Modeling lexical access in ACT-R

Jakub Dotlačil — University of Amsterdam
Adrian Brasoveanu — UC Santa Cruz
www.github.com/jakdot/conferences

1. Goal

Modeling lexical decision (LD) tasks in ACT-R:
the effect of frequency on reaction times & accuracies

- Data: Experiment of [2]
- Data explained by [2] through the Rank hypothesis; but they note that the effect of frequency on RTs could also be modeled by a power function

  $P = At^{-d}$ (P – performance; t – time; A, d – free params)

- Power function implemented in ACT-R, so could ACT-R model data?

Why relevant?

- ACT-R models processing & retrieval in processing ([1, 3, 4])
- Such models – complex, but retrieval underlyingly power law with standard free parameters
- Our contributions:

  More direct evidence for ACT-R retrieval in language

  More direct evidence for *which free parameters* should be used and *what values* they have (different from previous assumptions)

2. Murray et al. (2004)

LD task for 5-7 letter words from 16 frequency bands:

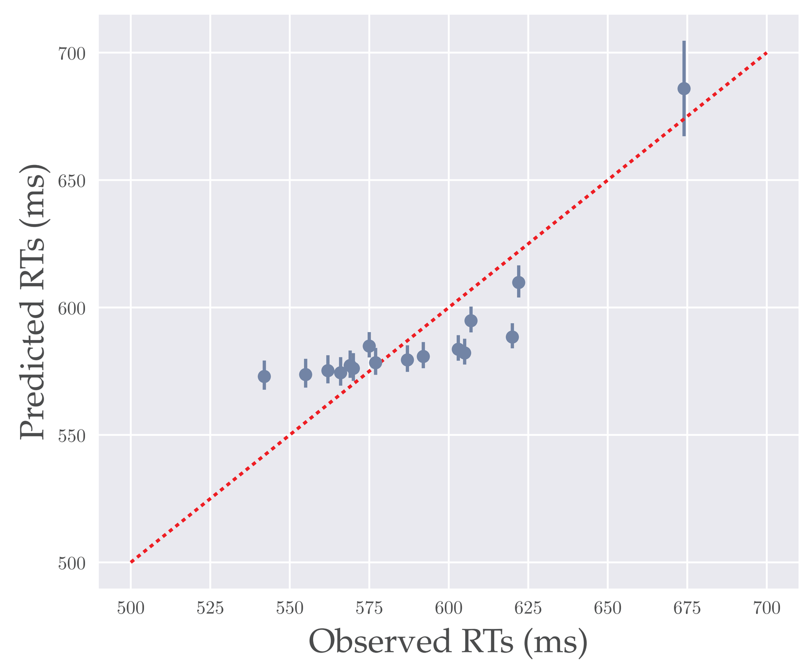
- highest frequency – 315 per 1 million
- lowest frequency – 1 per 1 million

