

Bridging the Gap between Training and Inference for Neural Machine Translation

Wen Zhang^{1,2} Yang Feng^{1,2*} Fandong Meng³ Di You⁴ Qun Liu⁵

¹Key Laboratory of Intelligent Information Processing
Institute of Computing Technology, Chinese Academy of Sciences (ICT/CAS)

²University of Chinese Academy of Sciences, Beijing, China
`{zhangwen, fengyang}@ict.ac.cn`

³Pattern Recognition Center, WeChat AI, Tencent Inc, China
`fandongmeng@tencent.com`

⁴Worcester Polytechnic Institute, Worcester, MA, USA
`dyou@wpi.edu`

⁵Huawei Noah's Ark Lab, Hong Kong, China
`qun.liu@huawei.com`

読み手: 三輪誠 (豊田工業大学)

一部, 図表は著者の論文より引用

読んだ動機: ACL The Best Long Paperの 説明 (ACL2019 organizing committeeより)

- The paper addresses the long-standing problem of **exposure bias** in sequence-to-sequence transduction.
→ よくある問題だよな。それでBest paperってすごいね。
- The proposed solution is to alternate between conditioning on words from the reference and **oracle-selected words** from the decoder's own output.
→ 何が新しいの？ Scheduled sampling?
- The approach works within the current teacher-forcing training paradigm and **improves over scheduled sampling**.
→ Scheduled samplingよりはいいらしい。
- The experiments are **solid**, the results are convincing and likely to **influence future work in MT**.
- The approach is also **applicable** to other seq2seq tasks.
→ 実験が素晴らしいのか？ 汎用性が高いのか？

Teacher forcingによるNeural MT

- 学習時
 - 前トークンの**正解**を
次トークンの入力に利用

Teacher forcingによるNeural MT

- 予測時
 - 前トークンの**予測**を
次トークンの入力に利用

Teacher forcingにおける問題と動機

- **exposure bias** (露出バイアス)

- 学習時と予測時で入力が違う
 - 学習時: ground truthを利用
 - 予測時: モデルの予測を利用

→ 予測時に学習していない環境にさらされる

- MTだと **Overcorrection** が起こる

reference: We should comply with the rule.

cand1: We should abide with the rule.

cand2: We should abide by the law.

cand3: We should abide by the rule.

Overcorrection recovery による解決をしたい

abideと予測してしまったものを無理にwithで正解に近づける (overcorrection)

