Automatic Assessment of Student Answers Consisting of Venn and Euler Diagrams

Diunuge B. Wijesinghe, Jcs Kadupitiya, Surangika Ranathunga and Gihan Dias
Department of Computer Science & Engineering, University of Moratuwa, Sri Lanka
Email: {diunuge.10, jcskadupitiya.16, surangika, gihan}@cse.mrt.ac.lk

***Abstract*—Venn & Euler diagrams are well-defined mathematical diagram types, which are the major representation methods of Set Theory. Venn and Euler Diagrams exist in various media types such as printed format in books, raster images and vector images in electronic media. Also Venn & Euler diagrams are part of major mathematics exams in secondary education such as London Ordinary Level and SAT Mathematics. Although computer representation and assessment methods of different diagram types such as ER diagrams have been addressed, no such research has been done for Venn and Euler diagrams. In this research, assessment in only for images in vector format. Methodology for set details extraction from a vector image is presented and Venn and Euler diagram assessment method is introduced, which can grade according to a provided marking rubric.**

***Index Terms*— Diagram Assessment, Venn Diagram, Euler Diagram, Set Theory, Diagram interpretation**

I. INTRODUCTION

Diagrams are a very important communication medium. This is especially the case with mathematical diagrams. Although humans can easily interpret these diagrams, mathematical diagram understanding is a complex challenge for researchers. Diagram understanding is an important prerequisite of various fields such as image database systems, and educational grading systems. Significant research has been done to understand mathematical diagrams in few domains such as coordinate graphs [2] as well as charts [5] (bar charts, pie charts). However, there is no research done to interpret Venn and Euler diagrams. Venn & Euler diagrams are a significant part of mathematics, especially in exams of secondary education, such as London Ordinary Level and SAT Mathematics. Since online based education popularity increases such as MOOCs (Massive open online courses), There is a need of an automatic assessment facility for these diagrams.

Venn diagrams can be described as diagrams that represent pictorial relations among sets. Venn diagrams are a specialized instance of a more general notation for representing relationships among a set of classes of concepts referred to as Euler diagrams [20]. Definition of Venn and Euler diagrams varies throughout the literature [9]. Generally, Venn and Euler diagrams can be defined as a finite set of labelled, closed curves. The closed curves in the diagram partition the plane into minimal regions, where each minimal region is a connected component of the plane inside a set of curves [20]. While a Venn diagram contains all the minimal regions, Euler diagrams omit the empty minimal regions. Normally two and three sets of Venn and Euler diagrams are being used in most of the educational contexts such as mathematics and science. Representation up to 6 sets is less complex in the graphical representation. Fig. 1 Shows a Venn diagram and the corresponding Euler diagram of two sets for which the set of $A \cap \sim B$ is empty.

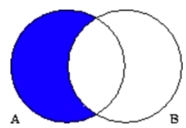
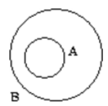
Set system	Venn diagram	Euler diagram
{∅,B,AB}		

Figure 1. Example of a Venn and an Euler Diagram

To address the problem of automatic assessment of Venn and Euler diagrams, we present a methodology that first extracts set details from a vector image and produces the output as an XML structure. This XML structure is assessed according to a marking rubric. Our system accepts the input images as vector images in SVG (scalable vector graphics) format. Under this representation, polygons are used to represent a vector image [21].

In vector images, sets are identified by labelled Jordan Curves that are non-self-intersecting continuous closed paths in the plane [20]. In SVG, sets can be represented using circles, ellipses, rectangles and closed curves. In this research, we only consider circles, ellipses and rectangles. However, same concepts can be extended for the arbitrary closed curves as well, though it is not currently handled. A set usually has an attached label. In the label identification, only nominal text labels are accepted as labels. Since the label identification is subjected to various ambiguity problems, various heuristics had to be used to identify correct labels.

After identifying the sets, minimal regions are identified. Then remaining text is classified as zone elements. After extracting all the Venn information, extracted data is output in a structured XML format that can describe any Venn diagram. It contains set details and the zone details. Arrangement difference in the vector image of the Venn diagram can produce several XML outputs that are equivalent.

Marking rubric structure is introduced to capture the marking rubric that is provided in an exam. A teacher has to provide a model diagram answer and a marking rubric to generate marking rules to grade student diagrams.

In the assessment, the structure of the student diagram is validated against the model answer. Corresponding sets from the model diagram and student diagram are matched by applying a sentence similarity matching algorithm to set labels. After identifying relevant sets student answer is graded against the marking rubric. Marking rubric may contain alternative marking methods for a same question/ sub-question. Maximum collective mark is extracted as the final mark for a student answer.

For the evaluation, A Venn and Euler diagram question paper with four questions taken from the Mathematics paper of the GCE O/L examination in the Sri Lanka to university undergraduate and postgraduate students and collected answers in handwritten format. Also, we have collected answer from a secondary school examination that had Venn and Euler diagram questions. The Sri Lankan ordinary level exams and the Landon ordinary level exams are similar. Based on the accessibility for a large enough test data set, we evaluated the system using data from the mathematics paper of the Sri Lankan ordinary level exam. We manually graded those answer sheets according to the marking rubric. Collected answer sheets were converted to SVG and graded using the system. Marking results produced from the system were validated against the manual marking, which showed an accuracy of 94.8%.

