

単語並び換えモデルを考慮した 統計的階層句機械翻訳システム

林 克彦^{†1,*1} 塚田 元^{†2} 須藤 克仁^{†2}
Kevin Duh^{†2} 山本 誠一^{†1}

階層句機械翻訳では言語間の大域的な単語並び換わりを同期文脈自由文法によって表現することができるが、日英のような文法の大きく異なる言語間の翻訳ではその単語並び換わりを評価するためのモデルが充分であるとは言えない。本稿では階層句機械翻訳におけるこの問題を解決するために単語並び換えモデルを特徴量に導入することを提案する。提案手法では同期文脈自由文法の各ルールに並び換えされた原言語の文字列を追加することで、探索中の単語並び換えモデルの計算を効率良く行っている。日英旅行会話データを用いた実験では従来手法に比べ、提案手法の方がより適切な単語並びの翻訳結果を選択することができ、より高い翻訳精度を達成することができた。

Hierarchical Phrase-based Machine Translation with Word-based Reordering Model

KATSUHIKO HAYASHI,^{†1,*1} HAJIME TSUKADA,^{†2}
KATSUHITO SUDOH,^{†2} KEVIN DUH^{†2}
and SEIICHI YAMAMOTO^{†2}

Hierarchical phrase-based machine translation can capture global reordering with synchronous context-free grammar, but has little ability to evaluate the correctness of word orderings during decoding. We propose a method to integrate word-based reordering model into hierarchical phrase-based machine translation to overcome this weakness. Our approach extends the synchronous context-free grammar rules of hierarchical phrase-based model to include re-ordered source strings, allowing efficient calculation of reordering model score during decoding. Our experimental results on Japanese-to-English basic travel expression corpus showed that the BLEU scores obtained by our proposed system were better than those obtained by a standard hierarchical phrase-based machine translation system.

1. はじめに

現在、階層句機械翻訳³⁾は有力な統計的機械翻訳²⁾手法の1つとなっている。階層句機械翻訳では対訳コーパスから自動獲得された同期文脈自由文法 (SCFG: Synchronous Context Free Grammar) に基づいて翻訳が行われ、原言語と目的言語間の単語の並び換わりは SCFG によって表現される。しかし、現在までこの SCFG によって作り出される単語の並び換えを明示的に制約するようなモデルは導入されておらず、現状のモデルでは翻訳候補の単語並びを十分に評価することができていない。

一方、通常の句に基づいた統計的機械翻訳では様々な語彙的並び換えモデル^{8),12),20)}の適用が提案されており、一定の成果を上げている。これらのような語彙的並び換えモデルを階層句機械翻訳に導入することができれば、より適切な単語並びをした翻訳結果を選択することが可能になると考えられるが、通常の句と異なり、階層句では非終端記号をルール内に持つため、これらのモデルを直接導入することは容易ではない。

大域的な単語並び換えを扱うため、通常の句に基づいた統計的機械翻訳では翻訳前処理として原言語の単語を目的言語側の単語並びに並び換えてから翻訳する手法も提案されている^{6),9),21),23)}。これらの前処理アプローチは階層句機械翻訳システムに導入することも原理的には可能であるが、並び換えに利用するルールやモデルが翻訳に用いている言語モデルや翻訳モデルなどの情報を活用できないという問題がある。また、階層句機械翻訳モデルにおける同期文脈自由文法から表現される単語並び換えの利点を損なうという問題も生じる。

本稿では上記した階層句機械翻訳の問題点を改善するためにモデルの素性値として単語並び換えモデルを導入するための手法を提案し、より適切な単語並びの翻訳結果を出力することができる翻訳システムの開発を行った。単語並び換えモデルとしては Tromble と Eisner によって提案された単語並び換えモデル²¹⁾を利用した。また、同期文脈自由文法の各ルール内に目的言語の文字列並びと同じ並びにした原言語文字列を追加することで、このモデルスコアを翻訳機の探索過程で効率良く計算を行った。本提案手法では前処理としてのモデ

†1 同志社大学

Doshisha University

*1 現在、奈良先端科学技術大学院大学

Presently with Graduate School of Nara Institute of Science and Technology

†2 NTT コミュニケーション科学基礎研究所

NTT Communication Science Laboratories

ルとは異なり、翻訳解の探索中にモデルを利用するため、階層句機械翻訳モデルや言語モデルといった他のモデル情報と統合した形で翻訳候補を評価することができるという利点がある。

