DOI: 10.47743/eufire-2025-1-16

QUANTITATIVE ANALYSIS OF THE RELATIONSHIP BETWEEN MATHEMATICAL AND COMPUTATIONAL THINKING IN THE CONTEXT OF AI INTEGRATION

ELENA KRAMER
Alexandru Ioan Cuza University of Iași
Iași, Romania
elenak@braude.ac.il

Abstract
This research explores the relationship between Mathematical Thinking (MT) and Computational Thinking (CT), two forms of reasoning that are often defined differently but share conceptual similarities. We propose a novel approach to studying this relationship by comparing the metalanguages – the general-purpose vocabulary and structures used to express ideas – across various fields in Mathematics and Computer Science.
Our main hypothesis is that if different fields share similar metalanguages, this reflects a deeper connection between them, which may influence understanding and success across domains. To test this, we analyzed multiple text corpora from a range of mathematical and computational disciplines. Using advanced Natural Language Processing (NLP) techniques, including lemmatization and tokenization, we filtered out domain-specific terms to reveal the underlying metalanguage.
We applied several clustering algorithms – K-Means, PAM, Density-Based Clustering, and Gaussian Mixture Models – to group fields based on linguistic similarity. Since clustering results can be sensitive to parameters and distance metrics, we further validated the outcomes using a Neural Network-based AI model. This AI integration helped assess the consistency of the clusters and provided a second layer of insight into the linguistic structures across fields.
To further evaluate the hypothesis – that fields with similar metalanguages may promote similar levels of comprehension – we combined this computational analysis with both quantitative and qualitative data from student participation. This paper presents the results of the quantitative component, highlighting how AI-assisted analysis can reveal meaningful connections between MT and CT through their shared linguistic foundations.
Keywords: *mathematical thinking; computational thinking; undergraduate students; education skills; metalanguage.*
JEL Classification: C6; A22.

1. INTRODUCTION

Computers and programming have significantly transformed the modern world, establishing technological literacy as an essential skill for academic and professional success in the digital age (Shute *et al.,* 2017). As a result, computational thinking (CT) – a mode of problem-solving applicable across diverse disciplines – has gained increasing attention in education.

There is a growing need to equip students, as future professionals, with complex thinking capabilities essential for addressing societal and business challenges. This necessitates the integration of both mathematical thinking (MT) and computational thinking (CT).

One of the major challenges lies in developing effective methods for evaluating thinking processes. We propose examining the linguistic dimension of thought, recognizing that thinking is embedded within language (De Saussure, 1916; Heidegger, 1927). Noam Chomsky, widely regarded as the father of modern linguistics, emphasizes the profound connection between language and thought (Chomsky, 2006).

Given that both Mathematics and Computer Science are structured around formal languages, and individuals employ metalanguage in cognitive processes, it becomes possible to analyze the relationships among these metalanguages to better understand thinking. As defined by Alfred Tarski, metalanguage refers to the language used to discuss the properties and rules of another language, including its expressions, sentence construction, syntax, and truth conditions (Tarski, 1944; Gruber, 2016).

To formulate a clearer hypothesis about the relationships between capabilities in specific domains of Mathematics and Computer Science, we adopted a novel approach – comparing the metalanguages used in various subfields within each discipline. In this Data Mining phase (Hand, 2007; Cheng, 2017), we analyzed numerous text files drawn from distinct areas of Mathematics and Computer Science.

The Mathematical domains examined included: Linear Algebra, Abstract Algebra, Combinatorics and Probability, Mathematical Analysis, Set Theory, and Logic. For Computer Science, the areas studied were: Functional Programming, Imperative Programming, Object-Oriented Programming, Data Structures and Algorithms, Automata and Formal Languages, and Operating Systems. These subjects represent core components of most Computer Science curricula.

The following steps were undertaken:

- Text classification and vectorization using the FastText model.
- Distance computation between vectors (Cohen *et al.*, 2009; Van Dongen and Enright, 2012).
- Application of clustering algorithms (Broder *et al.,*1997).

As a result, groups of fields in Mathematics and Computer Science with closed metalanguages were received. Because clustering results can be influenced by parameter choices and distance metrics, we used XLNet – an AI model based on a neural network to further validate the findings. This AI-based approach helped check the stability of the clusters and offered additional insights into the linguistic structures shared across different fields.

To test the hypothesis that fields with similar metalanguages may support similar levels of understanding, we combined this computational analysis with both quantitative and qualitative data gathered from student participation. This paper focuses on the quantitative findings, showing how AI-assisted analysis can uncover meaningful links between Mathematical Thinking (MT) and Computational Thinking (CT) through their common linguistic features.

