Shuffling Algorithms for Automatic Generator Question Paper System

Nor Shahida bt Mohd Jamail & Abu Bakar Md Sultan

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia, 43400 UPM SERDANG, Selangor, Malaysia

E-mail: shahida_jamail@yahoo.com

Abstract

Examination process is important activities for educational institutions to evaluate student performance. Thus the quality of the exam questions would determine the quality of the students produced by the institutions. Preparing exam questions is challenges, tedious and time consuming for the instructors. Usually the instructors keeping their own test bank in some form to help them prepare future exams. Current technologies help the instructors to store the questions in computer databases. The issue arise is how the current technologies would also help the instructors to automatically generate the different sets of questions from time to time without concern about repetition and duplication from the pass exam while the exam bank growing. This paper describes the usage of shuffling algorithm in an Automatic Generator Question paper System (GQS) as a randomization technique for organising sets of exam paper. The results indicate shuffling algorithm could be used to overcome randomization issue for GQS.

Keywords: Automatic generator question system, Randomization, Shuffling algorithms

1. Introduction

Producing a quality graduates is one of the main objective in any educational institution. The higher acceptance of their graduates in work market indicates the quality of the institutions. The quality of graduates produced by any institutions is determined by many factors. One of the factors is the quality of the evaluation system. Evaluation system could exist in many forms. Conventional evaluation system is normally based on the exam system. Before the exam could be given to the student, the instructors or lecturers must prepare the questions according to the topics covered for each of the subject. Preparing exam questions is challenges, tedious and time consuming for the instructors. Usually the instructors keeping their own test bank in some form to help them prepare future exams. Current technologies help the instructors to store the questions in computer databases. The issue arise is how the current technologies would also help the instructors automatically generate the different sets of questions from time to time without concern about repetition and duplication from the pass exam while the exam bank growing.

2. Current System of Generator Question Paper

There are a few systems in today market that offered the similar services like what describes in this paper. These systems are developed by different developer with different features. The randomization techniques employed by such system is also different. Below is one of current system.

2.1 ASTech's JARPPL System

Atlantic Simulation and Training Tecnologies (ASTECH, 2010) introduced automated examination management system known as JARPPL. JARPPL features as describes in ASTECH includes the entire basic requirement essential for managing examination process. The main function includes importing of the generating MCQ (Multiple Choice Questions) exams, automatic exam correction, storing results and producing reports.

ASTech's JARPPL System provides an intuitive interface and incorporates a main menu allowing access to any of the functionality at any stage. An On-Line help file, accessible from the main menu, ensures that instructions on using the system are always accessible.

The main functionality of the system as summarised in their web page is as follows:

i. Generate draft examinations by randomly or manually selecting questions from the central question databank.
ii. Sign-off and prepress.
iii. Print individualised exam papers and solution sheets
iv. Maintain complete records of applicants (exam results, dates, etc.)
v. Produce school, NAA and JAA pro forma reports.

3. Generator Question Paper System (GQS)

This paper introduced the usage of shuffling algorithm in Automatic Generator Question Paper System (GQS) to overcome the issue stated. The main role of the shuffling algorithms is to provide randomization technique in GQS thus different sets of question could be generated without repetition and duplication. Randomization technique is a method that has grown enormously over the past 20 years (Ian, 1998). This technique in widely use in gambling field. There have four categories of randomization which are generating a random permutation of a sequence, selecting a random sample of a population, generating random numbers and transforming a data stream using a scrambler in telecommunications (Wikipedia, 2010).

GQS is a special online system which generates question paper or test papers of user's choice just on button click based. GQS is the system that has features for keeping test bank question and produced exam paper. This is automatic process of exam paper generation together with answering scheme. It uses a huge question bank with based on the learning outcomes elements which refer to the Bloom's Taxonomy. Bloom's Taxanomy includes 6 elements of learning outcomes which are knowledge, comprehension, application, analysis, synthesis and evaluation. Bloom's Taxonomy (Barbara, 1994) divides the way people learn into three domains. One of these is the cognitive domain which emphasizes intellectual outcomes. This domain is further divided into categories or levels. The key words used and the type of questions asked may aid in the establishment and encouragement of critical thinking, especially in the higher levels. This system is very useful for lecturers. The system is available for users to generate the questions of all the subjects of exams. Lecturers also will free from tension of collecting question to generate question paper for examination.

