Comparison of a Computer-Based Grading Scheme to a Manual Rubric for Assessing Solid Models

Steven J. Kirstukas and Luz Amaya-Bower
Department of Engineering
Central Connecticut State University

Abstract

The same set of student solid models were graded manually with a rubric and by a computer program. The rubric was created and used by an instructor not familiar with the computer program. The different philosophies of the people behind the two grading schemes produced some different assessments. Best practices were shared and ideas for improvement of the computer program were generated.

Introduction

A computer program was recently presented that can assess and score solid models of student-created parts (Kirstukas, 2016). The output of this program compared favorably to previously generated hand-graded results scored by an algorithm that shared some of the same philosophies—that all sketches use dimensions from the original drawing and employ appropriate geometric constraints, and that part models be changeable in ways that honor the design intent expressed in the original drawing. In this study, the same computer program is used to assess a set of new solid models that were hand-graded by a different grader who had some similar and some different priorities.

Methods

The rubrics used for the manual grading of solid models focused on 9 or 10 categories, with each category being worth 6 to 16% of the total grade. Some of the typical categories are shown in Table 1.

#	Items from hand-graded rubric	Used by program?	If no, possible future implementation in program?
1	Location of part on coordinate frame	no	yes
2	Orientation of part on coordinate frame	yes	
3	Proper units	yes	
4	Proper use of geometric constraints	yes	
5	Correct dimensions	yes	
6	Presence of specific extrusions or revolves	no	yes
7	Presence of datum plane	no	yes

8	Presence of a particular feature, patterned features, or mirrored features	no	yes
9	Model saved in isometric view	no	yes
10	Model saved in specified rendering style	no	yes
11	Overall model	no	no

Table 1. Hand-graded rubric compared to the computer algorithm

Table 1 shows some of the differences between the philosophies of the manual grading rubric and the computer algorithm. The computer algorithm permits any choice of part origin and permits construction of the model by any means, provided there is a one-to-one correlation between dimensions in the original drawing and dimensions in the solid model. Both systems expect proper use of geometric constraints: In the rubric system this is done by inspection; in contrast, if the computer algorithm finds fully constrained sketches that only use the dimensions in the original drawing, it extrapolates that acceptable geometric constraints were used. The last category, Overall Model, which often carried fairly high weight, focused on a subjective feel of the solid model and would be difficult to implement into the computer algorithm.

The two grading systems were employed on a set of two different part-modelling exams. Exam 1 focused on unit specification, creating fully constrained sketches on appropriate planes, extrusions, revolutions, and rendering style. The exam consisted of two different parts (Figure 1). The first part was a moderately complex symmetrical bracket whose description contained 13 dimensions. Its solid model could be generated by three sketches and five extrudes. The second part was a fairly simple revolved part consisting of stacked cylinders and required 10 dimensions. Its solid model could be generated by one sketch and one revolve.

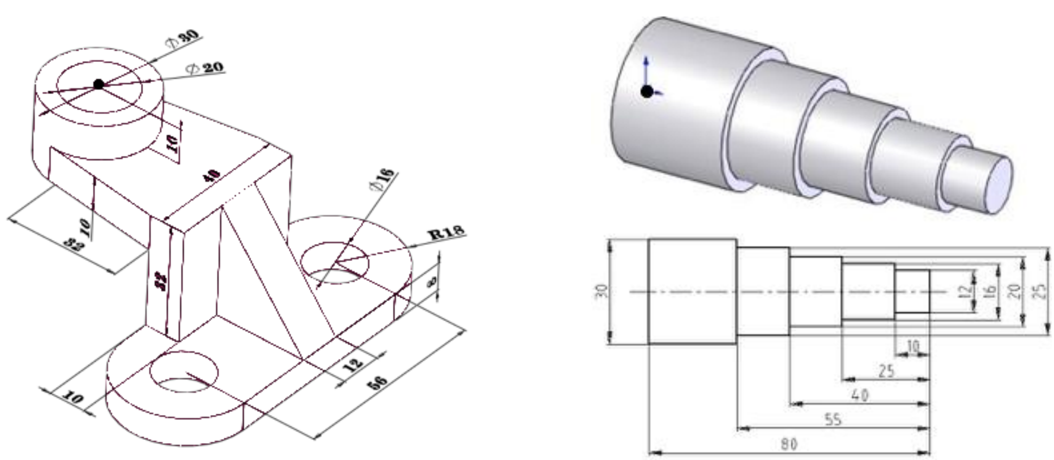

Figure 1. The two parts used in Exam 1

Exam 2 expanded on the goals of Exam 1 and incorporated features such as shell, tube, slot, groove, and threads, as well as patterned features and mirrored features. The exam consisted of three different parts (Figure 2), although creation of only two solid models was required. Students could choose to model either of the first two parts, and a model of the third part was required. The first part was a gear whose description contained 10 dimensions. Its solid model could be generated by two sketches, a datum plane, three extrudes (or two extrudes and a slot), a pattern of extrudes (or slots), and a set of mirrored features. The second part was a bottle that required 18 dimensions (not all are shown in Figure 2). Its solid model could be generated by one sketch, one revolve, a shell operation, and threads. Alternatively, a simpler sketch could be used and two features could be created using a circular groove and a U-groove. The third part was a paperclip whose model could be created with one sketch and a tube operation. It required the six shown dimensions and a wire diameter of 0.04 in.

