Motivation

Hard to annotate syntax in domains requiring subject matter expertise.
Paraphrases have similar syntax across alignments, easier to make than parse trees.

Task

Input: Sentence and paraphrase
Output: Parse of sentence
Approach: Encourage agreement between parses of sentence and paraphrase across alignments.

Dual Decomposition

(works with any arc-factored parser)

Idea: Parse sentence and paraphrase simultaneously, encourage agreement across aligned arcs using hard or soft dual decomposition.

Pair-finding

(works with any n-best parser)

Idea: Within pairs with smallest alignment cost, maximize the sum of parse scores.

Classifier

Idea: Pick between original and joint parse to better handle bad alignments.

Features: # crossing arcs, % non-proj. arcs, sentence lengths, word overlap, and parse structure templates

Evaluation details:

Parsers: Lei et al. (2014), Charniak and Johnson (2005), McClosky et al. (2006)
Corpora: 2,000 sentences annotated with their paraphrases.

Blue indicates significant improvement (p < 0.01) over parser baseline.