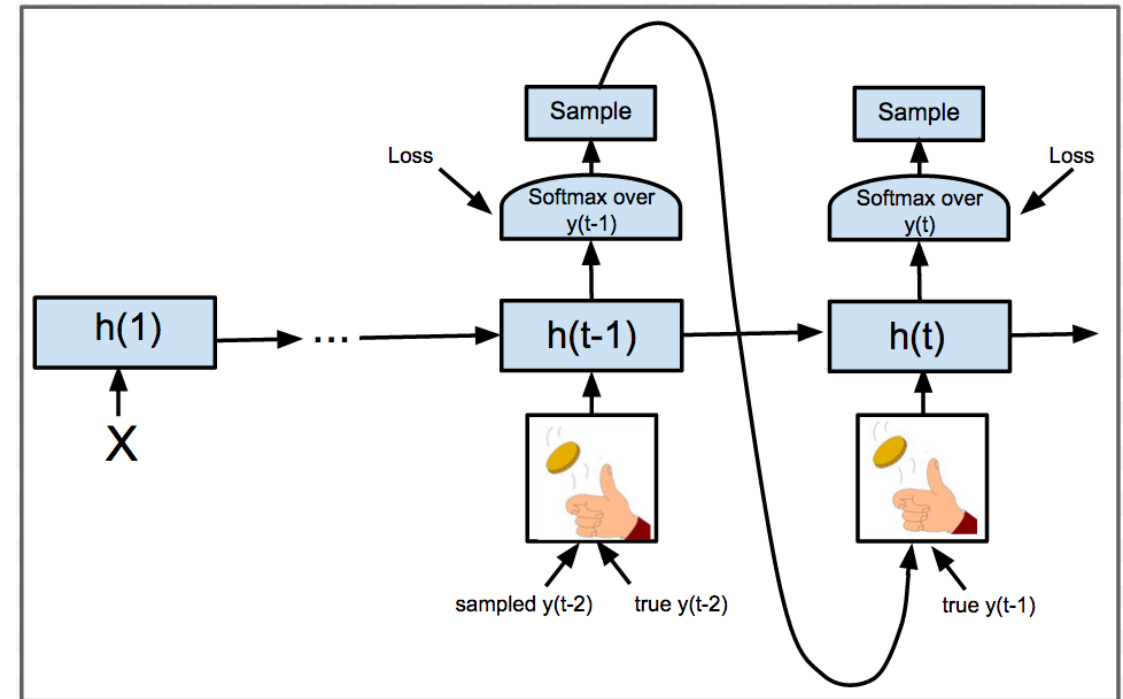
→ “abide with” という間違ったパターンを学習

abide byと予測してしまったためにthe lawを予測してしまう

abide byと予測してもwithを予測したと思ってthe ruleの予測に復帰したい → Overcorrection recovery

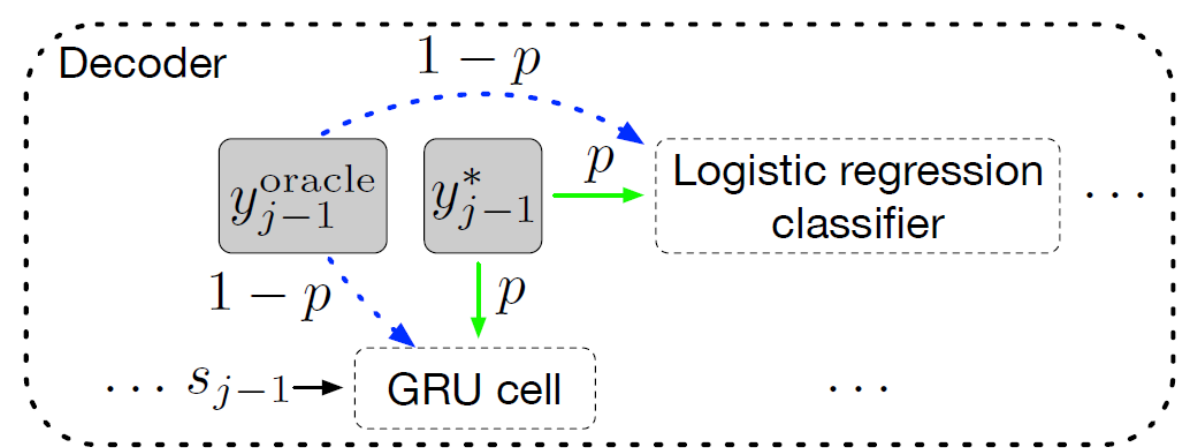
従来 of 解決法: Scheduled Sampling [Bengio et al., 2015]

- 学習時も確率的に前トークンの予測を使う
 - 学習当初は正解トークン
 - 学習後期は予測トークンを使うようにスケジューリング



[Bengio et al., 2015]より引用

提案手法



モデルへの入力を正解と
予測 (oracleと呼ぶ) から確率的に選んで利用

- Word-level oracleとSentence-level oracleの2種類のoracle wordの作り方を提案 (Oracle word selection)
 - Gumbel-maxを用いたノイズの追加
- 正解を選ぶ確率 p を徐々に減らす (Sampling with decay \doteq Scheduled sampling)

