

# Extensions of GOMS Analyses to Expert Performance Requiring Perception of Dynamic Visual and Auditory Information

Bonnie E. John

School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA, 15213  
412-268-7182, bej@cs.cmu.edu

## ABSTRACT

GOMS models of telephone toll and assistance operators (TAOs) are being constructed in an effort to provide theoretical predictions of expert performance on several dedicated workstations. This applied effort has pushed the development of GOMS modeling techniques into the area of speech input and output, and into a task where information is not always available when it is required by the TAO. This paper describes the task, heuristics for constructing the GOMS models, and parameters for making quantitative predictions of performance time.

**KEYWORDS:** user models, cognitive models, GOMS, Model Human Processor

## GOMS ANALYSES TO DATE

A recent review of cognitive modeling in human-computer interaction (HCI) indicates that GOMS analyses has been used to model many tasks in a diversity of domains [7]. These tasks have several common characteristics. Most of the stimuli are presented with a static visual display. Editing tasks were presented as red marks on a paper copy, choice reaction time stimuli were presented at a single position on a CRT screen, spreadsheet formulae and database queries were described on hard-copy, and typing tasks primarily involved transcription from hard-copy or from a well known position on a CRT screen. Also, the user was free to work at his or her own pace; the user did not have to wait for critical information to appear and critical information did not disappear.

These common task characteristics lead to very similar GOMS analyses. Because the input is primarily visual, static, and available before it is needed in the task, the user can employ strategies that take in information whenever necessary. Two strategies have been predicted and observed. The first strategy is sequential, where information is perceived right before it is needed for the next task. This leads to the *unit tasks* proposed by Card, Moran & Newell,

[2] where the user gets information that defines the next subtask and accomplishes that subtask, then gets the information for the next subtask, and so on. This strategy produces behavior that can be analyzed into a sequence of perceptual, cognitive, and motor processes. A second strategy involves looking ahead for information, as when a typist looks ahead of what his or her fingers are typing. This parallel strategy leads to observed behavior dominated by the duration of motor responses, with the perceptual processes always in the background. A strategy where the user overtly records information for future use is explicitly not predicted. Such a strategy would be inefficient in a task where information is always available, and it is not observed in actual task performance.

Although the tasks analyzed with GOMS to date include a large portion of HCI tasks, there are other important task characteristics that are not represented in these analyses. Even staying within the traditional GOMS domain of error-free, expert performance, there are HCI tasks with other perceptual input and response modes and other dynamic characteristics of perceptual input. The task described in this paper has fleeting auditory and visual input, that can occur before or after the information is needed in the task. It also includes speech input and output. In the remainder of this paper, the task will be described, heuristics for conducting the GOMS analysis will be presented, and new parameters will be introduced.

## THE TASK

The task under examination is that of a telephone company toll and assistance operator (TAO) responding to customer requests for assistance. The TAO sits at a dedicated workstation specifically designed for these tasks. He or she has had explicit training in the most efficient call-handling procedures, and has had sufficient practice to become expert at the task. There is typically no problem solving involved; the TAO simply recognizes the call situation and executes routine procedures associated with that situation. Some types of calls involve substantial system response times (e.g., while a credit card number is being verified).

Consider the following example. A customer dials 0 plus the desired telephone number. A "Call Arrival Tone" (CAT) sounds at the TAO's workstation. The TAO uses different

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish requires a fee and/or specific permission.

greetings for different types of calls, so he or she must scan relevant information automatically displayed on the screen to choose an appropriate greeting. In this example, 0+ appears on the screen and the TAO says "New England Telephone, may I help you?". The customer says, "Make this a collect call from Ann". The TAO presses the appropriate keys on the workstation keyboard, writes down the customer's name for future reference, and waits for verification that this number can accept collect calls. When the call is put through and the called party says, "Hello", the TAO says "I have a collect call to anyone from Ann. Will you pay for the call?". In this case, the called party says "Ah. Just a second. Babe, do you know an Ann.... Yeah, we'll accept it. Thank you." The TAO then presses a key to release the workstation and make it ready to accept the next call. The TAO also crosses out the name, in preparation for the next call.

The most striking difference between the TAO tasks and previously studied tasks is the use of the auditory channel. The temporal nature of speech makes the information fleeting; if the TAO is not attending to the customer when particular content words are voiced (e.g., "collect", "accept charges"), then the information is missed. (This information is fleeting, but not irretrievable, because the TAO could ask for the information to be repeated, but this is an inefficient strategy that increases call-handling time.) The visual input channel is also different from previously studied tasks because system response time makes some of the visual input unavailable when the TAO is ready to use it.

### A GOMS ANALYSIS OF THE TASK

TAO call handling is being analyzed using a GOMS approach, in an effort to provide a theoretical prediction of performance time that can be used in the evaluation of different TAO workstations [3]. We are in the early stages of using GOMS models for this task and have developed a tentative set of heuristics for use in the analyses.