This paper is arranged into following sections: Section II describes the related work. Details of the implemented system are included in Section III. Section IV provides an evaluation of the system. Finally, Section V concludes the paper with possible future extensions.

II. Related Work

In this section, we explore existing research related to mathematical diagram recognition, interpretation, diagram similarity measurements, diagram data representation, and diagram assessment methods.

Mathematical diagram recognition, understanding and evaluation is a relatively new research field. As an early attempt, Futrelle et al [1] presented a diagram understanding system to interpret diagrams based on constraint grammars. The system is capable of handling x-y graphs and gene diagrams on Biological domain [2]. Tsintsifas et al [8] developed a java based framework called DATsys that can be used to understand and evaluate diagrams in formative assessments. They were able to develop the diagram input system that can handle various diagram types. This system was developed as an extension to existing Ceilidh Computer Assessment System [17] and their primary aim was to assist the process of learning. Thomas et al [9, 10, and 12] developed a computer aided assessment system that can handle graph-based diagrams such as Entity-Relationship (E-R) diagrams and flow charts, where information can be represented as data nodes and relations. Diagrams are interpreted using a basic set of units called “Minimal Meaningful Units”. They also introduced a marking criteria based on “Minimal Meaningful Units” that can be applied to diagrams that are represented using nodes and their relationships. Brett et al [15] developed a feedback system that can work with graph-based diagrams by following the work of Tsintsifas et al [8]. They also discussed a simple evaluation method and developed a grammar for E-R diagrams. Batmaz et al [18] developed a diagram drawing tool that can be used for semi-automated database diagram assessment. Huang et al [4, 5] developed a system that can understand chart images. They were able to recognize and identify various types of chart images (both 2D and 3D) and interpret those images and produce an XML output that can be used for further processing. Research from Futrelle et al [1] and Huang et al [4] dealt with raw pixel images that were extracted from hard documents. They converted those pixel images to Vector Graphics format for further processing. Anderson et al [16] discussed the fundamental components of a diagrammatic processing system. In early research, they dealt with many types of input system types based on their research focus.

None of above research addresses the issue of Venn or Euler diagram computer representation and automatic assessment and we are not aware of any other general Venn or Euler diagram assessment research work.

Embedded text in an image is the key to recognize and interpret image information. Futrelle et al [2, 3] discussed the object association problem and heuristic approach development using special techniques in general that can apply for text association. Also, they discussed the ambiguity problems associated with image understanding, especially with text label processing. Huang et al [5] also introduced text classification techniques using machine learning approaches to resolve some ambiguities. When Computer Based Assessment (CBA) systems have their own custom input systems that use associated text in images, these ambiguity problems do not arise. However, those custom input systems provide less degree freedom to the user that will not simulate a real experience in the context of examinations where students get hard papers to answer and have more degree of freedom in drawing a diagram. Some diagram based assessment systems have added restrictions to the label choice in order to minimize the association problem. Diagram drawing tool by Batmaz et al [18] provides a set of labels that can be selected as inputs. DATsys [8] and OpenMark [9] provide label insertion areas but the user can decide the label she wants to enter. Since labelling has more freedom, text association is a hard problem.

In the diagram similarity, text label matching is critical that exists in the short answer grading. Jayal et al [13] discussed the label similarity problem from Natural Language Processing aspect. A research work done by Wael claims that the methodology they have used is the best for any unsupervised technique for short answer assessment [22]. For string based similarity, thirteen well-known algorithms (Damerau-Levenshtein, Jaro, Jaro–Winkler, N-gram, etc.) were used to select the best one. Most of these approaches are not suitable for Venn diagrams in the Sinhala language. Different semantic similarity measures were compared including Knowledge-based and Corpus-based algorithms [23, 24]. The research group has first worked on unsupervised techniques only using knowledge-based measures and corpus-based measures with WordNet hierarchy and

Wikipedia corpus [23]. Then they have tried "perceptron" machine learning algorithm to improve their results [24]. A research done by Li focused on Sentence similarity based on semantic nets and corpus statistics using WordNet [25]. WordNet is still under implementation for the Sinhala language, so we are unable to use their approaches.

Various marking methods have been developed for computer-based assessment of diagrams. Tselonis et al [6] discussed a marking method based on marking matrices for graph-based diagrams. Thomas et al [10] also developed a marking method based on "Minimal Meaningful Units" for graph-based diagrams. Brett et al [15] discussed the marking method and developed a grammar for E-R diagrams. Most of the marking techniques we are aware of cannot be directly adopted to Venn and Euler diagram marking due to the structure differences that exist between Venn/ Euler diagrams and graph equivalent diagrams. As for machine representation of marking rubrics, Richard [26] presented a way of representing a rubric document in a machine understandable format using XML. Li [27] developed a language for mathematical problem representation called Mathematics Assessment Markup Language (MAML), using XML, MathML and SVG. Researchers claim that they can represent any mathematical problem using MAML mark-up language.

III. Methodology

In this section we describe the implementation details of the Venn and Euler diagram parser, marking rubric model and the diagram assessment module.

