Learning Local Phonological Rules

Jane Chandlee and Cesar Koirala

Department of Linguistics and Cognitive Science, University of Delaware

This paper presents a learning algorithm for local phonological processes that is predicated on a tight computational characterization of these processes, meaning that it exploits a restriction on the expressive power needed to compute phonological patterns that apply locally. This restriction reflects the well-recognized idea that locality plays an important role in phonology (Kenstowicz 1994, Gafos 1996, among others). Both SPE-style rules and OT constraints encode locality via adjacency of the segments affecting and undergoing phonological change (with the possible exception of tonal and harmony patterns, which may still be 'local' at a more abstract level). This research thus makes both a technical and a linguistic contribution. The technical contribution is showing how Strictly Local (McNaughton and Papert 1971) sets (which have been shown by Heinz (2007) to describe phonotactic constraints defined over sequences of a finite length *k*) can be generalized to functions so as to describe local phonological processes. The linguistic contribution is to demonstrate the utility of a learning model that makes explicit use of this requirement of locality.

Oncina et al. (1993) present OSTIA, an algorithm that identifies the class of subsequential functions in the limit from positive data. Subsequential functions are those describable with subsequential transducers, which are deterministic weighted acceptors in which the weights are strings that are concatenated instead of multiplied. "Being subsequential" appears to be a necessary property of phonological processes (Chandlee et al. 2012, Gainor et al. 2012, Chandlee and Heinz 2012), but not a sufficient one, since many logically possible but phonologically bizarre processes are also subsequential. Similar to OSTIA, our learner first constructs a prefix tree transducer (see Figure 1), which is a finite representation of the input underlying/surface pairs; then the learner generalizes to words not found in the data set by merging states. Unlike OSTIA, the learner's criterion for state merging reflects a key property of the SL class – namely suffix substitution closure (see Rogers and Pullum 2011). Specifically, the learner merges states that have the same *input* suffix of length *k*-1. The algorithm was tested on a corpus of 51,723 underlying/surface pairs derived from the CELEX2 German lemma corpus by 'unapplying' the final devoicing rule to those forms that orthographically end with a voiced obstruent. The resulting FST is equivalent to the machine that describes the devoicing rule (see Figure 2). OSTIA, on the other hand, will not learn such rules without modification (Gildea and Jurafsky 1996). Furthermore, our learner learns only a *subclass* of the subsequential functions that are targeted by OSTIA, namely the class hypothesized to be delimited by the locality requirement of phonological processes. By using a key property of the SL class our learner directly reflects the hypothesis space and thus suggests a way in which humans generalize to learn phonological patterns. In turn, if humans do generalize in this way, we have an explanation for why local phonological patterns have this property.

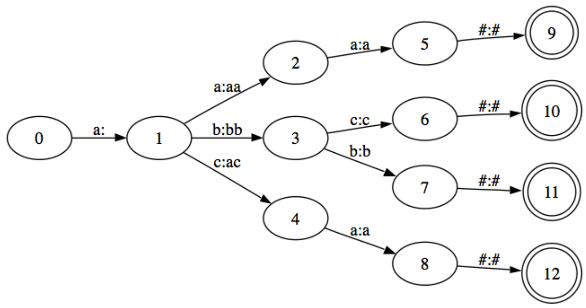

Figure 1. Prefix tree transducer for the set {(aaa,aaa), (abc,bbc), (abb,bbb), (aca,aca)}, which reflects the rule a → b / __ b

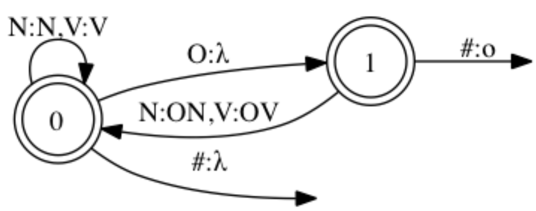

Figure 2. Target FST that describes the rule of German final devoicing (O = voiced obstruent, o = voiceless obstruent, N = sonorant consonant, V = vowel, # = word boundary)

Selected references

Chandlee, Jane and Jeffrey Heinz. (2012). Bounded copying is subsequential: implications for metathesis and reduplication. *12th Meeting of SIGMORPHON.*

Chandlee, Jane, Angeliki Athanasopoulou, and Jeffrey Heinz. (2012). Evidence for classifying metathesis patterns as subsequential. *29th WCCFL.*

Gainor, Brian, Regine Lai, and Jeffrey Heinz. (2012). Computational characterizations of vowel harmony patterns and pathologies. *29th WCCFL.*

Gildea, Daniel and Daniel Jurafsky. (1996). Learning bias and phonological-rule induction. *Computational Linguistics, 22*(4), 497-530.

Heinz, J. (2007). The inductive learning of phonotactic patterns. Ph.D. thesis, UCLA.

McNaughton, Robert and Seymour A. Papert. (1971). *Counter-Free Automata*. MIT Press.

Oncina, José, Pedro García, and Enrique Vidal. (1993). Learning subsequential transducers for pattern recognition interpretation tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence 15*(5), 448-457.

Rogers, James and Geoffrey K. Pullum. (2011). Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language, and Information 20*(3), 329-342.