A Study of Diagrammatic Ink in Lecture

Richard Anderson [a] Ruth Anderson [b] Crystal Hoyer [a]
Craig Prince [a] Jonathan Su [a] Fred Videon [a] Steven Wolfman [c]

[a] *Dept. of Computer Science and Engr., Univ. of Washington, Seattle, WA 98195*

[b] *Dept. of Computer Science, Univ. of Virginia, Charlottesville, VA 22904*

[c] *Dept. of Comp. Sci., U. of British Columbia, Vancouver, BC, Canada, V6T 1Z4*

Abstract

In this paper, we present a study of how instructors draw diagrams in the process of delivering lectures. We are motivated by wanting to understand challenges and opportunities for automatically analyzing diagrams, and to use this to improve tools to support the delivery of presentations and the viewing of archived lectures. The study was conducted by analyzing a large group of examples of diagrams collected from real lectures that were delivered from a Tablet PC. The main result of the paper is the identification of three specific challenges in analyzing spontaneous instructor diagrams: separating the diagram from its annotations and other surrounding ink, identifying phases in discussion of a diagram, and constructing the active context in a diagram.

1 Background and Study Details

Delivery of presentations using electronic tools is becoming prevalent. Advantages of electronic delivery include the high quality of displayed materials, ease of reuse, and the ability to share materials across machines and archive artifacts of the presentation. The technology to support presentation is rapidly advancing, in particular, there is growing use of digital ink with electronic slides. In many situations, the ability to draw spontaneous diagrams to support exposition greatly enhances communication. In this paper we explore

Email addresses: `anderson@cs.washington.edu` (Richard Anderson), `ruth@cs.virginia.edu` (Ruth Anderson), `clhoyer@cs.washington.edu` (Crystal Hoyer), `cmprince@cs.washington.edu` (Craig Prince), `jonsu@cs.washington.edu` (Jonathan Su), `fred@cs.washington.edu` (Fred Videon), `wolf@cs.ubc.ca` (Steven Wolfman).

properties of diagrams that are drawn naturally while delivering lectures. Our interest is to gain an understanding of drawing practices to inform the development of improved tools to support electronic presentations.

Our domain of study is university lectures delivered by an instructor writing on electronic slides with digital ink on a Tablet PC. We collected data using Classroom Presenter (Anderson et al., 2004a), a system that we developed. However, there are many other systems that provide similar functionality. These include university developed systems such as eFuzion (Peiper et al., 2004) and DyKnow (Berque et al., 2004) and commercial applications such as Microsoft's PowerPoint, OneNote and Journal.

Basic support for digital ink in lecture is provided by all of these systems — high quality ink over lecture slides displayed in real time to the audience — but these systems could all benefit from automated analyses of ink that enhance the in-class experience or interaction with archived materials. For example, identifying and removing ink with transient meaning would make it easier for students to view the ink with persistent, important meaning. Similarly, selectively rerendering the most critical ink to improve legibility (e.g., by exploiting spoken or displayed context) would aid students' learning. Once recognized, text and diagrams may also be rerendered to make them accessible for students with limited vision. Tidying messy text and simplifying diagrams would facilitate student note taking, both by removing extraneous detail and by freeing up space for students' own annotations. Together these and other analyses could be employed off-line to transform the live lecture experience into a static, browsable summary suitable for review (e.g., to be posted online after the lecture). Indexing and search are key use cases for both static summaries and standard archives.

All of these operations require understanding instructors' ink. Before pursuing automatic analysis of this ink, we wanted to understand the types of inking that occur. In previous work (Anderson et al., 2004c), we identified three key types of inking that instructors employ: "attentional" ink that ties spoken utterance to slide content (e.g., the underlines and circles shown in Figure 1), standard textual annotations, and diagrammatic ink. We have already explored styles of use and strategies for understanding attentional and textual ink (Anderson et al., 2004b). In this study, we address the key remaining annotation type by investigating how instructors use digital ink to construct diagrams in lectures. We examined a large number of diagrams that occurred naturally in university lectures, with the goal of identifying general patterns of diagrammatic writing that would indicate challenges and opportunities for automatic analysis.

