Asynchronous Binarization for Synchronous Grammars

John DeNero, Adam Pauls, and Dan Klein
{ denero, adpauls, klein } @ cs.berkeley.edu

Overview

In multi-pass decoding with synchronous grammars, rule binarization can be decoupled, rather than synchronized. Each unconstrained monolingual binarization can then be optimized for the relevant stage in decoding.

Parsing Stage: Source-Side Binarization

First, we project the synchronous grammar to the source language

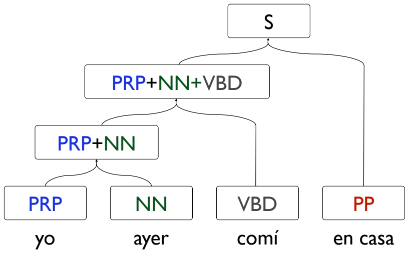
Root		Source Yield		Target Yield	Source Projection
S	→	PRP$_1$ NN$_2$ VBD$_3$ PP$_4$	;	PRP$_1$ VBD$_3$ PP$_4$ NN$_2$	S → PRP NN VBD PP
S	→	PRP$_1$ NN$_2$ VBD$_3$ PP$_4$	;	PRP$_1$ VBD$_3$ NN$_2$ PP$_4$	
PP	→	en casa	;	at home	PP → en casa
PP	→	en casa	;	indoors	

Next, we binarize the source-side projection (see our NAACL paper)

S → PRP NN VBD PP

(1) S → PRP+NN+VBD PP
(2) PRP+NN+VBD → PRP+NN VBD
(3) PRP+NN → PRP NN

Then, we build a source-binarized parse forest via CKY-style parsing

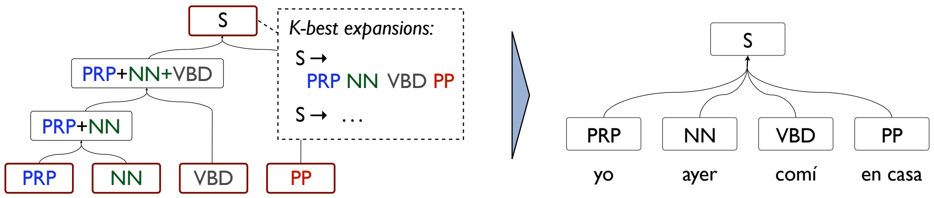


Highlights of our decoder:

- Binarization is chosen to minimize the total number of grammar symbols
- Coarse-to-fine parsing uses subsets of the monolingual grammar projetion
- Forests are pruned by thresholding node max-marginals (before LM)

Reranking Stage: Target-Side Binarization

Source-side binarization is collapsed out to create an n-ary forest

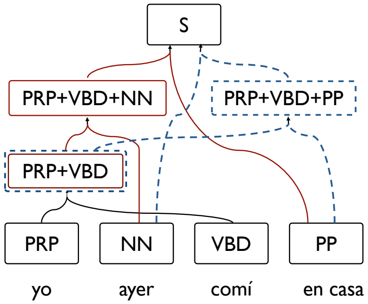


The parse forest is then re-binarized for target-side gap adjacency

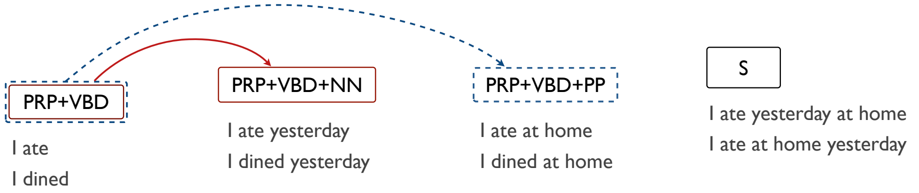
Target Yield Binarizations

S → PRP+VBD+NN PP
PRP+VBD+NN → PRP+VBD NN
PRP+VBD → PRP VBD

S → PRP+VBD+PP NN
PRP+VBD+PP → PRP+VBD PP
PRP+VBD → PRP VBD



Derivations are reranked efficiently with an n-gram language model



Common k-best list is shared across multiple parent nodes

Derivations are reranked after applying each binary sub-rule

Search error reduction relative to n-ary reranking gave 0.3 BLEU