Excel Grader and *Access Grader*

Dr. Thomas G. Hill
Northeast Mississippi Community College
Booneville, MS
(662) 720-7384
thill@necc.cc.ms.us
http://www.excelgrader.com

Abstract: ***Excel Grader*** *is an automated grader for use by instructors in Microsoft Excel and Microsoft Office application courses. Excel Grader performs static analysis on Microsoft Excel workbooks. The program compares a student workbook with the instructor's correct version of the workbook. The student workbook is marked and a grade report is embedded. Excel Grader includes reports and a tool for plagiarism detection.* ***Access Grader*** *is a companion product that grades Microsoft Access assignments. Access Grader embeds an "ERRORS" table in the student database. Access Grader grades database table and queries. Access Grader has similar reports and user interface as Excel Grader. The graders have been tested in thirty-nine Microsoft Office applications classes (with about 800 students) over two semesters. Thousands of assignments have been graded. According to a survey of instructors using the programs, the graders were extremely accurate and saved hours of grading time. Students that were surveyed felt the grading was fair and accurate.*

1. Excel Grader

Throughout the history of computer science education, a number of grading systems for evaluating programming assignments have been developed. These include Kassandra [von Matt], Ceilidh/CourseMaster [Foxley], BOSS [Joy], and Curator [McQuain].

Although program grading has been an active area of research, there is little evidence of research on automated grading systems for office applications. One commercially available system is SAM (Thomson Learning/Course Technology) [Thompson]. SAM grades Microsoft Office tasks performed in a simulated Office XP environment. SAM assigns students a task, and then records the students' mouse and keyboard events used in their response. Those events are compared with possible correct actions and the results are graded.

Excel Grader (see Figure 1) is another automated grader for use by instructors in Microsoft Excel and Microsoft Office application courses. In contrast to SAM, *Excel Grader* performs static analysis of Microsoft Excel workbooks.

Excel courses often require students to create worksheets based on a list of requirements. When the instructor has a correct solution to an assignment, this correct version can be compared to the student's worksheet. The differences detected can be used as a basis for marking and grading the student's worksheet.

Excel Grader has two goals. The first is to grade assignments more accurately and in greater detail when compared to hand grading via printed copies of assignments. The second is to reduce the time and effort required in grading Excel assignments.

Excel Grader is a Microsoft Access application (the grader requires both Excel 2002 (XP) and Access 2002). *Excel Grader* extracts attributes from a student workbook and the instructor's correct version of the workbook. These attributes are placed into Access database tables. The tables are compared. Missing and incorrect attributes are detected and recorded.
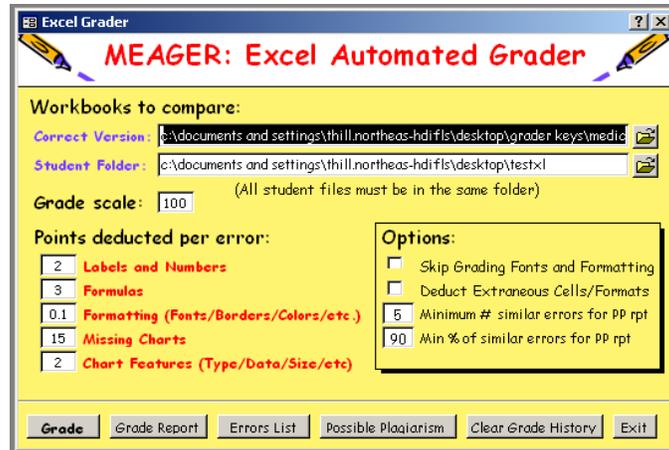
Figure 1 *Excel Grader* Screen Shot

Excel Grader grades the following worksheet attributes:

- text labels
- numbers
- formulas
- fonts (face, size, style: bold/italics/underline, color)
- cell alignment (left/right/center)
- cell fill color
- indent level
- cell orientation (tilt)
- number formats (currency, percent, accounting, scientific, etc.)
- borders (color, thickness, style: solid, broken, etc.)
- row height
- column width
- merged ranges
- worksheet names
- chart type, subtype, chart location, and chart source data

Excel Grader does not conditional formatting, embedded objects (such as pictures, clip art, organizational charts, word art, etc.). Neither does it grade drawing objects (shapes, lines, drop shadows, etc.) or controls (buttons, list boxes, text boxes, etc.).

