Computing Vowel Harmony

The Generative Capacity of Search and Copy

Samuel Andersson, Hossep Dolatian, and Yiding Hao

samuel.andersson@yale.edu,hossep.dolatian@stonybrook.edu,yiding.hao@yale.edu

Objectives & Introduction

Phenomenon: *Vowel harmony*, a common long-distance process
Theory: *Search & Copy*, a theory of long-distance dependencies where unspecified segments search for feature values
Question: What is the weak generative capacity of Search & Copy?
Results:

1. For unidirectional harmony, Search & Copy can only implement tier-based input strictly local (TISL) functions.
2. This improves Gainor et al.'s (2012) subsequential bound on unidirectional harmony processes.
3. Bidirectional Search & Copy is not TISL, and the composition of multiple harmony processes is not TISL in general.

Vowel Harmony

Traditionally, vowel harmony is thought to have the following properties.

- **Donor to Target:** Underlying forms contain underspecified vowels, which obtain feature values from fully-specified *donors*. We represent underspecified vowels with uppercase letters (e.g., A and I).
- **Directionality:** Features spread in one *direction*; i.e., donors for a particular harmony process either occur always to the left or always to the right of the underspecified target vowel.
- **Long-Distance:** Donor vowels and underspecified target vowels are often not adjacent to one another.

Example: Turkish backness harmony

	'house'	'horse'	'address'
SG	/ev/ ⇝ [ev]	/at/ ⇝ [at]	/adres/ ⇝ [adres]
PL	/ev-lAr/ ⇝ [ev-ler]	/at-lAr/ ⇝ [atlar]	/adres-lAr/ ⇝ [adres-ler]
PL.LOC	/ev-lAr-A/ ⇝ [ev-ler-e]	/at-lAr-A/ ⇝ [at-lar-a]	/adres-lAr-A/ ⇝ [adres-ler-e]

The Turkish plural and locative suffixes contain the vowel A, which carries the feature bundle $[\circ\text{back}, -\text{high}, -\text{round}]$. Since this vowel is underspecified for $[\pm\text{back}]$, it receives its backness value from the nearest preceding vowel.

Search & Copy

Search & Copy (Nevins, 2005, 2010; Mailhot & Reiss, 2007; Samuels, 2009a,b) is an approach to vowel harmony focusing on phonological computation.

- Underspecified segments *search* for a *target* matching *feature specification* F in a *direction* specified by a parameter δ.
- The underspecified segment receives the donor's feature values if a suitable donor is found.

Example: Turkish /adres-lAr-A/ → [adres-ler-e] 'at the addresses'

- $\delta = \leftarrow$, $F = [\pm\text{back}]$
- Begin with /adres/.
- Merge the plural suffix to get adres-lAr.
- /A/ searches for a target and finds /e/. /A/ then becomes [e].
- Merge the locative suffix to get adres-ler-A.
- /A/ searches for a target, finds /e/, and becomes [e].
- The result is [adres-ler-e].

Tier-Based Input Strictly Local Functions

Assume an alphabet Σ, a tier $\Gamma \subseteq \Sigma$, and a number $k > 0$. Let $\tau : \Sigma^* \to \Gamma^*$ be the homomorphism given by $\tau(\sigma) = \sigma$ if $\sigma \in \Gamma$ and $\tau(\sigma) = \varepsilon$ otherwise.
Definition. A function $f : \Sigma^* \to \Sigma^*$ is *k-input strictly local on tier* Γ if f is computed by a subsequential finite-state transducer (SFST) such that if $q_0 \xrightarrow[w]{y} q_1$ and $q_0 \xrightarrow[x]{z} q_2$ and $\#^k\tau(w)$ and $\#^k\tau(x)$ share a suffix of length at least $k-1$, then $q_1 = q_2$.

TISL Implementation of Search & Copy

- We use an SFST that reads the input in the direction opposite to δ.
- The tier consists of all target segments matching the features F.
- Underspecified segments are given feature values based on the most recent target segment.
- The TISL implementation runs in time $O(n)$, whereas Nevins's (2010) description of the Search & Copy procedure runs in time $O(n^2)$.

