The subatomic components of thought

Erik M. Altmann

Michigan State University

www.msu.edu/~ema

Issues

- Associative memory vs partial matching
- Math vs. process
  - Latency = *f*(Activation)
  - Error = *f*(Activation)
  - Competitive latency
- Base-level learning

$$B = \ln\left(\frac{n}{\sqrt{T}}\right) \longrightarrow B = \log\left(\sum_{i=1}^{m} t_{n-i}^{-d} + \frac{(n-m)\left(t_0^{1-d} - t_{n-m}^{1-d}\right)}{(1-d)\left(t_0 - t_{n-m}\right)}\right)$$

7 ± 2 sources of confusion

1.	Associative memory	Partial matching
2.	Context effects	Gradient effects
3.	Associative links	Similarity
4.	Diffuse priming	Constrained match
5.	Semantic/temporal	Psychophysical
6.	Arbitrary addressing	Content addressing
7.	Chunk as cue	Slot+value as cue

Learnability constraint

- How are associations learned?
  - Temporal co-occurrence of declarative items
    - C.f., Aristotle, Hume, etc.
  - Search for constraints on Sjis
    - Bayesian approach was strike 1
- How are similarities learned?
  - ACT-R: just another just-so story

Observations

- Misconception: Associative links are symbolic, clean, “sharp-edged”
  - Activation + noise + associative learning = gradient representations
- Experience (time) is effectively continuous
  - Semantic representations emerge from $10^{?}$ events
    - E.g., Latent semantic analysis
- Gradient effects with associative priming...

Cognitive arithmetic

(ACT 98, p. 78)

Answer

Problem

	0	1	2	3	4	5	6	7	8	Other
A&L 98 (100 runs):										
1+1	0	0.16	0.81	0.03	0	0	0	0	0	0
1+2	0	0.01	0.23	0.69	0.04	0.01	0	0	0	0.02
1+3	0	0	0.01	0.15	0.77	0.04	0	0	0	0.02
2+2	0	0	0.09	0.14	0.72	0.02	0	0	0	0.02
2+3	0	0	0.01	0.09	0.18	0.54	0.08	0.01	0	0.09
3+3	0	0	0	0.05	0.08	0.08	0.66	0.02	0	0.11
Siegler:										
1+1	0	0.05	0.86	0	0.02	0	0.02	0	0	0.06
1+2	0	0.04	0.07	0.75	0.04	0	0.02	0	0	0.09
1+3	0	0.02	0	0.1	0.75	0.05	0.01	0.03	0	0.06
2+2	0.02	0	0.04	0.05	0.8	0.04	0	0.05	0	0
2+3	0	0	0.07	0.09	0.25	0.45	0.08	0.01	0.01	0.06
3+3	0.04	0	0	0.05	0.21	0.09	0.48	0	0.02	0.11
EMA (100 runs):										
1+1	0	0.04	0.86	0.02	0.03	0.02	0.02	0	0	0
1+2	0	0.01	0.11	0.8	0.03	0.02	0.01	0	0	0.01
1+3	0	0.02	0.13	0.02	0.78	0.03	0.01	0	0	0.01
2+2	0	0	0.04	0.12	0.8	0.02	0.02	0	0	0
2+3	0	0	0.01	0.08	0.13	0.59	0.03	0.01	0	0.15
3+3	0	0	0	0.07	0.27	0.06	0.57	0.01	0.02	0

Partial matching:

RMSE = 0.050

R^2 = 0.94

No partial matching:

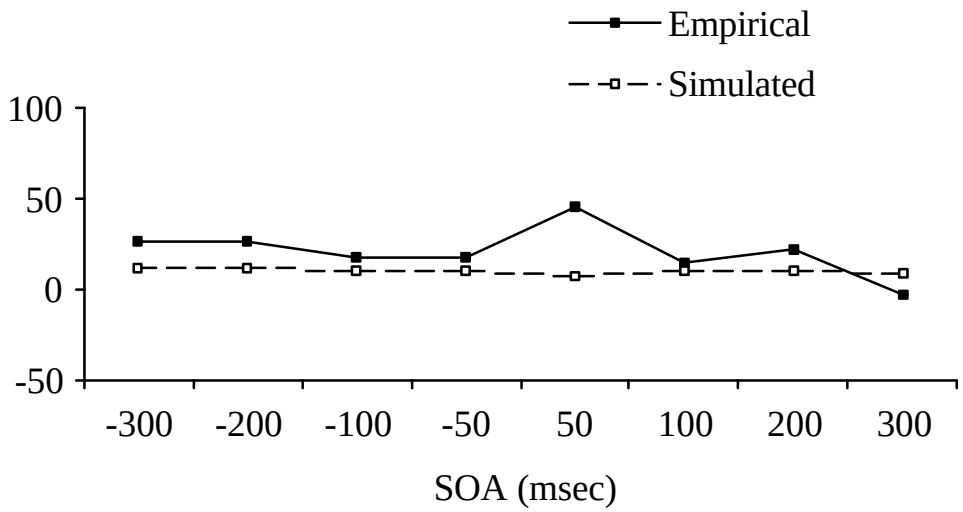
RMSE = 0.046

R^2 = 0.96

Semantic gradient

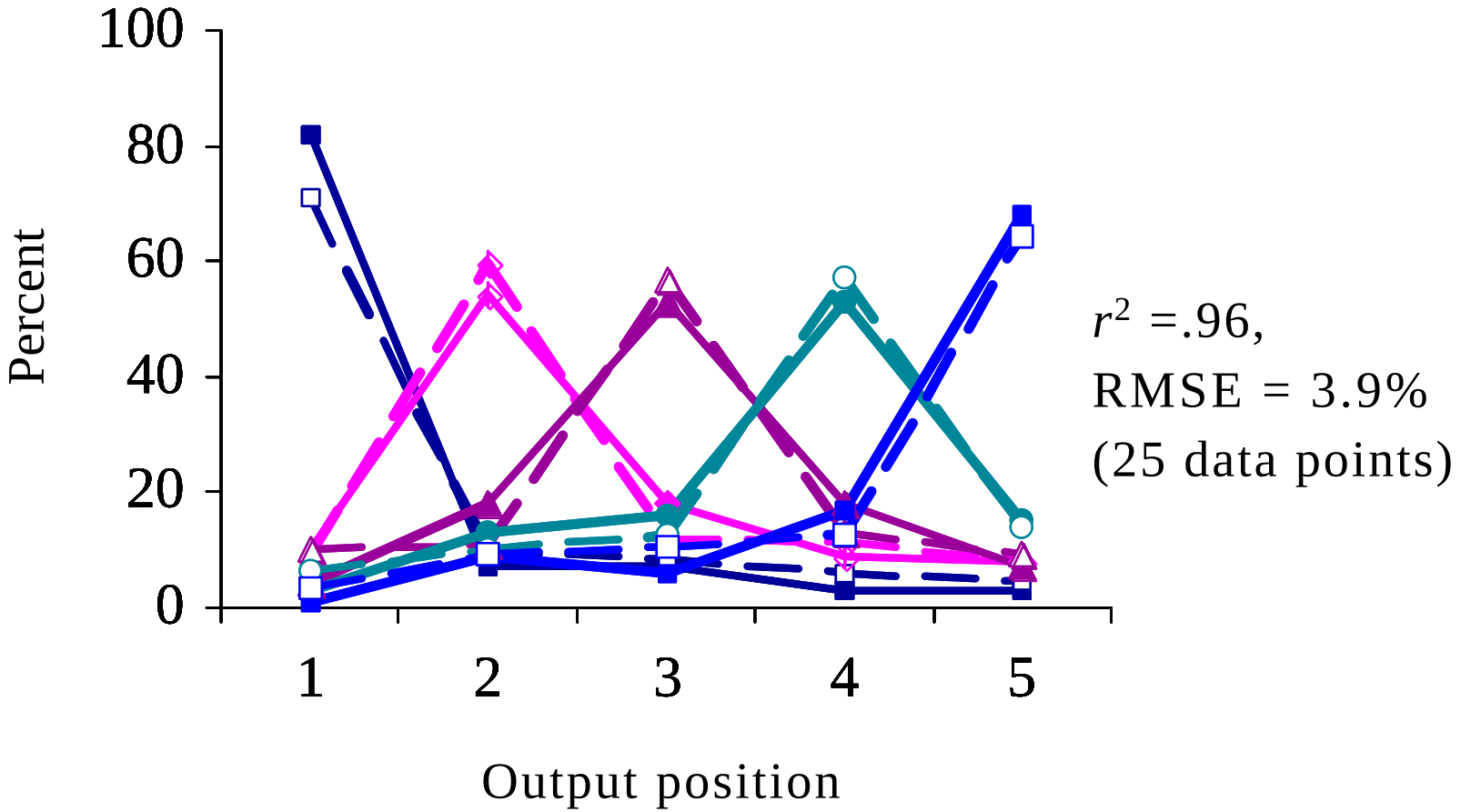
Lawn

Time



Temporal gradients

(Nairne, 92)



