

# eTutor: Online Learning for Personalized Education

Cem Tekin, Mihaela van der Schaar

Electrical Engineering Department, University of California, Los Angeles

Email: cmtkn@ucla.edu, mihaela@ee.ucla.edu

## Abstract

Given recent advances in information technology and artificial intelligence, web-based education systems have become complementary and, in some cases, viable alternatives to traditional classroom teaching. The popularity of these systems stems from their ability to make education available to a large demographics (see MOOCs). However, existing systems do not take advantage of the personalization which becomes possible when web-based education is offered: they continue to be one-size-fits-all. In this paper, we aim to provide a first systematic method for designing a personalized web-based education system. Personalizing education is challenging: *(i)* students need to be provided personalized teaching and training depending on their contexts (e.g. classes already taken, methods of learning preferred, etc.), *(ii)* for each specific context, the best teaching and training method (e.g type and order of teaching materials to be shown) must be learned, *(iii)* teaching and training should be adapted online, based on the scores/feedback (e.g. tests, quizzes, final exam, likes/dislikes etc.) of the students. Our personalized online system, e-Tutor, is able to address these challenges by learning how to adapt the teaching methodology (in this case what sequence of teaching material to present to a student) to maximize her performance in the final exam, while minimizing the time spent by the students to learn the course (and possibly dropouts). We illustrate the efficiency of the proposed method on a real-world eTutor platform which is used for remedial training for a Digital Signal Processing (DSP) course.

## Index Terms

Online learning, personalized education, eLearning, intelligent tutoring systems.

## I. INTRODUCTION

The last decade has witnessed an explosion in the number of web-based education systems due to the increasing demand in higher-level education [1], limited number of teaching personnel, and advances in

information technology and artificial intelligence. Nowadays, most universities have integrated *Massive Open Online Course* (MOOC) platforms into their education systems such as edX consortium, Coursera or Udacity [2]–[4], to give students the possibility to learn by interacting with a software program instead of human teachers. Several advantages of these systems over traditional classroom teaching are: (i) they provide flexibility to the student in choosing what to learn and when to learn, (ii) they do not require the presence of an interactive human teacher, (iii) there are no limitations in terms of the number of students who can take the course. However, there are significant limitations of currently available online teaching platforms. Since courses are taken online, there is no interaction between the students and the teacher as in a classroom setting. This makes it very difficult to meet the personalized needs of each student, which may arise due to the differences between qualifications, learning methods and cognitive skills of the students. It is observed that if the personalization of teaching content is not carried out efficiently, high drop-outs will occur [1]. For instance, the students that are very familiar with the topic may drop-out if the teaching material is not challenging enough, while the students that are new to the topic may get overstrained if the teaching material is hard.

Due to these challenges, a new web-based education system that personalizes education by learning online the needs of the students based on their contexts, and adapting the teaching material based on the feedback signals received from the student (answers to questions, quizzes, etc.) is required. For this purpose we develop the *eTutor* (illustrated in Fig. 1), which is an online web-based education system, that learns how to teach a course, a concept or remedial materials to a student with a specific context in the most efficient way. Basically, for the current student, eTutor learns from its past interactions with students with similar contexts, the sequence of teaching materials that are shown to these students, and the response of these students to the teaching materials including the final exam scores, how to teach the course in the most effective way. This is done by defining a *teaching effectiveness metric*, referred to as the *regret*, that is a function of the final exam score and time cost of teaching to the student, and then designing a learning algorithm that learns to optimize this metric. This tradeoff between learning (exploring) and optimizing (exploiting) is captured by the eTutor in the most efficient way, i.e., the average exam score of the students converge to the average exam score that could be achieved by the best teaching strategy. We illustrate the efficiency of the proposed system in a real-world experiment carried out on students in a DSP class.

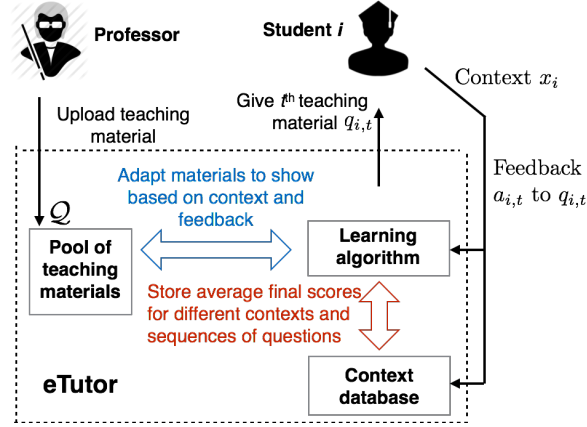


Fig. 1. eTutor, student and professor interaction.

### A. Related Work

Although web-based education systems have recently become popular, there is no consensus or standards on how to design an optimal web-based education system. A detailed comparison of our work with the related work in web-based education is given in Table I. Most of the recent works focus on the subfield of MOOCs, which are online courses with very large number of students [1]–[4]. Among these, several works examine students’ interaction with commercially available MOOC systems such as Coursera [2], [3] and edX [4].

Apart from these, two approaches exist in designing web-based education systems: adaptive education systems and intelligent tutoring systems. In an adaptive education system [5], [6], the teaching materials that are shown to each student are adapted based on the context of the student, but not based on the feedback the student provides during the course. This adaptation is based on numerous contexts including the student’s learning style, her knowledge, background, origin, grades, previously taken courses etc. In contrast, in an intelligent tutoring system adaptation is done based on the response of the student to the given teaching material [7]–[13], without taking into account contexts. Our work combines both ideas by adapting the sequence of teaching materials that is presented to a student based on both the context and the feedback of the student. However, our techniques are very different from both lines of research. Our goal is to learn the optimal way to teach a course *in a way that is most effective for each student*. To learn effectively, our method utilizes the past knowledge gained about the efficacy of the material from students with similar contexts who have taken the course before. This is different from [7]–[13], which only take into account the current student’s response to the previously shown teaching material.

