

概要

近年，大量の対訳学習文から自動的に翻訳規則を獲得し翻訳を行う統計翻訳が主流となっている．統計翻訳において，翻訳されない単語（未知語）が出力されることで翻訳精度の低下につながるという問題がある．この未知語を減少させるための手法として，対訳学習文などの学習データを新たに追加して行う未知語処理と，追加せずに行う未知語処理の2パターンが考えられる．前者の考えに基づいた日野らの研究 [1] では，新規の対訳句辞書を既存の対訳学習文に追加して未知語処理を行った．そして，未知語数の削減と翻訳精度向上が報告されている．しかし，日野らの手法では新規の学習データが必要になるためコストがかかる．後者の考えに基づいた藤原らの研究 [2] では，句に基づく統計翻訳において，フレーズテーブル作成時のヒューリスティックスを併用して未知語処理を行った．そして，未知語数の削減と翻訳精度向上が報告されている．しかし，藤原らの手法は，句に基づく統計翻訳器に限定されるため，他の統計翻訳器への応用が期待できない．

そこで，本研究では，他の統計翻訳器への応用を考慮した，新規の学習データを必要としない新たな未知語処理手法を提案する．具体的には，単語アライメントツール（GIZA++[3]）を用いて自動作成した対訳単語辞書による未知語処理を，一度生成した出力文に対する後処理の形で行う．提案手法は大きく分けて4つの手順で構成される．手順1では，任意の統計翻訳器における出力文から未知語を抽出し，未知語リスト1を作成する．手順2では，GIZA++を用いて対訳学習文から対訳単語を獲得し，対訳単語辞書を作成する．手順3では，手順2で作成した対訳単語辞書を用いて未知語リスト1を翻訳し，未知語リスト2を作成する．ここで，対訳単語辞書内に訳語が複数存在する場合は単語翻訳確率が最尤の単語を選択する．手順4では，未知語処理前の文に対して，未知語リスト2を適用する．以上の手法により，未知語数の削減と翻訳精度の向上を試みた．

実験の結果，1,644文中967文と1,847単語中1,134単語の未知語処理ができた．さらに，人手評価と自動評価の結果より文全体の翻訳精度が向上し，提案手法の有効性が確認できた．

目次

第1章	はじめに	1
第2章	日英統計翻訳	3
2.1	概要	3
2.2	句に基づく統計翻訳	3
2.3	言語モデル	4
2.4	IBM 翻訳モデル	5
2.5	句に基づく翻訳モデル	5
2.6	GIZA++	6
2.7	フレーズテーブル作成法	6
2.7.1	intersection	7
2.7.2	union	7
2.7.3	grow と grow-diag	8
2.7.4	final と final-and	8
2.8	デコーダ	9
2.9	パターンに基づく統計翻訳	9
2.10	評価手法	10
2.10.1	自動評価法	10
2.10.2	BLEU	10
2.10.3	METEOR	11
2.10.4	RIBES	12
2.10.5	人手評価法	12
第3章	先行研究	13
3.1	日野らの研究	13
3.1.1	日野らの手法の手順	13
3.1.2	日野らの手法の問題点	14

3.2	藤原らの研究	14
3.2.1	藤原らの手法の手順	14
3.2.2	藤原らの手法の問題点	15
第4章	提案手法	16
4.1	提案手法の概要	16
4.2	提案手法の手順	16
第5章	実験環境	18
5.1	統計翻訳器	18
5.2	実験データ	18
5.3	評価方法	19
第6章	実験結果	20
6.1	未知語数	20
6.2	未知語の翻訳精度	20
6.3	文全体の翻訳精度	21
6.3.1	人手評価結果	21
6.3.2	人手評価例	22
6.3.2.1	提案手法の例	22
6.3.2.2	提案手法×(差なし)の例	22
6.3.3	自動評価結果	23
第7章	考察	24
7.1	未知語の傾向	24
7.2	未知語の翻訳精度と文全体の翻訳精度の差	25
7.3	誤った未知語翻訳の解析	26
7.3.1	頻度1単語の問題	26
7.3.2	複合語の未知語の問題	27
7.4	IBM 翻訳モデルの比較	28
7.5	翻訳確率最尤の単語選択の効果	28
7.6	対訳句辞書を用いた未知語処理	29
7.7	デコーダ内とデコーダ外の未知語処理の差	32
7.8	追加実験1(句に基づく統計翻訳器での実験)	33

7.8.1	未知語数	33
7.8.2	未知語の翻訳精度	33
7.8.3	文全体の翻訳精度	33
7.8.3.1	人手評価結果	34
7.8.3.2	人手評価例	34
7.8.3.3	自動評価結果	36
7.8.4	Moses と PBSMT における文全体の翻訳精度の差	37
7.9	追加実験 2 (提案手法を拡張した実験)	38
7.9.1	拡張手法の手順	38
7.9.2	実験結果	40
7.9.2.1	未知語数	40
7.9.2.2	漢字未知語の翻訳精度	40
7.9.2.3	文全体の翻訳精度	41
7.9.2.4	人手評価結果	41
7.9.2.5	人手評価例	42
7.9.2.6	自動評価結果	44
7.9.3	拡張手法における文字翻訳の問題	44
7.9.3.1	複合名詞の問題	44
7.9.3.2	固有名詞の問題	44
第 8 章 おわりに		46

目次

2.1	日英統計翻訳の流れ	4
2.2	日英方向の単語対応	7
2.3	英日方向の単語対応	7
2.4	intersection の例	7
2.5	union の例	8
2.6	grow の例	8
2.7	grow-diag の例	8
2.8	grow-diag-final の例	9
2.9	grow-diag-final-and の例	9
2.10	パターンに基づく日英統計翻訳の流れ	9
3.1	日英統計翻訳における日野らの手法の流れ	13
3.2	日英統計翻訳における藤原らの手法の流れ	15
4.1	日英統計翻訳における提案手法の流れ	17
7.1	未知語の品詞分布	24
7.2	ループ処理の流れ	29
7.3	日英統計翻訳における拡張手法の流れ	39

表 目 次

2.1	フレーズテーブルの例	6
5.1	実験データ	18
5.2	単文コーパスの例	18
6.1	未知語数 (10,000 文中)	20
6.2	未知語の翻訳精度 (109 単語中)	20
6.3	正しく翻訳できた未知語の一例 (54 単語中)	21
6.4	正しく翻訳できなかった未知語の一例 (55 単語中)	21
6.5	提案手法の評価結果 (100 文中)	21
6.6	提案手法 の評価例 1	22
6.7	提案手法 の評価例 2	22
6.8	提案手法 の評価例 3	22
6.9	提案手法 × (差なし) の評価例 1	23
6.10	提案手法 × (差なし) の評価例 2	23
6.11	提案手法 × (差なし) の評価例 3	23
6.12	提案手法の自動評価結果. 精度が高い方を太字で示す	23
7.1	動詞と副詞の未知語例	24
7.2	仮に未知語の翻訳精度が 100% の場合に翻訳精度が向上する例	25
7.3	仮に未知語の翻訳精度が 100% の場合でも翻訳精度が向上しない例	25
7.4	頻度 1 単語の翻訳精度 (75 単語中)	26
7.5	正しい翻訳の一例 (33 単語中)	26
7.6	誤った翻訳の一例 (42 単語中)	26
7.7	頻度 2 以上の単語の翻訳精度 (34 単語中)	26
7.8	正しい翻訳の一例 (21 単語中)	27
7.9	誤った翻訳の一例 (13 単語中)	27

7.10	誤った未知語翻訳の内訳 (55 単語中)	27
7.11	複合語の未知語の翻訳失敗例	28
7.12	全対訳単語の精度 (ランダム)	28
7.13	頻度 1 単語の精度 (ランダム)	28
7.14	1 変数の単語レベル文パターンから作成した対訳句の精度	29
7.15	対訳句と 1 変数の句レベル文パターンの数	30
7.16	ループ処理で作成した対訳句の精度 (100 句中)	30
7.17	ループ数 1 の評価例	30
7.18	ループ数 2 の評価例	31
7.19	ループ数 8 の評価例	31
7.20	デコーダ外で処理した場合の未知語の翻訳精度 (213 単語中)	32
7.21	デコーダ内で処理した場合の未知語の翻訳精度 (約 210 単語中)	32
7.22	デコーダ外の未知語処理における人手評価結果 (100 文中)	32
7.23	デコーダ内の未知語処理における人手評価結果 (100 文中)	32
7.24	Moses の実験における未知語数 (10,000 文中)	33
7.25	Moses の実験における未知語の翻訳精度 (101 単語中)	33
7.26	提案手法 (Moses) の人手評価結果 (100 文中)	34
7.27	Moses の評価例 1	34
7.28	Moses の評価例 2	35
7.29	Moses の評価例 3	35
7.30	Moses × (差なし) の評価例 1	35
7.31	Moses × (差なし) の評価例 2	36
7.32	Moses × (差なし) の評価例 3	36
7.33	提案手法 (Moses) の自動評価結果. 精度が高い方を太字で示す	36
7.34	拡張手法の実験における未知語数 (10,000 文中)	40
7.35	拡張手法の実験における漢字未知語の翻訳精度 (104 単語中)	40
7.36	正しい翻訳結果の一例 (20 単語)	41
7.37	文字単位で正しい翻訳結果の一例 (16 単語)	41
7.38	誤った翻訳結果の一例 (64 単語)	41
7.39	拡張手法の人手評価結果 (100 文中)	41
7.40	拡張手法 の評価例 1	42
7.41	拡張手法 の評価例 2	42

7.42 拡張手法 の評価例 3	42
7.43 拡張手法 × (差なし) の評価例 1	43
7.44 拡張手法 × (差なし) の評価例 2	43
7.45 拡張手法 × (差なし) の評価例 3	43
7.46 拡張手法の自動評価結果. 精度が高い方を太字で示す	44
7.47 固有名詞の文字翻訳例	45

第1章 はじめに

機械翻訳において、人手で翻訳規則を定義し、翻訳を行うルールベース翻訳が一般的であった。しかし、人手で翻訳規則を定義するには、莫大なコストがかかる。また、言語毎に文法規則が異なるため、多言語への拡張が困難であった。そのため、近年では、大量の対訳学習文から自動的に翻訳規則を獲得し翻訳を行う統計翻訳が主流となっている。この統計翻訳において、翻訳されない単語（未知語）が出力されることで翻訳精度の低下につながるという問題がある。

この未知語を減少させるための手法として、対訳学習文などの学習データを新たに追加して行う未知語処理と、追加せずに行う未知語処理の2パターンが考えられる。前者の考えに基づいた日野らの研究[1]では、新規の対訳句辞書を既存の対訳学習文に追加して未知語処理を行った。そして、未知語数の削減と翻訳精度向上が報告されている。しかし、日野らの手法では新規の学習データが必要になるためコストがかかる。後者の考えに基づいた藤原らの研究[2]では、句に基づく統計翻訳において、フレーズテーブル作成時のヒューリスティックスを併用して未知語処理を行った。そして、未知語数の削減と翻訳精度向上が報告されている。しかし、藤原らの手法は、句に基づく統計翻訳器に限定されるため、他の統計翻訳器への応用が期待できない。

