Designing A Generating System Multiple Choice Questions

E Winarko¹, E Roesminingsih²
¹ Mathematic Departement, Airlangga University, Campus C, Mulyorejo Street Surabaya 60115
² Postgraduate Education Management Departement, Surabaya State University, Lidah Wetan Street Surabaya 60213

E-mail: edi_winarko@fst.unair.ac.id

Abstract. The purpose of this research is to build a system of generating questions that can make as many questions as possible. The Generating System of this Problem can be used starting from elementary/middle/high school/higher education level with the type of choice questions that have 3 components, namely the problem, choice and completion. In order for the system to function properly, a template is needed for the three components above that can identify the needs of variables that might appear on a problem. Several variables are obtained from the identification of the results, namely: Numerical Variables, String Variables, Image Variables. Templates needed on the system are 3 templates, namely the template, template selection and completion template. All those three contain text, variables and identification tags. Variables that are permitted in this template are Numerical Variables, String Variables and Image Variables. While the identification tags that are permitted are numerical [var] numeric [/var]; variable tag string [str]collectname[/str]; the variable tag image [img]filename[/img], the tag to calculate [eval]equation[/eval] to be executed and generating the value and tag equation [equ]equation[/equ] that will be displayed as an equation. It was found in the result of the implementation that one template able to build hundreds/thousands of questions, combination of choices and the completion of each question.

1. Introduction
1.1. Background

In the current Digital Shield era, the use of technology in all fields continues to grow, starting from the utilization in the world of banking, industry, public services including the world of education. Following the development of technology, the learning process began to utilize the digital world by utilizing E-Learning to help the learning process. With E-Learning technology, students and teachers are expected to be able to make optimal use so that they can support the development of the three intelligences, namely intellectual intelligence (IQ), emotional intelligence (EQ) and spiritual intelligence (SQ). In order for the potential of intelligence to grow and develop optimally, learning in schools must be based on multiple intelligences, with a learning approach that uses PAKEM, I2M3 (interactive, inspirational, challenging, fun and motivating) and accelerated learning. The emphasis of the learning process is on communication skills (public speaking), collaboration (collaborating) and problem solving (problem solving).

The essence of teaching is communication between teachers and students. In the conventional learning method, the relationship between the teacher and students is very close, because between the instructor and students, they can meet and communicate directly. But, along with the development of technology, the media and means of learning are increasingly developing, demanding more innovative...
models and learning methods where the learning process leads to the absence of boundaries of SPACE and TIME.

In today's technological development, E-Learning is a basic requirement for the world of Education. The emergence of the Internet has provided major changes in most human lives. The use of internet technology is expected to be able to improve the ability of students in honing their abilities and preparing themselves to face replications, be it daily, semester or national tests.

Fear that sometimes seems too excessive is not only seen in the school environment, such as teachers and students, but parents and even the city government themselves intervene to prepare the children by making a repeat training try-out. Even parents are also worry about the ability of their children to finally include their children in the Tutoring program. Though the main key to success in following the test is as much as practicing the Question.

For this reason, the Teachers must prepare as many questions as possible with varying qualities, from easy levels, to difficult levels. The more questions that are prepared, the more practice questions that can be prepared.

1.2. Aims and objectives

This study aims to develop a problem generating system that aims as follows: 1) Build a system that can make as many questions as possible. 2) As a place for students to learn and practice to their heart's content.

2. Methods

The method used in developing this problem generator system uses the Waterfall method, as mentioned by [5] that the most popular method of developing software is the Waterfall Model. According to [1] Waterfall is a method that is easy to implement and inexpensive, this is what makes the reason the waterfall method is widely used. This Waterfall method is commonly referred to as structured analysis. According to [6] Waterfall reflects basic development activities consisting of 5 stages, namely: requirement analysis and definition, system and software design, implementation and unit testing, integration and system testing, and operation and maintenance.

The Waterfall method diagram can be seen in the following Figure 1:

![Figure 1. The phases of a Waterfall Model (Sommervile, 2011)](image_url)

2.1. Requirement Analysis and Definition

At this stage, gather information on the user to get information about the system requirements. The following steps will be carried out:

- Determine the type of question
- Determine the components in the problem
- Determine the variables that exist in each component of the question
• Determine requirements if needed

2.2. System and Software Design
At this stage, form a system architecture based on the results of the Requirement Analysis with the following steps:
- Determining the Program Base.
- Determine the Language of Programs and supporting programs.
- Designing a Database.
- Designing Data Flow Diagrams that describe the system as a whole.
- Designing identification tags for each variable that will be used.
- Determine support programs.

