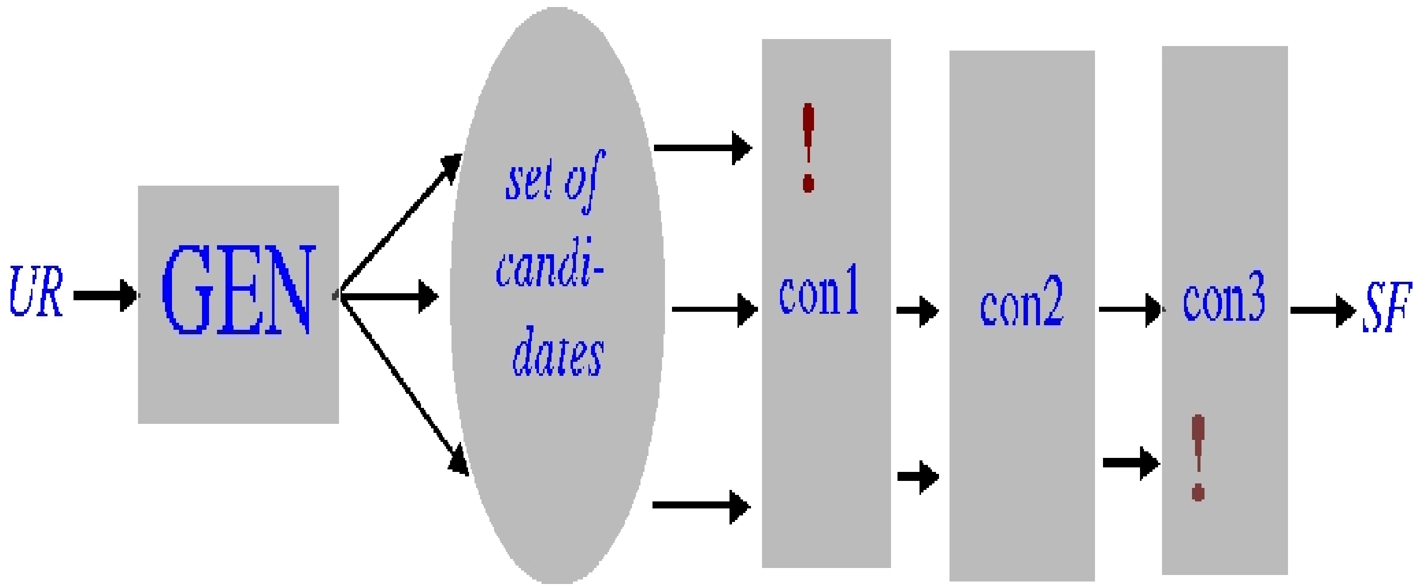
Quadratic Alignment Constraints and Finite State Optimality Theory

Tamás Bíró: `birot@let.rug.nl`

- Finite State approaches to Optimality Theory (phonology)
  (neither parallel, nor cascaded, but OT-like geometry for FS morphology)
- What kind of constraints can be encoded as a transducer?
- Example: metrical stress assignment in OT.
- Typology of constraints: some types cannot be encoded.

1. Optimality Theory



- Each constraint assigns violation marks to each candidate.
- Optimal wrt Con: no candidate has less violation marks.

2. *Feasibility of Finite State Optimality Theory*

```
Gen oo Con1 oo Con2 oo ...  oo ConN
```

- Formulate `Gen` as a FST.
- Encode constraints as FSTs.
- Formulate optimality operator `oo` turning constraints into filters.

Goal of this paper: What constraints can be encoded? Putting the emphasis mainly on phonology.

The approach to `oo` determines how to encode constraints.

3. Approaches to optimality operator: Counting approach

- Frank and Satta (1998), Karttunen (1998) (lenient composition)
- Distinguishing between 0 or 1 violation mark per word: define the **set of "absolute optimal" candidates**.

$$T_C := Ident_{\{w|w\in\Sigma^*, C(w)=0\}}$$

$$OT \textbf{ oo } C := (OT \textbf{ o } T_C) \cup (Ident_{\overline{domain(OT \textbf{ o } T_C)}} \textbf{ o } OT)$$

- Fixed k: 0, 1 ... k violation marks / word: series of filters

Approaches to optimality operator II.: Matching approach

- Create a set including the **“relative suboptimal candidates”**, identity transduction through its complement.
- Gerdemann and van Noord (2000):
  ∗ transducers for marking violations
  ∗ add extra violation marks (no max. no. of violations)
  ∗ permute violation marks (approximations)
- Jäger (2002) (generalized lenient composition)

T_C : maps a candidate to its suboptimal competitors

$$OT \ \mathbf{oo} \ C := OT \ \mathbf{o} \ Ident_{\overline{range(OT \ \mathbf{o} \ T_C)}}$$

4. Constraints are required thus to be encoded as...

- Regular expressions defining the non violating strings (Frank and Satta (1998), Karttunen (1998)).
- Transducers for assigning violation marks (Gerdemann and van Noord (2000)).
- Transducers mapping onto a set of suboptimal candidates (as well as non-competitors) (Jäger (2002)).

