PAPER

A Support System for Solving Problems of Two-Triangle Congruence Using 'Backward Chaining'*

Ryosuke ONDA†, Yuki HIRAI††, Kay PENNY†††, Bipin INDURKHYA††††, *Nonmembers*, *and* Keiichi KANEKO††a), *Member*

SUMMARY We developed a system called DELTA that supports the students' use of backward chaining (BC) to prove the congruence of two triangles. DELTA is designed as an interactive learning environment and supports the use of BC by providing hints and a function to automatically check the proofs inputted by the students. DELTA also has coloring, marking, and highlighting functions to support students' attempts to prove the congruence of two triangles. We evaluated the efficacy of DELTA with 36 students in the second grade of a junior high school in Japan. We found that (1) the mean number of problems, which the experimental group (EG) completely solved, was statistically higher than that of the control group on the post-test; (2) the EG effectively used the BC strategy to solve problems; and (3) the students' attempt to use both the forward chaining strategy and the BC strategy led to solving the problems completely.

key words: *secondary education, interactive learning environments, learning support system, problem-solving, backward chaining*

1. Introduction

Japanese students begin to practice explaining their thoughts to others in primary school, but most students do not practice generating logical proofs based on concrete reasoning until junior high school. Generating logical proofs is one of the most difficult tasks for junior high students [2]. In fact, according to the National Assessment of Academic Ability conducted by Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT), only about half of the students can correctly solve problems related to deductive proofs. For example, according to [3], only 49.2% of the responses to a proposition-based proof problem in a 2008 test were correct, and the non-response rate was 36.1%. Similarly, only 48.2% of the responses to a triangle congruence proof problem in a 2010 test were correct. MEXT identified the need for improved lecture methods to address the concepts for which the correct response rate on the tests was lower than 70%. MEXT also pointed out that the students are not able to correctly design deductive proofs. According to [4], MEXT suggested that students should practice making outlines of proof and connecting the preconditions and the conclusions of propositions in school. These needs and suggestions have been reiterated in later MEXT reports [5]–[7].

Forward chaining (FC) and backward chaining (BC) are processes that can be used to ascertain whether a given conclusion follows from a given set of preconditions. FC is a deductive process in which one generates inferences from the given preconditions, and further inferences from those inferences until the desired conclusion is derived. In BC, on the other hand, one starts with the given conclusion and generates pre-conditions that are required to be true in order for the conclusion to follow. Then these conditions are considered to be the conclusions and further pre-conditions are generated that would support these conclusions. The process is repeated until the necessary pre-conditions are found to be in the given set, at which point the original conclusion is considered to be proven. Both FC and BC have been used to develop intelligent tutoring systems [8] and to automatically prove mathematical theorems [9]–[11].

A typical human approach to generating a propositional proof using BC is to find sub-goals, which are preconditions of the conclusion [12], [13]. In [12] and [13], it was revealed that (i) all students were familiar with the FC strategy, (ii) students who solved problems well used the BC strategy, and (iii) most students who could not solve problems did not use the BC strategy. In addition, they found that (iv) a lack of BC strategy led to the inability to solve problems. The authors of [12] and [13] defined 'to use FC or BC for problem-solving' as 'to use the FC or BC strategy'. We use the same definition in this paper.

According to [4], it is important for students learning to generate proofs to consider the following factors: (a) What conditions are missing to derive the conclusion of a proposition? (b) What can be revealed from some preconditions? and (c) What is missing to connect (a) and (b)? The third factor (c) involves using the BC strategy. In [4], MEXT expects that teachers provide their students opportunities to practice considering (a) – (c) during classroom lessons. However, the classroom lecture format limits interaction between a teacher and each individual student because of time limitation. Hence, if students practice solving problems by following the above (a) – (c) outside the classroom lecture, they can learn more effectively.

Manuscript received August 30, 2016.
Manuscript revised April 24, 2017.
Manuscript publicized July 7, 2017.

†The author is with Faculty of Engineering, Tokyo University of Agriculture and Technology, Koganei-shi, 184–8588 Japan.

††The authors are with Institute of Engineering, Tokyo University of Agriculture and Technology, Koganei-shi, 184–8588 Japan.

†††The author is with The Business School, Edinburgh Napier University, Edinburgh, United Kingdom.

††††The author is with Institute of Philosophy, Jagiellonian University, Cracow, Poland.

*This paper is an extended version based on a conference paper [1].

a) E-mail: k1kaneko@cc.tuat.ac.jp (Corresponding author)

DOI: 10.1587/transinf.2016EDP7363

According to the curriculum provided by MEXT, the target age for learning to solve two-triangle congruence proof problems (TTCPPs) is junior high school. Several systems have been developed to support students in solving TTCPPs [14]–[17]. However, some of them do not support the students' self-learning, and the others support self-learning but do not support the use of BC strategy. Furthrmore, as far as we know, there is no existing system that supports the use of the BC strategy in solving TTCPPs.

Our goal in this research is to develop a self-learning system that supports junior high school students' use of the BC strategy to solve TTCPPs. We chose to name this system DELTA (Dialogue Environment for Learning Triangle Affairs). We also aim to measure the efficacy of DELTA's support for using the BC strategy. The DELTA system is standalone and works on a personal computer equipped with a web browser. Details are explained in Sect. 3.3.

The rest of this paper is structured as follows: In Sect. 2, related work and systems are introduced. DELTA and its behavior are explained in Sect. 3. The measurement experiment using DELTA and a baseline system is described in Sect. 4. Section 5 discusses the efficacy of DELTA's support. Finally, we present our conclusions and suggestions for future work in Sect. 6.

2. Related Work

Itoh [15] developed a system to support learners who are not good at solving TTCPPs. This system can log the steps by which learners solve the problems. However, it does not offer any appropriate support to learners who are not able to make any progress towards solving TTCPPs. This kind of support is critical in any self-learning system.

The Netlessonlab website [17] allows learners to study junior-high-school-level mathematics and other subjects in Japan. Learners can study the teaching materials by answering fill-in-the-blank problems. If a learner makes an error, some feedback is provided immediately. However, the website does not provide concrete support, which is the information needed to answer the problems. Hence, there is a possibility that the learners will not be able to make any progress towards answering the problems.

Based on the theory proposed by Wood et al. [18], Matsuda and VanLehn [16] developed a tool called Advanced Geometry Tutor (AGT). AGT applies scaffolding strategies, which adapt the tool to the learner's competence. In other words, the tool decreases support when the learner succeeds and increases support when the learner fails. Matsuda and VanLehn concluded that the FC strategy was more effective for solving proof problems than the BC strategy. Though they pointed out that the inferiority of the BC strategy is caused by the difficulty of setting sub-goals with the BC strategy, but AGT does not have any function to support the setting of sub-goals.

