WIP: Detection of Student Misconceptions of Electrical Circuit Concepts in a Short Answer Question Using NLP

Prof. James P Becker, Montana State University, Bozeman

James Becker is a Professor of electrical and computer engineering at Montana State University. His professional interests include microwave circuits, radio frequency electronics, nanoelectronics, pedagogical research, and distance education.

Dr. Indika Kahanda, University of North Florida

Dr. Indika Kahanda is an Assistant Professor in the School of Computing at the University of North Florida, where he directs the bioinformatics, biomedical informatics and medical informatics lab. Prior to that Dr. Kahanda worked as an Assistant Professor in the Gianforte School of Computing at Montana State University. He received his Ph.D. in Computer Science from Colorado State University in 2016 in the area of Bioinformatics, a Master of Science in Computer Engineering from Purdue University in 2010, and a Bachelor of Science in Computer Engineering from University of Peradeniya, Sri Lanka in 2007.

Nazmul H. Kazi, Montana State University

Nazmul Kazi is a master's student of Computer Science at Montana State University. His research interests include the application of Artificial Intelligence, Deep Learning, Natural Language Processing, and Parallel Computing.

WIP:

Detection of Student Misconceptions of Electrical Circuit Concepts in a Short Answer Question Using Natural Language Processing

Abstract

While the use of writing exercises in gateway STEM courses that focus on solving numeric problems is not widespread, there is evidence that students could benefit from the addition of such exercises [1]. Writing exercises may be effective in both uncovering student misconceptions that are not necessarily apparent with typical computation problems, and as tools to foster conceptual change and metacognitive skill.

In this paper, pilot studies of the use of two Natural Language Processing (NLP) techniques to identify common misconceptions in the writing of students in a course on electric circuit analysis are described. Performance on the writing exercise in question has been shown to correlate with a student's performance in the course [2]. This is of particular interest as the writing exercise has been administered during the fifth class period, sufficiently early to direct additional resources to the success of students appearing to be at-risk for failing the course. Realizing an automated software solution to analyze the responses to this exercise would remove burden on instructor time and open the door to immediate and personalized feedback to the student.

The first pilot study was run to determine how successful a simplistic rule-based approach would be in identifying the most common misconceptions found in a writing exercise requiring a student to speculate on the change in the power in the elements of a resistive circuit with a change to a single resistor value. An open-source NLP rule-based matching engine within spaCy [3] was used. The corpus consisted of one hundred and eighty-five unique responses to the question. Precision, recall, and F1-score [4] were used to assess the effectiveness of the rule-based NLP pipeline in comparison to that of a subject matter expert in identifying responses exemplifying seven misconceptions. Should this NLP pipeline be used in a system in which feedback is to be given to the student, a Directed Line of Reasoning (DLR) approach [5] would be beneficial in cases in which identification of a given misconception is in doubt. Considering this pilot study employed an extremely simplistic purely lexical-level rule-based classifier, the results are very promising and suggest the planned approach of developing a highly accurate, advanced rule-based classifier encompassing lexical/syntax/semantic driven rules is viable. As a compliment to the rule-based approach, this paper also describes a pilot study of the use of BERT (Bidirectional Encoder Representations from Transformers) [6], a machine learning approach that has shown tremendous promise in short-answer grading [7].

I. Introduction

Student struggles in gateway STEM courses may arise from a variety of factors. Two commonly identified impediments to student success in such courses include inadequate "prerequisite learning and thinking skills" [8] and the inaccurate prior knowledge [9]. Learning and thinking skills fall under the terms "self-regulated learning" [10] and "metacognition" [11] and involve skills such as accurate self-assessment of knowledge ("knowledge of cognition") and the ability to formulate and follow a plan for mastery ("regulation of cognition"). For mastery of the content studied in gateway courses, students must often demonstrate an understanding of concepts for which they may come to class with inaccurate models and frameworks. For example, it has been found that students entering courses on statistics often have intuitions regarding probability and statistics that are at odds with accepted reasoning [12]. Such erroneous mental frameworks and misconceptions are often challenging to correct [13].

