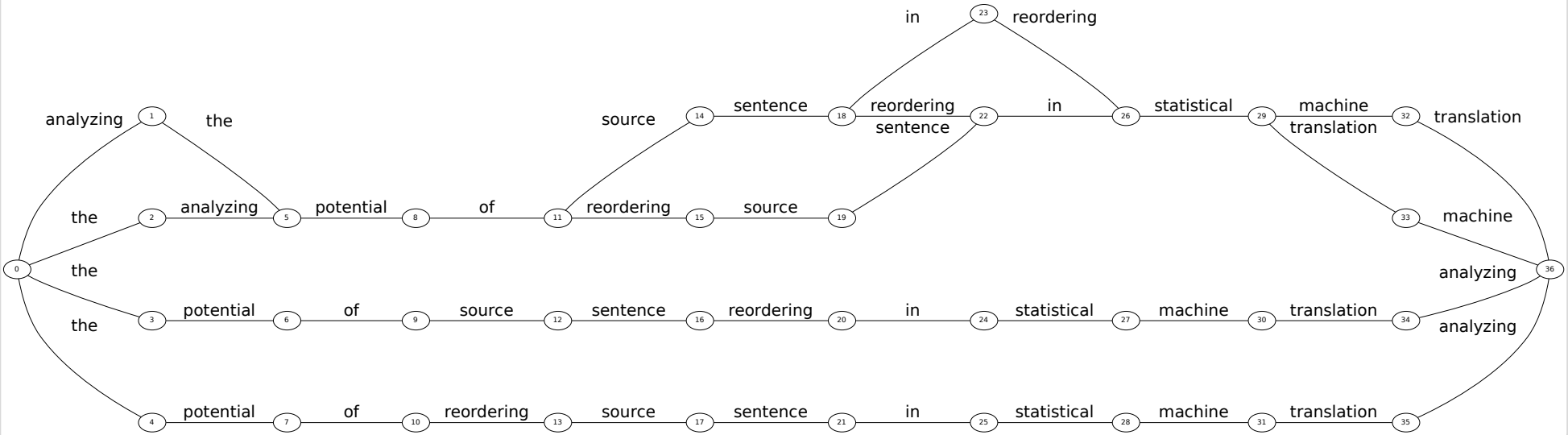
Karlsruhe Institute of Technology

interACT

Analyzing the Potential of Source Sentence Reordering in Statistical Machine Translation

Teresa Herrmann, Jochen Weiner, Jan Niehues, Alex Waibel

Institute of Anthropomatics, Interactive Systems Lab



Introduction

- **Word Reordering** is one of the main issues in Machine Translation
- Popular Approach:
  - Reordering the source sentence before translation
- Analyze the potential
- Oracle experiments
  1. How good can we get with an optimally reordered sentence?
  2. How well does the model approximate the optimally reordered sentence?
  3. How good are we at finding the best reordering?

Outline

- Reordering approach
  - Rule types
  - Training and application
- Reordering oracles
  - Optimally reordered source
  - Best model approximation of optimally reordered sentence
- Experiments
  1. Potential of reordering the source sentence
  2. Restriction of the search space
  3. Ranking different word orders

Reordering Approach

- POS-based reordering model
  - Short-range rules (Rottmann and Vogel, 2007)
  - Long-range rules (Niehues and Kolss, 2009)
- Tree-based reordering model (Herrmann et al., 2013)

- Rules for reordering the source sentence
  - automatically learned
  - non-deterministic
  - reordering variants stored in a lattice

- General rule
  - LHS: sequence of POS / constituents
  - RHS: indexed representation of target order for tokens
  - rule probability based on frequency of occurrence

Rule Types

- Short-range rules
  - continuous POS sequences
- Long-range rules
  - discontinuous POS sequences with placeholders
- Tree-based rules
  - head + sequence of children in syntactic parse tree constituents

