OTpimality? Interpreting performance errors in SA-OT

Tamás Bíró

Humanities Computing

Center for Language and Cognition Groningen

University of Groningen

`t.s.biro@rug.nl`

Third Workshop on OT and Interpretation

Groningen, November 7, 2008

Overview:
interpreting performance errors

1. Competence-performance

   (a) Fast speech forms and irregular forms

2. Interpretation in SA-OT

3. Efficient communication in a comunity of speakers

Overview:
interpreting performance errors

1. Competence-performance

   (a) Fast speech forms and irregular forms

2. Interpretation in SA-OT

3. Efficient communication in a comunity of speakers

Proposal: three levels

Level	its product	its model	the product in the model
Competence in narrow sense: static knowledge of the language	grammatical form	standard OT grammar	globally optimal candidate
Dynamic language production process	acceptable or attested forms	SA-OT algorithm	local optima
Performance in its outmost sense	acoustic signal, etc.	(phonetics, pragmatics)	??

# *Competence* vs. *performance*

N. Chomsky: "Linguistic theory is concerned primarily with an ideal speaker-listener, ... unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors ... in applying his knowledge of the language in actual performance." (1965:3)

P. Smolensky: "*competence* can be understood as an idealization of actual behavior—*performance*—in which we have removed the effects of limitations on computational resources: generally speaking, space, time, and precision." (*The Harmonic Mind*, 2006:I:228.)

Competence = grammar: $function$ input ↦ correct output/parse/struct. description

Performance: $algorithm$ that finds it. Or doesn't.

Competence **vs.** *performance*
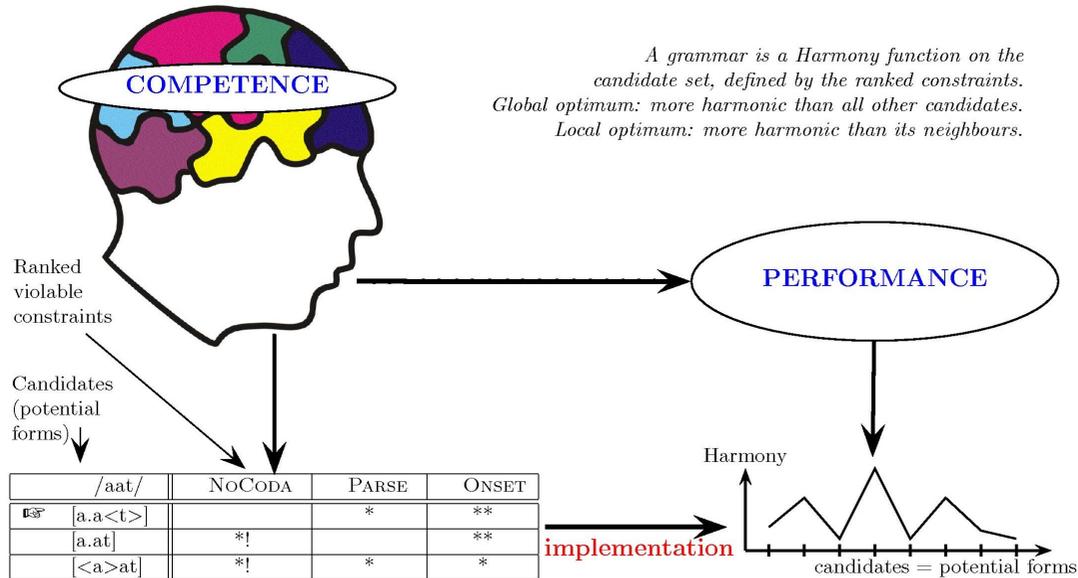
A grammar is a Harmony function on the
candidate set, defined by the ranked constraints.
Global optimum: more harmonic than all other candidates.
*Local optimum: more harmonic than its neighbours.*

COMPETENCE

PERFORMANCE

Ranked
violable
constraints

Candidates
(potential
forms)

Harmony

/aat/	NoCoda	Parse	Onset
☞ [a.a<t>]		*	**
[a.at]	*!		**
[<a>at]	*!	*	*

implementation

candidates = potential forms

Optimality Theory
grammar
competence model
grammatical form = ☞ (globally) optimal candidate

