Why High-Level Attention Constantly Oscillates: System-Based Explanation

Griselda Acosta[1], Eric Smith[2], and Vladik Kreinovich[3]
[1]Department of Electrical and Computer Engineering
[2]Department of Industrial, Manufacturing, and Systems Engineering
[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
gvacosta@miners.utep.edu, esmith2@utep.edu, vladik@utep.edu

Abstract

In many situations like driving, it is important that a person concentrates all his/her attention at a certain critical task – e.g., watching the road for possible problems. Because of this need to maintain high level of attention, it was assumed, until recently, that in such situations, the person maintains a constantly high level of attention (of course, until he or she gets tired). Interestingly, recent experiments showed that in reality, from the very beginning, attention level oscillates. In this paper, we show that such an oscillation is indeed helpful – and thus, it is necessary to emulate such an oscillation when designing automatic systems, e.g., for driving.

Keywords: parallel computing, attention oscillation, systems approach

1 Formulation of the Problem

In many real-life situations, high level of attention is crucial. In many practical situations, we concentrate on a certain task. For example, when a person drives a car, he/she needs to keep a close attention to the road, to make sure that if a problem appears, the driver will react as soon as possible – and thus, avoid a possible accident.

What researchers assumed. In critical situations, when the maximum attention is needed, psychologists assumed that the attention is consistently kept at the maximum possible level – of course, until the person becomes too tired to maintain this level of attention.

This assumption makes perfect sense: when a lot is at stake, including the person's own life, it makes sense to concentrate all the energy on avoiding possible catastrophic situations.

A recent surprising observation. Surprisingly, recent experiments showed that while the attention indeed remains high, the attention level – as measured, e.g., by the reaction time – constantly oscillates; see, e.g., [1, 2]. This level remains high, but the reaction time still oscillates between the smallest possible value and a much larger value. This larger value of reaction time is still good, but not as perfect as the smallest value.

The problem. It is not clear what is the reason for this observed phenomenon. Are they somehow needed for survival? Or are they due to an imperfection of human physiology?

This is not just an interesting theoretical question, it has practical applications:

- If the oscillations indeed improve the system's performance, then we should add similar oscillations to the self-driving cars and other automated vehicles and systems.
- On the other hand, if the oscillations are caused by imperfections of human physiology, then we should not emulate human drives in this; we should instead keep the computer's attention level constant.

What we do in this paper. In this paper, we show that oscillations do make the system more efficient – and thus, appropriate oscillations should be implemented in automatic control systems.

2 Analysis of the Problem

Need for a numerical model. To analyze the problem, to see whether constant attention of oscillating attention are more productive, we need to formulate this problem in precise numerical terms. Let us therefore describe a simple simplified model of this phenomenon.

Towards a simplified model. Let T denote the duration of the period during which we need to maintain high attention level. Without losing generality, we can start counting time from the beginning of this period. In this case, the corresponding time interval takes the form $[0, T]$.

There are natural limitations on how many observations we can process, whether in a computer or in our brains. For a high-performance computer, these limitations are higher than for a simple laptop, but they are still there. These limitations are real: e.g., when a conference speaker makes a presentation remotely (e.g., by skype), the system often does not catch up when the speaker's movements are too fast.

Let us assume that, because of these limitations, during a certain period of time T, we can process at most N observations. In crucial situations requiring high attention, it is important that the person concentrates on the corresponding task as much as possible – and thus, that this person processes as much information as possible. This means that in such situations, we should process the maximum possible number of observations: namely, we should process exactly N observations during the time T.

These observations correspond to, in general, different moments of time. Let us sort these moments of time in chronological order. For each i from 1 to N, let us denote the time of the i-th observation by t_i. Then, we have

$$0 \le t_1 \le t_2 \le \ldots \le t_N \le T.$$

We want to detect possible obstacles as early as possible, at the time when the corresponding signals are still weak. For weak signals, a single observation is not sufficient for reliable detection, since there is always some noise level: we are not sure that the observed signal is real or just a noise. Swerving every time when a speck appears which may be a car or a pedestrian is also a sure recipe for disaster: this means that a car would follow an unpredictable waving trajectory, like when the driver in drunk. We need to perform correcting actions only when we are reasonably sure that there is indeed a problem on the road.