2.3. Implementation and Unit Testing
At this stage, the implementation of the design results is carried out by creating the required Database and procedures with the following steps:
- Create a database to accommodate the question template.
- Making procedures for introducing identification tags in the template.
- Create a value generating procedure for each identification tag
- Make a procedure to display for each template

2.4. Integration and System Testing
At this stage is the stage of integrating all procedures into a single program and conducting trials as a whole system. Step up to the following test:
- Combining all procedures that have been made in the previous stage.
- Prepare the template for testing
- Test each component of the test to see the results of the program.

2.5. Operation and Maintenance
At this stage is the stage of installation and application of the system in the case to determine the reliability of the system and to find and correct errors that are not found at the stage of manufacture. If possible, add features or facilities to the system. At this stage, the following steps will be carried out:
- Installing the system on the server.
- Conduct a trial for every possible input.
- Doing an error error that might occur.
- Repair the program that has errors.

To conduct a trial program, the most important thing to do is to prepare the template. Following are the steps needed to prepare the template:
- Prepare the concept of questions to be tested.
- Determine the parameters / variables and formulations in the question concept.
- Make a question template based on the draft questions that have been set.
- Make an answer choice template based on the predetermined draft choices.
- Make a completion template based on a draft of a predetermined settlement.
System Development Flow Chart The generator can be seen in Figure 1 below:

![Flow Chart of the Generator System](image)

Figure 2. Flow Chart of the Generator System

3. Result
   In designing the Generating Problem system, Data Flow Diagrams (DFD) have been designed in 2 levels as follows:

![DFD Level 0](image)

Figure 3. DFD Level 0

At level 0 this only describes the process and Entity in general, namely the Teachers and Students involved in the system.
At level 1, it can be seen that there are several processes involved in the Entity of the Teacher, namely the template entry, the selected entry, the completion entry and the question display. For the processes involved in Entity Students only the process of displaying questions and checking answers. While the database involved is a question bank to save the template that was created.

According to [7], databases are defined as groups of data that are interconnected and organized in such a way that they can be reused quickly and easily. The following are the results of database design in the form of Conceptual Data Model (CDM) using Power Designer 6.62:
The Conceptual Data Model (CDM) design is generated into Physical Data Model (PDM), which will be the actual Database. The PDM diagram can be seen as follows:

![PDM Diagram]

**Figure 6. PDM Question Bank**

This PDM design will be a database as in the Question Master Table.

Based on the development stage, the results are obtained, that for the choice questions, 3 types of templates are needed, namely questions, answer choices and solutions. In order for the system to function properly, a template for all three is needed to identify the needs of the variables that might appear on a problem. Identification results obtained several variables, namely: Numerical Variables, String Variables, Image Variables. Template is needed in the system with 3 templates, namely the template, template selection and completion template. Each template can contain text, variables and identification tags. The variables that are permitted in this template are Numerical Variables, String Variables and Image Variables with identifiable tags that are permitted are numerical numerical [var] variable tags [/ var]; variable tag string [str] collectname [/ str]; variable image tag [img] filename [/ img]. Meanwhile, to calculate the results of the equation, use the eval equation which will produce the value and equ equation that will be displayed as an equation. These tags will be recognized by the program using the PHP script as follows:

- `preg_match_all (#\[var\] (. *?) \[/ Var\] \#", $ sentence, $ matches)`
  This script is used to identify the presence of the [var] [/ var] tag along with the number of [var] [/ var] tag pairs.

- `preg_match_all (#\[str\] (. *?) \[/ Str\] \#", $ sentence, $ matches)`
  This script is used to recognize the existence of the tag [str] [/ str] along with the number of pairs of tags [str] [/ str].

- `preg_match_all (#\[img\] (. *?) \[/ Img\] \#", $ sentence, $ matches)`
  This script is used to recognize the presence of the [img] [/ img] tag along with the number of pairs of tags [img] [/ img].

- `preg_match_all (#\[equ\] (. *?) \[/ Equ\] \#", $ soala, $ matches)`
  This script is used to recognize the existence of the tag [equ] [/ equ] along with the number of pairs of tags [equ] [/ equ].

- `preg_match_all (#\[eval\] (. *?) \[/ Eval\] \#", $ questions, $ matches)`
  This script is used to recognize the existence of [eval] [/ eval] tags along with the number of pairs of [eval] [/ eval] tags.