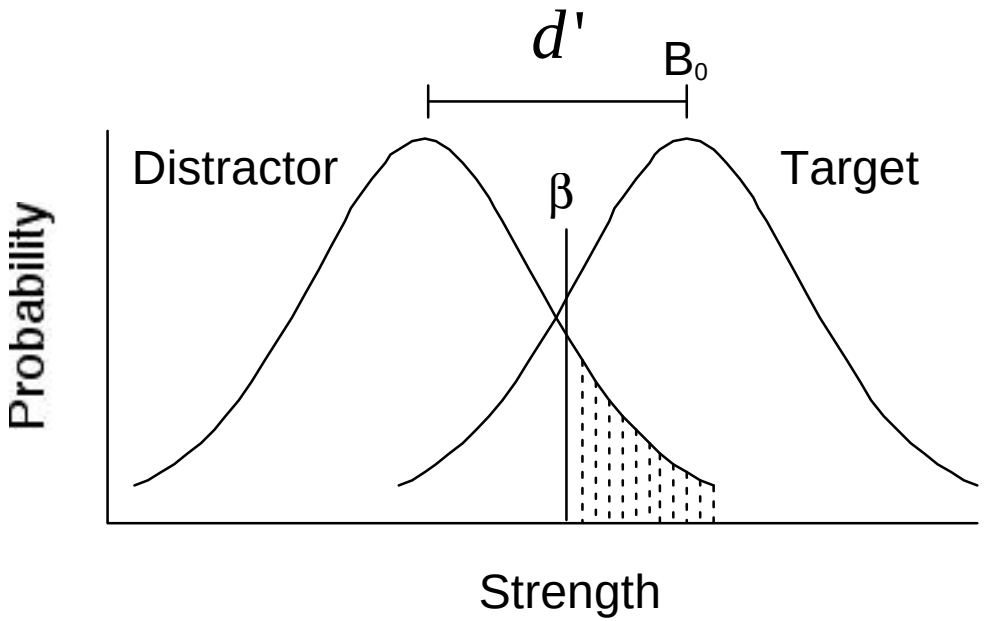
Comments

- Leave S_{ji}'s open (as similarities are now)
- Tackle psychophysical effects directly
  - Clock faces, hues, faces, …
  - Have we used partial matching on these?
- Listen to the architecture!
  - What can $10^{?}$ co-occurrences buy you?
- Throw away partial matching
  - Don't need it, don't want it, can't explain it

Know the equation, but ...

- What's the process linking activation to latency? To error?
  - Random walk models have an answer
- What process mediates the effect of distractors on the target?
  - Is there a competitive latency process?

Memory as signal detection



A retrieval process

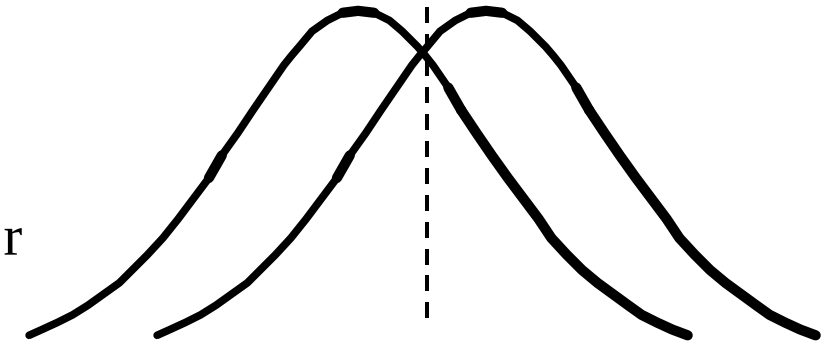
- Retrieve the most active item
- If you can recognize the target, and
  the retrieved item is not it, and
  there's time to try again, then
  attempt retrieval again
- Else stop and output item to next process

