Interactive Optimal Teaching with Unknown Learners

Francisco S. Melo, Carla Guerra, Manuel Lopes
INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
fmelo@inesc-id.pt, manuel.lopes@tecnico.ulisboa.pt

Abstract

This paper introduces a new approach for machine teaching that partly addresses the (unavoidable) mismatch between what the teacher assumes about the learning process of the student and the student's actual process. We analyze several situations in which such mismatch takes place and we show that, even in the simple case of a Bayesian Gaussian learner, the lack of knowledge regarding the student's learning process significantly deteriorates the performance of machine teaching: while perfect knowledge of the student ensures that the target is learned after a finite number of samples, lack of knowledge thereof implies that the student will only learn asymptotically (i.e., after an infinite number of samples). We propose *interactivity* as a means to mitigate the impact of imperfect knowledge and show that, by using interactivity, we are able to attain significantly faster convergence, in the worst case. Finally, we discuss the implications of our results in single- and multi-student settings.

1 Introduction

The use of computer and Internet technology in education has seen an enormous development in recent years, with the appearance of *massive open online courses* (MOOCs) and *intelligent tutoring systems* (ITS). MOOCs have created new opportunities for widespread education, while ITS provide personalized contents and/or exercises tailored for each user [Anderson *et al.*, 1995; Koedinger *et al.*, 1997; Nkambou *et al.*, 2010; Clement *et al.*, 2015; Mota *et al.*, 2015].

Typically, ITS do not consider the specific structure of the teaching goal and abstract the teaching experience as the process of providing a teaching item for which a response is received. In contrast, the field of *machine teaching* (MT) considers directly the task to teach, and seeks to select the smallest amount of information necessary for a student to learn that specific task [Balbach and Zeugmann, 2009; Zhu, 2013; 2015]. Such quantity is often known as the *teaching dimension* (TD) of that task [Goldman and Kearns, 1995; Shinohara and Miyano, 1991].

Machine teaching has the potential to strongly reduce the effort required from both learner and teacher. For instance, the structural properties of the chemistry domain can be used to accurately estimate the knowledge of a student [Davenport *et al.*, 2012] and use that to drive the selection of new teaching materials.

The main problem with machine teaching—which may partly explain its lack of use in actual ITS systems—is that the assumptions about what is known about the learner are often not realistic. Most methods assume a perfect knowledge about the learner, its learning process, and associated parameters. Such assumption is particularly unrealistic when dealing with human learners, but little work exists in the literature that explicitly addresses this unavoidable mismatch between what the machine teaching systems assume about the learners and the learners themselves.

In this work we take a step towards addressing such problem. Our contributions are two-fold:

- In the first part of the paper (Section 2), we show that, even in the simple situation of Bayesian Gaussian learners, the lack of knowledge regarding the student's learning process significantly deteriorates the performance of machine teaching: while perfect knowledge of the student ensures exact learning with a finite number of samples, incorrect information about the student implies that the latter will only learn asymptotically (i.e., after an infinite number of samples).

  The results of such conceptual exercise suggest that the lack of knowledge about the learner hampers the potential usefulness of machine teaching even in the simplest situations and we may expect this effect to be even more severe when interacting with human learners.

- In the second part of the paper (Section 3) we propose *interactive teaching* as a possible avenue to address the aforementioned mismatch between teacher and learner. In interactive teaching, the teacher interleaves the presentation of novel information with moments in which it seeks to assess the students current state. We show that, in the case of a Bayesian Gaussian learner, the use of interactivity significantly gains in terms of learning performance.

Finally, we conduct a simple study in which human students must learn the average price of an apartment in an un-

known city. Our results confirm the main conclusions of our conceptual analysis, namely that (i) the assumptions in the machine teaching system fail to reproduce the students' actual learning process; (ii) such mismatch renders machine teaching no different from a "random sampler"; (iii) the use of interactivity significantly improves the performance of the machine teaching system.