The functions in GQS are embedded with learning outcomes measure that would help lecturers produced quality exam question according to learning outcome objective for each course. This system is being tested for Faculty Computer Science and Information Technology (FSKTM) in Universiti Putra Malaysia. The questions are stored in huge Question bank. Lecturers can add new questions, delete the old questions and update the existing question in the Question bank at anytime. Lecturers could also generate different sets of question papers from the same database with just one click by selecting all the requirements needed. This system will prevent duplication of the questions by using the shuffling algorithms. The option to choose shuffling algorithm for randomization is because of simplicity. Apart of shuffling algorithm features are to prevent duplication and repetition of generated sets of question paper. All the questions are picks randomly from the database. Lecturers also can save generated papers for later use or delete after print.

4. Shuffling Algorithms Implementation.

Shuffling algorithms is very suitable and effective ways to implement for randomization of generated question. This algorithms check for duplication and repetition of the randomize question. The behaviour of the algorithm is as followed, for a set of N (the total number of question in the database) elements for generating a random permutation of the numbers 1–*N* goes as follows (wikipedia, 2010c)

1) Select the numbers from one to *N* in the database.

2) Pick a random number *k* between one and the number of unstruck numbers remaining (inclusive).

3) Counting from the low end, strike out the *k*th number not yet struck out, and mark it.

4) Repeat from step 2 until all the numbers have been struck out.

5) The sequence of numbers written down in step 3 is now a random permutation of the original numbers.

5. Findings and Discussions

The main objectives of the system are:

1) To applied randomization technique for the Automatic Generator Question Paper System by using shuffling algorithms.
2) To help lecturers preparing set of question paper based on the learning outcomes elements.

5.1 Randomization Achievement

The first objective has been achieved when generating 5 sets of question paper with same attributes. Result shown the questions are picking up randomly in the database according to the sequences. All 5 sets question papers generated by same lecturer using GQS. The lecturer generated 5 sets with same attributes and characteristics with 30 questions each set. The shuffling algorithms demonstrated that the entire question in 5 different sets is generated randomly. The result also indicates no duplication and repetition of the produced question. The shuffling algorithm employed shows it effectively support the randomization techniques especially in selecting the random data in one population.

In the implementation of GQS, The system successful generated random questions paper. Table 1 show that the first objective is achieved. (Note 1)

5.2 Randomization Based on Learning Outcomes Elements

GQS is an automatic generator question paper which is based on the learning outcomes elements. All the questions prepared are accompanied by their set of learning outcomes to be achieved by each of the respective topics.

The second objective for GSQ which is to help lecturers preparing set of question paper based on the learning outcomes elements. This objective has been achieved when generated 3 types of different question paper which are test 1, test 2 and final examination. All the question paper has 20 questions per set. The question paper that had been generated shown that all the questions generated are followed the rules that randomly selected based on the learning outcomes elements. Figure 2 shows all the questions generated are randomly selected based on the learning outcomes element. This indicates that second objective is achieved. Figure 3 shows percentage of Test 1 question paper generated according to stated learning outcomes. (Note 2)

5.3 Advantages of GQS

The system introduced brings several advantages to user compare to manual system. Listed below are the advantages of the system that brings to tourists:

i. Lecturers can generate question randomly by using GQS

ii. Save time in preparing paper for the examination.

iii. GQS can helps lecturers to generate the question based on the learning outcomes elements.

iv. Shuffling Algorithms helps randomization process in selecting questions in the database thus preventing duplication and repetition.

v. The question can be added to the bank at any time.

vi. Different sets of questions could be generated without any limitation.

vii. Implementing learning outcomes could classify the question and possible way to evaluate the level of achievement of each course.