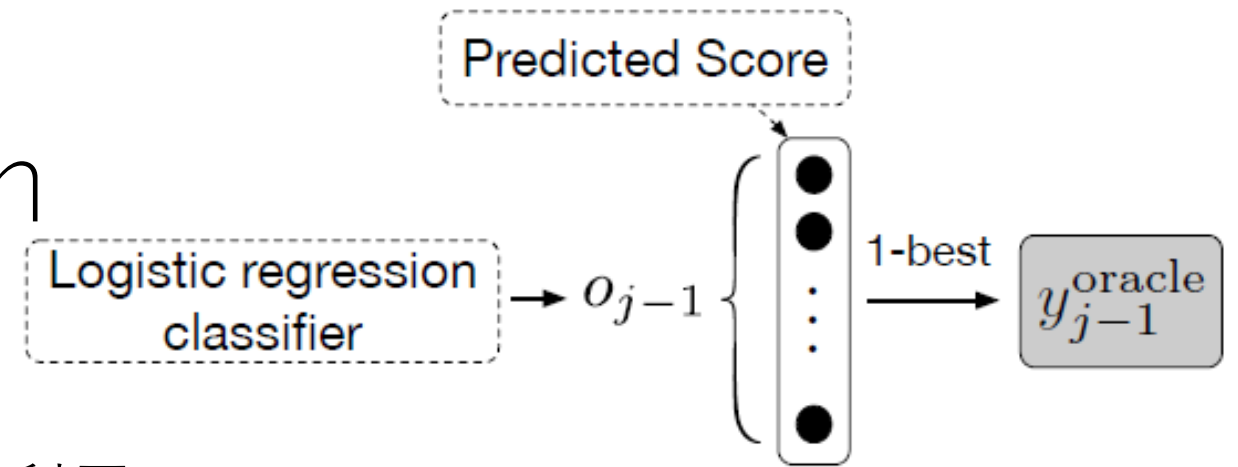
エポック

$$p = \frac{\mu}{\mu + \exp(e/\mu)}$$

ハイパーパラメタ
(実験では12)

※注意: teacher-forcingの入力を変えるだけで損失は同じ ← 予測は同じ

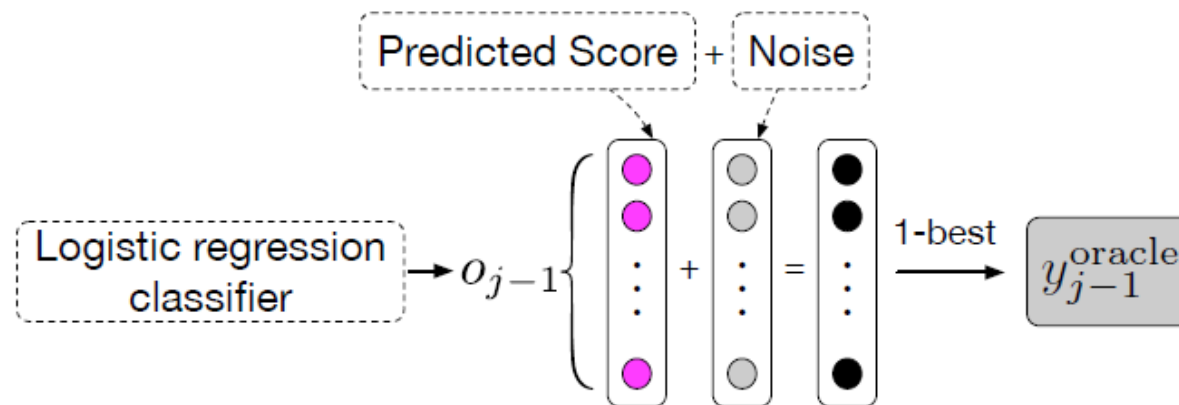
Oracle word selection



- Word-level oracle
 - トークンごとのモデルの予測を利用
- Sentence-level oracle
 - ビームサーチ (実験ではビーム幅は3) を用いてBLEUが最大となる系列の予測を利用
 - 文長が正解と予測系列で同じになるように文長が一致するところでEOSを出すように強制

Oracle word selectionにおける Gumbel noiseの追加

- Gumbel-maxにより確率的にmaxでない入力も利用
- Sentence-level oracleはデコード時にGumbel noiseを追加



$$u \sim U(0,1)$$

$$\eta = -\log(-\log u)$$

$$\tilde{o}_{j-1} = (o_{j-1} + \eta) / \tau$$

$$\tilde{P}_{j-1} = \text{softmax}(\tilde{o}_{j-1})$$

温度
(高いとランダムに近づく)

評価データと評価設定

- NIST Chinese→English (Zh→En)
 - 訓練1.25M文対, 開発878文 (NIST2002; MT02)
 - テスト919 (MT03), 1,788 (MT04), 1,082 (MT05), 1,664 (MT06)
 - 30K words (Chinese: 97.7%, English: 99.3%)
 - 訓練時の最大文長を50語に制限
 - BLEU (case-insensitive)で評価
- WMT'14 English→German (En→De)
 - 訓練4.5文対, 開発3,007文, テスト2,737文
 - BPE (English: 39,417 tokens, German: 40,274 tokens)
 - 訓練時の最大文長を128サブワードに制限
 - BLEU (case-sensitive)で評価