2 Teaching the Mean of a Gaussian Distribution

Let us consider the situation in which a teacher must select samples, $\{x_1, \ldots, x_N\}$, from which a learner then estimates the mean of a Gaussian distribution $\mathcal{N}(\mu^*, \tau)$, where μ^* represents the mean of the distribution and τ represents the corresponding precision [Zhu, 2013].[1]

Throughout our derivations, we assume that the precision of the distribution, τ, is known. We also assume the learner is a Bayesian learner with a (conjugate) prior $\mathcal{N}(\mu_0, \tau_0)$. In other words, upon receiving a sample x, the learner computes the updated estimate

$$\mu_1 = \frac{\tau_0}{\tau_0 + \tau}\mu_0 + \frac{\tau}{\tau_0 + \tau}x.$$

More generally, upon receiving the nth sample, x_n, the learner computes the updated estimate

$$\mu_n = \frac{\tau_{n-1}}{\tau_n}\mu_{n-1} + \frac{\tau}{\tau_n}x_n,$$

where $\tau_n = \tau_0 + n\tau, n \geq 0$.

2.1 Optimal Teaching with Perfect Knowledge

We start by considering the most favorable situation, in which the teacher is fully aware of the learner's model parameters, μ_0 and τ_0. In this case, if the teacher provides the sample

$$x_1 = \frac{\tau_1}{\tau}\mu^* - \frac{\tau_0}{\tau}\mu_0,$$

the learner's updated estimate is

$$\mu_1 = \frac{\tau_0}{\tau_1}\mu_0 + \mu^* - \frac{\tau_0}{\tau_1}\mu_0 = \mu^*.$$

In other words, if the teacher is perfectly aware of the learner's prior, a single sample is sufficient to ensure perfect learning. Any subsequent samples should take the value

$$x_n = \frac{\tau_n}{\tau}\mu^* - \frac{\tau_{n-1}}{\tau}\mu^* = \mu^*,$$

which ensure that $\mu_n = \mu^*$ for all $n > 0$.

In practice, however, the teacher will not have perfect knowledge regarding the learner. Let us then consider the situation in which the learner's parameters do not correspond exactly to the values that the teacher assumes.

2.2 Teaching with the Wrong Prior Mean

We first consider the situation where the learner's prior mean—henceforth denoted as μ_0^L—differs from the value assumed by the teacher—henceforth denoted as μ_0^T. Repeating the derivations in the previous subsection, the sample provided by the teacher will be

$$x_1 = \frac{\tau_1}{\tau}\mu^* - \frac{\tau_0}{\tau}\mu_0^T,$$

leading to the learner's updated estimate

$$\mu_1^L = \frac{\tau_0}{\tau_1}\mu_0^L + \frac{\tau}{\tau_1}x_1 = \frac{\tau_0}{\tau_1}\Delta\mu + \mu^*,$$

where $\Delta\mu = \mu_0^L - \mu_0^T$. Following the reasoning in the previous section, we again conclude the teacher will then select $x_n = \mu^*$, for all $n > 1$, leading to the learner's estimates

$$\mu_n^L = \frac{\tau_0}{\tau_0 + n\tau}\Delta\mu + \mu^*, \tag{1}$$

which asymptotically converges to μ^* at a rate $O(\frac{1}{n})$.[2]

We can see that the term $\Delta\mu$—corresponding to the mismatch between the parameters assumed by the teacher and their actual values—is responsible for the different behavior of the learner when compared to the previous subsection and, if $\Delta\mu = 0$, we recover our previous result.

2.3 Teaching with the Wrong Prior Precision

Let us now consider an alternative situation, in which it is the learner's prior precision, denoted that τ_0^L, that differs from the value assumed by the teacher, denoted as τ_0^T. Repeating again the derivations in the previous subsections, we get

$$x_1 = \frac{\tau_1^T}{\tau}\mu^* - \frac{\tau_0^T}{\tau}\mu_0, \tag{2}$$

where $\tau_1^T = \tau_0^T + \tau$. The sample x_1 leads to the learner's updated estimate

$$\mu_1^L = \frac{\tau_0^L}{\tau_1^L}\mu_0 + \frac{\tau}{\tau_1^L}x_1 = \frac{\mu_0}{\tau_1^L}\Delta\tau + \frac{\tau_1^T}{\tau_1^L}\mu^*,$$

where $\Delta\tau = \tau_0^L - \tau_0^T$. We write μ_1^L to emphasize the fact that the learner's estimate after the first sample differs from that assumed by the teacher—i.e., the teacher assumes that the learner's updated mean is $\mu_1^T = \mu^*$. Subsequent samples by the teacher will thus be $x_n = \mu^*$, for $n > 1$, leading to the general estimate

$$\mu_n^L = \frac{\mu_0}{\tau_0^L + n\tau}\Delta\tau + \frac{\tau_0^T + n\tau}{\tau_0^L + n\tau}\mu^*, \tag{3}$$

which again asymptotically converges to μ^* at a rate $O(\frac{1}{n})$.

