Discovering Real-World Use Cases for a Multimodal Math Search Interface

Keita Del Valle Wangari

Rochester Institute of Technology
152 Lomb Memorial Drive
Rochester, NY 14623
+1 951-963-4780
keita@thinkHCI.com

Richard Zanibbi

Rochester Institute of Technology
102 Lomb Memorial Drive
Rochester, NY 14623
+1 585-475-5023
rlaz@cs.rit.edu

Anurag Agarwal

Rochester Institute of Technology
85 Lomb Memorial Drive
Rochester, NY 14623
+1 585-475-7531
axasma@rit.edu

ABSTRACT

To use math expressions in search, current search engines require knowing expression names or using a structure editor or string encoding (e.g., LaTeX). For mathematical non-experts, this can lead to an "intention gap" between the query they wish to express and what the interface will allow them to express. m_{in} is a search interface that supports drawing expressions on a canvas using mouse/touch, keyboard and images. We present a user study examining whether m_{in} changes search behavior for mathematical non-experts, and to identify real-world usage scenarios for multimodal math search interfaces. Participants found query-by-expression using hand-drawn input useful, and identified scenarios in which they would like to use systems like m_{in} such as for locating, editing and sharing complex expressions (e.g., with many Greek letters), and working on complex math problems.

Categories and Subject Descriptors

H.3.3**. [Information Storage and Retrieval]**: Information Search and Retrieval – Query formulation, Search process; H.5.2. **[Information Interfaces and Presentation]**: User Interfaces – User-centered design.

Keywords

User interface design; multimodal input; search interfaces; Mathematical Information Retrieval; MIR; query-by-expression.

1. INTRODUCTION

Our study was designed to gain insight into whether expressions would be desirable and/or useful to non-experts (in our case, college students) when conducting math-related searches. We looked at relevant research in the areas of visual math perception and math input to understand the mental model and input preferences of math searchers and then surveyed the current state of math search interfaces to discover whether there are usability issues that might specifically impact the non-expert user.

What we found is that current math search interfaces limit users to expressing the math expression portion of their search need in the form of text, an encoding language, or via the use of some type of

equation editor. If users are visualizing a formally notated math expression before they search, it seems unfortunate to break this visual flow by forcing users to first convert the diagrammatic math expression into a coded, sentential one or build it through the tedious use of selection menus in order to initiate the search process. Further, since many math expressions are learned, consumed and recalled visually, being able to **query** using expressions (i.e., *query-by-expression*) seems a natural way for non-experts to search for mathematical information [12].

In a prior study of math professors and graduate students, participants could not identify scenarios where entering math expressions as search terms would be useful [14], and found it sufficient to lookup formulas, metrics and concepts textually, as they often knew them by name. However, the study did not consider non-experts, and participants were not provided with a prototype supporting non-textual math input for evaluation.
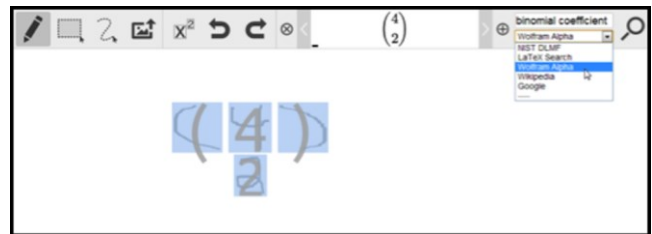

Figure 1**: A query combining a handwritten expression and keywords in m_{in}.**[1] **Different search engines may be selected from a drop-down menu (visible at top right).**

In contrast, for our study, non-expert participants used the math search interface m_{in} (see Figure 1 [6]). m_{in} supports drawing expressions as opposed to entering them textually or using template editors (e.g., Microsoft Equation Editor). The goals of our user study were: 1) to observe whether m_{in} changes user search behavior, and 2) document relevant use cases for math search interfaces by mathematical non-experts. Changes in user behavior included increased use of math expressions in search, and an increased capacity for some participants to represent expressions for search, and a number of realistic use cases were identified. We believe that our findings can inform the design of math search interfaces so that they better support mathematical non-experts.