Method	Context-based learning	Feedback-based learning	Learns from final exam	Regret bound
[5], [6]	Yes	No	No	No
[7]–[13]	No	Yes	No	No
Our work	Yes	Yes	Yes	Yes

TABLE I  
COMPARISON WITH RELATED WORK.

## II. FORMALISM, ALGORITHM AND ANALYSIS

In this section we mathematically formalize the online teaching/tutoring problem, define a benchmark tutor (i.e. the "ideal" tutor) and propose an online learning algorithm for the eTutor which converges in performance to the benchmark tutor that knows the optimal sequence of teaching materials to show for each student.

### A. Problem Definition

Consider a set of students participating in an online education system and a concept that should be learned by the students. The comprehension of the concept will be tested via a final exam (test). We assume that the students arrive sequentially over time and use index  $i$  to denote the  $i$ th student. Additionally, we assume that when a student first interacts with the online education system, she needs to answer a set of questions, which will form the *context* of the student. Context may include information about the student such as age, grades, whether she prefers visual or written instructions, etc. Denote the finite set of all possible contexts by  $\mathcal{X}$  and an element of  $\mathcal{X}$  by  $x$ . The concept will be taught by presenting a set of teaching materials (written or visual) to the student and asking a set of questions about these materials and providing their answers. Let  $\mathcal{Q}$  be the set of teaching materials (consists of text/images to learn from and questions) that can be given to the student. The number of elements of  $\mathcal{Q}$  is denoted by  $Q$ .

The materials that are shown to a student are chosen in an online way based on the context of the student, previous materials that are shown to the student, the student's response to shown questions (whether the answer is correct or not) and all the previous knowledge obtained from past students with contexts, responses and scores similar to the current student. It is also important to learn in which order the materials should be shown, since learning from one material may require knowledge of a concept which can be learned by understanding another material.

For each student  $i$ , we consider a discrete time model  $t = 1, 2, \dots, T_i$ , where time  $t$  denotes the sequence of events related to the  $t$ th material that is shown to the student.  $T_i$  denotes the number of teaching materials shown to student  $i$  before the final exam is given (depends on student's feedback).

Clearly,  $T_i \leq Q$ . The  $t$ th teaching material shown to student  $i$  is denoted by  $q_{i,t}$ . Let  $\mathbf{q}_i := (q_{i,1}, \dots, q_{i,T_i})$ , and  $\mathbf{q}_i[t] := (q_{i,1}, \dots, q_{i,t})$ .

We denote student  $i$ 's response to  $q_{i,t}$  by  $a_{i,t} \in \mathcal{A}$ , where  $\mathcal{A}$  is the set of possible feedbacks that the student can provide to a teaching material. We assume that  $\mathcal{A}$  is finite. For instance, an example can be the case when  $\mathcal{A} := \{-1, 0, 1\}$ . If the student does not provide any feedback on the teaching material, we have  $a_{i,t} = 0$ ; when the teaching material is a (multiple-choice) question,  $a_{i,t} = 1$  denotes a correct answer and  $a_{i,t} = -1$  denotes a wrong answer. Let  $\mathbf{a}_i := (a_{i,1}, \dots, a_{i,T_i})$  and  $\mathbf{a}_i[t] := (a_{i,1}, \dots, a_{i,t})$ ,  $t \leq T_i$ . In addition, let  $a_{i,0} := 0$ , which indicates that no feedback is available prior to 1st teaching material.

Let  $\mathcal{S}$  denote the set of all sequences of teaching materials that can be shown.<sup>1</sup> We have

$$|\mathcal{S}| = \sum_{t=1}^Q \binom{Q}{t} t! = Q! \sum_{t=1}^Q \frac{1}{(Q-t)!} = Q! \left( \frac{e\Gamma(Q+1, 1)}{\Gamma(Q+1)} - 1 \right) \geq Q!,$$

where  $\Gamma(y)$  is the gamma function and  $\Gamma(x, y)$  is the incomplete gamma function.

For a sequence of materials  $s \in \mathcal{S}$ , let  $\mathcal{A}(s)$  be the set of sequences of feedbacks a student can provide. The expected final exam score for a student with context  $x$ , sequence of questions  $s \in \mathcal{S}$  and sequence of feedbacks  $\mathbf{a} \in \mathcal{A}(s)$  is denoted by  $r_{x,s,\mathbf{a}}$ . We assume that the final exam score of a student with context  $x$ , the sequence of teaching materials  $s$  and the sequence of feedbacks  $\mathbf{a}$  is randomly drawn from a  $F_{x,s,\mathbf{a}}$  with expected value  $r_{x,s,\mathbf{a}}$ . Both  $F_{x,s,\mathbf{a}}$  and  $r_{x,s,\mathbf{a}}$  are unknown.

## B. The Benchmark Tutor

Due to the enormous number of possible sequences of teaching materials, it is not possible to learn the best sequence of teaching materials by trying all of them for different students. In this section we define a benchmark tutor, whose teaching strategy can be learned very fast. We call it the *best-first* (BF) benchmark. Due to limited space its pseudocode is given in Fig. 2, however, we describe it in detail below. In order to explain this benchmark, we require a few more notations.