そこで、本研究では、他の統計翻訳器への応用を考慮した、新規の学習データを必要としない新たな未知語処理手法を提案する。具体的には、GIZA++[3]を用いて自動作成した対訳単語辞書による2種類の未知語処理を、一度生成した出力文に対する後処理の形で行う。提案手法は大きく分けて4つの手順で構成される。手順1では、任意の統計翻訳器における出力文から未知語を抽出し、未知語リスト1を作成する。手順2では、GIZA++を用いて対訳学習文から対訳単語を獲得し、対訳単語辞書を作成する。手順3では、手順2で作成した対訳単語辞書を用いて未知語リスト1を翻訳し、未知語リスト2を作成する。ここで、対訳単語辞書内に訳語が複数存在する場合は単語翻訳確率が最尤の単語を選択する。手順3では、未知語処理前の文に対して、未知語リスト2を適用する。以上の手法により、未知語数の削減と翻訳精度の向上を目指す。

本論文の構成を以下に示す。第2章で日英統計翻訳の概要について説明する。第3章

では先行研究について説明し，第4章で提案手法について説明する．そして，第5章では実験環境を，第6章で実験結果を示し，第7章で本研究の考察を述べる．

第2章 日英統計翻訳

2.1 概要

現在，統計翻訳器として句に基づく統計翻訳が最も主流となっている．句に基づく統計翻訳では，学習データとして対訳学習文を与えることで翻訳が行える．そのため，翻訳にかかるコストが低い．さらに，対訳学習文から対訳単語と単語翻訳確率を自動的に取得することが可能である．

一方，機械翻訳の一手法としてパターン翻訳がある．パターン翻訳とは，句辞書と文パターン辞書を用いて翻訳を行う手法である．特徴として，入力文が適切な文パターンと照合した場合，翻訳精度の高い文が得られることが挙げられる．しかし，人手で辞書を作成する必要があるためコストがかかる．そこで江木らは，統計的手法を用いて自動的に辞書を作成して翻訳を行う“パターンに基づく統計翻訳 [4]”を提案した．現在，文献 [5] より，パターンに基づく統計翻訳の翻訳精度は句に基づく統計翻訳の翻訳精度よりも高いことが報告されている．

2.2 句に基づく統計翻訳

句に基づく統計翻訳は句対応の翻訳モデルを用いる．また，原言語文を目的言語文に翻訳する場合に，隣接する複数の単語を用いて翻訳を行う．本研究では日英翻訳を行うため，日英統計翻訳を説明する．日英統計翻訳では，日本語文 j を入力文とした場合，翻訳モデル $P(j|e)$ と言語モデル $P(e)$ の全ての組み合わせから，確率が最大となる英語文 \hat{e} を探索し，出力文とする． E を探索する翻訳器をデコーダと呼ぶ．以下に基本的なモデルを示す．また，日英統計翻訳の流れを図 2.1 に示す．

$$E = \arg \max_e P(e|j) \quad (2.1)$$

$$\simeq \arg \max_e P(j|e)P(e) \quad (2.2)$$

ここで， $P(e)$ が単語の場合は“単語に基づく翻訳翻訳”のモデル， $P(e)$ が句の場合は“句に基づく統計翻訳”のモデルとなる．

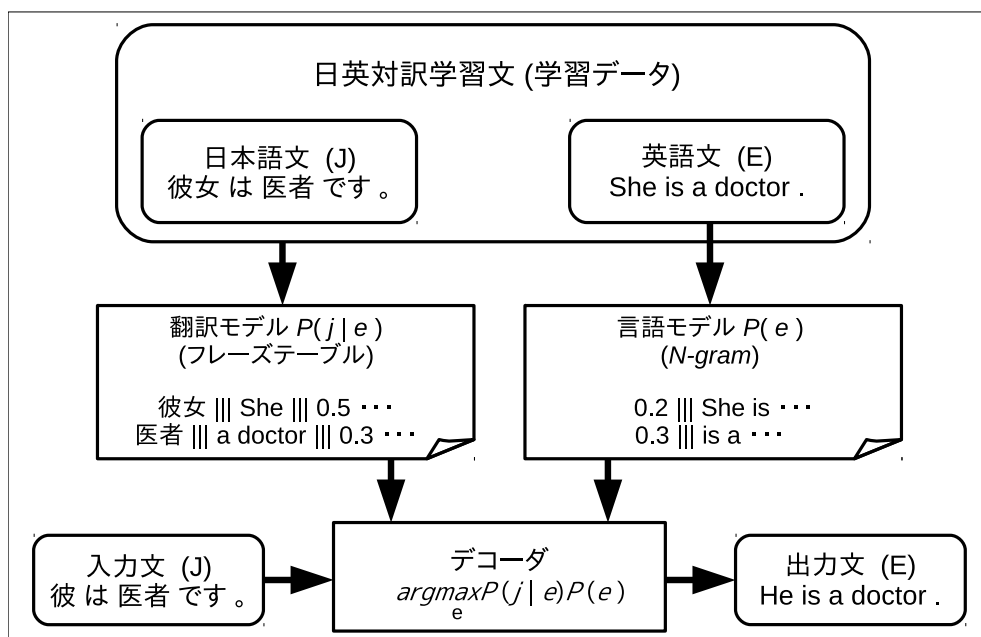


図 2.1: 日英統計翻訳の流れ

2.3 言語モデル

言語モデルは、単語列の生成確率を付与するモデルである。日英翻訳では、翻訳モデルを用いて生成された翻訳候補から、英語として自然な文を選出するために用いる。統計翻訳では一般的に、 N -gram モデルを用いる。 N -gram モデルとは“単語列 $w_1^n = w_1, w_2, w_3, \dots, w_n$ の i 番目の単語 w_i の生起確率 $P(w_i)$ は直前の $(N-1)$ の単語列 $w_{i-(N-1)}, w_{i-(N-2)}, w_{i-(N-3)}, \dots, w_{i-1}$ に依存する”という仮説に基づくモデルである。単語列 w_1^n の生起確率 $P(w_1^n)$ の計算式を以下に示す。

$$P(w_1^n) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \times \dots \times P(w_n|w_1^{n-1}) \quad (2.3)$$

$$\approx P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \times \dots \times P(w_n|w_{n-(N-1)}^{n-1}) \quad (2.4)$$

$$= \prod_{i=1}^n P(w_i|w_{i-(N-1)}^{i-1}) \quad (2.5)$$

また、 $P(w_i|w_{i-(N-1)}^{i-1})$ は以下の式で計算される。ここで $C(w_1^i)$ は単語列 w_1^i が出現する頻度を表す。

$$P(w_i|w_{i-(N-1)}^{i-1}) = \frac{C(w_{i-(N-1)}^i)}{C(w_{i-(N-1)}^{i-1})} \quad (2.6)$$

2.4 IBM 翻訳モデル

統計翻訳における単語対応を獲得するための代表的なモデルとして、IBM の Brown らによる仏英翻訳モデル [6] がある。IBM 翻訳モデルは仏英翻訳を前提としているが、本研究では日英翻訳を扱っているため、日英翻訳を前提に説明する。ここで、原言語の日本語文を J 、目的言語の英語文を E として定義する。また、IBM 翻訳モデルにおいて、日本語文 J と英語文 E の翻訳モデル $P(J|E)$ を計算するため、アライメント a を用いる。以下に IBM モデルの基本的な計算式を示す。

$$P(J|E) = \sum_a P(J, a|E) \quad (2.7)$$

ここで、アライメント a は、 J と E の単語の対応を意味している。IBM 翻訳モデルにおいて、各日単語に対応する英単語は 1 つであるのに対して、各英単語に対応する日単語は 0 から n 個あると仮定する。また、日単語と適切な英単語が対応しない場合、英語文の先頭に e_0 という空単語があると仮定し、日単語と対応させる。

また、IBM 翻訳モデルは、Model1 から Model5 までの 5 つのモデルから構成されている。各モデルの概要を以下に示す。

Model1 目的言語のある単語が原言語の単語に訳される確率を用いる

Model2 Model1 に加えて、目的言語のある単語に対応する原言語の単語の原言語文中での位置の確率（以下、permutation 確率と呼ぶ）を用いる（絶対位置）

Model3 Model2 に加えて、目的言語のある単語が原言語の何単語に対応するかの確率を用いる

Model4 Model3 の permutation 確率を改良（相対位置）

Model5 Model4 の permutation 確率を更に改良

2.5 句に基づく翻訳モデル

句に基づく翻訳モデルとは、確率的に日本語から英語の単語列へ翻訳を行うためのモデルである。統計翻訳において、句に基づく翻訳モデルとして、一般的にはフレーズテーブルが用いられている。フレーズテーブルは以下の手順で作成される。表 2.1 にフレーズテーブルの例を示す。

手順1 IBM 翻訳モデルを用いて，単語の対応を得る

手順2 ヒューリスティックなルールを用いて句に基づく対応を得る

手順3 手順2 で求めた句対応から，フレーズテーブルを作成する

表 2.1: フレーズテーブルの例

突然 天気 が		Suddenly , the weather		0.5	0.00217118	1	3.39949e-05	2.718	
0-0 0-1 2-2 1-3		2 1 1							
突然 天気 が 変わった		Suddenly , the weather changed		0.5	9.13961e-05	0.5	4.2075e-06	2.718	
0-0 0-1 2-2 1-3 3-4 4-4		2 2 1							

左から順に，日本語フレーズ，英語フレーズ，日英方向の翻訳確率 $P(j|e)$ ，日英方向の単語の翻訳確率の積，英日方向の翻訳確率 $P(e|j)$ ，英日方向の単語の翻訳確率の積，フレーズペナルティ，フレーズ内単語対応（日英方向）である．以後，フレーズペナルティは常に一定の値であるため省略する．

2.6 GIZA++

GIZA++[3] とは，日英方向と英日方向の対訳文から最尤な単語対応を得るための計算を行うツールである．IBM 翻訳モデルに基づいて，単語の対応関係の確率値を計算する．

2.7 フレーズテーブル作成法

フレーズテーブルの作成では，単語対応をとったあとに，ヒューリスティックなルールによって句の対応をとる．句レベルであるフレーズテーブルを得るには，まず，GIZA++を用いて，学習文から日英方向と英日方向の対訳文において最尤な単語アライメントを得る．例として，日本語文“風で火が消えた”と，その対訳英語文“The wind blew out the fire”を挙げる．図 2.2 に日英方向の単語対応を示す．また，図 2.3 に英日方向の単語対応を示す．なお，図 2.2 と図 2.3 において， \bullet は対応点を示す．

	The	wind	blew	out	the	fire
風		●				
で			●			
火						●
が			●			
消え			●			
た			●			

図 2.2: 日英方向の単語対応

	The	wind	blew	out	the	fire
風	●	●				
で					●	
火			●	●		●
が						
消え						
た	●					

図 2.3: 英日方向の単語対応

次に、両方向のアライメントから、両方向に1対多の対応を認めた単語アライメントをヒューリスティックなルールにより計算する。基本的なヒューリスティックとして、“intersection”、“union”、“grow”、そして“grow-diag”がある。