2. RELATED WORK

The appearance of math expressions affects our reasoning about them [6]. Landy & Goldstone ran a series of experiments where

---

[1] http://saskatoon.cs.rit.edu/min_instructions

participants were asked to judge whether given equations were valid. Equations were structured with either nonstandard spatial relationships, irrelevant or induced groupings, or manipulated spacing. This influenced participant responses even when given feedback, and when they recognized that they were likely being misled by their perceptions.

In another set of experiments, participants were presented with equations expressed textually, and asked to either handwrite or type them into an interactive website using a keyboard [5]. Participants systematically spaced the expressions even when it was unnecessary. Landy & Goldstone suggest that math notation may be grounded in visual structure, since people seem to be affected by small changes in symbol layout.

Typed input has been found not to be ideal for authoring math notation by many users [2]; equation editors fall short as searching for symbols is tedious and making changes in symbol layout can be difficult [11]. In an initial study, users asked to use and rate various input modalities for math, found handwriting to be the most natural and satisfactory [2]. Later, these results were found to generalize to middle- and high-schoolers and to simpler, easily typed equations [1].

If users have a math expression in mind, it may be disruptive for them to convert it into code or text to initiate the search process. This gap between a user's intentions and a system's allowable actions is known as the Gulf of Execution [9]. The point at which a user is unable to precisely express their search intent as a keyword query has been termed the "intention gap" [13]. We feel this intention gap is present in major text-based search engines like Google and in math search interfaces such as Wolfram Alpha, NIST Digital Library of Mathematical Functions, as well as in more recently developed math search engines like EgoMath [7] and WikiMirs [3], both of which enable math search for Wikipedia.

EgoMath, whose interface accepts math expressions in TeX format, text or both adapts keyword search for math. Keywords have sufficient semantics for search engines to provide users with a relevant and sensibly ranked results list but, on their own, math symbols do not as they usually derive meaning from how they are organized structurally in a math expression [4]. WikiMirs, whose interface accepts LaTeX input, is designed to look for textual and spatial similarities between the user's search query and the indexed information. Both EgoMath and WikiMirs require the searcher to know an encoding language and demonstrate the sentential form of input where support for the diagrammatic aspects of complex math expressions is missing.

The Math-Aware Search Engine (MASE) [8] is another recently developed math search engine. It is designed to support math question and answering (Q&A) systems and its interface allows users to input text and math expression queries using the equation editor CODECOGS as a front-end. While equation editors do lend support for the two-dimensional, spatial aspect of complex math expressions, building and modifying the expression can be tedious, as discussed earlier. Given the strong visual-spatial aspect of math notation and its impact on reasoning, and given the benefits of and preference for handwriting as a means of inputting math on a computer, it appears the current search interfaces may not provide the best user experience, particularly for non-experts, when searching for math-related information.

m_{in} [10] is a search interface that allows users to draw expressions on a canvas (see Figure 1). The system is multimodal, also accepting keyboard and image input. Recognition of user input is displayed on the canvas. Tools are provided for correction and manipulation including symbol selection, stroke selection, undo/redo, and optical character recognition (OCR) correction. Once the user is satisfied, they may submit the expression along with keywords to various search engines. m_{in} converts the expression into LaTeX, and then combines it with the keywords into a query string that can then be sent to various math-aware search engines that support LaTeX input.

3. METHODOLOGY

Participants. The 16 participants were 18 or older, currently enrolled in a first- or second-year college math course, self-rated as Beginner or Intermediate level in math knowledge, and self-rated as Comfortable or Very Comfortable using the internet. "Beginner" was defined as having knowledge consistent with Basic Math and Pre-Algebra concepts. "Intermediate" was defined as having knowledge consistent with Geometry, Algebra and Pre-Calculus concepts. Participants were recruited via email. Target recipients were current Rochester Institute of Technology (RIT) students in the College of Science, which contains the math department, and the College of Computing and Information Sciences.

Environment. The study was conducted in RIT's Usability Lab. The room was equipped with a desktop PC with a webcam, and a speakerphone connected to the observation room. The observation room was equipped with a desktop PC. Background information and post-study ratings were collected from each participant via online surveys. A moderator in the test room took notes during each session. For 14 of the 16 sessions, an observer in the observation room marked session recordings with task times and noted interesting participant comments and actions; for the 2 moderator-only sessions, task times were marked later.

