Learning and restrictiveness with Properties

Introduction. This paper develops the hypothesis that giving a learner access to certain kinds of typological structural information leads to restrictiveness and efficiency gains in computational learning algorithms. While previous work in computational learnability in Optimality Theory (OT) has shown that the theory's logical structure can be appropriated in algorithms to extract useful information (Tesar 1995 et seq., Merchant 2011), the higher order typological organization has not been used in this way. Through Property Theory (PT; Alber & Prince in prep., Alber, DelBusso & Prince 2016, DelBusso 2018), the current proposal modifies the Output-Driven Learner (ODL; Tesar 2014) to show the potential advantages of doing so. In a computer simulation of the property-modified algorithm, all languages are successfully learned without the use of other methods in the original algorithm for enforcing restrictiveness.

Restrictiveness in learning and OT. Grammars in a typology can differ in degree of *restrictiveness*: a more restrictive grammar neutralizes a potential contrast and thus restricts the kinds of forms that can appear at the surface relative to a less restrictive grammar where the contrast is maintained. A learner of a more restrictive language faces the 'subset problem' of ruling out the superset grammar even though the data is consistent with it.

In OT, markedness constraints (M), violated by certain features or structures, enable neutralization, and faithfulness constraints (F) enable contrast by preserving underlying features. For example, the markedness constraint m.NoLongV is violated by long vowels in outputs, while faithfulness constraint f.Length is violated in change of vowel length between input and output. If m.NoLongV dominates f.Length, the language neutralizes vowel length contrast, and has a subset of the output forms of a language that preserves the contrast, generated under the reverse ranking.

This long-recognized correlation between M>F rankings and degree of restrictiveness has been used in learning methods such as Biased Constraint Demotion (BCD; Prince & Tesar 2006), a variation of Recursive Constraint Demotion algorithm (RCD; Tesar 1995) that favors low ranking of faithfulness constraints. BCD implicitly encodes M>F rankings as a bias. ODL commits to ranking information by accumulating Elementary Ranking Conditions (ERCs) from data, and BCD invokes restrictiveness by choosing among hierarchies consistent with the ERCs, but does not store any restrictiveness commitments. As a result, the restrictiveness bias is not available for other learning tasks—such as finding underlying forms—that use the committed ranking information encoded in ERCs.

ODL. Tesar's (2014) ODL simultaneously learns a grammar (rankings) and lexicon (underlying input feature values). Using the error-driven method Multi-Recursive Constraint Demotion (Tesar 1997), the learner builds and maintains a support of ERCs that represent the current grammar. Concurrently with learning a ranking, underlying feature values of the lexicon are set using inconsistency detection and paradigmatic comparison. The ODL employs two main methods aimed at restrictiveness: BCD, and Fewest Set Features (FSF), a method that enforces restrictiveness in the lexicon. ODL is developed using a stress/length system called Paka. In a computer simulation, all 24 languages in the typology are successfully learned. The FSF method is required only to learn two left/right-mirror image languages that are paradigmatic subsets of other languages in the typology.

Learning with properties: PODL. The modified algorithm develops the hypothesis that the learner can make efficiency and restrictiveness gains by taking certain M>F rankings as ERCs at the onset of learning. This is similar to Smolensky's (1996) proposal that in the initial state structural markedness constraints rank above faithfulness constraints. However, in a typology, rankings aligning with restrictiveness involve *specific* M and F constraints rather than unstructured blocks of all. Determining these crucial rankings is non-trivial, but is precisely the kind of ranking information identified by a Property Analysis (PA) in PT. A PA analyzes an OT typology into a set of *properties* that describe conflicts between two sets of constraints. Each property value defines an ERC set and represents a key ranking relation holding in the grammars of the typology. The full set of property value combinations defines all languages of the typology. The PODL algorithm uses a subset of the properties that involve conflicts between M and F constraints, correlated with restrictiveness.

PODL modifies ODL in three central ways. First, rather than an empty support, the learner begins with a set of property value ERCs set to M>F values; these are labeled as PERCs. Second, it adds a method for retracting these PERCs from the support when they are inconsistent with the learning data. Unlike other ranking information evidenced by the data, PERCs are *defeasible* and can be removed in the case of inconsistency. Finally, PODL replaces BCD with unbiased RCD throughout the algorithm.

PODL was tested in a computer simulation of the same Paka system. All languages are successfully learned. The two languages requiring FSF in the original algorithm are learned without this method, as the information from the PERCs rules out the superset grammar in a way that the BCD bias cannot. Additionally, the PERCs reduce the number of learning steps for some languages that have a high degree of neutralization. In the most restrictive languages, the PERCs are maintained; in less restrictive grammars, the retraction method is called one or more times to remove the inconsistent PERCs.

Summary. By tapping into a level of theoretical structure not previously exploited in OT learning algorithms, the PODL algorithm results in successful learning of all grammars without BCD or FSF, which have known limitations with more complex systems (Prince & Tesar 2006, Tesar 2014, Moyer & Tesar 2019). Additionally, the final supports for some grammars learned with PODL contain more explicit ranking restrictions than the corresponding supports in ODL that rely on BCD. In this way, the use of properties and typological structure enables complete learning of grammars.

References. Alber, B. DelBusso, N. & Prince, A. 2016. From intensional properties to Universal Support. *Language: Phonological Analysis* 92(2): e88-116. // Alber, B. & Prince, A. In prep. *Typologies*. Ms, University of Verona & Rutgers University. // DelBusso, N. 2018. Typological structure and properties of Property Theory. PhD diss. Rutgers University. // Merchant, N. 2011. Learning ranking information from unspecified overt forms using the join. *Proceeding of CLS 47*, 261-275. // Moyer, M. & Tesar, B. 2019. Enforcing restrictiveness through ranking induction in the Output-Driven Learner. 93rd Annual Meeting of the Linguistic Society of America, New York, Jan. 3-6. // Prince, A. & Tesar, B. 2006. Learning phonotactic distributions. Tech. report, Rutgers University. // Smolensky, P. 1996. The initial state and 'Richness of the Base' in Optimality Theory. Tech. rep. JHU-CogSci-96-4. // Tesar, B. 1995. *Computational Optimality Theory*. PhD dissertation, University of Colorado, Boulder. // Tesar, B. 1997. Multi-Recursive Constraint Demotion. Tech. report, Rutgers University. // Tesar, B. 2004. Using inconsistency detection to overcome structural ambiguity in language learning. *Linguistic Inquiry* 35.2: 219-253. // Tesar, B. 2014. *Output-Driven phonology: Theory and learning*. Cambridge: Cambridge University Press.