

# Multiple-choice questions using STACK with partial score and feedback

*Kentaro YOSHITOMI*

yositomi@omu.ac.jp

Faculty of Liberal Arts, Sciences, and Global Education

Osaka Metropolitan University

Japan

**Abstract:** *Online testing systems in mathematics education are useful tools for both teachers and students. STACK is a system that is widely used in Europe and other parts of the world and it is a question plugin for Moodle, a Learning Management System (LMS). On the other hand, the use of multiple-choice questions in mathematics has been limited due to the randomization problem. The use of multiple-choice questions with STACK solves this problem, but the method described in the help document does not support partial score and feedback, which are important features of a learning tool. In this paper, we introduce a method for providing partial score and feedback by devising a coding method that utilizes the framework of STACK's multiple-choice questions, and present a template. In addition, we will present some examples in which multiple-choice questions, together with incorrect answer patterns, are considered to be important educational resources.*

## 1. Introduction

The online test system for mathematics (hereinafter referred to as "online test") is not yet fully functional when used for evaluation purposes in an environment where cheating is possible, such as using computational software like Wolfram Alpha or consulting with friends. In this paper, we focus on the use of online tests as a learning environment or for formative evaluation purposes. There are various types of online tests, but the most important feature is a feedback function based on students' input. The feedback function is a function that not only judges whether a student's answer is correct or incorrect, but also outputs appropriate messages according to the pattern of incorrect answers (or correct answers in some cases) and is the core function of online tests together with the grading function. Randomized online test is a very effective tool in a learning environment or for formative evaluation purposes. It is quite important to develop questions that can be solved faster by students' own thinking than by cheating and questions that induce students' awareness through feedback.

In this paper, we discuss STACK as a system for such online tests. STACK, which will be described in detail in the next section, is a question type plugin for Moodle. Moodle is a Learning Management System (LMS) and is widely used worldwide. STACK is also widely used in Europe (e.g. [1]) and other parts of the world (e.g. [2]) and is becoming popular in Japan (e.g. [3],[4]). STACK is a system that internally calls Maxima<sup>1</sup> to programmatically analyze input mathematical expressions and numerical values and gives scores and feedback through a binary tree flow.

The author has been working on the development of a flipped learning environment for linear algebra in the first year of college since 2014, and has developed a video and published it on YouTube. In addition, I have been actively developing questions using STACK since around 2015.

STACK is a great system and it is a system that can be used to develop a variety of questions in calculus, statistics, group theory, elementary number theory, and so on, in addition to linear algebra.

---

<sup>1</sup>Maxima is "a fairly complete computer algebra system written in Lisp with an emphasis on symbolic computation." (cf. <https://maxima.sourceforge.io/>)

Unfortunately, however, there are two problems. It is a heavy system load, and it is a burden to students depending on the input contents. As for the first point, in a large university, the system should be sufficiently powerful. As for the second point, appropriate tutorials are needed for first-year students to use the system. In addition, in Japan, where many students do not have their own PCs, problems often occur in inputting answers using smart phones.

In order to alleviate these two difficulties, the author has been working on the use of multiple-choice questions (hereinafter referred to as MCQs). Moodle comes standard with a variety of question types, including MCQs, but unfortunately, it does not support randomization. Moodle has a question bank where questions can be stored and categorized. The questions can be exported and imported as XML files. In addition, in the quiz activity of Moodle, there is a function to provide questions randomly from a set of questions in a category. Using this function, it is possible to provide pseudo-randomized questions. Based on this idea, the author and Nagasaka have been using *Mathematica* for automatic generation (cf. [5], [6]).

However, there were some difficulties with this method. The development procedure is as follows: import the generated file containing 50 or 100 questions into Moodle, check that the questions are correct by actually solving some of them, and if there is a bug, correct it again and start over. An example of a bug is a mistake in generating incorrect or correct answers, which results in two identical options being presented as one incorrect answer and one correct answer. Even if such bugs are included in several of the 100 quizzes, it is almost impossible to correct them individually, so if the quiz has already been opened to students and some of them have already taken the quiz, the quiz must be closed to students, and students who have already started taking the quiz must take it again. In other words, the students who are willing to take the quiz are forced to identify bugs in the quiz, which is a very serious problem. In general, teachers should always be aware of the fact that the frequent existence of such bugs is a serious obstacle to the learning environment. This is also true for STACK questions, but in the case of STACK, it is often possible to fix bugs on the spot and re-evaluate the test results and the deploy function makes it relatively easy to check the questions (quize) before they are opened to students.