There are many identified misconceptions among students beginning courses on electric circuit analysis [14]-[20]. For example, it is not uncommon for a student to fail to appreciate that an electric circuit is a system and rather believe that the "flow" of current is sequential. Colloquial phrasings such as "current flow" are often used by instructors for expediency but may strengthen student misconceptions. While evidence of certain misconceptions held by students in the electric circuit domain may be found in their computations, student writings offer the most vivid insight into a student's thinking. The value of writing as a tool for uncovering a student's misconceptions has been noted in other disciplines such as the medical field [21]. Unfortunately, grading and providing feedback to students on their written work is time consuming. This burden on instructor time may be a factor why, beyond common written works such as laboratory reports, courses such as electric circuit analysis or statics and dynamics are almost exclusively computation based. The authors of this paper do not suggest eliminating computation problems in gateway STEM courses, but rather to complement such problems with conceptual writing exercises as such exercises may be the key to effecting conceptual change particularly in the case of robust misconceptions.

The remainder of this paper focuses on a describing the results of pilot studies in the use of two techniques in natural language processing (NLP) to identify misconceptions in the responses of students to a writing quiz in an introductory circuits course, EELE 201, at Montana State University. The details of the writing quiz may be found elsewhere [2],[22]. In short, the question refers to a four-element circuit as depicted in Figure 1 and asks the student to argue what will happen to the power (increase, decrease, or remain the same) of each of the circuit's four elements, when the resistance of resistor R_2 decreases. Students are told to treat all elements as ideal including the independent voltage source and to thoroughly justify their responses.

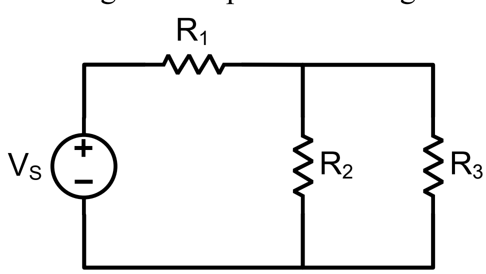

Figure 1: Circuit schematic diagram of the circuit students are to consider in a writing quiz.

This written quiz has been offered on the fifth period of class in EELE 201 and student performance on the quiz has been found to correlate with their performance in the class, making the writing exercise of significant value in identifying students likely to struggle in the course. It should be noted that prior to EELE 201, students do receive minor exposure to KVL, KCL, Ohm's Law, and series and parallel resistor combinations in a freshman course meant to introduce students to the electrical and computer engineering majors. The ultimate goal for this and other writing exercises used in EELE 201 is to embed a given writing exercise within a web-based application that without human intervention, evaluates a student's response and provides meaningful feedback to strengthen the student's conceptual understanding while promoting active metacognition.

II. The Rule-Based Matching Approach

While machine learning methods are common in natural language processing (NLP) applications [6],[7], they are most powerful when a large, labeled corpus is available and the task is simply to classify the responses, for example, to grade them. As the existing corpus of responses is modest (185) and a key goal in the current work is to identify examples of misconceptions in students' writing, we elected to begin by using a hand-crafted rule-based approach. The matching engine, "matcher", within the open-source NLP framework spaCy [3] was used. We divided the 185 responses into two groups, a training set consisting of 60% of the responses and a test set consisting of the remaining 40%. The training set was considered when generating rules to identify a given misconception. Each rule consisted of a search for an ordered set of between two and four key words, together which likely indicated a given misconception. Once the rules were established, they were applied to each response within the test set on a sentence-by-sentence basis. Measures of precision, recall, and F1-score were determined at the response level.

The *precision* of our rule-based NLP classifiers is the fraction of correct predictions of responses exemplifying a given misconception over the number of instances in which the classifier identified a response as exemplifying the considered misconception. The *recall* of a given classifier is the fraction of accurate identifications of a response exemplifying the misconception considered over the total number of responses that were tagged by the expert to exemplify the misconception. From measures of the precision and recall, the F1-score [4] was computed.

A. Sequential Misconception

A sequential misconception in terms of electric circuits is one in which it is believed that elements that are further "downstream" from a source (such as R_2 and R_3 in the example circuit of Figure 1) "receive" current after elements closer to the source (R_1 in the example circuit). With such a misconception, it is likely that a student will think that changes in R_2 have no effect on the potential difference and current associated with R_1 or V_s. The following two examples are drawn from the corpus of responses and indicative of a sequential misconception on the part of the writer.

Since R_1 is a component lying before R_2, its power absorbed should be unchanged.

The power at R_1 should be unaffected by changing R_2. R_1 is not affected as its current does not depend on R_2.

