A Correlative Analysis of Course Grades as Related to Curricular Prerequisite Structure and Inter-Class Topic Dependencies

Johnathon Nash, Laura E. Boucheron, and Steven J. Stochaj
Klipsch School of Electrical & Computer Engineering
New Mexico State University
Las Cruces, NM 88003
Email: {nashjohn, lboucher, sstochaj}@nmsu.edu

***Abstract*—This Research-to-Practice Work in Progress paper presents an analysis of student performance in courses in light of the prerequisite structure of those courses and the degree of topic-level dependencies between the courses. Prerequisite relationships, especially long sequences of relationships, are the largest contributing factor to curricular complexity, a measure of the ease of progression through a degree. While the effect of prerequisites on curricular complexity can be quantitatively measured, the relationship of student performance between the courses in a course sequence and the topic-level dependency between the courses have not been studied. Here we study the correlation in course grades in an electrical engineering curriculum. Additionally, we quantify the topic-level dependencies between courses by building a hierarchy of topics through the electrical engineering curriculum. These preliminary analyses indicate that there are likely other strong factors that affect class grades beyond specific course competencies and that the prerequisite structure of the curriculum may not necessarily accurately reflect a relationship in terms of course competencies.**

I. Introduction

We present an analysis of student performance in courses in light of prerequisite structure and the degree of topic-level dependencies between the courses. This work builds on the concept of curricular complexity [1] which quantifies the ease of progression through a degree and highlights courses that may be bottlenecks to progression. Degree programs with high curricular complexity raise concern about impediments to progression and increased time-to-graduation [2]. Prerequisite relationships, especially long sequences of relationships, are the largest contributing factor to curricular complexity.

Ostensibly, prerequisites exist because concepts in prerequisite courses are necessary for understanding concepts in subsequent courses. Many prerequisites, however, may exist for historical reasons that are no longer valid (e.g., redesign of material no longer relies on prerequisite topics) or for gatekeeper prerequisites (e.g., requiring all sophomore courses prior to junior courses to "keep students on track"). Unnecessary prerequisites add significant complexity to the curriculum, impeding student progress through the degree, but do not necessarily reflect actual topic dependencies between courses.

While the effect of prerequisites on curricular complexity can be measured [1], [2], the relationship of student performance between the courses and the topic-level dependency between courses are less frequently studied. These are two main contributions we discuss in this paper. First, we study correlation of course grades in an electrical engineering (EE) curriculum. Second, we quantify topic-level dependencies between courses by building a topic graph for the curriculum and counting topic dependencies between courses.

This analysis will provide insight into topics associated with critical prerequisite relationships. This could be used to redesign our EE courses to mitigate the long course sequences in the curriculum. For example, universities have had success implementing new models of engineering math to help with the calculus bottleneck [3]. Additionally, resources (e.g., teaching assistants or supplemental instruction) can be strategically invested in courses that support the widest reaching topics. Strategic investment of resources can positively affect student performance and curricular design can have a significant impact on time-to-degree by reducing curricular complexity.

II. Background and Theoretical Framework

The theoretical framework of this work is motivated by curricular analytics [4] which provides quantitative measures of the importance of courses to curricula and the complexity of curricula based on prerequisite structure of courses [1]. Curricular redesign, informed by measures such as curricular complexity, may provide efficient curricula in which students have more flexibility in degree progression [2], [5]. The measure of course cruciality [1] provides quantitative analysis of curricular bottlenecks, indicating courses that could have the largest impact on curricular efficiency and degree progression.

Course cruciality and curricular complexity do not, however, provide guidance on which *concepts* within the most crucial courses should be targeted for reconfiguration. In this paper, for conciseness, we use the term "postrequisite" to denote a course which depends on a prerequisite course. We propose a data-driven approach to study which concepts (or competencies or learning objectives) are necessary for competency in postrequisite courses. We first study the correlation between course grades in an a core EE curriculum. We expect that, if a postrequisite course has a strong dependence on a prerequisite

TABLE I
CORE BSEE COURSES WITH NUMBER OF TOPICS AND RELATIONSHIPS BY PRE-/CO-REQUISITES, TOPIC DEPENDENCIES, AND GRADE CORRELATION.

