Not Interfering: Simultaneous Typed Chat in COMPS Computer-Mediated Dialogues

Michael Glass
Valparaiso U.
michael.glass@valpo.edu

Anthony Nelson
U. of Southern Cal.
anthoncn@usc.edu

Chinedu Emeka
Valparaiso U.
chinedu.emeka1@valpo.edu

Jung Hee Kim
North Carolina A&T State U.
jungkim@ncat.edu

Abstract

COMPS computer-mediated typed-chat collaborative learning exercises permit the students to type at the same time. People can see and respond to each other's text in real time. Although everybody talking at the same time does not work in spoken conversation, students quickly discover they can type at the same time without interfering with each other. About 40% of typing occurs while other students have not yet formally ended a dialogue turn by pressing "enter." In COMPS dialogues normal conversational turn-taking often occurs when students pause to wait for each other, without pressing "enter." In this paper we estimate the polite delay that people use for deciding when the other person has relinquished a turn. Studies of educational dialogue will have to take into account the interactions that the new computer-mediated communication regime affords. This paper also characterizes the varieties of interaction that are observed during non-turn-taking simultaneous typing in COMPS dialogues. While students are typing together, they do not engage in tightly-interleaved two-way exchanges. Instead, each student individually responds to something that another student said earlier.

1 Introduction[1]

A feature of the COMPS (Computer-Mediated Problem Solving) online chat environment is that students can all type at once. They see each other's words in real time and can respond at the same time without interrupting. This adds an interactive dimension that spoken language does not support.

Most forms of human dialogue require that people take turns as they talk. Typed-chat lacks many of the signals such as prosodic effects that people use to regulate turn-taking in verbal conversation. COMPS follows the common convention that pressing <enter> signals the end of a dialogue turn. However the simultaneous typing feature means that pressing <enter> is not necessary for relinquishing the turn, other people can simply start typing. Hence people often do not press enter to end a turn. They simply pause, and other participants infer that a turn transition is possible. In this paper we estimate how long is the pause that the other participants wait before deciding the turn has ended.

Since overlapping dialogue for more than a short time is a phenomenon new to computer-mediated communication, this study examines overlapping dialogue manually to see how students are using it. One possibility is that the students are violating the turn-taking structure of regular dialogue, engaging in tightly intertwined dialogue where a person responds immediately to the words being typed by the other person. Manual characterization of 30 such interactions reveals that this doesn't happen. We have identified three common patterns of interaction, in all of them the overlapping students are effectively responding to earlier dialogue. Thus they follow conversation protocols akin to normal turn-taking even though several people are talking at once.

The two main results in this paper are then: a) an estimate of the pause time that signals a possible turn-switching point in typed chat, and b) showing, qualitatively, examples of the three main patterns of dialogue interaction that students employ when using the simultaneous chat, c) show that with these patterns of dialogue interaction students are still using conversational turn-taking.

These results can be used to inform the COMPS project's efforts at text analytics, which will identify characteristics of the dialogue in real time that will be indicative of students engaging in interactive conversation.

2 Background

2.1 COMPS dialogues

The COMPS project deploys and studies small group collaborative problem-solving exercises in college computer science and mathematics classes [Kim et al., 2016]. The exercises are designed to address student conceptual knowledge through group cognition [Stahl, 2004; Stahl, 2009]. Typically students work in groups of three, with a TA or instructor additionally participating intermittently. The exercise protocol for the exercises requires the students to solve a problem in steps, coming to agreement on each step. The students show the agreed-upon answer to the instructor, receive feedback or hints, then further discuss the step or proceed to the next item.

A	Labels 1, 2, 3, 4, 5, and 14	can be instantiated	anonymously. Because these do not have to be changed ◄
B		What about 6 and 7? ◄	

Figure 1. Example of simultaneous dialogue. End of turn is marked by "◄".

This is consistent with accepted practices for good collaborative exercise design, requiring creative interdependence [Eberly, 2016]. This protocol also discourages social loafing, all students must participate at multiple instances during the exercise.

Students discuss the problem through the COMPS software, a web-delivered chat interface that permits everybody to type and see each other's dialogue all at the same time. The software logs the chat for later analysis.