Characteristics

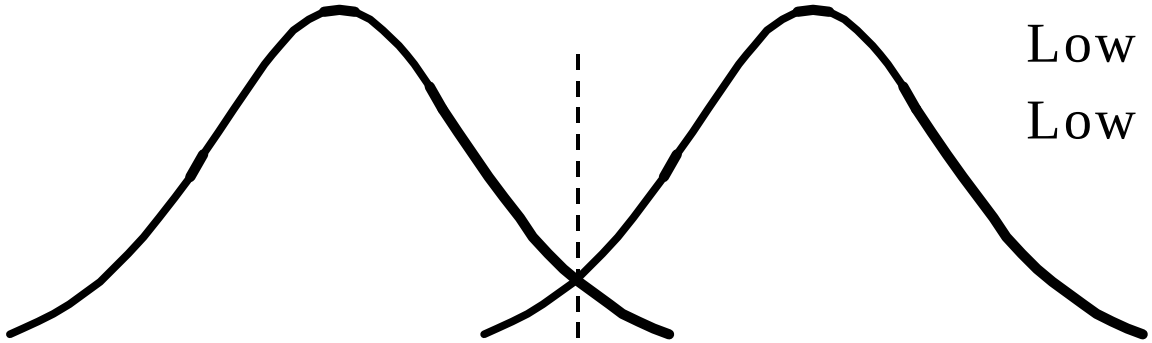
- Latency predicted by number of attempts
  - Each retrieval is constant time
- Errors predicted by intrusions
  - If you don’t know what you’re looking for
  - If you know, but run out of time
- Activation dynamics constrain parameters
  - Errors feed forward
  - Retrieval threshold and number of attempts

(Competitive) latency and error

The latency transfer function (e.g., Murdock, 65)



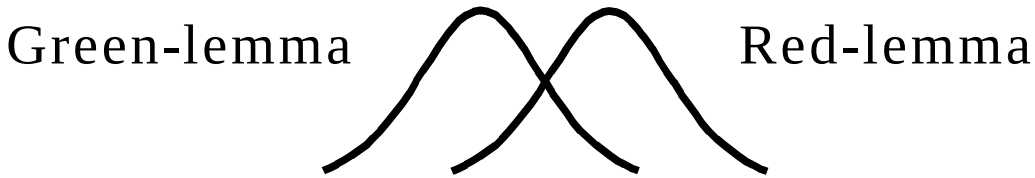
High latency,
High error



Low latency,
Low error

Target recognizable

- Speech production depends on lemmas
  - Word-sized syntactic units
- "Green" activates a lemma automatically
  - Green-lemma interferes with red-lemma

