Understanding Diagrammatic Ink in Lecture

Richard Anderson, Crystal Hoyer, Craig Prince, Jonathan Su, Fred Videon, Steve Wolfman
Department of Computer Science and Engineering
University of Washington
Seattle, WA 98195
{anderson, clhoyer, cmprince, jonsu, fred, wolf}@cs.washington.edu

Ruth Anderson
Department of Computer Science
University of Virginia
Charlottesville, VA 22904
ruth@cs.virginia.edu

Abstract

We are interested in understanding how digital ink and speech are used together in presentation. Our long range goal is to develop tools to analyze the ink and speech channels of recorded lectures. As a first step in this process, we are making a detailed study of instructors' digital ink usage in real university lectures. This work is being done in the context of a Tablet-PC based presentation system we have developed, but is applicable to other systems which record digital ink and speech. In this paper we concentrate on how instructors draw and use diagrams in the process of lecture delivery and identify phenomena which are important when automatically processing the diagrammatic ink.

Background

We are studying the use of digital ink and speech in university lectures. Our overall goal is to understand the use of these information channels in order to support the development of better tools for lecture presentation and for the analysis of recorded lectures. This work is being done in the context of presentation systems where the lecturer is using a Tablet PC to deliver a lecture with electronic slides and digital ink. The lecturer is able to write on top of the slides and have the slides and ink displayed to the audience. We have developed one such system for this: Classroom Presenter (Anderson *et al.* 2004a). There are many alternate system and approaches for this including university developed systems such as Classroom 2000 (Abowd 1999) and DyKnow (Berque, Bonewrite, & Whitesell 2004) and commercial applications such as Microsoft's PowerPoint, OneNote and Journal. Our results are not directed toward any particular system - we expect that very similar results will be observed for other Tablet PC-based presentation systems.

In order to automatically process digital ink used in lecture, we need to understand common usage patterns. We want to be able to work with digital ink as it is naturally created, and not to restrict the lecturer's behavior in order to generate an artifact that is easy to work with. The lecturing environment has a significant impact on how ink is used: the lecturer has much of his or her concentration on the exposition and the audience, the ink is often used in conjunction with speaking, and the physical setting for writing is often challenging. These factors lead to phenomena in drawing diagrams that are different from other domains. There has been substantial work in analyzing diagrammatic drawing in other domains (see for example (Alvarado & Davis 2001), (Landay & Myers 2001), (Gross & Do 1996), and (Mankoff, Hudson, & Abowd 2000).) The irregular and ambiguous nature of drawings such as the one shown in Figure 1 has been widely recognized. We consider the main contributions of this paper to be a discussion of phenomena which are specific to diagrams used in spoken communication.

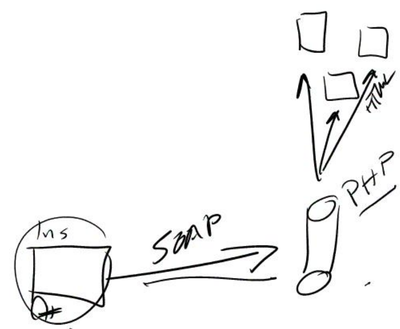

Figure 1: A typical diagram drawn during lecture, showing a software system labeled with technologies.

We have had roughly seventy deployments of Classroom Presenter in university courses. We have concentrated our study on a series of courses offered in our Professional Master's Program. These courses have been taught between two sites using internet based video conferencing. The instructor lectures using a Tablet PC, writing directly on the slides. Synchronized slides and writing are displayed to both the local and remote students. The audio, video, slides, and inking of these lectures are archived, and we have a replay tool available which has allowed us to study these lectures. We have found it critical to have the audio and the dynamic information about the ink for our study.

We have previous work studying ink usage in university lectures. In (Anderson *et al.* 2004c) we introduced a classification of ink into three types: textual ink, diagrammatic ink, and attentional ink. Attentional ink is used in conjunc-

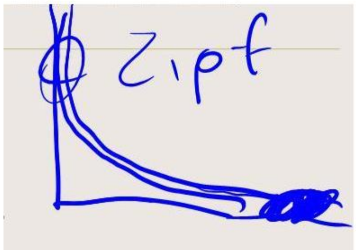

Figure 2: Diagram of a graph of the Zipf curve from a lecture.