Given a sequence  $s$  of teaching materials, let  $\mathcal{Q}_s$  be the set of remaining teaching materials that can be given to the student. Let  $\mathcal{S}[t] \subset \mathcal{S}$  be the set of sequences that consists of  $t$  teaching materials followed by the final exam. In order to explicitly state the number of teaching materials in a sequence of teaching materials, we will use the notation  $s[t]$  to denote an element of  $\mathcal{S}[t]$ . We will also use  $\mathbf{a}_{s[t]}[t']$  to denote

<sup>1</sup>In practice, it is possible to give  $\mathcal{S}$  as an input in addition to  $\mathcal{Q}$ . For instance, some sequences which are classified by the professor as unreasonable can be discarded, significantly reducing the size of  $\mathcal{S}$ .

```

1: Receive student context  $x$ 
2: Show  $q_{x,1}^* = \arg \max_{q \in \mathcal{Q}} y_{x,q,0}$ 
3: Receive  $a_1^*$ .
4: while  $1 < t \leq Q$  do
5:   if
      $r_{x,s^*[t-1],a^*[t-1]} \geq \max_{q \in \mathcal{Q}_{s^*[t-1]}} y_{x,s^*[t-1],a^*[t-1]} - c$  then
6:     Give the final exam. //BREAK
     else
8:        $q_{x,t}^* = \arg \max_{q \in \mathcal{Q}_{s^*[t-1]}} y_{x,s^*[t-1],a^*[t-1]}$ 
     end if
      $t = t + 1$ 
11: end while

```

Fig. 2. Pseudocode for BF.

the student's feedback to the first  $t'$  teaching materials in  $s[t]$ . Let

$$y_{x,s[t],a_{s[t]}[t-1]} := E_{a_t}[r_{x,s[t],(a_{s[t]}[t-1],a_t)}],$$

be the *ex-ante* final exam score of a student with context  $x$  which is given teaching materials  $s[t]$  and provided feedback to all of them except the last teaching material.

The BF benchmark incrementally selects the next teaching material to show based on the student's feedback about the previous teaching materials. The first teaching material it shows is  $q_{x,1}^* = \arg \max_{q \in \mathcal{Q}} y_{x,q,0}$ . Let  $\mathbf{q}_x^* = (q_{x,1}^*, q_{x,2}^*, \dots, q_{x,T}^*)$  be the sequence of teaching materials shown by the BF benchmark to a student with context  $x$ , where  $T$  is the random total number of materials shown to the student, which depends on the feedback of the student to the shown materials. In general, the  $t$ th teaching material to show, i.e.,  $q_{x,t}^*$ , depends on both  $\mathbf{q}_x^*[t-1]$  and  $a_{\mathbf{q}_x^*[t-1]}[t-1]$ . We assume that the following property holds for the expected final exam score given the sequence of materials shown by the BF benchmark and the feedback obtained from these shown materials:

*Assumption 1:* Consider a context  $x \in \mathcal{X}$  and any two sequences  $(\mathbf{q}^*, \mathbf{a})$  and  $(\mathbf{q}^*, \mathbf{a}')$ , where  $\mathbf{q}^*$  is the set of teaching materials shown by the BF benchmark and  $\mathbf{a}$  and  $\mathbf{a}'$  are two feedback sequences that are associated with this set of teaching materials. Then, we have

$$\arg \max_{q \in \mathcal{Q}_{\mathbf{q}^*[t]}} y_{x,(\mathbf{q}^*[t],q),\mathbf{a}[t]} = \arg \max_{q \in \mathcal{Q}_{\mathbf{q}^*[t]}} y_{x,(\mathbf{q}^*[t],q),\mathbf{a}'[t]}.$$

We have  $q_{x,t}^* = \arg \max_{q \in \mathcal{Q}_{\mathbf{q}_x^*[t-1]}} r_{x,\mathbf{q}_x^*[t-1],a_{\mathbf{q}_x^*[t-1]}^*}$ . For any  $t$ , if  $r_{x,\mathbf{q}_x^*[t],a^*[t]} \geq y_{x,(\mathbf{q}_x^*[t],q),a^*[t]} - c$  for all  $q \in \mathcal{Q}(\mathbf{q}_x^*[t])$ , then the BF benchmark will give the final exam after the  $t$ th teaching material. Here  $c > 0$  is the *teaching cost* of showing one more material to the student, which is the cost related to the time it

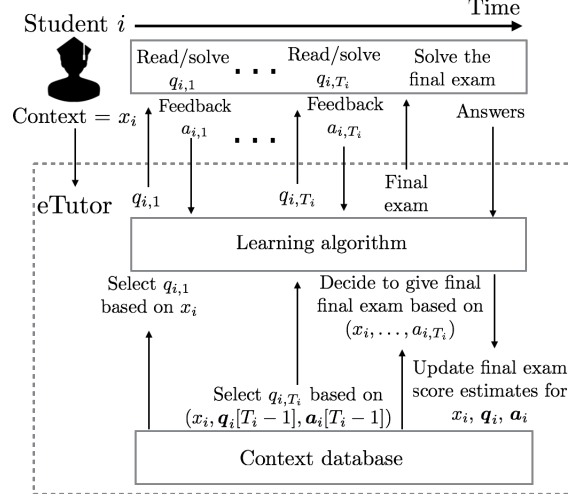


Fig. 3. Operation of the eTutor.