## *SA-OT*
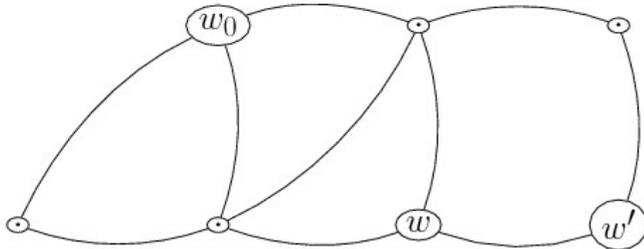**implementation**
performance model
produced forms = globally or locally optimal candidates

Simulated Annealing for OT – general idea
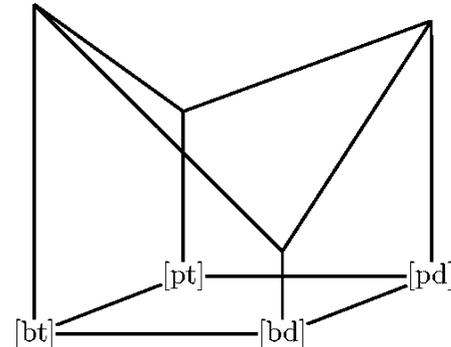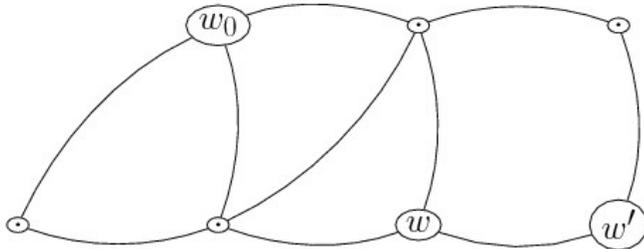
- Neighbourhood structure on the candidate set.

- Landscape's vertical dimension = harmony; random walk.

- If neighbour more optimal: move.

- If less optimal: move in the beginning, don't move later.

Simulated Annealing for OT – general idea
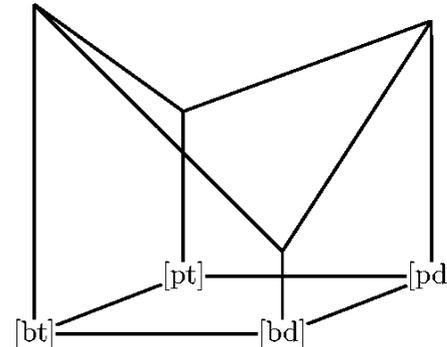
- Neighbourhood structure $\rightarrow$ local optima.

- System stuck in local optima: performance errors.

- Precision of the algorithm depends on its speed (!!).

- Many different scenarios.

Errors and irregularities

Divergence between competence and performance:

- ☞ *Grammatical forms* = globally optimal

- *! Performance errors*: frequency diminishes at slow (careful) production (as in traditional simulated annealing).

- ∼ *Irregularities*: frequency does not diminish at slow (careful) production (due to *strict domination*).

Not all forms in a language need be analysed as grammatical!

Adequacy of a performance model

Performance model: an algorithm that realizes (implements) the grammar (*i.e.*, the model of competence), which

- usually finds the form grammatical w.r.t. grammar (☞),

- but also makes the same errors as humans do,

- with a similar frequency

- under various conditions (speech rate, style, etc.).

Moreover, *runtime* and *complexity* of algorithm is plausible.

Example: string-grammar

$\mathcal{H}$: no0 $\gg$ ass $\gg$ Faith$_{\sigma=0000}$ $\gg$ ni1 $\gg$ ni0 $\gg$ ni2 $\gg$ ni3 $\gg$ nf0 $\gg$ nf1 $\gg$ nf2 $\gg$ nf3 $\gg$ no3 $\gg$ no2 $\gg$ no1 $\gg$ dis

Globally optimal form: ☞ 3333. But 13 local optima: 2222, $\{1,3\}^4$.

