Optimal Control Approaches to Language
How the Architecture "Shows Through"

Richard L. Lewis  Michael Shvartsman  Satinder Singh

Psychology, Computer Science
University of Michigan

Soar Workshop 32(!)
21 June 2012
All aspects of linguistic processing and behavior—from parsing strategies to production strategies to control of short and long-term memory to eye-movement control—may be understood as the solution to the constrained optimization problem posed by the external task environment, task structure, and internal processing structure/constraints (e.g. representation noise, knowledge).

Howes, Lewis & Vera (2009, *Psychological Review*)

We’ll pursue this via the application of state-of-the-art theoretical ideas ... from the 1940-60s: optimal control and optimal state estimation.
Overview

What determines the nature of eye-movements in linguistic tasks?
- The task and model
- Predictions vs. human behavior
- How architecture shapes adaptation

Conclusions and looking ahead
The List Lexical Decision Task

Version of this task first used by Schvanavedlt & Meyer (1973); Meyer & Schvanavedlt (1972)
Explicit payoffs: Motivating with cash

We evaluated both model and human participants according to three different payoffs (specified in Table 1). The payoffs were designed to impose different speed-accuracy tradeoffs for a given level of success, and were all defined in terms of a bonus for speed and penalty for incorrect responses. The bonus was continuous at the millisecond level, starting at zero points for responses longer than 5s and rising by a different number of points per second for each payoff.

An optimal control model

Main theoretical assumptions

We can now briefly state our three main theoretical assumptions:

1. Saccadic control is a “rise-to-threshold” system (Brodersen et al., 2008) conditioned on task-specific decision variables that reflect the accumulation and integration of noisy evidence over time. We model the dynamic evidence accumulation as Bayesian sequential sampling, and in our simple two-alternative task this is equivalent to a Sequential Probability Ratio Test (Wald & Wolfowitz, 1948).

2. The saccade thresholds are set to maximize task-specific payoffs, but this is one part of a joint optimization problem that includes all other policy parameters that determine behavior in the task. In our model of the LLDT, this consists of a separate decision variable and threshold that determines the task-level response to the entire trial (but does not include architectural parameters, which are fixed). These two thresholds together determine how long the model fixates on individual strings, how many strings it reads, and when and how it responds.

3. The shape of the payoffs surface (and thus its maxima) over the multi-dimensional policy space is determined jointly by the payoffs function and properties of the perceptual and oculomotor system, including saccade programming duration, eye-brain-lag, saccade execution duration, manual motor programming duration, and representational noise.

Overview

We provide a brief overview of a typical trial before focusing in on specific detail of each aspect of the model specification. See Figure 2 for a schematic diagram of the full model, and Figure 3 for simulated traces from two sample trials.

On a given trial, the first fixation starts on the leftmost string. During each fixation, noisy information about the fixated string is acquired at every timestep (with some delay, the eye-brain-lag, (VanRullen & Thorpe, 2001)). This noisy information is used for updating the model’s beliefs about the status of the current string as well as the trial as a whole. This continues until either the string-level or the trial-level belief reaches some threshold, at which point either a saccade is initiated (if the string-level threshold is reached), or a manual response is initiated (if the trial-level threshold is hit). We will refer to these thresholds as the saccade threshold and decision threshold. Information acquisition continues while the saccade or manual response is being programmed and until the saccade begins execution (with some visual persistence set). Once saccade programming and execution is complete, the model fixates on the following string (if there are strings remaining), or initiates a response otherwise. Once motor programming and execution is complete, the model receives point feedback (i.e. the payoffs) and the trial is over.

Dynamics assumptions: Oculo-motor architecture and noise

The model’s sequential perceptual inference mechanism is embedded in a simple oculomotor control machine, drawing upon modern mathematical models of oculomotor control in reading. The delays noted above (eye-brain-lag, saccade programming and execution times, and motor time) are drawn from gamma distributions (chosen for convenience because they are constrained to be positive and have been previously used to model these delays).
Explicit payoffs: Motivating with cash

<table>
<thead>
<tr>
<th>Incorrect penalty</th>
<th>Speed bonus (per second under 5s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-150</td>
<td>8</td>
</tr>
<tr>
<td>-50</td>
<td>6.7</td>
</tr>
<tr>
<td>-25</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Table 1: Quantitative payoffs given to both model and human participants. These payoffs translated into cash bonuses for the human participants.