As a matter of fact, STACK also has a function to provide multiple-choice questions. If one follows the method described in the help document, it is relatively easy to create multiple-choice questions. However, there is a difficulty in this method that partial scores and feedback cannot be returned. This is one of the reasons why the author has been generating a large number of Moodle standard multiple-choice questions automatically. However, the author has found a way to realize a partial score and feedback in STACK and this paper describes this method. In addition, we will introduce some learning materials that are considered to be useful and discuss in which cases multiple-choice questions are considered to be effective.

## 2. STACK and MCQ

### 2.1 Why STACK?

Moodle is the LMS used by 65% in Europe and more than 42% in Japan<sup>2</sup>. STACK is a question plugin that works with Moodle. For details about STACK, please refer to the official website<sup>3</sup> and the Github site<sup>4</sup>. STACK is widely used, especially in Europe, and as outlined in the previous section, it communicates with Maxima, which runs inside or outside the server, in order to programmatically evaluate the input answers. In STACK, the syntax of Maxima can be used almost as is. Actually,

---

<sup>2</sup>For the latest status, see <https://stats.moodle.org/>

<sup>3</sup><https://www.ed.ac.uk/mathstack>

<sup>4</sup>[https://github.com/mathstack/moodle-qtype\\_stack](https://github.com/mathstack/moodle-qtype_stack)

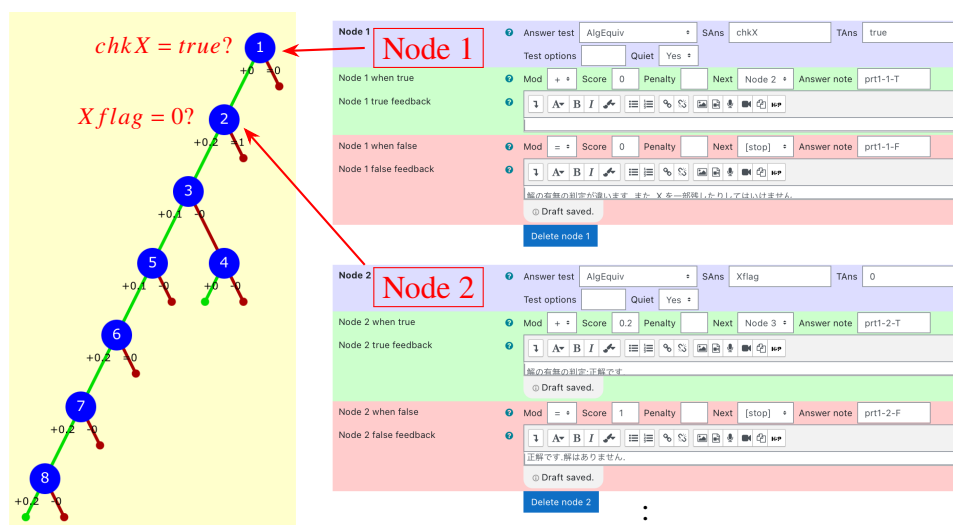


Figure 1: an example of PRT

certain functions are prohibited by STACK, and some functions are added originally by STACK. Parameters for random questions are set in the question variable at the beginning of the question edit screen. Randomization is an important and essential feature of online testing and STACK's own rand function allows one to randomly select from a list or to get a random integer value. If one has a maxima executable file which is preloaded and dumped with STACK macros, debugging work in the development of question variables can be done on the terminal. Since it is often difficult to locate error messages displayed on the STACK question edit screen, it is very efficient to debug the question on the terminal. The input is assigned to the answer variable (named ans1 by default, but the question author is free to change the name or add more variables), but it can be computed in the Feedback variables area while keeping the contents of the question variables. Using the results of the evaluation, partial score are added or subtracted according to the Potential Response Tree (PRT) algorithm and it is possible to set the feedback to be displayed for each branch. A sample PRT is shown in Fig 1.

