Introspective Forecasting

Loizos Michael

Open University of Cyprus

loizos@ouc.ac.cy

Abstract

Science ultimately seeks to reliably predict aspects of the future; but, how is this even possible in light of the logical paradox that making a prediction may cause the world to evolve in a manner that defeats it? We show how learning can naturally resolve this conundrum. The problem is studied within a causal or temporal version of the Probably Approximately Correct semantics, extended so that a learner's predictions are first recorded in the states upon which the learned hypothesis is later applied. On the negative side, we make concrete the intuitive impossibility of predicting reliably, even under very weak assumptions. On the positive side, we identify conditions under which a generic learning schema, akin to randomized trials, supports agnostic learnability.

1 Introduction

Several scientific disciplines seek to reliably forecast a future state of affairs. Often, utility is derived not simply by having one's predictions verified, but by acting upon the predictions prior to their verification. Being able to predict the stock market, for example, would be of little value if one were not to invest based on that prediction, or get paid to share it. In certain cases, taking such actions is accompanied by the distinct risk of altering the temporal evolution of the environment in a way that would lead to a different, than anticipated, future.

The problem is modeled as a forecaster in a possibly adversarial environment that becomes aware of the forecaster's prediction before the environment commits to the outcome that the forecaster attempts to predict. In this context, a forecaster is required to be *introspective*, and acknowledge the effect of its predictions on their realizability. The contributions of this work are four-fold: *(i)* it formalizes introspective forecasting (cf. Definition 4) as an extension of the PAC semantics [Valiant, 1984; Michael, 2011]; *(ii)* it identifies a realizability quantity (cf. Definition 3) that measures the optimal correctness that any introspective forecaster can achieve; *(iii)* it proposes a general schema for constructing introspective forecasters and identifies conditions (cf. Theorem 1) under which it applies; *(iv)* it establishes limitations on what can be introspectively forecasted (cf. Theorem 2) and discusses situations under which these limitations can be lifted (cf. Theorem 3).

Of course, introspective forecasting is not aimed to be used to make "self-fulfilling" predictions, which are realized simply because they were acted upon. Instead, having the ability to make such "self-fulfilling" predictions allows one to rationally choose whether it is warranted to act upon them, if this indeed happens to lead to a future that is preferred over one where no action is taken. We defer a discussion of the philosophical issues that inevitably arise for an extended version of this paper, and focus herein on the technical framework.

2 Overview

We start with an overview of certain key aspects of this work.

2.1 Causal Learnability

Causal learnability [Michael, 2011] extends the PAC learning semantics [Valiant, 1984] to a causal or temporal setting. Below we present a simplified version of that framework.

Given as inputs $\mathcal{T} \triangleq \bigcup_n \mathcal{T}_n$, $\mathcal{S} \triangleq \bigcup_n \mathcal{S}_n$, $\varepsilon, \delta \in (0, 1]$, where $\mathcal{S}_n$ is a set of states of size n, and $\mathcal{T}_n$ is a set of functions that map $\mathcal{S}_n$ to $\mathcal{S}_n$, the learning algorithm proceeds as follows: During a training phase, the learner observes pairs of states $\langle s, t(s) \rangle$, such that s is drawn independently at random from some arbitrary but fixed probability distribution $\mathcal{D}_n$, and such that t is a fixed target function in $\mathcal{T}_n$. After time upper bounded by a fixed polynomial in n, $1/\varepsilon$, $1/\delta$, the learner returns, with probability at least $1 - \delta$, a hypothesis function $h : \mathcal{S}_n \to \mathcal{S}_n$ such that the probability that $h(s) = t(s)$ on a state s randomly drawn from $\mathcal{D}_n$ is at least $1 - \varepsilon$. If the learner can meet the stated requirements for every choice of n, $\varepsilon, \delta \in (0, 1]$, $\mathcal{D}_n$, and $t \in \mathcal{T}_n$, then the class $\mathcal{T}$ is ***causally learnable***. Previous work [Michael, 2011] investigates which classes are causally learnable, by establishing, among others, connections to PAC concept learning [Valiant, 1984].

One may allow functions in $\mathcal{T}_n$ that map $\mathcal{S}_n$ to a set $\mathcal{L}_n$, when modeling scenarios where a learner seeks to predict not the entire *successor* state $t(s)$ of $s \in \mathcal{S}_n$, but only some aspect, or label in $\mathcal{L}_n$, of $t(s)$. When $|\mathcal{L}_n| = 2$, this resembles PAC concept learning, but with a key conceptual difference: the *causal* dependence of the label of $t(s)$ on s. Even when s is observed, the label of $t(s)$ is not determined until the successor state $t(s)$ materializes; if we were to somehow affect s *after* observing it, this would also affect the label of $t(s)$.

2.2 Recording Predictions

To make the aforementioned conceptual distinction between PAC concept learning and causal learning explicit, let $\mathcal{T}_n$ be as in the original causal learning framework, and introduce a new set $\mathcal{F}_n$ of functions that map $\mathcal{S}_n$ to $\mathcal{L}_n$. Like concepts in PAC learning, each $f \in \mathcal{F}_n$ maps a state to a label of that *same* state. Unlike PAC learning, however, class $\mathcal{F} \triangleq \bigcup_n \mathcal{F}_n$ does not aim to act as a hidden structure to be learned; the hidden structure is the temporal dynamics of the environment as captured by class $\mathcal{T}$. Instead, $\mathcal{F}$ is available for the learner to *choose* to apply any $f \in \mathcal{F}_n$ on the successor state $t(s) \in \mathcal{S}_n$ to produce a label (i.e., to extract a particular *feature*). For any chosen function $f \in \mathcal{F}_n$, then, the definition of causal learnability applies mutatis mutandis as before, except that the learner has access to $\mathcal{L}_n$ and f, and aims to return a hypothesis function $h : \mathcal{S}_n \to \mathcal{L}_n$ such that $h(s) = f(t(s))$. If the learner can meet the learning requirements for every choice of $n, \varepsilon, \delta \in (0, 1]$, $\mathcal{D}_n$, $t \in \mathcal{T}_n$, and $f \in \mathcal{F}_n$, then the class $\mathcal{T}$ is **extrospectively learnable** through the class $\mathcal{F}$.

