Use of an Evolutionary Approach for Question Paper Template Generation

Dimple V.Paul
Dept. of Comp. Sc.
D.M's College
Mapusa, Goa, India
dimplevp@rediffmail.com

Shankar B. Naik
Dept. of Comp. Sc.
Dhempe College
Miramar,Goa,India
xekhar@gmail.com

Priyanka Rane
M.Sc IT(student)
Chowgule College
Margoa,Goa,India
priyanka013@rediffmail.com

Jyoti D.Pawar
Dept.of Comp. Sc & Tech.
Goa University
Taleiegao,Goa,India
jyotidpawar@gmail.com

***Abstract*—This paper focuses on question paper template generation and its use in dynamic generation of examination question paper. Question paper template generation is a constrained based optimization problem. Choosing an efficient, scientific and rational algorithm to generate a template is the key to dynamic examination question paper generation. By using the evolutionary computational search technique of evolutionary programming and educational taxonomies, this paper analyses and experimentally proves that the generated question paper templates are best suited for dynamic examination paper generation. This new technique outperforms traditional algorithms in terms of coverage of topics, learning domains and marks distribution in the generated question paper.**

Keywords- blooms taxonomy; evolutionary programming; question bank; question paper template

I. INTRODUCTION

Every educational system consists of an examination system through which the qualities and abilities of the students are assessed in order to give grades and positions [1]. Preparing an examination question paper is challenging, tedious and time consuming for the paper setters. A high quality question/test paper must necessarily address the following issues –

- Are contents of question paper appropriate in terms of coverage of topics?
- Can questions (test items) examine student ability at different cognitive levels of Educational Taxonomy?
- Is it possible to avoid similarity among questions?
- Does the question paper satisfy both time and marks constraint?
- Can we generate question papers of different difficulty levels?

The efficiency and quality of dynamic examination paper generation depends entirely on its algorithm design. Traditional question paper generation algorithms have the disadvantages of slow convergence, low success rate and poor quality. Hence we propose an evolutionary programming technique for question paper template generation by considering the changing requirement of each paper setter. The question paper template can be used to generate examination question papers dynamically.

This paper is organized as follows. Section 2 presents the research background and related work, Section 3 presents the methodology used, Section 4 explains the problem formulation, Section 5 discusses the implementation details, Section 6 concludes the paper and Section 7 presents the future work.

II. RESEARCH BACKGROUND AND RELATED WORK

Three approaches commonly used till date for question paper generation are as under –

a. the question database is viewed as the questions set for selection by experienced paper setters;
b. previously generated question paper set is used for selection;
c. an intelligent algorithm is used for automatic generation of question paper.

The first and second approaches are able to guarantee quality. But the first one works similar to manual operation, while the second limits the number of papers. The third type is widely used in educational field at present and this approach is already implemented by many researchers [2-3]. But the main limitation of many of these existing implementations is that they use a random generation strategy and create thousands of question papers with repetitive questions lacking in coverage of topics and learning objectives.

In our approach, we generate a dynamic template and use it for generation of question paper that has proper weightage allotted to subject content, learning domain, type of question, etc. and can be used to generate several question papers almost without repetition depending on the paper setter's choice. The number of unique question papers (without any overlap) that can be prepared using a generated template depends on the quality and size of the Question Bank for that subject. The quality of a Question Bank is decided on the basis of the type of questions, such as questions of 1 mark, 2 marks, 5 marks etc., that exist for each unit under different levels of an educational taxonomy in the Question Bank.

The approach presented in this paper monitors the quality of question paper based on a wide range of paper setter requirements such as the average degree of difficulty, kinds of questions, selection of modules,

978-0-7695-4759-6/12 $26.00 © 2012 IEEE
DOI 10.1109/T4E.2012.24

selection of cognitive levels, etc.,. In order to incorporate the above requirements, we have applied the well established concept of Educational Taxonomies and evolutionary programming. Evolutionary programming offers a robust non-linear search technique that is particularly suited to problems involving large number of variables. The strength of evolutionary programming is derived from their ability to exploit, in a highly efficient manner, information about a population. This search method is modeled on natural selection by Holland in [4] and is being used to solve a variety of optimization problems.