比較対象

- RNNSearch
 - ベースライン. エンコーダ: Bi-GRU, デコーダ 2GRU+attention
- Transformer
 - もう一つのベースライン.
- SS-NMT (Scheduled Sampling) [Bengio et al., 2015]
 - Word-level oracleと同じ.
- MIXER [Ranzato et al., 2015]
 - 文レベルでのBLEUを報酬にした強化学習を用いる手法
- OR-NMT (Overcorrection Recovery)
 - Sentence-levelとGumbel noiseを用いる手法

学習設定

- RNNSearch
 - word embeddingと中間層は512
 - テスト時 (提案手法ではない) のビームサイズは10
 - minibatch sizeは80
 - Adadelta, dropoutを利用
- Transformer
 - fairseqをデフォルトパラメタで利用

評価

RNNSearchでもTransformerでも有効
(ノイズの言及はなし)

Systems	Architecture	MT03	MT04	MT05	MT06	Average
<i>Existing end-to-end NMT systems</i>						
Tu et al. (2016)	Coverage	33.69	38.05	35.01	35.77	35.40
Shen et al. (2016)	MRT	37.41	39.87	37.45	36.81	37.88
Zhang et al. (2017)	Distortion	37.93	40.40	36.81	35.77	37.73
<i>Our end-to-end NMT systems</i>						
this work	RNNsearch	37.93	40.53	36.65	35.80	37.73
	+ SS-NMT	38.82	41.68	37.28	37.98	38.94
	+ MIXER	38.70	40.81	37.59	38.38	38.87
	+ OR-NMT	40.40^{††*}	42.63^{††*}	38.87^{††*}	38.44[‡]	40.09
	Transformer	46.89	47.88	47.40	46.66	47.21
	+ word oracle	47.42	48.34	47.89	47.34	47.75
	+ sentence oracle	48.31[*]	49.40[*]	48.72[*]	48.45[*]	48.72

Systems	newstest2014
RNNsearch	25.82
+ SS-NMT	26.50
+ MIXER	26.76
+ OR-NMT	27.41[‡]
Transformer (base)	27.34
+ SS-NMT	28.05
+ MIXER	27.98
+ OR-NMT	28.65[‡]

