AC 2011-2060: MODE OF ERROR ANALYSIS OF STUDENT RESPONSES TO PRE-REQUISITE KNOWLEDGE ASSESSMENTS

David B Benson, Kettering University
Mode of Error Analysis of Student Responses to Pre-Requisite Knowledge Assessments

Abstract

In engineering education there are a number of central concepts and skills that form threads which connect one content area to another within a discipline. These threads form the core of an engineering education and are the scaffold upon which all future knowledge is built. An incomplete understanding in any of one of these concepts at an early stage in a student’s education can lead to a cascade of failures or difficulties that resonate throughout their academic career. Although a program of study is designed so that students entering a given class have successfully completed all of the pre-requisite course material to attempt the class, student recall and understanding of prior content varies. A longitudinal study is in progress to assess student abilities and growth in these key threads.

Current research has identified and mapped a number of central content and skill trajectories that are present in engineering education, focusing primarily on science and math content/skills essential to Mechanical Engineering. Several of these key content topics and skills have been assessed using pre-tests along their trajectories for selected classes. Common student errors (modes of failure) within these assessments have been identified and classified to generate a profile of the error modes for each topic. These modes of failure are indicative of cognitive issues such as transposition of skills, incomplete understanding of underlying concepts, and the disassociation of tool and concept. A severity of error scale has been developed and applied to student responses to facilitate the correlation of relationships between success and failure among the different trajectories. Results indicate that for many topics the types of errors encountered are shared by large numbers of students. Relationships between failures and the persistence of failures as students progress throughout the curriculum are presented. Validation of the error modes has been conducted through inter-rater reliability studies and student interviews.

Background

In engineering education there are a number of central concepts and skills that form threads which connect one content area to another within a discipline. These threads generally consist of basic or simple concepts and are central to a student’s engineering education because they form the scaffold upon which higher-order knowledge constructed. The recurrence of these threads throughout a curriculum or program of study is referred to as a trajectory.

One aspect of these recurring concepts and skills is that they may enter a given class or subject at a variety of different levels ranging from “central to the development of concepts” to “tool oriented”. In addition, when these trajectories are used to support advanced content their development need not follow a logical progression within an advanced course since they are considered pre-requisite knowledge: a student may use the concept in a sophisticated manner in one class and then simply as a tool in a later class. An example of a trajectory essential to mechanical engineering is the concept of the dot product, also known as the scalar inner product. Figure 1 shows a representative program of study in Mechanical Engineering at
University with the intersections of the dot product trajectory with various classes highlighted. Mechanical Engineering students (generally in the United States and, specifically, at Kettering University) typically receive their first full instruction on this topic in their intermediate-level Calculus classes. In this course the theory behind the dot product is developed from the perspective of vectors and vector multiplication. In their mathematics class, tools are developed to facilitate its implementation and, while the focus here is often applications of the topic, discussions and investigations also focus on developing the mathematical properties of the concept. Students are often next exposed to the dot product in their introductory Physics classes, primarily as a method for determining the component of a vector but they can also, depending on the level of instruction, be exposed to the dot product within the concept of flux if Gauss’ Law is discussed in the second semester of their introductory Physics class. In some circumstances, the instruction on the dot product in the Physics class precedes its instruction in the Calculus class.

In the representative mechanical engineering program of study, the dot-product concept is next encountered in the student’s Statics class where it is used almost exclusively to determine components of a vector. In some circumstances, the dot product element of selecting components is de-emphasized in favor of a “common sense” or “practical” approach using trigonometry. A similar circumstance exists in Dynamics, the follow-up course to Statics. Depending on a student’s course load and an institution’s curricular structure, the dot product may not re-appear until Thermodynamics where it is used exclusively to define a flux through a surface. This concept, however, occupies a small fraction of the intermediate-level course and, as such, it is a high-order topic that receives a low exposure. For the dot product, the emphasis in this course is on the motivation of conservation of energy from a control volume perspective but which is almost exclusively applied in a single-input/single-output form. In this form the dot product is implicitly applied and simplifying “rules” (inflows = negative, outflows = positive)
replace explicit assessment of the dot product. Within this class, the dot product is therefore not essential for understanding of the simplest or “practical” cases but is important in the generalized motivation of the topic and in advanced implementations of the subject. The dot product then reappears in a more advanced and central manner in their Fluid Mechanics class as both a flux concept and for determining components of vectors. It is also, however, used as a tool for re-writing equations of motion utilizing the del operator and other elements from vector calculus. In Fluid Mechanics concepts like the divergence also play a central role in the differential analysis subsections of the course.

