Introduction

Instant messaging has become a nearly ubiquitous mode of communication in many individuals' daily life. It provides a convenient and natural form of communication as well as an efficient medium for the transfer of digital information, such as links to online content. The reliance of text only is a form of communication that can be utilized in a variety of circumstances prohibitive to verbal communication. Because of its convenience, immediacy and accommodation of digital content text based messaging is likely only to become more prevalent.

Despite their virtues, solely text-based interfaces do possess major deficiencies as well. One primary limitation of text-based messaging is its inability to capture the rich detail of the non-verbal components of communication. In face-to-face communication multiple modalities of communication are simultaneously and effortlessly utilized to create meaning. Through the dynamic juxtaposition of disparate but complimentary modes of communication, such as gesture, inflection, and gaze, the interlocutors are able to contextually configure their understanding (Goodwin 2000). In a computer mediated text based communication these subtleties of conversation are completely lost.

Embodied Reference

In face-to-face communication participants can momentarily construct meaning in the milieu by referencing features of their local environment (Hindmarsh & Heath, 2000). The environment takes on significance relative to the discourse as it is annexed by the conversation, and in doing so the items referenced also serve to facilitate the discourse. In this sense, the environment can be seen as a device through which meaning is reflexively constituted and elaborated upon. This reflexive relationship between action and object is critical to the production and understanding of the participants' conduct in discourse (Heath & Hindmarsh,

2000). Due to the physical separation intrinsic to text-based communication, the interlocutors are incapable of using spatial cues to orient their interaction.

Asynchronous Communication

Another problem raised by text-based interfaces is the discrepancies that can arise between users. In face-to-face communication the rate of the conversation is interactionally constrained. The speaker is constantly receiving cues from the listener that provides them with immediate feedback that enables both parties to adjust the tempo of their speech for the other. However, the spatial remoteness of text-based communication excludes the use of feedback and can result in asynchronous rates of communication between speaker and listener. This inhibits the interlocutors' ability to engage in effective turn taking in their discourse (Sacks et al., 1974). As a result of this breakdown in the structure of communication, the participants experience difficulties making sense of the dialogue as it dynamically unfolds.

Importance of Topic

In face-to-face conversation, dialogue is aided by complex gestures, stress on certain words, or even external aids such as highlighters and pens which create a rich, multi-layered representation of the topic being discussed. The power of coding, highlighting, and graphical representations as discussed by Goodwin is very useful here (Goodwin, 1994), since the coding of certain voice inflections or physical usage of colors can be recognized; highlighting or indexical gestures can place emphasis on specific topics of interest; and physical papers or images become immutable inscriptions which interweave with the discourse. Currently, text-based communication does not support this sort of professional discourse.

Clusters of Relevant Information

Face-to-face communication through context in real-time creates a very dynamic environment; information is elicited and organized through semantic speech, intonation, and gestures interacting in a synchronized fashion so that information becomes naturally organized. Text chatting replaces this rich communication setting with unidirectional time-based organization, where any information posited automatically follows whatever came right before it in a serial list-like fashion ordered by timestamps, whether or not the two are relevant to each other. Semantic gaps are filled in by the users where they must mentally consider what the other user is saying and organizing information internally.

