Student Abstract & Poster Session AAAI-21

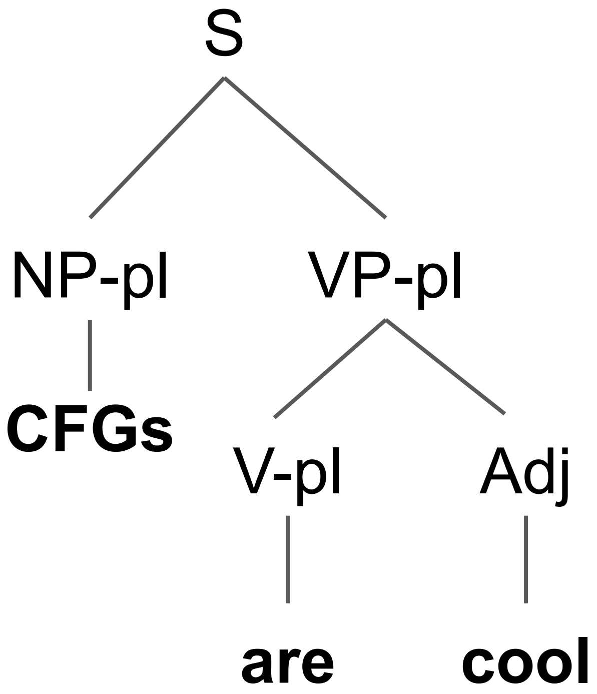
Comparing Symbolic Models of Language via Bayesian Inference

Annika Heuser & Polina Tsvilodub

Massachusetts Institute of Technology

UNIVERSITÄT OSNABRÜCK

Context-free grammars (CFGs) - Symbolic Models



Probabilistic CFGs (PCFGs)

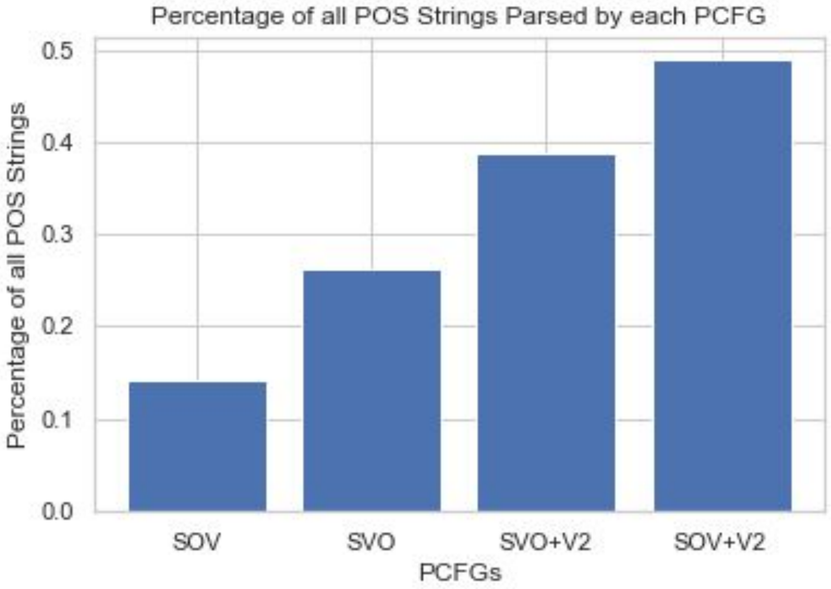
S -> NP-pl VP-pl 0.5
S -> NP-s VP-s 0.5

[1], [2], [5]

Our PCFG hypotheses cannot be compared directly

$$P(\text{CFG, Type} \mid \text{Data}) \propto P(\text{Data} \mid \text{CFG, Type})\, P(\text{CFG} \mid \text{Type})\, P(\text{Type})$$

Posterior Likelihood Priors



P(Data | CFG, Type) =

Product of each PCFG rule needed to parse each sentence

[4], [6], [8]

Challenges & Proposed Solutions

Errors in data: 31.65% of the total POS strings are ungrammatical

Different data subset sizes → Normalize by number of sentences in subset

PCFGs penalized for parsing more sentences because:

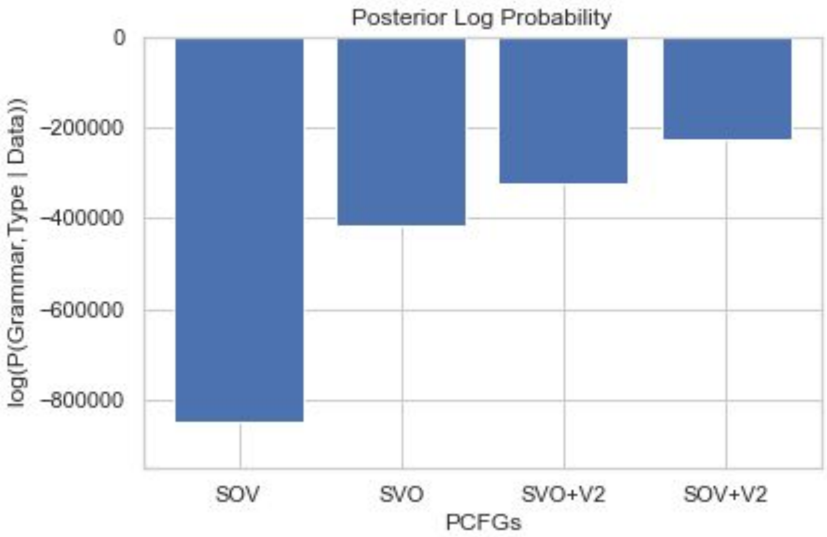
1. Longer sentences = lower likelihood → Normalize by sentence length
2. More generalizable = less probability mass per sentence type → Weight by proportion of parseable sentence types

[4], [6]

Contributions

- Revealed troubling number of errors resulting from transcription and part-of-speech (POS) tagging

- Method of comparing CFGs that account for different subsets of data



[6], [7]

References

[1] Chomsky, N. 1965. Aspects of the theory of syntax. *Cambridge, MA: MIT Press*(1977): 71–132.

[2] Crain, S.; and Nakayama, M. 1987. Structure dependence in grammar formation. *Language* 522–543.

[3] Lake, B. M.; Ullman, T. D.; Tenenbaum, J. B.; and Gershman, S. J. 2017. Building machines that learn and think like people. *Behavioral and brain sciences* 40.

[4] MacWhinney, B. 2000. *The CHILDES Project: Tools for analyzing talk. Transcription format and programs,* volume 1. Psychology Press.

[5] Pelletier, F. J. 1994. The principle of semantic compositionality. *Topoi* 13(1): 11–24.

[6] Perfors, A.; Tenenbaum, J. B.; and Regier, T. 2011. The learnability of abstract syntactic principles. *Cognition* 118(3): 306–338.

[7] Tenenbaum, J. B.; Kemp, C.; Griffiths, T. L.; and Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022): 1279–1285.

[8] Xu, F.; and Tenenbaum, J. B. 2007. Word learning as Bayesian inference. *Psychological review* 114(2): 245.