It is interesting to compare expressions (1) and (3), noting that they have, essentially, the same form: a vanishing term that weights the difference between the learner and the teacher, plus a term that converges to μ^*. And, once again, if $\Delta\tau = 0$, we recover the initial one-step convergence result.

---

[1]Given a normal distribution with mean μ and variance σ^2, the precision, τ, is the inverse of the variance, i.e., $\tau = \sigma^{-2}$. We use the precision throughout the document for ease of presentation.

[2]In other words, $\mu_n = \mu^* + O(\frac{1}{n})$.

2.4 Teaching with the Wrong Prior and Precision

We now consider the situation where both $\mu_0^T \neq \mu_0^L$ and $\tau_0^T \neq \tau_0^L$. The first sample provided by the teacher is

$$x_1 = \frac{\tau_1^T}{\tau}\mu^* - \frac{\tau_0^T}{\tau}\mu_0^T.$$

This sample results in the updated estimate

$$\mu_1^L = \frac{\tau_0^L}{\tau_1^L}\mu_0^L - \frac{\tau_0^T}{\tau_1^L}\mu_0^T + \frac{\tau_1^T}{\tau_1^L}\mu^*. \qquad (4)$$

Letting $\bar{\mu}_0 = \frac{1}{2}(\mu_0^L + \mu_0^T)$ and $\bar{\tau}_0 = \frac{1}{2}(\tau_0^L + \tau_0^T)$, we can rearrange (4) as

$$\mu_1^L = \frac{\bar{\tau}_0}{\tau_1^L}\Delta\mu + \frac{\bar{\mu}_0}{\tau_1^L}\Delta\tau + \frac{\tau_1^T}{\tau_1^L}\mu^*.$$

Since, as before, $\mu_1^T = \mu^*$, subsequent samples verify $x_n = \mu^*, n > 1$ and the updated estimates take the general form

$$\mu_n^L = \frac{\bar{\tau}_0}{\tau_n^L}\Delta\mu + \frac{\bar{\mu}_0}{\tau_n^L}\Delta\tau + \frac{\tau_n^T}{\tau_n^L}\mu^* \qquad (5)$$

Note that (5) subsumes (1) and (3). For example, if $\tau_0^L = \tau_0^T$, then $\bar{\tau}_0 = \tau_0^L = \tau_0^T$ and we recover (1). Similarly, if $\mu_0^L = \mu_0^T$, then $\bar{\mu}_0 = \mu_0^L = \mu_0^T$ and we recover (3). Therefore, as expected, (5) also converges to μ^* at a rate $O(\frac{1}{n})$.

The expression (5) also evidences the impact of the mismatch of each parameter on the performance of the learner.

2.5 Teaching with the Wrong Algorithm

Finally, we look at the case where the algorithm used by the learner does not exactly match the one assumed by the teacher. In particular, while the teacher assumes that the learner is a Bayesian learner, in reality the learner uses a standard Robbins-Monro stochastic algorithm, corresponding to the update rule

$$\mu_{n+1}^L = (1 - \alpha_n)\mu_n^L + \alpha_n x_{n+1},$$

where $\{\alpha_n\}$ is a non-negative step-size sequence satisfying the standard conditions

$$\sum_{n=1}^{\infty}\alpha_n = \infty, \qquad \sum_{n=1}^{\infty}\alpha_n^2 < \infty.$$

For our purposes, and following [Polyak, 1977], we consider a step-size sequence $\{\alpha_n, n \in \mathbb{N}\}$ such that $n\alpha_n \to K \leq 1$. As before, the sample provided by the teacher is

$$x_1 = \frac{\tau_1^T}{\tau}\mu^* - \frac{\tau_0^T}{\tau}\mu_0^T,$$

leading to the updated estimate

$$\mu_1^L = \mu_0^L - \alpha_0\Delta\mu + \alpha_0\frac{\tau_1^T}{\tau}(\mu^* - \mu_0^T). \qquad (6)$$

To facilitate future analysis, let $\tau_d = \frac{\tau_0^T}{\tau}$ and write (6) as

$$\mu_1^L = \mu^* + \alpha_0\tau_d\Delta\mu + (1 - \alpha_0(\tau_d + 1))(\mu_0^L - \mu^*).$$

Subsequent samples verify $x_n = \mu^*$, leading to the estimate

$$\mu_{n+1}^L = \mu^* + \alpha_0\tau_d\prod_{k=1}^{n}(1 - \alpha_k)\Delta\mu$$
$$+ (1 - \alpha_0(\tau_d + 1))\prod_{k=1}^{n}(1 - \alpha_k)(\mu_0^L - \mu^*). \quad (7)$$

We again get asymptotic convergence to μ^*, since the first and second terms converge to 0 as $n \to \infty$ (see Appendix A for details).