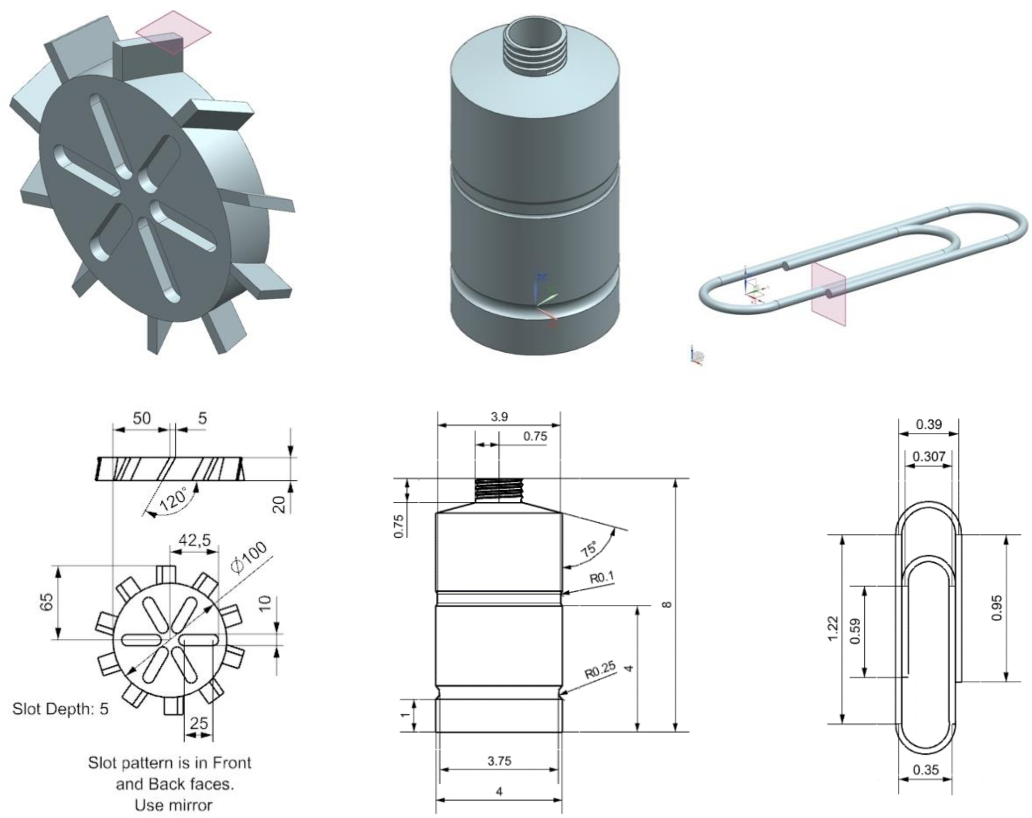

Figure 2. The three parts used in Exam 2

Results

A summary of the results is presented in Table 2. Each block of three columns represents one of the five parts from the exams. For example, the first block shows that 19 students generated solid models for the first part in Exam 1. Manual scoring ranged from 30 to 100, with a mean of 84. Computer scoring ranged from 26 to 90, with a mean of 73. The third column of each block shows the mean difference, the minimum difference, and the maximum difference observed in the 19 models. For instance, the minimum difference of –35 indicates that for one particular student model for Exam 1 – Part 1, the hand-graded rubric generated a score 35 points higher than the computer score. The maximum difference of +14 indicates that for one particular student model for Exam 1 – Part 1, the computer score was 14 points higher than the hand-graded rubric score. In general, the scoring ranges produced by the hand-graded rubric are similar to those of the computer algorithm, except for Exam 2 – Part 1, which only seven students attempted. The predominance of high negative minimum difference (min diff) values indicates that some student models greatly benefited from the hand-graded rubric. The generally smaller positive maximum difference (max diff) values indicates that some student models scored higher by the computer algorithm.

