# Interaction as a Diagnostic Resource in Tutoring

Barbara A. Fox

Technical Report #88-3

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

| | |
|---|---|
| 1a. REPORT SECURITY CLASSIFICATION<br>Unclassified | 1b. RESTRICTIVE MARKINGS |
| 2a. SECURITY CLASSIFICATION AUTHORITY | 3. DISTRIBUTION/AVAILABILITY OF REPORT<br>Approved for public release; distribution unlimited. |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S)<br>1530634 | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a. NAME OF PERFORMING ORGANIZATION<br>University of Colorado<br>6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION<br>Personnel and Training Research Programs<br>Office of Naval Research (Code 1142PT) |
| 6c. ADDRESS (City, State, and ZIP Code)<br>Campus Box 345<br>Boulder, CO 80309-0345 | 7b. ADDRESS (City, State, and ZIP Code)<br>800 North Quincy Street<br>Arlington, VA 22217-5000 |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION<br>8b. OFFICE SYMBOL (If applicable) | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER<br>DUNS-067-001412 |
| 8c. ADDRESS (City, State, and ZIP Code) | 10. SOURCE OF FUNDING NUMBERS<br>PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO. |

11. TITLE (Include Security Classification)
Interaction as a Diagnostic Resource in Tutoring

12. PERSONAL AUTHOR(S)

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Technical | FROM ______ TO ______ | 88, 01, 19 | 23 |

16. SUPPLEMENTARY NOTATION

| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | |
| | | | |
| | | | |

19. ABSTRACT (Continue on reverse if necessary and identify by block number)


According to much of the recent work in ICAI (see Sleeman and Brown, 1982; Anderson et al., 1987; Woolf, 1984; Oberem, 1987), the main job of a tutor is to discover and repair "buggy" procedures in a student's work. While I have argued elsewhere (Fox, 1987) that tutors appear to be oriented to a variety of tasks not centered on bugs, it is uncontroversial that tutors spend a great deal of time guiding students to reconstruct misconceptions and faulty procedures (see Arce, 1987, for a discussion of why the process should be viewed as construction and reconstruction rather than, say, simple correction).

How this diagnosis gets carried out is, however, quite controversial. In most research, tutors are said to derive their knowledge of student bugs by comparing the student's problem-solving procedures to their own; in other words, the student's steps in solving a particular problem is the sole source of diagnostic information. There thus (cont.)


| | |
|---|---|
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT<br>☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION<br>Unclassified |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL | 22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL |

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

19. Abstract (continued)

issues from this assumption a second assumption, namely that a successful ICAI system should be able to carry out debugging procedures just by examining the steps a student enters while solving a problem.

I would like to present here a somewhat different view of real-time diagnosis, and more generally, assessment. I do not intend to argue with the view that tutors base some of their "debugging" efforts on steps in solving a problem; what I would like to show here is that another source of diagnostic information is used at least as much, if not more than, the steps in problem-solving, namely, the tutoring interaction itself. This is to say that tutors use the timing of a student's response, and the way the response is delivered--in addition to what might be called the "literal content" of the response--as a source of diagnostic information.

The remainder of the paper is organized as follows. Section 1.1 describes the background of the study and outlines the transcription notations which the reader will find necessary for following the tutoring dialogues presented throughout. Section 2 presents evidence for the claim that interaction is a source of diagnosis; section 3 begins to establish a metric, based on the data examined in section 2, which in the future might allow an intelligent system to make use of the kinds of interactional cues utilized by human tutors. Section 4 presents some concluding remarks.

Interaction as a Diagnostic Resource in Tutoring

Barbara A. Fox
Dept. of Linguistics and
Institute of Cognitive Science
University of Colorado, Boulder

## 1 Introduction

According to much of the recent work in ICAI (see Sleeman and Brown, 1982; Anderson et al. 1987; Woolf, 1984; Oberem, 1987), the main job of a tutor is to discover and repair "buggy" procedures in a student's work. While I have argued elsewhere (Fox, 1987) that tutors appear to be oriented to a variety of tasks not centered on bugs, it is uncontroversial that tutors spend a great deal of time guiding students to reconstruct misconceptions and faulty procedures (see Arce, 1987, for a discussion of why the process should be viewed as construction and reconstruction rather than, say, simple correction).

How this diagnosis gets carried out is, however, quite controversial. In most research, tutors are said to derive their knowledge of student bugs by comparing the student's problem-solving procedures to their own; in other words, the student's steps in solving a particular problem is the sole source of diagnostic information. There thus issues from this assumption a second assumption, namely that a successful ICAI system should be able to carry out debugging procedures just by examining the steps a student enters while solving a problem.

