MANAGING ATTENTION BY PREPARING TO FORGET

Erik M. Altmann (altmann@gmu.edu)
Wayne D. Gray (gray@gmu.edu)
Human Factors & Applied Cognition
George Mason University

In dynamic task environments, human operators must update their memory for what is true of the current situation. This updating depends on forgetting old information, and this forgetting in turn places constraints on how an item is encoded in the first place — the cognitive system must prepare to forget. Functional decay theory accurately predicts how long this preparation takes — concentrated use of an item for about 5 sec requires an additional 1 sec for initial encoding. This quantitative prediction illustrates the potential of functional decay theory for evaluating cognitive workload.

INTRODUCTION

A common problem in many task environments is that they change — what is true of the world now may not be true in a moment. A key challenge facing the human operator in such environments is that of forgetting old states of the world in order to prevent them from interfering with memory for the current state of the world.

An everyday example is the "Where did I park my car?" problem, in which one has trouble remembering, at the end of the day, where one parked that morning. This phenomenon is generally attributed to proactive interference, in which old memory traces from past parking episodes interfere with memory for today's location. On this view, it would be helpful if all previous parking episodes would decay or fade away and leave a memory trace for today's location to stand alone. The memory-updating problem has been studied empirically in dynamic task environments as well. For example, in the keeping track paradigm (e.g., Venturino, 1997; Yntema, 1963) multiple environmental attributes change periodically and the operator are required to remember the current value of each.

We have examined the memory-updating problem using a serial attention paradigm, which simplifies the keeping-track paradigm to a single attribute with two values. In the serial attention paradigm, the operator serially attends to one value at a time, where the value is a task to perform. The task environment updates the current task periodically, but gives no clue to the current task between updates, requiring the operator to maintain the current task in memory.

A novel finding from the serial attention paradigm is that operators let themselves forget the current task, gradually, from one update to the next (Altmann & Gray, 1999b). The behavioral evidence for this forgetting is a small but steady performance decline over trials between updates, measured both in terms of response time (RT) and error. Our theoretical explanation for this effect is functional decay theory, which proposes that forgetting the current task is functional because it reduces interference — if the current task has already decayed by the time the next task comes along, it will cause less interference once the next task is current.

In this paper, we examine a key prediction of functional decay theory concerning how cognition prepares to forget the current task. The prediction is that in order to forget an item gradually, cognition must first encode it in memory with sufficient strength that it will not get any stronger through subsequent use — that the only direction its activation can take is down. The theory also predicts that this initial strength in memory is the product of a controlled and effortful process — that one must deliberately "pay attention" to an item to make it active in memory. With a formal model, we are able to make quantitative predictions about the duration of this deliberate encoding process. The implication is that in highly dynamic environments, a substantial fraction of an operator's time must be spent committing items to memory when they are not available from the environment.

THE SERIAL ATTENTION PARADIGM

Our serial attention paradigm (adapted from Gopher, Greenshpan, & Armony, 1996) involves giving participants two simple tasks and periodically issuing an instruction to switch from one task to the other. For example, for a stimulus like "222222", the correct response is either ">" or "<" depending on whether the task is to count the characters (">" means more than 5) or to rate the value of the repeated character ("<" means less than 5). Stimuli like those above are presented one after the other in runs of 7 to 13 (run length is chosen randomly). Each such stimulus is called a **choice trial,** because it requires a forced-choice response. A run of choice trials is preceded by an **instruction trial** giving the new task (e.g., "groupsize", meaning that the task is to count characters) Responses to all trials are self-paced and there is no calibrated inter-trial interval. All stimuli appear at the same location in the center of the screen. A typical session involves 3840 choice trials and 384 instruction trials, creating a large potential for interference in memory for the current task.