Research conducted in this study seeks to assess student performance along these trajectories both to generate a snapshot of student performance and to examine issues related to student growth and repair. The approach used in this research is similar to the rule-space model developed by Tatsuoka\textsuperscript{1,2}. In Tatsuoka research the rule-space model is used to analyze student capability with content and to address whether students have or are missing key pieces of knowledge or cognitive processing/thinking skills to accomplish basic problems. Within the rule-space approach classes of student errors are identified by analyzing student efforts in solving problems and broad classifications are used to characterize a level of performance. This rule-space method has been applied to both language skills\textsuperscript{3} as well as mathematics skills\textsuperscript{1}. In this research, however, this is referred to as “modes of error analysis” since this work differs from Tatsuoka by addressing the student body rather than an individual’s level of mastery at ‘sub-skills’.

As a tool for evaluating student learning within a given curriculum, the mode of failure analysis and the content/skill trajectories proposed are invariably compared to other existing programs and criteria: most notably ABET (Accreditation Board for Engineering and Technology) and SUCCEED (Southeastern University and College Coalition for Engineering Education). Within the SUCCEED structure two methods of curriculum analysis are provided which take structured and hierarchical views of the curriculum: the Knowledge/Skills Method and the Augmented Syllabus Method. The Knowledge/Skills Method\textsuperscript{4} encourages the establishment of a sequential process in which the “big picture” is first sketched in broad strokes and then filled in by detail to look at where in the curriculum specific knowledge and skills are addressed. Within this structure, skills are defined as learned capacities (as opposed to content) that are fundamental to engineering and cover broad areas of application (engineering design skills, engineering control skills, problem solving skills, organizational skills). With the Knowledge/Skills Method, a matrix for a given curriculum is generated to track student growth and development where the rows consist of the courses offered and the skills or knowledge areas are represented in the columns. These matrices are used to analyze a curriculum to determine how requirements are being met, how the knowledge elements and skills are integrated, or how groups of courses are “time-phased”.

The second method of evaluating a department’s curriculum proposed by the SUCCEED coalition is the Augmented Syllabus Method\textsuperscript{5}. The Augmented Syllabus Method addresses the curriculum at the degree to which goals are being met at a course level. It focuses on topical coverage and maps the path to mastery of a student by defining the level of accomplishment necessary at each stage, the pre-requisite knowledge for each topic and the anticipated use of the topics studied. One tool for use in the Augmented Syllabus method proposed by the
SUCCEED coalition is the use of temporal displays of a curriculum, also referred to as phase diagrams. These phase diagrams seek to map a pathway towards mastery and to identify a progression in the depth of instruction on a topic. This phase diagram approach ascribes a four level mastery scale to measure the expected level of student understanding for various topics. In evaluating a department’s curriculum, levels of implementation progress from “no exposure” to qualitative exposure (e.g. concepts) and then to quantitative exposure (e.g. homework problem) and finally to use in design or analysis. For example, a phase diagram for the concept of conservation principles published for Clemson University indicates that students will be exposed to the concept (Level 1) in their Sophomore year and then brought to Level 2 by the end of that year.

ABET’s approach to curriculum mapping and assessment is both more proscribed than the SUCCEED outcomes and, at the same time, more dependent on the institution for the manner in which student learning is assessed. The approach contained in the ABET, also differs from Project SUCCEED in the degree to which specific content are included for analysis. Within ABET the program outcomes (PO’s) set forth in Criterion 3 address eleven student objectives, (a) through (k), in which an institution must demonstrate that students have attained success. While ABET requires that there must be an assessment and evaluation process that periodically documents and demonstrates the degree to which the program outcomes are attained, within the ABET structure the methods by which these are addressed is up to each department to develop and pursue. For example, Wagner, et al state their hesitancy in using the FE exam (and their arguments against) as a standardized means of assessment and describe their own development of a multiple-choice test method for assessing ABET outcomes. Wang, et al address the application of the ABET assessment to their own classes by explaining how they used student surveys, quiz scores and test scores to evaluate ABET Outcomes (a) – (k). They use this data to conclude that student proficiency with mathematics is central to success in Dynamics. To evaluate their students the researchers looked at scores on individual problems and assessed math proficiency.