Course[a]	**# Topics**	**Pre-/Co-requisites**[b]	**# Topic Dependencies**[c]	**Grade Correlation**[d]
(1) DC Circuits & Digital Logic	83	Trigonometry & Precalculus[e]	-	-
(2) Calculus I	50	Trigonometry & Precalculus[e]	-	-
(3) Embedded Programming	46	(1) DC Circuits & Digital Logic	8	0.63
(4) Calculus II	26	(2) Calculus I	185	0.69
(5) Computer Architecture	44	(1) DC Circuits & Digital Logic	28	0.56
		(3) Embedded Programming (C)	0	0.65
		Trigonometry & Precalculus[e]	-	-
(6) Linear Algebra, Prob., & Stats	18	(3) Embedded Programming	0	0.59
		(4) Calculus II	14	0.62
(7) Engineering Physics I	45	(2) Calculus I	184	0.48
(8) AC Circuit Analysis	55	(1) DC Circuits & Digital Logic	584	0.32
		(4) Calculus II	58	0.36
		(11) Engineering Physics II (C)	0	0.64
(9) Differential Equations	20	(4) Calculus II	105	0.54
(10) Multivariate & Vector Calculus	49	(3) Embedded Programming	0	0.55
		(4) Calculus II	103	0.54
(11) Engineering Physics II	55	(7) Engineering Physics I	305	0.36
		(4) Calculus II	52	0.59
(12) Signals & Systems I	102	(6) Linear Algebra, Prob. & Stats	0	0.73
		(8) AC Circuit Analysis	144	0.55
		(9) Differential Equations (C)	288	0.58
(13) Electromagnetics & Optics	58	(8) AC Circuit Analysis	278	0.29
		(10) Multivariate & Vector Calculus	862	0.67
		(11) Engineering Physics II	0	0.56
(14) Signals & Systems II	81	(12) Signals & Systems I	2180	0.59
		(9) Differential Equations	480	0.50
(15) Semiconductors & Electronics	65	(8) AC Circuit Analysis	317	0.33

[a] A reference number included here is used in subsequent figures [b] Co-requisite courses are denoted with a (C) [c] The number of times a topic from a prerequisite course is referenced; this does not distinguish between multiple topic references versus multiple references to the same topic [d] All correlations reported here are significant to $p < 0.05$ [e] Competency assumed prior to the start of the BSEE

course, student grades should be more correlated between those courses. Related work has analyzed prerequisite competency [6]–[8] or course grades [9], [10] in specific course sequences, but has not focused on analysis of a curriculum.

We next study dependencies between pre- and post-requisite courses at the topic level by modeling the course topic structure as a graph, similar to the approaches in [11] which uses a curriculum graph to define course coverage and course interdependence and [12] which focuses on visualization of a curriculum graph. Different from [11], [12], we use the topic graph to analyze dependencies between courses in terms of number of topic dependencies. This is an important first step toward quantifying the dependency between courses and to better understand student performance across the curriculum.

III. DATA

A. BSEE Course and Topic Relationships

1) Course Pre- and Co-requisites: We focus our analysis on core EE, math, and science courses in our Bachelor of Science in Electrical Engineering (BSEE) degree as outlined in Table I. Since all BSEE students must complete the core, we can analyze student performance in the core courses across our entire student population. We outline the pre- and co-requisites for the core courses also in Table I.