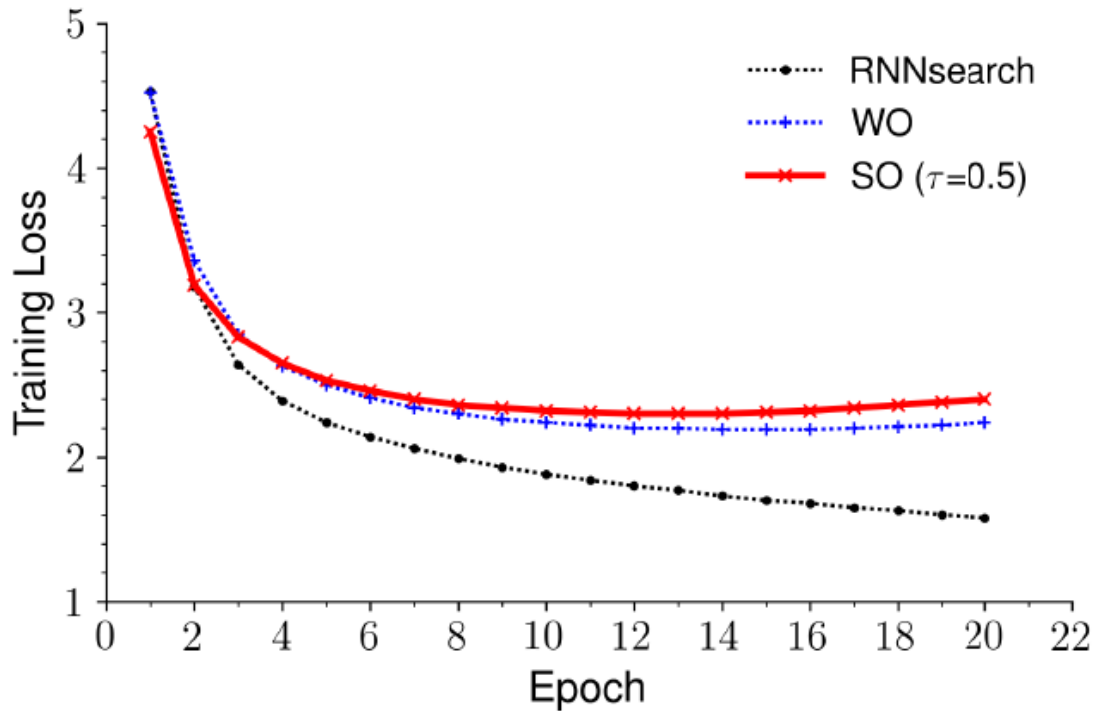
En→De

ノイズは有効
単語より文レベルの
oracleのほうが良い

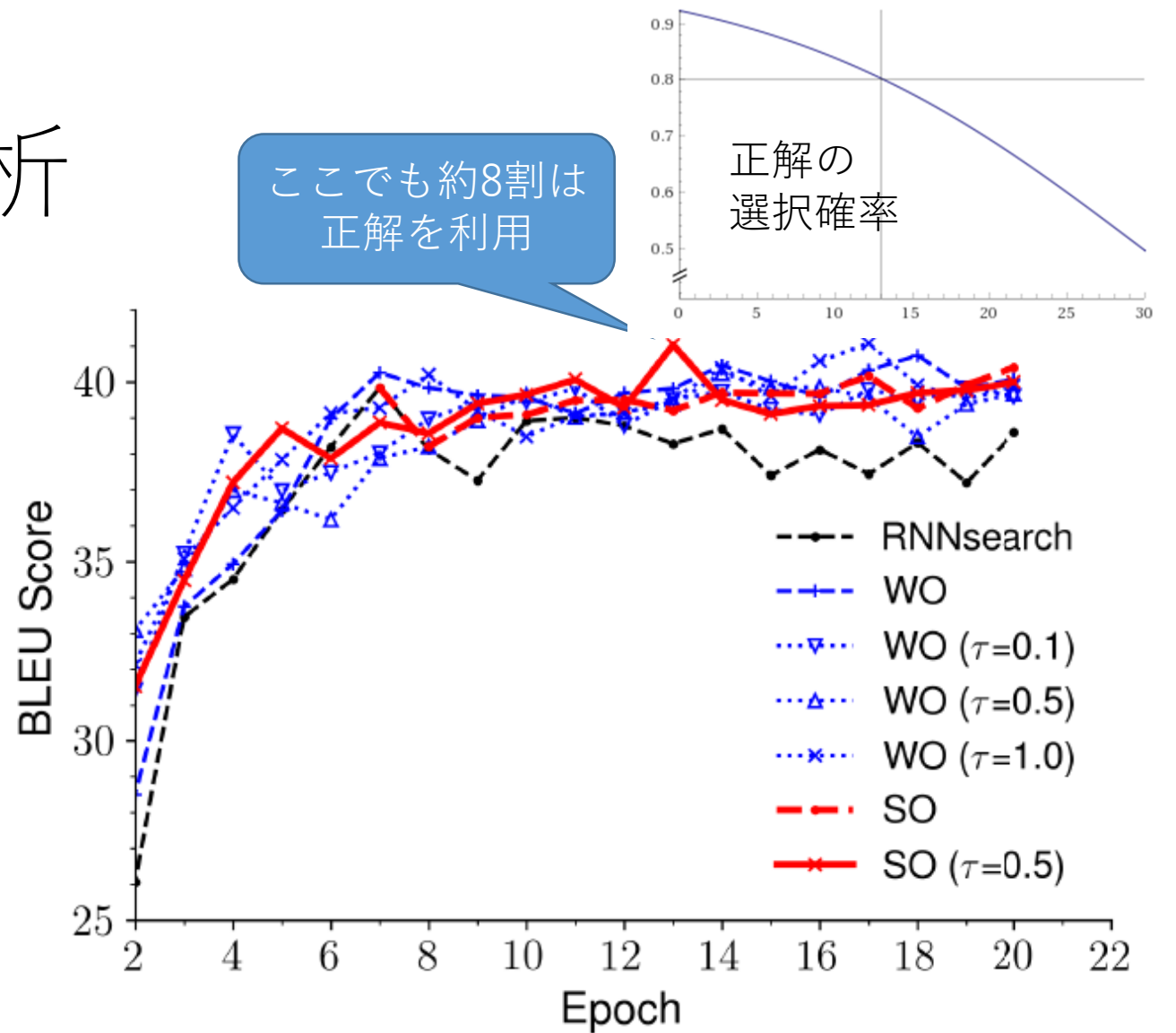
Systems	Average
RNNsearch	37.73
+ word oracle	38.94
+ noise	39.50
+ sentence oracle	39.56
+ noise	40.09