### The GOMS Methodology

A GOMS analysis of TAO tasks consists of defining the goals, operators, methods, and selection rules used to handle calls. These are inferred from detailed knowledge of the task obtained from real-time observation of expert TAOs, videotapes of expert TAOs handling calls placed by experimental confederates employed by the telephone company, and knowledge of the training undergone by new TAOs.

The overriding goal of the TAO is to complete all calls as efficiently as possible. Training material presents this as answering three questions: who pays for the call, at what rate, and is the call complete? TAO training teaches the most efficient method for completing each subgoal; each keystroke, each phrase to be spoken, each keyword to listen for, is explicitly taught. Since only the most efficient method is taught, there are no different methods to choose between and hence no selection rules. Since the goals are uniform across call types and there are no selection rules, the GOMS analysis reduces to determining the operators

necessary to complete a call, the sequence in which they are performed, and the performance time parameters associated with each operator.

We have chosen to represent the GOMS analysis in a *schedule chart* representation within a project management application program (see [5] for a detailed description of how these charts can be used for GOMS analyses). Each behavior is represented as an event box with an associated duration and information flow dependencies (Figure 1). The videotape is used to identify overt behaviors (e.g. keystrokes and words spoken), and provide actual start times, end times, and durations for each overt behavior. Unobserved behaviors (e.g. eye movements, cognitive operators) are also represented by event boxes. The information flow constraints, depicted by lines drawn between the event boxes, are the substance of the task analysis and reflect the analyst's knowledge of the task.

### Heuristics For Modeling

Heuristics for building GOMS schedule charts have emerged in the course of examining several specific calls. Most of these heuristics have theoretical foundations beyond the scope of this paper. Developing heuristics has been, and continues to be, an iterative process, with initial assumptions proved incorrect by observations of performance, or unanticipated observations leading to new assumptions. In order to provide a coherent explanation of how to build these charts, the iterative process has been suppressed. Examples of the schedule chart representations of these heuristics can be found in Figure 1.

*The Expertise Assumption.* It is assumed that the TAO is an expert and can anticipate likely call situations from preliminary information. Given the TAO's goal to complete the call quickly, it is assumed that the TAO will perform as many activities as possible in preparation for the arrival of expected information. For instance, the TAO will move his or her eyes to the field on the CRT screen where information will appear before it appears, and the TAO will move a finger to the next most likely key to be pressed before confirmation of the appropriateness of that key is received.

*System Response Time Heuristics (Fig. 1a,b).* System response time (RT) is the time it takes for the system to provide information to the TAO. For instance, the time between the end of the downstroke of a key and the initiation of the display of that key name on the CRT screen is considered to be system RT.

System RT has always been a component of GOMS modeling [2], but it has not played a major role in previously studied tasks and only one system has been involved in any one task. There are three systems in TAO tasks and each RT is recorded separately. The systems are the TAO workstation itself, which has screen display time (SDT) and CAT duration; a call-handling system (CHS), which provides the TAO workstation with information about the call and receives keypress information from the TAO workstation; and a billing-verification system (BVS), which receives information about the desired billing from

the call-handling system and provides verification of billing to the CHS.

To build system RTs into the GOMS schedule chart, begin with a dummy start event box of zero duration (Fig. 1a), from which subsequent system RTs can be measured. The CAT event box immediately follows from this dummy start box, with a duration set from the videotape. Parallel to the CAT, a CHS RT event box follows the dummy start box and ends when information begins to be displayed on the CRT screen. A SDT event box stems from this CHS RT event, beginning when the information begins to be displayed, with a duration set to the time it takes to become fully visible.

Information from the CHS is assumed to be sent simultaneously to the TAO workstation and to the BVS. Thus, the SDT event box showing that a signal has been sent to the BVS is represented as occurring in parallel to the BVS RT (Fig. 1b).

*Getting Auditory Input (Fig. 1a,c,e).* When auditory input is anticipated, the TAO performs a cognitive operator called "attend-to-x", where x is either the CAT or a customer's voice. Subsequent perceptual operators depend upon the completion of this operator.

The only non-speech auditory input necessary to the TAO tasks is the CAT. This tone is assumed to be perceived by a single, generic perceptual operator and verified by a cognitive operator (Fig 1a).

It is assumed that the TAO recognizes key-words in the customer's speech as soon as they are uttered. The perception of customer speech is represented as perceptual operations on phrases that end with key words.<sup>1</sup> In the example, "Make it collect from Ann" is parsed into a "Make it collect" perceptual event box and a "from Ann" perceptual event box (Fig 1e). The temporal dependency of this statement is preserved by making the duration of the perceptual operators equal to the observed duration of the phrases and by drawing a dependency line from "Make it collect" to "from Ann". However, actions that may be appropriate from information in a phrase can begin after the perception and cognitive verification of that phrase, without waiting for other phrases to complete.

