Brill Academic Publishers
P.O. Box 9000, 2300 PA Leiden
The Netherlands

Lecture Series on Computer and Computational Sciences
Volume 6, 2006, pp. 1-3

Providing shortcuts to the learning process

I.Varlamis[1] and S. Bersimis[2]
[1]Department of Computer Science,
Athens University of Economics and Business,
GR-104 34, Athens, Greece
[2]Department of Statistics and Insurance Science,
University of Piraeus,
GR-185 34, Piraeus, Greece

Abstract: The aim and scope of traditional education differs significantly from these of life long learning. Education is mainly addressed to young people. It defines a framework for acquiring knowledge and cultivates the fundamental principles of learning and self-improvement in all levels. On the other side life long learning is targeted to adults who have received or missed basic education but are still enthusiastic in improving themselves. Individuals participate in life long learning sessions in order to improve specific competencies and acquire targeted knowledge. As a matter of fact, life long learning programs must adapt to the different needs of trainees and their design should be flexible in order to help them achieve the maximum in the minimum amount of time.

Keywords: Runs, Scans, Waiting Time, Test-based learning, adaptive tests.
Mathematics Subject Classification: 68U35 Information systems (hypertext navigation, interfaces, decision support, etc.)

1. Introduction

As many other adult learners, one of the authors used a computer based learning method in order to learn Spanish, in his spare time. In brief, the method presents several keywords or phrases accompanied by four images and the user must choose the image that depicts the phrase. The level of difficulty increases gradually, though in a standard rate. The same phrases are repeated in several sections in order to test comprehension. This repetition is initially accepted positively as a good rehearsal method. However, after a few successful sections the repetition becomes boring and the only option available is to skip questions or whole sections. Although, this option is very helpful and definitely student-centric, it is not very flexible and adequate to keep learner's interest high. Skipping of questions was proved tiresome, whereas skipping whole sections was insecure, since valuable new knowledge could be lost.

This experience is common with unsupervised systems, which are based on a rich pool of reading material and evaluation tests, but lack of guidance and flexibility. This experience reveals the need for ***shortcuts in the learning process***. In a competitive learning process, where time matters, it is preferable for the "wandering learner" to take a shortcut towards the target with the risk to go back at some point, than following the complete path to the target. This work capitalizes on the concept of *facilitating students by interrupting the test process when the desired level of comprehension have been reached* and promote them directly to the next level of difficulty. The paper discusses on the available theoretical models and focuses on the selected approach. Through an application scenario, which is under development, we point at the critical points of this approach.

2. Related Work

In order to support the test-based methods of learning we should provide learners with automatically created shortcuts that shorten learning time, while not affecting the acquired knowledge. An analysis of existing research in recommendation systems, learning methods and related research fields has been performed in order to choose the appropriate solution to the problem.

In web-based learning systems, a recommender is a software agent that tries to "intelligently" recommend actions to a learner based on the actions of previous learners. The recommendation systems origin in e-commerce applications but have been successfully tried in e-learning. The suggestions to learners are deduced using web mining and machine learning techniques and are based on pattern analysis of previous learners' behavior. The suggestions comprise on-line learning activities, related course material [7] or even course redesign [6]. The main drawback of these techniques is that they need information from many users in order to be draw useful results, which is not the case with computer-based but not web-based learning. In addition to this, the behavior of a learner is assigned a pattern and is examined collectively and not individually.

In computerized adaptive tests [4] the students answer questions at a certain level of difficulty and when certain criteria are satisfied they move to a new set of more difficult questions. Computerized adaptive tests use large databases comprising questions of various topics and levels of difficulty [5]. The students are not obliged to answer all the questions in the database, thus variation in topic and level of difficulty is performed based on several criteria.

A very stimulating idea for such criteria derives from the field of psychometry and in particular from studies in learning and memorizing. Psychologists, who perform tests to their patients, usually must define a measure in order to decide whether to interrupt the test and mark it successful or not. The criterion must take into account the success ratio and the probability of random answers. The same paradigm applies in learning: a learner carries out a test comprising of successive questions and is expected to succeed (S) or fail (F). In an unsupervised procedure, a measure is needed in order to define: first whether answers are given randomly or not and second when to interrupt the test and advance in level or to redirect the user to easier questions. One of the most widely used criteria is the runs criterion defined by Grand [3], which is based on the sequential nature of questions. The main notion behind this criterion is that *the probability a person to succeed in random decreases when the number of consecutively successful answers increases*. In the following sections we will present a general framework for developing such criteria based on runs and generalization of them.

3. Motivating example

In Table 1 we present the use of Grand's [3] criterion. The table presents the success (S) or failure (F) of three students in the same sequence of questions. Based on the runs criterion, we decide that a student successfully completes a test when 10 consecutive correct answers are given. As a result the first student completes the test successfully after answering 19 questions, the second students succeeds after 28 questions, whereas the third student fails the test after answering the whole set of 50 questions.