Extrospective learnability adopts an assumption that is implicit in several supervised learning frameworks: the learner operates outside the environment, in that the label of a learning instance is determined irrespectively of the learner's prediction. This is so even when the learning instances are chosen adversarially with access to the learner's current hypothesis (e.g., in mistake-bounded learning [Littlestone, 1988]), since even then the label of each (adversarially) chosen learning instance is still independent of the learner. Closer in spirit to the causal nature of extrospective learnability, the stated assumption is also followed by certain statistical and learning approaches that are used to predict the temporal evolution of a sequence of states [Murphy, 2002; Box *et al.*, 2008].

To account for situations where the very act of making a prediction may potentially affect the correct / expected prediction, we introduce the notion of a **recorder** function $\lhd$ that maps $\mathcal{S}_n \times \mathcal{L}_n$ to $\mathcal{S}_n$; thus, a recorder function records into a state $s \in \mathcal{S}_n$ a given label $l \in \mathcal{L}_n$ and produces a new state $s \lhd l \in \mathcal{S}_n$. This new state belongs in the same set $\mathcal{S}_n$ of states as state s, since the manner in which the environment is represented is independent of whether predictions are made.

As in extrospective learnability, the goal is to make a prediction $h(s)$ on the label of the successor state when given only $s \in \mathcal{S}_n$. The successor state, however, is now $t(s \lhd h(s))$. Having observed s, the learner's prediction $h(s)$ is announced and recorded in s, so that the initial state becomes $s \lhd h(s)$. It is this affected initial state that brings about the successor state $t(s \lhd h(s))$, whose label $f(t(s \lhd h(s)))$ is to be predicted. For any recorder $\lhd$, then, the definition of extrospective learnability applies mutatis mutandis as before, except that the learner has access to $\lhd$, and aims to return a hypothesis function $h : \mathcal{S}_n \to \mathcal{L}_n$ such that $h(s) = f(t(s \lhd h(s)))$. If the learner can meet the learning requirements for every choice of $n, \varepsilon, \delta \in (0, 1]$, $\mathcal{D}_n$, $t \in \mathcal{T}_n$, $f \in \mathcal{F}_n$, and $\lhd$, then the class $\mathcal{T}$ is **introspectively learnable** through the class $\mathcal{F}$.

2.3 Introspective Setting

In the extrospective case (and other typical supervised learning models), the hypothesis function h that is sought is guaranteed to exist; simply set $h(\cdot) = f(t(\cdot))$. In the introspective case, however, the hypothesis function h appears in both sides of the equality $h(s) = f(t(s \lhd h(s)))$ that it seeks to obey; it is not necessarily possible to factor out h from the equation. This leads to the consideration of an agnostic setting, where h is required to obey the equality with probability $1 - \alpha - \varepsilon$, where $1 - \alpha$ is the optimal probability that any hypothesis can achieve. As is typical in agnostic or noisy learning models, the learner is allowed time polynomial in the usual learning parameters, but also in $1/(1 - \alpha)$, to compensate for the adversarial environment it faces. These extensions give rise to the final definition of introspective learnability that we adopt.

Under this final definition, we establish certain connections between extrospective (and, in particular, causal) learnability and introspective learnability. We first show that $\mathcal{T}$ is introspectively learnable through $\mathcal{F}$ if $\mathcal{T}$ is causally learnable (cf. Theorem 1), but causal learnability is not also a necessary condition (cf. Theorem 2). We then introduce a metric of how expressive set $\mathcal{F}_n$ is (cf. Definition 5) in terms of how many functions in the set are needed to invert the mappings they produce, and show that if running time polynomial in this **invertibility dimension** of $\mathcal{F}_n$ is also allowed, then the necessity of causal learnability can be established (cf. Theorem 3).

2.4 Randomized Trials

The sufficiency of causal learnability for introspective learnability comes through a process analogous to *randomized trials*, often used in clinical research and certain social sciences. Roughly, a randomized trial examines the effects of competing intervention strategies by applying a randomly chosen one on each instance of interest. In the case of clinical research, for example, one administers a randomly chosen drug among a certain set of such to each patient participating in the study.

The need for randomized trials arises from the fact that administering a drug to cure / prevent a predicted illness in the future may have inadvertent and unpredictable effects. What is needed, then, is making predictions that are introspectively accurate. Ideally, a doctor should administer drugs based not on the *extrospectively* predicted illness of a patient (i.e., "take this drug because without it you will remain / become sick"), but on the *introspectively* predicted wellness of the patient (i.e., "take this drug because with it you will remain / become well"). Randomized trials offer an empirical exploration of the space of possible introspectively accurate solutions.

Our setting differs non-trivially from that above. A prediction in clinical research is effectively always that the patient will not exhibit the illness specified by a certain study. Thus, $h(s)$ is thought to be a constant k, and one seeks to solve the equation $h(s) = f(t(s \lhd h(s)))$, or $k = f(t(s \lhd k))$, by choosing an appropriate recorder $\lhd$ (i.e., how to act upon the prediction) among a small set of such (e.g., administer a new drug, an existing one, or a placebo). By contrast, our results apply on all possible recorders, and our aim is not to identify one of them, but rather for each of them to identify the hypothesis h that solves the equation $h(s) = f(t(s \lhd h(s)))$. Further, the hypothesis is a *function*, and one cannot simply exhaustively consider its application on every state $s \in \mathcal{S}_n$ to empirically determine the value for $h(s)$. This last consideration necessitates the use of learning techniques along with randomized trials to *efficiently* and *reliably* solve the equation.

2.5 Reinforcement Learning

Although introspective learnability descends from a supervised learning model [Valiant, 1984], it blurs the dividing line between supervised (generalization and pattern recognition) and reinforcement (trial-and-error) learning [Sutton and Barto, 1998]. In the latter case, an agent (partially) observes a state of its environment, chooses an action, whose execution (stochastically) transitions the environment to another state, and presents the agent with an immediate reward. The agent's goal in this setting is to minimize its overall long-term regret.

The driving force behind the vast literature on reinforcement learning is primarily the tension between exploration to identify an optimal long-term strategy, and exploitation to get rewarded. Locally-optimal actions may transition the environment to parts of the state space that defeat the agent's long-term prospects. By contrast, an introspective forecaster observes only an initial and a successor state, and can opt for the locally-optimal prediction, ignoring long-term effects.