The TAO must avoid interrupting the customer, so it is assumed that the TAO waits an appropriate length of time to make sure the customer has stopped speaking. This is

---

<sup>1</sup> To be more rigorous, the analyst might want to include an "auditory event" to represent the physical presentation of auditory information, similar to the SDT event box that represents the physical presentation of visual information. Then there would be a separate perceptual operation for that auditory information. Since it is assumed that the perceptual operation is coincident in time with the physical presentation, this parallel representation is awkward in MacProject, and totally redundant, so I prefer to represent only the perceptual operations, as described in the body of the text.

represented by a "silence" perceptual event box after the customer's last phrase. A cognitive operator verifies this silence (Fig 1e). (Likewise, the customer waits for the TAO before speaking, indicated by a "customer pause" event after the TAO speaks, Fig. 1c.)

*Getting Visual Input From The CRT Screen (Figure 1b).* Information presented on the CRT screen is assumed to be well known to the TAO. Predictable information is presented in predictable fields. The task of the TAO is to recognize the information presented, as opposed to deciphering unexpected information. Given the task situation, the next most likely piece of information can be anticipated, so the TAO can prepare for its presentation.

When a visual input is anticipated, the TAO performs a cognitive operator called "attend-to-X", where X is the field where the information should appear (Fig. 1b). The TAO then directs his or her eyes to that field with a cognitive operator to initiate eye movement and a motor operator to move the eyes to the appropriate field. If the TAO has received visual information in that field immediately prior to the current situation, with no intervening visual or auditory perception or cognitive activity, then it is assumed that the eyes remained fixated on the appropriate field and these operators are not necessary (Figure 1b, second attend operator).

After the eyes are fixated on the appropriate field and the information is completely displayed on the CRT screen, a perceptual operator takes in the visual information. Finally, a cognitive operator verifies the existence of the expected information.

*Hand Movements (Figure 1d).* Observations of TAOs in a real work situation revealed that keys are pressed primarily with the right hand. The only exception is POS REL, the key that releases the workstation after a call is complete. POS REL is located on the extreme left of the keyboard and is almost exclusively pressed with the index finger of the left hand. This information, combined with the GOMS Model of Expert Transcription Typing [5,6] form the basis for the GOMS model of hand movements in TAO tasks.

Hand movement motor operators are described at the lowest level for which previous GOMS research provides performance parameters (Section 3.3.1). There are operators for moving between devices (between lap or tabletop, keyboard, and notepad), keystrokes, and writing. Keystrokes are further separated into three parts: horizontal movements to position the finger above the key, a down-stroke to press a key, an up-stroke to release a key.<sup>2</sup>

---

<sup>2</sup> Other hand movements may occur, like adjusting the screen and doodling. These movements are not predictable from the task goals or the TAO training, so are outside the realm of the GOMS analysis. However, as predicted by the assumption that the overriding goal of the TAO is to complete the call as efficiently as possible, such movements are observed to occur while the TAO is waiting for information and rarely have an effect on the duration of the call.

Figure 1a

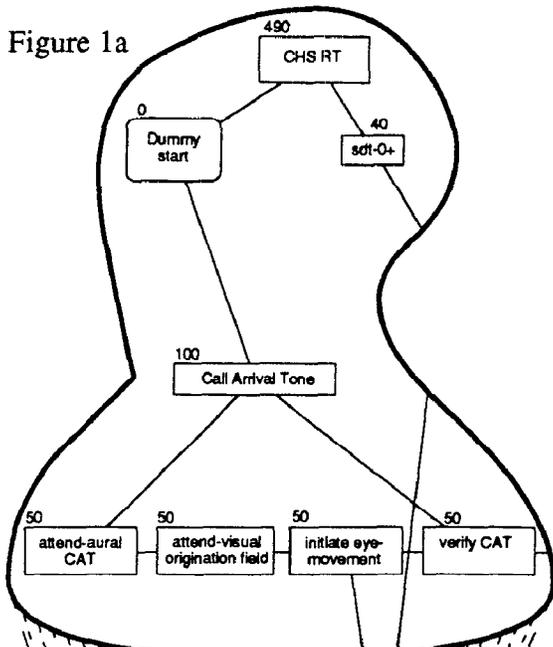
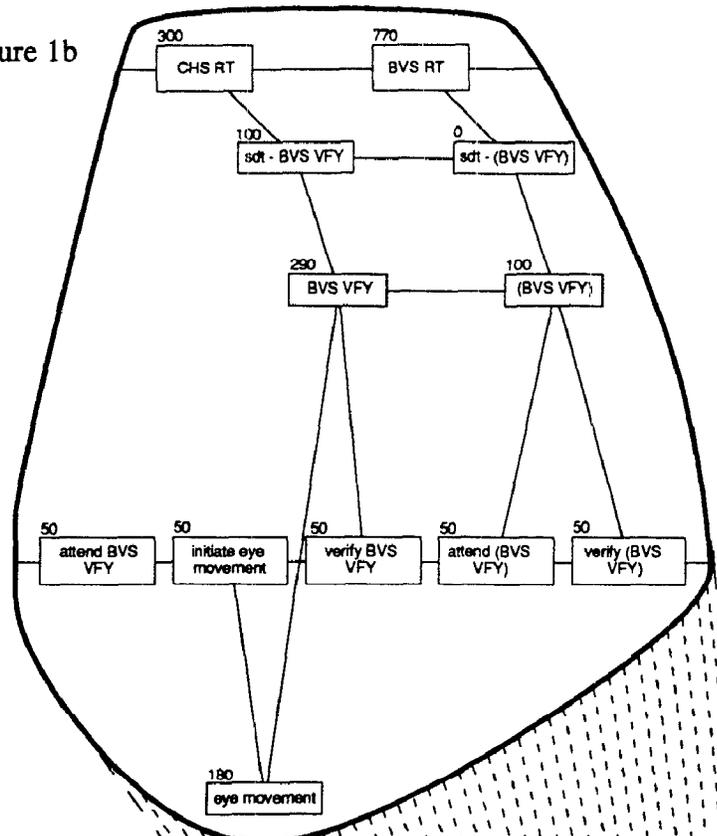