The more observations confirm that there is a problem, the higher our level of confidence that this problem is real. Let m denote the smallest number of observations that make us confident. Then, if a problem appears at time $t \in [0, T]$, we will detect it when m observations pass after this time t. Let $i(t)$ denote the first index i for which $t_i \ge t$. The problem can then we observed in observations made at times $t_{i(t)}$, $t_{i(t)+1}$, $t_{i(t)+2}$, etc. The problem will be detected after m such observations, i.e., at the moment $t_{i(t)+m-1}$. The difference $\Delta(t) \stackrel{\text{def}}{=} t_{i(t)+m-1} - t$ between the time when we detect the problem and the original time t is the main component of the reaction time.

For problems appearing at the end of the time period $[0, T]$, namely for problems corresponding to times $t > t_{N-m}$, there are not enough remaining observations to observe this problem.

Definition 1.

- *By an* high-attention situation*, we mean a tuple (T, N, m), where $T > 0$ is a real number, and m and N are integers for which $m < N$.*

- *For each high-attention situation, by a* strategy*, we mean an increasing sequence of real numbers $t_1, \ldots, t_N$ for which $0 \le t_1 \le t_2 \ldots \le t_N$.*

- *For a given strategy and for each moment $t \in [0, T]$, by the* reaction time *$\Delta(t)$, we mean the difference $t_{i(t)+m-1} - t$.*

Comment. As we have mentioned earlier, the reaction time is defined only for moments $t \le t_{N-m}$.

Which strategy should we prefer? We want to minimize reaction time. First of all, we want to make sure that no matter when the problem appears, we should be able to deal with it within a reasonable time r – and this time should be as small as possible. This means that for all the moments $t \le T - r$, we should have $\Delta(t) \le r$. This guaranteed reaction-time r should be as small as possible.

There may be several different strategies with the same worst-case reaction time. To select between then, it is reasonable to choose the strategy with the smallest possible *average reaction time*: the average value of $\Delta(t)$ over all the moments $t \in T - r$. Thus, we arrive at the following definition.

Definition 2.

- *For each strategy t_i, by its* worst-case reaction time $r_w(t_i)$, *we mean the smallest positive real number r for which* $\max\limits_{0 \le t \le T-r} \Delta(t) \le r$.

- *For a strategy t_i with worst-case reaction time r, by its* average reaction time $r_a(t_i)$, *we mean the value*

$$r_a(t_i) \stackrel{\text{def}}{=} \frac{1}{T-r} \cdot \int_0^{T-r} \Delta(t)\, dt.$$

3 Oscillations Are Better: Proofs

Discussion. Let us use the above model to check which strategy is better: the strategy is constant or the strategy in which attention is oscillating. Let us describe these strategies in precise terms.

Constant level of attention: how to formalize. Constant level of attention means that we have the exact same difference $\delta = t_{i+1} - t_i$ between the two consecutive observations, i.e., that $t_2 = t_1 + \delta$, $t_3 = t_2 + \delta = t_1 + 2\delta$, etc., all the way to $t_N = T$.

In this case, the worst-case reaction time is $r = m \cdot \delta$ that occurs if the problem appears right after each observation, at time $t = t_i + \varepsilon$ for some small positive $\varepsilon \ll \delta$. To maintain the same reaction time for $t = 0$, it is sufficient to take $t_1 = \delta$, thus, $t_i = i \cdot \delta$. So, $\delta = T/N$.

Since we ignore moments $t > T - r$, we can as well place all the moments t_i corresponding to these times at $T - r$.