Because of the central concern for identifying a long-term policy, reinforcement learning remains interesting even under the assumption that the immediate reward for each available action in each state is known to the agent. In this case, reinforcement learning effectively reduces to a planning problem in a known environment, with the agent's goal being to find an optimal sequence of actions from an initial state. In a sense, then, introspective forecasting is the dual problem to planning, since it focuses exactly on the learning aspect of reinforcement learning, ignoring any planning considerations.

Among the related literature on multi-armed bandits [Robbins, 1952], most relevant is the study of the contextual bandit problem [Li *et al.*, 2010], where an agent observes a context, takes an action, and is then rewarded accordingly. As in introspective forecasting, the problem becomes rather interesting when running time is allowed to grow only polynomially in the description size of the contexts (or states, in our case).

Even considering only the obvious difference that an introspective forecaster does not get rewarded, suffices to see that we assume the availability of a richer feedback mechanism for the agent. Attempting, for instance, to use binary rewards / bandits to simulate prediction fulfillment, would result in not seeing the label of the resulting state in case the prediction is not fulfilled. More to the point, in our setting we assume, in fact, that the agent gets to see *the entire successor state itself*.

Most importantly, the contextual bandit problem is viewed through the reinforcement learning paradigm, seeking to minimize regret by balancing exploration and exploitation. Our work approaches the problem through the supervised learning paradigm, seeking to provide guarantees on predictive accuracy after an initial training phase. At the end of the day, contextual bandits and our work could be seen as complementary attempts — moving towards the same point, but from opposing directions — to unify the two learning paradigms, by investigating learning algorithms that identify and exploit both the intra-state and inter-state structure of their environment.

3 Framework

The making of highly-accurate predictions can be viewed as a game between Nature and the Predictor. The game's parameters are determined by an ***environment*** $E \triangleq \langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$, which defines for each positive integer n: a set $\mathcal{L}_n \subset \mathcal{L}$ of label representations; a set $\mathcal{S}_n \subset \mathcal{S}$ of state representations; a set $\mathcal{F}_n \subset \mathcal{F}$ of functions from $\mathcal{S}_n$ to $\mathcal{L}_n$; a set $\mathcal{T}_n \subset \mathcal{T}$ of functions from $\mathcal{S}_n$ to $\mathcal{S}_n$; a set $\mathcal{H}_n \subset \mathcal{H}$ of all functions from $\mathcal{S}_n$ to $\mathcal{L}_n$. We assume that $\mathcal{L}_n, \mathcal{S}_n, \mathcal{F}_n, \mathcal{T}_n, \mathcal{H}_n$ are finite (but unbounded, as n grows), and that functions in $\mathcal{F}_n, \mathcal{T}_n$ can be evaluated on their inputs in time polynomial on the input size.

The game is played in three rounds. First, the Predictor chooses a function $f \in \mathcal{F}_n$, encoding the feature of states to be predicted. Second, Nature chooses: *(i)* a function $t \in \mathcal{T}_n$, encoding the laws governing the temporal evolution of states; *(ii)* a probability distribution $\mathcal{D}_n$ over $\mathcal{S}_n$, encoding the process of choosing states. Third, the Predictor chooses a function $h \in \mathcal{H}_n$, encoding the process that it uses to make a prediction. Once both players make their choices, the predictive ability of h is evaluated thus: A state $s \in \mathcal{S}_n$ is chosen at random from $\mathcal{D}_n$, and the prediction $h(s)$ is computed. State s is mapped to its temporal successor $t(s)$, and the label $f(t(s))$ of this latter state is computed. Function h is said to be ***extrospectively*** $(1-\varepsilon)$***-accurate w.r.t.*** $\langle f, t, \mathcal{D}_n \rangle$ if $h(s) = f(t(s))$ except with probability ε over the choice of state s from $\mathcal{D}_n$.

Next, we extend the framework to account for recordings.

Definition 1 (Recorder). *A **recorder** of labels $\mathcal{L}_n$ **in states** $\mathcal{S}_n$ is a function $\lhd : \mathcal{S}_n \times \mathcal{L}_n \to \mathcal{S}_n$ that can be evaluated on its inputs in time polynomial on the input size. The recorder $\lhd$ is **degenerate for** a pair of labels $l_1, l_2 \in \mathcal{L}_n$ if for every state $s \in \mathcal{S}_n$, it holds that $s \lhd l_1 = s \lhd l_2$; i.e., the recording of the two labels is not distinguishable. The recorder $\lhd$ is **degenerate** if it is degenerate for every pair of labels in $\mathcal{L}_n$.*

Our notion of a recorder is the broadest possible given the requirement of efficiency. It may fully or only partially record a label, so that distinct labels affect a state in distinct or similar ways. A recorder is degenerate if it records only the fact that some prediction is announced, but ignores the prediction itself. Among the degenerate recorders, the ***trivial*** recorder maps each state to itself, so predictions are not announced at all. Unless stated otherwise, we deal with arbitrary recorders.

The game is played in three rounds, as before, with an extra choice for the Predictor during the first round: the Predictor chooses a recorder $\lhd$ of $\mathcal{L}_n$ in $\mathcal{S}_n$. Once both players make their choices, the predictive ability of h is evaluated thus: A state $s \in \mathcal{S}_n$ is chosen at random from $\mathcal{D}_n$, the prediction $h(s)$ is computed and recorded in state s, giving rise to $s \lhd h(s)$. The resulting state $s \lhd h(s)$ is mapped to its temporal successor $t(s \lhd h(s))$, and the label $f(t(s \lhd h(s)))$ of this latter state is computed. Function h is said to be ***introspectively*** $(1-\varepsilon)$***-accurate w.r.t.*** $\langle \lhd, f, t, \mathcal{D}_n \rangle$ if $h(s) = f(t(s \lhd h(s)))$ except with probability ε over the choice of state s from $\mathcal{D}_n$.

Unlike in the extrospective setting, a highly-accurate hypothesis h may not (and typically does not) necessarily exist.

Proposition 1 (Non-Predictability Scenarios). *Assume that the Predictor chooses during the first round (i) any non-degenerate recorder $\lhd$ of $\mathcal{L}_n$ in $\mathcal{S}_n$, and (ii) any non-constant function $f \in \mathcal{F}_n$, such that the recorder $\lhd$ is non-degenerate for a pair of labels $l_1, l_2 \in \mathcal{L}_n$ in the image of f. Then, Nature can always choose during the second round (i) a function $t \in \mathcal{T}_n$, and (ii) a probability distribution $\mathcal{D}_n$ over $\mathcal{S}_n$, such*

that for every function $h \in \mathcal{H}_n$ that is introspectively $(1 - \varepsilon)$-accurate w.r.t. $\langle \vartriangleleft, f, t, \mathcal{D}_n \rangle$, it holds that $\varepsilon = 1$.