The concept/skill trajectory approach used in this research seeks to identify the dependence of future concepts on specific skills and content development gained earlier and to assess student performance in representative categories using a bug distribution model similar to Tatsuoka. In this respect, the level of detail is considerably higher than that accomplished by either the Augmented Syllabus or Knowledge/Skills method from SUCCEED. The model being developed within this research also differs from these methods because it does not presuppose that a student’s first exposure to the material leads directly to the advancement to the next level of mastery and the mode of error analysis seeks to expose. Finally, the content and skill trajectory approach differs from the ABET approach because it is designed to operate from an assessment standpoint as opposed to evaluation. With each skill or content topic addressed the goal is the identification as to common reasons for students not achieving success on that given topic and the identification of locations within the curriculum where transfer breaks down.

Methodology

The research being conducted under the NSF’s Course, Curriculum and Laboratory Improvement Program (CCLI) consists of “pre-test” assessment at the start of a term of student
capability in pre-requisite knowledge and skills (integration, differentiation, dot product, equilibrium conditions, etc.) across the curriculum. Student responses to these assessment questions are analyzed to determine the approach which each student took in addressing the problem and to identify aspects of their thought processes: this is especially important in those problems where the students answered the assessment questions incorrectly. The different methods by which students approach these problems are then categorized and a catalog of the most common error modes is developed.

The pre-test assessments are 20 minutes long and contain three questions, one each from the following categories: mathematics skill, general science skill, and engineering knowledge/skill. Questions used are taken from textbook example problems using resources that are familiar to the students and are chosen for their relative simplicity and to address targeted or potential error modes. During a given semester’s distribution, the same pre-test is delivered to each class regardless of the class location on the subject trajectory. Calculators are permitted for some problems and not for others, depending on the problem. Those without calculators during the pre-test are not provided with calculators and are instructed in the administration protocols and on the pre-test instructions to indicate the absence of a calculator and to “set up the problem to the best of their ability”.

Since the summer semester of 2009, pre-tests have been administered to 979 students over the course of six semesters (33 classes). The total number of student participants from each subject, as well as that subject’s location in the representative program of study, are indicated in Figure 2. Pre-tests have been distributed to Calculus III, Statics, Materials, Dynamics, Thermodynamics, Fluid Mechanics, and Heat Transfer classes. Factors influencing the inclusion of a course within are pre-test administrator course loads, schedule conflicts and instructor permission. Pre-tests are administered during the first 20 minutes of the first class for each subject to minimize potential exposure to instruction/re-instruction on the content. Students who have taken the pre-test in a previous course during a semester are instructed to indicate their prior participation and not complete a second assessment.

Figure 2: Aggregate Student Participation Numbers by Course and Semester in Representative Program of Study (n = 979)

Questions asked during the pre-tests are the same in each distribution regardless of the level of the class. The content and format of these questions have varied from semester to semester depending on topic of interest (Figure 3) or to address issues with previously observed modes, but changes in questions from one term to the next have been kept to a minimum. Questions have covered topics such as the dot product, partial derivatives, units conversion and static equilibrium.
Identification and faculty attribution of the error modes are conducted independently by the PI and co-I after each distribution and a comparison was made to standardize the observed modes. Assessment of student work was done on a problem by problem basis and new modes were allowed to form out of student responses. Modes were identified based on student work shown and evidence of intent in the problem solving. Comparisons were made after each term and conflicts in evaluation of student work were discussed and realigned.

For comparison of student capability from term-to-term or from class-to-class, a severity index was developed to permit correlations and to examine relationships. The initial severity index examined was a simple four level index: correct, minor errors, conceptual errors and major errors. The minor errors were often attributed to mathematical algebra ‘slips’ and major errors were often identified by problems which were solved using methods completely antithetical to the problem statement (integration where there should have been differentiation, etc.). Interviews with 19 students were conducted in the fall term of 2010 to assess accuracy of assigned modes and to add insight into level of awareness of student gaps as well as potential reasons for student errors.

