

QA レポート

システム情報工学研究科 200820634 越川満

研究題目：整数計画法を用いたフレーズ対応最適化による翻訳システムの改良
主任指導教官：山本幹雄 教授
発表日時：2008年10月2日

質問 1： BLEU を評価手法に使うのは妥当なのか？ 他の評価法で提案手法の有効性を示すことは出来ないのか？

発表時の回答：

統計的機械翻訳の分野では最も広く用いられている手法であるが、人手評価との相関は60%程度である。しかし現在のところ、BLEU に代わる評価手法がないため最もよく用いられている。したがって、BLEU を改善することは提案手法の有効性を示す上で大事であると考え。提案手法は BLEU を向上させなかったが、探索エラーを減少させたという主張は可能である。

補足：

本研究の目的は翻訳精度の向上であるので、BLEU による評価結果を示したこと自体は適切であると考え。しかし、提案手法が直接的に改善するのは各翻訳候補の確率であるため、ベースラインと提案手法とで各翻訳候補の確率をどれだけ改善したかすなわち探索エラーをどれだけ減少させることができたかを調査し発表していれば、提案手法の有効性をより明確に示すことができたと思われる。

質問 2： BLEU はどのようにして翻訳精度の評価を行うのか？

発表時の回答：

正解文と翻訳結果の一致率を表す。大まかに説明すると、1 単語、2 単語、3 単語、4 単語の一致率を求めその幾何平均を求める手法である。

改善した回答：

“単語”を“単語連鎖”に訂正。

質問 3： 歪み確率は何を表すのか？ 歪みや歪み確率の増加は翻訳精度の向上に繋がるのか？

発表時の回答：

歪み確率は語順変化の生じやすさを表す。正しい歪みほど歪み確率は高くなる。確率が高いほど翻訳精度は高い。

改善した回答：

歪み（語順変化）を一切許さない場合、翻訳後の語順は図 1 に示すように不自然になってしまう場合が多い。歪みを許す場合には、図 2 のように翻訳後の語順をより翻訳先の言語らしくすることができる。

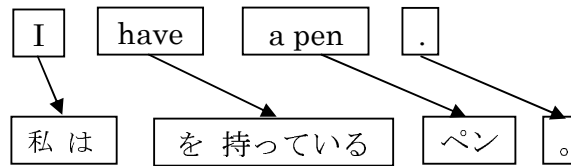


図 1：歪みを一切許さない場合

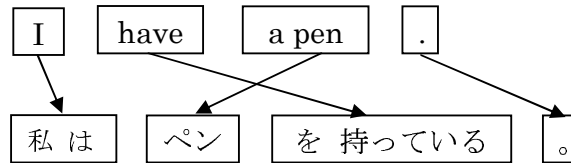


図 2：歪みを許す場合

歪みを増やすことで必ずしも翻訳精度が向上されるわけではないが、歪みを許すことで語順変化の生じる言語間（例えば日本語－英語）での翻訳精度向上が期待できる。歪みが正しいかどうかは、その歪みの生じやすさを表す指標である歪み確率を用いることにより評価可能であり、歪み確率が高いほどその歪みは尤もらしいと言える。

質問 4: 提案手法 1 (歪みモデルあり) とベースライン、提案手法 2 の比較は行わないのか？

発表時の回答：

提案手法 1 では歪み確率を一次の項として定式化に組み込めない。二次以上の項として提案手法 1 に組み込むことは可能だが、非線形となるため数理計画法の Solver で解を求めることが難しくなる。そのため提案手法 1 とベースラインとの比較は行っていない。

質問 5: 歪みモデルを考慮しない場合の提案手法 1 とベースラインを比較するとどのようになるのか？（結果が出ているのであればその理由も含めて）

発表時の回答：

本発表では割愛したが予備実験を行っており、提案手法 1 がベースラインに劣ることが分かっている。これはベースラインが歪みモデルを考慮しながら翻訳を行ったのに対して、提案手法 1 はフレーズ対応最適化時に歪みを考慮しなかったためであると考えられる。

