Equitable Access to Intelligent Tutoring Systems Through Paper-Digital Integration

Nirmal Patel, Mithilesh Thakkar, Bansri Rabadiya, Darshan Patel, Shrey Malvi, Aditya Sharma and Derek Lomas

April 6, 2022

Equitable Access to Intelligent Tutoring Systems Through Paper-Digital Integration

Nirmal Patel[1][0000−0003−1472−4029], Mithilesh Thakkar[1], Bansri Rabadiya[2], Darshan Patel[1], Shrey Malvi[1], Aditya Sharma[1], and Derek Lomas[3]

[1] Playpower Labs, India
[2] Teach for India
[3] Delft University of Technology

Abstract. Intelligent Tutoring Systems (ITS) can only respond adaptively to the digital learning activities of the students. If students are learning offline without any digital devices, they have little or no means to receive personalized learning materials with the help of intelligent systems. This paper proposes a Paper-Digital Integration System that can provide offline learners equitable access to ITS capabilities by looking at their work on paper and giving personalized printable feedback. We analyzed data from a paper algebra assessment of N = 17 students and found mistakes that may generalize and help us offer adaptive paper-based recommendations to students. Our analysis showed us some specific algebra mistakes that may help in providing intelligent feedback.

Keywords: Paper-Digital Integration, Offline Intelligent Tutor, Equitable Access, Offline Learning

1 Introduction

A UNICEF survey in November 2020 found that nearly two-thirds of the school-aged children in the world do not have access to the internet at home [1]. Data show that a significant number of students across the world do not have access to any digital devices. The 2021 Annual Status of Education Report of India showed that over 30% of the students in rural India might not have access to a smartphone at home [2] (India has 250M+ K-12 learners [9]). Paper is a pervasive medium for learning for students with low access to digital learning facilities. Intelligent tutoring systems cannot respond to student learning on paper as what happens on paper is barely accessible to digital systems. The more time students spend learning on paper, the less opportunity intelligent tutoring systems have to support and nurture them. To increase the impact of intelligent tutors on the students, we need to create a data-feedback loop between offline learning and intelligent systems. Once the digital systems have data about offline learning on paper, we can provide students online or offline adaptive feedback in response to their paper learning.

Paper is probably the most significant data source for offline learning in classrooms. Observational assessments also provide data for offline learning, but

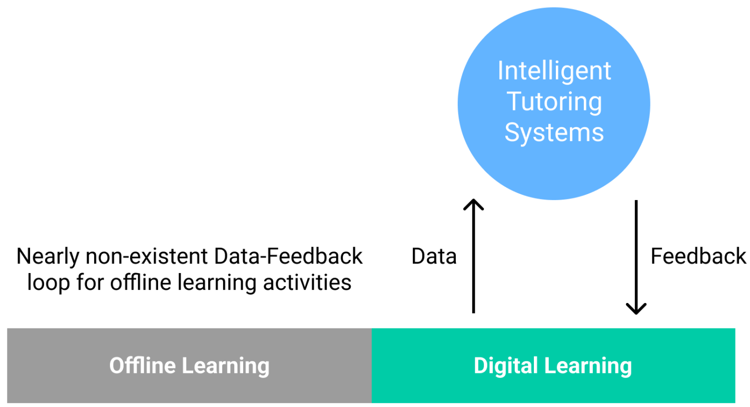

Fig. 1. Intelligent tutors are only equipped to support learning in the digital domain. Observational assessments are probably the only source of student data from the Offline Learning environment.

worksheets, workbooks, and notebooks are likely to contain a much higher quantity of offline learning data. The enormous amount of educational data on paper can enrich learning science. A workshop held at the 2020 NIPS Conference on Machine Learning for Education discussed 'ImageNets for Education', structured datasets that contain images of student work [6]. These structured datasets can help us better understand formative learning at scale. For STEM subjects, large-scale ImageNet for Education datasets can reveal distinct problem-solving patterns and help us identify common misconceptions or knowledge gaps. Large-scale image data also enable new possibilities for building intelligent tutors that can provide adaptive feedback to students by comparing their work with historical examples that have been reviewed by teachers or subject matter experts.