Design Suggestions

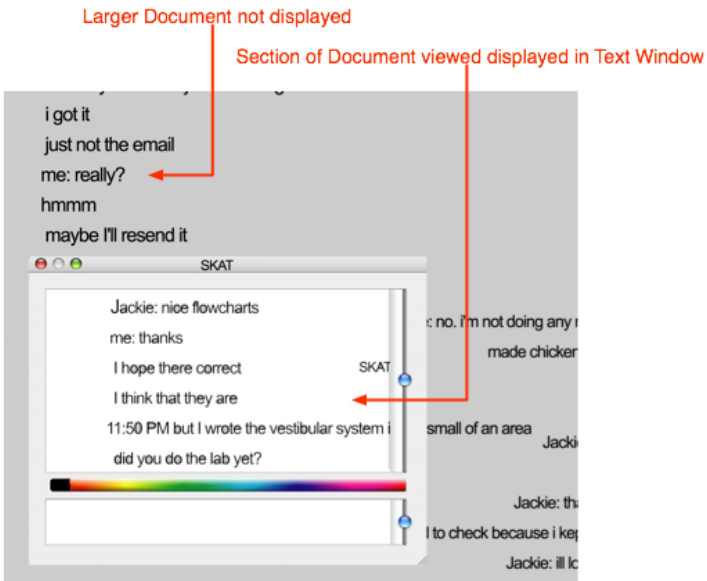
In our project we explore the possibilities of incorporating the extra linguistic components used in face-to-face speech into a text-based interface. The intention of this interface is to allow users to more readily convey the subtleties of verbal communication in a nonverbal device while retaining the advantages of text based communication. We also intend to improve upon current interfaces' ability to configure mutual orientation between those interacting with the interface. We believe that our design suggestions will support a more expressive platform for communication than previous instant messaging applications.

To do so, we introduce a new device designed after a metaphor of communication that is well suited for conditions where verbal communication is not available but still rich due to its ability to leverage nonverbal components of speech a metaphor of writing notes to another individual in class. In a classroom environment individuals can use relatively low bandwidth communication to communicate complex abstract ideas. In this scenario because their actions take on additional communicative value by referencing information with written text, gesture, and quick sketches. We believe this form of communication is highly adaptable to instant

messaging and allows for creativity and flexibility in communication while retaining the current desirable aspects of instant messaging. We named our new design SKAT to illustrate the fact that it aims to incorporate features of sketching and chatting into one interface.

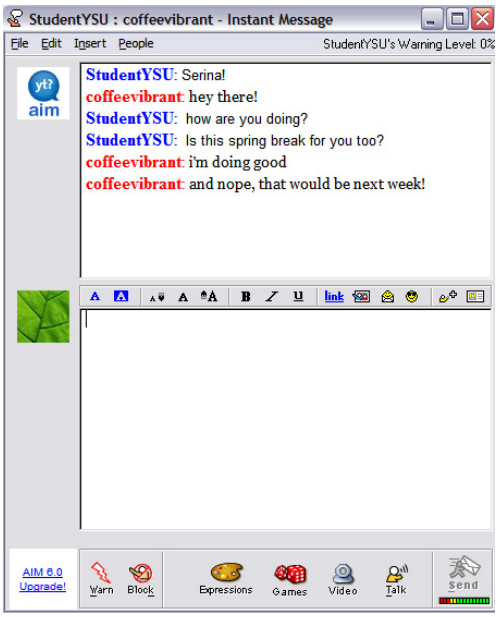
Interface

The interface for SKAT is a window that appears like many chat windows. However SKAT is best thought of as document that is filled out by the interaction of the users. The SKAT window displays the text that and images as they are being communicated but can be repositioned and resized to view the document sitting underneath. The document that is being filled out is fixed to an 8.5X11 size and users navigate the document by scrolling up and down, or using the Grab Tool or Zoom Tools. In this way navigating throught the conversation is similar to navigating a PDF. By allowing the users to resize and reposition themselves on the larger document through a small window the interface affords a variety of interaction prohibited by other chat interfaces while maintaining a small footprint on the screen.