以下、第2節では提案手法の概念について述べ、第3節において階層句機械翻訳の概要と提案手法のモデルについて述べる。第4節では実験結果について述べ、第5節では関連研究、第6節でまとめと今後の課題について記す。

2. 提案する翻訳手法の概念について

翻訳前処理での並び換えによるアプローチでは入力文を S 、並び換えされた入力文を S' 、目的言語文を T とすると

$$S \rightarrow S' \rightarrow T \quad (1)$$

として表現される。前処理アプローチでは一度並び換えされた原言語文を修正することができないため、前処理での誤りが翻訳に大きな問題を及ぼすことになる。Li はこれを解決するため、 K -best による前処理アプローチを提案したが⁹⁾、 K -best 出力を行ったとしても、並び換えモデルだけを利用して S' は生成されるため、翻訳モデルや言語モデルなど他のモデル情報を全く考慮できていないという問題がある。

本稿では並び換えモデルをモデルの素性値として導入することで、階層句機械翻訳の翻訳モデル (や他のモデル) から起こる単語並び換わりをそれらのモデルと統合した形で探索中に評価する。本稿で提案する手法の概念は以下のように表すことができる。

$$S \rightarrow (S', T). \quad (2)$$

3. 階層句機械翻訳モデルの拡張

3.1 階層句機械翻訳モデル

階層句機械翻訳³⁾では次のように表現される同期文脈自由文法に基づいて翻訳が行われる。

$$X \rightarrow \langle \gamma, \alpha, \sim, w \rangle. \quad (3)$$

ここで X は非終端記号、 γ は非終端記号と原言語側の終端記号から成る文字列、 α は非終端記号と目的言語側の非終端記号から成る文字列、 \sim は γ と α の中にある非終端記号の対応、 w はルールに与えられた重みとなる。

入力文 S が与えられたとき、このモデルによる翻訳は

$$\hat{T} = T \left(\operatorname{argmax}_{D: S(D)=S} w(D) \right) \quad (4)$$

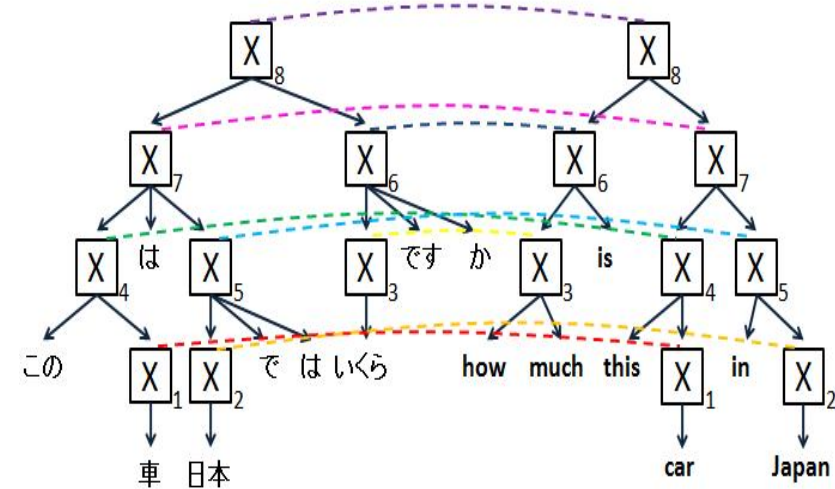


図1 階層句機械翻訳システムによる翻訳例

として行われる。ここで D は導出を表し、デコーダはスコア $w(D)$ の最も高い目的言語文 $T(D)$ を探索する。 $S(D)$ は導出 D の下での入力文を表す。Chiang (2007) では同期文脈自由文法における原言語側のルールに従って入力文を CKY アルゴリズムで構文解析することで目的言語文を探索する。また、枝刈りには Cube pruning³⁾ が用いられる。図1は階層句機械翻訳システムによる翻訳例を示したものである。

また、統計的機械翻訳ではモデルとして複数の素性値を重みで線形結合した線形対数モデルを用いることが一般的となっている¹⁶⁾。

$$w(D) \propto \prod_i w_i(D)^{\lambda_i} \quad (5)$$

λ は各特徴量に対する重み係数である。3.4節で詳細に触れるが、言語モデルは正しいスコアを仮説展開途中に計算することができないため、Chiang (2007) の提案している探索中での N -gram 適用手法を本稿では利用している。

