum rule of the gradient.

Next, we study the sensitivity of the distributions. Consider the distribution $P(\mathbf{X}^t(s_\pi, \theta))$ for any t. Its sensitivity to θ depends on to what extend the decisions will impact the attributes via the feedback loop. By assuming that the change of the distribution over the attributes in respond to the change of θ is bounded by a constant, we present following lemma. Please refer to the appendix for detailed proof.

Definition 4. *For any t, attributes $\mathbf{X}^{t+1}$ are c-sensitive if*

$$\| \sum_{Y^t} \nabla_\theta P_\theta(y^t|\mathbf{x}^t, s) P(\mathbf{x}^{t+1}|\mathbf{x}^t, y^t) \| \le c \sum_{Y^t} P(\mathbf{x}^{t+1}|\mathbf{x}^t, y^t).$$

Lemma 2. *For any t, suppose that $\mathbf{X}^{t+1}$ are c-sensitive, then distribution $P(\mathbf{X}^t(s_\pi, \theta))$ is ε-sensitive with $\varepsilon \le 2mc(t-1)$, where m is the maximum ground distance between two values of $\mathbf{X}^t$.*

After introducing the above two lemmas, we now present our main theoretical result.

Theorem 1. *Suppose that surrogated loss function $(\phi \circ h)(\cdot)$ is β-jointly smooth and γ-strongly convex, and suppose that $\mathbf{X}^{t+1}$ are c-sensitive for any t, then the repeated risk minimization converges to a stable point at a linear rate, if $2mc(t^* - 1) < \frac{\beta}{\gamma}$.*

The proof of Theorem 1 is based on Theorem 3.5 in (Perdomo et al. 2020). Please refer to the appendix for details. In practice, this theoretical criterion of convergence may be difficult to meet. However, our experimental results show that our algorithm can converge under reasonable conditions.

Experiments

We conduct experiments on both synthetic and semi-synthetic temporal datasets to evaluate the proposed algorithm. We show that our algorithm is effective in achieving both long-term and short-term fairness, while previous fair algorithms that do not consider the dynamics in sequential decision making actually do not mitigate or even exacerbate the short-term or long-term fairness. We consider three baselines in the experiments which treat the whole temporal dataset as a static dataset and train the decision model on it. Fairness constraints are added following the technique proposed in (Wu, Zhang, and Wu 2019).

- **Logistic Regression (LR)**: An unconstrained logistic regression model which takes user features and labels from all time steps as inputs and outputs.
- **Fair Model with Demographic Parity (FMDP)**: On the basis of the logistic regression model, fairness constraint is added to achieve demographic parity.
- **Fair Model with Equal Opportunity (FMEO)**: On the basis of the logistic regression model, fairness constraint is added to achieve equal opportunity.

Datasets

Synthetic Data. We simulate a process of bank loans following the time-lagged causal graph depicted in Figure 1, where S is the protected attribute like race, $\mathbf{X}^t$ represents the financial status of applicants, and Y^t represents the decisions about whether or not to grant loans. At $t = 1$, we generate samples where both values of S are sampled with the equal probability, and the values of $\mathbf{X}^1$ are sampled using two different Gaussian distributions according to the value of S. Then at each time t, we sample predicted decisions $\hat{Y}^t$ and the values of $\mathbf{X}^{t+1}$ as follows. Consider a ground-truth decision model $h_{\theta^*}(\cdot)$ for deciding the probability of whether an individual would default on a loan, given by $\sigma(h_{\theta^*}(\cdot))$ where $\sigma(\cdot)$ is the sigmoid function. Then, we sample the predicted decision $\hat{Y}^t$ (as well as the actual repayment Y^t which is sampled separately) from $\sigma(h_{\theta^*}(\cdot))$ as:

$$P(\hat{Y}^t) = \sigma(h_{\theta^*}(\mathbf{X}^t, S)), \ \hat{Y}^t \sim 2 \cdot \text{Bernoulli}(P(\hat{Y}^t)) - 1.$$

Then, $\mathbf{X}^{t+1}$ is generated according to the update rule below:

$$\mathbf{X}^{t+1} = \begin{cases} \mathbf{X}^t - \epsilon \cdot \theta^t + b & \hat{Y}^t = 1, Y^t = -1 \\ \mathbf{X}^t + \epsilon \cdot \theta^t + b & \hat{Y}^t = 1, Y^t = 1 \\ \mathbf{X}^t + b & \hat{Y}^t = -1 \end{cases} \quad (3)$$

where ϵ is a parameter that controls the sensitivity of the update to the predicted decisions, and $b = S \cdot b_1 + (1-S) \cdot b_0$ is a small base increment at each time step. In the simulation process, we generate a 5-step synthetic dataset with 5000 samples where parameters are set as $\epsilon = 0.5$, $b_0 = 0.2$, $b_1 = 1.0$.