takes for the student to complete the teaching material. The average final exam score minus the teaching cost achieved by following the BF benchmark for the first  $n$  students is equal to

$$RW_{BF}(n) = \sum_{i=1}^n \frac{E[Y_{x_i, Q_i^*, A_i^*} - c|Q_i^*|]}{n},$$

where  $Y_{x_i, Q_i^*, A_i^*}$  is the random variable that represents the final exam score of student  $i$ , where  $Q_i^*$  is the random variable that represents the sequence of teaching materials given to student  $i$  by the BF benchmark, and  $A_i^*$  is the random variable that represents the sequence of feedbacks provided by student  $i$  to the teaching materials  $Q_i^*$ . The BF benchmark is an *oracle* policy because we assume that nothing is known about the expected exam scores a priori. Any learning algorithm  $\alpha$  which selects a sequence of teaching materials  $Q_i^\alpha$  based on the sequence of feedbacks  $A_i^\alpha$  has a average regret with respect to the BF benchmark which is given by

$$R(n) = RW_{BF}(n) - \sum_{i=1}^n \frac{E[Y_{x_i, Q_i^\alpha, A_i^\alpha} - |Q_i^\alpha|]}{n}. \quad (1)$$

The next example illustrates that when the cost of showing new materials, i.e.,  $c$ , is high, BF benchmark is better than the best fixed sequence.

*Example 1:* Consider  $\mathcal{Q} = \{a, b\}$ ,  $\mathcal{A} = \{0, 1\}$  and  $\mathcal{X} = \{x\}$ . Assume that the expected rewards are given as follows:  $r_{x,a,0} = 0$ ,  $r_{x,b,0} = 0$ ,  $r_{x,a,1} = 12$ ,  $r_{x,b,1} = 6$ ,  $r_{x,(a,b),(1,1)} = 13$ ,  $r_{x,(a,b),(1,0)} = 12$ ,  $r_{x,(a,b),(0,1)} = 10$ ,  $r_{x,(a,b),(0,0)} = 9$ . Let  $P_x(a|q)$  denote the probability that feedback sequence  $a$  is given by a student with context  $x$  as a response to the sequence  $q$ . Assume that we have  $P_x(1|a) = 0.5$ ,  $P_x(0|a) = 0.5$ ,  $P_x((0,0)|(a,b)) = 0.3$ ,  $P_x((0,1)|(a,b)) = 0.2$ ,  $P_x((1,1)|(a,b)) = 0.4$ ,  $P_x((1,0)|(a,b)) = 0.1$ .

The BF benchmark shows  $a$  as the first teaching material. Then if feedback is 0 it also shows  $b$  before giving the final exam. Else, it gives the final exam just after showing  $a$ . Hence the expected reward of the BF is

$$RW_{\text{BF}}(n) = 0.5 \times 12 + 0.3 \times (9 - c) + 0.2 \times (10 - c) = 10.7 - 0.5c,$$

where  $c$  is the cost of showing the second material. The best fixed sequence always shows  $a$  first,  $b$  second, and then gives the final exam. Its expected reward is equal to

$$0.3 \times 9 + 0.4 \times 12 + 0.2 \times 10 + 0.1 \times 11 - c = 11 - c.$$

Clearly, the BF benchmark is better than the best fixed sequence given that  $c > 3/5$ .

### C. *eTutor*

In this section we propose *eTutor* (pseudocode given in Fig. 4), which learns the optimal sequence of teaching materials to show based on the student's context and feedback about the previously shown teaching materials (as shown in Fig. 3). In order to minimize the regret given in (1), *eTutor* balances exploration and exploitation when selecting the teaching materials to show to the student. Consider a student  $i$  and the  $t$ th teaching material shown to that student. *eTutor* keeps the following sample mean reward estimates: (i)  $\hat{r}_{x,t,q,a}(i)$  which is the estimated final exam score for students with context  $x$  that took the course before student  $i$  who are given the final exam right after material  $q$  is given as the  $t$ th material and feedback  $a$  is observed, (ii)  $\hat{y}_{x,a,t,q}(i)$  which is the estimated final exam score for students with context  $x$  that took the course before student  $i$  who are given the final exam right after material  $q$  is given as the  $t$ th material after observing feedback  $a$  for the  $t - 1$ th material. In addition to these, *eTutor* keeps the following counters: (i)  $T_{x,t,q,a}(i)$  which counts the number of times material  $q$  is shown as the  $t$ th material and feedback  $a$  is obtained for students with context  $x$  that took the course before student  $i$ , (ii)  $T_{x,a,t,q}(i)$  which counts the number of times material  $q$  is shown as the  $t$ th material after feedback  $a$  is obtained from the previously shown material for students with context  $x$  that took the course before student  $i$ .

