SCIENCE CHINA
Information Sciences

• **RESEARCH PAPER** •

Influence and Rehearsal Learning: Towards Avoiding Undesired Future

Zhi-Hua ZHOU

National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023*, China;*
School of Artificial Intelligence, Nanjing University, Nanjing 210023*, China*

Abstract While Machine Learning (ML) has achieved remarkable success in predictive tasks, many real-world scenarios require proactive decision-making. A pervasive and fundamental challenge is the Avoiding Undesired Future (AUF) problem: When an ML model predicts that something undesired is going to happen, what action(s) can be taken to avoid it? Statistical *association* is inadequate for this task since it lacks the forward-looking capability required for decision-making. *Causality*, though theoretically rigorous, is often hard to identify in practice and, even when identifiable, may not always yield effective guidance, as will be explained in this article. *Influence*, a pragmatic concept between association and causality, has been proposed and established recently as the real necessity of decision-making, and various *rehearsal learning* methods have been developed to discern influence relations. This article provides an overview of this line of research. It first elucidates the distinctions between influence, association and causality, and then briefly reviews rehearsal learning methods categorized according to whether interactions are needed and whether explicit graph structures are needed. Measures of influence strength are also introduced, followed by a discussion of open challenges and future research directions.

Keywords Influence, Rehearsal Learning, Machine Learning, Decision Making, Artificial Intelligence

Citation

1 Introduction

Machine Learning (ML) has achieved great success in predictive tasks [1, 2]. Note that, beyond mere prediction, many practical scenarios require effective decision-making. A pervasive and fundamental challenge in this context is the *Avoiding Undesired Future* (AUF) problem [3]: When an ML model predicts that something undesired is going to happen, what action(s) can be taken to avoid it?

Let $\mathbf{X}$ denote the observed variables used to predict the future outcomes $\mathbf{Y}$ with an ML model h. Given an instance $\boldsymbol{x}$, h outputs a warning signal $\hat{y} = h(\boldsymbol{x})$ which is outside the desired region $\mathcal{S}$ pre-specified by the user. Suppose that there is an intermediate stage where one can make actionable decisions to influence $\mathbf{Y}$; e.g., after a sales prediction made at the beginning of a month, there is an intermediate time window before the month-end for the sales manager to take action. Assume that the decisions are in the form of altering some variables with fixed values, e.g., setting the sales discount to 10%; such decisions are called *alterations* henceforth. Let $\mathbf{Z}$ denote the variables in the intermediate stage. The AUF problem is then how to alter variables in $\mathbf{Z}$ so that the future $\mathbf{Y}$ can be shifted to fall into $\mathcal{S}$.

Addressing the AUF problem requires identifying decision variables that can influence the decision outcome. Relying solely on statistical *association* is insufficient since it lacks the *forward-looking* capability necessary for effective decisions, i.e., enabling inference about the consequences of potential actions. For instance, while a crisp knocking sound of a watermelon usually indicates that it is quite sweet, artificially manipulating the knocking sound cannot alter the sweetness; this is because both the crisp knocking sound and the high sweetness are simply results of the watermelon being ripe. Consequently, statistical association patterns are crucial for prediction, but usually fail to provide effective decision guidance.

Email: zhouzh@nju.edu.cn

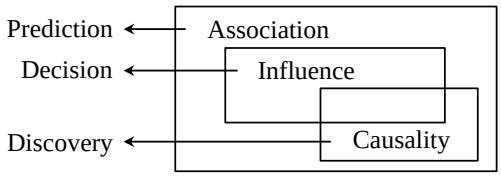


Figure 1 Association, causality, and influence.

A possible alternative is to leverage *causality*, which seeks to characterize underlying generating mechanisms for reasoning about causes and evaluating intervention or counterfactual effects [4,5]. If a variable has a causal impact on the decision outcome, altering it may effectively avoid the undesired future. However, causality is intrinsically hard to identify in practice as its identification typically relies on very strict assumptions that do not necessarily hold in real-world tasks [6]. In addition, causal variables are not always actionable, and they are of little help for AUF if they cannot be altered.

The above observations motivate a fundamental question: Do we really need full causal understanding to make effective AUF decisions? Recognizing that association is primarily suited for prediction and causality aims to reveal underlying mechanisms, Zhou [3, 7] proposed *influence*, a pragmatic concept between association and causality, tailored for decision-making. Discerning influence relations is essential for addressing the AUF problem, because they provide the forward-looking capability necessary for effective decisions.

To discern influence relations, *rehearsal learning* is proposed [3, 7]. This paradigm operates in an *induced world* to explore and identify actions that can effectively prevent undesired outcomes. A variety of rehearsal learning methods have been developed, demonstrating the feasibility of discerning influence relations that are helpful for guiding decision-making. Note that explorations in the induced world won't change the real world; this is fundamentally important for the AUF decision. Otherwise, if the real world has been improperly altered by inadequate operations, it cannot be restored to its original state, resulting in the miserable miss of adequate decisions that might have existed. For this reason, trial-and-error interaction with the real world, as used in online exploratory reinforcement learning, is less suitable for AUF decisions.

This article presents an overview of research about influence and rehearsal learning. Section 2 elucidates the distinctions between influence, association and causality. Section 3 briefly reviews rehearsal learning methods categorized according to whether interactions are needed and whether explicit graph structures are needed. Section 4 introduces influence measures. Section 5 clarifies the differences between rehearsal learning and seemingly relevant decision-making research. Finally, Section 6 presents concluding remarks.

2 Association, Causality, and Influence

Association serves as the foundation for predictive tasks, aiming to characterize statistically accompanying patterns. It can be quantified by various measures, ranging from Pearson correlation for linear relations to the Hilbert-Schmidt independence criterion (HSIC) [8] or mutual information [9] for non-linear relations. Association, however, is inherently inadequate for decision-making, as it captures mere statistical patterns without anticipating the consequences of actions.

Causality seeks to characterize the underlying generating mechanisms, suitable for scientific discovery of the laws of nature. It is typically quantified by measures such as average causal effect [10] and causal strength [11]. While causality provides forward-looking insights to some extent, identifying causal relations in practice is intrinsically hard, as it often requires assumptions whose suitability to a task is practically hard to be judged before the task is really solved [5].

Decision tasks, such as addressing the AUF problem, require a relation that provides a forward-looking capability. A deeper examination of the human decision-making process, however, reveals that it is not strictly tied to causality. Zhou [3] proposed this perspective in light of three primary considerations: First, in practice, humans can often make effective decisions without full causal understanding, or even when their understanding is flawed. Second, real decision tasks often involve open and dynamic environments, where relations among variables may change, so a causal model established in a snapshot of the changing environment may not remain adequate for decision-making. Third, even in a closed and static environment where the causality is known, it can still be ineffective for decision-making. For example, if the identified

causal factors are not actionable (i.e., they cannot be altered by the decision-maker), they evidently offer no leverage for preventing undesired outcomes.

Recognizing that neither association nor causality is essential for decision-making, Zhou [3,7] proposed the concept of *influence* relation. Generally, influence concerns whether altering Z can lead to a change in the future outcome Y in a decision task. As Figure 1 illustrates, prediction relies on association, scientific discovery relies on causality, whereas decision-making relies on influence. Specifically, if a variable Z has a causal or influence relation with the outcome Y, they are associated. An influence relation does not strictly require a comprehensive causal understanding, just like human beings make effective decisions without requiring a full understanding of the underlying causality. The most interesting point is that when Z has a causal impact on Y, it does not necessarily have an influence on Y. This is easy to understand when Z is unactionable, whereas it is also true even when Z is actionable, as what will be detailed in Section 4.

Note that the influence relation was initially called rehearsal relation in [3]. Later, Zhou [7] clarified that the property/relation is influence, while the process to discern/learn influence is rehearsal.