2.7.1 intersection

intersection（積集合）は、日英方向と英日方向の両方に単語対応が存在する場合、その単語対応を“対応点”とする方法である。intersectionの例を図2.4に示す。

	The	wind	blew	out	the	fire
風		●				
で						
火						●
が						
消え						
た						

図 2.4: intersection の例

2.7.2 union

union（和集合）は、日英方向と英日方向のどちらか一方に単語対応が存在する場合、その単語対応を“対応点”とする方法である。unionの例を図2.5に示す。

	The	wind	blew	out	the	fire
風	●	●				
で			●		●	
火			●	●		●
が			●			
消え			●			
た	●		●			

図 2.5: union の例

2.7.3 grow と grow-diag

grow, grow-diag は intersection と union の中間である．intersection からスタートし，既に採用した対応点の周りに union の対応点を加えていく．grow では縦と横の方向に，grow-diag では縦と横と対角の方向に union の対応点がある場合に，その対応点を用いる．図 2.6 に grow の例を，図 2.7 に grow-diag の例を示す．

	The	wind	blew	out	the	fire
風	●	●				
で						
火						●
が						
消え						
た						

図 2.6: grow の例

	The	wind	blew	out	the	fire
風	●	●				
で			●		●	
火						●
が						
消え						
た						

図 2.7: grow-diag の例

2.7.4 final と final-and

最終処理のヒューリスティクスとして，“final”と“final-and”を用いる．final は，少なくとも片方の言語における単語の単語対応がない場合に，union の単語対応を追加する．また，final-and は，両側言語における単語の単語対応がない場合に，union の候補対応点を追加する．図 2.8 に grow-diag-final の例を，図 2.9 に grow-diag-final-and の例を示された単語アライメントから，全ての矛盾しないフレーズ対を得る．このとき，そのフレーズ対に対して翻訳確率を計算し，フレーズ対に確率値を付与することで，フレーズテーブルを作成する．

	The	wind	blew	out	the	fire
風	●	●				
で			●		●	
火				●		●
が			●			
消え			●			
た	●		●			

図 2.8: grow-diag-final の例

	The	wind	blew	out	the	fire
風	●	●				
で			●		●	
火				●		●
が						
消え						
た						

図 2.9: grow-diag-final-and の例

2.8 デコーダ

デコーダは言語モデルと翻訳モデルの全ての組み合わせから、確率が最大となる翻訳候補を探索し、出力する。代表的なデコーダとして、Moses[7] が挙げられる。

2.9 パターンに基づく統計翻訳

パターン翻訳では句辞書と文パターン辞書を人手で作成する必要があったため、コストがかかるという問題があった。この問題に対し、江木らは、句辞書と文パターン辞書を統計的手法を用いて自動的に作成して翻訳を行う“パターンに基づく統計翻訳”を提案した。現在、文献 [5] より、パターンに基づく統計翻訳の翻訳精度は、句に基づく統計翻訳の翻訳精度よりも高いことが報告されている。

パターンに基づく日英統計翻訳は、大きく分けて 5 つの手順で構成される。パターンに基づく日英統計翻訳の流れを図 2.10 に示す。

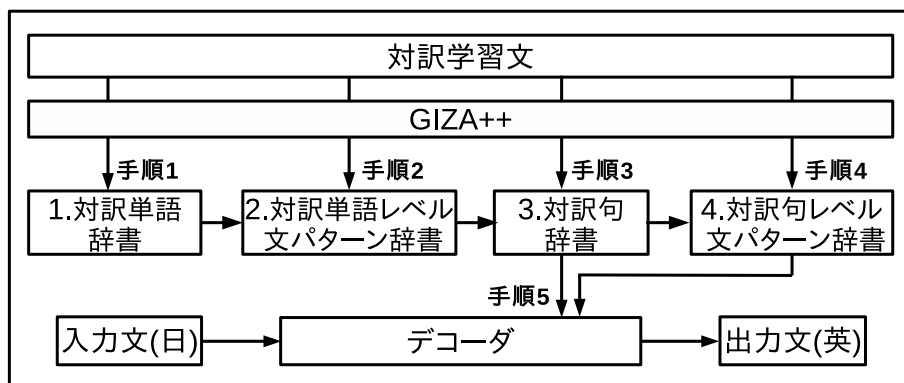


図 2.10: パターンに基づく日英統計翻訳の流れ

また、パターンに基づく日英統計翻訳における具体的な手順を以下に示す。

手順 1 対訳学習文と GIZA++ を用いて対訳単語辞書を作成

手順 2 対訳学習文と対訳単語辞書を用いて対訳単語レベル文パターン辞書を作成

手順 3 対訳学習文と対訳単語レベル文パターンを用いて対訳句辞書を作成

手順 4 対訳学習文と対訳句辞書を用いて対訳句レベル文パターン辞書を作成

手順 5 対訳句辞書と対訳句レベル文パターン辞書を用いて入力文を翻訳し、出力文を生成

2.10 評価手法

機械翻訳で生成した出力文を評価する方法として、自動評価法と人手評価法がある。以下にそれぞれの説明を示す。なお、本節は藤原らの論文 [11] を引用し、一部変更を加えて記述している。

2.10.1 自動評価法

翻訳精度を自動的に評価する手法として、用意された正解文と、機械翻訳器が出力した出力文とを比較する手法が一般的である。自動評価法には多くの手法が提案されているが、本論文では、BLEU (BiLingual Evaluation Understudy) [8]、METEOR (Metric for Evaluation of Translation with Explicit ORdering) [9]、RIBES (Rank based Intuitive Bilingual Evaluation Measure) [10] を用いる。

2.10.2 BLEU

BLEU [8] は、機械翻訳の分野において、最も一般的な自動評価基準である。BLEU は、n-gram マッチ率に基づく手法を用いている。以下に計算式を示す。

$$\text{BLEU} = \text{BP}_{\text{BLEU}} * \exp\left(\frac{1}{N} \sum_{n=1}^N \log p_n\right) \quad (2.8)$$

出力される値の領域は0~1であり，高いほど翻訳文が正解文に近いことを示す．式2.8の N には，一般的に $N = 4$ が用いられる． BP_{BLEU} は翻訳文が正解文と比較して短い場合に用いるペナルティ係数である．このペナルティは，翻訳文が正解文よりも短い場合に適合率が高くなる問題を回避するために用いられる． BP_{BLEU} は翻訳文の単語数を c ，正解文の単語数を r とし，以下の式で計算される．

$$BP_{BLEU} = \begin{cases} 1 & (c > r) \\ e^{1-\frac{r}{c}} & (c \leq r) \end{cases} \quad (2.9)$$

翻訳文の文字数が正解文よりも長い場合には， $BP_{BLEU} = 1$ となり，影響を及ぼさない．翻訳文の文字数が正解文よりも短い場合には，その長さに応じて， BP_{BLEU} が小さくなり，結果としてBLEU値が小さくなる．なお，単語数 c と r を算出するときの対象文は1文ではないため，各文の長さの自由度は残されている．また，式2.8の指数関数 p_n は以下の式で計算される．

$$p_n = \frac{\sum_{\text{全翻訳文}} \sum_{\text{正解文}} \text{正解文と一致する } N\text{-gram 数}}{\sum_{\text{全翻訳文}} \sum_{\text{正解文}} \text{全 } N\text{-gram 数}} \quad (2.10)$$

p_n は正解文と一致する N -gram数をカウントするとき，正解文の要素を重複して数えることを回避するための処理である．また，正解文に対して，翻訳文の方が長い場合には， p_n の値は小さくなるため，BLEU値も小さくなる．

2.10.3 METEOR

METEOR[9]は，再現率 R と適合率 P に基づく F 値に対して単語の非連続性に対するペナルティ関数 Pen を利用した評価基準である．以下に計算式を示す．

$$F = \frac{P \times R}{\alpha \times P + (1 - \alpha) \times R} \quad (2.11)$$

$$Pen = \gamma \times (c/m)^\beta \quad (2.12)$$

$$\text{METEOR} = F \times (1 - Pen) \quad (2.13)$$

式 2.12 のペナルティ関数 Pen にある m は機械翻訳の出力文と正解文との間で一致した単語数であり, c は一致した各単語を対象として語順が同じものを 1 つのまとまりとして統合した場合のまとまりの数である. したがって, 機械翻訳の出力文と正解文が完全一致の場合には $c = 1$ となり, 語順が全て逆の場合には $c = m$ となる. α, β, γ の値はパラメータである. METEOR は, 0 から 1 の値をスコアとして出力し, スコアが高いほど評価が良い.

2.10.4 RIBES

RIBES[10] とは, 順位相関係数を用いて共通単語の出現順序を評価する自動評価法である. 順位相関係数を用いることで, 文全体の語順に着目することが可能である. 自動評価と人手評価の相関について述べた松本らの研究 [12] において, 日英間翻訳の場合に比較的人手評価と相関が高い自動評価法であると報告されている. 以下に RIBES の計算式を示す.

$$\text{RIBES}(S) = \text{NSR} \times P^\alpha \quad (2.14)$$

$$\text{RIBES}(K) = \text{NKR} \times P^\alpha \quad (2.15)$$

式 2.14 における NSR はスピアマンの順位相関係数である. また, 式 2.15 における NKT はケンドールの順位相関係数である. そして, P はペナルティを表しており, 評価システムの翻訳が参照文に含まれる単語を含む割合である. α はペナルティに対する重みであり, $0 \leq \alpha \leq 1$ の値である. RIBES は, 0 から 1 の値をスコアとして出力し, スコアが高いほど評価が良い.

2.10.5 人手評価法

本研究では, 人手による評価として, 対比較評価を用いる. 対比較評価では, “入力文”, “正解文”, “ベースラインの出力文”, “提案手法の出力文” が与えられ, ベースラインの出力文と提案手法の出力文の比較を行う. 自動評価は正解文に近い文の評価が高い. 一方で, 人手評価は入力文に対する翻訳文として確からしい文の評価を行う. そのため, 自動評価と比較して評価の精度が高い [12].