I would like to present here a somewhat different view of real-time diagnosis, and more generaly, assessment. I do not intend to argue with the view that tutors base some of their "debugging" efforts on steps in solving a problem; what I would like to show here is that another source of diagnostic information is used at least as much, if not more than, the steps in problem-solving, namely, the tutoring interaction itself. This is to say that tutors use the timing of a student's response, and the way the response is delivered--in addition to what might be called the "literal content" of the response--as a source of diagnostic information.

The remainder of the paper is organized as follows. Section 1.1 describes the background of the study and outlines the transcription notations which the reader will find necessary for following the tutoring dialogues presented throughout. Section 2 presents evidence for the claim that interaction is a source of diagnosis; section 3 begins to establish a metric, based on the data examined in section 2, which in the future might allow an intelligent system to make use of the kinds of interactional cues utilized by human tutors. Section 4 presents some concluding remarks.

## 1.1 Background

The work presented here is part of a larger study of human-human tutorial dialogue. The study as a whole attempts to characterize, in extreme detail, the characteristics of tutorial interaction, how that interaction is changed when the participants are not face-to-face, and how student attitudes towards computers affect their interaction with them.

The data for the study were gathered by bringing together tutors and students from on campus. Graduate students from chemistry, physics, math and computer science were located and asked to participate as tutors in the study. All were recommended by their departments and all had worked as tutors before. They were paid for their work, in keeping with normal tutoring arrangements. Students were located by advertising in the student newspaper for people interested in being tutored as part of a research project. Each student-tutor pair met for an hour of face-to-face tutoring in a small office, and they were video-taped. The pairs also met for an additional 2 hours each, when they were asked to communicate using computer terminals hooked to a VAX 11/780. They were given no instructions regarding any aspect of the tutoring, either structure or content. The students represent a fairly wide spread of expertise, from nearly complete novice to intermediate.

Detailed transcripts of each face-to-face session were made, following the conventions of Sacks, Schegloff and Jefferson (1974). In order to help the reader recreate as much as possible the sound of each fragment used below, I provide here a brief description of each of the major transcription conventions used. Readers familiar with this method of transcription can skip to section 2.

A double slash (//) indicates the place at which a speaker's utterance is overlapped by talk from another speaker.

T: but I know on physics exams, .hh you have to strea:mli//ne.
S: Aha

Thus with this notation we can see that the student's utterance starts after the i in the tutor's streamline.

An utterance which has more than one double slash in it is overlapped at more than one place, and the utterances which do the overlapping are given in sequential order after the overlapped utterance.

S: A//dd them vector//ially, yeah.
T: You have no angles.
T: Right everything's on the line.

Here, T's utterance overlaps S's starting at add and again within vectorially.

A left-hand bracket at the beginning of two lines indicates that the two utterances begin simultaneously.

T: You were going to say (2.7) charge on the electron>
(0.2)
T: Right?
[
S: Aha

Tutor and student begin talking simultaneously.

The equals sign (=) indicates latching, that is, the next speaker begins without the usual "beat" of silence after the current speaker finishes talking. In this case there is an equals sign at the end of the current speaker's utterance and another equals sign at the beginning of the next speaker's utterance. If two speakers simultaneously latch onto a preceding utterance (that is, they begin talking simultaneously), this is indicated in the transcript with a left-hand bracket preceded by an equals sign.

T: You've got meters, kilograms, meters, seconds.=
S: Seconds.
=[
T: So this is in ess ay units.

Here S and T simultaneously latch onto T's utterance.

Numbers given in parentheses indicate elapsed silence, measured in tenths of seconds. Single parentheses with a dot between them represent a silence that is less then a tenth of a second but still longer than the usual beat of silence.

Certain facts about the production of the talk are given through the orthographic symbols used. Punctuation is used to suggest intonation; underlining indicates stress. A colon after a letter means that the sound represented by that letter is somewhat lengthened; a series of colons means that the sound is increasingly lengthened.

The letter h within parentheses indicates "explosive aspiration," and usually means some type of laughter is being produced. A series of h's preceded by a dot represents an inbreath (where number of h's is meant to correspond to the length of the inbreath), while the same series preceded by nothing represents exhaling.

Questionable transcriptions are enclosed within single parentheses; the transcribers thereby indicate that the exact form of the utterance is not clear. Speaker's initials given in single parentheses means that there is some question about the speaker's identity. Single parentheses with capitalized words enclosed --e.g.,(CLEARS THROAT)--represent non-transcribed material (i.e., noise which is non-linguistic).