A. High-Level Architecture

Fig. 2 shows the high-level architecture of the implemented system. The teacher has to provide a marking rubric and a Model diagram to the system for assessment. Then the model diagram is parsed and stored. When a student answer is given for assessment, student diagram is parsed using the parser and given to the assessment module. In the assessment module, diagram structure is validated and graded according to the marking rubric.

Input diagram is given in the SVG format. It contains geometric details of SVG primitive objects such as circles, ellipses and rectangles, as well as other details such as presentation details. Since it contains presentation and SVG specific data, SVG image has to be parsed and primitive shape details should be extracted. These details are needed to build Venn and Euler data information.

From the parsed SVG, sets are identified by finding the Jordan curves. After finding the sets, relevant set labels are identified using a heuristic algorithm ("Set Label Identification" section). After identifying the sets, text labels are associated with the correct minimal regions or zones (set of minimal regions) based on the Venn diagram domain knowledge. Several heuristic algorithms are used to deal with human errors ("Text Association Mapping" section). Heuristic parameter tuning depends on the source of the Venn and Euler diagrams such as distance closeness parameters depends on the font sizes based on the SVG drawing tool. After associating the text labels, Venn diagram is built using the extracted knowledge and output is created as structured XML, which contains only the Venn and Euler details without the initial image presentation details such as orientations, and set curve shapes.

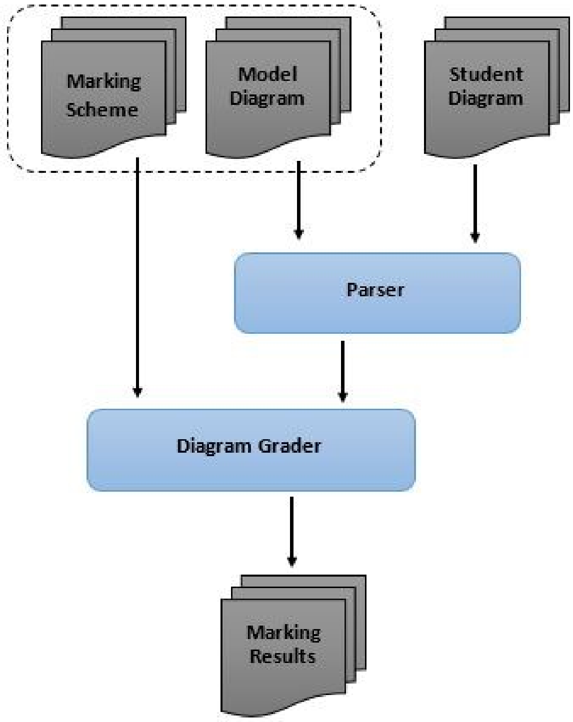


Figure 2. The High-Level Architecture of the system

In the Grader, student diagram is validated for structure similarity such as the number-of-sets. If the structures of two diagrams are matched, assessment is continued. Initially, the set labels in the model diagram and the student diagram are matched using a sentence similarity matching algorithm ("Text Label Matching" section). After set matching is completed, the diagram is graded according to the marking rubric, which is provided by a teacher.

B. SVG Parsing

Vector graphics are the main media of information graphics. Since the W3C standard of vector graphics is SVG, we accepted the input as SVG images. SVG has an XML structure, capable of supporting any Venn diagram representation, supports all major browsers and mobile devices, and is scalable. (Fig. 3 shows a Venn diagram drawn in SVG format).

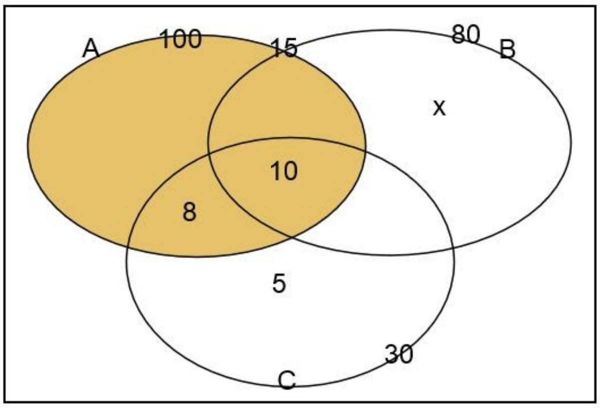


Figure 3. Venn diagrams as an SVG image

SVG of a Venn diagram contains basic mandatory SVG features and optional details such as size and file type, grouping details, image titles, primitive object (rectangle, line, circle …etc.) shape details, primitive object presentation details (fill, stroke details …etc.) and text label details. In SVG parsing, only details required to generate Venn information are retained.

C. Set identification

In Venn and Euler diagrams, sets are represented with Jordan curves (non-self-intersecting closed curves). In SVG diagrams, there are several possible primitive objects such as circles, rectangles, ellipses and closed paths that can be considered as Jordan curves. We identified those closed curves in the SVG diagram from the SVG objects. In this research, only circles, rectangles and the ellipses are considered since the majority of the Venn and Euler diagrams are drawn using those shapes [20]. However, the same methodology can be extended to other primitive objects that act as Jordan curves.

D. Set Label Identification

After identifying sets, associated set labels have to be identified. Sets can be labelled in two ways; with arrows and without arrows by putting the text label near the boundary of a set area. If labelled with arrows, arrow ending has to exist near the set area boundary. If associated arrow is found for any set, then the associated text with the arrow tail is considered as the set label.