2) Topic-Level Dependencies Between Courses: We define a topic-level relationship between courses; in future work this relationship will be defined in terms of learning objectives. Since not all faculty use learning objectives, we perform this preliminary analysis using topics as proxies for learning objectives and assuming that topics can be reformulated as learning objectives. We used sources available for all courses and from which topics can be easily mined, namely syllabi, course schedules, textbook sections, and assignments, to define topics for each course. Additional sources such as course notes or consultation with faculty will be considered in future work. We list the number of topics in each core course in Table I. We note a very wide spread in the number of topics per course with a minimum of 18 and maximum of 102. This may warrant closer inspection in future work to assure that topics are neither too broad nor too specific. Development of learning objectives and a course map for each course may help in this regard.

Next, we analyze each topic for topic(s) on which it relies. We define a mapping only to the immediately preceding topic since a hierarchy is defined by the graph structure. All authors and those noted in the Acknowledgements contributed to definition of topic dependencies; we thus leveraged the knowledge of BSEE faculty and graduates, including graduates of this specific program. Future work will involve consultation with faculty instructors to verify and adjust the topic dependencies.

Course topics are used to define a directed acyclic graph (DAG) G. The nodes V_i of G are topics and the directed edges $E_{ij} = (V_i, V_j)$ denote that topic V_i feeds into (is necessary for) topic V_j. We implement and analyze the DAG of course topics using the NetworkX network analysis library in python [13].

TABLE II
LETTER GRADE TO NUMERICAL GRADE MAPPING.

Letter	Numerical	Letter	Numerical
A+	4.00	C+	2.33
A	4.00	C	2.00
A-	3.67	C-	1.67
B+	3.33	D+	1.33
B	3.00	D	1.00
B-	2.67	D-	0.67
		F	0.00

B. Student Data

The use of student data was approved by our university's Institutional Review Board (IRB) and all data was de-identified as described below. We downloaded data for all declared EE majors from Fall 2016 to Spring 2020 from our university's records management system. These data provide all classes and grades enrolled in by EE students. The data are combined into a single comma separated variable file. Once collated, all personally identifiable information is removed and a unique alphanumeric string is assigned to each student using the SHA512 hashing algorithm [14] applied to the students' ID numbers. The deidentified dataset is used as the basis for all subsequent grade correlation analysis, which is conducted using the pandas data processing library in python [15].

From these data we analyze grades for courses listed in Table I for the 11 terms in our dataset. These data include 531 students and 3884 individual grades. We convert the assigned letter grade to a numerical grade using the mapping in Table II. In the case that a student re-takes a course, we use the average grade. Similarly, for courses where course and lab grades are assigned separately, we average the course and lab grades. This averaging of grades results in a total of 2845 grades which are further analyzed in Section IV-A; similar results are achieved when considering only the highest grade (results not shown).

IV. Results

A. Course Grade Correlations

We compute the Pearson correlation coefficient between grades in each pair of EE core courses; this includes all pairs, not just those with a pre- or co-requisite relationship. Correlation coefficients are displayed as a heatmap in Fig. 1. Similar results are obtained with the Spearman correlation coefficient (results not shown).

We note that correlations in Fig 1 are all in the range $[0.29, 0.77]$. This indicates, not surprisingly, that there are other factors that can explain grade correlation other than course competencies, e.g., general academic confidence of the student, difficulty of the semester load, or the grade distributions of different instructors. Interestingly, the grade correlation for pre- and post-requisite course pairs (see Table I) spans $[0.29, 0.73]$ which is almost entire range seen in Fig. 1.