Definition 3. *By a* uniform strategy, *we mean the strategy in which $t_i = i \cdot (T/N)$ for $i < N - m$ and $t_i = (N - m) \cdot (T/N)$ for $i \ge N - m$.*

Proposition 1. *For the uniform strategy, the worst-case reaction time is $r_w = m \cdot (T/N)$, and the average reaction time is*

$$r_a = \left(m - 0.5 - \frac{m \cdot (m-1)}{2(N-m)} \right) \cdot (T/N).$$

Proof. For the worst-case reaction time, the result is straightforward.

For the average reaction time, the interval $[0, T - r]$ is divided into $N - m$ intervals $[t_{i-1}, t_i]$ of equal width δ. Thus, to compute the average reaction time over the whole interval $[0, T-r]$, it is sufficient to compute the average reaction time over each of these small intervals, and then compute the arithmetic average of these averages.

For the first $N-2m+1$ intervals $[t_{i-1}, t_i]$, the reaction time changes between the maximal value $m\cdot\delta$ (attained close to t_{i-1}) and the smallest value $(m-1)\cdot\delta$ (attained at t_i), so the average over this interval is $(m-0.5)\cdot\delta$.

For $i = N-m+2$, we have reaction time changing from $(m-1)\cdot\delta$ to $(m-2)\cdot\delta$, with an average $((m-1)-0.5)\cdot\delta$. For the next interval, we have $((m-2)-0.5)\cdot\delta$, etc., all the way to $0.5\cdot\delta$ for the last interval.

In general, the average over each interval has the form $(j-0.5)\cdot\delta$, where in $N-2m+1$ cases, we have $j=m$, and then we have $m-1$ values $j=m-1$, $j=m-2,\ldots,j=1$. So, the average reaction time is equal to $r_a = (E[j]-0.5)\cdot\delta$, where $E[j]$ is the average value of j. Here,

$$E[j] = \frac{m\cdot(N-2m+1)+(m-1)+\ldots+1}{N-m} =$$

$$\frac{m\cdot(N-2m+1)+\dfrac{(m-1)\cdot m}{2}}{N-m}.$$

Since $N-2m+1 = (N-m)-(m-1)$, we get

$$E[j] = \frac{m\cdot(N-m)-m\cdot(m-1)+\dfrac{(m-1)\cdot m}{2}}{N-m} =$$

$$\frac{m\cdot(N-m)-\dfrac{(m-1)\cdot m}{2}}{N-m} = m - \frac{(m-1)\cdot m}{2(N-m)}.$$

Substituting this expression for $E[j]$ into the formula $r_a = (E[j]-0.5)\cdot\delta$, we get the desired result.

The proposition is proven.

Oscillations: how to formalize. Let us consider the extreme case of oscillations, where instead of having observations at uniformly distributed times, we bring observations in groups of m: no observations, then m of them in a row, then again no observations, then N of them in a row, etc., until we reach the last m values, i.e., the values starting with $k\cdot m+1$, where $k=\lfloor N/m\rfloor$:

$$t_1 = t_2 = \ldots = t_m = r, \quad t_{m+1} = \ldots = t_{2m} = 2r, \ldots,$$

$$t_{k\cdot m+1} = \ldots = t_N = k\cdot r. \tag{1}$$

Definition 3. *By a* maximally oscillating strategy*, we mean the sequence (1), where $k=\lfloor N/m\rfloor$ and $r=T/k$.*

Proposition 2. *For the maximally oscillating strategy, the worst-case reaction time is $r_w = T/k$, and the average reaction time is*

$$r_a = T/(2k).$$

Discussion. For the case when N is divisible by m, we get $k = N/m$. In this case, the worst-case reaction time $r_w = T/k = m \cdot (T/N)$ is the same as for the uniform strategy. However, the average reaction time is almost twice smaller. Thus, the oscillations indeed make the strategy more efficient.

