

Using a Structured Vocabulary to Support Machine Understanding of Student Work

by

Eryn Doris Maynard

B.S., Massachusetts Institute of Technology, 2012

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

Massachusetts Institute of Technology

February 2013

©2013 Massachusetts Institute of Technology

All rights reserved.

Author:

Department of Electrical Engineering and Computer Science
February 1, 2013

Certified by:

Dr. Kimberle Koile
Research Scientist, MIT CECI
Thesis Supervisor
February 1, 2013

Accepted by:

Prof. Dennis M. Freeman
Chairman, Masters of Engineering Thesis Committee

Using a Structured Vocabulary to Support Machine Understanding of Student Work

by

Eryn Doris Maynard

Submitted to the Department of Electrical Engineering and Computer Science

February 1, 2013

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

This thesis presents a method for machine interpretation of visual representations, including those that are hand-drawn, created by students solving elementary math problems. This interpretation system extends a pen-based wireless classroom interaction system called Classroom Learning Partner. The key idea behind the interpretation is to employ a *structured vocabulary* that provides students with tools that give them enough structure to facilitate machine interpretation, but not so much that they cannot be creative in making their own representations. This structured vocabulary consists of images, shapes, tiles, and stamps and enables students to create visual representations that are constructed more easily and quickly than with freehand drawing. A machine can construct an interpretation of the visual representation by finding the relationships between the objects used in the representation, focusing on object type, location, or additional “digital ink” lines that indicate grouping of objects. The interpretation methods were evaluated on examples of student work collected in classroom trials in fourth grade classrooms in the Boston area. The results indicate that the interpretation methods will enable teachers to easily and quickly view student work in real time in a classroom and will provide teachers with information about their students’ understanding of concepts underlying the visual representations.

Thesis Supervisor: Dr. Kimberle Koile

Title: Research Scientist, MIT CECI

Acknowledgements

I would like to thank everyone who has helped me throughout my time researching and writing this thesis. First and foremost I would like to thank my advisor, Dr. Kimberle Koile, for without whom none of this would have been possible. Her contagious passion drives the entire CLP research project while motivating me and my fellow researchers. Her willingness to give up her time and discuss and investigate the issues at hand was an invaluable resource. Her patience and understanding cannot be measured and are forever appreciated.

At the start, I had little understanding for the educational intricacies required in this project. Thank you, Andee Rubin and Lily Ko, the TERC counterpart to our MIT team, for providing this much needed educational insight. Also, thanks must be given to Sherry Nichols and Renee Lockwood for allowing us to enter and test our software in your classrooms. All of the members of the educational team gave much needed advice that shaped this research into something that would actually be of assistance in the classroom and that students would use.

I am very thankful for my fellow student Steve Chapman. His dedication to this project is without words. Without his knowledge of the code base and assistance, this project could have never been completed in its current time frame. I also would like to thank former colleague Kelsey Von Tish whose ink interpretation research I incorporated and who was always willing to act as a resource.

I am grateful for Chris Terman, my academic advisor throughout the last five years, whose encouraging words and advice helped me navigate my way throughout MIT and into a master's program. I also wish like to thank Luther Richardson, my high school teacher and continuous mentor, who took a young girl with no thoughts of a future in technology and redirected me into a world I would have never thought to discover. Finally, I would like to thank my parents Suzanne and Reuben Maynard. Without their support and encouragement, I would have never dreamed that I could accomplish so much.

Contents

1	Introduction.....	6
1.1	Motivation.....	7
1.2	Overview.....	7
1.3	Outline.....	8
2	Classroom Learning Partner Overview.....	9
2.1	CLP Interactions	9
2.2	Structured Vocabulary	10
2.2.1	Stamps.....	10
2.2.2	Tiles.....	12
2.3	Interpretation Process.....	13
3	Graphical Object Representation.....	16
4	Graphical Object Grouping.....	22
4.1	Object Type.....	22
4.2	Spatial Grouping	23
4.3	Ink Grouping.....	29
4.4	Containers	36
5	Results.....	42
5.1	Erroneous Results from Drawing.....	43
5.2	Counting of Objects, Tile, and Type of Object.....	44

5.3	Distance and Containers Identified from Distance Groups	47
5.4	Containers	52
5.5	Ink Delineated Groups	53
5.6	Other Grouping Techniques.....	58
6	Conclusion	59
6.1	Future Work	59
6.1.1	Object Representation.....	59
6.1.2	Improved Ink Grouping Interpretation.....	60
6.1.3	Alternate Types of Grouping	60
6.1.4	Post-Interpretation Integration.....	64
6.2	Contributions.....	64

1 Introduction

This thesis proposes a solution to the problem of machine interpretation of hand-drawn representations by introducing a *structured vocabulary*, which provides students with tools that give them enough structure to facilitate machine interpretation, but not so much structure that they cannot be creative in making their own representations. With this structured vocabulary, students can create representations such as the one shown in Figure 1-1, which contain multiple copies of a hand-drawn element, shown in the upper right corner. They can construct explanations, solve problems, and record observations that can be interpreted. Clustering techniques then can be used to group student work based on the interpretation, thus giving a teacher valuable information about what the students know.

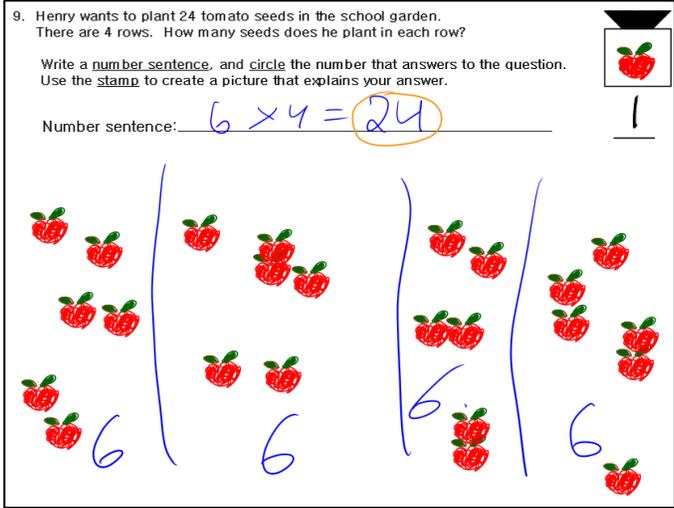


Figure 1-1: A student’s representation of four groups of six; the apple is drawn once in the upper right corner, then duplicated; a stamp object is in the upper right (discussed in Section 2.2.1)

1.1 Motivation

In upper elementary mathematics, there are many kinds of computational problems in which students may find it helpful to create multiple copies of an image, such as in Figure 1-1. Such visual representations serve as an important bridge between concrete objects and abstract mathematical concepts. Students use the representations to communicate their thinking to others and often working out a solution by manipulating elements of a representation as they would physical objects [1]. Such representations, while important, can be time-consuming for students to create: Imagine in the example in Figure 1-1 if the student had been required to draw all 24 apples. The point of the lesson was not in drawing that many apples, but in understanding how to use grouping to represent the concept of multiplication. Providing students with a means of quickly creating a representation with repeated elements enables them to focus on the math rather than just the drawing. For teachers, interpreting hand-drawn representations can be time-consuming and difficult; viewing groups of similar student representations helps teachers quickly gain insights into what representations their students have created. The methods developed in this thesis benefit students in creating representations, machines in interpreting representations, and teachers in viewing representations.

1.2 Overview

In this thesis, we developed methods that enable visual representations to be created and interpreted in our tablet-pc-based classroom interaction system called Classroom Learning Partner (CLP) [2, 3, 4, 5]. The key idea is to enable students to draw, but to structure their drawing by providing a set of graphical elements referred to as a *structured vocabulary* that can be annotated with “digital ink” and used to construct drawings. Even with this limitation on

drawing, understanding and evaluating the visual representation of student answers depends on a combination of two basic properties: *what the object “means”* and *how the object relates to the other objects* on the page. Understanding how the objects relate may depend upon *how many of each type of object* there are and *how the objects are arranged*. Understanding these properties can be difficult for both computer and human alike and is the focus of this thesis.

1.3 Outline

Chapter 2 provides background information about Classroom Learning Partner software and a basic understanding of the different software elements used in this thesis. It also explains where this thesis work fits into the overall architecture. Chapter 3 discusses how the system understands what an object represents for interpretation. Chapter 4 describes how the system creates an overall interpretation of the page using information about how different elements on the page interact with each other. Chapter 5 discusses evaluation of the methods developed in this thesis and presents results of interpreting examples of student work. Chapter 6 discusses future work and summarizes the contributions of this thesis.

2 Classroom Learning Partner Overview

CLP operates on tablet computers within a classroom, providing the same functionality as a paper workbook. These tablet-based electronic notebooks contain a series of pages created by the teacher. These pages can contain clip art and boxes for students to enter written responses, just as in paper workbooks. They also can contain interactive elements provided by CLP or multi-media elements such as photos taken using the tablet camera.

2.1 CLP Interactions

Students interact with a personal CLP notebook on a tablet via a tablet pen or their finger, depending on the activity, as seen in Figure 2-1. By strictly having this type of interaction with no keyboard usage, CLP aims to preserve and take advantage of student familiarity with paper notebooks and drawing while also enhancing the student activities [6]. The pen is especially important for the types of science, technology, engineering, and mathematics (STEM) activities that CLP targets, since many of these activities involve a mixture of text and drawing. Students may, for example, work a math problem and draw a picture to explain the answer, or collect and record data in a table and then graph the data. Upon completion of the task, they submit their work wirelessly to the teacher, who can select examples to share and discuss with the class.

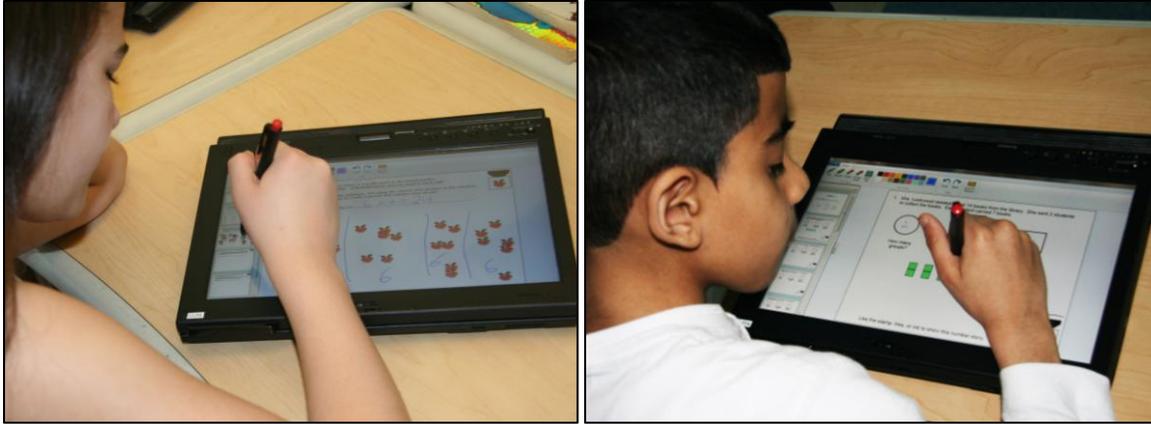


Figure 2-1: Fourth grade students using CLP to solve math exercises

2.2 Structured Vocabulary

CLP's graphical elements constitute what we call a structured vocabulary, providing structure to support students' creation of visual representations. The two types of graphical elements currently available for students are *stamps* and *tiles*. The student shown on the left in Figure 2-1 is using stamps to create a visual representation for four groups of six apples; the student on the right is using tiles to create a visual representation for seven groups of two.

2.2.1 Stamps

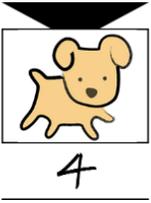
A *stamp* allows a user to quickly and easily create a visual representation containing multiple identical images. Its use is modeled after a physical stamp, with which one makes duplicate images using ink on paper. To create multiple copies of a CLP stamp, a student touches the pen to the stamp "handle", a black trapezoidal region that resembles a physical stamp handle, holds the pen down, and drags the pen to the desired location for the copy. Raising the pen then creates a copy of the stamp's image.

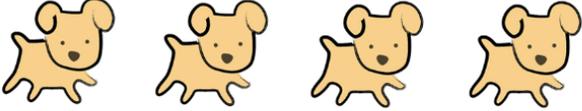
The CLP stamp has been used by elementary students during classroom trials to solve multiplication and division problems, although it is flexible enough to assist in solving other types of problems as well, e.g., fractions. Figure 2-2 shows the use of a stamp in solving a multiplication problem.

Nick, Kelvin, and Alex see 4 puppies playing in the mud.
Each puppy has 4 muddy paws.

How many muddy paws are there in all?

Use the stamp to create a picture of the puppies.
Then use the picture to help you answer the question.





$4 \times 4 = 16$

There are 16 muddy paws in all.

Figure 2-2: Visual representation made using a supplied image stamp to solve a multiplication problem

A stamp element consists of an image and what that image “means”. For our math problems, the meaning of an image corresponds to the number of things that a stamp represents. There are two types of CLP stamps which differ by the initial image: (1) a pre-drawn or image stamp and (2) a blank stamp that enables students to draw their own image. Shown in Figure 2-2 is an example of an image stamp that was provided by the teacher. The teacher also provided the number of things represented by the stamp by entering below the stamp the number “4” to represent four paws. The stamp shown earlier in the upper right corner of Figure 1-2 was originally blank; the student used the tablet pen to draw the apple shown on the stamp and then chose the desired numerical representation for the stamp in this case one. The importance of knowing the number of things represented by a stamp is discussed in Section 3.

Image stamps allow teachers to scaffold student interactions with stamps by giving students examples of how stamps can be used to create pictures, i.e., visual representations, and construct explanations. Currently only a teacher can place an image stamp. Blank stamps create an opportunity for even deeper student engagement and an opportunity for the student creativity and personal ownership seen in freehand drawing, while also facilitating machine interpretation [7, 8].

2.2.2 Tiles

A *tile* allows a user to create a visual representation of a tower of any number of tiles by stacking the tiles together, as displayed in Figure 2-3. The tiles and towers can be thought of as virtual manipulatives modeled on physical tiles or blocks [9]. Tiles can be connected together, by using the pen to drag one tile adjacent to another, to create one tower of two blocks. This interaction can then be repeated with towers to create one large tower that is the union of tiles in both towers. Towers can be easily copied to create identical towers. As students become more abstract thinkers, the tiles can be used to represent any number of things.

3. Kaelin built 7 towers of 6 blocks each. How many total blocks did she use?

a. Put x or + in the circle and numbers in boxes for the number sentence.
b. Circle the number that answers the question in the story.
c. Use tiles to create a picture that explains your answer.