補足：

この比較実験では、提案手法 1 により求めたフレーズ対応に対して後付けで歪み確率を計算している。最適化時の目的関数が翻訳確率のみであるため、翻訳確率+歪み確率を改善している保障がないことが歪み確率を考慮しない場合の提案手法 1 の問題点である。

質問6: フレーズ対応最適化前後で確率が変わっているのはなぜか？ ある候補の確率が上がることにより他の候補の確率は下がらないのか？

発表時の回答:

同じ翻訳結果を導くフレーズ区切り・対応は複数存在する。フレーズベースデコーダ Moses が各翻訳候補に対して求めたフレーズ対応はヒューリスティック探索で求めたものであるため、よりよい対応が存在する可能性がある。そこで最適なフレーズ対応を求めるシステムにより各翻訳候補のフレーズ対応を最適化する。フレーズ対応が変化することにより歪み確率やフレーズ翻訳確率が変化するため、フレーズ対応最適化前後で各翻訳候補の確率は変化する。各翻訳候補の確率はよくなることはあっても悪くなる事はない。

補足:

提案手法では Moses の翻訳候補上位 n 個を利用しているが、上位 n 個の確率の合計が 1 となるわけではない。統計的機械翻訳ではある原言語文 f に対する翻訳として、すべての目的言語で表現された文 e を考慮する。つまり $P(e|f)$ を全ての e について足しこむと 1 になる。したがって、提案手法によりフレーズ対応最適化が行われ確率が改善された翻訳候補がそうでない候補の確率に影響を及ぼすことはない。

質問7: 英→日、日→英以外の翻訳では、提案手法はどのような結果になると思われるか？

回答:

本研究の提案手法は特定の言語に依存した方法ではないため、どの言語間でもほぼ同様であると考えられる。ただし語順変化の少ない言語間、例えば英語→フランス語などではフレーズ対応の改善の余地が少なく、そのため提案手法の効果が薄れる可能性がある。日本語→英語以外の言語対についても実験を行うかどうかは検討中である。

質問8: アプリケーションとして Web 等に公開する予定はあるのか？

回答:

学習データとして NTCIR-7 特許文翻訳タスク[1]で配布される特許文集合[2]を用いており著作権等の問題から、学習データから学習した確率的翻訳規則を含む本研究のシステムそのものを Web 等で公開して使用してもらうことは難しいと思われる。ソースコードのみを公開することも考えられるが、最適なフレーズ対応を求める際に市販の数理計画問題 Solver である CPLEX[3]を用いており、実行環境がかなり制限されてしまう。上記の問題から、現在のところ本システムの公開は考えていない。

自分の発表に対する反省点:

全体を通してスクリーンの方を向いていることが多く、また教室の大きさを配慮せず音量が不十分だったため、聴衆が聞き取りづらかったと思われる。話すスピードについては適切であったという評価が多かったが、話の途中でつまるが多かった。話す速さだけでなく他の部分についても聴衆の聞き取りやすさを配慮する必要があると思われる。

る。

スライド 1 枚あたりの情報量が多いため、各スライドで注目すべき点がどこなのかわかりづらくなっていた。これは話し方次第でフォローできたかもしれないが、話し方に抑揚があまりなかったため重要な点を強調できていなかった。スライド中で提案手法の式を紹介したが、その説明が不十分であったと思われるので具体例を用いて直感的に説明するか、時間をかけてより詳しく解説すればよかったと思われる。

質問に対する回答は概ね適切だったと思われるが、やや質問者の意図を外した回答をすることも見受けられた。より質問の意味を適切に捉えることができるよう気をつけていきたい。

今後は上記の事柄について注意して、よりよい発表ができるよう心掛けていきたい。

参考文献：

[1] Atsushi Fujii, Takehito Utsuro, Mikio Yamamoto and Masao Utiyama. “Definition of Patent Translation Task at NTCIR-7”. 2007.

[2] Masao Utiyama, Mikio Yamamoto, Atsushi Fujii and Takehito Utsuro. “Description of Patent Parallel Corpus for NTCIR-7 Patent Translation Task”, 2007.

[3] ILOG. *ILOG CPLEX 11.0 User's Manual*. 2007.