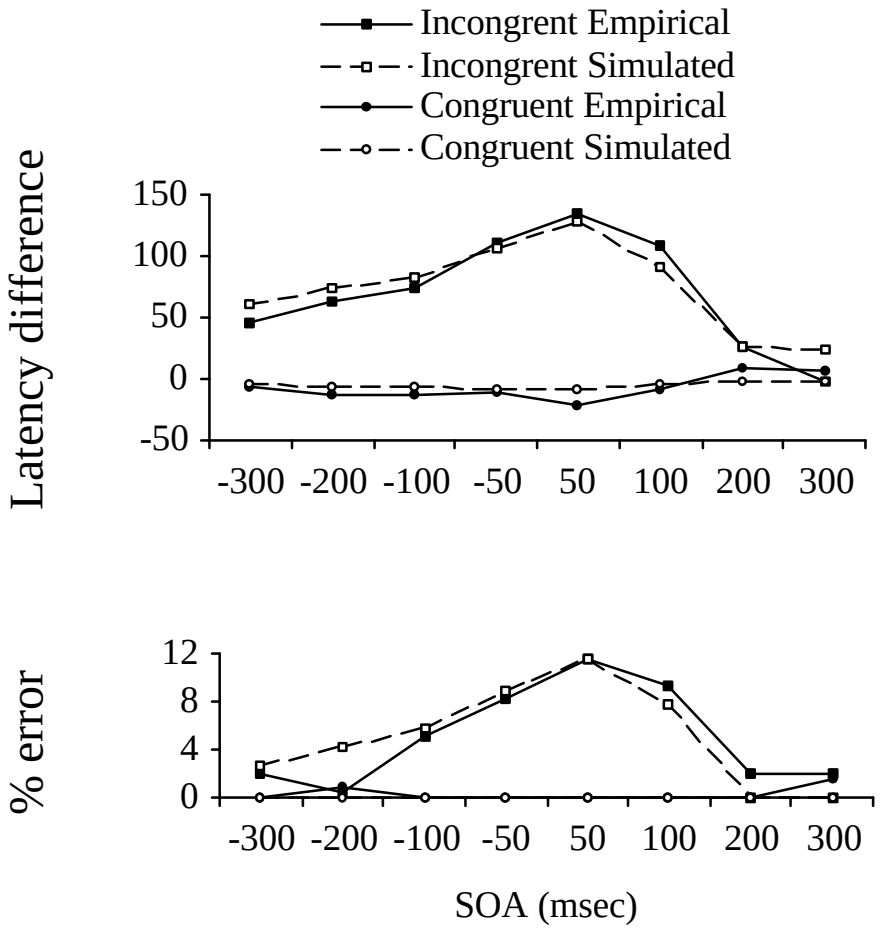


- Can compare the target lemma to the stimulus

A retrieval process

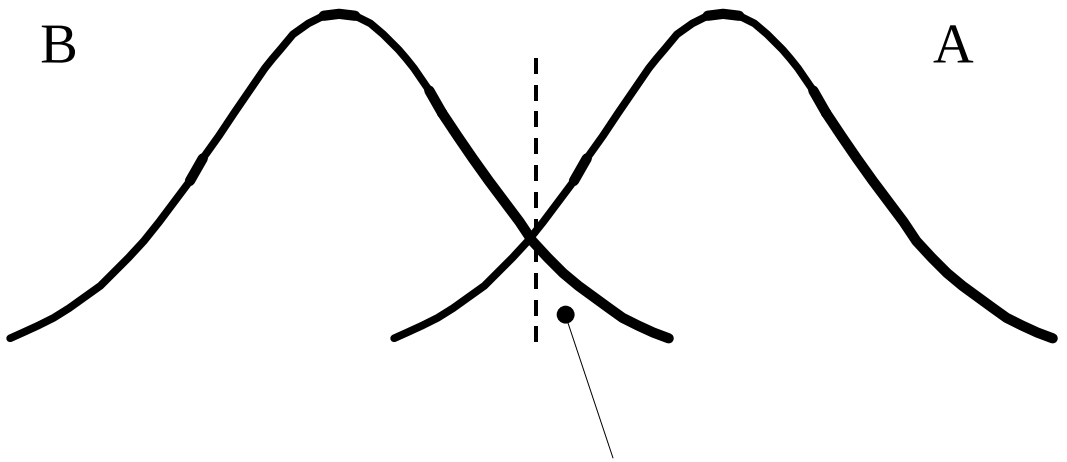
- Retrieve the most active item
- If you can recognize the target, and
  the retrieved item is not it, and
  there's time to try again, then
  attempt retrieval again
- Else stop and output item to next process
- *Prediction*: Error and latency should both increase with interference

Data from Glaser and Glaser (1989)