## 2 Interaction and Diagnosis

For the purposes of the present discussion, I propose that face-to-face tutoring consists mainly of two activities: description and explanation of some domain by the tutor; and working of problems by the student. Both of these activities are prominent in the sessions, with varying emphasis on one or the other.

In the second activity, the student participates by working through a collection of problems; in the first, the student participates mainly by showing that s/he understands what the tutor is explaining by completing the tutor's sentences, repeating the final phrases of the tutor's sentences, anticipating a crucial inference from what the tutor has just said (not only do I understand what you've just said, I can see in advance the conclusion you were going to draw from it), and by answering occasional questions. These are illustrated below.[1]

student completes tutor sentence

(1)

T: the rule is (1.4) most times it k- has to be just a number,
a constant If you're missing a variable
(0.2)
S: You can't do it.
[
T: you're in trouble.

(2)

T: how much energy does an electron have in a two volt
field. The answer being two electron volts.
[
S: two electron volts.

---

[1]Saying right or mhm does not really show that the student understands what the tutor is explaining; it merely claims understanding.

student repeats final phrase of tutor's sentence

(3)

T: you know you got this three hundred electron volts,
   and you go- and you always go, oh my God
   what (0.4) what is an electron vo//lt
S: and what is an electron volt

student supplies conclusion to tutor's statements

(4)

T: So (1.3) the main thing with the exponential function,
   (1.8) you have (0.2) ee to the ex (1.4) and its derivative
   is (2.3) ee to the ex // (    )
S: Right.
      (1.7)
T: So
  [
S: So.
      (1.6)
S: if you had ee:, to:, the:, ef of ex,
T: Right, exactly. That's (the) important case.=
S: =Right.

student answers tutor question

(5)

T: Okay, what's ohms
      (1.4)
S: It's resistance

Notice that the student's level of understanding in these passages is reflected in his/her participation in the developing conversation, rather than by simply solving steps of a problem.

Furthermore, the exact nature of the student's contribution gives the tutor an extremely delicate metric for establishing the student's level of understanding. For example, if the student completes the tutor's sentence, then that suggests one level of understanding; if the student actually **anticipates** that sentence, then we have an even greater level of understanding. Notice in passage (3) above, for example, how enthusiastic the tutor is in her response to the student's anticipation of her next move.

One can assume from these kinds of examples, then, that if participating in this way means that the student is displaying understanding, then if an opportunity to perform one of these actions arises and the student remains silent, the tutor can legitimately infer from this silence that the student does not understand fully. In this way, each opportunity passed by the student is a kind of metric for indicating how well the student is keeping up with what the tutor is saying.

Consider, for instance, the following passage, in which the tutor provides an explanation for a fact she has just mentioned:

(6)

T: Okay, and here's a time (0.2) when you can drop
the absolute value. (1.7) 'Kay .hh because,
again as long as the inside stuff is always positive
you can get rid of it, .hh and exponential functions
(0.7) are always po//sitive=
S: Right, yeah.
T: =.hh so you can if you want to: drop this out,
(1.5) since um exponentials (0.4) are always positive.

This passage follows a lengthy discussion concerning absolute value. If the student had understood that discussion, then he should have been able to either anticipate the tutor's utterances here or complete the first one once she had started. The student does neither--he in fact remains silent--which prompts the tutor to provide an explanation for what she has just said, since clearly the student does not understand. There are at least 3 obvious places where the student could have come in:

After "here's a time" (note the slight pause at the end of this phrase)
or somewhat later in the sentence

During the (1.7) second silence

At some point during the sentence beginning with "because.."

Since the student does not come in at any of these places with a turn that suggests he understands what she is getting at, the tutor continues ahead with an explanation. And, when the student finally says something, all he says is "right, yeah." which only **claims** understanding; it does not **display** it (see Schegloff, 1981, for a similar distinction). The tutor takes this as further evidence that the student's level of understanding is problematic, so she draws the "upshot" from what she has just said ("so you can if you want to drop this out"), and finally, since there is still no contribution from the student, repeats the explanation ("since um exponentials are always positive").

A similar example follows. In this passage, the tutor tries to elicit participation from the student; when she hears that he is not able (or willing) to join in, the tutor estimates his understanding to be insufficient for the task at hand and so offers the solution herself, laid out in detail:

(7)

T: we have one power so we'll call it (ey) over ex
minus four. (0.8) .hh Okay. We want these two things
to be equal (0.7) .hh so- the first thing we could
do is just make them look more alike (0.7) get the
common denominator. (0.3) .hh This is a step that
you will skip after a few times. .hh but this is (0.2)
where it comes from. Get the common denominator, so
ey ex minus four plus bee ex (0.2) .hh over ex times
ex minus four.