In mathematics, it is usual that the answer is not unique. For example, in the case of a question that requires the basis and dimension of a given subspace, the basis is of course not unique and it is necessary to determine whether or not the permutation (or a set) of the input vectors is actually the basis of the subspace in question. Although not covered in this paper, there are also questions of derivatives and integrals where the equivalence with the answer can be determined only by evaluating the formulas and cannot be determined by mere matching. Thus, the online test can determine the correctness of non-unique solutions. In addition, for example, in the case of the above basis problem, if the answer is incorrect, the system provides feedback on whether the solution does not belong to a vector space, or whether one or both of the linear independence and generation conditions of the basis are not satisfied, and if the latter is the case, the system displays a message asking the user to check the definition of the basis. In addition, it is often the case (in fact, it is not well known to teachers) that the correct answer is given for the basis but the wrong answer is given for the dimension, and in such cases, by displaying a message that prompts the student to check the definition of the dimension, it is possible to correct the student's perception of the definition of the dimension and to improve his/her understanding. In this way, the function to supply feedback is necessary and effective as a learning environment. In the case of paper-based exercises or quizzes, the burden of marking is quite heavy for teachers, and students may forget what they have learned by the time they receive the marked answers back a week later. On the other hand, online tests such as STACK do not require much work for the

teachers and students can receive the evaluation and messages instantly. We believe that this kind of immediate feedback is expected to be effective in preventing students from getting stuck in their learning and deepening their understanding. The fact that STACK is a system that is used as widely as possible in the viewpoint of sharing learning materials, that it is easy to debug in development and that it provides feedback and partial score, are the reasons why we use STACK as a system for learning materials.

## 2.2 STACK and MCQ

As described in § 2.1, STACK is a very attractive and effective system, but it is a problem that the load of calling Maxima is high. For this reason, there are some institutions of higher education in the field of science and mathematics that use Moodle but do not use STACK. From the students' point of view, another difficulty is that students often make mistakes in inputting numerical values and symbols into input forms, wasting their time. Especially in countries such as Japan, where 2-byte characters are used, there are many cases where students enter numbers, letters, or symbols in 2-byte code instead of 1-byte (i.e. ascii) code and the system does not process them as if they were not inputted. This so-called 2-byte character problem is usually solved in the early stage of the semester, but there is another input load problem. For example, there are many numerical values to be input for matrices and, as is often the case in calculus, problems related to the input method, such as not knowing how to describe a function, can still exist. Handwritten input of mathematical expressions is not yet generally available, and students will be required to become familiar with the input method.

Therefore, it is considered effective to use MCQ in addition to STACK's question input method, which requires input of mathematical expressions and numerical values, depending on the learning objectives. Unfortunately, MCQ with STACK does not solve the problem of system load, but it does solve the problems of input load, development load, and ease of debugging. In this paper, we consider a MCQ that has more than one correct answer and requires the user to select all the correct answers and discuss how to implement such a question in STACK. As we have mentioned, partial score and feedback are important in this case as well.

From the experience with online tests and paper-based quizzes, it is known that there are various patterns of incorrect answers. Feedback can be considered as a set of this empirical knowledge and a set of questions in which feedback is properly implemented can be called a set of educational empirical knowledge. Thus, feedback is important from the student's point of view in terms of learning efficiency, while from the teacher's point of view in terms of accumulation of error patterns. Thus the asset of multiple choice questions can be considered as an asset of educational empirical knowledge. Therefore, it is important to have a method to provide appropriate feedback in the implementation of MCQs in STACK. In the next section, we will review how multiple-choice questions work in STACK.

