

**MODEL-BASED PROCESS ANALYSIS OF  
PROBLEM POSING AS SENTENCE INTEGRATION  
IN ARITHMETIC WORD PROBLEMS**

(算数文章題の単文統合型作問についてのモデル  
ベースのプロセス分析)



**AHMAD AFIF SUPIANTO**

**D140720**

Graduate School of Engineering  
Hiroshima University  
September 2017



# DISSERTATION SUMMARY

---

Posing a problem comprises the generation of new problems as well as the reformulation of given ones. Problem-posing activities are suitable ways to promote mathematical thinking amongst learners. Providing learners with an opportunity to pose their problems can foster flexible thinking and enrich basic concepts on learning mathematics. Additionally, it has a positive effect on understanding mathematical concepts as well as on learners' views about what it means to know mathematics. Technology-enhanced approaches have great potential for promoting learning by problem posing. In particular, interactive learning systems provide enriched features where learners can pose the problems individually, assessed automatically, and, at the same time, given feedback to each posed problem. Monsakun, which means “Problem-posing Boy” in Japanese, is an interactive learning system that encourages learners in posing problems of arithmetic word problems. A slightly hidden potential of Monsakun practical uses that this thesis explores is that they can open new opportunities for capturing learner's digital traces of activity. These data can be analyzed through a variety of techniques, from statistical to data mining approaches that describe general aspects of learners' activity, learners' tendency in fulfilling problem requirements, and learners' impasse in thinking during pose the problems. Furthermore, the analysis of such data can provide means to help teachers and researchers inspect the process followed by learners and develop support functions and sophisticated feedbacks to learners.

Previous researches investigate learners pose the problems on Monsakun are intended to analyze the pre- and post- test scores, examine the first selected sentence, and concerning the completed posed problems. Investigations in such research are primarily limited to analyzing the result of the process. This thesis extends the analysis by involving the process of arranging the problem, which means every action of learners when they pose the problems is examined to understand the learning process of posing problems on Monsakun. Therefore, the purpose of studies in this thesis is to investigate and discover

the learning process of learners who learn problem-posing as sentence integration in arithmetic word problems through the log of interaction data on Monsakun.

This thesis consists of six chapters. In **Chapter 1**, the research context and outlines the goals, contributions, evaluation methods, and the general structure of the thesis are described. **Chapter 2** outlines relevant research on learning by problem posing, the use of learning system for posing problems, the area of log data analysis, and the exploration of previous work and the cognitive model related to the learning system mainly used in this thesis: Monsakun Touch. In **Chapter 3**, we describe the process analysis of log data of learners posing the problems on Monsakun to investigate whether learners consider the problem's structure in posing the required problems, focused on two different perspectives: satisfying the required problem and avoiding the violated constraints. The chapter presents correlation and difference analysis to investigate whether learners pose problems by attempting to satisfy as many required constraints as possible and to avoid as many violated constraints as possible. In case they violated the constraints, this chapter shows further analysis to investigate whether learners tended to have difficulty avoiding a particular type of constraint. Next, **Chapter 4** presents visual representations of learners' problem-posing process for discovering the bottlenecks in thinking while they pose the problems. The study in this chapter traces sequences of learners' actions and provide insight in what type of condition learners got stuck. In addition, the chapter describes the evaluation of the visualizations and the discussion of analysis result according to the cognitive model involved. In **Chapter 5**, we provide a detailed view of the reasoning behind the design of our computer-based scaffolding system. From the technology perspective, the system is the goal of the main studies in this thesis. The chapter presents the articulation of the problems based on the log data collected, system architecture, and operation procedures of the scaffolding system. Additionally, this chapter describes the discussion of the proposed system related to the problem constraints. Finally, the conclusion of this thesis and future work directions are given in **Chapter 6**.

# ACKNOWLEDGEMENTS

---

First of all, praise to Allah for his kindness for letting me finished this Dissertation. I would like to take this opportunity to extend my heartfelt appreciation to following persons whose have contributed directly or indirectly towards the completion of the study. Primarily, I would like to express my sincerest gratitude to my supervisors, Hirashima-sensei and Hayashi-sensei, for their unconditional guidance and support in this journey. I could not have had better supervisors. I learned so much from their knowledge in their respective areas of expertise, their discussion skills, and the ways they come up with new ideas. Besides his technical support, I have benefited from him positively in my personal life. Also, my sincerest gratitude to my co-supervisor Mukaidani-sensei for his advice and feedback to improve my research.

Secondly, I wish to extend my profound thanks to my colleagues in Brawijaya University, Malang who support my study in Japan. Thanks to Pak Herman Tolle for introducing me to Hirashima-sensei through Bu Saida Ulfa, State University of Malang. Also, Thanks to Pak Tibyani, Pak Tatang, Mas Budi, and Mbak Rekyan for always supporting me from Malang.

I am also grateful to the collaborators I had on this journey. A large part of my gratitude goes to my *senpai*, Nur-san. With her initial help and discussions, my works became smoothly and well-conducted. Additionally, my thanks go to all Learning Engineering Laboratory members, Ando-san as the representative of international students and Kitamura-kun as the representative of Japanese students. Also, I would also like to thank Karen for being my advisor in writing group. Micky and the writing group members for supporting and commenting to be a better writer.

I am also grateful to Indonesian friends. All members of *Keluarga Cijeruk*, who always be kindness, provide my family and me the best residence in Japan. Also, my sincere gratitude to all members of *Jama'ah Yasinta*. They always spiritually support me in praying together. They will always be a part of my life in Japan.

Most importantly, I would like to show my gratitude to all my family, who permanently loved me, motivated me, and supported me to do my best. My big thank to my wife, Eny Setiyowati, my children, Arsyah Dewi Lathifa and Previo Really Supianto. Also, thanks to my parents, *Bapak Matadji Supianto*, *Bapak Widji*, *Ibu Zulaikah*, and *Ibu Muryati* for their love and continual support through this challenging but amazing period of my life.

At last, I would like to acknowledge Monbukagakusho Scholarship, Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan for the financial support during the completion of my study to pursue a doctoral degree at Graduate School of Engineering, Hiroshima University.

Author

Ahmad Afif Supianto 印

# PREFACE AND NOTES

---

## Publications

### Journal articles

- RPTEL 2016: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016b). Visualizations of problem-posing activity sequences toward modeling the thinking process. *Research and Practice in Technology Enhanced Learning*, 11(14), 1-23. doi:10.1186/s41039-016-0042-4
- RPTEL 2017: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017b). Model-based analysis of thinking in problem posing as sentence integration focused on violation of the constraints. *Research and Practice in Technology Enhanced Learning*, 12(12), 1-21. doi:10.1186/s41039-017-0057-5
- IEICE 2017: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017c). An Investigation of Learner's Actions in Posing Arithmetic Word Problem on an Interactive Learning Environment. *IEICE Transactions on Information and Systems*, E100-D(11), article in press.

### Conference papers

- ICCE 2015: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2015a). Tracing problem-Posing activity sequences toward detection of trap-states in thinking. *The 23rd International Conference on Computers in Education*, (pp. 85-90). Hangzhou, China.
- ITS 2016: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016a). An Investigation of Learner's Actions in Problem-Posing Activity of Arithmetic Word Problems. *Intelligent Tutoring Systems: 13rd International Conference. LNCS, volume 9684*, pp. 478-479. Zagreb, Croatia: Springer International Publishing.
- ICCE 2016: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016c). Analysis of Steps in Posing Arithmetic Word Problem as Sentence-Integration on Interactive Learning Environment. *The 24th International Conference on Computers in Education*, (pp. 242-251). Mumbai, India.
- IOP 2017: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017a). Process-based Assignment-Setting Change for Support of Overcoming Bottlenecks in Learning by Problem-Posing in Arithmetic Word Problems. *The 3rd Mathematics, Science and Computer Science Education International*

*Seminar (MSCEIS 2016)*. Bandung, Indonesia. doi:10.1088/1742-6596/812/1/012004

### Workshop papers

WS TEL-PP 2015: Supianto, A. A., Hayashi, Y., & Hirashima, T. (2015b). Analysis of Problem-Posing Activity Sequences toward Modeling Thinking Process and Detection of Trap States. *The 8th Workshop on Technology Enhanced Learning by Posing/Solving Problems/Questions*, (pp. 558-567). Hangzhou, China.

### Doctoral Student Consortium

DSC-ICCE 2015: Supianto, A. A. (2015). Analysis of Problem-Posing Activity Sequences toward Modeling Thinking Process and Detection of Trap States. In *The 23rd International Conference on Computers in Education, ICCE 2015*. Hangzhou, China. Mentors: Chen-Chung Liu, Huei-Tse Hou, Danielle McNamara.

## Source and Original Work

Original material of my own from the above publications has been included in this thesis, with a citation to the appropriate publication appearing at the beginning of each chapter. Other external sources are cited, with the bibliography appearing at the end of the thesis. All figures are created originally by my own for supporting the above publications.

## Use of Work by Others

The dataset of Japanese first-grade elementary school students that are analyzed in this thesis was collected through a study conducted by the fellow researcher. It was gathered by Yamamoto et al. (2012). However, the pre-processing and analysis of these data are part of the original contributions of studies in this thesis.

# CONTENTS

---

Dissertation Summary.....	i
Acknowledgements.....	iii
Preface and Notes .....	v
Contents .....	vii
List of Figures.....	xi
List of Tables .....	xv
Chapter 1: Introduction.....	1
1.1 Context and Motivation.....	1
1.2 Thesis Statement .....	5
1.3 Thesis Goals .....	7
1.4 Thesis Contributions .....	9
1.5 Research Methodology and Validation Methods.....	9
1.6 Thesis Structure.....	12
Chapter 2: Background .....	15
2.1 Learning by Problem Posing.....	15
2.1.1 Problem Posing Ability and Processes .....	16
2.1.2 Classification of Problem Posing Task.....	17
2.2 The Use of Learning Systems for Posing Problems.....	18
2.2.1 Problem Posing Techniques on Learning Systems.....	19
2.2.2 Assessment and Investigation of the Posed Problems .....	21
2.3 The Problem-Posing Learning System based on the Triplet Structure Model.	22
2.3.1 The Triplet Structure Model .....	23

2.3.2	The Constraints based on the Task Model .....	26
2.3.3	Monsakun as a Problem-Posing Learning System.....	28
2.4	Analysis of Learning Activity on the Learning Systems.....	30
2.5	Scaffolding Systems .....	32
2.6	Chapter Summary .....	34
Chapter 3: Log Data Analytics in term of the Constraints .....		35
3.1	Introduction .....	35
3.2	Motivation .....	37
3.3	Research Questions.....	38
3.4	Context of the Study and Data Exploration.....	39
3.4.1	Participants and Procedure .....	39
3.4.2	Data Collection.....	40
3.4.3	Formulation of Problem-Posing Processes .....	42
3.5	Assessment of States .....	44
3.5.1	Satisfaction of the Constraints.....	44
3.5.2	Violation of the Constraints .....	47
3.6	Analyses in term of Satisfying the Required Constraints.....	49
3.6.1	RQ1: Do learners pose the problems by attempting to satisfy as many constraints as possible? .....	50
3.7	Analyses in term of Avoiding the Violated Constraints.....	54
3.7.1	RQ2: Do learners pose the problems by attempting to avoid as many violated constraints as possible? .....	55
3.7.2	RQ3: Do learners have difficulty avoiding a particular type of constraint?.....	59
3.8	Chapter Summary .....	61
Chapter 4: Visualizations of Learners Activity in Posing the Problems.....		63

4.1	Introduction .....	63
4.2	Motivation and Purpose .....	65
4.3	An Automatic Approach through Visualizations .....	67
4.3.1	Collecting Log Data.....	68
4.3.2	Formulating problem-posing process .....	68
4.3.3	Tracing activity sequences.....	70
4.3.4	Calculating the support and distance values .....	72
4.3.5	Visualizing the values using graphical representation.....	74
4.3.6	Finding trap-states as the bottlenecks .....	75
4.4	Implementation of the visualizations to a case study on the practical use of Monsakun.....	75
4.4.1	Participants and Materials.....	75
4.4.2	Visualization of Support and Distance .....	76
4.4.3	Visualization of Changes of the Trap State .....	78
4.5	Discussions.....	80
4.6	Chapter Summary.....	84
Chapter 5: Designing a Computer-Based Scaffolding on a Problem-Posing Learning System.....		85
5.1	Introduction .....	85
5.2	Problem Formulation .....	88
5.3	Proposed Scaffolding System .....	90
5.3.1	System Architecture and Components.....	91
5.3.2	Operation Procedures.....	92
5.3.3	New Setting of an Assignment .....	93
5.4	Chapter Summary.....	96
Chapter 6: Conclusions and Future Work.....		97

6.1	Summary of Studies.....	97
6.2	Future Directions .....	100
	References .....	103
	Appendix .....	115
	A.1. Detail Assignments and Sentence Cards in Level 5.....	115
	A.2. Rules on Deciding Validity of the Constraints.....	117

# LIST OF FIGURES

---

Figure 1-1	Overview of the context, goals, contributions and evaluation of this thesis.....	6
Figure 1-2	Research phases: associated milestones and thesis chapters.....	10
Figure 1-3	Structure of the research covered in this thesis and related published papers.....	12
Figure 2-1	Relation between arithmetic story and problem.....	24
Figure 2-2	Example of several compositions of sentence cards and their satisfaction of constraints.....	27
Figure 2-3	Interface of Monsakun.....	29
Figure 3-1	Context, goals, contributions and validation of Chapter 3.....	37
Figure 3-2	Procedure of practical use of Monsakun in a classroom.....	40
Figure 3-3	Example of log data of learners' activity on Monsakun.....	41
Figure 3-4	Example of states and the index of available sentence cards.....	44
Figure 3-5	Sequence of states generated from a learner's actions.....	44
Figure 3-6	Hierarchy of possible compositions of sentence cards and transitions among them.....	47
Figure 3-7	Derivative states of State 014 and the constraint violation in Level 5 Assignment 1.....	49
Figure 3-8	Scatterplot of correlation of transitions in Level 5 Assignment 1.....	52
Figure 3-9	Plot of correlation between the number of violated constraints and the occurrence frequency of states.....	56
Figure 3-10	Portion of the number of states in Assignment 10 to the occurrence frequency.....	59
Figure 4-1	Context, goals, contributions and validation of Chapter 4.....	64

Figure 4-2	Procedure of the study. ....	68
Figure 4-3	The graph of states space for combination story type with 6 available sentence cards. ....	71
Figure 4-4	The examples of the sequence of states. ....	71
Figure 4-5	An example of mapped a sequence to the problem states space. ....	72
Figure 4-6	An example of sequences of four learners for illustrating how to calculate support and distance values of State 014. ....	73
Figure 4-7	Support graph and distance graph of the first trial. ....	77
Figure 4-8	Trap graph of the first trial. ....	79
Figure 4-9	Trap graph of the second trial. ....	80
Figure 4-10	Trap graph of the third trial. ....	81
Figure 4-11	Portion of the number of states in Assignment 10 to the occurrence frequency. ....	83
Figure 5-1	Context, goals, contributions and validation of Chapter 5. ....	88
Figure 5-2	An example of processed log data of a learner’s problem-posing activity. ....	90
Figure 5-3	The system architecture of proposed problem-posing learning system. ....	91
Figure 5-4	The proportion of the number of states that violate the constraints. ....	94
Figure 5-5	An example of processed log data of a high-performance learner’s problem-posing activity. ....	95
Figure 6-1	Summary of the thesis goals matching to the thesis chapters 3-5. ....	98
Figure A-1	Rule validity of states in Level 0. ....	117
Figure A-2	Rule validity of states in Level 1 for Calculation and Story Type. ....	118
Figure A-3	Rule validity of states in Level 1 for Number, Object, and Sentence Structure. ....	119
Figure A-4	Rule validity of states in Level 2 for Calculation and Story Type. ....	120
Figure A-5	Rule validity of states in Level 2 for Number, Object, and Sentence Structure. ....	121
Figure A-6	Rule validity of states in Level 3 for Calculation. ....	122

Figure A-7	Rule validity of states in Level 3 for Story Type and Number. ....	123
Figure A-8	Rule validity of states in Level 3 for Object and Sentence Structure.....	124

*<empty page>*

# LIST OF TABLES

---

Table 2-1	Detailed level assignments in Monsakun. ....	30
Table 3-1	Descriptive statistics of learners' actions and mistakes when posing problems on Monsakun. ....	42
Table 3-2	Detailed assignments of Level 5 in Monsakun.....	43
Table 3-3	Example of several states and their satisfaction of constraints. ....	46
Table 3-4	Example of several states and their violation of constraints. ....	48
Table 3-5	Average actions and correlation result between frequency of states appearance and validity of states in Level 5.....	50
Table 3-6	Average actions and correlation result between frequency of states appearance and validity of states in Level 5.....	53
Table 3-7	Average actions and correlation result between frequency of states appearance and number of violated constraints in Level 5. ....	56
Table 3-8	Difference analysis between portions of number of states in assignment setting to its occurrence frequency. ....	58
Table 3-9	Ratio of the number of states in the assignment setting to the occurrence frequency according to the type of constraints.....	60
Table 4-1	Example of each type of states. ....	69
Table 4-2	Actual trap state for each trial.....	84

*<empty page>*

# INTRODUCTION

**Summary:** This chapter describes the research context of the thesis, the identified learning problem and the methodology followed to address it. This thesis builds on the intersection of three fields: Human-Computer Interaction, Educational Data Mining, and Learning Analytics. The thesis proposes an analysis of log data as a result of problem-posing activity that can be used to help understand learners' thinking while posing problems using a problem-posing learning system. The thesis shows how learners' interaction data can be exploited and analyzed to detect learners bottleneck in thinking so that the results can be used to enhance teacher's feedback become accurate and develop appropriate supports to the learners. This chapter outlines the goals, contributions, evaluation methods and the general structure of the thesis.

## 1.1 Context and Motivation

Posing a problem is a demanding process that requires critical thinking, evaluation, and reflection. It is recognized as a key component in the nature of mathematical thinking (Kilpatrick, 1987). Posing a problem involves generating new problems and questions aimed at exploring a given situation as well as reformulating a problem during solving a related problem (Silver, 1994). According to Devlin (2012), mathematical thinking is not the same as doing mathematics, which usually involves some symbolic manipulations. Mathematical thinking, by contrast, is a specific way of thinking about things in the world includes logical and analytic thinking as well as quantitative reasoning. Therefore, providing learners with an opportunity to pose their problems can foster flexible thinking, enhance problem-solving skills, broaden their perception of mathematics, and enrich and consolidate basic concepts (Brown and Walter, 1993; English, 1996).

In recent years, interest in integrating problem posing in mathematical instruction has continuously grown among mathematics education researchers and practitioners (Cai and Jiang, 2016; Ellerton, 2013; Norman, 2011; Singer et al., 2015). Investigations of problems posed by learners and teachers in classrooms have provided insight into the relationships between mathematical knowledge, skills, and processes (Chen et al., 2011; Kılıç, 2013; Stickles, 2011; Van Harpen and Presmeg, 2013). Given the importance of problem-posing activities in school mathematics, some researchers have investigated various aspects of problem-posing processes. One important direction is to examine thinking processes related to the posed problems (e.g., (Bonotto, 2013; Şengül and Katranci, 2015). In examining thinking processes, assessment of each problem and assistance based on it are necessary (Hirashima et al., 2007). Teacher assessment of posed problems encompasses learners' development of diverse mathematical thinking processes (English, 1997). Since learners are usually allowed to pose several kinds of problems in a broad range, it can be challenging for teachers to complete the assessment and feedback for the posed problems in classrooms.

Sophisticated technologies can be used to address this problem. Research on Human-Computer Interaction (HCI) has recently shown promise for the development of emerging intelligent learning systems for exercising posing problems individually and providing automatic assessment and feedback to each posed problem. Furthermore, a learning system named Monsakun, which uses a computerized assessment for posing arithmetic word problems in operations of addition and subtraction has been developed (Hirashima et al., 2007). The system has many problem-posing assignments, and requests learners pose the required problem by combining three simple sentences from given sentences until they successfully pose the required problem in each assignment. An assignment in here is a task for learners to pose a single problem. By using this system, the opportunity to pose the problems for learners increases, the feedback according to their mistakes is provided, and for teachers, checking the validity of the posed problems becomes easier. In practical use and long-term evaluation, it was confirmed that learning by problem posing with Monsakun is interesting and useful as a learning method (Hirashima et al., 2008a). Lectures and exercises with Monsakun improve not only learners' problem-posing skills but also their problem-solving skills (Yamamoto et al., 2012).

The basis of Monsakun is the triplet structure model (Hirashima et al., 2014) that defines the structure of an arithmetic word problem using sentence integration. This model deals with an arithmetic word problem that is solved using only one arithmetical operation. It is the fundamental unit of conceptual quantity representation, and much more complex the arithmetic word problems can be composed of the combination of the units. An arithmetic word problem in this model is an integration of three sentences representing numerical concepts. In addition to that, the model defines constraints for valid problems that must be satisfied. When a learner can pose the required problem in Monsakun, the problem certainly meets the constraints. In other words, posing problems in Monsakun is the division of the task to pose an arithmetic word problem into two sub-tasks: generation and integration of three sentences satisfying the required constraints and the replacement of the generation (sub-) tasks by selecting tasks of sentences to achieve valid problems. It is the same as the concept of the “kit-build concept map” and focuses learner’s thinking processes on the structure of the learning content (Hirashima et al., 2015).

Although the usefulness of Monsakun has been confirmed for learning by posing problems, it is necessary to investigate the validity of the learners' problem-posing process in Monsakun and conduct discovery learning in order to generate inferences of learners' thinking from their behavior in learning systems. Investigation the validity of posed problems reveals learners' consideration and thinking processes in regard to pose the problems. Discovery learning plays a role in increasing learners' motivation while creating more opportunities for learners to assess how well they could overcome obstacles, which may improve learning (Reiser, 1998).

The area of research that can provide with methods to exploit the rich contextual data that can be captured from the learner’s activity on interactive learning system is the emerging field of Educational Data Mining (EDM) (Baker and Yacef, 2009). While Learning Analytics (LA) focuses on the educational challenge, deals with the development of analytics for learning which concentrated on the perspectives of institutions needs such as grades and persistence, and the perspectives of learners related to their needs (Ferguson, 2012). Both LA and EDM reflect the emergence of data-intensive approaches to education and improve the quality of analysis of large-scale educational data, to support both basic research and practice in education (Siemens and

Baker, 2012). It can be used to find patterns of learner's activity associated with strategies or behaviors. These patterns may help produce information indicators of effective strategies or less desired learning outcomes.

Previous research in investigating learners pose the problems on Monsakun analyze the pre- and post- test scores, it suggests that the practice of problem-posing on the system improves the learners' ability not only in problem-posing but also in problem-solving (Yamamoto et al., 2012). Further analyses have been conducted on this topic by investigating the learners' thinking processes based on the first selected sentence in assignments (Hasanah et al., 2015a), concerning the completed posed problems (Hasanah et al., 2015b). This thesis extends the analysis by involving the process of arranging the problem, which means every action of learners when they pose the problems is examined in order to understand the learning process of posing problems on Monsakun. It analyzes problem-posing processes of Japanese first-grade elementary school students in an actual class.

The intersection of the three fields HCI, LA, and EDM, in this context, raises a challenging question. The question is **how can log data of problem-posing activity be automatically analyzed and exploited in the process level through data analytics techniques?** To address this, three sub problems arise from the previous question:

1. *Do learners consider the structure of arithmetic word problem in posing the problems using Monsakun?* This is particularly important for giving confirmation that learners did not pose the required problems randomly on Monsakun. Investigation of the Monsakun log data based on its model is conducted to confirm learners consider the constraints in posing valid problems. A first challenge is to ensure that learners pose problems by satisfying as many constraints as possible. A second challenge is to assure whether learners pose problems by attempting to avoid as many violated constraints as possible and whether learners have difficulty in avoiding a particular type of constraints.
2. *Do the history of learners' activity provide useful information regarding their difficulty in thinking while posing the problems on Monsakun?* The second issue is to find ways to produce key patterns of activity that can be discovered from learner's data. Notably, discovering a situation in which learners got a difficulty

in thinking while posing the problems, in the form of visualizations. A follow-up issue is to analyze the difficulties and its changes in three trials of an assignment.

3. *How can log of interaction data be analyzed and distilled to enhance learner's awareness of the problem's structure of arithmetic word problems on Monsakun?*

Addressing this question includes the analysis based on the activity data and the model to build a computer-based scaffolding system. The scaffolding detects individual learner's bottleneck in real-time while posing a problem and adaptively provide a personalized task based on the bottleneck found. It is aimed to develop the problem's structure knowledge for the future tasks.

## 1.2 Thesis Statement

This thesis aimed at analyzing the practical realization of the Monsakun system to address the three questions described above through the following statement. It embodies the approach of this thesis for supporting learner's awareness of the problem's structure while posing the problems. The thesis statement is stated as follows:

To investigate and discover problem-posing processes through the log of interaction data using statistical and educational data mining techniques for supporting learners enhance their awareness of the structure of arithmetic word problems in Monsakun.

Figure 1-1 (context) lists a set of keywords that can help understand the crossing of the three domains, HCI, LA and EDM, in terms of the thesis statement (in particular, and in no order, interaction log data, problem-posing learning system, statistical data analysis, visualizations, and computer-based scaffolding).

In order to provide support to learners in problem-posing activity, the interactive **problem-posing learning system** should be able to automatically capture learner's activity in a way that it does not interfere or restrict learner's interactions with the technology. It should provide usability and interaction affordances that at least do not produce an adverse impact on learner's posing. Additionally, the structure of **interaction log data** is necessary to be considered because the data structures should be easy to be processed and extracted to obtain the useful information. In this thesis, we used the concept of **statistical data analytics** to refer to the analysis techniques that can be applied

to investigate the interaction data captured from a problem-posing learning system. Moreover, we include **visual representations** as one of the educational data mining techniques. Visualizations could be fully leveraged to better understand via step-by-step data logs generated by learning systems. The last key term that is relevant to this thesis is scaffolding, especially **computer-based scaffolding** (Belland et. al., 2015). We use the results of learning analytics on learning systems to support learners through developing computer-based scaffolding system.

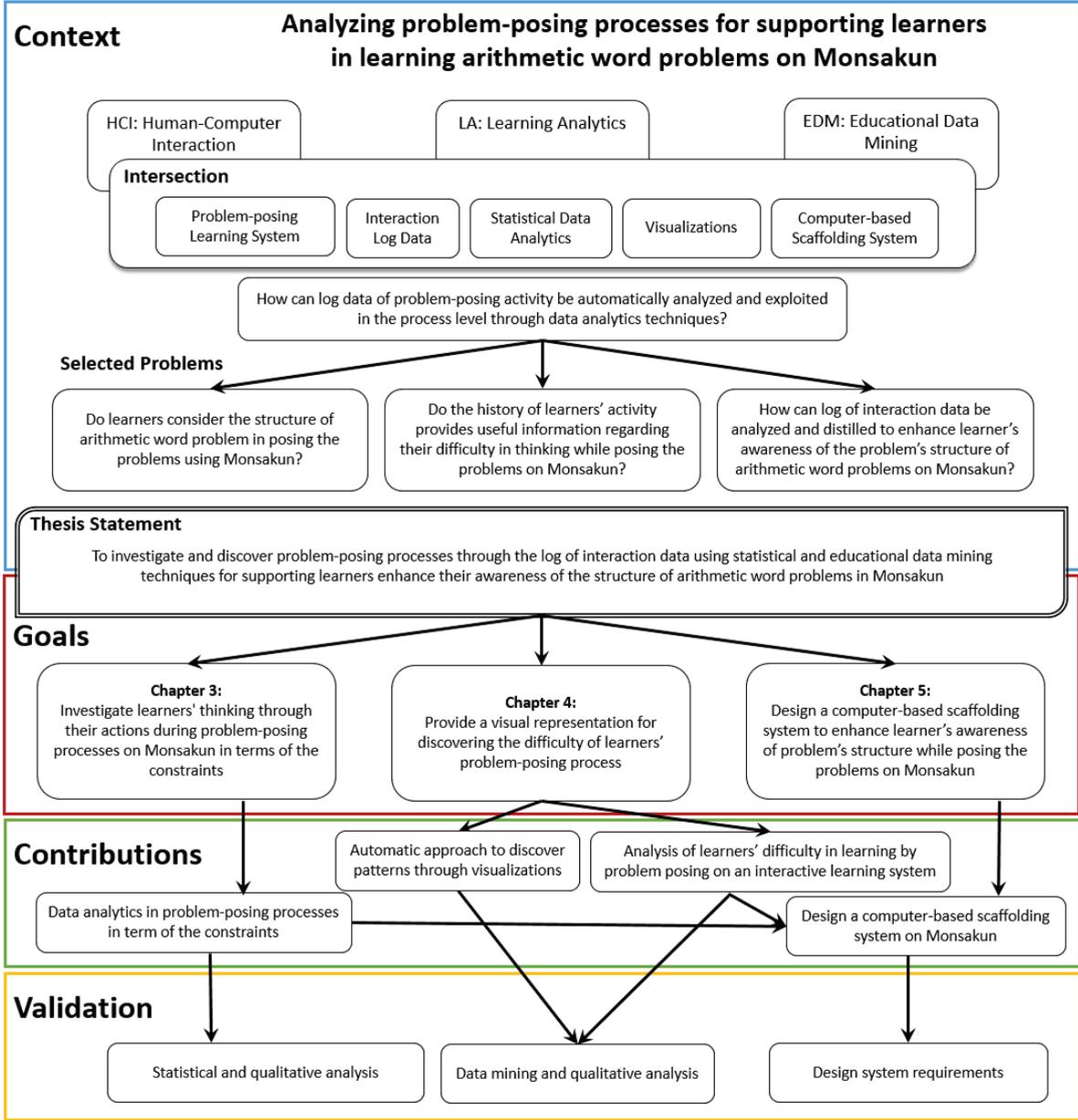


Figure 1-1 Overview of the context, goals, contributions and evaluation of this thesis.

Scaffolding is defined as providing assistance to a learner when needed and fading the assistance as the competence of the learner increases (Wood et al., 1976). This reduces the difficulty of complex learning and at the same time, let the learners focus on constructing knowledge and higher-order demands like thinking critically (Way and Rowe, 2008). The use of technology-based tools, learners received computerized scaffolds supporting their metacognitive activities in the learning process (Molenaar and Roda, 2008). The computerized scaffolding system generated the appropriate instance to send a scaffold based on the learner's attention focus.

### 1.3 Thesis Goals

Having described the research context and stated the thesis statement, we have formulated the main goals of the thesis (see Figure 1-1, Goals):

1. *Investigate learners' thinking through their actions during problem-posing processes on Monsakun in terms of the constraints.* There are previous works on examining satisfaction of the constraints learners in posing problems on Monsakun (Hasanah et al., 2015a, Hasanah et al., 2015b). However, investigations in such research are primarily limited in analyzing the first selected sentence and examining the posed problems as the result of the process. We extend the analysis by involving the process of arranging the problem. This goal addresses our first question: *Do learners consider the structure of arithmetic word problem in posing the problems using Monsakun?* The first study, we traced and investigated every action of learners in posing the problems on Monsakun focused on how learners satisfy the required constraints. Particularly, the goal of the first study addresses a question: *Do learners pose the problems by satisfying as many constraints as possible?* The analysis result is to demonstrate that the learners consider the constraints by satisfying them. In other words, they did not pose the required problems randomly. The second study investigated the problem-posing process and revealed the trends of the process, focusing on the violation of constraints. In particular, this goal addresses a question: *Do learners pose problems by attempting to avoid as many violated constraints as possible?* It was aimed to prove that learners tend to avoid as many violated constraints as possible in composing problems. Furthermore, the fact that learners

gave wrong answers illustrates that learners cannot avoid some mistakes. It emerges another question: *Do learners have difficulty avoiding a particular type of constraint?* For this reason, we detect the difficulty of learners regarding the violation of constraints. It was aimed to prove that learners have difficulty avoiding some specific types of constraints. This exploration and discussion are mainly described in Chapter 3 and in the following peer-reviewed papers (Supianto et al, 2016a; Supianto et al., 2016c; Supianto et al., 2017b; Supianto et al., 2017c).

2. *Provide a visual representation for discovering the difficulty of learners' problem-posing process.* It is challenging to define ways to present the information about learners' behaviors in a manner that is readily understood and useful for teachers and researchers. This goal addresses the second question: *Do the history of learners' activity provides valuable information regarding their difficulty in thinking while posing the problems on Monsakun?* For this, we presented a visual representation for analyzing the problem-posing processes of learners. These visualizations detected learners experience bottlenecks and misunderstanding while posing the problems on Monsakun. This research is described in Chapter 4, and the results of the research associated with this goal were published in the following papers (Supianto et al., 2015a; Supianto et al., 2015b; Supianto et al., 2016b).
3. *Design a computer-based scaffolding system to enhance learner's awareness of problem's structure while posing the problems on Monsakun.* This goal addresses our third question: *How can log of interaction data be analyzed and distilled to enhance learner's awareness of the problem's structure of arithmetic word problems on Monsakun?* This goal calls for the integration of the tracing procedure of learner's activity on problem-posing learning system, the analysis of log data based on the constraints, and the visualization of detecting learners experience bottlenecks in order to discover the individual learner's bottleneck in real-time while posing a problem and adaptively provide a personalized task based on the bottleneck found. It is aimed to support the learner in overcoming the bottleneck and fostering learner's consideration about the problem structure while posing arithmetic word problems on Monsakun. The design and implementation of the scaffolding system described in Chapter 5. Results of the study associated with this goal were presented in (Supianto et al., 2016c) and published in (Supianto et al., 2017a).

## 1.4 Thesis Contributions

The main contribution of this thesis is the investigation and discovering problem-posing processes through the log of interaction data using statistical and educational data mining techniques for developing a computer-based scaffolding system to provide support to learners by enhancing their awareness of the structure of arithmetic word problems in Monsakun.

The subsidiary contributions are listed in Figure 1-1 (Contributions) and can be described as follows:

1. *Data analytics in problem-posing processes in term of the constraints.* From the data analysis aspect, this thesis contributes to the deep exploration of learners thinking in posing problems by applying statistical techniques to confirm learners consider the constraints in posing valid problems.
2. *Visualization of learners' bottleneck in posing the problems.* Another contribution is the discovering learners' difficulty through visualizations of learners' interaction data that can be shown to 1) teachers, in order to provide more helpful and appropriate feedbacks and 2) researchers, in order to consider methods or develop systems that can help learners to overcome their difficulty.
3. *Design a computer-based scaffolding system on Monsakun.* We design a new version of Monsakun. This application presents novel affordances, including tracking interaction log data in the real-time, detecting learner's bottleneck in thinking through evaluating the data collected, and generate a personalized task based on the bottleneck found.