Best fit – Power law, Rank hypothesis

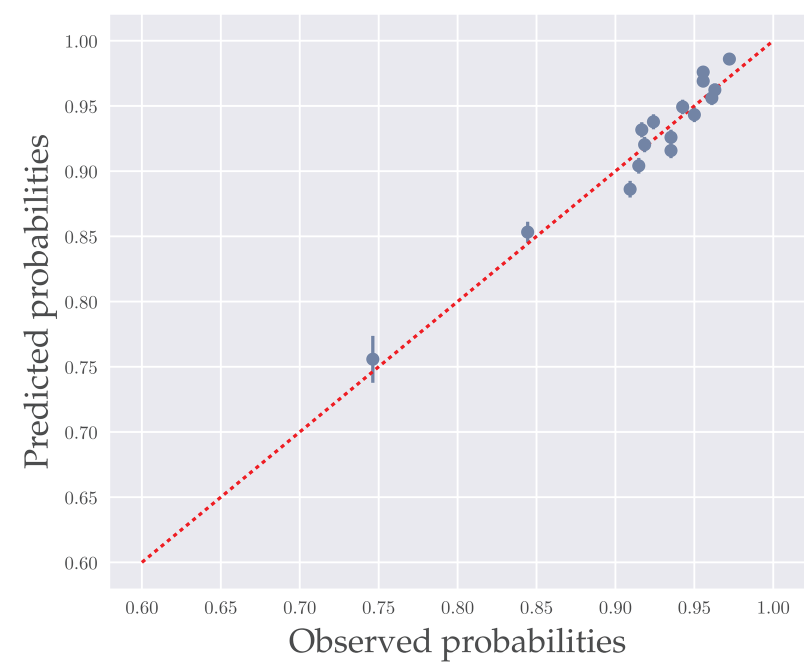
3. Modeling lexical decision task

1. Bayesian modeling using Python and pymc3
   Link between Accuracy and RTs given mappings in ACT-R (see box)
2. Evaluated directly against the data
3. Evaluated in a full ACT-R model that simulates the data, including visual and motor actions

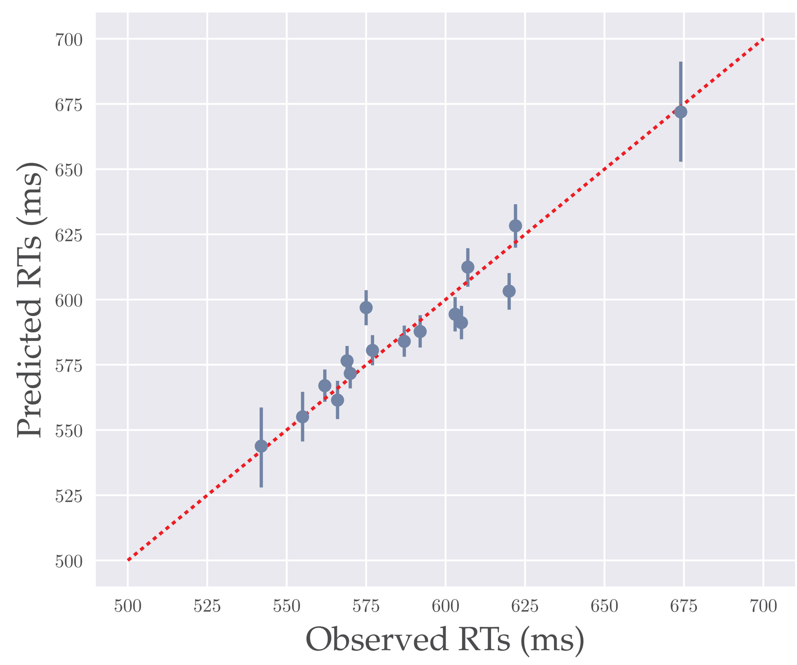
Reaction times (without f)



Accuracies (with/without f)



Reaction times (with f)



Estimates

- $f = 0.28[0.06 - 0.48]$
- $F = 0.45[0.1 - 0.86]$
- $d = 0.1[1e^{-6} - 0.24]$
- $\alpha = 0.5[0.3 - 0.56]$
- $\tau = 0.9[-1.9 - 2.9]$
- $s = 1.77[1.6 - 1.9]$