Example: Turkish Backness Harmony

- Scan the underlying form from left to right, since $\delta = \leftarrow$.
- Segments specified for $[\pm\text{back}]$ are projected to the tier.
- Segments with $[\circ\text{back}]$ (i.e., /A/) are given the same value of $[\pm\text{back}]$ as that of the last segment on the backness tier.
- All other segments are mapped faithfully to the surface form.
- This harmony process is 2-TISL because mappings are determined by the $2 - 1 = 1$ most recent segment on the backness tier.
- Sample FST run for /adres-lAr-A/ → [adres-ler-e]:

$$\# \xrightarrow[a]{a} a \xrightarrow[d]{d} a \xrightarrow[r]{r} a \xrightarrow[e]{e} e \xrightarrow[s]{s} e \xrightarrow[l]{l} e \xrightarrow[A]{e} e \xrightarrow[r]{r} e \xrightarrow[A]{e} e$$

Bidirectional Search & Copy

The directionality parameter δ can specify that the search must occur in *both* directions ($\leftrightarrow$), and both donors must provide the same feature values to the underspecified target.
Example: Woleaian (Howard, 1972)

	'drinking object'
Independent form	/ülüm/ ⇝ [üːl]
1SG	/ülüm-A-ji/ ⇝ [ülümej]
3SG	/ülüm-A-la/ ⇝ [ülümal]

- The /A/ receives a value for $[\pm\text{high}]$ from a donor to the left and to the right.
- The two donors must be equidistant from the /A/, and they need to have the same value of $[\pm\text{high}]$.
- If the search fails, then the /A/ receives a default value of [a].

Nevins (2005): "…when a target vowel is **equidistant** from two vowels, the **closest** source is **both** of them…"

- Detecting equidistance is **not finite-state**.
- If donors are *not* required to be equidistant from the target, then the resulting harmony process is weakly deterministic, but not TISL.
- Thus, Bidirectional Search & Copy exceeds the TISL functions in weak generative capacity.

Multiple Harmony Processes in One Language

Languages can have multiple *different* harmony processes.
Example: Turkish has both backness and rounding harmony.

	'stalk'			'end'		
SG	/sap/	⇝	[sap]	/son/	⇝	[son]
PL	/sap-lAr/	⇝	[sap-lar]	/son-lAr/	⇝	[son-lar]
GEN	/sap-In/	⇝	[sap-in]	/son-In/	⇝	[son-un]
PL.GEN	/sap-lAr-In/	⇝	[sap-lar-in]	/son-lAr-In/	⇝	[son-lar-in]

- The genitive suffix is /In/, where I $= [\circ\text{back}, +\text{high}, \circ\text{round}]$.
- In [son-un] 'of the end', the suffix value for round comes from the root.
- In [son-lar-in] 'of the ends', the suffix value for round comes from the plural: *[son-lar-un].

In Search & Copy, the two suffixes invoke separate harmony rules iteratively, as they are merged to the stem (Nevins, 2010).

- Plural backness harmony: $\delta = \leftarrow$, $F = [\pm\text{back}]$
- Genitive backness and roundness harmony: $\delta = \leftarrow$, $F = [\pm\text{back}, \pm\text{round}]$
- The /A/ and the /I/ conduct *separate* searches for a donor vowel.

The combination of roundness and backness harmony cannot be represented as a monolithic TISL function.

- The tier cannot contain /A/, since backness harmony needs to ignore it.
- However, rounding harmony requires /A/ to be on the tier, since its $[-\text{round}]$ feature can spread to /I/.
- The two harmonies can be represented as the composition of two separate TISL processes.

Conclusion and Future Work

We have shown that unidirectional Search & Copy, and all harmony processes analyzed within that framework, are TISL.

- Bidirectional Search & Copy, as described by Nevins (2005), is formally problematic due to its non-finite-stateness.
- In general, processes consisting of multiple Search & Copy rules cannot be compiled into a single TISL function.
- Can bidirectional harmony processes be analyzed without resorting to non-finite-state means?
- Are there formal restrictions on how Search & Copy rules may be combined with one another?
- How does Search & Copy compare formally to other accounts of vowel harmony, such as simultaneous and iterated rule-based accounts?

References

Gainor, Brian, Regine Lai & Jeffrey Heinz. 2012. Computational Characterizations of Vowel Harmony Patterns and Pathologies. In Proceedings of the 29th West Coast Conference on Formal Linguistics, 63–71. Somerville, MA: Cascadilla Proceedings Project.

Howard, Irwin. 1972. A Directional Theory of Rule Application in Phonology. Cambridge, MA: Massachusetts Institute of Technology PhD Dissertation.

Mailhot, Frédéric & Charles Reiss. 2007. Computing Long-Distance Dependencies in Vowel Harmony. BIOLINGUISTICS 1. 028–048.

Nevins, Andrew. 2005. Conditions on (Dis)Harmony. Massachusetts Institute of Technology PhD Dissertation.

Nevins, Andrew. 2010. Locality in vowel harmony, vol. 55. Mit Press.

Samuels, Bridget D. 2009a. Structure & specification in harmony. In Proceedings of the thity-eighth annual meeting of the North East Linguistics Society, vol. 2, 283–296. Amherst, MA: GLSA Publications.

Samuels, Bridget D. 2009b. The Structure of Phonological Theory. Cambridge, MA: Harvard University PhD Dissertation.