Proof. Since $\vartriangleleft$ is a non-degenerate recorder for the pair of (distinct) labels l_1, l_2, then by Definition 1 there exists a state $s_0 \in \mathcal{S}_n$ such that $s_0 \vartriangleleft l_1 \neq s_0 \vartriangleleft l_2$. Since l_1, l_2 are in the image of f, there exist states $s_1, s_2 \in \mathcal{S}_n$ such that $f(s_1) = l_1$ and $f(s_2) = l_2$. Define the probability distribution $\mathcal{D}_n$ to assign non-zero probability only to state s_0. For every state $s \in \mathcal{S}_n$, define $t(s)$ to equal s_1 if $s = s_0 \vartriangleleft l_2$ and to equal s_2 if $s \neq s_0 \vartriangleleft l_2$. The result follows by case analysis on $h(s)$. $\square$

The very weak conditions of Proposition 1 corroborate and explain the strong intuition that an announced prediction can be easily defeated. This ability derives from the possibility that the temporal dynamics $t \in \mathcal{T}_n$ may invariably respond to the recorded prediction itself in an adversarial manner. We seek conditions that quantify such an adversarial behavior — which presumably is limited in those real-world scenarios where announced predictions are not usually falsified.

Definition 2 (Response Function). *Consider a recorder $\vartriangleleft$ of $\mathcal{L}_n$ in $\mathcal{S}_n$, a function $f \in \mathcal{F}_n$, a function $t \in \mathcal{T}_n$, and a state $s \in \mathcal{S}_n$. The **response function for s under** $\langle \vartriangleleft, f, t \rangle$ is the function $r_s : \mathcal{L}_n \to \mathcal{L}_n$ such that for every label $l \in \mathcal{L}_n$, $r_s(l) = f(t(s \vartriangleleft l))$.*

Definition 3 (Realizability Degree). *Consider a recorder $\vartriangleleft$ of $\mathcal{L}_n$ in $\mathcal{S}_n$, a function $f \in \mathcal{F}_n$, a function $t \in \mathcal{T}_n$, and a probability distribution $\mathcal{D}_n$ over $\mathcal{S}_n$. Nature's choice $\langle t, \mathcal{D}_n \rangle$ is $(1 - \alpha)$-**realizable for** $\langle \vartriangleleft, f \rangle$ if the response function r_s for a randomly chosen state s under $\langle \vartriangleleft, f, t \rangle$ has a fixed-point except with probability α over the choice of state s from $\mathcal{D}_n$.*

Definition 3 captures precisely the likelihood α for Nature to present the Predictor with a state s that, under Nature's chosen temporal dynamics t, makes every prediction made by the Predictor (using any hypothesis) in state s self-defeating.

Proposition 2 (Condition for Predicting Introspectively). *Consider a recorder $\vartriangleleft$ of $\mathcal{L}_n$ in $\mathcal{S}_n$, a function $f \in \mathcal{F}_n$, a function $t \in \mathcal{T}_n$, and a probability distribution $\mathcal{D}_n$ over $\mathcal{S}_n$. Then, there exists a function $h \in \mathcal{H}_n$ that is introspectively $(1 - \alpha)$-accurate w.r.t. $\langle \vartriangleleft, f, t, \mathcal{D}_n \rangle$ if and only if Nature's choice $\langle t, \mathcal{D}_n \rangle$ is $(1 - \alpha)$-realizable for $\langle \vartriangleleft, f \rangle$.*

Proof. The claim follows directly from Definitions 2–3. $\square$

The Predictor's goal is, then, to *find* a function $h \in \mathcal{H}_n$ that is close to being introspectively $(1 - \alpha)$-accurate (i.e., to the optimal value that any possible hypothesis could achieve).

4 Learnability

The learning model formalized in this section operates under the following premises / principles: *(P1)* The Predictor *passively* observes the current state of affairs that is drawn from some distribution. *(P2)* A label is *recorded exactly once* and only according to the Predictor's chosen recorder, but recordings can also be simulated without affecting the actual state of affairs. *(P3)* The Predictor observes the *entire* state of affairs that results according to Nature's temporal dynamics, irrespectively of the prediction task that it undertakes. *(P4)* The

Predictor seeks introspectively accurate predictions *to the extent that they exist* by the choices of Nature, with no a priori assumptions. *(P5)* The Predictor has direct access to all of its choices (*predicted feature* and *recorder*), but not to those of Nature. *(P6)* Learning is computationally efficient.

Definition 4 (Introspective Learnability). *The environment $\mathrm{E} \triangleq \langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ is **information-theoretically introspectively learnable** if there exists an algorithm $\mathbb{A}$ such that for every positive integer n, every real number $\varepsilon \in (0, 1]$, every real number $\delta \in (0, 1]$, every recorder $\vartriangleleft$ of $\mathcal{L}_n$ in $\mathcal{S}_n$, every function $f \in \mathcal{F}_n$, every function $t \in \mathcal{T}_n$, every probability distribution $\mathcal{D}_n$ over $\mathcal{S}_n$, if $\langle t, \mathcal{D}_n \rangle$ is $(1 - \alpha)$-realizable for $\langle \vartriangleleft, f \rangle$ then there exists a function $h \in \mathcal{H}_n$ such that:*

(i) given E, n, $\vartriangleleft$, f, ε, δ as inputs, $\mathbb{A}$ enters a training phase at the end of which it returns h, and then terminates;

(ii) during the training phase the following are repeated: a state $s \in \mathcal{S}_n$ is drawn from $\mathcal{D}_n$ and given to $\mathbb{A}$; a label $l \in \mathcal{L}_n$ is chosen by $\mathbb{A}$; the state $t(s \vartriangleleft l)$ is given to $\mathbb{A}$;

(iii) h is introspectively $(1 - \alpha - \varepsilon)$-accurate w.r.t. $\langle \vartriangleleft, f, t, \mathcal{D}_n \rangle$, except with probability δ over the randomness employed by $\mathbb{A}$ and $\mathcal{D}_n$ during the training phase.