Rule Type	Example Rule	
Short	VVIMP VMINF PPER	→ 2 1 0
Long	VAINF * VVPP	→ 0 2 1
Tree	VP PTNEG NP VVPP	→ 0 2 1

Learning and Applying Reordering Rules

- Requirements
  - parallel training corpus with POS tags / parse trees for source side
  - word alignment
- Training
  - search for changes in word order between source and target sentences
  - if alignment links cross
    - monotonize alignment
    - extract rule that rearranges source words in the order of aligned target words
- Application
  - apply reordering rules to source sentence
  - store all produced reordering variants in the lattice
  - include original source sentence
  - edges are assigned transition probabilities based on the rule probability

Scoring Reordered Paths

- (Reordered) sentence
  - path in the reordering lattice
- Probability for given path
  - product of transition probabilities of traversed edges
- Highest scoring path should represent best reordering for the sentence
- Reordering lattice is one model in the log-linear model of translation system
- Model score is optimized with MERT

Outline

- Reordering approach
  - Rule types
  - Training and application
- Reordering oracles
  - Optimally reordered source
  - Best model approximation of optimally reordered sentence
- Experiments
  1. Potential of reordering the source sentence
  2. Restriction of the search space
  3. Ranking different word orders

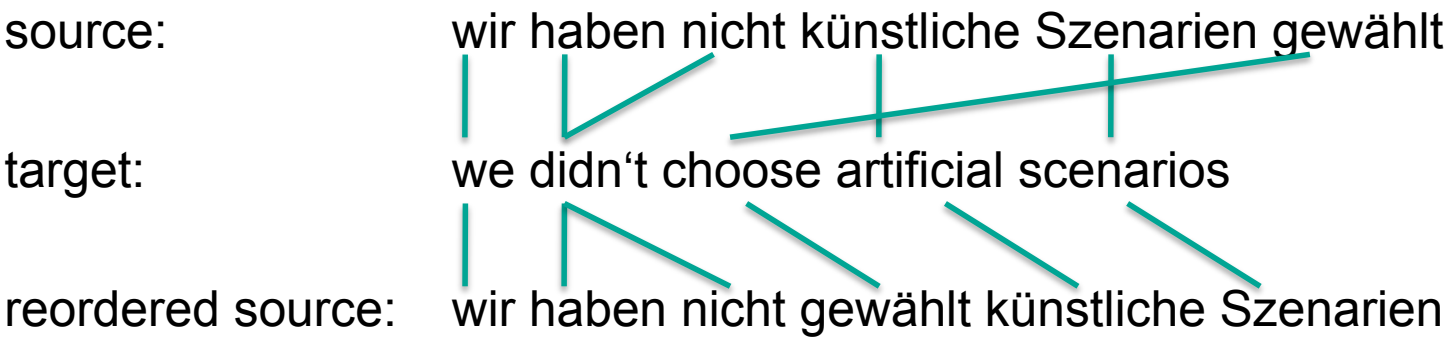
Oracle Reordering

- Investigate the impact of source reordering on translation quality
  - upper bound of the pre-ordering approach
  - upper bound of restriction of the search space by our reordering model
  - decoder performance at finding a good path

- Compare the actual system performance with two oracles
  - Optimally reordered sentence
  - Oracle path

Optimally Reordered Sentence

- alignment of source and reference sentence
- create reordering of the source sentence (Birch et al., 2010)
  - source words are assigned the position of their aligned target word