$\boxed{6} \circledast \boxed{7} = \boxed{42}$

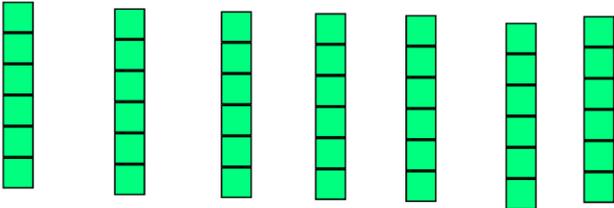


Figure 2-3: Using tiles to solve a multiplication problem

2.3 Interpretation Process

This thesis project developed a machine interpreter that analyzes all graphical elements on a page in order to gain an understanding of the visual representation created by the student. For interpretation to be a valuable resource to teachers, it must be capable of performing interpretation during classroom sessions so that the teacher can give immediate feedback to students when it will be most useful. This interpretation is performed on a student's tablet when the student submits a page to the teacher. The output of the interpretation is sent to the teacher along with the page. This combined submission has the disadvantage that more data must be sent across the network with each page, but this problem is far outweighed by the advantage of not forcing the teacher's machine to perform interpretation on every page received. Evaluation of submitted page sizes indicated that a larger page size is not a problem. [10, 11]

Several different types of interpreters exist within the CLP framework, e.g., for handwritten text, hand-drawn shapes, ink shading. It is unnecessary to run all of these interpreters everywhere on every page. It is much more efficient and makes for a more organized lesson if different regions of a page are designated for particular kinds of student input. For exercises in which students are asked to create a visual representation using stamps or tiles, a Grouping Interpretation Region is placed when a page is authored. This region is invisible to students and typically covers most of a page. All structured vocabulary elements and ink within the bounds of this region are used to construct an interpretation. Because the computational problems in elementary math that involve multiple identical images generally also involve students designating groups, e.g., four groups of six apples, this Grouping Interpretation Region

employs recursive routines that attempt to deduce groups of objects based on methods described in Chapter 4.

As an example, consider the two towers of three tiles shown in Figure 2-4. The process of interpretation starts with a tile “understanding” that it represents one tile, then the tower “understands” that it is made up of three tiles, and finally a Grouping Interpretation Region “understands” that it contains two towers of three tiles each. This modular, recursive approach enables each step to interpret only a small part of the whole problem and facilitates the writing and testing of different interpretation methods.

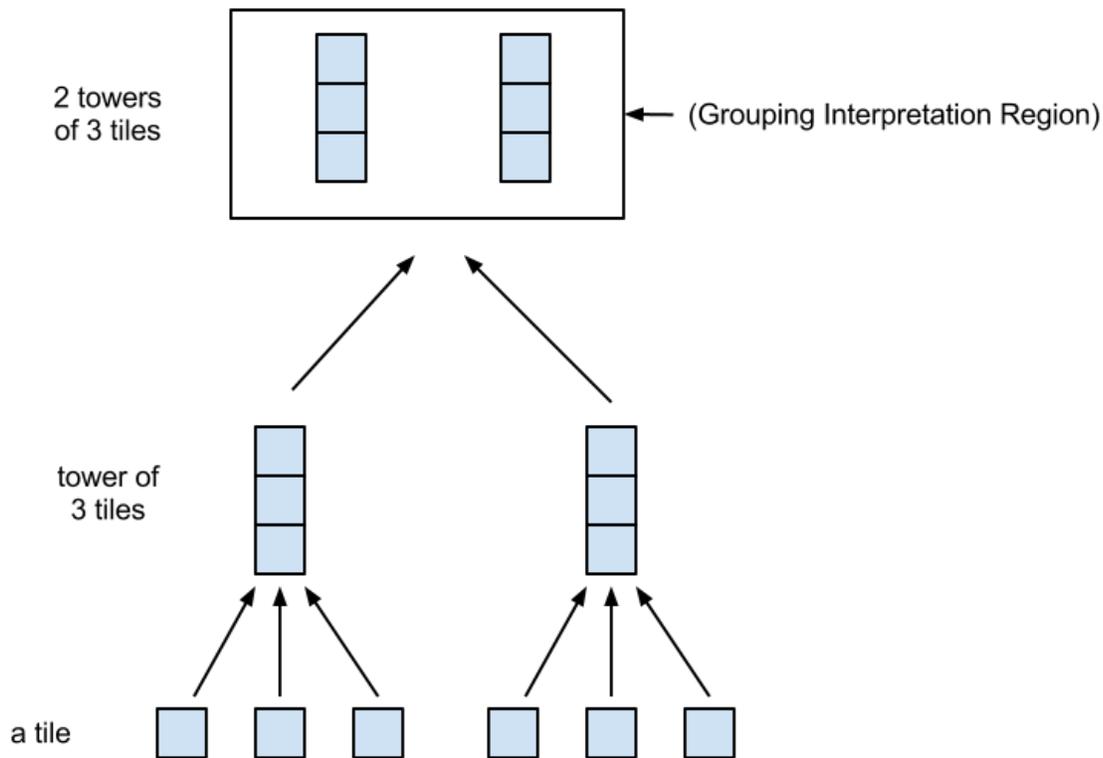


Figure 2-4: Layers of interpretation

As mentioned in Section 1.2, understanding a visual representation in a student answer depends on understanding two things: *what each object in the representation means* and *how each object relates to the other objects* on the page. It is an object’s responsibility to understand

what it means, which for our study is the number of things it represents, and the Grouping Interpretation Region's responsibility to determine how each object relates to the other objects on the page. The Grouping Interpretation Region then combines these two pieces of information to produce the interpretation for a visual representation. The topics of understanding how an object knows what number of things it represents and how the Grouping Interpretation Region determines how an object relates to other objects are discussed in detail in the next two chapters.

3 Graphical Object Representation

In order for a machine to return an interpretation that is more detailed than simply a count of the number of objects on a page, it must have information about what each object represents. With many of the mathematical exercises for which our graphical elements are useful, that information consists of the number of things represented by an element. In the case of tiles, that number is always one. In the case of a stamp, no default representation exists and the representation must be obtained from the student. For example, in the left image of Figure 3-1, the student chose to create a stamp representing a vase with six flowers and used five copies of the stamp's image in his drawing to represent five groups of six. Another student chose to answer the same problem by creating a stamp that represented a single flower and then created 30 copies of the stamp's image, as shown in Figure 3-1's right image. Both of these answers are correct, exemplifying that "correct" visual representations can take on a variety of forms, and understanding the representation requires knowledge of what each object represents.

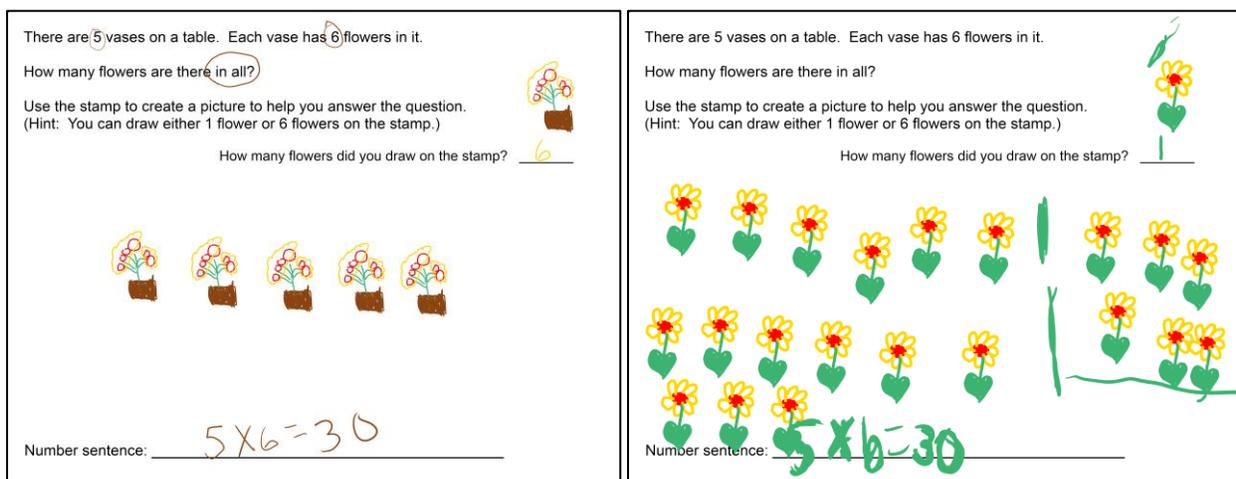


Figure 3-1: Multiple representations made using student drawn stamps to solve a multiplication problem

Asking students to designate the number of objects on a stamp allows students to demonstrate comprehension of the problem. Students may find this task difficult when they are unsure of what they are being asked to count. In the example of a puppy with four paws shown in Figure 2-2, when the numerical representation was not supplied by the teacher and students were asked how many things were on the stamp, some students answered “one puppy” rather than “four paws”. Knowing which way the student thought of the object can be helpful in identifying both understanding and misunderstanding.

Obtaining a numerical representation of a stamp from the student that was accessible by the CLP system presented a user interface (UI) challenge. When authoring the problems, teachers could limit students’ representation of the stamps by dictating what to draw on a stamp, e.g., telling the student that the stamp should represent “one flower”. Making the choice of a stamp for the student, however, would hinder learning by giving answers to parts of the problem and would prevent students from coming up with alternate equally correct representations.

In previous research, one attempt to satisfy knowing how many things a stamp represented presented students with the word “Parts” below a stamp and asked them to explicitly state the number of parts by using an on-screen keypad [12]. This method was a compromise between using keyboard input versus using a tablet pen and handwriting recognition routines, which were not highly reliable. This attempt used the UI seen in Figure 3-2 with a blank button appearing next to the word “Parts”. When the blank button was tapped with the tablet pen, a keypad with the numbers 1 through 12 appeared. The student could then select any of these 12 numbers as the numerical representation for the stamp, and that number would appear inside the button, as shown in Figure 3-2. The term “Parts” following the button was intended to act as a

cue to the students about what they were to fill in. Unfortunately, many students found the term “parts” to be confusing since this terminology did not appear in their curriculum.



Figure 3-2: UI using a button that popped up to a dialog followed by the word "Parts"

To avoid the terminology difficulty faced by the previous UI, the version developed in this thesis removed the term “parts” from the UI for the stamp. Instead, the teacher asks students to write the number of things being counted on their stamp. This technique allows the teacher to use whichever terminology her students are most comfortable. If a teacher did not include the instructions or if a student forgot to enter the number representation, a popup saying “What are you counting on the stamp? Please write the number on the line below the stamp.” appears when the student tries to create a copy of the stamp. The UI shows the student the machine’s interpretation of what number was written and provides a way for students to correct the number if the machine’s number is not what was intended: students can either erase their handwritten number and rewrite it or use a number pad to enter their number, as described below.

In keeping with the CLP goal of supporting interactions familiar from using paper notebooks, the button in the previous UI was replaced with a blank line appearing under the stamp, as shown in Figure 3-3, creating an appearance similar to what one might expect on a paper page. Behind this space lies a handwriting recognition region based on Microsoft's ink recognition software [13]. This region automatically detects every time a student writes anything within the region and updates the stamp's numerical representation accordingly. To avoid incorrect interpretation as a result of faults in handwriting recognition, a small version of the number interpreted from the student's writing appears in the top left corner of the region, showing the student what the machine thinks she wrote. If the handwriting interpreter could not determine a number, a question mark appears, as shown in Figure 3-3A. If the number is wrong or unable to be recognized, the student can either erase the number and write a new one, or tap the small printed number or question mark causing a number pad to appear that enables a student to tap on the correct number, as shown in Figure 3-3B. Once the student selects a number, the ink previously in the region is replaced with the new number, as shown in Figure 3-3C. This UI allows students to write the numbers as they would on a piece of page while ensuring that the interpreter accurately understands the numerical representation. If a student wants to change this new number, she simply needs to begin writing within the region, and the current number will be replaced with her handwritten one.

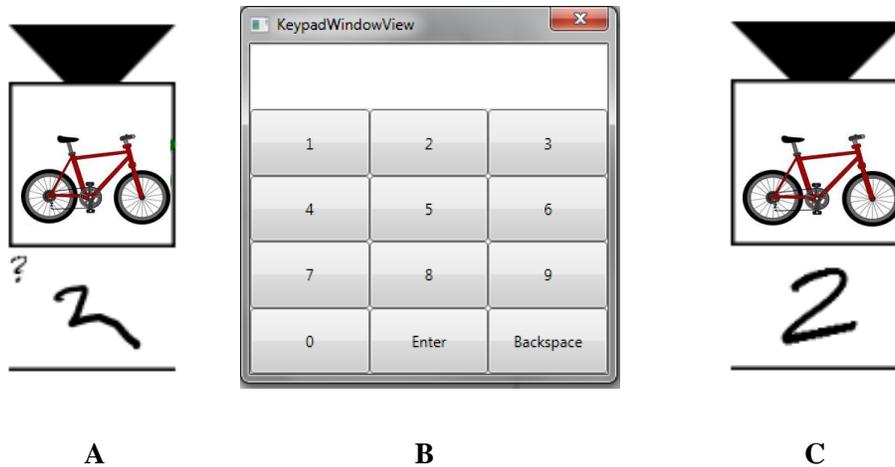


Figure 3-3: From left to right: A) A stamp for which the handwriting interpreter cannot recognize the handwritten number, B) Keypad that pops up when a student taps on the “?” appearing from the inability to recognize the written number, C) A stamp for which the number representation was entered by a student’s tapping on the keypad’s “2”.

While the interpreter requires student assistance to determine what an image or inked stamp represents, it does not require assistance when students add graphical elements directly to a stamp. In the example shown below in Figure 3-4, a student has drawn a stamp representing one orange, written the number “1” below the stamp, then added six copies of the orange image to the box stamp. The number of objects represented by the box stamp is incremented automatically by the orange’s numerical representation each time an orange image is overlaid on the box stamp. If one of the overlaid images is removed, the number below the box stamp is decremented by one, which is the orange’s numerical representation. Once the student is satisfied with the appearance and numerical representation of the box stamp, she then can make copies of the stamp in the same way that she does with other stamps.

Kelvin was helping put oranges in boxes. He had 30 oranges. Each box holds 6 oranges.
 How many **boxes** will he need?
 Draw an orange on the blank stamp, then put 6 oranges on the box stamp. Then use the box stamp to create a picture to help you answer the question.

Number sentence: $30 \div 6 = 5$

He will need 5 boxes.

