The Leaky, Competing Accumulator

Georgii Zhulikov

National Research University Higher School of Economics

July 12, 2019

Overview

Types of models

Two main types of stochastic information accumulation models are:

- Accumulator models
  - Possible to have many output options
  - Several cumulative processes
  - Less accurate
- Diffusion models
  - Only two choices for output
  - One diffusion process
  - More accurate

Classical diffusion model – the diffusion model without drift variance.
Diffusion-with-drift-variance (DDV) – the diffusion model with drift variance.

Experiment paradigms

Standard RT paradigm

"Answer as fast as possible and as accurately as possible"

Time-controlled paradigm

"Answer right after the response signal is presented"

Georgii Zhulikov

Asymptotic accuracy

Imagine a problem where the correct and incorrect stimuli are almost identical. In terms of a diffusion model this will mean that drift rate is very small.

What will the accuracy be if the model runs for a very long time? What will the accuracy be in an experiment if the participant has no time limit?

Asymptotic accuracy

What will the accuracy be if the model runs for a very long time?
For the *classical diffusion model* and some other classical models, the accuracy will be perfect given enough time.

What will the accuracy be in an experiment if the participant has no time limit?
The participant won't be able to distinguish between the two alternatives, no matter how much time is given.

Asymptotic accuracy

The *DDV model* solves this problem with drift variance.

Another option is to use leaky integrators. This way information decays over time, and if the rate of accumulation is slow, the decay will not allow information to be accumulated indefinitely.

Relative evidence

Diffusion models accumulate evidence of the two choices **in relation to each other**.

Accumulator models have several **independent** accumulators with an **absolute** criterion.

Relative evidence

The *diffusion models* are accurate and they can describe a range of aspects of RT data from two-choice experiments.
However, they don't generalize naturally to N alternatives, unlike *accumulator models*.

Lateral inhibition

Lateral inhibition is used to combine the best of these types of models.
As activation value of a unit increases, it inhibits activation values of all the other units.

- Added to the *accumulator model*, allowing N alternatives
- Includes relative evidence in the activation value
- Allows to use an absolute criterion for termination in the standard paradigm

2 Features

Recurrent excitation

- Gives the model a tendency to settle in an equilibrium
- Doesn't come from RT and time-accuracy studies
- Mainly included because of neurophysiological and other evidence

Nonlinearity

- Is required for many neural computations
- Mostly characterizes neural information processing

The Model

The Leaky, Competing Accumulator model (LCA) combines all the features described above. It has:

- Asymptotic accuracy
  - Via leakage. To provide accurate answers for long trials.
- Relative evidence
  - Via lateral inhibition. To combine relative evidence with N-alternative task.
- Recurrent excitation
  - As a simple recurrent term.
- Nonlinearity
  - Via the threshold-linear function, $f(x) = \max(x, 0)$.

Nonlinear Equation

The stochastic accumulator equation incorporating all the features of the LCA model is

$$dx_i = \left[\rho_i - \lambda x_i + \alpha f(x_i) - \beta \sum_{i' \neq i} f(x_{i'})\right] \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \tag{1}$$

Here, $\frac{dt}{\tau}$ is a scaled time interval,
ξ is the Gaussian noise term

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Nonlinear Equation

$$dx_i = \left[\rho_i - \lambda x_i + \alpha f(x_i) - \beta \sum_{i' \neq i} f(x_{i'})\right] \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \tag{2}$$

The terms represent the features of the model:

- ρ_i is the external input of the i-th accumulator unit,
- λx_i is the decay, reflecting leakage,
- $\alpha f(x_i)$ is the recurrent excitation,
- $\beta \sum_{i' \neq i} f(x_{i'})$ is the lateral inhibition,

Linear Equation

The analysis of the nonlinear equation is complex. However, within some conditions its behavior is linear. If $x_i > 0$ for all units, the equation turns into

$$dx_i = [\rho_i - (\lambda - \alpha)x_i - \beta \sum_{i' \neq i} x_{i'}]\frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \tag{3}$$

Now $\lambda - \alpha$ is just a single term, so we can reduce the number of degrees of freedom by introducing a parameter $k = \lambda - \alpha$.

However, the assumption $x_i > 0$ does not hold all the time, so if we use this equation, the process may evolve into a state $x_i < 0$. This brings problems such as positive lateral inhibition term, which is absolutely wrong.