3 Using Interactivity to Overcome Prior Mismatch

In this section we discuss how interactivity can be used to effectively mitigate the impact of a wrong learner model. We assume that the teacher is able to query the learner at any point during the learning process. When queried, the learner responds the value of its current estimate perturbed by noise. This situation can model, for example, the case where there is noise in the communication between teacher and learner, or when the teacher is interacting with a *population* of learners, not all of which have the same prior over the target parameter.

Formally, we assume that, when queried, the learner responds with a value $u_n = \mu_n^L + w_n$, where w_n is a random noise term that is independent of any other entities used in the interaction (namely, μ_n^L) and such that $w_n \sim \mathcal{N}(0, \sigma_n^2)$, where $\{\sigma_n, n \in \mathbb{N}\}$ is a positive, real-valued sequence such that $\sigma_n \to 0$. The teacher assumes that the learner is a Bayesian learner with prior parameters (μ_0^T, τ_0^T) and proceeds as follows:

1. Set $n = 0$.

2. Query the learner. The learner will then respond with the value $u_n = \mu_n^L + w_n$, where w_n is the Gaussian noise.

3. Set $\mu_n^T = u_n$.

4. Provide the sample

$$x_{n+1} = (\tau_d^T + n + 1)\mu^* - (\tau_d^T + n)\mu_n^L - (\tau_d^T + n)w_n$$

where, as before, $\tau_d^T = \frac{\tau_0^T}{\tau}$. Using the sample x_{n+1}, the learner updates its estimate of the mean, leading to

$$\mu_{n+1}^L = (1 - \alpha_n)\mu_n^L + \alpha_n x_{n+1},$$

where we again consider a learner following a standard Robbins-Monro stochastic algorithm. Replacing the expression for x_{n+1} yields

$$\mu_{n+1}^L = \mu^* + (1 - \alpha_n(\tau_d + n + 1))(\mu_n^L - \mu^*)$$
$$- \alpha_n(\tau_d + n)w_n.$$

5. Set $n = n + 1$ and repeat the process from step 2.

Unfolding the recursion, we get

$$\mu_n^L = \mu^* + \prod_{k=0}^{n}(1 - \alpha_k(\tau_d + k + 1))(\mu_0^L - \mu^*)$$
$$- \sum_{k=0}^{n}\alpha_k(\tau_d + k)\prod_{\ell=k+1}^{n}(1 - \alpha_\ell(\tau_d + \ell + 1))w_k. \quad (8)$$
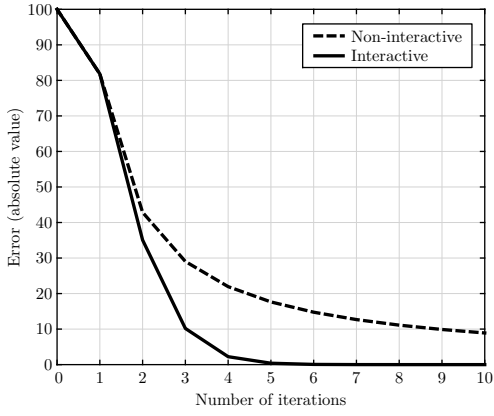
Figure 1: Comparison of the theoretical behavior of the three approaches (non-interactive, noise-free and sample-based), for the particular case where $\alpha_n = 1/(n+1)$, $\tau_d = 1$ and $\Delta\mu = 0$.

The main differences between (8) and (7) are: (i) the term in $\Delta\mu$ is absent in (8), while the term in w_n is absent in (7); (ii) the coefficient multiplying the term $(\mu_0^L - \mu^*)$ takes different forms in the two equations.