Next, we explain how exploration and exploitation is performed. Consider the event that *eTutor* asks question  $q_{i,t} = q$  and receives feedback  $a_{i,t} = a$ . It first checks if  $T_{x,t,q,a}(i) < D \log(i/\delta)$ , where  $D > 0$  and  $\delta > 0$  are constants that are input parameters of *eTutor*. If this holds, then *eTutor* explores by giving the final exam and obtaining the final score  $X(i)$ , by which it updates  $\hat{r}_{x,t,q,a}(i + 1) = (\hat{r}_{x,t,q,a}(i + 1) + X(i)) / (T_{x,t,q,a}(i) + 1)$ . Else if  $T_{x,t,q,a}(i) \geq D \log(i/\delta)$ , *eTutor* checks if there are any



```

1: Input  $D > 0, \delta > 0$ .
2: Initialize:  $\hat{r}_{x,t,q,a} = 0, \hat{y}_{x,a,t,q} = 0, T_{x,t,q,a} = 0, T_{x,a,t,q} = 0, \forall x \in \mathcal{X}, a \in \{-1, 0, 1\}, q \in \mathcal{Q}, t = 1, \dots, Q. a_{i,0} = 0, \mathbf{q}_i[0] = \emptyset, \forall i = 1, 2, \dots$ 
3: while  $i \geq 1$  do
4:   Receive student context  $x = x_i$ 
5:    $\mathcal{U}_1 = \{q \in \mathcal{Q} : T_{x,0,1,q} < D \log(i/\delta)\}$ 
6:   if  $\mathcal{U}_1 \neq \emptyset$  then
7:     Give  $q_{i,1}$  randomly selected from  $\mathcal{U}_1$ , get  $a_{i,1}$ .
8:     Give the final exam, get the score  $X(i), t^* = 1, //\text{BREAK}$ 
9:   else
10:    Give  $q_{i,1} = \arg \max_{q \in \mathcal{Q}} \hat{y}_{x,0,1,q}$ , get  $a_{i,1}$ .
11:   end if
12:    $t = 2$ 
13:   while  $2 \leq t \leq Q$  do
14:      $\mathcal{U}_t = \{q \in \mathcal{Q}_{\mathbf{q}_i[t-1]} : T_{x,a_{i,t-1},t,q} < D \log(i/\delta)\}$ 
15:     if  $T_{x,t-1,q_{i,t-1},a_{i,t-1}} < D \log i$  then
16:       Give the final exam, get the score  $X(i), t^* = t - 1, //\text{BREAK}$ 
17:     else if  $\mathcal{U}_t \neq \emptyset$  then
18:       Show  $q_{i,t}$  randomly selected from  $\mathcal{U}_t$  and get the feedback  $a_{i,t}$ .
19:       Give the final exam, get the score  $X(i), t^* = t - 1, //\text{BREAK}$ 
20:     else
21:       if  $\hat{r}_{x,t-1,q_{i,t-1},a_{i,t-1}} \geq \hat{y}_{x,a_{i,t-1},t,q'} - c, \forall q' \in \mathcal{Q}_{\mathbf{q}_i[t-1]}$  then
22:         Give the final exam, get the score  $X(i), t^* = t - 1, //\text{BREAK}$ 
23:       else
24:         Show  $q_{i,t} = \arg \max_{q \in \mathcal{Q}_{\mathbf{q}_i[t-1]}} \hat{y}_{x,a_{i,t-1},t,q'}$  and get the feedback  $a_{i,t}$ 
25:       end if
26:     end if
27:      $t = t + 1$ 
28:   end while
29:   Update  $\hat{r}_{x,t^*,q_{i,t^*},a_{i,t^*}}, \hat{y}_{x,a_{i,t^*}-1,t^*,q_{i,t^*}}$  using  $X(i)$  (sample mean update).
30:    $T_{x,t^*,q_{i,t^*},a_{i,t^*}} ++, T_{x,a_{i,t^*}-1,t^*,q_{i,t^*}} ++$ .
31:    $i = i + 1$ 
32: end while

```

Fig. 4. Pseudocode for eTutor.

questions  $q' \in \mathcal{Q}_{\mathbf{q}_i[t]}$  for which  $T_{x,a_{i,t},t+1,q'}(i) < D \log(i/\delta)$ . If there are such questions, then eTutor explores one of them randomly by showing that material to the student, obtaining the feedback, giving the final exam, and obtaining the final exam score. The obtained final exam score  $X(i)$  is used for both updating  $\hat{r}_{x,t+1,q',a_{i,t+1}}(i+1)$  and  $\hat{y}_{x,a_{i,t},t+1,q'}(i+1)$ . If none of the above events happen, then eTutor exploits at  $t$ . To do this it first checks if  $\hat{r}_{x,t,q,a_{i,t}}(i) \geq \hat{y}_{x,a_{i,t},t+1,q'}(i) - c$ , for all  $q' \in \mathcal{Q}_{\mathbf{q}_i[t]}$ . If this is the case, it means that showing one more teaching material does not increase the final exam score enough to compensate for the *teaching cost* of showing one more material. Hence, eTutor gives the final exam after its  $t$ th material. If the opposite happens, then it means that showing one more material can

improve final exam score sufficiently enough for it to compensate the cost of teaching. Hence, eTutor will show one more teaching material to the student which is  $q_{i,t+1} = \arg \max_{q' \in \mathcal{Q}_{q_i[t]}} \hat{y}_{x,a_{i,t},t+1,q'}(i)$ . The next decision to take will be based on the student's feedback to  $q_{i,t+1}$  which is  $a_{i,t+1}$ . This goes on until eTutor gives the final exam, which will eventually happen since  $\mathcal{Q}$  is finite.

#### D. Regret and Confidence Bounds For eTutor

The regret of eTutor can be written as the sum of two separate regret terms: regret for the experimental materials shown to students when eTutor explores, i.e.,  $R_e(n)$ , and regret for students that eTutor exploits, i.e.,  $R_s(n)$ . Hence we can write  $R(n) = E[R_e(n)] + E[R_s(n)]$ .

