Optimalitáselmélet formális megközelítésben

6. hét (2015. 03. 20/27.)

Biró Tamás

BBN-ENY-450SZ:F3, BMA-ENYD-321:F3, P/NY/ENY-10::F3, P/NY/ANY-8.02

biro.tamas@btk.elte.hu

http://birot.web.elte.hu/courses/2015-formOT/

Feladott cikkek

Feladott cikkek

- Bruce Tesar, Jane Grimshaw and Alan Prince (1999). Linguistic and cognitive explanation in Optimality Theory. In Ernest Lepore and Zenon Pylyshyn (eds.): *What is Cognitive Science?* 295–326. Malden, MA: Blackwell. `http://ruccs.rutgers.edu/~prince/hold/Introot.pdf`
- Prince, Alan, and Paul Smolensky (1997). Optimality: From neural networks to universal grammar. *Science* 275: 1604-1610.
  Újra kiadva: Paul Smolensky and Géraldine Legendre (eds.): *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar (Vol. 1: Cognitive architecture)*. MIT Press, 2006, chapter 4.
- Tamás Biró: Elephants and Optimality Again: SA-OT accounts for pronoun resolution in child language. In: B. Plank et al. (eds.). *Computational Linguistics in the Netherlands 2009*. LOT, 2009, pp. 9-24.

Ki vállal prezentációt?

Jespersen-ciklus

- Eszter adja elő Henriëtte de Swart *Expression and Interpretation of Negation: An OT Typology* c. könyve (Springer, 2010) 3. fejezetét.

- Alessandro Lopopolo and T. Biró. Language Evolution and SA-OT: The case of sentential negation. *Computational Linguistics in the Netherlands Journal*, vol. 1 (2011), pp. 21–40. `http://www.clinjournal.org/sites/default/files/Lopopolo.pdf`.

- `http://www.birot.hu/talks.php?abstract=CLIN-2011`.

- `http://www.birot.hu/publications/Lopopolo-Biro-CLIN2011.pdf`

- `http://www.birot.hu/publications/2011-OAP.pdf`

Optimalizáció a nyelvészetben

David Marr és a kognitív folyamatok leírási szintjei

1. **Komputációs szint:** az adott kognitív képesség, a folyamat (mint függvény, vagy mint komputációs feladat), amelyet az adott kognitív rendszer képes elvégezni.

2. **Algoritmikus szint:** (a) a reprezentációk (adatstruktúrák), (b) az azokon végzett műveletek, és (c) ezen műveletek sorozata, mint algoritmus, vagy általánosabban, egy „architektúra”.

3. **Fizikai szint:** az algoritmus implementációja valamely „hardveren” (vagy „wetware”-ben).

Az Optimalitáselmélet építőkockái (building blocks)

- Form (alak)
- Candidate (jelölt)
- Gen $\rightarrow$ az optimalizálás keresési tere (search space)
- Constraint (megszorítás, korlát): szűrők vagy elemi függvények
- Hierachia, constraintek rendezése
  $\rightarrow$ az optimalizálás célfüggvénye (target function)
- stb.

Komputációs szint: optimalizáció a nyelvészetben

$$\mathrm{SF}(u) = \underset{x \in \mathrm{Gen}(u)}{\arg\,\mathrm{opt}} \; H(x)$$

Harmony Grammar: $H(x) = \sum_{i=n}^{1} w_i \cdot C_i(x)$

opt: min for $<$ on $\mathbb{R}$.

Optimality Theory: $H(x) = (C_n(x), C_{n-1}(x), \ldots, C_1(x))$

opt: lexicographical order on $\mathbb{R}^n$.

Principles and Parameters: $H(x) = \bigvee_{i=n}^{1} (w_i \wedge C_i(x))$

opt: false „more optimal" than true.