## 2.3 Framework of MCQs in STACK

MCQs in STACK use checkboxes as input forms when multiple correct answers are to be selected. A checkbox is one of the features of HTML forms and is a form used for MCQs. In case only one choice is required (i.e., only one correct answer), a radio button or a drop-down form can be used. In authoring STACK questions, even if there is more than one correct answer in the answer field, it is necessary to specify one example of the correct answer when setting up the answer field. In the case of multiple-choice questions, instead of a correct answer, one should set a variable whose value is the list describing the contents of the options. The options are displayed in the order specified by this variable. If a random order is desired, the list could be rearranged randomly by using the `random_permutation` function in advance. Following the description in the document, we write this

list valued variable(called a list variable below) as ‘va’. The notation ‘va1’, ‘va2’, .... etc. are similar to the above.

Let  $c_1, c_2, \dots, c_m$  be the mathematical expressions or strings of correct options and  $w_1, w_2, \dots, w_k$  be the wrong(incorrect) options. The format of the list variables can be one of the following two forms. For the sake of simplicity, we will construct the question from two correct options ( $c_1, c_2$ ) and three wrong options ( $w_1, w_2, w_3$ ).

va1:random\_permutation([[c1,true],[c2,true],[w1,false],[w2,false],[w3,false]]);

→ e.g. [[w2,false],[c2,true],[w3,false],[w1,false],[c1,true]];

va2:random\_permutation([[1,true,c1],[2,true,c2],[3,false,w1],[4,false,w2],[5,false,w3]]);

→ e.g. [[4,false,w2],[2,true,c2],[5,false,w3],[3,false,w1],[1,true,c1]];

These lists are passed to STACK’s own function mcq\_correct as arguments to obtain a list of correct answers. The return values of va1 and va2 are denoted as Cans1 and Cans2.

Cans1:mcq\_correct(va1); Cans2:mcq\_correct(va2);

Cans1 becomes  $[c_2, c_1]$  and Cans2 becomes  $[2, 1]$ . The student’s answers are assumed to be stored in the variable ans1 (as mentioned before, the default name of the answer variable). If  $c_2, w_1$  is checked, In the case of va1, ans1 is  $[c_2, w_1]$  and in the case of va2, ans1 is  $[2, 3]$ . In the PRT, the correctness is determined by whether or not ans1 is equal to Cans1 or Cans2.

## 2.4 How to scoring of MCQ

We review the contents of the paper[7](written in Japanese). Let  $n_C$  be the number of correct options,  $n_W$  be the number of incorrect options, and  $N = n_C + n_W$ . As mentioned in the previous section, we suppose that  $N$  is fixed and  $1 \leq n_C < N$  is random. We consider the case that the examinee choose  $x_C$  correct options and  $x_W$  incorrect options. Let  $c = x_C + (n_W - x_W)$ , that is the number of correct choices.  $c = N$  holds if and only if the choices are the completely correct. As described in the paper, there are following methods to mark partial score.

- MR(Multiple response)-method :  $s_{MR} = \begin{cases} 1 & (x_C = n_C, x_W = 0) \\ 0 & (\text{otherwise}) \end{cases} \in \{0, 1\}$
- Count- $n$  method:  $s_{Cn} = \begin{cases} 1 & (c = N) \\ 0.5 & (c = N - 1) \\ 0 & (\text{otherwise}) \end{cases} \in \{0, 0.5, 1\}$
- MTF(Multiple-true-false) method:  $S_M = \frac{c}{N} = \frac{x_C + (n_W - x_W)}{N} \in [0, 1]$
- CFG(correction for guessing) method:  $S_{CFG} = \frac{c - (N - c)}{N} = 2S_M - 1 \in [-1, 1]$
- NM(negative-marking) method:  $S_N = \frac{x_C - x_W}{n_C} \in [-\frac{n_W}{n_C}, 1]$
- $\phi$ -coefficient method:  $s_\phi = \frac{x_C(n_W - x_W) - (n_C - x_C)x_W}{\sqrt{n_C n_W (x_C + x_W)(N - (x_C + x_W))}} \in [-1, 1]$
- Jaccard method:  $s_J = \frac{x_C}{n_C + x_W} \in [0, 1]$