2. Marking Errors

Excel Grader marks the original student workbook in two ways. A strike-through format is added to an error cell that has an invalid text label, number, or formula (see Figure 2). Comments are added to cells with error values and/or formatting errors. Errors (for all workbooks) are also recorded in the errors table. Figure 3 shows an entire graded workbook. Multiple comments are shown for illustration.
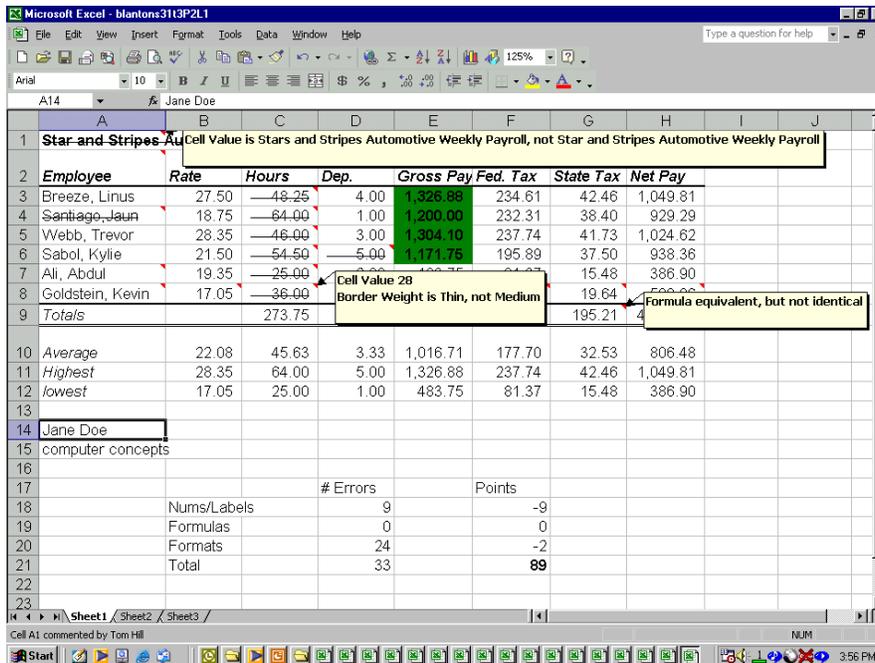

Figure 2 Cell Marking and a comment

Figure 3 A Student's Marked Up Worksheet

A grade summary is added to the bottom of the worksheet. This summary gives the number of problems missed and the number of points deducted.

3. Challenges: Grading Formulas / Renaming Worksheets

Excel formulas are similar to algebraic expressions, one difference being operands are often cell references. Some examples are =E15-E16, =C9+(10%*C9), =SUM(D8:D12). The last formula uses a built-in Excel function. *Excel Grader* first compares cell formulas as text strings. When the text strings are equal, the student formula is correct. When the strings are unequal, the formula may still be algebraically equivalent.