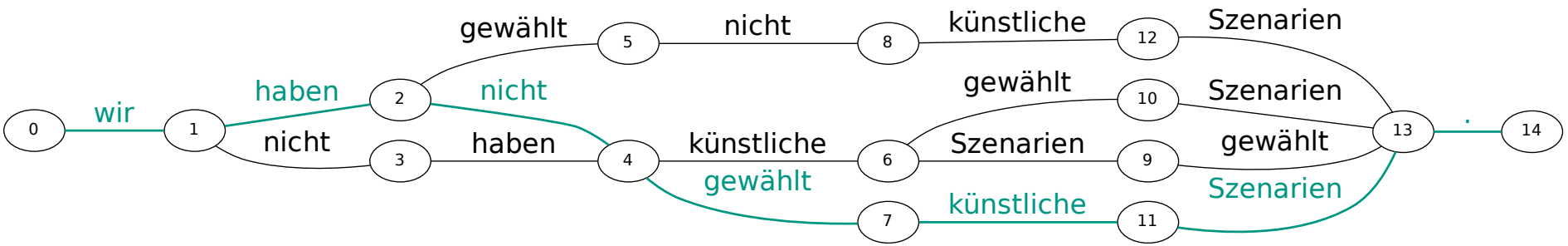


- optimally reordered sentence according to the alignment

→ **oracle reordering** of input sentence

Oracle Path

- decoder search space is restricted by applied reordering rules



- find the “best” path in the lattice
  - oracle reordering
  - closest path to oracle reordering
- calculate Kendall’s tau distance
  - reordering = permutation
  - transform one permutation into another
  - number of swaps between two adjacent symbols

→ **oracle path**: path with smallest distance to oracle reordering

Outline

- Reordering approach
  - Rule types
  - Training and application
- Reordering oracles
  - Optimally reordered source
  - Best model approximation of optimally reordered sentence
- Experiments
  1. Potential of reordering the source sentence
  2. Restriction of the search space
  3. Ranking different word orders

Experiments

- System
  - Phrase-based MT system
  - no optimization between translations using different rule types
  - no additional reordering in the decoder allowed
- Translation directions
  - German-English
  - English-German
- Domains
  - News texts
  - TED talks

Potential of Reordering the Source Sentence

- Compare 3 translation results
  - **no reordering**
    - monotone translation of source sentence
  - **oracle reordering**
    - translation of optimally reordered sentence
  - **system performance**
    - reordering lattices, decision during decoding

Potential of Reordering the Source Sentence

- German-English

Reordering Type	News	TED
Monotone	20.23	27.18
Lattice Reordering	22.45	30.87
Oracle	25.42	33.39

- English-German

Reordering Type	News	TED
Monotone	15.91	24.22
Lattice Reordering	16.34	24.95
Oracle	18.84	28.77

Potential of Reordering the Source Sentence

- gap between actual performance and oracle reordered translation
- possible reasons
  - lattice-based reordering
    - reordering variant not in the lattice
  - decoder search
    - in the lattice but decoder does not choose it

→ deeper analysis of those two aspects

Lattice-based restriction of the search space

- Lattice-based reordering
  - restriction of the search space of possible reorderings by the lattice
  - how much does this influence the drop in performace

- best translation reachable with reordering lattice: oracle path
- translations of oracle paths in lattices produced with different rule types
- compare to translation of oracle reordering

Lattice-based Restriction of the Search Space

- Translation of oracle path
  - German-English

Reordering Type	News BLEU	Size	TED BLEU	Size
Monotone	20.23		27.18	
Short	21.37	193K	29.98	68K
Short+Long	21.41	255K	30.66	163K
Tree	22.28	249K	30.22	82K
Short+Long+Tree	22.65	538K	31.12	213K
Oracle	25.42 (+ 2.8)		33.39 (+ 2.3)	

Lattice-based Restriction of the Search Space

- Translation of oracle path
  - English-German

Reordering Type	News BLEU	Size	TED BLEU	Size
Monotone	15.91		24.22	
Short	16.31	186K	25.83	76K
Short+Long	16.70	383K	25.99	170K
Tree	16.60	727K	25.49	237K
Short+Long+Tree	17.07	1M	26.38	373K
Oracle	18.84 (+ 1.8)		28.77 (+ 2.4)	

Ranking different word orders

- how good is the search
- decoder's search for the best path depends on
  - path probability
  - influence of the reordering model
  - interaction with other models

