



日英複文翻訳

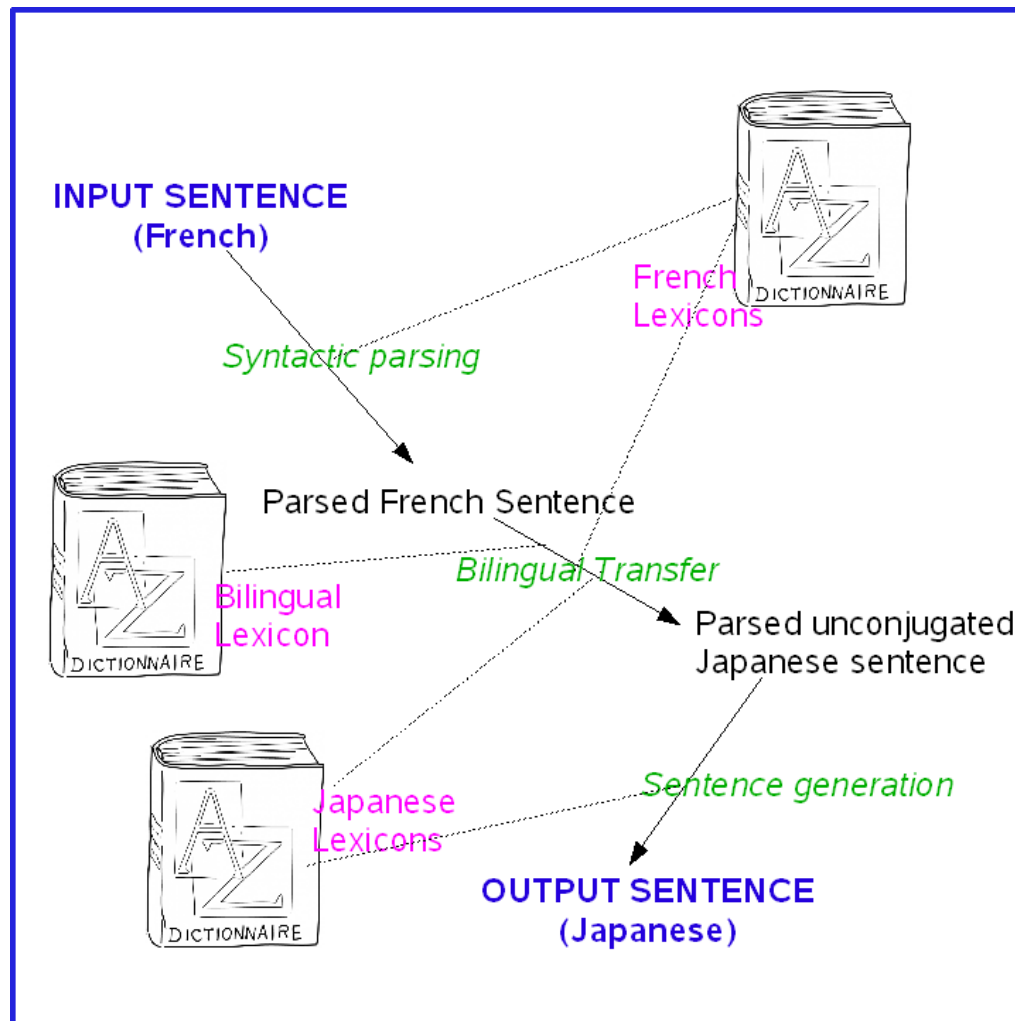
Alexis Kauffmann

目次

1. 方針:システムの概要
2. 複文翻訳:
 1. いろいろな文の構文
 2. 規則の定義
 3. 語彙のデータベースの改善
 4. 単語の選択の改善
 5. 規則の改善
3. 実研と効果
4. まとめ
5. 今後の課題



方針: システムの概要



方針:システムの概要(2)