## 1.5 Research Methodology and Validation Methods

In accordance with the research areas involved in each objective of the thesis, we started investigating problem-posing process log data to verify learners' thinking process. We were then exploring available dataset to discover compelling circumstances regarding learners' deadlock in thinking. Promising results and new areas for further development of problem-posing learning system emerged after that exploration, thus motivating the

conception of a novel support function and scaffolding system approach. We followed the next general phases proposed by Glass (1995), each associated with one or more thesis chapters (the order does not indicate strict time sequence, see Figure 1-2):

1. *The informational phase.* A literature survey of the current state of research was conducted on existing problem-posing learning systems, investigation of learners interaction log data, determine appropriate methods to dig up the data collected, and development computer-based scaffolding systems. In addition, we gathered information on the theoretical foundation of the cognitive model of the problem-posing learning system.
2. *The analytical phase.* In this phase, the conceptual solution was formulated based on the literature review and exploratory of previous research. Methods of technological approaches were conducted, and substantial dataset was collected that further analyzed mainly through statistical analysis and data mining techniques.

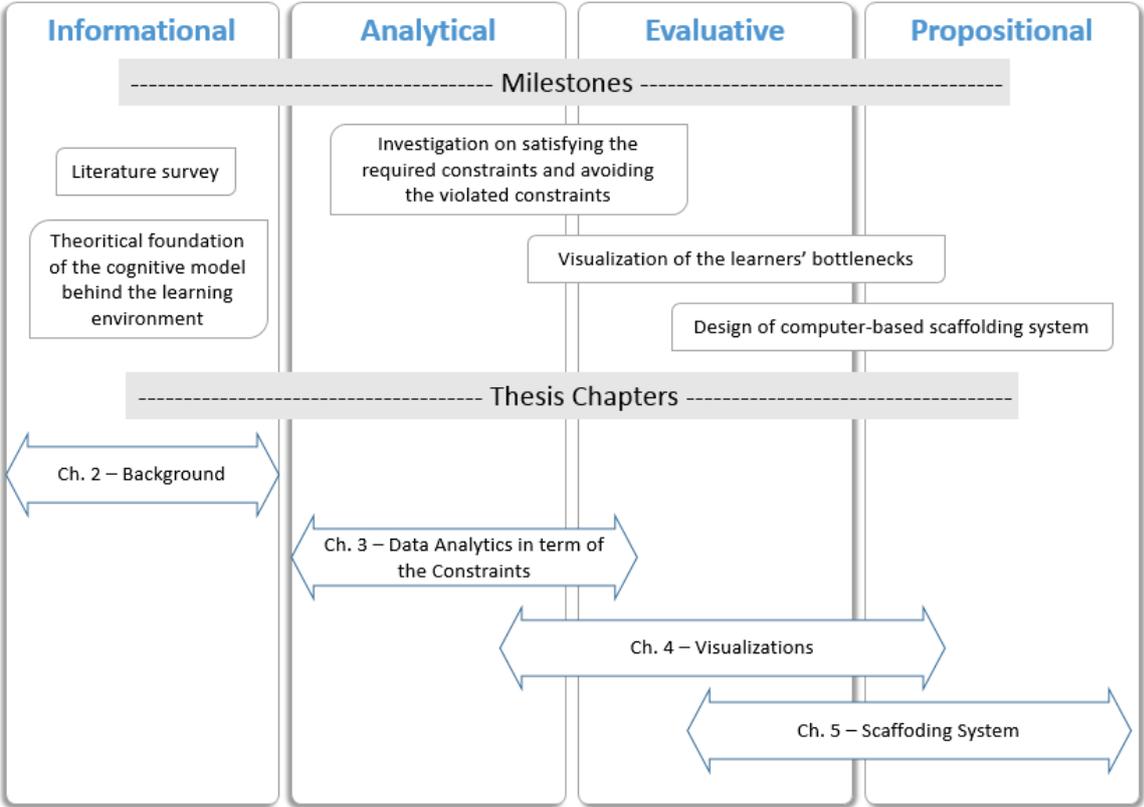


Figure 1-2 Research phases: associated milestones and thesis chapters.

3. *The evaluative phase.* The investigation and evaluation of the dataset were conducted in this phase. The validation of our approach was mainly carried out by performing quantitative evaluations with qualitative assessments to aspect specific research questions.
4. *The propositional phase.* The conceptual framework of support system is stated based on the analytical and evaluative phases that are described above. The design, implementation, and validation of the technological infrastructure in supporting learning were conducted. As a result, the computer-based scaffolding system was designed and developed.

Our main validation approach consisted of collecting, analyzing, and discovering data in a series of studies to provide a better understanding of the research problem (Creswell, 2013). In most of these studies we analyzed collected data quantitatively, and qualitative information discussed to confirm that the trends that might be automatically discovered are meaningful in problem-posing learning contexts. This mixed method served to evaluate the different aspects of our approach and the developed system. The specific validation methods and their relationships with the contributions of the thesis are listed in Figure 1-1 (Validation), and can be described as follows:

1. *Statistical and qualitative analysis.* Two studies were conducted to analyze the historical data of learners' posing activity. Correlation and difference analyses were performed to find the trends that can be associated with learners' strategies and problem-posing processes. Then, a qualitative analysis was used to validate the meaning of the results according to the cognitive model of the learning system.
2. *Data mining and qualitative analysis.* A study was proposed to explore learners' posing data. A data mining approach was implemented to find patterns that indicate learners difficult on understanding the structure of the problems. A qualitative analysis was also discussed to validate the meaningfulness of the patterns discovered.
3. *Design system requirements.* The scaffolding on problem-posing learning system for support learning was iteratively built and validated according to the system requirements before being used in the tablet PC setting.

## 1.6 Thesis Structure

This section describes the chapters of the thesis. Figure 1-3 illustrates the structure of the thesis and the publications associated with each chapter.

*Chapter 1* – Introduction - describes the research context and outlines the goals, contributions, evaluation methods, and the general structure of the thesis.

*Chapter 2* – Background - outlines relevant research on learning by problem posing, the growing usage of learning system for posing problems, the area of log data analysis, and the exploration of previous work and the cognitive model related to the learning system mainly used in this thesis: Monsakun Touch.

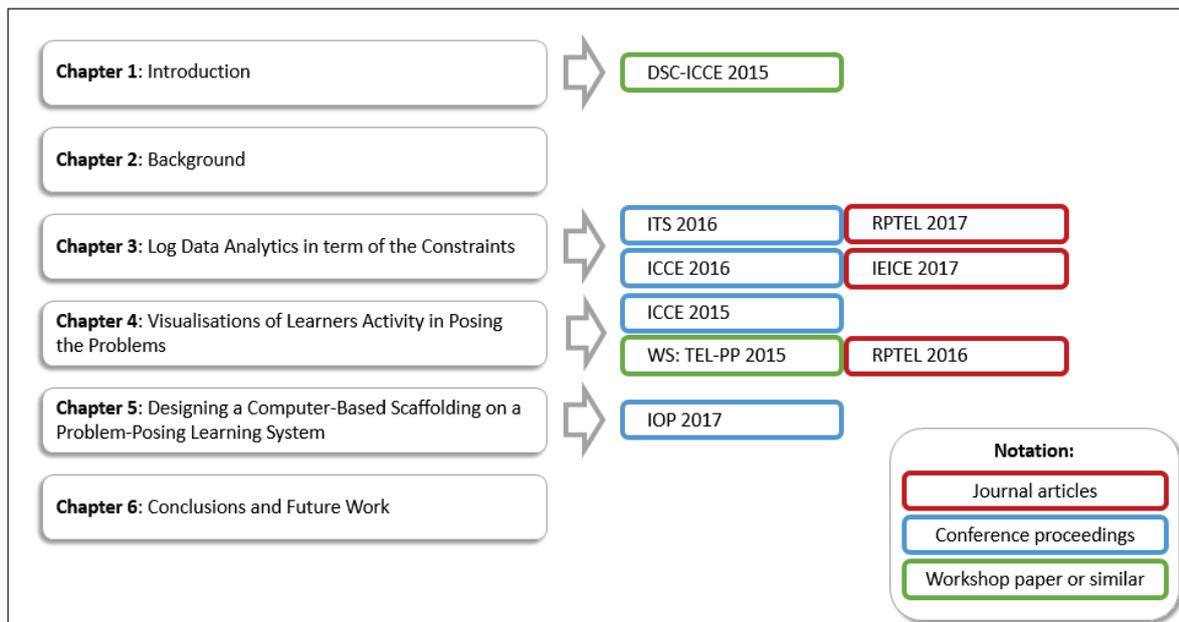


Figure 1-3 Structure of the research covered in this thesis and related published papers.

*Chapter 3* – Log Data Analytics of the Learners Thinking in term of the Constraints - describes the process analysis of log data of learners posing the problems on Monsakun. The chapter investigates whether learners did not think randomly in posing the required problems. This chapter focused on examining the problem constraints composed by learners. It presents statistical analysis according to the learners in which they attempt to satisfy the required constraints and to avoid the violated constraints. In case they violated the constraints, this chapter shows further analysis to investigate whether learners tended to have difficulty avoiding a particular type of constraint.

*Chapter 4 – Visualizations of Learners Activity in Posing the Problems* - presents visual representations of learners' problem-posing process for discovering the bottlenecks in thinking while they pose the problems. The study in this chapter traces sequences of learners' actions and provide insight in what type of condition learners got stuck. Additionally, the chapter describes the evaluation of the visualizations and the discussion of analysis result according to the cognitive model involved.

*Chapter 5 – Designing a Computer-Based Scaffolding on a Problem-Posing Learning System* - provides a detailed view of the reasoning behind the design of our computer-based scaffolding system. From the technology perspective, the system is the goal of the main studies in this thesis. The chapter presents the articulation of the problems based on the log data collected, system architecture, and operation procedures of the scaffolding system. Additionally, this chapter describes the discussion of the proposed system related to the problem constraints.

*Chapter 6 – Conclusions and Future Work* – revisits the studies presented in this thesis and describing the promising research avenues for future studies.

*<empty page>*

---

## BACKGROUND

**Summary:** This chapter reviews the state of research at the intersection of the multiple disciplines that are the focus of this thesis. First, we present the previous research work on learning by problem posing and how it relates to mathematical thinking and understanding. We discuss the importance of problem-posing activities in mathematics. Then, the current state of research of problem-posing learning systems is discussed. We review the current affordances and the learning effects of such technology to promote learning by problem posing. Furthermore, we present detail explanation of the triplet structure model which defines the structure of arithmetic word problems and Monsakun as the problem-posing learning system used at practical realization in posing arithmetic word problems, which is the target of analyses in this thesis. Next, we review work on the analysis of problem-posing activity through statistical or data mining techniques on the learners' log data. As the analysis results of studies in this thesis are used to help learners in mastering arithmetic word problems, we present the concept of scaffolding systems to support learning and the research upon them.

### 2.1 Learning by Problem Posing

Understanding, more than knowing or being skilled, has always been considered an important goal in learning mathematics. Understanding, as it happens, is a process occurring in the learner's mind; more often, it is based upon a long sequence of learning activities during which a great variety of mental processes occur and interact (Dreyfus, 2002). Activities which could provide us with valuable insights into children's understanding of mathematical concepts and processes, as well as their perceptions of, and attitudes toward mathematics in general, is problem-posing activities (Brown and

Walter, 1993). Problem posing is considered to be an essential part of mathematical activities. Problem posing can also promote a spirit of curiosity and more diverse and flexible thinking (English, 1997). It is recognized as a key component in the nature of mathematical thinking (Kilpatrick, 1987). Devlin (2012) asserted that mathematical thinking is not the same as doing mathematics, which usually involves the application of procedures and some heavy-duty symbolic manipulations so that it learns to think inside-the-box. In contrast, mathematical thinking is a specific way of thinking about things in the world includes logical and analytic thinking as well as quantitative reasoning so that it learns to think outside-the-box. Learners who are engaged in problem-posing activities become enterprising, creative and active learners. Therefore, the development of problem-posing skills for learners is one of the principal aims of mathematics learning, and it should occupy a central role in mathematics activities (Crespo, 2003).

### 2.1.1 Problem Posing Ability and Processes

There is an increased emphasis on providing learners with opportunities for posing problems in the mathematics classroom (Cankoy, 2014; Singer et al., 2011; Stoyanova, 2005). Research studies provided evidence that problem posing has a positive influence on learners' ability to solve word problems (Leung, 1996; Leung and Silver, 1997; Silver, 1994), and confirmed that learning by problem posing in classrooms is a promising activity in learning mathematics (English, 1997a; English, 1997b; English, 1998; Silver and Cai, 1996). Silver and Cai (1996) found that students generated a significant number of solvable mathematical problems, many of which were syntactically and semantically complex, and that nearly half the students made sets of related problems. English (1997a) asserted that problem posing improves learners' thinking and confidence in mathematics and contributes to a broader understanding of mathematical concepts. English (1997b) found that learners who participated in specific programs on problem posing created solvable problems and most of them created quite sophisticated problems using semantic relations in their problems. English (1998) investigated the problem-posing abilities of children who displayed different profiles of achievement in number sense and novel problem-solving.

Given the importance of problem-posing activities in mathematics, some researchers have investigated various aspects of problem-posing processes. One important direction is to examine thinking processes related to problem-posing activities (e.g., Bonotto, 2013; Şengül and Katranci, 2015). Other studies underline the need to incorporate problem-posing activities into mathematics classrooms to determine prospective teachers' problem-posing skills appropriate to selecting, translating, comprehending, and editing models and the possible difficulties they could encounter during the process in fraction problems (Işık et al., 2011), to explore students' creativity in mathematics by analyzing their problem-posing abilities in geometric scenarios (Van Harpen and Sriraman, 2013), and to examine the knowledge influences of learners' abilities in posing combinatorial problems (Melušová and Šunderlík, 2014). Furthermore, some studies provide evidence that problem posing has a positive influence on students' abilities in problem solving (e.g., Kar et al., 2010; Sengül and Katranci, 2012). Kar et al. (2010) asserted that the positive relation between posing and solving problems is an indicator of the acceptance of problem-posing skills as a phase in the development of problem-solving skills. In the analysis of the posed problems, the participants map the level of their own notions and concepts, understanding, and various interpretations and realize possible misconceptions and erroneous reasoning (Tichá and Hošpesová, 2009). Learning to pose problems might also enhance learning to understand mathematical concepts (Pirie, 2002). Pirie (2002) said that, in asking questions on mathematical concepts, students might come to understand those concepts in a more generalized, less context-dependent way. In addition, Toluk-Uçar (2009) emphasized that problem posing has a positive effect on understanding fractions as well as on learners' views about what it means to know mathematics.

### 2.1.2 Classification of Problem Posing Task

Although there is a wide variety of problem posing tasks (Silver and Cai, 1996), research so far indicates only a few ways to classify them. According to Hershkovitz and Nachmias (2009), the variety of analyses of problem-posing in educational research is classified into the subject of research and time reference. The subject of research might focus on the individual learner or a group of learners, and time reference depicts a learning process might be analyzed as end-point level (i.e., from a summarizing point of view) or at a

process level that describing various behaviors/activities that occur in the course of it. Moreover, Stoyanova (1998) identified three categories of problem-posing experiences that increase students' awareness of different situations to generate and solve mathematical problems:

- (1) *Free situations*, refer to situations in which learners pose problems without any restriction; they can generate a new problem utilizing a situation from daily life or a given subject.
- (2) *Semi-structured situations*, refer to ones where learners are asked to pose problems, which are similar to given problems or to pose problems based on specific illustrations. Learners are given an open-ended situation; then they are requested to pose problems that are similar to given problems or to write problems based on specific pictures and diagrams in order to utilize their knowledge, skills, and experience.
- (3) *Structured problem-posing situations*, refer to situations where learners pose problems by reformulating provided solved problems or by varying the conditions or questions of given problems. In structured problem-posing strategy; the known can be changed to pose a new problem or, the needed can be modified by maintaining the data presented.

## 2.2 The Use of Learning Systems for Posing Problems

With the rapid advancements in information technology and the discovery of the boosting effect of problem posing on problem-solving abilities and motivation in mathematical learning, researchers have attempted to develop problem-posing learning systems by applying information technology. A learning system called the Question Authoring and Reasoning Knowledge System (QuARKS) (Yu, 2009), which allows learners to contribute to and benefit from the process of question construction and peer feedback regarding composed questions in a cyclic manner, was developed. Essentially, QuARKS is comprised of two subsystems: question authoring and question reasoning. In the question authoring phase, learners generated questions activities, such as true-false, matching, fill-in-the-blank, multiple-choice, short-answer, and so on. All questions are kept in an item bank database, waiting to be evaluated by peers and redefined by the

author in the follow-up question reasoning phase. In the question reasoning phase, QuARKS was devised in such a way as to allow interacting counterparts to communicate easily, in writing, about specific question items to enhance interaction, collaboration, and negotiation of meaning between question-composers and their peers (assessors).

Another approach is called Animalwatch Web-based Environment (AWE) (Arroyo and Woolf, 2003), an on-line content authoring system that supports the creation of word problem “adventures” for mathematics. AWE is an intelligent tutoring system that teaches mathematics with word problems, which are integrated into narratives about endangered species to engage student interest and help them appreciate the value of learning math. Students authored one episode for their adventure and called it something along the lines of “Meet the (endangered species).” Students then authored addition, subtraction, and multiplication word problems related to the number facts that had collected for their species. Teacher supervision of the resultant word problems was then held in order to clarify students’ misconceptions.

On the other hand, a computer-based learning system that promotes the problem-posing activity of arithmetic word problem as sentence integration was developed (Hirashima et al., 2007). In posing problems via sentence integration, several simple sentences are provided to a learner. The learner, then, selects the necessary sentences and arranges them in an appropriate order. This approach makes straightforward and goal-oriented problem-posing tasks possible even for lower elementary students while maintaining its value as a viable learning method and practical approach to data collection in an interactive learning system. Studies in this thesis conducted an analysis of learners problem-posing activity uses this system.

### 2.2.1 Problem Posing Techniques on Learning Systems

Several problem-posing techniques on interactive learning systems have been conducted. One approach is using the question-posing technique. The systems allow students to generate different types of questions using various media formats with peer-assessment using one type of communication mode (Wilson, 2004) and multiple peer-assessment modes (Yu, 2011). The studies evaluated students’ abilities to pose questions and their processes in an online learning system. Lan and Lin (2011) developed a system

integrating a reward mechanism into assessment activities and analyzed student's abilities to pose questions in a web-based learning system. Moreover, Hung et al. (2014b) investigated the effect of promoting questioning ability in problem-based scientific inquiry activities. The research developed a ubiquitous problem-based learning system regarding learners' question-raising performance.

The second approach is learning from the example technique. This support system is developed to facilitate posing of diverse problems by learners using examples. Leikin (2015) described posing various types problems associated with geometry investigations using examples from a course with prospective mathematics teachers, while Hsiao et al. (2013) conducted examples across three homework exercises in which students were required to generate at least one applied problem. The studies showed that integrating worked examples into problem posing has a significant skill development effect on posing more oriented and complex problems. Moreover, Kojima et al. (2015) presented examples that are merely shown to the learners and prompted them to compare the base with their posed problems. They investigated the effects of learning from an example on solution composition for posing problems.

Another approach is learning by problem posing as sentence integration. Problem posing as sentence integration requires learners to interpret the sentence cards and integrate them into one problem. In an assignment, the system presents a requirement, which consists of a story type and a numerical expression. The system asks learners to arrange the provided sentence cards based on the requirement. The use of the sentence integration method was proven to improve learning by problem-posing for students in lower elementary school. A long-term evaluation of the system was carried out, and the study confirmed that it was interesting and useful for learning (Hirashima et al., 2008a). Moreover, the system also improved the problem-solving ability of low-performance students (Hirashima et al., 2008b). In 2011, a task model of problem-posing was developed that dealt with not only the forward-thinking problem but also the reverse-thinking problem (Hirashima and Kurayama, 2011). Yamamoto et al. (2012) implemented the system on a tablet-PC platform so that the teacher was able to use it in the usual classroom. Then, they enhanced by developing online connected media tablet and reported that utilized this approach improved learners' abilities not only in problem

posing but also in problem solving (Yamamoto et al., 2013). Finally, an interactive learning system for learning by problem-posing based on the triplet structure model was developed and practically used (Hirashima et al., 2014). In practical uses, it was confirmed that learning by the system was a useful learning method. Studies in this thesis used the learning system that the above studies used, which is problem posing of arithmetic word problems with sentence integration based on triplet structure model. It used a tablet-PC platform, which learners can practice the system in the classroom individually.

### 2.2.2 Assessment and Investigation of the Posed Problems

In problem-posing, assessment of each posed problem and assistance based on it are necessary (Hirashima et al., 2007). Self- and peer-assessed posed problems were examined to determine the effect of learners' self-assessment of their mathematical creativity (Shriki and Lavy, 2014), to explore learner's learning and knowledge sharing while engaged in an online question-posing and peer-assessment activity (Barak and Rafaeli, 2004), and to determine which peer-assessment mode(s) students perceive most positively using student generation of questions (Yu, 2011). Moreover, an online learning system with a focus on student-question generation strategy (Yu, 2009), was adopted and the effects of student question-generation on civics and citizenship (Yu and Pan, 2014) and English learning (Yu et al., 2015) have been reported. Student academic achievement, question-generation performance, learning satisfaction and learning anxiety, as well as learning motivation have been investigated. Self- and peer- assessed posed problems were also examined in these studies. In contrast, diagnosis functions that can assess and give automatic feedback to each posed problem, especially in learning arithmetic word problem, have been proposed (Nakano et al., 1999; Hirashima et al., 2000). This automated way of diagnosis-facility assessment is called agent-assessment. Furthermore, a computer-based learning system that uses practical agent-assessment has been developed (Hirashima et al., 2007).

Investigations of problem posing from the viewpoint of interactive learning systems promote active engagement in learning through the activities of learners. Chang et al. (2012) developed game-based problem-solving modules in a mathematics problem-

posing system and investigated the effects of the problem-posing system on students' abilities to pose and solve problems. Yamamoto et al. (2012) and Abramovich and Cho (2015) demonstrated how the appropriate use of digital technology tools could motivate problem-posing activities and evaluate the learner's performance by assessing the number of posed problems. Hung et al. (2014a) investigated the effects of an integrated mind mapping and problem-posing approach on learners' in-field mobile learning performance in an elementary school natural science course. Moreover, Majumdar and Iyer (2015) presented how an online visual analytic tool can be used to analyze clicker responses during an active learning strategy where the instructor poses a multiple-choice question. Studies in this thesis, an interactive learning system is used to encourage learners in posing arithmetic word problems. The system asks learners to arrange and integrate five or six presented sentence cards into a problem, which consists of three sentence cards. We analyze at the process level to find out the learners' tendencies while posing the problems in the system.

Several studies examined learners' behaviors through a collaborative problem-posing strategy. Beal and Cohen (2012) demonstrated that the mathematics problem-posing skill was improved when the activity was carried out over an online collaborative learning system. Mishra and Iyer (2015) implemented a collaborative problem-posing activity in which two learners collaborated as a team to generate questions. Sung et al. (2016) conducted a group collaborative problem-posing mobile learning activity. They found that such an approach could improve learning achievement and group learning self-efficacy. Studies in this thesis analyze log data of learners' activity collected from a tablet personal computer-based software for learning by posing arithmetic word problems.

### 2.3 The Problem-Posing Learning System based on the Triplet Structure Model

Having reviewed the use of learning systems for posing problems and described the techniques among them, this section describes one of problem posing techniques in more detail. We explain the triplet structure model, the task model, the constraints based on the task model, and the learning system that have been developed based on such model.

Following discussion is focused on the model in arithmetic word problems of addition and subtraction which is used in the studies of this thesis. One of the main reasons that arithmetic word problems are so crucial to learners learning mathematics is the way they show how mathematical concepts apply to real-world situations. This is in line with the purpose of mathematical thinking that involves logical and analytical thinking about things in the world. In addition, arithmetic word problems require higher order thinking because word problems require more than simply looking at numbers and symbols and figuring the solution. Learners must be able to read the problem, get the related information out of the problem, and make sure the answer makes sense in the context of the problem.

### 2.3.1 The Triplet Structure Model

Triplet structure model, as shown in Figure 2-1, describes the components of arithmetic word problems and the basic structure of them (Hirashima et al., 2014). In this model, an arithmetic word problem is defined that it consists of three sentences including different quantities. Each sentence has to represent only one quantity with the meaning of them in the story. The three sentences include two “independent quantity sentences” and one “relative quantity sentence.” Independent quantity sentences describe numbers of objects, for example, “There are 3 white rabbits.”, “There are 5 black rabbits.” and so on. Relative quantity sentences describe the relationship between the other independent quantity sentences, for example, “There are 8 white and black rabbits altogether.”, “2 white rabbits come.” and so on. Although an independent quantity sentence can be used in any story type, a relative quantity sentence is used only in one particular story type. Therefore it contains keyword determining the type of story, for example, “...altogether,” “...come,” “...go away,” “...less than...” or “...more than...”

There are four story types in arithmetic word problems of addition and subtraction: 1) combination, 2) increase, 3) decrease, and 4) comparison. The differences among them are defined as differences of integration of sentences. For instance, a combination story type problem, one independent quantity sentence describes the quantity of an object, and the other independent quantity sentence represents the quantity of another object. The relative quantity sentence expresses the total quantity of the two objects. The following

example is a typical arithmetic problem that is expressed by the triplet structure model in a combination story:

- (1) There are 3 white rabbits (first independent quantity sentence).
- (2) There are 5 black rabbits (second independent quantity sentence).
- (3) There are 8 white and black rabbits altogether (relative quantity sentence).

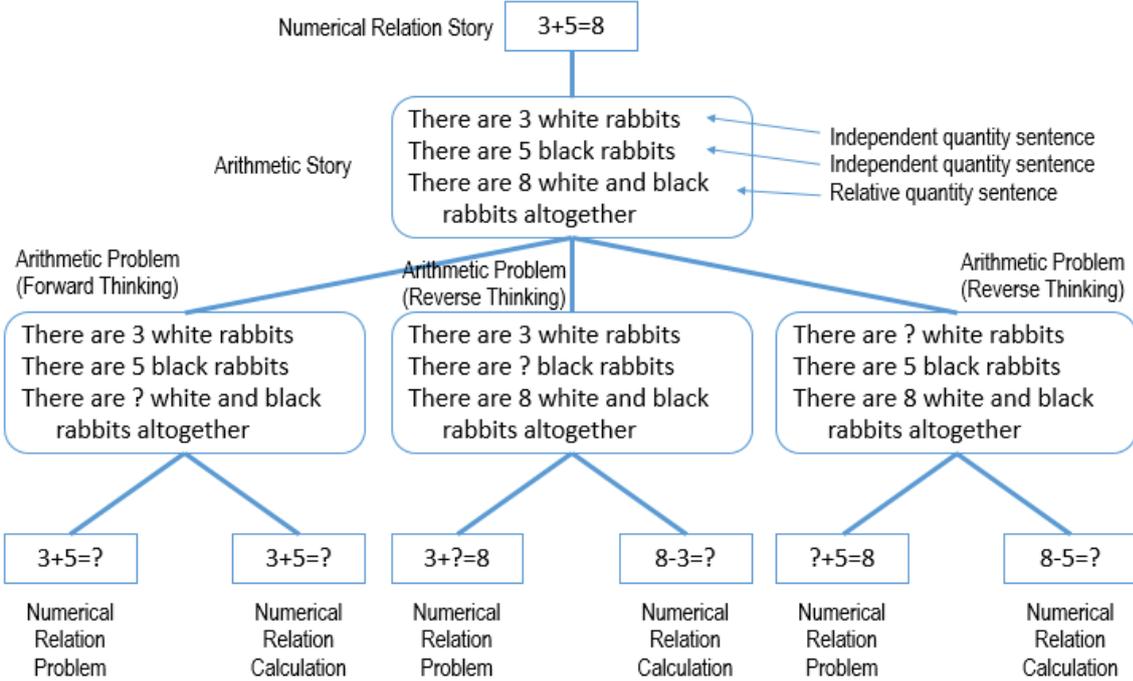


Figure 2-1 Relation between arithmetic story and problem.

In this story, the changes to the order of the sentences do not affect the problem story. For example, when the sentence “There are 8 white and black rabbits altogether,” is placed at the beginning, followed by “There are 3 white rabbits,” and ends with “There are 5 black rabbits,” this new composition still forms the problem story. This situation also applies to the comparison story type.

Although increase and decrease stories are also composed of two independent quantity and one relative quantity sentence, a different role is implemented. One independent quantity sentence describes the quantity before an increase or decrease, and the other independent quantity sentence represents the quantity after an increase or decrease. Each independent quantity sentence only describes the quantity of an object.

The relational sentence describes the quantity of the increase or decrease. The relative quantity sentence expresses the relation between before and after the quantity of the increase or decrease. The following is a typical arithmetic word problem that is expressed by the triplet structure model in a decrease story:

- (1) There are 5 white rabbits (first independent quantity sentence).
- (2) 2 white rabbits go away (relative quantity sentence).
- (3) There are 3 white rabbits (second independent quantity sentence).

In this case, any change in the sentence order will affect the problem story. For example, if the relative quantity sentence “2 white rabbits go away” is placed at the beginning, this composition cannot form the problem story. It is wrong in the story when suddenly two white rabbits go away without explaining their prior existence. The initial quantity of white rabbits is required.

In this model, depending on the combination of two independent quantity sentences and one relative quantity sentence, the role of each sentence is changed. The relation between an arithmetic story and other problems is shown in Figure 2-1. Based on the answer to the arithmetic word problem, it is possible to make a numerical relation and a cover story composed of all known numbers. We call this cover story as an “arithmetic story,” and this numerical relation as a “numerical relation story.” Then, the numerical relation in the problem including the unknown number is called a “numerical relation problem,” and a numerical relation used in the calculation is called a “numerical relation calculation.” When the numerical relation problem is the same as the numerical relation calculation, understanding the cover story is the same as solving the problem. We call such a problem a “forward-thinking problem.” However, when the numerical relation problem is different from the numerical relation calculation, it is necessary to transform the numerical relation problem to a numerical relation calculation after understanding the cover story. We call such problem a “reverse-thinking problem.” Because it is required to comprehend the relation between two structures, reverse-thinking problems are more severe than forward-thinking problems.

### 2.3.2 The Constraints based on the Task Model

The task model of problem-posing via sentence integration was developed based on the consideration of problem types in the triplet structure model (Kurayama and Hirashima, 2010). Based on the task model, we have devised five main constraints to be satisfied by each posed problem, which are:

- (1) *Calculation* – which is the numerical expression representing the story type. Calculation structure requires numbers assigned to the correct sentence structure, whether an independent quantity sentence or relative quantity sentence according to the story type.
- (2) *Story type* – which is one of the four available story types. They are combination story, increase story, decrease story, and comparison story. The story type should be identified in the problem requirement.
- (3) *Number* – which is the quantity in the sentence. Number structure requires the consistency of numbers in the problem. Each number in the problem must be derived from the other numbers.
- (4) *Objects* – which is the entity in the sentence. Object structure also requires the consistency of entities in the sentences. For example, if the story type is increase or decrease, the objects in the three sentences must be the same. On the other hand, if the story type is combination or comparison, objects in the independent quantity sentences are different, and both are in the relative quantity sentence.
- (5) *Sentence structure* – which is the composition of sentences. As defined in the triplet structure model, an arithmetic word problem must consist of two independent quantity sentences and one relative quantity sentence. The type of relative quantity sentence is related to the story types.

When a posed problem satisfies all five constraints, the required problem in the assignment is successfully posed. When a posed problem satisfies less than five constraints, the posed problem partially fills the requirements, and the unsatisfied constraints represent the cause of the inadequateness for the requirements. This also means that the posed problem is not meaningless because it still satisfies some constraints. When a posed problem satisfies no constraint, the problem is meaningless.

An example of several compositions of sentence cards and their satisfaction of constraints are presented in Figure 2-2. The requirement of this assignment is to *make a story problem about “How many are there overall” that can be solved by “8 - 3,”* which is an arithmetic word problem with a combination story type. There are six available sentence cards composed of the following:

- |   |  |
|---|--|
| (1) There are 3 white rabbits;                      | (4) There are 8 white rabbits;                             |
| (2) There are ? black rabbits;                      | (5) There are 3 more white rabbits than black rabbits; and |
| (3) There are 8 white and black rabbits altogether; | (6) There are 3 brown rabbits.                             |

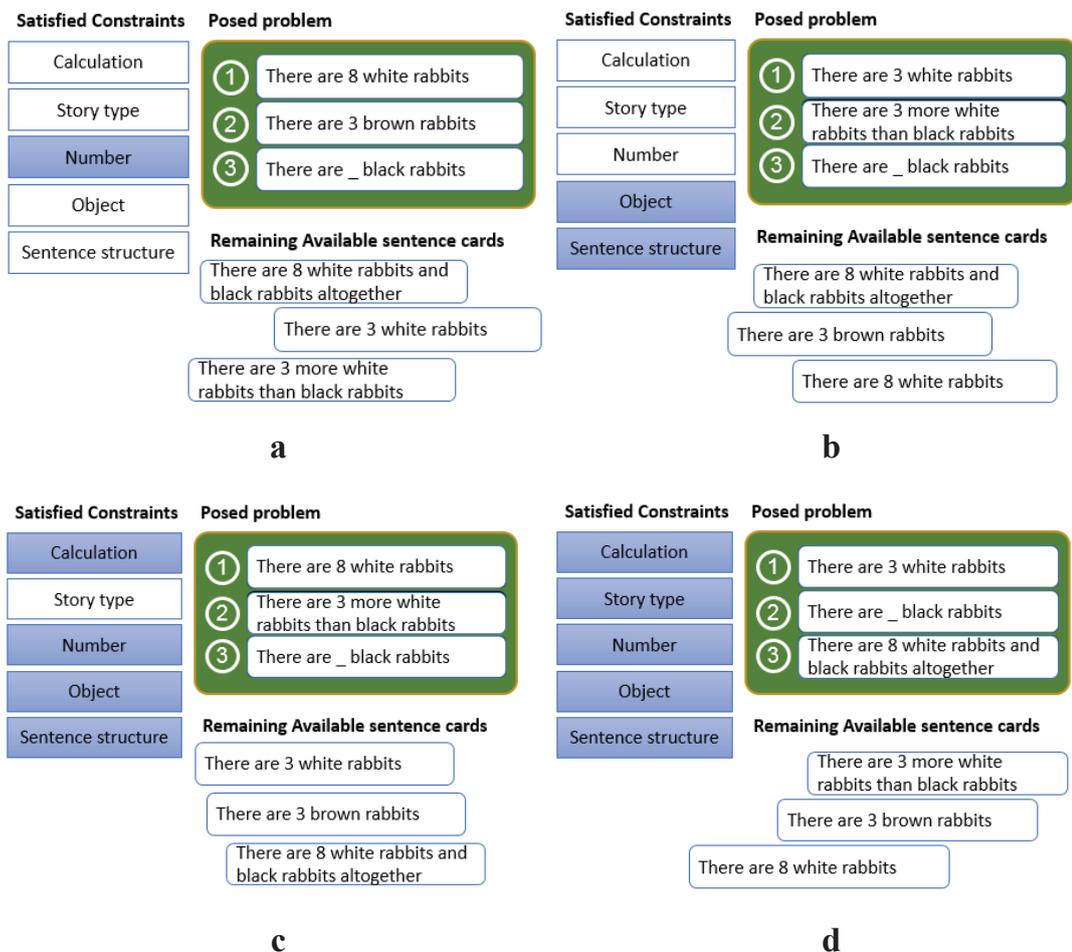


Figure 2-2 Example of several compositions of sentence cards and their satisfaction of constraints. a-c compositions that satisfy part of constraints. d composition that satisfies all constraints.