- reordering lattices produced by different rule types
- real decoding to find the path
- compare with oracle path and oracle reordering

Ranking different word orders

- Decoder Performance
  - German-English News

Reordering Type	DecoderPath BLEU	OraclePath BLEU	
Monotone		20.23	
Short	21.59	21.37	
Short+Long	21.35	21.41	
Tree	22.10	22.28	
Short+Long+Tree	22.45	22.65	+ 0.2
Oracle		25.42	

Ranking different word orders

- Decoder Performance
  - German-English News

Reordering Type	DecoderPath BLEU	OraclePath BLEU
Monotone		20.23
Short	21.59	21.37
Short+Long	21.35	21.41
Tree	22.10	22.28
Short+Long+Tree	22.45	22.65
Oracle	+ 3.0	25.42

Ranking different word orders

- Decoder Performance
  - German-English TED

Reordering Type	DecoderPath BLEU	OraclePath BLEU	
Monotone		27.18	
Short	30.00	29.98	
Short+Long	30.73	30.66	
Tree	29.96	30.22	
Short+Long+Tree	30.87	31.12	+ 0.3
Oracle		33.39	

Ranking different word orders

- Decoder Performance
  - German-English TED

Reordering Type	DecoderPath BLEU	OraclePath BLEU
Monotone		27.18
Short	30.00	29.98
Short+Long	30.73	30.66
Tree	29.96	30.22
Short+Long+Tree	30.87	31.12
Oracle	+ 2.5	33.39

Ranking different word orders

- Decoder Performance
  - English-German News

Reordering Type	DecoderPath BLEU	OraclePath BLEU	
Monotone		15.91	
Short	16.27	16.31	
Short+Long	16.31	16.70	
Tree	16.18	16.60	
Short+Long+Tree	16.34	17.07	+ 0.7
Oracle		18.84	

Ranking different word orders

- Decoder Performance
  - English-German News

Reordering Type	DecoderPath BLEU	OraclePath BLEU
Monotone		15.91
Short	16.27	16.31
Short+Long	16.31	16.70
Tree	16.18	16.60
Short+Long+Tree	16.34	17.07
Oracle	+ 2.5	18.84

Ranking different word orders

- Decoder Performance
  - English-German TED

Reordering Type	DecoderPath BLEU	OraclePath BLEU	
Monotone		24.22	
Short	24.83	25.83	
Short+Long	24.87	25.99	
Tree	24.50	25.49	
Short+Long+Tree	24.95	26.38	+ 1.4
Oracle		28.77	

Ranking different word orders

- Decoder Performance
  - English-German TED

Reordering Type	DecoderPath BLEU	OraclePath BLEU
Monotone		24.22
Short	24.83	25.83
Short+Long	24.87	25.99
Tree	24.50	25.49
Short+Long+Tree	24.95	26.38
Oracle	+ 3.8	28.77

Conclusions

- analysis of the potential of a common reordering approach
- German-English, English-German
- translation of News text and TED talks

- oracle reordering
  - improves the translation by up to 6.2 BLEU points

- search space restriction
  - best approximation of oracle reordering when more complex and complementary reordering rules are applied
  - gap between best oracle path and oracle reordering
    - German-English: 2.5 – 3.0 BLEU points
    - English-German: 2.5 – 3.8 BLEU points
  - need for better reordering rules

Conclusions

- decoder performance
  - German-English
    - decoder path quite close to oracle path
    - 0.2 – 0.3 BLEU points difference
  - English-German:
    - finding best path more difficult
    - 0.7 – 1.4 BLEU points difference
    - need for better ranking of reordering variants
- oracle experiments confirmed benefit of source reordering for translation quality
- potential improvements by up to 3.0 – 3.8 BLEU points
  - better rules
  - better ranking of reordering quality

Karlsruhe Institute of Technology

EU★BRIDGE

Thank you !

teresa.herrmann@kit.edu

http://isl.anthropomatik.kit.edu