ANCHOR:

A Memory-Based Model of
Category Rating

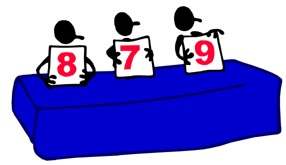
Alexander Petrov
John R. Anderson

Department of Psychology
Carnegie Mellon University

1

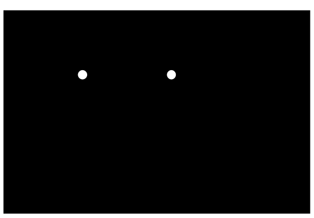
The Category-Rating Task

- “1”, “2”, …, “7”
- “very dissimilar”, …, “very similar”
- “strongly disagree”, …, “strongly agree”
- …



2

A Typical Setting



3

Motivation

- ubiquity of category rating
- organization of the “psychological continuum”
- reveals the dynamics of cognition

4

Plan of the Talk

- empirical phenomena
- psychophysical experiment
- ANCHOR model
- evaluation of the model
- conclusions

5

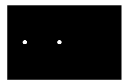
Empirical Phenomena

- Stevens’ power law
- sequential effects
- context effects
- memory effects
- range effects
- edge effects
- …

6

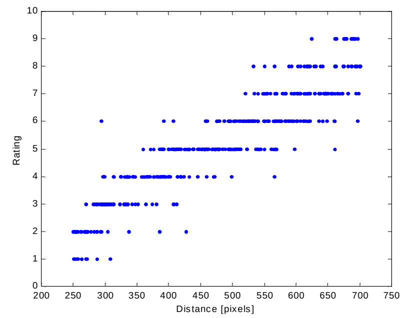
Experiment

- distances b/n 250 and 700 pixels
- 9-point scale
- 450 trials
- 17 demo trials; no feedback
- 48 participants
- randomized absolute position
- 4 seconds per trial, 30 min total



7

Typical Data Set



8

Two Main Steps

- perceive the stimulus
- report the subjective percept

S → perceptual subsystem → *M* → cognitive subsystem → *R*

9

ANCHOR Model

- stimuli represented as *magnitudes*
- response categories represented as *anchors*
- memory-based
- continuous learning
- based on the ACT-R architecture

10

Anchors

- magnitude-label associations
- serve as prototypes
- availability

Label	Magn.	Avail.
"1"	0.103	0.567
"2"	0.214	1.208
"3"	0.297	2.091
"4"	0.402	1.445
"5"	0.521	0.382

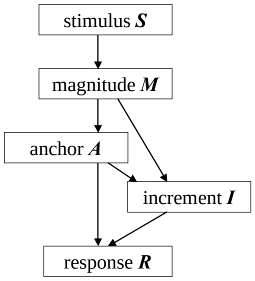
11

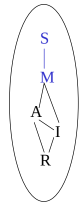
Talk-aloud Protocol

I see the dots…
The distance looks like a 7…
No, it's too short for a 7; I'll give it a 6.

12

Dependencies among Variables

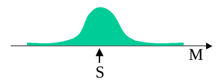




Perceptual Equation

$$M = k \cdot S \cdot \big(1 + \underbrace{\varepsilon}_{\text{multiplicative noise}}\big)$$

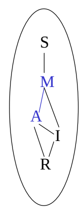
Each stimulus S defines a whole distribution of magnitudes.



Relation to Psychophysical "Laws"

$$M = k \cdot S \cdot (1 + \varepsilon)$$

- Stevens' law: $\boldsymbol{M = k.S^n}$
- Weber's law: ***dS/S = const***
- Ekman's law: ***dM/M = const***



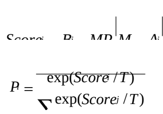
Anchor Selection

$$Score_i = B_i - MP.\left|M - A_i\right|$$

$$P_i = \frac{\exp(Score_i / T)}{\sum_j \exp(Score_j / T)}$$

Anchor Selection Highlights

- *general memory* mechanism
- stochastic (*softmax* rule)
- depends on the *similarity* b/n the target and each of the anchors
- depends on the *availability*
  - recency
  - strength



Connectionist Interpretation

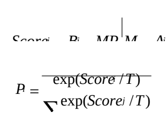
- *pattern completion* in an attractor network

OR

- *winner-takes-all* cluster

OR

- Kohonen network



ACT-R Interpretation

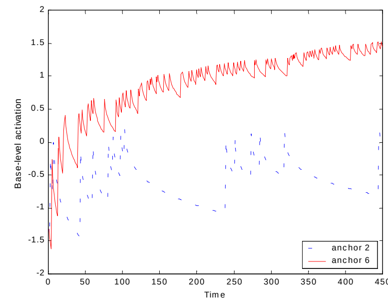
- each anchor is a *chunk*
- retrieval via *partial matching*
- *base-level activation* determines availability

$$Score_i = B_i - MP.|M - A_i|$$

$$P_i = \frac{\exp(Score_i / T)}{\sum_j \exp(Score_j / T)}$$

19

Dynamic Availability



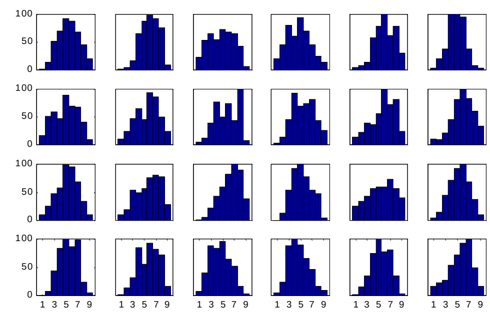
20

Sequential Effects

- assimilation towards previous response R_{t-1}
- contrast with previous stimulus S_{t-1}
- interaction between the two