Note especially the two silences of (0.7). "We want these two things to be equal" clearly invites collaboration from the student in how to make two things equal. The tutor even offers the next step in a vernacular formulation ("make them look more alike"), followed by more silence, within which the student could offer the technical formulation. He remains silent, and the tutor offers the desired formulation ("get the common denominator").

In cases where the student is being introduced to the material for the first time, the inferences drawn from lack of active student contribution may be slightly less negative. In the fragment given below, for example, the student does not complete any tutor utterances and does not anticipate a natural conclusion, so the tutor carries out the discussion making all of the conclusions clear, providing all the propositions the student might otherwise have provided, had he understood more.

(8)

T: And now in this case all of our derivative is there (1.1)
We're not missing anything .hh So you're actually
just integrating one over yu dee yu, .hh (1.5)
.hh which gives you back a natural log. okay>
.hh Um Something very nice happens (0.4) with this
integration in that you actually pick up more than you
(1.4) than you would expect. .hh You pick up natural
log of the absolute value. (1.6) Okay? .hh (0.3)
Why: is that:. (1.4) ( I'll just write this out)

Timing is also used as a metric of understanding in cases where the tutor is more explicitly eliciting information from the student, as with a question. The length of time the student takes to answer, in addition of course to what s/he is doing during that time, gives the tutor subtle cues as to how easily the student is able to answer. If the student stares blankly for one or two seconds, the tutor is almost certain to intervene with a hint, or another question; if the student appears to be "working on it," the tutor is likely to give him or her leeway until either the correct answer is achieved or the student gets stuck (see Fox, 1987, for a detailed discussion of the mechanisms of tutor intervention).

(9)

T: so for degree two we put which degree in the numerator.
(1.1)
T: You're down one
(4.8)
S: Y:eah::, but this is not (0.3) factorable any further.

(10)

T: that means that- you have some other technique to do it,
something that we've already learned how to do. (0.3)
.hh So, what's the only other thing we know how to do.
(0.8) Essentially.
(1.1)
T: .hh You let yu equal that inside part.
[
S: Oh, just
(1.9)

T: Substitution

(11)

T: and what's a first degree polynomial look like
(4.5)
S: I don't- I don't foll//ow. Just ey ex?
T: . (Okay)
(0.4)
T: Ey ex (.) plus
S: Oh. plus=
T: =Bee.=
S: =Oh. all right. yeah.

In the second activity, timing in the interaction still represents a crucial source of diagnostic information. although via a slightly different mechanism than in the earlier cases. Recall that, in the second activity, the student has set about to solve problems in the textbook (in one case the student had worked through the problems before coming in and had brought in only those problems which she had been unable to solve herself).

The basic structure of these interactions is something like question-answer, only instead of a one-step answer the student must provide a step-by-step procedure for working towards the answer. The timing of these steps is a major source of information for the tutor about the student's understanding of the problem.

Early on in one of the sessions, the student and the tutor negotiate that the tutor will read the problem out loud (presumably to familiarize herself with the problem) and then the student will show what she had done previously to work the problem. But note in the following fragment that after the tutor reads the problem the student does not come in with her approach to the problem.

(12)

T: A one ohm wire. Okay, what's ohms
(1.4)
S: It's resistance

T: Okay. Is drawn out to three times its original length.
What is the resistance now.
(2.1)
T: I think that's a hard question.

T: Write down the volume of this guy.
(1.2)
S: Okay, how do you do volume here, let's see: (2.1)
This is the area, and these are the lengths,
so volume is (1.4)
T: Area times length.
S: Area time- oh, that's good ((laugh))

In the following fragment, the tutor helps the student through the first few steps of the problem-solving process. During the (4.8) silence, the student is trying to carry out the second step of the process, namely to factor the denominator. After a certain point in the silence, he stops moving his pencil and displays "being stuck." The student's laughter indicates that he is stuck, and the tutor intervenes.

(13)

T: Second thing> i:s> (1.0) to>
(1.7)
T: fa//ctor>
S: factor the denominator.
T: Right.
(4.8)
S: heh heh heh // heh heh, heh
T: Okay, (0.7) here's the important thing to notice about
this one. You- you copied it okay. HA HA HA

The timing of the student's response (delayed with regard to their isolating of the next step), in addition to the kind of response (laughter), indicate to the tutor that the student's ability to proceed with the problem is questionable.

In passages involving student problem-solving, as in the first activity, the timing of response to questions is monitored closely for clues to the student's understanding:

(14)

T: what's the difference between weight and mass.
S: I used to know this let's see
(1.7)
T: I think, (0.8) I think what it is (0.8) is that.
(0.5) n- what is uh, when you do (0.6) uh:,
gravity problems.-
S: =Right.//It's that-
T: What do you always do
(1.2)
S: You have to multiply it be the-=
T: =by gee.=
S: =gee. Right.