3 Rehearsal Learning

Rehearsal is the process to discern influence relations. Recently, various rehearsal learning methods have been proposed. These methods typically focus on learning an *influence model* from data, which captures the influence relations and facilitates exploration in an *induced world*. By performing rehearsals, i.e., simulating the consequences of various decisions either explicitly or implicitly within this model, these approaches can identify actions for AUF that prevent undesired outcomes with a high success rate.

In the following, uppercase letters denote variables and lowercase letters denote their specific values. The objective is to identify alterations on the *alterable variable(s)* $\mathbf{Z}_A \subseteq \mathbf{Z}$, determined by domain expertise, that steer the future outcome $\mathbf{Y}$ toward the desired region $\mathcal{S}$ pre-specified by the user. Formally, this entails evaluating the *AUF probability* $\Pr(\mathbf{Y} \in \mathcal{S} \mid \mathbf{X} = \mathbf{x}, \mathbf{Z}_A \overset{a}{=} \mathbf{z}_A)$ given the observed context $\mathbf{x}$ and any candidate alteration $\mathbf{z}_A$, and then determining the alteration value that maximizes the AUF probability, where $\overset{a}{=}$ denotes a rehearsal alteration, which is an exploration in the induced world rather than the real world. The AUF probability can be simplified to $\Pr(\mathbf{Y} \in \mathcal{S} \mid \mathbf{x}, \mathring{\mathbf{z}}_A)$ for brevity, where $\mathring{\mathbf{z}}_A$ denotes altering $\mathbf{Z}_A$ to value $\mathbf{z}_A$ by rehearsal.

3.1 Interaction-based Rehearsal Learning with Explicit Graphs

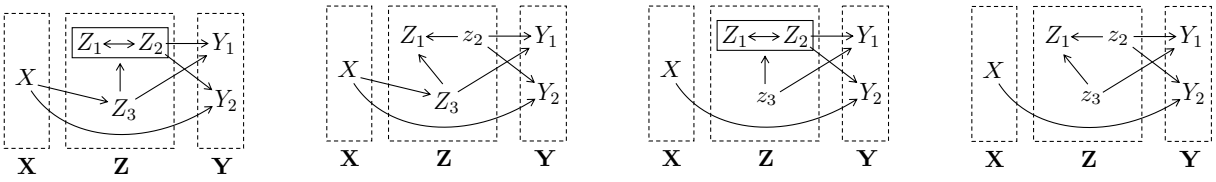
Qin et al. [12] proposed the first rehearsal learning approach which learns the influence model and makes decisions through interaction. Central to this approach are *Structural Rehearsal Models* (SRMs), a type of influence model that describes how variables are generated and how alterations propagate to affect downstream outcomes. An SRM consists of a set of potentially time-varying *rehearsal graphs*, denoted by G_t, and their associated structural equations, parameterized by $\boldsymbol{\theta}_t$. Rehearsal graphs allow both directional and bi-directional edges. Directional edges depict the generation ordering of variables and the bi-directional ones indicate interrelated variables that are mutually influenced. Figure 2(a) depicts an example where Z_1 and Z_2 are mutually influenced. The formal definition of a rehearsal graph follows, beginning with the foundational concepts of mixed graphs and bi-directional cliques.

Definition 1 (Mixed graph)**.** Let $G = (\mathbf{V}, \mathbf{E})$ be a graph, where $\mathbf{V}$ denotes the vertices and $\mathbf{E}$ the edges. G is a *mixed graph* if for any distinct vertices $u, v \in \mathbf{V}$, there is at most one edge connecting them, and the edge is either *directional* ($u \to v$ or $u \leftarrow v$) or *bi-directional* ($u \leftrightarrow v$).

Definition 2 (Bi-directional clique)**.** A *bi-directional clique* $C = (\mathbf{V}^c, \mathbf{E}^c)$ of a mixed graph $G = (\mathbf{V}, \mathbf{E})$ is a complete subgraph induced by $\mathbf{V}^c \subseteq \mathbf{V}$ such that any edge $e \in \mathbf{E}^c$ is bi-directional. C is *maximal* if adding any other vertex does not induce a bi-directional clique.

Definition 3 (Rehearsal graph)**.** Let $G = (\mathbf{V}, \mathbf{E})$ be a mixed graph. Let $\{C_i\}_{i=1}^{l}$ denote all maximal bi-directional cliques of G, where $C_i = (\mathbf{V}_i^c, \mathbf{E}_i^c)$. G is a *rehearsal graph* if and only if:

1. $\mathbf{V}_i^c \cap \mathbf{V}_j^c = \emptyset$ for any $i \neq j$.
2. $\forall i \in [l]$, $u \in \mathbf{V} \setminus \mathbf{V}_i^c$, if there is any edge pointing from u to $\mathbf{V}_i^c$, then $\forall v \in \mathbf{V}_i^c$, $u \to v$.
3. The directional edges permit a topological ordering for $\{C_i\}_{i=1}^{l}$.



(a) rehearsal graph G (b) $G^{\mathbf{Z}_2 \overset{a}{=} z_2}$ (c) $G^{\mathbf{Z}_3 \overset{a}{=} z_3}$ (d) $G^{\mathbf{Z}_2 \overset{a}{=} z_2, \mathbf{Z}_3 \overset{a}{=} z_3}$

Figure 2 Figure 2(a) is a rehearsal graph, and Figure 2(b)∼ Figure 2(d) illustrate the corresponding alteration graphs with different alterations. Note that when an alteration occurs to certain variables, all incoming arrows to those variables are removed while other graph structures are maintained.

Each vertex in a rehearsal graph corresponds to a variable. The variables are generated following the ordering depicted by the directional edges. The variables in a common maximal bi-directional clique, where only bi-directional edges exist, are mutually influenced instead of having a fixed generating order. Given a rehearsal graph G, structural equations accompanying the rehearsal graph are defined over cliques $\{(\mathbf{V}_i^c, \mathbf{E}_i^c)\}_{i=1}^l$. Let $\mathbf{PA}_i^G \triangleq \{u \mid \exists v \in \mathbf{V}_i^c, u \to v \text{ in } G\}$ denote parents of the i-th maximal bi-directional clique. Suppose that the structural equation describing the generation process of $\mathbf{V}_i^c$ is additive, i.e.,

$$\mathbf{V}_i^c := f_i(\mathbf{PA}_i^G; \beta_i) + \varepsilon_i, \tag{1}$$

where f_i is a multi-valued function parameterized by β_i and ε_i is a noise term. If ε_i is Gaussian, i.e., $\varepsilon_i \sim \mathcal{N}(0, \Sigma_i)$, the structural equation can be parameterized by $\{\beta_i, \Sigma_i\}_{i=1}^l \subseteq \mathbf{\Theta}$.

When $\mathbf{X} \overset{a}{=} \mathbf{x}$ is applied, the incoming arrows to $\mathbf{X}$ are removed from the original rehearsal graph G, yielding an *alteration rehearsal graph* $G^{\mathbf{X}=\mathbf{x}}$. The accompanying structural equations are modified according to the parental relations in $G^{\mathbf{X}=\mathbf{x}}$, where new parameters in $\mathbf{\Theta}$ can be introduced for previously unseen parental relations to provide greater modeling flexibility. The resulting distribution $\Pr(\mathbf{V} \mid \mathbf{X} \overset{a}{=} \mathbf{x})$ over all variables can then be derived from (1). Figure 2 (b) to (d) show the alteration graphs corresponding to $Z_2 \overset{a}{=} z_2$, $Z_3 \overset{a}{=} z_3$, and $Z_2 \overset{a}{=} z_2, Z_3 \overset{a}{=} z_3$, respectively.