3.2 同期文脈自由文法のルール拡張

提案手法では (2) で示したように目的言語文 T を生成すると共に、目的言語文と同じ単

Standard SCFG	$X \rightarrow \langle X1 \text{ は 人生 の } X2 \text{ だ, } X1 \text{ is } X2 \text{ of life} \rangle$
SCFG (move-to-front)	$X \rightarrow \langle X1 \text{ は 人生 の } X2 \text{ だ, は } X1 \text{ だ } X2 \text{ の 人生, } X1 \text{ is } X2 \text{ of life} \rangle$
SCFG (attach)	$X \rightarrow \langle X1 \text{ は 人生 の } X2 \text{ だ, } X1 \text{ は だ } X2 \text{ の 人生, } X1 \text{ is } X2 \text{ of life} \rangle$

表 1 提案手法におけるルール例. “move-to-front” は Tromble と Eisner (2009) のアルゴリズム, “attach” は AI-Onaizan と Papineni (2006) のアルゴリズムによる結果を示している.

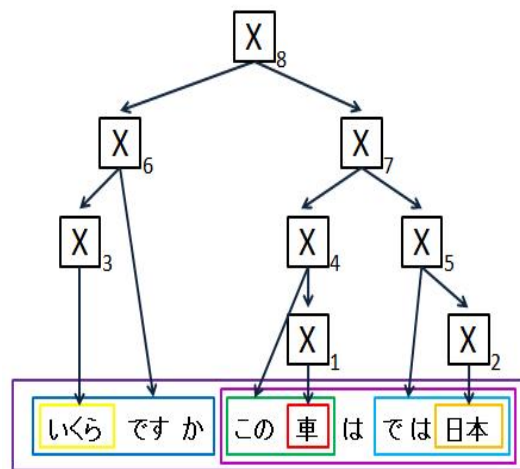


図 2 提案システムから出力される目的言語文と同じ構造に並び換えされた原言語文

語並びとなる原言語文 S' を生成する必要がある. このために (3) で表したルールを以下のように拡張する.

$$X \rightarrow \langle \gamma, \gamma', \alpha, \sim, w \rangle \quad (6)$$

γ' は γ 内の非終端記号と終端記号を α と同じ文字並びに換えた文字列である. このモデルに基づいた翻訳は

$$\hat{T} = (\hat{S}', \hat{T}) = (S', T) \left(\underset{D: S(D)=S}{\operatorname{argmax}} w(D) \right). \quad (7)$$

として表すことができる. この方式では目的言語文 T と同時に並び換えされた原言語文 S' を生成することができるので, 以下で説明する単語並び換えモデルを探索中に効率良く計算することが可能となる. 図 2 は図 1 の導出の例に対して提案システムが出力する並び換え

された原言語文の例を示している. 図 2 で示した並び換えされた原言語文を持つ構造は図 1 で出力された目的言語文と同じ構造になる.

ルールに γ' を追加するため, ルール獲得時に GIZA++¹⁴⁾ で学習した翻訳モデルから求めた Grow-diag-final⁸⁾ 単語アライメントを利用して γ を α と同じ文字並びに変換する. 並び換えのための手法としては Tromble と Eisner (2009) の手法と AI-Onaizan と Papineni (2006) の手法をルールに対して適用した (両手法とも原論文では文に対して適用している). 両手法ともに原言語側の単語に対して単語アライメントで対応付いた目的言語単語の文内のインデックスを与え, それに基づいてソートすることで並び換えを行う. 両手法の違いは対応関係がない原言語単語に対して, Tromble と Eisner (2009) では順序関係を保持したまま文の先頭に置く (“move-to-front”), AI-Onaizan と Papineni (2006) は直前の対応関係がある単語に添え付けた状態で移動させるという手法をとっている (“attach”).