In this paper we identify three key phenomena we have observed in instructors' use of diagrammatic ink during lecture. Our observations are from actual

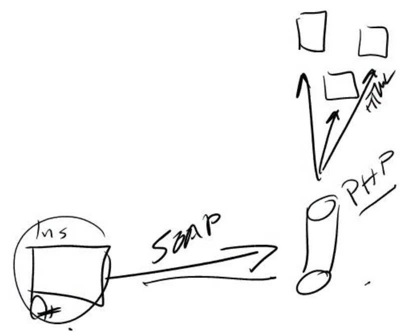


Fig. 1. A typical diagram drawn during lecture, showing a software system labeled with technologies. The underline beneath "PHP", the underline beneath the "SOAP" arrow, and the circle around the left-hand box are attentional ink.

classroom use of our system. We are not aware of other field studies of this sort, although a few researchers have collected a body of sketch data from other domains. The Classroom 2000 project (Abowd, 1999) collected a vast amount of data on ink use in lecture, but did not publish results on diagrammatic ink. Oltmans et al. (2004) used a Tablet PC to collect a corpus of sketch data including circuit diagrams, floor plans, family trees, and geometry. Similarly, Eisenstein et al. (2004) videotaped subjects describing mechanical devices at a whiteboard.

Classroom Presenter has been deployed in an estimated two hundred university courses. For this study we concentrate on a series of computer science courses offered in the Professional Master's Program at the University of Washington. These courses were taught between two sites using internet based video conferencing where the instructor lectured using a Tablet PC, writing directly on the slides. The Tablet PC was in slate mode (i.e., keyboardless) and instructors were standing or sitting in front of a podium that supported the tablet. Synchronized slides and writing were displayed to both the local and remote students. Audio, video, slides, and ink from these lectures were archived, allowing us to study ink use in great detail. In particular, the audio and the dynamic information about the ink has been critical for our study. Our collection consists of archives from six courses, totalling over 180 hours of classroom use.

In the remainder of the paper we elaborate on several behaviors we observed while analyzing instructors' use of diagrammatic ink in lecture. First we observe that instructors create diagrams that are difficult for automatic analysis because of the spontaneous, interactive nature of lectures - especially because such diagrams are intermixed with attentional ink. Second we observe that instructors frequently create diagrams in phases. Third we find that instructors reuse diagrams in ways that make sense when viewed locally in terms of the current active content; although, the diagram as a whole may become incomplete or inconsistent.

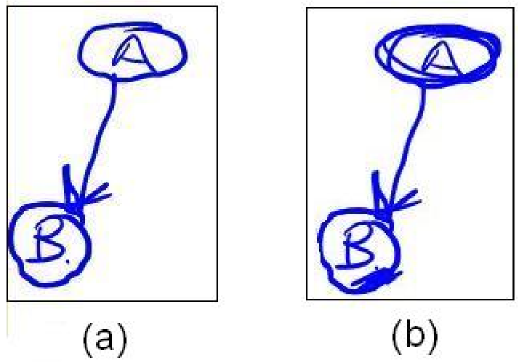


Fig. 2. Attentional marks drawn on diagrams. a) shows diagram before attentional marks, b) shows circling of node A, and underline of B during discussion.

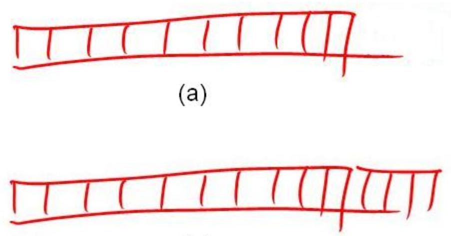
2 Impact of the Lecture Environment on Diagram Complexity

Diagrams created spontaneously by instructors during lecture are likely to be significantly different from the same diagram that might be drawn, for example, for a set of prepared lecture notes. Factors impacting the writing include:

- Lecture dynamics: instructors are concentrating on the exposition and are often nervous or excited. Diagrams are frequently drawn in reaction to student questions without much planning.
- Physical setting: Writing while attempting to maintain eye contact with students is difficult causing reduced attention to writing. The writing and viewing angles of the Tablet PC also contribute to difficulties in writing.
- 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 and the placement of other content can limit space. Writing near the edge of the tablet can be a problem because of degradation of the digitizer's accuracy and lack of space for resting the hand.