In the following example (Figure 4), the lower right corner sum may be obtained in different but algebraically equivalent ways. The spreadsheet comes from [*Shelly* 2].
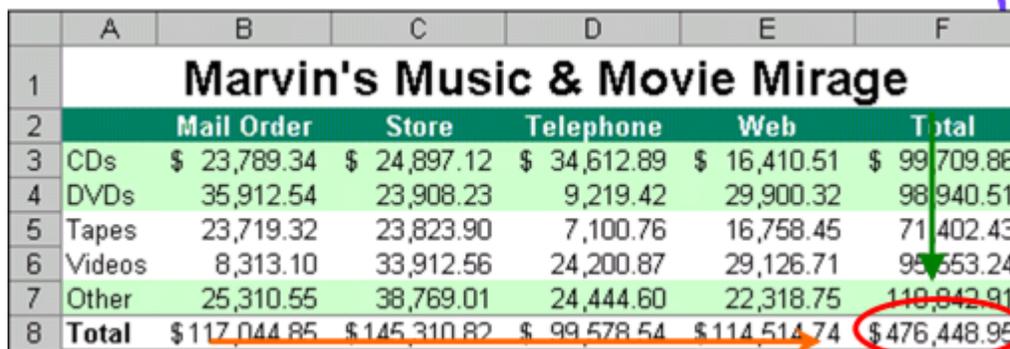

Figure 4 Two paths to the same sum [*Shelly* 2]

Since both column F and row 8 are totals, the sum of either is equivalent to =Sum(B3:E7), the numbers in the interior of the spreadsheet. So possible answers for the formula in cell F8 are =Sum(F3:F7), =Sum(B8:E8), or =Sum(B3:E7).

The above problem is one example in a category of formulas equivalent because of the order in which the operations are performed (sum by row versus sum by column). Since addition is commutative and associative, addition in either order gives the same answer.

Similar problems occur with algebraically equivalent formulas like =B9+(10%*B9) and =B9*(1+10%). Algebraic identities must be applied to convert one formula into the second.

Excel Grader solves this type of grading problem by employing the computer algebra system *Symbolic*. *Symbolic* is an interpreter for symbolic manipulation of mathematical expressions (like Maple or Mathematica) [Dolinksy].

Symbolic accepts identifiers consisting of at least one letter followed by several letters or digits. This means *Symbolic* can process cell references (A1 to IV65536) as variables. *Symbolic* also evaluates functions with the same syntax as Excel: SQRT(x), EXP(x), LN(x), SIN(x), COS(x), TAN(x), ASIN(x), ACOS(x), ATAN(x), SINH(x), COSH(x), ABS(x).

On the other hand, *Symbolic* does not handle recognize ranges, many Excel functions, including common ones like Sum and Average. *Symbolic* does not recognize Excel formula symbols like $ and %. These special cases are handled prior to calling *Symbolic*. Dollar sign symbols are stripped from the formulas, so *Excel Grader* does not currently mark a formula if they are missing.

But most importantly *Symbolic* includes a useful function: Simplify(instr, outstr). This function can simplify an algebraic expression. If both formulas are equivalent, we should be able to subtract one from the other and have an expression that reduces to zero.

Another problem that occurs when grading worksheets is a when a student misnames a worksheet. This problem occurs when an assignment has multiple worksheets and the student is instructed to rename each worksheet. For example, an assignment has three worksheets to be renamed: Payroll, AP, AR. If a student forgets to rename the worksheets, the sheet names retain their default values: Sheet1, Sheet2, Sheet3.

Cells cannot be accessed without first knowing the worksheet in which they are embedded. If corresponding worksheets are not matched, the student's worksheet contents cannot be graded. A minor error would result in major points being deducted.

Often the ordinal position of the worksheets will match, but this is unfortunately not always the case. *Excel Grader* uses a worksheet statistic to match worksheets: the number of non-empty cells in the worksheet. Worksheets with the closest number of non-empty cells are assumed to correspond. The worksheet name is then renamed in the Access tables used to model the worksheet (Cell, Format, Border, Formula, Row, Col). In order to mark the original error, the Worksheet table entry is not renamed.

4. Access Grader

Access Grader (Figure 5) is an automated grader for use by instructors teaching Microsoft Access database courses (or course components). *Access Grader* compares a correct version of an Access database with a student's version and records the errors in an Access table.

Access Grader grades the following database attributes: table data, table structure (field names, field types, field sizes), field attributes (default values, required fields, validation rules, validation text), indexes, primary keys, relations, SELECT queries (fields, table name, join clause, where criteria), total queries, DELETE queries (table name, where criteria), and UPDATE queries (table name, col=expr list, where critera).