Linear Equation

It is possible to achieve compromise between these two equations by introducing truncation:

$$dx_i = [\rho_i - kx_i - \beta \sum_{i' \neq i} x_{i'}] \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \tag{4}$$

$$x_i \rightarrow \max(x_i, 0)$$

This equation is different from the equation (1), but it closely approximates it while being easier to work with.

Two alternative choices

For two alternatives there are two accumulators described by equations:

$$dx_1 = [\rho_1 - kx_1 - \beta x_2]\frac{dt}{\tau} + \xi_1\sqrt{\frac{dt}{\tau}}$$

$$dx_2 = [\rho_2 - kx_2 - \beta x_1]\frac{dt}{\tau} + \xi_2\sqrt{\frac{dt}{\tau}}$$

Defining a variable $x = x_1 - x_2$, using the fact that $\rho_1 + \rho_2 = 1$ and subtracting the second one from the first one we get the equation:

$$dx = [(2\rho_1 - 1) - (k - \beta)x]\frac{dt}{\tau} + \xi\sqrt{\frac{dt}{\tau}}$$

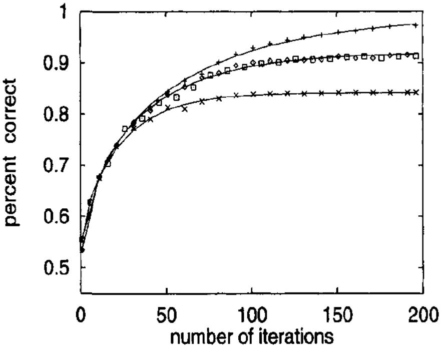
This equation describes the *Ornstein-Uhlenbeck* process.
The term $K = k - \beta$ is differential effective leakage or decay.

Time-accuracy curve

The top line: no decay, $K = 0$ same as the *classical diffusion process*.

The middle line consists of two lines: $K = 0.2$ and $K = -0.2$.

The bottom line: $K = -0.4$



Negative K means the process is decaying over time.
Positive K means the excitation is higher than the decay.
How do these opposite cases produce the exact same time-accuracy curve?

Distributions

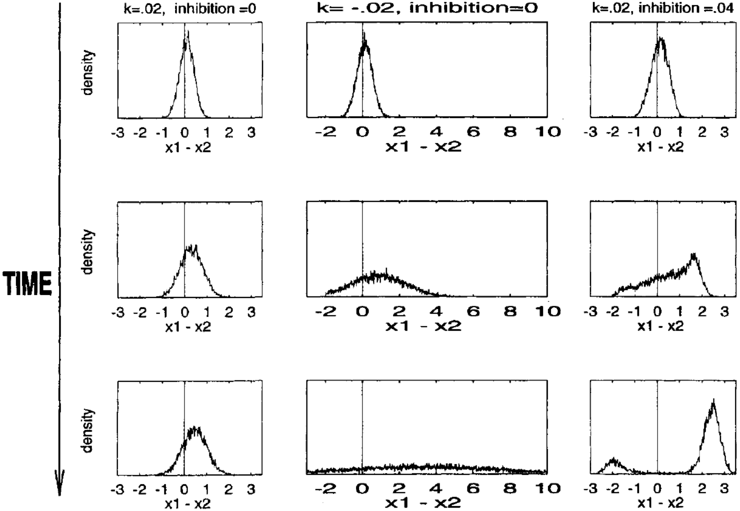
$|K| = 0.2$

Left: decay, no inhibition

Center: self-excitation, no inhibition

Right: decay, inhibition

Accuracy is the same



Results

Table 1

Parameters of Fit to Individual Participant Data From Experiment 1: Nine Parameters per Participant

Level of difficulty	$\frac{P(\text{data}\mid\text{OU})}{P(\text{data}\mid\text{DDV})}$	OU d'_{asy}	OU $1/K$	OU T_O	DDV d'_{asy}	DDV σ^2/σ_d^2	DDV T_O
Participant 1							
High	2.20	3.15	96	286	3.43	126	286
Medium	2.51	2.32	139	294	2.54	191	297
Low	1.54	1.11	393	337	1.47	1,175	339
Participant 2							
High	4.72	4.23	190	307	5.19	451	309
Medium	1.31	2.99	283	307	3.34	363	351
Low	1.29	1.12	209	302	1.14	120	363
Participant 3							
High	0.95	6.05	329	297	8.66	1,194	297
Medium	0.77	4.45	401	298	6.01	1,222	301
Low	1.25	1.47	205	306	1.65	298	328
All data	62.02						

Note. OU = Ornstein–Uhlenbeck; DDV = diffusion with drift variance; d'_{asy} = asymptotic value of the sensitivity statistic d'; $1/K$ = time constant for decay; T_O = time offset.