Data from this paradigm appear in Figure 1. The abscissa shows trials P0 through P7, where "P" means Position within the run. P0 is the instruction trial, P1 is the first choice trial of the run, and so on. The ordinate shows average RT for each position. The novel within-performance decline described

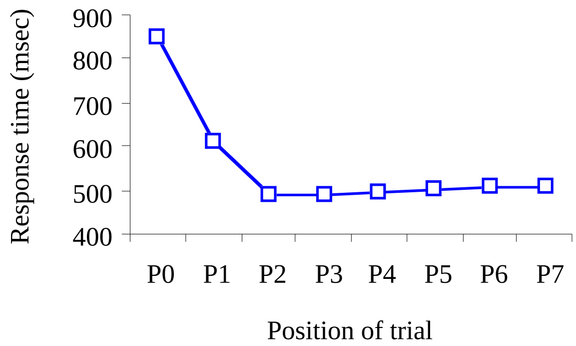

Figure 1: Serial attention response times. P0 is the instruction trial, and P1 to P7 are choice trials.

earlier is evident in the small but significant upward slope from P2 to P7, a trend also reflected in error rates. Both trends have been replicated repeatedly (Altmann & Gray, 1999b). For present purposes, the most important aspect of the data, which we address below, is that P0 and P1 take substantially longer than P2 to P7.

FUNCTIONAL DECAY THEORY

The premise of functional decay theory (Altmann & Gray, 1999a; Altmann & Gray, 1999b) is that the most recent instruction must be the most active in memory, if it is to be reliably retrieved on a given choice trial. This logic is illustrated in Figure 2, which shows the time course of activation of an instruction trace. Activation, as we use the term here, refers to the availability of an item in memory — the more active the item, the more quickly and accurately it is retrieved from memory by the cognitive system. In a speeded task, we assume that cognition takes the first element it can get, so the most active element effectively "masks" all less active elements.

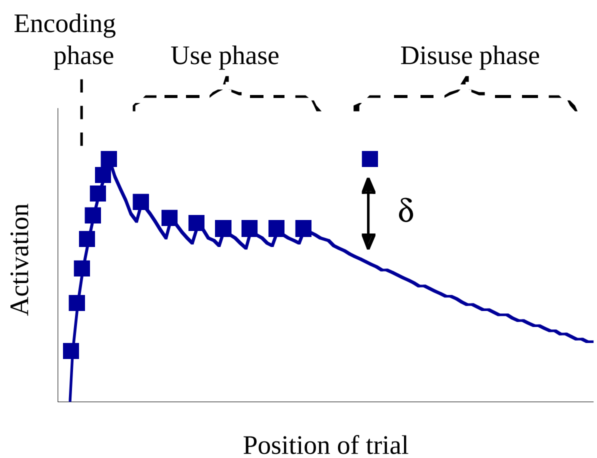

Figure 2: Activation of an instruction trace over time. Activation increases to a peak during the encoding phase, decays gradually from this peak during the use phase, then decays more rapidly when a new instruction supersedes it. Decay through use (by amount δ) ensures that the newest instruction is always the most active.

Our formal model for computing activation, and the function plotted in Figure 2, is $\ln(2n/\sqrt{T})$, where n is the total number of retrievals of the item and T is the length of the item's life. This equation is a core mathematical component of the ACT-R cognitive theory (Anderson & Lebiere, 1998).

Figure 2 shows three phases to the time course of an instruction's activation. During the encoding phase, cognition subjects the instruction to a period of massed use, causing n to increase rapidly in a short amount of time. During the use phase, the instruction is no longer visible and must be retrieved from memory on each choice trial to tell cognition what choice to make. Finally, during the disuse phase the (now) old instruction is no longer retrieved because a new one masks it.

The "functionality" of functional decay theory lies in the activation difference δ, the activation difference between an instruction and its predecessor. Assuming that each instruction is encoded to the same initial level of activation, decay from that point on (i.e., decreasing activation) means that δ will always be positive and that the newest instruction will always be the most active. Thus, the theory says that encoding and decay are critical to maintaining awareness of the current task when the task changes continually.

The critical initial test of the theory is that it explains the performance decline from P2 to P7 in Figure 1. As the current instruction decays, it becomes more difficult to retrieve from memory. This difficulty is realized both in terms of RT, as shown in Figure 1, and in terms of error (Altmann & Gray, 1999b). This performance decline is a novel effect in the memory and attention literature, and the explanation offered by functional decay theory is unique — to our knowledge no other existing theories of mental attention speak to this phenomenon.