Figure 3-4: Student entered “1” on the orange stamp that she drew; the “6” automatically appeared below the box stamp after she overlaid six copies of the orange stamp on top of the box stamp

A subtle issue about the example shown in Figure 3-4 concerns the initial representation of the box stamp. When an image stamp is added to a page by an author, the author has the option of setting a default value for the numerical representation for the stamp. In this example, the author set the value to zero. If a representation of zero is chosen, the number is not displayed to the students to avoid confusing them concerning what a stamp with a numerical representation of zero means. Zero was chosen to represent the box, rather than one, because the author expected that students would either drag other images onto the box stamp and the numerical representation would be automatically updated and displayed, as done in the example in Figure 3-4, or that students would write in the number themselves, as discussed in Section 5.2. The issue of whether the numerical representation should be zero or one in problems where one object is expected to “hold” other objects is one best decided by educational personnel since the choice depends on the context of the problem.

4 Graphical Object Grouping

In addition to understanding what individual objects on a page represent, the interpreter must also understand how the individual objects relate to each other to determine the meaning of the entire picture. Focusing on multiplication and division problems, the question of how objects relate to each other turns into a question of how the objects are grouped on the page. Students can designate groups in a variety of ways. This thesis identified and implemented four grouping techniques found in student examples: grouping objects by type, using space to create separation between groups, using ink to create separation between groups, and using a container object to “hold” other objects. When the Grouping Interpretation Region processes a page, each of these techniques is attempted in order to identify all possible groupings. Knowing which groupings are correct is difficult without knowing the type of problem or the correct answer, since certain types of problems lend themselves to particular types of groupings [12]. The interpreter developed in this thesis defers the question of which type of grouping is correct to another CLP module currently being implemented by another member of the CLP research team.

4.1 Object Type

Grouping by object type entails understanding how many times an object is duplicated on a page. For every object this means something slightly different, but at its most basic level, it means knowing how many identical looking objects appear on the page. For stamps this means knowing how many times a single stamp has been “stamped” onto the page, creating a copy of its image. In the example shown in Figure 4-1, two stamps exist on the page, the left stamp that depicts a cat is used to create four copies and the right stamp that depicts a spider is used to

create three copies. We can see from the student's scratch work while solving the problem that this type of grouping is what she intended.

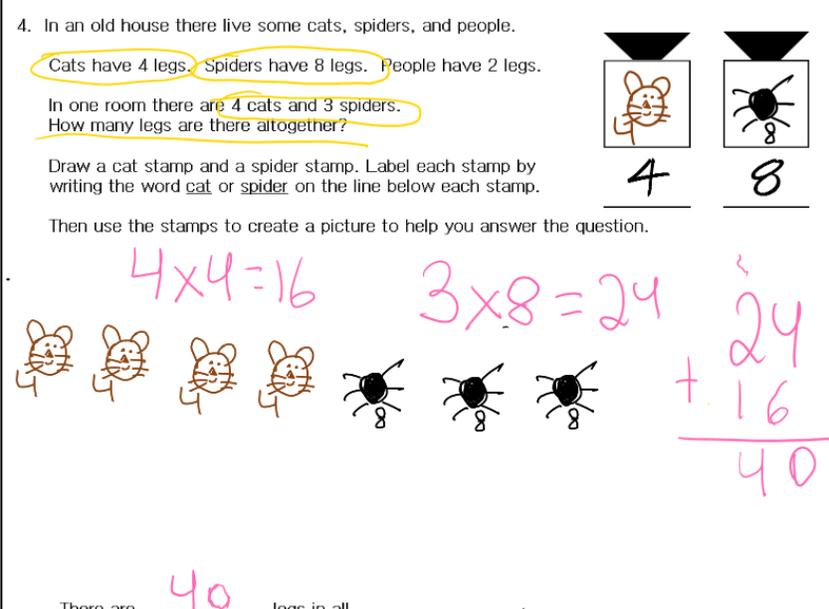
4. In an old house there live some cats, spiders, and people.

Cats have 4 legs. Spiders have 8 legs. People have 2 legs.

In one room there are 4 cats and 3 spiders.
How many legs are there altogether?

Draw a cat stamp and a spider stamp. Label each stamp by writing the word cat or spider on the line below each stamp.

Then use the stamps to create a picture to help you answer the question.



There are 40 legs in all.

Figure 4-1: Student using two stamps to solve a multiplication problem

While generally a trivial identity problem, all objects have at least one case that could now or in future versions cause incorrect values for the interpreter. Students can alter the size of all graphical objects except the tile. Though we did not see the behavior in our classroom trials, a student could alter the size of an object with the intention of creating a new type of object, and the difference in size would not be recognized as creating a different object. Currently, incorrect values could occur when a student edits a stamp after using it and believes that it now represents a different stamp because it displays a different image. Another instance of this problem will occur when the color of tiles and shapes can be changed by the student.

4.2 Spatial Grouping

Spatial grouping relies on the understanding of space as a separating factor between objects, as shown in Figure 4-2. This technique can be difficult for both machines and humans if

objects are not tightly packed together within a group or groups are too close to each other. The technical challenge comes from no guaranteed basic spatial structure: every page may use a different amount of space between groups. Even groups on the same page may use space differently.

To understand groups defined by the space between them, a hierarchical clustering algorithm with a Euclidean distance metric algorithm was implemented [14, 15, 16]. This technique uses an agglomerative approach, where every object represents its own group originally. The algorithm finds the two groups with the smallest distance between them and joins them to form one larger group. This process is then repeated until the groups are formed.

10. Elyse has 32 fish. If she has 4 fish tanks, how many fish are in each tank?

Write a number sentence, and circle the number that answers to the question.
Use the stamp or tiles to create a picture that explains your answer.

Number sentence: $32 \div 4 = 8$ fish

Figure 4-2: Spatial grouping representation of division problem

To find the smallest distance between two groups, the algorithm uses a Euclidean distance metric with single linkage. The Euclidean distance calculated is the absolute smallest line that can be formed between the two objects' bounds. Single linkage, also known as minimum linkage, looks at the Euclidean distance between every object on the page and every

other object on the page that is not in its group. The choice of what two groups to combine comes from the minimum distance found between all of these objects. Thus, a single object from each group determines if the two groups are joined, regardless of where the groups' other members are located.

In a traditional hierarchical clustering algorithm, groups are continuously joined together until either only one large group remains or a predetermined number of groups remain. One large group, however, would provide no useful information about the visual representation, and there is no way to know beforehand how many groups should remain without asking the student, defeating the purpose of running this algorithm. Therefore, a cutoff had to be found to determine when to discontinue joining groups [17]. This cutoff was determined by first averaging the distances of each object in the group to its nearest neighbor in the group (the distance that the groups had been combined with).

$$\text{Cutoff} = \text{Avg}(\text{Minimum Distances Connecting Each Object to the Group}) \times \text{Threshold}$$

This procedure creates a unique metric measuring how closely connected a group is. This metric is then multiplied by a fixed value to create the cutoff. If the distance between the two groups being joined is smaller than the cutoff calculated by each group, the groups are combined.

$$\begin{aligned} & \text{Join Groups if } \text{Min}(\text{Cutoff of Group A}, \text{Cutoff of Group B}) \\ & \geq \text{Distance Connecting Groups A and B} \end{aligned}$$

On inspection of student work, this metric could become skewed by a student placing objects on top of each other, producing a minimum distance of zero, requiring the distance of any new object joining the group to also be zero and making the threshold value useless. In addition, some of the occurrences where the interpreter believed the objects to be overlapping would not appear so to a human observer. For example in Figure 4-2, within each group, it appeared to the

interpreter that a fish was at least overlapping with one, if not more than one, other fish within its group even though no ink appears to overlap. Originally, a stamped image's bounds were the bounds of the entire space inside the stamp where a student could draw, including any white space. Figure 4-3A shows the original bounds of the object from Figure 4-2. A human interpreter or the student placing the stamp does not think to include this additional white space in their logic but only considers the area where ink or an object exists. Accordingly, the bounds of the object were adjusted for the interpreter to remove all unnecessary white space, as seen in Figure 4-3B. This step greatly minimizes the cases in which the interpreter believes objects are overlapping.

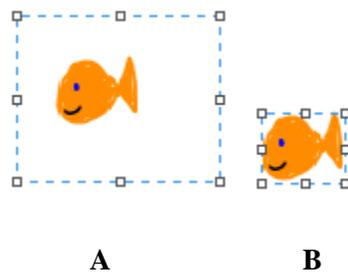


Figure 4-3: From left to right, A.) Original bounds of the object from Figure 4-2 which includes all white space surrounding the ink. B.) Adjusted bounds of the object eliminating any unnecessary white space

Two solutions were proposed to handle the remaining occurrences where a zero minimum distance prevented any new objects from being joined to a group even if it appeared to a human interpreter that they should be in the same group. The first was to change the distance metric to rely on the center of the object diminishing the number of occurrences two objects would be considered exactly on top of each other. This solution had the negative effect that objects with varying widths and heights might not be seen as close to each other even if to the human eye they were on top of each other. The second solution, which was the one implemented, requires a minimum distance to be multiplied by the threshold. This minimum distance chosen is approximately the size of a tile, the smallest manageable object used thus far within CLP.

Adjusted Cutoff

$$= \text{Max}(\text{Avg}(\text{Minimum Distances Connecting Each Object to the Group}), \text{Minimum Value}) \times \text{Threshold}$$

Let us illustrate this algorithm with an example. As shown below in Figure 4-4, there are five objects: three green squares and two blue circles. These five objects create two groups: one group of the three green squares and one group of the two blue circles, where color and shape have been used for to clarify the groups for the reader. For this example, we will use a threshold of 2.5 and a minimum distance of 2.



Figure 4-4: Original graphical representation on page

As the interpreter begins analyzing the visual representation, each of the objects constitutes its own group. The interpreter calculates the distance between every object and finds that the smallest distance is 1 between objects A and B as seen in Figure 4-5 with the red line.

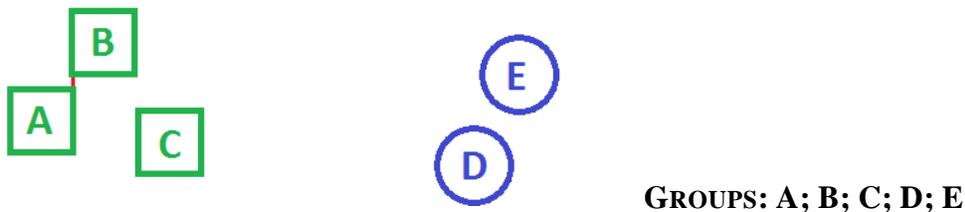


Figure 4-5: First iteration of spatial grouping algorithm

Finding that neither group has an average distance, the minimum distance 2 is used. Since $2 \times 2.5 \geq 1$, the two groups are joined to create one combined group containing A and B as designated by the grey outline enclosing A and B in Figure 4-6.



GROUPS: A & B; C; D; E

Figure 4-6: Second iteration of spatial grouping algorithm

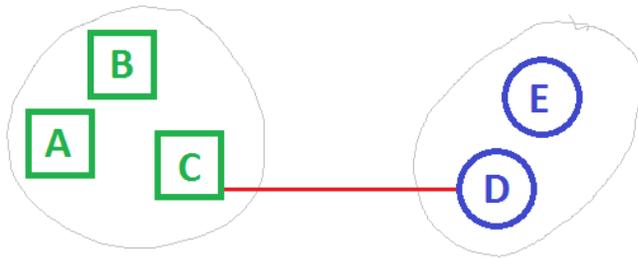
Now the process repeats itself, this time finding the smallest distance is 3 between D and E. Since yet again neither of these groups has an average distance, the minimum distance 2 is used, $2 \times 2.5 \geq 3$, and the two groups are joined. The process begins again and returns a minimum distance of 5 between B and C, shown in Figure 4-7. C has no average distance so it uses the minimum of 2, which meets the maximum bound to allow the two groups to merge. The group consisting of A and B has an average distance of 1, which is below 2 causing the minimum value of 2 is used and the groupings are combined.



GROUPS: A & B; C; D & E

Figure 4-7: Third iteration of spatial grouping algorithm

At this point, we have created the two groups that we expected from visual inspection, but we must still reiterate the process for a fourth time since we have yet to hit the cutoff point. This iteration, depicted in Figure 4-8, returns a minimum distance of 35, exceeding the grouping ‘A, B, C’ cutoff of 3×2.5 and stopping the algorithm.



GROUPS: A, B, & C; D & E

Figure 4-8: Fourth and final iteration of spatial grouping algorithm

4.3 Ink Grouping

Ink as a grouping mechanism requires the student to use ink to delineate the groupings. Groupings are typically marked by drawing a shape, such as a circle or square, around all of the grouped objects, as shown below in Figure 4-9, or by drawing vertical or horizontal lines between the groups as depicted in Figure 4-10. This thesis focused on these two commonly used cases.

10. Elyse has 32 fish. If she has 4 fish tanks, how many fish are in each tank?

Write a number sentence, and circle the number that answers to the question.
Use the stamp or tiles to create a picture that explains your answer.

Number sentence: $32 \div 4 = 8$

Figure 4-9: A division problem assisted by using both square and circle outlining around the groups.

2. Kaelin has 15 candies. She divides them among 3 bags. How many candies are in each bag?

$\boxed{15} \div \boxed{3} = \boxed{5}$
candies bags candies

a. Put numbers in the boxes for the number sentence.
 b. Put a word or words on the line below each number to say what the number means.
 c. Use red ink to circle the number that answers the question in the story.

d. Then use the stamp, tiles, or ink to show this number story.

Figure 4-10: Division problem using vertical lines to separate the groups

Other means of delineating groups, such as those seen in Figure 4-11, were noticed in our classroom trials. These cases present technical challenges, including when an outlining may be made up of multiple separate stroke objects, when one stroke may count as a member of two separate outlinings, and when an edge of the page may count as part of an outlining. A previous investigation explored using a new technique of checking every stamped image with every other stamped image and then searching for a stroke in the space between the two images. [12] Unfortunately, stray marks or scratch work by students easily renders these grouping results incorrect.

1. Kaelin has 3 bags, and she puts 5 candies in each bag. How many total candies are in the bags?

$$\boxed{5} \times \boxed{3} = \boxed{15}$$

Candies bags Candies

a. Put numbers in the boxes for the number sentence.
 b. Put a word or words on the line below each number to say what the number means.
 c. Use red ink to circle the number that answers the question in the story.

d. Then use the stamp, tiles, or ink to show this number story.

Figure 4-11: Student use of ink that this interpreter will not handle

In order to avoid the difficulties of understanding the delineations and the incorrect interpretations from our previous methods, this version opted to use a Shape Interpretation Region, which incorporates Microsoft’s ink interpretation library to analyze the ink [13]. The Shape Interpretation Region interpretation routine is run over the area of the Grouping Interpretation Region, described in Section 2.3, and returns all identified shapes including vertical and horizontal lines. The Grouping Interpretation Region can then look at the shape type and the bounds of the ink returned from the Shape Interpretation Region’s routines without ever needing to understand the intricacies of the ink.