Tasks. The math topics and expressions used in the tasks were selected based on the recommendations of the third author, a math professor who advised that they were consistent with information needs that students confront in freshman or sophomore college math courses. Each task intentionally included both keywords and math expressions to allow us to observe which one participants preferred to use in their searches. The tasks were written in peer-assist style, where participants imagine they are assisting a classmate, to reduce anxiety participants might experience by feeling their knowledge of math is being tested.

Task 1: Your classmate is having difficulty recognizing polynomials. Find one or more resources to help explain to your classmate why $x^2 - 7x + 2$ is a polynomial and why $\frac{x^2 - 7x + 2}{x+2}$ is not a polynomial.

Task 2: Your classmate has heard of Pascal's triangle but doesn't understand how it relates to math. Find one or more resources to help explain to your classmate how the equation $(x+y)^2 = x^2 + 2xy + y^2$ relates to Pascal's Triangle.

Task 3: Your classmate is struggling with binomial coefficients. Find one or more resources to help explain to your classmate how to find the value of $\binom{4}{2}$.

Task 4: Your classmate is having trouble understanding the prime counting function. Find resources that help explain why $\pi(2) = 1$.

Search Tool Conditions. Participants were asked to bring textbooks, class materials or notes from their current math class so math search behaviors that are tool-independent could be

identified. Participants were randomly assigned into four groups. The tasks were presented in a different order to each group allowing each task an equal opportunity to be first and to be presented uniformly across the search conditions. The four search conditions were presented to all participants in a fixed order: 1) using their choice of text books, notes, websites, and/or online search, 2) online search without the m_{in} interface, 3) online search using only m_{in}, and 4) online search with the option of using m_{in}.

Between the 2nd and 3rd conditions, participants were shown the m_{in} interface, asked brief questions to obtain their initial impressions, and then given a five-minute demonstration. The input modalities available and demonstrated were handwritten input on the canvas with a mouse, keyboard input on the canvas and keyboard input into the search box.

Post-study interview. Participants rated their experience via an online survey, were interviewed about the session by the moderator, and then paid $20 for their time.

4. RESULTS

In Search Condition 1, 14 participants used online search and 2 used their textbook. In Search Condition 4, 11 participants used m_{in} canvas and 5 chose instead to use online search with a web browser or m_{in}'s text search.

None of the participants were observed using math encoding languages (e.g., LaTeX) or template editors for expressing their math search need although a few claimed to be familiar with them. Admittedly, the structure of the expressions in 3 of the 4 tasks did not require any special encoding in order to be entered into a standard search engine text box. But for the one task that did, the Binomial Coefficient task containing the expression "4 choose 2", most participants not only did not attempt to use any special encoding languages, but many were not familiar with how to say the expression in words. They recognized it visually but beyond that, were unclear how to express it in sentential form – the only form afforded by the conventional search text box.

The biggest behavior change noted in participants using the m_{in} interface was the large increase in the use of expressions in their search query. In Search Condition 1, 2 of the 14 participants who chose online search used an expression in their initial query and in Search Condition 2 none of the 16 participants did. However, when required to use m_{in} for the first time in Search Condition 3, all 16 participants used an expression. In Search Condition 4, where m_{in} was optional, 10 of the 11 participants that chose to interact with the m_{in} canvas used an expression in their initial query.

Participants were asked to rate themselves as Successful, Somewhat Successful, or Not Successful for each task. There was little difference between non- m_{in} conditions (27 Successful) and m_{in} conditions (25 Successful).

There was a noticeable increase in average task time when participants used m_{in} (see Table 1). Across all tasks, task time for each non- m_{in} condition (Conditions 1 and 2) averaged roughly 2 to 2 ½ minutes. However, when required to only use m_{in} in Search Condition 3, average task time across all tasks jumped to an average of over five minutes (note the high standard deviation of 239.8 seconds). In Condition 4 where m_{in} was optional and participants were using it for the second time, task time averaged over 4 minutes – still about twice as long as the non- m_{in} conditions, with the standard deviation being almost double, as well. When we break task time down by task, we still find that the average task time was almost always higher using m_{in}. From

observing the on-screen videos collected during the study, participants spent a significant amount of time correcting symbol and layout recognition errors. In some cases participants appeared to be confused about the state of the recognized expression, and would start entering the expression again, or require assistance from the moderator.

