Scientific Research Journal *Vol. 6 No. 1, 15-30, 2009*

Assessing the Line-By-Line Marking Performance of n_Gram String Similarity Method

Arsmah Ibrahim[1], Zainab A Bakar, Nuru'l –'Izzah Othman and Nor Fuzaina Ismail

Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA (UiTM), Malaysia
[1]Email: arsmah@tmsk.uitm.edu.my

ABSTRACT

Manual marking of free-response solutions in mathematics assessments is very demanding in terms of time and effort. Available software equipped with automated marking features to mark open-ended questions has very limited capabilities. In most cases the marking process focuses on the final answer only. Few available software are capable of marking the intermediate steps as is norm in manual marking. This paper discusses the line-by-line marking performance of the n_gram string similarity method using the Dice coefficient as means to measure similarity. The marks awarded by the automated marking process are compared with marks awarded by manual marking. Marks awarded by manual marking are used as the benchmark to gauge the performance of the automated marking technique in terms of its closeness to manual marking.

Keywords: *Automated marking, string similarity, n_gram, Dice coefficient, free-response*

Introduction

Computerized marking of mathematics assessments is an actively researched area. Much of this research culminated in mathematics software packages capable of performing automated marking, however

ISSN 1675-7009

few are capabile of marking free-response answers. Those claiming to have this feature achieve so by exploiting the capabilities of a computer algebra system, while others fully utilize judged mathematical expression (JME) questions. Some examples of packages that utilize a computer algebra system as the underpinning marking engine are Maple TA [1], AIM [2], Question Mark Perception [3] and Wiley e-grade [4], and examples of those that utilize JME questions are CUE [5], Metric [6] and i-Assess [7]. A review of the automated marking features of these software packages and other popular packages revealed that these software are limited to marking a single-line entry of the free-response answers and are unable to mark solutions line-by-line as would a human assessor [8]. However these efforts are commendable and serve as a foundation for further research in this area.

The n_Gram Method

Zainab and Arsmah [9] adopted the n_gram string similarity method as the marking mechanism in the development of a computer program capable of implementing automated line-by-line marking of solutions for the following four (4) linear algebraic equations:

Question 1: $2x = 10$
Question 2: $3x - 15 = 9$
Question 3: $5x + 4 = 10 - 3x$
Question 4: Solve $\frac{x-4}{x} = 3$

The n_gram string similarity method works on the assumption that strings with highly similar structures have a high probability of having the same meaning [9]. In this approach, all mathematical terms are converted into mathematical tokens. A mathematical token is a group of characters which may comprise of numerals and (or) variables and is preceded by either a '+' or a '-' sign. The procedure is used to convert an algebraic equation into a string of mathematical tokens as follows:

i. All terms on the right-hand side of the '=' sign in an equation will be brought to the left-hand side leaving only 0 on the right-hand side.
ii. Every term in an equation will be grouped together with the preceding '+' or '-' sign and will be treated as single tokens. If a term is not preceded by any sign, then a default '+' sign will be assigned.

iii. Bracketed terms and terms with '/' are also regarded as single tokens.
iv. All '=' signs and '0's on the right-hand side will be ignored and not regarded as tokens.

Example 1: $\frac{(2x+1)}{3}+1=2 \Rightarrow +\frac{(2x-1)}{3}$, +1, –2 : Three tokens

Example 1 illustrates the conversion of a mathematical equation $\frac{(2x+1)}{3}+1=2$ into a string of three mathematical tokens $+\frac{(2x-1)}{3}$, +1, and –2 by the procedure.

The degree of correctness between any two mathematical equations is reflected by the degree of similarity between its respective equivalent mathematical token strings. The degree of similarity between two mathematical strings x and y is measured using the Dice coefficient:

$$Dice\,coefficient = \frac{2n\_gram\,(x \cap y)}{n\_gram\,(x) + n\_gram\,(y)}$$

The results of the study suggest that the method is feasible and that programs which this method have the potential to become a tool capable of performing automated marking of free-response mathematics assessments. However more tests need to be carried out to further ascertain the feasibility of the method.