Evolutionary programming applies genetically inspired operators to populations of potential solutions in an iterative fashion, creating new populations while searching for an optimum solution [5-6]. Despite their randomized nature, evolutionary programming is not a simple random search. It takes advantage of the old knowledge held in a parent population to generate new solutions with improved knowledge.

III. METHODOLOGY

The methodology adopted consists of the following main steps -

A. Select Units or Modules of a Subject

Examination is conducted for a subject of a course having pre-defined university specified syllabus file with unit-wise contents. It is necessary to mention whether the question paper template is designed for all units or modules of a subject or selected units of a subject.

B. Decide Cognitive Processing Levels

Paper setters always attempt to include questions that measure higher levels of cognitive processing. This is not a good approach to evaluate performance of students at different levels of learning such as Excellent, Good, Average, etc.,. It should be the goal of the paper setter to ensure that their questions have cognitive characteristics testing the understanding, problem-solving, critical thinking, analysis, synthesis, evaluation and interpretation rather than just declarative knowledge. There are many theories that provide frameworks on levels of thinking which have serious impact on framing good questions. Bloom's taxonomy [7-9] is often cited as a tool to use different cognitive levels in choosing the questions. Cognitive processing levels for a question paper template are decided on the basis of the taxonomy that is selected by the paper setter for each examination. Difficulty level of an examination paper is defined based on the type of template that is used for the same.

C. Design Question paper Template

The question paper template shown in Fig.1 can be thought of as a systematic design plan. It lays out exactly how the question paper will be created. The template specification also gives the flexibility for constructing many qualitatively good examination question papers using the same template. The template defines the scope of the paper with respect to syllabus contents and content of skills being measured by the examination.

We have generated templates with three major difficulty levels such as High, Medium and Low. High difficulty templates have high distribution of marks across higher/difficult levels of taxonomy. Medium difficulty templates have proportionate distribution of marks across all levels of taxonomy and low difficulty templates have low distribution of marks across higher/difficult levels of taxonomy. The question paper template so generated is then used for dynamic generation of a question paper. The quality of the question paper will depend on the quality of questions framed by paper setters and populated in the Question Bank [10]. The question paper generation from the template is outside the scope of this paper.

TABLE I. TERMINOLOGY USED

Term	Meaning
Course	Refers to a Degree/Diploma programme offered at a university. Example:- 1.Bachelor of Science (Computer Science) – B.Sc.(Comp. Sc), 2.Bachelor of Computer Application – BCA
Subject	Subject is a paper which is offered in different semesters of a course. Example:- Software Engineering (SE) in the 6th Semester of B.Sc (Comp. Sc).
Modules or Units	For each subject, there is a university prescribed syllabus which consists of different modules/units. A set of related topics are grouped as one unit/module. Each of these modules is allotted a particular weightage. Example:- Module on Software Requirement in SE subject is given a weightage of 30% in the 6th semester of B.Sc (Comp. Sc).
Paper setter	A paper setter is a faculty who is appointed to frame the question paper for a subject in a course for a particular examination.
Educational Taxonomy	It is a classification system of educational objectives based on the level of student understanding necessary for achievement or mastery. Example:- Blooms, Solo etc.
Educational Taxonomy Levels	Each Educational Taxonomy has its cognitive stages in learning and is called taxonomy levels. Example:- Blooms Taxonomy Levels: Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation
Question Paper Template	A matrix with rows representing modules, columns representing cognitive levels of taxonomy and cells representing weightage assigned to a level of a module. (Fig. 1)
Module-Level-Weight	Weight allotted to a cell of a template which represents the weight assigned to the particular level of that module.

Units	level 1	level 2	level 3	...	level n	Units weight
unit 1	x_{11}	x_{12}	x_{13}	…	x_{1n}	u_1
unit 2	x_{21}	x_{22}	x_{23}	…	x_{2n}	u_2
…	…	…	…	…	…	…
unit m	x_{m1}	x_{m2}	x_{m3}	…	x_{mn}	u_m
Level weight	l_1	l_2	l_3	…	l_n	N

Figure 1. Question Paper Template Format.