Figure 1b



System Response Time

Perceptual Operators

Cognitive Operators

Hand Movements

Eye Movements

Verbal Responses

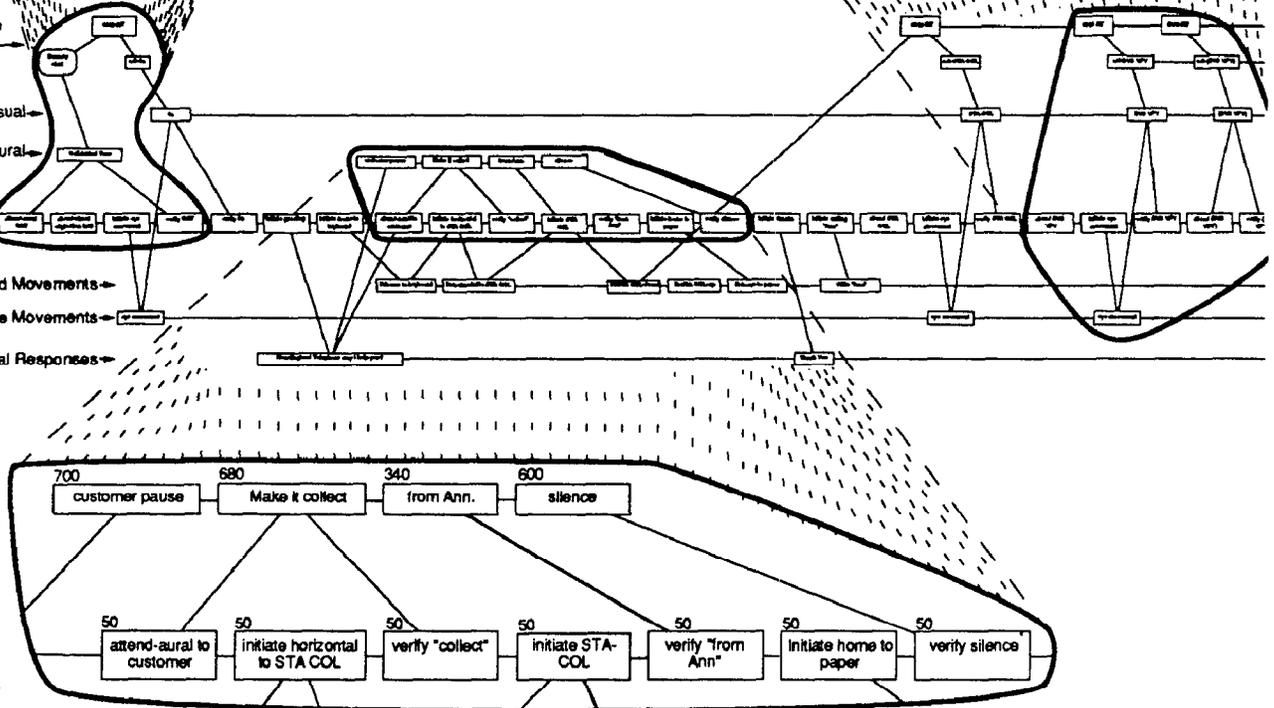


Figure 1c

Figure 1, a-e. The small chart stretching across the bottom third of the page is the complete schedule chart for the collect call used as an example in the text. It includes all perceptual, cognitive, and motor operators necessary to complete the call, and the system RT inherent in the call. Each operator is represented by a box; each information flow dependency, by a line connecting two boxes. All operators of the same type appear in the same horizontal row, as labelled at the extreme left of the chart. The estimated durations (in msec) appear above each blown-up box. Figure 1a demonstrates the heuristics for initial system RT and the perception of a simple auditory stimulus. Figure 1b demonstrates the heuristics for parallel system RT and the perception of visual information. Figure 1c demonstrates the heuristics for the production of well-practiced speech. Figure 1d demonstrates the representation of hand movements. Figure 1e demonstrates the heuristics for the perception of customer speech.