PREDICTIONS AND DATA

Our focus here is on the encoding-time predictions of functional decay theory. The question we would like to answer concerns how long the encoding phase must last to fit the constraints of the model. That is, some amount of encoding — some concentrated "paying attention" — is necessary to make the instruction "stick" in memory. How much encoding is enough?

Some simple assumptions generate remarkably accurate predictions about duration of encoding time. The assumptions concern the amount of decay needed during use. Implicit in the base-level activation equation is that decay through use is limited. This is evident in Figure 2, in that the slope of the activation curve levels off (measured across the squares, which represent retrievals) as the use phase nears its end. Were the use phase to continue longer than shown in Figure 2, all else being equal, the activation curve would begin to head upwards again. Instead, however, the disuse phase begins, and the instruction begins to decay more quickly because it is no longer being retrieved.

The critical assumption is that the rate of decay from one choice trial to the next reaches zero at about the end of the use phase as shown in Figure 2. The rationale for this **optimal minimum** assumption is cognitive economy. The underlying tradeoff is that the amount of decay during use depends on the amount of time invested in encoding. The more time invested in encoding, the higher the initial level of activation, and the longer the instruction trace decays (in terms of successive choice trials) before reaching an activation minimum. Once the minimum is reached, activation will begin to increase, and will continue to increase until the instruction falls into disuse. If decay levels off at the end of the use phase, this optimizes δ per unit of encoding. The optimal minimum assumption lets us predict the duration of the encoding process.

The duration of the encoding phase is equal to the number of **encoding cycles** multiplied by time per cycle. An encoding cycle is functionally equivalent to a memory retrieval in terms of its effect on activation — both increment *n* in the base-level activation equation. The difference is that encoding cycles occur more rapidly because there is no other process that requires attention. That is, on the instruction trial, the cognitive system can devote all its cycles to the encoding process, whereas on choice trials other processes must be performed as well — retrieval of the current instruction and selection of the correct response, for example. A quantitative bound of 100 msec per encoding cycle is implied by ACT-R theory (Altmann & Gray, 1999c). With this bound in hand, we need only estimate the number of encoding cycles in order to obtain a quantitative prediction for the duration of the encoding process.

The base-level activation equation provides the basis for estimating the number of encoding cycles. The original equation, $\ln(2n/\sqrt{T})$, can be simplified by denominating time in terms of number of retrievals such that one retrieval occurs per unit time. Thus, we can substitute n for T. We can now break n into its contributing terms, $n = E + R$, where E is the number of encoding cycles and R is the number of retrievals on choice trials. We assume one retrieval per choice trial, so $R = 10$, the average run length. This leaves E to be estimated.

To estimate E we apply the optimal minimum assumption; namely, the slope of the activation curve is flat at the end of the use phase. We would like to express this slope solely in terms of E, the number of encoding cycles. In terms of E and R, the base activation equation is $\ln(2(R+E)/\sqrt{R})$. Differentiating with respect to R and setting the resulting function to zero produces $E = R$. That is, for activation to reach a minimum at the Rth retrieval (i.e., at the Rth choice trial), the encoding process must be R cycles long. This is a simplifying assumption, because the marginal effect on δ tends to zero as E tends to infinity (Altmann & Gray, 1999c), but $E = R$ is a reasonable approximation for small E.

The predicted duration of the encoding process is thus roughly 1 sec, based on the estimate of 10 encoding cycles combined with the bound of 100 msec per encoding cycle. It still remains, however, to map the encoding process to empirical measures of RT. The obvious measure is duration of the instruction trial itself — the time during which the instruction text is perceptually available as input to the encoding process. However, there is reason to suspect that encoding "spills over" into P1, the first choice trial after the instruction trial. Since the Sperling result, theories of visual attention have generally predicted some persistence of a veridical trace of a stimulus after offset of the physical stimulus itself (e.g., Coltheart, 1999). Thus on P1, at least initially, a veridical trace of the instruction continues to be available as input to the encoding process. An efficient adaptation for the cognitive system is to apply the encoding process to this input as long as the input persists, because this improves performance accuracy without costing the system in terms of overall RT (Altmann & Gray, 1999a).