Figure 1 illustrates how simultaneous chat can differ from normal conversational turn-taking [Glass et al., 2015]. In this and subsequent figures the <enter> ending a turn is marked with "◄". Student A was listing widgets on a screenshot of a Java GUI, it was an answer to one of the exercise questions. Part way through A's dialogue turn student B inquired why certain widgets had been left out. Student A was still typing and had not typed <enter> yet. A, continuing to type in the same chat dialogue turn, then answered B's question. As the Figure 1 example illustrates, students indeed adapt to this simultaneous chat regime and engage in productive interactions while doing so [Glass et al., 2015].

2.2 Dialogue Turns

Potentially the instructor could be aided in knowing which discussion groups could benefit from instructor intervention and participation. The COMPS project is developing text analytics for this purpose.

Having the computer detect and measure the prevalence of transactive turns in the conversation might provide an estimate of conversation quality for an instructor's dashboard. Transactive dialogue is a key element of group cognition, where a dialogue turn a) contributes to the knowledge construction and b) responds to a previous dialogue move, usually by another person [Weinberger & Fischer, 2006]. There are a variety of different categories of transactive contribution, ranging from "eliciting" to "conflict-oriented consensus building."

We have simplified the computer-recognition task to two smaller steps: recognizing whether student dialogue turns are discussing the topic [Willis et. al, 2017] and whether dialogue turns are responding to other turns [Glass et al., 2014]. For the latter task, we adopted the conventions of Conversation Analysis, a framework from discourse linguistics [Sacks et al., 1974; Stubbs, 1983]. In Conversation Analysis dialogue is segmented, each segment starting with an "Initiate" turn and containing the other person's "Respond" turn plus possible "followup" turns. However COMPS project attempts to train classifiers to recognize Initiate and Respond turns had difficulty partly because of many turns where people type simultaneously [Glass et al., 2014a, 2014b]. As Figure 1 illustrates, where each dialogue turn responds to the other, there sometimes isn't even a clear separation of turns that can be tagged as which is responding to which.

One possible step toward disentangling the Initiate/Respond problem between overlapping dialogue turns would be to identify the Transition Relevance Places (TRPs) and Turn Allocational Components [Schegloff, 1990; Sacks et al., 1974]. In Conversation Analysis, a TRP is a place where turn reallocation is possible. TRPs can occur in the dialogue when a person stops speaking, but they can also occur, e.g., when a complete thought has been finished and another person can jump in and respond.

According to this analysis, the Figure 1 dialogue could be thought of as follows:

A: "labels 1, 2, 3, 4, 5, and 14" <TRP>"can be …"
B: "What about 6 and 7?"<TRP>
A: "Because they do not have to be changed"

In this analysis, B starts typing in response to A's first TRP. In speech this could have been an interruption. In typed-chat, B typing does not interrupt A's ability to type, so A and B can type simultaneously. Similarly, A's final sentence is a response to B's recently-ended turn.

Many TRPs will be determined semantically, as the first TRP in Figure 1 after "14." We observe that a pause in the typing represents a TRP, a possible place for the of a new turn for dialogue purposes. If student A paused to read B's question in Figure 1, for example, detecting that pause could be helpful both in manual and automated dialogue analysis.

Table 1. Statistics of Group Discussions.

Group discussions	28
Dialogue turns	2723
Keystrokes (not <enter>)	113,808
Participant/group	4 (3 students, 1 TA)
Duration: Swing exercise	35 min
Duration: Inheritance exercise	70 min
Pct of keystroke overlapping	17%

2.3 Data for this Study

Table 1 shows the descriptive statistics for the chat data used in this study. COMPS exercises have been deployed in two different quantitative literacy exercises for pre-service teachers and in four different object-oriented Java concept exercises in CS2 classes [Kim et al., 2016]. In this study we used two exercises from a CS2 class in Fall 2015. One exercise was about inheritance concepts in Java, the other about the concepts in a Java SWING graphical user interface. There were 14 discussions from each exercise

2.4 Method of Analysis

This study examined two phenomena: how long is a pause that represents a turn relevance place, and how do people use simultaneous typing to address each other.

For the timing study we used data on how long did person B wait to start typing after person A paused. This was applied to all the typing in Table 1.

For the varieties of interaction, we manually analyzed examples of simultaneous typing observed in the log files. The question was the relationship of each overlapping dialogue turn to the rest of the conversation: was it initiating or responding, and if responding where was the antecedent language. During this process we maintained a coding manual listing the varieties of interaction the annotators had found.

