A COMPARISON BETWEEN OT AND HG FROM A COMPUTATIONAL PERSPECTIVE

GIORGIO MAGRI

Optimality Theory (OT) and Harmonic Grammar (HG) differ because the former assumes a model of constraint interaction based on strict domination, while the latter assumes a weighted model of interaction. As Prince and Smolensky (1997) admit, "that strict domination governs grammatical constraint interaction is not currently explained". Yet, Legendre et al. (2006, 911-912) make two suggestions. The first suggestion is that OT's strict domination might have *algorithmic advantages*, in the sense that it "may enable quick-and-dirty optimization algorithms [...] to consistently find a single global [...] optimum, whereas arbitrarily weighted constraints typically lead such algorithms to produce widely varying solutions, each only a local optimum." The second suggestion is that OT's strict domination might have *learnability advantages*: "another possibility is that demands of learnability provide a pressure for strict domination among constraints", although they note that "it remains an open problem to formally characterize exactly what is essential about strict domination to guarantee efficient learning."

Both conjectures have been challenged in the recent literature. Pater (2009) compares OT and HG from an *algorithmic perspective*, and reaches the opposite conclusion. He advocates "the replacement of OT's ranked constraints with [HG's] weighted ones" based on the fact "that the resulting model of grammar [...] is compatible with well-understood algorithms for learning and other computations" and states that "the strengths of HG in this area are of considerable importance" (p. 1002). Furthermore, Riggle (2009) and Bane et al. (2010) compare OT and HG from a *learnability perspective*, using tools from Statistical Learning Theory. They show that the two frameworks pattern alike w.r.t. a classical measure of learning complexity, namely they have the same *Vapnik-Chervonenkis (VC) dimension*. And they thus conclude that, "though there may be factors that favor one model over the other, the complexity of learning [...] is not one of them". These recent papers show that computational phonology has entered a mature stage, characterized by stronger connections with the neighbouring field of Machine Learning.

Yet, In this talk, I challenge both of these recent conclusions, thus vindicating the initial conjecture of OT's optimality. I present a simple trick that allows algorithms for HG to be extended to OT. Thus, HG has no algorithmic advantages over OT, contrary to Pater's claim. Furthermore, I point out that the VC dimension is well known to be a rather course upper bound on learning-theoretic complexity (especially for the case of linear classifiers). And I build the case for a learnability advantage of OT over HG, based on some recent results in Koltchinskii et al. (2003b), Koltchinskii et al. (2003a) and Koltchinskii and Panchenko (2005).

References

Bane, Max, Jason Riggle, and Morgan Sonderegger. 2010. "The VC dimension of constraint-based grammars". *Lingua* 120.5:1194–1208.

Koltchinskii, Vladimir, and Dmitry Panchenko. 2005. "Complexities of convex combinations and bounding the generalization error in classification". *Ann. Statist.* 33.4:1455–1496.

Koltchinskii, Vladimir, Dmitry Panchenko, and Savina Andonova. 2003a. "Generalization bounds for voting classifiers based on sparsity and clustering". In *Lecture Notes in Artificial Intelligence 2777*, 492–505.

Koltchinskii, Vladimir, Dmitry Panchenko, and Lozano. 2003b. "Bounding the generalization error of convex combinations of classifiers: Balancing the dimensionality and the margins". *Ann. Appl. Probab.* 13:213–252.

Legendre, Geraldine, Antonella Sorace, and Paul Smolensky. 2006. "The Optimality Theory/Harmonic Grammar connection". In *The Harminic Mind*, ed. Paul Smolensky and Geraldine Legendre, 903–966. Cambridge, MA: MIT Press.

Pater, Joe. 2009. "Weighted Constraints in Generative Linguistics". *Cognitive Science* 33:999–1035.

Prince, Alan, and Paul Smolensky. 1997. "Optimality: From Neural Networks to Universal Grammar". *Science* 275:1604–1610.

Riggle, Jason. 2009. "The Complexity of Ranking Hypotheses in Optimality Theory". *Computational Linguistics* 35(1):47–59.