In the tool 'REasoning COntrol Matrix for the Proving Process (RECOMPP)' [14], learners can solve TTCPPs by filling their thoughts in the blanks in 'Sect. 1' to 'Sect. 6'. The targets of the authors' evaluation experiment were 15-year-old students. Their research concluded that the students who used RECOMPP developed better formal proof-writing skills than those who did not. According to [14], the learners' thought organization led to their developing better proof-writing skills. However, RECOMPP does not provide adequate BC support though it gives outlines of proof.

Miyazaki and colleagues [19], [20] have developed a web-based learning system based on 'conceptions of congruency' including the perceptual conception of congruency (PERC), the measure-preserving conception of congruency (MeaP), the correspondence conception of congruency (CORR), and the transformation conception of congruency (TRANS). They concluded that their web-based proof system might be used effectively during the introductory stage of proof learning because the tasks provided by the system are similarly designed to help learners to bridge between PERC or MeaP and CORR. In addition, they are following the MEXT suggestions like us. However, the system does not support the use of BC strategy explicitly.

As mentioned above, all the existing systems have some shortcomings for supporting self-learning. Students do not currently have access to systems or tools for acquiring the BC strategy needed to follow MEXT's suggestions for learning to solve TTCPPs. In our research, we developed a support system called DELTA to help students solve TTCPPs outside classroom lessons.

3. Design and Implementation of DELTA

We have developed DELTA as an interactive learning environment to give learning opportunities outside classroom lessons, which is an alternative to the teacher-based learning [21].

3.1 Content to be Learned

To solve TTCPPs, students need to understand the proof procedure and acquire the knowledge related to 'reasoning'. In this context, reasoning refers to applying the definitions and theorems in geometry that students have been studying since primary school. Since TTCPPs are studied in units during the second grade of junior high school in Japan, the scope of DELTA is restricted to the units studied prior to this grade (e.g., our system does not support units related to circles and auxiliary lines, which are studied during the third grade of junior high school in Japan).

There are several proof styles for solving TTCPPs. In our research, we follow the style shown in Fig. 1 because it is adopted in many textbooks in Japan. According to [16] and [22], a proof consists of a list of pairs, each of which is composed of an equation and its reasoning. In fact, each of rows 2 to 7 in Fig. 1 shows a pair of an equation and its reasoning step. In this paper, we define the columns for equations and reasoning steps as the 'equation column' and the 'reasoning column', respectively. We also define a pair consisting of an equation and its reasoning step as a proof

	Reasoning Column	Equation Column	
1	In △ABC and △EDC		
2	From one of given preconditions,	BC = DC	…(a)
3	From one of given preconditions,	AC = EC	…(b)
4	Because of vertical angles,	∠ACB = ∠ECD	…(c)
5	From (a) to (c), because of SAS,	△ABC ≡ △EDC	
6	Because of corresponding angles,	∠ABC = ∠EDC	
7	Because alternative angles are equal,	AB // DE	

Fig. 1 A proof-writing style for TTCPPs. Note that SAS means "side, angle, side".

Table 1 Proof-writing patterns in TTCPPs and their examples.

Pattern	Example of proof-writing	# of proof steps
A	In ΔABC and ΔADC, From one of given preconditions, AB = AD· · ·(a) From one of given preconditions, BC = DC· · ·(b) Because of common segments, AC = AC· · ·(c) From (a) to (c), because of SSS, ΔABC≡ ΔADC	4
B	In ΔABC and ΔDCB, From one of given preconditions, AB = DC· · ·(a) From one of given preconditions, ∠ABC = ∠DCB· · ·(b) Because of common segments, BC = CB· · ·(c) From (a) to (c), because of SAS, ΔABC≡ ΔDCB Because of corresponding segments of two congruence triangles, AC = DB	5
C	In ΔAOD and ΔBOC, From one of given preconditions, AO = BO· · ·(a) From one of given preconditions, DO = CO· · ·(b) Because of vertical angles, ∠AOD = ∠BOC· · ·(c) From (a) to (c), because of SAS, ΔAOD ≡ ΔBOC Because of corresponding angles of two congruence triangles, ∠ADO = ∠BCO Because alternative angles are equal, AD // CB	6

Note: SAS = "side, angle, side", SSS = "side, side, side"

step.

As shown below, TTCPPs are classified into three patterns (A) – (C) based on the number of proof steps needed to solve them. The number of proof steps depends on a given problem:

(A) Prove the congruence of two triangles: 4 steps
(B) Prove equality of segments or angles after proving by the pattern (A): 5 steps
(C) Prove a geometric property after proving by the pattern (B): 6 steps

The TTCPPs on which we are focusing in our research fit into these patterns. Table 1 shows examples of these proof-writing patterns. For example, the problem shown in Fig. 1, namely to prove the parallelism of two lines, is classified as a pattern-(C) problem.

3.2 System Design

3.2.1 Support of Backward Chaining

On the basis of findings mentioned in [16], we considered that a lack of adequate BC support leads to the difficulty of setting sub-goals, and consequently to the failure in solving TTCPPs. Hence, specific instruction and support for using the BC strategy are needed. In our design, when DELTA deems that a student is not making any progress towards solving the problem, it encourages the student to use the BC strategy. The detailed design of this support function is explained in Sect. 3.3.4.

3.2.2 Towards More Flexible Proof-Writing

MEXT suggested that it is important for students to design proofs in their own words to master the art of proof writing [23]. Therefore, we considered it to be important that DELTA encourages students to design proofs in their own words. Hence, we introduced free-writing forms for solving TTCPPs in DELTA.

While students solve TTCPPs, the order of proof steps may change (e.g., the rows 2 and 3 shown in the pattern (A) of Table 1 are exchangeable.), the names of triangle may also change (e.g., ΔBAC instead of ΔABC), and so on. DELTA should judge whether the students' proof is correct or not even if such changes occur. Hence, we designed DELTA such that it accepts various proof-writing styles. For example, students can input each of the rows 2 to 4 in Fig. 1 in random order and they can use ΔBAC instead of ΔABC at row 1 in Fig. 1. DELTA judges all these inputs to be correct.

Thus, it is necessary to allow students to change the order of proof steps and the expressions of triangles/segments/angles. However, in Japanese mathematics education, the latter change is sometimes not allowed so that the relationship of the corresponding segments or angles can be clarified. Therefore, there are some restrictions in the students' input to DELTA, as explained in Sect. 3.3.3.

In our initial design, students were to only use computer keyboards, but because second-grade students in junior high schools are not always used to keyboards, we also adopted mouse input to make DELTA easy to use. In our final design, students can input with keyboards and mice.

3.2.3 Other Necessary Functions for Solving TTCPPs

When students solve TTCPPs on paper, they usually mark or color equal segments, equal angles, parallel lines, two triangles to be proved congruent, and so on. Some of the existing systems provide TTCPPs that include these marks and/or colors [15], while others do not have this feature [16], [17]. In some systems, no figures are provided in solving TTCPPs [14], requiring students to draw figures on their own piece of paper. DELTA includes features that enable students to place marks and put colors. The details are explained in Sect. 3.3.2.