For a sequence of numbers  $\{r\}_{r \in \mathcal{R}}$ , let  $\min 2(\{r\}_{r \in \mathcal{R}})$  be the difference between the highest and the second highest numbers. Consider any sequence of materials  $\mathbf{q}_x^*[t] \in \mathcal{S}[t]$  and feedback  $\mathbf{a}[t] \in \mathcal{A}(\mathbf{q}_x^*[t])$ , where  $\mathbf{q}_x^*[t]$  is the sequence of materials shown by the BF benchmark to a student with context  $x$ . Let

$$\Delta_{\min,1} := \min 2(\{y_{x,q,0}\}_{q \in \mathcal{Q}}),$$

and

$$\Delta_{\min,t} := \min 2(r_{x,\mathbf{q}_x^*[t]}, \mathbf{a}[t], \{y_{x,(\mathbf{q}_x^*[t],q),\mathbf{a}[t]}\}_{q \in \mathcal{Q}_{\mathbf{q}_x^*[t]}}),$$

for  $1 < t < Q$ . Let  $\Delta_{\min} := \min_{t=1,\dots,Q-1} \Delta_{\min,t}$ . Given that the constant  $D$  that is input to eTutor is such that  $D \geq 4/\Delta_{\min}^2$ , where  $\Delta_{\min} = (\min_{x \in \mathcal{X}, s \in \mathcal{S}, \mathbf{a} \in \mathcal{A}(s)} r_{x,s,\mathbf{a}})^2$ , Assuming that the maximum final exam score is equal to 1, we have the following bounds on the regret.

*Theorem 1:* Setting the parameters of eTutor as  $D = 4/\Delta_{\min}^2$  and  $\delta = \sqrt{\epsilon}/(Q\sqrt{2\beta})$ , where  $\beta = \sum_{t=1}^{\infty} 1/t^2$ , we have the following bounds on the regret of eTutor. The regret of eTutor for the first  $n$  students is bounded as follows:

- (i)  $R_e(n) \leq 2|\mathcal{X}||\mathcal{A}|Q^2D \log(n/\delta)/n$  with probability 1.
- (ii)  $R_s(n) = 0$  with probability at least  $1 - \epsilon$ .
- (iii)  $R(n) \leq 2|\mathcal{X}||\mathcal{A}|Q^2D \log(n/\delta)/n + \epsilon$ .

*Proof:* Our proof involves showing that when the eTutor estimates the final exam scores for the sequences of materials it gives to the students such that they are within  $\Delta_{\min}/2$  of the true final exam score, then it will always show the same set of teaching materials as the BF benchmark does.

To proceed, we define the following sets of students. Let  $E_1(n)$  be the set of students in  $\{1, \dots, n\}$  for which eTutor explores the teaching material in the first slot, i.e.,  $i \in \{1, \dots, n\}$  for which  $T_{x_i,0,1,q}(i) < D \log(i/\delta)$  for some  $q \in \mathcal{Q}$  such that after the material is shown, the final exam is given. Let  $E_t(n)$ ,

$1 < t \leq Q$  be the set of students in  $\{1, \dots, n\}$  for which eTutor explores in the  $t$ th slot, i.e., the set of time slots for which eTutor exploited up to the  $t - 1$ th slot and  $\bar{r}_{x_i, t-1, q_{i, t-1}, a_{i, t-1}}(i) \leq D \log(i/\delta)$  or  $\hat{y}_{x_i, a_{i, t-1}, t, q} \leq D \log(i/\delta)$  for some  $q \in \mathcal{Q}_{q_i[t-1]}$  such that the final exam is given either after the  $t - 1$ th slot or the  $t$ th slot depending on which teaching material is under-explored. Let  $\tau_1(n)$  be the set of students in  $\{1, \dots, n\}$  for which eTutor exploits for the first slot, i.e.,  $T_{x_i, 0, 1, q} \geq D \log(i/\delta)$  for all  $q \in \mathcal{Q}$ . Let  $\tau_t(n)$  be the set of students in  $\{1, \dots, n\}$  for which eTutor exploits for the  $t$ th slot, i.e.,  $\bar{r}_{x_i, t-1, q_{i, t-1}, a_{i, t-1}}(i) \geq D \log(i/\delta)$  and  $\hat{y}_{x_i, a_{i, t-1}, t, q} \geq D \log(i/\delta)$  for all  $q \in \mathcal{Q}_{q_i[t-1]}$  such that eTutor has not given the final exam before slot  $t - 1$ . Let  $Z_t(n) := \tau_t(n) - \tau_{t+1}(n) - E_{t+1}(n)$ ,  $1 \leq t < Q$  denote the set of students in  $\{1, \dots, n\}$  for which eTutor stops and gives the final exam after the  $t - 1$ th teaching material is shown at times when it exploits. Let  $Z_Q(n) := \tau_Q(n)$ . The set of all students for which eTutor explores until the  $n$ th student is equal to  $E(n) := \bigcup_{t=1}^Q E_t(n)$ , where  $E_t(n) \cap E_{t'}(n) = \emptyset$  for  $t \neq t'$ . The set of all students for which eTutor exploits until the  $n$ th student is equal to  $Z(n) := \bigcup_{t=1}^Q Z_t(n)$ , where  $Z_t(n) \cap Z_{t'}(n) = \emptyset$ . We also have  $Z(n) := \tau_1(n)$ ,  $\tau_t(n) = \tau_{t+1}(n) \cup E_{t+1}(n) \cup Z_t(n)$  for  $1 \leq t < Q$ .