In this work, examples such as the first that refer to the relative locations of R_1 and R_2 to argue that the power in R_1 will not change with a decrease in the value of R_2 are termed "explicit" examples of the sequential misconception. Such examples are easily identified using rules looking for two-word combinations such as "before R_1" and "after R_2". Examples such as the second appear to imply the sequential misconception and so are termed "implied" examples. Extracting such examples typically required a three-word combination with the words separated by other "tokens". For example, the second example provided could be extracted using a rule looking for the word sequence "R_1 not affected" but allow another word or words to exist between "R1", "not", and "affected". Tokenization is one of the first steps in an NLP pipeline and often includes more complex tasks such as stemming or lemmatizaton [23]. In this work a basic tokenization scheme was followed, namely, separating words by spaces and punctuation marks; in some instances the lemma (i.e. base form) of a given word was invoked to make a given rule more widely applicable. While extracting explicit examples of the sequential misconception was straightforward, attempting to identify the implied cases of the sequential misconception resulted in some false positives. In this case, a false positive is a response identified by the rule-based algorithm as a sequential misconception that was not tagged as a sequential misconception by the content expert. A set of rules were developed considering the training set; the rules were then applied to the test set. As identified by the content expert, approximately 16% of the 185 responses considered included evidence of either explicit or implicit sequential thinking.

The precision, recall, and F1-score were found to be 0.632, 1, and 0.77, respectively. The value of unity for the recall indicates that the rules were sufficient to extract <u>all</u> instances in the unseen test set tagged by the content expert as exemplifying either an explicit or implicit sequential misconception. Of the seven false positives, four exemplified another misconception, what in this work is termed a "localized" misconception. Students exemplifying this misconception express the belief that only quantities associated with R_2, such as its current or power would change. All three of the remaining false positives indicated that they believed the power associated with R_1 either increased or decreased, conclusions that are in opposition to the sequential misconception. As described in [22], a web-based application for the writing question has been created though it has yet to be powered with NLP. One of the features of the existing application is a drop-down selection for each of the four circuit elements for which the user must indicate whether their response supports that the power associated with a given element *increases*, *decreases*, or *stays the same*. The original intent of including the drop-down was simply to remind students they were to address the power associated with each element as it was observed from the first deployment of the handwritten quiz that many students failed to address all four elements. Since we will have unambiguous information regarding what a student feels are the correct answers, specific rules based on a student's drop-down selections can be applied once certain responses are removed from consideration. In this case, the rules would be applied after the three responses indicating a change in the power of R_1 were removed. So too, responses that correspond to drop-down selections in which ONLY the power of R_2 was believed to change could be removed to avoid lumping localized with sequential misconceptions. Even without using drop-down selection to eliminate the potential for false positives, the results are promising considering only a very simplistic word-matching approach was used.

B. Constant Voltage Errors

The second most common error that showed up in in student responses, found in approximately 15% of responses, was the belief that the voltage drop across R_2 and/or R_3 did not change as the resistance of R_2 decreased. Students making this error often did not justify why they felt the voltage drop across one or both of these resistors would not change as the value of R_2 decreased. Perhaps these students recalled the fact that the potential difference across the ideal voltage source would not change and erroneously extended the thought to the other components, failing to appreciate how the potential difference redistributes across the resistors with changes in the value of R_2. An example response follows. Note that in addition to the constant voltage error, the given response suggests localized thinking on the part of the student.

The total resistance is found by combining R_2 and R_3 in parallel and adding R_1. V_s does nothing different. R_1 has the same potential difference as before. R_2 is less resistive but the potential difference has to be the same so to make up for the loss in resistance. The current is stronger therefore the power related to R_2 increases due to the stronger current and the same voltage. Power equals voltage times current. R_3 remains the same.

Once again, a series of rules were generated based on an analysis of the training set and implemented using matcher. The rules were then applied to the unseen test set. The precision, recall, and F1-score were found to be 0.69, 0.82, and 0.75, respectively. In examining the false positives, it was found that augmenting the basic capability of matcher with a function that excluded a match if there existed the term "voltage source" in the tagged sentence, the precision could be increased to one. The false negatives were found to be cases in which evidence for the misconception was implied as suggested in the following example.

... but R_3 is not directly affected in any way by R_2, thus I_3 stays the same. If I_3 remains the same then the power must also remain the same.

In the above example, specific mention of the potential difference across R_3 is not made. The example was tagged by the content expert with the constant voltage error based on the implication of having I_3 and the power of associated with R_3 being unchanged.