2. DATA MINING RESEARCH STAGE WITH AI INTEGRATION

This section presents the clustering process and the use of XLNet – an AI model based on neural networks. The aim of both processes is to derive a division of fields in Mathematics and Computer Science according to the similarity of their metalanguages.

2.1. Clustering process

After converting the texts into vectors and calculating the distances between them, four clustering algorithms were applied: K-Means, PAM, Density-Based Clustering, and Gaussian Mixture Models. The process began with K-Means, as it is widely regarded as one of the most commonly used and effective algorithms. However, a key challenge is that K-Means performs optimally with the Euclidean distance metric. Since our analysis focuses on correlations, we opted to use the Pearson distance metric instead.

Although K-Means remains a valuable method due to its relative simplicity, the results were not entirely logical or consistent, possibly due to its inherent reliance on Euclidean assumptions, as noted earlier. After running all four algorithms, the PAM (Partitioning Around Medoids) algorithm produced the most coherent and meaningful clusters. The results are presented in Figure 1.

Operation Systems	63	23	9	5
Data Structures and Algorithms	59	10	30	2
Functional Programming	17	21	58	5
Imperative Programming	65	5	9	20
OOP	58	14	14	15
Automata and Computation Theory	20	7	63	10
Abstract Algebra	16	67	9	8
Analysis	8	9	8	75
Combinatorics And Probability Theory	62	12	14	12
Linear Algebra	15	65	10	10
Logic	17	10	52	21
Set Theory	30	9	51	10

Figure 1. Division of fields into clusters

The analysis resulted in the following four clusters:

Cluster 1: Operation Systems, Data Structures and Algorithms, Imperative Programming, OOP, Combinatorics and Probability.

Cluster 2: Linear Algebra, Abstract Algebra.

Cluster 3: Functional Programming, Automata and Computation Theory, Logic, Set Theory.

Cluster 4: Analysis.

2.2. Using XLNet – a neural network-based AI model

XLNet is a transformer-based language model that captures the meaning and context of words within text (Yang *et al.*, 2019). In this study, XLNet is used to analyze the metalanguages found in texts from different fields in Mathematics and Computer Science. By converting these texts into numerical representations (vectors), XLNet allows us to compare and group fields based on similarities in their metalanguages. This helps reveal underlying connections between disciplines through their shared or distinct ways of expressing concepts.

After training the XLNet model on our domain-agnostic text segments, we achieved an overall accuracy of 86% on the test set. The results obtained using the model are presented in the Figure 2.

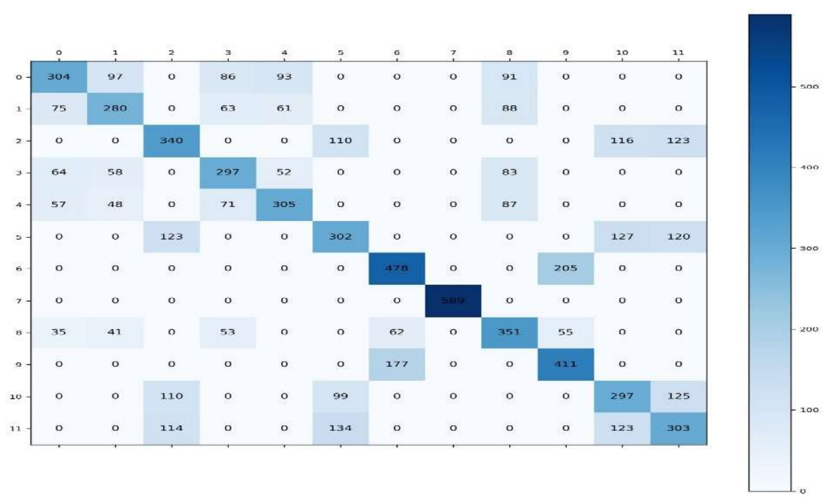

Figure 2. Outcomes of applying the XLNet Model

All the disciplines were labeled from 0 to 11, with the Computer Science domains labeled from 0 to 5 (e.g., Operating Systems is labeled as 0), and the

Mathematics domains labeled from 6 to 11 (e.g., Abstract Algebra is labeled as 6), following the order shown in Figure 1. If we examine the first line, which represents the data labeled as Operating Systems, we find a total of 671 samples. Approximately 45% of this data was correctly identified as related to Operating Systems, while the remaining samples were classified under domains 1, 3, 4, and 8. This distribution suggests a similarity in the metalanguages of Operating Systems and these domains. The result is consistent with the previously shown cluster division, supporting the reliability of the clustering process. A similar pattern can be observed across the remaining lines, indicating a consistent regularity in the data classification.