3.3 System Implementation

3.3.1 System Overview

We assumed that students would use DELTA outside class-

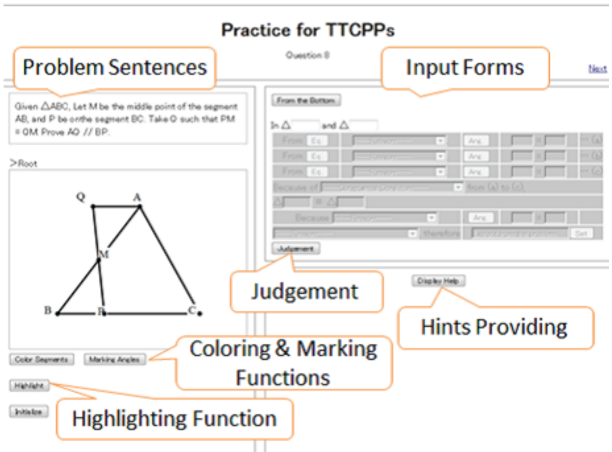

Fig. 2 Architecture of DELTA.

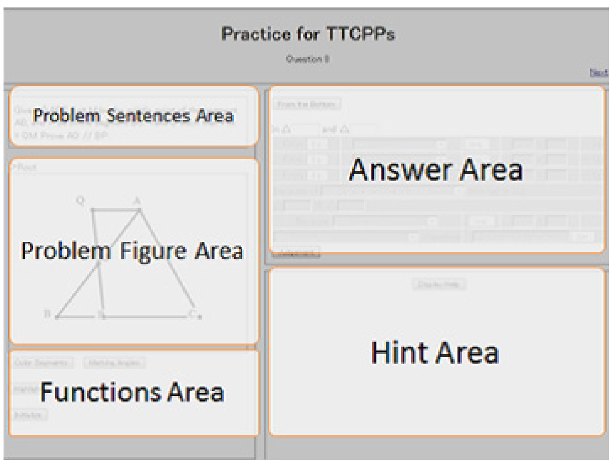

Fig. 3 Screen layout (area names) of DELTA.

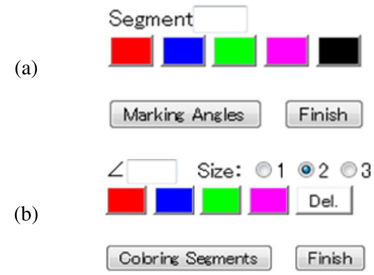

Fig. 4 Coloring and marking functions. (a) Interface for segment coloring. (b) Interface for angle marking.

Fig. 5 Interface for highlighting triangle(s).

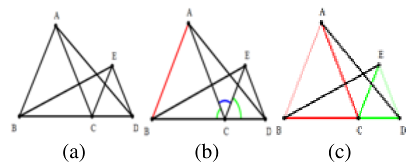

Fig. 6 Use of coloring and marking functions (from (a) to (b)) or highlighting function (from (a) to (c)).

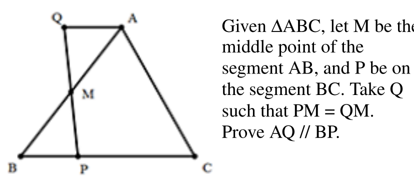

Fig. 7 An example problem.

room lessons, such as in their homes. Therefore, assuming that our target users are accustomed to using Web browsers, we implemented DELTA using HTML5 and JavaScript. Note that DELTA is a stand-alone system and it can be run without Internet connection. Figures 2 and 3 show the architecture of DELTA and the screen layout. The role of each area on the screen is as follows:

- Problem Sentences Area: A problem is provided.
- Problem Figure Area: A figure is provided. When students use support (= coloring, marking, or highlighting) functions (see Sect. 3.3.2), the results appear in this area.
- Functions Area: Support functions are aggregated as buttons here. The 'Initialize' button returns to the initial settings.
- Answer Area: Students can write proofs here.
- Hint Area: Some hints to solve TTCPPs are provided here.

3.3.2 Coloring, Marking, and Highlighting Functions

If a student presses the 'Coloring Segments' button provided in the Functions Area, the interface shown in Fig. 4 (a) is displayed. If the student presses the 'Marking Angles' button, the interface switches to that shown in Fig. 4 (b). Using these interfaces, the student can place some marks and/or put some colors on the figure provided in the Problem Figure Area. If a student presses the 'Highlighting' button in the Functions Area, the interface shown in Fig. 5 is displayed. The student can then put colors on each segment of the specified triangle(s) in the given figure.

Figure 6 shows the results after using the coloring and marking functions (from Figs. 6 (a) to 6 (b)), or the highlighting function (from Figs. 6 (a) to 6 (c)).

3.3.3 Using DELTA

We explain here how DELTA is used based on the example problem shown in Fig. 7. When a student starts DELTA, he/she will see the given problem sentences and figure (something similar to the left side of Fig. 2). He/she can start to construct a proof by filling in the input forms (Fig. 8) in the Answer Area one by one. These forms are provided based on the patterns (A), (B), and (C).

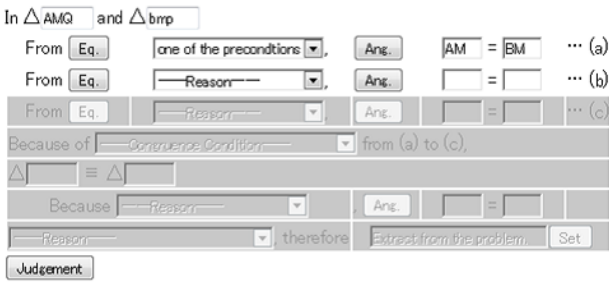

Fig. 8 Proof-writing forms for the problem shown in Fig. 7.

Fig. 9 Form to input angles in equation column.

Fig. 10 Form to input equations in reasoning column.

In the first row of the proof-writing form, the student decides which two triangles are to be proved. For example, if the student focuses on ΔAMQ and ΔBMP, he/she inputs ‘AMQ’ and ‘BMP’ into the form. If DELTA judges that the student’s input is correct, it will make the next input form available. The following forms behave in the same way as this first form. In judging whether a student’s input is correct or not, DELTA does not distinguish between uppercase and lowercase letters. In addition, even if the student inputs ‘MQA’ and/or ‘PBM’ in the above example, DELTA judges that these inputs are correct. (The following forms behave in the same way.) However, as for forms of the equation column, the student needs to input corresponding segments and/or angles (that are involved in the triangle which the student input into the left side of first row) into the left side of each equation form. For example, in Fig. 8, the student needs to input segments and/or angles related to ΔAMQ in the left side of each equation form. We expect that this limitation will encourage students to write proofs that other people can read easily.

In the second, third, and fourth rows, the student inputs three equations as reasoning steps for the congruence condition of the two triangles input by him/her. If the student presses a ‘Seg.’/‘Ang.’ button, he/she can input segments/angles (Fig. 9). Also, if the student presses a ‘St.’/‘Eq.’ button, he/she can input a statement/equations (Fig. 10). To prevent input mistakes, students can only type two letters in the segment input and three letters in the angle input. To facilitate DELTA’s automatic checking, we adopted pulldown forms for the reasoning input.