C. Misconception with an Ideal Independent Voltage Source

In a typical course on electric circuit analysis, students are introduced to the model of an ideal independent voltage source. Ideal independent voltage sources provide a constant potential difference across their terminals. The current and thus the power associated with such a source depend on the circuit to which the source is connected. For simplification purposes, batteries are often modeled as ideal sources, though the limitations of such a model for batteries should be explained. As has been noted elsewhere [14],[15] a common misconception regarding batteries (independent voltage sources) is that they are sources of constant current. This misconception, along with the notion that the power associated with a voltage source is constant turned up in approximately 12% of the responses. Two examples follow.

Vs would stay the same power wise because it is independent from the other elements of the circuit.

Given that the voltage source remains unchanged, the voltage and current through it remain the same, so the power does not change.

Considering the training set, rules were developed to catch misconceptions regarding the ideal independent voltage source. Again, these rules were based on either a two- or three-word set in order, with tokens allowed between the key words. The precision, recall, and F1-score were found to be 0.50, 1.0, and 0.67, respectively. Therefore, while the rules caught all responses tagged by the content expert to suggest an error in applying the concept of an ideal independent source, the rules picked out an equal number of false positives. Seventy percent of the false positives could be ascribed to another misconception. If knowledge from the use of drop-down selections as previously described were employed, the remaining 30% of the false positives would not occur, thus improving the performance (precision = 0.59, recall = 1.0, F1 = 0.74) of the simple matching algorithm. It is interesting to note that considering the full corpus (185 responses), the terms "ideal source" or "independent source" when referring to V_s showed up in just over 8% of the responses. Of the responses specifically identifying V_s as an ideal or independent source, nearly half (~47%) misused the term.

D. Resistor Combination Errors

In responding to the writing quiz, students scoring at the upper end of the scale recognized that thinking about the equivalent resistance "seen" by the source was an effective starting point. Such students identified that R_2 and R_3 are in parallel and this parallel combination is in series with R_1. From there, successful students would note that in decreasing R_2, the effective resistance seen by the source decreases. Approximately 9% of the responses in the corpus of 185 unique answers revealed at least one error regarding claims made in terms of resistance; three examples follow.

R_1 and R_3 are in parallel therefore...

...if R_2 goes to zero, then R_1 and R_3 are in series

The equivalent resistance of the circuit increases as R_2 decreases.

The first response demonstrates a clear misunderstanding of what is required for resistors to be in parallel. The error in the second statement likely stems from a misunderstanding of the result of placing a short circuit in parallel with a resistance. The third statement reveals a lack of awareness that the net resistance of two resistors in parallel is less than the smaller of the constituent resistors. It should be remembered that the writing quiz was given early in the semester (fifth class period); nevertheless, promptly correcting such errors is critical. Once again rules to pick out resistance combination errors were composed based on an examination of the training set (60% of the corpus). The rules were then applied to the test set (remaining 40% of the corpus) to establish values of precision (1.0), recall (0.71) and F1-score (0.83). It is interesting to note that the precision was perfect, indicating that zero false positives were identified. It appears then that one would not have to invoke a specific set of rules using knowledge of the drop-down selections when attempting to identify resistance combination errors. That fact that recall was not perfect reflects that additional rules are necessary to capture all the examples of resistor combination errors in the existing corpus. The following is a somewhat obscure example of a "resistance combination error" that was not caught by the rules created when considering the training set.

The equivalent resistance must stay constant, so as the resistance of R_2 decreases, the resistance of R_3 must increase.

An advantage of rule-based systems is that the rules can grow without substantial change in the system. Considering this most recent example, a simple rule can be added to catch such a rare misconception (1 of 185 responses suggested that the equivalent resistance must remain constant).

E. Localized Misconception

As noted in the discussion of the sequential misconception, in terms of the work here, a localized misconception is one in which the responder believes that for a quantity such as current or power to change in an element, that element must itself change. In the problem considered, only the resistance value of R_2 changes and so a student falling prey to the localized misconception would express that only the power of R_2 would change. Just over 4% of the responses were found to exhibit evidence of the localized misconception. In considering the training set when developing the rules to catch the localized misconception, it was noticed that our chosen procedure of analyzing responses one sentence at a time, while effective for the other misconceptions noted, would not be in the case of the localized misconception. The following example illustrates the need for a somewhat more sophisticated approach.

The power associated with Vs with the resistance of R_2 decreasing will be the same. The power of R_1 will remain the same as the current and resistance stay the same. The power of R_2 will decrease for as R decreases, so will power. The power of R_3 will remain the same as current and resistance remain the same.