In the following we will bound  $R_s(n)$ . Next, we define the events which correspond to the case that the estimated final exam scores for the sequences followed by the BF benchmark are within  $\Delta_{\min}/2$  of the expected final exam scores. Let

$$\text{Perf}_1(n) := \{|\hat{y}_{x_i, 0, 1, q} - y_{x_i, q, 0}| < \Delta_{\min}/2, \forall q \in \mathcal{Q}, \forall i \in \tau_1(n)\},$$

and

$$\begin{aligned} \text{Perf}_t(n) := & \left\{ |\hat{r}_{x_i, t-1, q_{i, t-1}, a_{i, t-1}} - r_{x_i, \mathbf{q}_{x_i}^*[t-1], \mathbf{a}_{\mathbf{q}_{x_i}^*[t-1]}[t-1]}| < \Delta_{\min}/2, \right. \\ & \left. |\hat{y}_{x_i, a_{i, t-1}, t, q} - y_{x_i, (\mathbf{q}_{x_i}^*[t-1], q), \mathbf{a}_{\mathbf{q}_{x_i}^*[t-1]}[t-1]}| < \Delta_{\min}/2, \forall q \in \mathcal{Q}_{\mathbf{q}_{x_i}^*[t-1]}, \forall i \in \tau_t(n) \right\} \end{aligned}$$

Let

$$\text{Perf}(n) = \bigcap_{t=1}^Q \text{Perf}_t(n).$$

On event  $\text{Perf}(n)$ , eTutor selects teaching materials for the students in the same way as BF benchmark does. Hence the contribution to the regret given in (1) on event  $\text{Perf}(n)$  is zero.

Next, we lower bound the probability of event  $\text{Perf}(n)$ . Using the chain rule we can write

$$\begin{aligned} \text{P}(\text{Perf}(n)) &= \text{P}(\text{Perf}_Q(n), \text{Perf}_{Q-1}(n), \dots, \text{Perf}_1(n)) \\ &= \text{P}(\text{Perf}_Q(n) | \text{Perf}_{Q-1}(n), \dots, \text{Perf}_1(n)) \times \text{P}(\text{Perf}_{Q-1}(n) | \text{Perf}_{Q-2}(n), \dots, \text{Perf}_1(n)) \\ &\quad \times \dots \times \text{P}(\text{Perf}_2(n) | \text{Perf}_1(n)) \times \text{P}(\text{Perf}_1(n)). \end{aligned} \quad (2)$$

For an event  $E$ , let  $E^c$  denote its complement. Note that we have

$$\begin{aligned} \text{P}(\text{Perf}_1(n)^c) &\leq \sum_{i \in \tau_1(n)} \sum_{q \in \mathcal{Q}} \text{P}(|\hat{y}_{x_i,0,1,q} - y_{x_i,q,0}| < \Delta_{\min}/2) \\ &\leq \sum_{i \in \tau_1(n)} 2Q \exp(-2D \log(i/\delta) \Delta_{\min}^2/4) \end{aligned} \quad (3)$$

$$\leq \sum_{i \in \tau_1(n)} 2Q \delta^2 / i^2 \leq 2Q \beta \delta^2, \quad (4)$$

since  $D \geq 4/\Delta_{\min}^2$  and  $\beta = \sum_{i=1}^{\infty} 1/i^2$ . Hence, we have

$$\text{P}(\text{Perf}_1(n)) \geq 1 - 2Q \beta \delta^2.$$

On event  $\text{Perf}_1(n)$ , it is always the case that the first teaching material that is shown is chosen according to the BF benchmark, independent of whether the eTutor explores or exploits the second slot. Hence given  $\text{Perf}_1(n)$ , the sample mean estimates that are related to  $\text{Perf}_2(n)$  are always sampled from the distribution in which the first teaching material is shown according to the BF benchmark. Because of this, we have

$$\text{P}(\text{Perf}_2(n) | \text{Perf}_1(n)) \geq 1 - 2Q \beta \delta^2.$$

Similarly, it can be shown that

$$\text{P}(\text{Perf}_t(n) | \text{Perf}_{t-1}(n), \dots, \text{Perf}_1(n)) \geq 1 - 2Q \beta \delta^2.$$

Combining all of this and using (4) we get

$$\begin{aligned} \text{P}(\text{Perf}(n)) &\geq (1 - 2Q \beta \delta^2)^Q \\ &\geq 1 - 2Q^2 \beta \delta^2. \\ &= 1 - \epsilon, \end{aligned} \quad (5)$$

since  $\delta = \sqrt{\epsilon}/(Q\sqrt{2\beta})$ .

Next we bound  $R_e(n)$ . From the definition of  $E_t(n)$ ,  $t = 1, \dots, Q$ , we know that  $|E_1(n)| \leq |\mathcal{X}||\mathcal{A}|QD \log(n/\delta)$ . Similarly, for  $E_t(n)$ ,  $t = 2, \dots, Q$ , we have  $|E_t(n)| \leq 2|\mathcal{X}||\mathcal{A}|QD \log(n/\delta)$ . Hence, we have  $|E(n)| \leq 2|\mathcal{X}||\mathcal{A}|Q^2D \log(n/\delta)$ . Since the worst-case reward loss due to showing a suboptimal set of teaching materials to a student is at most 1, we have

$$R_e(n) \leq 2|\mathcal{X}||\mathcal{A}|Q^2D \log(n/\delta)/n.$$

Finally, the regret bound on  $R(n)$  holds by taking the expectation. ■

Theorem 1 implies that the average final exam score of students tutored by eTutor converges to the average final exam score of students tutored by BL (with probability at least  $1 - \epsilon$ ) which knows the expected final exam scores, and hence, how students learn for each sequence of teaching materials perfectly. Moreover, the regret gives the convergence rate, and since it decreases with  $\log n/n$ , eTutor converges very fast.

