Ice Cream or Chocolate Cake?
Why does it take us so long to choose between two equally-good options?

Jan Drugowitsch
Harvard Medical School
Department of Neurobiology

Cowles Economic Theory Conference 2016
June 6, 2016
Choices and reaction times are driven by value difference

Decision speed depends only on value difference?

2D information: is $1(\frac{1}{2})D$ diffusion models are optimal?
Outline

Value-based decision setup

How option values influence the optimal policy

The limits of diffusion models

Preview: choosing between more than two options

Summary
Outline

- Value-based decision setup

  How option values influence the optimal policy

  The limits of diffusion models

  Preview: choosing between more than two options

Summary
A value-based decision setup

Decision maker learns about values by evidence accumulation

2 items, a-priori reward belief

Collect noisy evidence about reward

Evidence accumulation: uncertainty about true rewards decreases with more evidence

Aim: maximize total expected reward

Consisting of evidence accumulation costs (e.g. effort) $c_t$, grows linearly with time

Final reward $\langle z_j \rangle$ for consuming item $j$

Fast choices  ← speed/accuracy trade-off  → slow choices

Low accumulation cost might choose lower-rewarding option

High accumulation cost certain to pick higher-rewarding option
Finding the optimal decision strategy

Sufficient statistics

after having accumulated evidence for some time $t$:

- mean estimates, $\hat{z}_1$
- current time, $t$

3 required statistics:
- reward estimate $\hat{z}_1$
- reward estimate $\hat{z}_2$
- uncertainty($t$)

Decide now or accumulate more evidence?

- decide now
- accumulate more evidence

branching points for expected reward

branching points for expected reward

expected reward vs. mean estimate difference

expected reward vs. mean estimate difference

branching points for accumulation cost

branching points for accumulation cost
Finding the optimal decision strategy

Sufficient statistics

after having accumulated evidence for some time $t$:

3 required statistics:
- mean estimates, $\hat{z}$
- current time, $t$

Decide now or accumulate more evidence?

- decide now
- accumulate more evidence
Optimal decision-making with collapsing boundaries

- Mean estimate 1: $t = 250\text{ms}$
- Mean estimate 2: $t = 1\text{s}$
- Mean estimate 2: $t = 2\text{s}$

- Decision boundaries remain parallel
- Rotate by $-45^\circ$
- Only mean estimate difference matters
- Reward certainty increases with time
- Higher certainty allows faster choices
- Bound collapses in mean estimate difference
Optimal decision-making with diffusion models

Diffusion models are a straightforward mechanism to implement the optimal strategy for simple, value-based decisions.

- Certainty increases with time.
- "Diffusion" variance decreases with time (unlike in "real" diffusion models).
- Diffusion variance stays constant (as in "real" diffusion models).

Time-dependent rescaling.

"Observed" reward.
Outline

Value-based decision setup

→ How option values influence the optimal policy

The limits of diffusion models

Preview: choosing between more than two options

Summary
So far: decisions depend *only* on value difference

Decision boundaries depend on *a-priori value uncertainty*

...but *not* on absolute a-prior values
Moving from simple choices to sequences thereof

So far: single, isolated choices

- availability of choice options
- choice
- time $t$

maximize
expected reward - accumulation cost

More natural: long sequence of choices

- unlimited choices, limited time
- wait
- wait
- time $t$

maximize
total expected reward of all choices
$=$ reward rate
Loss of future reward increases time pressure

*deciding early because of*

- Single, isolated choices: cost of accumulating evidence
- Long sequence of choices: cost of accumulating evidence, loss of future reward \( \propto \) avg. reward across choices

---

**Graphs:**

- **Single, isolated choices or fixed number of choices:**
  - Particle location vs. time \( t \)

- **Long sequence of choices within limited time:**
  - Particle location vs. time \( t \)
  - Average prior reward:
    - 0
    - 4
Loss of future reward increases time pressure

*deciding early because of*

<table>
<thead>
<tr>
<th>Single, isolated choices</th>
<th>cost of accumulating evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long sequence of choices</td>
<td>cost of accumulating evidence, loss of future reward ( \propto \text{avg. reward across choices} )</td>
</tr>
</tbody>
</table>

If we have limited time to make choices, higher average rewards should lead to faster choices
Decision speed depends on relative and absolute rewards

For fixed a-priori average rewards

Why does it take so long to choose between high-valued items?

In lab settings, only the value difference matters.
In more natural settings, high absolute values should trigger faster choices (assuming the same a-prior value uncertainty).

Across different a-priori average rewards

different a-priori reward distributions
average reward

boundaries depend on average rewards, higher average rewards → faster choices
Outline

Value-based decision setup

How option values influence the optimal policy

➡️ The limits of diffusion models

Preview: choosing between more than two options

Summary
Why diffusion models work

Decision boundaries remain parallel over time → can reduce dimensionality

Why do decision boundaries remain parallel?

expected stochastic evolution of reward estimates independent of current estimates
(is there a more intuitive explanation?)
Can we always use diffusion models?

No. For example, diminishing returns predict faster choices for higher rewards which cannot be captured with diffusion models (require two-dimensional representation).
Outline

Value-based decision setup

How option values influence the optimal policy

The limits of diffusion models

→ Preview: choosing between more than two options

Summary
Models for multi-alternative choices

more than one value difference

\[ x(t) \]

\[ \text{accumulated evidence, option } 2 \]

\[ \text{accumulated evidence, option } 3 \]

\[ \text{accumulated evidence, option } 1 \]

Use race models instead?

more than one value difference

\[ \text{can't use diffusion models} \]
Urgency and divisive normalization

Urgency signal

(Churchland, Kiani & Shalden, 2008; Drugowitsch et al, 2011)

Divisive normalization

(Louie, Grattan, and Glimcher, 2011)
Normative choices with multiple alternatives

Accumulated evidence for option 1, $x_1$

Accumulated evidence for option 2, $x_2$

Accumulated evidence for option 3, $x_3$

Choose 1

Choose 2

Choose 3

From 2-AFC to 3-AFC

Lower-dimensional projection

Accumulate evidence

Choose 1

Choose 2

Choose 3

Time $t$
A mechanism to approximate the optimal policy

Network model well-approximates optimal decision boundaries

Network model well-approximates optimal decision boundaries

Used to model data in, e.g.: LoFaro, Louie, Webb & Glimcher (2014), Hunt et al. (2012)
Influence of urgency and normalization

Optimal decision-making mandates divisive normalization (and urgency signal)
Replicated empirical findings

Data

- Firing rate (normalized) for target in RF
  - Data points showing the relationship between the total value of targets outside RF (μl) and firing rate (normalized).

Model

- Unit activity (a.u.)
  - Graph showing the relationship between the total value of targets outside RF (a.u.) and unit activity (a.u.).

...and others
Summary

How option values influence the optimal policy
Single or sequence of choices, no time pressure: only value difference, a-prior uncertainty
Sequence of choices in limited time: absolute values start to matter

Optimality of diffusion models
Holds for homogeneous “evidence” structure across values
Breaks down in case of inhomogeneity, for example, diminishing returns

Multi-alternative decisions
Require complex decision boundaries in higher-dimensional space
Optimal strategy well approximated by race model with normalization and urgency
Matches neurophysiology
Thanks

Satohiro Tajima
Alexandre Pouget

The Pouget Lab
Ruben Coen Cagli
Guillaume Dehane
Daniel Robles Llana
Tamás Madarász

UNIVERSITÉ DE GENÈVE