The following passage beautifully illustrates the tutor's use of interaction to estimate how much the student understands. Just prior to the beginning of this fragment, the tutor constructed an analogy between a person crawling through two different pipes and the resistance of a wire after it has been stretched to three times its length.

(15)

T: Would it be: easier for you to go through thi(hh)s
pipe, (0.9) or this pipe.
(1.6)
T: 'Kay, 'cause we // stretched it out to three times its length.
S: Yeah.
S: Yeah.
T: It'd be a hell of a lot easier to crawl through this//
pipe (than if it had) to go through this pipe.=
S: (P-)
S: =Right.

The tutor tries to elicit an answer from the student at a minimum of 3 places before she provides a hint: when the student still fails to initiate an answer, the tutor finally answers her own question. Notice that the student does not even collaborate in the construction of the answer once the tutor has started it, but merely claims to understand the answer once the tutor has completed it.

I would like to examine one more piece of data before moving on to looking at the metric in more detail. This fragment illustrates clearly that the tutor responds not only to what the student says but how s/he says it:

(16)

T: Are you familiar with um> (1.6) integration by parts?
(0.2)
S: .hh Yeah I've=
T: =Sort of.=
S: =Yea//h I've done it and, but
T: Okay.
T: Okay.

In this passage, the delay in the production of his answer, in addition to the inbreath that proceeds the answer itself, gives the lie to the content of what the student is saying. Even though his answer on the surface says that he is familiar with integration by parts, the way he delivers it hedges the degree of familiarity. The tutor picks up on this hedge immediately and makes it overt: "sort of" (and the student accepts her formulation of what he was doing by answering that way).[2]

## 3 Metric of Interaction

The data presented above provide clear evidence that tutors monitor not only what students say but when and how they say it; furthermore, what is **not** said (and when it is not said) is just as informative to an astute tutor as what is said. How can this metric of interaction be described?

In everday interaction when people describe something as not having happened, we take it to mean that this something was relevant at a particular moment, at which moment it did not take place. It would be extremely odd for me to utter (17), for instance, in July in Hawaii, since snow is not expected to occur under those conditions.

[2]Another framework within which one can address these issues is the preference organization (Pomerantz, 1975). According to preference organization, certain responses are socially preferred over other alternative responses. For example, if one is invited to a social event, it is socially preferred to accept the invitation and dispreferred to reject it. Similarly, if one is asked a yes/no question with positive polarity, the preferred response is yes. Furthermore, while the preferred response is usually done immediately and without much elaboration, the dispreferred response is delayed with regard to the thing it responds to, is usually accompanied by an explanation, excuse or some other kind of elaboration, and regularly starts off with a version of the preferred response. The student's response in this passage is therefore done as a classical dispreferred: it is delayed, it has the beginnings of an explanation, and it starts with a version of the preferred response ("yeah"). The dispreferred format indicates that the student is really saying "not really" or even "no" when on the surface he is saying "yes."

(17) It didn't snow today.

while it might be completely common for me to utter the same thing at the end of Christmas day in Boulder. Since there is an infinite array of things which are currently not happening, when one speaks of one particular thing as not happening it must be because that thing happening was relevant at that moment (Givon, 1979; Sacks, Schegloff and Jefferson, 1974).

So, when I speak of a student not having contributed at a particular point in the interaction, it must be because a student contribution was relevant, or expected, at that point. That is, I must have, either implicitly or explicitly, a theory of expectedness, or relevance (in this limited sense), in interaction from which I can determine when a student has relevantly not done something. Of course, I do not have a full, explicit theory of anything, much less of expectations in tutoring interactions; nonetheless, this research on tutoring, in conjunction with current work on everyday interaction, does provide a way of approaching a metric of interaction.

Let us suppose that an interaction consists of opportunities for both parties to speak. The tutor keeps track of the opportunities which the student passes and takes and plans her own behavior using these passed and taken opportunities as a kind of metric.

So, we need to sketch out when it is relevant for a student to contribute, or when there is an opportunity to contribute.

The most obvious slot in which a student contribution is relevant is after a question from the tutor. If the tutor asks a question, it is clearly appropriate for the student to answer. If the student remains silent in this situation, or says something that does not count as an answer to the question, we can speak of the student **not answering**, in other words, s/he has passed a relevant place to answer.