Note that SRMs differ from standard SCMs (*Structural Causal Models*) that are widely used to characterize causal relations in the following two aspects. First, SRMs can explicitly represent *mutual-influence relations*, whereas standard SCMs do not represent such mutual influence. Although bidirected edges also occur in causal graphs, they hold different meanings [4, 13]. For instance, in a competitive retail environment, if one store reduces the price P of a product to attract customers, a competitor usually will decrease the price Q accordingly; conversely, Q also influences P by symmetry. This is illustrated in Figure 2: in the rehearsal graph, $Z_1 \leftrightarrow Z_2$ denotes a mutual-influence relation and becomes $Z_1 \leftarrow Z_2$ when applying $Z_2 \overset{a}{=} z_2$, capturing the interrelated *influence* between the two variables. By contrast, bidirected edges in a causal graph typically represent unobserved confounding rather than mutual influence [4, 13], where intervening on one variable in the bidirected edge does not change the value of the other variable through the edge itself.

Second, SRMs use evolving rehearsal graphs and corresponding parameters to characterize time-varying influence relations for AUF decision-making. At time t, the model characterized by $\langle G_t, \boldsymbol{\theta}_t \rangle$ can differ from previous ones, because variable relations may evolve and appropriate decisions may vary over time. By contrast, SCMs generally use a fixed causal graph to describe how observations are generated in the modeled environment, even when the variables are indexed by time. Note that dynamic environments should not be confused with the conventional notion of dynamic systems, where the relations among variables simply repeat over time. Recall observing $\mathbf{x}$, the AUF objective is to find alterations to maximize the AUF probability that $\mathbf{Y}$ falls in a desired set $\mathcal{S}$. The AUF probability is estimated using data sampled from the SRM represented by $\langle G_t, \theta_t \rangle$ at each time index t. The process can repeat consecutively for T rounds. The overall goal is then to maximize the cumulative success probability $\sum_{t=1}^T \Pr(\mathbf{Y}_t \in \mathcal{S} \mid \mathbf{x}_t, \mathring{\mathbf{z}}_t)$, where $\mathbf{Y}_t, \mathbf{X}_t, \mathbf{Z}_t$ denote the outcome variables, the variables that are observed for prediction, and the variables that are altered in the t-th round, respectively. The modeling of the SRM $\{\langle G_t, \theta_t \rangle\}_t$ over T rounds could effectively capture the non-stationarity.

If the true SRM were known, the problem would reduce to searching for the best alteration. In practice, however, the SRM is unknown and needs to be learned from data. This leads to an exploration-exploitation trade-off: Actions that are likely to successfully change the undesired outcome may be less informative for learning the SRM, while informative actions may be riskier. Based on the established SRM and rehearsal graph, Qin et al. [12] further proposed a preliminary method for learning SRMs and

making decisions through interaction. In this preliminary method, a static environment is assumed, i.e., $G_t = G, \theta_t = \theta$. To handle the trade-off between making good decisions and learning the SRM, the information gain $\mathrm{I}(\cdot\,;\,\cdot)$ in the t-th round subject to the AUF probability constraint is optimized:

$$\max_{\mathbf{z}_t} \mathrm{I}(\mathbf{G}, \mathbf{\Theta}; \mathbf{Z}_t, \mathbf{Y}_t \mid D_t, \mathbf{x}_t, \mathbf{Z}_t \overset{a}{=} \mathbf{z}_t) \quad \text{s.t.} \quad \mathbb{E}_{G,\theta|D_t}\Big[\mathbb{I}\Big(\mathbf{Y}_t \in \mathcal{S} \mid G, \theta, \mathbf{x}_t, \mathbf{Z}_t \overset{a}{=} \mathbf{z}_t\Big)\Big] \geqslant \tau, \tag{2}$$

where $\mathbb{I}(\cdot)$ is the indicator function. The objective is the mutual information between the SRM $\langle \mathbf{G}, \mathbf{\Theta} \rangle$ and the post-alteration data, which quantifies the additional information about the SRM that can be revealed by the post-alteration outcome. The threshold τ controls how conservative the alteration is: larger τ favors actions more likely to avoid the undesired future in the current round, and the posterior $\mathbb{E}_{G,\theta|D_t}$ summarizes uncertainty over graphs and parameters given accumulated data D_t. Exact computation is intractable due to the large space of graphs and parameters, so this approach maintains a Bayesian posterior approximation and maximizes mutual information via Bayesian optimization. Further, a probably approximately correct (PAC) bound is established to quantify the associated risk of a decision based on (2), offering a high-confidence guarantee for the AUF probability.

3.2 Interaction-free Rehearsal Learning with Explicit Graphs

In most practical applications, the use of interaction is constrained since the environment may be altered by the interaction process, which potentially leads to the permanent miss of optimal decision states that originally existed. Hence, it is more demanded to establish *interaction-free rehearsal learning* methods.

3.2.1 *Linear Influence Model*

Du et al. [14] explored rehearsal learning within the linear Gaussian structural rehearsal model, assuming a pre-specified rehearsal graph in a static environment. This approach operates in an interaction-free manner, aiming to identify an alteration that maximizes the AUF probability. In contrast to (2), where the objective is merely to ensure that the AUF probability meets a predefined threshold, this approach focuses on identifying the *globally optimal alteration*. This shift addresses the inherent limitations of threshold-based strategies as (2), which are highly sensitive to the hyperparameter τ and cannot guarantee that the identified alteration can lead to the optimal AUF probability. Additionally, as the AUF constraints in (2) lack analytical solutions, these threshold-based methods are compelled to resort to sampling-based estimation, which suffers from poor scalability in high-dimensional alteration spaces.

A primary advantage of linear systems for global optimization of the alteration lies in the fact that any alteration $\mathbf{z}$ merely induces an additive shift in the mean of the outcome distribution without altering its covariance matrix, i.e., $\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z}^{\xi} + \mathbf{C}\boldsymbol{\varepsilon}$, where $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are constant matrices, $\mathbf{z}^{\xi}$ are alteration values, and $\boldsymbol{\varepsilon}$ are noises. This allows the consequences of an alteration to be modeled as a simple location shift of the probability mass within the outcome space.

Despite the preservation of the integrity of the distribution's shape under alterations, directly maximizing the AUF probability remains generally intractable, because calculating the integral of a multivariate Gaussian distribution over an arbitrary region $\mathcal{S}$ lacks a closed-form analytical expression. To explore the boundary that the AUF probability can be optimized efficiently, Du et al. [14] established the CARE (CAnonical Rectangle) condition, which characterizes a kind of geometric relation between the desired region $\mathcal{S}$ and the noise structure such that $\mathcal{S}$ is a canonical rectangle relative to the noise covariance matrix, i.e., the boundaries of the desired region become perfectly aligned with the axial planes of the hyperellipsoid induced by the noise distribution. When CARE condition holds, the dependencies among the outcome variables $\mathbf{Y}$ can be effectively decoupled, such that the multivariate joint probability can be analytically factorized into a product of univariate Gaussian Cumulative Distribution Functions (CDFs)

$$\Pr\big(\mathbf{Y} \in \mathcal{S} \mid \mathbf{X} = \mathbf{x}, \mathbf{Z}_A \overset{a}{=} \mathbf{z}_A\big) = \prod_{i=1}^{|\mathbf{Y}|} \Big\{\Phi(-\mathbf{k}_i^{\mathrm{T}}\mathbf{z}_A - b_i) - \Phi(\mathbf{k}_{i+|\mathbf{Y}|}^{\mathrm{T}}\mathbf{z}_A + b_{i+|\mathbf{Y}|})\Big\}, \tag{3}$$

where Φ denotes the standard Gaussian CDF, and the coefficients $\mathbf{k}_i$ and b_i can be directly obtained from the structural parameters of $\mathcal{S}$.