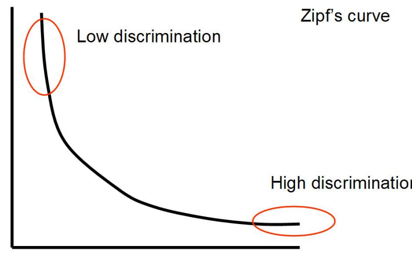

Figure 3: A summary of the diagram in Figure 2, with the attentional ink removed from the curve and the two regions labeled.

tion with speech to tie the spoken utterance to slide content. Examples of attentional ink include the underlines and circles shown in Figure 1. The meaning of this ink is dependent upon the spoken context. Attentional ink represented a significant fraction of the total writing we observed. In lectures that we analyzed in detail, attentional ink comprised 50 to 75 percent of the writing. One of the major themes of the paper was the relationship between how people use attentional marks and linguistic analysis of hand gestures (McNeill 1992). In (Anderson *et al.* 2004b) we took a more detailed look at textual and attentional ink, evaluating opportunities for automatic analysis and identifying links between speech and use of ink. The current paper turns its attention to diagrammatic use of ink. Diagrammatic ink is likely to be the most challenging of the three types of ink to work with because of its wide range and the abstract nature of many diagrams. In this paper we examine a collection of diagrams that arose in real lectures, and identify issues in what it would take to automatically analyze diagrammatic ink.

Ink Understanding

'Ink Understanding' is a broad term. Our approach to the problem of ink understanding is *task oriented* in that understanding means being able to perform specific operations on the ink. Understanding textual ink is being able to convert the ink to text. For attentional ink, the goal is to identify the link between speech and slide content. For diagrammatic ink there is a much wider choice in the operations that we might want to apply. To guide our analysis we picked two hypothetical lecture-specific applications (building static summaries, and simplifying diagrams for note taking) that would be very useful to have. Figure 2 shows an actual diagram from a lecture and Figures 3 and 4 show manually created results of our two hypothetical applications. Our first step in the study was to take a number of examples of diagrammatic ink from lectures, and manually perform the analysis tasks that we would like our hypothetical applications to be able to perform. The goals of this evaluation were to make concrete the potential utility of applications that understand diagrammatic ink and to expose technical challenges in automatically performing these tasks.

Static Summaries Instructors and students often want to have a record of lectures showing the final content of the slides and whiteboard. However, a snapshot of the final ink has a number of drawbacks. First of all, the quality of the writing and drawing is often poor because of the environment it was created in. Second, the diagrams often contain many extraneous marks drawn during the discussion to draw attention to particular points. A static summary would involve cleaning up ink, geometry, and text, removing attentional ink and possibly providing extra context from the audio channel.

Student Note Taking Making the instructor's ink available to students in real time could improve note taking. Instead of copying the material written by the instructor, the students could make their own annotations augmenting the instructors writing. Unfortunately there are a number of difficulties with providing all instructor ink directly to students. First of all, attentional ink is probably not useful in the notes, since it has only temporary value and would just clutter the other ink. Secondly, ink is written continuously, sometimes with erasures, so providing an exact copy of instructor ink to students would not preserve temporal aspects of the ink. Finally, we have observed that instructor's diagrams often are drawn in distinct phases, so instead of providing a single summary, it might be more valuable to create a group of key frames for students to take notes on.

Our methodology for the study was to analyze particular diagrams by manually performing the above tasks. We began by going through our corpus of recorded lectures and identifying diagrams used by instructors. See (Anderson *et al.* 2004c) for background on the data set that we used. We chose representative diagrams from five lecturers and then narrowed the set to ten diagrams. Authors independently analyzed the diagrams, and produced the results of the applications. Lectures were replayed multiple times during the construction of the summaries in order to understand the nuances of some of the ink strokes. The resulting cleaned up diagrams were produced in Visio, PowerPoint, and on paper.

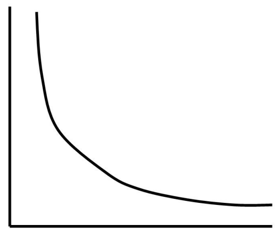

Figure 4: A version of the diagram in Figure 2 that has been simplified for note taking. In this case, all that is made available is the curve and axes so that students can add their own annotations.

Basic Diagramming Behavior

We begin by documenting the basic usage that we observed in the drawing of diagrams, and in later sections discuss some of the more complicated behaviors that demonstrate the richness of the domain. The basic behavior is not surprising: diagrams take on a wide range of appearance and drawing quality is often severely degraded.