Access Grader does not grade data access pages, macros, or modules. *Access Grader* can detect whether a form or report is absent, but it does not grade the form or report in detail.
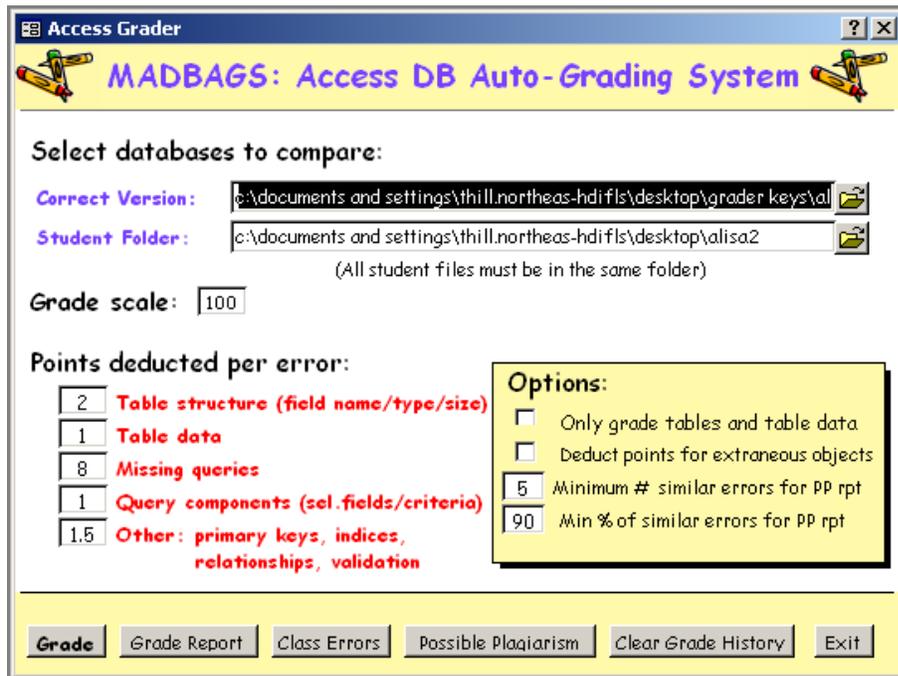

Figure 5 *Access Grader* Screen Shot

To give students a documentation of the grading, *Access Grader* places a table named "E R R O R S" in the students' databases. The "E R R O R S" table includes the following information: error type, table in which error occurred, record (primary key value or first field value), field in which error occurred, the correct answer, and the student's incorrect answer (see Figure 6). A grade report is added to the bottom of the "E R R O R S" table. This report includes a count of the type of errors found, the number of points deducted for each error type, and a grade. In the example below, the student had five mistyped (or forgotten) field values. The student also had two incorrect field attributes (default field value and validation text). The primary key was also omitted from the Customer Table.

Type	Table	Record	Field	CorrectValue	StudentValue
Field Value	Driver Table	75	YTD Earnings	23417	0
Field Value	Driver Table	75	Hourly Rate	17	0
Field Value	Driver Table	75	Zip Code	37613	37631
Field Value	Driver Table	75	State	FL	F
Field Value	Driver Table	75	Last Name	Ortiz	Oritz
Fields	Customer Table	ValidationText	Amount Paid	Must be between $0.00 and $90,000.00	Must be between o.00 and 90000.00
Fields	Customer Table	DefaultValue	Customer Type		SER
Indexes	Customer Table	Required	PrimaryKey	-1	0
Indexes	Customer Table	Primary	PrimaryKey	-1	0
	# Errors	Points			
Field Def	0	0			
Table Data	5	5			
Query	0	0			
Other	5	7.5			
Grade	10	88			

Figure 6 Example "E R R O R S" Table

5. Evaluations

In Fall of 2003, three instructors who were beta testing *Excel Grader* were surveyed. The results of the survey follow.