In reference to Table I, we find that 19 of the 24 correlations are in the range $[0.48, 0.73]$ while the remaining 5 are in the range $[0.29, 0.36]$. The correlation of 0.29, 0.33 between (8) AC Circuit Analysis and (13) Electromagnetics & Optics, (15)

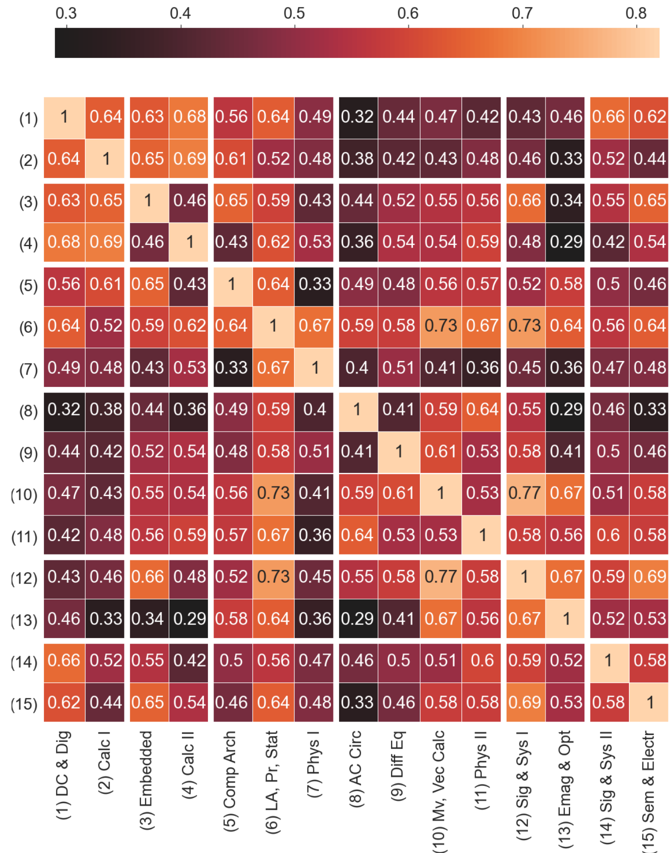

Fig. 1. Correlation between grades in BSEE core courses. All correlations are significant to $p < 0.05$ except between (2) Calculus I and (13) Electromagnetics & Optics which has $p = 0.08$. Courses are ordered from top to bottom (and left to right) according to the term typically taken and the thick white lines separate terms. Thus, courses toward the upper left are taken earlier in the degree, and courses toward the lower right are taken later.

Semiconductors & Electronics, respectively, is surprising since both postrequisite courses appear to rely heavily on concepts from the prerequisite (see Section IV-B). We expect that this is likely due to instructor differences in grade distribution and reinforcement of necessary prerequisite concepts within the postrequisites. Other small correlations are likely due to similar factors and possibly more tenuous relations between content, e.g., that (8) AC Circuit Analysis relies on, albeit heavily, only a few specific topics from (4) Calculus II.

On the other hand, the largest correlation of 0.73 between (6) Linear Algebra, Probability, & Statistics and (12) Signals & Systems I is in line with anecdotal evidence that (6) is one of the most problematic bottleneck courses in the curriculum. We hypothesize that the mathematical rigor necessary in both (6) and (12) contribute to the large grade correlation. The highest correlation of 0.77 between (10) Multivariate & Vector Calculus and (12) (which do not have a pre-requisite relationship) is additional evidence toward this hypothesis.

These preliminary results present intriguing questions for future work. Specifically, we will seek to correct for the effect of instructor and other factors that may contribute to grade correlation. These preliminary results seem to indicate, however, that academic performance is "recoverable" in the sense that students are not necessarily destined to receive a corresponding grade in a postrequisite course.

B. Course Topic Dependencies

We analyze the DAG G on a per-node basis and generate a histogram of the total number of times class c uses a topic from class r. This takes into account the hierarchy of relationships in G such that topic dependencies are followed back to the root course. Note, however, that this analysis cannot distinguish between a single topic being referenced multiple times and multiple topics being referenced a single time.

A heatmap of topic dependencies is shown in Fig. 2, where a logarithmically spaced colorbar is used due to the large dynamic range. We note that courses earlier in the curriculum have a higher count indicating that the topics in those classes are propagating through the curriculum. Additionally, courses in which there is the expectation of a large dependency, e.g., (12) Signals & Systems I and (14) Signals & Systems II, do demonstrate a strong dependency. We also note some surprising results, including zero references from postrequisite courses to prerequisite courses (see also Table I). These zero references between pre- and post-requisites may not be surprising to the faculty teaching the courses, but this analysis provides quantitative evidence toward the need to reconsider the prerequisite structure of the curriculum.