IV. PROBLEM FORMULATION

A. Steps for applying Evolutionary Progamming

1) Generate Population: Population consists of P (paper setter input) different question paper templates which are either generated at initialization or at successive iterations. A template is formed by randomly assigning module-level-weight to each cell of it in such a way that it satisfies the unit weightage and level weightage.

2) Calculate Fitness: Calculate the fitness score of each template. The details of Fitness calculation are explained in section B.

3) Selection: Apply selection operation to the generated population. It is carried out based on the criteria that the set of templates with fitness value in the range of 0.5-1.0 are to be identified and selected.

4) Mutation: Among the selected templates, identify the ones that can be mutated to increse their fitness value . Perform muatation operation on these identified templates.

5) Termination: Repeat step 1 till step 4 for the paper setter specified number of iterations or until optimal solution is found (whichever is earlier).

B. Problem statement

Let N be the total marks allotted for the question paper. Let m be the number of units selected by paper setter and n be the number of taxonomy levels. Let $U=\langle u_1, u_2,...,u_m\rangle$ be the vector of weights where u_i is the weight assigned to the i^{th} unit, and, $L=\langle l_1,l_2,\ldots,l_n\rangle$ be the vector of weights where l_i is the weight assigned to the i^{th} level.

Let $X=\langle x_{11}, x_{12}, x_{ij},\ldots,x_{mn}\rangle$ be the set of module-level-weights where x_{ij} is the module-level-weight assigned to the j^{th} level of i^{th} unit.

For a unit i, $\sum_{j=1}^{n} x_{ij}=u_i$, and

For a level j, $\sum_{i=1}^{m} x_{ij}=l_i$

The problem is to assign module-level-weights, $X=\langle x_{11},x_{12},x_{ij},\ldots,x_{mn}\rangle$, so as to get the optimum value for the fitness function (F). Let w_1 be the percentage of importance assigned to unit coverage and let w_2 be the percentage of importance assigned to taxonomy level coverage. The problem can be mathematically stated as follows-

$$\textbf{Max } F= \left(\left(w_1*\sum_{i=1}^{m} \left(1-\left(\sum_{j=1}^{n} |(x_{ij}*N)/u_i -l_j|/N\right)\right)/m\right) + \left(w_2*\sum_{j=1}^{n} \left(1-\left(\sum_{i=1}^{m} |(x_{ij}*N)/l_j -u_i|/N\right)\right)/n\right)\right) / (w_1+ w_2)$$

subject to the constraints

$$\sum_{j=1}^{n} x_{ij}=u_j$$
$$\sum_{i=1}^{m} x_{ij}=l_i$$
$$x_{ij}>0$$

In order to define fitness, we have defined the following terms-

1) The weakness of a unit (WU_i): For a unit i, $X_i=\langle x_{i1},x_{i2},\ldots,x_{in}\rangle$ and X_i X. Before calculating WU_i, normalize X_i to obtain $X_i=\langle x_{i1},., x_{in}\rangle$ such that $x_{ij}= x_{ij} *N/u_i$.

$$WU_i = \left(\sum_{j=1}^{n} |x_{ij} -l_j|\right) /N \qquad (1)$$

2) The fitness of a unit (Funit):

$$F_{unit} = \sum_{i=1}^{m} (1-WU_i)/m \qquad (2)$$

3) The weakness of a level (WL_j): For a level j, $X_j=\langle x_{1j},x_{2j},\ldots,x_{mj}\rangle$ and X_j X. Before calculating WL_j, normalize X_j to obtain $X_j = \langle x_{1j},., x_{1m}\rangle$ such that $x_{ij}= xij *N/l_j$.

$$WL_j= \left(\sum_{i=1}^{m} |x_{ij} -u_i|\right) /N \qquad (3)$$

4) The fitness of a level (Flevel):

$$F_{level} = \sum_{j=1}^{n} (1- WL_j)/n \qquad (4)$$

5) The overall fitness(F), of the template:

$$F = \left(w1*\sum_{i=1}^{m} F_{unit} +w2* \sum_{j=1}^{n} F_{level}\right) / (w_1+ w_2) \qquad (5)$$

V. IMPLEMENTATION DETAILS

A. Hardware and Software Platform Used:

Implementation is done using Microsoft Visual Basic .NET as Front End Tool and SQL Server as Back End Tool on a 2 GHz processor with 1GB RAM.