It is important to point out here that the answer is relevant even before the question is possibly complete; if the student is keeping up with what the tutor is saying, s/he will be able to predict, after a certain portion of the tutor's question, what the rest of the question will be and when it will be possibly complete. As soon as the student "sees" in this way what the question will be asking, s/he can offer an answer. The ability to correctly answer a question before that question has even been completely formulated is a clue to the tutor that the student is understanding the material fairly well.

(18)

S: And then, the velocity would be in (5.7) uhm
    (6.0)
T: (You) think of this over here.
    (2.4)
S: Something per second
T: Right. And what's th- what's the di//stance>
S: Meters>
T: Meters, right.

Student participation is thus clearly relevant after a tutor question. But even when the tutor does not so actively invite the student to collaborate, there are certainly places where it is relevant for the student to contribute. Following Sacks, Schegloff and Jefferson (1974), I will refer to these places as **transition relevance places**, inasmuch as transition to another speaker is relevant; TRP's are said to occur at the end of Turn Constructional Units--linguistic units from word to sentence.

TCU=word

S: And then this is the same as that one.=
T: =Right.
    (0.3)
S: Oh: okay.

TCU=phrase

S: Not of both places, just of one plate.=
T: =Just of one plate.
S: Okay.

TCU=clause

S: 'cause I don't think I had a dialectric (0.6) wait.
    no, I do:
    (0.9)

T: Ts, 'cause you have paraffin.

TCU=sentence

S: I just assumed that one was positive and one was negative.=
T: =Yeah, that's the only way it makes sense.

This system works as follows in most interaction: each speaker is alloted one TCU; at the end of a TCU by person A, person B may speak; if B chooses not to speak, A may speak again.

This normal turn-taking system can be temporarily suspended, however, so that one speaker can take what is called a multi-unit turn, for example to tell a story or explain a complicated aspect of differentiation. This suspension does not of course imply that no one else can speak at all; the recipient of the multi-unit turn nods, makes comments and in general collaborates in the production of the turn, these productions being located at or near the end of **clause boundaries**. The clause boundary, and especially the sentence boundary, of a multi-unit turn from the tutor is thus a relevant place for students to contribute.

The approach of the end of a clause and/or sentence boundary in a mult-unit turn raises several opportunities for the student, as we saw in section 2. The student can (1) complete the current clause, (2) repeat the end of the clause once it has come to a possible completion, (3) anticipate the tutor's next clause/sentence, or (4) ask a relevant question. These can be ranked in order of cognitive difficulty 3 > 4 > 1 > 2; the option chosen by the student, if s/he opts to do more than a simple continuer (e.g. mhm, right), gives the tutor some indication of how well the student is following.

A careful examination of the transcripts reveals that students insert their utterances within a multi-unit tutor turn within what I will call a **span of relevance**, lasting roughly from the last constituent of clause (a) to the first constituent of clause (b), as shown schematically below.

--------------/---. --/---------

Within this span, student contribution is relevant. In fact, it is relevant at two different structural slots within the span, namely within clause a and in the transition space between clause a and b (for the purposes of this paper we can treat the beginning of clause b as part of this transition space). Since within this span of relevance active participation from the student is relevant, then when the student does not say anything, or produces only a continuer, then the tutor can legitimately infer certain things about the student's understanding.

We have seen what happens when the tutor is the main speaker. I turn now to a case in which the student is the main speaker.

Within the activity of a student solving a problem, the student verbalizes steps of the procedure to display his/her understanding of each step. In some cases the student will produce a sentence which displays a correct understanding of one step, or of a preliminary to a step, and then will fail to produce the step that must logically follow, indicating that s/he does not see how the step s/he did produce is a resource for continuing the rest of the problem.

(19)

S: Ey plus cee equals zero.
(0.3)
T: Right, so that tells // you
S: Ey equals cee.
(0.4)
T: Minus cee.

(20)

S: So it's got to be: in our fourth quadrant.
T: Right.
(1.4)
S: Aha. (0.8) Aha, o//kay.
T: Which means?
S: Then I have to come back down here and I'm- you're
asking me to choose a sign, righ//t?
T: Right
S: .hh Okay.
T: For your cosine.

In each of these cases a correct preliminary move is followed by a bit of silence, wherein the tutor hears that the student is not going to produce the substantive move which logically follows from the preliminary move. Here again, a further student contribution is relevant, and when it is not forthcoming, the tutor overtly solicits it.

We thus have three foci within which a metric of interaction for tutoring is necessary: (1) after a tutor question, (2) in the span of opportunity within a multi-unit tutor turn, and (3) after a student utterance whose implication for the next step of solving a problem has not been elaborated. Of course, a tutor can always elicit student participation in other contexts, using other devices--the laughter in passage (15) above, for example (see Jefferson, 1979, for a discussion of laughter inviting participation)--but these are the most common spaces of student relevance.