第3章 先行研究

3.1 日野らの研究

日野らは、対訳学習文などの学習データを新たに追加して行う未知語処理手法を提案した。過去に、Maja Popovićらがセルビア語英語間とスペイン語英語間において、学習データに対訳句辞書を追加し統計翻訳を行った [13]。その結果、セルビア語英語方向の翻訳とスペイン語英語方向の翻訳の両方において翻訳精度の向上が確認された。そこで、日野らは、日本語英語間の翻訳において同様の手法を用いた場合の調査を行った。具体的には、句に基づく統計翻訳器 (Moses) において、新規の対訳句辞書を既存の対訳学習文に追加して翻訳を行った。そして未知語数の削減と翻訳精度向上が報告されている。

3.1.1 日野らの手法の手順

日野らの手法の流れを図 3.1 に示す。

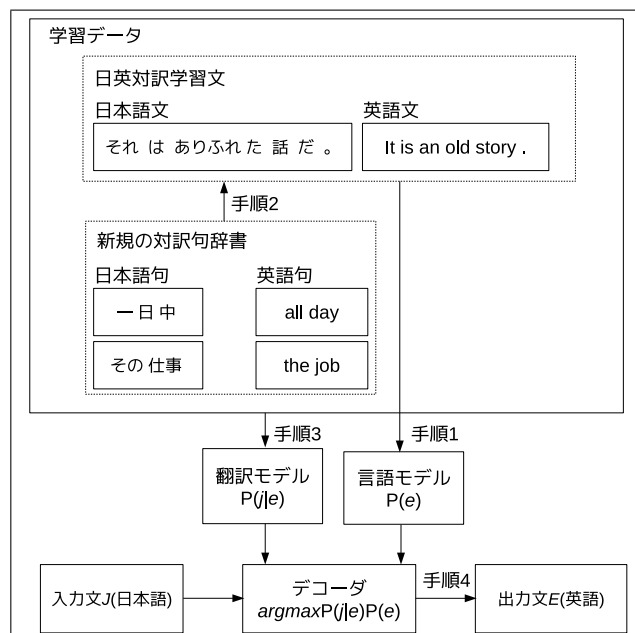


図 3.1: 日英統計翻訳における日野らの手法の流れ

また、以下に日野らの手法の具体的な手順を示す。

手順 1 日英対訳学習文を学習データとして言語モデルを作成する

手順 2 日英対訳学習文に新規の対訳句辞書を追加する

手順 3 手順 2 で作成したコーパスを学習データとして翻訳モデルを作成する

手順 4 手順 1 と手順 3 で作成したモデルを用いて統計翻訳を行う

3.1.2 日野らの手法の問題点

日野らの手法の問題点は、新規の学習データを用意するためのコストがかかる点である。

3.2 藤原らの研究

藤原らは、句に基づく統計翻訳器 (Moses) において、対訳学習文などの学習データを追加せずに行う未知語処理手法を提案した。Moses ではフレーズテーブルを参照して翻訳を行う。そして、フレーズテーブルを作成する際には “grow-diag-final-and” というヒューリスティックスを用いるのが一般的である。この “grow-diag-final-and” を用いて作成したフレーズテーブルの問題として、1 単語対 1 単語の対訳データが作成されにくいことが挙げられる。そのため、1 単語の翻訳を行う場合に未知語として出力されることがある。一方で、“intersection” というヒューリスティックスを用いてフレーズテーブルを作成した場合は、1 単語対 1 単語の対訳データのみ作成される。そこで、藤原らは、“grow-diag-final-and” を用いて作成したフレーズテーブルと、“intersection” を用いて作成したフレーズテーブルの 2 種類を組み合わせたフレーズテーブルを用いて翻訳を行った。そして、未知語数の削減と翻訳精度向上が報告されている。

3.2.1 藤原らの手法の手順

藤原らの手法の流れを図 3.2 に示す。

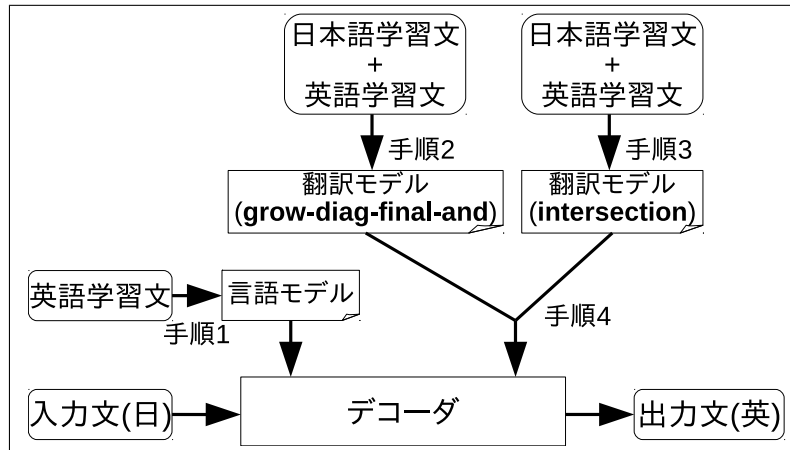


図 3.2: 日英統計翻訳における藤原らの手法の流れ

また，以下に藤原らの手法の具体的な手順を示す．

準備 日英対訳学習文を準備する．

手順 1 英語学習文を用いて言語モデルを作成する．

手順 2 英語学習文と日本語学習文を用いて翻訳モデルを作成する．また，ヒューリスティックスは“grow-diag-final-and”を用いる．

手順 3 手順 2 と同様にして翻訳モデルを作成する．また，ヒューリスティックスは“intersection”を用いる．

手順 4 手順 3 で作成されたフレーズテーブルから未知語が含まれるフレーズ対を抽出し手順 2 で作成されたフレーズテーブルに直接追加する．このフレーズテーブルを用いて翻訳を行う．

3.2.2 藤原らの手法の問題点

藤原らの手法の問題点は，Moses 内で未知語処理を行なっているため，他の統計翻訳器への応用が期待できない点である．

第4章 提案手法

4.1 提案手法の概要

本研究では，他の統計翻訳器への応用を考慮した，新規の学習データを必要としない新たな未知語処理手法を提案する．日野らの手法では，人手で作成した対訳句辞書を追加するため，新規の学習データを準備する必要があった．一方で，提案手法は，GIZA++により自動作成した対訳単語辞書を未知語処理に用いる．GIZA++は，既存の対訳学習文さえ準備できれば自動的に対訳単語辞書を作成することが可能である．したがって，提案手法は新規の学習データを必要としない．また，藤原らの手法では，新規の学習データを必要としなかったが，句に基づく統計翻訳器（Moses）内で未知語処理を行っているため，他の統計翻訳器への応用が期待できなかった．一方で，提案手法は，一度生成した出力文に対して未知語処理を行うため，統計翻訳器の種類に依存しない．したがって，提案手法は，あらゆる統計翻訳器に実装することができる．

以上のことから，提案手法の特徴をまとめると以下ようになる．

- 自動作成した対訳単語辞書を用いた未知語処理のため，新規の学習データが不必要
- 一度生成した文に対する後処理のため，あらゆる統計翻訳器に実装可能

4.2 提案手法の手順

提案手法は4つの手順で構成される．日英統計翻訳における提案手法の流れを図4.1に示す．また，以下に提案手法の具体的な手順を示す．

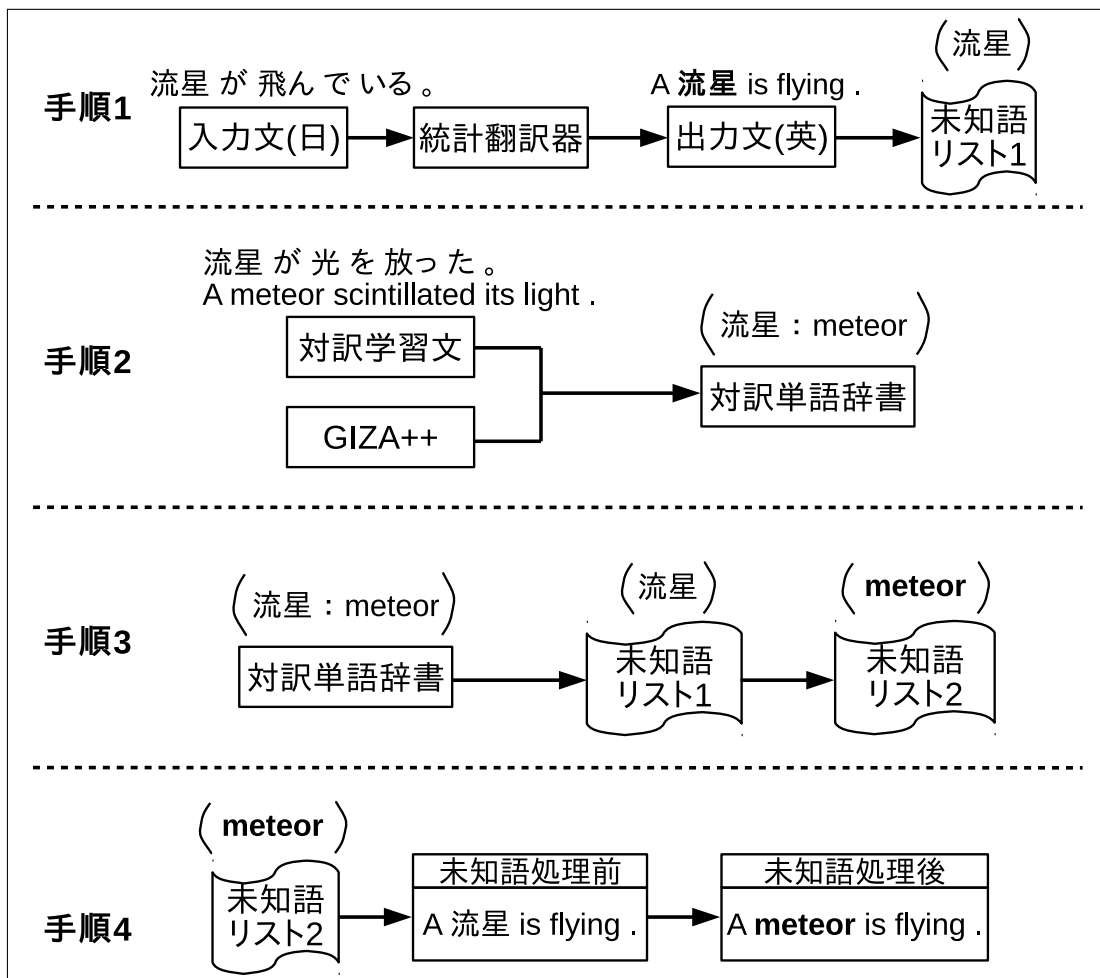


図 4.1: 日英統計翻訳における提案手法の流れ

- 手順 1 任意の機械翻訳器における出力文から未知語を抽出し未知語リスト 1 を作成する
- 手順 2 GIZA++ を用いて対訳学習文から対訳単語を獲得し、対訳単語辞書を作成する
- 手順 3 手順 2 で作成した対訳単語辞書を用いて未知語リスト 1 を翻訳し未知語リスト 2 を作成 (対訳単語辞書内に訳語が複数存在する場合は単語翻訳確率が最尤の単語を選択)
- 手順 4 未知語処理前の文に対し未知語リスト 2 を適用

第5章 実験環境

5.1 統計翻訳器

本実験では，図 4.1 中の統計翻訳器として，パターンに基づく統計翻訳器（PBSMT）を用いる．また，PBSMT における対訳単語辞書の枝刈り条件を以下に示す．

- 単語翻訳確率が $\log_2(0.20)$ より高い単語
- 日本語単語と英語単語が両方同時に含まれる文の数が 1 つの単語
- 単語翻訳確率の順位が日本語・英語ともに 8 位以上の単語

5.2 実験データ

本研究では，電子辞書などの例文より抽出した単文コーパス [14] を用いる．使用するデータの内訳を表 5.1 に示す．

表 5.1: 実験データ

対訳学習文	160,000 文
入力文	10,000 文

また，統計翻訳の前処理として，各コーパスの日本語文に対して，“MeCab[15]”を用いて形態素解析を行う．そして，英語文に対して“tokenizer.perl[16]”を用いて分かち書きを行う．表 5.2 に単文コーパスの例を示す．

表 5.2: 単文コーパスの例

日本語文	水が腐っている。
英語文	The water is foul .
日本語文	素行を改めなさい。
英語文	You should mend your ways .