Type and Drag Feature

Current standardized chat applications afford text to be placed into text space by 'sending' it into the display window, where the text is automatically temporally organized , placing input underneath the one immediately before it. Consider the conversation below:



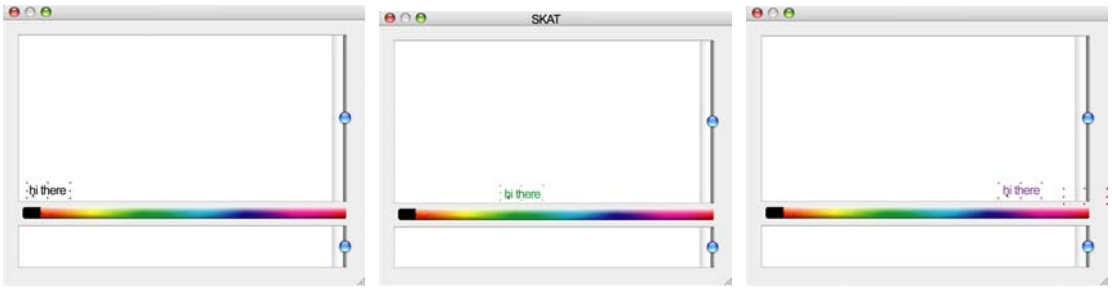
In this example where StudentYYY and coffeevibrant are chatting with each other, the text is ordered by whoever enters text last, and therefore one user can input multiple lines of text as long as they output it before the other user does, highlighting the temporal relationship of dialogue between two users.

In SKAT, the input area and display window will remain structurally similar to current chat applications, but integrates a 'Type and Drag' Feature which allows users to type text in a text space and physically drag the text into the text display window where they are free to place it in any area of the space they wish. This new feature removes the emphasis of temporal relationships between a user's comments and the next and replaces it with emphasis on semantic consistency across users by affording spatial organization as desired by the user. Articulatory and

semantic gaps are therefore effectively closed by allowing directness via spatial organization (Hutchins, Hollan, Norman, 1986).

Color Dipper

The color dipper is a gradient of interactive colors located directly above the standard text space of the chat window. In the process of dragging the text from the text space to the text display window, users can pass through this color gradient and 'dip' their text through a desired color, in which the text is updated to the color of the area the mouse is passing through. The images below illustrate the same text with different colors resulting from dipping through various areas of the color gradient:



Traditional current chat interface constrains users to inputting text from the top-left corner of the text space. We considered this when implementing the color gradient into our design by making the left of the gradient begins black before it fades to other colors. A user who desires to use plain black text can type in their text, simply drag it directly above their box, and place it onto the text display space. They can also press enter on their keyboard, which would appear by default black in the text display window.

In this design implementation, we also attempt to address three major areas of discursive practices (Goodwin, *ibid)* lost in current designs of text-chatting: first, personality can be graphically represented through colors, enabling a user to quickly personalize their text by

dipping in a color. Secondly, we attempt to partially rebuild the meanings created through intonation and inflection by color coding words. Words that are emphasized in speech, for example, can be colored in text to emphasize a word. An example below illustrates this, where the word ‘cute’ is encoded by color to represent different meanings:

Thirdly, through techniques similar to our previous point, text can be colored differently from others to highlight its importance. The location of the color dipper (between the text space and the text window display) conveniently allows the user to dip their text through the colors as they pass over from the text space to the text window display, thereby also sparing the user the inconvenience of opening a menu and having to choose from a grid of colors, as is often the approach used by current text chat applications.

TOOL KIT

One of the main reasons that communication can still be effective even in the absence of verbal cues is from peoples ability to make use of cues in their environment. At remote distances this may seem an like impossibility, but if the participants are oriented to the same document, then making use of the environment is a viable way to aid in communication. In face-to-face communication to orient a person to a feature of document the participants can merely gesture to

document and then the document can serve as a resource for sense making (Heath & Hindmarsh 2000). To allow for a similar type of communication to occur in an instant messaging environment we introduce a set of tools to aid in more effective remote asynchronous communication.

The Tool Kit can be accessed by right-clicking the mouse; while the mouse is depressed the five tools appear. Starting from the upper right side going clockwise the tools are:

- Shape Tool-Adds shapes to the Text Display Area
- Pain Brush- Adds permanent markings
- Grab Tool- Repositions where the text is displayed
- Zoom Tool- Resizes the view in the text display window
- Trail Tool- Adds nonpermanent markings

Problem Solutions

Asynchronous Relationship

One of the main advantages of SKAT over other instant messaging programs is its ability to sketch inside of the text display window.