Az Elvek és Paraméterek (P&P) alapaxiómája

- Konsztréntek = elvek. $C_i(x) = \begin{cases} t, & \text{ha } x \text{ megsérti } C_i\text{-t,} \\ f, & \text{ha } x \text{ nem sérti meg } C_i\text{-t.} \end{cases}$
- Nyelvtan = az egyes C_i elvekhez tartozó w_i paraméterek értéke:

$$\begin{array}{ll} w_i = t \text{ (igaz)}, & \text{ha } C_i \text{ paraméter bekapcsolva, és} \\ w_i = f \text{ (hamis)}, & \text{ha } C_i \text{ paraméter kikapcsolva.} \end{array}$$

- Az optimalizálandó célfüggvény: $H(x) = \bigvee_{i=n}^{1}(w_i \wedge C_i(x))$.
- Ezen nyelvtan szerint az u bemenethez tartozó grammatikus alak: $\mathrm{SF}(u) = \text{vagy} \in \arg\,\mathrm{opt}_{x\in \mathrm{Gen}(u)} \bigvee_{i=n}^{1}(w_i \wedge C_i(x))$, ahol f jobb, mint t. Implicite fel szokás tenni, hogy legalább egy jelölt mindig létezik, amelyik teljesíti a bekapcsolt paramétereket. Ekkor: $\mathrm{SF}(u) \in \left\{x \in \mathrm{Gen}(u) \,\middle|\, \forall i : (w_i = f) \vee (C_i(x) = f)\right\}$.

Az Optimalitáselmélet (OT) alapaxiómája

- Nyelvtan = konsztréntek véges sorozata:
  $(F_n, F_{n-1}, \ldots, F_1)$, avagy $(C_n, C_{n-1}, \ldots, C_1)$.

- Az optimalizálandó célfüggvény: $H(x) = \big(C_n(x), C_{n-1}(x), \ldots, C_1(x)\big)$.

- Ezen nyelvtan szerint az u bemenethez tartozó grammatikus alak:

  $\mathrm{SF}(u) =$ vagy $\in F_1\big(\ldots F_{n-1}\big(F_n\big(\ \mathrm{Gen}(u)\ \big)\big)\big)$, ill.

  $\mathrm{SF}(u) =$ vagy $\in \arg\ \mathrm{opt}_{x \in \mathrm{Gen}(u)} \big(C_n(x), C_{n-1}(x), \ldots, C_1(x)\big)$.

- . . . a *lexikografikus rendezés* („szótárak ábécésorrendje") szerint:
  a vektorok balról első különböző eleme alapján rendezzük őket.

A Harmónianyelvtan (HG) alapaxiómája

- Nyelvtan = valósértékű konsztréntekhez rendelt súlyok rendszere: $(C_n, C_{n-1}, \ldots, C_1)$-hez tartozik $(w_n, w_{n-1}, \ldots, w_1)$.
- $C_i(x) \geq 0$. $w_i \geq 0$ Erősebb konsztrént nagyobb súlyt kap.
- Az optimalizálandó célfüggvény: $H(x) = \sum_{i=n}^{1} w_i \cdot C_i(x)$.
- Ezen nyelvtan szerint az u bemenethez tartozó grammatikus alak: $\mathrm{SF}(u) =$ vagy $\in \arg\min_{x \in \mathrm{Gen}(u)} \sum_{i=n}^{1} w_i \cdot C_i(x)$ a valós számok körében értelmezett „kisebb mint” reláció szerint.
- A szűrő-perspektívának nincs értelme.
- Másoknál $\arg\min$ helyett $\arg\max$, és vagy $w_i \leq 0$, vagy $C_i(x) \leq 0$.

Az OT és a HG viszonya

Kumulativitás: alacsonyabbra rendezett sértések összeadódnak, és súlyosabbak, mint egy magasabbra rendezett sértés. Két fajtája:

- Counting cumulativity: alacsonyabbra rendezett constraint többszörös sértése.
- Ganging-up cumulativity: két alacsonyabb constraint sértése összeadódik.