After identifying and identifying variables in the template, a process is then carried out to access the values of each variable.
For the Numeric variable, the script $ dt [$ contents [$ i]] = rand (lower limit, limit) is used which will produce a number between the lower limit and the limit. String variables can be divided into several values that can be created by the user themselves, such as people's name strings, city name strings, country name strings and so on. The random value of this string variable is very dependent on the number of databases for each string variable. The [img] tag can be divided into several values that can be created by the user, such as gabor worms, fruit images, mountain images, triangular images, square images, circular images and so on. The random value of this string variable is very dependent on the number of databases for each image variable.

For the [equ] tag will display the equation placed between the [equ] and [/equ]. To make the appearance of the equation better, the Latex format is used, which can be accessed online through the script <img src='http://latex.codecogs.com/gif.latex?$equ$' border='0' /> with $equ$ is the equation obtained after the value of the numerical variable is included in the equation. The random value of this string variable is very dependent on the number of databases for each image variable.

For the [eval] tag the value that will be raised depends on the equation placed between the [eval] and [/eval] marks. The equation to be calculated is obtained after the value of the numerical variable is included in the equation.

From the results of the implementation, it was found that one template was able to build hundreds or even thousands of questions, a combination of choices and the completion of each of the questions depending on the number of variables contained in the question template. Where for the chosen template you will get a random sequence of choices.

4. Discussion

Based on the results of the database design to accommodate the question template, the Master Question table is created whose structure is as follows:

| No. | Atribut  | Type of Atribut | Information |
|-----|----------|----------------|-------------|
| 1   | Kdquestion | Integer(11)    | Key         |
| 2   | Kdlevel   | Integer(11)    |             |
| 3   | Kdclass   | Integer(11)    |             |
| 4   | Kdlesson  | Integer(11)    |             |
| 5   | Kddiscussion | Integer(11)  |             |
| 6   | Contquestion | Longtext     |             |
| 7   | Selectquestiona | Varchar(256) |             |
| 8   | Selectquestionb | Varchar(256) |             |
| 9   | Selectquestionc | Varchar(256) |             |
| 10  | Selectquestiond | Varchar(256) |             |
| 11  | Selectquestione | Varchar(256) |             |
| 12  | Keyquestion | Char(1)        |             |
| 13  | Solution   | Longtext       |             |
| 14  | Owner      | Varchar(100)   |             |

Three components which include the questions, answer choices and solutions are accommodated in the MasterSoal table above. Noting that the questions that are built are choice questions and existing variables, producing values from random processes including sorting the answer choices a, b, c, d, e also randomly, so you can make sure that there are 5! Possible answers, while for the possibility of the number of questions depends on the number of variables in the question template. Suppose that the
number of variables in a question has 2 numerical variables, 1 name variable and 2 city variables, then the possible questions that can be formed are:

Possible Numeric Variables x Possible Name Variables x Possible City Variables.

If numerical variables are 1-3 digits, then the probability of the problem for the Numeric Variable is 999 x 2 = 1998. For Name Variables, there are 100 database names, so there are probably 100 different names. For City Variables, the number of city databases is 100, so the possibility of 2 different names is 100 x 99 = 9900. So the total possibility for one template can make a question as much as 1998 x 100 x 9900 = 1,978,020,000 questions.

With regard to item analysis for choice questions, based on classical test theory or so-called classical pure score theory [2] based on an additive model, the observation score is the sum of the actual score and the measurement error score. This classical theory is used as the basis for developing formulas in determining the validity and reliability of the tests used to determine the quality of the test. Another criterion that can be used to determine the quality of the test is the index of difficulty and differentiation.

Using the analysis of items, the power index value is said to be different if it is greater or equal to 0.3. If the value of the distinguishing index of an item is of little value it will cause the item cannot be used to distinguish students with high abilities and students with low abilities.

With the system built we can make the items of choice become more different, because we ourselves determine the template of choice by making traps that might be made, so that we can create quality questions from the level of difficulty and different power. If the student's ability to work on the problem exceeds the difficulty level of the item, then the student's response is expected to be correct, and if the student's ability is less than the difficulty level of the item, then the student's response is expected to be wrong [3]. Future development of the system is that this generator can be developed to be able to take advantage of the concept of meaningful learning as conveyed by David Paul Ausubel (1918-2008) is an educational psychologist from America. The concept of meaningful learning is widely used, as stated by [4] that meaningful learning is learning with clearer goals, this learning enables people involved in it to do more meaning to the world around them learning about more realistic things that are characterized by more active, deliberate, constructive, authentic and cooperative learning.

The following is a sample case of a draft question, choice and completion of the template.

Draft questions:
If Bambang travels from the city of Kertosono to the city of Lumajang, which is 5 km away, it will take 5 hours, so Bambang's speed in driving his car is ...