In the fifth row, the student inputs the congruence condition of two triangles using a pulldown form. In the sixth row, the student inputs two congruent triangles. If the student needs to prove the equality of segments or angles by using two congruent triangles, she/he inputs them and the reasoning steps in the seventh row. If the student needs to prove a geometric property by using segments or angles input in the seventh row, he/she inputs the property and the reasoning steps in the eighth row.

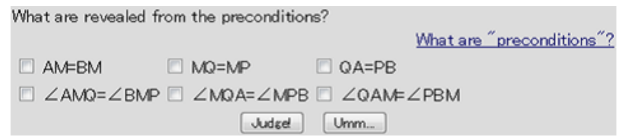

Fig. 11 Example of a hint.

Table 2 Hints for FC support on problem shown in Fig. 7.

No.	Hint	Space to Answer
1	Let us follow the conditions shown in the problem sentences one by one. Where is the segment AM equal to?	AM = [ ]
2	Where is the segment PM equal to?	PM = [ ]
3	To prove AQ // BP, what should we prove?	∠[ ] = ∠[ ] or ∠[ ] = ∠[ ]
-	We have finished following the conditions. Let us find equal segments/angles to be proved. If necessary, let us use the highlight function and/or your textbook.	

Note that a squared bracket ‘[ ]’ represents a textbox.

3.3.4 Provision of Hints and Support of FC and BC Strategies

When the student cannot make any progress to solve a given problem, he/she can obtain some hints by pressing the ‘Display Help’ button in the Hint Area. As shown in Fig. 11, each hint is provided by using a hint box, which consists of a question sentence and an array of checkboxes, textboxes, or radio buttons. The student inputs answers to the question in the hint box, then presses the ‘Judge!’ button. DELTA judges whether the input answers are correct or not. If they are correct, the student can obtain the next hint. Otherwise, DELTA gives feedback as a dialogue and urges the student to correct the answer. If the student cannot respond, DELTA encourages him/her to use the coloring function, the marking function, the highlighting function, and/or textbooks. DELTA provides some hints to support students’ use of FC and/or BC strategies.

In supporting the use of FC strategy, DELTA provides some hints so that the student can fill in the proof-writing form from top to bottom. Table 2 shows an example set of hints for the use of FC strategy. At first, DELTA provides a hint so that the student can find the given preconditions and input them into the proof-writing forms. Then, DELTA encourages the student to find segments or angles necessary to prove the congruence of two triangles. Finally, DELTA encourages the student to input equal segments/angles for the conclusion of the given problem.

If the student cannot make any progress to solve a given problem for 20 seconds, DELTA provides the message: ‘Try to consider from bottom to top’. When the student presses the ‘From the Bottom’ button in the upper side of the An-

Table 3 Hints for BC support on problem shown in Fig. 7.

No.	Hint	Space to Answer
1	Let us consider a proof of this proposition from the conclusion. What is the conclusion in this problem?	[ ]
2	To prove the conclusion, what should we prove?	∠[ ] = ∠[ ] or ∠[ ]= ∠[ ]
3	Let us find a pair of triangles, including the segments/angles selected in the last hint, that may be congruent.	△[ ] = △[ ]
4	What are revealed from the preconditions?	□ AM = BM □ MQ = MP □ QA = PB □ ∠AMQ = ∠BMP □ ∠MQA = ∠MPB □ ∠QAM = ∠PBM
5a	Based on the equations selected in the last hint, what is possible congruent condition(s)?	□ SSS □ SAS □ ASA
5b	Which condition do you use?	○ SSS ○ SAS
6a	(This hint is shown when a student selects SSS in the hint #5b.) To adopt the congruent condition, which equation(s) should we prove?	□ AM = BM □ MQ = MP □ QA = PB □ ∠AMQ = ∠BMP □ ∠MQA = ∠MPB □ ∠QAM = ∠PBM
6b	(This hint is shown when a student selects SAS in the hint #5b.) To adopt the congruent condition, which equation(s) should we prove?	□ AM = BM □ MQ = MP □ QA = PB □ ∠AMQ = ∠BMP □ ∠MQA = ∠MPB □ ∠QAM = ∠PBM
-	Let us prove the equation(s) selected. If necessary, let us refer to your textbook.	

Note that a squared bracket '[ ]', a square □, and a circle ○ represent a textbox, a checkbox, and a radio button, respectively.

swer Area, DELTA supports the learner's use of the BC strategy. At that time, he/she solves the problem by working from the conclusion backward; thus, the student inputs from the bottom to the top of the proof-writing form. In this case, if DELTA judges that the student's input is correct, it will make the input form in the previous row available. When the student is working in this case, the 'From the Bottom' button changes to a 'From the Top' button. Students can press the button to switch modes anytime.

In the 'From the Bottom' mode, DELTA supports the BC strategy. Table 3 shows an example set of hints for the use of BC strategy. DELTA provides hints along the following six steps:

(1) Confirming the conclusion of a given problem: First, DELTA makes the student extract the conclusion from the given problem sentences.

(2) Defining a sub-goal (for the pattern (C) only): The student should recognize a sub-goal to derive the conclusion.

(3) Finding a pair of triangles (for the patterns (B) and (C)): To solve the given problem, DELTA makes the student select a pair of triangles on which to focus.

(4) Considering what are revealed from the given preconditions: DELTA asks the student to clarify which segments or angles are equal and input the corresponding equations, based on the given preconditions.

(5) Presuming an appropriate congruent condition: Based on the equations input in the step (4), DELTA asks the student to presume an appropriate congruent condition. For example, if the equality of two pairs of segments is revealed, possible congruent conditions are 'side, side, side (SSS)' and 'side, angle, side (SAS)'.

(6) Finding equations necessary to adopt the congruent condition: DELTA then asks the student to find the pairs of segments and/or angles whose equality should be proved and to find the appropriate equations for that proof.

4. Efficacy of DELTA: Empirical Evaluation

4.1 Procedure

We conducted an evaluation experiment to measure the efficacy of the BC strategy support provided by DELTA. The experiment included 36 students who were in the second grade of a public junior high school. We conducted two description-style examinations, a pre-test and a post-test, and a questionnaire survey.

After conducting a 20-minute pre-test, we divided the students into two groups such that mean and variance of the pre-test scores were balanced across both groups. In other words, we tried to match pairs of students with equal scores and put one into each group. Each group consisted of 18 participants. One group was the experimental group (EG) who used DELTA. The other group was the control group (CG) who used a baseline system (BS), which excluded the BC strategy support from DELTA. Participants in each group practiced solving TTCPPs for 30 minutes by using the assigned system. Immediately after the practice, we carried out a 20-minute post-test and the questionnaire.