What kind of constraints
do we eventually encounter in OT?

5. A typology of constraints in OT

Cf. McCarthy (2002)

- Maximally 1 violation mark for each word
- Maximally 1 violation mark for each locus (substring)
- More violation marks for each locus (gradience): bounded or unbounded number of violations / locus.

Max. no. of viol. marks (@) assigned to a word σ can be:

- constant for any word
- proportional to the length of the word (upper-bounded linearly in the length of the word): $\exists k \in \mathbf{N}^+ : \#(@, \sigma) < k \mid \sigma \mid$
- growing more quickly: non-linear in the length of the word (non-linear constraints: e.g. unbounded number of viol. marks per locus).

6. Example: Metrical Stress assignment in OT

- Syllables parsed into feet. One foot is the head [main] foot:

$$\sigma(\sigma\ \sigma)[\sigma\ \sigma]\sigma(\sigma)$$

- Each foot has a head syllable, that will bear stress.
- Primary stress in the main foot, secondary stresses in non main feet:

$$\sigma(\sigma\ \sigma2)[\sigma1\ \sigma]\sigma(\sigma2)$$

- Gen produces all possible parses of a word.

7. Constraints for assigning stress: One violation mark per word

- Word-Foot-Left: Align the left edge of the word with the left edge of some foot.
- Word-Foot-Right: Align the right edge of the word with the right edge of some foot.
- Word-Non-Final: Do not foot the final syllable of the word.

Transducers assigning violation marks: easy to formulate.

No problem for either the counting or the matching approach.

One violation mark / locus

- Parse-syllable: Each syllable must be footed.
- Iambic: Align the right edge of each foot with its head syllable.

Maximal number of violation marks per string is linear in the string's length.

Transducers assigning violation marks: easy to formulate.

Both counting and matching approaches: usually only approximations are possible.

8. Gradient alignment constraints: Unbounded number of violation marks assigned to each locus.

- Main-Foot-Left: Align head-foot with word, left edge.
- Main-Foot-Right: Align head-foot with word, right edge.

Gradience: the head foot receives as many violation marks as the number of syllables intervening between the relevant edges.
E.g. 4 violation marks assigned by MFR to

$$\left\{_{wd}\ \sigma[\sigma 1]\sigma\sigma\sigma\sigma\ \right\}_{wd}$$

But: possibility to reformulate them as non-gradient constraints:

- Assign one violation mark to each syllable intervening between the relevant word edge and the relevant foot edge.

Thus:

- One violation mark per locus.
- Maximal number of violation marks per string: linear in the string's length.
- Transducers assigning violation marks: easy to formulate.
- Hard for counting approach, but easy for matching approach.

Further alignment constraints: AFL, AFR

- All-Feet-Left: Align each foot with the word, left edge.
- All-Feet-Right: Align each foot with the word, right edge.

They are gradient again: each foot receives as many violation marks as the number of syllables intervening between the relevant edges.

- Approximation possible (cf. MFL, MFR).
- But no exact formulation:
  Double cycle needed, not possible with FS technologies.
- How to prove that mathematically? Note that...

AFL and AFR are Quadratic Alignment Constraints

Number of violation marks: not linear in the word's length!

$[\sigma\sigma](\sigma\sigma)(\sigma\sigma)$ gets $2 + 4 = 6$ violation marks.

$[\sigma](\sigma)(\sigma)(\sigma)(\sigma)(\sigma)$ gets $1 + 2 + ... + 5 = 15$ violation marks.

A word of n syllables can be assigned $\frac{n(n-1)}{2}$ violation marks, which is quadratic in the word's length.

Thus: no linear upper bound on the number of violation marks assigned, in function of the string's length.

9. No FST possible for non-linear constraints

No functional transducer assigning violation marks:

> **Theorem**: Let **T** be a functional finite state transducer. Then there exists a linear upper bound on the length of the output, *i.e.* there exists a positive integer N such that for any input string σ (for which there exists an output $T(\sigma)$) the following holds:
>
> $$\mid T(\sigma) \mid \leq N \mid \sigma \mid \qquad \Box$$

Proven by *Pumping Lemma*.

10. Conclusion

Is it possible to realize AFL and AFR in FS OT?

- Transducer assigning violation marks: not
- Transducers producing suboptimal candidates: probably not

Feasibility of FS OT:

- Some wide-spread used constraints in phonology (such as quadratic alignment constraints) cannot (probably) be encoded.
- Are they really needed in phonology? McCarty argues: not.

Thank you for your attention...

...and enjoy your stay in Budapest!

Tamás Bíró: birot@let.rug.nl