Table 1: Task time by search tool condition (in seconds)

	Mean (μ)	Stdev (σ)
Cond 1 (free), n=15 (n<16 due to a system crash)	147.40	88.87
Cond 2 (online), n=15 (n<16 due to a system crash)	118.87	62.80
Cond 3 (m_{in}), n=16	315.19	239.80
Cond 4 (choice), n = 16 (online/ m_{in}), n=11 Participants using the m_{in} canvas	251.00	168.23
(online/ m_{in}), n=5 Participants not using the m_{in} canvas	51.80	17.69

Average task times for Tasks 2 and 3 are helpful for understanding where m_{in} is useful and where it can be improved. The largest increase in average task time occurred for the Pascal's Triangle task (Task 2), whose expression contained several terms and had the most superscripts in the study. It averaged 134.86 seconds ($\sigma \pm 74.25$) across both non- m_{in} conditions and jumped to an average of 461 seconds ($\sigma \pm 273.28$) across both m_{in} conditions. In contrast, the Binomial Coefficient task (Task 3) whose expression has a smaller number of terms and a relatively simple visual layout, averaged 121.57 seconds ($\sigma \pm 71.28$) across both non- m_{in} conditions and increased to an average of 134.29 ($\sigma \pm 45.05$) seconds across both m_{in} conditions. Note the standard deviations which are lower relative to the same conditions in Task 3 and comparable to non- m_{in} conditions.

In post-study, Likert-scale surveys, 11 out of 16 of participants (69%) agreed or strongly agreed that m_{in} made it easy to enter mathematical expressions, that m_{in} is a valuable tool, and that they would use m_{in} frequently when working on a math problem. 13 out of 16 (81%) agreed or strongly agreed that being able to search using mathematical expressions made it easier to find what they needed. For example, although the expression in the Binomial Coefficient task has a simple visual layout, several participants had difficulty expressing it verbally and textually. Participant quote: "Like 4 choose 2 – that's really hard to 'write' but it knew what I meant and it accurately translated what I was trying to say to it."

In post-study interviews, 12 out of 16 of participants (75%) identified scenarios where they would use m_{in} or could have used m_{in} in the past. This included studying for math tests (in particular, working with Calculus, integrals, complex math problems and expressions with lots of Greek letters), taking notes, collaborating with remote students on assignments, and exporting expressions as image files or LaTeX for use in reports. Also, even though the "upload image" functionality of m_{in} was not demonstrated, several participants expressed really liking that feature, as well as suggesting real-world usage scenarios where they could imagine using that function of m_{in} for distance collaboration on math equations or sharing notes in class.

5. DISCUSSION

We feel that that the increased usages of expressions in queries for Search Conditions 3 and 4 can be attributed to three things:

1. *Affordance of the interface.* The m_{in} interface is designed with a large open, blank area with several tools located horizontally across the top. Upon viewing the interface for the first time, without using it or knowing its function, several participants mentioned how similar in appearance it was to a smartboard. When asked what the purpose of the interface was, the majority said "drawing."

2. *Novelty of the interface.* When provided with a demo of the interface, several participants showed visible and audible signs of being fascinated and impressed and seemed eager to try it themselves. In the post-study interview, about a quarter of the participants who chose drawing over typing cited "new and intriguing" as being their reason for making that choice.

3. *Bridging the "intention gap".* Although participants may have chosen to draw their expressions on the m_{in} canvas because the open space was inviting and because the functionality was new and intriguing, the Binomial Coefficient task contained an expression that was difficult to enter correctly in a standard text search box. In the post-study interview, a few of the participants expressed not knowing how they would have entered it in a text search.

The most surprising result was that participants found query-by-expression useful and identified real-world usage scenarios despite increased task time and little difference in task success when using m_{in}. Again, this finding may be attributed to the possibility that m_{in} helps to bridge the query "intention gap" for non-experts, allowing them to search for expressions as they appeared in the task descriptions. We feel this should motivate developers working on query-by-expression technology to continue their work as it seems query-by-expression may be a missing piece that closes the intention gap in existing math search interfaces.