Figure 1c

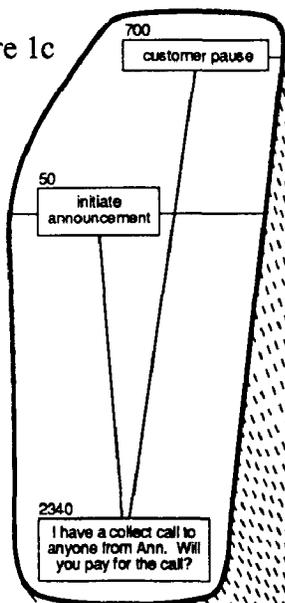
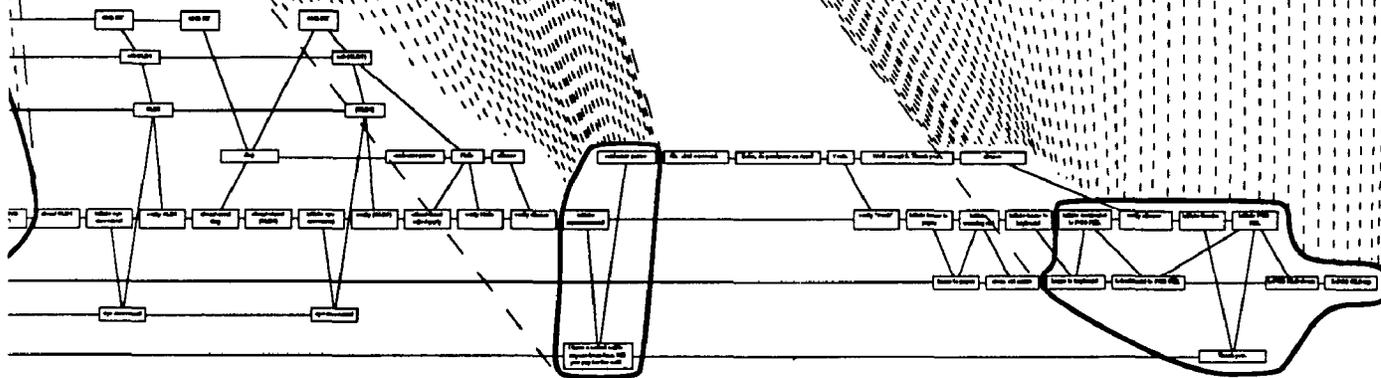
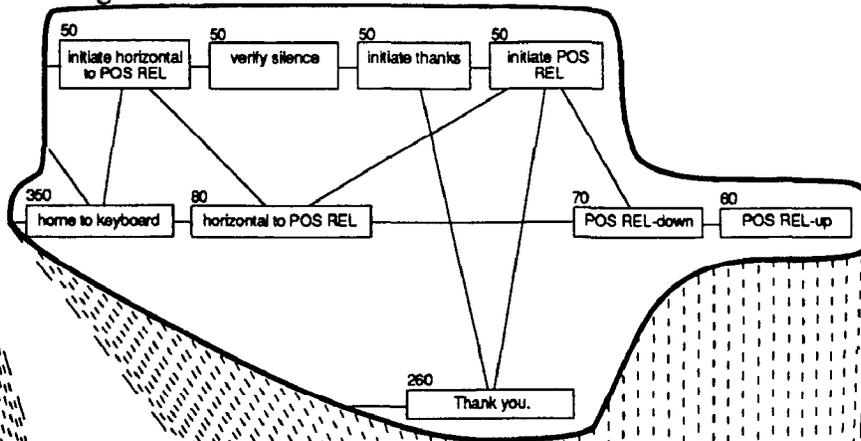


Figure 1d



Each hand movement chunk is initiated with a cognitive operator and implemented by one or more motor operators. The size of the movement chunk is determined by knowledge of the call situation. For instance, the POS REL key releases the workstation, making it available for the next call (Figure 1d). This key cannot be pressed until the call is complete. Thus, the POS REL-down hand movement cannot be initiated until after the TAO says, "Thank you" to the called party. However, the end of the call can be anticipated well before this acknowledgement, so the finger can be positioned over POS REL before the completion of all the acts necessary for the downstroke. Thus, the hand movements to press POS REL separate into two chunks: the horizontal movement to position the finger above POS REL, and the press and release of the key. Other situations do not allow anticipation (e.g., keying in a customer's credit card number), so the movement chunk includes the horizontal movement, the down-stroke and the up-stroke. In general, to include cognitive operators that initiate hand movement, place one cognitive operator before each horizontal movement motor operator. If pressing the key depends on additional information being presented to the TAO, then place another cognitive operator before the down-stroke motor operator. The up-stroke motor operator immediately follows the down-stroke motor operator with no intervening cognitive operator.