Writing is less likely to be abandoned in favor of digital typing for classroom learning - probably because writing improves outcomes. A recent meta-analysis showed that writing is an effective way to learn science, social studies, and mathematics [7]. A study from 2014 showed that when college students took written notes, they were better at answering conceptual questions [11]. Writing is also the preferred modality for solving math problems. Anthony et al. found that students preferred writing-based input for interacting with intelligent tutors [4]. In a small classroom study, Hinkley et al. observed that when students received digital devices for math learning, they still attempted to interact with those devices as they do with paper [8]. Writing on digital devices with pens is affordable to only a tiny portion of students, and it immediately poses equity challenges as different devices have different processing capabilities. For example, it is easier to solve multi-step math problems on a large-screen tablet than on mobile screens. Paper-digital integration can help us preserve the benefits of writing while keeping digital systems informed about student learning.

The rest of the paper is structured as follows: Section 2 describes prior work on digitizing and analyzing handwritten data. In section 3, we outline our proposed paper-digital integration system. Section 4 describes the algebra assessment we used for our data analysis. Sections 5 and 6 contain image data and their analysis. Later sections discuss the need and impact potential of paper-digital integration.

2 Related Work

Some work has been done to create systems that can give intelligent feedback to student learning on paper. A recent study presented Homework Helper, a system to provide feedback on addition problem-solving [5]. To build the system, the authors collected a training dataset of how students solve multi-digit addition problems on paper and used that data to give new students feedback about their specific types of addition mistakes. Researchers have also described systems to cluster handwritten mathematical equations [13]. Clustering handwritten work can help us identify generalizable problem-solving patterns that can directly inform pedagogical practices.

Several systems have been created to incorporate online handwriting input into intelligent tutoring systems. Anthony et al. conducted a series of studies to develop handwriting input interfaces for cognitive tutors [3]. A recent study found no significant difference in student outcomes when comparing the input modality in a digital intelligent tutor. The researchers in this study compared online handwriting input with the typing-based input [10].

3 Proposed Paper-Digital Integration System

The inability to receive intelligent feedback for offline learning puts marginalized students at a disadvantage for receiving the benefits of Intelligent Tutoring Systems (ITS). We propose a Paper-Digital Integration System to make ITS accessible to students learning offline on paper. To give students personalized feedback for offline paper learning, an ITS needs input data about student learning. We can collect images of offline student work on paper and extract data from them using computer vision algorithms and models. Teachers, volunteers, community members, or school staff can collect images of student work via phones or scanners. We can upload these images to the cloud asynchronously (i.e., immediately or when the internet becomes available). Computer vision programs can extract question-level data from the photos of student work and link them with student data models in the ITS. Handwriting recognition of item response photographs can give us step-by-step problem-solving processes of the students. Based on the student response, the ITS can produce feedback in a printable document form that we can deliver to offline students by printing it first where the internet is available, followed by physical delivery. The feedback can contain corrective remarks on students' mistakes in the photographs, personalized instruction based

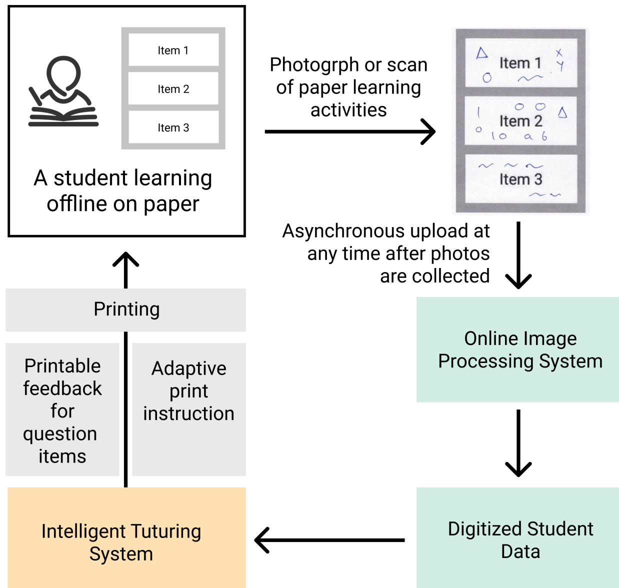

Fig. 2. Proposed Paper-Digital Integration System where ITS can support students who are learning offline.