Both the pre- and post-tests were description-style examinations, where one point was assigned to each correct equation derivation and each correct reasoning step in each problem. If an equation was incorrect, no points were given either for the equation or the reasoning step. No points were given for the conclusion. In summary, full marks were six points for a pattern-(A) problem, eight points for a pattern-(B) problem, and 10 points for a pattern-(C) problem. Each examination consisted of one pattern-(A) problem as P1, two pattern-(B) problems as P2 and P3, and one pattern-(C) problem as P4. Hence the full marks for each examination were 32 points. When we set the examination problems, we referred to two textbooks and a study-aid website [24]–[26]. The difficulty of the problems was similar to the practice problems or end-of-chapter problems in these textbooks. The post-test problems were more difficult than the pre-test problems, because probably it was harder for the participants to find two triangles that should be proved to be congruent in the post-test.

In the practice session, each participant in the two groups addressed the same problems in the same order. We prepared 10 problems, which consisted of four pattern-(A) problems, four pattern-(B) problems, and two pattern-(C) problems. The BS supported only the FC strategy. It did not provide the 'From the Bottom' button and provided hints

Table 4 Mean score on pre- and post-tests.

Group (N)	Pre-test *Mean±S.D.* (*Median*)	Post-test *Mean±S.D.* (*Median*)
EG (18)	18.17±10.84 (22)	18.11±11.50 (22)
CG (18)	18.33±10.73 (22)	17.05±9.54 (19.5)
Mann-Whitney's U-test	p=0.91	p=0.50

Table 5 Mean score on pre- and post-tests except for results of FMNP (full-mark and no-point) participants.

Group (N)	Pre-test *Mean±S.D.* (*Median*)	Post-test *Mean±S.D.* (*Median*)
EG (14)	21.07±7.63 (22)	21.07±9.13 (22.5)
CG (14)	19.00±8.54 (22)	17.36±8.67 (19)
Mann-Whitney's U-test	p=0.56	p=0.17

Table 6 Mean number of problems on which a participant got full marks (on pre- and post-tests).

Group (N)	Pre-test *Mean±S.D.* (*Median*)	Post-test *Mean±S.D.* (*Median*)
EG (14)	1.93±0.88 (2)	1.50±1.12 (1)
CG (14)	1.36±0.97 (2)	0.64±0.72 (0.5)
Mann-Whitney's U-test	p=0.18	p=0.04

only for the FC strategy. Concerning the ethics of putting participants in the CG (that is, depriving them of the opportunity of getting the benefit by using DELTA), we would like to note that they would have been deprived anyway because DELTA is not yet generally available [27]. However, we addressed this issue by allowing the CG participants the opportunity to use DELTA after our experiment.

4.2 Result

4.2.1 Results of Pre- and Post-Tests

Table 4 shows the mean scores of EG and CG on the pre- and post-tests. A Mann-Whitney's U-test revealed no significant difference in pre- and post-tests. However, some participants in both groups obtained full marks (FM) or obtained no points (NP) on the pre-test. We called these participants 'FMNP participants'. There was a possibility that the ability of FMNP participants could not be correctly measured in the experiment: because the post-test scores of the FM participants cannot exceed their pre-test scores, we cannot measure the effect obtained by their use of DELTA. Similarly, because we cannot expect that the NP participants have equivalent performance on the pre-test, we cannot measure their increase from the pre-test to the post-test. Therefore, FMNP participants were excluded from the analysis, leaving 14 participants for the following analyses. Table 5 shows the mean scores of each test with the FMNP participants excluded. A Mann-Whitney's U-test revealed no significant difference in this grouping.

Tables 4 and 5 include the partial points received to measure whether or not the participants could consider the factors (a) and (b) of MEXT's suggestion. In addition, we needed to measure whether or not they could consider the factor (c). Hence, we analyzed the mean number of problems on which students got full marks on the pre- and post-tests. Table 6 shows these means. A Mann-Whitney's U-test revealed no significant difference between EG and CG in pre-test. On the other hand, the same test revealed that there was a significant difference between EG and CG in post-test (U = 56, p = 0.04). From these results, we can observe that the participants obtained different effects depending on the system used. We discuss this effect in Sect. 5.

M is the middle point of the segment AB. The segments AC and DB are parallel. Prove ΔACM ≡ ΔBDM.
(a) Because M is the middle of the segment AB, AM = BM.
(b) Because of AC // DB, ∠CAM = ∠DBM.
(c) Which congruent condition can I use in this problem?
(d) <u>What is lacking to prove this problem?</u>
(e) Because of vertical angles, ∠AMC = ∠BMD.
(f) Others (   )

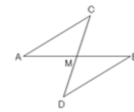

Fig. 12 Example questionnaire (Underlined sentence is BC strategy for this problem.)

Table 7 Number of problems on which each EG or CG participant tried to use BC strategy (on post-test)

Participant's ID	EG (N = 14)	CG (N = 14)
1	2 problems	1 problem
2	2	4
3	4	0
4	0	1
5	3	0
6	2	4
7	2	1
8	0	4
9	4	1
10	4	1
11	2	4
12	3	1
13	0	0
14	0	0
Mann-Whitney's U-test	p=0.45	

Table 8 Number of participants who tried to use FC and/or BC strategies on each problem given in post-test

Attempt to Use Strategies	P1 EG - CG	P2 EG - CG	P3 EG - CG	P4 EG - CG	Total
Neither	0 - 0	3 - 1	0 - 3	4 - 5	16
FC only	8 - 10	5 - 8	8 - 7	0 - 0	46
BC only	0 - 0	0 - 1	0 - 0	0 - 0	1
Both	6 - 4	6 - 4	6 - 4	10 - 9	49

4.2.2 Questionnaire

Next, we investigated the strategies that the participants tried to use for each problem in the post-test. We enumerated the possible thinking processes to solve the problems and classified them as using FC or BC strategies based on [12]. One of the questionnaires is shown in Fig. 12. In the questionnaire, we asked the participants to check the processes that they tried to use. There are 4 problems provided on the post-test. Table 7 shows the number of problems on which each EG or CG participant tried to use BC strategy. Table 8 shows the number of participants who attempted to use the FC and/or BC strategies on each problem given in the post-test.

Finally, we investigated whether or not there was a relationship between the participants who solved the problem and those who attempted to use the FC and/or BC strategies. Table 9 shows the relationship between the participants' attempt to use the FC and/or BC strategies and their

Table 9 Relationship between participants' attempt to use FC and/or BC strategies and their performance in post-test

Attempt to Use Strategies	Attempts in Total Shown in Table 8	Score Rate (%) of Post-test±S.D.
Neither	16	15.63±25.54
FC only	46	67.84±34.50
BC only	1	0.00±0.00
Both	49	72.50±31.23

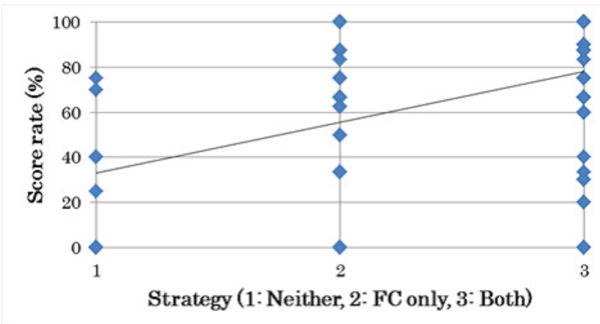

Fig. 13 Relationship between participants' post-test performance and their attempt to use neither FC nor BC strategies (1), only to use FC strategy (2), or to use both FC and BC strategies (3).

performance in the post-test; and in Fig. 13 this relationship is shown as a scatter plot. Note that because the number of participants who attempted to use the BC strategy only was quite low, we excluded such case in Fig. 13. After we calculated the Pearson product-moment correlation coefficient with the data shown in Fig. 13 (N = 111), the correlation coefficient was 0.43 (the p-value of the test for no correlation was $p < 0.01$). From this result, we can see that the more the participants attempted to use both strategies, the higher were their scores.