- Rate the accuracy of the grader:
    - *Extremely accurate (3 for 3)*
- Rate the difficulty learning to use the grader:
    - *Easy (3 for 3)*
- Give the total number of student complaints:
    - *4 student complaints for all instructors out of 392 total students (1%)*
- Give the number of files skipped (the grader was unable to process):
    - *2 total files for all instructors*
- Estimate of the number of hours the grader saved you per week:
    - *Instructor 1: 6 hours, Instructor 2:15 hours, Instructor 3: 10 hours*

A survey of students was also conducted during the Fall 2003 semester. Sixty students were surveyed. Fifty-nine students (98%) rated *Excel Grader* as accurate or extremely accurate (see Figure 7.) Forty-eight students (80%) felt the grading was fair. Five students (8%) felt the grading process was harsh (see Figure 8.) Twenty-four students (40%) preferred the automated grader to an instructor. Seven students (12%) preferred the instructor, and 29 (48%) students had no preference (see Figure 9.)
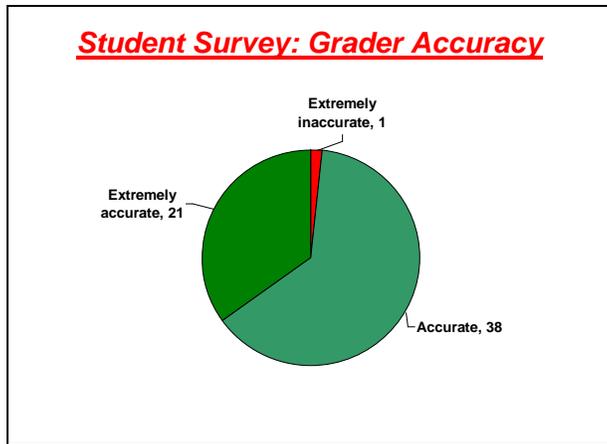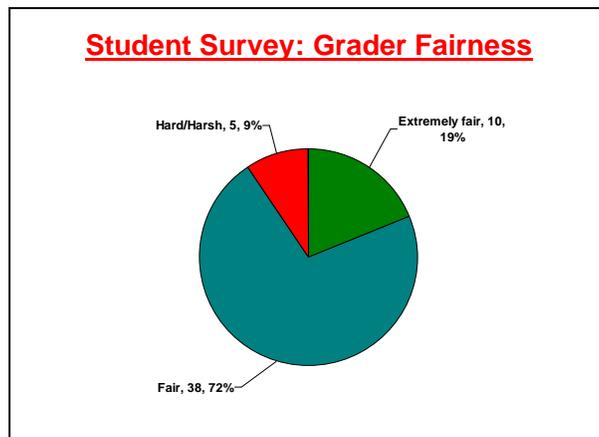
Figure 7 Student Survey: Grader Accuracy

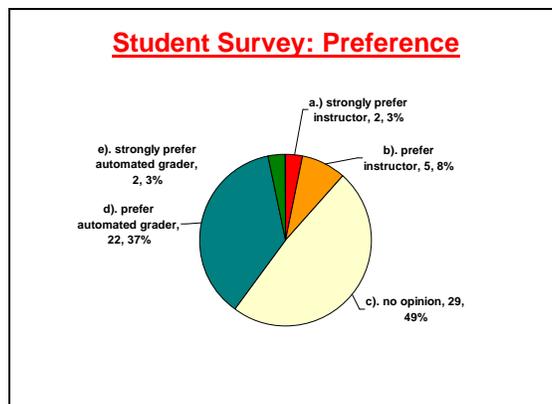

Figure 8 Student Survey: Grader Fairness

Figure 9 Student Survey: Grading Preference

To evaluate *Excel Grader*, Excel Grader processed several assignments previously graded by hand by two instructors. The comparison yielded the following observations on the strengths of the grader:

- **The graders mark in greater detail than human graders.**
  An instructor will rarely mark all the details missed by a student who has many errors on an assignment. Details ignored (or missed) by human graders included: alignments, column widths, row heights, missing borders, number formats, merged cells, and bold formatting

- **The grader can detect errors a human grader cannot.**
  Excel Grader can grade negative number formats for currency cells when the value of the cell is positive. *Excel Grader* can grade font color and fill color, which may be difficult or impossible if the instructor has a grayscale printout of the worksheet. *Excel Grader* can differentiate between fonts which look similar but are different.

