

InkChat: A Collaboration Tool for Mathematics

Rui Hu and Stephen M. Watt

The University of Western Ontario
London Ontario, Canada N6A 5B7
{rhu8, Stephen.Watt}@uwo.ca

Abstract. We investigate the question of how multimodality can be used in computer-mediated mathematical collaboration. To demonstrate our ideas, we present InkChat, a whiteboard application, which can be used to conduct collaborative sessions on a shared canvas. It allows participants to use voice and digital ink independently and simultaneously, which has been found useful in mathematical collaboration.

Keywords: Pen computing, multimodal computing, computer aided collaboration, InkML

1 Introduction

A number of computer applications have been developed over the past years to accommodate the needs of collaboration. One general category of these is whiteboard systems, where multiple users can interact over a shared canvas using a particular input method. For example, the use of pen input allows participants to write or draw naturally, which has great potential to increase productivity, especially in mathematics [1]. Existing systems, however, typically allow only one input method to be used at one time. For example, in Microsoft OneNote [2], one can either type or draw, but not simultaneously. This places strong limitations on what can be done. For example, it becomes quite awkward to explain an activity while it is being performed.

Mathematical collaboration software can be most useful when it supports input from multiple modalities, as it allows collaborators to interact more richly and participants to use the input methods that are most suitable. For example, in mathematics, some points can be most efficiently made through the spoken word while others can best be communicated by a hand-drawn diagram or equation. Software for collaboration also should allow users to perform various editing operations to revise or tidy up jointly created drawings, text and so on. Finally, collaborative work usually includes exploring ideas in discussions that can have false starts and dead ends so the ability to record and roll-back the discussion to prior points is important.

The primary objective of this article is to analyze the design issues in incorporating multimodal interactions in this kind of mathematical collaboration. Considerable related work has been conducted, some of which we highlight here. In the 1990s, QuickSet [3] was a multimodal framework used by the US Navy

and US Marine Corps to set up training scenarios and to control virtual environments. It accepted voice and pen input, communicating via a wireless LAN through an agent architecture to a number of systems. The system could recognize voice input with certain responses. If voice interaction was not feasible, it could still analyze digital ink and then give several possible interpretations. This demonstrated that multimodal interactions could enable efficient communication. Classroom 2000 was an application [4] whose primary purpose was to create an environment to capture as many activities as possible from the classroom experience. It included tools to automate the production of lecture notes and to assist students in reliving lectures. This application did not support real-time distributed collaboration, however. In 2004, the InkBoard [5] network-shared whiteboard application was released. It allowed graphical collaboration and design, including network-shared ink strokes and audio/video conferencing. Because it integrated the Microsoft Conference XP technology, it was limited to Windows platforms. In the same year, Electronic Chalkboard [6] was developed. Its goal was to integrate distance education tools with the traditional blackboard experience. It could load images and interactive programs from a file system or the Internet and could interact with computer algebra systems and display computation results. Because content was saved as images, it was not possible to later edit or perform semantic operations on saved sessions.

To explore ideas for mathematical collaboration, Regmi and Watt [7] developed a whiteboard application that provided collaborative sessions with synchronized voice and digital ink on a shared canvas. This system could save sessions with the digital ink in InkML [8] format and the voice as an MP3 sound file. A significant drawback, however, was that the client interface implementation varied from platform to platform: The client for Windows was implemented in C#, using the .NET framework, while the client for Linux and Mac OS X was implemented in Python. Although Python supports cross-platform portability, the client was constructed using Linux-specific and Mac OS X-specific APIs and thus could not be ported to other platforms.

To address the portability issue, Hu, Mazalov, and Watt [9] proposed a streaming digital ink framework for multi-party collaboration. The framework was portable across multiple platforms and consisted of a number of extensions. These extensions could work independently and simultaneously, serving as plug-ins for the host collaboration software. It is portable across multiple platforms including Windows, Linux, and Mac OS X. It currently uses the popular Skype and Google Talk services as the backbone to deliver data streams, but other transport mechanisms could be used. The digital ink data is represented in InkML, allowing flexible manipulations for different content types, such as mathematics and diagrams. The collaborative sessions can be recorded and stored for playback, analysis or annotation. InkChat is available for download at <http://www.orcca.on.ca/InkChat/>.

The present article explores support for collaboration in this framework. It is based on the same InkChat software infrastructure as [9], but where that article focused on the handling of streaming digital ink, we now focus on the

multimodal aspects. The remainder of the article is organized as follows. In Section 2, we examine how multimodal input can best be used in collaborative environments to improve efficiency. Section 3 recalls InkChat, the whiteboard application for multimodal collaboration. In Section 4, we explain how InkChat supports multimodality. Section 5 describes the collaborative aspects of InkChat. In Section 6, we conclude the article.