5.3 評価方法

本研究では，未知語処理前と処理後における未知語数と文全体の翻訳精度の比較評価を行う．文全体の翻訳精度の評価として人手評価と自動評価を行う．人手評価には対比較評価を用いる．そして，自動評価にはBLEU，METEOR，RIBESを用いる．また，未知語処理前の出力文をベースラインとし，未知語処理後の出力文を提案手法とする．

第6章 実験結果

未知語数と未知語の翻訳精度および文全体の翻訳精度の結果を以下に示す。

6.1 未知語数

ベースラインと提案手法において、未知語数の調査を行った。調査結果を表 6.1 に示す。

表 6.1: 未知語数 (10,000 文中)

翻訳手法	未知語を含む文数	未知語数
ベースライン	1,644 文	1,847 単語
提案手法	677 文	713 単語

表 6.1 より、提案手法において 967 文と 1,134 単語の未知語処理ができた。

6.2 未知語の翻訳精度

表 6.1 において未知語処理ができた 967 文からランダムに抽出した 100 文において、未知語は 109 単語存在した。この未知語の翻訳精度を調査した。結果を表 6.2 に結果を示す。

表 6.2: 未知語の翻訳精度 (109 単語中)

正しい翻訳	誤った翻訳
54 単語	55 単語

また、表 6.2 において、正しく翻訳できた未知語の一例を表 6.3 に示す。正しく翻訳できなかった未知語の一例を表 6.4 に示す。

表 6.3: 正しく翻訳できた未知語の一例 (54 単語中)

翻訳前	翻訳後
表玄関	entrance
貴い	valuable
匿名	anonymous
心房	atrium
光化学	photochemical

表 6.4: 正しく翻訳できなかった未知語の一例 (55 単語中)

翻訳前	翻訳後
現存	few
不純	custody
多事	Political
省ける	trouble
ひとつひとつ	invests

6.3 文全体の翻訳精度

未知語処理ができた文における文全体の翻訳精度を調査するため、人手評価と自動評価を行った。以下に評価結果を示す。

6.3.1 人手評価結果

未知語処理ができた 967 文からランダムに抽出した 100 文を用いて人手評価を行った。評価の基準を以下に示す。また、評価結果を表 6.5 に示す。

- 提案手法 : 入力文の意味が読み取れる
- 提案手法 × (差なし) : 入力文の意味が読み取れない

表 6.5: 提案手法の評価結果 (100 文中)

提案手法	提案手法 × (差なし)
19 文	81 文

表 6.5 より、提案手法による文全体の翻訳精度向上が確認できた。

6.3.2 人手評価例

表 6.5 における“提案手法”の評価例と“提案手法 × (差なし)”の評価例を以下に示す。

6.3.2.1 提案手法 の例

“提案手法”の評価例を表 6.6～表 6.8 に示す。

表 6.6 では、提案手法により“ベニス”が“Venice”に翻訳された。“Venice”は、“ベニス”に対する正しい翻訳となっている。そのため、文全体を通して入力文の意味が読み取れるようになったため、“提案手法”とした。

表 6.6: 提案手法 の評価例 1

入力文	彼は 晩年 を ベニス で 過ごした。
参照文	He spent the afternoon of his life in Venice .
ベースライン	He spent his late years in <u>ベニス</u> .
提案手法	He spent his late years in <u>Venice</u> .

表 6.7 では、表 6.6 の評価例と同様に判断し、“提案手法”とした。

表 6.7: 提案手法 の評価例 2

入力文	母親 は 子供 の ズボン の 破れ目 を 継いだ。
参照文	She sewed up the tear in her son's pants .
ベースライン	The mother inherited child's <u>破れ目</u> of trousers .
提案手法	The mother inherited child's <u>tear</u> of trousers .

表 6.8 では、表 6.6 の評価例と同様に判断し、“提案手法”とした。

表 6.8: 提案手法 の評価例 3

入力文	その 政治家 は 暴力団 と ぐる になっていた。
参照文	The politician worked hand in glove with the gangsters .
ベースライン	The politician was <u>ぐる</u> with gangs .
提案手法	The politician was <u>league</u> with gangs .

6.3.2.2 提案手法 × (差なし) の例

“提案手法 × (差なし)”の評価例を表 6.9～表 6.11 に示す。

表 6.9 では、提案手法により“方策”が“measures”に翻訳された。“measures”は、“方策”に対する正しい翻訳となっている。しかし、文全体を通して入力文の意味が読み取れないままであるため、“提案手法×(差なし)”とした。

表 6.9: 提案手法×(差なし)の評価例 1

入力文	適当な方策が思い浮かばない。
参照文	I have yet to come up with suitable measures .
ベースライン	not occurred thought out A occasion <u>方策</u> .
提案手法×(差なし)	not occurred thought out A occasion <u>measures</u> .

表 6.10 では、表 6.9 の評価例と同様に判断し、“提案手法×(差なし)”とした。

表 6.10: 提案手法×(差なし)の評価例 2

入力文	失敗も貴い経験である。
参照文	Failures can result being precious experiences .
ベースライン	He is <u>貴い</u> experience for his failure .
提案手法×(差なし)	He is <u>valuable</u> experience for his failure .

表 6.11 では、提案手法により“申し込む”が“3715-4671”に翻訳された。“3715-4671”は、“申し込む”に対する誤った翻訳となっている。そして、入力文の意味が読み取れないままであるため、“提案手法×(差なし)”とした。

表 6.11: 提案手法×(差なし)の評価例 3

入力文	あの会社に嚴重な抗議を申し込むべきだ。
参照文	We should enter a strong protest against that company .
ベースライン	He should <u>申し込む</u> protest rigidly that company .
提案手法×(差なし)	He should <u>3715-4671</u> protest rigidly that company .

6.3.3 自動評価結果

未知語処理ができた 967 文を用いて自動評価を行った。表 6.12 に結果を示す。

表 6.12: 提案手法の自動評価結果. 精度が高い方を太字で示す

翻訳手法	BLEU	METEOR	RIBES
ベースライン	0.068	0.314	0.661
提案手法	0.072	0.342	0.680

表 6.12 より、提案手法による翻訳精度向上が全ての自動評価において確認できた。

第7章 考察

7.1 未知語の傾向

第6.1節で獲得した未知語(1,847単語)の傾向を調べるため品詞の分布を調査した。品詞の解析には“MeCab”を用いた。結果を図7.1に示す。

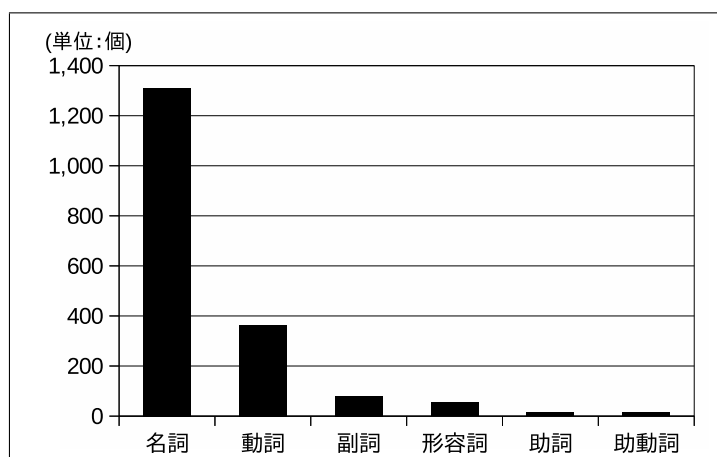


図 7.1: 未知語の品詞分布

図7.1より、上位3位を名詞・動詞・副詞が占めていることが分かった。具体的には、1位の名詞が1,311単語、2位の動詞が364単語、3位の副詞が80単語存在した。また、最も多かった名詞の中でも特に一般名詞が多く存在し、名詞全体の約7割を占めていた。残り3割の内、サ変接続名詞が約2割を占めており、固有名詞が約1割を占めていた。また、上位3位における動詞と副詞の未知語例を表7.1に示す。

表 7.1: 動詞と副詞の未知語例

品詞	例1	例2	例3
動詞	償う	偲ん	吹き替える
副詞	徐徐に	なにやら	あらためて

7.2 未知語の翻訳精度と文全体の翻訳精度の差

表 6.2 より、未知語の翻訳精度は約 50%であり、表 6.5 より、文全体の翻訳精度は約 20%であることが分かった。これらの結果より、未知語の翻訳精度と文全体の翻訳精度に差が生じていることが分かる。この主な原因として、ベースライン（未知語処理前の文）の翻訳精度が考えられる。提案手法は、一度生成された出力文に対する後処理の形で未知語処理を行っている。そのため、未知語処理前の文の翻訳精度に大きく左右される。したがって、ベースラインの翻訳精度が低い場合には、未知語が正しく翻訳できたとしても文全体の翻訳精度に影響しなかったと考えられる。この具体例として、表 6.9 と表 6.10 の評価例が挙げられる。表 6.9 と表 6.10 は、どちらも未知語の翻訳は正しくできているが、未知語部分以外の翻訳精度が低く意味が読み取れない状態であるため、最終的に文全体を通して入力文の意味が読み取れない結果となった。さらに、表 6.2 の未知語の翻訳精度が 100%だと仮定して人手評価を行った場合でも、表 6.5 の文全体の翻訳精度は約 35%までしか向上しない。ここで、未知語の翻訳精度が 100%だと仮定した場合に翻訳精度が向上する出力文と向上しない出力文の例を表 7.2 と表 7.3 に示す。表 7.2 では、“冷夏”という未知語が誤った翻訳となっているが、仮に“cool summer”に正しく翻訳された場合には文全体の翻訳精度が向上することが分かる。また、表 7.3 では、“省ける”という未知語が誤った翻訳となっているが、仮に“save”に正しく翻訳されたとしても文全体の翻訳精度が向上しないことが分かる。

表 7.2: 仮に未知語の翻訳精度が 100%の場合に翻訳精度が向上する例

入力文	ことしは冷夏になりそうだ。
参照文	It seems as if we are going to have a cool summer this year .
ベースライン	We'll have <u>冷夏</u> this year .
提案手法 × (差なし)	We'll have <u>summer</u> this year .

表 7.3: 仮に未知語の翻訳精度が 100%の場合でも翻訳精度が向上しない例

入力文	洗濯機によってかなりの労力が省ける。
参照文	A washing machine saves considerable labor .
ベースライン	A considerable effort <u>省ける</u> planes by washing .
提案手法 × (差なし)	A considerable effort <u>trouble</u> planes by washing .