The first example only satisfies one constraint, number (see Figure 2-2a). In this case, the learner only focuses on the number. No story can be built from this composition, nor the calculation and sentence structure. It can be calculated and well-structured when it consists of two independent quantity sentences and one relative quantity sentence, instead of only independent quantity sentences. Also, there is no relation between objects in the composition. They are independent objects that consist of white, black, and brown rabbits. Figure 2-2b shows that the composition satisfies two constraints: object and sentence structure. There is a relationship between objects (*white and black rabbits*), and the structure of the sentence cards consists of two independent quantity sentences and one relative quantity sentence. However, the calculation, the story type, and the number are not fulfilled. There is no number “8,” as it causes a calculation process that cannot be done and the number constraint is not satisfied. Concerning the story type, this is a comparison story, instead of a combination story. The third example is an example of a state that satisfies all constraints. The third example satisfies almost all constraints (see Figure 2-2c). Only the story type is not satisfied. It is because the story type of the composition is a comparison story type. In contrast, the requirement is to pose the problem in a combination story type. Hence, the valid composition that is satisfying all constraints is shown in Figure 2-2d.

### 2.3.3 Monsakun as a Problem-Posing Learning System

Monsakun has been developed as an interactive learning system for learning by problem-posing with sentence integration based on the triplet structure model. The interface of Monsakun is shown in Figure 2-3. It consists of three components: the problem composition area, the sentence cards, and the diagnosis button. Monsakun gives learners problem-posing exercises, in which they pose a problem that satisfies certain requirements. The problem composition area on the left side of the interface consists of the requirement part and the card slot part to pose a problem. The requirement part includes two types of requirements: (1) a story type that the required problem to be posed must belong to and (2) a numerical expression that must represent the numerical relation in the required problem. Learners try to formulate the required problem to put sentence cards in the card slot at the bottom of this area. Sentence cards are presented on the right side of the interface. Learners can move the cards by dragging and dropping them freely

to the card slots. There are more than three cards provided to the learners, which means that not all the cards included are necessary to complete the required problem. We call such cards as “dummy cards,” and they are intentionally included to test learners and check for their understanding. For example, an overlooking of the story type or confusion about the formula may lead to problems completing the exercise correctly. The last component is a button located under the problem composition area called the diagnosis button. The diagnosis button is used to check the composition of sentence cards chosen by the learner.

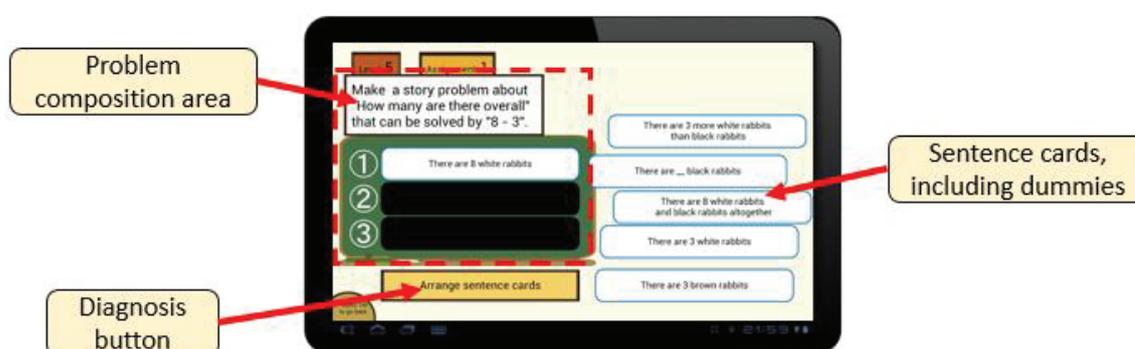


Figure 2-3 Interface of Monsakun.

Monsakun has five levels of assignments (the sixth level is random), that require different thinking approaches. All levels are the same in terms of posing problems from a card set, but they have different requirements. Table 2-1 shows the setting of assignments at each level. Levels 1–4 provide the numerical expression of the story, while Level 5 is required to consider the unknown number. An assignment is completed when learners pose the problem correctly. As a feature of Monsakun, each time a learner makes a mistake, the system will provide explanation feedback according to the error. This feedback will stimulate the learner to think about the other solutions and lead them to the correct answer.

In the problem-posing activity using Monsakun, learners do not create their own problem statements; however, they are required to interpret the sentence cards and integrate them into one problem in the card slot part. Therefore, this activity is called “problem-posing as sentence-integration” (Hirashima et al., 2007). The system provides

a set of sentence cards and a numerical expression in the requirement part, and then learners pose an arithmetic word problem based on the triplet structure model using the numerical expression by selecting and arranging appropriate sentence cards. Monsakun records learners' problem-posing activity as a combination of sentence cards in the card slots. The log data that stored in the system is the result of selecting and arranging a sentence card in the card slot or removing a sentence card from the card slot in a sequence. Studies in this thesis mainly investigate and discover the learners' problem-posing activity while posing the problems through the system's log data.

Table 2-1 Detailed level assignments in Monsakun.

<b>Level</b>	<b>Number of assignments</b>	<b>Type of thinking problem</b>	<b>Type of numerical relation</b>	<b>Story types</b>
1	12	Forward	Problem	Combination, increase, decrease, comparison
2	3	Forward	Problem	Increase and combination
3	12	Reverse	Problem	Combination, increase, decrease, comparison
4	3	Reverse	Problem	Increase and combination
5	12	Reverse	Calculation	Combination, increase, decrease, comparison
6	Random			

## 2.4 Analysis of Learning Activity on the Learning Systems

Even though students seem to be highly engrossed in learning activities using computer or tablet, Dynarski et al. (2007) show not much evidence of the software influence on the higher performance of math and reading in the students. One of the few research studies that have been found in the analysis of learning activity on learning systems is about analyzing the results. Conducting pre- and posttest is the most common way to evaluate a learning system as seen in Beal et al. (2010), Chang et al. (2012), and Oliveira Chaves et al. (2015). Another way is to conduct a deep analysis of the students' behavior as seen in Biswas et al. (2005; 2010).

In term of learning by problem posing, a direction of analysis is about analyzing the results of the posed problems. Hirashima et al. (2007) examined whether learners could pose the problems, showing and discussing the number of posed problems and correct problems based on the system log data. Hirashima and Kurayama (2011) analyzed the learning effects by comparing pre- and post-test problem-solving and problem-posing scores. Further analyses have been conducted on this topic by investigating the learners' thinking processes based on the first selected sentence in assignments (Hasanah et al., 2015a) concerning the completed posed problems (Hasanah et al., 2017). There is a dearth of research that investigates every action of learners in posing the problems to understand the learning process of problem posing on an interactive learning system. Moreover, no significant research has been found that examines the intermediate products while posing the problems.

There has been considerable thorough and fine-grained investigation of the activities of learners in interactive learning systems to reveal their behavior throughout the learning process. Fournier-Viger et al. (2010) developed a virtual learning system for learning how to operate the Canadarm2 robotic arm on the International Space Station. The study extracted patterns from learners' solutions to problem-solving exercises for automatically learning a task model that can then be used to aid and guide them during problem-solving activities. Hou (2012) utilized an online discussion activity adopting a role-playing strategy and conducted an empirical analysis to explore and evaluate both the content structure and behavioral patterns in the discussion process. The study adopted a new method of multi-dimensional process analysis that integrates both content and sequential analyses, whereby the dimension of interaction and cognition are analyzed simultaneously. Hsieh et al. (2015) identified higher and lower engagement patterns to represent students' learning processes in a game-based learning system. The study investigated a possible connection between students' verbal (asking themselves, expressing frustration, etc.) and nonverbal (smiling, focusing, moving closer to the screen, moving away from the screen, etc.) behaviors. However, the central issue in such research is limited to solving problems and does not include posing problems.

## 2.5 Scaffolding Systems

Scaffolding is defined as providing assistance to a student when needed and fading the assistance as the competence of the student increases (Wood et al., 1976). Research indicates that scaffolding facilitates learning as it supports learners in tasks that they cannot accomplish successfully by themselves, as well as developing knowledge for future learning (Hmelo-Silver and Azevedo, 2006; Pea, 2004; Sharma and Hannafin, 2007). Metacognition is defined as knowledge about and regulation of cognitive activities (Flavell, 1979). Metacognitive scaffolding aims to help learners to adequately control and monitor their learning (Azevedo et al., 2008; Molenaar et al., 2010; Veenman et al., 2005). Conceptually, scaffolding means teachers, instructors, or other knowledgeable persons providing learners with guidance or support before slowly shifting the responsibility to them as they develop their own understanding and skills. This reduces the difficulty of complex learning and at the same time, let the learners focus on constructing knowledge and higher-order demands like thinking critically (Way and Rowe, 2008). As technology extends learning from classrooms to learning systems, the scaffolding is no longer implemented via face-to-face instruction that literally exists between a teacher and learners in a classroom.

Recently, the form of supports that emerges to learners is facilitated through technology. The scaffolding metaphor has been used by researchers to describe features and functionality of educational software that help users to complete certain tasks (Sherin et al., 2004). Learners received computerized scaffolds supporting their metacognitive activities in the learning process (Molenaar and Roda, 2008). These scaffolds were given when metacognitive activities are typically executed in the learning process. The computerized scaffolding system generated the appropriate instance to send a support based on the learner's attention focus. Hannafin et al. (1999) categorized four types of scaffolding strategies in computer-based learning systems:

- (1) *Conceptual scaffolds* – guide learners in what to consider, and help them reason through complex problems and concepts. Conceptual scaffolds can be made available through explicit hints and prompts, and through structure maps and content trees.

- (2) *Metacognitive scaffolds* – provide guidance on how to think about the problem under study. They can be either domain-specific, such as where enabling contexts are externally induced, or more generic where the enabling context is not known in advance.
- (3) *Procedural scaffolds* – provide guidance on how to utilize available resources and tools. They orient learners/performers to system features and functions, or aid them while navigating the system. The scaffolds can be achieved by providing tutoring on system functions and features, or by providing a “balloon” or “pop-up” help to define and explain system properties.
- (4) *Strategic scaffolds* – suggest alternative approaches to analysis, planning, strategy, and tactical decision-making. They can be achieved by enabling intelligent responses to system use, suggesting alternative methods or procedures, providing start-up questions to be considered, and providing advice from experts.

Among the four types, metacognitive scaffolding was the most explored by researchers. It promotes higher order thinking (Way and Rowe, 2008) for it assists students to reflect on what they have learned (self-assess) and assesses their progress (Teo and Chai, 2009). As a result, it allows students to plan ahead.

A number of different metacognitive scaffolding techniques have been used in computer-based learning systems, to facilitate novices in activating, deploying, and monitoring the success of self-regulated learning. Azevedo and Hadwin (2005) used adaptive scaffolding based on on-going evaluation and calibration to the individual learner which may include some degree of fading. Others have attempted to engage the student in self-diagnosis with no other form of individualized support or fading (Choi et al., 2005; Dabbagh and Kitsantas, 2005; Puntambekar and Hubscher, 2005). Self-explanation prompts (Aleven and Koedinger, 2002) have also been used to facilitate problem solving and help determine the adequacy of one’s understanding of the topic.

Other studies demonstrated that learners who received scaffolding moved towards more sophisticated mental models and increased the frequency of use of self-regulated learning strategies when compared to those who received no scaffolding (El Saadawi et al., 2010). Previous research has indicated that when students learn about complex topics with computer-based learning systems in the absence of scaffolding, they show limited

ability to regulate their learning, leading to a failure of conceptual understanding (Azevedo et al., 2004; Green 2000; Hill and Hannafin, 1997).

## 2.6 Chapter Summary

This chapter corresponds to the informational phase (see Section 1.5) of the engineering method that our approach follows to address the research questions (Section 1.1) of this thesis. This literature survey described the principles of problem-posing learning systems and the scaffolding system among them that are necessary for our research. These include the definition of problem posing and how learning by problem posing is used, the ability, process and classification of problem posing. Then, Section 2.2 presented a review of the current state of the use of learning system for posing problems. We described the assessment and investigation of posed problem on learning systems and several techniques in posing problems from learning systems point of view. Section 2.3 explained the theory behind the learning system that is used for this thesis: Monsakun. In Section 2.4, we described current approaches to analyze learning activity on the systems. The smaller size of this section reflects the limited current state in analyzing at the process level, especially in the problem posing research area. Finally, the result of learning analytics on learning systems can be used to support learners through developing computer-based scaffolding system. Section 2.5 described the concept of scaffolding system and the current state of research among them.

---

## LOG DATA ANALYTICS IN TERM OF THE CONSTRAINTS

**Summary:** This chapter presents two studies of analysis the log data from learners' activity on Monsakun. The purpose of this chapter is to explore how data analytic techniques can be used to investigate whether learners pose the problems by considering the structure of arithmetic word problems in terms of the constraints. The studies examine what learners think while posing problems as sentence integration concerning intermediate states as well as the posed problems states as the resultant of problem-posing activity on Monsakun. Problem posing as sentence integration on Monsakun defines the arithmetic word problem structure, and posing a problem is a task to satisfy all the constraints and requirements to build a valid structure. The first study focused on how learners satisfied the required constraints and the second study focuses on the violation of constraints in the process of posing the problems. The chapter concludes with a discussion of the findings and the promising future works.

### 3.1 Introduction

The purpose of Monsakun as a problem posing learning system is to encourage learners to understand the structure of arithmetic word problems while they pose the problems. Monsakun provides learners with a different way to promote learning by problem posing, and it has distinct aspects of other practice of problem posing activity. The usefulness of Monsakun has been confirmed for learning by problem posing through previous research (Hirashima et al., 2008a; Hirashima et al., 2008b; Kurayama and Hirashima, 2010; Yamamoto et al., 2012). On the other hand, problem posing tasks in Monsakun is

conducted by making a combination of given sentences, which seems not to require deep thinking. In other words, learners can potentially pose the problems in a random way, and they can achieve the correct answer through trial-and-error. It means that they might not consider anything when posing the problems. In contrast, the purpose of development of Monsakun is to promote learners' thinking ability through the posing. Therefore, this leads to a gap investigation into how the validity of problem posing in each process of learners when they were posing the problems.

As was stated in Chapter 1, and discussed in detail in Chapter 2, there is little research that investigates every action of learners in posing the problems to understand the learning process of posing problems on an interactive learning system. Our literature review showed that even though there is a growing interest in investigating learners' activity on learning systems, there are not many explorations on investigating at the process level (Sections 2.4).

This chapter presents a study of investigation of learners' thinking through their actions during problem-posing processes on Monsakun. Figure 3-1 shows the goals, contributions and validation methods addressed in this chapter, particularly for the goal presented in Section 1.3: **investigating problem-posing processes on Monsakun in term of the constraints**. This chapter addresses the first research question (Section 1.2): *Do learners consider the structure of arithmetic word problem in posing problems using Monsakun?* We conduct a model-based analysis of problem-posing activity from Monsakun log data of 39 first-grade elementary school students to investigate their methods of thinking when they pose arithmetic word problems. Two studies have been conducted in order to confirm that learners consider the problem's structure while posing the problems. The first study presents a particular research question, whether learners pose problems by attempting to fulfill as many satisfied constraints as possible. Fulfilling satisfied constraints represents that learners consider to the problems' structure. The second one analyzes using a different perspective. The difference is the way to assess the composition of sentence cards and how to analyze the data. The main contribution of studies in this chapter is the approach of data analytics in problem-posing processes in term of the constraints.

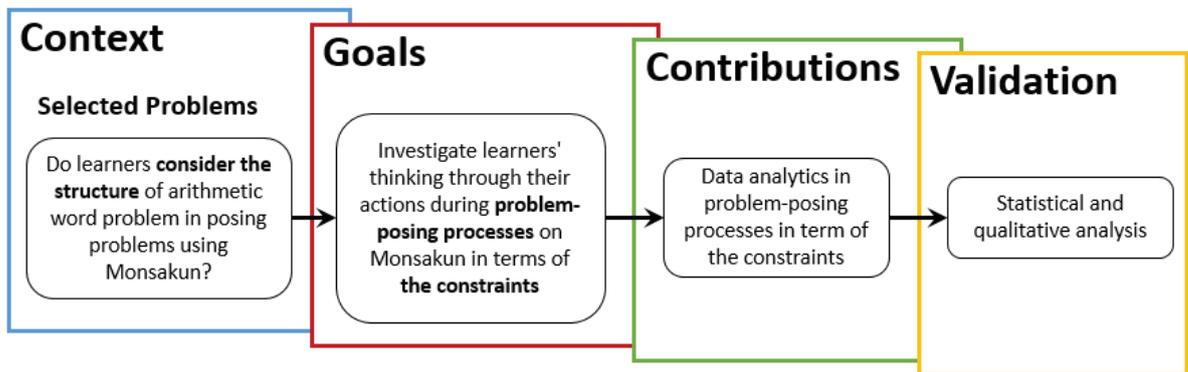


Figure 3-1 Context, goals, contributions and validation of Chapter 3.

The next section elaborates the motivation of the study. Section 3.3 describes the research questions. Section 3.4 presents the context of the study and data exploration. Section 3.5 explains assessment of problem-posing compositions in term of satisfaction and violation the constraints. Section 3.6 discusses the first study followed by the second study in Section 3.7. At last, Section 3.8 summarizes this chapter.

## 3.2 Motivation

Using Monsakun as a problem-posing learning system, learners' abilities to solve problems as well as to understand them are promoted. The basis of Monsakun is the triplet structure model (Hirashima et al., 2014) that defines the structure of an arithmetic word problem using sentence integration. Additionally, the model defines constraints for valid problems that must be satisfied. When a learner can pose the required problem in Monsakun, the problem certainly meets the constraints.

Although the usefulness of Monsakun has been confirmed for learning by problem posing, it is necessary to investigate the validity of the learners' problem-posing process in Monsakun. Two main points explain the necessity of the investigation in this study. First, the previous study reported that although learners gave many wrong answers to get the correct answer in some assignments, they did not pose the required problems randomly, and their many wrong answers are not meaningless as the results of thinking (Hasanah et al. 2015b). In the study, problem posing as sentence integration is presumed from the trends of posed problems as the result of the process. However, we extend the analysis by involving the process of arranging the problem. We assume that learners must

think the constraints form a valid problem throughout the problem-posing process. We conducted the first study to demonstrate that learners attempted to pose problems in satisfying as many constraints as possible based on their own understanding. In contrast, this second study investigates the problem-posing process and reveals the trends of the process, focusing on the violation of constraints. We conducted the study to prove that learners tend to avoid as many violated constraints as possible in composing problems.

Second, the fact that learners gave wrong answers illustrates that learners cannot avoid some mistakes. Therefore, it is essential to understand the learners' difficulties while posing the problems. Supianto et al. (2016b) detected important circumstances in the situation in which learners experience difficulties and misunderstanding of the structure of the problems. The study proposed a method to visualize learners' actions from Monsakun log data. In addition, the second study broads the analysis of the problem-posing process based on the constraints. It investigated learners' thinking through their actions and showed that learners have difficulty avoiding some specific types of constraints.

### 3.3 Research Questions

According to the motivation discussed in the previous section, we conducted a process analysis of problem-posing activity using Monsakun based on three research questions. The first question investigates the correlation between the number of satisfied constraints and the occurrence frequency of states performed by learners. We evaluated both intermediate and posed problem states for each arranged state.

1. Do learners pose the problems by attempting to satisfy as many constraints as possible?

The second question investigates the correlation between the number of violated constraints and the occurrence frequency of states performed by learners.

2. Do learners pose the problems by attempting to avoid as many violated constraints as possible?

The third question investigates the difficulty of learners regarding the violation of constraints. We assume that although learners tend to avoid compositions containing violated constraints, they have difficulty avoiding a particular type of constraint.

3. Do learners have difficulty avoiding a particular type of constraint?

## 3.4 Context of the Study and Data Exploration

### 3.4.1 Participants and Procedure

In this study, we analyze the Monsakun log data of 39 Japanese first-grade elementary students who participated in the practical use of Monsakun; their average age was six years old. Learners had already studied problem structure on the blackboard using several sentence cards that were parts of problems. These cards were provided to the learners as a request to pose the problems. We used the latest version of Monsakun, called Monsakun Touch (Yamamoto et al. 2012), to allow the learners to use Monsakun not only for the exercises but also for lessons on problem structure as part of their regular classes.

As described by Yamamoto et al. (2012), in practical use (see Figure 3-2), Monsakun was introduced as a problem-posing system of arithmetic word problems at the beginning of class (5–10 mins). The teacher distributed tablets containing Monsakun to learners and explained how to operate the system. Then, the teacher taught problem structures by simulating an assignment on the blackboard (20–35 mins). The teacher provided several sentence cards from Monsakun problems and conducted a lesson that resembled the Monsakun problem-posing process. The teacher encouraged participation and active discussion from all learners to pose the correct answer together. First, the teacher presented one sentence card to the learners from the prepared sentence cards, and the teacher explained the elements of the sentence card. The sentence card was composed of object(s), a value of the object(s), and the predicate. Second, the teacher presented another sentence card from the prepared sentence cards. Then, learners gave responses about whether the presented sentence card was necessary to pose the problem or not, and they also expressed the reason why the card was needed. Finally, the teacher explained the problem structure based on their answer. Through this teaching, the learners understood the following things: (1) a problem was composed of two independent quantity sentences

and one relative quantity sentence; (2) a sentence represented a story type and a relation among sentences, and there was a required sentence order based on the specific story type; and (3) a numerical formula was needed to represent the story type and find the answer. In that time, learners actively engaged in the lesson to express their opinions and ideas about posing arithmetic word problems provided on the blackboard. Finally, at the end of class, learners used Monsakun to complete an exercise in posing the problems individually (5–10 mins). We collected the log data from the activity at this time.

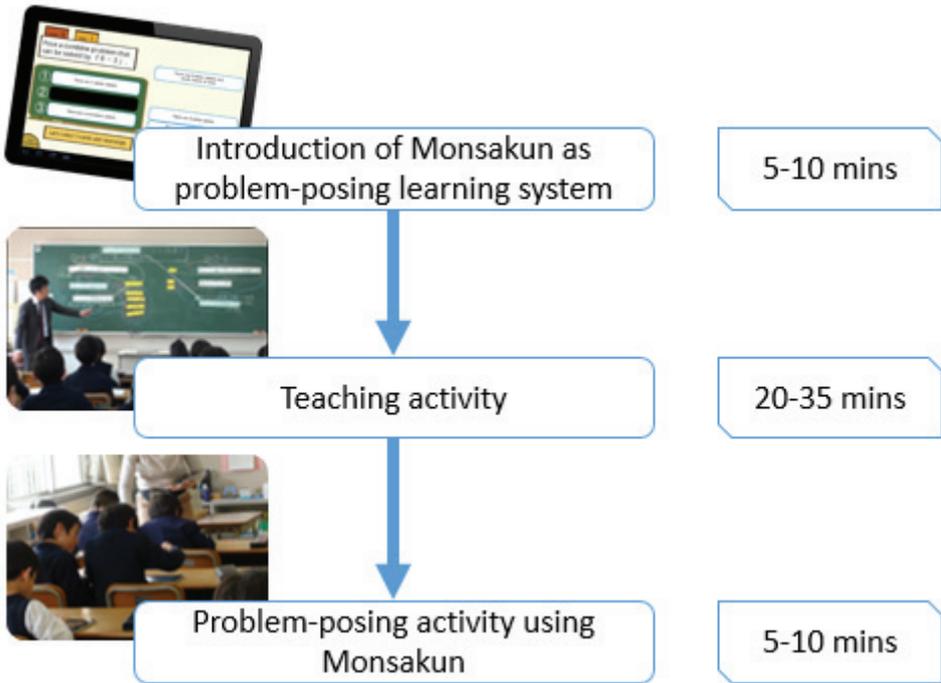


Figure 3-2 Procedure of practical use of Monsakun in a classroom.

### 3.4.2 Data Collection

The data of this study is gathered from the log files of learners’ problem-posing activities on Monsakun. The log file records the actions of learners during the learning activity using Monsakun, such as placing the sentence card, removing the sentence card, and clicking the “diagnosis” button (Section 2.3.3 – Problem-Posing Activity on Monsakun). The log file consists of the learner Id and information about the activities performed in Monsakun. They are labeled lvl, lid, asg, stp, act, crd, slt, stt, and jdg. The label “lvl” is the level of assignment that determines the difficulty of the problem-posing task, and “lid” shows the learner ID. The label “asg” is the number of the assignment, and “stp” is the

sequence number of the step. The label "act" consists of two possible values, set or remove, which represents placing the sentence card or removing the sentence card, respectively. The label "crd" is the index of the sentence card that is put in the slot, which is denoted by the label "slt." The label "stt" indicates the state generated from the combination of three sentence card indexes that is placed in the slot. The last code, "jdg," shows the type of action, for example, *incomplete state* (n), *wrong answer* (f), or *successful state* (s). We present a sample of the log data in Figure 3-3.

lvl	lid	asg	stp	act	crd	slt	stt	jdg
5	1	1	1	set	1	2	010	n
5	1	1	2	set	4	1	410	n
5	1	1	3	remove	1	2	400	n
5	1	1	4	set	6	2	460	n
5	1	1	5	remove	6	2	400	n
5	1	1	6	set	5	2	450	n
5	1	1	7	set	2	3	452	f
5	1	1	8	remove	2	3	450	n
... ..								
5	1	1	44	remove	4	1	012	n
5	1	1	45	set	3	1	312	s
5	1	2	1	set	3	1	300	n
5	1	2	2	remove	3	1	000	n

Figure 3-3 Example of log data of learners' activity on Monsakun.

The analysis of learners' performance by examining the average steps and mistakes in posing the problems on Monsakun has been reported in past research (Hasanah et al. 2015b). The average of the steps and mistakes shows how many steps a learner required to give a correct answer in one assignment and how many mistakes the learner made during the process, focusing on the posed problem states, respectively. Ideally, a learner would only need three steps to pose a correct answer because a problem in Monsakun consists of the arrangement of three simple sentence cards. The average of actions and mistake in Level 5 were significantly different compared to the others as can be seen in Table 3-1.

According to Table 3-1, the mean and standard deviation of actions of Level 5 are 42.01 and 28.59, whereas the mean and standard deviation of mistakes of Level 5 are 8.16 and 6.95. It shows that the average actions and mistakes at Level 5 were very high compared to the others, and it means that Level 5 was very challenging for learners. In

this study, we collected the log data in Level 5 in order to investigate learner’s actions at Level 5. The information of assignments at Level 5 is presented in Table 3-2 (see Appendix A.1 in more detail). There are twelve assignments in total, and each story type consists of three assignments.

Table 3-1 Descriptive statistics of learners' actions and mistakes when posing problems on Monsakun.

Level	Actions		Mistakes	
	Mean	Standard Deviation	Mean	Standard Deviation
1	7.53	2.28	0.48	0.33
2	6.42	2.86	0.49	0.42
3	6.92	2.31	0.64	0.49
4	6.05	1.62	0.57	0.23
5	42.01	28.59	8.16	6.95

### 3.4.3 Formulation of Problem-Posing Processes

As was stated in Chapter 2, problem-posing activity in Monsakun is called "problem-posing as sentence-integration", where the system provides a set of sentence cards and a numerical expression in the requirement part, and then learners pose an arithmetic word problem based on the triplet structure model using the numerical expression by selecting and arranging appropriate sentence cards. They are required to interpret the sentence cards and integrate them into one problem according to the given requirement.

Monsakun records problem-posing activity as changes in compositions of sentence card(s). We called the composition as a “state.” When the state is composed of three sentence cards, then it is known as the “posed problem state,” which is the card slots are completely arranged. An example of a posed problem state is shown in Figure 3-4c. Whereas when the arrangement is not composed of three sentence cards, then it is called the “intermediate state,” which is in the process of posing the problem. The examples of the intermediate states are shown in Figure 3-4a, b.

The sentence cards are encoded with an indexing number shown in Figure 3-4d. When the slot is still empty, index = 0 is implemented. For instance, when learners pose the

problem by selecting sentence card #1 and arrange it into the second slot, state 010 has been obtained, which is shown in Figure 3-4a. Another example of a state is shown in Figure 3-4b; state 410 happens when learners pose the problem by selecting sentence Card #4 and then arranging it into the first slot and selecting sentence card #1 and then arranging it into the second slot.

Table 3-2 Detailed assignments of Level 5 in Monsakun.

Assignment	Story type	Number of sentence cards	Requirement
1	Combination	6	Make a word problem about 'How many are there overall' that can be solved by '8 - 3'
2	Combination	6	Make a word problem about 'How many are there overall' that can be solved by '12 - 8'
3	Combination	5	Make a word problem about 'How many are there overall' that can be solved by '8 - 6'
4	Increase	5	Make a word problem about 'How many are there after increased' that can be solved by '12 - 8'
5	Increase	5	Make a word problem about 'How many are there after increased' that can be solved by '11 - 9'
6	Increase	5	Make a word problem about 'How many are there after increased' that can be solved by '10 - 2'
7	Decrease	5	Make a word problem about 'How many are left' that can be solved by '6 + 4'
8	Decrease	5	Make a word problem about 'How many are left' that can be solved by '8 + 6'
9	Decrease	5	Make a word problem about 'How many are left' that can be solved by '8 + 1'
10	Comparison	5	Make a word problem about 'How many are the difference' that can be solved by '12 - 5'
11	Comparison	6	Make a word problem about 'How many are the difference' that can be solved by '14 - 6'
12	Comparison	6	Make a word problem about 'How many are the difference' that can be solved by '9 + 4'

To complete an assignment, the learners attempt to arrange various combinations of sentence cards to generate a particular state according to what they set. They arrange the composition until they reach the composition of the correct answer. We formulated their actions as a sequence of states. For instance, several steps performed by a learner are

shown in Figure 3-5. First, state 000 is generated as the initial state. In the first step, the learner begins with state 010; this means that the learner has selected the first sentence card and arranged it into the second slot. In the second step, state 410 was composed, which means the learner selected the fourth sentence card and arranged it into the first slot. In the next step, the learner removed the first sentence card from the second slot; this condition changes the state to 400. Then, the learner tries to pose the problem resulting in state 450, and so on, until the correct state is reached.

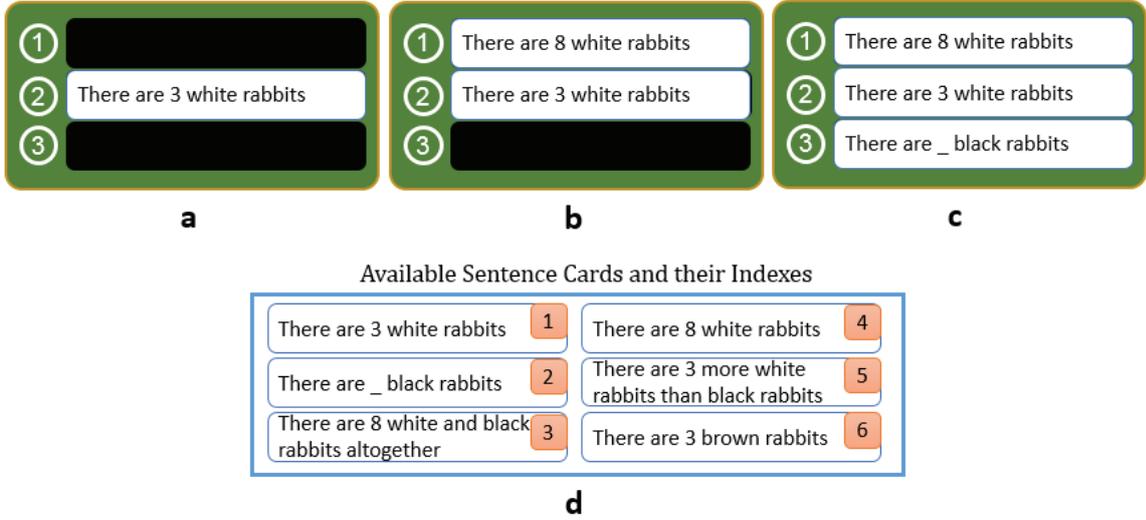


Figure 3-4 Example of states and the index of available sentence cards. a-c States. d Available sentence cards and their indexes.



Figure 3-5 Sequence of states generated from a learner's actions.

### 3.5 Assessment of States

#### 3.5.1 Satisfaction of the Constraints

The task model of posing problems via sentence integration has been developed based on the consideration of problem types in the triplet structure model (Kurayama and

Hirashima 2010). Based on the task model, five main constraints must be satisfied by each posed problem; they are 1) calculation, 2) story type, 3) number, 4) objects, and 5) sentence structure (discussed in detail in Section 2.3.2). When all five constraints are satisfied, the learner has succeeded in posing a correct problem according to the assignment requirements. When less than five constraints are satisfied, the posed problem is not valid; that is, the problem cannot be solved, or it is not the required one. We define two values for each constraint: 1, and -1. The value of 1 indicates the constraint is satisfied and the value of -1 indicates the constraint is not satisfied. The validity is measured based on the number of satisfied constraints, which is obtained by counting how many constraints are satisfied.

The example of several states and their satisfaction of constraints in Level 5 Assignment 1 is presented in Table 3-2. The requirement is to *make a story problem about "How many are there overall" that can be solved by "8 - 3,"* which is an arithmetic word problem with a combination story type. Learners can use six available sentence cards to pose the problem. The sentences for each card are composed of the following:

- (1) There are 3 white rabbits;
- (2) There are ? black rabbits;
- (3) There are 8 white and black rabbits altogether;
- (4) There are 8 white rabbits;
- (5) There are 3 more white rabbits than black rabbits; and
- (6) There are 3 brown rabbits.

The first example represented as State 462 (see Table 3-3, No. 1) has the validity equal to 1 because the state only satisfies one constraint (number). The state consists of numbers that fit the requirement; they are 8, 3, and the unknown number (?). However, the calculation cannot be made because it is necessary to transform the numerical expression, "8 - 3," into the numerical expression representing a combination story, "3 + ? = 8." In that formula, the number "3" and the unknown number "?" should be assigned as existence sentence cards, and the number "8" should be assigned to a relational sentence card, but the number "8" is an existence card on that state. Regarding the story type constraint, no relational sentence card indicates a combination story type. Then, the object also does not satisfy the constraints because all three objects are different, and they

are not connected to each other. Finally, to satisfy the sentence structure constraint, the state must consist of two independent quantity sentence cards and one relative quantity sentence card, but the state is composed of three independent quantity sentence cards.

Table 3-3 Example of several states and their satisfaction of constraints.

No	State	Composition of sentence cards	Constraints					Validity
			C1	C2	C3	C4	C5	
1	462	There are 8 white rabbits There are 3 brown rabbits There are ? black rabbits	-1	-1	1	-1	-1	1
2	152	There are 3 white rabbits There are 3 more white rabbits than black rabbits There are ? black rabbits	-1	-1	-1	1	1	2
3	123	There are 3 white rabbits There are ? black rabbits There are 8 white and black rabbits altogether	1	1	1	1	1	5

*C1 calculation, C2 story type, C3 number, C4 object, C5 sentence structure*

The second example, State 152, satisfies two constraints: object and sentence structure. There is a relationship between objects (white and black rabbits), and the structure of the sentence cards consists of two independent quantity sentences and one relative quantity sentence. However, the calculation, the story type, and the number are not fulfilled. There is no number “8,” as it causes a calculation process that cannot be done and the number constraint is not met. Concerning the story type, this is a comparison story, instead of a combination story. The third example is an example of a state that satisfies all constraints.