表 1 は提案手法におけるルール例を示している. 階層句機械翻訳では既存の人手で作成された文法を用いずに対訳コーパスから自動で同期文脈自由文法を獲得するため, 非終端記号は開始記号 S と X のみであり, また CKY 構文解析を適用するため, 非終端記号はルール内に 2 つまでとしている. 原言語側の文字列を ($X1$ は 人生 の $X2$ だ), 目的言語側の文字列を ($X1$ is $X2$ of life), これらの単語アライメントを $a=((\text{人生}, \text{life}), (\text{の}, \text{of}), (\text{だ}, \text{is}))$ としたとき, 原言語側の対応関係のある単語と非終端記号は目的言語単語のインデックスに従ってソートすることで並び換えされる. 原言語側の対応関係がない単語 (は) は先頭に置か (“move-to-front”), 直前の対応関係のある単語か非終端記号 (例では直前の非終端記号 $X1$) に添え付けた状態で移動させる (“attach”).

3.3 単語並び換えモデル

前述したように, 統計的機械翻訳は一般に対数線形モデルによってモデル化される. 本稿で利用する単語並び換えモデルは以下で示す Tromble と Eisner によって提案されたモデルであり, そのモデルスコア $score(S')$ を対数線形モデルの素性値に加える.

Uni-gram 素性	Bi-gram 素性
$s_r, s\text{-}pos_r$	$s_r, s\text{-}pos_r, s_l, s\text{-}pos_l$
s_r	$s\text{-}pos_r, s_l, s\text{-}pos_l$
$s\text{-}pos_r$	$s_r, s_l, s\text{-}pos_l$
$s_l, s\text{-}pos_l$	$s_r, s\text{-}pos_r, s\text{-}pos_l$
s_l	$s_r, s\text{-}pos_r, s_l$
$s\text{-}pos_l$	s_r, s_l
	$s\text{-}pos_r, s\text{-}pos_l$

表 2 単語並び換えモデルに使用する素性の種類

周辺形態素素性
$s\text{-}pos_r, s\text{-}pos_r + 1, s\text{-}pos_l - 1, s\text{-}pos_l$
$s\text{-}pos_r - 1, s\text{-}pos_r, s\text{-}pos_l - 1, s\text{-}pos_l$
$s\text{-}pos_r, s\text{-}pos_r + 1, s\text{-}pos_l, s\text{-}pos_l + 1$
$s\text{-}pos_r - 1, s\text{-}pos_r, s\text{-}pos_l, s\text{-}pos_l + 1$

表 3 文脈素性の一例

$$score(S') = \sum_{i,j:1 \leq i < j \leq n} B[s'_i, s'_j] \quad (8)$$

$$B[s'_i, s'_j] = \theta \cdot \phi(s'_i, s'_j) \quad (9)$$

ここで n は原言語文 $S' (= (s'_1 \dots s'_n))$ の長さ, θ は重みベクトル, ϕ は素性ベクトルを示す (対数線形モデルの重みや素性とは異なる). $B[s'_i, s'_j]$ は s'_i が s'_j よりも前にあった場合のスコアであり, 直感的には単語 s'_i が s'_j よりも前に来る頻度が高ければ高いほど, その値も大きくなる.

本稿では素性関数 ϕ に Tromble と Eisner (2009) 同様に McDonald ら (2005) で係り受け解析のために提案された素性を利用する. ただし, 単語 (または品詞) の係り受け関係ではなく, 単語 (または品詞) の並び関係に対して作られる素性値である点が McDonald ら (2005) のものとは異なる. 本稿で使用した素性の種類は表 2 に示す. 一方で, 表 3 に示した周辺形態素素性は予備実験の結果からあまり効果がなかったため, 本稿では使用していない. また, 周辺形態素素性のような文脈性を考慮した素性を利用しないことで 3.4 節において提案する単語並び換えモデルの実装が可能となる.

パラメータ θ の訓練にはパーセプトロン^{5),7)}を用いた. 訓練では正解となる S^* が必要となるが, これは第 3.2 節で説明した単語並び換えの手法を訓練用の対訳データに適用することで自動生成した. 単語並び換えモデルに基づいて生成する正解 S' は Tromble と Eisner