We would therefore expect the predicted 1 sec of encoding time to be distributed over P0 and P1. Figure 1 shows RT data for these two trials. RT on P0 is 850 msec, representing a substantial fraction of the predicted 1 sec. To estimate how much encoding spills over from P0 to P1, we take P2 RT as an estimate of pure choice RT and subtract this time from P1 RT. This calculation yields an estimate of 124 msec available for encoding on P1 over and above the minimum time needed for ordinary processing on a choice trial. The total empirical estimate of encoding time is 974 msec, which is strikingly close to the 1 sec predicted by theory. This prediction has borne out in multiple data sets to be reported in a future paper.

Moreover, the 974 msec estimate is strikingly distant from the prediction of the obvious alternative model. In the absence of an encoding process, the cognitive system would only need to identify the current stimulus as an instruction to make the correct response (pressing the space bar). As a point of comparison, choice trials involve several stages in addition to stimulus identification, including memory retrieval (of the current instruction) and response selection (of ">" or "<"). Thus, instruction trials (on this encoding-free model) should be faster than choice trials. Instead, they are twice as slow, suggesting that some other, effortful process is at work.

DISCUSSION

Our question concerned how operators maintain awareness of the current state of the world when the world changes often. We proposed that forgetting helps by preventing stale information from interfering with fresh information. However, forgetting is difficult when the world changes faster than memory can decay. Cognition responds to this limitation by starting to forget an item as soon as it is encoded. This eager forgetting requires the system to make the item so active initially that it has no place to go but down. In short, cognition must prepare itself to forget.

A qualitative prediction of our analysis is that human operators in dynamic task environments need time for encoding and decay — without time to pay adequate attention to an update, and without enough time to let an item fade, situation awareness could degrade catastrophically.

A rough quantitative prediction is that if an attribute is updated every five seconds (a choice trial takes roughly 500 msec, and there are 10 per run on average), the cognitive system must have a second to encode each new value. These quantities must be heavily qualified, of course. To obtain precise predictions one would need to define "concentrated use" operationally, for example, and would need to account for the effects of any environmental cues indicating the task at time of need. Clearly, a more precise mapping of our analysis to a complex task environment is an important goal for future research. However, even if these quantities are off by a factor of two, they suggest that "paying attention" can be expensive.

In conclusion, functional decay theory formally represents key constraints on memory and attention in the human cognitive architecture. When applied to dynamic task environments, the theory makes quantitative predictions for human performance, in particular that encoding and decay are critical processes in the maintenance of situation awareness.

ACKNOWLEDGEMENTS

This work was funded by AFOSR grant F49620-97-1-0353.

REFERENCES

Altmann, E. M., & Gray, W. D. (1999a). The anatomy of serial attention: An integrated model of set shifting and maintenance. *Manuscript submitted for publication.*

Altmann, E. M., & Gray, W. D. (1999b). Functional decay in serial attention. *Manuscript submitted for publication.*

Altmann, E. M., & Gray, W. D. (1999c). Serial attention as strategic memory, *Proceedings of the twenty first annual meeting of the Cognitive Science Society* (pp. 25-30). Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Lebiere, C. (Eds.). (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.

Coltheart, V. (Ed.). (1999). *Fleeting memories*. Cambridge, MA: MIT Press.

Gopher, D., Greenshpan, Y., & Armony, L. (1996). Switching attention between tasks: Exploration of the components of executive control and their development with training, *Proceedings of the HFES 40th annual meeting* (pp. 1060-1064). Philadelphia, PA: Human Factors and Ergonomics Society.

Venturino, M. (1997). Interference and information organization in keeping track of continually changing information. *Human Factors, 39*(4), 532-539.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7-17.