These factors help explain the "less than perfect" appearance of many diagrams. Figure 1 is typical of diagrams we observed. The geometric constructs and arrows are somewhat crude, but not difficult for a human to recognize. The diagram labels are harder to read: SOAP, PHP, and HTML are recognizable for someone with appropriate context. The label "C#" below the box is difficult to identify as text even for someone who is familiar with the name. Ad hoc abbreviations (e.g. "Ins" for "Instructor Application") add to the confusion.

A feature common to lecture diagrams is their annotation with attentional ink. We have previously noted that attentional ink represented a significant fraction (50–75%) of the total writing during lecture (Anderson et al., 2004c).



(a)

(b)

Fig. 3. Instructor extending a diagram of an array. a) shows the initial array, and b) the segment added immediately after the completion of the first part.

Attentional marks are writing used to provide a link between spoken utterance and slide content. They can reference either static content on slides or instructor's writing, although in this paper, we are primarily concerned with the attentional marks that occur on hand drawn diagrams. The circle around the "Instructor Application" box and the underline under the SOAP arrow in Figure 1 are both attentional marks used during discussion of the diagram. These marks would be difficult to distinguish from the diagrammatic ink based purely on geometric considerations. Figure 2 gives another example of attentional marks on diagrams. In this case, the instructor redrew the upper circle while discussing the example, and then underlined the "B" in the lower circle. Again, this extra ink is very different from the initial diagrammatic content, but similar in geometry. We argue below that it is important to be able to separate diagrammatic from attentional ink, and pose this as one of our challenge problems.

Another feature we have observed in instructor drawn diagrams is that structures are often drawn in a compound fashion, with multiple strokes being used to create a single line and parts of a diagram extended as they are drawn. Figure 3 shows an example of this. The instructor was drawing an array. The instructor first drew the cells on the left and then extended the array with the four cells on the right, adding an additional segment to the top of the array. The diagram was drawn as a single episode. These types of drawings create difficulties for naive recognition approaches. The natural approach for recognizing an array is to look for two roughly parallel lines with perpendicular segments creating cells. However, in this case multiple segments are used to create the upper line, and vertical segments were not always connected to horizontal segments. Similar issues have been observed in other types of naturally drawn diagrams by Oltmans et al. (2004).

We now discuss an example diagram in greater depth to highlight some of the behaviors described above and how they link to speech. Figure 4 shows a diagram with a curve drawn on an x and y axis. 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

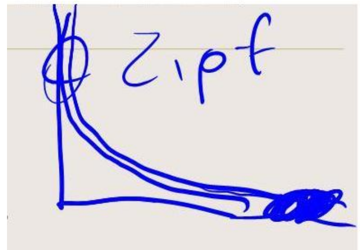


Fig. 4. Diagram of a graph of the Zipf curve from a lecture.

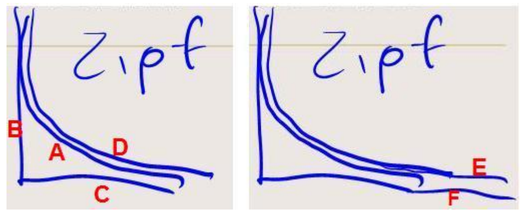


Fig. 5. Instructor drawing the Zipf distribution curve. The instructor draws the curve (A), highlights it (D), and later extends it (E).

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 y (B) then x (C) 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), arriving at the final diagram (Figure 4).

This 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 attentional ink used for emphasis. The circles drawn in Figure 4 are also problematic in that they act as both attentional and diagrammatic ink. They are attentional in that they resolve deixis, linking a visual region in the diagram to its verbal description. However, they are also diagrammatic in that they refine the diagram by establishing meaningful new regions of the graph.