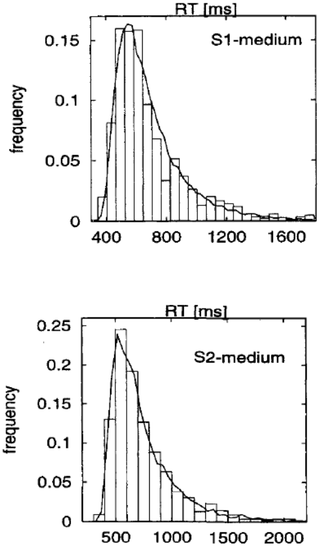
Results

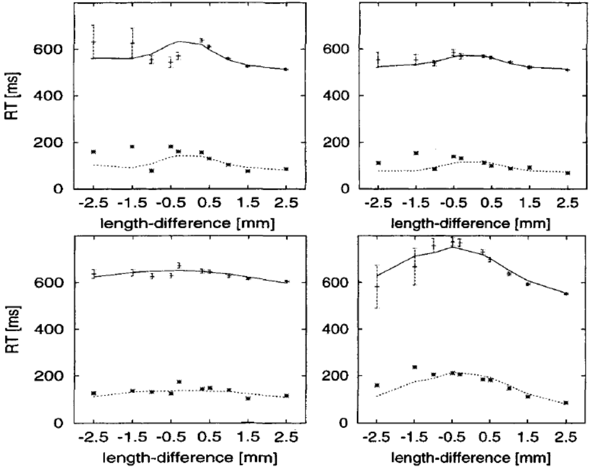
Table 2
Parameters of Model Fits to Individual Part Data From Experiment 1: Five Parameters per Participant

		Parameter									
		OU					DDV				
Participant	$\frac{P(\text{data}\mid\text{OU})}{P(\text{data}\mid\text{DDV})}$	d'_1	d'_2	d'_3	$1/K$	T_O	d'_1	d'_2	d'_3	σ^2/σ_d^2	T_O
1	6.25	3.34	2.28	0.80	135	284	3.73	2.54	0.90	211	287
2	4.76	4.41	2.82	1.14	223	307	5.34	3.43	1.38	503	308
3	1.45	6.04	4.13	1.60	325	297	8.41	5.74	2.19	1,101	298
All	43.14										

Note. OU = Ornstein–Uhlenbeck; DDV = diffusion with drift variance; $1/K$ = time constant for decay; T_O = time offset.

Results


RT [ms]
S1-medium
frequency
0.15
0.1
0.05
0
400
800
1200
1600
RT [ms]
S2-medium
frequency
0.25
0.2
0.15
0.1
0.05
0
500
1000
1500
2000


RT [ms]
600
400
200
0
-2.5 -1.5 -0.5 0.5 1.5 2.5
length-difference [mm]

Hick's Law

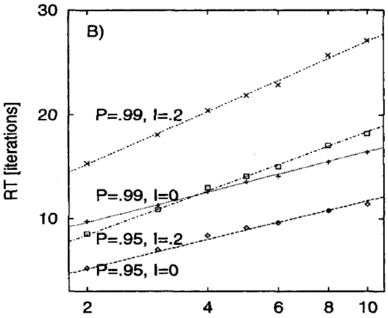
The reaction time is proportional to the logarithms of the number of alternatives.

The input to the correct accumulator is a signal ($\rho_c > 0$) plus noise.
The input to other accumulators is a smaller signal ($0 \leq \rho_i < \rho_c$) plus noise.
The input does **not** depend on the number of alternatives. The reaction time dependency should result from the structure of the model itself.

Results

The leaky, competing accumulator and two versions of the classical accumulator model show the linear relationship.

The results depend on the constant accuracy policy. If the criterion is not adjusted to keep the accuracy constant, the Hick's law behavior disappears.



Discussion

The leaky, competing accumulator combines the best elements of the drift diffusion model and the accumulator models.

It results in comparable fit for time-accuracy curves to the diffusion-with-drift-variance.

It allows for the new type of investigation: which information is more important in the decision - early (one accumulator quickly inhibits all the others and doesn't leave them any chances) or late (leakage erases the early information).

In the original paper there are few tests for the multialternative choice. However, they show potential.

Thank you