Only a small sample of about 30 of the instances have been manually annotated at this time according to the current manual. The three coders periodically compared annotations and arrived at a consensus. Interrater reliability thus has not been tested. The result reported here is qualitative, a description of the categories that were found and coded.

3 Measuring the Pause That Signals Turn-Taking

3.1 Turn Allocation Using <enter>

The explicit component for turn allocation in typed-chat is the <enter> which ends a turn. COMPS also has a button for ending a turn, the effect is the same. We start by defining a delimited turn as the time from the first keystroke until the <enter> delimiter that ends the turn. A simultaneous keystroke occurs if person B types while any other person A is still within a delimited turn. This is similar to the usual definition of interruption in spoken dialogue, except that B typing does not interrupt A's ability to type. Using this definition of a turn, between 25% and 50% of keystrokes in COMPS dialogues are simultaneous typing.

3.2 Ending a Turn by Pausing

Just as in spoken conversation a turn can end when a person stops speaking, a chat turn can also end without explicit marking if the student stops typing. In spoken conversation, 3 seconds is an awkwardly long pause [McLaughlin, 1984]. In COMPS dialogues there are many pauses. In one extreme, participant A may pause typing without pressing <enter>, waiting for the other participants to respond. Everyone else can see what A has written so far, and can observe that A has stopped. Another participant B can then type without interference. After B pauses, A can resume, inserting a second logical turn in the single <enter>-delimited turn. In this typing regime, there may not be much incentive to type <enter> to formally end a dialogue turn.

What causes people to formally end their turns is the COMPS scrolling text box behavior. Turns properly work within the interleaved dialogue scroll only if they have been <enter> delimited. Furthermore, a single turn containing several logical dialogue turns concatenated together in a single stretch of text becomes hard to read. The result is that turn-taking by pausing rarely continues past two logical turns. It is, however, common for a turn to end with a long pause followed by the <enter> before the new turn.

3.3 Measuring a Release Time

We hypothesize that there is some time that will be recognized as meaning that the typist has given up the turn, allowing somebody else to speak. We extended the definition of simultaneous typing to include a release time. A participant is *not inside* a dialogue turn if a) the last thing typed was <enter> (explicitly ending the turn) or b) a release time has elapsed (implicitly releasing the turn).

How long after people pause typing before the other people recognize it as a typed-chat release of turn? We can look at typing behavior. If there is a well-recognized time that releases the turn, starting typing without waiting for the release time to elapse would be a form of interrupting. Under this hypothesis, we would expect to see many fewer keystrokes from other persons during the interval of the release time than after it has passed. After the release time the rate of keystrokes from other typists will reflect the average rate for conversation in general.

Figure 4 (at end) shows the percent of keystrokes classed as overlapping according to different trial values of release time. As expected, for larger times the curve is approximately linear, representing the rate of new dialogue added into the conversation. However for the first 2 seconds after a person pauses typing the other participants are much more reluctant to type.

Accordingly, for computer analysis we use 2 seconds to declare the end of a turn and subsequent typing is non-simultaneous.

3.4 Other Means of Turn Allocation

We have not considered other mechanisms for turn allocation. In conversation a turn can be ended, for example, by asking a question. The same holds true for typed chat. Spoken dialogue also employs prosodic features and other mechanisms for turn allocation that are not available in typed chat. The paucity of conventions for turn allocation in computer-mediated typed chat produces quite different behavior than spoken conversation. One experiment using a chat system similar to ours observed about 30% of turns were overlapping with other turns, an amount of overlap not possible in speech. It also observed long pauses where nobody was typing, another feature rarely observed in speech [Anderson, et al., 2010].

An issue occurs if person A pauses shortly after person B begins simultaneous typing. A's ability to type unimpeded is not compromised, however A sometimes pauses, perhaps to read what B is saying. Clearly B's first keystrokes constitute simultaneous typing. After A pauses for a while, it becomes

A	It's only one answer for this one. It's not A. Sorry About that◄	
B		It's just E.◄

Figure 2. Example of overlapping response

A	i think its B D and E◄	
B		I agree with B,D, and E◄
C		Actually no im changing to just E◄

Figure 3. Example of simultaneous response.

clear that they are no longer typing simultaneously. If A were to pause for only the time for few keystrokes it should probably still be call simultaneous. In our work we use the same release time for this determination. Up until the 2 second release time has elapsed, B's keystrokes are still counted as simultaneous typing.