The Grouping Interpretation Region creates an internal representation of the page and the relationships of the discovered shapes by building a tree constructed from the shape bounds. The root of the tree is the bounds of the entire region. The algorithm iterates through all of the discovered shapes adding them to the tree. If shape A’s bounds can be contained in another shape B’s bounds than shape A is a child of shape B. If the shape is a line, two rectangles are

created on each side of the line using the line and the edge of the region as two opposite sides of the rectangle. The remaining two sides of the rectangle connect to the two endpoints of the line and the respective edge of the region. Once the tree has been completed, the algorithm iterates through the page objects and places them within the smallest node of the tree that can contain the page object's bounds. The resulting groups are each node's child page objects.

To illustrate this algorithm, we will walk through an example using the objects displayed in Figure 4-12. The outer black rectangle represents the Grouping Interpretation Region's bounds. Within the region are two vertical lines and a square. The pages objects are not displayed because this algorithm only involves the ink shapes' relations to each other. Deciding the groupings of the page objects is a trivial final step that identifies objects' locations with respect to the shapes.

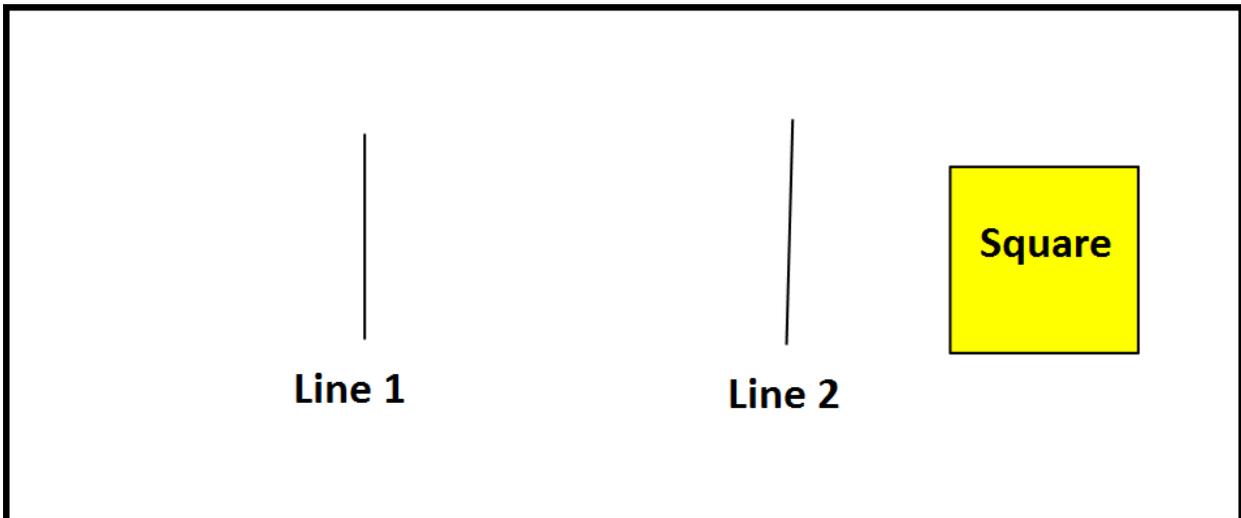


Figure 4-12: Original picture with two vertical lines and a square delineating the groups. No objects that are actually being grouped are shown.

The algorithm begins by creating its internal tree representation with the bounds of the entire region as the tree's root, as depicted in Figure 4-13.

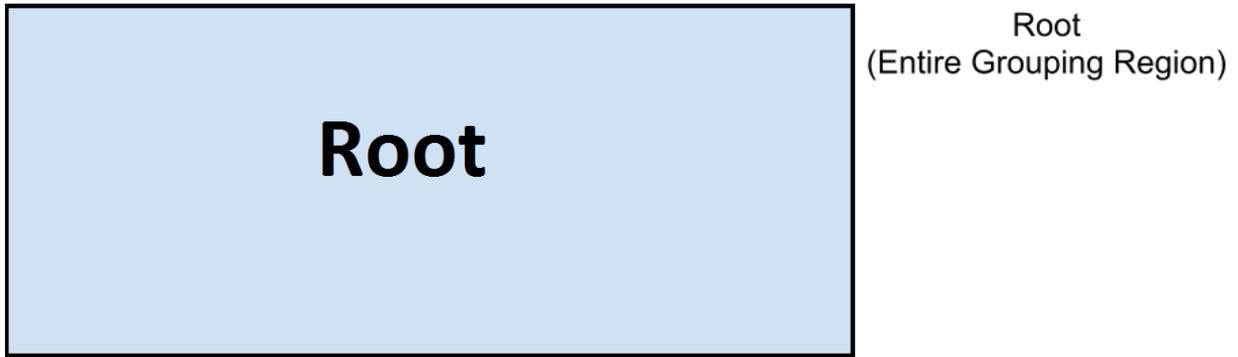


Figure 4-13: Original representation recognized by interpreter with the root as the entire grouping region

The first shape added to the representation is the vertical line referenced as Line 1 in Figure 4-12. Two rectangles are added as children of the root node. Rectangle A uses Line 1 as its right side and the left edge of the grouping region as its own left side. The height of the rectangle is derived from the height of Line 1, and the width of the rectangle is whatever distance is needed to connect the chosen left and right side. Rectangle B repeats this process except using Line 1 as its left side and the right side of the region as its right side.

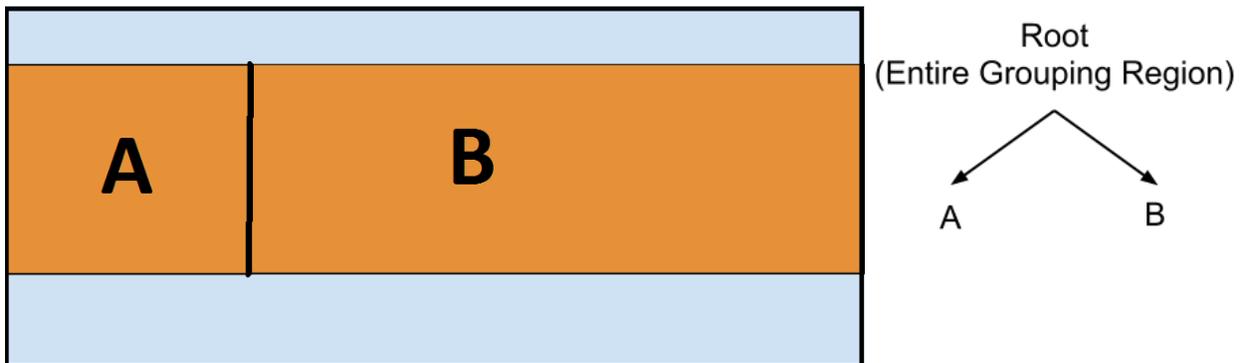


Figure 4-14: Line 1 creates Rectangles A and B. These rectangles are added as children to the root node.

In the next iteration, the interpreter attempts to add Line 2's resulting rectangles; however, when the rectangle C, resulting from the left side of Line 2, checks if it is a child or parent of A, it finds that it intersects A, as shown in Figure 4-15. If the intersection occurs because a bound specified by a shape intersects the edge of the grouping default bound, the node with the default

bound adjusts its bounds to use the intersecting shape as a bound. Since C's intersection with A results from the presence of Line 1, C's bounds are adjusted to use Line 1 as a left side and Line 2 as a right side, eliminating the intersection.

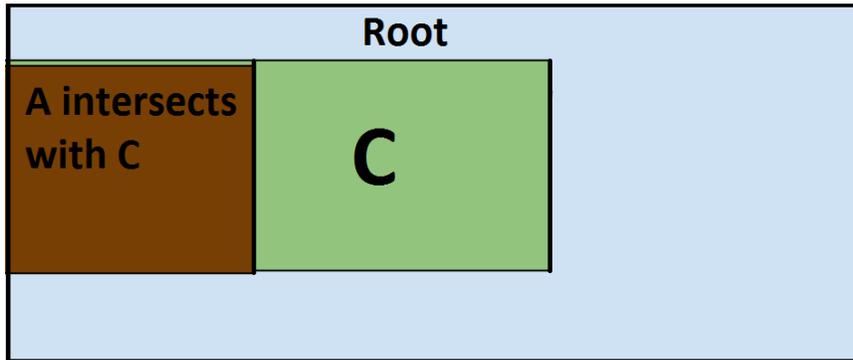


Figure 4-15: The rectangle C formed to the left of Line 2 intersects with A because C uses the default left side of the grouping region edge. Rectangle B is not shown but still exists

When C checks to see if it is a child or parent of B, it discovers another intersection as depicted in Figure 4-16. This time B's default right side causes the intersection with C's right side of Line 2, and B's bounds are adjusted.

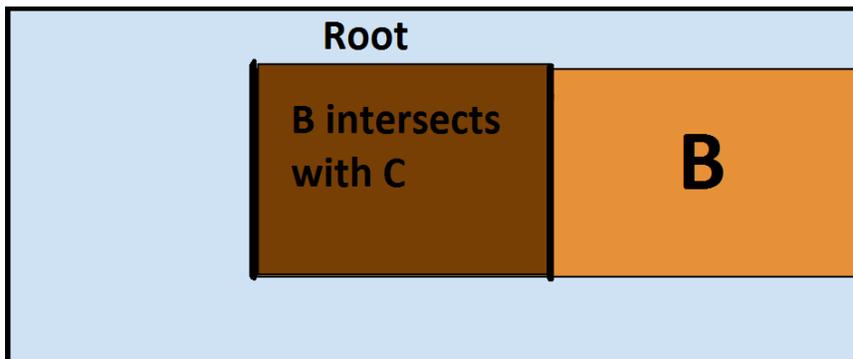


Figure 4-16: The rectangle C formed to the left of Line 2 intersects with A as a result of B's right side using the grouping region right edge. A is not depicted in this visualization but still exists.

At this point, B and C discover that they have the same bounds and are combined, creating one node BC shown in Figure 4-17. Instead of finding an intersection, if Line 2 had been shorter than Line 1, C would have appeared to be a child of B. Before adding itself as a child of B, C would check to see if it shares all of B's explicit bounds (not the default grouping region edge). Seeing

that this is the case and they both share the explicit bound of using Line 1 as a left bound, if C has any additional explicit bounds declared, the two nodes are combined. Since C also has the explicit bound of Line 2 as a right side, the interpreter realizes that the appearance of a child was from uneven lines and B and C combine to form BC.

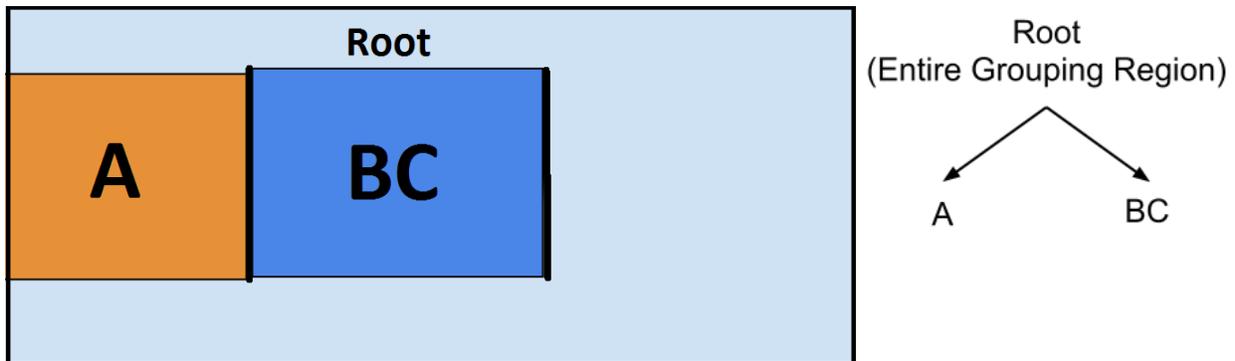


Figure 4-17: Representation of the page as seen by the interpreter following the addition of rectangle C

The rectangle D formed to the right of Line 2 is now added to the tree. Since it is not a child or parent nor does it intersect any other node within the tree, it is added as a child of the root node.

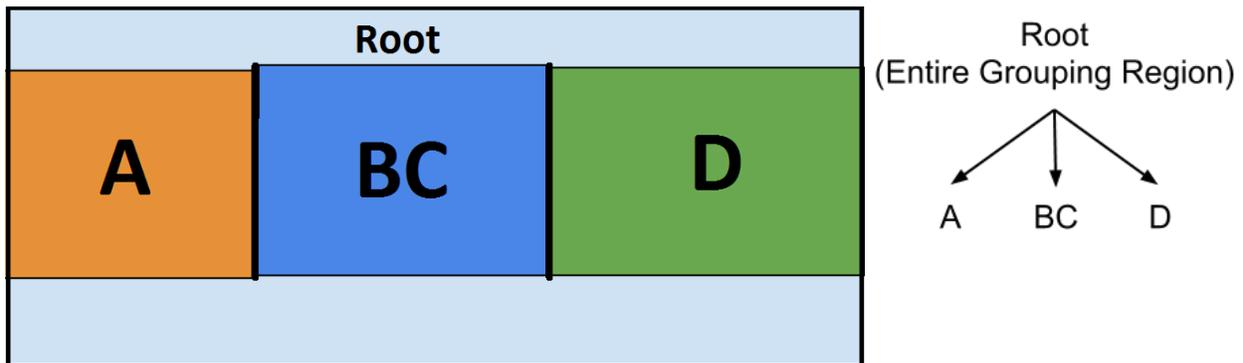


Figure 4-18: The interpreter's vision of the page following the addition of Line 2's right rectangle D

The final object to be added to the interpreter's representation of the page is the square E. E detects that its bounds are contained within D. When checking to see if the D and E should be combined, E discovers D's left side is Line 2, but its own left side is the left side of the square. The two nodes are thus not combined, and E is added as a child of D as viewed in Figure 4-19.

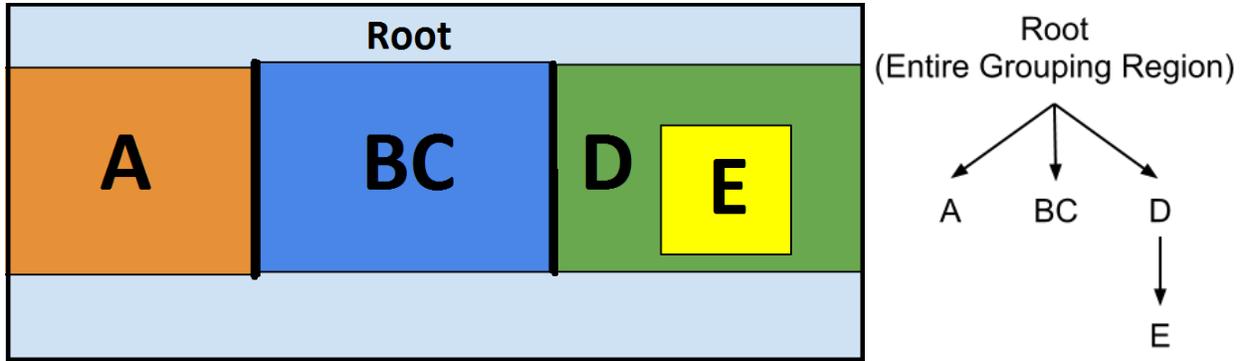


Figure 4-19: Final interpreter representation of the page

Now that the page’s representation has been created for the interpreter, it will iterate through all of the page objects, placing each object in the lowest layer of the tree possible that can contain the page object’s entire bounds. The final groupings are the page objects that each node contains at the end of this iteration.