Semi-synthetic Data. We use the Taiwan credit card dataset (Yeh and Lien 2009) as the initial data at $t = 1$. To form a balanced dataset, we extract 3000 samples and choose two features PAY_AMT1 and PAY_AMT2 that are appropriate in fitting into our update rule. Then, we generate a 4-step dataset using the same update rule as shown above.

Training and Evaluation

We conduct the training process following the RRM algorithm. At each iteration, we sample the data according to the current decision model and the causal graph. Similar to the data generation process, predicted decisions are sampled according to the probability given by $\sigma(h_\theta(\cdot))$, and the feature values are sampled according to Eq. (3). In our experiments, we assume that the true update rule is known in order to remove errors introduced by causal graph construction. In practice, the causal graph learned from data may introduce additional errors.

We then design an evaluation process which simulates the real model deployment procedure and feedback loops. At each time step t, we use the trained decision model $h_\theta(\cdot)$ to make decisions $\hat{Y}^t$, and use the ground-truth model $h_{\theta^*}(\cdot)$ to determine the repayment Y^t. The accuracy is measured by comparing $\hat{Y}^t$ and Y^t, the long-term fairness is measured based on the distribution of $\hat{Y}^{t^*}$ in the evaluation, and the short-term fairness is measured based on the distribution of $\hat{Y}^t$ at different time steps according to proposed definitions.

Table 1: Accuracy, short-term and long-term fairness of different algorithms on the synthetic dataset.

Alg.	Metric	Time steps				
		$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
RL	Acc	0.912	0.894	0.917	0.921	0.917
	Short	0.152	0.160	0.166	0.164	0.174
	Long	0.058	0.117	0.173	0.246	0.340
FMDP	Acc	0.735	0.706	0.704	0.708	0.725
	Short	0.212	0.216	0.224	0.220	0.232
	Long	0.180	0.306	0.376	0.431	0.481
FMEO	Acc	0.829	0.790	0.795	0.800	0.814
	Short	0.010	0.010	0.010	0.014	0.020
	Long	0.080	0.122	0.190	0.276	0.352
Ours	Acc	0.801	0.754	0.729	0.707	0.692
	Short	0.012	0.008	0.012	0.008	0.002
	Long	0.040	0.024	0.020	0.012	0.002

Table 2: Accuracy, short-term and long-term fairness of different algorithms on the semi-synthetic dataset.

Alg.	Metric	Time steps			
		$t=1$	$t=2$	$t=3$	$t=4$
RL	Acc	0.828	0.826	0.841	0.816
	Short	0.015	0.018	0.021	0.012
	Long	0.038	0.088	0.243	0.433
FMDP	Acc	0.830	0.843	0.846	0.841
	Short	0.063	0.066	0.075	0.069
	Long	0.038	0.076	0.223	0.397
FMEO	Acc	0.824	0.830	0.830	0.813
	Short	0.072	0.075	0.087	0.078
	Long	0.006	0.045	0.156	0.295
Ours	Acc	0.648	0.648	0.680	0.687
	Short	0.006	0.006	0.003	0.006
	Long	0.064	0.043	0.016	0.003

Implementation Details.[1] For baselines FMDP and FMEO, they are formulated as constrained optimization forms which are directly solved by the CVXPY package (Diamond and Boyd 2016). For our algorithm, we use the logistic loss function for the surrogate function ϕ and the linear model for the decision model. All algorithms use the l_2-regularization which can equip the logistic loss function with strong convexity. In our algorithm, ReLU activation function is adopted to ensure that the fairness constraints are always non-negative, and we adopt PyTorch (Paszke et al. 2019) to implement optimization with Adam optimizer.