The diagrams that we studied were drawn by instructors giving live lectures to students. The instructors were lecturing from a Tablet PC in slate mode and were standing or sitting in front of a podium that supported the Tablet PC at a slight angle. A number of factors impact the writing environment:

- Lecture dynamics: instructors are concentrating on the exposition and are nervous or excited.
- Physical setting: Writing while attempting to maintain eye contact with students is difficult. If the instructor is standing, the writing and viewing angle might be bad. Glare from the lights can be a problem.
- Natural writing: The instructor is writing ink as ink, and does not have the attention to receive any feedback from a recognizer. It is difficult for instructors to perform mode switches or use gestures while lecturing.
- Tablet challenges: Writing on a Tablet PC can be more difficult than on paper because of its slippery surface or unusual pen. The screen area can be too small for writing. Writing near the edge of the tablet can be a problem because of lack of space for resting the hand.

These factors help explain the basic appearance of many diagrams. As a starting example, consider Figure 1. This fairly simple diagram shows many key features of classroom diagrams. The geometric constructs and arrows are somewhat crude, but not difficult to recognize. The diagram labels are harder to read: SOAP, PHP, and HTML are readable for someone with appropriate context. The meaning of “Ins” is unclear (it’s “Instructor Application”) and the writing below the box is “C#”. The circle around the Instructor Application box and the underline under the SOAP arrow are both attentional marks used during the discussion of the diagram. It would probably be difficult to distinguish them from the diagrammatic ink based purely on geometric considerations.

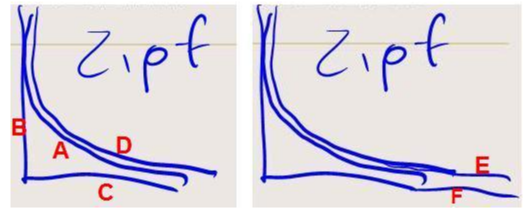

Figure 5: Instructor drawing the Zipf distribution curve. The instructor draws the curve (A), highlights it (D), and later extends it (E).

We now look at a second example of a diagram from lecture, Figure 2 above. This is fairly typical diagram showing a curve with an x and y axis. However, the diagram shows the complexity one is faced with in attempting to automatically analyze diagrams. The instructor was lecturing about word distribution in the English language, and introduced the Zipf distribution. He began by writing “Zipf” and then drew the curve and the axis. Figure 5 shows the curve drawing in more detail, with individual curves labeled A-F. His speech and writing were:

> Zipf was a mathematician who studied curves of this form [Draws the Zipf curve (A)] [Draws x (C) and y (B) axes] - curves with very long tails [Draws curve above(D)]. If you look at the frequency of words in any natural language the frequency follows this kind of curve. The most common words like “a” occur very often and then as you go out more and more rare words [Draws extension of Zipf curve (E)] [Draws extension of x axis (F)] there are fewer and fewer of them.

The instructor then went on to identify the regions of rare words (solid blob) and common words (circle). The diagram presents a number of interesting challenges for analysis. One of these is the distinction between the Zipf curve (A), which was the key part of the diagram, and a later tracing above the curve (D) which was for emphasis. Another challenge were the extensions (E) and (F) of the curve and the x axis. This example shows how the drawing of a very simple object can be quite complex to analyze.

We now describe several observations about diagrams that emerged in our study. We noted that diagrams often occur in phases, have changing focus, and present difficulties in discriminating between attentional and diagrammatic ink. We illustrate our findings with examples from lectures.

Phases

The first of three important observations about diagrammatic ink in presentation is that it is often drawn in phases. By phases we mean that the diagram progresses through several episodes of drawing during a presentation where the diagram takes on different meanings between episodes.

The basic phasing behavior that we observed is that instructors would use diagrams as an evolving collection of static diagrams as opposed to a continuously evolving system. This has a significant impact on algorithms for processing diagrams, raising problems such as how to identify phases and how to analyze the incremental contributions to a diagram. The manual solutions to our benchmark applications often exhibited a breakdown into phases. The coders showed significant consistency in their identification of phases, giving evidence that the phases are natural and well defined.

Our definition of phases is rather broad so we will clarify by analyzing four different examples of diagrams containing phases. These examples were each chosen because they show the diverse circumstances under which phases occur in diagrams.