3.1 Rates of Convergence

To analyze the rate of convergence of the interactive approach, we rewrite (7) and (8) in their incremental form. For the non-interactive approach, we have that

$$\mu_{n+1}^L = \mu^* + (1 - \alpha_n)(\mu_n^L - \mu^*).$$

Letting $\delta_n = \mu_n^L - \mu^*$, we get

$$\delta_{n+1}^{NI} = (1 - \alpha_n)\delta_n^{NI}, n > 0,$$

where the NI superscript stands for "non-interactive" and

$$\delta_1^{NI} = (1 - \alpha_0(\tau_d + 1))\delta_0^{NI} + \alpha_0 \tau_d \Delta\mu.$$

Considering the interactive approach, we have, for all n,

$$\delta_{n+1}^I = (1 - \alpha_n(\tau_d + n + 1))\delta_n^I - \alpha_n(\tau_d + n)w_n.$$

Since $n\alpha_n \to K \leq 1$, we have that

$$\frac{1 - \alpha_n}{1 - \alpha_n(\tau_d + n + 1)} \to \frac{1}{1 - K} > 1,$$

and it follows that, in expectation, δ_n^I converges to 0 faster than the non-interactive approach. The difference between the two approaches is illustrated in Fig. 1, for the particular case where $\alpha_n = 1/(n+1)$, $\tau_d = 1$ and $\Delta\mu = 0$. The plot clearly illustrates our main conclusion: the interactive approach significantly outperforms the non-interactive one.

4 Results

This section illustrates how the analysis from the previous section translates in actual learning performance. We present two sets of results, the first involving "simulated" students and the second a learning task involving human students. Our results compare interactive and non-interactive learning methods in different scenarios, and showcase the impact that incorrect information about the students can have in the performance of machine teaching.

$\mu_0^T = 1, \mu_0^L = 2, \tau_0^T = \tau_0^L = 1$
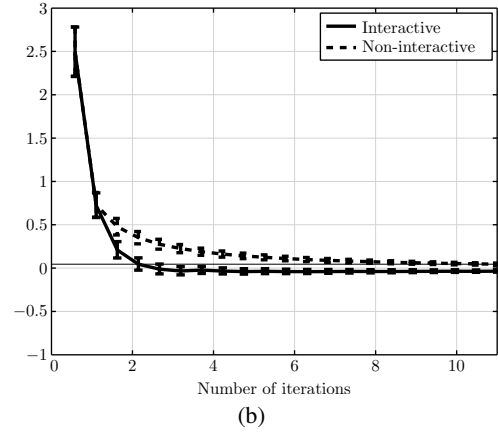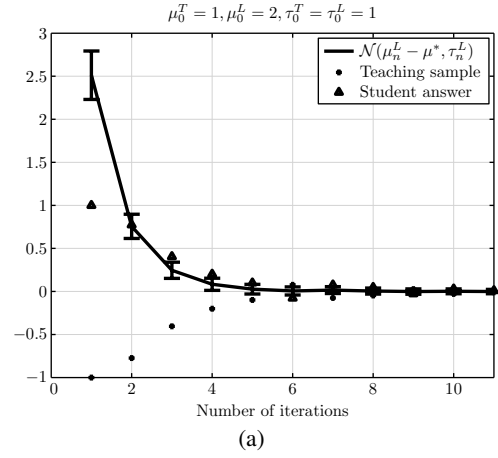

(a)

(b)

Figure 2: Comparison between an interactive (sample-based) and non-interactive teacher. (a) Trace of an interactive session. The solid line corresponds to the estimation error of the student. The triangles correspond to the successive responses of the student, while the dots correspond to the samples provided by the teacher. (b) Evolution of the estimation error for both interactive and non-interactive teaching.

4.1 Simulation Experiments

We consider the Bayesian Gaussian learning setting, in which the student is a Bayesian learner with a prior that differs from that of the teacher. Figure 2 presents a comparison between interactive versus non-interactive learning. The error of the non-interactive approach after 20 samples is reached after only 4 samples using interaction. The figure also show the trace of one run of the algorithm to illustrate what happens during the interaction. We can observe the effect of *overshooting* [Zhu, 2013], where the teaching sample always compensates the student's error: if the student over or underestimates, the teaching sample will take a value that is below or above the target value, respectively. Note also that the answer of the student is not the mean estimated by the student, but a noisy sample thereof, and such noise does not affect the superior performance of the interactive teaching.

4.2 User Study

We conducted a simple user study to assess whether our conclusions from Sections 2 and 3 hold (to some extent) when

using a machine teaching system with actual human students.