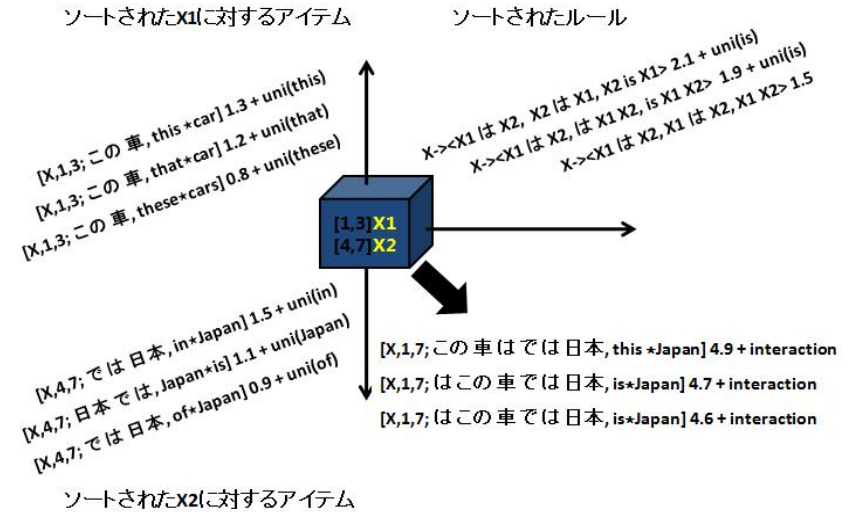


図 3 N-gram を考慮した Cube Pruning による探索の一例

(2009) で提案されている $O(n^3)$ で計算可能なボトムアップアルゴリズムを用いて生成した^{*1}.

3.4 探索過程での単語並び換えモデルスコアの計算法

前述したように階層句機械翻訳システムの翻訳探索は SCFG の原言語側のルールに従って入力文を構文解析 (CKY 構文解析) することで行われる. 図 3 は図 1 の例で用いた入力文 (日本語文) の部分文字列 (この車は日本では) に対して行われる探索の一例を示したものである. 探索は基本的に各アイテムとルールが現在持つコストを足すだけで良いが, N-gram 言語モデルを考慮する場合, 各アイテムとルールが結合されるまでそのスコアを知ることにはできない. よって, 各アイテムには結合時まで未計算となる N-gram スコア (例では 2-gram としている) を結合時に計算するための目的言語文字列の先頭と末尾の文字が記

*1 アルゴリズムの詳細は Tromble と Eisner の論文に記載されているのでここでは触れないが, アルゴリズムは CKY 構文解析と同様の手続きで動作し, 入力文字列のある範囲と範囲の置き換えを行うか行わないかをモデルスコアの差分で調べながらもっともスコアの良くなる文字列の並び換えを探索する (置き換ええない場合は差分が 0 となるので置き換えた場合のモデルスコアが正となれば置き換えが起こる).