4 Varieties of Interaction

The main observation is that overlapping dialogue usually addresses earlier dialogue. Overlapped dialogue never in our sample of 30 simultaneous typing incidents directly addressed or responded to the text that was being typed at approximately the same time. People were not simultaneously reading, thinking, and writing.

There are three patterns of overlapping dialogue, characterized by what parts of earlier dialogue are being addressed by the simultaneous participants:

- Overlapping response, where participant B responds to something that A has recently uttered, while A continues the same turn.
- Simultaneous response, where two participants respond to an earlier turn by a different participant.
- Simultaneous initiation, where two participants simply have different ideas to insert into the conversation and happen to type them in overlapped fashion. Both utterances fit into the conversation as of the point when they started.

Examples of the first two varieties follow.

4.1 Overlapping Response

Figure 2 shows an overlapping response. The students were asked to pick one or several from the lettered answers to a multiple choice question. The first student, who has suggested multiple letters, corrects that by saying "It's only one answer for this one. It's not A. Sorry about that." Another student starts typing in the middle of this turn, responding to the assertion that there is only one answer letter. That student says "Its just E." According to our analysis, the end of A's first sentence represents a Transition Relevance Place.

The Figure 1 example shows overlapped responses. Student B responds after the first part of A's answer, noticing that some of the numbered SWING components were not mentioned. Later, Student A responds to B after A has paused a short moment to read B's question.

4.2 Simultaneous Response

Figure 3 shows an example of a simultaneous response, student B agreeing with student A's suggestion and student C disagreeing. Simultaneous responses occur at TRPs.

One interesting aspect of this example is there is a three second gap between the end of A's turn and the beginning of both B and C. This gap is visible in the keystroke timing diagram of the same dialogue turns shown in Figure 5 (at end). Student A is the bottom line, B and C above it in that order. We think this gap represents cognitive processing, rather than waiting for the release time, since A's turn had been terminated with <newline>.

5 Discussion and Future Work

There are considerable differences between computer-mediated typed chat and spoken communication. The ability for multiple participants to talk simultaneously for extended periods is new to computer communication. It has not been well-studied how people use this ability as they engage in problem-solving dialogues.

Regarding turn allocation, it has been observed by other researchers turn allocation is not always controlled by simple <enter> and pause times. People have been observed to type <enter> in the middle of a logical thought, for example, and then quickly continue typing [Markman, 2013]. This circumvents the turn allocation mechanism, permitting a single person to hold the floor for longer stretches of dialogue.

Regarding simultaneous interaction, cognitive science suggests it is unlikely that when several people are typing simultaneously they are multitasking, viz: reading other students' dialogue, thinking about it, and writing responses. Generally people cannot multitask between two tasks that require attention and cognition without switching back and forth between them and degrading performance [Bermúdez, 2014]. Our sampling of simultaneous dialogue events is consistent with this. When people chat simultaneously, they are interacting with events in the dialogue that occur before they initiated their chat turn.

The manual analysis and annotation of simultaneous events has proven to be difficult. This will have to be addressed in order to get meaningful statistics on the prevalence of categories. The main issue is that Conversation Analysis was developed for two-person dialogues. Categorizing utterances as initiate or respond (some versions include a third category "followup") seems to break down when there are more than two people in the group. In a group dialogue it becomes necessary to determine whose utterance is being responded to. Once it becomes possible to have multiple responses to one utterance, the same analysis produces chains of responses. Did C respond to B's response to A? Or did C respond to A directly? The same utterance can often be read both ways. The methods of linguistic analysis of multi-party conversations hinge on this distinction. But when the dialogue turns overlap, it becomes harder to resolve the ambiguities. Although there are many cases where three independently-working annotators readily agree, there are a similar number where agreement comes only after consensus discussion or not at all. In the future we will try categories that do not depend on identifying the antecedent of a response.