Concrete Process Diagram We begin with an example of a concrete process diagram, meaning a diagram used to demonstrate the steps in a concrete process. The process being demonstrated was a method for drawing a hexagon. The instructor was illustrating a technique introduced by Sutherland in his seminal 1963 paper (Sutherland 1963).

Figure 6 illustrates the various phases in the demonstration, where each phase represents the result of another step in the hexagon-drawing process. Of particular note are the phases shown in Figures 6d and e, notice that the lecturer erases the red circle between these phases - corresponding to the erasure of the circle in Sutherland's process. This is interesting because by only looking at the final diagram (Figure 6e) there is no way to tell that a circle was there. This means that for diagrammatic understanding the entire drawing process as a whole must be analyzed and understood - it is not enough to just analyze the end result.

As one might expect the phases of a process diagram closely follow the steps in the process. The reason that we consider each step a different phase is because the meaning of the diagram changes between each phase. Specifically the meaning is no longer to demonstrate the previous step of the process, but to demonstrate the current step. The initial Sutherland paper illustrated this process with a group of diagrams which matched the phases quite closely - the difference is that in the lecturer's presentation the phases were temporally separated, while in print a spatial separation was used.

Abstract Process Diagram Our second example shows that phasing behavior also occurs in diagrams used to illustrate abstract processes. Figure 7 shows the four distinct phases of a diagram used to illustrate how to calculate the conditional probability of a node in a Bayes' net. Notice how the lecturer begins the example by labeling the nodes A and B (see Figure 7a). At this point, the lecturer decides to move away from the generic A and B labels to a more specific example involving "Fire" and "Smoke" (see Figure 7b). This starts a new phase because there is a move from the generic example using A and B to the specific example using "Smoke" and "Fire", changing the meaning of the diagram.

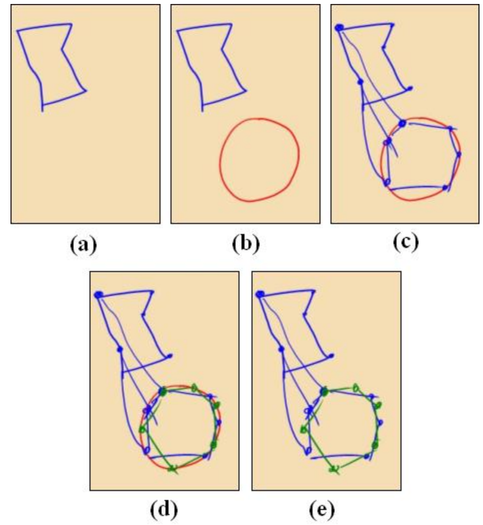

Figure 6: Snapshots of a diagram depicting the five phases in the hexagon-drawing process.

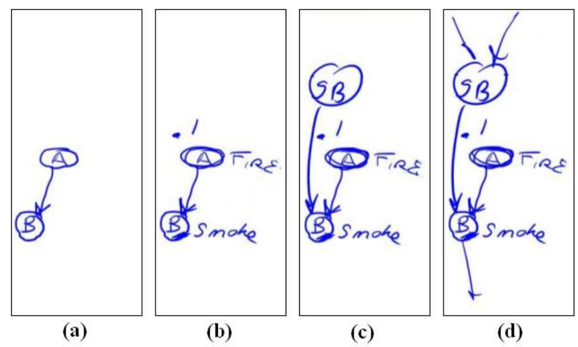

Figure 7: Snapshots of the four phases of a diagram.

The third phase (see Figure 7c) simply explains a different step in the abstract process. The last phase is very interesting since the final arrows drawn on the diagram represent arcs that do not have to be considered, while earlier phases showed values that did need to be considered. The only way we can make this distinction is because the lecturer explicitly states this. If the diagram were viewed outside the scope of the lecture it would be impossible to know the meaning of the arrows in this last phase since the meaning comes from the lecturer's speech. This observation shows the importance and difficulty of understanding the context of diagrammatic ink.

Alternatives Diagram In order to show that it is not just process diagrams where phases occur, our third example gives a different instance of phases in diagrams that has not been discussed yet: how one diagram can have different phases representing different alternatives. Figure 8 shows the various phases of a diagram that a lecturer used while describing three different variations of a speech recognition

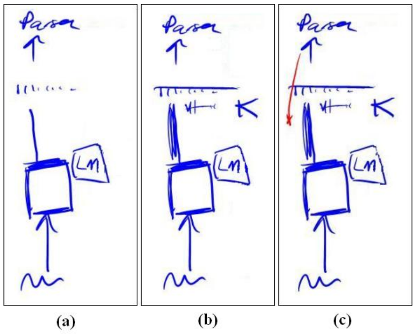

Figure 8: The three phases of a diagram describing a speech recognition system.

