Efficiently Computing OT Typologies

We present a set of algorithms and a software package designed to create, manipulate, and extract information from sets of OT tableaux. The practical aim of this tool is to aid in the management and visualization of complex bodies of data. The theoretical goals of this project are to establish that it's possible to efficiently map all the ranking arguments derived from a set of tableaux to an equivalent Minimal set of Ranking Arguments (MRA) without losing information and to determine how the size of the least minimizable MRA relates to |CON|. Minimization is necessary because, as we show, building typologies from stratified hierarchies is infeasible for large models.

Minimizing ranking arguments

To represent ranking arguments derived from OT tableaux we use Prince's (2002) Elementary Ranking Conditions (ERCs). Every ERC-set corresponds to a unique MRA that encodes exactly the same information. Formally, the MRA for *R* is the set *S* generated via transitive closure of the *fusion* operation (by which inferences are derived from groups of ranking statements) coupled with the elimination of any ERC in *S* that is entailed by another ERCs in *S*. The naïve approach to building an MRA involves searching through the power set of *R*, which is computationally infeasible for even moderate collections of tableaux. Instead, we build *S* by doing pairwise fusion of ERCs in *R* while discarding less informative (entailed) ERCs. This algorithm is quite fast, rarely needing more than a second to minimize ERC-sets in our test suite of 340 tableaux of 6,298 candidates over ten constraints.

The necessity of MRAs

With ten constraints, 28,501 ERCs are consistent with each ranking yet the largest MRA has only 189 (a 150-fold reduction). With thirty constraints the ratio is a million to one. To build typologies, we collect candidates across tableaux into sets with compatible ERCs. Stratified hierarchy building (Tesar and Smolensky 1993) is excellent for checking whether a set of candidates is compatible, but without minimization, the ERC-sets associated with the hypotheses can grow far too large to feasibly work with.

Maximum MRA cardinality for *k* constraints: $m = \sum_{x=0}^{\lfloor k/4 \rfloor} \binom{k-1-x}{k-\lceil k/2 \rceil - 2x}$

Though MRA construction is efficient in |*R*|, *m* still grows exponentially with *k*. This has serious ramifications for learning and typology-building because it will be impossible to perfectly store and manipulate all ERCs when *k* is large.

An online tool for checking analyses and generating typologies

To make our MRA-generating and typology-building algorithms widely available we created an open source web-based tool that allows users to create, load, manipulate, and save tableaux in a format compatible with other packages like OTSoft (Hayes et al. 2003). Our software checks/generates ranking arguments, detects harmonic bounding, generates (partial) rankings from (partial) sets of winners (and vice versa) and generates images of tableaux, Hasse diagrams, and stratified hierarchies that can be dragged directly from the browser into a word processor. Finally, by running the software on a central server we hope to create a forum for phonologists to discuss and directly share constraints, tableaux, and data for OT analyses.

word count: 499