on student performance, and other paper-based recommendations. The generated print materials can go through the same paper photographing/scanning process and help us provide further individualized feedback to the students. Figure 2 shows an illustration of the described system.

For our proposed system to work, the ITS must reliably collect student learning data from page photographs. To provide students automated individualized feedback for their paper learning activities, we can use the results of the handwriting detection and match them with historical data to find potential automated responses. If the new work does not match any known pattern or historical instances, it can be sent for human review.

4 Research Questions

We attempted to build a proof-of-concept program that mimicked our system's data collection and feedback generation parts. We collected paper assessment data of N = 17 students from a Grade 7 classroom in Ahmadabad, India. We analyzed the scans of the classroom assessment to answer two key research questions:

1. **RQ1**: Can we reliably digitize handwritten question responses on paper (using handwriting recognition)?

2. **RQ2**: Are we able to see any generalizable patterns in the data that can be used to give students automated individualized feedback?

The handwritten algebra test we used for our analysis consisted of 12 questions. There were 5 one-point questions, 3 two-point questions, 2 three-point questions, and 2 four-point questions. Some questions gave students an option to answer either part A or part B. One question asked the students to mention two things they learned in the ongoing topic. We decided to analyze 4 questions that only asked students to solve specific algebra problems that had steps. The test had four pages, and each question had space below it for the response. One student ended up consuming all of the provided space and wrote their answer on a separate sheet of paper. The four problems that we analyzed are written below.

– Q6 - Subtract $3x^2 – 5x + 7$ from $5x^2 – 7x + 9$
– Q7 - Multiply: $(2y + 5)(2y + 5)$
– Q9 - Add $8x^2 + 7xy – 6y^2$, $4x^2 – 3xy + 2y^2$, and $-4x^2 + xy – y^2$
– Q12 - Multiply $(x^2 + 2y)$ by $(x^3 – 2xy + y^3)$ and find the value of the product for $x = 1$.

5 Handwriting Data

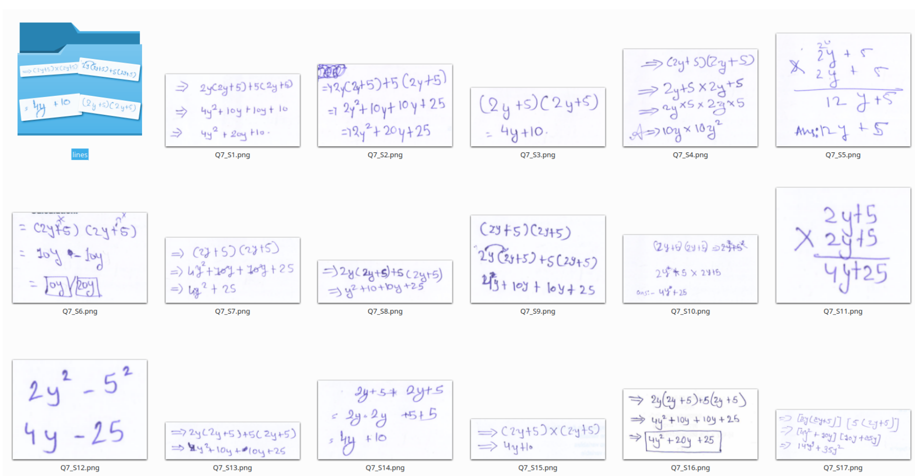

Fig. 3. N = 17 handwritten responses to the question "Multiply: $(2y + 5)(2y + 5)$". The "lines" folder contained the individual lines of each response.