6. CONCLUSION

We have presented a first study of math search behavior in math non-experts. This study was designed to target the non-expert math searcher as the feeling was that this is not a population familiar with LaTeX or other textual encoding languages, nor with template editors and expression names. Our observations support that view. In our study, search time, expression use, user confidence, user comfort, and user satisfaction were all impacted by introducing the ability to draw expressions in addition to using keywords for search. Participants also indicated that such technology would be useful to them.

We believe that improving the recognition algorithms and results visualizations used in m_{in} can substantially reduce the increased entry times seen in our experiment. Based on study videos and questionnaires, improvements to m_{in} have already been made that may reduce the observed increases in search task completion time when using m_{in} vs. text-based search. For example, we now allow operator shorthands (e.g. '^' for exponents) in text typed directly on the canvas, which are then rendered. Both text and handwritten symbols on the canvas are now rendered into a 'clean' expression using MathJax to produce a simpler and more visually pleasing symbol layout. Handwritten strokes are hidden after recognition to address participant comments about visual clutter in the interface.

Finally, as most participants found query-by-expression through drawing valuable, search interfaces supporting handwritten entry for other diagram types (e.g., chemical diagrams) are worth exploring in the future.

8. REFERENCES

[1] Anthony, L., Yang, J. & Koedinger, K. R. Adapting handwriting recognition for applications in algebra learning, *Proc. Int'l Work. Educational Multimedia & Multimedia Education (Emme'07),* (2007), 47-56.

[2] Anthony, L., Yang, J., and Koedinger, K. R. Evaluation of multimodal input for entering mathematical equations on the computer, *Proc. SIGCHI (CHI'05),* (2005), 1184-1187.

[3] Hu, X., Gao, L., Lin, X., Tang, Z., Lin, X., & Baker, J. *Proc. of Joint Conference on Digital Libraries* (2013)*,* 11-20.

[4] Kamali, S., & Tompa, F. Retrieving documents with mathematical content. *Proc. ACM SIGIR,* (2013), 353-362.

[5] Landy, D., & Goldstone, R. Formal notations are diagrams: Evidence from a production task*, Memory & Cognition 35*, 8 (2007), 2033-2040.

[6] Landy, D., & Goldstone, R. How abstract is symbolic thought*?, J. Experimental Psychology: Learning, Memory, and Cognition 33*, 4 (2007), 720-733.

[7] Misutka, J., & Galambos, L. System Description: EgoMath2 as a Tool for Mathematical Searching on Wikipedia.org, *Int. Computer Mathematics*, LNCS Vol. 6824 (2011), 307-309.

[8] Nguyen, T., Chang, K., & Hui, S. C. A math-aware search engine for math question answering system, *Proc. ACM CIKM,* (2012), 724-733.

[9] Norman, D. User Centered System Design: New Perspectives on Human-computer Interaction. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1986.

[10] Sasarak, C., Hart, K., Pospesel, R., Stalnaker, D, Hu, L., LiVolsi, R., Zhu, S., & Zanibbi, R. min: A multimodal web interface for math search, Symp. HCI and IR (HCIR), Cambridge, MA (online), (2012).

[11] Smithies, S. Novins, K., & Arvo, J. Equation Entry and Editing via Handwriting and Gesture Recognition, *Behavior and Information Technology 20*, 1 (2001), 53-67.

[12] Zanibbi, R. & Blostein, D. Recognition and Retrieval of Mathematical Expressions. *Int'l. Journal on Document Analysis and Recognition 15*, 4 (2012), 331-357.

[13] Zha, Z.J., Yang, L., Mei, T., Wang, M., Wang, Z., Chua, T.S., & Hua, X.S. Visual query suggestion: Towards capturing user intent in internet image search, *ACM Trans. TOMCCAP 6,* 3 (2010).

[14] Zhao, J., Kan, M., & Theng, Y. Math Information Retrieval: user requirements and prototype implementation, *Proc. Joint Conference on Digital Libraries, (2008)*, 187-196.