We evaluated both model and human participants according to three different payoffs (specified in Table 1). The payoffs were designed to impose different speed-accuracy tradeoffs for a given level of success, and were all defined in terms of a bonus for speed and penalty for incorrect responses. The bonus was continuous at the millisecond level, starting at zero points for responses longer than 5s and rising by a different number of points per second for each payoff.
Explicit payoffs: Motivating with cash

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Balanced</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect penalty</td>
<td>-150</td>
<td>-50</td>
<td>-25</td>
</tr>
<tr>
<td>Speed bonus (per second under 5s)</td>
<td>8</td>
<td>6.7</td>
<td>5.7</td>
</tr>
</tbody>
</table>
An adaptive model that performs the complete task

See Legge et al. (1997); Bicknell & Levy (2010); Ratcliff & McKoon (2008); Norris (2009).
“Random walk” of Bayesian posterior update

Graph from Bogacz (2009), TINS.
Bayesian priors and stimulus representation

Model maintains belief probabilities over:

(a) probability distribution over all possible strings in the currently fixated position;

(b) the probability of a nonword in each position; probability current trial is a word trial is $1 - \text{sum over these}$.

The prior over (a) is based on Brown Corpus frequency; prior over (b) is probability of a nonword trial (0.5) divided by the # of positions (6).

The string stimulus is represented with a simple indicator vector coding (length $26 \times 4$) (Norris, 2009). At each sample (10ms), Guassian noise of mean zero and $SD = 1.2$ (more on this) is added to the true representation.
Sample model behavior

CORRECT WORD Trial

brag none mean bloc hilt hair

FIXATION prog prog prog prog prog prog
EBL sampling sampling sampling sampling sampling sampling

MOTOR All words!!

0 500 1000 1500 2000

Saccade threshold = 0.92000, Decision yes = 0.99900, Decision no = 0.99900, Noise = 1.15, Sac. prog = 125 ms
Sample model behavior

INCORRECT WORD Trial

lard sane helm toot east lost

Not all words!!
Generating predictions

- Selecting a policy (pair of thresholds) “programs” the machine to perform the task.
- Policy selected through payoff optimization—not data fitting.

<table>
<thead>
<tr>
<th>Payoff condition</th>
<th>Optimal Saccade Threshold</th>
<th>Optimal Response Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.99</td>
<td>0.999</td>
</tr>
<tr>
<td>Balanced</td>
<td>0.97</td>
<td>0.999</td>
</tr>
<tr>
<td>Speed</td>
<td>0.92</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Generating predictions

- Selecting a policy (pair of thresholds) “programs” the machine to perform the task.
- Policy selected through payoff optimization—not data fitting.

<table>
<thead>
<tr>
<th>Payoff condition</th>
<th>Optimal Saccade Threshold</th>
<th>Optimal Response Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.99</td>
<td>0.999</td>
</tr>
<tr>
<td>Balanced</td>
<td>0.97</td>
<td>0.999</td>
</tr>
<tr>
<td>Speed</td>
<td>0.92</td>
<td>0.99</td>
</tr>
</tbody>
</table>

I.e, \( \pi_{\text{speed}}^* = (0.92, 0.99) \) and so on. With fixed policy, machine generates dozens of behavioral measures (e.g. RTs, errors, RTs for accurate vs./inaccurate, fixation durations for words, nonwords, frequency effects, . . .)
Finding the sweet spot: Payoff as function of thresholds

$$\pi^* = \arg \max_{\pi \in \Pi} \mathbb{E}_{\text{trial} \sim \text{Experiment}} U(\pi, \text{trial})$$  (1)
Payoff as a function of behavior

Payoffs vs. SFD
(MODEL, noise=1.20)

Payoffs vs. Trial RT
(MODEL, noise=1.20)

Payoffs vs. Frequency Effect
(MODEL, noise=1.20)
Model and human at level of trial

Response Time for Word Trials
(Model, noise=1.20)

Response Time for Nonword Trials
(Model, noise=1.20)

Percent Correct
(Model, noise=1.20)
Model and human at level of trial

- **Response Time for Word Trials** (MODEL, noise=1.20)
- **Response Time for Nonword Trials** (MODEL, noise=1.20)
- **Percent Correct** (MODEL, noise=1.20)

- **Response Time for Word Trials** (HUMAN)
- **Response Time for Nonword Trials** (HUMAN)
- **Percent Correct** (HUMAN)
Model and human at level of word/string

- **Single Fixation Duration**
  - (MODEL, noise=1.20)
  - Accuracy, Balance, Speed

- **SFD by Frequency Class**
  - (MODEL, noise=1.20)
  - Accuracy, Balance, Speed

- **Frequency Effect on SFD**
  - (MODEL, noise=1.20)
  - Accuracy, Balance, Speed