If there is no associated arrow for a set, then the closest nominal label near the boundary of the set area is considered as the set label. Figure 5 shows the boundary condition check for an ellipse. In the ellipse boundary condition check, only the centre (C) coordinates and horizontal radius (a) and vertical radius (b) are given in the SVG. Using these details, focal points (F_1, F_2) and the sum of the distance to any point on the ellipse from focal points are calculated (Eq. 1, 2 and 3) from mathematical properties of an ellipse (Fig. 4 shows the required mathematical properties).

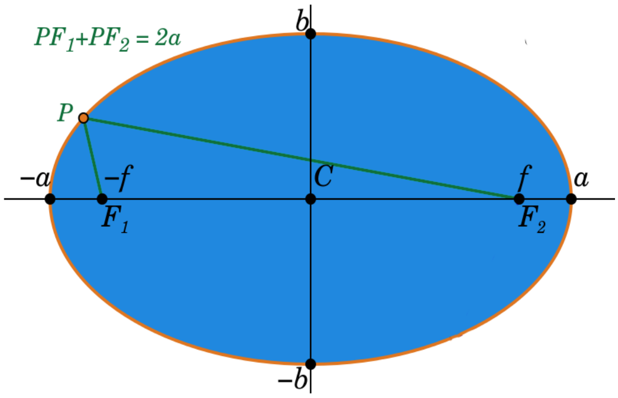


Figure 4. Mathematical properties of an ellipse.

$f = sqrt(a^2 - b^2)$ --- Eq. (1)

$F1 = (Xc - f, Yc)$ --- Eq. (2)

$F2 = (Xc + f, Yc)$ --- Eq. (3)

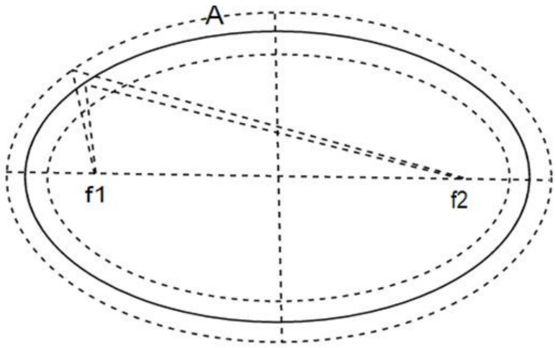


Figure 5. Boundary condition check for an ellipse.

These are used to check the boundary conditions. Text labels appearing in the narrow area near a set boundary are considered as possible set labels for the considered set. From text labels that exist only in the considered set area boundary, the closest nominal label is selected as the label for the corresponding set. Sets that do not have text labels are given computer-generated labels for further processing.

E. Text Association Mapping

Figure 6 shows the algorithm of text label association. In a Venn or Euler diagram, elements associated with minimal regions are marked on the minimal region area. Since the minimal regions are constructed from combinations of set boundary parts, minimal regions normally have complex boundaries. Since the identification of minimal region boundaries is a complex task, minimal region centroid and the area of the minimal region is approximated by counting coordinates belonging to each region.

```
Calculate the centroid and the area of all minimal
     regions
Create pool of text labels except set labels
For (pool of text labels)
If (is text label on a border of a set)
     Text association with corresponding minimal
          region
Else
     Text association with relevant zone
Normalization
```

Figure 6. Pseudo code of the text association mapping

Then, the possible text that can be an element of a minimal region is filtered using the centroid and minimal region area. From the filtered text elements, correct text elements that are in the region are identified.

For an example, Labels “A”, “B”, “C” in Figure 3 are identified as set-labels. “x”, “10”, “8” and “5” within minimal regions are identified as minimal region elements. “100”, “15”, “80” and “30” labels are identified as zone elements that are on the boundary of sets and “15” is associated with “A.B” zone.

F. Venn Data Structure

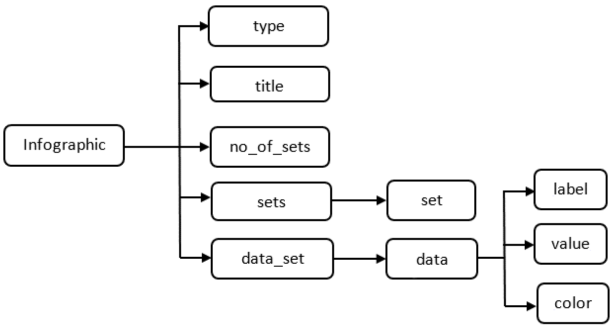


Figure 7. Venn/ Euler data output XML structure

Figure 7 shows the XML schema of the Venn and Euler data output. This Venn Structure is influenced by the XML structure provided by Huang et al [5] for Charts. In the Venn data XML structure, there are five top-level tags; (1) type: Type of the diagram (E.g.: - “Venn Diagram”), (2) Title: Diagram Title Label (If any), (3) no_of_sets: Number of sets in the diagram, (4) sets: set of sets, (5) data_set: minimal region and zone data.

Set order is ignored in the Venn information XML structure. Depending on the set order, different Venn or Euler diagrams may produce equivalent solutions. When reading the output XML ignoring the order of sets, it will produce the same Venn or Euler data.