Table 1. Test status based on Grand's runs criterion

Student No	Sequence of answers	Status
1	SFSSFSSSF**SSSSSSSSSS**	Success
2	SFSSFSSSFSSFFFSSSF**SSSSSSSSSS**	Success
3	SFSSFSSSFSSFFFSSSFSSSSSFFFSFSSSSFSFSFFSSSSSSFSSSFS	Failure

It is obvious that this approach is very simplistic. In practice, more complex cases are met. A typical case is a test, which combines questions that evaluate more than one attributes of learning, for example "the level of understanding" and the "level of expressing" (e.g. the ability to read, write or speak a foreign language). The learner is assigned a grade (i.e. low-0, middle-1, high-2) for each distinct feature and the test completes when the desired level is achieved in all attributes. In computer based adaptive tests, the runs criterion [3] is inadequate and this is obvious in the next example (Table 2).

Table 2. Test status based on Grand's runs criterion

Student No	Sequence of answers	Status
1	FSFFF**SSSSS**	Success
2	FFSFFFSFSFF**SSSSS**	Success
3	SSFSSSSFSSFSSSSFSSSS	Failure

Considering that the rule for success is to give five consecutive correct answers, we notice that the first student succeeds the test in the 10^{th} question with a 70% success rate. The second student similarly succeeds, however with a lower rate (50%). The third student fails the test after 20 questions although the success rate reaches 80%. Although the criterion of runs mark the attempt as failed, the third student has fulfilled successfully many criteria, even from the 7^{th} question (i.e. a scan of length 5 containing at least 4 successes), and is obvious that can be promoted to the next level. Even in more complex tests that run in parallel and have more than two states (i.e. low, middle, high instead of fail and succeed) such criteria prove to be more reliable. The model presented in [2] is the basis for defining interrupting criteria for two parallel tests.

4. Criteria based on runs and scans

In the above examples, we extensively used the notion of '***run***'. Thus, before advancing to a more mathematical example we will proceed with a definition of run and related concepts. Generally, a run may be defined as an uninterrupted sequence of similar elements bordered at each end by elements of different type. The number of elements in the sequence is referred as the ***length of run***. By extending the concept of run we may define the notion of '***scan***' of length ***k*** as a sequence of *k* elements in which ***at least r*** are of the same type. Taking it a step further, we use runs or scans in order to define the notion of ***waiting time until the first appearance*** of a success run or a scan *r*/*k* (for a detailed presentation of the theory of runs and related literature, the interested reader may consult the work of Balakrishnan and Koutras [1]). Moreover, the waiting time until the first occurrence of a run of successes is the number of trials until the first completion of a criterion based on runs. To provide an example, in Table1, the waiting time for the first student to complete a run of successes was 19, whilst the corresponding time for the second student was 28.

The waiting time until the first completion of a criterion is a tool of great importance, because of the fact that may be used in order to define more sensitive criteria than those based on the appearance of one or more runs or scans. For example, if we denote the waiting time as T (T is the total number of questions involved in a subject's test) until the subject completes a criterion of k consecutive trials containing at least r successes, then a reasonable decision scheme is the following:

> "If $T \leq$c then the student has succeeded in a section and can be transferred in another section of the test. On the contrary,
>
> if $T>$c then the student has failed the test and must study and repeat the section".

It is obvious that c will be chosen according to the maximum allowable level (probability) of reaching a wrong decision through the aforementioned rule.

In statistical terminology, the success or the failure of a subject according to the abovementioned rule may be seen as a hypothesis test of the form

$$H_0 : p \leq p_{random} \\ H_1 : p > p_{random}$$.

In a numerical example, in which a student answers consecutively questions with 5 possible choices, the probability is p_{random}=1/5=0.2. Also, let a criterion using a value for parameter k equal to 5 and a parameter r equal to 4 (4 correct answers in a series of 5 questions). Then, using the distribution of the waiting time random variable under the null hypothesis (Figure 1 and Figure 2), we find a critical value of 17 for c ($\Pr[T \leq 17 \mid H_0] = 0.052705$). Thus, if the value of T for a person is greater than or equal to c=17, we reject the null hypothesis using statistical significance $\alpha = \Pr[H_1 \mid H_0] = 0.05 = 5\%$.