$w =$	C_2 2	C_1 1	H
HG ☞ x	1!	0	2
OT ☞ y	0	3	3

$w =$	C_3 4	C_2 3	C_1 2	H
HG ☞ x	1!	0	0	4
OT ☞ y	0	1	1	5

Standard OT-ben egyik sincs. HG-ban mindkettő lehetséges.
Boersma-féle sztochasztikus OT: ganging-up van, counting nincs.

V.ö. Jäger, G., & Rosenbach, A. (2006). The winner takes it all–almost: Cumulativity in grammatical variation. *Linguistics*, 44(5), 937-971.

Az OT és a HG viszonya

$w =$	C_2 w_2	C_1 w_1	H
HG ☞ x	1!	0	w_2
OT ☞ y	0	k	$k \cdot w_1$

$w =$	C_3 w_3	C_2 w_2	C_1 w_1	H
HG ☞ x	1!	0	0	w_3
OT ☞ y	0	1	1	$w_2 + w_1$

- Counting cumulativity: alacsonyabbra rendezett constraint többszörös sértése. HG-ben, ha $w_2 < k \cdot w_1$, azaz $k > w_2/w_1$.
- Ganging-up cumulativity: két alacsonyabb constraint sértése összeadódik. HG-ben, ha $w_3 < w_2 + w_1$.
- HG és OT ekvivalens, ha HG súlyok olyanok, h. nem lehetséges kumulativitás.
- Tegyük fel, hogy maximálva van a constraintek sértéseinek a száma: $\exists\ q$ korlát, hogy $\forall x \in \mathrm{Gen}(u)$-ra és $\forall i (\in \mathbb{Z})$-re: $C_i(x) \leq q - 1$. Ekkor használjunk hatványsor szerinti súlyokat: $w_i \geq q^i$. (NB: $q > 1$.)

Rang / rendezés / hierarchia, HG $\rightarrow$ OT

- (Megj.: $\forall u$-hoz $\exists q_{(u)}$, amely. . . ; vagy $\exists q$, amely $\forall u$-hoz. . . ?)

- Legyen C_i constraint **rangja** *(rank)* r_i, és súlya w_i.

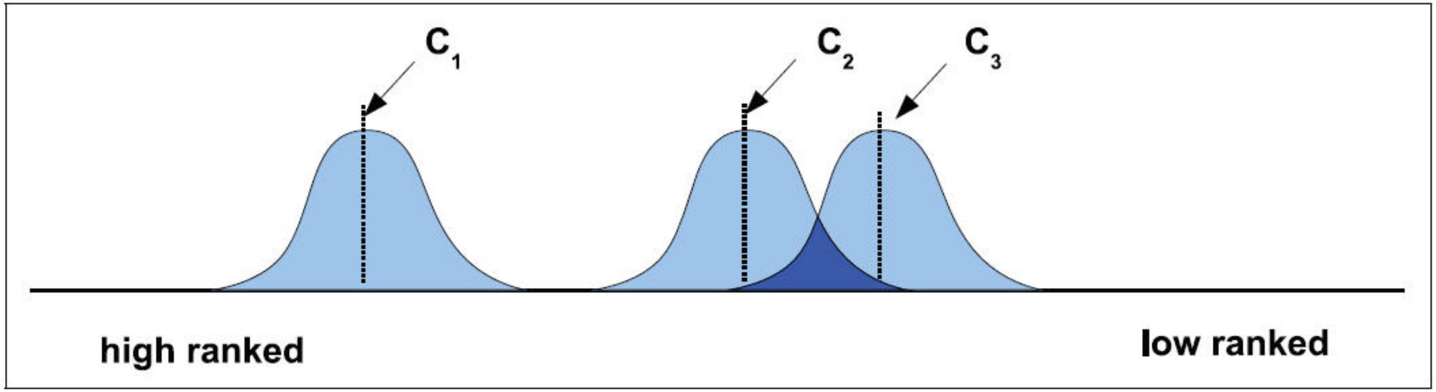
OT:	$C_i \gg C_j$ („magasabbra van rendezve"), a.cs.a. ha $r_i > r_j$.
	Eddig $r_i = i$. De lehet más is, pl. tanulóalgoritmusokban.
Lineáris HG:	$w_i := r_i$.
Exponenciális HG:	$w_i := e^{r_i}$ (ahol $e = 2,7182\ldots$; vagy más, rögzített alap).
q-HG:	$w_i := q^{r_i}$, és q változtatható.