Once the students finished the test, we collected the pages and scanned them using a feed scanner. We scanned 64 pages in total, out of which 4 pages had unacceptable skews produced by the machine. After collecting scans of each page, we manually cropped responses to the selected questions for each student. Afterward, we arranged the data so that all responses to one question were within one folder. This data arrangement immediately allowed us to look at the variance in the item response patterns of the students. Next, we took the images of individual item responses and manually segmented them into lines. Open-source line segmentation algorithms are available through software packages like OpenCV, but they need to be tuned to work on real-world datasets. We did the line segmentation step manually. Once the response data were at the line level, we used a commercially available math handwriting recognition service *Mathpix*. This service provides an API that can take images of equations and return the recognized math script in various formats, including LaTeX and ASCII Math. We used ASCII Math output for our analysis as that was something that we could read and sort easily. Figure 3 shows our collected responses for the question "Multiply: $(2y + 5)(2y + 5)$".

6 Analysis

The Mathpix API for handwritten equation recognition returned a detection confidence metric that allowed us to determine whether the digital handwriting in ASCII Math should be considered for analysis or not. The distribution of the detection confidence in Figure 4 on Page 7 shows that majority of the values had high confidence. Specifically, 62% of the equations had detection confidence of over 90%. We analyzed the confidence values for their face validity and found them reliable, although a bigger and systematic analysis would be required for usage at scale. We also found that equations with scratches had low confidence values, even though most of the symbol recognitions were accurate. Based on our observations, we concluded that the commercially available handwriting recognition capability was likely reliable for a majority of the students, but a systematic analysis is required before using it at scale (**RQ1**).

To answer RQ2, we looked at the step-level data of student responses. We first attempted to cluster the student responses by using string edit distance. We produced one string of response for each student by concatenating the individual steps and then calculated the pairwise distance matrix between the students for each question. Using hierarchical clustering over the distance matrices for the four unique questions did not yield interesting clusters, even though some students had arrived at the same answer in the test. We found the edit distance metric problematic because if one student had written a short response that had the same final answer as another student who wrote a more extended response, the edit distance remained large.

Given that we had a small sample, we decided to tabulate and sort the final steps of the students and see whether any students had given similar answers or had shown similar errors. By sorting the data, we quickly found that for Q6, three students had added the constants of the quadratic terms instead of subtracting them (i.e., students did $5+3$ instead of $5-2$). For Q7, we found that

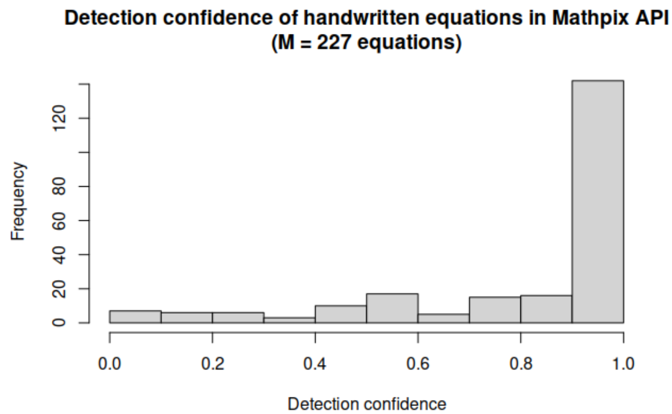

Fig. 4. Distribution of handwritten equation recognition confidence using Mathpix.

