Maximal feature specification is feasible; minimal feature specification not so much

Kyle Gorman (Graduate Center, CUNY) & Charles Reiss (Concordia University)

Georgian has a standard five-vowel system and two laterals in complementary distribution. Plain (or clear) [l] occurs before the front vowels [i, e], and the velarized [ɫ] occurs elsewhere.

(1) Georgian (Robins and Waterson 1952):

leɫo ‘goal’ k’b̥i**ɫ**s ‘tooth (dat.)’ ts^{h}o**li** ’wife’ **ɫ**amazad̥ ‘prettily’

(2) $ɫ \rightarrow l \,/\, \_\_ \{i, e\}$

If we wish to specify (2) more formally, then how general or concise should the rule be? Clearly the rule should not be so general as to convert /ɫ/ to [l] in all contexts. But even if we restrict ourselves to rules that produce the appropriate output, should the right context of the structural description target [−BACK], [−BACK, −LOW], [−BACK, −LOW, −ROUND], or an even more fully specified set of segments? There is no empirical way to decide, since Georgian has no [−BACK] vowels that are not also redundantly specified [−LOW], [−ROUND], [−NASAL], and so on. This is not a purely philosophical matter: it is a problem for a child acquiring Georgian just as it is a problem for the linguist.

Our impression is that most phonologists would prefer the minimal feature specification consistent with the data, and would teach students to do the same. *SPE* (Chomsky and Halle 1968:§8.1), for instance, proposes a feature-counting evaluation metric which favors the most concise—in terms of feature specifications—extensionally adequate grammar. However, minimal feature specification was challenged by Hale and Reiss (2003, 2008) on logical grounds related to the subset principle and the non-uniqueness of minimal solutions (cf. Chen and Hulden 2018). The present paper provides a further argument against minimal, non-redundant feature specification on the basis of computational tractability, and shows that one simple alternative is feasible.

Cognitive scientists have argued that the hypothesis space for plausible computational-level theories of cognitive operations should be constrained to exclude computationally intractable algorithms. In the narrowest form (e.g., Frixione 2001), this excludes any algorithms which lack a polynomial-time solution—i.e., a solution which can be reached is bounded by a polynomial function of *n* where *n* is the size of the input to the algorithm—whereas others (e.g., van Rooij 2008) do not exclude supra-polynomial solutions when the size of the inputs is small and fixed, or if heuristics are generally successful. This *tractable cognition thesis* appears to have been adopted—albeit implicitly—by phonologists. For instance, witness the debate as to whether one can compute an optimal candidate in Optimality Theory in polynomial time (see, e.g., Heinz et al. 2009); similarly Heinz (2010) and Chandlee et al. (2014) emphasize that their algorithms for acquiring long-distance phonotactic generalizations and phonological mappings, respectively, are polynomial time.

Chen and Hulden (2018) recently proved that selection of the minimal specification of natural classes is a computationally intractable problem—it is $\mathcal{NP}$-complete—which under standard assumptions means that there is no polynomial time algorithm for feature minimization, whether features are binary (equipollent) or privative (univalent). Chen and Hulden further show that greedy and branch-and-bound heuristics fail to find the minimal feature specification even for simple problems (e.g., such as finding the minimal specification for a single phone). Chen and Hulden’s results are a serious problem for proposals to perform phonological acquisition by searching for rules and representations which minimize the description length (e.g., Rasin et al. 2021). We respond by showing that determining the maximal feature specification is feasible via (generalized) intersection, which we now illustrate for a natural class with two members, as in the trigger of our Georgian rule.

Without loss of generality, let us assume binary feature specification, and consider the problem of computing a feature specification consistent with a set of two segments s and t. Let F be a universal, finite feature set. A binary feature specification is a pair (α, f) such that $\alpha \in \{+, -\}$ and $f \in F$. Let S, T be the full feature specifications, both of size $|F|$, for s and t respectively. Any candidate feature specification for the set $\{s, t\}$ will necessarily be a subset of $S \cap T$ where $\cap$ denotes intersection.

Algorithm 1 Intersection of binary feature sets $R = S \cap T$.

$R \leftarrow \emptyset$
for $(\alpha, f) \in S$ **do**
  if $(\alpha, f) \in T$ **then**
    $R \leftarrow R \cup \{(\alpha, f)\}$
  end if
end for

This algorithm produces the maximal feature set in $O(|F|)$, which is polynomial time and therefore tractable under even the most stringent definition. We note this algorithm can also be used to iteratively compute the generalized intersection over more than two sets—as intersection is associative—or can be trivially modified to operate over privative features.

However, as shown by Chen and Hulden, the only way to compute the minimal feature specification is via brute-force search, i.e., by enumerating all subsets of $S \cap T$, of which there are $2^{|S \cap T|}$. Even under modest assumptions about the size of F, there may be hundreds of thousands of possible solutions to consider. Furthermore, there is no guarantee the solution will be unique. We submit: why bother? Children need only to compute the maximal specification via intersection, which is empirically adequate and can be computed in polynomial time.

References

Chandlee, Jane, Rémi Eyraud, and Jeffrey Heinz. 2014. Learning strictly local subsequential functions. *Transactions of the Association for Computational Linguistics* 2:491–503.

Chen, Hubie, and Mans Hulden. 2018. The computational complexity of distinctive feature minimization in phonology. In *Proceedings of NAACL-HLT 2018*, 542–547.

Chomsky, Noam, and Morris Halle. 1968. *Sound Pattern of English.* Harper & Row.

Frixione, Marcello. 2001. Tractable competence. *Minds and Machines* 11:379–397.

Hale, Mark, and Charles Reiss. 2003. The subset principle in phonology: why the tabula can't be rasa. *Journal of Linguistics* 39:219–244.

Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise.* Oxford University Press.

Heinz, Jeffrey. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41:623–661.

Heinz, Jeffrey, Gregory M. Kobele, and Jason Riggle. 2009. Evaluating the complexity of Optimality Theory. *Linguistic Inquiry* 40:277–288.

Rasin, Ezer, Iddo Berger, Nur Lan, Shefi, and Roni Katzir. 2021. Approaching explanatory adequacy in phonology using Minimum Description Length. *Journal of Language Modelling* 9:17–66.

Robins, R. H., and Natalie Waterson. 1952. Notes on the phonetics of the Georgian word. *Bulletin of the School of Oriental and African Studies* 141:55–72.

van Rooij, Iris. 2008. The tractable cognition thesis. *Cognitive Science* 32:939–984.