5. Discussion

In this section, we discuss the efficacy of DELTA and its limitation.

5.1 Efficacy of BC Support

5.1.1 Attempt to Use BC Strategy

We expected that the EG participants would try to use the BC strategy to solve TTCPPs on the post-test. However, the results shown in Table 7 reveal that there was no significant difference between EG and CG in their attempts to use the BC strategy. Nevertheless, Table 6 shows that a significant difference can be seen in the post-test results. We conclude that EG participants were better able to use the BC strategy for solving TTCPPs.

On the other hand, we investigated the correlation between the number of problems in which the participants attempted to use the BC strategy and the total scores of participants' post-test (Fig. 14): the Spearman's Rho of this correlation was 0.38 ($p = 0.04$). We also computed the correlation between the number of problems in which the participants attempted to use the BC strategy and the number of problems which participants were able to solve completely in the post-test, and the Spearman's Rho for this came out to be 0.30 ($p = 0.12$) (Fig. 15). From these results, we can see that the former correlation is statistically significant, that is, there is a weak correlation between increased attempts to use of the BC strategy to solve problems and increased post-test scores.

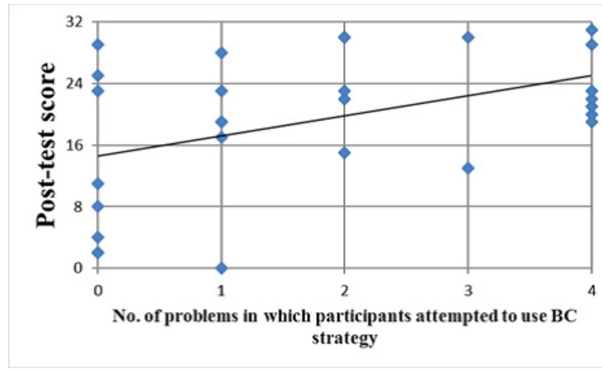

Fig. 14 Relationship between participants' post-test scores and the number of problems in which they attempted to use BC strategy.

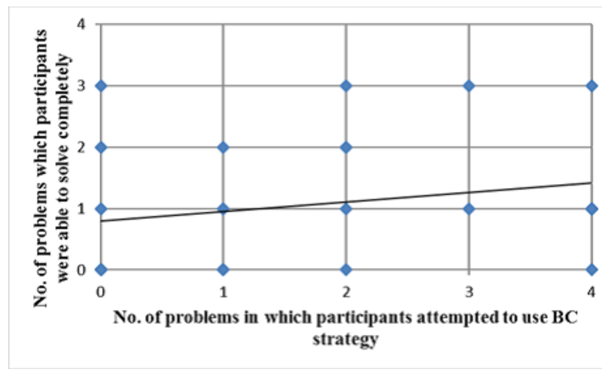

Fig. 15 Relationship between the number of problems which participants were able to solve completely and the number of problems in which they attempted to use BC strategy.

From Table 6, we can see that the participants who received FC and BC support showed better performance than those who received only FC support. On the other hand, from Tables 7 and 8, we can see that there was not any significant difference in the degree of BC strategy use between EG and CG participants. Therefore, those results indicate that the efficacy of DELTA is not to facilitate students' use of BC strategy but to improve the quality of their BC strategy use. However, this indication needs to be explored further in future research. Because our knowledge of the strategies that the participants used is based on the questionnaires, we cannot determine whether the participants actually used the FC or BC strategies. Observation of the participants' proof-writing process might help to reveal the strategies they actually used, but observing trial and error in thought processes is obviously difficult.

Table 10 Mean score rate on each problem of pre- and post-tests.

Group	Pre-test				Post-test			
	P1	P2	P3	P4	P1	P2	P3	P4
(Proof steps)	(4)	(5)	(5)	(6)	(4)	(5)	(5)	(6)
EG(N = 14)	100(%)	69.6	69.6	39.3	81.0	73.2	73.2	45.0
CG(N = 14)	84.5	69.6	50.9	42.9	57.1	64.3	64.3	36.4
Total	92.3	69.6	60.3	41.1	69.0	68.8	68.8	40.7

Table 11 Number of participants who got full marks on each problem of pre- and post-tests.

Group	Pre-test				Post-test			
	P1	P2	P3	P4	P1	P2	P3	P4
(Proof steps)	(4)	(5)	(5)	(6)	(4)	(5)	(5)	(6)
EG(N = 14)	14	7	6	0	7	4	7	3
CG(N = 14)	10	7	1	1	1	1	6	1
Total	24	14	7	1	8	5	13	4

5.1.2 Relationship between the Number of Proof Steps and Attempts to Use BC Strategy

We can suppose that as the number of proof steps necessary to solve a problem increases, the success rate of the students using only the FC strategy will decrease. Tables 10 and 11 show the results of each problem on pre- and post-tests. From Table 10, the score rates of problem P1, whose number of proof steps is smaller, are higher than other problems in both pre- and post-tests. On the other hand, those of P4, whose number of proof steps is larger, are lower. However, in Table 11, we can see that the results are contrary to our supposition. Hence, this issue needs further investigation.

Though a 30-minute practice session might not be enough for the participants to acquire the BC strategy, EG participants' performance is better than CG in all the problems of post-test. Especially, the difference between the P1 score rates of EG and CG is larger. In this experiment, the reason behind this difference cannot be identified. If the participants practiced using DELTA for a longer time, it might increase their score on the post-test.

We adopted the fundamental TTCPPs in the experiment because of the curriculum of the school participating in the experiment. If the problem range was larger and/or more difficult problems were included, it is possible that the use of the BC strategy would be even more effective. We plan to investigate and discuss this possibility in future research.

5.2 Comparison of Our Results with Related Work

According to [12] and [13], the authors found that (a) all students were familiar with the FC strategy; (b) students who successfully solved problems utilized the BC strategy in their problem-solving; (c) many students who could not solve problems did not use the BC strategy; and (d) lack of the BC strategy led to a failure in problem-solving.

From the results of P1 shown in Table 8, all participants tried to use the FC strategy, which is compatible with (a). The result is natural, because many textbooks and study-aid books show proof-writing procedures based on the FC strategy.