The presented study is an extension of a previous study [10]. It involves marking a sample of another four (4) algebraic equations of different forms and levels of difficulty. This paper presents the results for the further evaluation of the line-by-line marking performance of the n_gram string similarity method using manual marking benchmarking.

The Similarity Measure

The same program for the automated marking procedure used in the previous research has been used in this study. The program is written in C and is still undergoing verification. Implementation requires schemes for all possible solutions to each question and all the respondents' solutions to be keyed in and saved as data files. The Dice coefficient is used to evaluate similarity and the degree of correctness of a respondent's solutions in comparison to the solutions in the answer scheme. For the purposes of this study the Dice coefficient may be mathematically expressed as:

$$d_{i,j} = \frac{2n\_gram\,(x \cap y)}{n\_gram\,(x) + n\_gram\,(y)}, \quad 1 \le i \le m, 1 \le j \le n \tag{1}$$

where x_i is the i-th row string in the respondent's solution scheme and y_j is the j-th row string in the prepared answer scheme, i and j being positive integers. The measure of the degree of correctness of each line of solution is D_j, which is the best Dice coefficient or maximum Dice score chosen from the list of Dice coefficients calculated according to equation (1), where

$$D_j = \max_{1 \le j \le n} d_{i,j} \tag{2}$$

The degree of correctness of the whole question is measured using the average Dice score, which is calculated according to:

$$Average\ Dice\ Score = \frac{\sum_{j=1}^{n} D_j}{n} \tag{3}$$

Data Collection

A sample test consisting of four (4) questions, which require solving different forms of algebraic equations, was given to secondary school students in Shah Alam and Kepong for which there were 78 respondents. The questions are as follows:

Question 1: $\frac{x}{3} - \frac{x}{2} = \frac{1}{6}$

Question 2: $y + 4 = -2(2y + 3)$

Question 3: $3(4 - x) - 5(x - 1) = 3x$

Question 4: $\frac{(2x-1)}{3} = -\frac{x+2}{2}$

Methodology

The respondents' solutions were entered into a computer and saved as data files. A scheme of possible solutions for each question to be used by the automated marking system was prepared (Refer to the Appendix) and entered into the computer, also as data files. The scripts were then marked by the automated technique and manually. The n_gram scores for the automated marking were recorded. The manual marking of the test scripts was performed using a scoring rubric based on the mathematical skills needed to answer the questions. The automated marking scores recorded as maximum Dice scores (MDS) were compared against the manually marked scores, which were recorded as manual mark (MM) and used as the benchmark for the comparison. The measure of closeness between the two scores indicates the accuracy of the automated marking system. The total marks awarded by the automated marking system for each respondent was recorded as an average Dice score (ADS) and the total marks for the manual marking was recorded as the total manual marking score (TMM). The automated marking will be judged as comparable to manual marking if the ADS is equal to the TMM.

Results and Discussions

Table 1 presents the percentage of respondents whose ADS is equal to the TMM and the percentage of respondents with discrepancies between the ADS awarded by the automated marking system and the TMM awarded by manual marking. The results are tabulated in terms of:

Case 1: Similarity in marks given by both automated and manual marking in which ADS is equal to TMM.

Case 2: Totally correct solutions, but given 0 or partial marks by the automated marking system, where totally correct solution refers to a perfect score of 1.00 awarded by the manual marking and a partial mark refers to a score of between 0.00-1.00.

Case 3: Solutions that are awarded a total mark of 0.00 by the manual marking, but given a full (1.00) or partial marks by the automated marking system.

Case 4: Solutions awarded partial marks by both the manual marking and the automated marking system.