Experimental design   In order to conduct our study, we designed an activity which, to some extent, replicates the Gaussian learning scenario discussed throughout the paper. The task consisted of estimating/learning the average monthly rent of an 1-bedroom apartment in an undisclosed city somewhere in the USA. The information provided was selected to balance the impact of the student's prior knowledge in the task.

The task proceeds in rounds where, in each round, the student is asked its current "guess" for the average monthly rent, after which it receives an example of a rent presented to the student as a real rent practiced in the city under consideration. The task continues until the error in the student's guess is below 50\$ up to a maximum of 30 rounds.

Each student repeats the activity three times, each time with a different city:

- "City A" corresponds to a target value (average monthly rent) of 500\$;

- "City B" corresponds to a target value of 1000\$;

- "City C" corresponds to a target value of 1500\$.

The order by which each city is presented to the students is randomized between students.

We investigated the learning performance of the students in 3 different conditions:

- *Condition 1:* The samples are presented to the student following the *interactive teaching* approach described in Section 3. The answers provided by the student are used to compute the next "example rent" to be shown. In order to make the examples more believable to the student, we perturb the value prescribed by the algorithm by a small random amount between 1\$ and 10\$ and enforce that such value is never below 100\$.

- *Condition 2:* The samples presented to the student are selected *randomly* from a Gaussian distribution with the prescribed mean and a standard deviation of 300\$, ignoring the students' responses.

- Condition 3: The samples are presented to the student following the *non-interactive* machine teaching approach outlined in Section 2. We use the first three responses from the student to estimate the prior parameters used in the algorithm, but subsequent samples ignore the responses by the student.

The study involved a total of 62 engineering students divided uniformly among the three conditions (see Table 1). The average age was 23, with 68% males. In analyzing each activity/condition, we disregarded situations in which the students first estimate corresponded to the target value.

Results

The results of our study are summarized in Fig. 3 and prompt several relevant observations:

- The students were able to learn the target value after 30 rounds in all conditions.

	Condition 1	**Condition 2**	Condition 3
"City A"	20	20	19
"City B"	17	18	16
"City C"	22	20	19
Total	22	20	20

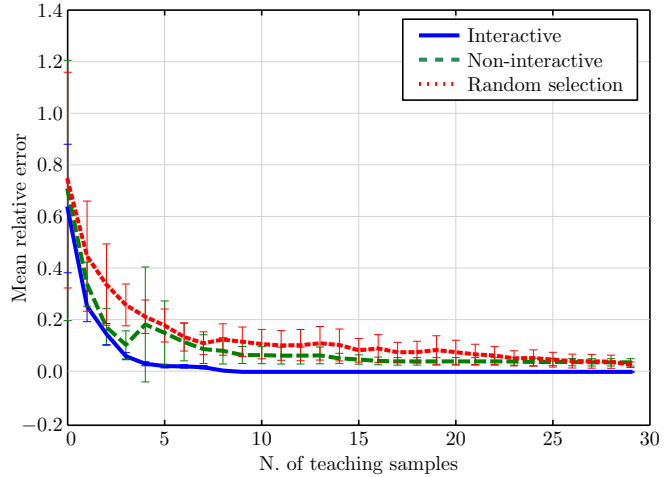Table 1: Total number of participants per condition/activity.

Figure 3: Comparison of the learning performance in all conditions, averaged across all subjects and activities.

- The interactive teaching approach clearly outperforms the other two approaches—both in terms of the average error and in terms of standard deviation of the estimates. In particular, we performed Mann-Whitney U-tests to compare the distributions obtained by the students after 8 and 30 rounds, and observed that the differences were significant in all cases.[3] This result is in line with our conclusion in Section 3: even if the teacher's assumptions regarding the student are wrong (which is most likely the case in our study), interactivity somewhat mitigates such mismatch, rendering. Similar results are observed if we consider the three activities ("City A", "City B" and "City C") individually, as illustrated in Fig. 4.

- Finally, the observed difference in performance between the interactive and non-interactive conditions is larger than the observed difference in performance between the non-interactive and random. This result is also in line with our conclusions in Section 2: when the teacher's assumptions regarding the student are wrong, the performance of the non-interactive machine teaching system may not be significantly different than that of random sampling.