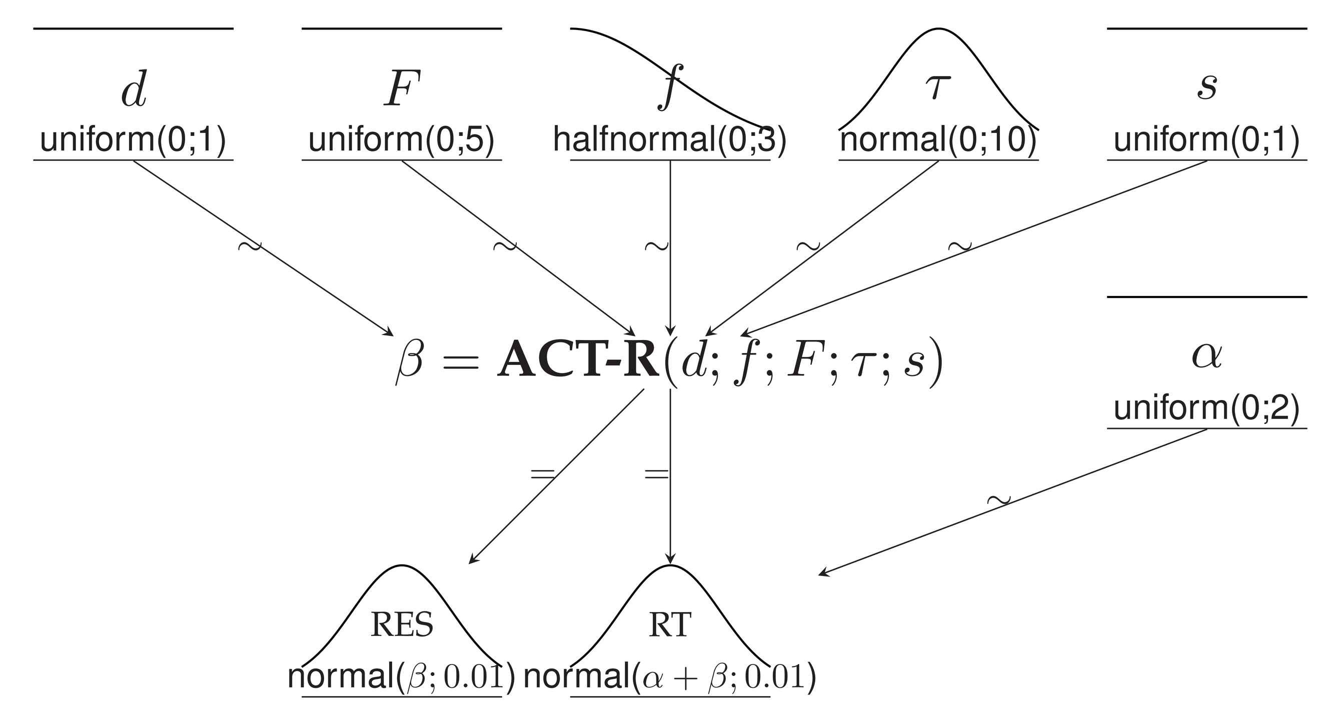
Relevant mappings in ACT-R

$$\text{Activation: } A_i = \log\left(\sum_{k=1}^{n} t_k^{-d}\right) \; (d: \text{decay}; n: \text{rehearsals}) \quad (1)$$

$$\text{Retrieval prob.: } P_i = \frac{1}{1+e^{-\frac{A_i-\tau}{s}}} \; (s: \text{noise}, \tau: \text{threshold}) \quad (2)$$

$$\text{Latency: } T_i = Fe^{-fA_i} \; (F, f: \text{latency factor and exp.}) \quad (3)$$

Bayesian model (with and without f)



- ACT-R can model the role of frequency in lexical decision tasks very well
- the params d, τ, s needed to model accuracies
- the latency exponent f is essential for modeling RTs, but all psycholinguistic ACT-R models approximate retrieval latencies by manipulating only the F parameter (see [1, 3, 4], a.o.; cf. [5])
- using statistical techniques to estimate ACT-R params explicates what sub-symbolic properties are needed
- this contrasts with the standard practice (arbitrary+default values), which obscures the ingredients of the model/the comparative quality of fit

[1] Lewis, R., and S. Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *CogSci* 29:1–45. ∗ [2] Murray, W. S, and K. I Forster. 2004. Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review* 111:721. ∗ [3] Reitter, D., F. Keller, and J. D. Moore. 2011. A computational cognitive model of syntactic priming. *CogSci* 35:587–637. ∗ [4] van Rij, J. 2012. *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults*. Groningen. ∗ [5] West, R., A. Pyke, M. Rutledge-Taylor, and H. Lang. 2010. Interference and ACT-R: New evidence from the fan effect. In *Proceedings of the 10th International Conference on Cognitive Modeling*, ed. D. D. Salvucci and G. Gunzelmann, 211–216. Philadelphia, PA: Drexel University.