Target unknown

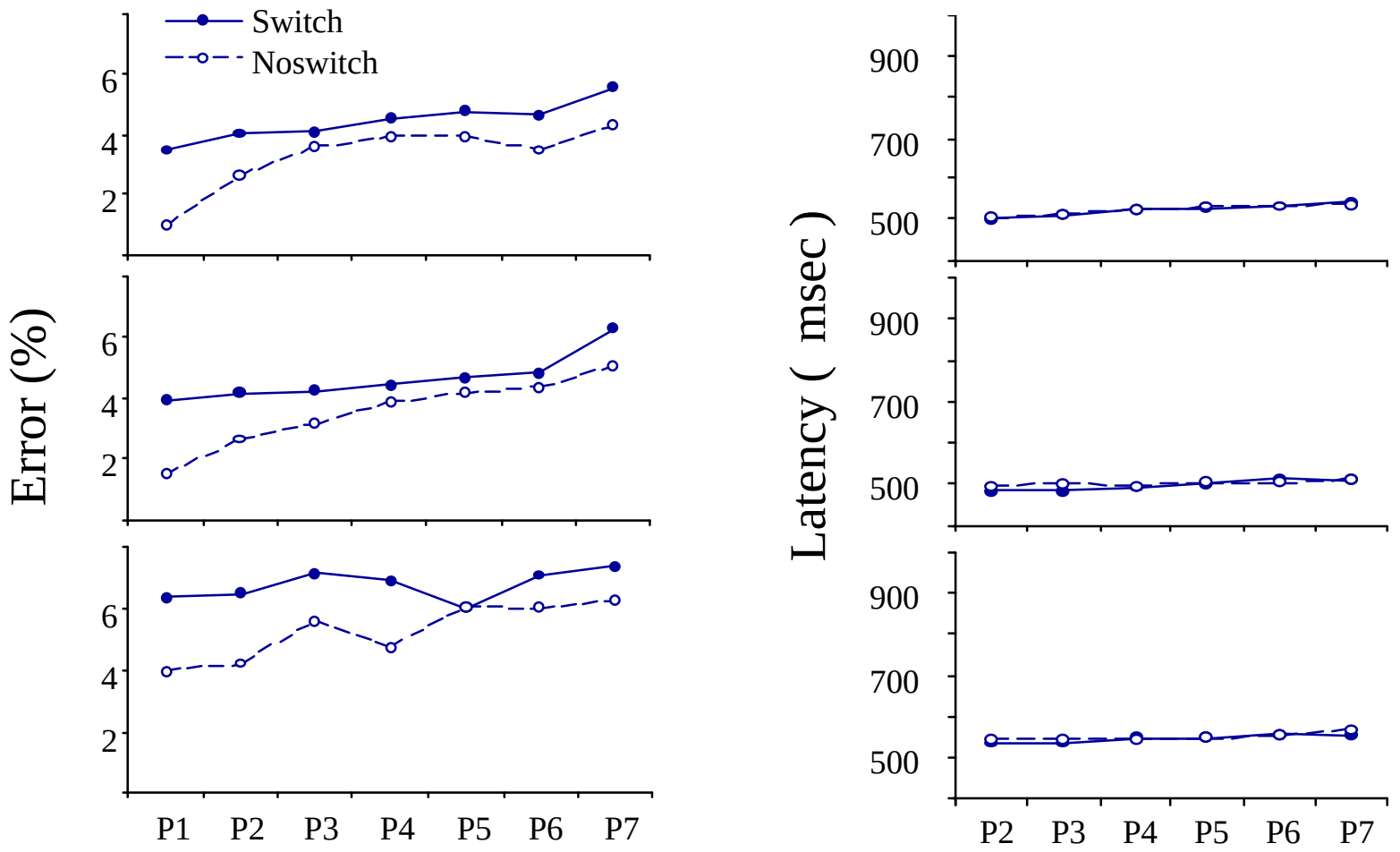
AaaaaaaBbbbbbAaaaaaaAaaaaaa ...



A retrieval process

- Retrieve the most active item
- If you can recognize the target, and
  the retrieved item is not it, and
  there's time to try again, then
  attempt retrieval again
- Else stop and output item to next process
- *Prediction*: Error but not latency should increase with interference