Cognitive operators initiating movement are dependent on the completion of all previous cognitive operators and previously initiated motor operators using the same hand. However, motor operators for the right and left hands can work in parallel.

*Voice Output (Figure 1c).* The phrases used by TAOs are taught during training. It is assumed that these phrases are well learned and require no problem-solving or unique speech generation.

A cognitive operator is placed before each voice output event to initiate the speech. A motor operator performs each voice output event. This motor operator contains the entire phrase used by the TAO.

### Estimating Operator Durations

After all the operators assumed necessary to complete the task are placed in a schedule chart, and all dependency lines are drawn, the time course of events can be predicted if durations can be assigned to each operator. Estimates of duration come from several sources: previous GOMS research, independent research done in other domains, the videotapes of TAO call handling, and aggregate data about actual calls.

*Estimates From Previous GOMS Research.* Previous research involving the Model Human Processor and GOMS modeling provides estimates of the duration of several operators used in TAO tasks. Although different researchers have come up with slightly different estimates of some of these operators [7] they are in the same ballpark. Where possible, we have chosen to use estimates based on experiments specifically designed to produce such values.

Where that is not possible, we revert to the estimates given by Card, et. al. [2]. Table 1 presents the estimates of duration obtained through previous GOMS research, and the source of each estimate.

*Estimates From Independent Research In Speech Generation.* Previous GOMS research has not dealt with speech, either as input to the user or as output generated by the user. However, there is substantial literature of patterns and parameters of human speech that provide estimates for duration. When saying novel sentences, people speak at 170 msec/syllable; with highly-practiced sentences like phrases used by the TAO, this reaches asymptote at 130 msec/syllable [1]. The pause between speakers in a dialogue is about 700 msec [4]. GOMS assumes this pause to include the silence signalling the end of the speaker's turn, the verification of that silence by the listener, and the cognitive initiation of the listener's reply. These parameters can be used to provide estimates of speech duration (Table 2).

Table 1. Estimates of operator duration provided by previous GOMS research

<u>Operator type</u>	<u>Duration estimate</u>	<u>Source</u>
<b>Perceptual</b>		
Simple binary signal	100 msec	Card, et. al., 1983
Complex visual signal (word or code)	290 msec <sup>3</sup>	John & Newell, 1989
<b>Cognitive</b>		
	50 msec	John & Newell, 1989
<b>Motor</b>		
Homing between devices	350 msec <sup>4</sup>	Card, et. al., 1983
Horizontal movement within function keys	80 msec	derived from John, 1988
Horizontal movement within numeric keys	40 msec	derived from John, 1988
Down-stroke	60 msec	John, 1988
Up-stroke	60 msec	John, 1988
Eye movement	180 msec <sup>5</sup>	Card, et. al., 1983

<sup>3</sup> John & Newell record 340 msec for perception of a 6-letter word. Here, we separate this perception into two parts: a perceptual operator (290 msec) which perceives and encodes the information, and a cognitive operator (50 msec) that verifies the information.

<sup>4</sup> Card, et al. record 400 msec for homing between devices. Here, we separate this action into two parts: a cognitive operator (50 msec) that initiates the action, and a motor operator (350 msec).

<sup>5</sup> Card, Moran & Newell record 230 msec for eye movements. Here, we separate this action into two parts: a cognitive operator (50 msec) that initiates the action, and a motor operator (180 msec).

Table 2. Estimates of speech-related operations provided by speech research.

<u>Operation</u>	<u>Duration estimate</u>
Say a syllable in an unpracticed sentence (customer speech)	170 msec
Say a syllable in a highly practiced sentence (TAO speech)	130 msec
TAO's pause after a customer's utterance is the sum of:	
silence signalling the end of the customer's turn	600 msec
cognitive verification of silence	50 msec
cognitive initiation of TAO's response	50 msec
Customer pause before speaking (composite of silence, verification, & initiation)	700 msec

Sample duration estimates for TAO phrases (130 msec x no. of syllables):

May I have your name, please?	780 msec
Go ahead, please.	520 msec
They do not answer. Will you try your call later, please?	1690 msec