Results

The results of the accuracy and fairness of the baselines and our algorithm on the synthetic dataset are shown in Table 1. As can be seen, our algorithm achieves the short-term fair-

---

[1]The code and hyperparameter settings are available online: https://github.com/yaoweihu/Achieving-Long-term-Fairness.
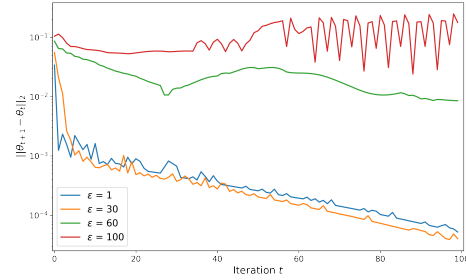

Figure 2: The convergence results for different values of ϵ on the synthetic dataset.

ness at all time steps. More importantly, the long-term fairness is improved with time and approaches zero at $t = 5$. For other baselines, there is a clear trend that the long-term fairness continuously accumulates with time. This demonstrates that static fairness notions may harm fairness in the long run. The short-term fairness remains stable with time as it shows the bias in the model that is related to the protected attribute. The experiments on the semi-synthetic dataset produce similar results as shown in Table 2. We also observe a trade-off between accuracy and fairness meaning that some accuracy needs to be sacrificed in order to achieve fairness.

We also plot in Figure 2 the convergence results of our algorithms for different ϵ values. As mentioned earlier, the value of ϵ controls the sensitivity of $\mathbf{X}^{t+1}$ to the update of θ. Figure 2 shows that our algorithm converges when the value of ϵ is reasonably small, which is consistent with the results in (Perdomo et al. 2020). We observe similar results on the semi-synthetic dataset.

Conclusions and Future Work

We proposed a framework to achieve long-term fairness in sequential decision making. The decision-making process was modeled by a time-lagged causal graph, in which the hard intervention was performed on the protected attribute and soft interventions were performed on the decisions. We measured both long-term and short-term fairness as path-specific effects. The problem of fair sequential decision making was formulated as a performative risk optimization problem, and repeated risk minimization is adopted to train the model on the datasets sampled from post-intervention distributions. The convergence of the proposed algorithm is analyzed theoretically. Finally, we verify the effectiveness of the proposed framework and algorithm by comparing it with the baselines on two synthetic datasets.

Path-specific effects may be unidentifiable from the observational data if the "kite structure" presents in the causal graph (Avin, Shpitser, and Pearl 2005). The long-term fairness loss term $l_l(\theta)$ cannot be accurately estimated if $P(\mathbf{X}^{t^*}(s_\pi^+, \theta))$ is unidentifiable if the paths in π form a "kite structure". We will adopt the bounding technique proposed in (Zhang, Wu, and Wu 2018) for unidentifiable path-specific quantify, compute the lower and upper bounds of $P(\mathbf{x}^{t^*}(s_\pi^+, \theta))$ for each $\mathbf{x}$, and then obtain the upper bound of $l_l(\theta)$. We leave this to our future work.

Ethics Statement

Our research could benefit decision makers for achieving long-term fairness and balancing the trade-off between fairness and accuracy. The proposed method relaxes the constrained optimization problem (Problem Formulation 1) to an unconstrained optimization problem (Problem Formulation 2). There may be gaps between the two problems, i.e., the optimal solution to the unconstrained optimization problem may not be the optimal solution to the original constrained one. This may potentially result in solutions that achieve compromised fairness which is lower than user requirements.

Acknowledgments

References

Avin, C.; Shpitser, I.; and Pearl, J. 2005. Identifiability of path-specific effects.

Bower, A.; Kitchen, S. N.; Niss, L.; Strauss, M. J.; Vargas, A.; and Venkatasubramanian, S. 2017. Fair pipelines. *arXiv preprint arXiv:1707.00391*.

Caton, S.; and Haas, C. 2020. Fairness in Machine Learning: A Survey. *ArXiv*, abs/2010.04053.

Correa, J.; and Bareinboim, E. 2020. A calculus for stochastic interventions: Causal effect identification and surrogate experiments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 10093–10100.

Creager, E.; Madras, D.; Pitassi, T.; and Zemel, R. 2020. Causal modeling for fairness in dynamical systems. In *International Conference on Machine Learning*, 2185–2195. PMLR.

D'Amour, A.; Srinivasan, H.; Atwood, J.; Baljekar, P.; Sculley, D.; and Halpern, Y. 2020. Fairness is not static: deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 525–534.

Diamond, S.; and Boyd, S. 2016. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83): 1–5.

Dwork, C.; and Ilvento, C. 2018. Fairness under composition. *arXiv preprint arXiv:1806.06122*.

Ge, Y.; Liu, S.; Gao, R.; Xian, Y.; Li, Y.; Zhao, X.; Pei, C.; Sun, F.; Ge, J.; Ou, W.; et al. 2021. Towards Long-term Fairness in Recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, 445–453.