*If, in addition, algorithm $\mathbb{A}$ runs in time polynomial in $|\mathcal{L}_n|$, $\log |\mathcal{S}_n|$, $\log |\mathcal{F}_n|$, $\log |\mathcal{T}_n|$, $\log |\mathcal{H}_n|$, $1/\varepsilon$, $1/\delta$, $1/(1 - \alpha)$, then environment E is **(efficiently) introspectively learnable**.*

Condition *(i)* follows the typical Learning Theory requirement of having certain problem parameters available as inputs (cf. Principle *(P5)*). Condition *(ii)* follows passive supervised learning (cf. Principles *(P1)* and *(P3)*), while embracing aspects of active learning by having the algorithm partly affect, through recording, the input state (cf. Principle *(P2)*). Condition *(iii)* follows the typical PAC learning semantics in insisting on predictive guarantees, while also incorporating aspects of agnostic learning by acknowledging that the optimal accuracy might be exogenously restricted (cf. Principle *(P4)*).

Efficiency (cf. Principle *(P6)*) adopts the usual polynomial dependence of the running time on the problem parameters. Time linear in $\log |\mathcal{S}_n|$, $\log |\mathcal{F}_n|$, $\log |\mathcal{T}_n|$, $\log |\mathcal{H}_n|$ is needed even to read or write an element of $\mathcal{S}_n$, $\mathcal{F}_n$, $\mathcal{T}_n$, $\mathcal{H}_n$. Time linear in $1/\varepsilon$ and logarithmic in $1/\delta$ is needed to compensate for the arduous requirement to achieve arbitrarily high accuracy and confidence in the returned hypothesis [Ehrenfeucht *et al.*, 1989]. Time linear in $1/(1 - \alpha)$ is needed to compensate for the (possibly small) probability $1 - \alpha$ of drawing states for which a correct prediction exists; an analogous treatment applies for noisy instances [Angluin and Laird, 1988]. To support a general, task-independent formulation, the introspective learning algorithm is a priori oblivious to the actual value of α (as determined by Nature's choices), even if its running time may depend on it. Lastly, time linear in $|\mathcal{L}_n|$ (as opposed to $\log |\mathcal{L}_n|$) is needed even to enumerate the labels in $\mathcal{L}_n$ during the training phase, which is shown next to be necessary. In effect, the following result shows the necessity of employing some form of randomized trials for introspective learning, at least when the labels lack any usable internal structure.

Proposition 3 (Linear Dependence on Number of Labels). *Any algorithm that information-theoretically introspectively learns an environment $\langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ runs in time $\Omega(|\mathcal{L}_n|)$.*

Proof. Consider a presumed counterexample algorithm to the claim. Let $\mathcal{D}_n$ assign probability 1 to a state s, so the response function r_s under $\langle \triangleleft, f, t \rangle$ has one fixed-point. For $\varepsilon, \delta < 1$, the algorithm identifies a fixed-point of r_s, except with probability δ over its randomness. For large n, however, the algorithm's running time is less than $(1 - \delta) \cdot |\mathcal{L}_n|$, and probing r_s for its fixed-point fails with probability more than δ. □

Given the allowed computational resources, then, an introspective learner $\mathbb{A}$ seeks to find a fixed-point of the response function r_s for s under $\langle \triangleleft, f, t \rangle$, for states s of sufficient probability mass under $\mathcal{D}_n$. To appreciate the central aspects of the learning model, we next consider a possible training strategy for algorithm $\mathbb{A}$. Following Condition *(ii)* of Definition 4, algorithm $\mathbb{A}$ first obtains a state $s_0 \in \mathcal{S}_n$ drawn from $\mathcal{D}_n$. Having no indication as of yet on what the sought function h might look like, the algorithm may choose an arbitrary label $l \in \mathcal{L}_n$, and compute the state $s_1 \triangleq s_0 \triangleleft l_0$ that results from recording l_0 in s_0. After the label is recorded, the algorithm receives the state $s_2 \triangleq t(s_1)$, and may compute the label $l_2 \triangleq f(s_2)$. At this point the algorithm may seek to exploit the information that has been made available so far.

We examine the algorithm's behavior by case analysis. Assume first that $l_0 = l_2$, so that the algorithm predicts correctly. It can, then, safely set $h(s_0) := l_0$. Assume now that $l_0 \neq l_2$, so that the algorithm predicts incorrectly. Note that, unlike in typical supervised learning models, the correct prediction is not revealed to the algorithm. Indeed, label l_2 is not necessarily the appropriate value to assign to $h(s_0)$; for the same state s_0, Nature may have responded to a prediction of l_2 with l_2', to a prediction of l_2' with l_2'', and so on, without any of these labels corresponding to a fixed-point of the response function r_{s_0} for s_0 under $\langle \triangleleft, f, t \rangle$. Thus, algorithm $\mathbb{A}$ can safely infer in this case only that $h(s_0) \neq l_0$, but not that $h(s_0) = l_2$.

In either case, the algorithm has access to a pair $\langle s_1, s_2 \rangle$ comprising the initial / successor states of the environment according to the temporal dynamics t. The algorithm may seek, then, to solve the auxiliary learning problem of approximating t by an auxiliary hypothesis g (that agrees with t on "sufficiently many" inputs) [Michael, 2011]. If algorithm $\mathbb{A}$ obtains such a hypothesis g, it can compute a highly-accurate h: consider the response function $\tilde{r}_s(l) = f(g(s \triangleleft l)) \approx f(t(s \triangleleft l))$ for s under $\langle \triangleleft, f, g \rangle$; systematically test it on each label in $\mathcal{L}_n$ to identify the least fixed-point l; if l exists, set $h(s) := l$. Since the algorithm has access to recorder $\triangleleft$ and function f, it can compute $f(g(s \triangleleft l))$. Note that the computation of h in this manner cannot be done explicitly during the training phase, but it is done lazily when hypothesis h is applied during the testing phase to make a prediction on state s.

Why is the problem not solved by learning t and computing h? Although, the algorithm obtains a pair $\langle s_1, s_2 \rangle$ of initial / successor states even when $l_0 \neq l_2$, in that case state s_1 is not drawn from the "right" probability distribution (in which only correct predictions are recorded)! Thus, while $\langle s_1, s_2 \rangle$ can be used to learn t, it might offer little in ensuring that g approximates t on those states that are important (for computing h).