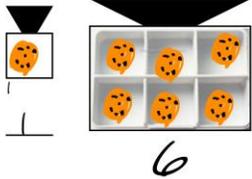
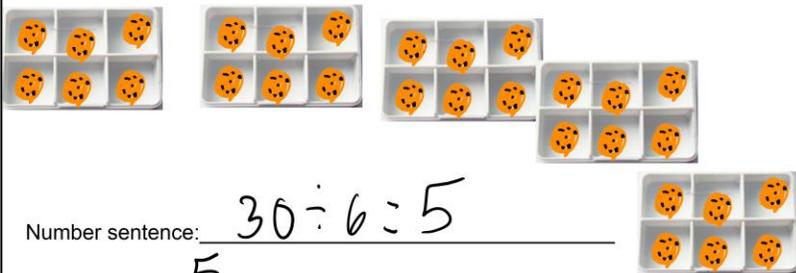
4.4 Containers

A container uses one object to “hold” other objects. Containers in CLP can either be explicit or implicit. An explicit container “captures” the objects that it holds, allowing it and the objects it is holding to be treated as one collective object that can be copied and grouped. One such example was shown in Figure 3-4 and is repeated here.

Kelvin was helping put oranges in boxes. He had 30 oranges. Each box holds 6 oranges.

How many **boxes** will he need?

Draw an orange on the blank stamp, then put 6 oranges on the box stamp. Then use the box stamp to create a picture to help you answer the question.

Number sentence: $30 \div 6 = 5$

He will need 5 boxes.

Figure 4-20: The box stamp acts as a container for the orange stamp's images

Explicit containers are typically used for grouping in *quotative* division problems, which take on the form, “You have 35 flowers, and 5 flowers make a bouquet. How many bouquets can you make?” The student is told in the problem what the groupings are; the question is how many groups exist. In this sort of problem the student can create one explicit container group and then create as many copies as he needs in order to solve the problem and create the correct number of groups. Two types of explicit containers exist in CLP: the tower and the stamp. The tower is a container of single tiles as seen in Figure 2-3 and described in Section 2.3. A stamp acts as a container whenever other graphical elements overlay it, as shown in Figure 4-20.

An implicit container occurs when the object that is supposed to be the container is placed onto the page and all other objects that it is supposed to hold are placed near it in some type of grouping. In Figure 4-9, the basket stamp is an implicit container that contains four beach balls. *Partitive* division problems commonly use this type of container. These problems generally take the form, “You have 35 flowers and 7 people. How many flowers does each person

receive?” In a partitive problem, the student knows how many groups are needed but does not know how many objects will belong in each group. He can put an implicit container on the page for each group and then deal out the objects “into” the container until the appropriate number of objects has been placed on the page.

After discussion with the educational researchers on our project, we determined that students should not define that an object is a container because the concept of a “container” is unfamiliar to elementary students. Consequently, the interpreter must deduce that a container exists and identify which object is the container. Identification of an implicit container is performed following the completion of the other grouping techniques by iterating through every grouping interpretation and checking to see if these three rules are upheld for any object:

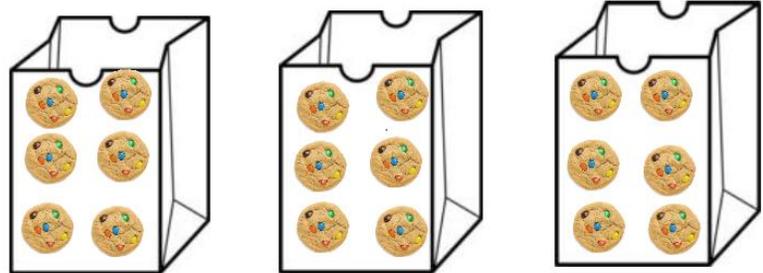
- 1.) Object appears in every group
- 2.) Only one instance of the object can be present in a group
- 3.) Object must represent one thing

If all of the rules are upheld then the object is identified as a container.

Kaelin baked 18 cookies. She wants to put them in 3 different bags.

How many cookies will be in each bag?

Use the stamp to create a picture of the cookies in the bags to help you answer the question.

Number sentence: 18 ÷ 3 = 6

Figure 4-21: Implicit container where the bag is the container and it is “holding” cookies

While analyzing the results of our evaluation, we observed that the stamped images of the container, if placed sufficiently close together, might be interpreted as a single group by the distance algorithm described in Section 4.2. When a human views the student work shown in Figure 4-22, he or she understands that the basket is supposed to act as a container and anything on top of this basket is within a group formed by that single container. The machine, however, views the balls and basket stamped images in the same way initially and must deduce that the baskets should be viewed as their own separate groups.

Henry and Nick were bringing 16 balls out to the playground for recess. They wanted to put them in baskets. Each basket holds 4 balls.

How many baskets did they need?

Use the stamps to create a picture to help you answer the question.

Number sentence: $16 \div 4 = 4$

Figure 4-22: Division problem where the containers are placed very close to each other

The distance algorithm employed by the interpreter believes all of the objects are in one giant grouping because of how close the baskets are placed together, as can be seen in Figure 4-23. When a human looks at the same picture without the balls, as in Figure 4-23, the human will also observe the baskets form one group. This comparison suggests that the difference between a human and a machine is a perceptual one: a human understands that the basket is being used as a container, which allows the basket images to be placed closer together while still maintaining distinct groups, one group per basket.

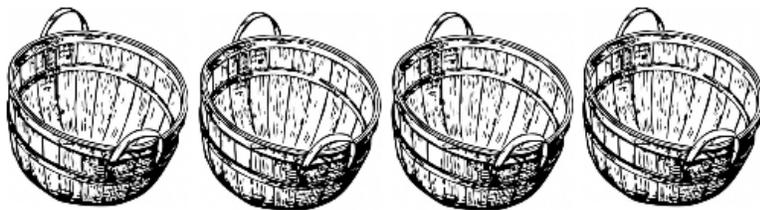


Figure 4-23: The containers from Figure 4-22 without the ball that they hold

Two methods were proposed to handle this situation. One suggestion was to test every object that represents one thing, against every other object and check for a certain percentage of

overlap. This test will produce a group for each object that contains all of the objects that it overlaps. These groups would then need to be filtered, keeping only the groups with a viable container. The advantage of this method is that the percentage of overlap between two objects needed to combine them into a single group can be adjusted, preventing incorrect results from objects that barely overlap. This method has been saved for future work.

The second method utilizes the existing distance algorithm with the modification that the metric determining the combinations of groups does not use a minimum distance. This change ensures that if any objects overlap, then their groups can only contain objects that overlap with some member of the group. This second method was implemented with only minimal changes to the code.

5 Results

This interpretation system was evaluated by testing its effectiveness on students' responses to math problems and comparing the system's interpretations to a human's interpretations. Two fourth grade classes in Waltham, MA, over the course of twelve days throughout the 2011-2012 school year solved 58 distinct math problems creating a total of 293 pages, on which we evaluated the performance of our interpreter. Chart 5-1 below displays the different methods that students utilized to solve the problems. The interpretation results are very promising: Of the 293 student pages, the interpreter supported 81% of the grouping techniques used by the students. Out of these samples the interpretation was correct 84% of the time with the intended student technique and 86% of the time using any technique. These results demonstrate the value of our structured vocabulary in balancing freehand drawing with structured drawing.

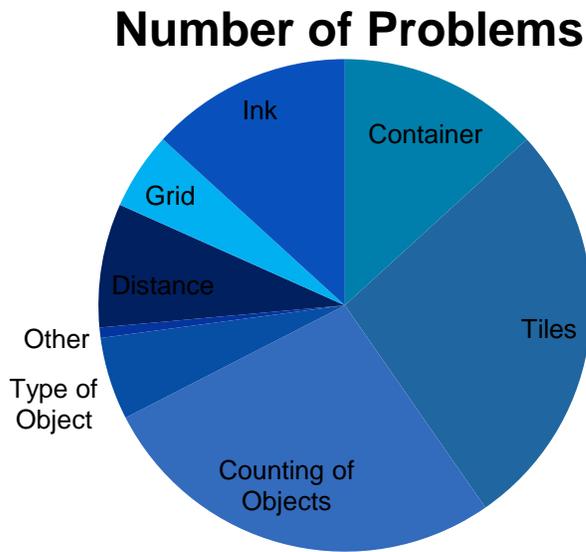


Chart 5-1: Breakdown of the different methods students used in their graphical representations

5.1 Erroneous Results from Drawing

Every type of grouping technique except for responses using only tiles returned at least one incorrect interpretation resulting from students drawing objects onto the page.¹ We decided that interpreting drawn objects that are not inside the bounds of a stamp do not produce enough information about the ink's representation, and all ink that is not inside a stamp or recognized as a shape is ignored by the interpreter. As a result, the grouping techniques are incapable of providing the same interpretation as a human in all cases because the interpreter does not “see” some of the objects present. In Figure 5-1 on the left, the student uses distance to create three groups of six cookies contained within a bag. The student, however, has drawn eight of the eighteen objects, causing the interpreter to only see ten stamped objects. As a result, it returns an interpretation of two groups of three within a bag and one group of four within a bag. In the right image of Figure 5-1, the student has placed six donut boxes but has drawn the donuts on top of the donut box images. The donut box image has a numerical representation of zero, as discussed in Chapter 3, and since the grouping region has no knowledge of any of the donuts, the interpreter returns a value of six groups of zero.

¹ In many of these cases, students accidentally deleted their stamp, and in the version of CLP used at the time, the option of creating a new stamp was not available to students, so they had to finish their representation by drawing. We have since added functionality that enables students to add new stamps to their pages.

Kaelin baked 18 cookies. She wants to put them in 3 different bags.

How many cookies will be in each bag?

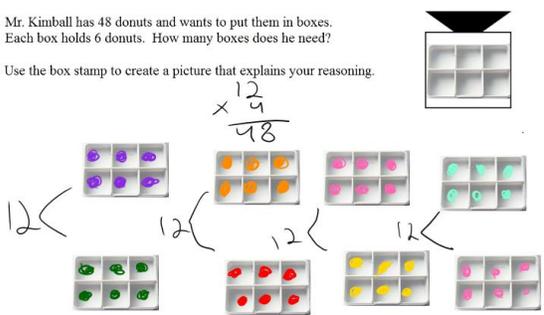
Use the stamp to create a picture of the cookies in the bags to help you answer the question.



Number sentence: $6+6+6=18$, $18 \div 3 = 6$

4. Mr. Kimball has 48 donuts and wants to put them in boxes. Each box holds 6 donuts. How many boxes does he need?

a. Use the box stamp to create a picture that explains your reasoning.



b. Write a number sentence and circle the number that is the answer to the question.

$6 \times 8 = 48$

Figure 5-1 Examples where the student has drawn on top of the intended container

5.2 Counting of Objects, Tile, and Type of Object

The counting of objects, the use of tiles, and the differentiation by type of object can appear to be three separate methods of explanation but are all understood by the interpretation mechanism in the same way. As Section 4.1 describes, the interpreter returns how many of each type of object exist on the page along with how many things the object represents. The interpreter always returns the correct result as long as the student has used elements from the structured vocabulary and not drawn freehand on the page, and if a stamp is being used, the student has entered the correct representation for the number of things on the stamp. This technique does not require the interpreter to make any assumptions about the way the objects are grouped.

Student work interpreted by counting objects, using tiles, and differentiating by the type of object comprised over sixty percent of the problems examined. These three techniques are the easiest for the interpreter to understand, and they are also flexible enough to adapt to students' understanding at many levels. Students who are just learning multiplication may break down the problem by counting while adding an additional object to the picture, as demonstrated in Figure 5-2, or turn the multiplication problem into a more familiar addition problem, as seen in the

examples in Figure 5-3. These students used the structured vocabulary to discover the correct answer. Other students may already know the answer and only use the picture to explain their answer, such as those students whose work is shown in Figure 5-4.²

3. Michael used 30 blocks total to build 5 towers. How many blocks are in each tower? 30

$5 \times 6 = 30$
 How many groups? towers
 How many objects in a group? blocks
 How many objects total? blocks

5 10 15 20 25 30

5. Mrs. Nichols is making 8 flower bouquets for a friend's party. Each bouquet has 6 flowers in it. How many flowers does Mrs. Nichols need?

a. Create a picture that explains your reasoning.

b. Write a number sentence and circle the number that is the answer to the question

$6 \times 8 = 48$ flowers

Figure 5-2: Student examples explicitly counting the objects to find the answer

4. Mr. Kimball has 48 donuts and wants to put them in boxes. Each box holds 6 donuts. How many boxes does he need?

a. Use the box stamp to create a picture that explains your reasoning.

$6 + 6 + 6 + 6 + 6 + 6 + 6 + 6$

$8 \times 6 = 48$

b. Write a number sentence and circle the number that is the answer to the question.

$8 \times 6 = 48$

There are 8 people in our class. Ms. Lockwood wants to give each person 3 star stickers.

How many stickers does she give out?

Then use the blank stamp to create a picture to help you answer the question.

$3 + 3 + 3 + 3 + 3 + 3 + 3 + 3$

$3 \times 8 = 24$

Ms. Lockwood gives out 24 stickers in all.