Results: Error Mode Analysis

Individual class error mode assessments have been assessed for all of the classes studied and selected data is presented here to illustrate the trends and methods used in this study. For example, data from the summer 2009 distribution (Appendix A) of the mode of error analysis to a Fluid Mechanics class (MECH-322) indicates some very clear trends regarding student baseline performance. On the partial differentiation problem (Problem 1, Appendix A) there were only two dominant modes of error (Figure 4) relating to how students addressed the ancillary terms in the problem: those terms which were to be considered as constant with respect to the derivative operation. The most dominant mode of error (23%) was where students maintained the ancillary (non-x) terms in their solution: in other words, the ancillary terms were treated as a constant and were not operated on by the differentiation.
In the dot product problem (Problem 2, Appendix A) the situation was more complicated (Figure 5). Only 5% of the students taking the pre-test answered the problem correctly. The dominant mode of error, “Vector Output”, showed students completing the mechanics of the dot product correctly but reporting the answer as a vector. A sizable percentage of the students surveyed did not attempt the problem or produced errors that could not be grouped in to any of the recognized modes or placed in their own mode. The remaining students, 38% of the sample, used one of two methods determined by the reviewers to be associated with the cross product to address the problem.

Assessment conflicts for the summer term of 2009 the observed modes of error were also studied. Assessments were observed to be closely aligned, with 18 conflicts observed in the dot product problem (41%) and 4 conflicts observed in the partial differentiation problem (9.3%).
For the dot product problem, the majority of the conflicts (9) arose because of similarities between reviewer descriptions of two of the modes of error and were attributed to differing levels of resolution. Of the remaining 9 conflicts associated with the dot product problem, only two discrepancies (4.6%) did not relate to the ‘miscellaneous errors’ category and were resolved under further review. In the Fall term of 2009 the failure modes were more clearly identified and only two conflicts (0.55%) arose which needed resolving and this trend has been continued in subsequent terms.

Throughout each of the terms studied, student performance has yielded similar results. The dominant modes of error have persisted with isolated or unique modes of error arising sporadically throughout the study. Examination of relationships between student responses on the science skill questions (units conversion) and performance on the mathematics skill questions (partial derivatives) for the Spring 2010 data been addressed using the simple four-point severity index. Figure 6 illustrates the assignment of severity to each of the observed modes that were arrived at by the PI and co-I. Each mode in the figure is accompanied by a brief description of the error and an indication of the total impact of that mode on the entire distribution (all classes). The severity index is +2 for correct, +1 for algebra errors, -1 for conceptual and process errors, and -2 for severe errors.

<table>
<thead>
<tr>
<th>Student Responses</th>
<th>Percent Observed</th>
<th>Error Mode</th>
<th>Severity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>95</td>
<td>50.8%</td>
<td>Correct</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Correct Differentiation, Minor Algebra Errors</strong></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.1%</td>
<td>Algebra Errors with Expansion</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>2.7%</td>
<td>Chain Rule Mechanics Errors</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Conceptual Issues</strong></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>9.6%</td>
<td>Issues with Mixed-Term Only</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>2.7%</td>
<td>Chain Rule Mechanics Errors</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Inconsistent Issues</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.6%</td>
<td>Correct Treatment of X-Terms —— Non-X Terms Maintained</td>
<td>-2</td>
</tr>
<tr>
<td>10</td>
<td>5.3%</td>
<td>Inconsistent Treatment --- kept some non-X or didn't differentiate some x</td>
<td>-1</td>
</tr>
<tr>
<td>6</td>
<td>3.2%</td>
<td>Multiple Errors on Derivatives</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>4.3%</td>
<td>Derivatives without Expansion</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>1.6%</td>
<td>Applied Derivatives on All Terms</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Serious Knowledge Error</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.6%</td>
<td>Substitution with Division Attempt</td>
<td>-2</td>
</tr>
<tr>
<td>5</td>
<td>2.7%</td>
<td>Integration Approach</td>
<td>-2</td>
</tr>
<tr>
<td>6</td>
<td>3.2%</td>
<td>Attempted Division</td>
<td>-2</td>
</tr>
<tr>
<td>4</td>
<td>2.1%</td>
<td>Unknown Effort</td>
<td>-2</td>
</tr>
<tr>
<td>5</td>
<td>2.7%</td>
<td>Did Not Attempt</td>
<td>-2</td>
</tr>
</tbody>
</table>

Figure 6: Severity Assessment of Modes for Partial Differentiation Question (Q1) from Spring 2010 Pre-Test Distribution (all respondents)

With this initial four-point severity index, results show a strong relationship between the severities of the errors made on each problem. Figure 7 indicates the relationships between student errors on the mathematic skill and science skill questions for both Statics (SO I) and Fluids/Thermo (JR I and JR II) courses. With this scale the data was strongly correlated for each grouping with a Pearson correlation of 0.93 and 0.92, respectively.
Conclusions