I would like now to sketch out a metric of interaction that might begin to capture some of the regularities to be found in tutoring interactions regarding diagnosis and assessment. As I have presented it, the metric would not be strictly accurate for human-human interaction, since it is almost certainly not the case that humans keep an exact count of the number of missed opportunities (although they might well keep vague count-- more than 2, for example, would be a natural category to expect). Nonetheless, while people may not actively count and store values, it is clear that they make use of perceptual cues constructed by means of attending to patterns of temporal flow; inasmuch as these patterns are based in part, at least, on factors of timing and passed and taken opportunities, then the metric offered below affords a useful starting point for thinking about timing and diagnosis.

The metric of interaction I suggest is the following. At each structural slot of student relevance a timer will be turned on; when the structural slot has passed, the timer will turn off and a counter will be updated by a value of 1. Each kind of structural slot will have its own timer and its own counter, since the tutor's response will be heavily influenced by the kind of context and so we might want to differentially weight the counters (we might want the question counter to have more weight in the overall metric than the multi-unit counter, for example, so adjusted weightings can be determined in a separate function). These two pieces of information, together with what kind of turn the student does produce (e.g. anticipation of the tutor's next move vs. completing her current utterance), give us a way of estimating the student's level of understanding.

Why have both a timer and a counter? The counter is needed to determine how many slots the student has passed up; the timer is needed to see where within a slot the student responds. For example, the student's understanding will be judged differently if s/he responds after 1.4 seconds in the third structural slot than if s/he responds after 0.2 seconds in the same slot.

Let's see how this metric works in passage (15) from above, repeated here as (21).

(21)

T: Would it be: easier for you to go through thi(hh)s
pipe, (0.9) or this pipe.
(1.6)
T: 'Kay, 'cause we // stretched it out to three times its length.
S: Yeah.
S: Yeah.
T: It'd be a hell of a lot easier to crawl through this//
pipe (than if it had) to go through this pipe.=
S: (P-)
S: =Right.

A timer would be turned on at "thi(hh)s pipe," where the tutor has invited the student, with her laughter, to collaborate, turned off at the end of the 9/10 of a second silence, and turned on again after "this pipe," turned off at the end of the silence, turned on again near the end of line 4, turned off at the student's line 6, turned on towards the end of the first embedded clause at lines 7-8, turned off, and turned on again towards the end of the line 8. The counter would be updated at each of these structural slots: 3 before the tutor finally offers a hint at line 4; 4 (or even perhaps 5) before the tutor finally gives the answer. In this case the student's "yeah" is treated virtually as silence, since, like silence, it passes on the structural opportunity to do what is relevant--in this passage, answer.

Another example follows.

(22)

[T and S are talking about absolute value]

T: if ex (.) were negative to start with (0.6) now minus
ex would be positive. .hh So:=
S: =Oh so you have to ah (1.2) know what the number is
befo//re you can say
T: Right. exactly. Before you can simplify.

The timer would turn on towards the end of the first clause, turn off at the end of the clause, back on through the (0.6) silence, off at the beginning of the next clause, back on at the end of the sentence, through the inbreath to what looks to be the beginning of the "upshot" of this whole explanation. The student comes in, at the very end of the span of relevance, with his guess at the upshot the tutor was about to produce. Note that because the student passed the one relevant opportunity to show he could anticipate what the tutor was going to say, and looked as though he was going to pass a second opening which the tutor has provided (by lengthening the transition space with an inbreath), the tutor prepared to produce the upshot herself. This is quite different from the situation illustrated by passage (21) above, where the student passed 3 opportunities, with fairly long silences, before the tutor spoke again.

The differences between these two passages are not of mysterious origin; rather, they follow from the fact that in passage (21) the tutor has selected the student to speak next, while in passage (22) the tutor has "the floor" even through the span of relevance. Since in a sense it is a violation of social norms for the tutor to speak again in the first case, the student must produce a greater number of signals that s/he is not understanding before the tutor will take another turn; whereas if the tutor already holds the floor, then it is no violation for her to speak again. The exact effects of the metric on the tutor's tutoring strategy is thus context-dependent.

While it is apparent from the data fragments presented here that both the number of opportunities passed and the length of each passed opportunity, in addition to the kind of participation the student eventually produces, are extremely reliable clues to the level of a student's understanding, I do not believe that there are precise mathematical algorithms which can take the values from these variables as input and produce either a value for the student's level of understanding or the tutor's behavior as output.