Figure 5-3: Examples of students using stamps to break multiplication into an addition problem

² Other work in our research group has investigated the order in which students create their pictures and write their text, so that we can identify cases in which students write their answers first. [11]

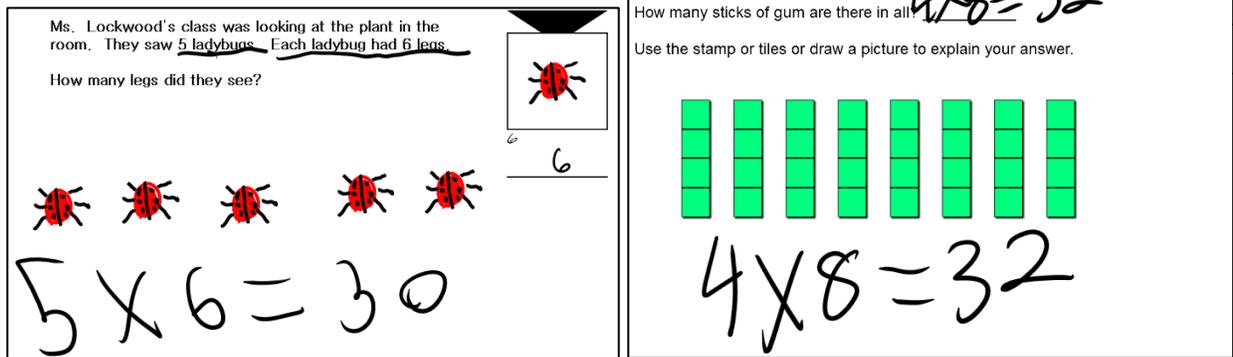


Figure 5-4: Students using tiles or stamps to depict the mathematical sentence

In each of the above examples, all objects in the visual representations were identical, i.e., students only used one type of object. The same interpretation methods, however, can be used when all objects are not identical, i.e., when multiple types of objects appear in a representation, as shown in the problems in Figure 5-5. In the left image of Figure 5-5, the student uses two different objects to create two separate groups and then performs addition on the multiplication result represented by the groups. The interpreter similarly outputs that there are four objects that represent four things and three objects that represent three things. It is up to another CLP component currently under development to decide how the numbers relate to the math problem and to know to combine the two groups as the student has to return a final value of forty. These examples illustrate the flexibility and power of this relatively simple interpretation technique.

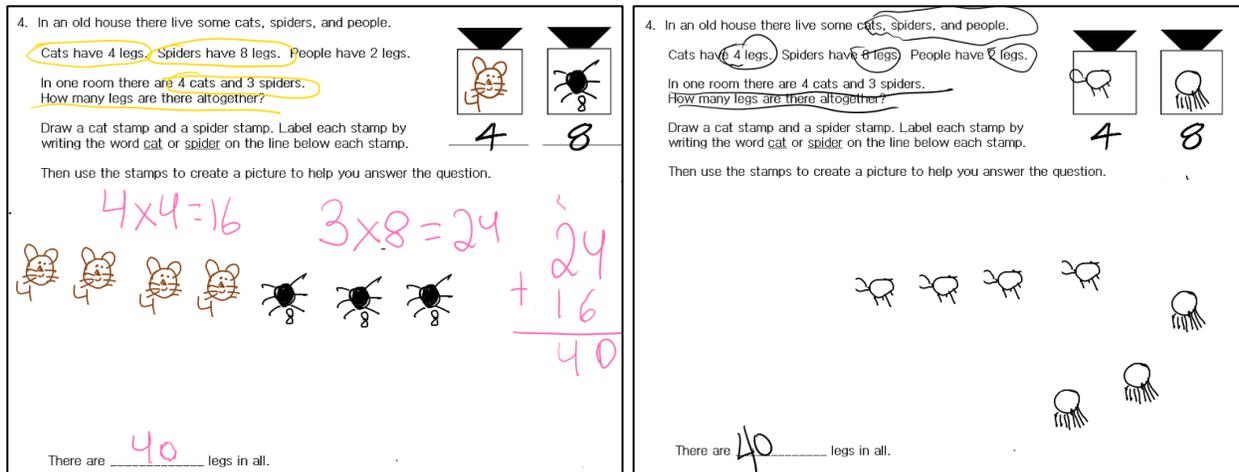


Figure 5-5: Left.) Student thinks of the objects by their object type and then adds the two groups together. Right.) Different object types are treated as one group while depicting the math problem

5.3 Distance and Containers Identified from Distance Groups

In this section we discuss responses that used distance to identify groups with and without containers. Forty-two problems used distance as a group separation tactic with the interpreter analyzing over fifty percent of the problems correctly, as outlined in Table 1. A few of the examples analyzed correctly are shown in Figure 5-6. These examples generally exhibited a relatively large distance between groups and a close and consistent distance between objects within each group. The incorrect interpretations were due to one or more of the following: items being drawn by hand, extraneous items on the page altering groups, items within the same group being too far apart, or groups appearing too close to each other.

	Number of Problems	Number of Correct Problems	Percent Correct
Distance Groups	25	13	52%
Containers from Distance Groups	17	11	65%
Overall	42	24	57%

Table 1: A summary of the interpreters results when evaluating groups using distance to separate them

Kelvin was helping put oranges in boxes. He had 30 oranges. Each box of oranges holds 6 oranges.

How many boxes did he fill?

Draw an orange on the blank stamp, then use the stamps to create a picture to help you answer the question.




Number sentence: $5 \times 6 = 30$

Henry and Nick were bringing 16 balls out to the playground for recess. They wanted to put them in baskets. Each basket holds 4 balls.

How many baskets did they need?

Use the stamps to create a picture to help you answer the question.




Number sentence: $4 \div 4 = 16$

Kaelin baked 18 cookies. She wants to put them in 3 different bags.

How many cookies will be in each bag?

Use the stamp to create a picture of the cookies in the bags to help you answer the question.




Number sentence: $18 \div 3 = 6$

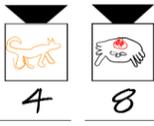
4. In an old house there live some cats, spiders, and people.

Cats have 4 legs. Spiders have 8 legs. People have 2 legs.

In one room there are 4 cats and 3 spiders. How many legs are there altogether?

Draw a cat stamp and a spider stamp. Label each stamp by writing the word cat or spider on the line below each stamp.

Then use the stamps to create a picture to help you answer the question.



4 8



There are 40 legs in all.

Figure 5-6: Student responses using distance as a grouping technique that the interpreter understood correctly

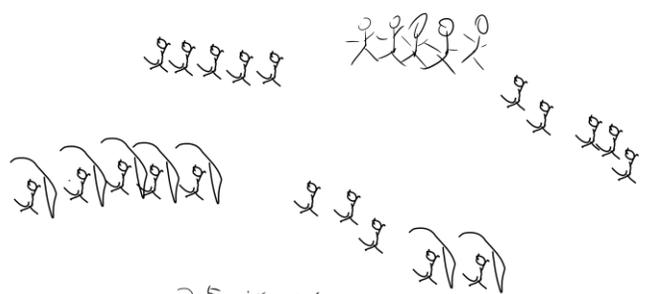
Gaps between objects within the same intended group can prove difficult for the interpreter to correctly understand. The interpreter must determine if the gap signifies a new group or is a misplaced item of an existing group. This decision is a function of the distance cutoff chosen in Section 4.2 and may not reflect the groups a human understands. For instance when viewing the example in Figure 5-7, a human sees five groups of five as depicted in the left image of Figure 5-8. The top right group is circled in red to indicate that the elements are drawn, rendering the interpreter incapable of identifying the objects of the group. The interpreter's grouping, shown in the right image of Figure 5-8, accurately identifies the two groups on the left,

but each of the two groups on the right is seen as a group of three and a group of two. A human recognizes the two groups on the right because she understands from the statement of the problem that the students are asked to represent groups of five. Without information that the groups should contain five items, the interpreter sees a large gap and believes that this gap marks the separation of groups. Similar cases of objects within the same group having too large of a distance between them and being mistaken as separate cases are shown in Figure 5-9. These examples illustrate the difficulty of finding an accurate distance metric without having the knowledge about a problem that a human does.

Nael wants to play basketball. There are 25 students.
Each team has 5 students.

How many teams will there be?

Use the stamp to create a picture of the students. Draw a circle around each team of 5 students.



Number sentence: $25 \div 5 = 5$

Figure 5-7: Gaps within the intended groupings cause the computer's interpretation to differ from a human's interpretation; the student accidentally deleted his stamp and finished his representation by drawing freehand

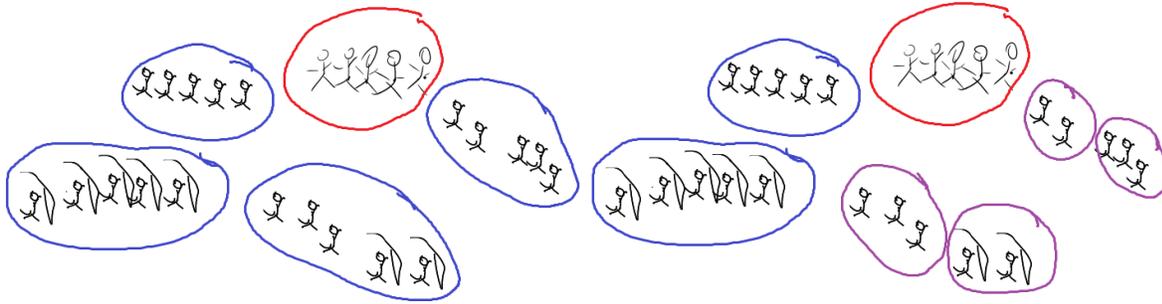


Figure 5-8: Left.) A human's interpretation of the depiction in Figure 5-7 Right.) The interpreter's interpretation of the depiction in Figure 5-7

Nick and Kelvin have 15 plants for the school garden. They are putting them in 3 rows.

How many plants will be in each row?

Draw a plant on the stamp, then use the stamp to create a picture to help you answer the question.

Number sentence: 15 ÷ 3 = 5

Michael has 4 packs of gum with 8 sticks in each pack.

How many sticks of gum are there in all? 32

Use the stamp or tiles or draw a picture to explain your answer.

$4 \times 8 = 32$

Figure 5-9: Incorrect grouping from items being too far apart within a group

Extraneous items also can cause incorrect groupings, as displayed in Figure 5-10. In the example on the left, the third group contains all but two of the extraneous cookies. In the example on the right, despite all of the boxes containing the correct number of donuts, the two extraneous donuts in the middle merge the two middle groups and the two far right groups. Since the time that this student work was created, an easier method to erase unwanted elements was introduced into the CLP system, which has resulted in students producing cleaner pages.

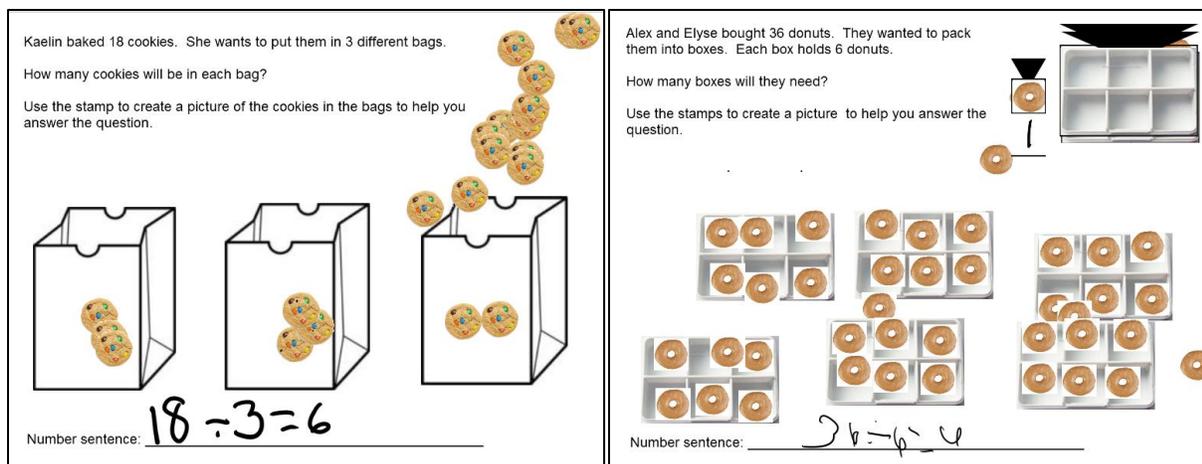


Figure 5-10: Erroneous groupings from extraneous items

Groups also may be incorrectly formed when one item of a group falls too close to another group, as demonstrated in Figure 5-11. In the left example, the student has drawn three pencils over every student container stamp. The distance algorithm does not recognize the student's intention and first creates groups from the overlapping stamps, such as the top left student container with the top pencil from the group below it. In the right example, no erroneous overlaps occur; instead, wrong groupings result from the minimum distance times the threshold being too large and causing the third and fourth groups to be combined. This case suggests that this minimum value or threshold is set too high. Other examples, however, such as those in Figure 5-9, suggest that these values may be too low. A more dynamic breaking point taking into account how many objects the group contains and the separation metrics of the other existing groups on the page could potentially increase the accuracy of the interpreter.

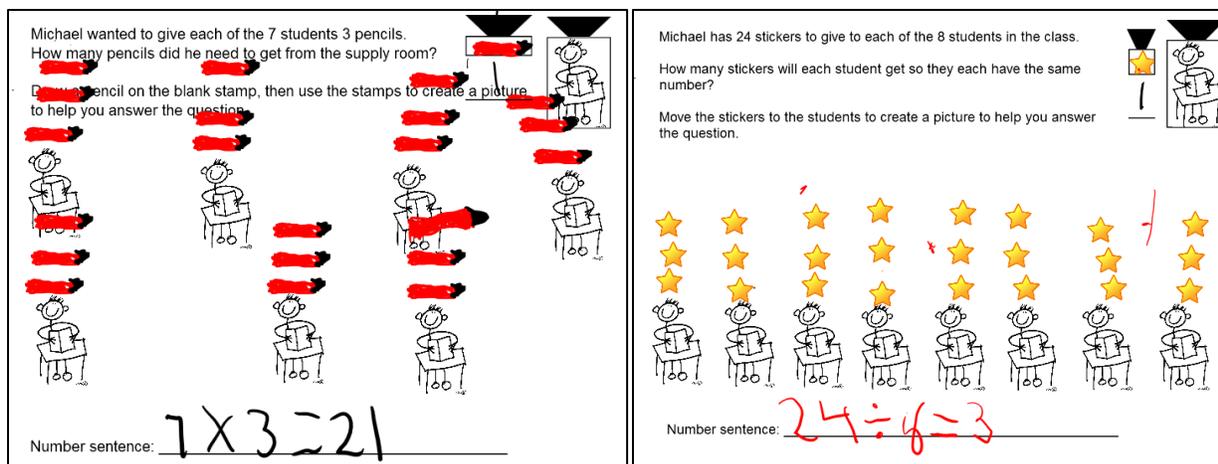


Figure 5-11: Incorrect groupings resulting from groups too close to each other

5.4 Containers

A container as a grouping mechanism occurs when all of the objects a container “holds” are placed on top of the container in such a way that the objects overlap the container. This grouping mechanism only runs when no containers are found within the returned distance groups, as the current implementation of containers as a grouping mechanism is a subcase of the distance algorithm that uses no minimum distance, as described in section 4.4.1. Nine student answers had containers not identified by the distance algorithm; two returned accurate interpretations and are shown in Figure 5-12. In the seven remaining cases, the error resulted from the containers overlapping, as shown in Figure 5-13. Generally, the overlap is extremely small suggesting the other proposed method mentioned in 4.4.1 that required a certain threshold of overlap to occur before combining objects into a group would eliminate the majority of these incorrect cases. That method, however, could have the unintended side effect that objects not completely overlapping their container, such as the bottom left orange in the third container in Figure 5-13, would not be included in the container’s group.