According to the triplet structure model, we only can measure the validity of the posed problem states, which is based on the number of satisfied constraints. Therefore, in order to cover the measurement of the intermediate states, we measure the validity by calculating the average of their descendant states. Figure 3-6 illustrates the hierarchy of possible compositions of sentence cards and transitions among them. For instance, in the dotted red rectangle at Figure 3-6 shows the descendant states of State 120. Consequently,

the validity of State 012 is obtained from the average of the validity of State 123, State 124, State 125, and State 126.

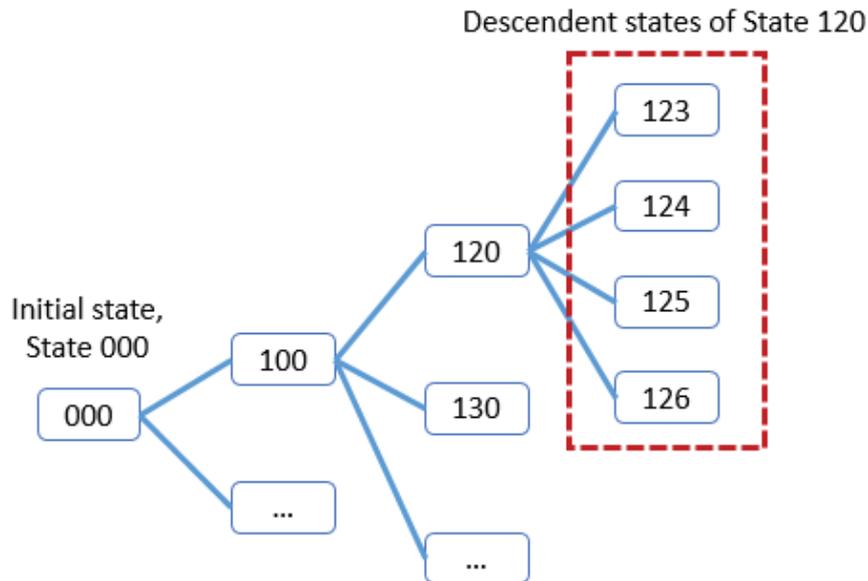


Figure 3-6 Hierarchy of possible compositions of sentence cards and transitions among them.

### 3.5.2 Violation of the Constraints

To cover the measurement of the intermediate states according to the violation of the constraints, we extend the definition of values for each constraint become:  $-1$ ,  $0$ , and  $1$ . The value of  $-1$  indicates the constraint is violated, and the value of  $0$  indicates the constraint is not violated, while the value of  $1$  indicates the constraint is satisfied. The number of violated constraints is obtained by counting how many constraints are violated.

Regarding the violated constraint, three states of Level 5 Assignment 1 shown in Table 3-4 are explained. The requirement and detail available sentence cards of the assignment have described in the previous subsection. The assignment is a task to pose an arithmetic word problem with a combination story type. In this story, there is no required order of sentence cards in the state (Section 2.3.1), which means the changes to the order of the sentence cards do not affect the problem story. For instance, the role that applies to State 001 (Table 3-4, No 1) also applies to State 100 and State 010. However, examples in Table 3-4 are presented in the order of indexes.

Table 3-4 Example of several states and their violation of constraints.

No	State	Composition of sentence cards	Constraints					Number of violated constraints
			C1	C2	C3	C4	C5	
1	001	– – There are 3 white rabbits	0	0	0	0	0	0
2	014	– There are 3 white rabbits There are 8 white rabbits	-1	0	0	0	0	1
3	246	There are ? black rabbits There are 8 white rabbits There are 3 brown rabbits	-1	-1	1	-1	-1	4

*C1 calculation, C2 story type, C3 number, C4 object, C5 sentence structure*

There is no satisfied constraint, nor violated constraint at the first example, State 001. This condition allows the calculation constraint not to be violated. The story type, number, object, and sentence structure are also not violated. Therefore, all constraints in this state are assigned to 0. The second example is State 014, which violates the calculation constraint. Based on the numerical expression in the requirement, the number "8" should be on the relative quantity sentence. However, sentence card #4 is an independent quantity sentence card, containing the number 8. Therefore, this state violates the calculation constraint, and this constraint is assigned to -1. The story type, number, object, and sentence structure are not violated nor satisfied because we still cannot determine them. Thus, the four constraints are assigned to 0. Four states can be derived from State 014, and each will not meet the correct state. The problems facing the derivative states are illustrated in Figure 3-7 and the detail rule for calculating the validity values presented in Appendix A.2. Like State 014, all derivative states will at least violate to the calculation constraint. The first derivative state is State 124. Besides the calculation constraint, this state violates the story-type constraint due to the lack of story and sentence structure because all arranged sentence cards are existence cards. The second and third derivative states are State 134 and State 145. Both violate calculation and number constraints. The difference between the states lies in the relative quantity sentence card. State 134 contains a relative quantity sentence card that fits the required story type (combination story),

while State 145 forms the comparison story type. Lastly, none of the constraints are satisfied in the derivative State 146. Four constraints are violated, and one constraint is not violated or satisfied.

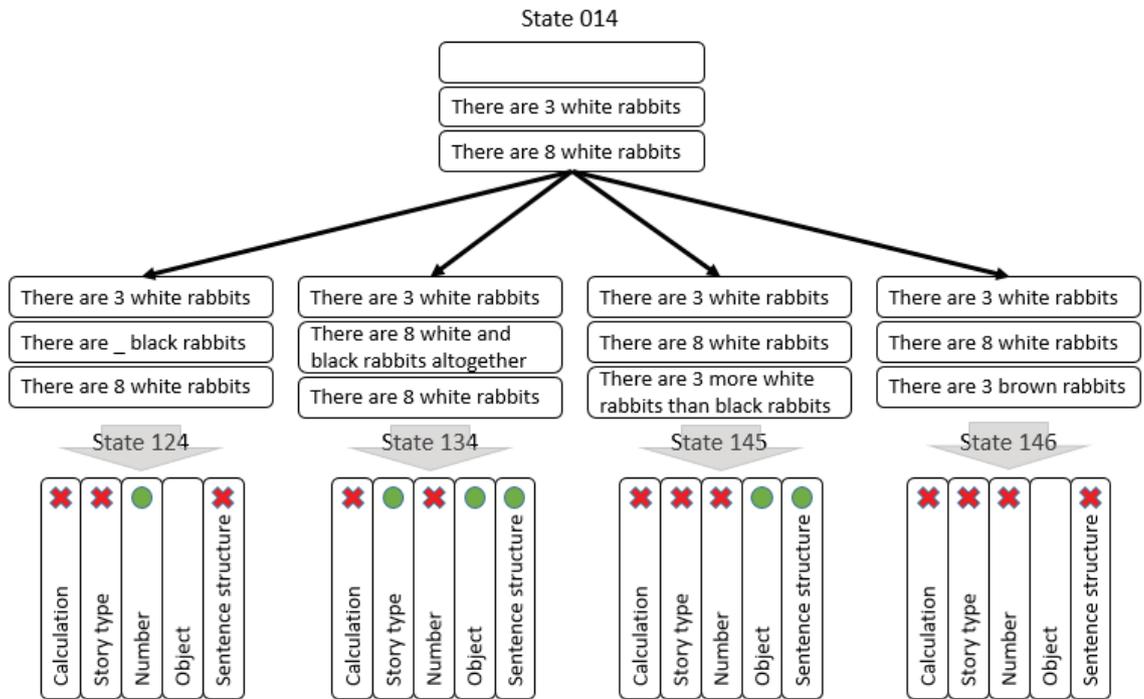


Figure 3-7 Derivative states of State 014 and the constraint violation in Level 5 Assignment 1.

The last example in Table 3-4, State 246, satisfies only one constraint, the number constraint. No story can be built from this composition, nor can the calculation and sentence structure be built. It can be calculated and well-structured when it consists of two existence sentences and one relational sentence, instead of all sentence cards being existence cards. Besides, there is no relation between objects in the composition of the sentence cards. They are independent objects consisting of white, black, and brown rabbits. This condition causes the number of violated constraints to be four because four constraints are violated.

### 3.6 Analyses in term of Satisfying the Required Constraints

We conducted two analyses of learners sentence card compositions focused on satisfying the required constraints at Level 5. We roughly analyzed every action of the learners. The

first analysis is to investigate learners' actions based on compositions they made, and the second one is to analyze the transition between compositions. The goal of both analyses is to address the first research question (RQ1).

### 3.6.1 RQ1: Do learners pose the problems by attempting to satisfy as many constraints as possible?

#### Analysis of States

The analysis results of Level 5 including the average number of actions to pose problems correctly and the correlation result between the frequency of states appearance and the validity of states are shown in Table 3-5. From the average actions, it denotes that Level 5 is more challenging than the other levels. Ideally, learners only need three steps to reach the correct answer, because a problem consists of 3 sentence cards. However, in average, they need at least 9 actions (Assignment 3) and at most 97 actions (Assignment 4).

Table 3-5 Average actions and correlation result between frequency of states appearance and validity of states in Level 5.

Assignment	Average actions	Pearson's correlation	p-value
1	50.08	0.087	0.584
2	13.32	0.507 **	0.001
3	9.24	0.581 **	0.002
4	97.18	0.224 **	0.009
5	32.00	0.375 **	7.040e-06
6	60.57	0.270 **	0.001
7	92.39	0.238 **	0.005
8	36.10	0.344 **	4.105e-05
9	54.71	0.384 **	3.878e-06
10	26.88	0.417 *	0.034
11	12.41	0.784 **	8.520e-10
12	19.27	0.631 **	7.617e-06

\*\* significant correlation ( $p < 0.01$ ), \* significant correlation ( $p < 0.05$ )

We conduct a Pearson's correlation test between the frequency of states appearance and the validity of states. The frequency shows how many numbers of unique states have

been reached by learners in order to pose a required problem in one assignment. We assume that the degree of correlation is related to the degree of learners' understanding. Therefore, there is a positive correlation between the validity and the frequency when learners understand the structure and aware of it in posing the problems.

Significant correlation at  $p < 0.05$  in eleven out of twelve assignments is found, which shows that many actions performed by learners had an inclination to satisfy as many constraints as possible. Hence, this the finding confirms that learners consider thinking about the structure of arithmetic word problem during pose the required problem in Monsakun.

In addition, we pay attention to the difference of degree of correlations among assignments. Significant correlation shows that learners are aware of the structure of arithmetic word problems. However, the correlations in Assignments 4 – 9, which are *increase* and *decrease* story types, are weak. Although the minimum actions to get the correct answer is three actions, learners take the average actions ten times or more from the minimum actions. It means their understanding about *increase* and *decrease* story types is worse than *combination* and *comparison* story types. Actually, *increase* and *decrease* story types have severe constraints and require a strict order of sentences in posed problems (Supianto et al., 2016b). We consider that this is a reason why the assignments of *increase* and *decrease* story types have a significant but weak correlation. That is, learners had misunderstood and attempted to pose problems with the misunderstanding for a long time. This finding shows that the learners were trying to satisfy the constraints when they constructed their answer because they produced many mistakes while doing the required problem.

Furthermore, the result of correlation in Assignment 1 shows no significant correlation ( $p > 0.05$ ). In order to examine more detail in Assignment 1, further analysis is observed in the next sub-section.

### *Analysis of Transitions*

In this analysis, we discuss the correlation between the change of satisfied constraints and the transition of states. Here, a transition is an action of learners from one state to another



In satisfying the particular constraints, learners attempted to pose more valid intermediate states. Based on the validity of several states listed in Table 3-6, if the learners have a comprehensive understanding of all constraints, then the validity of State 014 and State 046 is 1.00. Comparing the value with the other states, the validity is relatively low. However, when the learners have a partial understanding, the narrowly focused on some constraints, then the validity becomes different. For example, when a learner takes particular note of “number,” “object,” and “sentence structure,” there is a big difference between these two states and the others. On the other hand, if they mainly focus on “calculation” or “story type,” the values are not so different. In such case, it is difficult for learners to distinguish the State 014 and State 046 from the others. Also, the possibility to choose them is not low. This means that the learners’ intention in composing State 014 and State 046 is reasonable when it is viewed from several constraints they want to satisfy.

Table 3-6 Average actions and correlation result between frequency of states appearance and validity of states in Level 5.

States	All constraints	C1	C2	C3	C4	C5
004	1.60	0.10	0.00	0.40	0.50	0.60
014	1.00	0.00	0.00	0.25	0.25	0.50
024	2.50	0.25	0.00	0.75	1.00	0.50
034	1.50	0.00	0.00	0.25	0.50	0.75
045	2.00	0.25	0.00	0.50	0.50	0.75
046	1.00	0.00	0.00	0.25	0.25	0.50

C1 calculation, C2 story type, C3 number, C4 object, C5 sentence structure

The difficulty of this assignment is that learners are confused about the gap between the required story type (*combination*) and the numerical expression of subtraction (8-3) in the requirement (see Section 3.5.1). Although subtraction implies story type of decrease and comparison, in this case, learners must pose a problem of combination story type. The above assumption that they might mainly focus on “calculation” or “story type”

is related to this difficulty. This conflict of “calculation” and “story type” implied from the other can be considered as the cause of the difficulty.

At previous levels, there was no conflict at the required story type and numerical expression. Also, the order of numbers in sentences was the same as the numerical expression. Learners pose the required problem by arranging sentences according to the order of numbers in the numerical expression. However, this is not valid for Level 5 because the numerical expression does not express the order of numbers in the required story, but the solution is to evaluate unknown number. Learners meet this situation in the first assignment at Level 5. This is considered a reason for the unexpected behaviors where there is no significant correlation, and there are some frequent transitions with negative differences in the validity. Despite the difficulty encountered in Level 5, the analysis shows that the learners had a tendency to enhance the validity of the intermediate states to achieve the correct answers. Therefore, this analysis confirms that learners consider the constraints when they pose arithmetic word problem on Monsakun.

### 3.7 Analyses in term of Avoiding the Violated Constraints

Three analyses from the log files of learners’ problem-posing activity on Monsakun are conducted. We analyze their sentence card compositions. The first and second analyses provide the answer to the second research question (RQ2), while the third analysis gives the answer for the third research question (RQ3).

In the first analysis, we investigate the states and conduct a bivariate correlation analysis between the occurrence frequency of the states and the number of violated constraints. The occurrence frequency shows how many states have been arranged, while the number of violated constraints shows how many constraints are violated based on the state. We assume that the degree of correlation is related to the degree of the learners’ understanding. If the number of violated constraints has a negative correlation to the frequency, then the high number of violated constraints will be followed by the lower number of actions. It means that the high number of violated compositions of sentence cards has a small number of learners’ actions. Therefore, this correlation test will provide an answer to the second research question.

The second part of the analysis investigates the portion of states in the assignment setting to the occurrence frequency. We observe differences between the number of states in the assignment setting and the occurrence frequency. Moreover, we examine the differences for each number of violated constraints. In the low-frequency violated constraints, if the portion of occurrence frequency is higher than the number of states in the assignment setting, then it expresses that learners arrange states that have low error rates. In addition, in the high-violated constraints, if the portion of occurrence frequency is lower than the number of states in the assignment setting, then it shows that learners avoid solutions that potentially have a high error rate. Hence, this analysis will support providing the answer to our second research question.

Although two previous analyses show that learners tend to avoid mistakes, they still cannot avoid some mistakes, which demonstrates their difficulty in understanding the problem structure. Therefore, the third part of the analysis inspects the difficulty of learners according to the violation of constraints. We determine the ratio of the number of states in the assignment setting to the occurrence frequency. We examine the relative number of states for each type of constraint and their actual occurrence. If the number of occurrences is high, then the ratio is low. Thus, the minimum ratio in a constraint indicates that learners have difficulty avoiding such types of constraints while posing the problems. This analysis will confirm our third research question regarding whether learners have difficulty avoiding some particular type of constraints.

### 3.7.1 RQ2: Do learners pose the problems by attempting to avoid as many violated constraints as possible?

In this analysis, we conducted a Pearson's correlation test between the number of violated constraints and the occurrence frequency of states. We evaluated both intermediate and posed problem states for each arranged state. The result is shown in Table 3-7. A significant correlation ( $p < 0.05$ ) in 11 out of 12 assignments was found. Many actions performed by learners showed an inclination to avoid as many violated constraints as possible. The highest coefficient is in Assignment 10 ( $\rho = -0.5619$ ,  $p < 0.01$ ), and the scatterplot of this assignment is shown in Figure 3-9a.

Table 3-7 Average actions and correlation result between frequency of states appearance and number of violated constraints in Level 5.

Assignment	Pearson's correlation	p-value
1	-0.3701 *	0.0158
2	-0.4928 **	0.0014
3	-0.3879 +	0.0745
4	-0.2565 **	0.0033
5	-0.2778 **	0.0051
6	-0.3460 **	4.51E-05
7	-0.4006 **	1.35E-06
8	-0.3552 **	0.0001
9	-0.3990 **	4.43E-06
10	-0.5619 **	0.0028
11	-0.5570 **	0.0011
12	-0.4486 **	0.0054

\*\* significant correlation ( $p < 0.01$ ), \* significant correlation ( $p < 0.05$ ), + marginal correlation ( $p < 0.1$ )

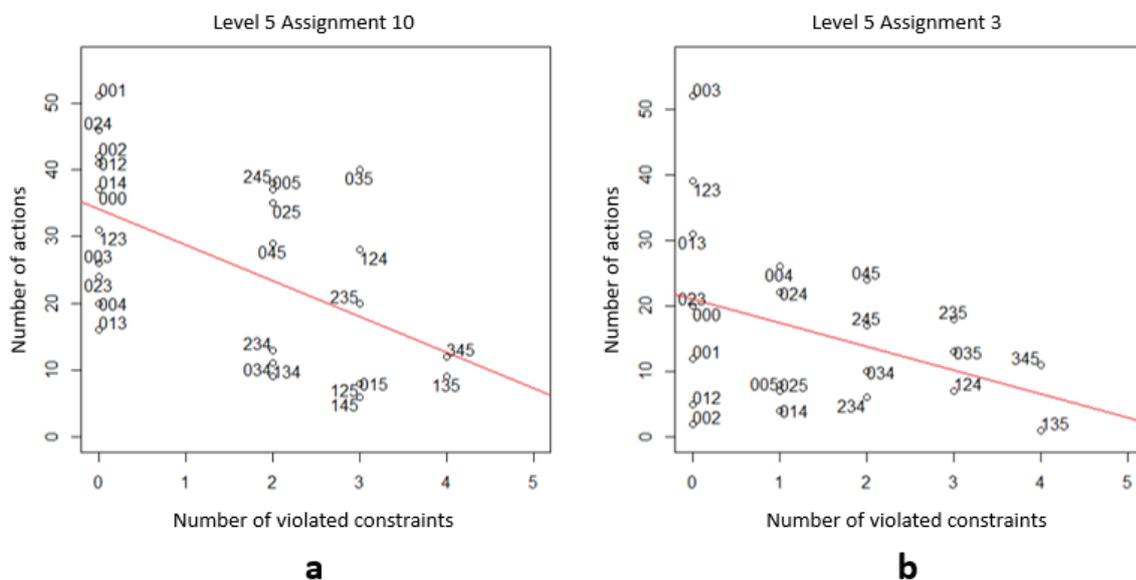


Figure 3-9 Plot of correlation between the number of violated constraints and the occurrence frequency of states. a Assignment 10. b Assignment 3.

The red line in the scatterplot shows the regression line of the data. Based on this information, the frequency of each state has a negative correlation with constraints violated in it. It is supposed that learners attempted to arrange the problem to avoid violating

more constraints. If the learners posed the problem randomly, the distribution of the number of the learners' actions would not have a significant correlation compared to the violated constraints. This finding shows that learners were inclined to pose more valid compositions.

Furthermore, the result of correlation in Assignment 3 shows marginal correlation ( $p < 0.1$ ). The scatterplot of correlation in Assignment 3 illustrated in Figure 3-9b indicates that there is no significant difference in this assignment. Therefore, the chi-square test was conducted to determine the trends in the details. We identified the portion of the number of states in the assignment setting to the occurrence frequency based on the number of violated constraints. Here, the number of states in the assignment setting means the space of all possible compositions that can be arranged by the learners. We check the number of states that are categorized in each number of violated constraints and the occurrence frequency. We show that although the correlation between the number of violated constraints and the occurrence frequency of states is not significant, there is a significant difference between the number of states in the assignment setting and its occurrence frequency. The results of the difference analysis and the detail portion of the assignments investigated in this study are presented in Table 3-8.

We found a significant difference in 11 out of 12 assignments ( $p < 0.01$ ), which shows that learners made a conscious attempt to avoid more violated constraints in the assignments. In addition, we pay attention to the portion of the number of states in the assignment setting to its occurrence performed by the learners according to the violated constraints. We found that the occurrence frequency in the high-violated constraints is lower than the number of states in the assignment setting, while the occurrence frequency in the low-violated constraints is higher than the number of states in the assignment setting. This implies learners were trying to avoid making a composition of sentence cards with a high number of violated constraints. Moreover, we show the portion of Assignment 10 (see Figure 3-10), which has a marginal difference. The portion of occurrence frequency, which is more than the number of states, happens at zero violated constraints, which means that learners tried to arrange the least instances of compositions of sentence cards that could potentially have many violated constraints. This finding strengthens the previous statement that many actions of learners were aimed to avoid as many violated

constraints as possible in arranging the posed problem states and the intermediate states as well.

Table 3-8 Difference analysis between portions of number of states in assignment setting to its occurrence frequency.

Asg		Number of violated constraints												Setting vs Occurrence	
		0	p	1	p	2	p	3	p	4	p	5	p	Chi-Sq	p
1	S	0.238	▼	0.190	▲	0.143	△	0.190	▼	0.214	▼	0.024	▽	< 0.01	**
	O	0.230		0.370	**	0.217	+	0.103	*	0.077	**	0.003	+		
2	S	0.286	▲	0.214	▲	0.190	▼	0.214	▼	0.048	▼	0.048	▼	< 0.01	**
	O	0.523	**	0.316	*	0.080	**	0.071	**	0.003	**	0.006	**		
3	S	0.308	▲	0.192	▽	0.192	▽	0.231	▼	0.077	▽			< 0.01	**
	O	0.510	**	0.189		0.161		0.107	**	0.034					
4	S	0.324	▲	0.147	△	0.176	▼	0.265	▼	0.088	▼			< 0.01	**
	O	0.541	**	0.196		0.087	*	0.167	*	0.009	**				
5	S	0.324	▲	0.191		0.265	▼	0.132	▽	0.088	▼			< 0.01	**
	O	0.649	**	0.191		0.044	**	0.103		0.013	**				
6	S	0.324	▲	0.147	△	0.176	▽	0.265	▼	0.088	▼			< 0.01	**
	O	0.520	**	0.190		0.131		0.151	*	0.007	**				
7	S	0.324	▲	0.147	△	0.176	▽	0.265	▼	0.088	▼			< 0.01	**
	O	0.500	**	0.189		0.149		0.139	**	0.024	*				
8	S	0.434	▲	0.081	▽	0.265	▼	0.132	▽	0.088	▽			< 0.01	**
	O	0.696	**	0.077		0.117	**	0.077		0.034	+				
9	S	0.324	▲	0.147	△	0.176	▽	0.265	▼	0.088	▼			< 0.01	**
	O	0.546	**	0.170		0.106	+	0.153	*	0.025	*				
10	S	0.423	▲			0.269	▽	0.231	▽	0.077	▽			< 0.10	+
	O	0.552	*			0.256		0.161		0.031	+				
11	S	0.500	▲	0.071	▼	0.143	▼	0.048	▼	0.143	▼	0.095	▼	< 0.01	**
	O	0.912	**	0.018	*	0.055	*	0.004	**	0.011	**	0.000	**		
12	S	0.333	▲	0.119	▽	0.143	▽	0.262	▽	0.095	▼	0.048	▼	< 0.01	**
	O	0.585	**	0.118		0.094		0.180	+	0.016	**	0.007	*		

S Setting, O Occurrence;

\*\* significant difference (p<0.01), \* significant difference (p<0.05), + marginal difference (p<0.1);

occurrence is more than setting (▲ significant, △ not significant);

occurrence is less than setting (▼ significant, ▽ not significant)

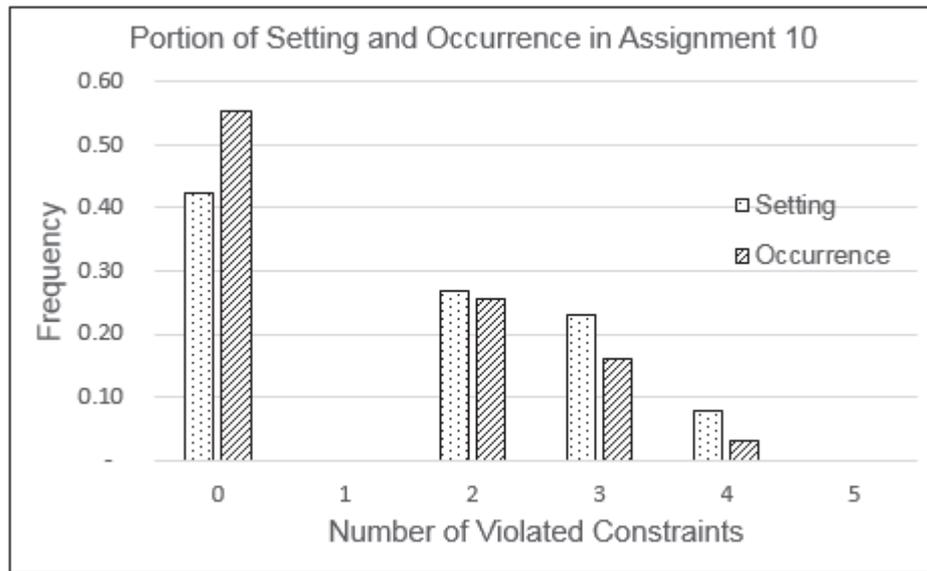


Figure 3-10 Portion of the number of states in Assignment 10 to the occurrence frequency.

### 3.7.2 RQ3: Do learners have difficulty avoiding a particular type of constraint?

In this analysis, we detect the difficulty of learners regarding the violation of constraints. We assume that although learners tend to avoid compositions containing violated constraints, they have difficulty avoiding a particular type of constraints. To prove it, we calculated the ratio of the number of states in the assignment setting to its occurrence frequency according to the type of constraints. The result is shown in Table 3-9. This ratio shows the relative sizes of the states in the assignment setting and the actual occurrence. The minimum ratio indicates that learners performed many actions, which indicates they have difficulty avoiding such constraints while posing the problem. We found that, in 8 out of 12 assignments, learners have difficulty avoiding the story constraint. In addition, in 4 out of 12 assignments, they have difficulty avoiding the calculation constraint. Based on this result, we confirm that they have difficulty avoiding a particular type of constraint.

As previously described, Level 5 is required to consider the unknown number because it is not given in the requirement. It is challenging for learners, especially in considering the story constraint. At the previous levels, there is no conflict at the required story type

and numerical expression. In addition, the order of numbers in sentences is the same as the numerical expression. For instance, the requirement of Level 3 Assignment 1 make a word problem about “How many are there overall” that can be solved by “ $4 + ? = 10$ ,” learners can pose the required problem by arranging sentences according to the order of numbers in the numerical expression. However, this is not valid for Level 5 because the numerical expression does not express the order of numbers in the required story but the solution is to evaluate the unknown number. To complete assignments at this level, for example, in the first assignment, learners need to transform the numerical expression “ $8 - 3$ ” into the numerical expression representing a combination story, “ $3 + ? = 8$ .” Then, learners could assign the existence sentence cards to the number “3” and the unknown number “?”.

Table 3-9 Ratio of the number of states in the assignment setting to the occurrence frequency according to the type of constraints.

Assignment	Type of constraints				
	C1	C2	C3	C4	C5
1	0.0151 *	0.0230	0.0317	0.0469	0.0350
2	0.0971 *	0.2308	0.3191	0.1852	0.2308
3	0.1060 *	0.1224	0.2857		0.1190
4	0.0532	0.0458 *	0.1341		0.0509
5	0.2145	0.2113 *	0.6857		0.2113 *
6	0.0633	0.0467 *	0.1472		0.0616
7	0.0709	0.0559 *	0.1277		0.0726
8	0.2800	0.2778 *	0.3750		0.2778 *
9	0.1484	0.1297 *	0.2609		0.1449
10	0.0498	0.0448 *	0.1081		0.0459
11	0.9474 *	3.3333	0.9474 *	1.5000	3.3333
12	0.1705	0.1258 *	0.2083	0.7143	0.2500

C1 calculation, C2 story type, C3 number, C4 object, C5 sentence structure

\* the minimum value of ratio

Investigation of learners’ activities at the process level promotes an opportunity to discover the learners’ behavior in detail. Moreover, when it is associated with a cognitive load, then the learners’ thinking processes can be explored. Particularly, what conditions

learners face difficulties in attempting to pose problems could be detected. With such detections, we could define learning support depending on the learners' mistakes and develop an adaptive function to overcome learners' bottlenecks in attempting to pose problems.

### 3.8 Chapter Summary

We conduct a model-based analysis of problem-posing activity including posed problem and intermediate products while posing problems from Monsakun log data of first-grade elementary school students to investigate their actions while posing arithmetic word problems. The assumption in this first study is that they consider the structure of arithmetic word problems as sentence-integration through satisfying as many constraints as possible. The result shows that significant correlation in 11 out of 12 assignments is found, which indicates that many actions performed by learners had an inclination to satisfy as many constraints as possible. Further investigation shows that learners had a tendency to enhance the validity of the intermediate states to achieve the correct answers. It means that learners tended to pose the required problem with an awareness of the structure of arithmetic word problems throughout the problem-posing process.

The second study focuses on the violation of the constraints. Correlation between the numbers of violated constraints and the frequency of each intermediate product that the learners actually made was reported. Moreover, to determine the detail trends of learners' actions, a chi-square test between the number of states in the assignment setting and the occurrence frequency was conducted. Significant correlation and difference in 11 out of 12 assignments was found, which shows that many actions performed by learners had the inclination to avoid as many violated constraints as possible. It indicates that they tended to avoid as many mistakes as possible. It means that learners tended to pose the problem with some consideration to the structure of arithmetic word problems throughout the problem-posing process. Furthermore, although learners tended to avoid the violated constraints, they could not avoid some mistakes. However, most of the learners' mistakes violated at most two constraints. Further analysis shows that, in 12 assignments, learners generally have difficulty fulfilling 2 out of 5 constraints, which are "story" and "calculation" constraints. Based on this analysis, it would be possible to detect the

difficulty of learners' actions from the model perspective. Hence, accurate feedback and appropriate support can be provided.

---

# VISUALIZATIONS OF LEARNERS ACTIVITY IN POSING THE PROBLEMS

**Summary:** This chapter presents a study proposes a method to offer a visual representation for analyzing the problem-posing activity sequence in Monsakun, a digital learning system for posing arithmetic word problems via sentence integration. The system writes every single action into logs as sequences of problem-posing activity. The sequences are considered to represent the thinking processes of learners. The thinking process reflects their understanding and misunderstanding about the structure of the problems. The study in this chapter created visualizations of learners' problem-posing processes from the data obtained through the practical use of Monsakun, including the states in which many learners had difficulties finding the correct answer. In this study, we refer to such states as “trap states.” In Monsakun, a trap state is a combination of simple sentences where many learners tend to make and need relatively more actions to obtain the correct answer. As the result of the visualization and analysis of the data, some trap states have been identified, and they changed for each trial in the same problem.

## 4.1 Introduction

Previous chapters have established the ground for a study of learners' thinking and considerations when they pose the problems on Monsakun. Section 3.6 presented a study that showed learners consider thinking about the structure of arithmetic word problem while posing the required problem through satisfying as many constraints as possible. Then, in Section 3.7, we presented a study that showed many actions of learners were aimed to avoid as many violated constraints as possible in arranging the required problem.

Both sections involve the log data and the constraints in Monsakun’s model to evaluate the problem-posing process. In contrast, this chapter presents a study in discovering the problem-posing process based on the log data only. As the interactive problem-posing learning system can be used to provide new ways to capture traces of learning activities, this little explored and somewhat hidden potential of these devices can be exploited to enhance learners’ awareness of their thinking progress based on unusual patterns from the captured traces of actions. These learner’s data can make key aspects of highlight possible problems.

This chapter demonstrates a study that it is possible to automatically discover patterns of learning activity on Monsakun that can help teachers and researchers provide more helpful feedback to learners. Figure 4-1 shows the goals, contributions and validation methods addressed in this chapter, particularly for the goal presented in Section 1.3: **providing a visual representation for discovering the bottlenecks of learners’ problem-posing process.**

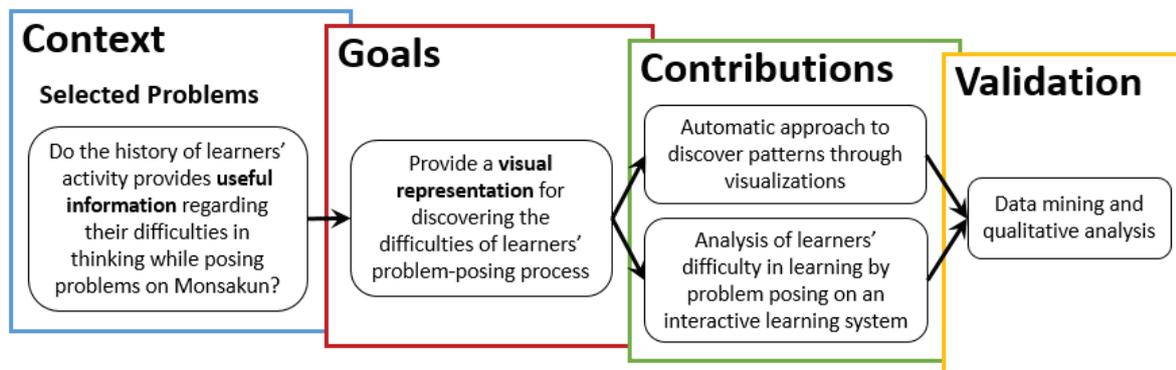


Figure 4-1 Context, goals, contributions and validation of Chapter 4.

To achieve this goal, we traced every single movement on learners posing the problems and presented their patterns using data mining technique. We validate the approach by conducting two sequence implementations to a case study on the practical use of Monsakun:

1. The implementation of visualizations to discover patterns, which are the key aspects of learners’ difficulty in thinking.
2. The implementation to analyze learners' changes in thinking to pose the same problem during three trials in term of the difficulty they faced.

The chapter is organized as follows. Section 4.2 describes motivation and research purpose. Section 4.3 explains the method used as the approach to distilling patterns in term of visualizations. Section 4.4 presents the implementation of the visualizations to a case study on the practical use of Monsakun. Section 4.5 presents discussions of unusual patterns found. Section 4.6 concludes this chapter with some sort summaries.