B. Datasets used:

Experimental study was conducted using the syllabus prescribed for Software Engineering (SE) offered in the third year of the three year bachelor's degree course of computer science (B.Sc Computer Science) at Goa University. Blooms taxonomy with six levels such as knowledge, understanding, Application, Analysis, Synthesis and Evaluation were considered as cognitive processing levels.

C. Sample Input screen:

Paper setter is given the flexibility to choose some/all units of a subject and also some learning objectives of the specified subject. Question paper template for different examinations such as in-semester (20 marks), end-semester (80 marks), practical (50 marks), etc. can be generated. Provision is also made to prepare question paper template on different difficulty levels such as Low, Medium and High. Fig.2 below shows the sample input screen for question paper template generation.

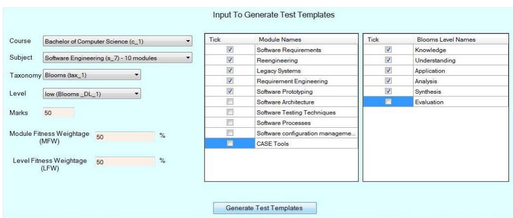

Figure 2:Input for question paper template generation

D. Results Obtained:

Fig. 3 below shows sample screen shost of the iterative stages of SE question paper template generation. It is generated by selecting only five units namely Software Requirement, Reengineering, Legacy Systems, Requirement Engineering and Software Prototyping of SE syllabus and also by choosing only five levels of blooms taxonomy namely Knowledge, Understanding Application, Analysis and Synthesis.

Figure 3 : Iteratively generated question paper templates with calculated fitness values

Tab. II below shows the experimental results obtained after iteratively generating SE templates with 5 units and 5 levels as input.The population size P was set to 10 and the number of iterations was set to 100. The number of templates generated in the worst case for this input could be 10*100=1000.The results indicates that the algorithm terminated with an optimal solution at the 82^{nd} iteration .

TABLE II. CALCULATED FITNESS FOR SE TEMPLATE WITH 5 UNITS AND 5 LEVELS

template_no	itearion	fitness
$i_{1,1}$	1	0.5031
$i_{1,2}$	1	0.5064
$i_{1,3}$	1	0.5099
$i_{1,4}$	1	0.5104
$i_{1,5}$	1	0.5159
$i_{1,6}$	1	0.5206
$i_{1,7}$	1	0.5279
$i_{1,8}$	1	0.5315
$i_{1,9}$	1	0.5397
$i_{1,10}$	1	0.5434
$i_{2,1}$	2	0.5498
$i_{2,2}$	2	0.5568
.	.	.
$i_{3,1}$	3	0.5922
.	.	.
$i_{4,1}$	4	0.6047
.	.	.
$i_{28,1}$	28	0.7003
$i_{28,2}$	28	0.6913
.	.	.
$i_{52,1}$	52	0.7588
$i_{52,2}$	52	0.7588
.	.	.
$i_{64,1}$	64	0.7204
$i_{64,2}$	64	0.7399
.	.	.
$i_{70,1}$	70	0.7802
.	.	.
$i_{77,1}$	79	0.8802
$i_{77,2}$	79	0.8905
.	.	.
$i_{82,1}$	82	0.9617
$i_{82,2}$	82	0.9902

Tab. III below shows the experimental results obtained after iteratively generating SE templates with 8 units and 6 levels as input.The population size P was set to 10 and the number of iterations was set to 100. The results indicates that optimal solution is not achieved even after 100 iterations.So the algorithm terminates after 100 iterations in this case.