To deduce that the student exhibits the "localized" misconception, three of the four sentences (sentence 1, 2, and 4) would need to be tagged, as individually they show the author thinks that the power of V_s, R_1, and R_3 will remain the same, respectively. While certainly this could be done, recalling that the existing web-based application has drop-down selections to indicate the nature of a change in power of each of the elements, any response that indicated only the power of R_2 changed could be identified as a localized misconception. Such an approach is the most efficient means to tag responses for the localized misconception in the given question.

F. Precedence of Current Misconception

Consider the following response from the corpus.

If the equivalent resistance for this circuit decreases than [sic] the current will increase in order for the voltage to remain constant.

A common misconception regarding electric circuits has to do with the relationship between potential difference and current. While the correct understanding is to appreciate that it is the potential difference that causes current (i.e. the flow of charge), many operate under the erroneous idea that current causes potential difference. Such a misconception may lead to errors in applying Ohm's law for example, as well as not appreciating that a potential difference can exist across an open circuit [24]. Just under 4% of responses provided direct evidence of this misconception. Rules to catch this misconception were created based on the training set and found to successfully capture all instances of the misconception in the test set with zero false positives. This suggests that perfect precision, recall, and F1-score were achieved. The number of instances of this misconception (4 in the training set and 3 in the test set) was so small, that the results should only be taken to suggest that capturing examples of this misconception may be done relatively easily with the chosen approach.

G. Conservation of Energy Errors

A search was performed to identify all responses that included either the term *conservation of energy* or *conservation of power* and their variants. Of the 185 responses, fourteen (~8%) included such a reference. In ten of these fourteen, the use of the term is either incorrect or ambiguous. Consider the following examples.

The power associated with V_s and R_1 will stay the same because Vs is the only provider of power in the circuit and because power is conserved it must stay the same.

I know that R_2 and R_3 are connected in parallel so they won't have the same currents. But, I would need to combine all my resistors first to find the total current. After finding all the currents, figure out [sic] power equations before and after R_2 (keeping in mind that power is supposed to be conserved but in this case, it's not).

The first statement attempts to use the notion of conservation of power to argue (incorrectly) that the power of the voltage source must remain constant as it is the only source and thus only provider of energy in the circuit. In the second case, while mentioning that, "power is supposed to be conserved," the student believes he/she has found an example in which it is not. Students who used conservation of energy correctly considered the power (energy per unit time) of each element in the circuit before coming to a conclusion. As the method pursued in this pilot study looked only at the sentence level and did not try to tie the meaning from one sentence in a response to another within the response, no attempt was made to determine precision, recall, and F1-score. It is quite possible that values for those quantities would be considerably higher than warranted due to the fact that it was trivial to extract examples of the use of the concept of conservation of energy and that most of the uses were not applied properly.

III. A Machine Learning Approach

As a compliment to the rule-based approach, a machine learning approach was investigated. We modeled the task of predicting responses with sequential misconceptions as a supervised binary classification problem in which examples are the whole responses and the labels are positive if the response contains a sentence conveying the misconception. We used a BERT model pre-trained on a large corpus of general English text data in a self-supervised fashion. Then, we fine-tuned the BERT model for our task using the labeled training data. We used the exact same data and experimental setup used to evaluate the rule-based approach; a comparison of key metrics between the rule-based and BERT approaches are provided in Table 1.

Table 1: Comparison of models for sequential misconception prediction

Model	Precision	Recall	F1
Rule-based	0.63	1.0	0.77
BERT	0.90	0.75	0.82

As depicted in the table, BERT's precision is excellent but its recall, which we claim is more important for misconception prediction, is relatively lower than that of the rule-based approach. With more training data and domain-specific pre-training [25], data-hungry models such as BERT will likely outperform rule-based models in both metrics. For our purposes, a reasonable approach in developing a misconception detection system for a given conceptual-based writing problem would be to create a combined system in which the rule-based method is initially active until sufficient gold-standard data is generated from multiple offerings of the quiz at which point the BERT model may start assisting for improved accuracy. Regardless of the chosen NLP approach to misconception detection, the ultimate aim is to correct the misconception through some form of feedback. In the following section, the proposed method of feedback for the conceptual-based writing problem is described and demonstrates that perfect interpretation of the student's response is not necessary.

