

Quantitative Comparison for Generative Theories: Embedding Competence-Performance Linguistic Theories into Bayesian Models

Adrian Brasoveanu (UC Santa Cruz) & Jakub Dotlačil (ILLC, University of Amsterdam)

I. Overview We introduce a new framework in which (i) qualitative generative grammar hypotheses can be embedded into performance/processing models, and (ii) these integrated competence-performance models are embedded into Bayesian models, which enables us to fit them to experimental data and do quantitative model comparison for qualitative theories.

Following in the footsteps of Lewis and Vasishth (2005), step (i) will take generative grammar theories and embed them in processing theories formulated in the ACT-R cognitive architecture (Anderson and Lebiere 1998). This will enable us to model the behavior of human participants in standard experimental tasks like lexical decision, forced-choice, self-paced reading or eye-tracking tasks. This embedding of competence theories into performance ACT-R models is computationally implemented in a new Python3 library.

This reimplementaion of ACT-R enables us to do step (ii), namely interface competence-performance linguistic theories (in the form of ACT-R models) with standard statistical estimation methods. In particular, we use ACT-R models as the likelihood component of full Bayesian models, and fit the model parameters to experimental data. We are also able to consider alternative models, that is alternative generative grammar theories, and quantitatively compare how well they fit experimental data collected with standard methodologies. The ability to do quantitative comparison for generative theories on this scale is unprecedented, as far as we know. Even in ACT-R, subsymbolic/quantitative parameters are usually set by hand instead of estimated from the data using standard statistical methods.

II. Case study We focus on a simple case study – the lexical decision task in Murray and Forster (2004) (M&F) – to highlight how different symbolic/qualitative models can be devised and quantitatively compared. The data we use is the response accuracy and response times for words from 16 frequency bands reported in M&F’s Exp. 1. We focus on one limited model comparison here, but the talk/paper will consider a wider range of models and compare them based on the M&F data. Building on the ACT-R theory of declarative memory, we take the activation of a lexical item i to be a power-function of the time since n -many presentations of a word – (1). The activation is used to compute accuracy and latency of word retrieval – (2) and (3) (the free parameters are enumerated in parentheses).

$$(1) \quad A_i = \log \left(\sum_{k=1}^n t_k^{-d} \right) \quad (d: \text{decay})$$
$$(2) \quad P_i = \frac{1}{1 + e^{-\frac{A_i - \tau}{s}}} \quad (s: \text{noise}, \tau: \text{threshold})$$
$$(3) \quad T_i = F e^{-f A_i} \quad (F: \text{factor}, f: \text{exponent})$$

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 in (3) 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). Model 1 fitted retrieval accuracy well, but latencies were poorly modeled. In contrast, Model 2 had a good fit for both accuracy and latency.

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 (latency factor) parameter.

References

- Anderson, John R., and Christian Lebiere. 1998. *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hart, Betty, and Todd R Risley. 1995. *Meaningful differences in the everyday experience of young american children..* Baltimore: Paul H Brookes Publishing.
- Lewis, Richard, and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:1–45.
- Murray, Wayne S, and Kenneth I Forster. 2004. Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review* 111:721.