Simulated Annealing in Syntax

Author: Daniel Gallagher
Date: February 24, 2025

1 SA-OT-Syntax: A strain too far?

In Bíró 2006, the optimisation strategy known as Simulated Annealing (SA) is applied to Optimality Theory (OT) with the aim of accounting for the differences seen in *competence* - what the framework predicts as the ideal candidate - and *performance* - what is actually seen in speech data. Bíró focuses on applying SA-OT to three empirical phenomena in order to aid his argument: (1) stress in Dutch fast speech, (2) dutch voice assimilation and (3) the cliticisation of the definite article in Hungarian. He thus focuses primarily on phonological variations pertaining to the structure of individual words or syllables. However, he does not explore the possibility of using SA with OT-Syntax and therefore we do not observe whether syntactic alternations such as word order or singular/plural agreement variations can be modelled in terms of probabilities through this approach. This would be particularly interesting as it can be speculated that variations may tend to be less likely than phonological variations and the two forms would provide an interesting juxtaposition.

How could syntax be approached in such a manner? Syntactic variations can be observed where multiple candidates occur in real-world speech and these include both the global optimum as well as multiple local optima. On the other hand, there may be multiple ideal candidates in OT and it is not obvious how to deal with these within the framework. For instance, the example provided in Table 1 shows one such example with multiple optimal candidates. A topology or neighbourhood structure can be determined in a similar manner to how it was for phonological phenomena: the neighbour *w'* of candidate *w* will contain typically one operation OP of difference and this is represented across a horizontal plain. These operations could be MOVE or MERGE, for instance. Since simulated annealing is typically modelled as trying to reach a minimum, the vertical axis represents a *violation vector* where higher means that more highly-ranked constraints have been violated. We see three examples of syntactic constraints which could be violated in Table 1 - namely RECOVERABILITY, LEFT EDGE (CP) and TELEGRAPH. Candidate *d* is the global *maximum* here in that it has the worst violation vector and therefore represents a higher point on the three-dimensional plane in Figure 1. The other candidates fare better and thus represent a lower point on this plane and any of the three may be reached by the Simulated Annealing algorithm depending on the specific parameters set such as the *temperature T*. The algorithm for SA-OT seen in Bíró 2006: p. 64 is not made specific to phonology and can therefore be comfortably re-applied to candidates which comprise of syntactic alternations.

Bíró's insightful research into this implementation of an optimisation algorithm to explain phonological alternations in the OT framework could therefore have been further enriched with a case study in syntactic variations. This could further be used in research into the merging of Minimalism and Optimality Theory as seen in Broekhuis & Woolford 2013.

		REC	LE(CP)	TEL
a.	☞ the man [who I know t]		*>	
b.	☞ the man [that I know t]			*<
c.	☞ the man [I know t]		*>	
d.	the man [who that I know t]		*>	*<

Table 1: Relative clauses with preposed relative pronoun (Broekhuis & Woolford 2013: p. 129)

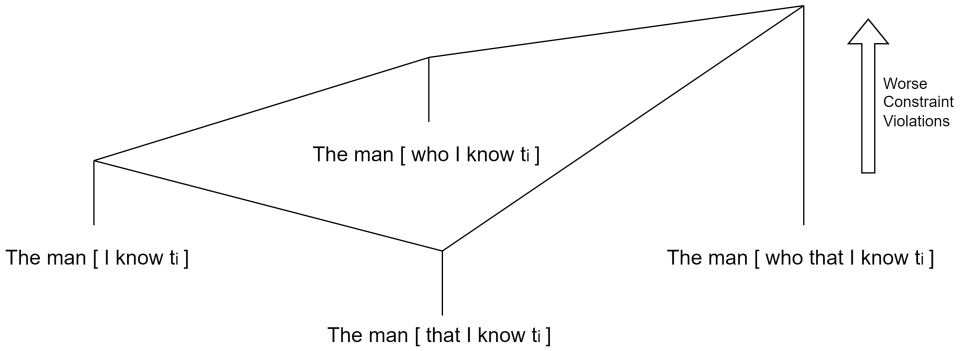

Figure 1: Potential Syntactic Neighbourhood Structure

References

Bíró, Tamás Sándor. 2006. *Finding the right words: implementing optimality theory with simulated annealing*. English. date$_s$ubmitted:2006 Rights: University of Groningen. University of Groningen dissertation.

Broekhuis, Hans & Ellen Woolford. 2013. Minimalism and optimality theory. In https://api.semanticscholar.org/CorpusID:30736197.