Table 12 Relationship between attempts to use BC strategy and solved/unsolved problems on post-test

	Solved	Unsolved	Total
Attempt	16	35	51
No Attempt	14	47	61
Total	30	82	112

As for (b), we define successfully solving a TTCPP as getting full marks. Here we use the concepts of Support, Confidence and Lift, defined in research on Association Rule Mining [28], to assess the strength of findings (b) – (d) above. In the rule: 'A ⇒ B', the count of B divided by the count of whole event (=P(B)) is the right-hand-side Support. The Confidence is calculated by P(A∪B)/P(A). The Lift is calculated by P(A∪B)/P(A)P(B). A Lift value greater than 1.0 suggests that there is some usefulness to the rule. The larger the Lift value, the greater is the strength of the association between A and B. With the Lift value, the importance of a rule can be validated in an effective manner [28]. In our experiment, the total number of problems in the post-test was 112 (=28×4). The total number of problems on which the participants got full marks on the post-test was 30, and 15 of these problems were solved by trying to use the BC strategy. The total number of problems solved by trying to use the BC strategy was 50. In other words, the right-hand-side Support of the rule: 'a participant successfully solved a TTCPP ⇒ he/she tried to use the BC strategy' is 0.45 (= 50/112). The Confidence of the rule is 0.50 (=15/30). The Lift value of the rule is 1.12 (=(15/30)/(50/112)). Thus, the rule: 'a participant successfully solved a TTCPP ⇒ he/she tried to use the BC strategy' is useful because the Lift value of the rule is greater than 1.0. We could not find a negative factor of (b) in our experiment.

As for (c), we define not being able to solve a problem as not getting full marks. As shown in Table 12, among 112 problems, 82 were not solved successfully. Among these 82 problems, the participants did not try to use the BC strategy on 47 problems. On the other hand, the total number of problems on which they did not try to use the BC strategy was 61. Therefore, the right-hand-side Support of the rule: 'a participant did not solve a TTCPP ⇒ he/she did not try to use the BC strategy' is 0.54 (=61/112). The Confidence of the rule is 0.57 (=47/82). The Lift value of the rule is 1.05 (=(47/82)/(61/112)); that is, the rule is useful. We could not find a negative factor of (c) in our experiment.

As for (d), we define 'lack of the BC strategy' as never trying to use the BC strategy on the post-test. Among 28 participants, 8 participants never tried to use the BC strategy on the post-test and 4 of these 8 participants never got full marks on any problem. The total number of participants who never got full marks on any problem was 10. Therefore, the right-hand-side Support of the rule: 'a participant never tried to use the BC strategy on the post-test ⇒ he/she never got full marks on any problem' is 0.36 (=10/28). The Confidence of the rule is 0.50 (=4/8). The Lift value of the rule is 1.40 (=(4/8)/(10/28)); that is, the rule is useful. We could not find a negative factor of (d) in our experiment, and agreed with the rule.

Chi et al. reported that experts used both the FC and BC strategies for solving physics problems [29]. This report relates to the aforementioned MEXT suggestion: It is important for students to consider the following when designing proofs: (a) What conditions are lacking to derive the conclusion of a proposition? (b) What are revealed from some preconditions? and (c) What are lacking to connect (a) and (b)? As for Chi et al.'s report, the total number of problems on which the participants tried to use both the FC and BC strategies on the post-test was 48, and 15 of these were solved with full marks. The total number of problems solved with full marks was 29. Therefore, the right-hand-side Support of the rule: 'a participant solved a TTCPP by trying to use both strategies ⇒ he/she solved the TTCPP with full marks' is 0.26 (=29/112). The Confidence of the rule is 0.31 (=15/48). The Lift value is 1.21 (=(15/48)/(29/112)); that is, the rule is useful. We cannot find a negative factor in Chi et al.'s report.

We note that results and/or reports provided in our research and related work cannot be fully compared, simply because the contents and/or difficulty of problems adopted in each study were different. However, we could generally reconfirm the findings described by [12], [13], [29]. In addition, we confirmed that an attempt to use both FC and BC strategies led to solving TTCPPs.

5.3 Limitation of DELTA

In this research, we implemented a system with which students can learn how to solve TTCPPs outside of their classrooms. However, to check the students' inputs automatically, we adopted multiple-choice forms and fill-in-the-blank proof-writing in DELTA. It is possible that these forms provide hints for solving TTCPPs. We suggest that students first learn to solve TTCPPs with multiple-choice or fill-in-the-blank forms for their self-learning. Afterwards, if they can get used to answering TTCPPs with free-description forms, the non-response rate on these problems will decrease in public examinations such as the National Assessment of Academic Ability conducted by MEXT.

6. Conclusion and Future Work

In this research, we designed and implemented a system, called DELTA, that supports students as they use the BC strategy to prove the congruence of two triangles. In supporting the use of the BC strategy, DELTA provides hints so that the student can fill in the proof-writing form from bottom to top. DELTA also provides 'coloring', 'marking', and 'highlighting' functions to support students.

We conducted an evaluation experiment to measure the efficacy of the support of the BC strategy by DELTA. The participants were divided into an experimental group (EG) and a control group (CG). The participants in the EG addressed TTCPPs using DELTA, while those in the CG did so by using the baseline system, which excluded the support function for the BC strategy. Each group consisted of 18 students who were in the second grade at a public junior high school in Japan. The experiment consisted of a pre-test, a post-test and a questionnaire survey to measure efficacy.

We analyzed the results of the experiment, excluding students who got full marks or got no points on the pre-test. Our analysis revealed the following points:

- Support of the BC strategy in DELTA is effective, because the mean number of problems on which the EG got full marks was higher than that of the CG.
- The participants in the EG were better able to use the BC strategy to solve TTCPPs.
- Attempts to combine both the FC and the BC strategies led to solving TTCPPs.

We would like to introduce free-description forms into DELTA instead of multiple-choice and/or fill-in-the-blank forms. Integration of a database for supporting teachers in making problems is also left for the future work. In addition, we plan to devise a method to understand the strategies and tactics students use for generating proofs, and to clarify students' use of DELTA. For this purpose, we plan to introduce a logging function into DELTA and/or make video recordings. It is necessary to verify that the efficacy of DELTA is not to facilitate students' use of BC strategy but to improve the quality of their BC strategy use.

Acknowledgments

We appreciate the anonymous reviewers for their insightful comments and constructive suggestions. This study is supported in part by a Grant-in-Aid for Young Scientists (B) of the Japan Society for the Promotion of Science (JSPS) under Grant No. 15K20980 and by a Grant-in-Aid for Scientific Research (C) of JSPS under Grant No. 17K01118.