### III. ILLUSTRATIVE RESULTS

We deployed our eTutor system for students who have already studied digital signal processing (DSP) one or more years ago, and the goal of this implementation of the eTutor is to have them refresh the material about *discrete Fourier transform* (DFT) in the minimum amount of time. Student contexts belong to  $\mathcal{X} = \{0, 1\}$ , where for a student  $i$ ,  $x_i = 0$  implies that she is not confident about her knowledge of DFT, and  $x_i = 1$  implies that she is confident about her knowledge of DFT.  $\mathcal{Q}$  contains three (*remedial*) materials: one text that describes DFT and two questions that refreshes DFT knowledge. If a question is shown to the student and if the student's answer is incorrect, then the correct answer is shown along with an explanation. For each  $q \in \mathcal{Q}$ , we set the cost to be  $c_q = 0.04 \times \theta_q$ , where  $\theta_q$  (in minutes) is the average time it takes for a student to complete material  $q$ . The value of  $\theta_q$  is estimated and updated based on the responses of the students. Performance of the students after taking the remedial materials are tested by the same final exam.

We compare the performance of eTutor with a *random rule* (RR) that randomly selects the materials to show and a *fixed rule* (FR) that shows all materials (text first, easy question second, hard question third). The average final score achieved by these algorithms for  $n = 100$  and  $n = 500$  students are shown in Table II. From this table we see that eTutor achieves 15.7% and 1.1% improvement in the average final score for  $n = 500$  compared to RR and FR, respectively. The improvement compared to FR is small

# of students	$n = 100$	$n = 500$
eTutor	(66.4, 8.7)	(75.8, 8.5)
RR	(62.4, 10.2)	(62.5, 10.2)
FR	(75.5, 17.0)	(75.0, 17.0)

TABLE II

COMPARISON OF eTUTOR WITH RR AND FR: FOR EACH ENTRY  $(x, y)$ ,  $x$  DENOTES THE AVERAGE FINAL SCORE (MAXIMUM = 100) AND  $y$  DENOTES THE TIME SPENT IN MINUTES TAKING THE COURSE.

because FR shows all the materials to every student. It is observed that the average final score of eTutor increases with  $n$ , which is expected since eTutor learns the best set of materials to show for each context as more students take the course. In contrast, RR and FR are non-adaptive, hence their average final exam scores do not improve as more students take the course. For  $n = 500$ , the average time spent by each student taking the course is 8.5 minutes for eTutor which is 16.7% and 50% less than the average time it takes for the same set of students by RR and FR, respectively. eTutor achieves significant savings in time by showing the best materials to each student based on her context instead of showing everything to every student.

#### IV. CONCLUSION

In this paper, we proposed a novel online education system called eTutor. While in this paper, eTutor was used to learn the best sequence of materials to show a specific student, eTutor can also be easily adapted to learn the best teaching methodology such as what types of materials/examples to show (visual or not), what style of teaching to use etc.

#### REFERENCES

- [1] C. Brinton and M. Chiang, "Social learning networks: A brief survey," in *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, March 2014, pp. 1–6.
- [2] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, "Engaging with massive online courses," in *23rd International World Wide Web Conference*, 2014.
- [3] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: analyzing learner subpopulations in massive open online courses," in *Third Conference on Learning Analytics and Knowledge*, 2013, pp. 170–179.
- [4] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. T. Seaton, "Studying learning in the worldwide classroom research into edx's first mooc," *Research and Practice in Assessment*, pp. 13–25, 2013.
- [5] S. Guven, "Mltutor: a web-based educational adaptive hypertext system," in *Lost in the Web - Navigation on the Internet (Ref. No. 1999/169)*, *IEE Colloquium*, 1999, pp. 4/1–4/3.
- [6] N. Henze and W. Nejdl, "Adaptation in open corpus hypermedia," *International Journal of Artificial Intelligence in Education*, pp. 325–350, 2001.

- [7] A. Mitrovic, "An intelligent sql tutor on the web," *International Journal of Artificial Intelligence in Education*, pp. 171–195, 2003.
- [8] T. Heift and D. Nicholson, "Web delivery of adaptive and interactive language tutoring," *International Journal of Artificial Intelligence in Education*, pp. 310–324, 2001.
- [9] K. Forbes-Riley, D. Litman, and M. Rotaru, "Responding to student uncertainty during computer tutoring: A preliminary evaluation," in *Proceedings 9th International Conference on Intelligent Tutoring Systems (ITS)*, 2008.
- [10] S. Schiaffino, P. Garcia, and A. Amandi, "eteacher: Providing personalized assistance to e-learning students," *Computers and Education*, vol. 51, no. 4, pp. 1744 – 1754, 2008.
- [11] A. Zouhair, E.-M. En-Naimi, B. Amami, H. Boukachour, P. Person, and C. Bertelle, "Intelligent tutoring systems founded of incremental dynamic case based reasoning and multi-agent systems (its-idcbr-mas)," in *Advanced Logistics and Transport (ICALT), 2013 International Conference on*, May 2013, pp. 341–346.
- [12] S. Piramuthu, "Knowledge-based web-enabled agents and intelligent tutoring systems," *Education, IEEE Transactions on*, vol. 48, no. 4, pp. 750–756, Nov 2005.
- [13] S. Nafiseh and M. Ali, "Evaluation based on personalization using optimized firt and mas framework in engineering education in e-learning environment," in *E-Learning and E-Teaching (ICELET), 2013 Fourth International Conference on*, Feb 2013, pp. 117–120.