For a given set, output XML can have a limited number (n!) of equivalent formats depending on the arrangement of sets. Variation of Venn or Euler diagram for 2 sets produces 2 equivalent XML formats. Figure 8 shows the equivalent formats for Venn diagrams with two sets.

```
<infographic>
    <type>Venn Diagram</type>
    <title>Untitled</title>
    <no_of_sets>
        2
    </no_of_sets>
    <Sets>
        <set>A</set>
        <set>B</set>
    </Sets>
    <data_set>
        <zone><label>~A.~B</label><value>none</value>
            <color>none<color></zone>
        <zone><label>~A.B</label><value>18</value>
            <color>none<color></zone>
        <zone><label>A.~B</label><value>20</value>
            <color>none<color></zone>
        <zone><label>A.B</label><value>82</value>
            <color>none<color></zone>
        <zone><label>A</label><value>102</value>
            <color>none<color></zone>
        <zone><label>B</label><value>100</value>
            <color>none<color></zone>
    </data_set>
</infographic>
```

```
<infographic>
    <type>Venn Diagram</type>
    <title>Untitled</title>
    <no_of_sets>
        2
    </no_of_sets>
    <Sets>
        <set>B</set>
        <set>A</set>
    </Sets>
    <data_set>
        <zone><label>~B.~A</label><value>none</value>
            <color>none<color></zone>
        <zone><label>~B.A</label><value>20</value>
            <color>none<color></zone>
        <zone><label>B.~A</label><value>18</value>
            <color>none<color></zone>
        <zone><label>B.A</label><value>82</value>
            <color>none<color></zone>
        <zone><label>B</label><value>100</value>
            <color>none<color></zone>
        <zone><label>A</label><value>102</value>
            <color>none<color></zone>
    </data_set>
</infographic>
```

Figure 8. Equivalent XML representations for Venn diagrams presenting same set data with two sets

G. Marking Rubric

Figure 9 shows the marking rubric structure for Venn and Euler diagrams. Since the marking rubric for the same diagram type can change with the time, the tag has a version attribute that is intended for future changes. Each rubric contains type and question tag. Type tag specifies the diagram category such as “Venn Diagram” and “Euler Diagram”. Question tag contains all the specific marking details. Question tag has “id” (specify the relevant question) and “totalMarks” attributes. Each part of the question is described under “sub_question” tag since the question can have any number of sub-parts with dedicated marking-rules. Each sub_question can have multiple “mark_set” tags, each specifying one possible solution. Each mark set includes three major tags; (1) “mandatory”: specifying the pre-requirements in order for the given solution to be true, (2) data: condition for the evaluation, (3) feedback: suggested feedback by the teacher. Mark set has “method” attribute specifying the way to give marks such as “all of the conditions have to be true”, “at least 3 conditions have to be true” and “any condition has to be true”. “data” part contains the condition for the evaluation. It can check sets, zones, labels and colors regardless of the drawing orientation.

```
<?xml version="1.0" encoding="UTF-8"?>
<rubric version="1.0">
    <type>Venn Diagram</type>
    <question id="q01" totalMarks="2">
        <sub_question id="q01-01"
            model_diagram="dm01" totalMarks="2">
            <mark_set id="1" mark="2" method="all">
                <mandatory></mandatory>
                <data>
                    <zone>
                        <label>~A.B.C</label>
                        <value>5</value>
                        <color>true<color>
                    </zone>
                </data>
                <feedback>
                    <success></success>
                    <fail>Zone element is not
                        marked!</fail>
                </feedback>
            </mark_set>
        <sub_question>
    </question>
</rubric>
```

Figure 9. Example marking rubric for a question

H. Text Label Matching

```
1.  T= tuning parameter
2.  labelValidation (question, studentLabel,
        modelLabel, T)
3.  jointVector[] = createBagOfWords (question,
        studentLabel, modelLabel)
4.  similarityMetrix[][] = createSimilarityM (
        jointVector[], studentLabel, modelLabel)
5.  sim_max = getMaxCosineSim (similarityMetrix)
6.  editDistance = levenshteinEditDist (
        studentLabel, modelLabel)
7.  maxEditDistance = maxLevenshteinDist (
        studentLabel, modelLabel)

8.  if(minEditDistance=<maxEditDistance*sim_max
        *T)
9.      validatedSuccessfully
10. else
11.     validationFailed
```

Figure 10. Pseudo code of the label validation module.

Labels in the diagrams are written in the Sinhala language (the official language in Sri Lanka) and it can be a single word or a short phrase. Figure 10 shows the pseudo code of the label validation module. First, it checks if the labels in the student’s answer directly match with one of the labels found in the model answer labels. If it matches, the system evaluates student label as correct. If not, using the question and the labels provided in the model, this module forms a word joint vector (Bag of Words).

Then Similarity matrix is formed using model answer labels and the student answer labels. Next, the system calculates cosine similarity between pairs of column vectors in the matrix. Finally, it selects the maximum similarity value and the relevant model answer and it calculates the Levenshtein (edit) distance (LD-min). Then it calculates the maximum edit distance (LD-max) between the selected label and the label in the student’s answer. Finally, it checks whether minimum edit distance is less than or equal to the maximum edit distance between student unit and the selected unit multiplied by a tolerance value (T_{val}). Using the training data collection, the tolerance value is tuned to be 0.51.