The factorization of the AUF probability into univariate functions is not merely a mathematical convenience. Indeed, it reveals a fundamental property that guarantees the reliability of the resulting decisions. Specifically, while the raw probability function is not inherently convex, its negative logarithm

$l(\mathbf{z}_A) = -\log \Pr\big(\mathbf{Y} \in \mathcal{S} \mid \mathbf{x}, \mathring{\mathbf{z}}_A\big)$ is proven to be convex with respect to the alteration $\mathbf{z}_A$. This discovery is a pivotal mathematical property for the optimization, as it transforms the search for the best action into a constrained convex program $\min_{\mathbf{z}_A \in [\mathbf{z}_l, \mathbf{z}_r]} l(\mathbf{z}_A)$, where $\mathbf{z}_l$ and $\mathbf{z}_r$ represent the lower and upper bounds of feasible alterations, respectively. Because the exact gradient and Hessian of $l(\mathbf{z}_A)$ can be explicitly computed using standard Gaussian probability functions, this approach can leverage a projected Newton algorithm, which iteratively refines the alteration using second-order information and projects it onto the boundary constraints, achieving a superlinear convergence rate to the global optimum.

The procedure described above provides a tractable framework for identifying the optimal alteration when the CARE condition holds. For scenarios where the desired region $\mathcal{S}$ is irregular and fails to meet the condition, an inner CARE embedding strategy can be employed to identify the largest canonical rectangle $\mathcal{I}_{cr}$ inscribed within the original region $\mathcal{S}$. Optimizing the surrogate probability $\Pr\big(\mathbf{Y} \in \mathcal{I}_{cr} \mid \mathbf{x}, \mathring{\mathbf{z}}_A\big)$ offers a reliable convex lower bound for the original intractable objective. This approach ensures that even when the desired region is complex, this approach can identify a robust and mathematically sound decision that effectively steers the outcome into the desired region.

3.2.2 *Linear Influence Model Extensions: Long-term Objective and Non-Stationary Environment*

In many practical applications, decision-makers are not solely concerned with one single isolated outcome; instead, the objective is to regulate the aggregated effect of a sequence of decisions to ensure a desired long-term outcome. For example, managing a long-term supply chain such that the average annual yield falls within a desired region $\mathcal{S}$. Notably, the long-term AUF objective is not a mere repetition of single-round decisions over T periods. Unlike isolated decision-making, long-term objectives involve complex temporal dependencies where initial alterations may continuously exert an influence on future variables.

Based on a pre-specified rehearsal graph, Du et al. [15] tackled this challenge through seeking a sequence of alterations $\mathbf{z}_{0:T}$ over a time horizon T that maximizes the AUF probability of the aggregated outcome $\bar{\mathbf{Y}} = \frac{1}{T+1}\sum_{t=0}^{T} \mathbf{Y}_t$ falling into $\mathcal{S}$, denoted as $\Pr(\bar{\mathbf{Y}} \in \mathcal{S} \mid \mathbf{x}_0, \mathring{\mathbf{z}}_{0:T})$. Compared to single-round decision-making, the fundamental challenge lies in the fact that the influence can propagate and accumulate over time, rendering the exact evaluation of the AUF probability computationally intractable. Theoretical analysis, however, shows that the aggregated outcome $\bar{\mathbf{Y}}$ can be decomposed into two distinct components, i.e., a deterministic expected value that is linearly controlled by the alteration sequence $\mathbf{z}_{0:T}$, and a zero-mean stochastic term representing the accumulated noise. This linear decomposition simplifies the problem significantly, allowing the optimal alteration sequence to be derived through an optimization procedure analogous to the solutions developed for single-round linear cases [14].

Another important issue is about decision-making in an inherently non-stationary environment, and variable alteration can incur varying costs. For example, a manager adjusts price discount (Z_t) monthly to prevent a predicted decline in sales (Y_t) based on business indicators ($\mathbf{x}_t$). In this context, the influence of pricing on sales may fluctuate seasonally, and implementing discounts directly affects profit margins, thereby introducing an alteration cost.

To tackle this problem, Du et al. [16] reformulated the AUF problem in the t-th round as cost-minimization subject to an AUF probability constraint, i.e.,

$$\min_{\mathbf{z}_t} (\mathbf{z}_t - \mathbf{z}_t^0)^{\mathrm{T}} \mathbf{W} (\mathbf{z}_t - \mathbf{z}_t^0) \quad \text{s.t.} \quad \Pr(\mathbf{Y}_t \in \mathcal{S} \mid \mathbf{x}_t, \mathring{\mathbf{z}}_t) \geqslant \tau, \tag{4}$$

where $\mathbf{z}_t^0$ is the default zero-cost state and $\mathbf{W}$ is a pre-specified weight matrix reflecting the unit cost of every alterable variable. The primary challenge lies in the fact that the influence learned from historical data may become obsolete under environmental drift, which directly compromises the accurate estimation of the AUF probability in the constraint of (4). Hence, rapid adaptation to current dynamics is essential for maintaining decision reliability. For this purpose, the *online ensemble* mechanism [17] is employed, where multiple base online learners are maintained such that the dynamical environment adaption can be captured. When the desired region $\mathcal{S}$ is a convex polytope and the influence model is characterized by linear-Gaussian relations, the intractable AUF probability constraint $\Pr(\mathbf{Y}_t \in \mathcal{S} \mid \mathbf{x}_t, \mathring{\mathbf{z}}_t) \geqslant \tau$ can be transformed into a deterministic constraint. Consequently, the AUF problem in (4) is explicitly reformulated as a standard convex quadratically constrained quadratic program (QCQP), which enables a polynomial-time algorithm solvable in $\mathcal{O}(|\mathbf{V}|^3)$ via interior-point methods.

3.2.3 *Non-linear Influence Model*

While linear-Gaussian methods provide a tractable foundation for rehearsal learning, real-world applications usually exhibit complex non-linear relations and non-Gaussian noise. Evaluating the probability that the outcome $\mathbf{Y}$ falls within the desired region $\mathcal{S}$ requires solving high-dimensional integrals over non-linear densities, which is generally computationally prohibitive. In addition to the challenge of probability evaluation, the optimization process itself encounters a significant barrier. The AUF objective is inherently defined via a non-smooth indicator function that dictates whether $\mathbf{Y}$ falls into the desired region $\mathcal{S}$. This creates a "flat" optimization landscape where traditional gradient-based algorithms usually fail, as gradients are either zero or undefined at the boundaries. Naive approaches resorting to intensive Monte Carlo sampling to estimate the AUF probability scale poorly as the dimensionality of decision variables increases. Therefore, navigating non-linear environments requires a transition from *exact* analytical integration toward flexible *approximated* gradient-based strategies.

Qin et al. [18] proposed the Grad-Rh method which replaces the exact AUF probability maximization with an approximated differentiable optimization scheme. First, it leverages the Conditional Normalizing Flows (CNF) [19, 20] to model the post-alteration distribution $\Pr(\mathbf{Y} \mid \mathbf{x},\ \mathring{\mathbf{z}}_A)$. A CNF operates by mapping a simple base distribution through a sequence of smooth, invertible transformations, effectively warping it into an expressive representation of the outcome $\mathbf{Y}$. This architecture enables the model to capture intricate, non-Gaussian densities and provides a highly flexible representation for non-linear influence relations.