It can be observed that the distribution of groups closely resembles the one produced by the clustering process. The purpose of this dual approach is to verify that the outcome of the clustering process is not random. Based on this comparison, it can be concluded that the grouping derived from the clustering process is both reliable and meaningful.

3. EXPLORING STUDENT RESPONSES: A QUANTITATIVE APPROACH

After obtaining the results of text clustering in the previous stage of the research, the next step was to test the hypothesis that students would demonstrate a similar level of understanding in areas that share similar metalanguages. To examine this, a study was conducted involving student participation.

3.1. Participants profile

Both primary and secondary data were utilized in this study (Hox *et al.*, 2005). The secondary data consist of students' grades in Mathematics and Computer Science courses offered throughout the duration of the engineering degree at a college of engineering. These data were obtained from the faculty secretary, in accordance with ethical guidelines and confidentiality regulations. The research participants were approximately 200 students who had completed their studies in Software Engineering at a College of Engineering.

To gather primary data, questionnaires were constructed containing questions from each of the 12 topics (6 from Mathematics and 6 from Computer Science). These questionnaires were administered in class with the participation of about 150 final-year students studying software engineering at a college of engineering. These students were chosen in their final academic year of studies, so they had already studied the courses taken as representative fields in Mathematics and Computer Science. They also bring a varied background level in Mathematics and Computer Science from school education.

Participants received comprehensive information about the research's purpose, and they were informed that their participation was voluntary and that their anonymity would be maintained. Additionally, the research was conducted with the approval of the college ethics committee.

3.2. Questionnaires and data collection

The questionnaire was designed to assess understanding in Mathematics and Computer Science domains in order to evaluate thinking. Before distribution, it underwent expert validation. The questionnaires consist of a series of problem-solving exercises, such as developing new algorithms, writing code in various programming languages, improving and debugging existing code, and proving the correctness of algorithms. We chose to use questionnaires because they allow for the systematic collection of data to assess understanding in specific domains and evaluate higher-order thinking

Questionnaires are widely used in research involving Mathematics and Computer Science for several key reasons. First, they enable the collection of large-scale quantitative data efficiently, which is essential for identifying patterns and drawing statistically significant conclusions (Creswell, 2014). Second, questionnaires allow researchers to assess participants' perceptions, experiences, and conceptual understanding across diverse topics, supporting cross-domain comparisons (Rach *et al.,* 2013). Third, structured questionnaires can be replicated or adapted for longitudinal studies, enhancing the reliability and generalizability of findings (Cai and Wang, 2010).

As a result, we received data consisting of 12 grades for each of the 150 students. For each discipline, students received a grade ranging from 0 (lowest) to 100 (highest), with 55 considered a passing grade. The 0–100 grading scale is widely used due to its clarity, flexibility, and fine level of granularity, allowing educators to differentiate student performance with precision. It provides a familiar and intuitive framework for both instructors and students, making it easier to interpret results and compare outcomes across assessments and institutions (Brookhart, 2011; McMillan, 2014).

4. RESULTS

This chapter presents the results of a quantitative analysis to test the hypothesis that fields with similar metalanguages may support similar levels of understanding.

4.1. Secondary data analysis

The analysis of secondary data did not yield significant results. A possible reason for this is that grading does not always reflect knowledge. Sometimes instructors use factors and adjustments, making it difficult to conduct a clean analysis of the relationship between mathematical thinking and computational thinking.

4.2. Primary data analysis

The analysis includes interpreting the results of questionnaires with exercises in the fields of Mathematics and Computer Science.

To examine if the grades of 12 fields can be divided into clusters, principal-component exploratory factor analysis with orthogonal rotation was conducted (Jolliffe and Cadima, 2016; Brown, 2009). The analysis yielded three factors, explaining 65.9% of the variance. The results are presented in Table 1.

Table 1. Rotated Component Matrix

	Component		
	1	2	3
Logic	.783		.103
Set Theory	.681	-.138	.113
Automata and Computation Theory	.635		
Combinatorics and Probability Theory	.624	-.141	-.157
OOP	-.163	.682	
Data Structures and Algorithms	.107	.647	
Imperative Programming	-.133	.614	
Functional Programming		.561	
Operation Systems		.497	
Linear Algebra		.105	.821
Modern Algebra			.690
Analysis			.686

The table shows a division of the fields into three groups based on color. This division matches 75% of the clustering made according to the meta-languages. This is a very important result that can certainly support the main hypothesis that fields with similar metalanguages may support similar levels of understanding.