Returning to some of the specific cases discussed in Section IV-A, (13) Electromagnetics & Optics references topics in (8) AC Circuit Analysis 278 times, commensurate with the 300 self references. Similarly, (15) Semiconductors & Electronics references topics in (8) 317 times, commensurate with the 285 self references. This provides additional quantitative evidence toward the reliance of (13) and (15) on topics from (8). Surprisingly, we find that there are zero references to concepts in (6) Linear Algebra, Probability, & Statistics from (12) Signals & Systems I, despite that pair having one of the largest grade correlations and a prerequisite relationship. This necessitates the questions of whether (6) should be a prerequisite to (12) and why grades in those courses are so highly correlated.

The results in Fig. 2 are preliminary and need to be refined as noted earlier, but provide additional intriguing possibilities for future work. First, the results in Fig. 2 do not take into account elective courses that may depend on the core; this is certainly the case for (5) Computer Architecture. Second, and of critical importance, future work will consider not only of the number of topic dependencies, but *which* specific topics are most critical in postrequisite courses.

V. Conclusions and Future Work

While preliminary, the analyses we present here provide intriguing possibilities for future analysis of a curriculum at a more granular level. We find less of a correlation in performance (grades) in pre- and post-requisite courses than expected, indicating that there may be other factors of more

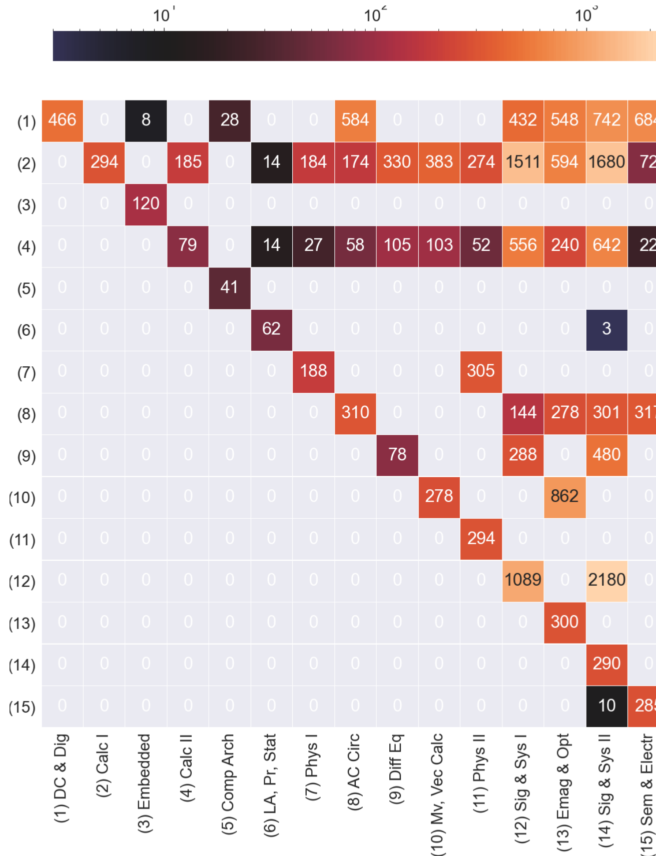

Fig. 2. Relationship between topics in BSEE core courses as a heatmap with a logarithmically scaled colorbar. The (r, c) location in the heatmap tallies the total number of times class c uses a competency from class r. Courses are ordered from top to bottom (and left to right) according to the term typically taken. Thus, courses toward the upper left are taken earlier in the degree, and courses toward the lower right are taken later in the degree.

importance to student success such as the pedagogical approach of the instructor. We find some interesting results in the topic-level dependencies between the core courses which warrant further investigation, particularly in those courses with a prerequisite relationship but zero or small topic dependency.