In the paper[7], the expected value of the score for a solution by guesswork is computed for  $N = 8, n_C = 3$ . Comparing the methods to convert negative scores to zero, the  $\phi$ -coefficient method is

the lowest at 0.156, followed by the Jaccard-coefficient method at 0.285 and the NM method at 0.061. The paper also examines each method from the viewpoint that partial scores should not be given in both cases of selecting all choices and not selecting any choice, and that negative scores are usually not given. The methods of converting scores to the range of  $[0,1]$  are also discussed, and it is concluded that the Jaccard-coefficient method is the most appropriate. Note that the standard MCQs in Moodle only allow you to add  $\pm \frac{1}{n} (n \in \mathbb{N})$  to the score for each correct or incorrect answer, so that none of the above scoring methods can be realized. If you want to score 0 when all the choices are checked, you have to compute  $\frac{x_C}{n_C} - \frac{x_W}{n_W} \in [-1, 1]$ . In this case, however, the standard MCQs of Moodle can not convert the scores and negative scores may result. In this paper, we denote by  $s$  the score obtained by one of the above methods, adjusted so that  $s \in [0, 1]$ .

### 3. Improved coding of MCQ with STACK

#### 3.1 Question variables

Suppose that the score  $s \in [0, 1]$  in a MCQ has been determined as described in §2.4. In STACK, a comparison between the list of answers and the list of correct answers generally yields either a full score or a zero score, but here we consider a method to give a partial score. This method can also provide feedback.

The value of the variable `va`, which is set as an example of the correct answer to the MCQ answer variable `ans1`, was a list of options. For our goal, we use the second method (in the form of variable `va2`) of the methods described in §2.3. That is, in the following form:

“ `[[num1,true/false,option1],[num2,true/false,option2],...]`”

where the integers `num1,num2,...`, which are the first elements of each list constituting `va2` are defined as follows.

Put `va2[i]= [ni, ti, si]`. If  $s_i$  is correct (equivalently  $t_i = \text{true}$ ), then  $n_i \in \mathbb{Z}, n_i < 100$ . In the case that there are some patterns among the correct answers, e.g.  $s_i$  is correct of pattern  $k = 1, 2, \dots$ , then  $n_i \in \mathbb{Z}$  and  $n_i \in (10(k-1), 10k)$ , that is separated by 10. If  $s_i$  is incorrect ( $t_i = \text{false}$ ) of pattern  $k = 1, 2, \dots$ , then  $n_i \in (100k, 100(k+1))$ . The concrete description is as follows.

`va2=[[1,true,1st option of correct pattern1],[2,true,2nd option of correct pattern1],...,[11,true,1st option of correct patterns2],...,[21,true,1st option of correct patterns3],...,[101,false,1st option of incorrect pattern1],...,[201,false,1st option of incorrect pattern2],...]`

The code of question variables of Example 1 (cf. § 4.2 ) is available at: [http://www.las.osakafu-u.ac.jp/~yositomi/ATCM2022/Example1\\_qvar.txt](http://www.las.osakafu-u.ac.jp/~yositomi/ATCM2022/Example1_qvar.txt)

#### 3.2 Feedback variables and PRT

As explained in §2.1, feedback variables can be computed as a preprocessing of the PRT. Here, the contents of the question variables are also preserved.

Suppose that `ans1= [103, 2, 11, 201, 1, 102]`. From this list, we see that the student has selected two correct options for pattern 1 ( $2, 1 \in \text{ans1}$ ), one correct option for pattern 2 ( $11 \in \text{ans1}$ ), two options for incorrect answer pattern 1 ( $103, 102 \in \text{ans1}$ ), and one option for incorrect answer pattern 2 ( $201 \in \text{ans1}$ ). These can be obtained computationally by a list operation (Fig 2). In fact, for each number  $x$  in the list, if the integer part (floor) of  $x/100$  is 0, it is a correct answer, and if it is an integer