## 4.2 Motivation and Purpose

In general problem-posing exercises where learners pose problems freely; it is hard for learners to pose problems and for teachers to analyze the posed problems. The learning system for problem-posing exercises, Monsakun, resolves this difficulty with the problem-posing by sentence integration. In this method, learners pose problems meeting certain requirements by combining three simple sentences from the given sentences. By using this method, the opportunities for learners to pose problems increase, feedback to learners according to their mistakes is provided, and for the teacher, checking the validity of posed problems becomes easier.

From the results of previous studies with Monsakun, lessons and exercises with it improve not only learners' problem-posing abilities but also their problem-solving skills (Yamamoto et al., 2012). In addition, from the preliminary analysis of sentence selection, learners change their approach to problem-posing after they have experienced posing the same type of story (Hasanah et al., 2015a). Although posing problems in the learning system is considered to contribute to the understanding of the problem's structure, it is not clear how learners finally could understand it through the activity. Therefore, it is essential to conduct discovery learning and to generate inferences of learners' thinking from their behavior in the system. Discovery learning plays a role in increasing learners' motivation while creating more opportunities for learners to assess how well they could overcome obstacles, which may improve learning (Reiser et al., 1998).

There had been considerable works analyzing learners' activities to get a general view of learners' learning. Statistics and visualization information are the two main techniques that have been most widely used for this task (Romero and Ventura, 2010). For statistical techniques, Hadwin et al. (2007) examine logs of trace data affords opportunities to

explore the intersection between what students perceive about their studying, and what they do when they study; and Zorrilla et al. (2005) exploring the learners' behavior and time distribution of network traffic over time. While visualization information techniques, they oriented toward visualizing educational data such as learner tracking data regarding social, cognitive and behavioral aspects of learners (Mazza and Dimitrova, 2003); and tracking learners' answer that reflects on their behavior in an adaptive tutorial (Ben-Naim et al., 2009).

The purpose of visualization is to “amplify cognition” about data (Card et al., 1999). Visualization could be fully leveraged to get a better understanding from step-by-step data logs generated by learning systems. Anscombe (1973) suggested both calculations and graphs should be used by a computer, both sorts of output should be studied due to each of them would contribute to understanding. Visualization could help to avoid misinterpretation of data. Shneiderman (2002) claimed that integration of both data mining and information visualization to invent discovery tools could enable more effective exploration and promote responsibility.

There has also been considerable work exploring the importance of visualizations to externalize the activity of learners. Some of them have conducted design and visualize learning process in a computer supported collaborative learning system (Janssen et al., 2007a; Tan et al., 2008), visualize and externalize the activity of groups working together on collaborative learning participation (Janssen et al., 2007b; Rabbany et al., 2011), and visualize the learning interaction with respect to collaborative and learning attitudes of each participant (Hayashi et al., 2013). On individual learning, systems that collect detailed real-time data on learner behavior and interpret those data by drawing on behavioral research have been developed (Macfadyen and Sorenson, 2010). In 2011, the adaptive learning system was developed (Anjewierden et al., 2011). This system could monitor learner behavior through the actions they perform and identify patterns that point to systematic behavior using visual representation. Moreover, the visualization uses a tree structure to provide an overview of class performance also has been developed to allow easy navigation and exploration of student behavior (Johnson et al., 2011).

A study in this chapter offers a visual representation for analyzing the problem-posing processes of learners in Monsakun. These visualizations could detect important

circumstances according to the changes in their thinking processes. Ben-Naim et al. (2008; 2009) propose trap-states as pre-defined error answer and have identified actual trap-states as a solution trace based on the learners' answer. In addition to their definition of trap-states, we broaden the definition to include the process of arranging the answer. By this detection, we would be able to provide information regarding a situation in which many learners experience difficulties and misunderstanding of the structure of the problems. The people who benefit from the results of this approach are teachers and researchers. If teachers understand the difficulties faced by learners, they can provide more helpful feedback. If researchers understand the difficulties experienced by learners, they can consider and develop functions to overcome them. In the current state of this study, we propose a method to visualize learner's actions from the log data of Monsakun, to extract some information from it, and to analyze the results.

### 4.3 An Automatic Approach through Visualizations

This study explores the log data of learners' activities collected from Monsakun as an interactive learning system for problem-posing activity. The goal of this study is to provide visual representations to discover and analyze learners' thinking in posing the problems. Specifically, this study tries to discover the bottleneck of learners' problem-posing process. Figure 4-2 shows the procedure of study to reach our goal.

In order to achieve our research goal, we first collected log data of learners' activity from Monsakun. Then, we designed a formulation to encode the problem-posing process. At this stage, a transformation from a sentence card to the number representation was applied. After that, we traced the problem-posing activity sequences recorded in the database. In the next stage, we calculated two kinds of values: support and distance values. These two values were used to measure the frequency of learners who attempted an action and to measure the distance from an action to the correct answer, respectively. Then, we visualized the values through states as graphical representations. Finally, we found trap-states as the bottlenecks of learners' thinking and analyzed our findings based on the visualizations.

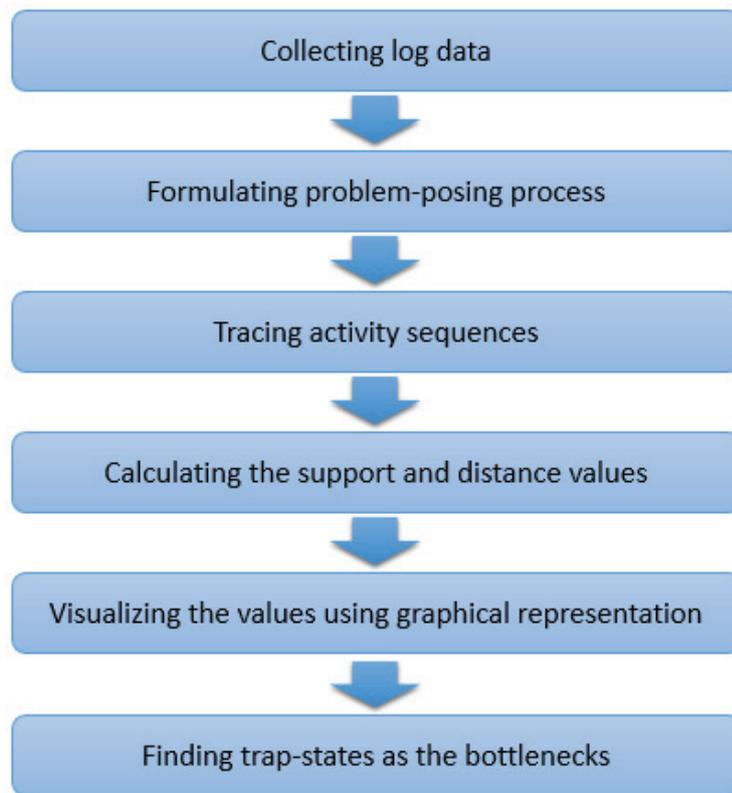


Figure 4-2 Procedure of the study.

#### 4.3.1 Collecting Log Data

Each learner's actions on Monsakun were logged into a database. The raw data was coded as a series of events. The events were coded as lvl, lid, asg, stp, act, crd, slt, stt, and jdj, where have described in Section 3.4.2. Participants tried to pose the problem three times. For instance, there are 12 assignments in Level 5. When a learner finished all of 12 assignments in the level, they could complete the challenge for that level. As a result, some learners were able to attempt the same problem a maximum of three times. Therefore, we collected the data from three trials of the same problem by the same learners.

#### 4.3.2 Formulating problem-posing process

Monsakun records learners' problem-posing activities as combinations of cards set in the card slots. The resultant of activity is a combination of sentence cards, which is called the "state" in the problem-posing process by learners. We defined four types of states shown

in Table 4-1. Type 1 is the uncompleted state, which is composed of less than three sentence cards. This type includes, at least, one empty card slot. The empty card slots are represented by zero (the state shown in brown). Type 2 is the completed state, which is composed of three sentence cards. However, the learner did not ask the system to diagnose the posed problem. This type is coded by a combination of three sentence card indexes and followed by the string “[u]” (the state is shown in black). Type 3 is the failed state. Although this is also composed of three cards, the system diagnosed it as a failure. This type is coded by a combination of three sentence card indexes and followed by the string “[f]” (the state is shown in red). The last one is the successful state (type 4), which the system diagnosed as the correct answer. This type is coded by a combination of three sentence card indexes without being followed by any other string code.

Table 4-1 Example of each type of states.

Type	Definition	Example	Description
1	Uncompleted state	010	Slot 1 is empty, slot 2 is occupied by sentence card #1, slot 3 is empty
2	Complete state without push button	413 [u]	Slot 1 is occupied by sentence card #4, slot 2 is occupied by sentence card #1, slot 3 is occupied by sentence card #3, and without checking the answer
3	Complete state and gets wrong answer	315 [f]	Slot 1 is occupied by sentence card #3, slot 2 is occupied by sentence card #1, slot 3 is occupied by sentence card #5, and check the answer then gets the wrong answer
4	Complete state and gets correct answer	312	Slot 1 is occupied by sentence card #3, slot 2 is occupied by sentence card #1, slot 3 is occupied by sentence card #2, and check the answer then gets success

Based on the model, all possible combinations of sentence cards and transitions among them could be clearly defined as a network of states. We call this network as the “problem states space.” All the actions of learners could be mapped to a transition from one state to another in this network. All possible states can be obtained by combining all the available sentence cards, including the empty slot. When a state represents a basic unit of thinking, then the problem states space provides the range of thinking in a problem-posing assignment. Therefore, this network becomes the basis of our proposed visualization approach.

The possible combination starts from State 000, the initial state, which means that all slots are empty (root state). It continues with combinations of one card slot filled, two card slots filled, and three card slots filled (complete state). There is a constraint that must be satisfied to generate all possible states. Each sentence card can be used only once. For example, it is impossible to create State 131, which means that the first sentence card is used twice, in the first slot and the third slot. However, it becomes possible to make a combination of more than one empty slot, for example, 001, 002, 003, and so on.

In the next step, we connected each state in accordance with the proper conditions. The proper condition is one where there is a relation between the situation before and after an action. For example, we connected a situation where all slots are empty with a situation where one card slot is filled. It was impossible to connect a situation where all slots are empty with a situation in which two slots are filled with sentence cards because there is one situation that elapsed. As a concrete example, we could not connect State 000 to State 012 because there is one step that elapsed before the State 012. The state that may be done before State 012 is State 001 or State 002.

The example of problem states space from six available cards with a combination story is shown in Figure 4-3. Since there is no required order of cards in the state, states that have the same composition are combined into one state (Section 2.3.1). For instance, States 013, 031, 103, 130, 301, and 310 are combined into State 013. As a result of combining states, we get 42 states in total for combination and comparison stories. While for the increase and decrease story types, we get 136 states in total from five available sentence cards.

### 4.3.3 Tracing activity sequences

In order to complete an assignment, the learners tried varying compositions of the sentence cards, in order to generate a particular state according to what they thought. They continued to change the composition of the sentence cards until they reached the correct composition. Every state that occurred for learners was stored by the system. Thus, we had an order of each state called a “sequence of states.” A sequence of states is a collection of states that are sorted based on the sequence of learners’ activity. This sequence reflects the learners’ thinking processes. Examples of sequences are shown in Figure 4-4.

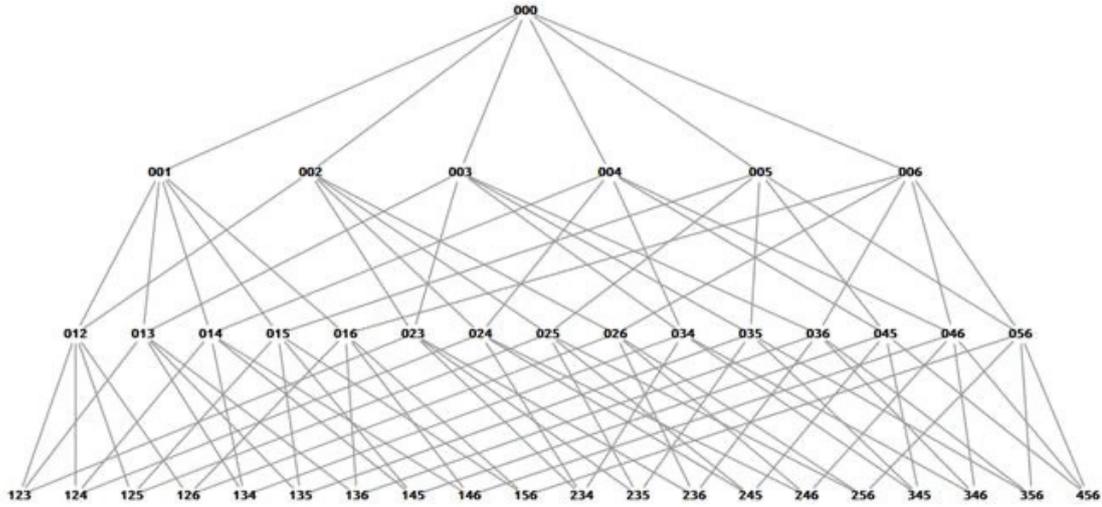


Figure 4-3 The graph of states space for combination story type with 6 available sentence cards.

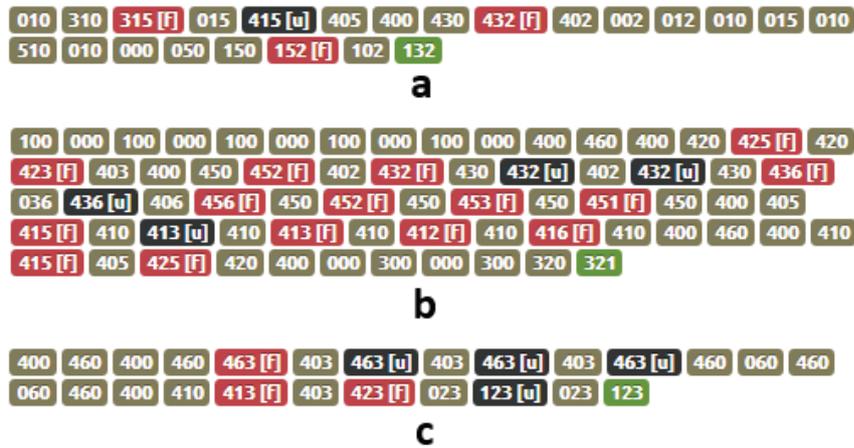


Figure 4-4 The examples of the sequence of states. a The sequence of learner 1. b The sequence of learner 2. c The sequence of learner 3.

There are three examples of sequences that have a different number of states, as shown in Figure 4-4. The first sequence had 23 states as shown in Figure 4-4a; this means that the sequence comprises 23 steps to reach the correct answer. The second and third sequences had 67 steps (see Figure 4-4b) and 25 steps (see Figure 4-4c), respectively. The first sequence begins with the State 010; this means that the learner put the first sentence card into the second slot. The next state is State 310. In this state, the learner put the third sentence card into the first slot. Next, the learner put the fifth sentence card into the third slot, followed by pressing the diagnosis button. At this point, the learner achieves the

wrong answer. Representation of this state is State 315[f]. The learner tried to correct the error by taking the third sentence card from the first slot; this condition turns into State 015. The complete steps of the first sequence (see Figure 4-4a) could be mapped to the problem states space, as shown in Figure 4-5. The blue nodes indicate visited states. The yellow links show the relations between the visited states, and the thickness represents how many steps the learner followed to get there. On the other hand, the gray nodes show the states that were never arranged by the learners. Blue nodes and yellow links represent what the learner considered before he/she arrived at the correct answer.

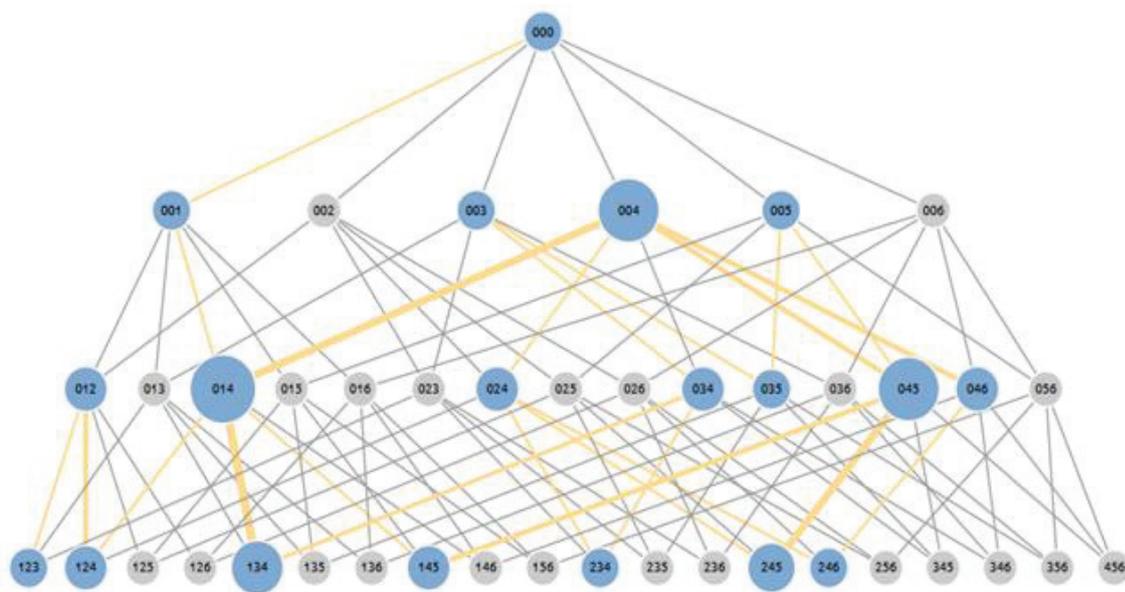


Figure 4-5 An example of mapped a sequence to the problem states space.

#### 4.3.4 Calculating the support and distance values

We propose two kinds of values to be visualized: support value and distance value. In this study, support value denotes the number of learners who attempted to pose a state. This value aims to show how many learners arranged the state. On the other hand, distance value denotes the average number of steps it took for learners to reach a correct answer from a state. This value aims to show how far or close the learner came to the correct answer; thus, it is called the distance of state. Figure 4-6 shows sequences of states of four learners for illustrating how to calculate support and distance values of State 014.

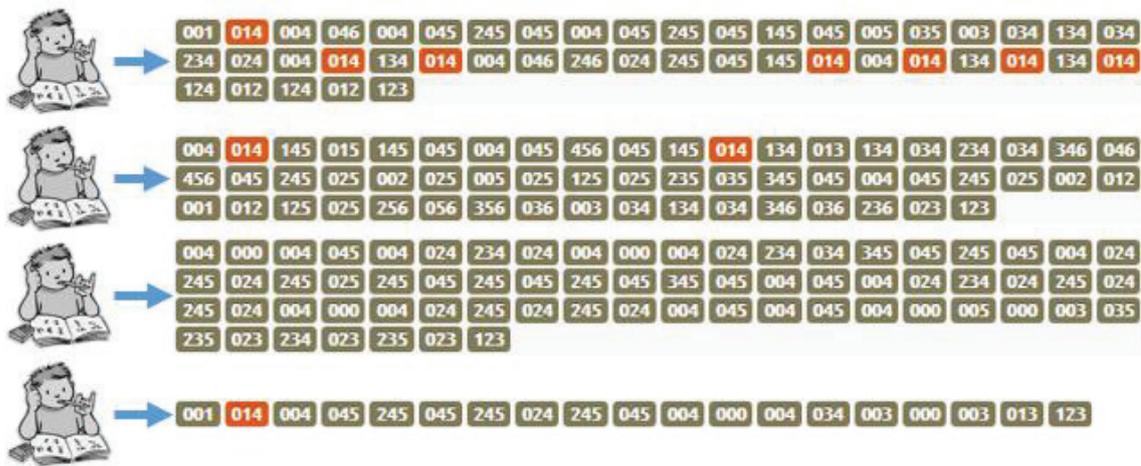


Figure 4-6 An example of sequences of four learners for illustrating how to calculate support and distance values of State 014.

As defined above, support value is obtained by calculating the number of learners arranged a state. In the case of State 014, we can see in Figure 4-6 that there are three learners perform the state. Therefore, the support value of State 014 is 3. The distance value is obtained by calculating the average number of steps from a state to the correct answer. Since it is possible for a learner to pose the same state for more than once, we also calculate the average number of steps in one learner before calculating the average for all learners. Let say State 014 of the first learner occurs seven times to achieve the correct answer; they are at the 2nd step, the 24th step, the 26th step, the 34th step, the 36th step, the 38th step, and the 40th step for total 45 steps. As results, it causes distances for each step is 43, 21, 19, 11, 9, 7, and 5, respectively. Thus, the distance value of State 014 of the first learner is 16 steps. With the same procedure, we calculate the distance value for all learners as follows:

- Learner1's distance (014)
 
$$= \{(45-2) + (45-24) + (45-26) + (45-34) + (45-36) + (45-38) + (45-40)\} / 7$$

$$= (43 + 21 + 19 + 11 + 9 + 7 + 5) / 7$$

$$= 115 / 7$$

$$= 16.43$$
- Learner2's distance (014)
 
$$= \{(57-2) + (57-12)\} / 2$$

$$= (55 + 45) / 2$$

$$= 100 / 2$$

$$= 50.00$$

- Learner3's distance (014)

$$= 0.00$$

- Learner4's distance (014)

$$= \{(19-2)\} / 1$$

$$= 17 / 1$$

$$= 17.00$$

At last, we calculated the average distance of learners, which is  $(16.43 + 50.00 + 0.00 + 17.00) / 4 = 83.43 / 4 = 20.86$ . Since the unit of distance is the step, we omit the decimal places. Therefore, the distance value of State 014 is 20 steps.

#### 4.3.5 Visualizing the values using graphical representation

A sequence has several states as objects that are linked by ordered steps. The first step linked to the second step, the second step linked to the third one, and so on. For this reason, we proposed the graph visualization, which shows the states and their relations. Moreover, information visualization is best represented in graph structures, which act as bridges between the visualization and the graph drawing field (Gröllner, 2002). Rabbany et al. (2011) used graph representation to visualize overall snapshots of the students' participation in the discussion forums and give the instructor a quick view of what is under discussion in online courses. In this study, such as in Johnson et al. (2011), we designed a graph where each node represents a state, and each link represents an action that takes a learner from one state to the next one. The graph gives an overview visualization of all relations between the previous state and the next state in a sequence.

We described two kinds of graphs, “support graph” and “distance graph,” based on their values. A support graph displays the support value of the states performed by learners. This graph aimed to visualize states that have the number of supports shown by the size of the node. The node with a larger size has a higher number of supports than the node with smaller sizes. A distance graph displays the proximity of states to the correct answer. This graph aimed to visualize the average number of steps of a state to the correct answer indicated by the size of the node. The larger the node size, the greater the number

of average steps. It means that a large-sized node had a longer distance to the correct answer on average.

#### 4.3.6 Finding trap-states as the bottlenecks

A state performed by learners is the result of their thinking. When learners choose to put one sentence card into an empty slot due to its possible potential, it has a consequence. Similarly, when learners tried to take out a sentence card that has been installed in one slot, it will lead to consequences too. The consequences could move learners away from the correct answer. In this case, the learners are stuck in a condition where they would have to do more steps to reach the correct answer. In other words, the learners are trapped in the state that distanced them from the correct answer. Moreover, many learners found themselves in this state.

We defined two important circumstances related to learners who were stuck in these conditions. The first circumstance is related to learners who were in a condition where there was no dummy card contained in the state. The second circumstance is related to learners who were in the condition in which there was at least one dummy card found in the state. Thus, we defined a state with no dummy card that could lead learners to doing many steps to get to the correct answer and reached by many learners as “confusing state,” while a state contained at least one dummy card where it could lead learners to do many steps to get to the correct answer and reached by many learners as an “actual trap state.”

### 4.4 Implementation of the visualizations to a case study on the practical use of Monsakun

#### 4.4.1 Participants and Materials

In this implementation, we visualize log activities of 39 Japanese first-grade elementary students who participate in a practical use of Monsakun in class. As the focus of the study, we processed the log data not only the completed states but also the uncompleted states. In other words, we visualized the learners’ activities, including every action, in order to extract some information through data mining techniques, which is represented by

support and distance values. We believe that the visualization and analysis could be used to support learners due to their difficulty in thinking.

Level 5 that is focused on the study in Chapter 3 consists of twelve assignments. In this study, we focused on one assignment in order to apply our proposed method and analyzed it. We picked the first assignment of Level 5. This assignment covered a combination story type problem. The requirement of this assignment was to make a word problem about “How many are there in all” that can be solved by “8–3.” There were six sentence cards provided that could be used by learners. The sentences for each card, from the first card to the sixth card, were the following:

1. There are 3 white rabbits.
2. There are ? black rabbits.
3. There are 8 white and black rabbits altogether.
4. There are 8 white rabbits.
5. There are 3 more white rabbits than black rabbits.
6. There are 3 brown rabbits.

In this assignment, the correct answer consisted of sentence card #1, sentence card #2, and sentence card #3. We collected the data from three trials of the first assignment by the same participants described above.

#### 4.4.2 Visualization of Support and Distance

Figure 4-7a, b shows examples of support graphs and distance graphs, respectively. The value of each state in both types of graphs was normalized by scaling 0 to 1. We discarded the node that had a value of zero, which means the learners had never done that state. We focused on the state that had already been made by learners. We also implemented two different colors for nodes. The color was determined based on the top five support and distance values. The top five nodes were colored red, and the rest were blue. We did this on the grounds that the top five nodes are (1) states that had a high-value of support, as shown in Figure 4-7a, and (2) states that had long distances from the correct answer, as presented Figure 4-7b. For that reason, we focused on the red states for further analysis.

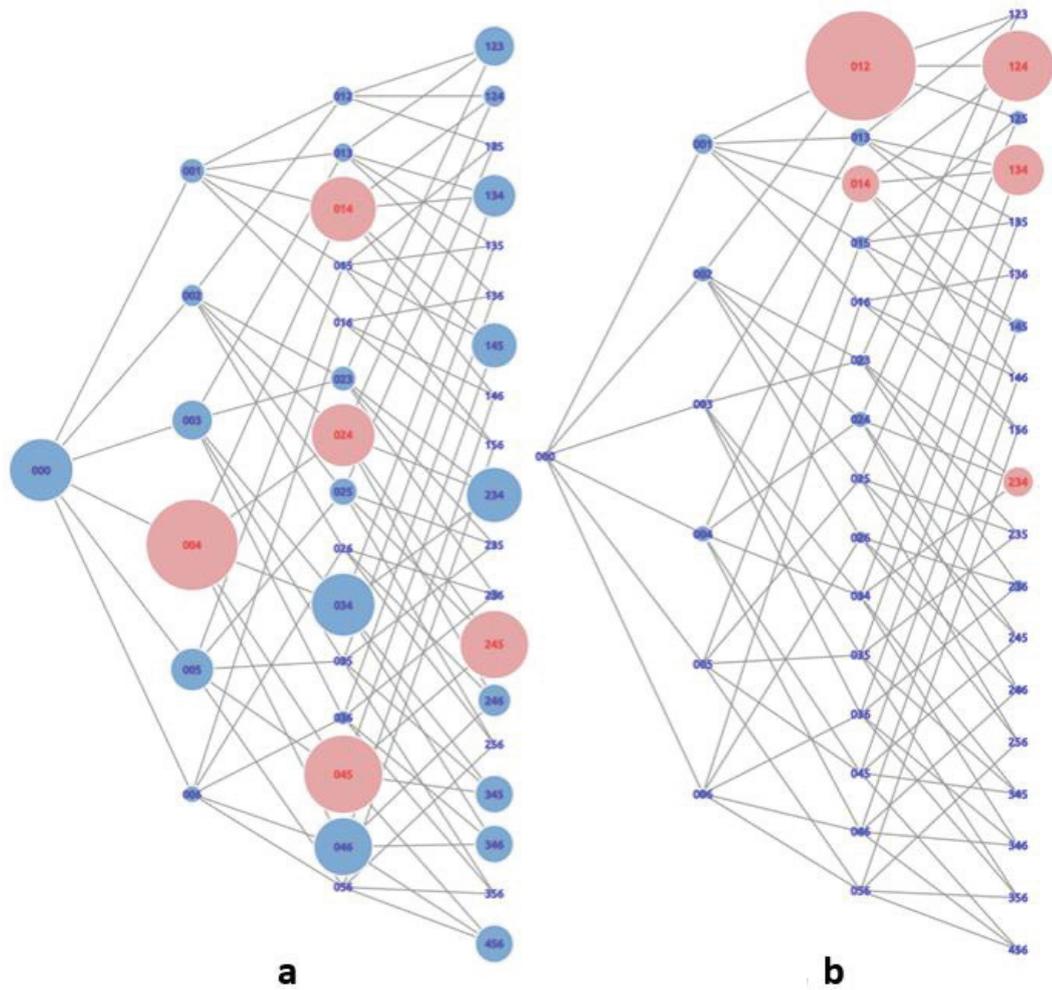


Figure 4-7 Support graph and distance graph of the first trial. a The support graph. b The distance graph.

We argue that using these two graphs, based on the red states for both sides, valuable information could be discovered. For instance, State 004 is one of the principal states from the support graph. However, it does not have a high-distance value. This means that many learners made the state but they easily became aware of its incorrectness. On the other hand, State 012 is one of the principal states from the distance graph. However, it does not have a high-support value. It means that very few learners were able to make it and felt that it was difficult. We need to find states where many learners feel challenged. For this purpose, we combined both graphs.

Based on the value of its support, the red states on the support graph shown in Figure 4-7a are 004, 045, 245, 014, and 024, which have support values of 37, 32, 28, 27, and 26

respectively. This means that State 004 was attempted by 37 learners, State 045 was attempted by 32 learners, etc. On the other hand, the red states on the distance graph shown in Figure 4-7b are 012, 124, 134, 014, and 234, which had the average distance to the correct answers of 68, 44, 32, 24, and 19, respectively. This means that when learners were in State 012, they took 68 steps to reach the correct answer on average. Similarly, State 124 required 44 steps, State 134 required 32 steps, etc. Highlighted states on the distance graph are strong candidates for trap states because learners got stuck and had to do many steps to reach the correct answer. However, this was not enough to identify that a state was a trap state. As defined previously, that trap state is a state that not only leads learners to do many steps but also requires learners to be supported by others. Therefore, by combining the distance graph with the support graph, the actual trap states were revealed.

The most distinct state in the distance graph is State 012 (the largest node is shown in Figure 4-7b). This situation indicates that the State 012 could potentially be a trap state for many learners. However, when we looked at the support value, this state was only supported by a few learners (shown by the little blue node in Figure 4-7a). Although State 012 was one that required many steps to reach the correct answer in this data, there were not many learners who could arrange this state. The same thing occurred in State 124, State 134, and State 234. The rest state with the red color shown in Figure 4-7b is State 014. When learners were in this state, they were required to complete 24 steps to reach the correct answer on average. Moreover, this state was also reached by many learners, as shown in red in Figure 4-7a. This means that, for many learners, they tended to do more steps and move further away from the correct answer when they were in State 014. This state was supported by 27 learners. Besides, State 014 contained a dummy card. Thus, this state could be said to be an actual trap state. In other words, by using visualization, it was revealed that the red colored state shown in the support and distance graphs was an actual trap state.

#### 4.4.3 Visualization of Changes of the Trap State

In this section, we present another visualization called the “trap graph.” It was obtained by multiplying the support value and the average distance value together. We refer to this

multiplication result as the trap value. We also normalized the trap value by scaling 0 to 1. The trap graph shown in Figure 4-8 provides information about the actual trap state in the first trial of the assignment. The actual trap state is displayed by a red node in the graph. From this graph, we also obtained the result that the actual trap state was State 014 as the highest trap value in the first trial. The result of the actual trap state in the trap graph showed the same result when we used support and distance graphs. We argue that the trap graph could be used to represent both support and distance graphs. For this reason, we use the trap graph to detect the actual trap states in the second and third trials as the changes of the trap state.

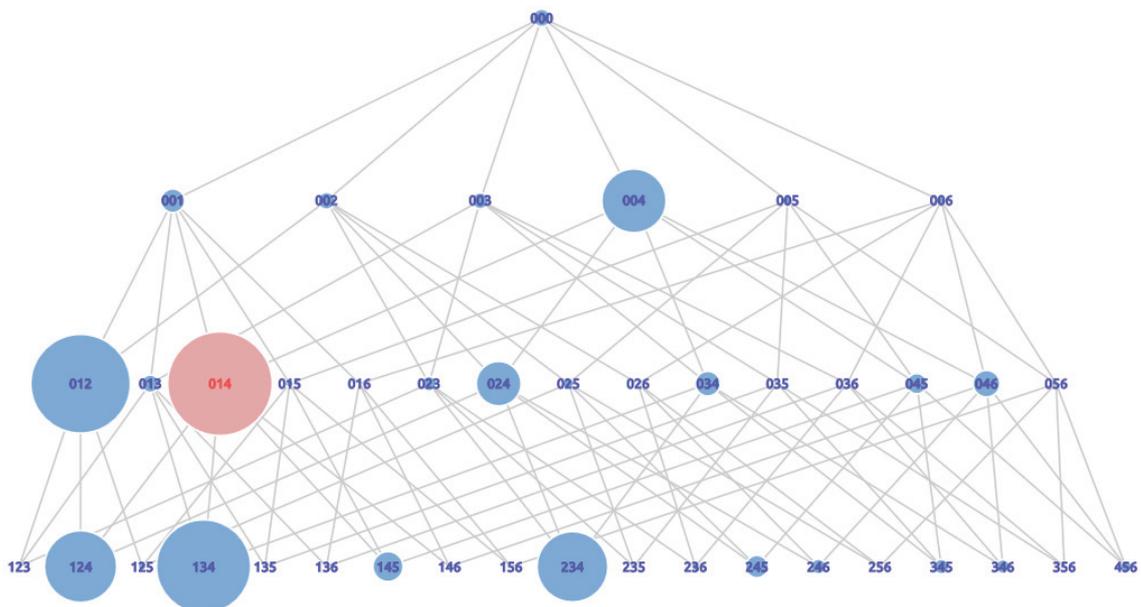


Figure 4-8 Trap graph of the first trial.

In the second trial, we found that the two highest trap values were State 023 and State 013 shown in Figure 4-9, which did not contain any dummy cards. State 023 contained sentence card #2 and sentence card #3, both of which were correct cards. Also, state 013 contained sentence card #1 and sentence card #3, both of which were correct cards. Because these states did not contain any dummy cards, we refer to these states as confusing states. In these states, the learners are close to the answer. They have been on the correct path and only need one more step to reach the answer, meaning that learners need only one correct sentence card to complete the problem. However, they did many steps to reach the correct answer. They needed 22 steps on average to get to the correct

answer. However, there is nothing wrong with their actions to achieve the State 023 or State 013. Therefore, we argue that they were only confused, instead of trapped.