Lewis (University of Michigan)
Model and human at level of word/string

**Model vs. human**

- **Optimal Control Approaches**

21 June 2012

Lewis (University of Michigan)

---

**Single Fixation Duration**

- Model, noise=1.20

**SFD by Frequency Class**

- Model, noise=1.20

**Frequency Effect on SFD**

- Model, noise=1.20

---

**Single Fixation Duration**

- Human

**SFD by Frequency Class**

- Human

**Frequency Effect on SFD**

- Human
Model and human: Words vs. nonwords, position effects

Word SFD by correctness
(Model, noise=1.20)

Nonword SFD by correctness
(Model, noise=1.20)
Model and human: Words vs. nonwords, position effects

Word SFD by correctness
(MODEL, noise=1.20)

Nonword SFD by correctness
(MODEL, noise=1.20)

Word SFD by correctness
(HUMAN)

Nonword SFD by correctness
(HUMAN)
Model and human: Words vs. nonwords, position effects

Word SFD by correctness

Nonword SFD by correctness

Word SFD by Position in List

Optimal Control Approaches
Model and human: Words vs. nonwords, position effects

**Word SFD by correctness** (MODEL, noise=1.20)

**Nonword SFD by correctness** (MODEL, noise=1.20)

**Word SFD by Position in List** (MODEL, noise=1.20)

---

**Word SFD by correctness** (HUMAN)

**Nonword SFD by correctness** (HUMAN)

---
Model and human: Words vs. nonwords, position effects

Word SFD by correctness
(MODEL, noise=1.20)

Nonword SFD by correctness
(MODEL, noise=1.20)

Word SFD by Position in List
(MODEL, noise=1.20)

Word SFD by correctness
(HUMAN)

Nonword SFD by correctness
(HUMAN)

Word SFD by Position in List
(HUMAN)
1. What determines the nature of eye-movements in linguistic tasks?
   - The task and model
   - Predictions vs. human behavior
   - How architecture shapes adaptation

2. Conclusions and looking ahead
Does the processing architecture matter?

The theoretical claim here is that eye-movement control is jointly shaped by both task payoff and architecture. What is the evidence for this?

Through modeling we can explore adaptation to different architectures than the one hypothesized for the human oculomotor system.
The “minimal model”

Experiment Environment

lead hilt robe helm guru east

Oculomotor System

Initiate saccade program

saccade program

~ Gamma(125ms)

saccade

~ Gamma(40ms)

eye-brain lag ~ Gamma(50ms)

Noisy sample $s$ from position $k$

Manual response

~ Gamma(250ms)

Initiate press

Posterior Update

$Pr_{\text{new}}(T = W | s^k)$, ...

(see Appendix)

(Bounded) Optimal Control

optimal thresholds

saccade decision

trial decision

Reward Function $I(\cdot)$

Lexicon + Model of experiment

Button press indicating trial response

Feedback

RT dist. for RN.mismatch

SD = 2.00, threshold = 0.95

RT dist. for NRN.mismatch

SD = 3.50, threshold = 0.92

RT dist. for YES.match

SD = 1.75, threshold = 0.95

Hillsdale, NJ: Lawrence Erlbaum.
The “minimal model”

Experiment Environment

Button press indicating trial response

Oculomotor System

Initiate saccade program

Manual response

~ Gamma(250ms)

Posterior Update

Pr_{new}(T = W|s^k),... (see Appendix)

(Bounded) Optimal Control

Reward Function

I(·)

Noisy sample s from position k

Eye-brain lag ~ Gamma(50ms)

Penultimate, we need the probability that the string at position k is a word or nonword, i.e.,

In order to make decisions, in addition to the probability that the string at position k is a word or nonword, i.e.,

Finally, we update the trial level beliefs:

Next, we update the string-level beliefs:

References


Payoff structure & predictions for the minimal model

**ACC payoff vs. Saccade Threshold**

(MODEL, noise=1.90)

- Payoff values: (0.9990, 0.9999)

**BAL payoff vs. Saccade Threshold**

(MODEL, noise=1.90)

- Payoff values: (0.9950, 0.9990)

**SPD payoff vs. Saccade Threshold**

(MODEL, noise=1.90)

- Payoff values: (0.9900, 0.9990)

**ACC payoff vs. Decision Threshold**

- Payoff values: (0.9990, 0.9999)

**BAL payoff vs. Decision Threshold**

- Payoff values: (0.9950, 0.9990)

**SPD payoff vs. Decision Threshold**

- Payoff values: (0.9900, 0.9990)