The results in Table 1 show that the performance of the n_gram method is fairly satisfactory when marking question 2 with 51.3 % similarity with manual marking. However the performance is unsatisfactory in marking the other questions especially when marking question 3 in which

Table 1: Case 1: Percentage of Similarity and the Discrepancy in Marks Awarded

	Similarity in marks awarded		Discrepancy in marks awarded					
	Case 1		Case 2		Case 3		Case 4	
	ADS = TMM		0.00 ≤ ADS < 1.00 TMM = 1.00		0.00 < ADS ≤ 1.00 TMM = 0.00		0.00 < ADS < 1.00 0.00 < TMM < 1.00	
	No. of Respondents	%	No. of Respondents	%	No. of Respondents	%	No. of Respondents	%
Q1	11	14.1	47	60.3	6	7.7	14	17.9
Q2	40	51.3	19	24.4	8	10.2	11	14.1
Q3	0	0.00	51	65.4	17	21.8	10	12.8
Q4	9	11.5	28	35.9	34	43.6	7	9.0

there is no similarity in the marks given. In order to explain the discrepancies in the marks awarded, the line-by-line performance of the n_gram method for the automated marking has been evaluated. The evaluation consisted of analyzing the maximum Dice scores (MDS) for each line of the respondents' solutions and comparing them to the respective manual marks. Tables 2-4 record the line-by-line maximum Dice scores and manual marks for selected respondents in each case, from Cases 2-4. The resultant contributing factors for the discrepancies have been determined.

Table 2 presents the line-by-line maximum Dice scores and manual marks for selected respondents for each question in Case 2. In the case of Q1R68, the contributing factor for the low MDS for L1 and L2 is the *inability of the program to recognize tokens (2x - 3x) 6* in L1 and –(*1/6) 6* in L2 even though both tokens are available in the answer scheme (Refer to the Appendix). The contributor for the 0.50 MDS is the presence of token -*1/6* in L1. For L2 the MDS should have been 0.50 instead of 0.00, since the token - *(1/6)6* is present in L12, L14, L19 and L20 of the answer scheme. The reason for the inability of the program to recognize the tokens could be due to some *flaws in the tokening algorithm* that implements the program. Another factor that caused the lowering of the MDS is *the absence of necessary tokens.* The absence of token *6(-x/6)* in L2 and tokens *-x/-1* and *1/-1* in L3 of the answer scheme is the reason for a MDS of 0.00 for L2 and L3 when manual marking of L2 and L3 awards each a perfect score of 1.00. These factors consequently lowered the average dice score (ADS) for Q1R68 to only 0.38 when they deserved a score of 1.00.

Table 2: Case 2: $0.00 \leq ADS < 1.00$ but TMM = 1.00

Solution line	Q1		Q2		Q3		Q4	
	Respondent 68		Respondent 60		Respondent 13		Respondent 58	
	MDS	MM	MDS	MM	MDS	MM	MDS	MM
L1	0.50	1.00	0.57		1.00	1.00	0.00	1.00
L2	0.00	1.00	0.67		0.00	1.00	0.25	
L3	0.00		0.50	1.00	0.50	1.00	0.25	1.00
L4	1.00	1.00	1.00				0.33	
L5			0.00	1.00	0.00	1.00		
L6	0.00	1.00	0.50	1.00				
ADS/TMM	0.38	1.00	0.46	1.00	0.50	1.00	0.22	1.00

Key: MDS: Maximum Dice Score   MM: Manual Mark
ADS: Average Dice Score   TMM: Total Manual Mark

In the case of Q2R60, L1 contains the question expression for question 2 (Refer to the Appendix). *Rewriting the question* has caused the MDS to be reduced, as the answer scheme does not contain the question expression. However, due to the presence of tokens *+y* and *+4* as part of the question, which are also present in the answer scheme, a MDS of 0.57 is awarded when L7 and L9 are compared to the answer scheme. In this case, the *inclusion of the question expression* has not only increased the number of solution lines for Q2R60, but also the number of the lines with MDS < 1.00. The net effect of these two factors is the lowering of the ADS. However, the inclusion of the question expression in solutions that are totally wrong will ensure some MDS due to the presence of some matching tokens, thus ensuring some values for the ADS.