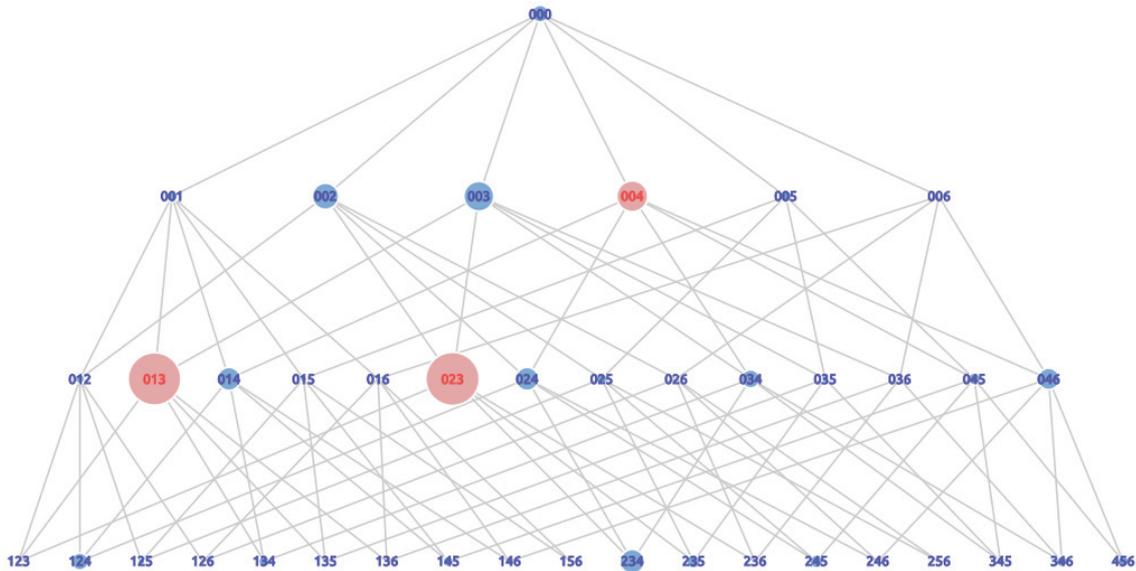


Figure 4-9 Trap graph of the second trial.

The next high value is State 004, the red state shown in Figure 4-9. The state is the top-3 state, which was supported by many learners and required many steps to reach the correct answer. The difference with state 023 was that this state contained a dummy card. In this case, the learners chose the wrong sentence card. The mistake of choosing this sentence card caused the learners to be trapped in a condition that increased the distance to the correct answer. Thus, we detected State 004 as an actual trap state in the second trial of the assignment. The last trap graph generated from the third trial is shown in Figure 4-10. State 035, the highest trap value on the graph, was detected as an actual trap-state.

## 4.5 Discussions

The difficulty in this assignment was that learners were confused about the gap between the required story type of combination and the numerical expression of subtraction (8–3). Although subtraction implies the story type of decrease and comparison, in this case, learners must pose a problem of the combination story type. Before this assignment, in previous levels, learners had done assignments in which they could obtain the correct answers by arranging cards according to the order of numbers in the numerical expression.

However, this was not valid for this assignment because the numerical expression did not express the story, but expressed the solution to evaluate the unknown number. Even if they made a strategy to arrange the cards according to the numerical expression from previous assignments, this strategy did not work in this assignment. Learners tended to make such a strategy (Hasanah et al., 2015a). To complete this assignment, for example, the learners needed to transform the numerical expression, “8–3,” into the numerical expression representing a combination story, such as “3 + ? = 8.” Moreover, learners could assign existence sentence cards to “3” and “?”.

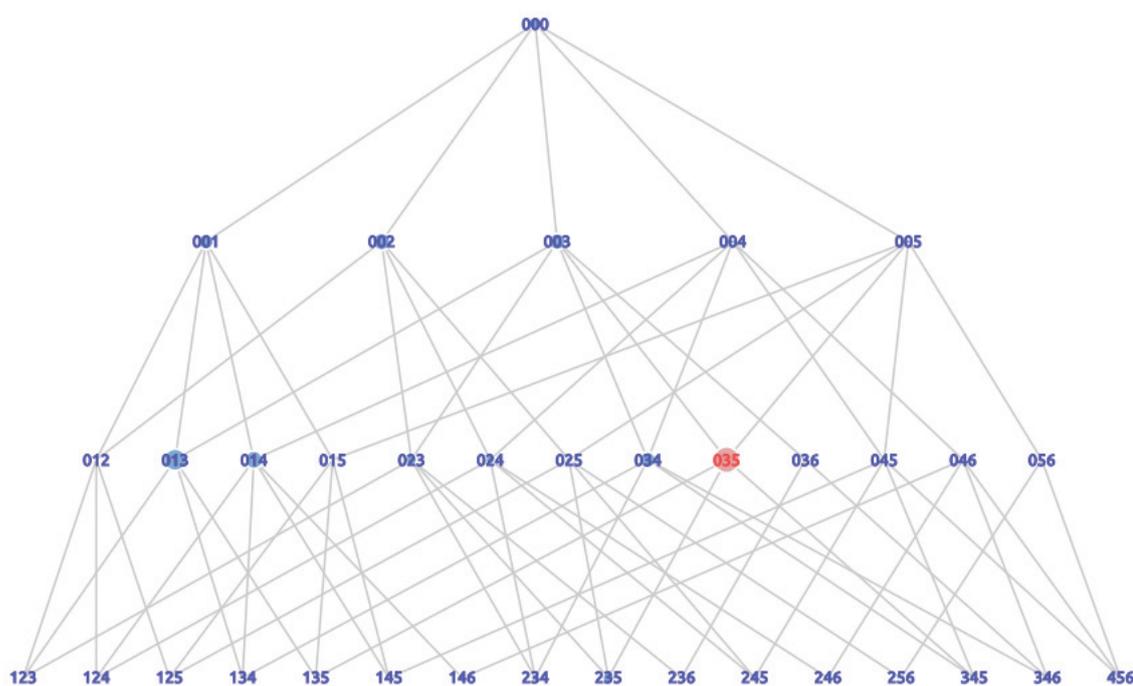


Figure 4-10 Trap graph of the third trial.

The trap state in the first trial that tended to make learners do more steps and distanced them from the correct answer was State 014. State 014 consists of sentence card #1 (*There are 3 white rabbits*) and sentence card #4 (*There are 8 white rabbits*). This tendency demonstrated that learners tried to use the given numerical expression directly, “8–3,” and to assign sentence card #1 and sentence card #4 to the number “3” and “8,” respectively. Based on the available cards, it was reasonable that those sentence cards had been chosen instead of sentence card #2, which contained an unknown number (*There are ? black rabbits*), and sentence card #6, which contained a different type of object from the others (*There are 3 brown rabbits*). In this situation, most of the learners became stuck

because they thought the correct answer was the number “8” on the numerical expression belongs to the independent quantity sentence card (*There are 8 white rabbits*), but actually they should have chosen the number “8” on the relative quantity sentence card (*There are 8 white and black rabbits altogether*). Thus, State 014 could also be explained as an actual trap state based on the triplet structure model. We thus confirm that, by using these visualizations, the actual trap state for learners could be detected.

Based on the trap graph in each trial, it was demonstrated that the principal trap state changed from the first trial to the second and third trials. In the first trial, the principal trap state was State 014. In the second trial and the third trial, the principal trap states were State 004 and State 035, respectively. More precisely, the changes of the actual trap states are indicated by the changes of the trap value for each trial, as shown in Figure 4-11. The chart in Figure 4-11 shows that State 014 had a trap value of 2621 in the first trial. This value is the highest value that enables State 014 to be detected as an actual trap state in the first trial. However, in the second trial, State 014 only has a trap value of 377. This value is lower than State 004, which has trap value of 662. Therefore, in the second trial, the actual trap state moved to State 004. In the third trial, the actual trap state changed again to State 035. This state has a trap value of 699, which increased from the previous trial where it only had a trap value of 136. On the other hand, the trap values of States 014 and State 004 decreased to 266 and 101, respectively. We confirm that by using these visualizations, the actual trap state for learners in every trial could be detected. Moreover, the actual trap state changed from the first trial to the third trial, as summarized in Table 4-2. Table 4-2 also shows the number of learners who were in the state that was detected as an actual trap state. In general, it could be said that the number of learners in that state decreased.

State 014 was found to be an actual trap state in the first trial. As discussed before, in this trial, learners used strategy to pose the problem by directly using the given numerical expression, “8-3.” The actual trap state in the second trial changed to State 004. State 004 only consisted of sentence card #4 (*There are 8 white rabbits*). This was to encourage learners to try to change strategies to pose the problem and encourage them to move away from using numerical expressions. However, the strategy did not work because learners still kept trying to use sentence card #4. There was still a decided tendency based on some

sort of thinking (Hasanah et al. 2015b). They assumed that the number “8” in the numerical expression was considered to be an independent quantity sentence. In this situation, most of them were still confused and stuck because the correct answer with the number “8” was the number in the relative quantity sentence (*There are 8 white and black rabbits altogether*). Thus, State 004 could also be explained as an actual trap state based on the triplet structure model. In the third trial, the actual trap-state changed and was detected as State 035. State 035 consisted of sentence card #3 (*There are 8 white and black rabbits altogether*) and sentence card #5 (*There are 3 more white rabbits than black rabbits*). This composition did not make a particular type of story. They put sentence card #3, which reflected a combination story, and sentence card #5, which reflected a comparison story, together in one state. Based on the number of learners shown in Table 4-2, we could say that a small number of learners still did not have a good understanding of the base structure of the problem.

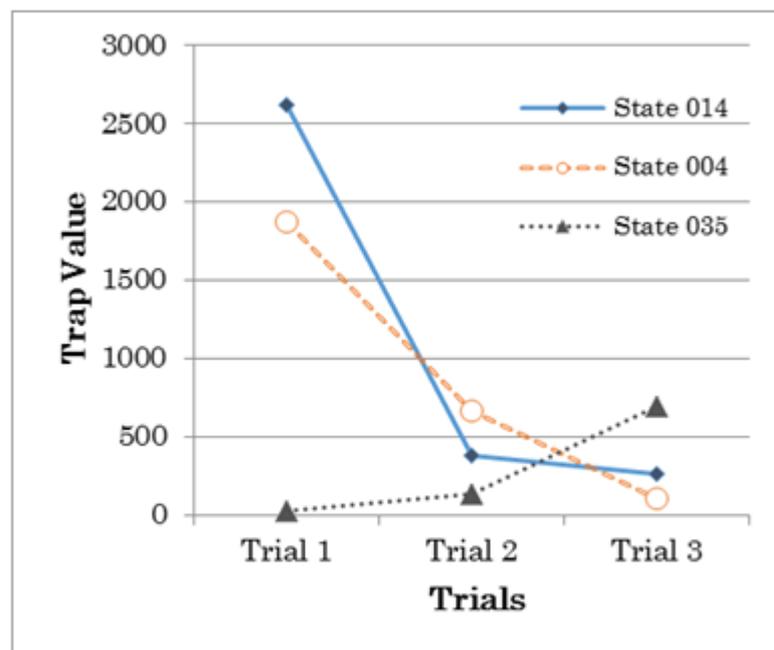


Figure 4-11 Portion of the number of states in Assignment 10 to the occurrence frequency.

Finally, we confirm that by using these visualizations, trap states for learners could be detected and they changed for each trial. This means that trap states in the previous trial did not remain trap states in the next trial. Moreover, the number of learners and actions decreased from one trial to the next. This indicates that the number of learners

who had a good understanding of the problem structure increased, although there were small numbers of learners who were still confused.

Table 4-2 Actual trap state for each trial.

<b>Trial</b>	<b>Actual trap state</b>	<b>Number of learners</b>
1	014	27
2	004	22
3	035	6

### 4.6 Chapter Summary

The study in this chapter presented visualizations that externalized the activity of learners in a problem-posing learning system, where learners posed problems based on the requirements of an assignment. The support and distance graphs trace different aspects of the learners’ activity, and the combination of both visualizations produce the trap graph, which could detect the trap states as the bottleneck for learners. Through this detection, we are able to receive information regarding situations in which many learners have some misunderstandings on the structure of the problems. The detected trap state is obtained based on the highest trap value. The current trap value is detected via tentative analysis, and thus, we need to sophisticate the calculation. At the minimum, we need to distinguish between the high-supported state that was arranged by many learners and the low-supported state that was arranged by few.

In addition, we conducted three trials to investigate their progress of thinking in posing the same type of problem. From the analysis, we found that the trap state changed from the first trial to the third trial. Some learners had a tendency to change their strategy to pose the problem for each trial. The changes in their strategies caused the number of learners who were trapped to decrease from one trial to the next. We infer that the number of learners who had a good understanding of the problem structure increased as a result of this activity.

---

# DESIGNING A COMPUTER-BASED SCAFFOLDING ON A PROBLEM-POSING LEARNING SYSTEM

**Summary:** This chapter presents a study proposes a scaffolding in a problem-posing learning system, called Monsakun, to increase learner's awareness of the problem's structure of arithmetic word problems. Scaffolding defined as supports for learners in order to enhance learning and Monsakun is an interactive problem-posing learning system to facilitate learning in arithmetic word problems. A study in this chapter presents a design of a computer-based scaffolding system. The proposed system detect individual learner's bottleneck through problem-posing processes in the real-time when they pose a problem and adaptively provide a personalized task based on the bottleneck found. It is aimed to support the learner in overcoming the bottleneck and gaining a better understanding of the problem's structure. In Monsakun, the bottlenecks are often detected as a frequently repeated use of a specific dummy. If such dummy can be detected, it is the key factor to support learners to overcome their difficulty. This support system is expected to foster learner's consideration about the problem structure while posing arithmetic word problems on Monsakun.

## 5.1 Introduction

As technology extends learning from classrooms to learning systems, the scaffolding is no longer implemented via face-to-face instruction that literally exists between a teacher and learners in a classroom. Recently, the form of supports that emerges to learners is facilitated through technology. Learners received computerized scaffolds supporting their

metacognitive activities in the learning process (Molenaar and Roda, 2008). These scaffolds were given when metacognitive activities are typically executed in the learning process. The computerized scaffolding system generated the appropriate instance to send a scaffold based on the learner's attention focus.

There has been considerable work developing the scaffolding modules to enhance students in learning. Carr et al. (2011) developed Ecolab, an interactive learning system designed to support 10 - 11 years old learners' understanding of ecology. The system offers help at different levels of specificity and invites users to consider what level of help they need – a form of metacognitive assistance. Teo and Chai (2009) focused on the use of sentence openers as the main form of a scaffold in an online discussion activity. Molenaar et al. (2012) focused on supporting students in their virtual collaboration system with experts, called *Ontdeknet* (Molenaar, 2003). The experts provide students with information about their subject of expertise, give the assignment and monitor students' progress through a 3-D agent David with a particular expression. Using a similar way, Tiantong and Teemuangsai (2013) conducted scaffolding modules which represented by a 3-D animation experts cartoon to attract students in giving concepts and using logics and problem structure to fight for solutions. We developed scaffolding with a different way by providing an assignment based on learner's bottleneck automatically to enhance their awareness of the problem's structure in the process of posing arithmetic word problems.

Intelligent tutoring systems can dramatically increase learners' comprehension by adapting the learning activity to the learners' needs, based on an intelligent assessment of their level of knowledge. A system for personalization based on hybrid recommendation strategy and learning style identification has been developed (Klašnja-Milićević et al., 2011). Ben-Naim et al. (2008) propose the Solution Trace Graph from learners' activities on Adaptive eLearning Platform that shows learners' behavior during the Adaptive Tutorial. The analysis of adaptive tutorials is always performed with the purpose of refining and improving them for the next time they run. Using the result of the analysis, the feedback of the tutoring system that did not help the students to understand their mistake can be changed to be more specific and helpful. Moreover, Ben-Naim et al. (2009) defined trap states as pre-defined error answers and specified an overly general

trap states based on the learners' answers. When researching why such a mistake was so prominent, it was noticed that learners' mistake was that they used a particular strategy. The system will then simply add a new trap state targeting the misconception of the particular strategy. Supianto et al. (2015a) conducted detection of trap states by tracing problem-posing activity sequences from the system's log data in problem-posing learning system (mostly discussed in Chapter 4). In addition to Ben-Naim et al. (2009), Supianto et al. (2016b) proposed the definition of trap states not only based on the learners' answers but also including the process of arranging the answer. Moreover, the trap state was found in the study is conducted by technology in order to detect the bottlenecks of learners based on their activity, instead of pre-defined errors, in the group of learners. In contrast, the study in this chapter, detecting trap state is implemented for each learner in order to support individually based on the process of learning by problem-posing in arithmetic word problems.

Having described how learning by problem posing via sentence integration using Monsakun is working in the previous chapter (this is mainly explored in Section 2.3), it was confirmed that this approach makes simple and goal-oriented problem-posing tasks even for lower-grade of elementary school students while maintaining its value as a viable learning method. Nevertheless, in some cases, the system provides an assignment that makes learners difficult and got stuck in posing the problem. Therefore, there is a need to support learners in problem-posing processes for helping them to pose meaningful problems. This can be achieved by generating an assignment automatically based on the history of their difficult, and hence, the assignment would meet the learner's level of understanding. This chapter presents a study to design a computer-based scaffolding system to enhance learner's awareness of problem's structure while posing the problems on Monsakun. Figure 5-1 shows the goals, contributions and validation methods addressed in this chapter. The study in this chapter describes the system architecture, system components, and details of the procedures for the proposed scaffolding on Monsakun.

This chapter is structured as follows. Section 5.2 presents problem formulation to show how the bottleneck occurred through learner's problem-posing process in Monsakun. Section 5.3 presents the design of proposed scaffolding system consists of the system

architecture, procedures, and process of generating new assignment. Finally, the chapter is summarized in Section 5.4.

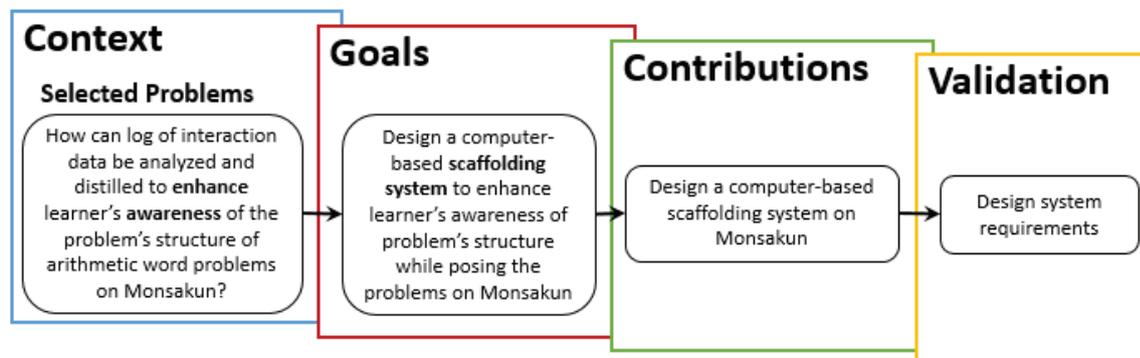


Figure 5-1 Context, goals, contributions and validation of Chapter 5.

## 5.2 Problem Formulation

In the study of this chapter, each composition of sentence cards in Monsakun is treated as a "state" in the problem-posing process by the learner. All possible combinations of sentence cards and transitions among them could be clearly defined as a network of states (Supianto et al., 2016b). All the actions of a learner could be mapped to a transition from one state to another in this network. All possible states can be obtained by combining all the available sentence cards. Each state represents a basic unit of thinking, and a problem state space provides the range of thinking in a problem-posing assignment. Based on the task model in Monsakun (Kurayama and Hirashima, 2010), five main constraints to be satisfied by each posed problem have been devised, which are: calculation, story type, number, objects, and sentence structure. When all five constraints are satisfied, the learner has succeeded in posing a valid problem according to the requirement. When less than five constraints are satisfied, it shows that the learner has acquired a level of understanding of the problem's structure but the final problem does not yet satisfy the requirements. If there are no constraints satisfied by the learner, it shows that the learner is unable to understand the structure of the arithmetic word problem.

Although learners have a tendency to keep or enhance the meaningfulness of posed problems even in the middle of the problem-posing process (Supianto et al., 2016a), the average of steps and mistake in the top-level was very high (Hasanah et al., 2015b), which

shows that the top-level was indeed very challenging for learners. The average of steps and mistakes shows how many steps a learner needed in order to pose a correct answer in one assignment and how many mistakes the learner made during the process, respectively. Ideally, a learner would only need 3 steps to pose a correct answer, because a problem in Monsakun consists of the arrangement of 3 simple sentence cards. However, when a learner combines the sentence cards due to its possible potential, it has consequences. The consequences could move learners away from the correct answer. In this case, the learners are stuck in a condition where they would have to do more steps to reach the correct answer. In other words, the learners are trapped in the composition that distanced them from the correct answer. Then, we call such state as an actual trap-state as the bottleneck in thinking. In this study, a method for detection of that bottlenecks and support system for overcoming them is proposed.

The example of learners' log data of Level 5 Assignment 1, which is designed as combination story problem and presented in Chapter 4.4.1. It will be used for explaining the procedure suggested in this study. The requirement is to *make a word problem about "How many are there in all" that can be solved by "8-3."* and the sentence cards are presented as follow:

1. **There are 3 white rabbits.**
2. **There are ? black rabbits.**
3. **There are 8 white and black rabbits altogether.**
4. There are 8 white rabbits.
5. There are 3 more white rabbits than black rabbits.
6. There are 3 brown rabbits.

In this assignment, the correct answer consists of sentence card #1, sentence card #2, and sentence card #3 (the sentences appearing in bold), while sentence card #4, sentence card #5, and sentence card #6 are identified as dummy cards.

Figure 5-2 depicts an illustrative example of processed log data of a learner's problem-posing activity of the assignment above. The brown and red rectangles (see Figure 5-2a) indicate the uncomplete states and the mistaken states, respectively. In this example, the total number of steps is 64 times and the total number of mistakes is 18 times, which

indicates that the learner feels difficult to reach the correct answer, and hence, need to be supported. The yellow nodes in Figure 5-2b indicate the states that had already been made by the learner. The green lines show the path of the learner's steps and the thickness of the link indicates the number of steps from a node to another. Based on this information, it looks the learner got stuck in State 004, 014, 134, and 034. The learner seems to struggle to achieve the correct answer, which is State 123.

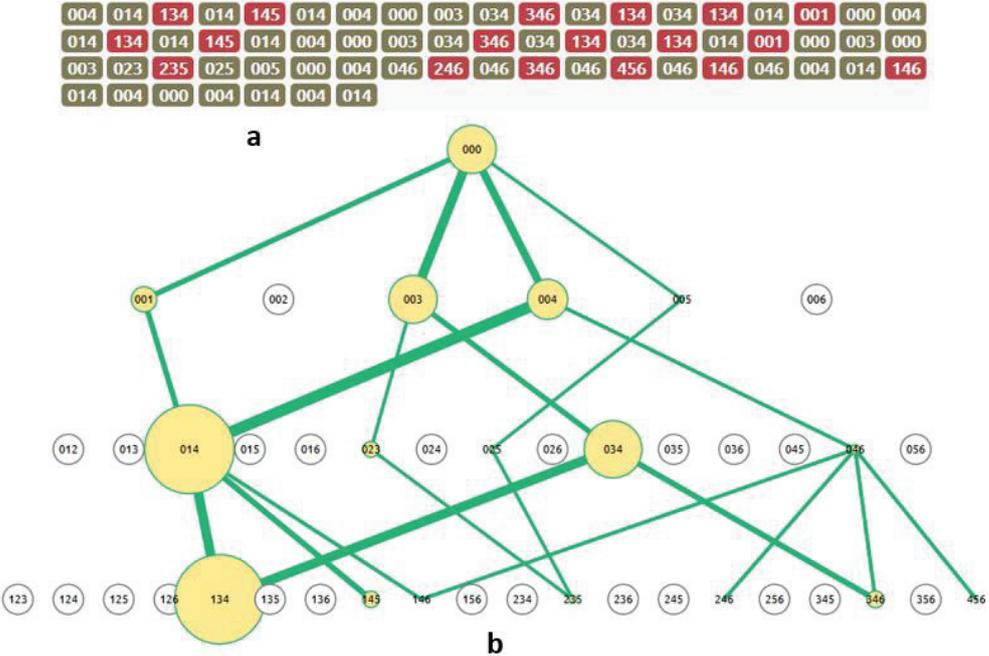


Figure 5-2 An example of processed log data of a learner’s problem-posing activity. a The sequence of states. b Graphical representation.

According to the problem above, a new method implemented to the learning system that can support learners according to their bottleneck is conducted in order to help them out of the impasse. The proposed system presents adaptive assignment setting based on the detected bottleneck. This setting expects the learner’s awareness of the problem's structure in the problem-posing process. Therefore, the learner has a better understanding to their bottleneck and could accomplish the next assignment smoothly.

### 5.3 Proposed Scaffolding System

This section describes the system architecture, system components, and details of the procedures for the proposed problem-posing learning system.

### 5.3.1 System Architecture and Components

The system architecture is shown in Figure 5-3. The system consists of Monsakun and proposed systems. The Monsakun system consists of four functional modules: 1) interface module, which manages the appearance of an interactive and interesting learning system, so the learner can operate the learning system easily and smoothly; 2) problem-posing assignment module, which presents the storage for all essential learning assignments, such as requirements, sentence cards, and rules; 3) application module, which delivers the assignment and performs an evaluation of learner's answer based on the rules, this evaluation responsible for providing a feedback message to the learner regarding his/her mistakes; and 4) log data module, which collects every learner's steps and records it into the files.

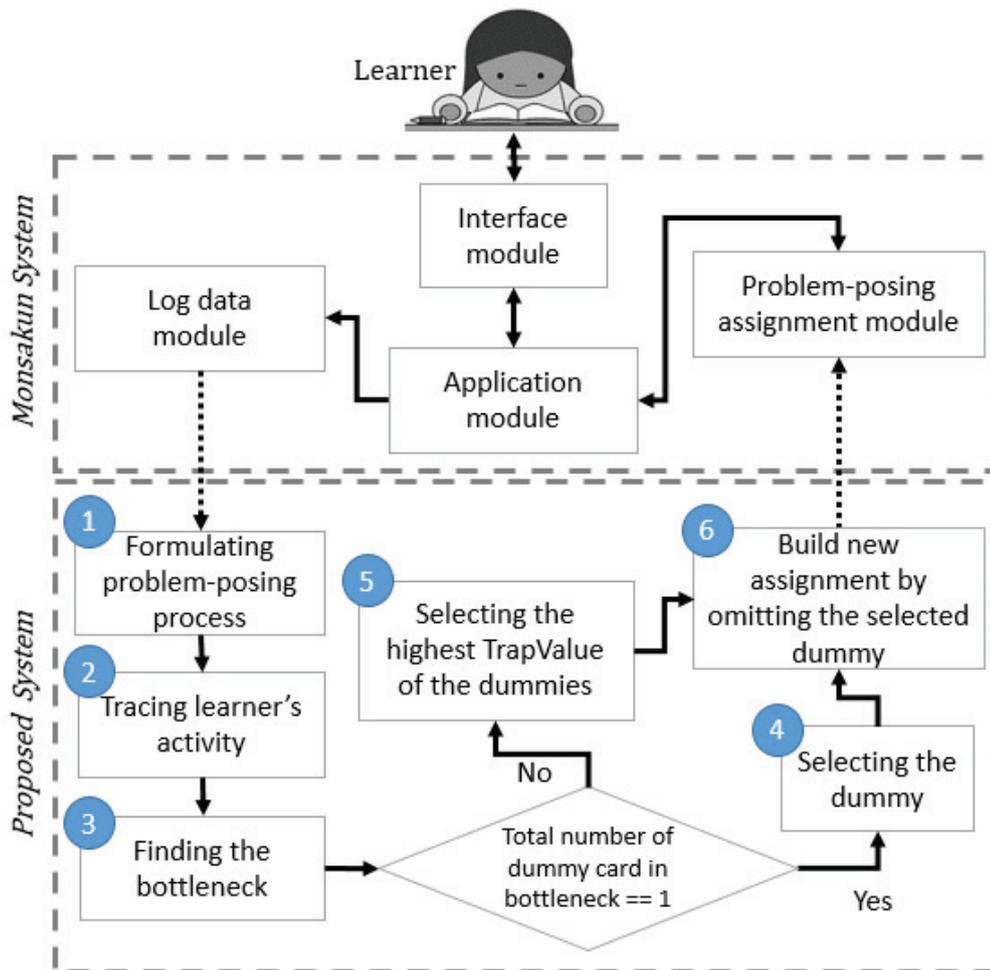


Figure 5-3 The system architecture of proposed problem-posing learning system.

The proposed system consists of two functional modules: 'detection of bottleneck' and 'recommendation setting'. The detection of bottleneck module responsible for finding the learner's bottleneck. It processes the log data of learner's activities and provides bottleneck, which is represented by Module 1 to Module 3. The recommendation setting module generates new assignment setting to a learner based on the bottleneck found in the previous stage. It is handled by Module 4 to Module 6. As the result, we remove a specific dummy sentence card that makes learner stuck while posing the problem and build new assignment by omitting this dummy.

### 5.3.2 Operation Procedures

Based on the system architecture mentioned above, the proposed system operation procedures are briefly described as follows:

- Stage 1.* Formulating problem-posing processes by encoding states performed by the learner using indexing number. When the slot is still empty, index = 0 is implemented. For instance, when the learner poses the problem by selecting sentence card #1 and arranges it to the second slot, State 010 is generated.
- Stage 2.* Tracing learner's activity and processed into a sequence of states. Every state stored by the system sequentially produce ordered states which represent the sequenced learner's steps. This steps reflect the learner's thinking processes. The sequence is then used in the next processes to detect the bottlenecks. An example of a sequence that consists of 64 states is shown in Figure 5-2a.
- Stage 3.* Finding the bottleneck state based on the highest trap value obtained from the support and distance values (Supianto et al., 2016b). The size of nodes presented in Figure 5-2b depicts the trap values. The node with a larger size has a higher value of trap than the node with smaller sizes. According to that, State 134 with the highest score is selected as the bottleneck state. When the highest score happens in more than one state, then all states with the highest score will be selected as trap-states. For instance, when State 134 and 014 have the same value then both states will be selected. Next, the system will check the total number of dummy card in the bottleneck state(s). If the dummy card in the bottleneck is only one card, then go to Stage 4. Otherwise, go to Stage 5.

- Stage 4.* Selecting the dummy card in the bottleneck state, which is obtained from Stage 3. In case the bottleneck state is State 134, then the selected dummy card is the sentence card #4.
- Stage 5.* Selecting the highest trap-value of the dummies and select the dummy having the highest score. For example, let say State 145 with the highest trap value is selected as a bottleneck. This state has two dummy cards, sentence card #4 and sentence card #5. The system sums-up the total trap values of states which containing each dummy. As result, total trap value of sentence card #4 and sentence card #5 are obtained. If sentence card #4 has the highest value then the selected dummy card is the sentence card #4, and it also applies to the sentence card #5.
- Stage 6.* Using the selected dummy obtained from the previous stage, the system generates the new setting of an assignment by removing the dummy on the available sentence cards. Then, the system provides the next assignment to the learner.

### 5.3.3 New Setting of an Assignment

Removing the dummy, which is from the beginning intended as a distractor and eventually becomes the cause of the bottlenecks, might be able to overcome their impasse. We propose a method to change the characteristic of assignment based on the learners' bottlenecks. The changes of the assignment are generated from the previous one in order to guide the learners. The characteristic of current assignment has caused learners difficult to get the correct answer. Therefore, by changing the characteristic of assignment to be less difficult, having a good understanding of the problem's structure is expected. In this study, the characteristic of the assignment is defined as the violation of the constraints.

Figure 5-4 shows the proportion of the number of states that violate the constraints. It can be seen that the new setting when sentence card #4 is removed will help learners having difficulty in calculation constraint. The number of states in the calculation constraint of this setting is lowest than the other new settings. Thus, this setting is fit for learners that have bottlenecks in states related to the calculation constraints. On the other hand, the new setting when sentence card #5 is removed will help learners having

difficulty in the story constraint because of the possibility of the violation of constraint about the story is decreased. While when the new setting in which sentence card #6 is removed will support learners that have difficulty in objects constraint because of the possibility of the violation of constraint about objects is decreased. Therefore, depending on the characteristic of the difficulty that faced by learners, the new setting would be generated.

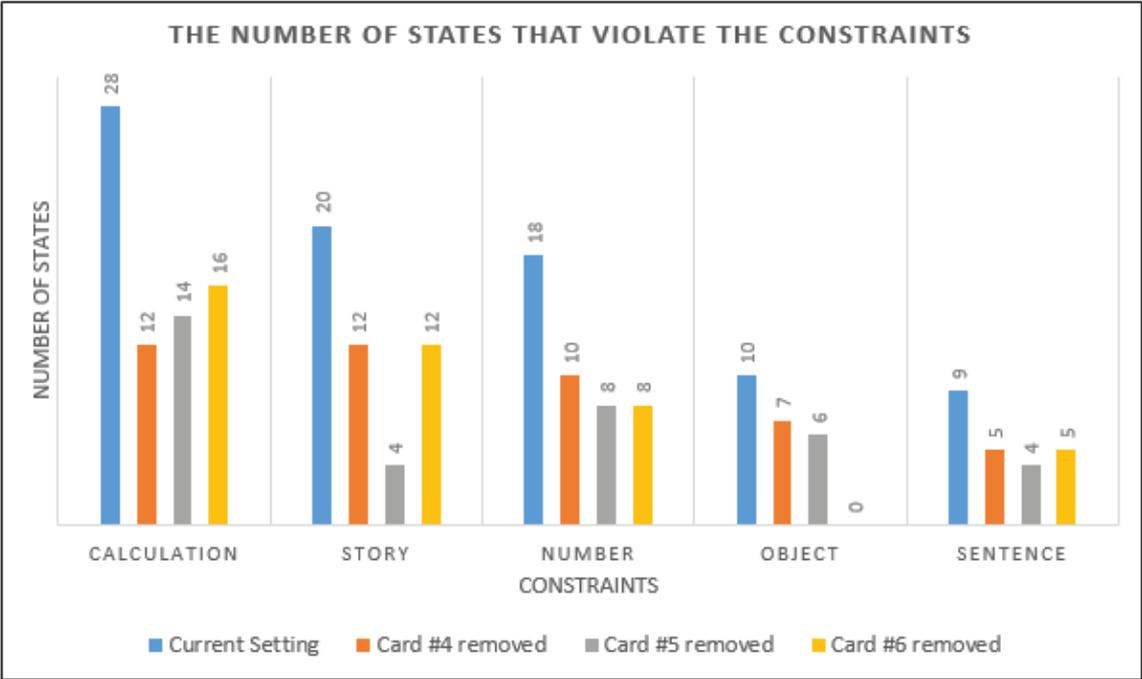


Figure 5-4 The proportion of the number of states that violate the constraints.

Based on the example discussed above, the bottleneck state is State 134. According to the requirement, the difficulty in the assignment is that there is a gap between the required story type of combination story problem and the numerical expression of subtraction, “8-3.” Although subtraction usually implies the story type of decrease and comparison, in this case, the requirement is to pose a problem of the combination story type. To complete the assignment, it is needed to transform the numerical expression, “8-3,” into the calculation expression representing a combination story, such as “3 + ? = 8,” or “? + 3 = 8.” Thus, the learner could assign existence sentence cards to “3” and “?” (Unknown number).

