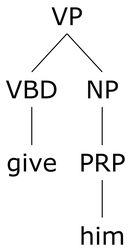
Judging Grammaticality with Count-Induced Tree Substitution Grammars

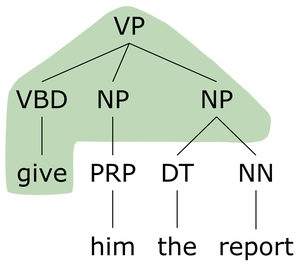
human language technology center of excellence

Francis Ferraro, Matt Post and Benjamin Van Durme • Johns Hopkins University

Motivation

Syntax should indicate whether a sentence is grammatical. PCFGs used in practice are too permissive and make unrealistic independence assumptions and thus are too permissive in modeling language.





Tree substitution grammars (TSGs) have an extended domain of locality and can capture long-range grammatical dependencies:

✗ *They* ***give him****.*

✓ *They* ***give him the report****.*

Post (2011) showed that non-parametric Bayesian TSGs perform well on grammaticality judgments, but do we need these complex inference techniques?

Learning Probabilistic TSGs (PTSGs)

DOP

Extract all fragments from observed corpora; heuristics are needed to scale. Learned fragments can be very large and not generalizable.

Bayesian Methods

A DP Prior encourages compact fragments (Cohn et al., 2009; Post et al., 2009). But they are non-deterministic and complex to implement.

Are PTSGs needed for *classification*?

Count-Induced Unweighted TSGs

We deterministically count compact TSG fragments via dynamic programming: iteratively extract *K* most frequent subtrees of size *R*. These parameters enforce sparsity and help temper exponential growth.

ExtractFragments(R,K)

```
S ← ∅
F(1,K) ← top K CFG rules used
for r = 2 to R do
  S ← {F ∈ F(r-1, K), extended by 1 rule}
  F(r,K) ← top K elements of F(r-1, K) ∪ S
end for
```

Classification Tasks

We view grammaticality judgments as a discriminative learning problem. We experiment on two granularities:

Coarse (Okanohara & Tsujii, 2007; Cherry & Quirk, 2008)
Grammatical sentences are from BLLIP (LDC2000T43):

The most troublesome report may be the August merchandise trade deficit due out tomorrow.

Ungrammatical ones are Kneyser-Ney trigram sequences

To and , would come Hughey Co. may be crash victims , three billion .

Fine-grain (Foster & Andersen, 2009)
Grammatical examples are from the WSJ; errors are automatically injected to force ungrammaticality:

The league's promoters hope retirees and tourists will join die-hard fans like Mr. de Castro and pack then stands to see the seniors .

Sentences are of equal length and each has an automatic "gold standard" parse. Train/dev/test splits are predefined.

Model and Baselines

We build binary feature vectors by matching fragments against the parse. We use `liblinear` to train an L-2 regularized SVM. We optimize on `dev`.

Baselines	
Bigrams	All bigrams and unigrams
CFG	Berkeley parser most probable parse
TSG	Post (2009)'s Viterbi derivation of sentence
Model	
COUNT+CFG	ExtractFragments(R, K), and **CFG** features

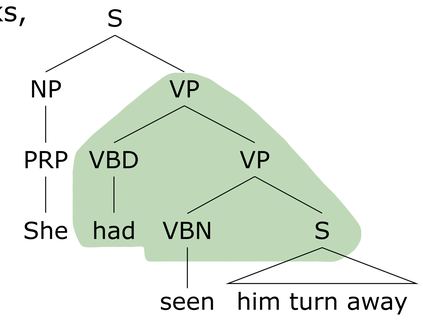
Results: Test Accuracy

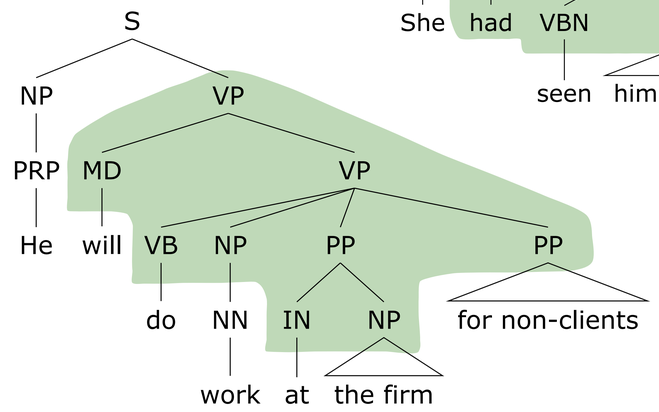
We tune R and K on `dev`; best was achieved with R=3, K=50k. Fragments extracted with R=15, K=50k are comparably sized to Post (2009)'s TSGs.

		coarse	fine
Chance		50.0	50.0
Bigrams		68.4	61.4
CFG		86.3	64.5
TSG		**89.1**	67.0
COUNT+CFG	R=3, K=50K	**89.1**	**67.2**
	R=15, K=50K	88.2	66.6

Grammatical Fragments

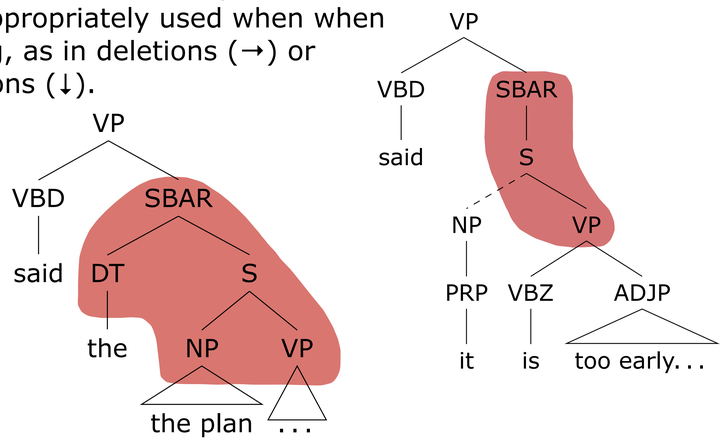
For both coarse and fine tasks, our fragments learn simple structure and complex phenomena, including perfective (→) phrases and modal (↓) constructions.





Ungrammatical Fragments

Grammatical errors may force rules to be inappropriately used when when parsing, as in deletions (→) or insertions (↓).



Conclusions

Using a simple, deterministic counting method, we can achieve the gains of Post (2011)'s Bayesian-induced TSGs for grammaticality judgments.

References

C. Cherry, C. Quirk. Discriminative, syntactic language modeling through latent SVMs. AMTA 2008.

T. Cohn, S. Goldwater, P. Blunsom. Inducing compact but accurate tree-substitution grammars. NAACL 2009.

J. Foster, O.E. Andersen. GenERRate: generating errors for use in grammatical error detection. NAACL: BEA 2009.

D. Okanohara, J. Tsujii. A discriminative language model with pseudo-negative samples. ACL 2007.

M. Post, D. Gildea. Bayesian learning of a tree substitution grammar. ACL 2009.

M. Post. Judging grammaticality with tree substitution grammar derivations. ACL 2011.