The "right" probability distribution is not $\mathcal{D}_n$ from which Nature draws states, but a probability distribution $\mathcal{D}_n^+$ over states $s \triangleleft l$ induced by drawing a state $s \in \mathcal{S}_n$ from $\mathcal{D}_n$, and

then recording the least (or any other prescribed tie-breaking function) fixed-point l of the response function r_s for s under $\langle \triangleleft, f, t \rangle$, if one exists. It is on states of sufficiently high probability mass under probability distribution $\mathcal{D}_n^+$ that g should approximate t. However, sampling from $\mathcal{D}_n^+$ seems to presuppose a highly-accurate h to identify an appropriate label l for state s. Are we back to square one? Fortunately not quite, since it suffices to sample from $\mathcal{D}_n^+$ only sufficiently often.

Theorem 1 (Sufficient Learnability Condition w.r.t. $\mathcal{T}$). *The environment $\langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ is introspectively learnable if the class $\mathcal{T}$ is causally learnable.*

Proof. Let $\mathbb{B}$ be a causal learning algorithm for class $\mathcal{T}$. Consider an environment $\mathrm{E} \triangleq \langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ and algorithm $\mathbb{A}$:

> Given inputs E, n, $\triangleleft$, f, ε, δ, algorithm $\mathbb{A}$ starts by simulating the execution of algorithm $\mathbb{B}$ on inputs $\mathcal{S}_n$, $\mathcal{T}_n$, $\varepsilon(1 - \alpha)/|\mathcal{L}_n|$, $\delta/2$; we assume for now that α is known, and deal with the task of identifying it later. Whenever algorithm $\mathbb{B}$ requests a training instance, algorithm $\mathbb{A}$ samples a state $s \in \mathcal{S}_n$ from the probability distribution $\mathcal{D}_n$, selects a label $l \in \mathcal{L}_n$ uniformly at random, computes $s_1 \triangleq s \triangleleft l$, receives $s_2 \triangleq t(s_1)$, and passes $\langle s_1, s_2 \rangle$ to algorithm $\mathbb{B}$. Once algorithm $\mathbb{B}$ returns a hypothesis g and terminates, algorithm $\mathbb{A}$ constructs the hypothesis $h \in \mathcal{H}_n$ defined so that on any input $s \in \mathcal{S}_n$, $h(s)$ equals the lexicographically first fixed-point of the response function $\tilde{r}_s(l) = f(g(s \triangleleft l))$ for s under $\langle \triangleleft, f, g \rangle$; if no such fixed-point exists, then $h(s)$ equals $\perp$ (or any arbitrary label). Algorithm $\mathbb{A}$ returns h, and terminates.

Let $\mathcal{D}_n^+$ be the probability distribution defined as in our discussion above. Since a fixed-point exists with probability $1 - \alpha$, and since algorithm $\mathbb{A}$ records a uniformly at random label in s, it follows that with probability $(1 - \alpha)/|\mathcal{L}_n|$ the chosen label is the least fixed-point of r_s. Overall, then, algorithm $\mathbb{A}$ obtains a pair $\langle s_1, s_2 \rangle$ satisfying $s_2 = t(s_1)$, so that state s_1 is sampled from $\mathcal{D}_n^+$ with probability $(1 - \alpha)/|\mathcal{L}_n|$, and from some other probability distribution, say $\mathcal{D}_n^-$, with the remaining probability. It follows that if $g(s_1) = t(s_1)$ except with probability $\varepsilon(1 - \alpha)/|\mathcal{L}_n|$ on states s_1 sampled from the probability distribution induced by $\mathcal{D}_n^+$ and $\mathcal{D}_n^-$, then $g(s_1) = t(s_1)$ except with probability ε on states s_1 sampled from $\mathcal{D}_n^+$. The correctness of algorithm $\mathbb{A}$ follows.

The running time is as required by Definition 4 even if α is replaced with any upper-bound $\bar{\alpha}$ of α such that $1 - \bar{\alpha}$ is polynomially related to $1 - \alpha$. Such an $\bar{\alpha}$ can be efficiently computed iteratively as done when learning with an unknown noise rate [Laird, 1988; Angluin and Laird, 1988]. □

Although algorithm $\mathbb{A}$'s inputs resemble a causal learner's, the condition of introspective learnability is not necessary.

Theorem 2 (Counterexample to Necessity w.r.t. $\mathcal{T}$). *There exists an environment $\langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ that is introspectively learnable, but the class $\mathcal{T}$ is not causally learnable. (The result holds even if classes $\mathcal{F}$ and $\mathcal{F} \circ \mathcal{T}$ are causally learnable.)*

Proof. Consider an environment $\mathrm{E} \triangleq \langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ such that $\mathcal{F}$ comprises constant functions only, and the class $\mathcal{T}$ is not causally learnable (see, e.g., [Michael, 2011] for such a

class). Observe, however, that the environment E is (trivially) introspectively learnable: given $f \in \mathcal{F}_n$ as input, the learning algorithm can return the hypothesis $h(\cdot) := f(\cdot)$, for which it holds that $h(s) = f(s) = f(t(s \triangleleft h(s)))$ for every state $s \in \mathcal{S}_n$. With regards to the parenthetical constraint, $\mathcal{F}_n \circ \mathcal{T}_n = \mathcal{F}_n$, and $\mathcal{F}$ is causally learnable. $\square$

Theorem 2 illustrates a scenario where the temporal dynamics $\mathcal{T}$ are unlearnable, but the features $\mathcal{F}$ of the successor states are simple enough to nullify this unlearnability. Similar in spirit results can be established for more expressive classes $\mathcal{F}$, all attesting to the following point: If the features $\mathcal{F}$ that the Predictor can reliably predict are "insufficiently expressive", then it does not follow that the Predictor is able to learn the temporal dynamics $\mathcal{T}$ of Nature. We formalize this below.