以上のことから、ベースラインの翻訳精度が向上した場合には提案手法の効果がより現れると考えられる。したがって、今後、提案手法において文全体の翻訳精度を向上させるためには、まずベースラインの翻訳精度を向上させなければならないと考える。

7.3 誤った未知語翻訳の解析

表 6.2 より，109 単語中 55 単語は誤った未知語翻訳となった．この原因として，対訳学習文中に出現する頻度が 1 回の単語（頻度 1 単語）の問題と，複合語の未知語の問題が考えられる．

7.3.1 頻度 1 単語の問題

未知語の翻訳精度低下につながる主な原因として頻度 1 単語の問題が挙げられる．頻度 1 単語は，翻訳確率の統計的信頼性が低いため精度が低い傾向にある．表 6.2 において頻度 1 単語は 109 単語中 75 単語存在していた．この 75 単語における翻訳精度を表 7.4 に示す．また，表 7.4 における正しい翻訳例を表 7.5 に，誤った翻訳例を表 7.6 に示す．

表 7.4: 頻度 1 単語の翻訳精度 (75 単語中)

正しい翻訳	誤った翻訳
33 単語	42 単語

表 7.5: 正しい翻訳の一例 (33 単語中)

翻訳前	翻訳後
鞘	sheath
還流	reflux
盆栽	bonsai
友誼	friendship
大塚	Otsuka

表 7.6: 誤った翻訳の一例 (42 単語中)

翻訳前	翻訳後
旧来	clinging
印加	voltage
空爆	Milosevic
膨らむ	heaved
リウマチ	Since

また，対訳学習文中に出現する頻度が 2 回以上の単語（頻度 2 以上の単語）における翻訳精度を表 7.7 に示す．また，表 7.7 における正しい翻訳例を表 7.8 に，誤った翻訳例を表 7.9 に示す．

表 7.7: 頻度 2 以上の単語の翻訳精度 (34 単語中)

正しい翻訳	誤った翻訳
21 単語	13 単語

表 7.8: 正しい翻訳の一例 (21 単語中)

翻訳前	翻訳後
交戦	war
マス	trout
運輸	transport
スモッグ	smog
マドリード	Madrid

表 7.9: 誤った翻訳の一例 (13 単語中)

翻訳前	翻訳後
冷夏	summer
音感	pitch
適合	compatibility
まじっ	hairs
一両日	or

表 7.4 と表 7.7 より，頻度 1 単語の翻訳精度は約 40%であり，頻度 2 以上の単語の翻訳精度は約 60%だった．したがって，頻度 1 単語の翻訳精度は頻度 2 以上の単語の翻訳精度よりも低いことが分かる．ここで，誤った未知語翻訳となった 109 単語中 55 単語の内訳を表 7.10 に示す．

表 7.10: 誤った未知語翻訳の内訳 (55 単語中)

頻度 1 単語	頻度 2 以上の単語
42 単語	13 単語

表 7.10 より，誤った未知語翻訳における約 80%は頻度 1 単語であることが分かる．つまり，精度の低い頻度 1 単語の未知語が多く存在していることが，表 6.2 の翻訳精度低下につながっていると考えられる．

7.3.2 複合語の未知語の問題

未知語の翻訳精度低下につながる別の原因として，複合語の未知語の問題が挙げられる．提案手法では，対訳単語辞書を用いて未知語処理を行っている．対訳単語辞書は 1 対 1 対応になっているため，1 単語の日本語に対して 1 単語の英語でしか翻訳することができない．したがって，日本語 1 単語に対して複数の英単語を必要とするような複合語が未知語の場合には，意味が欠落し翻訳に失敗する．実際に，誤った未知語翻訳となった 55 単語の内，複合語の未知語は 10 単語存在した．表 7.11 に複合語の未知語の翻訳失敗例を示す．表 7.11 より，翻訳後の単語は全て意味が欠落しており，翻訳に失敗していることが分かる．そして，翻訳前の単語に対する正しい表現は全て句単位であることが分かる．以上のことから，日本語 1 単語に対して複数の英単語を必要とするような複合語の未知語は，翻訳精度低下につながると考える．

表 7.11: 複合語の未知語の翻訳失敗例

翻訳前	翻訳後	正しい表現
冷夏	summer	cold summer
白雲	white	white cloud
今秋	autumn	this autumn
金婚式	golden	golden wedding
中南米	Latin	Latin America

7.4 IBM 翻訳モデルの比較

IBM 翻訳モデルには Model 1 ~ Model 5 までの 5 つのモデルが存在する。本研究では、最も基本的な Model 1 を用いて実験を行った。しかし、Model 2 ~ Model 5 を用いた実験は行なっておらず、対訳単語辞書の精度が変化する可能性が残っている。したがって、今後は他のモデルを用いた場合の対訳単語辞書の精度調査および、その辞書を用いた場合の未知語処理結果の比較が必要であると考える。

7.5 翻訳確率最尤の単語選択の効果

本研究では、未知語処理における訳語選択時に、翻訳確率が最尤の単語を選択している。また、中村らの研究 [17] により、自動作成した対訳単語辞書からランダムに 100 単語抽出して様々な精度調査を行った結果が報告されている。その中でも、全対訳単語の精度と、頻度 1 単語の精度の調査結果を表 7.12 と表 7.13 に示す。

表 7.12: 全対訳単語の精度 (ランダム)

正しい翻訳	誤った翻訳
17 単語	83 単語

表 7.13: 頻度 1 単語の精度 (ランダム)

正しい翻訳	誤った翻訳
14 単語	86 単語

表 7.12 と表 7.13 より、ランダムに単語を選択した場合は、翻訳精度が非常に低いことが分かる。一方で、翻訳確率最尤の単語を選択した場合は、表 6.2 より、全対訳単語の精度が約 30% 向上していることが分かった。さらに、表 7.4 より、頻度 1 単語の精度も約 30% 向上していることが分かった。

7.6 対訳句辞書を用いた未知語処理

本研究では、対訳単語辞書を用いて未知語処理を行った。しかし、自動的に作成した対訳句辞書を用いて未知語処理を行うことも考えられる。具体的には、1変数の単語レベル文パターン辞書を元に対訳句辞書と対訳句レベル文パターン辞書をループ処理で作成し、最後に作成した対訳句辞書を未知語処理に用いる方法である。ループ処理の流れを図7.2に示す。ここで、図中の“単語パターン”は対訳単語レベル文パターン辞書を指し、“句パターン”は対訳句レベル文パターン辞書を指す。

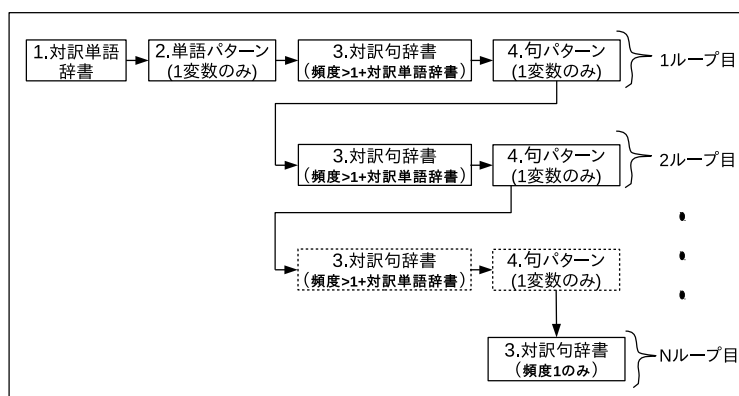


図 7.2: ループ処理の流れ

また、1変数の単語レベル文パターン辞書は 343,566 パターン存在しており、このパターンを使用して作成された句の数は 4,308 句だった。この句の精度調査結果を表 7.14 に示す。表中の“ ”は、意味が欠落または不必要な意味が付与されていることを意味する。

表 7.14: 1変数の単語レベル文パターンから作成した対訳句の精度

		×
78	22	0

表 7.14 より、1変数の対訳単語レベル文パターン辞書から作成した対訳句は精度が高いことが分かる。この対訳句辞書から対訳句レベル文パターン辞書をループ処理で作成した結果を表に示す。また、ループ処理で作成した対訳句の精度調査結果を表 7.16 に示す。

表 7.15: 対訳句と1変数の句レベル文パターンの数

ループ数	対訳句の数	1変数の句レベル文パターンの数
1	3,112	467,663
2	9,274	545,524
4	95,737	624,900
8	257,763	801,523

表 7.16: ループ処理で作成した対訳句の精度 (100 句中)

ループ数			×
1	76	24	0
2	62	38	0
4	61	39	0
8	60	40	0

また、表 7.16 におけるループ数が 1 の場合の評価例を表 7.17 に、ループ数が 2 の場合の評価例を表 7.18 に、ループ数が 8 の場合の評価例を表 7.19 に示す。

表 7.17: ループ数 1 の評価例

日本語句 英語句	せみ cycada	夕食の後にトランプ cards after dinner
日本語原文 英語原文	せみ が かえる。 The cycada emerges .	夕食のあとでトランプをしよう。 Let's play cards after dinner .
日本語文パターン 英語文パターン	N00 が かえる。 The N00 emerges .	N00 を しよう。 Let's play N 00 .

表 7.18: ループ数 2 の評価例

日本語句 英語句	歴史的 課題 historical question	きみは 席 seat
日本語原文 英語原文	これは 歴史的 課題 である。 This is a historical question .	きみは 席 に 戻り なさい。 Go back to your seat .
日本語文パターン 英語文パターン	これは N00 である。 This is a N00 .	N00 に 戻り なさい。 Go back to your N00 .

表 7.19: ループ数 8 の評価例

日本語句 英語句	風 が やん だ wind has stopped	は 服毒 した took poison
日本語原文 英語原文	風 が やん だ。 The wind has stopped .	彼女 は 服毒 した。 She took poison .
日本語文パターン 英語文パターン	N00 。 The N 00 .	彼女 N00 。 She N 00 .