5   Discussion

In this work we investigate the impact that the strong assumptions usually made in the machine teaching literature can have

---

[3]p-values at round 8: 4.3×10^{-5} (I vs R) and 0.03 (I vs NI). p-values at round 30: 0.015 (I vs R) and 0.006 (I vs NI).
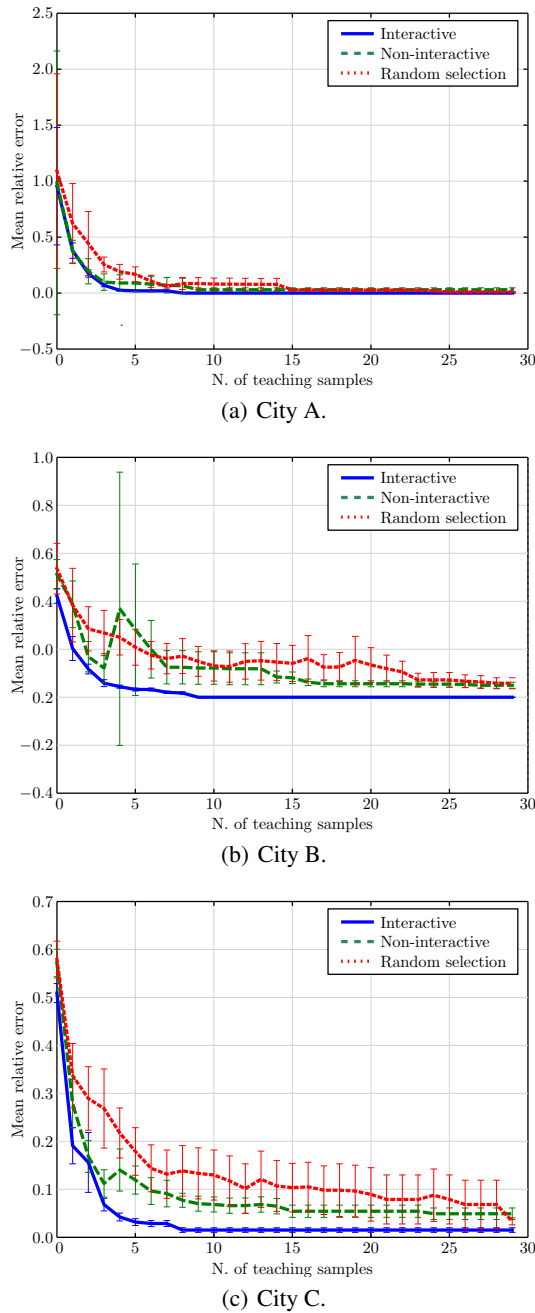
(a) City A.

(b) City B.

(c) City C.

Figure 4: Comparison of the learning performance in all conditions per activity.

in the learning performance of the student.

We conduct a conceptual exercise in which we show that, even in a simple Bayesian Gaussian learning setup, a mismatch between the values of the parameters of the learner and the values assumed by the machine teaching system for those values can have a significant impact on the learning performance of the student. In particular, it renders the machine teaching approach similar to a random sampler.

We then show that, by "closing the loop" between learner and teacher—i.e., by allowing the teacher to *interactively assess* the state of the learner, the impact of the aforementioned mismatch is significantly mitigated.

A simple user study involving human students confirm our conclusions, showing that—even with a simple learner model—an interactive machine teaching system is able to significantly outperform non-interactive alternatives.

The idea of interactive teaching is not new. However, previous works either continued to have too strong assumptions about the learners or do not provide guarantees on how good interactive learning is [Zilles *et al.*, 2008]. For instance [Liu *et al.*, 2017] provides bounds on the teaching dimension for known learners but not for unknown ones. [Cakmak and Thomaz, 2011] as well as [Suh *et al.*, 2016] suggest a combination of machine teaching and active learning but without formal guarantees. [Milli *et al.*, 2017] explore interpretability in machine teaching.

Our discussion also opens the door for situations with multiple students and uncertainty about the students. For example, [Zhu *et al.*, 2017] considers the teaching dimension (TD) for teaching multiple students. They show that the TD grows sub-linearly with the number of students, which roughly means that the number of teaching examples required to teach N students grows sub-linearly with N. Extending our discussion of interactivity to multi-student settings raises a number of novel very interesting questions—e.g., how to interact with multiple students? How to deal with the individual differences between students? How much does the feedback from one student inform the teacher about the state of the class?—thus opening many challenging avenues for future research.