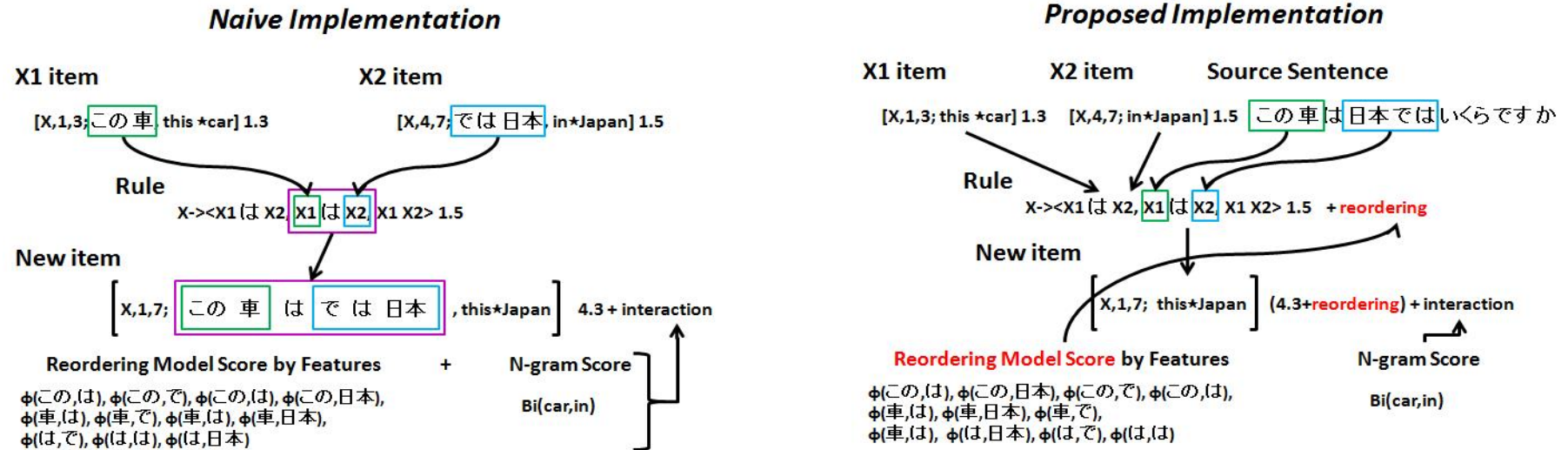


図 4 単語並び換えモデルを導入するための素朴な実装法と提案実装法の比較

憶されており、また、各アイテムとルールでは未計算部分の N -gram スコア (uni) をヒューリスティクス値として与えている。interaction は各アイテムとルールの結合時にわかるコストであり、未計算であった正しい N -gram スコアに当たる。

3.3 節で説明した単語並び換えモデルを導入するため、本稿では 2 種類の実装法を提案する。一つ目は素朴な実装法であり、図 4 の左図が示す実装となる。素朴な実装法では単語並び換えモデルを N -gram モデルと同様に interaction として扱うことで、ルールとアイテムが結合されたときに計算が行われるように実装している。

図 4 の右図は提案実装法を示している。提案実装法は単語並び換えモデルの素性として文脈素性を用いていない場合 (や単語間の距離を考慮しない場合) にのみ実装が可能となる。単語の前後 (や距離) を考慮する必要がない場合、結合するアイテムが持つスパンでの並び換えされた原言語文字列の並びに関わらず、結合されたときの単語並び換えモデルスコアは一定である。よって、結合に使われる全てのルールに対して、各アイテムが持つ並び換えされた原言語文字列の並びとは無関係に正しい単語並び換えモデルスコアを計算することが可能である。提案実装法では結合に使われるルールをソートする前に単語並び換えモデルの

スコアを各ルールに計算しておくという実装をとる。

4. 実験

4.1 目的

実験では単語並び換えモデルを階層句機械翻訳システムのモデルに導入することの有効性を示すために以下のシステムを比較した。

- ベースライン: 階層句機械翻訳 (Hiero) システム
- 前処理: Tromble と Eisner の前処理を行って、Hiero システムで翻訳を行う。
- Hiero システム + 単語並び換えモデル: Hiero システムに Tromble と Eisner(2009) のモデルを素性の 1 つとして用いる。

実験で使用した Hiero システムは特徴量として N -gram 言語モデル尤度 $P_{LM}(T)$, 単語翻訳確率 $P_w(\gamma|\alpha)$, $P_w(\alpha|\gamma)$, ルール翻訳確率 $P(\gamma|\alpha)$, $P(\alpha|\gamma)$, 単語ペナルティ, arity ペナルティ (ルールの適用回数) を持つ。“Hiero システム + 単語並び換えモデル” では前述した “move-to-front” と “attach” による 2 種類の単語並び換えモデルを適用した。前処理では

System	Metrics	
	BLEU	PER
ベースライン (Hiero)	28.09	39.68
前処理	17.32	45.27
Hiero + move-to-front	28.89	39.90
Hiero + attach	29.12	39.40

表 4 テストデータに対する BLEU と PER の値

“attach” モデルのみを適用した。

4.2 データセット

本実験は日英旅行会話データ (BTEC:Basic travel expression corpus) を用いて行った。日本語の単語並びは英語とは文法的に大きく異なるため、単語並び換えモデルの有効性を適切に評価できるタスク設定であると考えられる。訓練データは 208000 文で階層句の抽出、言語モデルの学習、単語並び換えモデルの学習に用いた。対数線形モデルは 1000 文 (参照訳各 1 文) の開発データを用いて誤り最小訓練法 (MERT:Minimum error rate training) で学習を行った¹⁵⁾。テストデータは 1000 文 (参照訳 1 文) とした。

4.3 実験結果

表 4 は各システムから得られた翻訳結果の BLEU¹⁷⁾ と PER (Position-Independent Word Error Rate)¹³⁾ の値を示している。BLEU 値は翻訳結果と参照訳の間の N -gram の重なりを幾何平均した値であり、PER は文章中での単語の位置を考慮しない単語誤り率である。BLEU は値が大きいほど、PER は値が小さいほど良い翻訳結果であることを示す。実験結果からは従来法のシステムに比べ、単語並び換えモデルを導入した提案システムの方が BLEU 値において明らかに有意な結果となった一方で、前処理による翻訳方式では著しく BLEU 値が悪化した。PER に対する実験結果は BLEU に対する実験結果とは異なり、従来法と提案法の間で大きな違いは見られなかった。この結果より提案法は主に単語の並びを改善したことがわかる。