Definition 5 (Invertibility Dimension). *Consider a set $\mathcal{C}_n$ of functions with domain $\mathcal{X}_n$ and codomain $\mathcal{Y}_n$. The **invertibility dimension of** $\mathcal{C}_n$ equals the minimum non-negative integer d for which there exist functions $c_1, c_2, \ldots, c_d \in \mathcal{C}_n$ and a function $u : \mathcal{Y}_n^d \to \mathcal{X}_n$ such that (i) all of the $d+1$ functions can be evaluated on their inputs in time polynomial on the input size, and (ii) for every element $x \in \mathcal{X}_n$, it holds that $u(c_1(x), c_2(x), \ldots, c_d(x)) = x$. If no such d exists, the invertibility dimension of $\mathcal{C}_n$ is ∞.*

Recall that each $f \in \mathcal{F}_n$ captures a feature of some state of affairs. d is the minimum number of such features that would be required to get full information for a given state of affairs.

Theorem 3 (Necessary Learnability Condition w.r.t. $\mathcal{T}$). *The environment $\langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$ is introspectively learnable only if the class $\mathcal{T}$ is causally learnable in non-uniform time polynomial in the usual learning parameters and also in the invertibility dimension d of $\mathcal{F}_n$. For uniform time, the necessary condition needs to account also for the time required to identify $d, u, c_1, c_2, \ldots, c_d$ as given in Definition 5.*

Proof. Let $\mathbb{A}$ be an introspective learning algorithm for environment $E \triangleq \langle \mathcal{L}, \mathcal{S}, \mathcal{F}, \mathcal{T}, \mathcal{H} \rangle$. Consider algorithm $\mathbb{B}$:

> Given inputs $\mathcal{L}_n, \mathcal{S}_n, \mathcal{F}_n, \mathcal{T}_n, n, \varepsilon, \delta$, algorithm $\mathbb{B}$ starts by simulating the execution of d independent copies of algorithm $\mathbb{A}$, so that the i-th copy is simulated on inputs $E, n, \triangleleft, c_i, \varepsilon/d, \delta/2d$, where $\triangleleft$ is the trivial recorder, d is the invertibility dimension of $\mathcal{F}_n$, and $u, c_1, c_2, \ldots, c_d$ are as in Definition 5 for $\mathcal{C}_n = \mathcal{F}_n, \mathcal{X}_n = \mathcal{S}_n, \mathcal{Y}_n = \mathcal{L}_n$. Whenever algorithm $\mathbb{A}$ requests a training instance, algorithm $\mathbb{B}$ samples a state $s \in \mathcal{S}_n$ and $t(s)$ from the probability distribution $\mathcal{D}_n$, passes s to algorithm $\mathbb{A}$ and receives a label $l \in \mathcal{L}_n$, sets $s_1 \triangleq s \triangleleft l$ and $s_2 \triangleq t(s_1)$, and passes s_2 to algorithm $\mathbb{A}$. Once each copy of algorithm $\mathbb{A}$ returns a hypothesis h_i and terminates, algorithm $\mathbb{B}$ constructs the hypothesis $h \in \mathcal{H}_n$ defined so that on any input $s \in \mathcal{S}_n$, $h(s)$ equals $u(h_1(s), h_2(s), \ldots, h_d(s))$. Algorithm $\mathbb{B}$ returns h, and terminates.

Correctness follows by observing that algorithm $\mathbb{A}$ is simulated given the trivial recorder $\triangleleft$, which, in particular, implies $\alpha = 0$. The computational complexity is non-uniform, since given $\mathcal{F}_n$ we assume access to d and $u, c_1, c_2, \ldots, c_d$. $\square$

Analogous results to Theorems 1-2 can be obtained w.r.t. the class $\mathcal{F} \circ \mathcal{T}$ (the composition of functions from $\mathcal{F}$ and $\mathcal{T}$).

5 Conclusions

Understanding how the realizability of predictions is affected by the actions effected by the predictions themselves, is key in designing intelligent and autonomous machines. We have put forward a learning-theoretic model within which the problem can be formally investigated. Further connections to the general issue of usefully accommodating predictions in the process of learning [Michael, 2008; 2014] remain to be studied.

Future work can extend the model to settings where: state sensing is noisy [Angluin and Laird, 1988; Shackelford and Volper, 1988] or partial [Michael, 2010]; only statistical information is available [Kearns, 1998]; the successor state is not always available [Balcan and Blum, 2010]; temporal dynamics are stochastic [Kearns and Schapire, 1994]; the initial state is partially chosen / fixed by the learner [Angluin, 1988; Michael, 2011] or a teacher [Goldman and Mathias, 1996]; the labels exhibit learnable internal structure that allows introspective learnability with sub-linear (and even logarithmic) time-dependence on the number of labels (cf. Proposition 3). Further work can be done on developing *unbiased* introspective learners that return a uniformly at random chosen hypothesis among all highly-accurate ones, addressing, thus, concerns of manipulation by a forecaster [Simon, 1954, p. 251].

On a second front, future work can seek to approach as special instances of introspective forecasting, tasks whose understanding would prima facie benefit from such a treatment: *(i)* accommodating for the side-effects of announcing poll results [Simon, 1954]; *(ii)* personalizing the education of students without risking to overwhelm and discourage them [Michael, 2012]; *(iii)* anticipating how spam senders adapt to the spam filtering tools in use [Fawcett, 2003][1]; and *(iv)* playing in general game competitions [Genesereth *et al.*, 2005], where a technique known as Monte Carlo Tree Search [Browne *et al.*, 2012], which effectively reacts to a randomly predicted strategy for an agent's opponent, was found to be very effective. Approaching randomized trials in clinical research as special instances of introspective forecasting could conceivably yield optimal interventions more efficiently and less intrusively.

A third front relates to investigating the deeper connection of this work to fixed-points. We have sought to identify fixed-points to the extent they happen to exist (as given by α), side-stepping concerns on the conditions under which they exist [Brouwer, 1911], and without a need to make multiple queries to anyone of Nature's response functions [Sperner, 1928]. We have considered the identification of *exact* fixed-points for a *set* of nominal-valued response functions, and sought to efficiently ε-approximate the subset of functions for which this was possible. One could examine the case of identifying approximately the fixed-points themselves [Etessami and Yannakakis, 2010], either by introducing a distance metric between labels, or by clustering them in equivalence classes. Investigating learnability in such coarser settings could be informed by work on learning to approximate real-valued functions [Balcan and Harvey, 2012], while also accommodating

---

[1]Offered as a challenge to Knowledge Discovery and Data Mining: "[this perspective is] foreign to most data mining researchers: the data are mined and the results are deployed, but the data environment is not considered to be an active entity that will react in turn".

environments with sets $\mathcal{L}_n, \mathcal{S}_n, \mathcal{F}_n, \mathcal{T}_n, \mathcal{H}_n$ of infinite size.