Moreover, Grad-Rh attempts to approximate the non-differentiable indicator function $\mathbf{Y} \in \mathcal{S}$ with smooth surrogate losses. It first identifies the Chebyshev center $\mathbf{y}_c$, which is the center of the largest inscribed ball within $\mathcal{S}$. Let r_c denote the radius of the ball. Then, the optimization of the indicator of whether Y falls into $\mathcal{S}$ can be reformulated as a differentiable optimization of the distance between the expected outcome $\mathbf{Y}$ and the center $\mathbf{y}_c$, with the following surrogate losses:

- **Center MSE (Mean Squared Error):** $\mathcal{L}_{\text{MSE}}(\mathbf{z}_A) = \mathbb{E}\big[\|\mathbf{Y} - \mathbf{y}_c\|_2^2\big]$. This loss function penalizes the expected squared Euclidean distance from the outcome to the center, providing a strong incentive for the distribution to stay concentrated near the target point.
- **Center MAE (Mean Absolute Error):** $\mathcal{L}_{\text{MAE}}(\mathbf{z}_A) = \mathbb{E}\big[\|\mathbf{Y} - \mathbf{y}_c\|_1\big]$. This loss function penalizes the expected Manhattan distance, exhibiting greater robustness against extreme values.
- **r_c-insensitive loss:** $\mathcal{L}_r(\mathbf{z}_A) = \mathbb{E}\big[\max(0, \|\mathbf{Y} - \mathbf{y}_c\|_2 - r_c)\big]$. This loss function ignores outcomes that fall within the inscribed ball $\mathcal{B}(\mathbf{y}_c, r_c)$ and only penalizes those outside.
- **Huber loss:** $\mathcal{L}_{\text{Huber}}(\mathbf{z}_A) = \mathbb{E}\big[H_{r_c}(\|\mathbf{Y} - \mathbf{y}_c\|_2)\big]$, where $H_{r_c}(a) = a^2$ if $|a| \leqslant r_c$, and $r_c(2|a| - r_c)$ otherwise. This loss function provides a smooth combination of squared and linear penalties inside and outside the ball, respectively, preventing extreme outliers from dominating the optimization.

By using the above surrogate losses, the non-differentiable binary indicator is replaced with a smooth, distance-based objective that enables the use of efficient gradient descent strategies. In practice, the expectations $\mathbb{E}[\cdot]$ for these surrogates are empirically approximated via Monte Carlo sampling from the learned CNF models. By minimizing the differentiable surrogates through standard gradient descent, the post-alteration distribution of the outcome $\mathbf{Y}$ is steered toward $\mathcal{S}$; this avoids the exhaustive sampling typically required by black-box optimization methods, thereby significantly reducing the computational overhead. Note that the optimization landscape here is inherently non-convex, and the overall performance is dependent on both the quality of the trained CNF model and the geometric consistency between the employed surrogate loss and the underlying AUF probability.

3.3 Interaction-free Rehearsal Learning without Explicit Graphs

The explicit graph structure required by the above approaches can be obtained from domain experts based on their prior knowledge. Such graphs, however, are not always available and may also be susceptible to inaccuracies in many practical tasks. Recently, it has been shown that, under mild assumptions, a compatible *variable order*, though containing less structural information than an explicit graph, is sufficient for evaluating the AUF probability from observational data [21].

3.3.1 *Order-based Rehearsal Learning*

Tao et al. [21] showed that knowing the topological influence *order* of variables is sufficient for evaluating the AUF probability. This is the first study in this direction, and thus, it works on a simple acyclic

Table 1 Comparison of assumptions needed for learning a full graph structure and learning a compatible order (in an acyclic setting with independent noises and no latent variables).

Aspect	Representative graph-structure learning	OLEM order learning [21]
Learning target	The set of directed edges	A variable order
Structural information	Detailed graph structure	Order structure, which is less detailed than graph structure
Functions	Identification often requires a particular function class, such as linear or restricted non-linear functions	Additive, without further restrictions on the specific forms of the functions
Noise distributions	Identification guarantees often require a particular noise type, such as Gaussian or non-Gaussian noise	No restriction on the specific noise-distribution family
Additional condition	Method-specific conditions for recovering edges	The information-theoretic condition in Assumption 3.1 of [21]

additive model with independent noises, and no latent variables. Specifically, as long as the variable sequence remains topologically consistent, i.e., an "upstream" variable A in the influence relations is positioned before a "downstream" variable B, the resulting AUF probability remains identical no matter whether there is a direct edge between them. Intuitively, consider the flow of a river. If one knows that point A is upstream of point B, one knows the possible direction in which an alteration may propagate, even without knowing the complicated twists or turns between A and B.

Mathematically, the AUF probability can be estimated from observational data using the variable order. Let $\mathbf{V} = \mathbf{X} \cup \mathbf{Z} \cup \mathbf{Y}$ denote the set of all variables. Under the context $\mathbf{X} = \mathbf{x}$ and the alteration $\mathbf{Z}_A \stackrel{a}{=} \mathbf{z}_A$, the probability of $\mathbf{V} = \mathbf{v}$ is given by

$$\Pr\left(\mathbf{V} = \mathbf{v} \mid \mathbf{X} = \mathbf{x}, \mathbf{Z}_A \stackrel{a}{=} \mathbf{z}_A\right) = \delta_{\mathbf{X}=\mathbf{x}}(\mathbf{v})\delta_{\mathbf{Z}_A=\mathbf{z}_A}(\mathbf{v}) \prod_{V_i \notin \mathbf{X} \cup \mathbf{Z}_A} \Pr\left(V_i = v_i \mid \mathrm{pred}(V_i)\right), \tag{5}$$

where $\delta_{\mathbf{X}=\mathbf{x}}(\mathbf{v})$ is a delta function that equals one if the sub-vector of $\mathbf{v}$ corresponding to $\mathbf{X}$ is $\mathbf{x}$, and zero otherwise; $\delta_{\mathbf{Z}_A=\mathbf{z}_A}(\mathbf{v})$ is defined analogously; $\mathrm{pred}(V_i)$ denotes the values of all variables that precede V_i in the topological order. Thus, the AUF probability $\Pr(\mathbf{Y} \in \mathcal{S} \mid \mathbf{x}, \mathbf{Z}_A \stackrel{a}{=} \mathbf{z}_A)$ can be derived from (5) by marginalizing out all the variables other than $\mathbf{Y}$. If the learned order is incompatible with the model or relevant variables are latent, the resulting AUF probability is not guaranteed to be valid.

To learn the compatible order from observational data, Tao et al. [21] proposed the OLEM (Order Learning via conditional Entropy Maximization) method, which recovers the order by repeatedly finding and removing the "last" variable, called as *sink*, which does not influence any other remaining variables. It works with an information-theoretic criterion. In each iteration, OLEM selects the variable that maximizes the conditional entropy $h(V_i \mid V_{-i}^{\Lambda})$, that is, the variable that is least predictable from the remaining variables, denoted by V_{-i}^{Λ}. Under the mild condition stated in Assumption 3.1 of [21], the method is guaranteed to recover a compatible order.

Table 1 compares the assumptions needed for learning a compatible order and learning a full graph structure in the acyclic setting with independent noises and no latent variables. OLEM assumes an additive model but does not further restrict the specific forms of the functions or the family of noise distributions. Its soundness relies on the information-theoretic condition in Assumption 3.1 of [21]. Indeed, by learning a compatible order rather than the full set of directed edges, OLEM can avoid many additional assumptions commonly required for recovering the full graph.

This approach can be implemented in polynomial time relative to the number of variables. A scalable solution for modeling post-alteration distributions and identifying optimal decisions that maximize the AUF probability can then be obtained by leveraging the learned ordering with the differentiable optimization framework proposed by Qin et al. [18], leading to a rehearsal learning method that can work with complex and non-linear systems.