**SFD by Frequency Class**

- Frequency classes: LOW, HIGH

**Log Frequency Effect on SFD**

- Frequency values: 20, 21, 22, 23, 24, 25

**Word SFD by Position in List**

- Position values: 2, 3, 4, 5

**Review word-level results**

Lewis (University of Michigan)
Payoff structure & predictions for the minimal model

ACC payoff vs. Saccade Threshold
BAL payoff vs. Saccade Threshold
SPD payoff vs. Saccade Threshold

Saccade Threshold
SFD by Frequency Class
Frequency Effect on SFD
Word SFD by Position in List

Single Fixation Duration (ms)
Model fit for the alternative architectures

Model Error vs. Noise for Architectural Variants

- Complete architecture
- Architecture with only saccade programming
- Architecture without saccade programming
- Minimal model (noise only)

RMSE against Human SFD

Noise

Lewis (University of Michigan)
1. What determines the nature of eye-movements in linguistic tasks?

2. Conclusions and looking ahead
We applied *optimal control* and *state estimation* techniques to pursue, computationally, an interesting theoretical idea. Doing so yielded two things:

1. What determines the nature of eye-movements in linguistic tasks?  
   **Answer:** Eye-movement control is the solution to a constrained optimization problem posed by task structure and payoff, linguistic knowledge, and oculomotor processing architecture.

2. A novel empirical demonstration: Humans adapt their oculomotor control at the level of *single fixation durations* to maximize payoff in linguistic tasks, and do so in ways sensitive to the specific contingencies of the task at hand.
Stronger ties between psycholinguistics and linguistics and other areas of cognitive science?

Bayesian memory & perception
reinforcement learning
decision making
rational analysis
bounded rationality
psycholinguistics:
    classic parsing strategies
    rational approaches
    language-as-action
syntactic theory
language evolution
The question we can pose is: What optimization problem (specifically, bounded optimal control problem) is human language the solution to?

- This offers a perspective on language evolution/emergence that complements existing approaches by placing emphasis on how the details of cognitive architecture and utility shape language, abstracting away from processes of evolution.
- It perhaps offers another way to pursue the “Strong Minimalist” thesis of optimality in language (recent work by Chomsky).

For more, see Bratman, Shvartsman, Lewis & Singh (2010) on my website.
# Grammar as bounded optimal policy

Bratman, Shvartsman, Lewis & Singh (2012)

<table>
<thead>
<tr>
<th>ENVIRONMENT</th>
<th>AGENT MEMORY</th>
<th>LEXICON SIZE (S)</th>
<th>PROPERTIES OF EMERGENT LINGUISTIC SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two Rooms</td>
<td>one symbol working memory + one symbol long-term memory</td>
<td>3</td>
<td>Association and systematic order, where in addition single symbols uttered in isolation denote specific box-key combinations. Can only achieve 75% success.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>Association and systematic symbol order. SPEAKER first describes the box, then the key (see Figure 2b).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>Highly context-dependent and idiosyncratic symbol meanings. For example key 2 is represented by symbol 4 if uttered before box, but symbol 5 after.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>Each symbol denotes a box-key combination. For example symbol 5 means key 1 and box 1.</td>
</tr>
<tr>
<td>Two rooms</td>
<td>two symbol working memory (no long-term memory)</td>
<td>3</td>
<td>Similar to case with 3 symbols above.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>Complex lexical forms. Describes entire box-key combination with two symbols which can be observed simultaneously by LISTENER effectively creating a 2-symbol length word (see Figure 3b).</td>
</tr>
<tr>
<td>One room</td>
<td>one symbol working memory + one symbol long-term memory</td>
<td>3</td>
<td>Symbols act as direct orders to LISTENER, but otherwise policy is similar to the cases of 3 symbols above.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>Association and symbol order, but no storing or retrieving from long-term memory is necessary because LISTENER can act immediately upon hearing a symbol (see Figure 4b).</td>
</tr>
</tbody>
</table>
Collaborators and sponsors

Michael Shvartsman  
U of Michigan Psychology

Satinder Singh  
U of Mich Computer Sci

Andrew Howes  
U of Birmingham Comp Sci

Marc Berman  
U of Toronto Psych

Julie Boland  
U of Michigan Psychology

Sam Epstein  
U of Michigan Linguistics

Miki Obata  
Mie University Linguistics

John Jonides  
U of Michigan Psychology

Language & Cognitive Architecture Lab: Bryan Berend, Yasaman Kazerooni, Emmanuel Kumar, Craig Sanders, Mehgha Shyam

National Science Foundation