Modeling lexical access in ACT-R

Adrian Brasoveanu (UC Santa Cruz) & Jakub Dotlačil (ILLC, University of Amsterdam)
abrsvn@gmail.com

I. Introduction. One very robust parameter affecting latencies and accuracies in lexical decision tasks is frequency. Frequency effects have been found not just in lexical decision tasks, but in many if not all tasks that involve some kind of lexical processing (Forster, 1990; Monsell, 1991). Since Howes and Solomon (1951), it is accepted that lexical access can be approximated as a log-function of frequency. However, log-frequency provides a good but not perfect fit to the data. In their detailed study of (log-)frequency, Murray and Forster (2004) (M&F) collected responses and response times in a lexical decision task using words from 16 frequency bands (their Exp. 1) and argued that their findings support a specific retrieval mechanism, the Rank Hypothesis. They note that frequency effects could also be modeled as skill learning (basically, practiced memory retrieval), which is commonly assumed to be a power function of time, in the same way that memory performance is (Anderson 1982 a.o.). Skill learning is implemented in the ACT-R cognitive architecture (Anderson and Lebiere 1998 a.o.), so it is a natural testing ground for modeling the lexical decision task and the rich data set in M&F. This detailed ACT-R modeling of lexical decision is one of the two contributions of this paper. The second is that we use general-purpose Bayesian modeling libraries to model the ACT-R account of lexical decision data: the `Python3` libraries `pymc3` and `theano`. This enables us to provide a general way of estimating crucial ACT-R parameters that are usually set by hand.

II. Modeling lexical access in ACT-R. In ACT-R, the base activation of a lexical item is a power-function of the time since n-many practice trials / 'rehearsals' of a word – (1). The activation is used to compute accuracy and latency for retrieval – (2) and (3). The free parameters are enumerated in parentheses.

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

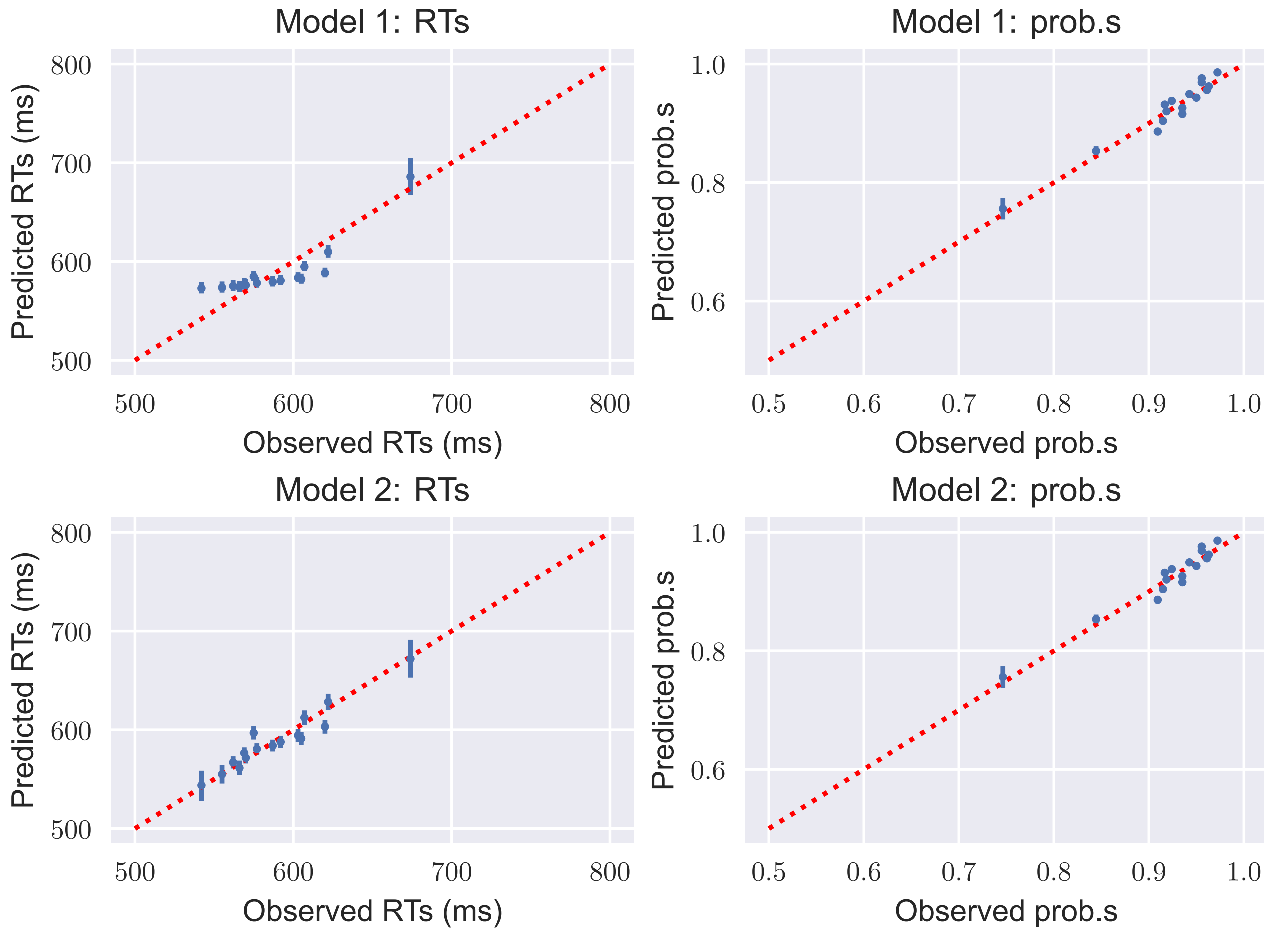
(2) $P_i = \frac{1}{1+e^{-\frac{A_i-\tau}{s}}}$ (s: noise, τ: threshold)