Separating attentional from diagrammatic ink is a key first step in analyzing diagrams. Many approaches to recognition are severely compromised by extraneous lines or symbols, and the geometric appearance of attentional marks make them particularly likely to confuse the diagram recognition algorithms. The separation of attentional from diagrammatic ink is also important for many tools to support lecture delivery and analysis. Attentional marks often have ephemeral meaning, and a number of researchers have suggested rendering the ephemeral ink in a different manner than other ink, for example, by using a different color, or by having the ephemeral ink fade over time (Qian and Gross, 1999). If separate pens are given to the speaker for the different types of inking, then it is straightforward to handle the rendering differently, however, in domains such as lecturing, users have substantial difficulties in switching modes for different types of writing, so automatic segmentation is preferred.

Previous work on ink classification has primarily addressed separating diagrammatic ink from textual ink (Shilman et al., 2003). Results on that problem are moderately good and classification routines are available in standard commercial packages [1], although there are still inherently ambiguous cases, such as an isolated 'O', which could be either the letter Oh, or a circle. A separate context for recognizing attentional ink is writing with respect to static slide content, where attentional ink will include check marks, underlines, and circling of existing content. Recognizing attentional ink in this case is made substantially easier by taking slide content into account (Anderson et al., 2004b). Our problem of interest is separating attentional ink from the diagrammatic and textual ink. The central challenge is the geometric ambiguity between attentional ink and diagrammatic ink. To resolve this, it will be necessary to take advantage of other information. One possibility is to use information about the domain of the diagram to place global constraints on the drawing. Another very important source of information is timing of the strokes, since attentional marks are often isolated in time and drawn after the diagram. Another interesting direction to pursue is to take advantage of speech information to do a multimodal analysis.

3 Phases

Our second observation 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 a single physical diagram to represent an evolving set of static diagrams each with different meanings and/or purposes. This has a significant impact on algorithms for processing diagrams, raising problems such as how to identify

[1] For example, routines for separating textual ink from diagrammatic ink were included in the 1.5 release of the Tablet PC SDK.

phases and how to analyze the incremental contributions to a diagram. When manual analysis of diagramming episodes were performed they almost always exhibitted a natural breakdown into phases. This was especially true when such analyses were done with our example applications in mind. Different authors performing these manual analyses independently showed significant consistency in their identification of phases, giving evidence that the phases are natural and well defined.

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

3.1 Concrete Process Diagram

We begin with an example of a *concrete process diagram*, meaning a diagram used to demonstrate an actual process taking place. 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 on pen computing (Sutherland, 1963).

Figure 6 shows the phases in the demonstration, where each phase is 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 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. Identifying and understanding such mappings between spatial separation and temporal progression will be key to some ink understanding applications (like generation of static summaries).

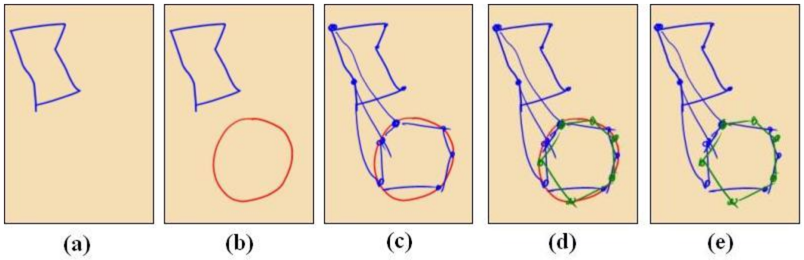


Fig. 6. Snapshots of a diagram depicting the five phases in the hexagon-drawing process.

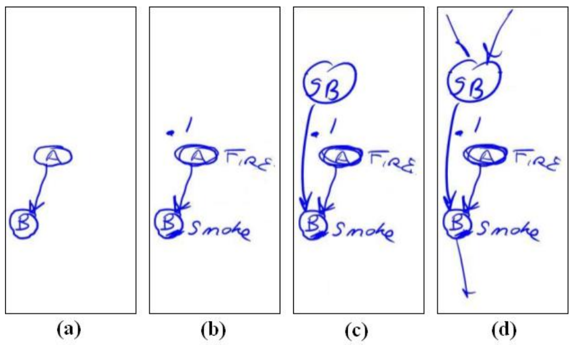


Fig. 7. Snapshots of the four phases of a diagram.

