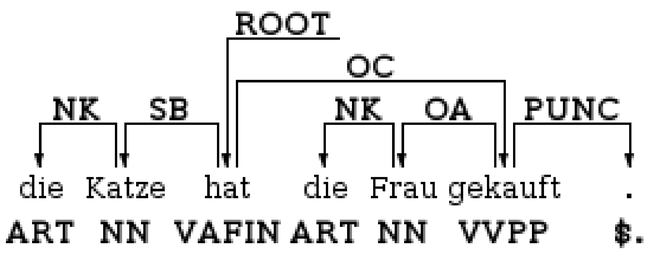
Non-projective Dependency-based Pre-Reordering with Recurrent Neural Network for Machine Translation

Antonio Valerio Miceli-Barone Giuseppe Attardi
Università di Pisa

Reorder sentences in a target-like word order

die Katze hat die Frau gekauft .
die Frau hat gekauft die Katze .
the woman has bought the cat .

Based (but not constrained) on the source dependency parse



Using recurrent neural network models:
Base RNN-RM
Base GRU-RM
Fragment RNN-RM

Results

- Experiments on *de-en* Europarl, news2009, news2013
- All models improve over baseline (Moses)
- Base GRU-RM and Fragment GRU-RM equal or improve over Collins et al. (2005) rules
- Translation BLEU (Europarl):

Moses:	33.00		Base RNN-RM:	33.41	+0.41
Collins:	33.52	+0.52	**Base GRU-RM:**	**34.15**	**+1.15**
"Oracle":	41.80	+8.80	Fragment RNN-RM:	33.54	+0.54

Reordering transition system

State:
(curNode, emittedNodes, lastAction)

Actions:
EMIT, UP, DOWN$_{childNode}$ *, RIGHT*

Fragment:
Sequence of actions between word emissions

System

- Align parallel training corpus with giza++
  Symmetrize alignment with *grow-diag-final-and* heuristic
- Generate heuristic reference reordering of source training corpus
  Al-Onaizan and Papineni (2006) heuristic
- Parse source side with DeSR (Attardi and Ciaramita 2007)
  Parse tree may be non-projective
- Train one of these neural network models:
  Base RNN-RM or *Base GRU-RM*
  Compute permutation as sequence of words
  Fragment RNN-RM (two-level hierarchical RNN)
  Compute permutation as sequence of words and sequence transition actions between each word.
  all these model support *non-tree-local* reordeing
- Reorder source training corpus with the neural network model
  Beam search
- Train *Moses* on the reordered parallel corpus