Target unknown



Comments

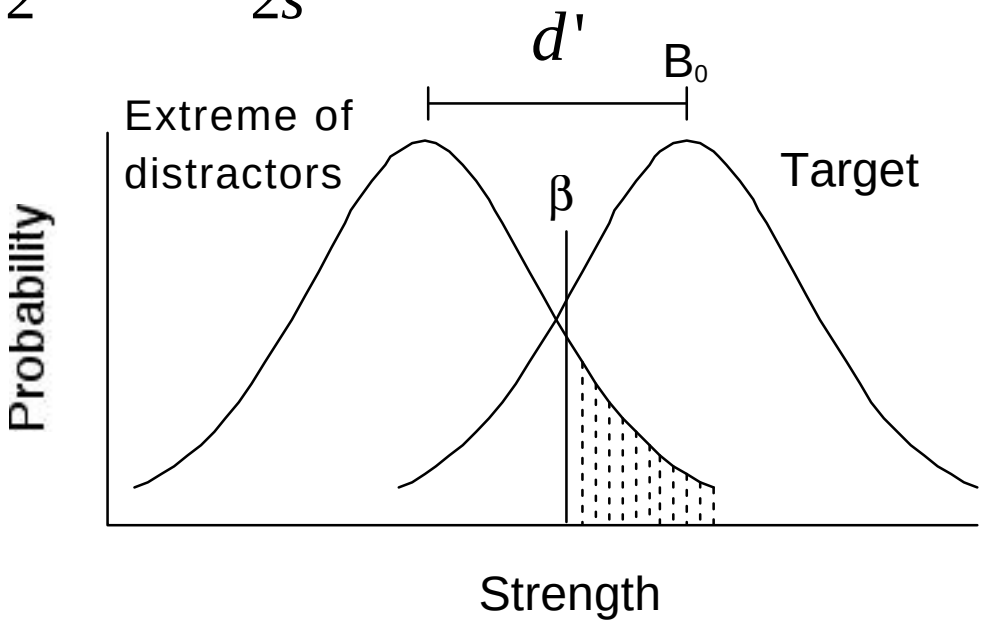
- Competitive latency for analytical models
- A retrieval process for process models
  - Do the math
  - Do distributional analysis

How to compute activation?

$$B = \ln(n) - \frac{\ln(t)}{2} - s\ln(\frac{1}{2s} - 1)$$

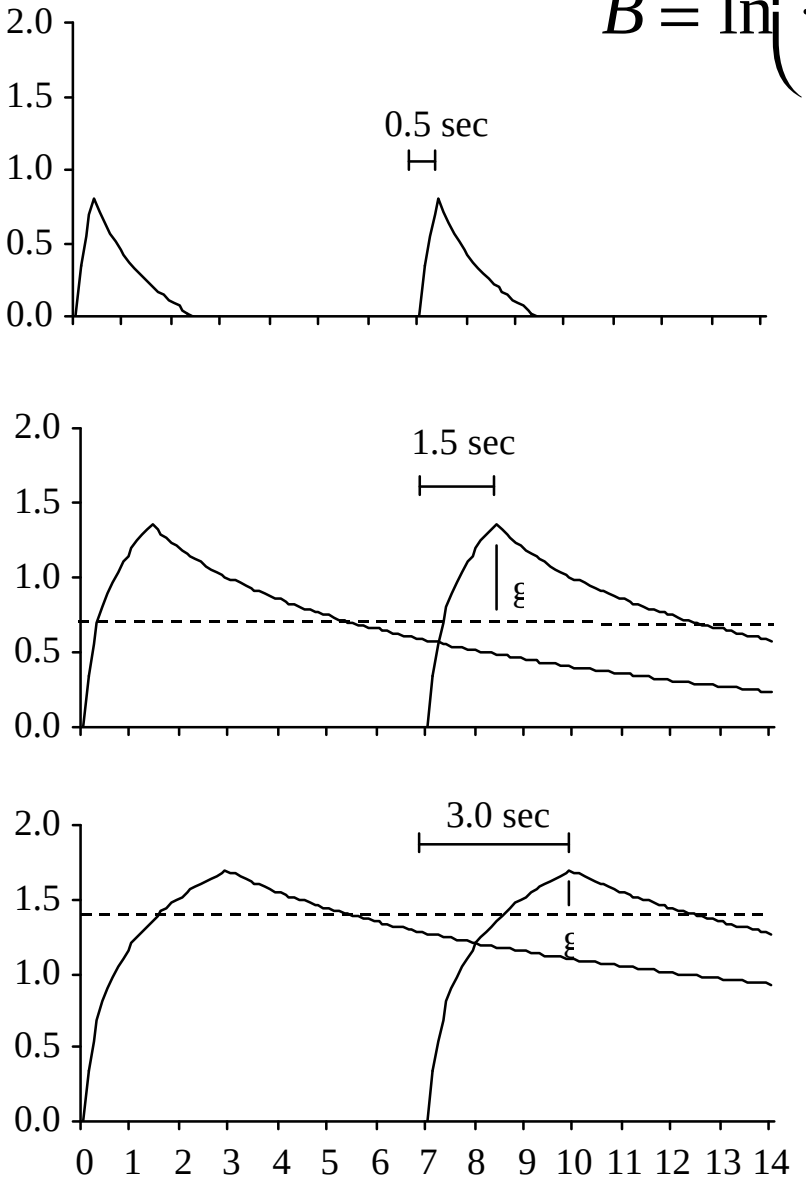
$$B = \ln\left(\frac{n}{\sqrt{T}}\right)$$