The determination of the release time will be refined by statistical analyses of the gaps between the end of one turn and the start of the next. There are other phenomena that potentially affect pauses in the conversations. We could try to disambiguate between a cognitive reason and a social reason. Conversational gaps might be expected to be more prevalent in problem-solving dialogues due to cognitive processing. The Figures 3 and 5 example of simultaneous response exhibits just such a gap. However there is possibly a social explanation: with three or more participants, no one person is responsible for filling the gap in the conversation after one speaker pauses. We plan to study this by two methods. One is by separating the conversations into segments where students are attending to the problem vs. attending to other matters. We hypothesize longer pauses in the problem-solving segments. Another method will be to check for correlation between the length of a pause and the length of the dialogue turn which follows the pause. Longer thinking pauses might result in longer utterances subsequently.

Another aspect of simultaneous chat that could be studied is whether the degree of simultaneity correlates with student engagement or other positive measures of dialogue quality. Anecdotally, students report being engaged by this facility. If degree of simultaneity correlates with some positive or negative outcomes of group problem-solving chat, it would be useful to put this measure on the instructor dashboard.

Acknowledgments

Partial support for this work was provided by the National Science Foundation's Improving Undergraduate STEM Education (IUSE) program under Award No. 1504917. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

References

Anderson, J. F., F. K. Beard, and J. B. Walther. (2010). Turn-Taking and the Local Management of Conversation in a Highly Simultaneous Computer-Mediated Communication System. Language@Internet vol. 7 article 7.

Bermúdez, José Luis. (2014) Cognitive Science: An Introduction to the Science of the Mind. 2nd ed. Cambridge University Press.

Eberly Center (2016). Carnegie Mellon, Eberly Center for Teaching Excellence and Innovation. What are best practices for designing group projects? http://www.cmu.edu/teaching/designteach/design/instructionalstrategies/groupprojects/design.html. Retrieved March, 2016.

Glass, Michael, Jung Hee Kim, Kelvin Bryant, Melissa Desjarlais, Micayla Goodrum, and Thomas Martin. (2014a). Toward Measurement of Conversational Interactivity in COMPS Computer Mediated Problem Solving. Proceedings of the 25th Modern Artificial Intelligence and Cognitive Science Conference (MAICS-14), Spokane, WA.

Glass, Michael, Jung Hee Kim, Kelvin Bryant, and Melissa Desjarlais. (2014b). Indicators of Conversational Interactivity in COMPS Problem-Solving Dialogues, Third Workshop on Intelligent Support for Learning in Groups (ISLG) at Twelfth International Conference on Intelligent Tutoring Systems, Honolulu, Hawaii, June.

Glass, Michael, Jung Hee Kim, Kelvin Bryant, and Melissa Desjarlais (2015). Come Let Us Chat Together: Simultaneous Typed-Chat in Computer-Supported Collaborative Dialogue. Journal of Computing Sciences in Colleges, 31 no. 2, Dec.

Kim, Jung Hee, Michael Glass, Taehee Kim, Kelvin Bryant, Angelica Willis, Ebonie McNeil, Zachery Thomas. 2016. Student Understanding and Engagement in a Class Employing COMPS Computer Mediated Problem Solving. Modern AI and Cognitive Science Conference, Dayton, OH. pp. 69–74

Markman, K. M. (2013) Conversational Coherence in Small Group Chat. In Herring, S.,D. Stein.,and T. Virtanen, eds. Pragmatics of Computer-Mediated Communication., Berlin: De Gruyter Mouton, pp. 539-564C

McLaughlin, M. (1984). Conversation: How Talk is Organized. Newbury Park, CA: Sage Publication.

Sacks, H., E. A. Schegloff, and G. Jefferson. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. Language, vol. 50, pp. 696-735.

Schegloff, E. A. (1992). On the Organization of Sequences as a Source of "Coherence" in Talk-in-Interaction. In: B. Dorval, ed. Conversational Organization and Its Development, Norwood NJ: Ablex, pp. 51-77.

Stahl, Gerry (2004). Building Collaborative Knowing: Elements of a Social Theory of CSCL. In J. W. Strijbos, P. Kirschner and R. Martens, eds., What We Know About CSCL and Implementing It in Higher Education, Boston, MA: Kluwer Academic Publishers, pp. 53–86.

Stahl, Gerry, ed. (2009). Studying Virtual Math Teams, Springer, pp. 57–73.

Stubbs, Michael. (1983). Discourse Analysis, U. of Chicago Press.