3.3.2 *Non-parametric Rehearsal Learning*

To overcome the limitation of parametric models that rely on rigid functional forms, Du et al. [22] proposed a non-parametric approach using kernel methods, where influence relations are captured through conditional mean embeddings (CMEs) that map conditional distributions onto a reproducing kernel Hilbert space (RKHS). This formulation enables the AUF objective to be expressed as a kernel inner product between a desirability function and an action-dependent distribution embedding. Such a representation effectively bypasses the need for explicit functional forms, facilitating the modeling of highly non-linear environments and non-additive noise.

Table 2 Summary of representative rehearsal learning methods.

Method	Data	Structure	Main assumptions	Guarantee or result
Interaction-based SRM [12]	Observations, interactions	A learned SRM	A specified model family; interaction is available	A PAC bound for the risk of selected alteration
Linear graph-based methods [14–16]	Observations	A rehearsal graph	Linear-Gaussian relations; various method-specific conditions on the desired region, distribution shift, or temporal setting	Global optimum under the CARE condition; results for non-stationary and temporal settings
Non-linear graph-based method [18]	Observations	A rehearsal graph	An adequately learned post-alteration distribution; adequate agreement between the surrogate loss and the AUF objective	A differentiable surrogate enabling gradient-based decision optimization
Order-based method [21]	Observations	A compatible variable order learned from data	Acyclic additive model with independent noises and no latent variables; Assump. 3.1 of [21] for order learning	Recovery of a compatible order and its sufficiency for AUF evaluation
Non-parametric kernel method [22]	Observations	A compatible variable order	Continuous, bounded, positive-definite, and decomposable kernels; the smoothed desirability function lies in the outcome RKHS	Identifiability and consistent estimation of the smoothed AUF objective

Although some of the aforementioned rehearsal methods do not restrict functional forms [18, 21], they typically rely on surrogate losses that may sometimes deviate from the true goal of maximizing the AUF probability. Du et al. [22] unified the direct optimization of the AUF probability and the non-parametric influence modeling within the kernel-based framework. The primary technical challenge lies in the discontinuity of the desirability indicator $\mathbb{I}(\mathbf{Y} \in \mathcal{S})$, which lacks gradient signals and does not reside in standard continuous kernel spaces. Here, the indicator is approximated via a smooth Probit surrogate. In particular, consider a desired region $\mathcal{S}$ characterized as a convex polytope defined by l linear constraints

$$\mathcal{S} = \{\mathbf{y} \in \mathbb{R}^{d_y} \mid m_k^T \mathbf{y} \leqslant b_k, k = 1, \ldots, l\}. \tag{6}$$

The Probit surrogate $w_\eta(\mathbf{y})$ approximating the indicator is

$$w_\eta(\mathbf{y}) = \prod_{k=1}^{l} \Phi(\eta(b_k - m_k^T \mathbf{y})), \tag{7}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution, and $\eta > 0$ is a scaling parameter controlling the sharpness of the approximation. This surrogate ensures differentiability and provides informative signals even for "near-miss" outcomes that fall slightly outside the desired region $\mathcal{S}$. It is worth noting that this surrogate loss differs fundamentally from the point-based surrogates introduced in Section 3.2.3. Rather than relying on a single interior point such as the Chebyshev center, the Probit surrogate accounts for the entire geometry of $\mathcal{S}$. Furthermore, it is formally shown that $w_\eta(\mathbf{y})$ converges to the indicator $\mathbb{I}(\mathbf{y} \in \mathcal{S})$ as $\eta \to \infty$ for any $\mathbf{y}$ not on the boundary of $\mathcal{S}$.

To model the distribution of $\mathbf{Y}$ under alterations, a kernel ridge regression (KRR) based nested estimator is employed. Following an order-based structural decomposition of variables, i.e., $\mathbf{Z} = \mathbf{U} \cup \mathbf{Z_A} \cup \mathbf{U}_{post}$, where $\mathbf{U}$ are pre-alteration variables and $\mathbf{U}_{post}$ are post-alteration variables, the AUF objective is decomposed into a two-stage estimation according to

$$\mathbb{E}_{\mathbf{y} \sim \Pr(\cdot|\mathbf{x}, \hat{\mathbf{z}}_A)}[w_\eta(\mathbf{y})] = \mathbb{E}_{\mathbf{u} \sim \Pr(\cdot|\mathbf{x})}[\mathbb{E}_{\mathbf{y} \sim \Pr(\cdot|\mathbf{x}, \mathbf{u}, \mathbf{z}_A)}[w_\eta(\mathbf{y})]], \tag{8}$$

which consists of an inner expectation over outcomes $\mathbf{Y}$ and an outer expectation over pre-alteration variables $\mathbf{U}$. This formulation yields a closed-form surrogate objective that can be optimized via gradient-based methods. To achieve robust convergence, Du et al. [22] employed a multi-start optimization strategy and selected the optimal alteration across multiple initializations.

Finally, Table 2 summarizes the required data, structural information, main assumptions, and results of representative rehearsal learning methods. The summary shows that these methods provide different types of results, including risk quantification, decision optimization, order recovery, and estimation of the AUF objective, under their respective conditions.

4 Influence Power

Influence has been established as the fundamental relation supporting decision-making. The previous sections present rehearsal learning methods that can identify the influence relation in different scenar-

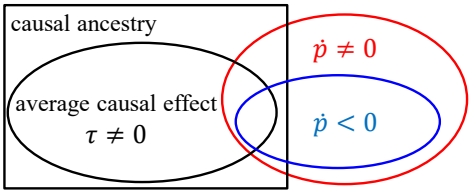


Figure 3 Illustration of relations between causal ancestry, average causal effect, and influence power. Here, $\tau \neq 0$ denotes nonzero ACE, $\dot{p} \neq 0$ denotes nonzero InP, and $\dot{p} < 0$ denotes negative InP.

ios. This section is devoted to answering the fundamental question of how to quantitatively measure the strength of influence. Influence may admit different quantitative measures, just as association and causality do.

Tao et al. [23] proposed *influence power* (abbr. InP, with symbol notation $\dot{p}$), a quantitative measure of influence for AUF decision-making. Influence power considers that observations become available and further alterations can be selected in future. A current alteration may change these future observations and thereby affect which subsequent alterations are selected. Thus, influence power measures the extent to which control over a variable can improve the attainable probability of future success over the *entire* decision process. This differs from causal measures [10, 11], which quantify the effect of a specified intervention as it propagates to the outcome while subsequent variables unfold without further decision-dependent alterations.

Specifically, the influence power of an actionable variable Z on the outcome variable Y is defined as

$$\dot{p}(Z;Y) \;=\; \max_{z\in\Delta_Z} \Pr^{reh}(Y \in \mathcal{S} \mid Z \overset{a}{=} z) \;-\; \mathbb{E}_{z\sim\Pr(Z)}[\Pr^{reh}(Y \in \mathcal{S} \mid Z = z)], \tag{9}$$

where Δ_Z denotes the feasible alteration domain of variable Z, and $\Pr^{reh}(\cdot)$ denotes the *rehearsal probability*, in which variable values may be either observed, as in common probability terms, or specified through rehearsal alterations. Tao et al. [23] instantiated the probability $\Pr^{reh}(\cdot)$ with the *maximum expected probability (MEP)* via a Bellman-style recursion [24], rehearsing future possibilities of subsequent variables following the principle of maximum expected utility [25]. When no future alteration is considered, rehearsal probability reduces to the standard probability $\Pr(\cdot)$. InP compares the probability of reaching the desired set with/without rehearsal alteration on Z, which provides a holistic valuation of an alteration by capturing both its direct impact and its downstream, decision-mediated impact on the outcome. By construction, influence power is bounded within $[-1, 1]$, where a positive value indicates a beneficial alteration, whereas a value near zero or below implies helpless or even harmful alteration.