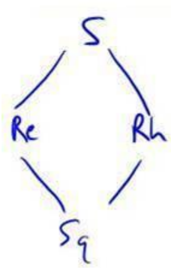

Figure 9: The class hierarchy base diagram.

system. The first phase (Figure 8a) was used to show the basic system, which produced a series of words (the small tick-marks) for the parser. After discussing this basic system, the lecturer went on to discuss a variation, which can return the K most likely series of words, ending with the diagram in Figure 8b. Finally, the lecture drew the red arrow downward to illustrate a type of architecture that has not been widely explored - one that is top-down instead of bottom-up. Notice that the lecturer changed color for this third phase both to show that this variation was unusual and to emphasize that the top-down architecture was very different than the previous two variations.

The biggest challenge in algorithmically analyzing this type of diagram is determining when one alternative ends and the next begins. This challenge is helped by the fact that the lecturer usually spends time discussing one alternative before moving onto the next and so verbal cues could be used for this task.

Reused Base Diagram The final example has a diagram which was used to make several different points. Figure 9 was drawn first by the instructor as the "base diagram" for the remainder of the inking episode. This diagram was a representation of a class hierarchy showing the "diamond inheritance" problem which comes up when implementing multiple inheritance. In the diagram, the 'S', 'Re', 'Rh', and 'Sq' represent different classes in the hierarchy.

The lecturer used Figures 10a, b and c to talk about the ambiguity of three different methods in the class hierarchy.

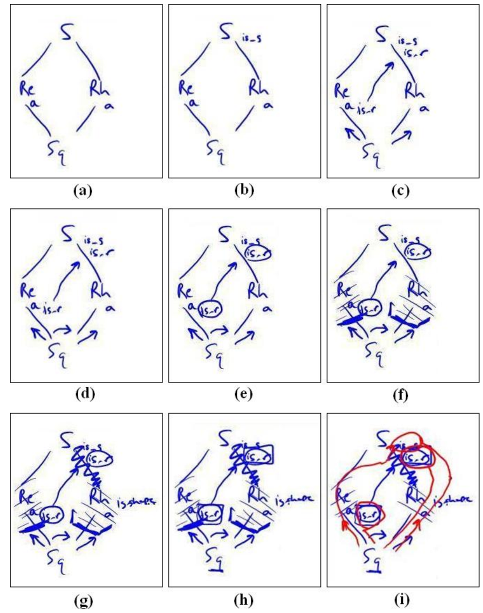

Figure 10: Snapshots of the nine phases of the class hierarchy diagram.

Each of these discussions were distinct and thus belong in separate phases. Figure 10d was the result of a student question, which is another way in which a diagram's meaning can be changed. Next, Figures 10e, f, h and i were each used to discuss four different implementations of multiple inheritance and how they handle the "diamond inheritance" problem. Figure 10g shows the ink resulting from a somewhat different phase that did not give a new implementation of multiple inheritance, but instead showed an additional implication of the implementation discussed in Figure 10f. Another interesting note is that Figures 10h and i, were drawn after the instructor had moved on to other slides and referred directly to content on these other slides. This means that analysis tools cannot only take into account additional content spatially close to the diagram.

This style of diagram reuse occurs very often in the lecture context. Sometimes the "base diagram" is part of the static slide content and other times, like in the previous example, the "base diagram" is drawn by the lecturer. In the latter case, slide content often becomes very cluttered and messy - a difficulty for automatic analysis.

We believe identifying the phases of diagrammatic ink is necessary for ink understanding. In both of our hypothetical applications finding key frames is vital. In the case of static summaries, it is important that the summary be understandable after-the-fact. This requires an understanding of the purpose of the diagram and thus is linked directly to understanding the phases of the diagram. Similarly, in the case of student note taking, the phases provide guidance to students on when a new step or concept is being discussed.

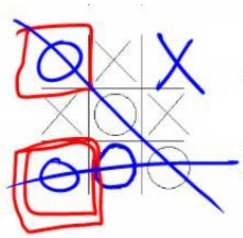

Figure 11: Example of a contradictory diagram.