- Nyelvtan = constraint ranking = az r_i vagy w_i értékek rendszere.

- q-HG-ból OT, ha strict domination limit: $q \rightarrow \infty$ (vagy $q = \omega$).

Sztochasztikus OT, sztochasztikus HG (Paul Boersma)

- Nyelvtan = constraint ranking = az r_i vagy w_i értékek rendszere.
- Minden egyes EVAL-alkalmazás során („at evaluation time"): minden C_i konsztrénthez: $r'_i = r_i + \epsilon_i$, ahol az ϵ_i zaj értékek valamely (például egyenletes vagy normális) valószínűségi eloszlás(ok)ból származnak.
- Sztochasztikus OT: a konsztrénteket az r'_i értékek szerint rendezzük, és ezen pillanatnyi hierarchia szerint választjuk ki az optimális jelölt(ek)et.
- Sztochasztikus HG: az r'_i-kből számítjuk a w'_i súlyokat, $H(x) = \sum w'_i C_i(x)$.



Variáció az Optimalitáselméletben

Egy u-hoz több grammatikus alak: $\mathrm{SF}_1 \sim \mathrm{SF}_2$. Több megközelítés:

- $\mathrm{Gen}(U)$ több eleme azonos „violation profile"-lal.
- Nem csak $\mathrm{Gen}(U)$ legjobb elemét generáljuk:
  - Coetzee: A legjobb elemek: mindazok, amelyek ugyanúgy sértik a „kritikus pont" feletti konsztrénteket. A kritikus pont alatti konsztréntek mintha nem is léteznének, ill. csak a relatív gyakoriság meghatározásához kellenek.
  - Biró: Lokális optimumok.
- Egy mentális nyelvtan = több „elemi nyelvtan" kombinációja.
  - Anttila: rendezetlen konsztréntok lehetséges permutációi.
  - Nagy és Reynolds: „lebegő konsztréntet" különböző pontokon szúrhatjuk be a többi közé, így különböző rendezéseket kapunk $\longrightarrow$ ezek kombinációi.
  - Boersma (v.ö. Jarosz): elemi nyelvtanok sztochasztikus kombinációja.

OT és a kognitív folyamatok leírási szintjei

1. **Komputációs szint:** az adott kognitív képesség, a folyamat (mint függvény, vagy mint komputációs feladat), amelyet az adott kognitív rendszer képes elvégezni.
2. **Algoritmikus szint:** (a) a reprezentációk (adatstruktúrák), (b) az azokon végzett műveletek, és (c) ezen műveletek sorozata, mint algoritmus vagy „architektúra".
3. **Fizikai szint:** az algoritmus implementációja valamely „hardveren" (vagy „wetware"-ben).

1. **Komputációs szint:** a nyelv, mint kognitív képesség, egy függvény: az u bemenethez (pl. mögöttes alakhoz) hozzárendeli $\mathrm{SF}(u) = \arg\,\mathrm{opt}_{c \in \mathrm{Gen}(u)} H(c)$ kimenetet (felszíni alakot).

2. **Algoritmikus szint:** eljárás, amely (jól-rosszul; determinisztikusan vagy sztochasztikusan) kiszámítja u-hoz, Gen-hez, H-hoz, stb. $\mathrm{SF}(u)$-t.

3. **Fizikai szint:** az optimalizáció implementációja „hardveren".

Következő alkalomra:

- Ki vállal prezentációt?

Viszlát áprilisban!