It is worth mentioning that InP shows that a variable with high influence on the outcome does not necessarily have a strong (or even nonzero) *average causal effect* (ACE) on the outcome. Consider anesthesia Z, a subsequent surgery W, and patient recovery Y. If a randomized, double-blind, and controlled trial were conducted, anesthesia alone would not have an impact on recovery; it merely sedates the patient without treating the disease, so its ACE on Y is zero. Yet anesthesia can have strictly positive InP, because it serves as a necessary precursor that creates the clinical conditions for the surgery W to be effectively performed. Moreover, it has been shown that a variable with nonzero ACE can have zero InP, and neither nonzero ACE nor causal ancestry guarantees nonnegative InP [23]. These qualitative relationships are summarized in the Venn diagram in Figure 3.

Many practical tasks require multiple steps of decision-making, where evidence revealed in earlier steps refines the evaluation of influence for later actionable variables. For example, consider a clinical scenario. If earlier evidence indicates that a patient is allergic to penicillin (e.g., revealed by a prior decision on skin test), later treatment decisions should avoid penicillin because its impact is likely harmful. If no such evidence is provided, penicillin may be considered because its impact is beneficial on most patients. This motivates evaluating influence at each step while conditioning on evidence revealed so far. That is, rather than evaluating the influence of an actionable variable Z on an outcome Y without evidence, one can evaluate the influence of Z on Y given evidence *evid* from previous steps. For this purpose, (9) can be extended to

$$\dot{p}(Z;Y \mid evid) \;=\; \max_{z\in\Delta_Z} \Pr^{reh}(Y \in \mathcal{S} \mid Z \overset{a}{=} z, evid) \;-\; \mathbb{E}_{z\sim\Pr(Z|evid)}[\Pr^{reh}(Y \in \mathcal{S} \mid Z = z, evid)], \tag{10}$$

where the evidence *evid* comprises variables observed or altered up to the current step.

Table 3 Comparison between rehearsal learning and seemingly relevant decision-making research.

Research area	Starting point	Key requirement	Main output	Main difference from rehearsal learning
Rehearsal learning	Actionable variables and observational or interaction data	Influence relations exist	Influence relations and an alteration for achieving the AUF objective	N/A
Offline reinforcement learning	Recorded state-action-reward transitions	$P(s', r\|s, a)$ remains invariant under the evaluated policy	A policy $\pi(a \mid s)$	Rehearsal learning requires no recorded actions, does not prespecify invariant relations under an alteration, and can provide interpretability through the explicitly identified influence relation
Policy learning	Recorded covariates, actions, and outcomes	The action-outcome relation is identified and remains invariant under the learned rule	A decision rule $\pi(X)$	Rehearsal learning requires no recorded actions or prespecified invariant relations under alteration and can provide interpretability through the explicitly identified influence relation
Model-based planning	A given or learned action model	The model adequately represents candidate actions and their consequences	An action or action sequence	Rehearsal learning requires no action model or variables predefined as actions
Algorithmic recourse	A predictor and feasible input changes	The proposed input change is feasible and the predictor remains applicable	A prediction-changing input x'	Rehearsal learning targets changes in future outcomes rather than model predictions
Causal recourse and causal decision methods	A causal model or identified causal quantities	Causal assumptions and intervention semantics appropriate to the decision	An intervention or decision rule	Rehearsal learning can find influence relations that are not causal, which cannot be found by causal learning methods

Note that, evaluating InP conditioned on evidence requires updating the MEPs to incorporate *evid*. Such updates can be computationally demanding when the number of variables is large. In addition, the updates must be performed online as new evidence may arrive sequentially. An efficient closed-form update is available for linear Gaussian models [26]. Moreover, it has been theoretically shown in [26] that under mild conditions, the evidential inference is more accurate than non-evidential inference even when the structural influence model is misspecified; this supports the practical reliability of influence evaluation in open and dynamic environments where the model may deviate from true structure.

5 Rehearsal Learning and Seemingly Relevant Decision-making Research

Rehearsal learning was proposed as a research direction for decision-making problems such as AUF. Several other research areas may appear related because they also concern decision-making, but they adopt different formulations and assumptions. The main distinctions are summarized in Table 3, briefly compared according to their typical formulations.

Offline reinforcement learning and policy learning. Offline reinforcement learning (RL) starts from recorded transitions (s, a, r, s'), where s is the current state, a is an action, r is the reward, and s' is the next state [27]. Its goal is generally to learn or evaluate a policy $\pi(a \mid s)$. Policy learning starts from recorded covariates, actions, and outcomes, and learns a decision rule [28]. Both select among actions represented in decision data and require the corresponding action-response relation to remain invariant under the evaluated policy or rule. For offline RL, this relation is represented by $P(s', r \mid s, a)$. Different from these topics, rehearsal learning can start from purely observational data, consider alterations to variables that were never actions, and does not prespecify which relations remain invariant after alteration. More importantly, offline RL/policy learning outputs a policy/decision rule, which does not explicitly identify which variables influence the outcome. For example, an offline RL controller may learn a sequence of operating commands that avoids product defects, but the learned policy does not reveal which variables, such as temperature or pressure, have influence on the defect outcome. Rehearsal learning instead explicitly discerns such influence relations. This explicit result provides interpretability beyond the learned policy or decision rule.

Model-based planning. Model-based planning searches for an action or action sequence among predefined actions [24, 25]. Its goal is to optimize these actions after their consequences have been represented in an action model. By contrast, rehearsal learning does not require an action model as input, and the variables to be altered need not be predefined as actions.

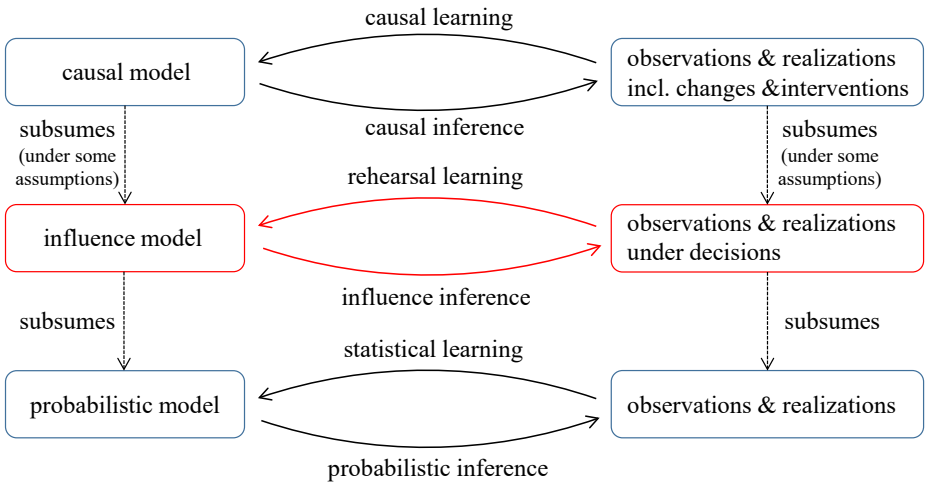


Figure 4 A probabilistic view of learning and inference with causal/influence/association frameworks.

Algorithmic recourse and causal decision methods. Algorithmic recourse searches for a feasible input change that makes a prediction change [29]. Its result is a change in the model prediction, which does not imply a change in the decision outcome. Causal recourse and related causal decision methods seek interventions or decision rules based on causal relations or causal quantities [4,30–32]. Rehearsal learning seeks influence relations, which need not be causal relations. For example, for an outcome Y, a variable that is not a causal ancestry can have positive influence power, as illustrated in Figure 3. A causal learning method seeking causal relations cannot find such an influence relation, whereas rehearsal learning is developed precisely for this purpose.