4.4 考察

表 5 は単語並び換えモデルを考慮した提案システムが従来システムに比べて大局的な単語並び換えにおいて効果的であった翻訳結果の一部を示している。表 5 からは提案システムによる出力結果の方が従来システムによる翻訳結果よりも適切な単語並びの文を選択できていることがわかる。

さらなる分析のため、訓練データによって訓練した単語並び換えモデルを開発データに適用

S	アメリカで製品の販売を始める計画がありますか。
T_B	sales of product in america are you planning to start ?
T_P	are you planning to start products in the u.s. ?
R	do you plan to begin selling your products in the u.s. ?
S	この手紙は航空便で日本までいくらかかりますか。
T_B	this letter by airmail to japan . how much is it ?
T_P	how much does it cost to this letter by airmail to japan ?
R	how much will it cost to send this letter by air mail to japan ?

表 5 S は入力文、 T_B は従来システムの翻訳結果、 T_P は提案システム (attach) の翻訳結果とする。 R は参照訳とする。下線部は同じ意味となる対応箇所であり、大局的な単語並び換わりを示している。

した際の出力 S' (Tromble と Eisner のボトムアップアルゴリズムによって出力) と GIZA++ のアライメントから作り出した正解 S'^* を BLEU 値で比較すると 26.2 となった。単語並び換えモデルの精度がよくない原因の 1 つとして日英では対応関係の付かない単語が多いということが考えられる。実際、本実験で使用した訓練データの日本語文には目的言語文の単語に対応が付かない単語が約 34.7%含まれていた。対応が付かない単語が増えると単語並び換えモデルでそれらを扱うことは難しく、モデル精度が下がる大きな要因になると言える。前処理アプローチではこの低い精度のモデルを単独使ったため翻訳精度が著しく低下したが、我々の提案法は低い精度のモデルもその他のモデルと併用することで翻訳精度を改善することができた。

5. 関連研究

本稿での提案手法は原言語文の単語を目的言語文の単語と同じ単語並びにするという点で前処理アプローチ^{6),9),21),23)} と類似している。しかし、本稿の提案手法は翻訳機の探索過程で並び換えを行うという点で単語並び換えを翻訳の前に行う前処理アプローチとは異なる。通常の句を用いた統計的機械翻訳では多くの並び換えモデル^{8),12),20)} が提案されている。しかし、通常の句と階層句という違いがあるため、これらの研究で提案されたモデルを階層句機械翻訳に直接適用することはできない。

Shen ら (2008,2009) は単語の係り受け構造を階層句のルール上に導入した。Shen らの研究は係り受け構造を導入することで単語の並びにある種の制約を加えているという点で本稿の提案手法と共通する。しかし、提案手法では係り受けのような構文のアノテーションが必要ないという点で Shen らの手法とは異なる。また、構文のアノテーションが必要ない

ため、任意の原言語並び換えモデルを訓練することが可能である。

6. まとめと今後の課題

本稿では階層句機械翻訳システムに単語並び換えモデルを導入するための手法を提案した。階層句機械翻訳では SCFG に基づいて言語間の大局的な単語並び換わりを表現することができるが、単語並び換わりを評価するための明示的なモデルは導入されてこなかった。本稿では日英旅行会話タスクを用いて実験を行った結果、従来システムや前処理アプローチに比べ、提案手法の方がより高い翻訳精度を達成することができた。

本稿では階層句機械翻訳システム上での単語並び換えモデルの影響のみを対象として実験を行ったが、さらなる拡張としてルール翻訳確率 $P(\gamma'|\gamma)$, $P(\gamma|\gamma')$ を素性に追加することが考えられる。また、原言語文の並び換えを考慮するだけではなく、並び換えした目的言語文字列 α' をルールに加えることで原言語と同じ並びをした目的言語文を評価することも可能である。さらには、 S' の並びが良くなるほど、 T の並びはよくなると考えられるため、誤り最小訓練法時の目的関数に S' の BLEU を考慮することを検討している。今後の課題としては日英では対応関係の付かない単語が多くであるため、これを解決することが精度を向上させる上で重要なこととなる。