I. Assessment Module

In the assessment, the diagram is initially checked for mandatory requirements such as the number of sets. If the number of sets matches, similar sets are identified by matching the set names. If a diagram has N sets, at least N-1 set names have to be matched correctly for the assessment.

After matching relevant sets, according to the marking rubric, for each sub-question, all mark-sets are evaluated until at least one mark set gains maximum marks allocated for that sub-question. Then the mark set with maximum marks for that sub-question is collected as final marks for that sub-question.

Similarly, each sub-question in the marking rubric is graded and total mark is calculated from collective marks for sub-questions.

IV. Implementation and Analysis

A. Implementation

The system is developed using Java. Input Venn or Euler diagram images are in SVG format. In the parser, conversion to the relevant set objects from an SVG file is done in the initial phase using Java XPath.

Circles, rectangles, ellipses are identified as sets in the set identification phase from the SVG model. The minimum font size of an SVG image is extracted using all the text labels. In the set label mapping, closeness to the boundary of a set has an upper limit of T_1 * MINIMAL_FONT_SIZE. Nominal label closest only to a set boundary is considered as the set label when arrow labelling is not present. Arrows are presented as lines in SVG. If an arrow ending exists near a boundary of a set, closest nominal label (up to the upper limit of T_2 * MINIMAL_FONT_SIZE) of the other end of the arrow is considered as the set label. After set identification, sets without any label are given a generated label name. In the experiment, T_1 boundary parameter is selected as 1.5 and T_2 boundary parameter is selected as 5 based on the tuning phase results since closeness depends on the image scale that can be tuned. After text label association with the relevant minimal regions and zones, object-oriented model of the Venn data is built and the output is generated as an XML format, which is passed onto the assessment module.

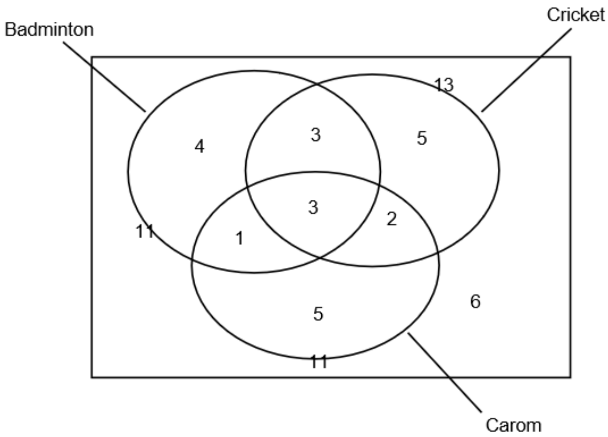


Figure 11. Model answer for a question

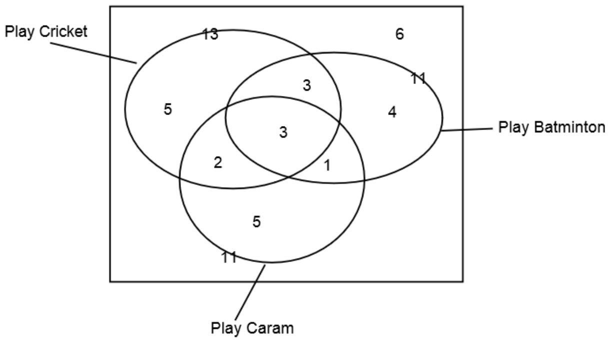


Figure 12. Student answer for a question

Figure 11 shows a model answer for a Venn diagram question that has 3 sets presenting sport details of a class of students (Even though research validated using the Sinhala Language, have use an example in English so that wider audience can understand). A student answer is given in Figure 12, which represents the same set details. Set orientation is different in two diagrams, as well as label positions. Set labels have differences and have spelling mistakes (the third set is labelled as “Caram” instead of “Carom”). In the set label matching phase, these sets are matched. {“Badminton”, “Play Batminton”}, {“Cricket”, “Play Cricket”} and {“Carom”, “Play Caram”} is identified as matching pairs. Then the student diagram is marked according to the marking rubric by comparing the data tuples in the parsed XML and the results will be presented. Since in the set theory set order details are not significant, by omitting the sets order in the XML Venn data, the equivalency of two diagrams can be measured.

B. *Analysis*

Assessment module was and tuned using the Venn diagrams gathered from university undergraduates, which are drawn directly using an SVG editor. For the evaluation, we collected hand-written answer scripts, which collectively contained 3 Venn and 4 Euler diagrams from university undergraduates and grade 10 school students. In total, we had 77 Venn & Euler diagrams. Those diagrams were hand-drawn. Therefore they were converted to Vector format using an SVG editor before given to the system. The model answer for each question was drawn similar to the marking scheme and the marking rubric was generated according to the marking scheme provided originally in the exams. Grader successfully graded 73 of those diagrams, having a collective accuracy of 94.8%.