In addition to the future work discussed throughout Section IV, we will map course assignments to course topics. This will allow us to analyze student performance on course topics (or learning objectives). The ability to analyze not only a hierarchy of topics in the BSEE curriculum, but also the effects of that hierarchy on student performance will provide a powerful tool for curricular redesign and proactive advising toward student success.

Acknowledgments

The authors acknowledge US Department of Education Grant P120A170064 which supported this work. The authors also gratefully acknowledge the efforts of Anthony DeRocchis, Raksha Gala, Mark Garcia, Mario Renteria-Piñon, and David Rodriguez in formulating the course topic mapping.

REFERENCES

[1] A. Slim, J. Kozlick, G. L. Heileman, and C. T. Abdallah, "The complexity of university curricula according to course cruciality," in *2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems*. IEEE, 2014, pp. 242–248.

[2] J. Wigdahl, G. L. Heileman, A. Slim, and C. T. Abdallah, "Curricular efficiency: What role does it play in student success," in *2014 ASEE Annual Conference & Exposition, Indianapolis, Indiana*, 2014.

[3] N. W. Klingbeil and A. Bourne, "The Wright State model for engineering mathematics education: A longitudinal study of student perception data," in *121st ASEE Annual Conference & Exposition*, 2014, pp. 1–11.

[4] G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman, "Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations," *arXiv preprint arXiv:1811.09676*, 2018.

[5] A. Slim, G. L. Heileman, E. Lopez, H. Al Yusuf, and C. T. Abdallah, "Crucial based curriculum balancing: A new model for curriculum balancing," in *2015 10th International Conference on Computer Science & Education (ICCSE)*. IEEE, 2015, pp. 243–248.

[6] S. Valstar, W. G. Griswold, and L. Porter, "The relationship between prerequisite proficiency and student performance in an upper-division computing course," in *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 2019, pp. 794–800.

[7] N. A. Smith, R. Y. Myose, S. J. Raza, and E. Rollins, "Correlating mechanics of materials student performance with scores of a test over prerequisite material," in *ASEE North Midwest Section Annual Conference*. South Dakota State University, 2020, p. 10.

[8] S. Krause-Levy, S. Valstar, L. Porter, and W. G. Griswold, "Exploring the link between prerequisites and performance in advanced data structures," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 2020, pp. 386–392.

[9] J. Simpson and E. Fernandez, "Student performance in first year, mathematics, and physics courses: Implications for success in the study of electrical and computer engineering," in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*. IEEE, 2014, pp. 1–4.

[10] R. Wright, S. Cotner, and A. Winkel, "Minimal impact of organic chemistry prerequisite on student performance in introductory biochemistry," *CBE—Life Sciences Education*, vol. 8, no. 1, pp. 44–54, 2009.

[11] J. A. Pavlich-Mariscal, M. Curiel, and G. Chavarro, "CDIO curriculum design for computing: A graph-based approach," in *Proceedings of the 15th International CDIO Conference, Aarhus, Denmark*, 2019.

[12] T. Auvinen, J. Paavola, and J. Hartikainen, "STOPS: a graph-based study planning and curriculum development tool," in *Proceedings of the 14th Koli Calling International Conference on Computing Education Research*, 2014, pp. 25–34.

[13] A. Hagberg, D. Schult, and P. Swart, "NetworkX reference, release 2.6rc1.dev0," https://networkx.org/documentation/latest/_downloads/networkx_reference.pdf, Tech. Rep.

[14] Q. H. Dang, "Secure hash standard," Federal Inf. Process. Stds. (NIST FIPS), National Institute of Standards and Technology, Tech. Rep., 2015-08-04 2015.

[15] The pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020. [Online]. Available: https://doi.org/10.5281/zenodo.4524629