Research conducted under the NSF’s Course, Curriculum and Laboratory Improvement Program (CCLI) consists of pre-instruction assessment of student capability with support topics and skills (integration, differentiation, dot product, equilibrium conditions, etc.) across the curriculum. Three question pre-tests focusing of mathematics, science and basic engineering pre-requisite knowledge were given at the beginning of each class in the study for a total of six semesters. Student responses to these assessment questions are analyzed to determine the approach which each student took in addressing the problem and to identify aspects of their thought processes. The different methods by which students approached these problems were categorized and a catalog of the most common error modes was developed.

Results indicate that there are major gaps in student recall of pre-requisite knowledge and that these gaps persist throughout their education. A mode of error analysis indicates the degree to which students misapply concepts or entangle methods and concepts as they progress throughout a program of study in Mechanical Engineering. With a simple four-level severity index the relationships between student errors on both the mathematics skill and science skill questions show a strong correlation and persistence throughout the program of study. Current efforts are directed at refining the severity index and developing alternate tools for assessing severity as well as continuing the analysis of wealth of data on student performance.

Coupled with content and skills trajectory research, this mode of error analysis seeks to identify and categorize the methods of content and skills failure within the trajectory structure. Unlike other efforts to evaluate student learning this project looks at failure of learning points rather than success. This analysis will be used to identify where learning breaks down or where gaps exist in student prerequisite knowledge. In this way an institutional profile of student competency in selected background skills across the curriculum can be developed. Student responses on pre-assessments can also be compared both within a class and as students progress through the curriculum over the two years of the study. Aggregate data from this longitudinal study will permit assessment of student growth in these areas and it is anticipated that insights will be gained in the solidification of student knowledge as they progress through their academic
career. Long term goals related to this research are the development of tools that will permit an institution to: profile student competency in key content and skills, identify and develop targeted instruction or other remedial efforts to improve student learning, and to have a measure for demonstrating the effectiveness of their remedial/re-instruction efforts.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 0837310.

Bibliography


Appendix A: Mechanical Engineering Pre-Test (Summer 2009 Distribution)

Pre-Test: MECH-322 Summer 2009

You have 20 minutes: No calculators for the first two problems
No talking: Show all of your work

1) \( g(x, y, t) = 3x^2 + 2xt + 3y^2 + t^3 \)

Find \( \frac{\partial g}{\partial x} \)

\( \vec{A} = (4x)\hat{i} + (3xy)\hat{j} + (5x)\hat{k} \)
\( \vec{B} = (3)\hat{i} + (2y)\hat{j} + (7z)\hat{k} \)

Find \( \vec{A} \cdot \vec{B} \)

Please indicate by circling YES or NO whether you have a calculator

<table>
<thead>
<tr>
<th>Calculator?</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
</table>

If you do not have a calculator then set up the following problems as best you can.

3) The column is used to support the floor which exerts a force on 3,000 lb on the top of the column. The effect of the soil pressure along the column’s side is distributed as shown. Replace this loading by an equivalent resultant force and specify where it acts along the column, measured from its base A.
Pre-Test - Spring 2010

You have 15 minutes. No calculators for the first problem. No talking during the pre-test. Please show all of your work.

If you have taken a pre-test for another class during the Spring of 2010 then write your name in the space provided and leave the remainder of the pre-test blank.

1) For the function \( g(x, y, z, t) \) given below, find \( \frac{\partial g}{\partial x} \):

\[
g(x, y, z, t) = 3x^2 + (x^2 + y^3)^2 + 3yt^2 + \pi
\]

find \( \frac{\partial g}{\partial x} \)

Indicate by circling YES or NO whether or not you have a calculator for the pre-test:

| Calculator? | YES | NO |
---|---|---|

If you do not have a calculator, set-up problems #2 and #3 as best you can.

2) A tank has a volume of 750 cm\(^3\). Determine the volume of the tank in Liters.

\[
V = 750 \text{ cm}^3
\]

\[
1 \text{ L} = 0.001 \text{ m}^3
\]

3) A pipe is held in place by a vise. If the bolt exerts a force of 50 lb on the pipe in the direction shown, determine the forces \( F_A \) and \( F_B \) that the smooth contacts at A and B exert on the pipe.