$n$  greater than 1, then it is an incorrect answer pattern  $n$ . Furthermore, in the case of a correct answer, we can also determine the pattern of the correct answer by the integer part of  $x/10$ . If there is a pattern in the correct answer options, the assumption that the number of options for each pattern is less than 10 at most is a realistic assumption. In the examples in this paper, the map function is used to compute the list variable ansL for correct or incorrect answers and the list variable ansL2 for patterns in the case of correct answers. If ansL matches a list of only 0, we know that the student has chosen only the correct answer, and if ansL matches a list of the length of the number of correct answers consisting only of 0, we know that the student's answer is completely correct. If ansL contains 1 or 2 (it can be checked by using the member function), then it contains incorrect answer options of patterns 1 and 2.

```

ansL:=map(lambda(u,floor(u/100)),ans1);/* Incorrect pattern list 100,101,...,200,201,...,301,302,... */
ansL2:=map(lambda(u,floor(u/10)),ans1); /* correct pattern 1,2,...,11,12,...,21,22,... <100 */

alen:=length(ans1);
Calen:=cn;
CansL:=makelist(0,1,Calen);
count0(L):=block([ct,i],ct:0,for i:1 thru alen while i <= alen do if L[i] = 0 then ct:ct+1,return(ct));

ct0:=count0(ansL);
sc0:=max(0,ct0/cn-(alen-ct0)/wn); /* original */
scL:=ct0/(cn+ct1);
sc:=scL;
/* scL : the list for partial points with PRT */
scL:=makelist((is(sc>=0.1*j)),1,10);
fig0:=is(ct0=cn); /* All correct options are checked ? */
/* fig1,fig2,... : flag to detect patterns of checked incorrect options */
figL:=makelist(member(k,ansL),k,1,9);
figCL:=makelist(member(k,ansL2),k,0,9);

```

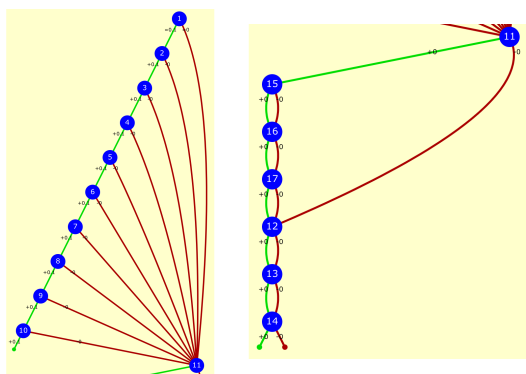


Figure 2: Partial score and input patterns      Figure 3: PRT Flow:Scoring(left)/Feedback(Right)

On the other hand, for the score  $s$ , unfortunately, only concrete numbers can be described in the score description field in the current PRT of STACK and a variable cannot be described. Therefore we define an array of boolean values scL of length 10 as follows:

$$scL[j] = \begin{cases} \text{true} & (s \geq 0.1j) \\ \text{false} & (\text{otherwise}) \end{cases} \quad (j = 1, 2, \dots, 10)$$

Using this array, partial score can be realized by approximating it in units of 0.1 points(Fig 3).

In the PRT, after processing the score, feedback is returned based on the information about which incorrect answers were selected, obtained by the member function (Fig 2). In the case of the PRT shown in this figure, no individual feedback is shown for the correct choices, but it is possible to change the node connections so that individual feedback is shown for the correct ones. If the correct answer also has a pattern, it is possible to display information about the pattern as a confirmation for the correct choices and to give a hint about the correct patterns that was not selected.

## 4. Application

### 4.1 Effectiveness of MCQ

There is a view that MCQ, which requires the selection of all the correct options, is equivalent to multiple-true-false questions. However, the author believes that multiple-true-false questions are completely different in the sense that the context of the question is not the same as that of random parameters and thus it is not possible to include various incorrect answer patterns for a single context. The educational effects of using such MCQs in university mathematics are considered as follows.

- Accurate understanding of definitions is obtained through repetition.