In L2 only part of the equation, which is the result of the manipulation of terms on the right-hand side of the equation, was written. In manual marking, L2 is acceptable, but no marks are allocated for this line. In automated marking, the tokening algorithm will convert L2 into +4y+6=0. Since tokens *+4y* and *+6* are available when compared to L7 and L9 of the answer scheme, the MDS awarded is 0.67 even though it is mathematically incorrect. Even though all the tokens in L3 perfectly matched the tokens in L7, L8 and L9 of the answer scheme, L3 is only awarded a score of 0.50 when it deserved a full score of 1.00 according to manual marking. This situation could be caused by some *computation flaws in the program* itself since the manual calculation of the MDS is 1.00. The same situation occurs in L5 in which the tokens are similar to L10 of the answer scheme. The *absence of tokens -2* and *–y* in L6, and the *inability to judge the mathematical equivalence* between the expressions in L6 of R60's solution and the expression in L11 of the answer scheme are the contributing factors that resulted in a MDS of 0.00, wheras manual marking awards a full 1.00 to these solution lines.

Another contributing factor that can reduce the ADS is *the number of solution lines with a MDS of 0.00*. The more lines with MDS = 0.00 the lower the ADS. All the above factors have resulted in a reduced ADS of 0.46 for Q2R60 compared to a full mark of 1.00 with respect to the manual marking. In the case of Q3R13, the *inability of the program to recognize token 17/11* has resulted in a MDS of 0.50 in L3. This score is due to the presence of +*x* in L3. The other contributing factor for lines with MDS < 1.00 in Q3 is the *absence of necessary tokens* in the answer scheme. This reason is also evident in L2 of Q3R13 in which the tokens +*11x* and -*17* are not available in the answer scheme. In the case of Q4R58, the line-by-line answer given does not match any of the automated answer schemes; hence this has been classified as *answer scheme similar to the respondent's solution had not been considered*. This implies that all other tokens except for +*x* in L6 in the respondent's solution are not available in the answer scheme. This accounts for the 0.00 MDS for L1 and L6, and also for the low MDS for lines L2, L3 and L4. However, considering that none of the needed tokens are available in the answer scheme the expected MDS should have been 0.00 instead of 0.25, 0.25 and 0.33 respectively. This again could be attributed to some *computation flaws in the program*. With respect to L6, the 0.50 score, corresponds to the presence of token +*x* in L6.

Table 3 presents the marks for respondents with partial ADS, but with a TMM of 0.00. Analysis of the marks has elucidated that the MDS in all cases are due to the *presence of some matching tokens* in the solution lines of the respondent and the answer scheme. For example, in L1 of Q3R71 the presence of tokens +*12*, -*3x* and -*5x* are the contributing factors for a MDS of 0.80 even though the solution is incorrect according to the manual marking. The same explanation is valid for solution lines L4 and L5 of Q3R71. In fact, the whole solution of R71 did not reflect a true understanding of the relevant concept and a substantial mastery of the necessary skills needed to solve the problem. Another example is that for Q4R53. L1 and L2 of Q4R53 are not even written as equations and do not reflect in any way an understanding of the concept. Therefore in manual marking these lines of solution are not awarded any marks. However, since most of the necessary tokens are available, L1 and L2 are given a MDS of 0.75 and 0.50 respectively. The same can be said

Table 3: Case 3: $0.00 < \text{ADS} \leq 1.00$ but $\text{TMM} = 0.00$

Solution line	Q1		Q2		Q3		Q4	
	Respondent 53		Respondent 65		Respondent 71		Respondent 53	
	MDS	MM	MDS	MM	MDS	MM	MDS	MM
L1	0.75	0.00	0.75	0.00	0.80	0.00	0.75	0.00
L2	0.50	0.00	1.00	0.00	1.00		0.50	0.00
L3			1.00		1.00			
L4			1.00		0.67	0.00		
L5					1.00	0.00		
L6					0.50			
L7					0.50			
ADS/TMM	0.67	0.00	0.90	0.00	0.78	0.00	0.63	0.00

Key: MDS: Maximum Dice Score MM: Manual Mark
ADS: Average Dice Score TMM: Total Manual Mark

about the rest of the respondents in Case 3. Therefore in the case of solutions that are judged as totally wrong by the manual marking standards (TMM = 0), the presence of some matching tokens will result in a relatively higher average Dice score using the automated marking system.