Much work remains to be done in exploring the intricacies of introspective forecasting and its utilization in various settings. Yet, its demonstrable amenability to a task-independent formalization already offers a glimpse into its possible role as a convergence point across disciplines, contra to arguments offered against the unity of science [Fodor, 1974]. The gist of an oft-cited argument is that prediction "tend[s] to change the initial course of developments" in the social sciences, whereas "[t]his is not true of prediction in fields which do not pertain to human conduct" [Merton, 1936, p. 903]. Our work responds by demonstrating that to the extent that prediction is possible in any given setting and discipline, the same methodology can yield optimal, in a defined sense, predictions.

References

[Angluin and Laird, 1988] Dana Angluin and Philip Laird. Learning From Noisy Examples. *Machine Learning*, 2(4):343–370, 1988.

[Angluin, 1988] Dana Angluin. Queries and Concept Learning. *Machine Learning*, 2(4):319–342, 1988.

[Balcan and Blum, 2010] Maria-Florina Balcan and Avrim Blum. A Discriminative Model for Semi-Supervised Learning. *Journal of the ACM*, 57(3):19:1–19:46, 2010.

[Balcan and Harvey, 2012] Maria-Florina Balcan and Nicholas Harvey. Learning Submodular Functions. In *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2012)*, pages 846–849, 2012.

[Box et al., 2008] George Box, Gwilym Jenkins, and Gregory Reinsel. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Inc., Hoboken, New Jersey, U.S.A., 4th edition, 2008.

[Brouwer, 1911] Luitzen Brouwer. Über Abbildungen von Mannigfaltigkeiten. *Mathematische Annalen*, 71(1):97–115, 1911.

[Browne et al., 2012] Cameron Browne, Edward Powley, Daniel Whitehouse, Simon Lucas, Peter Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A Survey of Monte Carlo Tree Search Methods. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(1):1–43, 2012.

[Ehrenfeucht et al., 1989] Andrzej Ehrenfeucht, David Haussler, Michael Kearns, and Leslie Valiant. A General Lower Bound on the Number of Examples Needed for Learning. *Information and Computation*, 82(3):247–261, 1989.

[Etessami and Yannakakis, 2010] Kousha Etessami and Mihalis Yannakakis. On the Complexity of Nash Equilibria and Other Fixed Points. *SIAM Journal on Computing*, 39(6):2531–2597, 2010.

[Fawcett, 2003] Tom Fawcett. "In Vivo" Spam Filtering: A Challenge Problem for KDD. *SIGKDD Explorations Newsletter*, 5(2):140–148, 2003.

[Fodor, 1974] Jerry Fodor. Special Sciences (or: The Disunity of Science as a Working Hypothesis). *Synthese*, 28(2):97–115, 1974.

[Genesereth et al., 2005] Michael Genesereth, Nathaniel Love, and Barney Pell. General Game Playing: Overview of the AAAI Competition. *AI Magazine*, 26(2):62–72, 2005.

[Goldman and Mathias, 1996] Sally Goldman and David Mathias. Teaching a Smarter Learner. *Journal of Computer and System Sciences*, 52(2):255–267, 1996.

[Kearns and Schapire, 1994] Michael Kearns and Robert Schapire. Efficient Distribution-Free Learning of Probabilistic Concepts. *Journal of Computer and System Sciences*, 48(3):464–497, 1994.

[Kearns, 1998] Michael Kearns. Efficient Noise-Tolerant Learning from Statistical Queries. *Journal of the ACM*, 45(6):983–1006, 1998.

[Laird, 1988] Philip Laird. *Learning from Good and Bad Data*. Kluwer Academic Publishers, Norwell, Massachusetts, U.S.A., 1988.

[Li et al., 2010] Lihong Li, Wei Chu, John Langford, and Robert Schapire. A Contextual-Bandit Approach to Personalized News Article Recommendation. In *Proc. 19th International Conference on World Wide Web (WWW 2010)*, pages 661–670, 2010.

[Littlestone, 1988] Nick Littlestone. Learning Quickly when Irrelevant Attributes Abound: A New Linear-Threshold Algorithm. *Machine Learning*, 2(4):285–318, 1988.

[Merton, 1936] Robert Merton. The Unanticipated Consequences of Purposive Social Action. *American Sociological Review*, 1(6):894–904, 1936.

[Michael, 2008] Loizos Michael. *Autodidactic Learning and Reasoning*. Doctoral Dissertation, Harvard University, U.S.A., 2008.

[Michael, 2010] Loizos Michael. Partial Observability and Learnability. *Artificial Intelligence*, 174(11):639–669, 2010.

[Michael, 2011] Loizos Michael. Causal Learnability. In *Proc. 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, pages 1014–1020, 2011.

[Michael, 2012] Loizos Michael. "Primum Non Nocere" for Personalized Education. In *Proc. Workshop on Personalizing Education With Machine Learning (PEwML 2012)*, 2012.

[Michael, 2014] Loizos Michael. Simultaneous Learning and Prediction. In *Proc. 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014)*, pages 348–357, 2014.

[Murphy, 2002] Kevin Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. Doctoral Dissertation, University of California, Berkeley, U.S.A., 2002.

[Robbins, 1952] Herbert Robbins. Some Aspects of the Sequential Design of Experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

[Shackelford and Volper, 1988] George Shackelford and Dennis Volper. Learning k-DNF with Noise in the Attributes. In *Proc. 1st Annual Workshop on Computational Learning Theory (COLT 1988)*, pages 97–103, 1988.

[Simon, 1954] Herbert Simon. Bandwagon and Underdog Effects and the Possibility of Election Predictions. *The Public Opinion Quarterly*, 18(3):245–253, 1954.

[Sperner, 1928] Emanuel Sperner. Neuer Beweis für die Invarianz der Dimensionszahl und des Gebietes. *Abhandlungen aus dem Mathematischen Seminar der Universitt Hamburg*, 6(1):265–272, 1928.

[Sutton and Barto, 1998] Richard Sutton and Andrew Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, Massachusetts, U.S.A., 1998.

[Valiant, 1984] Leslie Valiant. A Theory of the Learnable. *Communications of the ACM*, 27(11):1134–1142, 1984.