Kelvin was helping put oranges in boxes. He had 30 oranges. Each box of oranges holds 6 oranges.

How many boxes did he fill?

Draw an orange on the blank stamp, the use the stamps to create a picture to help you answer the question.



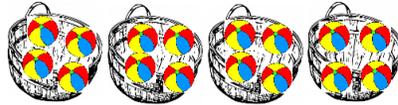


Number sentence: 30 ÷ 6 = 5

Henry and Nick were bringing 16 balls out to the playground for recess. They wanted to put them in baskets. Each basket holds 4 balls.

How many baskets did they need?

Use the stamps to create a picture to help you answer the question.

Number sentence: 16 ÷ 4 = 4

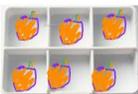
Figure 5-12: Correctly detected containers and groups using the container grouping technique

Kelvin was helping put oranges in boxes. He had 30 oranges. Each box of oranges holds 6 oranges.

How many boxes did he fill?

Draw an orange on the blank stamp, the use the stamps to create a picture to help you answer the question.





Number sentence: 30 ÷ 6 = 5

Alex and Elyse bought 36 donuts. They wanted to pack them into boxes. Each box holds 6 donuts.

How many boxes will they need?

Use the stamps to create a picture to help you answer the question.




Number sentence: _____

Figure 5-13: Examples where the containers overlap causing all of the objects to be seen as one giant group by the interpreter

5.5 Ink Delineated Groups

Ink delineated groups occur when a student creates a shape such as a circle or square around the intended group or places a horizontal or vertical line between the desired groups. The system uses the Microsoft ink interpretation software to identify shapes on the page and then creates an internal representation of the ink groupings on the page, as described in Section 4.3. Twenty-six student solutions use ink delineation with thirteen of these interpreted correctly.

Figure 5-14 depicts some of the student responses for which the interpreter returned the correct groupings by recognizing circles, rectangles, and lines.

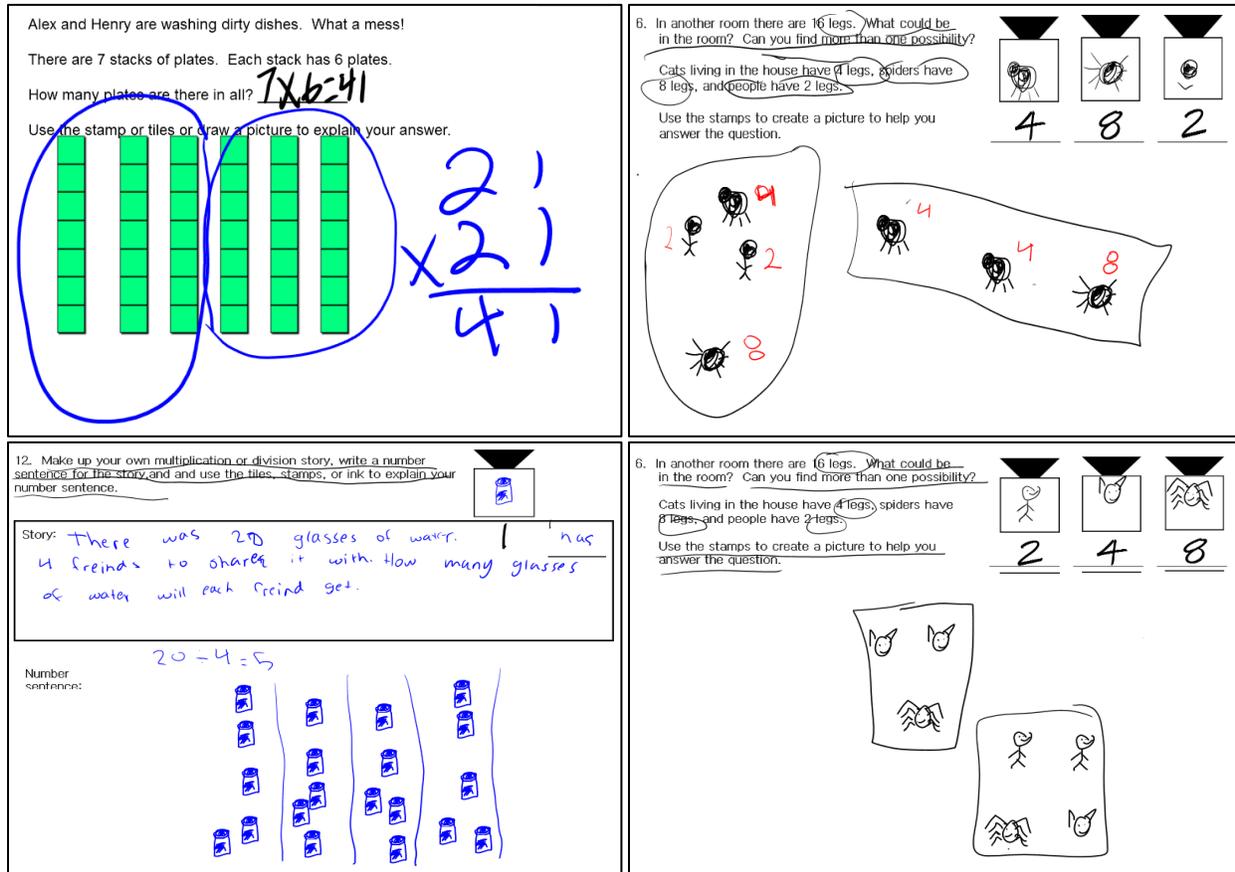


Figure 5-14: Student examples using ink as a grouping technique that the interpreter understood correctly

Half of the incorrect answers were the result of the Microsoft's ink interpretation software not recognizing the student's marks as shapes. In the left image of Figure 5-15, none of the three outlinings were detected, presumably because they were not closed shapes. In the right image, only the rectangles marked with 3, 6, and 15 were recognized by the system. These examples suggest that a more customized shape interpretation system is needed. With such a system, the accuracy of the groups identified via ink may improve greatly.

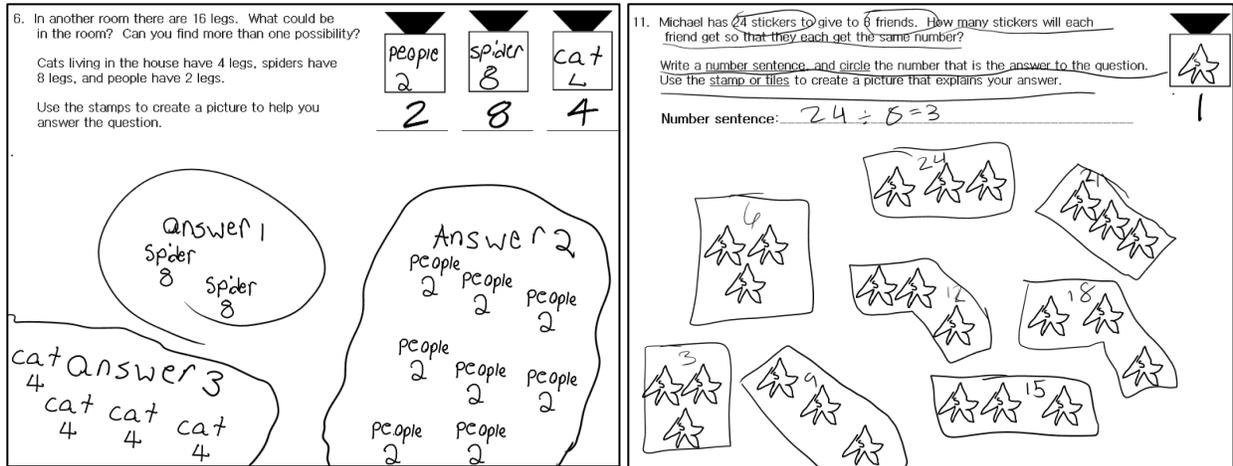


Figure 5-15: Examples where the Microsoft ink interpreter does not recognize the ink as a shape

This thesis extended the Microsoft ink interpretation software to identify horizontal and vertical lines. Lines were detected if they passed two rules:

1. Must be one ink stroke
2. The width or height respective to vertical and horizontal lines cannot be greater than a certain threshold

The goal was to allow some flexibility in the line that a student could draw as a separation technique while disallowing diagonal lines to be used, which do not fit into the box scheme used by the grouping region interpreter's code. One fourth of the remaining incorrect responses resulted from lines not being detected because they did not meet the width/height threshold. In both examples shown in Figure 5-16, the ink interpretation code detects the leftmost vertical line, but ignores the other two lines because they were deemed too diagonal. These results suggest that a more appropriate threshold should be determined when a larger set of examples in which students have drawn grouping lines can be collected.

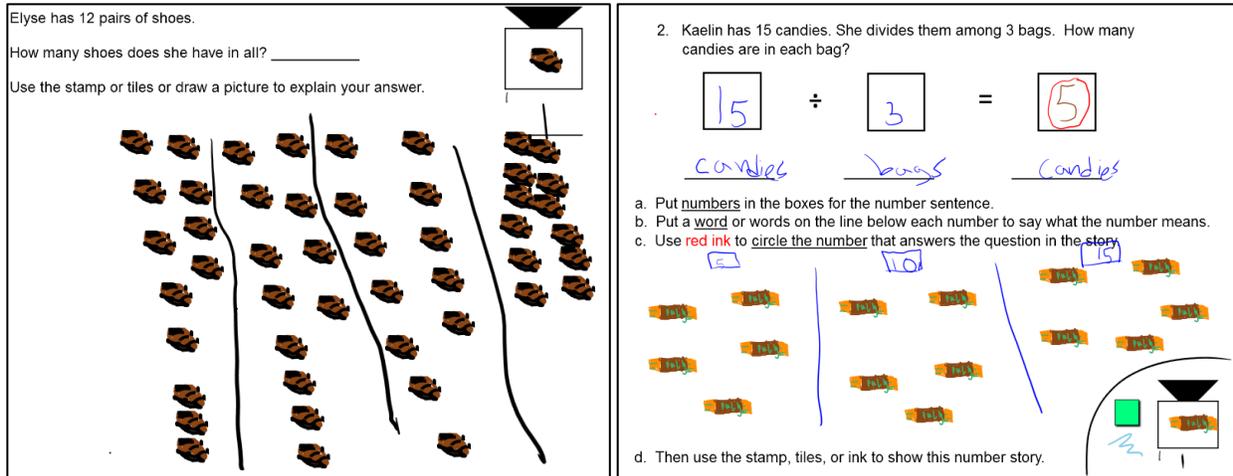


Figure 5-16: Examples of student work utilizing lines to separate their groups that the CLP system does not recognize as vertical lines.

The remaining incorrect results occurred from the interpreter's determination that the bounds of a shape overlapped with another shape on the page, even though they do not necessarily appear overlapped to a human. In Figure 5-17 one of the objects on each page is put into a separate group because of this overlap. While a human may see a slight overlap in the image on the left and no overlap in the image on the right, the interpreter creates its internal representation using the outermost bounds of each object, which creates the bounds seen in in Figure 5-18. Any overlap is treated as its own separate outlined box capable of grouping objects. The only solution for such a problem would depend on improvements in understanding ink strokes.

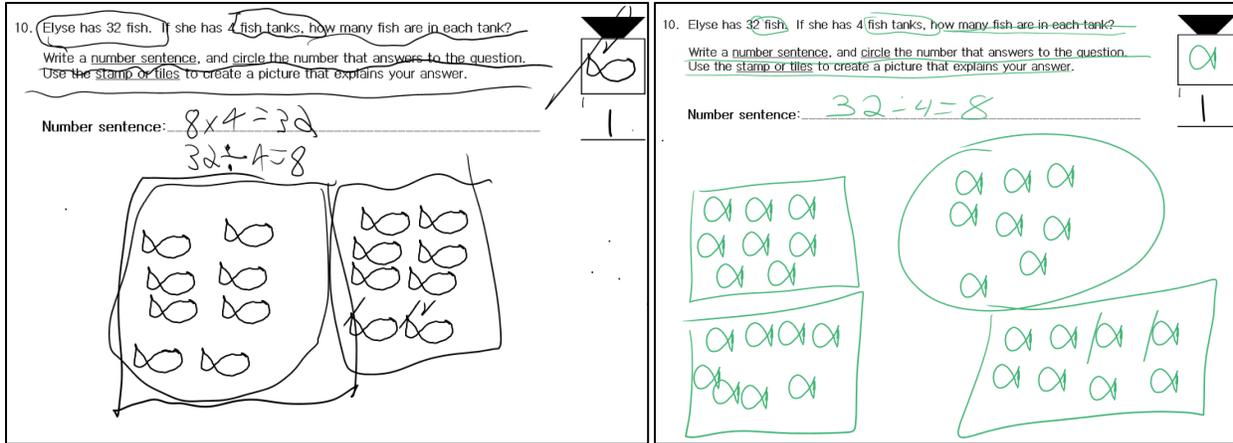


Figure 5-17: Student examples where overlapping ink shapes produce incorrect results

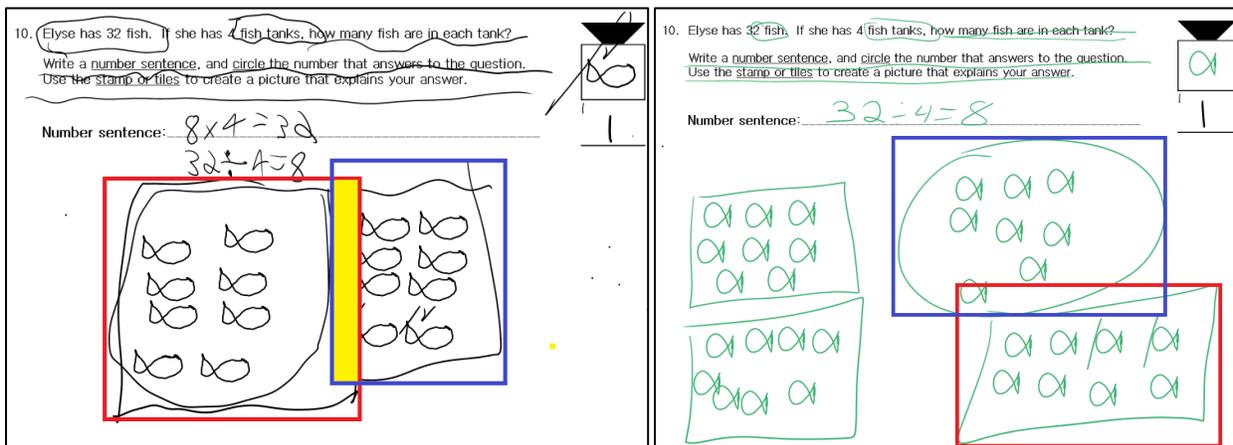


Figure 5-18: Outlines of ink shapes as seen by the Microsoft shape interpretation software

In many cases where grouping by ink fails, the correct groupings are still found through the distance grouping technique because the student sufficiently spaces the groups to allow space for the ink. Despite the Microsoft ink interpretation software not detecting the given ink shapes, both examples from Figure 5-15 are correctly analyzed using the distance grouping technique, supporting why it is advantageous to output all possible groups.