Table 4 displays cases in which both the automated marking system and manual marking awarded partial ADS and partial TMM respectively. The results in Table 4 reveal that in cases where the MDS has some value and the manual mark is 0.00, the scores are again contributed by *the presence of some matching tokens* in the students' solution lines. For example, L2 and L3 of Q3R52 are each awarded a full 1.00 even though they are mathematically unacceptable since they are not written as equations. The whole solution by this respondent did not reflect a substantial mastery of skills needed to solve the problem, which explains why the manual marking awarded no marks at all. In the cases where the MDS is 0.00, but the MM is 1.00, the contributing factor for the lowering of the MDS is the *inability of the program to recognize tokens* $+\ 2(2x\text{-}1)$ and $+3(x+2)$ even though they are available in the answer scheme as in L1 of Q4R69.

Conclusion and Recommendation

The analysis of data in all four cases presented identified six factors that contributed to deviation between the automated line-by-line marking system, which uses the n_gram string similarity method, and manual

Table 4: Case 4: 0.00 < ADS < 1.00 and 0.00 < TMM < 1.00

Solution line	Q1		Q2		Q3		Q4	
	Respondent 5		Respondent 53		Respondent 52		Respondent 69	
	MDS	MM	MDS	MM	MDS	MM	MDS	MM
L1	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
L2	0.67	0.00	1.00	0.00	1.00	0.00	0.75	0.00
L3			1.00	0.00	1.00	0.00	1.00	0.00
L4			0.50	0.00	0.40	0.00	1.00	0.00
L5							1.00	0.00
ADS/TMM	0.83	0.33	0.88	0.33	0.85	0.33	0.75	0.25

Key: MDS: Maximum Dice Score MM: Manual Mark
ADS: Average Dice Score TMM: Total Manual Mark

marking. These factors influenced the maximum dice score (MDS), which in turn affected (either increase or decrease) the average dice score (ADS). The six factors identified are as follows:

i. Inefficiency in terms of correctness of the tokening technique for the conversion of mathematical terms to mathematical tokens, which leads to an inability of the program to recognize certain tokens.
ii. Inability of the program to judge the mathematical equivalence of the expressions between a student's solution and the answer scheme.
iii. The inclusion of the question expression, which leads to a low average dice score in the cases of a totally correct solution (total manual mark = 1.00). However in cases of wrong solutions (total manual mark = 0.00) the inclusion of the question expression increases the average dice score.
iv. Number of solution lines and lines of solution with a maximum dice score = 0.00. The more lines with a MDS = 0.00 the lower the average dice score.
v. The presence or absence of some relevant tokens in the lines of a student's solution that match the tokens in the answer scheme. In cases of a totally correct solution (total manual mark = 1.00) the unavailability of certain tokens reduces the average dice score, whereas in cases of totally wrong solutions (total manual mark = 0.00) the availability of certain tokens increases the average dice score.
vi. The quality of the answer scheme prepared, for which in the case of this study not all possible solutions were considered in the question answer schemes. The answer scheme for a more successful

implementation of an automated marking system must be more extensive and consider all possible solutions.

To improve the maximum Dice scores awarded some improvements and refinements to the technique that implements the automated marking procedure need to be performed. The following are a list of recommendations to improve and refine the automated marking procedure:

i. Refine the tokening technique.
ii. Incorporate a technique to identify the question expression. Once identified, the program should be able to ignore this particular line if it is written by the student and not attribute any marks.
iii. Add a function that takes into account other forms of numbers such as decimals, mixed fractions and exponents.
iv. Incorporate another level of intelligence aside from string similarity to enable the program to judge the mathematical equivalence of expressions.
v. Consider more possible solutions within the answer scheme.
vi. Improve the computation technique used to measure the similarity of expressions.
vi. Consider other similarity measures besides the Dice coefficient; one that is able to award marks that are more reflective of the solution's correctness. The automated marks obtained by computation using Dice tend to be lower than they should be, for example, if one out of four tokens in the solution line is found to be wrong then the mark assigned should be reflective of the solution's correctness, *i.e.* 0.75, which is sometimes not the case upon implementation of the Dice coefficient.