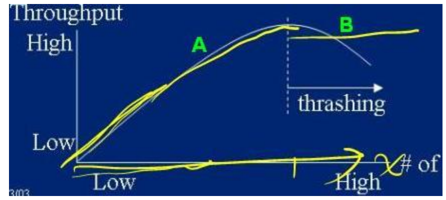

Figure 12: A difficult case for distinguishing between attentional and diagrammatic ink. A is attentional, used for tracing the curve, and B is diagrammatic, since it shows an extension of the original curve.

Locality of Focus

Our second main observation is that discussion can focus on just a portion of the diagram. This allows a speaker to use a diagram to discuss multiple points. The diagram will make sense locally, but may become contradictory or illogical when viewed as a whole.

An excellent example of this is shown in Figure 11 where the lecturer was describing the rules for Tic-Tac-Toe. The instructor first completed the bottom row for O, then also completed the diagonal row.

Clearly according to the rules of Tic-Tac-Toe Figure 11 is an illegal board configuration; however, the point was to show that there are multiple winning moves for O. Locally, when only one of the rows of O's is considered the diagram makes sense, but if viewed as a Tic-Tac-Toe game, it violates the rules. This sort of locality of focus is common in diagrams where parts of the diagram become obsolete. The audience easily understands when part of a diagram is no longer of interest, but for a computer this would be a tremendous challenge since it would require an understanding of the context in which the diagram was drawn as well as the diagram's purpose.

While this locality of focus is often intentional as in the previous example, oftentimes it is the dynamic nature of diagrammatic ink that results in this locality of focus. Consider the last example in the previous section. In the seventh phase (Figure 10g) above, the lecturer drew a squiggle, crossing out a line on the diagram. Locally this made sense because the lecturer wanted to describe a scenario where the link didn't exist. Notice, later as shown in Figure 10i, the lecturer then uses the red ink and draws arrows from the bottom of the diamond to the top, one up the right side of the diamond and one up the left side. When drawing these arrows the lecturer assumed that all the original links in the diamond still existed. This is in direct contradiction to the still visible squiggle, which got rid of one of the links.

For a final example of locality of focus, consider Figure 7d. In this example the final arcs are drawn to illustrate arcs that need not be considered. A verbal distinction is made between two types of objects which are visually indistinct. This pattern has been observed fairly frequently, where negative examples are drawn with positive examples, with only a verbal phrase such as "this can't happen" given to indicate the distinction.

Attentional Ink

Finally, attentional ink plays a key role in exposition using diagrams. Instructors frequently reference components of diagrams by circling or underlining. This creates obvious challenges in recognizing the geometry of the underlying diagrams. Another behavior we observed is retracing components of the diagram for emphasis. A specific challenge is to distinguish between diagrammatic and attentional ink. In this section we look at attentional ink and present several hard cases for distinguishing between attentional and diagrammatic ink.

We begin with an example where it would be difficult to automatically distinguish between attentional and diagrammatic ink strokes, although there is no ambiguity in the stroke types. Figure 12 shows a diagram that an instructor was using to talk about system behavior at capacity. The diagram shows that throughput falls off after full capacity is reached, instead of flattening out at maximum capacity. The instructor demonstrates this by drawing curve B, which shows a hypothetical flattening out of the graph. This is diagrammatic ink. However, the overwriting of the curve (A) is attentional ink used to trace the curve during discussion. The distinction between which lines are attentional and which are diagrammatic is made by having an understanding of their meaning through the associated speech.

In our earlier work, we took the view that ink could be classified into the three types: attentional, diagrammatic, and textual, and that the boundary between these types was well defined, even if there were challenges in algorithmically computing the separation. Our experience in this study has caused us to retreat somewhat from that view, as we have observed ink that occupies a middle ground - partially attentional and partially diagrammatic. Figure 13 shows ink that falls between two types. The explanation mark showed the focus of attention, and also linked together three separate bullet points - both satisfying the definition of attentional ink. However, the mark also had lasting meaning, independent of the speech, indicating the importance of the bullet points. The circles drawn in the Zipf diagram in Figure 2 are also intermediate between attentional and diagrammatic. They can be viewed as attentional in that they are pointing to a region in a diagram, and they can be viewed as diagrammatic because they are defining regions in the diagram. Our coders were not consistent on whether or not to show the circles in the summaries and notes. One of the difficulties

Figure 13: The exclamation mark exhibits characteristics of both attentional and non-attention ink: it draws attention and links together bullet points, but it also has a persistent meaning.