The Asynchronous relationship between the participants can create complications in communication as the conversation unfolds when the relationship between statements of the

participants becomes unclear. In the following section of dialogue between Jodorowsky and Coixet, we can see their discourse breakdown as Coixet misinterprets which comment Jodorowsky is responding to.

Transcript 1

1. Jodorowsky: I hate when I'm trying to say something and I just can't spell it so I have to go back and rewrite the sentence to accomodate my bad spelling
2. accommodate for instance
3. Coixet: but as long as the message gets across, thats ok
4. i just have issues
5. Jodorowsky: that's not the OCD Coixet that I know
6. right
7. hmmmm
8. Coixet: what?
9. !?!?!
10. Jodorowsky: never mind

In this example, we can see that the communication breakdown occurred because Jodorowsky responded to a question outside the window of time where Coixet can easily interpret his response. In this instance, the breakdown resulted from their asynchronous communication. In the following sequence we can see an example of how the SKAT tools can be used to rectify the problem, while allowing the user to respond in a time frame of their choosing:

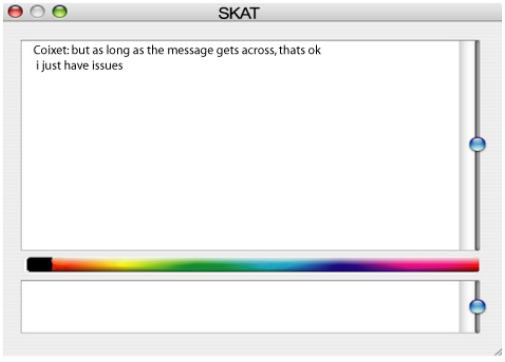


Image 1

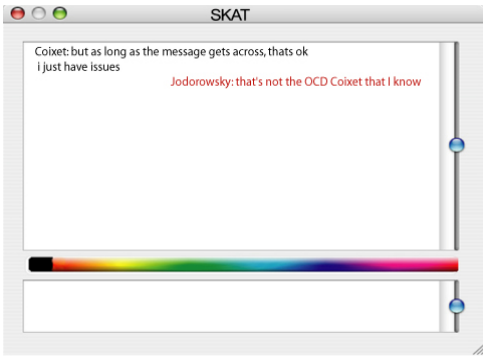


Image 2

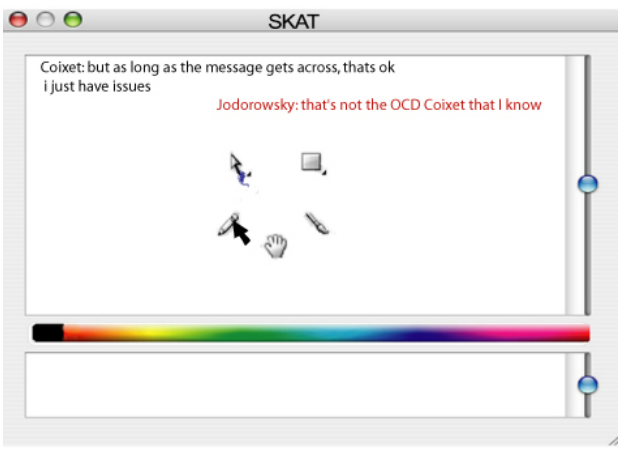


Image 3

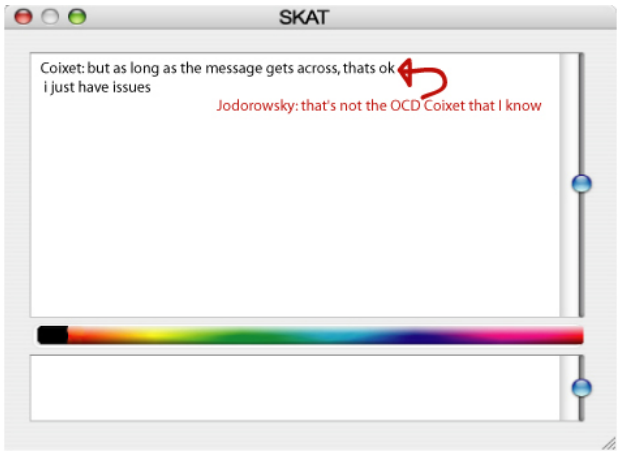


Image 4

This example illustrates two methods by which SKAT can be used to disambiguate references in conversation. First, by allowing the interlocutors to directly reference the line to which they are responding with drawings, confusion can be prevented. Secondly in SKAT, the participants are able to spatially arrange their text to their choosing, thereby allowing participants to group responses or ideas by location to prevent the asynchronicity from being imposed onto the conversation structure.

The other tools in the SKAT tool kit are equally useful for such purposes. In the following example we can see how the Trails Tool in conjunction with pasting images can be used to refer to features of the conversation without adding permanent markings. Consider the transcript below:

1. Jodorowsky: oh I found it
2. Coixet: huh
3. Jodorowsky: one moment
4. ok well the in the lab it says:
5. sorry i'm still looking on the website
6. Coixet: At test, there are three cue conditions: no cue, same cue, or different cue.

What determines your ability to recall a particular target is the interaction between how you processed it at encoding and how you process it at test. If the cue changes, then it is likely that you will process the item in a slightly different way than when the same cue is presented again.

7. i guess thats it then
8. Jodorowsky: I found something different
9. that is more specific
10. Annoying
11. it won't let me copy the text
12. do you see where it says
13. Wait
14. Coixet: where????
15. Jodorowsky: what it the url?
16. nevermind
17. F*%k it

In this excerpt we can clearly see that the participants are having difficulty orienting themselves to the same domain and therefore are not able to use the environment as a resource as a communicative device. Below, we can see how the Tool Kit could help them orient their understanding:

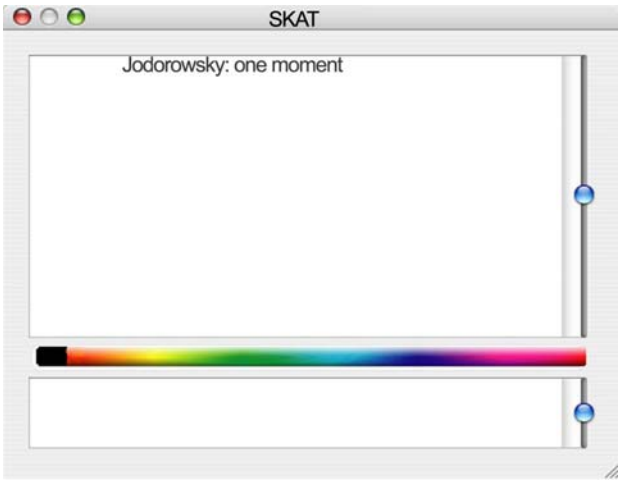


Image 1

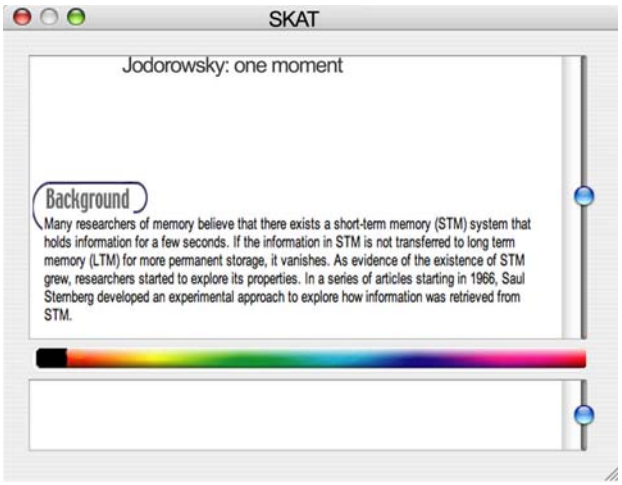


Image 2

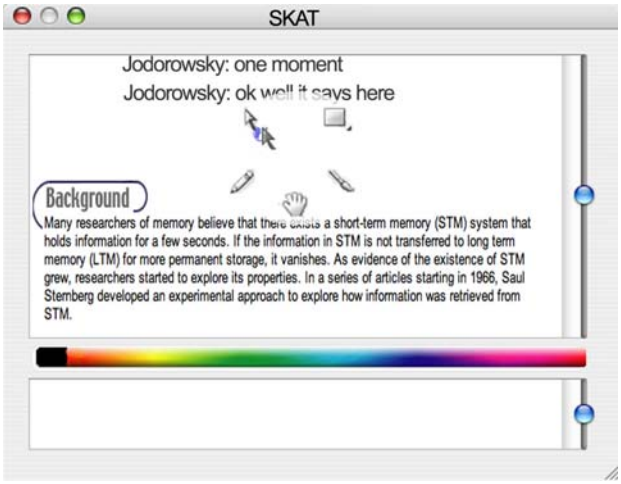


Image 3

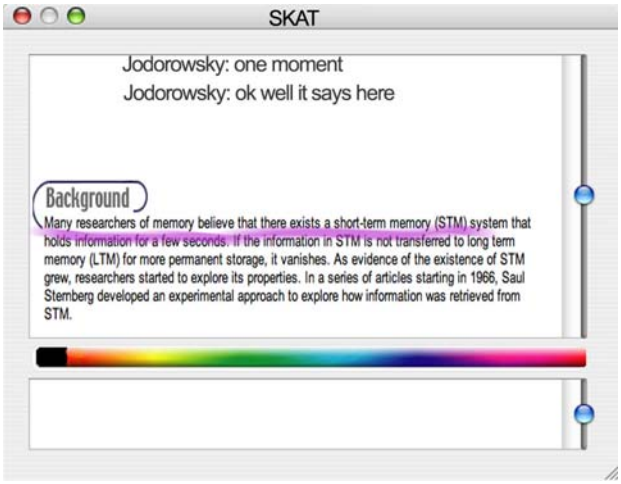


image 4

In the sequence we can see Jodorowsky pasting a screenshot from the webpage and then selecting the Trail Tool to gesture to the relevant information. In this example, the Trails Tool was chosen because it produces an impermanent marking fading in time, so that items can be referenced without providing visual clutter.

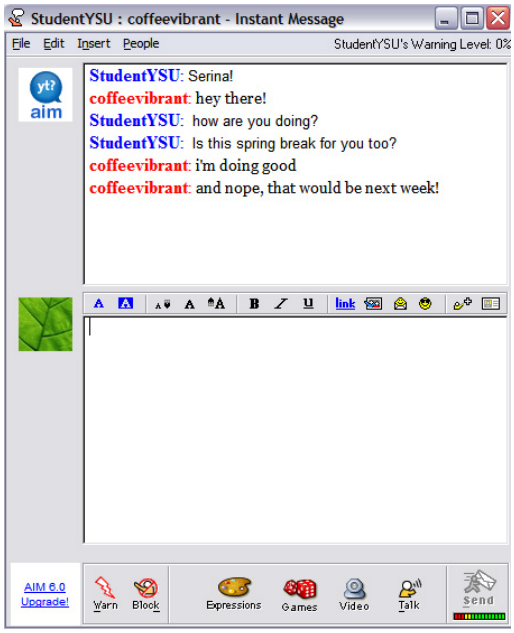
Topic Importance

The inability of current text-chatting interfaces in supporting text organization by topic importance in the course of dialogue is addressed in SKAT: through the combination of the spatial arrangement, use of the Color Dipper and tools available on the Toolkit, a variety of different options are presented for the user to mark their text. A user may choose to convey important information to another user by taking advantage of the spatial organization of our interface, such as placing it in the center of the text display window. They may also use the Color Dipper to separate text by their color scheme, drawing out desired segments of their text by marking it with a color so as to highlight it, or take advantage of the Paint tool on the Toolkit and highlight information that way. These options are illustrated in the image below:

These techniques are also combinable with one another to create a rich overlapping ability to highlight information and effectively draw the attention of another user to certain parts of their text.

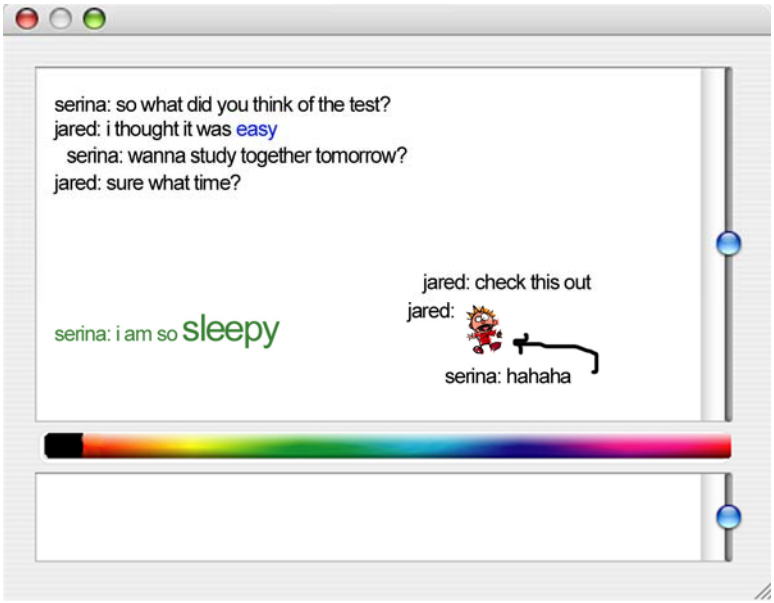
Clusters of Relevant Information

Due to the temporal nature of current chat applications, information displayed in the text display window is unable to be organized by its relevance, unless they just so happen to be ordered together by the time of output. In the example earlier of the chat dialogue between StudentYYY and coffeevibrant (shown again below), StudentYYY asks two different questions successively without allowing enough time for coffeevibrant to finish replying to the first question – subsequently, coffeevibrant works around this by answering the questions successively in a similar fashion so that these answers will disambiguate each other by the way they are returned to StudentYYY.



This particular example may be a simple workaround; however, as more information is passed between two chatters, relevant information (and relevant replies to these) become more ambiguous.

SKAT addresses this issue by affording spatial organization in the text display window. Multiple topics which arise through the course of natural chat conversation can be separated by the interlocutors in their placement of text in the area of space available to them in the text display window: two users can be engaged in a topic of discussion in one corner of the text display window while one user may ask a question irrelevant to this topic in another corner.



Similar to the way notes are passed from one individual to another in class, clusters of topics can be separated by the way they are mapped out in space, creating an organizational structure with text that is arguably intuitive.

Conclusion

Text-based chatting programs have become a very ubiquitous form of communication. Although providing advantages such as convenience, immediacy, and accommodation of digital content, disadvantages such as the removal of vocal cues and gestures arise. Due to these deficiencies, users are required to do a larger amount of cognitive work in order bridge the semantic gaps between what they wish to express and what the computer will accommodate; they also must bridge the gap between the text they receive and the true meaning behind the text. We have geared our design with the intention of closing these semantic gaps: through spatial organization and inscription tools, we hope to incorporate the shared, external modes of representation outside of text communication.