2 Multimodal Collaboration

Multimodal input is useful as it provides versatile means for users to interact with computers. These input modalities include keyboard, mouse, voice, pen, and video and so on. We will focus on the voice input and pen input in this section, and explore their capabilities in mathematical collaboration.

Voice Input Voice communication is a fast and natural way to interact with other people and computers. Most people speak faster than they can type or manipulate a mouse. Notably, certain people with physical disabilities prefer operating their computers simply by speaking. Voice input is hands-free, which is useful if one is driving. Also, voice input is flexible. One does not have to sit in front of a computer: it is possible to use voice input while active or while sitting, standing, or reclining. Most modern devices come with built-in microphones.

Handwriting Input With the widespread availability of pen-based devices such as Tablet PCs, PDAs and even cell phones, pen input starts playing an important role in human computer interaction. Pen input is a natural and powerful input modality since everyone learns to write in school. It is versatile as it provides more gestures and motions available compared to mouse and keyboard input. Many devices without pens support touch input, which may also be used to capture handwriting using a finger, but at a lower resolution.

Handwritten input is expressive. Modern devices capture digital ink traces in a two dimensional writing plane, and may support pressures angles and non-contact pen height. This may be used to capture mathematics, as most mathematical notations are two dimensional, with similarities to both text and drawing. Mathematical formulae are hard to understand by means of voice, keyboard or mouse, but can be expressed easily in handwriting.

Voice and Handwriting Multimodal Input We chose voice and pen as the input modalities for collaboration. Together they provide more than either individually, and indeed we find in this combination the whole is more than the sum of its parts. The advantages of voice and pen multimodal collaboration include:

Portability Most computing platforms support both voice and pen input.

Ease of use Writing and voice are familiar and require little training to use.

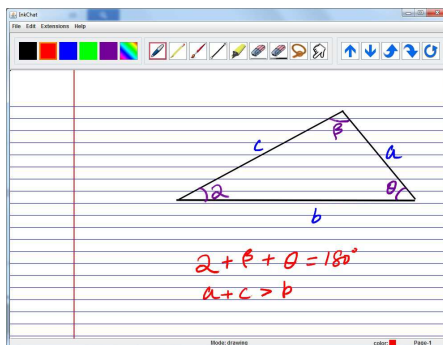


Fig. 1: InkChat user interface

Complementarity Both input modalities can work independently and simultaneously. Speaking and writing at the same time gives two communication channels, allowing one to be used to explain or amplify the other. Either may give the main message, with the other supporting it, or both be equally important. Two channel communication avoids the clutter and confusion that arises when two related messages are multiplexed through one modality. For example, text with footnotes or parenthetical remarks or formulae with many annotations.

3 InkChat

As stated earlier, we have developed a platform-independent version of InkChat to evaluate and demonstrate our ideas. This is built on top of the portable framework presented in [9], whose primary purpose is to collect digital ink across a variety of platforms and to provide a platform-independent, consistent interface for digital ink applications. As a result, InkChat is available on these platforms and can process ink data without knowing the underlying details. To support digital ink data portability, InkChat uses InkML to represent digital ink data as it provides digital ink streaming and archival support independent of platforms. This allows flexible interchange of digital ink data in collaborative environments and, in addition, allows cut and paste of digital ink between applications, e.g. between Microsoft Office 2010 and InkChat.

Figure 1 shows the user interface of InkChat. A number of control buttons are located at the top of the canvas and grouped together to minimize the distance pen movement. InkChat provides a set of pre-defined colors and a palette to create new colors, if needed. To accommodate the needs of different writing activities, we have developed a few brush types. For example, one can choose the pen or pencil for diagramming and the tear drop brush for digital painting or calligraphy. Editing is also supported. This includes redo, undo, and select, cut, copy, and paste of different kinds of content, such as images, typed text, and digital ink.

4 InkChat Support for Multimodality

InkChat is a multi-year ongoing project. Its primary design objective is to enhance mathematical collaboration by incorporating multiple modalities. This allows participants to flexibly choose the input methods that are most suitable for a particular topic. Below we describe the modalities that are supported by InkChat.

Ink Traces Handwriting is one of the most natural ways to input mathematics, as most mathematical notations are two dimensional, with elements of both writing and drawing. InkChat captures handwriting as ink traces and exchanges the data with other clients using InkML.

Voice InkChat also supports voice input. The voice stream is paired with the ink stream to improve the efficiency of collaboration. For example, one can verbally explain the underlying meaning of a complex diagram while drawing it on a shared canvas. This avoids the clutter and confusion that may arise when either input method is used individually.

Floating Pointer To support collaboration, InkChat provides users with floating pointers that can be used to point at target objects on the shared canvas without leaving any ink mark. Together with the voice channel, participants can point to and discuss aspects of the common canvas.