Weinberger, A., and F. Fischer. (2006). A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. Computers & Education, vol. 46 no. 1, pp. 71–95.

Willis, Angelica, Ashana Edwards, Jung Hee Kim, Kelvin Bryant, and Michael Glass. (2017). Topic Modeling to Detect Student Expressions of Understanding in Collaborative Problem-Solving Dialogues. Poster abstracts, Thirtieth Florida Artificial Intelligence Research Symposium (FLAIRS-30). (to appear)

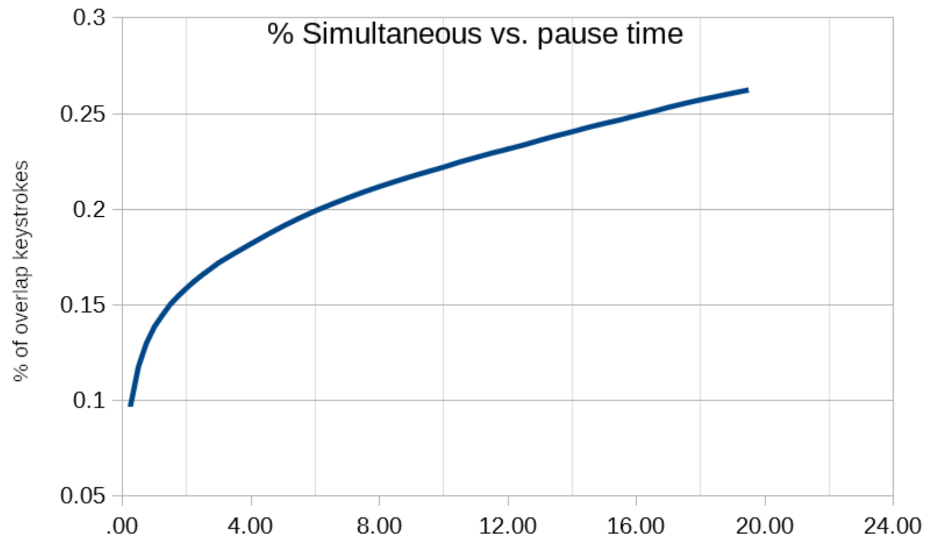

Figure 4. Percent of turns classified as simultaneous vs. pause time that would start a new turn

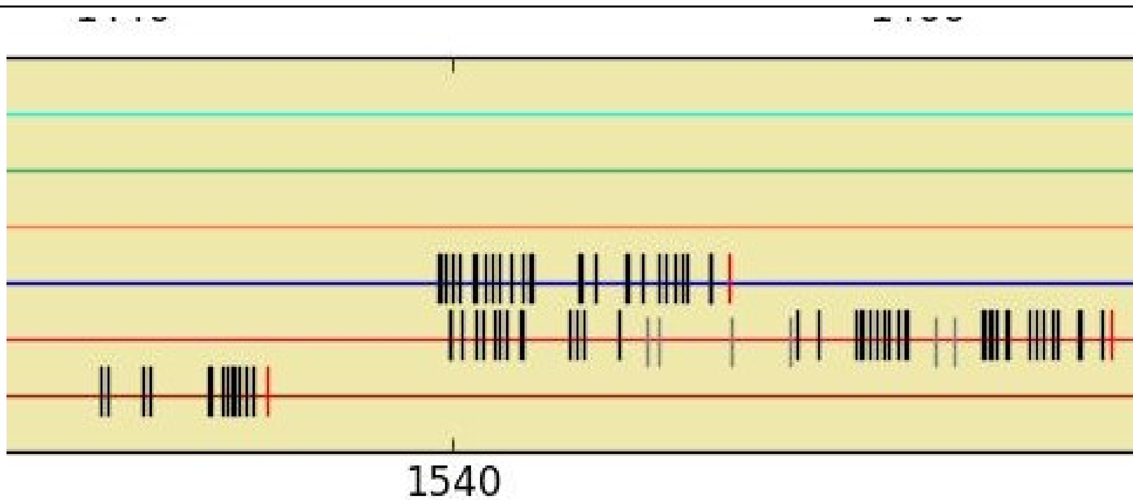

Figure 5. Keystroke timing diagram for dialogue in Figure 3. Student A on bottom line, Students B and C above.