The Computing Brain: Focus on Decision-Making

Cog 1

March 5, 2009

Angela Yu

ajyu@ucsd.edu
Understanding the Brain/Mind?

Behavior

Cognitive Neuroscience

≈

Neurobiology

A powerful analogy: the computing brain
An Example: Decision-Making
An Example: Decision-Making

What are the computations involved?

2. Uncertainty

3. Costs:

1. Decision

THUR

74°/61°

Health?

Coolness?
Monkey Decision-Making

Random dots coherent motion paradigm
Random Dot Coherent Motion Paradigm

30% coherence  
5% coherence
How Do Monkeys Behave?

Accuracy vs. Coherence

<RT> vs. Coherence

(Roitman & Shadlen, 2002)
What Are the Neurons Doing?

Saccade generation system

MT tuning function

(Britten & Newsome, 1998)

MT: Sustained response

99.9%

Preferred

Anti-preferred

25.6%

(Britten, Shadlen, Newsome, & Movshon, 1993)
What Are the Neurons Doing?

Saccade generation system

LIP: **Ramping** response

(Roitman & Shalden, 2002)  
(Shadlen & Newsome, 1996)
A Computational Description

1. Decision: Left or right? Stop now or continue?

\( \pi(t) \): input(1), \ldots, input(t) ⇒ \{left, right, wait\}
Timed Decision-Making

$x_t$  

wait  

$x_{t+1}$  

wait  

$x_{t+2}$
A Computational Description

2. **Uncertainty**: Sensory noise (coherence)
Timed Decision-Making

$x_t$  

$\text{wait}$  

$\text{wait}$  

$\text{wait}$  

$L$  

$L$  

$L$  

$R$  

$R$  

$R$
A Computational Description

3. Costs: Accuracy, speed

\[ \text{cost}(\pi) = \text{Pr(error)} + c*(\text{mean RT}) \]

**Optimal** policy \( \pi^* \) minimizes the cost
Timed Decision-Making

$X_t \xrightarrow{\text{wait}} X_{t+1}$

$+\text{: more accurate}$

$-\text{: more time}$

$L \xrightarrow{\text{wait}} R$
Mathematicians Solved the Problem for Us

Wald & Wolfowitz (1948):

**Optimal policy:**

- accumulate evidence over time: Pr(left) versus Pr(right)
- stop if total evidence exceeds “left” or “right” boundary
Model ⇔ Neurobiology

Saccade generation system

Model

LIP Neural Response

Motion strength
- 51.2
- 25.6
- 12.8
- 6.4
- 3.2
- 0

Time (ms)

Firing rate (sp/s)
Model ⇔ Behavior

Saccade generation system

Model

RT vs. Coherence

<table>
<thead>
<tr>
<th>Coherence</th>
<th>Reaction time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000</td>
</tr>
<tr>
<td>10</td>
<td>900</td>
</tr>
<tr>
<td>100</td>
<td>800</td>
</tr>
</tbody>
</table>

boundary
Putting it All Together

Saccade generation system

- MT neurons signal motion direction and strength
- LIP neurons accumulate information over time
- LIP response reflect behavioral decision (0% coherence)
- Monkeys behave optimally (maximize accuracy & speed)
Understanding the Brain/Mind?

Behavior

Neurobiology

Cognitive Neuroscience

A powerful analogy: the **computing brain**
Sequential Probability Ratio Test

Log posterior ratio \( r_t \triangleq \log \frac{q_t}{1-q_t} \), \( q_t = \frac{p(x_t|s=1)q_{t-1}}{Z_t} \)

undergoes a random walk:

\[ r_t = \log \frac{p(x_t|s=1)}{p(x_t|s=0)} + r_{t-1} \quad r_0 = \log \frac{P(s=1)}{P(s=0)} \]

\( r_t \) is monotonically related to \( q_t \), so we have \( (a', b') \), \( a'<0, b'>0 \).

In continuous-time, a drift-diffusion process w/ absorbing boundaries.
Generalize Loss Function

We assumed loss is *linear* in error and delay:

\[ L(\pi) = P_\pi(\hat{s} \neq s) + c\langle\tau\rangle_\pi \]

What if it’s *non-linear*, e.g. maximize reward rate:

\[ \frac{1 - P(\hat{s} \neq s)}{\langle\tau\rangle} \]

(Bogacz et al, 2006)

Wald also proved a *dual* statement:

Amongst all decision policies satisfying the criterion \( P(\hat{s} \neq s) \leq \alpha \)

SPRT (with some thresholds) minimizes the expected sample size \( \langle\tau\rangle \).

This implies that the SPRT is optimal for *all* loss functions that increase with inaccuracy and (linearly in) delay (proof by contradiction).
Neural Implementation

Saccade generation system

LIP - neural SPRT integrator? (Roitman & Shadlen, 2002; Gold & Shadlen, 2004)

LIP Response & Behavioral RT

LIP Response & Coherence
Caveat: Model Fit Imperfect
(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)
Fix 1: Variable Drift Rate
(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)
(idea from Ratcliff & Rouder, 1998)
Fix 2: Increasing Drift Rate
(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)

Accuracy  Mean RT  RT Distributions

Accuracy vs. Motion Strength
- % correct vs. Motion strength [% coh]
- Data and model comparison

Mean RT vs. Motion Strength
- Mean RT (+/- SE) vs. Motion strength [% coh]
- Correct trials and error trials comparison

RT Distributions
- Correct trials and Error trials distributions
- Model and data comparison
A Principled Approach

Trials aborted w/o response or reward if monkey breaks fixation

• More “urgency” over time as risk of aborting trial increases
• Increasing gain equivalent to decreasing threshold

(Data from Roitman & Shadlen, 2002)
(Analysis from Ditterich, 2007)
Imposing a Stochastic Deadline
(Frazier & Yu, 2007)

Loss function depends on error, delay, and whether deadline exceed

\[ L(\pi) = P_{\pi}(\hat{s} \neq s)P(\tau < D) + c\langle \tau \rangle_\pi + dP(\tau \geq D) \]

Optimal Policy is a sequence of *monotonically decaying* thresholds
Timing Uncertainty

(Frazier & Yu, 2007)

Timing uncertainty $\rightarrow$ lower thresholds

Theory applies to a large class of deadline distributions: $\textit{delta, gamma, exponential, normal}$
Some Intuitions for the Proof

(Frazier & Yu, 2007)
Summary

• Review of Bayesian DT: optimality depends on \textit{loss} function

• Decision \textit{time} complicates things: infinite repeated decisions

• Binary hypothesis testing: SPRT with fixed thresholds is optimal

• Behavioral & neural data \textit{suggestive}, but …

• … imperfect fit. \textit{Variable/increasing} drift rate both ad hoc

• Deadline (+ timing uncertainty) \rightarrow \textit{decaying} thresholds

• Current/future work: generalize \textit{theory}, test with \textit{experiments}