Title: Using the fusional closure to assist in learning ranking information

A well-known learnability problem both within an OT/HG framework and without is that a phonological learner must construct a lexicon that matches observed surface forms and must produce a grammar that maps those constructed lexical items to those observed surface forms. Adopting Optimality Theory narrows the problem: grammatical learning consists solely of ranking determination. Even with this narrowing, the learner must produce a lexicon concomitantly with a ranking. The logic of Optimality Theory has allowed researchers to begin to address the simultaneous nature of the task by positing a number of differing techniques for extracting ranking information from unspecified and underspecified forms (Merchant and Tesar 2006, Jarosz 2006, Apoussidou 2007, amongst others). These techniques are known to fail in a number of situations stemming from a number of causes, often related to the form of the information generated from the learning algorithms. For example in CPR (Merchant 2008) ranking information is produced using MRCD from not fully specified forms producing sets of ERCs that are missing entailed and crucial ERCs, gaps in the set, that cause the algorithm to fail to capture relevant ranking arguments.

In this poster I present a means of filling these gaps by, for a set of ERCs, producing the set's *fusional closure*. The fusional closure of a set is the *fusional closure* of the set in the mathematical sense: it is that set that is *closed* under the operation of fusion (Prince 2002), so that the result of fusing any two ERCs in the set is also in the set. This set is useful for learning for a number of reasons: a set and its fusional closure are mutually entailing, allowing the learner to compute over fusional closures instead of base ERC sets. Furthermore, the fusional closure of a set of ERCs has the property that every ERC entailed by the original set is entailed by exactly one ERC in its fusional closure.

This final property of individual entailment is shown to be necessary for extracting ranking information from sets of ERCs using an algorithm such as CPR. CPR works by generating sets of ERCs using MRCD, one for each possible lexicon for a pair of overt forms and determining the shared ranking information across these sets of ERCs using the join (Merchant 2011) to extract shared information. The algorithm, computing over MRCD produced sets, misses ranking information because there are entailed ERCs that are not entailed by individual ERCs in the ERC sets, but by combinations of ERCs. Computing over fusional closures shores up this lack of individual entailment.

The ideas presented here emphasize that the logic of the linguistic system plays a crucial role in learning. Determination of ranking information from lexical items occurs because there is an entailment relationship between sets of ERCs consistent with given lexical hypotheses. Gaps in the information produced can be filled using computational tools built up from the linguistic system itself. Open questions still obtain: how are these computational tools manifested in an ICS architecture and what constraints does an ICS system impose on learning theory? Even with these questions it is clear that the logic of the linguistic system can play a crucial role in linguistic, and possibly other types, of learning.

Keywords: Phonology, Learnability, ICS

References:

Apoussidou, Diana. 2007. The Learnability of Metrical Phonology. PhD. dissertation, University of Amsterdam, Amsterdam.

Jarosz, Gaja. 2006. Rich Lexicons and Restrictive Grammars - Maximum Likelihood Learning in Optimality Theory. PhD. dissertation, The Johns Hopkins University, Baltimore, MD. ROA-884.

Merchant, Nazarré. 2008. Discovering underlying forms: Contrast pairs and ranking. PhD. dissertation, Rutgers University, New Brunswick. ROA-964.

Merchant, Nazarré, and Tesar, Bruce. 2008. Learning underlying forms by searching restricted lexical subspaces. In *Proceedings of the Forty-First Conference of the Chicago Linguistics Society (2005), vol. II: The Panels*, 33-47. ROA-811.

Merchant, Nazarré, 2011. Learning ranking information from unspecified overt forms using the join. To appear in: *The Proceedings of the Forty-Seventh Conference of the Chicago Linguistics Society*.

Prince, Alan, 2002. Entailed Ranking Arguments. ROA-500.