The values for these variables cannot be taken as precisely describing the exact level of the student's understanding for a variety of reasons. First, if the tutor is explaining material that is genuinely new to the student, it is not necessarily an indication of lack of understanding if the student does not anticipate every major conclusion or complete every utterance; second, not all utterances in a long tutor turn are equally inviting of student collaboration; and third, no lack of understanding should be inferred if it takes a student 5 seconds to calculate the answer to a tutor question (although in this case we might want to have the timer turned off as soon as the student begins writing calculations down).

Nor can these values be taken as direct predictors of the tutor's response. The tutor's exact response reflects a variety of factors, including

* the nature of her own previous utterance,
* whether the topic under discussion has been explained before (and therefore ought to be part of the student's "bag of tricks"),
* what the student is doing during a particular silence (e.g. calculating the answer vs. "staring blankly") and not just the length of the silence,
* in a multi-unit tutor turn, the importance of the projected next tutor turn (whether it will describe the upshot of what they have been talking about or whether it is another piece of preliminary material).

In spite of a lack of mathematical precision, however, this metric, when taken together with contextual information, can provide a fairly delicate guide to tutor-student interaction. Here are just a few of the recurring patterns which the metric allows us to make explicit.

1. If the student passes 3 or more structural opportunities to answer a question, the tutor is very likely to answer the question herself.

2. If the student passes one structural opportunity to answer a question and fails to respond after 2 seconds within the second opportunity--and if the student is displaying signs of looking confused or not in control--then the tutor is very likely to offer a hint or clarify the question.

3. Within a multi-unit tutor turn, if the student passes 2 or more opportunities, the tutor can assume the student's understanding is not at a very high level and may further motivate the point she has just made.

4. If the student passes on only one (or no) place of relevance and then anticipates an upshot, the tutor is likely to respond with greater than average enthusiasm.

## 4 Conclusions

The goal of this paper has been twofold: to demonstrate that tutor's base their assessment of student understanding on more than steps in a problem-solving task; and to introduce a simple metric for beginning the process of formulating how tutors assess student understanding on the basis of the tutoring interaction itself.

The relevance of the metric to tutoring systems should be apparent. While a simulation of a human's use of the metric is almost certainly not to be automated, the principle of monitoring the timing of student participation (and lack thereof) must be transferable to machine tutoring.

It is worth pointing out here that from the student's perspective there is also a metric of tutor response. If the tutor starts nodding and agreeing with the student's turn even before that turn is possibly complete, then the student can infer that the answer displays a high level of understanding, whereas if the tutor delays responding after a student's turn, then the student will hear that to mean that the last turn was in some way less than fully correct or appropriate. The following fragments illustrate the tutor metric.

(23)

[T and S are talking about absolute value]

T: if ex (.) were negative to start with (0.6) now minus

ex would be positive. .hh So:=
S: =Oh so you have to ah (1.2) know what the number is
befo//re you can say
T: Right, exactly. Before you can simplify.

(24)

T: We actually have (0.6) an extra factor of=
S: Two.
T: Two.

(25)

S: That should be twelve shouldn't it?
(1.0)
T: No: I think it should be, twenty four.

Furthermore, this work emphasizes a perspective which treats all dialogue as deeply interactional. This perspective includes thinking of interaction not only as what people do say, but what they don't say, and when they don't say it, that is, as passed and taken opportunities; and thinking of what appears to be monologic material, such as long tutor explanations, as in fact interactional, produced in delicate response to minute facts about the recipient's behavior. Tutoring interactions should thus never be treated as pre-planned, since even in the most monologic situations the talk is still fundamentally interactional, responding as it does to things said and not said by the recipient, things whose presence or absence could not have been predicted in advance.

## REFERENCES

Anderson, J., Boyle, C.F., Corbett, A. and M. Lewis. 1987. "Cognitive modelling and intelligent tutoring." MS.

Arce, M. 1987. "On the nature of tutorial dialogue and its implications for computer tutoring systems." MS.

Fox, B. 1987. Cognitive and interactional aspects of correction in tutoring dialogues. MS.

Givon, T. 1979. On Understanding Grammar. New York: Academic Press.

Oberem, G. 1987. "ALBERT: a physics problem-solvinng monitor and coach." Paper presented at the Third International Conference on Artificial Intelligence and Education.

Sacks, H., Schegloff, E. and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking. Language 50, 696-735.

Schegloff, E. 1981. "Discourse as an interactional achievement." In D. Tannen (ed.) Text and Talk. Washington, D.C.: Georgetown University Press.

Sleeman, D. and Brown, J.S. (eds.). 1982. Intelligent Tutoring Systems. New York: Academic Press.

Woolf, B. 1984. "Context-dependent planning in a machine tutor." Ph.D. dissertation, University of Massachusettts, Amherst.