Feature Reactivation in Minimalist Parsing

Aniello De Santo (University of Utah)

Overview A top-down parser for Minimalist grammars [MGs; 9] can successfully predict a variety of off-line processing preferences, via metrics linking parsing behavior to memory load [6, 2, 4]. Given the close association between this model and modern minimalist syntax, it is important to extensively evaluate its empirical coverage. In this abstract we propose new metrics for the MG parser, that take into account the set of features triggering movement steps in a derivation — thus implementing a notion of memory reactivation. As a case study of how these metrics improve the empirical coverage of the MG approach, we successfully model the processing preferences for stacked relative clauses (RC) in [11], and a variety of previously modeled RC contrasts.

MG Parsing The MG parsing model systematically links syntactic structure to processing difficulty by connecting the stack states of a (deterministic) top-down parser [9] to memory burden. Memory usage [3, 7] is measured based on how long a node is kept in memory (**tenure**). Consider the MG derivation in Fig. 1. The index of a node n encodes the moment n was predicted and put in memory by the parser. The outdex encodes the moment n is confirmed and flushed out of memory. Tenure for n is measured as $outdex(n) - index(n)$, and can then used to define a set of off-line metrics of processing difficulty (e.g., **max.** or **avg.** tenure across all nodes in the derivation [4]).

Implementing Feature Reactivation We want to make the parsing model sensitive to structural repetition. Inspired by previous literature on syntactic priming, we stipulate that if a moved element has been recently stored in memory, storing the next item of the same kind (e.g., triggered by a wh-feature) should be less costly (feature *reactivation*). Note that these items are not in memory at the same time, so this is different from interference effects. We implement this procedure by counting the number of parsing steps between movements of the same type. Consider the derivation in Fig. 2, with two NP movers associated to a feature f. Practically, reactivation for NP_2 is measured by subtracting from its index the outdex of the previous node associated to f (NP_1; so $w - y$). Finally, since reactivation is supposed to encode facilitatory effects induced by structural repetition, we operationalize it as: $R(m_i) := 1 - \frac{1}{i(m_i) - o(m_{i-1})}$. Additionally, we weight the tenure of a node by its reactivation value (*boost*, $BT := Tenure(m_i) * R(M_i)$), to investigate the interaction between reactivation and notions of storage previously employed by the MG parser. We then derive metrics that use reactivation and boost to compute processing costs over full derivations (e.g., **max.** R).

A Case Study We consider stacked RC constructions, in which a noun phrase (*the reporter*) is modified by two relative clauses. Zhang [11] explores the processing of stacked RCs in English (1) and Mandarin Chinese (2), in a 2×2 design crossing extraction type (subject or object) with the position of the RC (RC1 or RC2). She reports faster reading times when RC1 and RC2 are of the same type, than when they are of different types (i.e. SS $>$ OS and OO $>$ SO). Crucially, none of the metrics used in the previous MG parsing literature is able to account for this effect. We model these contrasts as in (1) and (2), and we also consider a classical contrast between subject (SRC) and object (ORC) RCs both in English and Mandarin, which has been focus of much MG processing work in the past [4, 11, a.o.]. Since the parser is sensitive to detailed grammatical information, we consider two analyses for the RC construction: a promotion analysis [5], and a wh-movement analysis [1]. Our simulations show that the parser now successfully captures the facilitatory effect associated to consecutive processing of similar movement types (ORC-ORC; SRC-SRC), as well as the more classical SRC-ORC contrasts. We discuss how these results relate to the way different reactivation metrics are sensitive to differences between syntactic analyses. This extension to the computational model will clearly require extensive empirical evaluation. However, these results provide a valuable proof-of-concept in favor of a careful exploration of how ideas from the priming literature can be incorporated in formal models of structural processing.

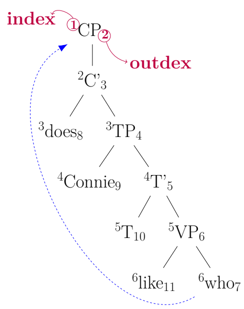

Figure 1: MG derivation tree with parse steps.

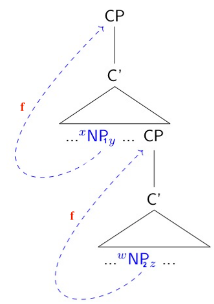

Figure 2: Example tree for memory reactivation.

(1) Test sentences for English Stacked Relative Clauses

a. The horse that kicked the wolf on Tuesday that patted the lion just now went home **SS**
b. The horse that the wolf kicked on Tuesday that patted the lion just now went home **OS**
c. The horse that kicked the wolf on Tuesday that the lion patted just now went home **SO**
d. The horse that the wolf kicked on Tuesday that the lion patted just now went home **OO**

(2) Example of test sentences for Mandarin Chinese Stacked Relative Clauses

a. Nage zai xingqier tile xiaoma haojici de zai jintian zhuile daxiang
Dem on Tuesday kick-perf horse several-times de on today chase-perf elephant
de gongniu likaile jia
De bull leave-perf home
'The bull that kicked the horse for several times on Tuesday that chased the elephant earlier today left home.' **SS**

b. Nage zai xingqier xiaoma tile haojici de zai jintian zhuile daxiang
dem on Tuesday horse kick-perf several-times de on today chase-perf elephant
de gongniu likaile jia
De bull leave-perf home
'The bull that the horse kicked for several times on Tuesday that chased the elephant earlier today left home.' **OS**

Language	**Processing Contrast**	⟨**MaxR′**, **AvgBT**⟩		⟨**MaxBT**, **MaxR′$_R$**⟩	
		Promotion	Wh-movement	Promotion	Wh-movement
English	$OO < SO$	✓	✓	✓	✓
	$SS < OS$	✓	×	✓	✓
Mandarin	$OO < SO$	✓	✓	✓	✓
	$SS < OS$	✓	✓	×	✓
English	$SRC < ORC$	✓	×	✓	✓
Mandarin	$ORC < SRC$	✓	✓	×	✓

Table 1: Summary of results of ranked metrics by contrast and RC construction.

[1] Chomsky, N. (1977). On wh-movement. *Formal syntax*. **[2]** Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language ,brain*. **[3]** Graf T., J. Monette, and C. Zhang. (2017). Relative clauses as a benchmark for Minimalist parsing. *Journal of Language Modelling*. **[4]** Kayne, R.S. (1994). The antisymmetry of syntax. *MITPress*. **[5]** Kobele, G.M., S. Gerth, and J. Hale. (2013) Memory resource allocation in top-down minimalist parsing. *Formal Grammar*. **[6]** Rambow, O. and Joshi, A.K. (2015). A processing model for free word-order languages. *Perspectives on sentence processing*. **[7]** Reitter, D., Keller, F., and Moore, J. D.. (2011). A computational cognitive model of syntactic priming. *Cognitive science*. **[8]** Stabler, E. P. (2013). Two models of minimalist, incremental syntactic analysis. *Topics in cognitive science*. **[9]** Troyer M., O'Donnell, T.J., Fedorenko, E., and Gibson E. (2011). Storage and computation in syntax: Evidence from relative clause priming. *Proc. of the Cognitive Science Society*. **[10]** Zhang, C. (2017). *Stacked Relatives: Their Structure, Processing and Computation*. PhD thesis, Stony Brook U.