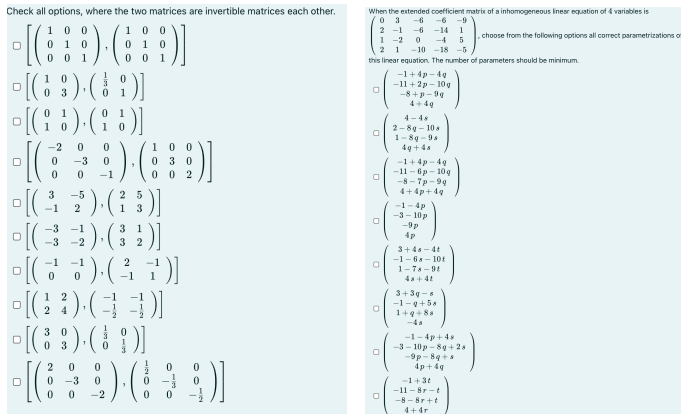


Figure 4: Preview of Example1/Example2



Figure 5: Question variables (Excerpt)

- Awareness of basic properties is raised through repetition.
- Deepening of thinking, since it is difficult to answer correctly by guesswork.

In the following subsections, examples of actual learning tasks that are expected to have the above effects will be presented, along with specific incorrect answer patterns and feedback.

## 4.2 Example1: Definition of Inverse Matrix

The first example is a question to check the definition of an inverse matrix with its basic properties, as shown in the left-hand figure of Fig 4. The basic property is, for example, that a matrix is not regular if there is a constant-multiplicative relationship between two rows or columns, especially if the two rows or columns are coincident. Or the same is true if the rows or columns are zero-vectors. This is intended to further deepen the students' understanding of the definition of the product of matrices, while keeping in mind the concept of rank to be studied later. We believe that the students will be able to grasp the concept of inverse matrices in their consciousness by solving such questions repeatedly.

In the second semester linear algebra curriculum, students learn about diagonalization, which is described as a representation matrix of bases. The matrix formed by the basis (which we call diagonalizing matrix) is, of course, regular. If the square matrix in question is not diagonalizable, then there is no diagonalizing matrix. However, there are a few students who construct the diagonalizing matrix using the same eigenvectors or including a 0 vector in place of the missing eigenvector. The purpose of the inverse matrix question above is to prevent such a situation as much as possible. We believe that this will also deepen the students' understanding of diagonalizability.<sup>5</sup>

Fig 5 shows an overview of question variables<sup>6</sup>.

## 4.3 Example2: Parametrization of the solution of linear equations

The right-hand figure of Fig 4 is a sample of a question in which students are asked to choose all the correct parametric solutions of a simultaneous linear equations. The aim of this question is to

<sup>5</sup>In fact, since there is a considerable time gap between the two learning periods, we have not yet been able to verify whether this is effective or not, but we would like to do so in the future.

<sup>6</sup>The full codes is available at <http://www.las.osakafu-u.ac.jp/~yositomi/ATCM2022/Example1.xml>



make the students understand the importance of checking whether or not the parameter expressions of the solutions of the equations they have found are correct. Most of the time, students obtain the parameter representation from the row reduced echelonform in the specified way, and they often obtain it correctly, but sometimes they make mistakes in obtaining it. A more common case is to compute the elementary transformation incorrectly and thus it is important to check whether the obtained solution is correct as a solution of the original equation or not.

In addition, there are the following questions that require a more advanced understanding. The simultaneous linear equations themselves are not given, but when one of the parametric expressions of the solution is given as  $\mathbf{x} = \mathbf{x}_0 + p\mathbf{u} + q\mathbf{v}$ , the question is to select all other correct solutions. The aim of this question is to understand how  $\mathbf{x}_1$  and  $\mathbf{a}, \mathbf{b}$  can be obtained from the parameter representation of the given solutions for the other solution  $\mathbf{x} = \mathbf{x}_1 + p\mathbf{a} + q\mathbf{b}$ . That is, given some  $p_0, q_0 \in \mathbb{R}$ ,  $\mathbf{x}_1 = \mathbf{x}_0 + p_0\mathbf{u} + q_0\mathbf{v}$  and a regular square matrix  $M = (m_{ij})$  of order 2,  $(\mathbf{u}' \ \mathbf{v}') = (\mathbf{u} \ \mathbf{v})M$ . In fact, at this stage, the concept of linearly independence has not yet been taught, so a rigorous understanding of the latter equation is a little difficult, but it is expected that a certain level of understanding can be obtained in the case of up to two parameters.