Zh→En

収束についての解析

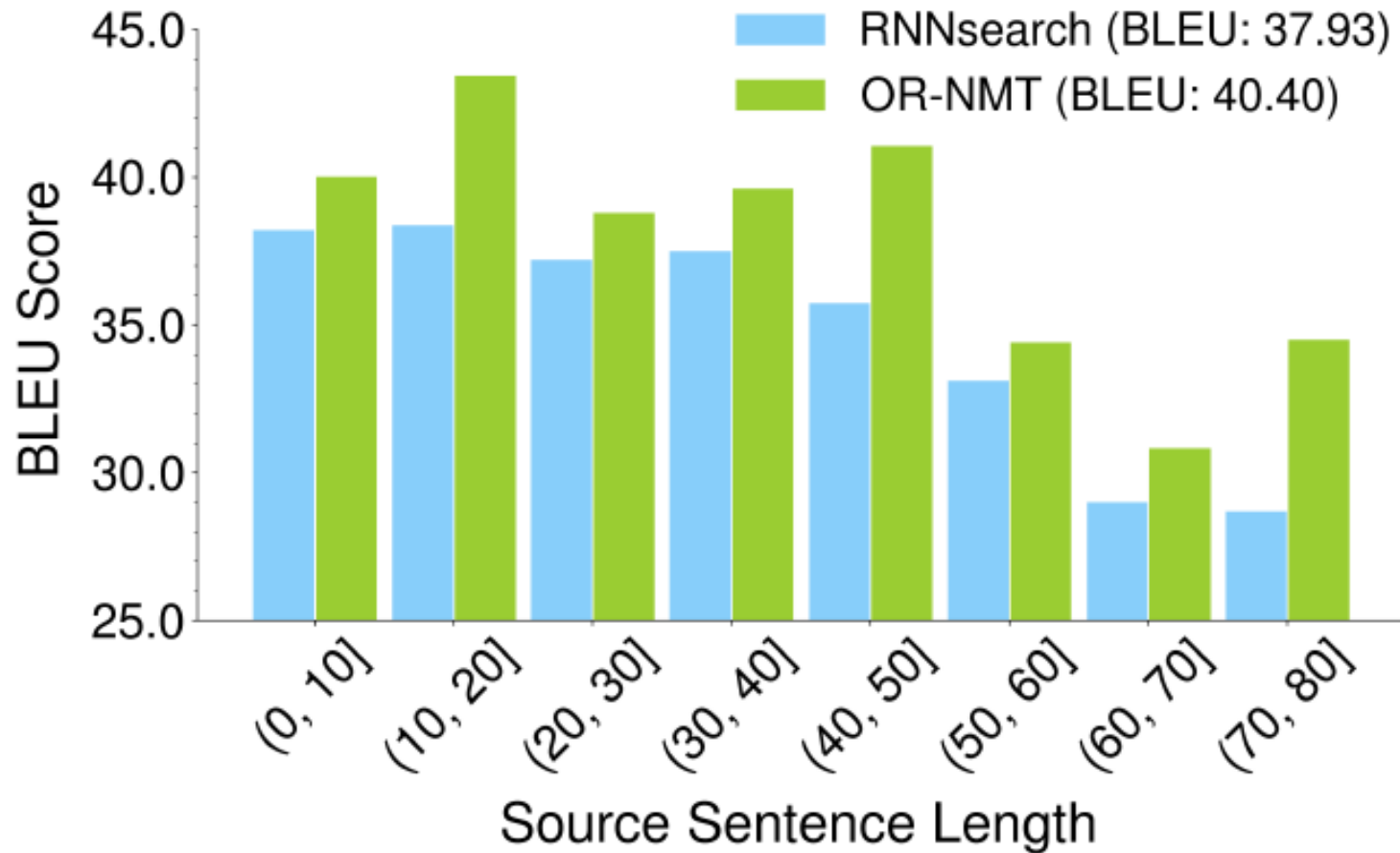


Zh→Enでの訓練ロス
提案手法はオーバーフィットしない



Zh→En開発データでのBLEU
RNNSearchは早めに収束
提案手法は伸び続けている

文長についての解析



Zh→EnテストデータでのBLEUの比較
提案手法は全文長で精度向上

まとめ

- Gumbelノイズの載った文レベルでの予測をTeacher forcingの入力と確率的に入れ替えることで2つのデータセット・モデルいずれについても、BLEUを有意に(1.3-2.3ポイント)向上