<u>Comparison between predicted and observed durations</u>	<u>Predicted (msec)</u>	<u>Observed (msec)</u>	<u>Percent Difference</u>
New England Telephone. May I help you?	1300	970	-34%
Make it collect from Ann.	1020	1620	37%
Thank you.	260	200	-30%
I have a collect call to anyone from Ann. Will you pay for the call?	2340	2500	6%
Hello	340	460	26%
Ah. Just a second.	850	600	-42%
Babe, do you know an Ann?	1020	1010	-1%
Yeah.	170	400	58%
We'll accept it. Thank you.	1020	940	-9%
Thank you.	260	600	<u>57%</u>

average % error = -7%

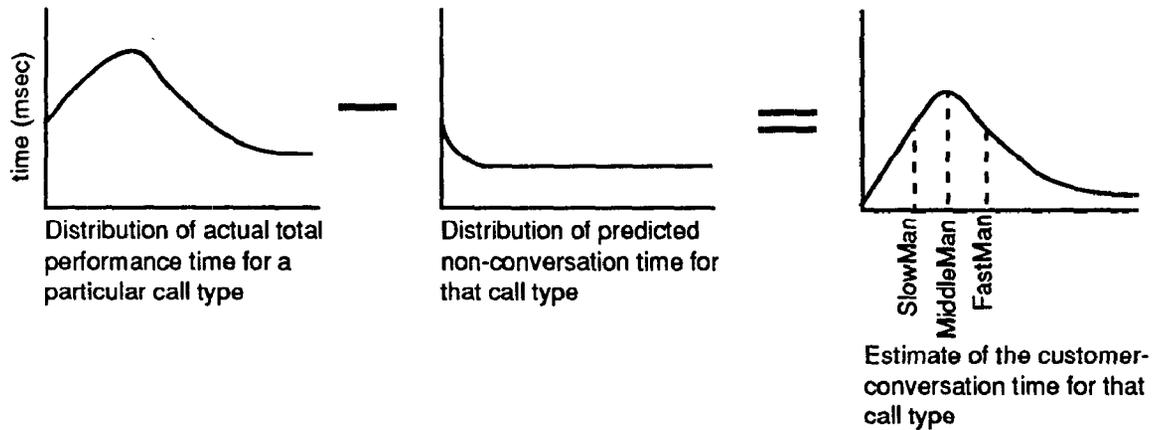


Figure 2. Process to make estimates of customer conversation time

*Estimates From Videotapes Of Sample Calls.* We have recorded TAOs using existing workstations and systems to handle twenty different types of calls placed by an experimental confederate employed by the telephone company. These videotapes can verify estimates gotten by other means. For example, the speech duration rules-of-thumb stated above, overpredict the duration of the specific phrases spoken in the example call by an average of 7% (Table 2). These videotapes also supply estimates of times unavailable from other sources.

Consider system RT. The example call showed actual SDT to be about 50 msec, CHS RT for display of relevant information to be between 50 and 500 msec with a median value of 300 msec, CHS RT to initiate ringing to be over 3000 msec, and BVS RT to be 770 msec. These estimates will be continually updated as more videotape segments are analyzed.

*Estimates From Aggregate Data Of Actual Calls.* Some elements of actual calls cannot be estimated on the basis of abstract task analysis or recording of controlled experiments. Most obvious is the duration of customer conversation. However, the range and distribution of customer conversation can be estimated by analysis of the distribution of time for actual calls (automatically collected for other purposes). If all other durations have been determined by means described above, they can be subtracted from the distributions of actual calls to leave an estimate of the distribution of customer conversation time (Figure 2). This distribution can be used to make estimates akin to the FastMan, MiddleMan and SlowMan parameters used by Card, et. al. [2] and explorations can be made to reveal the implications of different distributions.

More than one unknown duration (e.g. calling party conversation time, called party conversation time, and ringing time) can be estimated from these distributions by making simple assumptions of additivity and minimum and maximum ranges. For example, the duration of a single ring is a well known quantity in telephony, which serves as a lower bound on the duration of ring time. An upper bound on ring time is the duration of six rings, the maximum

number allowed in TAO training. Thus the range of ringing time can be estimated independently and a simple normal distribution might be assumed. This could then be subtracted from the total distribution to leave customer conversation time.<sup>6</sup>

### Resulting GOMS Analysis

Figure 1 gives the schedule chart resulting from the task analysis for the collect call situation, heuristics and duration estimates described above.

This schedule chart serves two purposes. The first is to act as a check on the task analysis and duration estimates implicit in the chart. When actual start and end times observed on the videotape are enforced in this schedule chart, some times may be impossible to meet. This usually indicates that there is something wrong with the task analysis.