References

[1] R. Onda, Y. Hirai, and K. Kaneko, "Development of a learning system for proving the congruence of two triangles by supporting 'backward chaining'," Proceedings of the 3rd ICT International Student Project Conference, pp.1–4, 2014.
[2] S.L. Senk, "How well do students write geometry proofs?," The Mathematics Teacher, vol.78, pp.448–456, 1985.
[3] National Institute for Educational Policy Research, Summary of expected future works based on 4-years experimental results of the national assessment of academic ability — Toward improvement and fulfilling of lecture to students (Junior high school part) (in Japanese), Kyoiku-shuppan Press, 2012.
[4] National Institute for Educational Policy Research, Report of the national assessment of academic ability in FY2014 (Mathematics part of junior high school) (in Japanese), 2014.
[5] National Institute for Educational Policy Research, Report of the national assessment of academic ability in FY2010 (Mathematics part of junior high school) (in Japanese), pp.178–307, 2010.
[6] National Institute for Educational Policy Research, Report of the national assessment of academic ability in FY2012 (Mathematics part of junior high school) (in Japanese), pp.210–341, 2012.
[7] National Institute for Educational Policy Research, Report of the national assessment of academic ability in FY2013 (Mathematics part of junior high school) (in Japanese), 2013.
[8] S. Autexier, D. Dietrich, and M. Schiller, "Towards an intelligent

tutor for mathematical proofs," Proceedings of Theorem Proving Components for Educational Software, vol.79, pp.1–28, 2012.
[9] S.-C. Chou, X.-S. Gao, and J.-Z. Zhang, "Automated generation of readable proofs with geometric invariants. ii: theorem proving with full-angles," Journal of Automated Reasoning, vol.17, pp.325–347, 1996.
[10] N. Matsuda and K. VanLehn, "Gramy: A geometry theorem prover capable of construction," Journal of Automated Reasoning, vol.32, no.1, pp.3–33, 2004.
[11] S. Wilson and J. Fleuriot, "Combining dynamic geometry, automated geometry theorem proving and diagrammatic proofs," Proceedings of the European Joint Conferences on Theory and Practice of Software/Satellite Workshop on User Interfaces for Theorem Provers, 2005.
[12] S. Karimata, "A study on junior high school student's problem-solving using act (in japanese)," Bulletin of Educational Research and Training Center, vol.3, pp.1–11, 1995.
[13] T. Makino, "Student's cognitive difficulties in 'bridging process' of proving (in japanese)," Technical Report of Japan Society for Science Education, vol.26, pp.25–30, 2012.
[14] G. Dimakos, E. Nikolusdakis, S. Ferentinos, and E. Choustoulakis, "Developing a proof-writing tool for novice lyceum geometry students," The Teaching of Mathematics, vol.10, pp.87–106, 2007.
[15] T. Itoh, "A system for understanding proof procedures related to congruence of triangles (in japanese)," Graduation thesis, Tokyo University of Agriculture and Technology, Tokyo, Japan, 2002.
[16] N. Matsuda and K. VanLehn, "Advanced geometry tutor: An intelligent tutor that teaches proof-writing with construction," Proceedings of the 12th International Conference on Artificial Intelligence in Education, pp.443–450, 2005.
[17] "Netlessonlab." http://netlessonlab.com/. (Online, Available on July 12th, 2016).
[18] D. Wood, H. Wood, and D. Middleton, "An experimental evaluation of four face-to-face teaching strategies," International Journal of Behavioral Development, vol.1, no.2, pp.131–147, 1978.
[19] K. Jones, T. Fujita, and M. Miyazaki, "Learning congruency-based proofs in geometry via a web-based learning system," Proceedings of the British Society for Research into Learning Mathematics, pp.31–36, 2013.
[20] M. Miyazaki, K. Jones, and T. Fujita, "A web-based learning system for congruency-based proofs in geometry in lower secondary school," Proceedings of the 11th International Conference on Technology in Mathematics Teaching, pp.206–212, 2013.
[21] A. Takeuchi and T. Okamoto, "Knowledge-base-oriented teaching and learning support system (in japanese)," in Educational information technology of the Internet age 2, ch. 5, Morikita Publishing, 2001.
[22] W.K. Wong, S.K. Yin, H.H. Yan, and Y.H. Cheng, "Using computer-assisted multiple representation in learning geometry proofs," Educational Technology & Society, vol.14, pp.43–54, 2011.
[23] Ministry of Education, Culture, Sports, Science and Technology (MEXT), Commentary of courses of study for junior high school mathematics (in Japanese), Kyoiku-syuppan Press, 2008.
[24] T. Fujii and H. Matano, New Mathematics 2 (in Japanese), Tokyo Shoseki Press, 2013.
[25] "Mathematics website." http://math.005net.com/mondai.php. (Online Available: July 12th, 2016).
[26] K. Okamoto, K. Koseki, M. Morisugi, and T. Sasaki, Gateway to the future math 2 (in Japanese), Keirinkan Press, 2013.
[27] B.E. Whitley Jr. and M.E. Kite, Principle of research in bahavioral science, 3 ed., Routledge, 2012.
[28] M.K.N. Vaani and E. Ramaraj, "An optimal approach to derive disjunctive positive and negative rules from association rule mining using genetic algorithm," IOSR Journal of Computer Engineering, vol.13, no.1, pp.73–82, 2013.
[29] M.T.H. Chi, P.J. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," Cognitive Science, vol.5, no.2, pp.121–152, 1981.

Ryosuke Onda was an undergraduate student at Tokyo University of Agriculture and Technology. Now, he is a teacher of the 6th junior high school of Chofu-city in Tokyo, Japan. His main research area was educational technologies. He received the B.E. degree from Tokyo University of Agriculture and Technology in 2014.

Yuki Hirai was an Assistant Professor at Tokyo University of Agriculture and Technology in Japan. Now, he is a Lecturer at Shinshu University in Japan. His main research area is educational technologies. He received the B.L.A. and M.Ed. degrees from Tokyo Gakugei University in 2007 and 2009, respectively. He also received the Ph.D. degree from University of Tsukuba in 2012. He is a member of IPSJ and JSAI.

Kay Penny is a Lecturer in Research Methods and Statistics at Edinburgh Napier University in the UK. Her main research interests include applied statistics, survey methods, data mining and methods for dealing with missing data. She was awarded a B.Sc. (Hons) in 1992 and a Ph.D. in 1995 at the University of Aberdeen. She is a Fellow of the Royal Statistical Society and a Senior Fellow of the Higher Education Academy.

Bipin Indurkhya is a Professor of Cognitive Science at the Jagiellonian University, Cracow, Poland. His main research interests are social robotics, usability engineering, affective computing and creativity. He received his Master's degree in Electronics Engineering from the Philips International Institute, Eindhoven (The Netherlands) in 1981, and Ph.D. in Computer Science from University of Massachusetts at Amherst in 1985.

Keiichi Kaneko is a Professor at Tokyo University of Agriculture and Technology in Japan. His main research areas are functional programming, parallel and distributed computation, partial evaluation, and fault-tolerant systems. He received the B.E., M.E., and Ph.D. degrees from the University of Tokyo in 1985, 1987, and 1994, respectively. He is a member of ACM, IEEE CS, IPSJ, and JSSST.