IV. Correcting the Misconceptions through Feedback

The original intent of the writing quiz described above was to identify students at-risk to fail the course [2]. Currently, additional writing quizzes are being explored to correct student misconceptions and to foster metacognitive skill. The manner in which students are given feedback is no doubt crucial. Our conception of feedback in terms of the writing quizzes is not primarily one of simply articulating correctness or incorrectness of a response but rather is itself a form of instruction as suggested in [26]. As noted by Hattie [27], feedback, "is most powerful when it addresses faulty interpretations, not a total lack of understanding. Under the latter circumstance, it may even be threatening to a student." Hattie goes on to note that, "a key theme arising from this review of the literature is the importance of ensuring that feedback is targeted at students at the appropriate level." Simply describing to the class as a whole, a correct approach to answering the question is therefore not the most effective for it does not target the individual. Our desire is to use a web-based application to evaluate student responses and to immediately provide them individualized feedback based on a Directed Line of Reasoning (DLR) algorithm [5].

In the previous sections, we have described how employing drop-down selections in the web application and simple word matching using an open-source tool, or using a machine learning approach such as BERT, we can begin to assemble a model of a given student in terms of their understanding of the writing quiz. Based on any identified misconceptions, or solid conceptions such as the value of considering the equivalent resistance, a starting point in the DLR could initiate. For example, should a student mention conservation of energy, the algorithm could congratulate the student on considering the term in the context of the problem and then lead them through simple question and answering (perhaps through multiple choice selections) to a proper understanding of the term. In a similar fashion, should a response be tagged with a sequential misconception, the DLR algorithm could help the student come to appreciate the importance of looking at a circuit as a system. The key point with such an approach is that it can provide a level of individualized feedback that is immediate as opposed to either a class discussion (not individualized) or instructor feedback given days later (not immediate) due to grading demands.

V. Conclusions

As a means to identify misconceptions with regard to basic concepts related to electric circuit analysis within the writings of electrical and computer engineering students, a simple word matching approach applied at the sentence level, has been described. Even though a very rudimentary NLP approach was implemented, the results have shown considerable promise for the task. For example, the precision, recall, and F1-score were 0.632, 1.0, and 0.77, respectively when searching for evidence of a sequential misconceptions. Other misconceptions and errors, such as that associated with resistor combinations were found with similar metrics. Misconceptions such as "localized thinking" would not be effectively found by analyzing the word choices of single sentences, but rather need a more sophisticated algorithm that looks at the meaning (as revealed through word search) of multiple sentences within a response. More simply, using the drop-down selections of the existing web-based application would permit the localized misconception to be identified without resorting to an NLP algorithm altogether. Preliminary results from BERT, a machine learning approach apparently well-suited to short answer evaluation, were described. The value of such an approach grows as the corpus of responses used for training grows. Whether the evaluation is done by a human or via an NLP-based algorithm as described, there is often ambiguity for the very reason that students were often found to fail to adequately justify their responses to the considered conceptual writing quiz. This is where a Directed Line of Reasoning approach to providing feedback would be most useful.

Acknowledgements

This material is based upon work supported of the National Science Foundation under Grant No. 1504880. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Works Cited

[1] Writing to Develop Mathematical Understanding, David K. Pugalee, Christopher-Gordon Publishers, Inc. Norwood, Massachusetts, p. 111, 2005.

[2] Becker, J. P., & Plumb, C. (2018, June), Board 8: "Identifying At-Risk Students in a Basic Electric Circuits Course Using Instruments to Probe Students' Conceptual Understanding," paper presented at 2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah. 10.18260/1-2—30110.

[3] https://spacy.io/usage/rule-based-matching; last accessed 12/17/2020

[4] D.M.W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation", Journal of Machine Learning Technologies, Vol. 2, Issue 1, pp. 37-63, 2001.

[5] Bruce Mills, Martha Evens, Reva Freedman, "Implementing Directed Lines of Reasoning in an Intelligent Tutoring System Using the Atlas Planning Environment", IEEE Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04).

[6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, "Pre-training of Deep Bidirectional Transformers for Language Understanding," Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (2019) Association for Computational Linguistics; Minneapolis, Minnesota.

[7] Sung C., Dhamecha, T.I., Mukhi, N. (2019) Improving Short Answer Grading Using Transformer-Based Pre-training. In: Isotani S., Millán E., Ogan A., Hastings P., McLaren B., Luckin R. (eds) Artificial Intelligence in Education. AIED 2019. Lecture Notes in Computer Science, vol 11625. Springer, Cham.