- Its2はLBMTのシステムだ。RBMTに似ている。
- 構文解析モジュール。
- 構文木の変換のモジュール。実装した言語的な規則を使う。
- 出力文の生成のモジュール。
- 単一言語と二言語のデータベースを使う。

=>複文翻訳のために作った規則とデータベースの改善の提示。



目次

1. 方針:システムの概要
2. 複文翻訳:
 1. いろいろな文の構文
 2. 規則の定義
 3. 語彙のデータベースの改善
 4. 単語の選択の改善
 5. 規則の改善
3. 実研と効果
4. まとめ
5. 今後の課題



複文翻訳: いろいろな文の構文

- 従属節
- 副次節:
 - I will come because you will prepare a good dinner.
 - I will come if you prepare a good dinner.
- 官僚層の従属節:
 - I think that you will prepare a good dinner.
- 関係詞節:
 - I like the good dinners that you always prepare .



複文翻訳:

いろいろな文の構文(2)

- 等位構造接続:
 - I will come and you will prepare a good dinner.
 - I will come or you will prepare a good dinner.
- 並列構造:
 - I will come, you will prepare a good dinner.
- ミックス:
 - I will come, and you will prepare a good dinner, if it doesn't rain.



複文翻訳: 規則の定義

○ 副次節:

- I will come because you will prepare a good dinner.
- I will come if you prepare a good dinner.

○ 官僚層の従属節:

- I think that you will prepare a good dinner.

=>従属節は、主節の主語の後で置く。従属節の後で、接続詞か接続助詞。

たとえば、“山田さんは田中さんがいい定食を作ります。”

○ 関係詞節:

- I like the good dinners that you always prepare.

=>関係詞節は名詞の前に置いて、英語の代名詞の翻訳しない。



複文翻訳: 規則の定義(2)

○ 等位構造接続:

- I will come and you will prepare a good dinner.
- You will prepare a good dinner so I will come.
- I will come but you will not prepare a good dinner.

=>接続詞によって:

- 項 動詞一て、(接続詞) 項 .

たとえば、“山田さんが来て、田中さんはいい定食を作る。”

- 項 動詞 接続助詞(、) 項 .

たとえば、“山田さんが来るけど、田中さんはいい定食を作らない。”

- 項。接続詞(、)項 .

たとえば、“田中さんはいい定食を作る。だから山田さんが来る。”

○ 並列構造:

- I will come, you will prepare a good dinner.

=> 項 動詞一て、(接続詞) 項 .



複文翻訳:

語彙のデータベースの改善

- 接続詞(と接続助詞と接続助詞)を類纂するために。
- ダーターで、接続詞の接続詞と接続詞の位置がわかる。
- ダーターと言語的な話しで、
 - =>単一言語と二言語のデータベースを類纂する。
 - =>規則も類纂する。



複文翻訳: 単語の選択の改善

- たとえば: “if” => 動詞+なら、動詞+か。
- たとえば: “so” => だから、そして。



複文翻訳: 規則の改善

- たとえば: “provided that” →ただし...動詞一れば
- たとえば: “if” →動詞+なら,
名詞+“だ”+なら。 →名詞+なら。



目次

1. 方針:システムの概要
2. 複文翻訳:
 1. いろいろな文の構文
 2. 規則の定義
 3. 語彙のデータベースの改善
 4. 単語の選択の改善
 5. 規則の改善
3. 実研と効果
4. まとめ
5. 今後の課題



実研と効果

- 旅行の英語のBLOG (主に複文) の機械翻訳
=>48文で4文が改善して、1文が悪くなった。
=>文が6 %が改善した。
- 旅行のphrasebook (主に短い文)の機械翻訳
=>138文で6文が改善した。
=>文が4%が改善した。



まとめ

- 機械翻訳の改善は、わるくはないですけど、また複文の翻訳と、問題がある。
- もっと他切な問題はモダリチの翻訳と正しくないパーシング。



今後の課題

- モダリチの翻訳
- 動詞のsubcategorisationの改善