Output frequencies for different t_{step} (=inverse speed) values:

output	t_step $= 1$	t_step $= 0.1$	t_step $= 0.01$	t_step $= 0.001$
☞ 3333	0.1174 ± 0.0016	0.2074 ± 0.0108	0.2715 ± 0.0077	0.3107 ± 0.0032
$\sim$ 1111	0.1163 ± 0.0021	0.2184 ± 0.0067	0.2821 ± 0.0058	0.3068 ± 0.0058
$\sim$ 2222	0.1153 ± 0.0024	0.2993 ± 0.0092	0.3787 ± 0.0045	0.3602 ± 0.0091
! 1133	0.0453 ± 0.0018	0.0485 ± 0.0038	0.0328 ± 0.0006	0.0105 ± 0.0014
! 3311	0.0436 ± 0.0035	0.0474 ± 0.0054	0.0344 ± 0.0021	0.0114 ± 0.0016
! others	0.5608	0.1776	< 0.0002	–

$L = P = 4$, $T_{max} = 3$, $T_{min} = 0$, $K_{step} = 1$. Each candidate 4 times as w_0.

Summary of SA-OT

(Without algorithmic details...)

- Universal GEN: also includes neighbourhood structure.

- Therefore notion of *local optimality* introduced.

- Grammatical form: globally optimal.
  Attested form: locally optimal (includes globally optimal).

Overview:
interpreting performance errors

1. Competence-performance

   (a) Fast speech forms and irregular forms

2. Interpretation in SA-OT
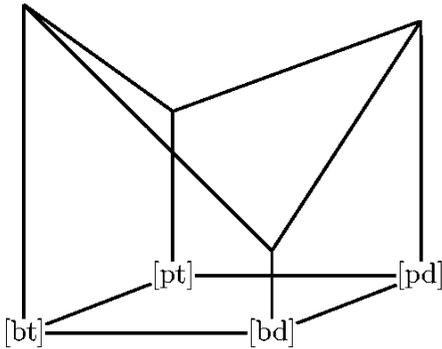
3. Efficient communication in a comunity of speakers

Interpretation in SA-OT?

- Suppose speaker and hearer share hierarchy $\mathcal{H}$ (no learning), beside universal Gen (with neighbourhood structure).

- Observed form: o.

- (Mental) implementation: SA-OT in both directions.

- Interpretation: search for i such that

  - o is local optimum for $\mathrm{Gen}(i)/\mathcal{H}$.
  - Heuristic: probably also global optimum.

Bidirectional SA-OT??

- $\mathrm{Gen}_p(i)$: set of possible forms,
  and $\mathrm{Gen}_i(o)$: set of possible meanings.

- $\mathcal{H}_p$: hierarchy of production constraints
  and $\mathcal{H}_i$: hierarchy of interpretation constraints.

- (i, o) is attested if

  - o is locally optimal for $\mathrm{Gen}_p(i)/\mathcal{H}_p$, and
  - i is locally optimal for $\mathrm{Gen}_i(o)/\mathcal{H}_i$.

The elephant is hitting him: an alternative account?

/ pd /	/him/	C1	C2	C3
pt	e			*
pd	ae		*	...
bd	a			
bt	-		*	...
...@^{k}	...@^{k}	k	...	...

Overview:
interpreting performance errors

1. Competence-performance

    (a) Fast speech forms and irregular forms

2. Interpretation in SA-OT

3. Efficient communication in a comunity of speakers

Efficient communication full of errors: Performance in a virtual speech community

- Multi-agent simulations:
  Agent: competence (OT) + mental computation (SA-OT).

- Communication between two agents: errors affecting communication efficacy?

- New-born agents learn language from elders: correctly? errors inducing language change?

- Theoretical work (computational and mathematical) + checking theories against concrete phenomena.

A note on language change

- Iterative learning: change due to imperfect learning, which can also be due to performance errors in production and interpretation.

- Speed vs. precision: evolutionary advantage of efficiency?

Conclusion

- SA-OT: model of performance.

- Performance vs. interpretation?

- Performance erros in interpretation?

- Effect on communication and language change?

- An answer hopefully soon...

Thank you for your attention!