Hardt, M.; Price, E.; and Srebro, N. 2016. Equality of Opportunity in Supervised Learning. In *NIPS*.

Hu, L.; and Chen, Y. 2018. A short-term intervention for long-term fairness in the labor market. In *Proceedings of the 2018 World Wide Web Conference*, 1389–1398.

Hu, Y.; Wu, Y.; Zhang, L.; and Wu, X. 2020. Fair Multiple Decision Making Through Soft Interventions. *Advances in Neural Information Processing Systems*, 33.

Jabbari, S.; Joseph, M.; Kearns, M.; Morgenstern, J.; and Roth, A. 2017. Fairness in reinforcement learning. In *International Conference on Machine Learning*, 1617–1626. PMLR.

Johnson, K. D.; Foster, D. P.; and Stine, R. A. 2016. Impartial predictive modeling: Ensuring fairness in arbitrary models. *Statistical Science*, 1.

Kannan, S.; Roth, A.; and Ziani, J. 2019. Downstream effects of affirmative action. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 240–248.

Kusner, M. J.; Loftus, J. R.; Russell, C.; and Silva, R. 2017. Counterfactual Fairness. In *NIPS*.

Liu, L. T.; Dean, S.; Rolf, E.; Simchowitz, M.; and Hardt, M. 2018. Delayed impact of fair machine learning. In *International Conference on Machine Learning*, 3150–3158. PMLR.

Löwe, S.; Madras, D.; Zemel, R.; and Welling, M. 2020. Amortized causal discovery: Learning to infer causal graphs from time-series data. *arXiv preprint arXiv:2006.10833*.

Nabi, R.; and Shpitser, I. 2018. Fair inference on outcomes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Pamfil, R.; Sriwattanaworachai, N.; Desai, S.; Pilgerstorfer, P.; Georgatzis, K.; Beaumont, P.; and Aragam, B. 2020. Dynotears: Structure learning from time-series data. In *International Conference on Artificial Intelligence and Statistics*, 1595–1605. PMLR.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; and Chintala, S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Wallach, H.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 32*, 8024–8035. Curran Associates, Inc.

Pearl, J. 2009. *Causality*. Cambridge university press.

Perdomo, J.; Zrnic, T.; Mendler-Dünner, C.; and Hardt, M. 2020. Performative prediction. In *International Conference on Machine Learning*, 7599–7609. PMLR.

Runge, J. 2018a. Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7): 075310.

Runge, J. 2018b. Conditional independence testing based on a nearest-neighbor estimator of conditional mutual information. In *International Conference on Artificial Intelligence and Statistics*, 938–947. PMLR.

Runge, J. 2020. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In *Conference on Uncertainty in Artificial Intelligence*, 1388–1397. PMLR.

Runge, J.; Nowack, P.; Kretschmer, M.; Flaxman, S.; and Sejdinovic, D. 2019. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5(11): eaau4996.

Wen, M.; Bastani, O.; and Topcu, U. 2021. Algorithms for Fairness in Sequential Decision Making. In *International Conference on Artificial Intelligence and Statistics*, 1144–1152. PMLR.

Wu, Y.; Zhang, L.; and Wu, X. 2019. On convexity and bounds of fairness-aware classification. In *The World Wide Web Conference*, 3356–3362.

Yeh, I.-C.; and Lien, C.-h. 2009. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2): 2473–2480.

Zemel, R.; Wu, Y.; Swersky, K.; Pitassi, T.; and Dwork, C. 2013. Learning fair representations. In *International conference on machine learning*, 325–333. PMLR.

Zhang, L.; Wu, Y.; and Wu, X. 2017a. Achieving Non-Discrimination in Data Release. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Zhang, L.; Wu, Y.; and Wu, X. 2017b. A Causal Framework for Discovering and Removing Direct and Indirect Discrimination. In *IJCAI*.

Zhang, L.; Wu, Y.; and Wu, X. 2018. Causal modeling-based discrimination discovery and removal: Criteria, bounds, and algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 31(11): 2035–2050.

Zhang, X.; and Liu, M. 2020. Fairness in learning-based sequential decision algorithms: A survey. *arXiv preprint arXiv:2001.04861*.

Zhang, X.; Tu, R.; Liu, Y.; Liu, M.; Kjellström, H.; Zhang, K.; and Zhang, C. 2020. How do fair decisions fare in long-term qualification? *arXiv preprint arXiv:2010.11300*.