According to the state 134, the learner seems kept trying to use sentence card #4 (*There are 8 white rabbits*). There was a decided tendency based on some sort of thinking

that the number “8” in the numerical expression was considered to be an independent quantity sentence. However, the learner also used sentence card #3 (*There are 8 white and black rabbits altogether*) which represents that the number “8” was considered as a relative quantity sentence. In this situation, the learner confused and stuck in keeping sentence card #4 (*There are 8 white rabbits*) to pose the correct answer. Such kind of difficulty, it would be helpful for learner when the assignment is changed to the new setting when sentence card #4 is removed. In addition, according to the number of states that violate the constraints, the calculation constraint of the new setting, which sentence card #4 is removed, is lowest than the other new settings. Our expectation is illustrated in Figure 5-5.

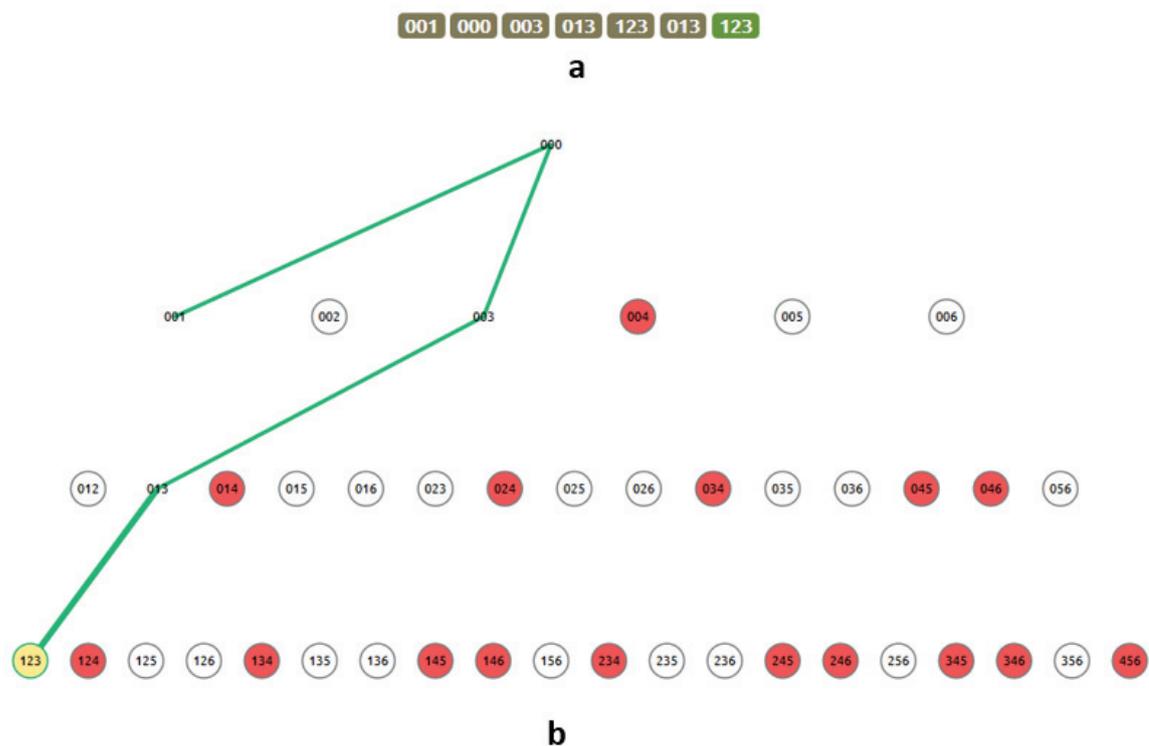


Figure 5-5 An example of processed log data of a high-performance learner’s problem-posing activity. a The sequence of states. b Graphical representation.

Figure 5-5 shows an example of problem-posing activity from a high-performance learner in the same assignment as described in previous. The learner poses the problem smoothly within 7 steps to reach the correct answer (see Figure 5-5a). The learner has a good understanding and avoids compositions that include the dummy sentence card #4, which is shown as the red nodes in Figure 5-5b. In the third step, the learner composed

State 003, which means the learner used sentence card #3 (*There are 8 white and black rabbits altogether*). The learner realized that the number “8” was considered as a relative quantity sentence, instead of an independent quantity sentence. Therefore, the learner evaded in composing State 004, a state that contains sentence card #4 (*There are 8 white rabbits*). Such awareness about the problem structure is expected to the low-performance learners by giving support to them while posing the problems. In this paper, the support is implemented by providing the assignment based on their bottleneck of thinking.

## 5.4 Chapter Summary

We designed a scaffolding system to increase learner’s awareness of the problem structure in problem-posing process on Monsakun, a problem-posing learning system of arithmetic word problems as sentence integration. The proposed system detect individual learner's bottleneck in real-time while posing a problem and adaptively provide the next assignment setting based on the bottleneck found in the previous assignment. It is aimed to support the learner in overcoming the bottleneck and gaining a good understanding of the problem's structure. The proposed system responsible for formulating problem-posing process from log data of learner’s activities, finding a bottleneck condition, and generating a new assignment to the learner. For the future work, we plan to do the practical use in the classroom and analyze the result.

---

## CONCLUSIONS AND FUTURE WORK

### 6.1 Summary of Studies

Learning Analytics and Educational Data Mining are emerging fields that aim to exploit learner's data to discover and make visible several aspects the learning processes. These can include learner's strategies, patterns of behavior, quantitative indicators of activity and quality of learner's products. Our research on problem-posing learning system offer approaches to investigate and discover log data activity that can help teachers, learners, and researchers enhance learning by problem posing, especially on Monsakun, an interactive learning system that promotes learning by problem posing of arithmetic word problem as sentence integration.

This thesis set out to investigate how log data of problem-posing processes can be automatically analyzed and exploited through data analytics techniques, in order to confirm learners' consideration of the problem's structure and to discover learners' bottleneck in thinking when they pose the problems for supporting them enhance their awareness of the problem's structure in posing arithmetic word problems. To achieve this, we proposed an approach grounded on the intersection of the three fields: Human-Computer Interaction (HCI), Learning Analytics (LA) and Educational Data Mining (EDM). We defined the problematic in terms of three main thesis questions:

1. Do learners consider the structure of arithmetic word problem in posing problems using Monsakun?
2. Do the history of learners' activity provide useful information regarding their bottlenecks in thinking while posing problems on Monsakun?

3. How can log of interaction data be analyzed and distilled to enhance learner's awareness of the problem's structure of arithmetic word problems on Monsakun?

To address these questions, we conducted studies presented in previous chapters. Figure 6-1 presents a summary of the thesis goals that address the questions described above. The figure also illustrates how each chapter of the thesis is associated with those. Chapter 3 presented studies to investigate learners' thinking through their actions during problem-posing processes on Monsakun in term of the constraints. Chapters 4 described a visual representation as educational data mining approach in order to discover the bottlenecks of learners' problem-posing processes. Building on top of the previous chapters, Chapters 5 expressed study in designing a computer-based scaffolding system to enhance learner's awareness of problem's structure while posing the problems on Monsakun.

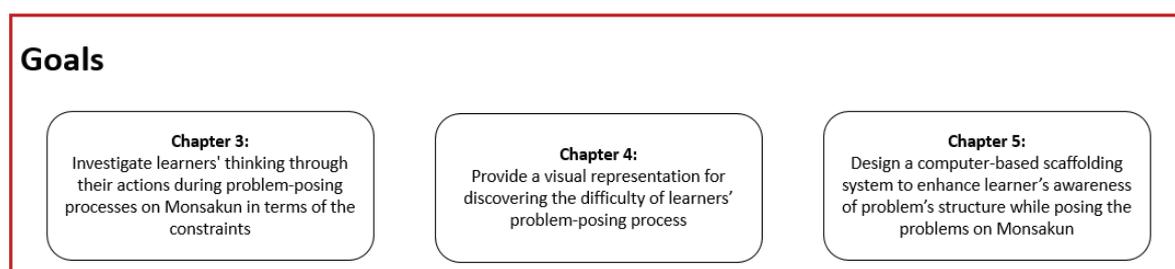


Figure 6-1 Summary of the thesis goals matching to the thesis chapters 3-5.

### *Chapter 3*

In this chapter, we present two studies that conducted a model-based analysis of problem-posing activity including posed problem and intermediate products while posing problems from Monsakun log data of first-grade elementary school students to investigate their actions while posing arithmetic word problems. The assumption in the first study is that they consider the structure of arithmetic word problems as sentence-integration through satisfying as many constraints as possible. The result shows that significant correlation in 11 out of 12 assignments is found, which shows that many actions performed by learners had an inclination to satisfy as many constraints as possible. Further investigation shows that learners had a tendency to enhance the validity of the intermediate states to achieve the correct answers. It means that learners tended to pose

the required problem with an awareness of the structure of arithmetic word problems throughout the problem-posing process.

The second study focused on the violation of the constraints. The analysis involves intermediate products to prove that the learners attempt to avoid invalid intermediate products. Correlation between the numbers of violated constraints and the frequency of each intermediate product that the learners actually made was reported. Moreover, to determine the detail trends of learners' actions, a chi-square test between the number of states in the assignment setting and the occurrence frequency was conducted. Significant correlation and difference in 11 out of 12 assignments was found, which shows that many actions performed by learners had the inclination to avoid as many violated constraints as possible. It indicates that they tended to avoid as many mistakes as possible. Furthermore, although learners tended to avoid the violated constraints, they could not avoid some mistakes. However, most of the learners' mistakes violated at most two constraints. Further analysis shows that, in 12 assignments, learners generally have difficulty fulfilling 2 out of 5 constraints, which are "story" and "calculation" constraints. Based on this analysis, it would be possible to detect the difficulty of learners' actions from the model perspective. Hence, accurate feedback and appropriate support can be provided.

#### *Chapter 4*

The study in this chapter presented visualizations that externalized the activity of learners in a problem-posing learning system, where learners posed problems based on the requirements of an assignment. The support graph provides the number of states that were visited by learners. The distance graph depicts the number of steps to the correct answer. These visualizations trace different aspects of the learners' activity, and the combination of both visualizations produce the trap graph, which could detect the trap states as the bottleneck for learners. Through this detection, we are able to receive information regarding situations in which many learners have some misunderstandings on the structure of the problems. In addition, the study conducted three trials to investigate their progress of thinking in posing the same type of problem. From the analysis, we found that the bottleneck changed from the first trial to the third trial. Some learners had a tendency to change their strategy to pose the problem for each trial. The changes in their strategies caused the number of learners who were trapped to decrease from one trial to

the next. We infer that the number of learners who had a good understanding of the problem structure increased as a result of this activity.

### *Chapter 5*

In this chapter, study on designing a scaffolding system to increase learner's awareness of the problem structure in problem-posing process on Monsakun is presented. The proposed system detect individual learner's bottleneck (according to the study in Chapter 4) in real-time while posing a problem and adaptively provide the next assignment setting based on the bottleneck found. It is aimed to support the learner in overcoming the bottleneck and gaining a good understanding of the problem's structure. The proposed system responsible for formulating problem-posing process from log data of learner's activities, finding a bottleneck condition, and generating a new assignment to the learner.

## 6.2 Future Directions

In this thesis, we focused on showing that our approach can be effectively applied to help teachers and researchers understanding of learners' thinking through their log of problem-posing activity. The information can be used to provide appropriate feedbacks to learners in order to enhance their awareness of the problem's structure in posing arithmetic word problems on the learning system. However, the investigation can be extended to a wide range of areas of research and practice. Moreover, a scaffolding system also can be developed to support learners in enhancing their learning. The description of our future work is aligned with current educational technology trends and can be condensed into two main objectives: to analyze learning processes of learners' activity on learning systems and provide personalized/adapted learning experiences.

*Exploring and analyzing other learning situations and dimensions of learner's activities on the problem-posing learning system.*

We plan to analyze in more detail the characteristics of learners' thinking process. We also would like to explore methods to identify other significant actions. In particular, according to Chapter 3 and Chapter 4, it is necessary to investigate which constraints learners actually pay attention in a variety of situations. Furthermore, we would like to

use other data mining techniques in extending study in Chapter 5, such as sequential data mining to discover learners' action sequences while posing the problems and use the clustering method for grouping learners' thinking processes.

*Practical use of the computer-based scaffolding system.*

According to study in Chapter 6, we plan to implement the developed new system including the scaffold in practical uses in the actual classes. Next, we would like to investigate what is the learning effect on implementing the scaffolding system.

*<empty page>*

## REFERENCES

---

- Abramovich, S., & Cho, E. K. (2015). Using digital technology for mathematical problem posing. In F. M. Singer, N. F. Ellerton, & J. Cai (Eds.), *Mathematical Problem Posing* (pp. 71-102). New York: Springer.
- Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*, 147-181.
- Anjewierden, A., Gijlers, H., Saab, N., & de Hoog, R. (2011). Brick: mining pedagogically interesting sequential patterns. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Ed.), *The 4th International Conference on Educational Data Mining (EDM 2011)* (pp. 341-342). Eindhoven, the Netherlands: International EDM Society.
- Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*(1), 17-21.
- Arroyo, I., & Woolf, B. P. (2003). Students in AWE: changing their role from consumers to producers of ITS content. *The 11th International AIED Conference, Workshop on Advanced Technologies for Math Education*. Sydney, Australia: University of Sydney.
- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition-implications for the design of computer-based scaffolds. *Instructional Science*, *33*, 367-379.
- Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate student's ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, *29*(3), 344-370.
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development*, *56*(1), 45-72.
- Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM-Journal of Educational Data Mining*, *1*(1), 3-17.
- Barak, M., & Rafaeli, S. (2004). On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. *International Journal of Human-Computer Studies*, *61*(1), 84-103.
- Beal, C. R., & Cohen, P. R. (2012). Teach ourselves: Technology to support problem posing in the STEM classroom. *Creative Education*, *3*(4), 513-519.

- Beal, C. R., Arroyo, I., Cohen, P. R., & Woolf, B. P. (2010). Evaluation of AnimalWatch: an intelligent tutoring system for arithmetic and fractions. *Journal of Interactive Online Learning, 9*, 64-77.
- Belland, B. R., Walker, A. E., Olsen, M. W., & Leary, H. (2015). A Pilot Meta-Analysis of Computer-Based Scaffolding in STEM Education. *Educational Technology & Society, 18*(1), 183-197.
- Ben-Naim, D., Bain, M., & Marcus, N. (2009). A user-driven and data-driven approach for supporting teachers in reflection and adaptation of adaptive tutorials. In T. Barnes, M. Desmarais, C. Romero, & S. Ventura (Ed.), *The 2nd International Conference on Educational Data Mining (EDM 2009)*, (pp. 21-30). Cordoba, Spain.
- Ben-Naim, D., Marcus, N., & Bain, M. (2008). Visualization and analysis of student interaction in an adaptive exploratory learning environment. *International Workshop on Intelligent Support for Exploratory Environment, EC-TEL, Vol. 8*.
- Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning, 5*(2), 123-152.
- Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching: a new agent paradigm for educational software. *Applied Artificial Intelligence, 19*, 363-392.
- Bonotto, C. (2013). Artifacts as sources for problem-posing activities. *Educational Studies in Mathematics, 83*(1), 37-55.
- Brown, S. I., & Walter, M. I. (1993). *Problem posing: reflections and applications*. Hillsdale: Lawrence Erlbaum Associates.
- Cai, J., & Jiang, C. (2016). An analysis of problem-posing tasks in Chinese and US elementary mathematics textbooks. *International Journal of Science and Mathematics Education, 1-20*. doi:10.1007/s10763-016-9758-2
- Cankoy, O. (2014). Interlocked problem posing and children's problem posing performance in free structured situations. *International Journal of Science and Mathematics Education, 12*(1), 219-238.
- Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Carr, A., Luckin, R., Avramides, K., & Yuill, N. (2011). Scaffolding Metacognitive Processes in the Ecolab: Help-Seeking and Achievement Goal Orientation. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Ed.), *The 15th International Conference on Artificial Intelligence in Education (AIED 2011)*. 6738, pp. 432-434. Auckland, New Zealand: Springer-Verlag Berlin Heidelberg.
- Chang, K. E., Wu, L. J., Weng, S. E., & Sung, Y. T. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education, 58*(2), 775-786.

- Chen, L., Van Dooren, W., Chen, Q., & Verschaffel, L. (2011). An investigation on Chinese teachers' realistic problem posing and problem solving ability and beliefs. *International Journal of Science and Mathematics Education*, 9(4), 919–948.
- Choi, I., Land, S. M., & Turgeon, A. Y. (2005). Scaffolding peer-questioning strategies to facilitate metacognition during online small group discussion. *Instructional Science*, 33, 483–511.
- Crespo, S. (2003). Learning to pose mathematical problems: exploring changes in preservice teachers' practices. *Educational Studies in Mathematics*, 52(3), 243–270.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Dabbagh, N., & Kitsantas, A. (2005). Using web-based pedagogical tools as a scaffolds for self-regulated learning. *Instructional Science*, 33, 513-540.
- Devlin, K. J. (2012). *Introduction to mathematical thinking*. Palo Alto: Keith Devlin.
- Dreyfus, T. (2002). Advanced Mathematical Thinking Processes. In D. Tall (Ed.), *Advanced Mathematical Thinking* (Vol. 11, pp. 25-41). Springer Netherlands.
- Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Emery, D., & Sussex, W. (2007). *Effectiveness of reading and mathematics software products: findings from the first student cohort*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.
- El Saadawi, G. M., Azevedo, R., Castine, M., Payne, V., Medvedeva, O., Tseytlin, E., ... & Crowley, R. S. (2010). Factors affecting feeling-of-knowing in a medical intelligent tutoring system: the role of immediate feedback as a metacognitive scaffold. *Advances in Health Sciences Education*, 15(1), 9-30.
- Ellerton, N. F. (2013). Engaging pre-service middle-school teacher-education students in mathematical problem posing: development of an active learning framework. *Educational Studies in Mathematics*, 83(1), 87–101.
- English, L. D. (1996). *Children's problem posing and problem solving preferences* (J. Mulligan & M. Mitchelmore ed.). Adelaide: The Australian Association of Mathematics Teachers Inc.
- English, L. D. (1997a). The development of fifth-grade children's problem-posing abilities. *Educational Studies in Mathematics*, 34(3), 183–217.
- English, L. D. (1997b). Development of seventh-grade students' problem posing. *The 21st Conference of the International Group for the Psychology of Mathematics Education*. 2, pp. 241-248. Lahti, Finland: E. Pehkonen (Ed.).
- English, L. D. (1998). Children's problem posing within formal and informal contexts. *Journal for Research in Mathematics Education*, 29(1), 83-106.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6), 304-317.

- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American psychologist*, 34(10), 906-911.
- Fournier-Viger, P., Faghihi, U., Nkambou, R., & Nguifo, E. M. (2010). Exploiting sequential patterns found in users' solutions and virtual tutor behavior to improve assistance in ITS. *Educational Technology & Society*, 13(1), 13-24.
- Glass, R. L. (1995). A structure-based critique of contemporary computing research. *Journal of Systems and Software*, 28(1), 3-7.
- Green, B. A. (2000). Project-based learning with the world wide web: A qualitative study of resource integration. *Educational Technology Research and Development*, 48(1), 45-66.
- Gröller, E. (2002). Insight into Data through Visualization. In P. Mutzel, M. Jünger, & S. Leipert (Ed.), *The 9th International Symposium on Graph Drawing*. 2265, pp. 352-366. Vienna, Austria: Springer-Verlag Berlin Heidelberg.
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2-3), 107-124.
- Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. *Instructional-design theories and models: A new paradigm of instructional theory*, 2, 115-140.
- Hasanah, N., Hayashi, Y., & Hirashima, T. (2015a). Analysis of Problem-Posing Process of Arithmetical Word Problem as Sentence Integration: Viewpoint of First Selected Sentence. *Emerging issues in smart learning* (pp. 85-88). Springer Berlin Heidelberg.
- Hasanah, N., Hayashi, Y., & Hirashima, T. (2015b). Investigation of students' performance in Monsakun problem posing activity based on the triplet structure model of arithmetical word problems. *The 23rd International Conference on Computers in Education*, (pp. 27-36). Hangzhou, China.
- Hasanah, N., Hayashi, Y., & Hirashima, T. (2017). An analysis of learner outputs in problem posing as sentence-integration in arithmetic word problems. *Research and Practice in Technology Enhanced Learning*, 12(9), 1-16.
- Hayashi, Y., Ogawa, Y., & Nakano, Y. I. (2013). Visualization system for analyzing collaborative learning interaction. *Procedia Computer Science*, 22, 1379-1387.
- Hershkovitz, A., & Nachmias, R. (2009). Learning about online learning processes and students' motivation through web usage mining. *Interdisciplinary Journal of E-Learning and Learning Objects*, 5(1), 197-214.
- Hill, J. R., & Hannafin, M. J. (1997). Cognitive strategies and learning from the world wide web. *Educational Technology Research and Development*, 45(4), 37-64.
- Hirashima, T., & Kurayama, M. (2011). Learning by problem-posing for reverse-thinking problems. *The 15th International Conference on Artificial Intelligence in Education*. 6738, pp. 123-130. Auckland, New Zealand: Springer-Verlag Berlin Heidelberg.

- Hirashima, T., Nakano, A., & Takeuchi, A. (2000). A diagnosis function of arithmetical word problems for learning by problem posing. In R. Mizoguchi, & J. Slaney (Ed.), *The 6th Pacific Rim International Conference on Artificial Intelligence (PRICAI-2000)* (pp. 745-755). Melbourne, Australia: Springer, Berlin Heidelberg.
- Hirashima, T., Yamamoto, S., & Hayashi, Y. (2014). Triplet Structure Model of Arithmetical Word Problems for Learning by Problem-Posing. *The 16th International Conference on Human Interface and the Management of Information. Information and Knowledge in Applications and Services*, 8522, pp. 42-50. Heraklion, Crete, Greece: Springer International Publishing Switzerland.
- Hirashima, T., Yamasaki, K., Fukuda, H., & Funaoi, H. (2015). Framework of kit-build concept map for automatic diagnosis and its preliminary use. *Research and Practice in Technology Enhanced Learning*, 10(17), 1-21. doi:10.1186/s41039-015-0018-9
- Hirashima, T., Yokoyama, T., Okamoto, M., & Takeuchi, A. (2007). Learning by problem-posing as sentence-integration and experimental use. *Artificial Intelligence in Education, 2007*, 254–261.
- Hirashima, T., Yokoyama, T., Okamoto, M., & Takeuchi, A. (2008a). Long-term use of learning environment for problem-posing in arithmetical word problems. *The 16th International Conference on Computers in Education*, (pp. 817–824). Taipei, Taiwan.
- Hirashima, T., Yokoyama, T., Okamoto, M., & Takeuchi, A. (2008b). An Experimental Use of Learning Environment for Problem-Posing as Sentence-Integration in Arithmetical Word Problems. *The 9th International Conference on Intelligent Tutoring Systems*, 5091, pp. 687-689. Montreal, QC, Canada: Springer-Verlag Berlin Heidelberg.
- Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *The Journal of the learning sciences*, 15(1), 53-61.
- Hou, H. T. (2012). Analyzing the learning process of an online role-playing discussion activity. *Educational Technology & Society*, 15(1), 211-222.
- Hsiao, J. Y., Hung, C. L., Lan, Y. F., & Jeng, Y. C. (2013). Integrating worked examples into problem posing in a web-based learning environment. *Turkish Online Journal of Educational Technology-TOJET*, 12(2), 166-176.
- Hsieh, Y. H., Lin, Y. C., & Hou, H. T. (2015). Exploring elementary-school students' engagement patterns in a game-based learning environment. *Educational Technology & Society*, 18(2), 336-348.
- Hung, C. M., Hwang, G. J., & Wang, S. Y. (2014a). Effects of an integrated mind-mapping and problem-posing approach on students' in-field mobile learning performance in a natural science course. *International Journal of Mobile Learning and Organisation*, 8(3), 187-200.
- Hung, P. H., Hwang, G. J., Lee, Y. H., Wu, T. H., Vogel, B., Milrad, M., & Johansson, E. (2014b). A problem-based ubiquitous learning approach to improving the

- questioning abilities of elementary school students. *Education Technology & Society*, 17(4), 316-334.
- Işık, C., Kar, T., Yalçın, T., & Zehir, K. (2011). Prospective teachers' skills in problem posing with regard to different problem posing models. *Procedia-Social and Behavioral Sciences*, 15, 485-489.
- Janssen, J., Erkens, G., & Kanselaar, G. (2007a). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior*, 23(3), 1105-1125.
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007b). Visualization of participation: does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49(4), 1037-1065.
- Johnson, M., Eagle, M., Joseph, L., & Barnes, T. (2011). The EDM vis tool. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Ed.), *The 4th International Conference on Educational Data Mining (EDM 2011)* (pp. 349-350). Eindhoven, the Netherlands: International EDM Society.
- Kar, T., Özdemir, E., İpek, A. S., & Albayrak, M. (2010). The relation between the problem posing and problem solving skills of prospective elementary mathematics teachers. *Procedia-Social and Behavioral Sciences*, 2(2), 1577-1583.
- Kılıç, Ç. (2013). Pre-service primary teachers' free problem-posing performances in the context of fractions: an example from Turkey. *The Asia-Pacific Education Researcher*, 22(4), 677-686.
- Kilpatrick, J. (1987). Problem formulating: Where do good problems come from. *Cognitive science and mathematics education*, 123-147.
- Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education*, 56(3), 885-899.
- Kojima, K., Miwa, K., & Matsui, T. (2015). Experimental study of learning support through examples in mathematical problem posing. *Research and Practice in Technology Enhanced Learning*, 10(1), 1-18.
- Kurayama, M., & Hirashima, T. (2010). Interactive Learning Environment Designed Based on A Task Model of Problem-Posing. *The 18th International Conference on Computers in Education*, (pp. 98-100). Malaysia.
- Lan, Y. F., & Lin, P. C. (2011). Evaluation and improvement of student's question-posing ability in a web-based learning environment. *Australasian Journal of Educational Technology*, 27(4), 581-599.
- Leikin, R. (2015). Problem posing for and through investigations in a dynamic geometry environment. In F. M. Singer, N. F. Ellerton, & J. Cai (Eds.), *Mathematical Problem Posing* (pp. 373-391). New York: Springer.
- Leung, S. K. (1996). Problem posing as assessment: reflections and reconstructions. *The Mathematics Educator* 1, 159-171.

- Leung, S. S., & Silver, E. A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers. *Mathematics Education Research Journal*, 9(1), 5-24.
- Macfadyen, L., & Sorenson, P. (2010). Using LiMS (the learner interaction monitoring system) to track online learner engagement and evaluate course design. In R. Baker, A. Merceron, & P. Pavlik Jr. (Ed.), *The 3rd International Conference on Educational Data Mining (EDM 2010)* (pp. 301-302). Pittsburgh, PA, USA: International EDM Society.
- Majumdar, R., & Iyer, S. (2015). Beyond clickers: tracing patterns in students' response through iSAT. *The 23rd International Conference on Computers in Education*, (pp. 612-621). Hangzhou, China.
- Mazza, R., & Dimitrova, V. (2003). Informing the design of a course data visualisator: an empirical study. *The 5th International Conference on New Educational Environments (ICNEE)*, (pp. 215-220). Lucerne, Switzerland.
- Melušová, J., & Šunderlík, J. (2014). Pre-service teachers' problem posing in combinatorics. *Proceedings of international conference 12th Mathematics Conference in Nitra held under the auspices. Acta Math.* (pp. 115-122). Nitra: Department of Mathematics FNS CPU.
- Mishra, S., & Iyer, S. (2015). An exploration of problem posing-based activities as an assessment tool and as an instructional strategy. *Research and Practice in Technology Enhanced Learning*, 10(5), 1-19.
- Molenaar, I. (2003). Knowledge exchange from citizens to learners through online collaboration. In D. Lassner, & C. McNaught (Ed.), *EdMedia: World Conference on Educational Media and Technology 2003* (pp. 894-899). Honolulu, Hawaii, USA: Association for the Advancement of Computing in Education (AACE).
- Molenaar, I., & Roda, C. (2008). Attention management for dynamic and adaptive scaffolding. *Pragmatics & Cognition*, 16(2), 224-271.
- Molenaar, I., Roda, C., van Boxtel, C., & Slegers, P. (2012). Dynamic scaffolding of socially regulated learning in a computer-based learning environment. *Computers & Education*, 59(2), 515-523.
- Molenaar, I., Van Boxtel, C. A., & Slegers, P. J. (2010). The effects of scaffolding metacognitive activities in small groups. *Computers in Human Behavior*, 26(6), 1727-1738.
- Nakano, A., Hirashima, T., & Takeuchi, A. (1999). Problem-making practice to master solution-methods in intelligent learning environment. *The 8th International Conference on Computers in Education* (pp. 891-898). Japan: University of Electro-Communications & Chiba Institute of Technology.
- Norman, I. (2011). Secondary school students' problem posing strategies: implications to secondary school students' problem posing performances. *Journal of Edupres*, 1, 1-8.
- Oliveira Chaves, R., Gresse von Wangenheim, C., Costa Furtado, J. C., Ronaldo Bezerra Oliveira, S., Santos, A., & Favero, E. L. (2015). Experimental evaluation of a

- serious game for teaching software process modeling. *Education, IEEE Transactions*, 58(4), 289-296.
- Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The journal of the learning sciences*, 13(3), 423-451.
- Pirie, S. E. (2002). Problem posing: what can it tell us about students' mathematical understanding? *The 24th Annual Meeting [of the] North American Chapter of the International Group for the Psychology of Mathematics*. 1-4, pp. 927-958. Athens, Georgia: Eric.
- Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40, 1-12.
- Rabbany, R., Takaffoli, M., & Zaïane, O. R. (2011). Analyzing participation of students in online courses using social network analysis techniques. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Ed.), *The 4th International Conference on Educational Data Mining (EDM 2011)* (pp. 21-30). Eindhoven, the Netherlands: International EDM Society.
- Reiser, B. J., Copen, W. A., Ranney, M., Hamid, A., & Kimberg, D. Y. (1998). *Cognitive and motivational consequences of tutoring and discovery learning*. Northwestern University: ARI Research Note.
- Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601-618.
- Şengül, S., & Katranci, Y. (2012). Problem solving and problem posing skills of prospective mathematics teachers about the 'sets' subject. *Procedia-Social and Behavioral Sciences*, 69, 1650-1655.
- Şengül, S., & Katranci, Y. (2015). Free problem posing cases of prospective mathematics teachers: difficulties and solutions. *Procedia-Social and Behavioral Sciences*, 174, 1983-1990.
- Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive learning environments*, 15(1), 27-46.
- Sherin, B., Reiser, B. J., & Edelson, D. (2004). Scaffolding analysis: Extending the scaffolding metaphor to learning artifacts. *The Journal of the Learning Sciences*, 13(3), 387-421.
- Shneiderman, B. (2002). Inventing discovery tools: combining information visualization with data mining. *Information Visualization*, 1(1), 5-12.
- Shriki, A., & Lavy, I. (2014). Students' self-assessment of creativity: benefits and limitations. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Ed.), *The 38th Conference of the International Group for the Psychology of Mathematics Education*. 5, pp. 177-184. Vancouver, Canada: PME.

- Siemens, G., & d Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. *The 2nd international conference on learning analytics and knowledge* (pp. 252-254). Vancouver, BC, Canada: ACM.
- Silver, E. A. (1994). On mathematical problem posing. *For the learning of mathematics*, 14(1), 19-28.
- Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for Research in Mathematics Education*, 27(5), 521-539.
- Singer, F. M., Pelczer, I., & Voica, C. (2011). Problem posing and modification as a criterion of mathematical creativity. *The 7th Conference of the European Society for Research in Mathematics Education* (pp. 9-11). Poland: University of Rzeszów.
- Singer, F.M., Ellerton, N.F., & Cai, J. (2015). *Mathematical problem posing: from research to effective practice*. New York: Springer.
- Stickles, P. R. (2011). An analysis of secondary and middle school teachers' mathematical problem posing. *Investigations in Mathematics Learning*, 3(2), 1–34.
- Stoyanova, E. (1998). Problem posing in mathematics classrooms. In A. McIntosh, & N. Ellerton, *Research in Mathematics Education: a contemporary perspective* (pp. 164-185). Perth, Australia: Mathematics, Science and Technology Education Centre Edith Cowan University.
- Stoyanova, E. (2005). Problem-posing strategies used by years 8 and 9 students. *Australian Mathematics Teacher*, 61(3), 6-11.
- Sung, H. Y., Hwang, G. J., & Chang, Y. C. (2016). Development of a mobile learning system based on a collaborative problem-posing strategy. *Interactive Learning Environments*, 24(3), 456-471.
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2015a). Tracing problem-Posing activity sequences toward detection of trap-states in thinking. *The 23rd International Conference on Computers in Education*, (pp. 85-90). Hangzhou, China.
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2015b). Analysis of Problem-Posing Activity Sequences toward Modeling Thinking Process and Detection of Trap States. *The 8th Workshop on Technology Enhanced Learning by Posing/Solving Problems/Questions*, (pp. 558-567). Hangzhou, China.
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016a). An Investigation of Learner's Actions in Problem-Posing Activity of Arithmetic Word Problems. *Intelligent Tutoring Systems: 13th International Conference. LNCS, volume 9684*, pp. 478-479. Zagreb, Croatia: Springer International Publishing.
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016b). Visualizations of problem-posing activity sequences toward modeling the thinking process. *Research and Practice in Technology Enhanced Learning*, 11(14), 1-23. doi:10.1186/s41039-016-0042-4

- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2016c). Analysis of Steps in Posing Arithmetic Word Problem as Sentence-Integration on Interactive Learning Environment. *The 24th International Conference on Computers in Education*, (pp. 242-251). Mumbai, India.
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017a). Process-based Assignment-Setting Change for Support of Overcoming Bottlenecks in Learning by Problem-Posing in Arithmetic Word Problems. *Journal of Physics: Conference Series*, 812. doi:10.1088/1742-6596/812/1/012004
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017b). Model-based analysis of thinking in problem posing as sentence integration focused on violation of the constraints. *Research and Practice in Technology Enhanced Learning*, 12(12), 1-21. doi:10.1186/s41039-017-0057-5
- Supianto, A. A., Hayashi, Y., & Hirashima, T. (2017c). An Investigation of Learner's Actions in Posing Arithmetic Word Problem on an Interactive Learning Environment. *IEICE Transactions on Information and Systems*, article in press.
- Tan, W., Lin, S., Yang, Y., & Zhao, X. (2008). Design on Collaborative Virtual Learning Community and Learning Process Visualization. In F. Li, J. Zhao, T. K. Shih, R. Lau, Q. Li, & D. McLeod (Ed.), *The 7th International Conference on Web-Based Learning, Advances in Web Based Learning - ICWL2008*. 5145, pp. 424-433. Jinhua, China: Springer-Verlag Berlin Heidelberg.
- Teo, Y. H., & Chai, C. S. (2009). Scaffolding Online Collaborative Critiquing for Educational Video Production. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 1(1), 51-66.
- Tiantong, M., & Teemuangsai, S. (2013). The four scaffolding modules for collaborative problem-based learning through the computer network on moodle lms for the computer programming course. *International Education Studies*, 6(5), 47-55.
- Tichá, M., & Hošpesová, A. (2009). Problem posing and development of pedagogical content knowledge in pre-service teacher training. *The 6th Conference of the European Society for Research in Mathematics Education*, (pp. 1941-1950). Lyon, France.
- Toluk-Uçar, Z. (2009). Developing pre-service teachers understanding of fractions through problem posing. *Teaching and Teacher Education*, 25(1), 166-175.
- Van Harpen, X. Y., & Presmeg, N. C. (2013). An investigation of relationships between students' mathematical problem-posing abilities and their mathematical content knowledge. *Educational Studies in Mathematics*, 83(1), 117-132.
- Veenman, M. V., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science*, 33(3), 193-211.
- Way, J., & Rowe, L. (2008). The role of scaffolding in the design of multimedia learning objects. *The 11th International Congress on Mathematical Education*. Mexico.
- Wilson, E. V. (2004). ExamNet: Asynchronous learning network: augmenting face-to-face courses with student-developed exam questions. *Computers & Education*, 42(1), 87-107.

- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17(2), 89-100.
- Yamamoto, S., Kanbe, T., Yoshida, Y., Maeda, K., & Hirashima, T. (2012). A case study of learning by problem-posing in introductory phase of arithmetic word problems. *The 20th International Conference on Computers in Education*, (pp. 25–32). Singapore.
- Yamamoto, S., Kanbe, T., Yoshida, Y., Maeda, K., & Hirashima, T. (2013). Learning by Problem-Posing with Online Connected Media Tablets. *The 15th International Conference on Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business*. 8018, pp. 165-174. Las Vegas, NV, USA: Springer-Verlag Berlin Heidelberg.
- Yu, F. Y. (2009). Scaffolding student-generated questions: design and development of a customizable online learning system. *Computers in Human Behavior*, 25(5), 1129-1138.
- Yu, F. Y. (2011). Multiple peer-assessment modes to augment online student question-generation processes. *Computers & Education*, 56(2), 484-494.
- Yu, F. Y., & Pan, K. J. (2014). The effects of student question-generation with online prompts on learning. *Journal of Educational Technology & Society*, 17(3), 267-279.
- Yu, F. Y., Chang, Y. L., & Wu, H. L. (2015). The effects of an online student question-generation strategy on elementary school student English learning. *Research and Practice in Technology Enhanced Learning*, 10(1), 1-16.
- Zorrilla, M. E., Menasalvas, E., Marin, D., Mora, E., & Segovia, J. (2005). Web Usage Mining Project for Improving Web-Based Learning Sites. In M. E. Zorrilla, E. Menasalvas, D. Marin, E. Mora, & J. Segovia (Ed.), *The 10th International Conference on Computer Aided Systems Theory*. 3643, pp. 205-210. Las Palmas de Gran Canaria, Spain: Springer-Verlag Berlin Heidelberg.

*<empty page>*

# APPENDIX

## A.1. Detail Assignments and Sentence Cards in Level 5

Assignment	Requirement	Sentence Cards	
		Correct cards	Dummy cards
1	Make a word problem about 'How many are there overall' that can be solved by '8 - 3'.	There are 3 white rabbits	There are 8 white rabbits
		There are ? black rabbits	There are 3 brown rabbits
		There are 8 white and black rabbits altogether	There are 3 more white rabbits than black rabbits
2	Make a word problem about 'How many are there overall' that can be solved by '12 - 8'.	There are 6 red balloons	There are 10 red balloons
		There are ? yellow balloons	There are 6 blue balloons
		There are 10 red and yellow balloons altogether	There are 6 more red balloons than yellow balloons
3	Make a word problem about 'How many are there overall' that can be solved by '8 - 6'.	I make 8 white flower arrangements	I make 12 white flower arrangements
		I make ? red flower arrangements	I make 8 yellow flower arrangements
		I make 12 white and red flower arrangements altogether	I make 8 more white flower arrangements than red flower arrangements
4	Make a word problem about 'How many are there after increased' that can be solved by '12 - 8'.	There are ? sparrows	There are 8 sparrows
		8 more sparrows come	? sparrows fly away
		There are 12 sparrows	
5	Make a word problem about 'How many are there after increased' that can be solved by '11 - 9'.	There are ? kids	There are 9 kids
		9 more kids come	? kids go away
		There are 11 kids	
6	Make a word problem about 'How many are there after increased' that can be solved by '10 - 2'.	I have ? goldfish	I have 2 goldfish
		I receive 2 more goldfish	I give away ? goldfish
		I have 10 goldfish	

7	Make a word problem about 'How many are left' that can be solved by ' $6 + 4$ '.	I have ? bottles of juice	I have 6 bottles of juice
		I give away 6 bottles of juice	I buy 4 more bottles of juice
		I have 4 bottles of juice	
8	Make a word problem about 'How many are left' that can be solved by ' $8 + 6$ '.	There are ? strawberry cakes	There are 8 strawberry cakes
		8 strawberry cakes are sold	6 strawberry cakes are purchased
		There are 6 strawberry cakes	
9	Make a word problem about 'How many are left' that can be solved by ' $8 + 1$ '.	I have ? pencils	I have 1 pencil
		I give away 1 pencil	I buy 8 more pencils
		I have 8 pencils	
10	Make a word problem about 'How many are the difference' that can be solved by ' $12 - 5$ '.	There are 12 guests coming today	There were 12 guests coming yesterday
		There were ? guests coming yesterday	There are 5 guests coming tomorrow
		There were 5 less guests coming yesterday than today	There are ? guests coming in total of today and yesterday
11	Make a word problem about 'How many are the difference' that can be solved by ' $14 - 6$ '.	There are 14 yellow flowers	There are 14 white flowers
		There are ? white flowers	There are 6 red flowers
		There are 6 less white flowers than yellow flowers	There are ? yellow and white flowers altogether
12	Make a word problem about 'How many are the difference' that can be solved by ' $9 + 4$ '.	There are 9 language exercises	There are 9 arithmetic exercises
		There are ? arithmetic exercises	There are 4 science exercises
		There are 4 less arithmetic exercises than language exercises	There are ? arithmetic and language exercises altogether

## A.2. Rules on Deciding Validity of the Constraints

This Appendix section provides details about the rules on how we calculated the validity values of states. Level = 0 denotes states in which all card slots are empty. Level = 1 shows states in which there is one sentence card in the card slots. Level = 2 shows states in which there are two sentence cards in the card slots and Level = 3 indicates states in which there are three sentence cards in the card slots.

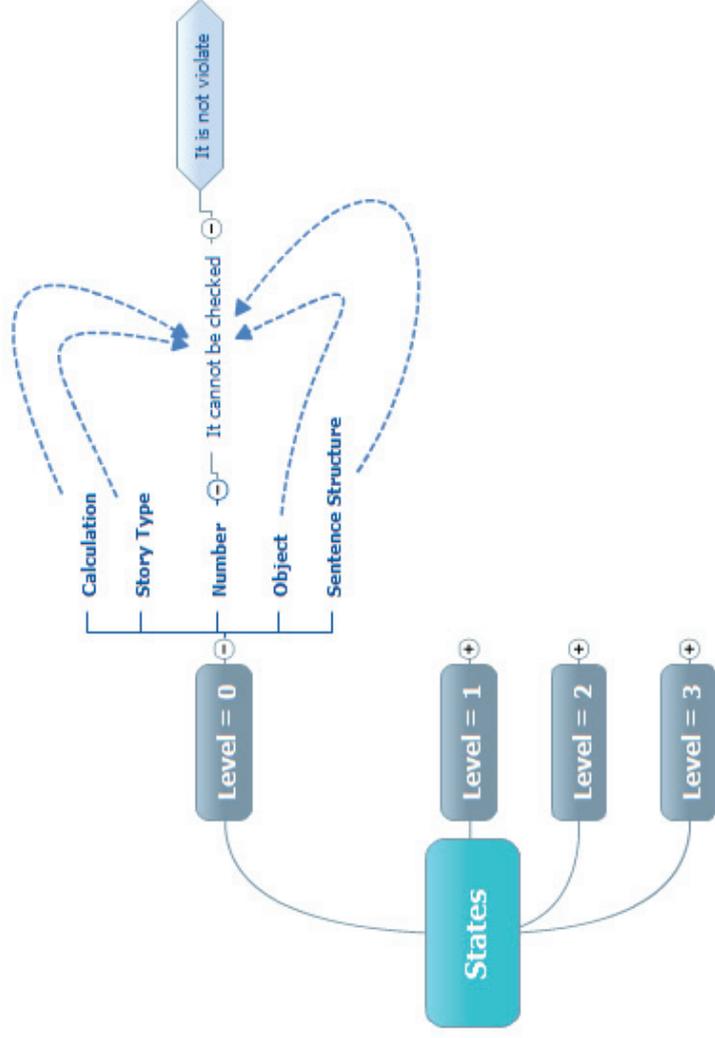


Figure A-1 Rule validity of states in Level 0.

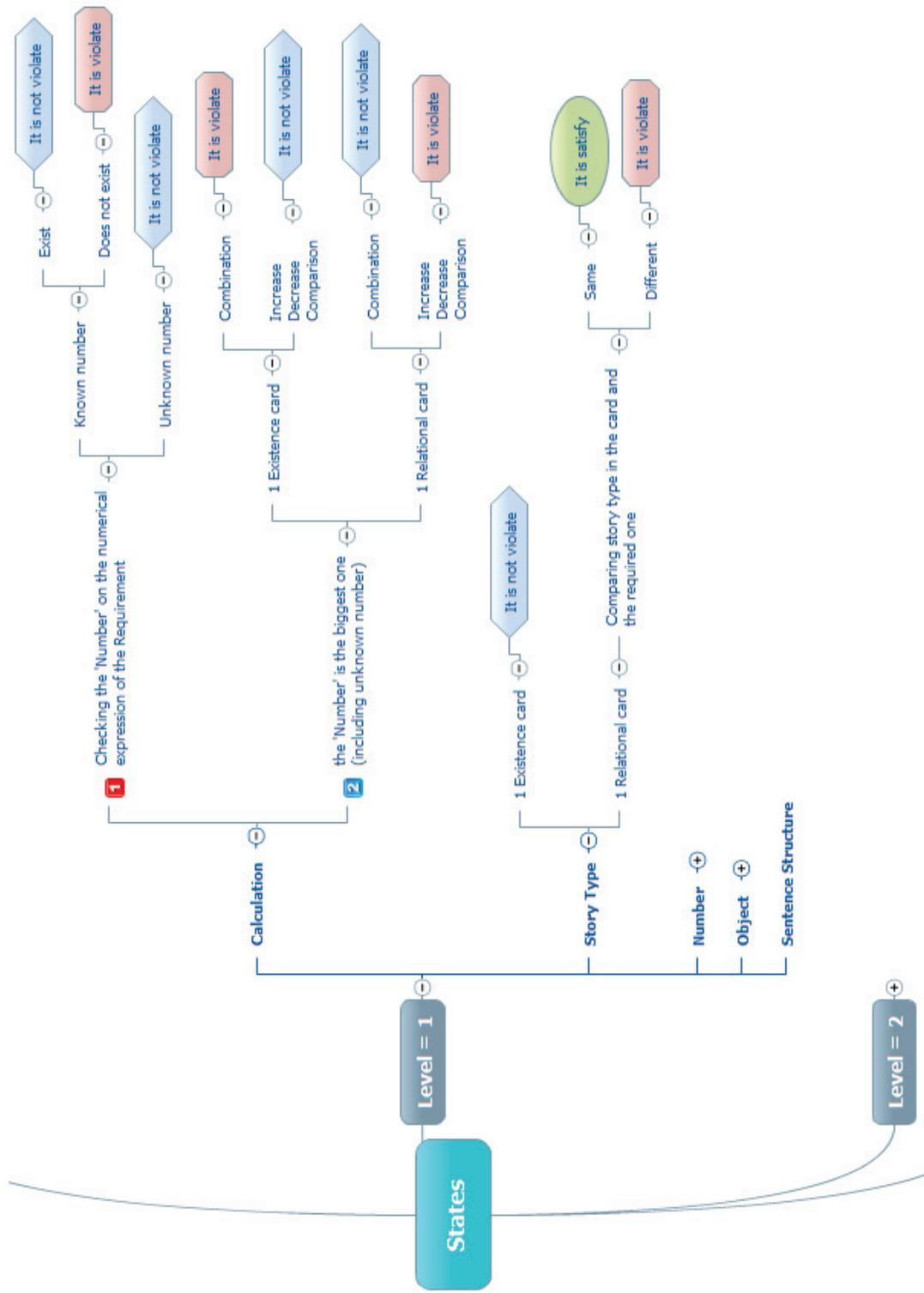


Figure A-2 Rule validity of states in Level 1 for Calculation and Story Type.

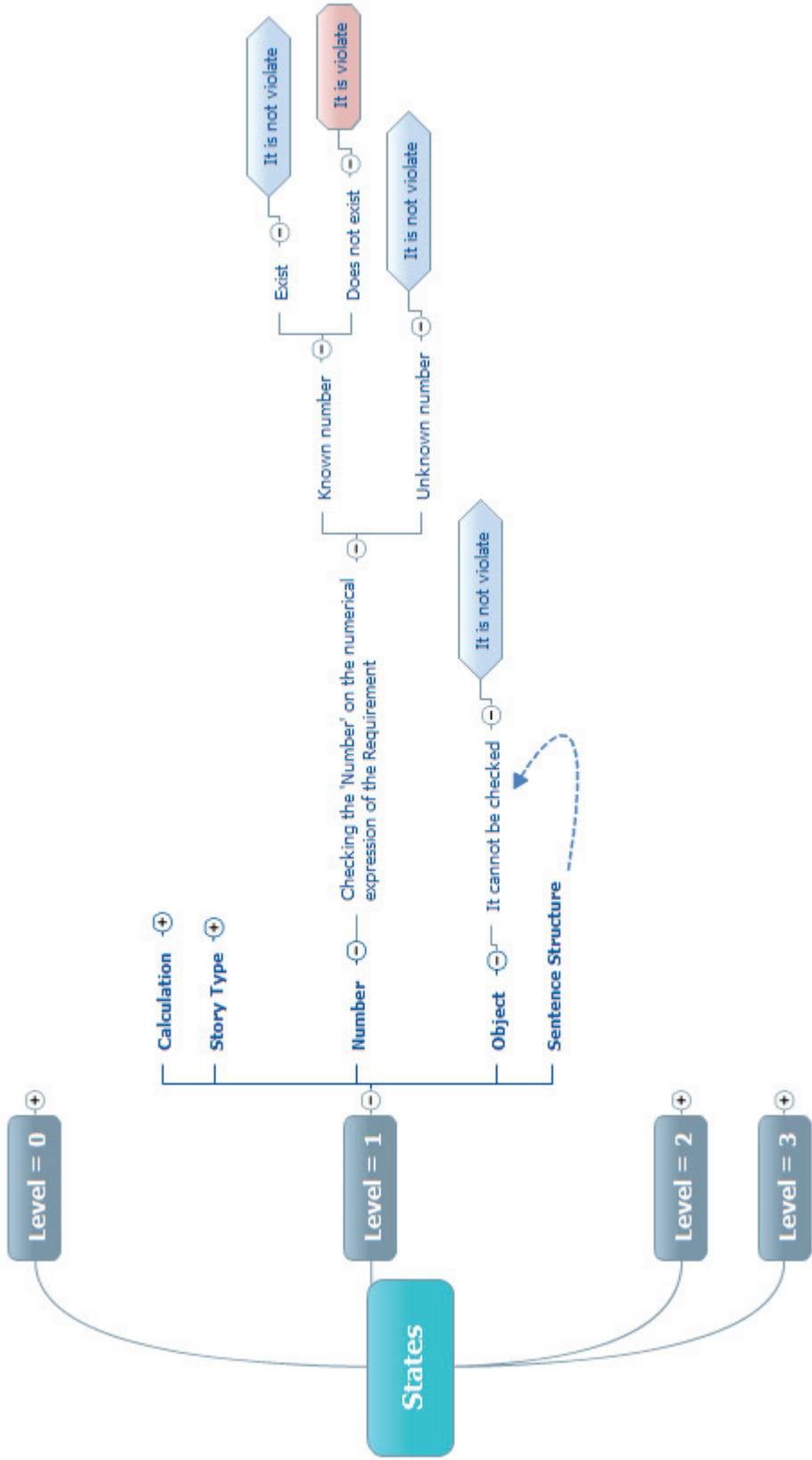


Figure A-3 Rule validity of states in Level 1 for Number, Object, and Sentence Structure.

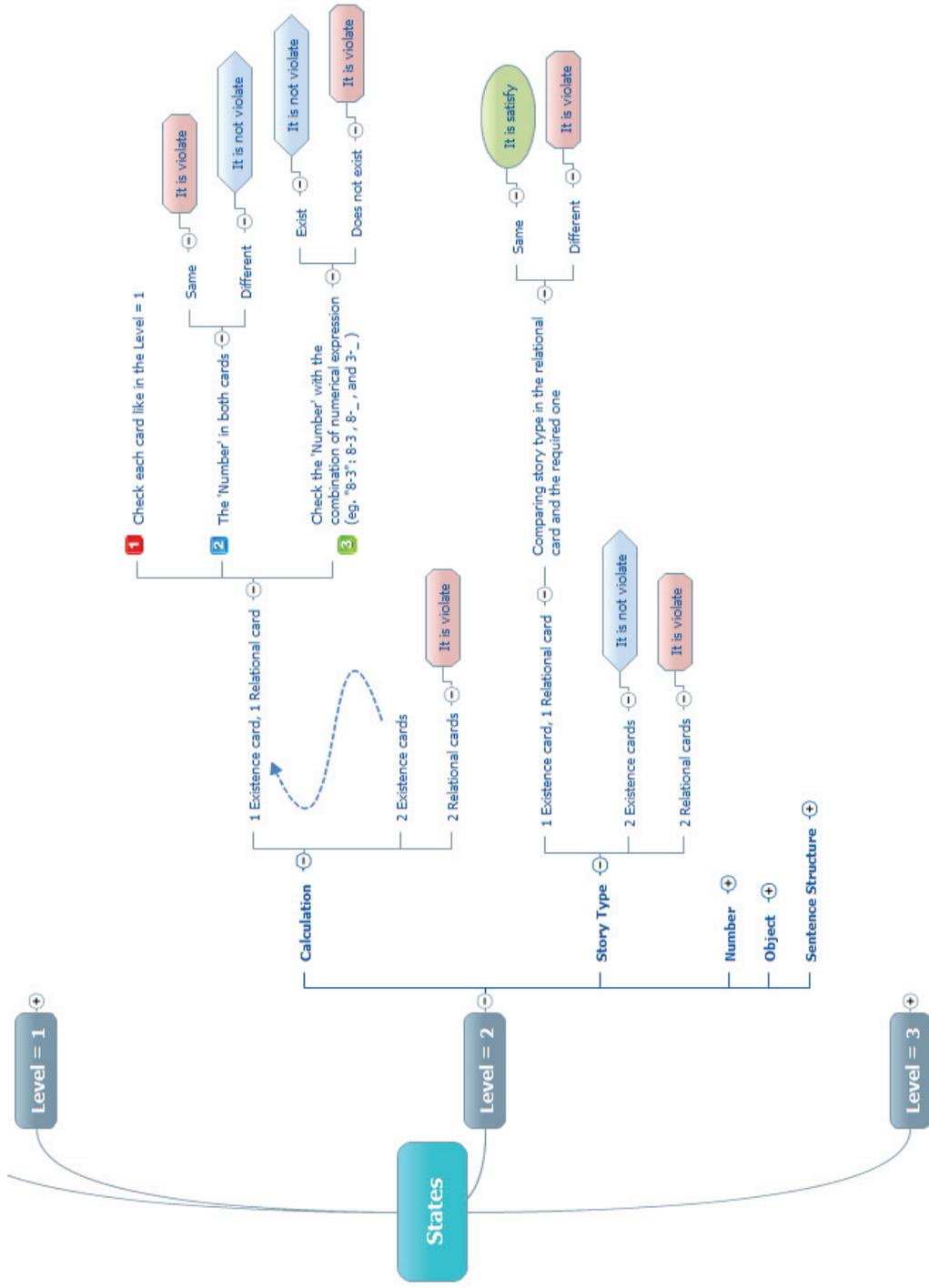


Figure A-4 Rule validity of states in Level 2 for Calculation and Story Type.

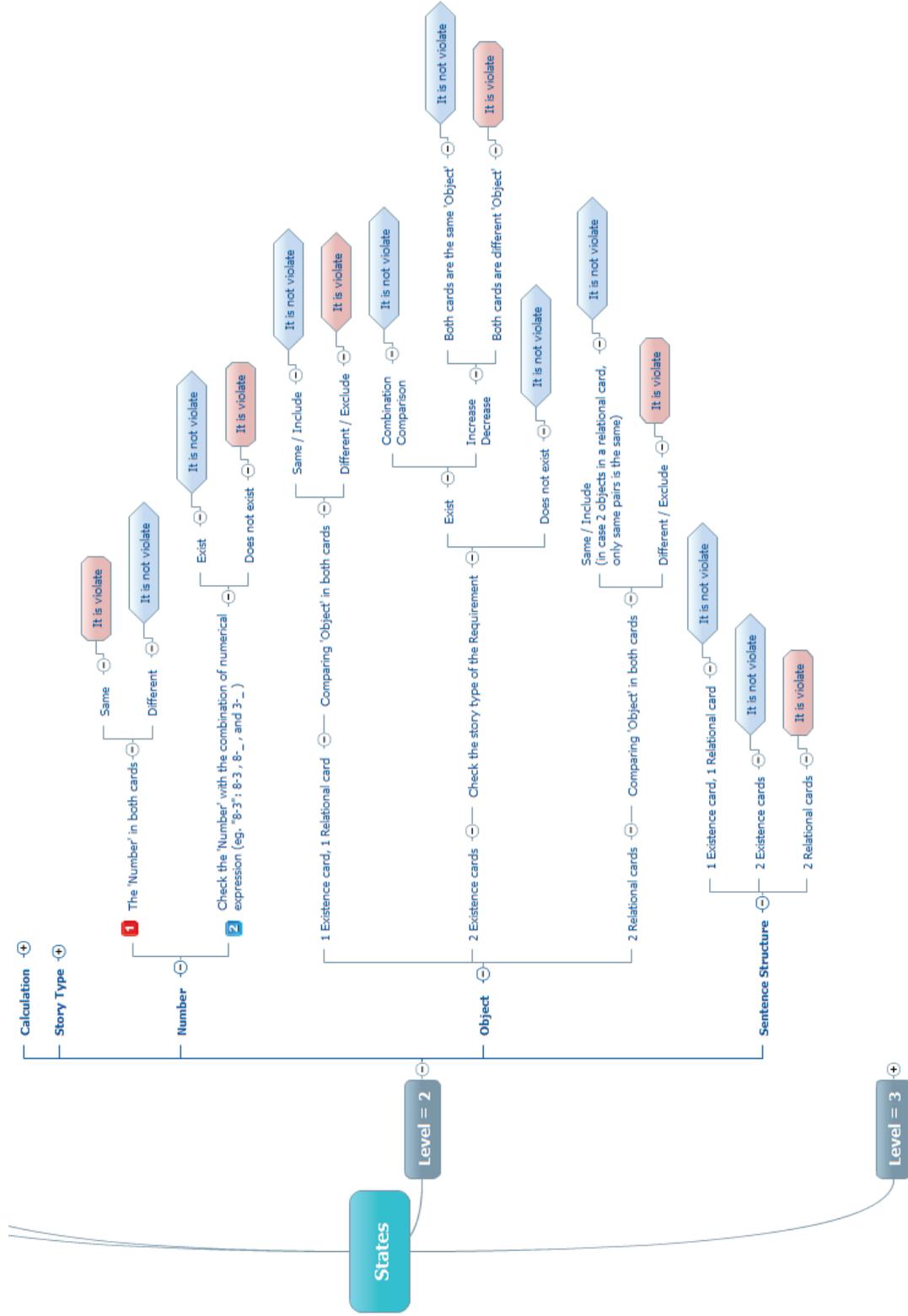


Figure A-5 Rule validity of states in Level 2 for Number, Object, and Sentence Structure.

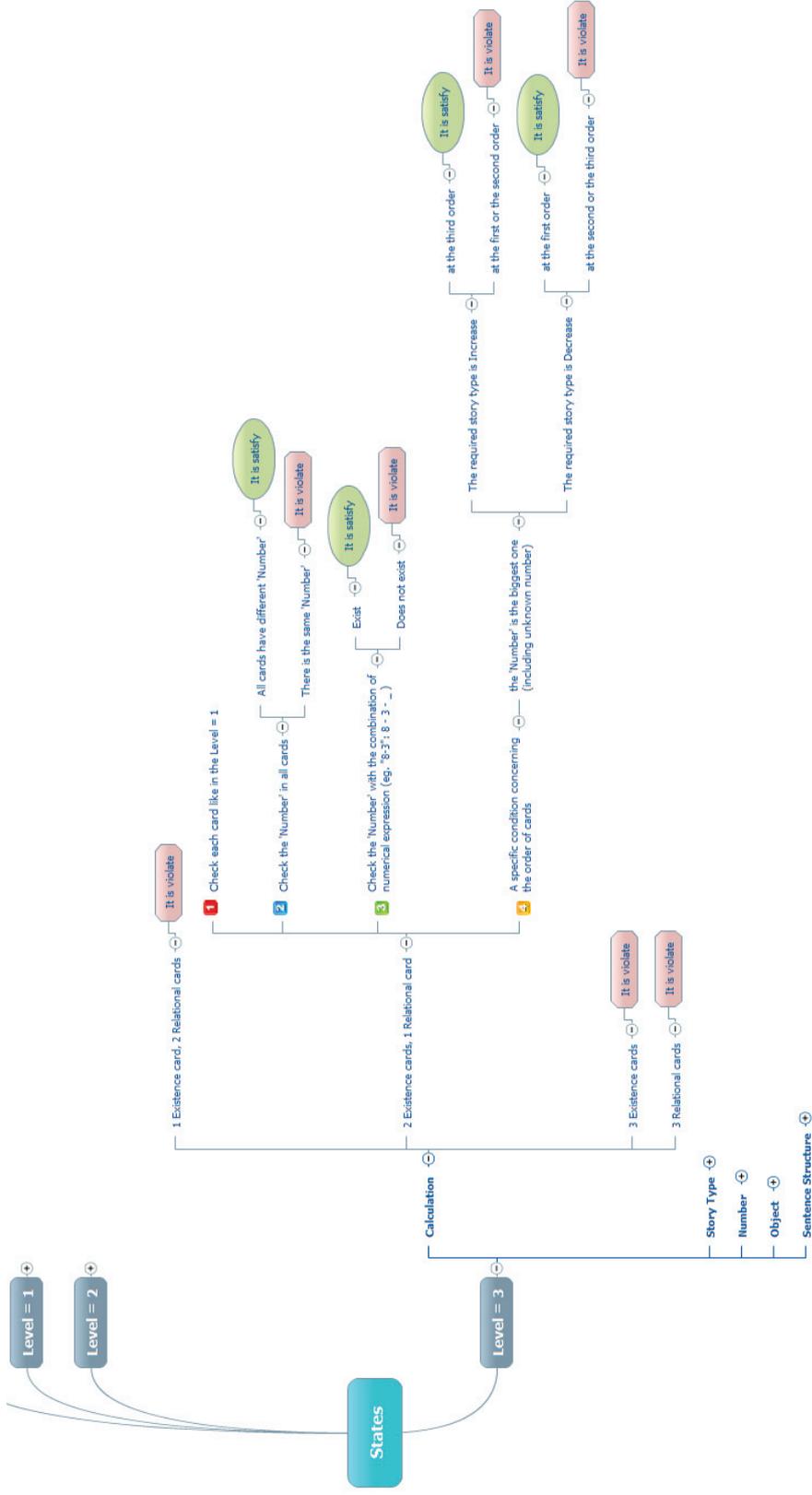


Figure A-6 Rule validity of states in Level 3 for Calculation.

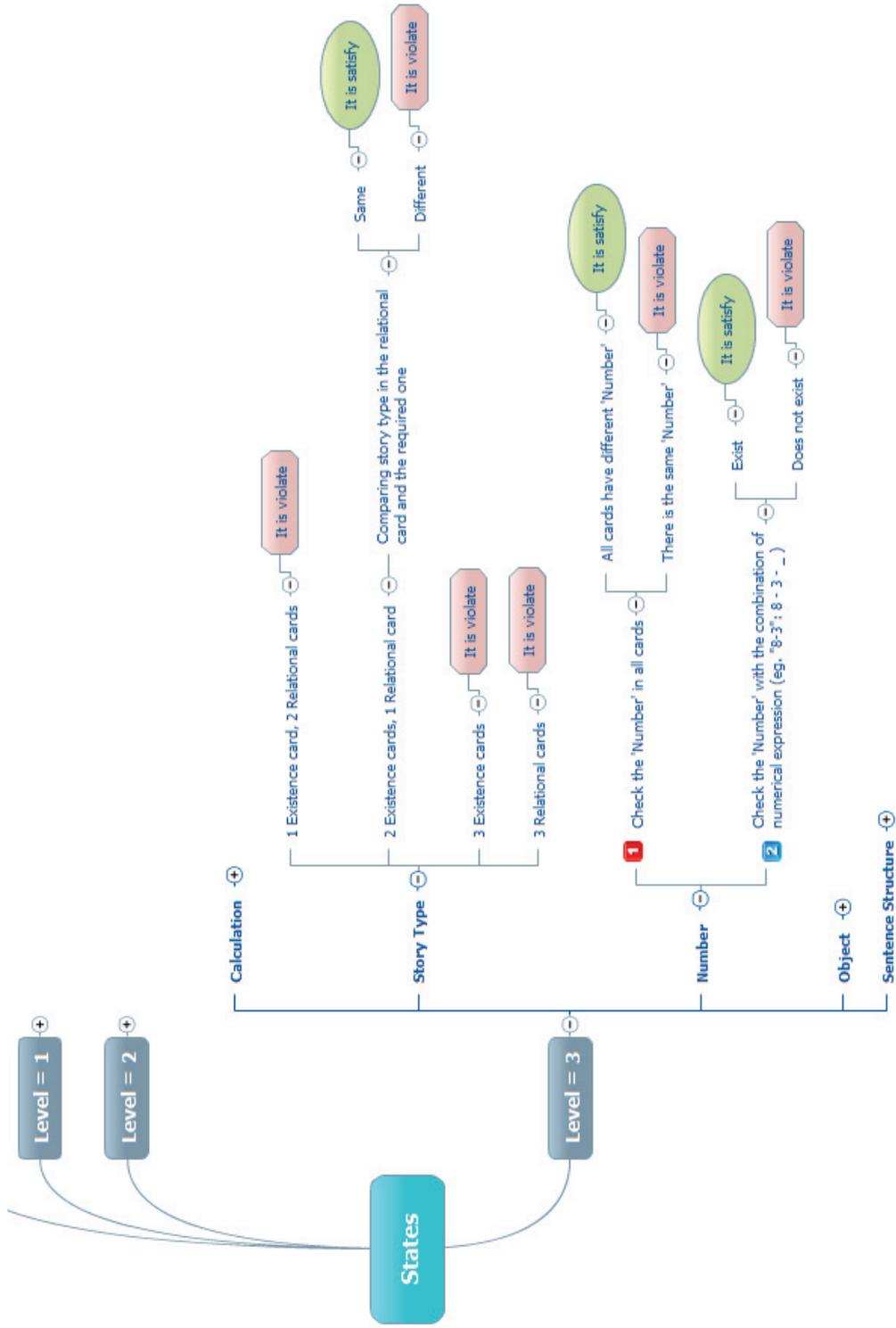


Figure A-7 Rule validity of states in Level 3 for Story Type and Number.

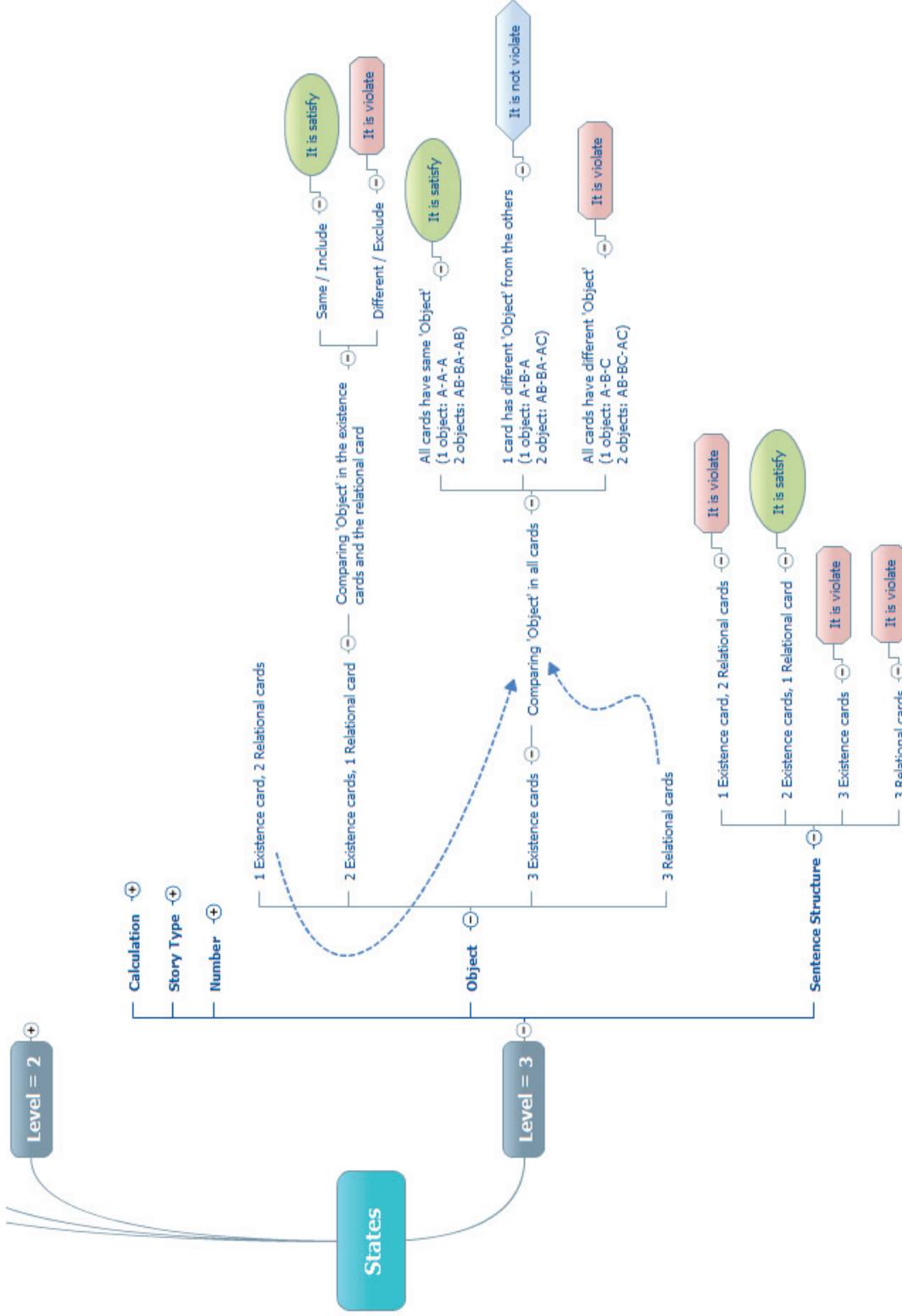


Figure A-8 Rule validity of states in Level 3 for Object and Sentence Structure.