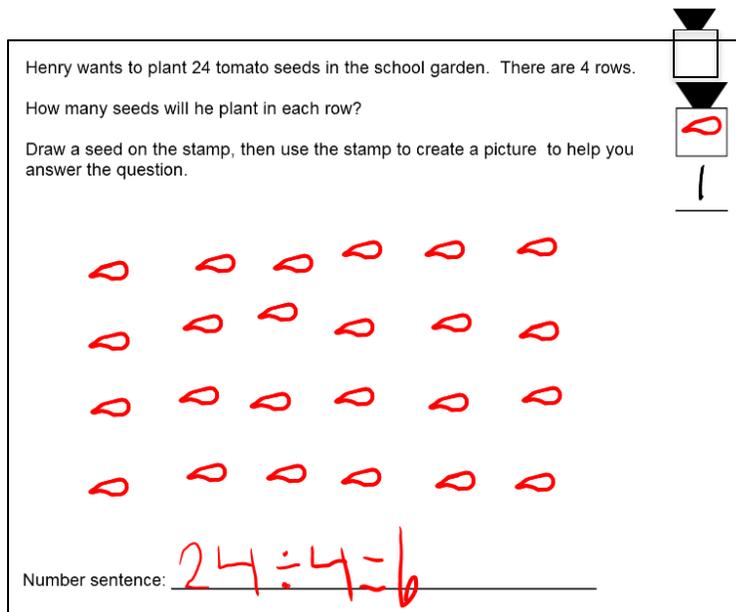
5.6 Other Grouping Techniques

Nineteen percent of the student answers used a grouping technique that is not currently implemented. Eighty-five percent of those unhandled responses created a grid with the objects on the page to express their groupings, such as the response shown in Figure 5-19. Other techniques included using a line to connect all objects within a group, similar to connect the dot, or a modified version of the grid where alternating the columns with two different types of objects is intended to create multiple groupings within the row or vice versa. These techniques are discussed in Section 6.1.3.

Henry wants to plant 24 tomato seeds in the school garden. There are 4 rows.

How many seeds will he plant in each row?

Draw a seed on the stamp, then use the stamp to create a picture to help you answer the question.



Number sentence: $24 \div 4 = 6$

Figure 5-19: Student representation using a grid representation

6 Conclusion

6.1 Future Work

The machine interpretation methods developed in this thesis correctly interpreted 68% of the examples of student work on which they were evaluated. Erroneous and sometimes surprising interpretations provide us with promising new directions for improvement. Future developments fall into the following categories: representation for graphical elements, ink grouping algorithms, additional grouping algorithms, and integration with teacher UI.

6.1.1 Object Representation

In this thesis, we were only concerned with what number a stamp represents, but in future classroom implementations it may be important to understand more about what an object represents. Two possible representations of the stamp in Figure 2-2, for example, are “puppy” or “four paws”. To include this information would most likely require further input from the student in a manner similar to obtaining information about what number a stamp represents. Asking a student to write the word “basket” below a stamp might help the interpreter in identifying the stamp as a container stamp, if the interpreter knew that baskets often acted as containers.

Using word descriptions of what a stamp represents may ease interpretation in some cases, but the question of which words best describe an object is an educational choice more than a technological one. With the puppy example, allowing a student to label a stamp as a single “puppy” may obscure the fact that the puppy is really a representation for four things, namely paws. More scaffolding may be needed by the software or the teacher to lead the student in the correct direction in ambiguous cases such as this one. A teacher, for example, might encourage

students to specify “1 puppy with 4 paws” as a way of introducing a conversation about numbers, groups, and multiplication.

In future releases of CLP, other types of graphical elements besides stamps may benefit from additional information about their meaning. Currently tiles, for example, are always identical, but providing different colors or shapes of tiles may broaden the kinds of visual representations that students create.

6.1.2 Improved Ink Grouping Interpretation

The ink grouping algorithm presented in this thesis is reasonable, but steps could be taken to improve its performance. Currently all delineations are treated as squares by the interpretation region, even if the shape interpreter outputs that the shape is a circle or triangle. Improvement to the bounds of shapes could provide a more accurate grouping mechanism. This modification would require revision of the tree being used and how the interpreter “sees” the page.

If ink strokes cannot be identified as a shape, they are ignored in the current implementation. Many times these invalid ink strokes could be identified as a shape by a human but not by Microsoft’s ink interpretation software. Extensions to the shape interpretation code to recognize the imperfect shapes used by elementary school students would improve the ink grouping methods.

6.1.3 Alternate Types of Grouping

Two additional grouping mechanisms were seen in student work. The first of these mechanisms is the connection of all objects in a group via a line, as shown in Figure 6-1, The line connecting the objects has no restrictions and could be straight, horizontal, vertical, bent, jagged, or any other shape that involves two distinct end points. Stray marks can cause erroneous

results, however, such as the two diagonal lines in Figure 6-1. Filtering of ink could possibly minimize these errors..

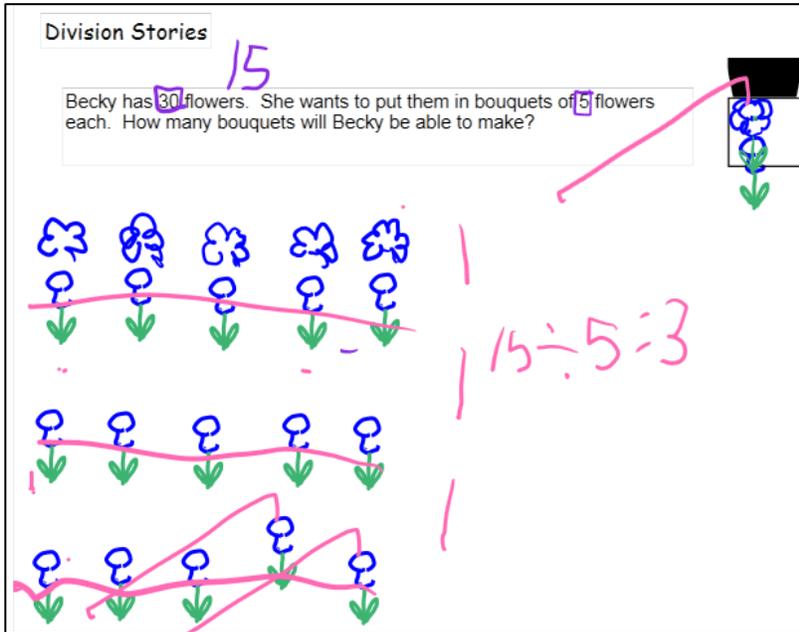


Figure 6-1: Division problem using a line to connect all objects within a group

The second additional grouping mechanism is a grid representation of objects, as shown in Figures 5-19 and 6-2. The student places objects in a grid with either the rows or columns of the formation representing the separate groups. In Figure 6-2, the columns create four groups of six. Difficulties arise from the grid grouping technique because the grid is not perfectly formed, making identification as a grid arduous. In the same example, the columns slowly tilt and the bottom two elements in the far most right column appear to be in the wrong rows, making both the rows and columns ill-formed.

9. Henry wants to plant 24 tomato seeds in the school garden. There are 4 rows. How many seeds does he plant in each row?

Write a number sentence, and circle the number that answers to the question. Use the stamp to create a picture that explains your answer.

Number sentence: $6 \times 4 = 24$

Figure 6-2: A grid representation of the objects to solve a division problem

Students also use grids with containers by making one column or row contain elements of the container, as displayed in Figure 6-3, where each image of a student represents a container for the stars or pencils. The interpreter will need to consider both grouping by row and grouping by column in the returned output since either may be a valid choice. In this problem, the student intends for the bottom row to be treated as a container and for the groups to be found by column, creating eight groups of three. However, one could imagine another problem that wished to output three groups of eight stars and one group of eight students.

Michael has 24 stickers to give to each of the 8 students in the class.

How many stickers will each student get so they each have the same number?

Move the stickers to the students to create a picture to help you answer the question.

Number sentence: $24 \div 8 = 3$

Michael wanted to give each of the 7 students 3 pencils.

How many pencils did he need to get from the supply room?

Draw a pencil on the blank stamp, then use the stamps to create a picture to help you answer the question.

Number sentence: $21 \div 3 = 7$

Figure 6-3: Using containers with the grid grouping technique

A modified version of grids and containers together occurs by placing either alternate rows or columns of the same type of objects, creating a striped effect as seen in Figure 6-4. The student intends for each set of the alternating columns to be considered as a group. This turns each row into a set of four groups with each group consisting of one student container and five letters.

12. Make up your own multiplication or division story, write a number sentence for the story, and use the tiles, stamps, or ink to explain your number sentence.

Story: Sherry sent 5 letters to each 12 friends. How many letters did Sherry send in all?

Number sentence: $12 \times 5 = 60$ letters

Figure 6-4: Alternating vertical stripes of objects create four groups within a single row

6.1.4 Post-Interpretation Integration

CLP currently has a UI that enables a teacher to view all student submissions sorted by student name or time of submission. Current work on the UI is extending this sorting to include other information, including the interpretation information output from the methods in this thesis. In order to not overwhelm the teacher with all possible groupings identified by the interpretation methods, either only one grouping should be returned or the groupings should be ranked. One idea for the rankings is providing a confidence interval of how likely the grouping represents the visual representation on page. Possible strategies would be including information about the type of problem being authored and mapping different kinds of groupings to the problem type [18].

6.2 Contributions

This thesis created an interpretation system capable of understanding visual representations made by elementary students in explaining multiplication and division problems. The key to its success is the introduction of what we call a structured vocabulary: We provide students with tools that give them enough structure to facilitate machine interpretation, but not so much that they cannot be creative in making their own representations. This structured vocabulary enables students to create visual representations that are constructed more easily and quickly than with freehand drawing. Our system constructs an interpretation of the visual representation by finding the relationships between the objects used in the representation, focusing on object type, location, and additional hand-drawn ink lines that indicate grouping of objects. The interpretation methods were evaluated on examples of student work collected in classroom trials in fourth grade classrooms in the Boston area and performed very well. The results indicate that our structured vocabulary and the accompanying interpretation methods will be valuable additions to a classroom interaction system because they provide students with tools for easily and quickly

creating visual representations, and teachers with valuable information about their students' understanding of concepts that underlie the visual representations.

Bibliography

- [1] K. R. Woleck, "Listen to their pictures: an investigation of children's mathematical drawings," in *The Roles of Representation in School Mathematics*, A. A. Cuoco and F. R. Curcio, Eds., Reston, VA, National Council of Teachers of Mathematics, 2001, pp. 215-227.
- [2] K. Koile, K. Chevalier, C. Low, S. Pal, A. Rogal, D. Singer, K. Wu, J. Sorensen and K. S. Tay, "Supporting Pen-Based Classroom Interaction: New Findings and Functionality for Classroom Learning Partner," in *First International Workshop on Pen-Based Learning Technologies*, May 24-26, 2007.
- [3] K. Koile, K. Chevalier, M. Rbeiz, A. Rogal, D. Singer, J. Sorensen, A. Smith, K. S. Tay and K. Wu, "Supporting Feedback and Assessment of Digital Ink Answers to In-Class Exercises," in *Proceedings of the Nineteenth Conference on Innovative Applications of AI*, July 22-29, 2007.
- [4] K. Koile and D. Singer, "Development of a Tablet-PC-based System to Increase Instructor-Student," in *The Impact of Pen-based Technology on Education; Vignettes, Evaluations, and Future Directions*, D. Berque, J. Gray and R. Reed, Eds., Purdue University Press, 2006.
- [5] K. Koile and D. Singer, "Assessing the Impact of a Tablet-PC-based Classroom Interaction System," in *Proceedings of Workshop on the Impact of Pen-Based Technology on Education (WIPTE) 2008*, D. Berque, J. Gray and R. Reed, Eds., Purdue University Pres, 2008.
- [6] K. Koile, D. Reider and A. Rubin, "INK-12: A Pen-based Wireless Classroom Interaction System for K-12," in *The Impact of Tablet PCs and Pen-based Technology on Education: Evidence and Outcomes*, D. A. Berque and R. H. Reed, Eds., Purdue University Press, 2010, pp. 93-102.
- [7] H. Dibeklioglu, T. M. Sezgin and E. Ozcan, "A Recognizer for Free-Hand Graph Drawings," in *First International Workshop on Pen-Based Learning Technologies, 2007.*, 4-5 May 2007.

- [8] M. A. Hearst, M. D. Gross, J. A. Landay and T. F. Stohovich, "Sketching intelligent systems," *Intelligent Systems and their Applications, IEEE*, vol. 13, no. 3, pp. 10-19, May/Jun 1998.
- [9] P. S. Moyer, J. J. Bolyard and M. A. Spikell, "What are virtual manipulatives?," *Teaching Children Mathematics*, vol. 8, no. 6, pp. 372-377, 2000.
- [10] J. L. Mueller, "Data Compression, Storage, and Viewing in Classroom Learning," Master's Thesis, MIT, 2012.
- [11] C. DeRosa, "Dynamic Multimedia in Classroom Learning Partner," Master's Thesis. MIT, 2012.
- [12] E. Maynard, "Understanding Stamps and the Logic They Represent in Classroom Learning Partner," Undergraduate Advanced Project, MIT, 2012.
- [13] K. V. Tish, "Interpretation and Clustering of Handwritten Student Responses.," Master's Thesis, MIT, 2012.
- [14] A. Jain and R. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, Inc., 1988.
- [15] S. C. Johnson, "Hierarchical Clustering Schemes," *Psychometrika*, vol. 2, pp. 241-254, 1967.
- [16] S. P. Borgatti, "How to explain hierarchical clustering," *Connections*, vol. 17, no. 2, pp. 78-80, 1994.
- [17] K. Daniels and C. Giraud-Carrier, "Learning the Threshold in Hierarchical Agglomerative Clustering," in *5th International Conference on Machine Learning and Applications, 2006. ICMLA '06.* , Dec. 2006.
- [18] K. S. Tay and K. Koile, "Improving Digital Ink Interpretation through Expected Type Prediction and Dynamic Dispatch," in *Proceedings of the 19th International Conference on Pattern Recognition*, December 8-11, 2008.