Experience using both graders has also shown anecdotal evidence that:

- **The graders can reduce the time required to grade assignments.**
  When using *Excel Grader*, the Excel instructor will want to verify the results of the program's grading. Objects not graded (such as charts) must be graded manually. The instructor must also record the results. Still, anecdotal evidence suggests these two tasks require less time than grading by hand. This timesaving comes primarily from not having to search for errors. Since *Excel Grader* marks worksheet errors, the instructor only has to evaluate the severity of the errors that have been marked. It is the practice of beta testers to check documents with low grades. Other workbooks are returned to students without being opened.

Weaknesses of the graders include:

- **The graders require instructors to gather and redistribute Excel workbook files.**
  Instructors may choose to have students give them floppy disks, or submit files via ftp. If assignments are submitted as email attachments, Attachment Executive (http://www.madsolutions.com/AE/Main.htm) is a commercially available product for processing attachments and moving them to directories. Blat (http://www.interlog.com/~tcharron/blat.html) is an excellent freeware command line tool for mailing graded Excel files as attachments. GetMail (http://www.interlog.com/~tcharron/getmail.html) is the companion tool for downloading and extracting files from an email account. An overview of the process used by beta testers for gathering, grading, and returning student files is found in Figure 10. This process can be adapted for those with access to distance learning environments. The grader includes a batch file that can be scheduled for overnight grading.
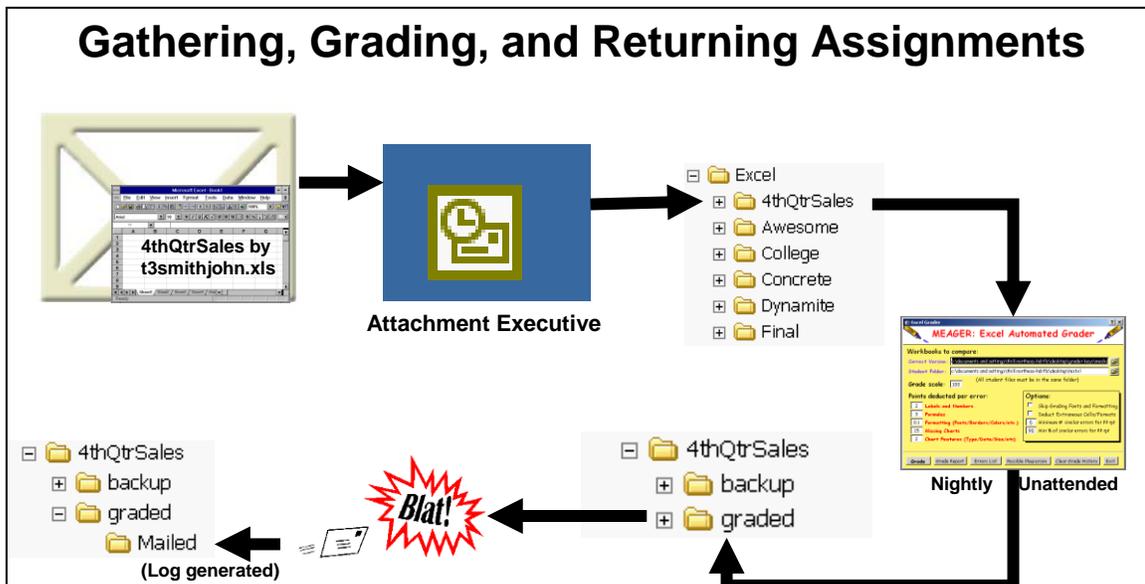
Figure 10 Example scenario for acquiring and returning workbooks

- **The graders' marking may seem excessive.**
  Since *Excel Grader* accepts the instructor's worksheet attributes as correct, non-equivalent student values are marked as incorrect. Extra spaces in student labels are marked. *Excel Grader* can differ between number formats and fonts that look almost identical (for example, Currency with no dollar sign versus Accounting with no dollar sign). *Excel Grader* grades negative number formatting for cells with Currency and Accounting formats.