3.2 Abstract Process Diagram

Our second example shows that phasing behavior also occurs in diagrams that illustrate abstract processes. Figure 7 shows the four distinct phases of a diagram illustrating 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). In the second phase, the lecturer switches from a generic discussion to a specific example using "Fire" and "Smoke" (see Figure 7b). The third phase (see Figure 7c) shows another step in the abstract process. The last phase is interesting because 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 expression of this distinction was in the lecturer's verbal comments. If the diagram were viewed without the context of the verbal commentary the arrows in the last phase would become meaningless or, worse, misleading. This observation shows the importance and difficulty of understanding the context of diagrammatic ink.

3.3 Alternatives Diagram

Our third example of phases illustrates alternative (i.e., parallel) choices rather than sequential steps. Figure 8 shows the various phases of a diagram that a

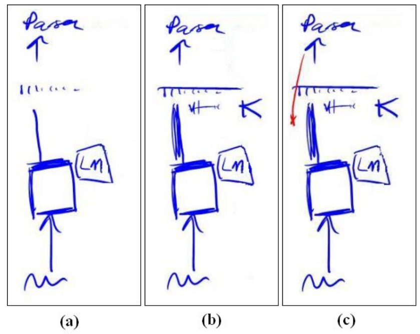


Fig. 8. The three phases of a diagram describing a speech recognition system.

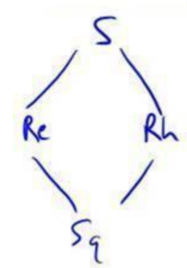


Fig. 9. The class hierarchy base diagram.

lecturer used while describing three variants of a speech recognition system. The first phase (Figure 8a) shows the basic system, which produces a series of words (the small tick-marks) for the parser. After discussing this basic system, the lecturer went on to discuss a variant which returns the K most likely series of words, ending with the diagram in Figure 8b. Finally, the lecturer 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 variant was unusual and to emphasize the difference between the top-down architecture and the previous variants.

The biggest challenge in algorithmically analyzing this type of diagram is determining when one alternative ends and the next begins. A key consideration is that the lecturer usually spends time discussing each alternative; so, verbal and temporal cues may inform the segmentation task.

3.4 Reused Base Diagram

The final example is a diagram that was reused 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 that comes up when implementing multiple inheritance. In the diagram, the 'S', 'Re', 'Rh',

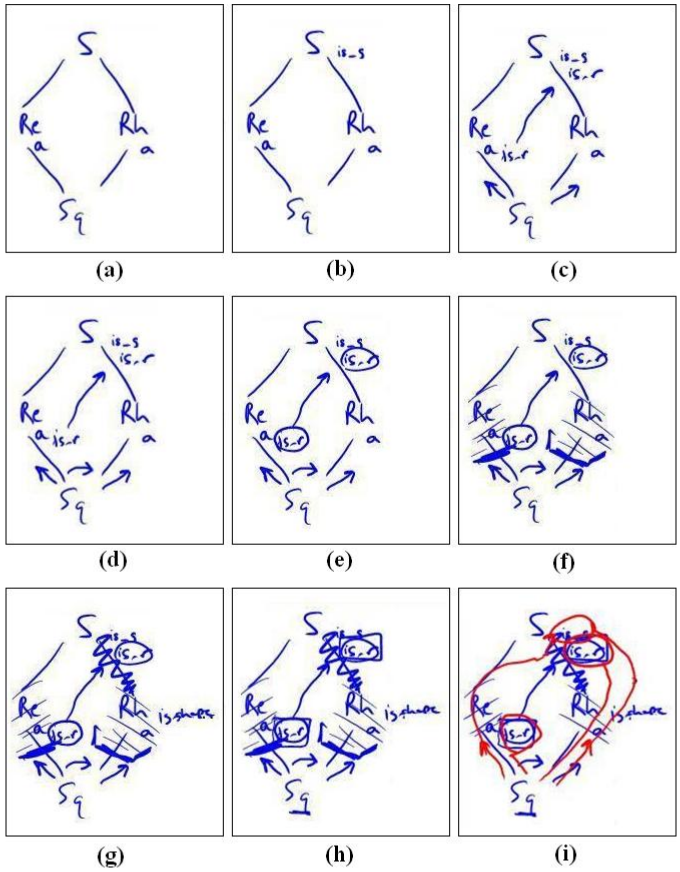


Fig. 10. Snapshots of the nine phases of the class hierarchy diagram.