plus an instance representation



$$B = \log\left(\sum_{i=1}^{m} t_{-i} + \frac{(n-m)(t_0^{1-d} - t_{n-m}^{1-d})}{(1-d)(t_0 - t_{n-m})}\right)$$

$$B = \ln\left(\frac{n}{\sqrt{T}}\right)$$



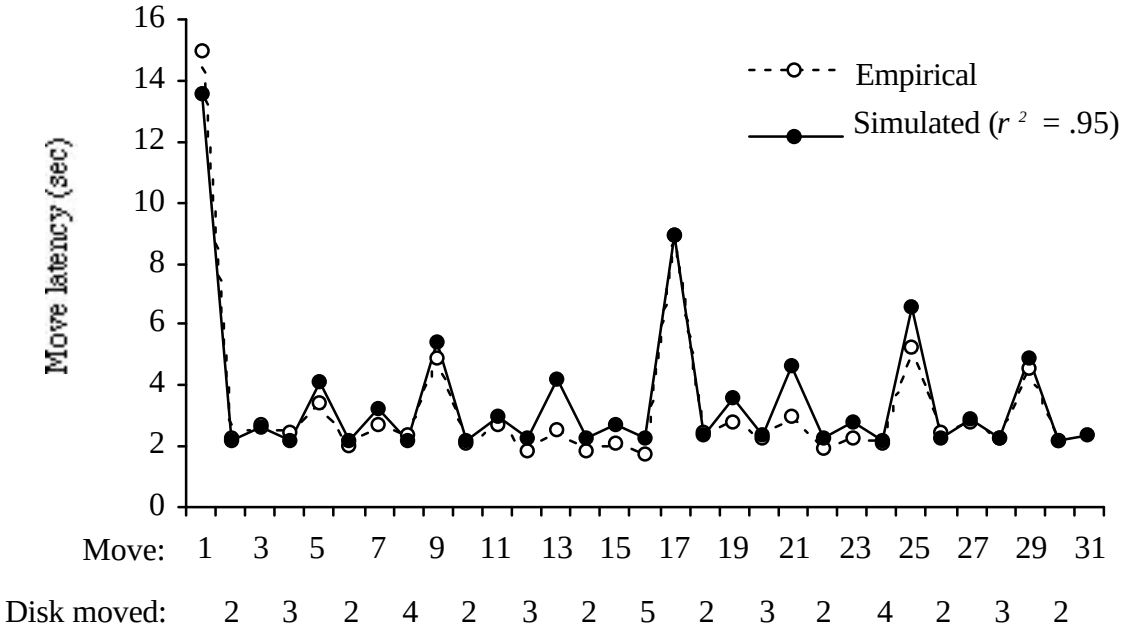
Implications:

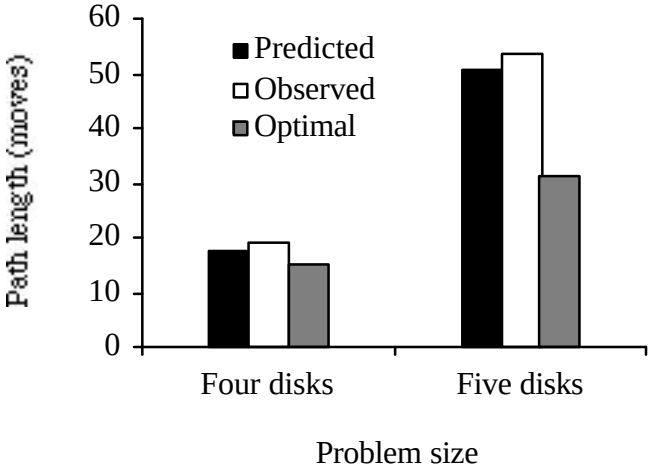
Short-term sensitivity

Encoding time predictions

PAS is unnecessary

Data from Anderson et al (1993)





Comments

- Optimized learning may be the better model
  - Computationally, analytically, pedagogically tractable
  - More accurate
- Instance-based representation has other useful implications
  - Time to strengthen an instance