TABLE I. EVALUATION RESULTS

Diagram No.	Diagram Type	Total Diagrams	Correctly Marked against manual assessment	Accuracy
1	Venn	19	18	94.7%
2	Venn	12	12	100%
3	Euler	11	11	100%
4	Euler	11	11	100%
5	Euler	5	4	80.0%
6	Euler	8	7	87.5%
7	Venn	11	10	90.9%

Diagram 1 and 2 contain 2 sets and remaining diagrams contains 3 sets. Diagram 1 collected from school student answer scripts and remaining from university undergraduates. Diagram 5 and 6 had fewer answers because of the higher complexity of the question.

All of the grading errors are due to the completely different set label names & absence of set labels. In the exam, when a Venn diagram is provided for completion, set labelling is ignored for the marking. We had 2 diagrams without set labels which have marks when evaluating manually but assessment module was unable to understand those diagrams. We had another 2 diagrams where students used only the starting letter of the correct set label as set labels that would not be matched with the model answer. But in the exam, if the teacher can differentiate sets using labels, marks are given.

V. Conclusions

J. *Future Work*

A number of possible further extensions can be identified with respect to the implemented system.

In this research, sets can be drawn from few shapes such as rectangles, circles and ellipses. This method can be extended to apply to sets drawn with any type of Jordan curves. In some Venn and Euler representations, one set label can have more than one Jordan curve. This method can be extended to address those representations.

Machine learning approaches can be used to improve the label classification and clustering. Currently, machine learning methods are not used since we do not have a significantly large vector image database.

Set label matching algorithm can be improved. In the test similarity matching, we currently do not have a method to match words with similar meaning for the Sinhala Language due to the unavailability of natural language processing resources such as WordNet for Sinhala. By integrating a text corpus for Sinhala grader module, accuracy can be improved.

The grader is optimized for the Diagrams given in O/L exams in the Sri Lanka. Assessment module can be extended for other languages and different exams.

This method can be easily extended into other vector formats of images since conversion methods of pixel images format into vector format are already developed.

K. Concluding Remarks

Diagram assessment is a complex problem in the computer research field due to the complexity of diagram understanding. Structured diagrams can be dealt with using the domain ontology. However, unstructured drawing understanding is a very complex problem. In particular, dealing with human generated diagrams is difficult due to the ambiguity problems and human errors that exist in the diagrams. Moreover, for some diagrams, diagram definitions are not standardized. Having various structured and unstructured notations for the same diagram types makes it more complex to address diagram understanding.

In this research, we have successfully established the required methods to interpret Venn or Euler diagrams represented in SVG vector format, introduced a generic format to represent a Venn or Euler diagram, and introduced a marking rubric modeling method and an assessment method for Venn & Euler diagrams. We believe that the solution we have introduced will help to develop educational learning systems and system that require Venn and Euler diagram similarity measurements.

ACKNOWLEDGMENT

This research is funded by DL4D 2016 research grant.

REFERENCES

[1]. R. P. Futrelle, I. Kakadiaris, J. Alexander, C. M. Carriero, N. Nikolakis, and J. M. Futrelle, “Understanding diagrams in technical documents,” in IEEE Computer, 1992, Vol. 25(7), pp. 75-78.

[2]. R. P. Futrelle and N. Nikolakis, “Efficient analysis of complex diagrams using constraint-based parsing. In Document Analysis and Recognition,” in Proceedings of the Third International Conference on IEEE, 1995, Vol. 2, pp. 782-790.

[3]. R. Futrelle, “Ambiguity in visual language theory and its role in diagram parsing”, in Proceedings 1999 IEEE Symposium on Visual Languages, 1999.

[4]. W. Huang, C. Tan, and W. K. Leow, (2005, August). “Associating text and graphics for scientific chart understanding. In Document Analysis and Recognition,” in Proceedings of the Eighth International Conference on IEEE, 2005, pp. 580-584.

[5]. W. Huang, C. Tan, “A System for Understanding Imaged Infographics and Its Applications,” in Proceedings of the 2007 ACM symposium on Document engineering - DocEng '07, 2007.

[6]. C.Tselonis, J. Sargeant, and M. M. Wood, "Diagram matching for human-computer collaborative assessment," Proceedings of the 9th CAA Conference, 2005.

[7]. R. F. Gunstone and R. T. White, (1986). “Assessing understanding by means of Venn diagrams,” Science Education, Vol 70(2), pp. 151-158, 1986.

[8]. A. Tsintsifas, "A framework for the computer based assessment of diagram based coursework," in Computer Science and Information Technology, 2002.

[9]. N. Smith, P. Thomas, and K. Waugh, "Interpreting imprecise diagrams," Diagrammatic Representation and Inference. Springer Berlin Heidelberg, 2004. 239-241.

[10]. P.Thomas, K. Waugh, and N. Smith, "Experiments in the automatic marking of ER-diagrams," ACM SIGCSE Bulletin, Vol 37(3), pp. 158-162, 2005.

[11]. P. Thomas, K. Waugh, and N. Smith, "Generalised diagramming tools with automatic marking," Innovation in Teaching and Learning in Information and Computer Sciences Vol 10(1), pp. 22-34, 2011.

[12]. P. Thomas, K. Waugh, and N. Smith, "Using patterns in the automatic marking of ER-diagrams," ACM SIGCSE Bulletin. Vol. 38(3), 2006.