6. Conclusion

The main purpose of this paper is to describe an automatic generator question paper using Shuffling Algorithm for randomization. This system is web-based application system with several features mainly producing unduplicated sets of exam paper. The result indicates the potential evidences of employment such algorithm for this type of system. Our future effort is to employed different types of randomization to see the most effective randomization method for such system.

References

Astech. (2010). Atlantic Simulation and Training Technologies. [online] Available: http://www.astech.ie/

Barbara,F. (2004). Bloom's Taxonomy and Critical Thinking [online] Available http: //mcckc.edu/ longview / ctac/ blooms.htm

Barbara,F. (2004). Bloom's Taxonomy and Critical Thinking [online] Available http://mcckc.edu/ longview/ctac/blooms.htm

Ian, H. (1988). Telepathy: Origins of Randomization in Experimental Design. Isis (A Special Issue on Artifact and Experiment) 79 (3): pp. 427-451.

Wikipedia. (2010a). Randomization. [online] Available: http://en.wikipedia.org/wiki/Randomization

Wikipedia. (2010b). Random Permutation. [online] Available: http: //en.wikipedia.org/ wiki/ Random_ permutation

Wikipedia. (2010c). Shuffling Algorithm. [online] Available: http:// en.wikipedia.org/ wiki/Shuffle #Shuffling_algorithms

Table 1. Randomization of the question selected using Shuffling Algorithms

		No of Question set									
		1	Status	2	Status	3	Status	4	Status	5	Status
Question Number	1	34		10		8		32		56	
	2	71	FALSE	37	FALSE	29	FALSE	51	FALSE	149	FALSE
	3	94	FALSE	78	FALSE	43	FALSE	89	FALSE	222	FALSE
	4	105	FALSE	121	FALSE	60	FALSE	116	FALSE	236	FALSE
	5	108	FALSE	635	FALSE	130	FALSE	138	FALSE	653	FALSE
	6	111	FALSE	705	FALSE	133	FALSE	202	FALSE	659	FALSE
	7	139	FALSE	736	FALSE	155	FALSE	212	FALSE	671	FALSE
	8	148	FALSE	754	FALSE	206	FALSE	224	FALSE	693	FALSE
	9	208	FALSE	755	FALSE	217	FALSE	619	FALSE	708	FALSE
	10	229	FALSE	1173	FALSE	640	FALSE	642	FALSE	747	FALSE
	11	652	FALSE	1203	FALSE	641	FALSE	675	FALSE	1162	FALSE
	12	701	FALSE	1213	FALSE	646	FALSE	691	FALSE	1200	FALSE
	13	758	FALSE	1255	FALSE	665	FALSE	704	FALSE	1202	FALSE
	14	764	FALSE	1260	FALSE	676	FALSE	761	FALSE	1222	FALSE
	15	778	FALSE	1262	FALSE	711	FALSE	788	FALSE	1233	FALSE
	16	1186	FALSE	1285	FALSE	768	FALSE	1176	FALSE	1236	FALSE
	17	1211	FALSE	1303	FALSE	782	FALSE	1179	FALSE	1248	FALSE
	18	1230	FALSE	1306	FALSE	789	FALSE	1196	FALSE	1305	FALSE
	19	1231	FALSE	1316	FALSE	1176	FALSE	1267	FALSE	1332	FALSE
	20	1274	FALSE	1317	FALSE	1198	FALSE	1298	FALSE	1707	FALSE
	21	1281	FALSE	1332	FALSE	1227	FALSE	1307	FALSE	1708	FALSE
	22	1290	FALSE	1702	FALSE	1231	FALSE	1334	FALSE	1733	FALSE
	23	1303	FALSE	1710	FALSE	1322	FALSE	1726	FALSE	1785	FALSE
	24	1304	FALSE	1736	FALSE	1717	FALSE	1767	FALSE	1787	FALSE
	25	1311	FALSE	1760	FALSE	1765	FALSE	1798	FALSE	1804	FALSE
	26	1316	FALSE	1765	FALSE	1815	FALSE	1817	FALSE	1815	FALSE
	27	1734	FALSE	1773	FALSE	1817	FALSE	1830	FALSE	1821	FALSE
	28	1771	FALSE	1789	FALSE	1825	FALSE	1861	FALSE	1847	FALSE
	29	1818	FALSE	1796	FALSE	1850	FALSE	1868	FALSE	1861	FALSE
	30	1880	FALSE	1858	FALSE	1859	FALSE	1879	FALSE	1878	FALSE

Notes

Note 1

Lecturer's Name : Dr Abu bakar Md Sultan
Subject : SAK5090
Type of Test : Test 1
Hours : 2H 30M
Marks : 100%
Date : 20/11/2009
Semester : 1
Year : 2009
No of Question : 30
Chapter : 1, 2, 3

Figure 1. The attributes used in generated 5 of question paper.