[8] Robert A. Blanc, Larry E DeBuhr and Deanna C. Martin, "Breaking the Attrition Cycle, The Effects of Supplemental Instruction on Undergraduate Performance and Attrition," Journal of Higher Education, Volume 54, Number 1, 1983, pp. 80-90.

[9] Susan Ambrose, et al. "How Does Students' Prior Knowledge Affect Their Learning?" in How Learning Works: Seven Research-Based Principles for Smart Teaching, John Wiley & Sons, pp. 10-39, 2010.

[10] Deborah L. Butler and Philip H. Winne, "Feedback and Self-Regulated Learning: A Theoretical Synthesis," Review of Educational Research, Vol. 65, No. 3, pp. 245-281, 1995.

[11] John Flavell, "Metacognitive Aspects of Problem Solving," in The Nature of Intelligence, Lauren B. Resnick ed., Lawrence Erlbaum Associates, Hilsdale, N.J., 1976.

[12] Clifford Konold (1995) Issues in Assessing Conceptual Understanding in Probability and Statistics, Journal of Statistics Education, 3:1, , DOI: 10.1080/10691898.1995.11910479.

[13] Michelene T.H. Chi, "Commonsense Conceptions of Emergent Processes: Why Some Misconceptions Are Robust," Journal of the Learning Sciences, 14(2), pp. 161-199, 2005.

[14] Paula Vetter Engelhardt and Robert J. Beichner, "Students' Understanding of Direct Current Resistive Electrical Circuits," American Journal of Physics, Vol. 72 (98), pp. 98-115, 2004.

[15] Tatiana V. Goris and Michael J. Dyrenfurth, "How Electrical Engineering Technology Students Understand Concepts of Electricity. Comparison of Misconceptions of Freshmen, Sophomores, and Seniors," Proceedings of the 2013 American Society for Engineering Education Annual Conference and Exposition. Paper ID 5849.

[16] David P. Tallant, "A Review of Misconceptions of Electricity and Electrical Circuits," The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, August 1-4, 1993.

[17] Deepika Sangam and Brent K. Jesiek, "Conceptual Understanding of Resistive Electric Circuits Among First-Year Engineering Students," Proceedings of the 2012 American Society for Engineering Education Annual Conference and Exposition.

[18] Richard Gunstone, Brian McKittrick, Pamela Mulhall, "Textbooks and their authors: another perspective on the difficulties of teaching and learning electricity," in *Research and the Quality of Science Education*, Kerst Boersma, Martin Goedhart, Onno de Jong, Harrie Eijkelhof,.eds. Springer, 2005.

[19] Deepika Sangam and Brent K. Jesiek, “Conceptual Gaps in Circuits Textbooks: A Comparative Study,” IEEE Transactions on Education, Vol. 58, August 2015, pp. 194-202.

[20] Brian J Skromme, “Addressing Barriers to Learning in Linear Circuit Analysis,” 122nd ASEE Annual Conference and Exposition, Paper ID #14125, June 2015.

[21] Olde Bekkink et al., “Uncovering students’ misconceptions by assessment of their written questions,” BMC Medical Education (2016) 16:221 DOI 10.1186/s12909-016-0739-5

[22] Becker, J. P., & Sior, E., & Hoy, J., & Kahanda, I. (2019, June), Board 11: “Predicting At-Risk Students in a Circuit Analysis Course Using Supervised Machine Learning,” presented at 2019 ASEE Annual Conference & Exposition, Tampa, Florida. 10.18260/1-2--32185

[23] *Natural Language Processing in Action*, Hobson Lab, Cole Howard, and Hannes Hapke, Manning Publications, 2018.

[24] R. Cohen, B. Eylon, and U. Ganiel, “Potential difference and current in simple electric circuits: A study of students’ concepts,” American Journal of Physics 51, 407 (1983)

[25] Chul Sung, Tejas Dhamecha, Swarnadeep Saha, Tenghai Ma, Vinay Reddy, and Rishi Arora, “Pre-Training BERT on Domain Resources for Short Answer Grading,” Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 6071-6075.

[26] R.W. Kulhavy, “Feedback in written instruction.” Review of Educational Research, 47(1), 211–232 (1977).

[27] John Hattie and Helen Timperley, “The Power of Feedback,” Review of Educational Research March 2007, Vol. 77, No. 1, pp. 81–112.