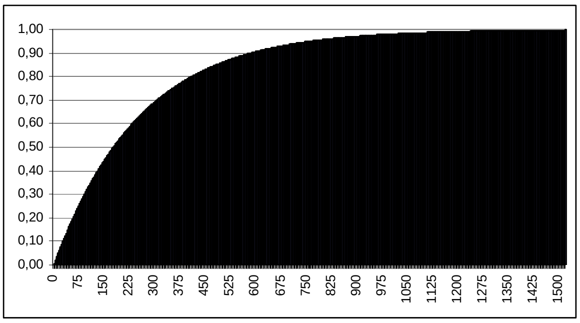

Figure 1. Cumulative Distribution of the Random Variable T with parameters k=5, r=4, and p_{random}=0.2

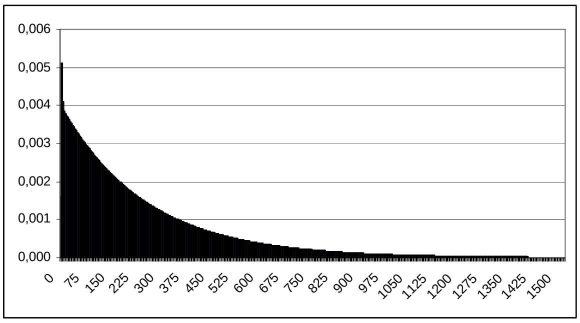

Figure 2. Probability Distribution of the Random Variable T with parameters k=5, r=4, and p_{random}=0.2

5. The adaptive test scenario

In order to test the aforementioned criteria based on waiting time random variables, we decided to incorporate the ability to interrupt tests, in favor of students' time, in a prototype adaptive learning system we develop. The system (an early version of it has been presented in [6]) provides suggestions according the content and structure of a course based on the students' preferences and performance in tests. The tests consist of series of questions (of close-type, i.e. multiple choice questions) which are presented to the user before, during and after the study of a topic (three phases).

The first step of the evaluation scenario for the criteria is to create series of questions of different difficulty level and topic. The next step is to present the tests to the students before they read the material, and make them to answer the questions of a level. A wrong answer or a skipped question equals to failure, while a correct answer is added to the run of successes. It is expected that novice students will answer the whole set of questions and will finally fail the section. As a consequence, they will proceed with reading the course material and rerun the test, until the interruption criterion is satisfied. In this case, the students will proceed to the next level. On the other side, the advanced students will have the ability to directly proceed to the next level, if they manage to fulfill the interruption criterion from the first phase (e.g. if the achieve a successful run in a short period of questions).

The added value of this approach is that students sooner arrive at the level of difficulty that best matches their expertise. Even in the case of failure, they have the ability to read and rerun the test in parallel (second phase), however with more strict criteria. This will increase the comprehension of a certain topic in the minimum amount of time and will allow students to self-improve their knowledge through reading. The second chance which is given in phase two, will remove the barrier of consecutive failures in the same test, since the reading material will allow readers to succeed the test. Finally, students will have the ability to validate the acquired knowledge by running the test without reading the course material (phase three).

It is essential to use different criteria in all three phases of testing through the learning process. The selection of criteria and the tuning of parameters (mainly thresholds) will be performed on the applied system and is on our next plans.

6. Conclusions – Future Work

This paper presents an ongoing work, which aims in facilitating learners by inferring shortcuts in test-based methods used for learning. The theoretic models which are based on success runs, time until the first appearance of a run or a scan can be a powerful tool for defining criteria for interrupting or creating shortcuts between topics or levels of comprehension in computerized tests. The theoretical framework has been defined and tested in many use cases and the scenario of the test procedure has been decided. The next step is to develop a prototype application that will allow the testing and tuning of the recommendation engine in the scope of a real course.

References

[1] Balakrishnan N. and Koutras M.V., (2002). *Runs and Scans with Applications*, Wiley, New York.

[2] Balakrishnan, N., Bersimis, S. and Koutras, M.V. (2006). *Waiting until the first occurrence of patterns in a bivariate sequence of trinomial trials with some applications in quality control*, Submitted paper.

[3] Grand, D. (1946). *New statistical criteria for learning and problem solution in ex-periments involving repeated trials*, Phycologika Bulletin, 43, 272-282.

[4] Lord F. (1971) *Tailored testing, an application of stochastic approximation*, Journal of the American Statistical Association, 66, 707-711.

[5] Henze, N., & Nejdl, W. (2001). *Adaptation in open corpus hypermedia*. International Journal of Artificial Intelligence in Education, 12(4), 325-350. Available online at http://cbl.leeds.ac.uk/ijaied/abstracts/Vol_12/henze.html

[6] Varlamis, I., Apostolakis, I., Karatza, M. (2005). *A Framework for Monitoring the Unsupervised Educational Process and Adapting the Content and Activities to Students' Needs*. WISE Workshops, 124-133.

[7] Zaiane, O.R. (2001). *Web usage mining for a better web-based learning environment. Proc. Conference on Advanced Technology for education* CATE 2001, 60-64.