Note 2

No.	Type	Question	No of Question Selected	No of Chapter
1	C1	A yellow excalmatory mark oftem shows in the center of network connection which on the screen right-down corner. network will be stopped to inquiry then re-got IP while the yellow mark shows. Whats happend with the computer?	1	1
2	C5	is it possible to IPL for individual LPAR or its related only for total server?	48	1
3	C1	Why demanding applications that can compensate for variations in bandwidth and delay with large receive buffers, which is often possible for example in video streaming.	63	2
4	C2	Why Internet peering arrangements are already complex, and there appears to be no enthusiasm among providers for supporting QoS across peering connections, or agreement about what policies should be supported in order to do so.	85	2
5	C5	Software quality can be defined as 'conformance to requirements' and/or 'fitness of use'. Please explain.	110	2
6	C5	By achieving CMM Level 2, projects can set realistic expectations, commit to attainable deadlines and avoid the Level 1 "death marches" on nights and weekends that produce excessive defects. Please explain.	217	3
7	C1	The definition of DQ that is often used is "...fit for use...". Explain.	619	1
8	C4	From the study carried out, we plan to introduce new framework to enhance the available quality model for web portal which is PSP/DQ model. Explain.	653	1
9	C4	Sun certification also offers a natural progression to support your career goals. Explain	715	2
10	C2	The process is then documented, including project scope, process flow, Work Breakdown Structures (WBS). Explain.	752	3
11	C1	This architecture allows greatly simplified software implementation. Peer-to-peer architecture was originally based on the Client-Queue-Client concept. Explain.	1158	1
12	C4	Notable distributed networks that use DHTs include BitTorrent's distributed tracker, the Kad network, the Storm botnet, YaCy, and the Coral Content Distribution Network. Explain.	1197	1
13	C6	Flooding also causes a high amount of signaling traffic in the network and hence such networks typically have very poor search efficiency. Most of the popular P2P networks are unstructured. Explain.	1212	1
14	C1	An other type of hybrid P2P network are networks using on the one hand central server(s) or bootstrapping mechanisms, on the other hand P2P for their data transfers. Explain.	1220	2
15	C3	The distributed nature of P2P networks also increases robustness, and—in pure P2P systems—by enabling peers to find the data without relying on a centralized index server. Explain.	1239	2
16	C6	Today, vector spaces are applied throughout mathematics, science and engineering. Explain.	1277	2
17	C2	"Topology," its English form, was introduced in 1883 in the journal Nature to distinguish "qualitative geometry from the ordinary geometry in which quantitative relations chiefly are treated". Explain.	1294	3
18	C4	Different forms of such reflection on reasoning occur in different fields. In philosophy, the study of reasoning typically focuses on what makes reasoning efficient or inefficient, appropriate or inappropriate, good or bad. Explain.	1732	1
19	C6	Aristotle points out that by understanding the reasoning involved in this type of argument, we can know that whatever the As and Bs are, we can reach the same conclusion about the relationship between them. Explain.	1754	1
20	C2	Islamic logic not only included the study of formal patterns of inference and their validity but also elements of the philosophy of language and elements of epistemology and metaphysics. Explain.	1771	2

Figure 2. Test 1 question paper with 20 random questions.

Rules for Test 1

%	Learning Outcomes Elements	Calculation	Total
50	C1, C2, C3	50%/20	10
35	C4, C5	35%/20	7
15	C6	15%/20	3
Total Question			20

Figure 3. Generated Test 1 based on Learning Outcomes