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. 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. Such behavior further complicates analysis by drawing in content that is distant from 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 cluttered and messy, complicating automatic analysis.

We believe identifying the phases of diagrammatic ink is necessary for ink understanding. In all our potential applications finding key frames is vital. For example, 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. Furthermore, as noted above, there is a natural correspondence between the temporal progression used for dynamic, phased diagrams and spatial separation used in static, phased diagrams.

4 Locality of Focus

Our final major 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. Unlike the previous examples this example was not from a course lecture but came from an experimental study where instructors were asked to prepare and give a lecture using the Tablet PC from the same set of slides. While this scenario is less natural, this behavior is not unusual.

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. However, the lecturer later (Figure 10i) uses red ink and draws arrows from the bottom of the diamond

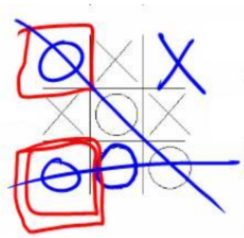


Fig. 11. Example of a contradictory diagram.

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 despite the still-visible squiggle indicating removal of one link.

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.

The challenge that arises is automatically identifying which ink is relevant at a given time, or identifying the “active context”. This appears to be a very hard problem - although it is something that people can do fairly well. In observing instructors talking about diagrams, we had little difficulty in focusing on the relevant ink. It is possible that timing, speech, and deictic cues can be used to help identify which content is being referred to.

5 Conclusions

In this paper we have reported our observations on the use of diagrams in a series of computer science courses. The diagrams were often irregular, and were used in a complex process of exposition. From the analysis we identified three core problems for the analysis of naturally occurring diagrams. These problems are very challenging, but progress on them will allow automatic analysis of lecture diagrams, which will enable improved tools for supporting lecturing and viewing of archived lectures.

- Ink classification: The spontaneous and interactive nature of lectures makes diagram understanding difficult in this domain. A particularly important problem is separating the diagram from its annotations and other surrounding ink. This problem is challenging since the different types of ink often are geometrically similar.

- Diagram phasing: Instructors frequently use a diagram in multiple phases, with additional diagrammatic ink added at each phase. The problem is to identify the "key frames" and extract a sequence of diagrams corresponding to exposition.
- Locality of focus: Often, only portions of a diagram are used to make individual points, while the overall diagram becomes inconsistent. The problem is to identify the local portion of the diagram that relates to the particular point being made.

6 Acknowledgments

We thank the instructors who made their lecture materials available to us for study. We also thank Microsoft Research for its 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.

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

Anderson, R. J., Hoyer, C., Prince, C., Su, J., Videon, F., Wolfman, S. A., Oct. 2004b. Speech, ink and slides: The interaction of content channels. In: Proceedings of twelfth ACM International Conference on Multimedia. pp. 796–803.

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

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

Eisenstein, J., Adler, A., Guttentag, L., 2004. A trimodal dialogue corpus: Speech, gesture, and sketching. In: Student Oxygen Workshop.

Oltmans, M., Alvarado, C., Davis, R., 2004. Etcha sketches: Lessons learned from collecting sketch data. In: Making Pen-Based Interaction Intelligent and Natural. AAAI Fall Symposium, pp. 134–140.

Peiper, C., Chan, E., Campbell, R., Bresler, J., Al-Muhtad, J., 2004. Expanding education through active space collaboration. In: Second IEEE Annual

Converence on Pervaise Computing and Communications Workshops. pp. 236–240.

Qian, D., Gross, M., 1999. Collaborative design with netdraw. In: CAAD Futures 1999 Conference. pp. 213–226.

Shilman, M., Wei, Z., Raghupathy, S., Simard, P., Jones, D., 2003. Discerning structure from freeform handwritten notes. In: Seventh International Conference on Document Recognition and Analysis.

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