A   Auxiliary Computations

In Section 2.5, it was stated that the two last terms in the right-hand side of (7) converged. We now formally establish such statement.

We have that

$$\lim_{n\to\infty} \prod_{\ell=1}^{n}(1-\alpha_\ell) = \lim_{n\to\infty} \exp\left[\sum_{\ell=1}^{n}\log(1-\alpha_\ell)\right].$$

Using the fact that $\log(1-x) \geq -x, 0 < x < 1$, we have that

$$\lim_{n\to\infty}\sum_{\ell=1}^{n}\log(1-\alpha_\ell) = -\infty,$$

and hence

$$\lim_{n\to\infty}\prod_{\ell=1}^{n}(1-\alpha_\ell) = 0.$$

This establishes the desired convergence result.

Acknowledgments

References

[Anderson *et al.*, 1995] John Anderson, Albert Corbett, Kenneth Koedinger, and Ray Pelletier. Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, 4(2):167–207, 1995.

[Balbach and Zeugmann, 2009] Frank Balbach and Thomas Zeugmann. Recent developments in algorithmic teaching. In *Proceedings of the International Conference on Language and Automata Theory and Applications*, pages 1–18, 2009.

[Cakmak and Thomaz, 2011] Maya Cakmak and Andrea Thomaz. Mixed-Initiative active learning. In *Proceedings of the ICML Workshop on Combining Learning Strategies to Reduce Label Cost*, 2011.

[Clement *et al.*, 2015] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining*, 7(2):20–48, 2015.

[Davenport *et al.*, 2012] Jodi Davenport, Anna Rafferty, Michael Timms, David Yaron, and Michael Karabinos. ChemVLab+: Evaluating a virtual lab tutor for high school chemistry. In *Proceedings of the 2012 International Conference on Learning Sciences*, 2012.

[Goldman and Kearns, 1995] Sally Goldman and Michael Kearns. On the complexity of teaching. *Journal of Computer and Systems Sciences*, 50(1):20–31, 1995.

[Koedinger *et al.*, 1997] Kenneth Koedinger, John Anderson, William Hadley, and Mary Mark. Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8:30–43, 1997.

[Liu *et al.*, 2017] Weiyang Liu, Bo Dai, Ahmad Humayun, Charlene Tay, Chen Yu, Linda Smith, James Rehg, and Le Song. Iterative machine teaching. In *Proceedings of the 34th International Conference on Machine Learning*, pages 2149–2158, 2017.

[Milli *et al.*, 2017] Smitha Milli, Pieter Abbeel, and Igor Mordatch. Interpretable and Pedagogical Examples. *CoRR*, abs/1711.00694, 2017.

[Mota *et al.*, 2015] Pedro Mota, Francisco S. Melo, and Luísa Coheur. Modeling students self-studies behaviors. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 1521–1528, 2015.

[Nkambou *et al.*, 2010] Roger Nkambou, Riichiro Mizoguchi, and Jacqueline Bourdeau. *Advances in Intelligent Tutoring Systems*. Springer, 2010.

[Polyak, 1977] Boris Polyak. Convergence and rate of convergence of iterative stochastic algorithms. II. The linear case. *Automation and Remote Control*, 38(4):537–542, 1977.

[Shinohara and Miyano, 1991] Ayumi Shinohara and Satoru Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.

[Suh *et al.*, 2016] Jina Suh, Xiaojin Zhu, and Saleema Amershi. The label complexity of mixed-initiative classifier training. *Proceedings of the 33rd International Conference on Machine Learning*, pages 2800–2809, 2016.

[Zhu *et al.*, 2017] Xiaojin Zhu, Ji Liu, and Manuel Lopes. No learner left behind: On the complexity of teaching multiple learners simultaneously. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 3588–3594, 2017.

[Zhu, 2013] Xiaojin Zhu. Machine teaching for Bayesian learners in the exponential family. In *Advances in Neural Information Processing Systems 26*, pages 1905–1913, 2013.

[Zhu, 2015] Xiaojin Zhu. Machine teaching: An inverse problem to machine learning and an approach toward optimal education. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, pages 4083–4087, 2015.

[Zilles *et al.*, 2008] Sandra Zilles, Steffen Lange, Robert Holte, and Martin Zinkevich. Teaching dimensions based on cooperative learning. In *Proceedings of the 21st Annual Conference on Learning Theory*, pages 135–146, 2008.