Proof of Proposition 2. For the worst-case reaction time, the proof is straightforward. On each interval of width r, the reaction time changes from 0 to r. For each value t from 0 to r, the reaction time is $r - t$. Thus the average reaction time is

$$\frac{1}{r} \cdot \int_0^r (r-t)\,dt = \frac{1}{r} \cdot \left(r \cdot t - \frac{t^2}{2} \right) \Big|_0^r = \frac{1}{r} \cdot \left(r^2 - \frac{r^2}{2} \right) = \frac{1}{r} \cdot \frac{r^2}{2} = \frac{r}{2}.$$

The proposition is proven.

Discussion. It is possible to show that not only the maximally oscillating strategy is better than the uniform strategy, it is actually the best possible.

Definition 4. *Let an high-attention situation (T, N, m) be given. We say that a strategy t_i is* optimal *of for every other strategy t'_i, we have:*

- *either $r_w(t_i) < r_w(t'_i)$,*
- *or $r_w(t_i) = t_w(t'_i)$ or $r_a(t_i) \le r_a(t'_i)$.*

Proposition 3. *For each high-attention situation, the maximally oscillating strategy is optimal.*

Proof. Let us assume that we have an optimal strategy, and that its worst-case reaction time is equal to $r = r_w(t_i)$. For the maximally oscillating strategy, we have $r_a(t_i) = 0.5 \cdot r$. Let us show that, vice versa, we cannot have $r_a < 0.5 \cdot r$, and that if $r_a = 0.5 \cdot r$, then the corresponding strategy is maximally oscillating. This will prove that the maximally oscillating strategy is indeed optimal.

Indeed, the fact that the worst-case reaction time is equal to r means there exists a moment t_0 for which $\Delta(t)$ is as close to r as possible. This, in turn, means that between the moments t_0 and $t_0 + r - \varepsilon$, there are m values t_i, namely the values $t_{i(t_0)}, t_{i(t_0)+1}, \ldots, t_{i(t_0)+m-1}$. If all these m values are equal to each other, then for each moment t between t_0 and the common value of t_i, we get $\Delta(t) = t_{i(t_0)} - t = t_0 + r - t$, and thus, the average value of $\Delta(t)$ over the corresponding interval is equal to $0.5 \cdot r$.

In general, the next m values t_i cannot be earlier that $t_0 + r$, thus we have $\Delta(t) \ge t_0 + r - t$. If for some t, we get strict equality, then the average reaction time over the corresponding interval is $> 0.5 \cdot r$. The only possibility to have this part of t_a equal to $0.5 \cdot r$ is when for all t, we have $\Delta(t) = t_0 + r - t$.

Let us show that in this case, we have at least m values t_i equal to $t_0 + r$. Indeed, let j be the last value for which $t_j < t_0 + r$. Then, any t between t_j and $t_0 + r$, the fact that we have $\Delta(t) = t_0 + r - t$ means that the next m values t_i must be $\le t_0 + r$. Since the only value t_i between t and $t_0 + r$ is the value $t_0 + r$, this means that we have at least m values equal to $t_0 + r$. Thus, for the optimal solution, we have a group of at least m equal values, then another group of at least m equal values, etc.

If we group t_i into groups of size $> m$, then we would be divide the interval $[0, T]$ into fewer pieces than in the case when each group has exactly m values t_i. So, in this case, the distance between two consecutive groups will be larger than in the case when we have the division into groups of m; thus, this arrangement cannot be optimal. Hence, in the optimal arrangement, we should have m indices in each group of equal consecutive values t_i. This is exactly the oscillating arrangement. The proposition is proven.

Acknowledgments

This work was partially supported by the US National Science Foundation via grant HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] R. F. Helfrich, I. C. Fiebelkom, S. Z. Szczepanski, J. J. Lin, J. Parvizi, R. T. Knight, and S. Kastner, "Neural mechanisms of sustained attention are rhythmic", *Neuron*, 2018, Vol. 99, pp. 854–865.

[2] P. Jaekl, "Pay attention", *The Scientist*, 2018, No. 12, pp. 15–17.