In conclusion, the results of this study have confirmed that the implementation of the n_gram string similarity method in a marking mechanism to automate the marking of free-response mathematics expressions is viable. The discrepancies observed between the average dice score and the total manual score are not due to the n_gram string similarity method, but rather due to the technique that implements the method, especially the computation used to perform the similarity measure and the quality and exhaustiveness of the answer scheme.

References

[1] A. Heck, 2004. Assessment with Maple T.A.: Creation of Test Items. Retrieved from http://www.adeptscience.co.uk/products/mathsim/mapleta/MapleTA_whitepaper.pdf

[2] C. Sangwin, 2004. Assessing mathematics automatically using computer algebra and the internet. *Teaching Mathematics and its Applications*. vol. 23. no. 1. pp. 1-14.

[3] QuestionMark, 2005. Retrieved from http://www.questionmark.com/us/index.aspx

[4] C. Sangwin, 2003. Computer aided assessment with eGrade. *MSOR Connections*. vol. 3. no. 2. pp. 40-42. Retrieved from http://ltsn.mathstore.ac.uk/newsletter/may2003/pdf/egrade.pdf

[5] J. S. Paterson, 2002. The CUE Assessment System. *Maths CAA Series April 2002*. Retrieved from http://www.mathstore.ac.uk/articles/maths-caa-series/apr2002/index.shtml

[6] P. Ramsden, 2004. Fresh Questions, Free Expressions: METRIC's Web-based Self-test Exercises. *Maths CAA Series June 2004*. Retrieved on 5th August 2004. http://ltsn.mathstore.ac.uk/articles/maths-caa-series/june2004/

[7] D. Lawson, 2003. An Assessment of I-assess. *MSOR Connections*. vol. 3. no. 3, pp. 46-49. Retrieved from http://www.mathstore.ac.uk/headocs/33iassess.pdf

[8] Nuru'l–'Izzah Othman and Arsmah Ibrahim, 2005. Automated marking of mathematics assessment in selected CAA packages. *Prosiding Seminar Matematik 2005*, pp. 28-29 Disember. FTMSK, UiTM, Shah Alam. (ISBN 983-43151-0-4).

[9] Zainab Abu Bakar and Arsmah Ibrahim, 2003. Evaluating Automated Grading on Linear Algebraic Equations. *Prosiding Simposium Kebangsaan Sains Matematik ke-XI*, 22-24 Disember 2003. Kota Kinabalu (ISBN 983-2643-27-9), pp. 57-65.

[10] Zainab Abu Bakar and Arsmah Ibrahim, 2003. Experimenting n_gram Method on Linear Algebraic Equations for Online Grading. *International Conference on Research Education in Mathematics,* 2-4 April 2003, INSPEM. Serdang.

Appendix

Question, answer scheme and solutions of selected respondents

Question 1		Student's Solutions		
x/3 - x/2 = 1/6		0 ≤ ADS < 1.00 TMM = 1.00	0 ≤ ADS < 1.00 TMM = 0.00	0 ≤ ADS < 1.00 0 ≤ TMM < 1.00
Answer Scheme		R68	R53	R5
L1	: (2x-3x)/6=1/6	: (2x-3x)/6=1/6	: 2x-3x=1/6	: 2x/6-3x/6=1/6
L2	: (2x-3x-1)/6=0	: 6(-x/6)=(1/6)6	: 12x-18x=1	: 2x-3x=1/6
L3	: (-x-1)/6=0	: -x/-1=1/-1		
L4	12x-18x=6	: x=-1		
L5	: 2x/6-3x/6=1/6			
L6	: 2x/6-3x/6-1/6=0			
L7	: 2x-3x=(1/6)6			
L8	: 2x-3x=1			
L9	: 2x-3x-1=0			
L10	: 6(2x-3x)/6=1			
L11	: 6(2x-3x)=6			
L12	: 6(x/3)-6(x/2)-6(1/6)=0			
L13	: 6(x/3-x/2)=1			
L14	: 6(x/3-x/2)=6(1/6)			
L15	: -6x=6			
L16	: x(2)/3(2)-x(3)/2(3)=1/6			
L17	: -1x/6=1/6			
L18	: -x/6=1/6			
L19	: -1x=(1/6)6			
L20	: -x=(1/6)6			
L21	: -x=6/6			
L22	-x-1=0			
L23	: -x=1			
L24	: x=-1			