five students used an incorrect process for bracket expansion and multiplied the first two numbers, and added/multiplied the last two numbers. Three of these students answered the $(2y + 5)(2y + 5)$ as $(4y + 10)$, while one answered it as $(4y + 25)$ and another as $(4y - 25)$ (this can be verified through Figure 3). Q9 had much variance in the answers, and no students had a similar solution. Still, we saw that some students had mistakenly added the constants of the quadratic terms without considering their signs. This 'misconception' led to five students making a mistake in calculating the first term of the answer (which should be $8x^2$, but two students got $16x^2$ while two got $-16x^2$). For Q12, we found that three students had given the same accurate answer while others had given various types of responses, none of them being true. In summary, we recognized at least one generalizable misconception about bracket opening in Q7 and common mistakes in Q6 and Q8 about adding/subtracting the constants of the quadratic terms. These findings provided a partial answer to **RQ2** and led us to hypothesize that on a larger sample of data, we may see similar errors again and find new generalizable error patterns.

7 Discussion

Our proposed system produced an 'ImageNet for Education' dataset, where we had item responses in the form of handwriting images. We built this dataset for four questions a small number of students. On a larger scale, such a dataset can help us build problem-solving process models that capture the step-by-step process of solving problems typically presented in STEM disciplines. These models can contain aggregated information about how students respond to individual question items. For mathematics and related subjects like physics, where students typically solve problems step-by-step, we may be able to build models that can accurately predict the student's next step given the previous steps. Such models can provide real-time feedback to students solving problems on screens using digital pens.

Nearly all intelligent tutoring systems available today operate in the digital domain. While technology and digital devices may be the future, we need to consider the realities of the present and create ITS that can deliver benefits to all students. If intelligent tutoring systems can be nearly as effective as human tutors [12], then expanding their reach through Paper-Digital Integration is likely to increase the overall impact of ITS research.

8 Conclusion and Future Work

Our study shows how Paper-Digital Integration can extend the benefits of ITS to offline learners. Education technology should be designed to address the needs of all learners. We plan to collect a bigger dataset in the future that contains question items from various topics and subjects and devise algorithms that can give data-driven automated feedback on paper learning.

References

1. Two thirds of the world's school-age children have no internet access at home, new unicef-itu report says (2020), https://www.unicef.org/press-releases/two-thirds-worlds-school-age-children-have-no-internet-access-home-new-unicef-itu
2. India's annual status of education report (rural) 2021 p. 28 (2021), http://img.asercentre.org/docs/aser2021forweb.pdf
3. Anthony, L., Mankoff, J., Mitchell, T., Gross, M.: Developing handwriting-based intelligent tutors to enhance mathematics learning. Unpublished doctoral dissertation, Carnegie Mellon University, USA (2008)
4. Anthony, L., Yang, J., Koedinger, K.R.: A paradigm for handwriting-based intelligent tutors. International Journal of Human-Computer Studies **70**(11), 866–887 (2012)
5. Davis, S.R., DeCapito, C., Nelson, E., Sharma, K., Hand, E.M.: Homework helper: Providing valuable feedback on math mistakes. In: International Symposium on Visual Computing. pp. 533–544. Springer (2020)
6. Garg, K., Heffernan, N., Meyers, K.: Workshop: Advances and opportunities: Machine learning for education (2022), https://nips.cc/virtual/2020/public/workshop_16104.html
7. Graham, S., Kiuhara, S.A., MacKay, M.: The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. Review of Educational Research **90**(2), 179–226 (2020)
8. Hinkley, W., Heffernan, N., Lee Bouygues, H.: The benefits of using pencil and paper in math (2020)
9. Jain, P., Dhawan, A., Ishwaran, G., Jhingan, A.: Private sector's contribution to k-12 education in india (2014)
10. de MORAIS, F., JAQUES, P.A.: Does handwriting impact learning on math tutoring systems? Informatics in Education (2021)

11. Mueller, P.A., Oppenheimer, D.M.: The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. Psychological science **25**(6), 1159–1168 (2014)
12. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist **46**(4), 197–221 (2011)
13. Vu, T.M., Phan, K.M., Ung, H.Q., Nguyen, C.T., Nakagawa, M.: Clustering of handwritten mathematical expressions for computer-assisted marking. IEICE Transactions on Information and Systems **104**(2), 275–284 (2021)