7.7 デコーダ内とデコーダ外の未知語処理の差

本研究では、デコーダ外で未知語処理を行ったが、デコーダ内で未知語処理を行うことも可能である。本研究において、デコーダ外で未知語処理を行った大きな理由は、他の統計翻訳器への応用を可能にするためである。また、川原らの研究 [18] より、デコーダ内で未知語処理を行う方法よりも、デコーダ外で未知語処理を行う方法の方が翻訳精度が高いということも報告されている。表 7.20 と表 7.21 にそれぞれの未知語の翻訳精度を示す。また、表 7.22 と表 7.23 にそれぞれの文全体の翻訳精度を示す。

表 7.20: デコーダ外で処理した場合の未知語の翻訳精度 (213 単語中)

翻訳成功	正しい翻訳	誤った翻訳
186 単語	96/186 単語	90/186

表 7.21: デコーダ内で処理した場合の未知語の翻訳精度 (約 210 単語中)

翻訳成功	正しい翻訳	誤った翻訳
60 単語	20/60 単語	40/60

表 7.22: デコーダ外の未知語処理における人手評価結果 (100 文中)

デコーダ外	デコーダ外 ×	差なし	同一出力
4 文	2 文	79 文	15 文

表 7.23: デコーダ内の未知語処理における人手評価結果 (100 文中)

デコーダ内	デコーダ内 ×	差なし	同一出力
4 文	3 文	60 文	33 文

以上の結果より、デコーダ内で未知語処理を行った場合の未知語の翻訳精度は約 30% であるのに対して、デコーダ外で未知語処理を行った場合の未知語の翻訳精度は約 50% だった。したがって、今後の未知語処理の研究では、本研究のようにデコーダ外で未知語処理を行う方法を採用した方が良いと考える。

7.8 追加実験1 (句に基づく統計翻訳器での実験)

提案手法は、パターンに基づく統計翻訳器 (PBSMT) に限らず、あらゆる統計翻訳器に応用可能である。そこで、追加実験として、図 4.1 における統計翻訳器に、句に基づく統計翻訳器 (Moses) を用いた場合の実験を行った。以下、Moses における未知語処理前の出力文を“ベースライン (Moses)”と呼ぶ。また、Moses における未知語処理後の出力文を“提案手法 (Moses)”と呼ぶ。実験結果を以下に示す。

7.8.1 未知語数

未知語数の調査を行った。調査結果を表 7.24 に示す。

表 7.24: Moses の実験における未知語数 (10,000 文中)

翻訳手法	未知語を含む文数	未知語数
ベースライン (Moses)	941 文	1,032 単語
提案手法 (Moses)	783 文	843 単語

表 7.24 より、提案手法 (Moses) において 158 文と 189 単語の未知語処理ができた。

7.8.2 未知語の翻訳精度

表 7.24 において未知語処理ができた 158 文からランダムに抽出した 100 文において、未知語は 101 単語存在した。この未知語の翻訳精度を調査した。結果を表 7.25 に結果を示す。

表 7.25: Moses の実験における未知語の翻訳精度 (101 単語中)

正しい翻訳	誤った翻訳
48 単語	53 単語

7.8.3 文全体の翻訳精度

未知語処理ができた文における文全体の翻訳精度を調査するため、人手評価と自動評価を行った。以下に評価結果を示す。

7.8.3.1 人手評価結果

未知語処理ができた 158 文からランダムに抽出した 100 文を用いて人手評価を行った。評価の基準を以下に示す。また、評価結果を表 7.26 に示す。

- Moses : 入力文の意味が読み取れる
- Moses × (差なし) : 入力文の意味が読み取れない

表 7.26: 提案手法 (Moses) の人手評価結果 (100 文中)

Moses	Moses × (差なし)
3 文	96 文

表 7.26 より、提案手法による文全体の翻訳精度向上がわずかながら確認できた。

7.8.3.2 人手評価例

表 7.26 における “Moses ” および “Moses × (差なし)” の場合における評価例を以下に示す。

A) “Moses ” の評価例

“Moses ” の評価例を表 7.27 ~ 表 7.29 に示す。

表 7.27 では、提案手法 (Moses) により “煮え返っ” が “boiling” に翻訳された。“boiling” は、“煮え返っ” に対する正しい翻訳となっている。そのため、文全体を通して入力文の意味が読み取れるようになったため、“Moses ” とした。

表 7.27: Moses の評価例 1

入力文	やかんの湯が煮え返っていた。
参照文	The water in the kettle was boiling briskly .
ベースライン	The water in the kettle was <u>煮え返っ</u> .
Moses	The water in the kettle was <u>boiling</u> .

表 7.28 では、表 7.27 の評価例と同様に判断し、“Moses ”とした。

表 7.28: Moses の評価例 2

入力文	嵐がニレの大枝を揺さぶった。
参照文	The storm waved the heavy branches of the elm .
ベースライン	The storm rocked the branch of <u>ニレ</u> .
Moses	The storm rocked the branch of <u>elm</u> .

表 7.29 では、表 7.27 の評価例と同様に判断し、“Moses ”とした。

表 7.29: Moses の評価例 3

入力文	作業員が道路に鉄管を埋めた。
参照文	The workers laid iron pipes in the road .
ベースライン	The workers filled the <u>鉄管</u> on the road .
Moses	The workers filled the <u>pipes</u> on the road .

B) “Moses × (差なし)” の評価例

“Moses × (差なし)” の評価例を表 7.30 ~ 表 7.32 に示す。

表 7.30 では、提案手法 (Moses) により “スカンク” が “skunk” に翻訳された。“skunk” は、“スカンク” に対する正しい翻訳となっている。しかし、文全体を通して入力文の意味が読み取れないままであるため、“Moses × (差なし)” とした。

表 7.30: Moses × (差なし) の評価例 1

入力文	スカンクが悪臭を放った。
参照文	The skunk gave off a terrible odor .
ベースライン	The <u>スカンク</u> .
Moses × (差なし)	The <u>skunk</u> .

表 7.31 では、表 7.30 の評価例と同様に判断し、“Moses × (差なし)”とした。

表 7.31: Moses × (差なし) の評価例 2

入力文	彼ら は 大 攻勢 を 来月 まで 延ばす だろ う 。
参照文	They will hold off their major offensive until next month .
ベースライン	They will be a major assault on the <u>延ばす</u> until next month .
Moses × (差なし)	They will be a major assault on the <u>extend</u> until next month .

表 7.32 では、提案手法 (Moses) により“生き返る”が“live”に翻訳された。“live”は、“生き返る”に対する誤った翻訳となっている。そして、文全体を通して入力文の意味が読み取れないままであるため、“Moses × (差なし)”とした。

表 7.32: Moses × (差なし) の評価例 3

入力文	春 に は 万物 が 生き返る 。
参照文	In spring everything returns to life .
ベースライン	The assimilates <u>生き返る</u> in spring .
Moses × (差なし)	The assimilates <u>live</u> in spring .

7.8.3.3 自動評価結果

未知語処理ができた 158 文を用いて自動評価を行った。表 7.33 に結果を示す。

表 7.33: 提案手法 (Moses) の自動評価結果. 精度が高い方を太字で示す

翻訳手法	BLEU	METEOR	RIBES
ベースライン	0.092	0.351	0.695
提案手法 (Moses)	0.097	0.366	0.700

表 7.33 より、提案手法による翻訳精度向上が全ての自動評価において確認できた。

7.8.4 Moses と PBSMT における文全体の翻訳精度の差

表 7.25 より，Moses においても未知語の翻訳精度は約 50% だった．一方で，表 7.26 より，Moses における文全体の翻訳精度は約 5% しか向上しなかった．また，表 6.5 より，PBSMT における文全体の翻訳精度が約 20% 向上したことから，Moses と PBSMT における文全体の翻訳精度には約 15% の差が生じていることが分かる．未知語の翻訳精度が同等にも関わらず，文全体の翻訳精度において差が生じた理由として，ベースラインの文構造の差が考えられる．Moses は，フレーズテーブルから句を参照し，複数の句を結合することで文を生成するため，文構造が明白ではない傾向にある．一方で，PBSMT は，入力文に対してパターン照合を行い，適合したパターン内の変数部分を対訳句辞書を用いて翻訳することで文を生成する．したがって，文構造が Moses よりも明白である傾向にある．つまり，PBSMT の方が Moses よりも，正しい未知語翻訳ができた場合に文全体の翻訳精度が向上しやすくなる．したがって，Moses と PBSMT における文全体の翻訳精度に差が生じたと考えられる．

7.9 追加実験 2 (提案手法を拡張した実験)

第 6.1 節より, 1,644 文中 677 文の未知語を含む文と 1,847 単語中 713 単語の未知語が残る結果となった。これらの未知語数をさらに削減する方法として, 単語を文字単位に分割して再度翻訳する処理が考えられる。そこで, 追加実験として, 提案手法に単語を文字単位に分割して再度翻訳する処理を加えた場合の実験を行った。以下, 提案手法に単語を文字単位に分割して再度翻訳する処理を加えた手法を“拡張手法”と呼ぶ。また, 本実験は第 5 章と同じ実験環境で行った。

7.9.1 拡張手法の手順

拡張手法は大きく分けて 4 つの手順で構成される。提案手法との違いは, 手順 3 において, 翻訳できなかった未知語を文字単位に分割して再度翻訳する処理が加わっている点である。日英統計翻訳における拡張手法の流れを図 7.3 に示す。また, 以下に拡張手法の具体的な手順を示す。

手順 1 任意の機械翻訳器における出力文から未知語を抽出し未知語リスト 1 を作成する

手順 2 GIZA++ を用いて対訳学習文から対訳単語を獲得し, 対訳単語辞書を作成する

手順 3.1 手順 2 で作成した対訳単語辞書を用いて未知語リスト 1 を翻訳 (単語翻訳) し未知語リスト 2 を作成する (訳語が複数存在する場合は翻訳確率が最尤のものを選択)

手順 3.2 未知語リスト 2 において漢字だけで構成される 2 文字以上の単語 (漢字未知語) を文字単位に分割し未知語リスト 3 を作成する

手順 3.3 手順 2 で作成した対訳単語辞書を用いて未知語リスト 3 を翻訳 (文字翻訳) し未知語リスト 4 を作成する (訳語が複数存在する場合は翻訳確率が最尤のものを選択)

