Goldilocks meets the subset problem: Evaluating Error Driven Constraint Demotion for OT language acquisition

Joshua Tauberer
University of Pennsylvania

The Error Driven Constraint Demotion (EDCD) algorithm for learning Optimality Theory grammars (Tesar & Smolensky 2000) was perhaps abandoned in the literature too early, not because it doesn't suffer from the subset-problem-type problems explored in the literature (e.g. Tesar and Smolensky 2000, Smith 2000), but because there is something still to be learned from a careful look at where EDCD fails and how extensive the failure is. A Markov model approach, in the spirit of Niyogi and Berwick's (1996) work in parametric learning spaces, is used to characterize the complete behavior of a learner employing EDCD. This is done for all possible target languages resulting from the permutation of eight constraints proposed in the literature for several phonological alternations.

The EDCD learning algorithm is a memoryless OT grammar learning algorithm in which the learner maintains a single hypothesis for what the target grammar is. The learner begins with an 'initial' hypothesis grammar. After each input, here a word from the target language, the learner may update his hypothesis by demoting one or more constraints. This procedure is carried out indefinitely, but the learner is guaranteed to eventually settle on a hypothesis. The part of EDCD called Robust Interpretive Parsing plays an important role in the learnability results. EDCD relates to UG in that the space of possible initial hypotheses must be restricted so that certain constraint orders can be learned (e.g. Smith 2000).

The constraints used in the simulations described in this paper related to an oral/nasal vowel alternation in English (*V_{oral}N, *V_{nas}, IDENT-IO$_{V}$(NAS)), the voiced/unvoiced alternation of the English plural suffix (*VOICE, *UNVOICEDVOICED, IDENT-IO(VOICE)), and consonant denasalization that occurs in other languages (*NC̥, IDENT-IO$_{C}$(NAS)).

We show that not all of the languages from permuting these constraints can be learned under the algorithm, given plausible assumptions about UG. (A language we say is learnable from an initial hypothesis if the learner has chosen with high probability any grammar which has the same extension (yield) as the target grammar. We made several assumptions about how to compute the yield of an OT grammar, based on assuming all underlying representations are possible and equally probable.) It is not a new finding that a language may be unlearnable, but a complete characterization of the learner's behavior has not been presented before, including: time to convergence to the *correct* grammar, the number of intermediate hypotheses the learner passes through, what proportion of languages are learnable, and which initial states make all languages learnable. We find, for instance, that no single fully stratified hierarchy in which all markedness constraints dominate all faithfulness constraints — a bias proposed in the past to circumvent the subset problem — is a possible initial state from which a learner could acquire any target language.

From here there are two possibilities. The unlearnable grammars though predicted by OT may not correspond to attested languages, and this may be evidence that the EDCD algorithm is on the right track. However, from the languages that are known to exist, we can eliminate grammars as possible initial states when the language cannot be acquired from an initial state. Hopefully we will be left with a space of initial states that could plausibly be encoded in UG.

References

Niyogi, Partha and Robert C. Berwick. 1996. A language learning model for finite parameter spaces. *Cognition*, 61:161-193.

Smith, Jennifer L. 2000. Positional Faithfulness and Learnability in Optimality Theory. In Rebecca Daly and Anastasia Riehl, eds., Proceedings of ESCOL 99. Ithaca: CLC Publications, 203-214.

Tesar, Bruce and Paul Smolensky. 2000. Learnability in Optimality Theory. MIT Press.