[13]. J. Ambikesh and M. Shepperd, "The problem of labels in E-assessment of diagrams," Journal on Educational Resources in Computing (JERIC), Vol 8(4), pp. 12, 2009.

[14]. K. Waugh, P. Thomas, and N. Smith, "Toward the automated assessment of entity-relationship diagrams," 2004.

[15]. B. Bligh and C. A. Higgins, "Formative computer based assessment in diagram based domains." ACM SIGCSE Bulletin, Vol 38(3), pp. 98-102, 2006

[16]. M. Anderson and R. McCartney, "Diagram processing: Computing with diagrams," Artificial Intelligence, Vol 145(1), pp. 181-226, 2003.

[17]. E. Foxley, C. Higgins, T. Hegazy, P. Symeonidis, and A. Tsintsifas. "The coursemaster cba system: Improvements over ceilidh," 2001.

[18]. F. Batmaz and C. J. Hinde, "A diagram drawing tool for semi–automatic assessment of conceptual database diagrams," 2006.

[19]. G. Hoggarth and M. Lockyer, "An automated student diagram assessment system," ACM SIGCSE Bulletin, Vol. 30(3), pp. 122-124, 1998.

[20]. P. A. Rodgers, “A survey of Euler diagrams,” Journal of Visual Languages & Computing, vol 25(3), pp.134-155, 2014.

[21]. (2014) The Vector graphics [Online]. Available: https://en.wikipedia.org/wiki/Vector_graphics

[22]. W. H. Gomaa and A. A. Fahmy, “Short Answer Grading Using String Similarity And Corpus-Based Similarity,” on International Journal of Advanced Computer Science and Applications (IJACSA), 2012, vol 3(11), pp. 122-124.

[23]. M. Mohler, “Text-to-text Semantic Similarity for Automatic Short Answer Grading,” Department of Computer Science, University of North Texas, 2009.

[24]. M. Mohler, “Learning to Grade Short Answer Questions using Semantic Similarity Measures and Dependency Graph Alignments,” Department of Computer Science, University of North Texas, 2011

[25]. Y. Li, D. McLean, Z. A. Bandar, J. D. O'shea and K. Crockett, "Sentence similarity based on semantic nets and corpus statistics," on IEEE transactions on knowledge and data engineering, 2006, vol 18(8), pp1138-1150.

[26]. R. F. Boehme, P. G. Fairweather, U. Farooq, D. Lam and K. Singley, "Method, apparatus and computer program code for automation of assessment using rubrics," U.S. Patent Application No. 10/722,926, 2008

[27]. G. Li, W. Su, Y. Zhao, and L. Li, "Design and implementation of MAML," in Proceedings of ATCM, 2006.

Eng. Diunuge B. Wijesinghe was born in Matara, Sri Lanka. He received his B.Sc. (Hons) (in Engineering) degree from the Department of Computer Science and Engineering, University of Moratuwa, Sri Lanka in 2015. He is currently a full time Masters student (Computer Science) at University of Moratuwa, Sri Lanka.
At present he works as a Visiting Instructor of Faculty of Engineering, University of Moratuwa, Sri Lanka on part-time basis. His areas of research interest include Computer Aided Assessment, Automation, Robotics, Machine Learning, and Artificial Intelligence.

Eng. Jcs Kadupitiya was born in Ragama, Sri Lanka. He received his B.Sc. (in Engineering) degree from the Department of Electrical and Information Engineering, University of Ruhuna, Sri Lanka in 2015. He is currently a full time Masters student (Computer Science) at University of Moratuwa, Sri Lanka.
He worked as a Contract Basis Lecturer in University of Ruhuna prior to being enrolled as a full-time Masters student. At present he works as a Visiting Instructor of Faculty of Engineering, University of Moratuwa, Sri Lanka on part-time basis. His areas of research interest include Computer Aided Assessment, natural Language Processing, and Artificial Intelligence.

Dr. Surangika Ranathunga was born in Panadura, Sri Lanka. She received her Bachelor's degree in Computer science and Engineering in 2005 and her Masters degree in Computer Science in 2009 from University of Moratuwa, Sri Lanka. She received her PhD from University of Otago, New Zealand in 2013.
She is currently working as a Senior Lecturer in the Department of Computer Science and Engineering, University of Moratuwa. Her areas of research interests include multi-agent systems, pattern recognition, complex event processing, and natural language processing.

Prof. Gihan V. Dias graduated from the University of Moratuwa in 1985 and was awarded a doctorate in Electrical Engineering and Computer Science from The University of California, Davis in 1992. In addition to being a Professor in Computer Science and Engineering at the University of Moratuwa, Sri Lanka, he serves as the Domain Registrar of the LK Domain Registry and as a Trustee of the Internet Society and President of its Sri Lanka Chapter. He has been active in several areas of computing, including computer networking, language technologies, security, distributed event processing, Internet identity and personal cloud
Prof. Dias is a member of the Asia-Pacific Top-Level Domain Association (APTLD) board, and the President of the Federation of Information Technology Industry Sri Lanka – Software Chapter. He has assisted many public and private organizations to build, secure, and make best use of their computer systems. He was inducted to the Internet Hall of Fame in 2013.