Template soal:
If [str]nameperson:1[/str] travels from the city of [str]tbkotadn:1[/str] to the city of [str]tbcity:2[/str], which is [eval][var]int2d:1[/var]*[var]int2d:2[/var][/eval] km away, it will take [var]int1d:1[/var] hours, so [str]nameperson:1[/str]'s speed in driving his car is...

Note:
time -> [var]int2d:1[/var]
speed -> [var]int2d:2[/var]

Template Choice:
A. distance-time
   [eval][var]int2d:1[/var]*[var]int2d:2[/var]- [var]int2d:1[/var][/eval]
B. distance – speed
   [eval][var]int2d:1[/var]*[var]int2d:2[/var]- [var]int2d:2[/var][/eval]
C. time + speed
   [eval][var|int2d:1[/var]+[var]int2d:2[/var][/eval]
D. distance-time-speed
\[ \text{E. Speed} \]
\[ \text{[var]int2d:2[/var]} \]

Settlement:
To calculate the motion speed of the car can be used formula:
\[ \text{V} = \frac{\text{Distance}}{\text{Traveling time}} \]
\[ \text{V} = \frac{\text{[var]int2d:1[/var]}}{\text{[var]int1d:1[/var]}} \]
\[ \text{V} = \text{[eval][var]int2d:1[/var]/[var]int1d:1[/var]} \]

So \text{[str]nameperson:1[/str]} travels from the city of \text{[str]tb.city:1[/str]} to the city of \text{[str]tb.city:2[/str]} which is \text{[eval][var]int2d:1[/var]*[var]int2d:2[/var]} km away, it will take \text{[var]int1d:1[/var]} hours, then the speed \text{[str]nameperson:1[/str]} in driving the car is: \text{V} = \text{[eval][var]int2d:1[/var]/[var]int1d:1[/var]} km/hours.

Program results:
If Dian Ambarwati travels from the city of Sidoarjo to the city of Bojonegoro which is 35 km away within 5 hours, the speed of Dian Ambarwati in driving her car is ...

A. 35 km/hour
B. 7 km/hour
C. 12 km/hour
D. 0.7143 km/hour
E. 2 km/hour

5. Conclusions and recommendations

5.1. Conclusions
Based on the results of the analysis, identification and discussion can get the following results:
1. A template is needed to be able to make a large number of questions.
2. The template needed includes the question template, the selected template and the completion template.
3. In the template, several variables are needed as supporting templates, namely Numerical Variables, String Variables and Image Variables.
4. To recognize a variable and other commands, an identification tag is needed, namely: Numerical [var] numeric [/ var] tag; tag string [str] collectname [/ str]; tag the image [img] filename [/ img], the tag for evaluting the equation, and the tag to display the [equ] equation [/ equ] equation.
5. The program is built using a WEB base, so a PHP program script is needed.
6. The generator system of the formed question can be accessed from anywhere and at any time.

5.2. Recommendations
1. Generating System Questions can be developed again for lessons other than physics.
2. Generating System This problem can be developed with an offline system.
3. Variables and parameters can be developed again, so they can meet broader needs.
4. The system can be expanded by building an Android-based application.

6. Funding
This research was jointly funded by Airlangga University, Mathematic and Information System Department.
7. acknowledgements

The authors would like to thank the Head of the Faculty of Science and Technology of Airlangga University.

References

[1] Adetokunbo A.A. Adenowo, Basirat A. Adenowo. 2013, Software Engineering Methodologies: A Review of the Waterfall Model and Object-Oriented Approach., International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013
[2] Allen, M. J. & Yen, W. M. 1979. Introduction to measurement theory. Monterey, (CA: Brooks/Cole Publishing Company). p 57
[3] Keeves, J.P. dan Alagumalai, S. 1999. New approaches to measurement. In Masters, G.N. dan Keeves, J.P.(Eds). Advances in measurement in educational research and assessment. (Amsterdam: Pergamon) p 24.
[4] Miles Berry. 2012 [ONLINE]: http://milesberry.net/2009/09/meaningful-learning-and-ict/
[5] Pfleeger, S.L. & Atlee, J.M. 2006. Software Engineering: Theory and Practice, 3rd Edition. US: Prentice Hall.
[6] Sommerville, Ian. 2011. Software Engineering (Rekayasa Perangkat Lunak). Jakarta: Erlangga.
[7] Winarko, Edi, 2006, Perancangan Database dengan Power Designer 6.32, (Jakarta: Prestasi Pustaka) p 1.