参 考 文 献

- 1) Al-Onaizan, Y. and Papineni, K.: Distortion models for statistical machine translation, *Proc. the 44th ACL*, pp.529–536 (2006).
- 2) Brown, P.F., Pietra, S. A.D., Pietra, V. D.J. and Mercer, R.L.: The mathematics of statistical machine translation: Parameter estimation, *Computational Linguistics*, Vol.19, pp.263–312 (1993).
- 3) Chiang, D.: Hierarchical Phrase-based Translation, *Computational Linguistics*, Vol.33, pp.201–228 (2007).
- 4) Chiang, D., Knight, K. and Wang, W.: 11,001 New Features for Statistical Machine Translation, *Proc. NAACL*, pp.216–226 (2009).
- 5) Collins, M.: Discriminative training methods for hidden Markov models, *Proc. of EMNLP* (2002).
- 6) Collins, M., Koehn, P. and Kucerova, I.: Clause restructuring for statistical machine translation, *Proc. the 43rd ACL*, pp.531–540 (2005).
- 7) Freund, Y. and Schapire, R.E.: Experiments with a New Boosting Algorithm, *Proc. of the 13th ICML*, pp.148–156 (1996).
- 8) Koehn, P., Axelrod, A., Mayne, A.-B., Callison-Burch, C., Osborne, M. and Tal-

- bot, D.: Edinburgh System Description for 2005 IWSLT Speech Translation Evaluation, *Proc. the 2nd IWSLT* (2005).
- 9) Li, C.-H., Zhang, D., Li, M., Zhou, M., Li, K. and Guan, Y.: A Probabilistic Approach to Syntax-based Reordering for Statistical Machine Translation, *Proc. the 45th ACL*, pp.720–727 (2007).
- 10) Li, Z. and Khudanpur, S.: A Scalable Decoder for Parsing-based Machine Translation with Equivalent Language Model State Maintenance, *Proc. ACL SSST* (2008).
- 11) McDonald, R., Crammer, K. and Pereira, F.: Spanning tree methods for discriminative training of dependency parsers, *Technical Report MS-CIS-05-11, UPenn CIS* (2005).
- 12) Nagata, M., Saito, K., Yamamoto, K. and Ohashi, K.: A Clustered Global Phrase Reordering Model for Statistical Machine Translation, *COLING-ACL*, pp.713–720 (2006).
- 13) Niesen, S., Och, F., Leusch, G. and Ney, H.: An Evaluation Tool for Machine Translation: Fast Evaluation for MT Research, *Proc. the 2nd International Conference on Language Resources and Evaluation* (2000).
- 14) Och, F. and Ney, H.: A systematic comparison of various statistical alignment models, *Computational Linguistics*, Vol.29, pp.19–51 (2003).
- 15) Och, F.J.: Minimum error rate training in statistical machine translation, *Proc. the 41st ACL*, pp.160–167 (2003).
- 16) Och, F.J. and Ney, H.: Discriminative training and maximum entropy models for statistical machine translation, *Proc. the 40th ACL*, pp.295–302 (2002).
- 17) Papineni, K.A., Roukos, S., Ward, T. and Zhu, W.-J.: Bleu: a method for automatic evaluation of machine translation, *Proc. the 39th ACL*, pp.311–318 (2001).
- 18) Shen, L., Xu, J. and Weischedel, R.: A new String-to-Dependency Machine Translation Algorithm with a Target Dependency Language Model, *Proc. ACL*, pp.577–585 (2008).
- 19) Shen, L., Xu, J., Zhang, B., Matsoukas, S. and Weischedel, R.: Effective Use of Linguistic and Contextual Information for Statistical Machine Translation, *Proc. EMNLP*, pp.72–80 (2009).
- 20) Tillman, C.: A unigram orientation model for statistical machine translation, *Proc. HLT-NAACL*, pp.101–104 (2004).
- 21) Tromble, R. and Eisner, J.: Learning Linear Ordering Problems for Better Translation, *Proc. the 47th ACL*, pp.1007–1016 (2009).
- 22) Watanabe, T., Suzuki, J., Tsukada, H. and Isozaki, H.: Online Large-Margin Training for Statistical Machine Translation, *Proc. EMNLP-CoNLL*, pp.764–773 (2003).
- 23) Xia, F. and McCord, M.: Improving a statistical mt system with automatically learned rewrite patterns, *Proc. the 18th ICON*, pp.508–514 (2004).