TABLE III. CALCULATED FITNESS FOR SE TEMPLATE WITH 8 UNITS AND 6LEVELS

template_no	itearion	fitness
$i_{1,1}$	1	0.5001
$i_{1,2}$	1	0.5012
$i_{1,3}$	1	0.5034
$i_{1,4}$	1	0.5075
$i_{1,5}$	1	0.5093
$i_{1,6}$	1	0.5108
$i_{1,7}$	1	0.5127
$i_{1,8}$	1	0.5135
$i_{1,9}$	1	0.5149
$i_{1,10}$	1	0.5198
$i_{2,1}$	2	0.5202
$i_{2,2}$	2	0.5209
.	.	.
$i_{3,1}$	3	0.5223
.	.	.
$i_{4,1}$	4	0.5291
.	.	.
$i_{28,1}$	28	0.5703
$i_{28,2}$	28	0.5716
.	.	.
$i_{52,1}$	52	0.6102
$i_{52,2}$	52	0.6113
.	.	.
$i_{64,1}$	64	0.6313
$i_{64,2}$	64	0.6326
.	.	.
$i_{70,1}$	70	0.6741
$i_{70,2}$	70	0.6758
.	.	.
$i_{77,1}$	79	0.7005
$i_{77,2}$	79	0.7011
.	.	.
$i_{100,9}$	100	0.7318
$i_{100,10}$	100	0.7346

VI. CONCLUSION

This paper focused on a new approach for dynamic question paper generation by using question paper templates that are obtained using evolutionary algorithm. The primary objective of this study was to generate question paper template, using evolutionary programming that can be used in dynamic generation of examination question paper. The main advantage of this new approach is the application of strengths of evolutionary programming for use in dynamic question paper generation. Complexity of this template generation algorithm is proportional to the number of units and the number of levels.

VII. FUTURE WORK

In our work, we have focused on generation of the question paper template by using evolutionary programming. This template can be used to generate question paper, either by manually selecting the questions from the question bank or using intelligent algorithms for the selection of questions from the question bank. Our future work will focus on developing algorithms to automate question paper generation which will require the minimum number of inputs from paper setters.

REFERENCES

[1] Ahmad,Z. “Existing System of Examination and the Need of Reform” ,Journal of Elementary Education vol .1,3,56,1993.

[2] Adeli, H., & Cheng, N.,Concurrent evolutionary programmings for optimization of large structures. ASCE Journal of Aerospace Engineering, 7(3), 276–296, 1994.

[3] Ting Long, Houjun Wang ,Shulin Tian,” Test Generation Algorithm for Linear Systems Based on Evolutionary programming”, Journal of Electronic Testing: Theory and Applications, Volume 26 Issue 4, August 2010.

[4] Anbo Meng, Luqing Ye, Daniel Roy Pierre Padilla,”Evolutionary programming based multi-agent system applied to test generation”, Computers & Education 49, 1205–1223, 2007.

[5] K.G.Srinivasa,K.R,Venugopal,L.M.Patnaik,”A Self-Adaptive Migration Model Evolutionary programmings for Data Mining Applications”, In Proceedings of Inf. Sci.. 4295-4313,2007.

[6] Bloom (1956). Taxonomy of educational objectives: The classification of educational goals. In B. Bloom (Ed.),Handbook 1, cognitive domain. New York: Longman.

[7] Krathwohl, D. R., “A revision of Bloom’s taxonomy: An overview, Theory Into Practice”, Vol.41 (4), pp 212-219,2002.

[8] Yeh L. C. and LIN S. P., “The research of Bloom’s taxonomy for educational objects,” Journal of Education Research , Vol.105, pp. 94-106, 2003.

[9] Khairuddin, N. N., and Hashim, K.,” Application of Bloom’s Taxonomy in Software Engineering Assessments”, Proceedings of the 8th WSEAS International Conference on Applied Computer Scince, 2008.

[10] Guang Cena, Yuxiao Donevolutionary programming, Wanlin Evolutionary programmingoa,”A implementation of an automatic examination paper generation system”, Mathematical and Computer Modeling 51 1339-1342, 2010.