5 InkChat Support for Collaboration

Communication The primary goal of InkChat is to allow users on different computers to collaborate on a shared canvas. It currently uses the popular Skype and Google Talk services for the communication channel, but other transport mechanisms could be used. The primary design principle is to give users the freedom to choose the communication mechanism without too much configuration. If one service is not available in a particular location, it is easy to switch to another. Conference mode is supported, where more than two participants can be involved in one conversation. Depending on the chosen underlying communication service, InkChat adopts different mechanism to exchange data with other participants. For example, when a P2P backbone is used, the conference is initiated by the host that has a connection with every other participant. Digital ink routing shares the same mechanism as audio routing, each piece of ink stroke will be broadcast by the host to all participants except the initiator.

Page Navigation InkChat also supports page navigation. This has been found useful when participants wish to cover multiple topics in one session or to load previous work in the middle of a conversation. In both cases, the current page will first be saved to the file system as an InkML file. Then the Ink Canvas will send a page request to the file system to check if the next page already exists. If so, the Ink Canvas will parse the InkML file and load the content so that users can continue to work on that page. Otherwise, a blank page will be created. Figure 2 illustrates the communications used in page navigation.

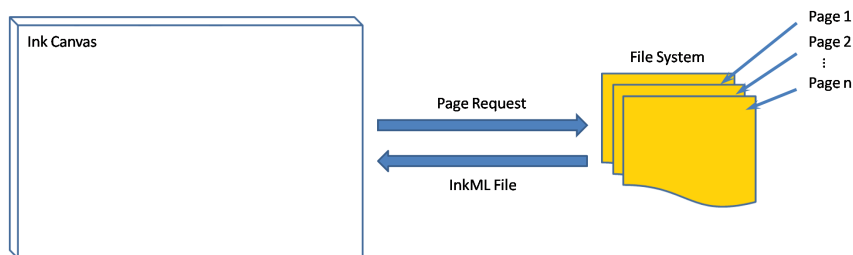


Fig. 2: InkChat page model

Ink Editing In collaboration it is useful to edit or modify a work in progress, and in order to edit, it is necessary to be able to erase digital ink. InkChat provides two ways to erase ink: either by erasing whole strokes or parts of strokes. We call these “stroke-wise” and “point-wise” erasing. Stroke-wise erasing uses a hit testing method to detect whether a particular stroke is selected. If so, it removes the stroke from the canvas and re-renders other strokes that may be affected. Point-wise erasing erases part of a stroke instead of removing the whole from the canvas. A stroke may be split into pieces when using point-wise erasing, and this requires the application to detect where the stroke is broken up. Point-wise erasing uses a hit testing method as well, and this returns a collection of ink points that need to be removed from the target stroke. It then groups the remaining ink points into new strokes and calculates the properties for each, including starting time and duration. The new strokes are then placed in sequence by starting time.

Drag and Drop InkChat allows existing ink to be moved on the canvas using drag and drop. This uses a special lasso cursor to select the content to be moved. This lasso is a free selection tool that allows users to create a selection by encircling a region with a pen. The Lasso is useful in mathematical domains as users may often wish to select a portion of a mathematical expression. Figure 3 shows an example of using Lasso. Notably, as the bounding boxes of character “a” and “+” overlap, a rectangular selection is not suitable for this operation.

Real-Time Mirroring InkChat is able to animate the drawing of ink strokes, and uses this to render the ink of collaborators as it is being written. To avoid jarring and distraction that are caused by large visual changes, InkChat splits long ink traces into small pieces and send each individually. This is to allow smooth rendering on each participant’s canvas. The representation of digital ink data is the key to the success of this animation. Collaboration sessions often take place in heterogeneous environments, where participants may work on different platforms and use various pen devices. These pen devices typically have different settings such as sampling rate, sensitivity, channel properties and so on, and consequently output digital ink data with different characteristics. This requires digital ink data to be represented in a flexible, platform- and vendor-independent

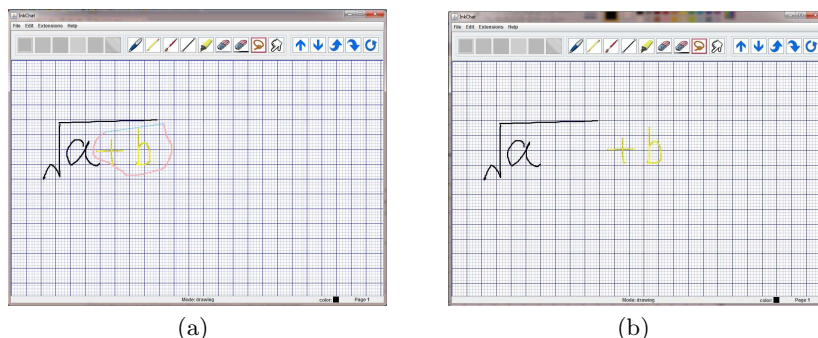


Fig. 3: An example of using Lasso in InkChat. (a) Selection (b) Drag and Drop

format so that the animation is possible across different platforms. Meanwhile, ink strokes must be organized in time order in order to support smooth rendering and synchronization with other modalities.