	E1 - P1, n= 19			E1 - P2, n = 19			E2 - P1, n = 7			E2 - P2, n = 12			E2 - P3, n = 19		
	R	C	diff	R	C	diff	R	C	diff	R	C	diff	R	C	diff
mean	84	73	-11	79	59	-19	98	74	-24	77	70	-6	86	74	-12
min	30	26	-35	30	33	-43	92	50	-50	43	41	-42	46	16	-84
max	100	90	14	100	100	3	100	90	-1	100	97	13	100	97	12

Table 2. Scores via hand-graded rubrics (R) and by the computer algorithm (C)

In plotting the two schemes against one another, it is seen that while many data points lie on or near the line that indicates perfect correlation, one-third to one-half of all data points are significantly far from the line, and these tend to be below the perfect correlation line (Figure 3). Student models corresponding to the data points far below the line are graded much more harshly by the computer than by the hand-graded rubric.

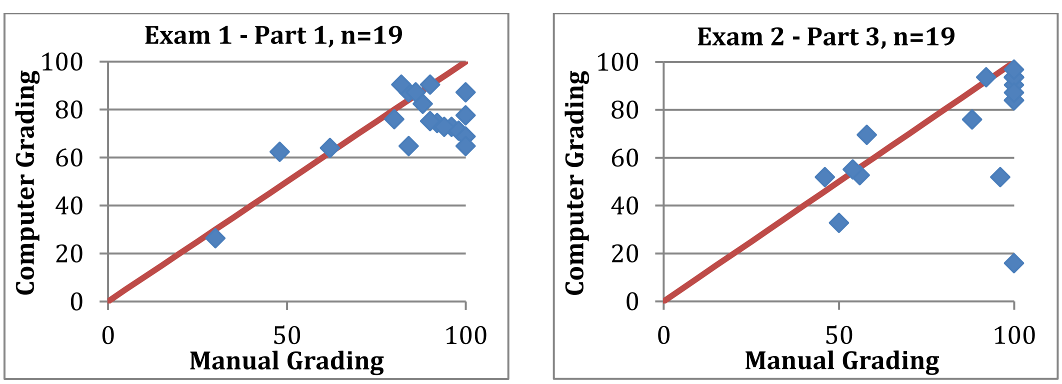

Figure 3. Comparison of grading schemes for two of the parts used in the exams

The hand-graded rubric was concerned primarily with the as-submitted model geometry. In contrast, the computer program valued changeability so it penalized repetitious dimensions such as a 7-mm deep hole in a 7-mm thick plate. Instead of repeating a numerical dimension that interferes with changeability, phrases such as “through all” or “until next” are better used.

Some of the discrepancy in scoring was due to missing or unwanted dimensions. For things like symmetrical extrudes with a total width of 10, the instructor using the manual rubric would accept either “10/2” or 5 for symmetric extrude depths as they both produce the same solid geometry. In contrast, the computer program required that all drawing dimensions be fully incorporated into the model, so if the student specified 5 instead of “10/2”, then two deductions would occur: one for a missing dimension (the 10) and one for an unwanted dimension (the 5).

The hand-graded rubric was relatively soft on non-constrained sketches as the maximum deduction was limited to 16%, and deductions on the high end were only employed when there was a trend of disregarding constraints on multiple sketches.

Conclusions

It has previously been noted that the manual checking of part files is a time-consuming task, and impractical for large classes (Ault et al., 2013). For this study, manual scoring of the total 38 models for each exam took about 4 minutes apiece for a total grading time of 152 minutes – just over two and a half hours. In contrast, the computer could process the same number of models in just a few seconds. The human grader is subject to time pressures and fatigue and sometimes misses issues such as non-united bodies, or absent or inappropriate dimensions and geometric constraints. Other studies (e.g., Branoff et al., 2016) have shown sometimes large variability in scoring when different untrained graders use the same hand-graded rubric in assessing solid models of parts. The computer algorithm aims to eliminate the variability in scoring and enable fast feedback.

This has been a valuable exercise for both instructors, as it initiated a discussion of assessment philosophies. Some emphasis in teaching and assessment will change going forward. Additionally, ideas for improved computer assessment were noted, such as optionally requiring the presence of specific features, such as revolves, holes, or slots. Some bugs in the computer algorithm were noted, such as in evaluating part orientation of imperfect models. Moment of inertia magnitudes are currently used to aid in evaluating orientation. If a student model has the expected geometry and expected orientation, then moments of inertia about the lower bounding box axes will match those of the instructor’s model, even when different part origins are employed. However, if the student model has some minor flaws in geometry but nevertheless has correct orientation, then moment of inertia computations will be different. Work is ongoing on adjusting inertial tolerances and evaluating other schemes to better judge model orientation.

References

Ault, H. K., & Fraser, A. (2013), *A Comparison of Manual vs. Online Grading for Solid Models* Paper presented at 2013 ASEE Annual Conference, Atlanta, Georgia. https://peer.asee.org/19045

Branoff, T. J., & Devine, K. L., & Brown, J. (2016, June), *Evaluating a Rubric for Assessing Constraint-Based Solid Models*, Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. https://peer.asee.org/26747

Kirstukas, S. J. (2016, June), *Development and Evaluation of a Computer Program to Assess Student CAD Models*, Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. https://peer.asee.org/26781