Question 2	Student's Solutions

y + 4 = -2(2y + 3)		0 ≤ ADS < 1.00 TMM = 1.00	0 ≤ ADS < 1.00 TMM = 0.00	0 ≤ ADS < 1.00 0 ≤ TMM < 1.00
Answer Scheme		R60	R65	R53
L1	: 5(y+2)=0	: y+4=-2(2y+3)	: y+4=4y-6	: y+4=-4y-6
L2	: 5y/5=-10/5	: =-4y-6	: y+4y=4-6	: y+4+4y+6
L3	: 5y+10=0	: 4+6=-4y-y	: 5y=-10	: 5y+10
L4	: 5y+4=-6	: 10 = -5y	: y=-10/5	: y=2
L5	: 5y=-10	: -2 = y	: y=-2	
L6	: y+2=0			
L7	: y+4=-4y-6			
L8	: y+4y+4+6=0			
L9	: y+4y=-4-6			
L10	: y=-10/5			
L11	: y=-2			
L12	: y=-4y-10			

Question 3		Student's Solutions		
3(4 - x) - 5(x - 1) = 3x		0 ≤ ADS < 1.00 TMM = 1.00	0 ≤ ADS < 1.00 TMM = 0.00	0 ≤ ADS < 1.00 0 ≤ TMM < 1.00
Answer Scheme		R13	R71	R52
L1	: -11x=-17	: 12-3x-5x+5=3x	: 12-3x-5x-5=3x	: 12-3x-5x+5=3x
L2	: 12-3x=3x+5x-5	: 11x=17	: 12+5-3x-5x=3x	: 12+5-3x-5x-3x
L3	: 12-3x=8x-5	: x=17/11	: 17-8x=3x	: 17-8x-3x
L4	: 12-3x-5x+5=3x		: -8x-3x=17	: =17-11x
L5	: 12-3x-5x+5-3x=0		: -11x=-17	
L6	: 17/11=x		: x=17/11	
L7	: 17-11x=0		: x=17/11	
L8	: 17-8x=3x			
L9	: 17-8x-3x=0			
L10	: 3(4-x)=3x+5(x-1)			
L11	: -3x-8x=-5-12			
L12	: -8x-3x=-17			
L13	: x=-11x-17			
L14	: x=17/11			
L15	: x=-17/-11			

Question 4	Student's Solutions

(2x - 1)/3 = -(x + 2)/2		0 ≤ ADS < 1.00 TMM = 1.00	0 ≤ ADS < 1.00 TMM = 0.00	0 ≤ ADS < 1.00 0 ≤ TMM < 1.00
	Answer Scheme	R58	R53	R69
L1	: (4x+3x)/6=-2/3	: (-2)2x-1=x+2(3)	: 4x-2-3x+6	: 2(2x-1)=-3(x+2)
L2	: 2(2x-1)=-3(x+2)	: -4x+2=3x+6	: x+4	: 4x-2=-3x+6
L3	: 2x/3+x/2=-1+1/3	: -4x+2-3x-6=0		: 4x+3x=-6+2
L4	: 2x/3-1/3=-x/2-2/2	: -7x-4=0		: 7x=-4
L5	: 4x+3x=-6+2	: -7x=4		: x=-4/7
L6	: 4x+3x-2+6=0	: x=4/-7		
L7	: 4x-2=-3x+(-6)			
L8	: 4x-2=-3x-6			
L9	: 6(2x-1)/3=-6(x+2)/2			
L10	: 7x+4=0			
L11	: 7x=(-2/3)6			
L12	: 7x=-4			
L13	: x=-4/7			