Another interesting result emerged from the correlation analysis between prior knowledge in Computer Science and grades in core programming courses. Figure 3 clearly shows a strong correlation, indicating that prior knowledge in Computer Science significantly influences interest and abilities in the field.

	Prior High CS Level
Analysis	0.041895
Set Theory	-0.152997
Linear Algebra	0.084848
Modern Algebra	-0.020629
Logic	-0.053952
Combinatorics and Probability Theory	-0.108896
Imperative Programming	0.841845
OOP	0.242819
Data Structures and Algorithms	0.152565
Operation System	0.059551
Automata and Computation Theory	-0.019092
Functional Programming	0.256233

Figure 3. Outcomes of applying the XLNet Model

5. CONCLUSIONS

As shown in Table 1, the results of grade distribution across different fields were divided into three groups based on similarity in success. 75% of the data match the clusters obtained in the previous stage, according to the fields division based on similarity in metalanguages. There is also a reasonable explanation for fields that were grouped differently. For example, two types of programming (Imperative Programming and Functional Programming) are expected to differ in concept and clearly have different metalanguages. However, since both fall under the category of programming, it is possible that someone who is good at programming performs well in all its types. Another example is Analysis: although the data suggest that its metalanguage is relatively unique and unlike the others, students study Analysis and the two Algebra courses (which were grouped into the same cluster) in the same semesters. Therefore, the level of effort in these courses may be similar, which could result in similar levels of success. Still, the 75% alignment with the metalanguage-based grouping strongly supports the validity of the main hypothesis that fields with similar metalanguages may support similar levels of understanding.

A possible continuation of this research could be to design a questionnaire with exercises in a specific field where a student struggles, but framed in the style of a field where the student performs well. If performance improves, it would suggest a strong connection between thinking and language.

Another interesting finding is the correlation between prior knowledge in Computer Science and success in programming courses, which are considered core courses. This can be seen as a strong recommendation for those who want their children to pursue a popular and in-demand profession like Computer Science. The suggestion is to teach children programming in school, as this may increase their interest and, according to the research, also improve their success in the field.

References

1) Broder, A., Glassman, S., Manasse, M. and Zweig, G. (1997). Syntactic clustering of the web. *Computer Networks and ISDN Systems,* 29(8), pp. 1157–1166.
2) Brookhart, S.M. (2011). *Grading and Learning: Practices That Support Student Achievement*. Bloomington, Solution Tree Press.
3) Brown, J.D. (2009). Choosing the right type of rotation in PCA and EFA. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 13(3), pp. 20–25.
4) Cai, J. and Wang, T. (2010). Conceptions of effective mathematics teaching within a cultural context: Perspectives of teachers from China and the United States. *Journal of Mathematics Teacher Education*, 13(3), pp. 265–287.
5) Cheng, J. (2017). Data-mining research in education. *arXiv preprint* arXiv:1703.10117.
6) Chomsky, N. (2006). *Language and Mind*. Cambridge University Press.
7) Cohen, I., Huang, Y., Chen, J. and Benesty, J. (2009). Pearson correlation coefficient. In: *Noise reduction in speech processing*, 2, pp. 1-4.
8) Creswell, J.W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed. SAGE Publications.
9) De Saussure, F. (2011). *Course in general linguistics*. Columbia University Press.
10) Gruber, M. (2016). *Alfred Tarski and the "Concept of Truth in Formalized Languages": A Running Commentary with Consideration of the Polish Original and the German Translation*, 39. Berlin: Springer.
11) Hand, D.J. (2007). Principles of data mining. *Drug safety,* 30, pp. 621-622.
12) Heidegger, M. (2010). *Being and Time*. Albany: SUNY Press.
13) Jolliffe, I.T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), p. 20150202.
14) McMillan, J.H. (2013). *Classroom Assessment: Principles and Practice for Effective Standards-Based Instruction*. 6th ed. Boston: Pearson.
15) Rach, S., Ufer, S. and Heinze, A. (2013). Learning from errors: Effects of teacher and student responses in mathematics. *ZDM*, 45(4), pp. 503–515.
16) Shute, V.J., Sun, C. and Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, pp. 142-158.
17) Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and phenomenological research,* 4(3), pp. 341-376.
18) Van Dongen, S. and Enright, A.J. (2012). Metric distances derived from cosine similarity and Pearson and Spearman correlations. *arXiv preprint* arXiv:1208.3145.
19) Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. and Le, Q.V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems*, pp. 5754-5764.