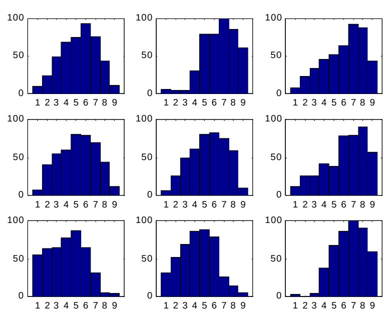
21

Category Strength (Human data)



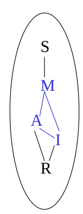
22

Category Strength (Model)



23

Correction Mechanism



- explicit strategy
- discrepancy reference points [-2*d*, -*d*, 0, *d*, 2*d*]

$$CorrScore_k = \left| d_k - (M - A) \right|$$

k = [-2,-1,0,1,2]

24

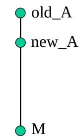
Question

The response has
been produced.
Is this the end of
the trial?

25

Updating Anchor Magnitudes

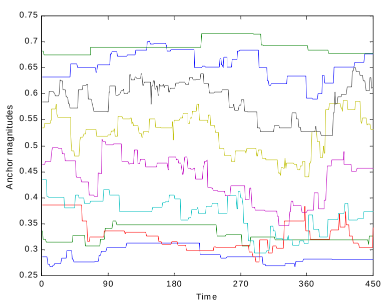
$$new\_A = \alpha.M + (1-\alpha).old\_A$$



- a form of competitive learning
- anchors act as *prototypes*
- consistency of responses
- similarity among the category labels
- redistribution of strength
- no need for extensive initial practice

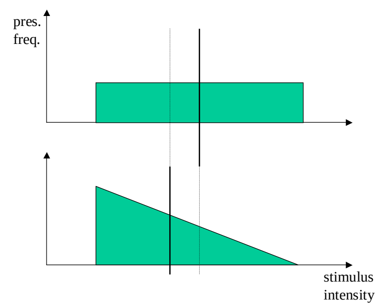
26

Dynamics of Anchor Magnitudes



27

Context Effects



28

Parameter Search

29

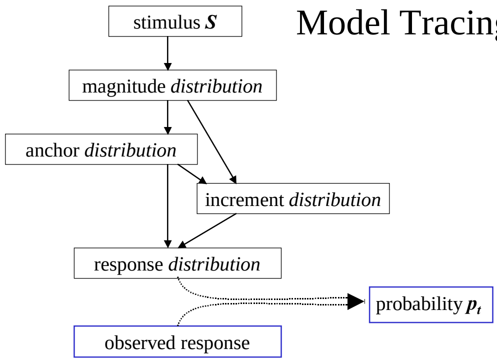
5+1 Parameters

- anchor temperature
- increment temperature
- mismatch penalty
- alpha
- correction reference point *d*

- perceptual noise

30

Model Tracing



31

Maximize the Log Likelihood

$$L = -\sum_{t=1}^{450} \ln[p_t]$$

where p_t is the probability that the model produces on trial t the response that was produced by the human subject

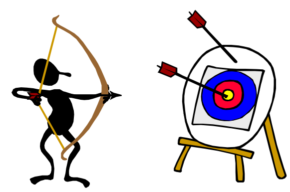
32

Parameter Search

- Gradient descent on L
- Individual parameter set optimized for each subject
- Reliability of the method tested on synthetic data
- Sensitivity analyses
- Idea: stepwise elimination

33

Testing the Model



34

Simulation Experiment

- 48 Ss → 48 parameter sets
- 48 stimulus sequences; the same that were shown to the human subjects
- 48 x 300 runs of the model
- the basic unit of analysis is the subject
- suite of statistical measures

35

Model Fits

Statistic	**empir**	**model**	R^2
Accuracy (R^2)	0.78	0.78	.86
Transm. info (T)	2.28	2.22	.92
Mean response	5.43	4.86	.11
Standard deviation	1.89	1.91	.83
Entropy (informtn)	2.86	2.81	.88
ACF(residuals)	0.32	0.20	.33
Context eff (d)	-.40 ; +.21	-.51 ; +.50	.55
Context eff (B_{osc})	-.12 ; +.15	-.18 ; +.14	.57

36

Future Work

- Half-split and rand-permutation validation
- Experiment with “lesions” of the model
- Compare with instance-based formulation
- Cover more empirical phenomena: range effects, number of categories
- Extend to absolute identification
- Extend to response times

37

The End

38

Base-Level Activation

$$B = \ln\left[\sum_{l=1}^{n} t_l^{-d}\right]$$

$$B = \ln\left[t^{-d} + \frac{n.\left(T^{1-d} - t^{1-d}\right)}{(1-d).(T-t)}\right]$$

$$B = \ln\left[t^{-.5} + \frac{2n}{(\sqrt{T} + \sqrt{t})}\right]$$

39