In this example, one impossible date emerges. For this call handled by this TAO, the TAO crosses out the name "Ann" before the called party gives any indication that he will accept the call. Thus, the assumption in the task analysis that the TAO requires confirmation of acceptance before crossing out the name is incorrect. This may be due to a general misconception about the use of the written name. The original assumption was that the name was written for future reference in case the call was rejected and the calling party wished to place another collect call, but it may be the case that the name is written only to allow the TAO to unerringly remember it for the announcement of the first call. Another explanation is that the original task assumption is correct in general, but this TAO is idiosyncratic in her use of the written name. Either explanation can be checked with subsequent analysis of videotapes of this and other TAOs handling similar calls.

<sup>6</sup> This estimate procedure has not yet been followed for any call types. Therefore, for the GOMS analysis of the collect call detailed in this paper, the durations of customer conversation time and ringing time observed on the videotape are used.

The second use of this schedule chart is to provide a theoretical explanation for the time course of events. If all durations are estimated and no start or end times are enforced beyond the initial starting time, then the schedule chart gives an estimate of the total performance time for a call, and for the start times of intermediate events. This procedure yields an estimate of 21350 msec for the total performance time, 10% under the observed performance time of 24040 msec, and an average absolute percent error of 14% for the anchor times of 15 observable intermediate events (Table 3).

### WORK IN PROGRESS

This paper represents a preliminary stage in an ongoing project to demonstrate GOMS modeling to be a viable technique for system evaluation in the real world. To that end, work continues in several directions.

Other call types are being analyzed to test the heuristics detailed above and to discover new heuristics.

The heuristics are being taught to several researchers familiar with TAO operations and independently developed GOMS models are being compared to evaluate inter-analyst reliability. This will measure the amount of training necessary before acceptable reliability is obtained.

Table 3. Comparison of observed events to GOMS predictions.

Event	Observ (msec)	Pred (msec)	% Diff
"New England Telephone..." starts	1330	920	31%
R-home to keyboard starts	1800	970	46%
"Make it collect..." starts	2550	2920	-15%
R-horizontal to STA COL starts	3900	2320	41%
-STA COL down ends	4540	4110	9%
R-home to paper starts	4600	4170	9%
"Thank you" starts	4940	4640	6%
Write "Ann" starts	5500	4690	15%
"Hello" starts	13980	13520	3%
"I have a collect call..." starts	14880	14560	2%
"Ah. Just a second..." starts	17950	17600	2%
R-crossing out name starts	21280	20140	5%
L-home to POS REL	23180	20810	10%
"Thank you" starts	23080	21360	7%
L-POS REL down ends	24040	21740	10%
average absolute percent error = 14%			

"What-if" analyses are being made. That is, the GOMS models are manipulated to show the impact of different changes to the task. For instance, if the CHS RT was decreased by 20%, the impact on total performance time for different call types can be predicted. Such analyses can lead to recommendations about where to most profitably invest time and money: changing the systems or the workstation hardware, or perhaps mounting a massive advertising campaign to get customers to be more succinct in stating their desires.

Lastly, zero-parameter predictions are being made about what would change in total performance time if a new workstation was used. These predictions will be checked against call performance data collected in a field study comparing the existing and proposed workstations. In addition to the accuracy of the predictions, the relative cost of doing the GOMS modeling versus the cost of conducting the field study are being recorded to allow cost/benefit analyses.

It is our hope that the promise of GOMS modeling that has been demonstrated in several academic laboratories will be fulfilled in this transfer to industrial use.

### ACKNOWLEDGEMENTS

This research represents a collaborative effort between the author and the Intelligent Interfaces Group at the NYNEX Science and Technology Center. Support from that organization, specifically Wayne Gray and Mike Atwood, and from New England Telephone, specifically Karen O'Brien, is gratefully acknowledged. Special thanks to Allen Newell and Sandy Esch for their comments on the content of this paper.

### REFERENCES

- Boff, K. R., & Lincoln, J. E. *Engineering data compendium: Human perception and performance*. Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, 1988.
- Card, S. K., Moran, T. P., & Newell, A. *The psychology of human-computer interaction*. Lawrence Erlbaum, Associates, Hillsdale, NJ, 1983.
- Gray, W. D., John, B. E., Lawrence, D., Stuart, R. & Atwood, M. E. GOMS meets the phone company, or can 8,400,000 unit-tasks be wrong? Poster presented at CHI, 1989 (Austin, Texas, April 30-May 4, 1989).
- Jaffe, J., & Feldstein, S. *Rhythms of dialogue*. Academic Press, New York, 1970.
- John, B. E. *Contributions to engineering models of human-computer interaction*. Doctoral dissertation, Carnegie Mellon University, 1988.
- John, B. E. & Newell, A. Cumulating the science of HCI: From S-R compatibility to transcription typing. In *Proceedings of CHI* (Austin, Texas, April 30-May 4) ACM, New York, 1989, pp.109-114.
- Olson, J. R., and Olson, G. M. The growth of cognitive modeling in human computer interaction since GOMS. To appear in *Human Computer Interaction*.