Such an understanding of the meaning of the parameter representation of the solution of simultaneous linear equations is a basic knowledge necessary for future use of computational software (*Mathematica*, Wolfram Alpha, Maple, etc.) to solve the equations. Human errors, such as input errors, can occur anywhere and it is necessary to understand how to check the computation when accuracy is required. In addition, if the parameter representation of the computed solution is not necessarily optimal for practical use, it may be necessary to find another parameter representation. We believe that MCQs, as described above, are those that facilitate this kind of understanding. If students are only asked to actually input their answers, they are not likely to be aware of the properties of the solutions. Therefore, MCQ may be effective as a question format in such a case.

## 5. Conclusion and Future work

As described in the help document of STACK, partial scores and feedback are not available for multiple-choice questions in STACK. However, it is found that by appropriately utilizing the framework of multiple-choice questions in STACK, and by making some coding efforts, it is possible to provide partial scores and feedback. In this paper, we show and describe such coding of STACK as a template.

By using a string array with feedback messages corresponding to incorrect answer patterns and, in some cases, correct answer patterns, the above template can be used almost as is for the PRT part. Note, however, that in the case of multi-byte codes such as Japanese, the messages are not displayed correctly, so it is necessary to write the messages directly in the PRT nodes.

One of our future works is to use the STACK code for mass generation, as the authors have done with *Mathematica*, so that such MCQs can be used even in Moodle where STACK is not installed. This is because all the necessary data are generated in the question variables of STACK and it should be theoretically possible to do so. Since STACK still has the problem of being a bit heavy as a system, this kind of usage will also reduce the system load. Development of a method to automatically generate MCQs from the code of questions that are confirmed to be bug-free in STACK is a future task.

The author's basic belief in educational research is that educational resources should be shared globally. As mentioned at the beginning of this paper, incorrect answer patterns of MCQs are an educational assets and the developed STACK questions, including MCQs, will be a useful resource for both teachers and students. Although the system will continue to mutate in various ways, it is desirable that the incorrect answer patterns and feedback, which are the essence of such a question resource, be shared, including general questions other than MCQs. The author proposes to share such question data as an educational resource as a mathematics e-Learning question specification (MeLQS[8]). It

is expected that MeLQS data will be widely shared and become a valuable asset for mathematics education.

## 6. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP21H00921 that is the research project “Development and application of online mathematics questions which require deep thought to solve even with cheating” by K. Yoshitomi with K. Nagasaka, M. Kameda, Y. Ichikawa, S. Higuchi and K. Kanenishi.

## References

- [1] A. Rasila. (2016). *E-assessment material bank abacus*. In Proceedings of EDILEARN16, 8th Annual International Conference on Education and New Learning Technologies
- [2] Ichikawa, Y. (2021). *Case study for STACK in Thai KOSEN*. Symposium on the case study of math online test systems. <https://stack.phys.cs.i.nagoya-u.ac.jp/kaken2021/symposium/symposium01.html>
- [3] Aiba, D., Yamada, T., et.al, (2017). *Possibilities for using STACK in mathematics education at the National Institute of Technology*, National Institute of Technology, Fukui College.
- [4] Ichikawa, Y. (2017). *Practical Report on Utilization of STACK*. RIMS Kokyuroku, Vol.2022.
- [5] Yoshitomi, K., (2018). *Generation of Abundant Multi-choice or STACK Type Questions Using CAS for Random Assignments*. Mathematical Software ICMS 2018, 492-497.
- [6] Nagasaka, K., (2020). *Multiple-choice questions in Mathematics: automatic generation, revisited*. Electronic Proceedings of the 25th Asian Technology Conference in Mathematics, 21785-1-15
- [7] Arai, S., Miyano, H., (2017). *Scoring Method for “Select All the Choices That are True” Items*, 6(1), 101-112, Bulletin of Data Analysis of Japanese Classification Society(in Japanese).
- [8] Nakamura, Y., Yoshitomi, K., et.al, (2018). *Effective Use of Math E-Learning with Questions Specification: International Trends in Research and Development*, 133-148, ICME-13 Monographs, Chapter 8.