We have found InkML suitable for these animation purposes. It is platform- and vendor-independent and allows complete and accurate representation of digital ink by capturing and recording information such as the device characteristics, pen tilt, pen pressure and so on. Most importantly, it provides a wide range of features to support smooth rendering and synchronization.

Session Recording and Playback Collaboration sessions may be recorded and stored for later playback, analysis or annotation. InkChat stores digital ink in InkML archival style which keeps the contextual information and ink traces separately in order to achieve compact representation. When playback is desired, the digital ink data can be efficiently converted into streaming style which organizes ink strokes along with contextual information in time order [10]. In addition, each ink trace and its constituent ink points can be timestamped in order to support accurate synchronization with content input by other modalities. Figure 4 shows an example of playback in InkChat.

6 Conclusion and Future Work

We have shown how InkChat has been made to support multimodal interaction and communication. We have found, informally, that these features greatly enhance the InkChat's effectiveness for collaboration. This seems to arise primarily by separating the creation and manipulation of the objects of discourse (diagrams, equations, and so on) from the discussion about the objects and the manipulations. It is an ongoing question of investigation to quantify these findings.

We would like to thank Michael Friesen, Vadim Mazalov, Amit Regmi, and Coby Viner for their contributions to the implementation of InkChat. We would also like to thank James Wake for investigating how InkChat may be integrated in other environments, including Google Hangouts.

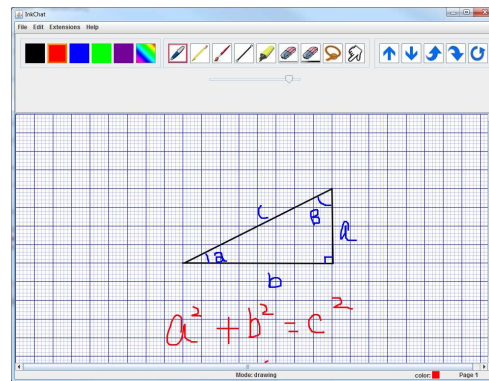


Fig. 4: InkChat animation

References

1. Anthony, L., Yang, J., Koedinger, K.R.: Evaluation of multimodal input for entering mathematical equations on the computer. In: CHI '05 Extended Abstracts on Human Factors in Computing Systems. CHI EA '05, ACM (2005) 1184–1187
2. Microsoft Inc.: Onenote 2010
3. Cohen, P.R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., Clow, J.: QuickSet: Multimodal Interaction for Distributed Applications. In: Proceedings of the fifth ACM international conference on Multimedia. MULTIMEDIA '97, ACM (1997) 31–40
4. Abowd, G.D., Brotherton, J., Bhalodia, J.: Classroom 2000: A system for capturing and accessing multimedia classroom experiences. In: CHI'98: CHI 98 conference summary on Human factors in computing systems, ACM (1998) 20–21
5. Ning, H., Williams, J.R., Slocum, A.H., Sanchez, A.: InkBoard - Tablet PC Enabled Design Oriented Learning. In: Proc. of the 7th International Conference on Computers and Advanced Technology in Education. CATE 2004 154–160
6. Friedland, G., Knipping, L., Rojas, R., Tapia, E.: Teaching with an intelligent electronic chalkboard. In: ETP'04: Proceedings of the 2004 ACM SIGMM workshop on Effective telepresence, ACM (2004) 16–23
7. Regmi, A., Watt, S.M.: A Collaborative Interface for Multimodal Ink and Audio Documents. In: Proceedings of the 2009 10th International Conference on Document Analysis and Recognition. ICDAR '09 (2009) 901–905
8. Watt, S.M., Underhill, T. (Editors): Ink Markup Language (InkML) W3C Recommendation. <http://www.w3.org/TR/InkML/> (September 2011)
9. Hu, R., Mazalov, V., Watt, S.M.: A Streaming Digital Ink Framework for Multi-Party Collaboration. In: Proceedings of the 11th international conference on Intelligent Computer Mathematics. CICM'12 (2012) 81–95
10. Keshari, B., Watt, S.M.: Streaming-archival inkml conversion. In: Proc. 2007 International Conference on Document Analysis and Recognition, (ICDAR 2007), Curitiba, Brazil (September 23-26 2007) pp. 1253–1257