(3) $T_i = Fe^{-fA_i}$ (F:factor, f: exponent)

For any word, the number of rehearsals that contribute to its base activation was determined by its frequency and the number of words per year average humans are exposed to (estimated using Hart and Risley 1995). Two models were built for the M&F data, the crucial difference between them being whether the latency exponent f was ignored, i.e., set to its default value of 1 as it is in most ACT-R models (Model 1), or was estimated (Model 2). Both models also had a latency intercept, meant to capture the latency of the operations other than actual memory retrieval necessary to complete the task, e.g., motor or visual operations, encoding in the goal or imaginal buffers etc.

Model 1 fitted retrieval accuracy well, but latencies were poorly modeled – see the top two panels in the figure on the next page. In contrast, Model 2 had a good fit for both accuracy and latency – see the bottom two panels in the same figure. The plots show posterior estimates (on the y axis) against the observed values (on the x axis). The blue points plot observed data against predictions; the blue segments provide 95% credible intervals (CRIs) for the predictions; the diagonal red lines show the subspace where observations and predictions are identical (perfect fit). The paper provides full details of the estimated Bayes+ACT-R models, including the low-information priors we used as well as the posterior means and 95% CRIs for all the ACT-R parameters.

III. Conclusion. The results show that ACT-R can be used to model the role of frequency in lexical decision tasks very well, and that frequency effects can be understood as skill learning / practiced memory retrieval. The fact that Model 1 fails in modeling latencies, unlike Model 2, makes an important point for previous psycholinguistic ACT-R work (Lewis and Vasishth 2005 a.o.): the latency exponent f is essential, but to the best of our knowledge, all psycholinguistic ACT-R models approximate retrieval latencies by manipulating only the F parameter (this is true for ACT-R models in general, not only in psycholinguistics; West et al. 2010 is a notable exception). Our results cast doubt on the received way of modeling retrieval in ACT-R psycholinguistic models, or at the very least show that different parameters might modulate different cases of retrieval in language processing.



References

Anderson, John R. 1982. Acquisition of cognitive skill. Psychological review 89:369.

Anderson, John R., and Christian Lebiere. 1998. The atomic components of thought. Hillsdale, NJ: Lawrence Erlbaum Associates.

Forster, Kenneth I. 1990. Lexical processing. In Language: An invitation to cognitive science, ed. Daniel Osherson and Howard Lasnik, 95–131. Cambridge, MA: MIT Press.

Hart, Betty, and Todd R Risley. 1995. Meaningful differences in the everyday experience of young american children.. Baltimore: Paul H Brookes Publishing.

Howes, Davis H, and Richard L Solomon. 1951. Visual duration threshold as a function of word-probability. Journal of experimental psychology 41:401.

Lewis, Richard, and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science 29:1–45.

Monsell, Stephen. 1991. The nature and locus of word frequency effects in reading. In Basic processes in reading: Visual word recognition, ed. D. Besner and G. W. Humphreys, 148–197. Hillsdale, NJ: Erlbaum.

Murray, Wayne S, and Kenneth I Forster. 2004. Serial mechanisms in lexical access: the rank hypothesis. Psychological Review 111:721.

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.