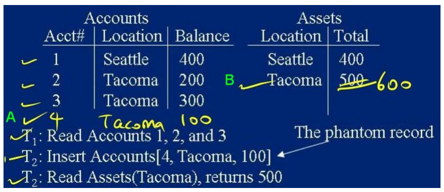

Figure 14: Ink changing meaning: the check marks were initially drawn as being attentional, but later, marks A and B were removed to show data begin unlocked.

is that in a diagram a reference to a component can become a part of the diagram. In Figure 5, the stroke that extends the curve (E) is also a difficult stroke to classify. Based on the instructors speech, it was classified as attentional, since the instructor was talking about going further out along the curve, and this was drawing attention to the process. However, it could also be viewed as directly extending the curve, a view that is supported by the x-axis being extended directly. Another interpretation is that the stroke started out as being attentional, but then became diagrammatic after it was written.

Figure 14 shows a fascinating example where the ink changes type. The instructor is demonstrating a result in transaction processing where a series of locking operations lead to an unexpected state. The setup involves setting a collection of locks, and then releasing the locks. In setting the locks, the instructor makes check marks one by one, marking off the operations being performed, without identifying the check marks as locks. The check marks are applied both to the data items, and to the transactions. When transaction two completes, the check marks in front of Acct 4 (A) and Tacoma (B) are erased, indicating that the locks are released. The check marks started out as being attentional, but were erased as if they were diagrammatic ink. The interesting aspect of this is that the instructor opportunistically took advantage of the ink, changing its meaning to illustrate the act of unlocking.

Conclusions

This paper presents the results of our detailed examination of instructors' use of digital ink while diagramming during lecture. We see this study as a first step towards both: revealing the potential utility of applications that are able to 'understand' diagrammatic ink and identifying specific difficulties inherent in arriving at that understanding. Three particular practices were identified in our study:

- Drawing diagrams in phases: Instructors would frequently use a diagram in multiple phases, where additional diagrammatic ink was added at each phase. This raises the problems of identifying these phases and understanding incremental use of ink.
- Locality of focus: Often, only portions of a diagram are used to make individual points. In the process, a diagram will become globally inconsistent.
- Rich use of attentional ink: Attentional ink used in explanation can take on additional meaning, or even change meaning in the process of exposition. This increases the challenge of the ink classification tasks which are often the first step in recognition.

Acknowledgments

We thank the instructors who made their lecture materials available to us for study. We also thank Microsoft Research for support in this collaborative research.

References

Abowd, G. D. 1999. Classroom 2000: an experiment with the instrumentation of a living educational environment. *IBM Systems Journal* 38(4):508–530.

Alvarado, C., and Davis, R. 2001. Resolving ambiguities to create a natural computer-based sketching environment. In *Proceedings of IJCAI-01*.

Anderson, R.; Anderson, R.; Simon, B.; Wolfman, S. A.; VanDeGrift, T.; and Yasuhara, K. 2004a. Experiences with a Tablet PC based lecture presentation system in Computer Science courses. In *Proceedings of SIGCSE'04*, 56–60.

Anderson, R. J.; Hoyer, C.; Prince, C.; Su, J.; Videon, F.; and Wolfman, S. A. 2004b. Speech, ink and slides: The interaction of content channels. In *Proceedings of twelfth ACM International Conference on Multimedia*.

Anderson, R. J.; Hoyer, C.; Wolfman, S. A.; and Anderson, R. 2004c. A study of digital ink in lecture presentation. In *Proceedings of CHI'04*, 567–574.

Berque, D.; Bonewrite, T.; and Whitesell, M. 2004. Using pen-based computers across the computer science curriculum. In *Proceedings of SIGCSE'04*, 61–65.

Gross, M. D., and Do, E. Y.-L. 1996. Ambiguous intentions: a paper-like interface for creative design. In *9th annual UIST*, 183–192. ACM Press.

Landay, J. A., and Myers, B. A. 2001. Toward more human interface design. *IEEE Computer* 34(3):56–64.

Mankoff, J.; Hudson, S. E.; and Abowd, G. D. 2000. Providing integrated toolkit-level support for ambiguity in recognition-based interfaces. In *Proceedings of CHI'00*, 368–375.

McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. The University of Chicago Press.

Sutherland, I. 1963. Sketchpad: A man-machine graphical communication system. In *Proceedings of the 1963 Spring Joint Computer Conference*, 329–346. Spartan Books.