6 Concluding Remarks

Machine learning has achieved remarkable success in prediction, whereas many real-world scenarios require proactive decision-making. Recently, it has been disclosed that neither *association* nor *causality* is crucial for decision-making, while the pragmatic concept *influence* is the real necessity, offering a principled foundation for practical and effective decision-making, particularly for the AUF problem. Various *rehearsal learning* methods have been developed to discern influence relations, and the influence power (InP) has been proposed for measuring the strength of influence. Python codes of the relevant methods are publicly available at https://www.lamda.nju.edu.cn/rehearsal.

Now it is clear that effective decision-making does not necessarily require the mechanistic information provided by causality. Previous studies have demonstrated that influence and causality are decoupled in several critical ways. First, a variable which has a causal impact on an outcome does not have influence if the variable is not actionable [3]. Second, even when a variable is causal and actionable, it can still lack influence [23]; this occurs at least when altering the variable fails to make a significant or desired change in the outcome within the specific context of the AUF task.

By positioning association, influence, and causality together, their respective roles become clearer. Association captures statistical patterns for prediction but remains *blind* to the consequences of actions. Causality seeks the full underlying mechanism, which often introduces excessive information irrelevant to decision goals, but is not concerned with *what can be done*. In contrast, influence serves as the pragmatic middle ground. It does not seek to uncover the entire underlying mechanism of the environment but directly concerns decision effectiveness. The requirement for solving decision tasks is fundamentally distinct from those of traditional causal learning, and the influence can provide the forward-looking capability necessary to guide actions without the heavy overhead of uncovering the mechanism. Influence is very different from *classic* causality. Considering that causality research has been extended to various directions, and quite different things have been indiscriminately put under the umbrella of causality, influence research has overlaps with some special kind of *generalized* causality studies from some views; however, this is just like that causality overlaps with special kind of association studies. Such overlap does not erase the conceptual distinction. In particular, rehearsal learning seeks influence relations, which need not be causal relations. Shifting from causality and association to influence provides a more principled

and practical framework for decision-making. Figure 4 presents a probabilistic view of the hierarchical relations between the causal/influence/association frameworks.

InP is the first measure for evaluating the strength of influence relation. Just like that association and causality can be quantified by various measures, it is interesting to explore other influence measures that enable scenario-dependent choices. Note that, these measures can be closely tied to the learning process, where an influence model is typically needed. Such a model can take various forms, such as graph-based structures [12] and order-based structures [21]. It is interesting to explore adequate model structures that align well with the chosen influence measure, and to develop rehearsal learning methods capable of recovering the influence model from observational data effectively and efficiently. Once an influence model corresponding to a specific measure is established, the subsequent focus is *influence inference*, i.e., forecasting the outcomes under various candidate actions. This connects the influence model directly to decision. The development of efficient optimization methods to identify actions that maximize the AUF probability is vital, especially when the action space is prohibitively large. Furthermore, the development of benchmark tasks and data sets that can be made publicly available is very important.

Acknowledgements This work was supported by the MoE FIDBP (JYB2025XDXM118).

References

1 I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
2 Z.-H. Zhou, *Machine Learning*. Springer, 2021.
3 ——, "Rehearsal: Learning from prediction to decision," *Frontiers of Computer Science*, vol. 16, no. 4, p. 164352, 2022.
4 J. Pearl, *Causality*. Cambridge University Press, 2009.
5 J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press, 2017.
6 P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, Prediction, and Search*. MIT Press, 2000.
7 Z.-H. Zhou, "Rehearsal: Learning from prediction to decision," Keynote at the CCF Conference on Artificial Intelligence, Urumqi, China, 2023.
8 A. Gretton, O. Bousquet, A. J. Smola, and B. Schölkopf, "Measuring statistical dependence with Hilbert-Schmidt norms," in *Proceedings of the 16th International Conference on Algorithmic Learning Theory (ALT)*, Singapore, 2005, pp. 63–77.
9 C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
10 P. W. Holland, "Causal inference, path analysis and recursive structural equations models," *ETS Research Report Series*, vol. 1988, no. 1, pp. 1–50, 1988.
11 D. Janzing, D. Balduzzi, M. Grosse-Wentrup, and B. Schölkopf, "Quantifying causal influences," *Annals of Statistics*, vol. 41, no. 5, pp. 2324–2358, 2013.
12 T. Qin, T.-Z. Wang, and Z.-H. Zhou, "Rehearsal learning for avoiding undesired future," in *Advances in Neural Information Processing Systems 36 (NeurIPS)*, New Orleans, LA, 2023, pp. 80 517–80 542.
13 T. Richardson, P. Spirtes *et al.*, "Ancestral graph markov models," *The Annals of Statistics*, vol. 30, no. 4, pp. 962–1030, 2002.
14 W.-B. Du, H.-Y. Lei, L. Tao, T.-Z. Wang, and Z.-H. Zhou, "Enabling optimal decisions in rehearsal learning under CARE condition," in *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, Vancouver, Canada, 2025, pp. 14 536–14 561.
15 W.-B. Du, T. Qin, T.-Z. Wang, and Z.-H. Zhou, "Variance-reduced long-term rehearsal learning with quadratic programming reformulation," in *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2025, pp. 113 245–113 275.
16 ——, "Avoiding undesired future with minimal cost in non-stationary environments," in *Advances in Neural Information Processing Systems 37 (NeurIPS)*, Vancouver, Canada, 2024, pp. 135 741–135 769.
17 Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, 2nd ed. CRC Press, 2025.
18 T. Qin, T.-Z. Wang, and Z.-H. Zhou, "Gradient-based nonlinear rehearsal learning with multivariate alterations," in *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI)*, Philadelphia, PA, 2025, pp. 26 859–26 867.
19 L. Ardizzone, C. Lüth, J. Kruse, C. Rother, and U. Köthe, "Guided image generation with conditional invertible neural networks," arXiv:1907.02392, 2019.
20 C. Winkler, D. Worrall, E. Hoogeboom, and M. Welling, "Learning likelihoods with conditional normalizing flows," arXiv:1912.00042, 2019.
21 Y.-X. Tao, T.-Z. Wang, and Z.-H. Zhou, "Order-based rehearsal learning," arXiv:2605.04955, 2026.
22 W.-B. Du, T.-Z. Wang, H.-J. Ye, and Z.-H. Zhou, "Non-parametric rehearsal learning via conditional mean embeddings," arXiv:2605.08999, 2026.
23 L. Tao, T.-Z. Wang, Y. Jiang, and Z.-H. Zhou, "On measuring influence in avoiding undesired future," in *Proceedings of the 14th International Conference on Learning Representations (ICLR)*, Rio de Janeiro, Brazil, 2026.
24 R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.
25 S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2020.
26 L. Tao, T.-Z. Wang, Y. Jiang, and Z.-H. Zhou, "Avoiding undesired future with sequential decisions," in *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)*, Montreal, Canada, 2025, pp. 6245–6253.
27 S. Levine, A. Kumar, G. Tucker, and J. Fu, "Offline reinforcement learning: Tutorial, review, and perspectives on open problems," *arXiv preprint arXiv:2005.01643*, 2020.
28 S. Athey and S. Wager, "Policy learning with observational data," *Econometrica*, vol. 89, no. 1, pp. 133–161, 2021.
29 B. Ustun, A. Spangher, and Y. Liu, "Actionable recourse in linear classification," in *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*)*, 2019, pp. 10–19.
30 A.-H. Karimi, B. Schölkopf, and I. Valera, "Algorithmic recourse: From counterfactual explanations to interventions," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2021, pp. 353–362.

31 S. Lee and E. Bareinboim, "Structural causal bandits: Where to intervene?" in *Advances in Neural Information Processing Systems 31 (NeurIPS)*, 2018, pp. 2573–2583.
32 V. Aglietti, X. Lu, A. Paleyes, and J. González, "Causal bayesian optimization," in *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020, pp. 3155–3164.