- **The graders may occasionally skip a file.**
  Excel Grader still does not handle all types of exceptions. When *Excel Grader* encounters an error, it skips the workbook and continues grading the remainder of the files. Skipped workbooks are given a grade of -1 in the grade report. The instructor must grade these workbooks manually.

6. Future Work

Future plans for *Excel Grader* include continued testing and corrections. The program has been used for four semesters grading thousands of assignments from two different textbooks ([*Shelly* 1] and [*Shelly* 2]). Exposing the program to more assignments from other textbooks will improve the program's handling of exceptional conditions.

There is also a need for tools to facilitate assignment gathering and redistribution. Students frequently misname documents that are submitted. Student email addresses are also incorrectly submitted. Web pages that require students to submit assignments in a proper manner would be useful. Integrating the program with distance learning environments is also needed.

Once an assignment is received, the ultimate goal is to have the graders perform their work without any human intervention. The current practice of beta testers is manually checking all assignments that receive a grade less than 80%. Hopefully, the graders will improve to the point that instructors will be content to receive reports of the grading results.

Both graders could grade more document attributes. In particular, grading Excel conditional formatting would be useful. Grading more Excel chart attributes and Access form and report features would improve the program. Having the programs deduct points for late submissions has also been suggested.

Student and instructor surveys at other educational institutions would advance the research. More studies contrasting the grader results with hand grading would also be constructive.

Information about the programs can be found at http://www.excelgrader.com and http://www.accessgrader.com..

References

Dolinsky, Jens-Uwe and Pawletta, Thorsten, Interpreter for symbolic manipulation of mathematical expressions, Internet WWW page, at URL <http://www.mb.hs-wismar.de/~pawel/Uwe/casE.html> (version current at September 12, 2001).

Foxley, E., Higgins, C. A , Tsintsifas, A and Symeonidis, P, The Ceilidh CourseMaster System, An Introduction, 4th Java in the Curriculum Conference, South Bank University, UK, 24th Jan 2000.

Forsythe, George, and Wirth, Niklaus, Automatic grading programs; *Commuications of the ACM* 8, 5 (May. 1965), Pages 275 – 278

Hill, Thomas G., MEAGER & MADBAGS: Automated Graders for Microsoft Excel and Access Assignments, Unpublished doctoral dissertation, University of Mississippi, Oxford.

Joy, Mike, and Luck, Michael, The BOSS System for On-line Submission and Assessment. *Monitor: Journal of the CTI Centre for Computing* 10 pp. 27-29, 1998

McQuain, William, Curator Introduction. Internet WWW pages and documentation at http://ei.cs.vt.edu/~eags/Curator.html> (version current at September 8, 2000).

Pierce, Matt, MDB Diff. Internet WWW page, at URL <http://jupiter.drw.net/matpie/PBSystems/products/retired/MDBDiff.html> (version current at October 7, 2000).

Shelly, Gary B., Cashman, Thomas J., and Vermaat, Misty, Microsoft Office 2000 Introductory Concepts and Techniques, Course Technology, Cambridge, MA, 2000.

Shelly, Gary B., Cashman, Thomas J., and Vermaat, Misty, Microsoft Office XP Introductory Concepts and Techniques, Course Technology, Cambridge, MA, 2002.

Thomson Learning/Course Technology. Course Technology – SAM. Internet WWW page, at URL <http://www.course.com/assessment/sam/default.cfm> (version current at November 29, 2002).

von Matt, Urs. Kassandra: the automatic grading system. *SIGCUE*(22):26-40, 1994.