手順 4 未知語処理前の文に対し未知語リスト 4 を適用

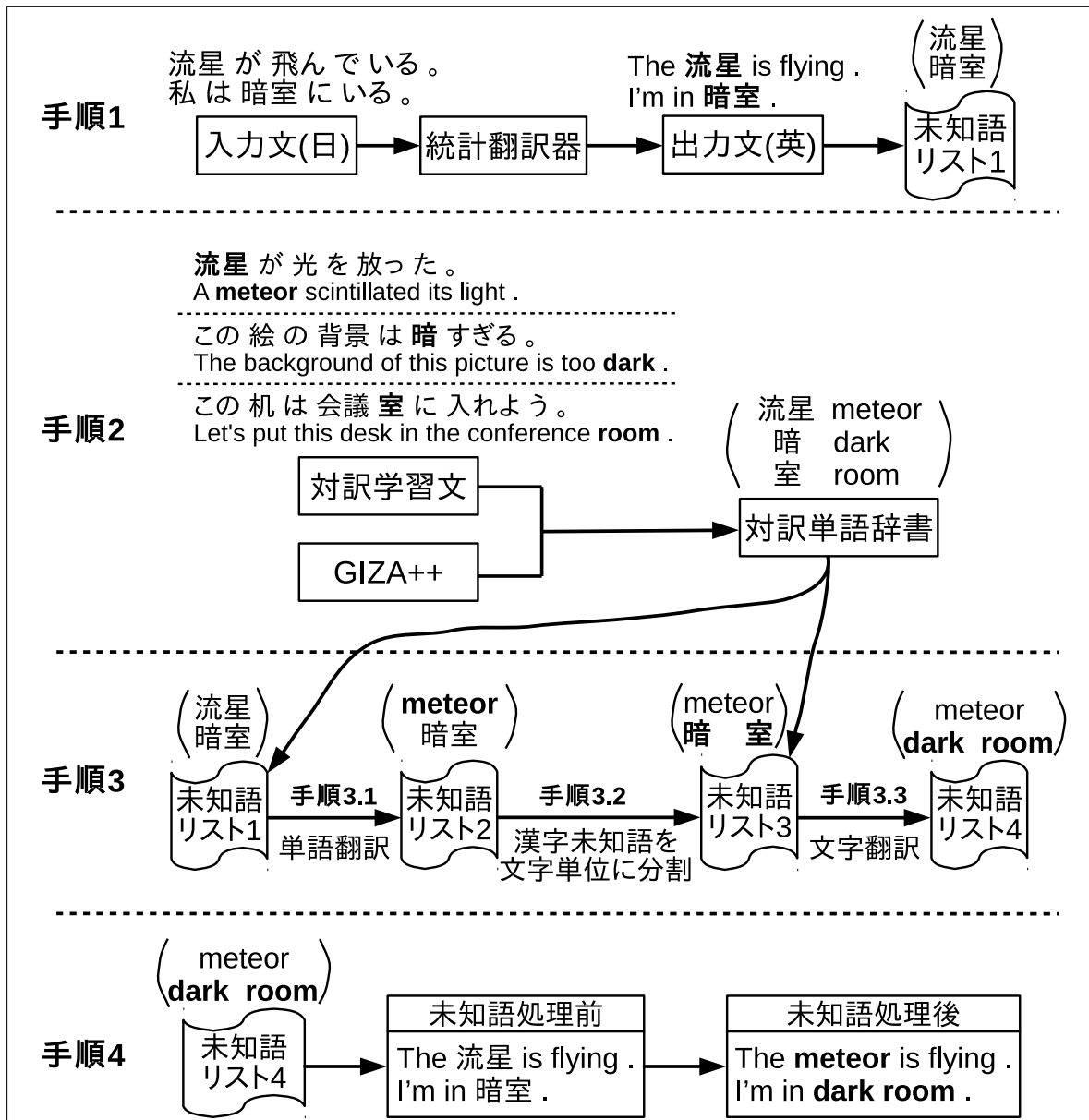


図 7.3: 日英統計翻訳における拡張手法の流れ

7.9.2 実験結果

拡張手法の実験における未知語数と未知語の翻訳精度および文全体の翻訳精度を以下に示す。

7.9.2.1 未知語数

ベースラインと提案手法，拡張手法において未知語数の調査を行った。調査結果を表 7.34 に示す。

表 7.34: 拡張手法の実験における未知語数 (10,000 文中)

翻訳手法	未知語を含む文数	未知語数
ベースライン	1,644 文	1,847 単語
提案手法	677 文	713 単語
拡張手法	501 文	521 単語

表 7.34 より，拡張手法において，新たに 176 文と 192 単語の未知語処理ができた。

7.9.2.2 漢字未知語の翻訳精度

表 7.34 において，新たに未知語処理ができた 176 文からランダムに抽出した 100 文において，漢字未知語（漢字だけで構成される 2 文字以上の単語）は 104 単語存在した。この漢字未知語の翻訳精度を調査した。結果を表 7.35 に結果を示す。

表 7.35: 拡張手法の実験における漢字未知語の翻訳精度 (104 単語中)

正しい翻訳	文字単位で正しい翻訳	誤った翻訳
20 単語	16 単語	68 単語

また，表 7.35 における正しい翻訳結果の一例を表 7.36 に，文字単位で正しい翻訳結果の一例を表 7.37 に，誤った翻訳結果の一例を表 7.38 に示す。

表 7.36: 正しい翻訳結果の一例 (20 単語)

翻訳前	翻訳後
初 球	first ball
大 功	great success
泥 道	mud road
文 頭	sentence head
古 刀	old sword

表 7.37: 文字単位で正しい翻訳結果の一例 (16 単語)

翻訳前	翻訳後
正 夢	positive dream
緑 青	green blue
甘 党	sweet party
音 階	sound floor
日 傘	day umbrella

表 7.38: 誤った翻訳結果の一例 (64 単語)

翻訳前	翻訳後
威 圧	triumphantly compression
魔 性	devil importance
英 気	English about
仲 人 口	other people mouth
鋳 工 業	miners plumber business

7.9.2.3 文全体の翻訳精度

拡張手法において、新たに未知語処理ができた 176 文における文全体の翻訳精度を調査するため、人手評価と自動評価を行った。以下に評価結果を示す。

7.9.2.4 人手評価結果

新たに未知語処理ができた 176 文からランダムに抽出した 100 文を用いて人手評価を行った。評価の基準を以下に示す。また、評価結果を表 7.39 に示す。

- 拡張手法 : 入力文の意味が読み取れる
- 拡張手法 × (差なし) : 入力文の意味が読み取れない

表 7.39: 拡張手法の人手評価結果 (100 文中)

拡張手法	拡張手法 × (差なし)
7 文	93 文

表 7.39 より、拡張手法による文全体の翻訳精度向上がわずかながら確認できた。

7.9.2.5 人手評価例

表 7.39 における“拡張手法”の評価例と“拡張手法×(差なし)”の評価例を以下に示す。

A) 拡張手法 の例

“拡張手法”の評価例を表 7.40～表 7.42 に示す。

表 7.40 では、拡張手法により“大功”が“great success”に翻訳された。“great success”は、“大功”という 1 単語に対する正しい翻訳となっている。そのため、文全体を通して入力文の意味が読み取れるようになったため、“拡張手法”とした。

表 7.40: 拡張手法 の評価例 1

入力文	国家に大功があった。
参照文	There was a great act in the state.
ベースライン	There was a <u>大功</u> to nation .
拡張手法	There was a <u>great success</u> to nation .

表 7.41 では、表 7.40 の評価例と同様に判断し、“拡張手法”とした。

表 7.41: 拡張手法 の評価例 2

入力文	損料 5、000 円で車を借りた。
参照文	I paid 5,000 yen for the hire of the car .
ベースライン	I borrowed a car in <u>損料</u> 5,000 yen .
拡張手法	I borrowed a car in <u>lost fees</u> 5,000 yen .

表 7.42 では、表 7.40 の評価例と同様に判断し、“拡張手法”とした。

表 7.42: 拡張手法 の評価例 3

入力文	泥道がひどくぬかった。
参照文	The dirt road has poached badly .
ベースライン	The <u>泥道</u> was muddy .
拡張手法	The <u>mud road</u> was muddy .

B) 拡張手法 × (差なし) の例

“拡張手法 × (差なし)”の評価例を表 7.43 ~ 表 7.45 に示す。

表 7.43 では、拡張手法により“大敵”が“great enemy”に翻訳された。“great enemy”は、“大敵”という 1 単語に対する正しい翻訳となっている。しかし、文全体を通して入力文の意味が読み取れないままであるため、“拡張手法 × (差なし)”とした。

表 7.43: 拡張手法 × (差なし) の評価例 1

入力文	恐ろしい大敵に勇敢に立ち向かった。
参照文	They faced fearful odds courageously .
ベースライン	I bravely 大敵 terrible .
拡張手法 × (差なし)	I bravely <u>great enemy</u> terrible .

表 7.44 では、拡張手法により“楽曲”が“easy music”に翻訳された。“easy music”は、文字単位では全て正しい翻訳であるが、“楽曲”という 1 単語に対しては誤った翻訳となっている。そして、入力文の意味が読み取れないままであるため、“拡張手法 × (差なし)”とした。

表 7.44: 拡張手法 × (差なし) の評価例 2

入力文	その楽曲の創作年代は正確にはきめられない。
参照文	We cannot fix the exact date of the composition .
ベースライン	There is the cause precisely pairing generation 楽曲 creative writing .
拡張手法 × (差なし)	There is the cause precisely pairing generation <u>easy music</u> creative writing .

表 7.45 では、拡張手法により“斜視”が“sulfur importance”に翻訳された。“sulfur importance”は文字単位で全て誤った翻訳であるため“斜視”という 1 単語に対しても誤った翻訳となっている。そして、入力文の意味が読み取れないままであるため、“拡張手法 × (差なし)”とした。

表 7.45: 拡張手法 × (差なし) の評価例 3

入力文	彼女の片方の目は少し斜視であった。
参照文	One of her eyes had a slight squint .
ベースライン	She was a little 斜視 his eye .
拡張手法 × (差なし)	She was a little <u>sulfur importance</u> his eye .

7.9.2.6 自動評価結果

拡張手法において，新たに未知語処理ができた 176 文を用いて自動評価を行った．表 7.46 に結果を示す．

表 7.46: 拡張手法の自動評価結果. 精度が高い方を太字で示す

翻訳手法	BLEU	METEOR	RIBES
ベースライン	0.067	0.284	0.654
文字翻訳	0.069	0.299	0.667

表 7.46 より，拡張手法による翻訳精度向上が全ての自動評価において確認できた．

7.9.3 拡張手法における文字翻訳の問題

拡張手法では，提案手法で翻訳できなかった未知語に対する更なる未知語処理として文字翻訳を行った．この文字翻訳における問題として，複合名詞と固有名詞の問題が考えられる．

7.9.3.1 複合名詞の問題

文字翻訳において，3 文字以上の漢字未知語における複合名詞の問題が存在する．例えば，本実験で得られた漢字未知語において“基本給”が存在するが，文字翻訳の結果は“based book salary”となり失敗している．このような場合，本来ならば“基本”と“給”に分割した状態で翻訳する必要があるため，文字翻訳では正しく翻訳することが困難であることが分かる．したがって，単語翻訳の段階で複合名詞を分割する処理が今後必要になると考えられる．実際に，対訳単語辞書において，“基本”と“basic”の対訳が最尤の翻訳確率で存在し，“給”と“salary”の対訳が最尤の翻訳確率で存在しているため，仮に実装できた場合には精度向上が期待できると考える．

7.9.3.2 固有名詞の問題

文字翻訳において，固有名詞の問題が存在する．具体例を表 7.47 に示す．ここでは，固有名詞“苗場”が未知語として出力されている．文字翻訳の結果が“seedlings spot”となっており，文字単位では正しい翻訳であるが単語単位では誤った翻訳である．つまり，未知語が固有名詞である場合，表 7.47 のような単語単位で誤った翻訳にしかならない

め，精度向上につながらない．ただし，漢字未知語における固有名詞のみローマ字に変換する処理を追加することで解消できると考える．

表 7.47: 固有名詞の文字翻訳例

入力文	私は苗場へスキーに行った。
参照文	I went skiing at Naeba .
ベースライン	I went to 苗場 to skiing .
文字翻訳	I went to <u>seedlings spot</u> to skiing .

第8章 おわりに

近年主流となっている統計翻訳において、未知語が出力されることで翻訳精度の低下につながるという問題があった。この未知語を減少させるため、日野らの研究では、新規の対訳句辞書を既存の対訳学習文に追加して未知語処理を行った。そして、未知語数の削減と翻訳精度向上が報告されている。しかし、日野らの手法では新規の学習データが必要になるためコストがかかるという問題が残った。また、藤原らの研究では、Mosesにおいて、フレーズテーブル作成時のヒューリスティックスを併用して未知語処理を行った。そして、未知語数の削減と翻訳精度向上が報告されている。しかし、藤原らの手法は、Moses内で未知語処理を行なっているため、他の統計翻訳器への応用が期待できないという問題が残った。

そこで、本研究では、他の統計翻訳器への応用を考慮した、新規の学習データを必要としない新たな未知語処理手法を提案した。具体的には、GIZA++を用いて自動作成した対訳単語辞書による未知語処理を、一度生成した出力文に対する後処理の形で行った。

実験の結果、PBSMTにおいて1,644文中967文と1,847単語中1,134単語の未知語処理ができた。さらに、人手評価と自動評価の結果より、文全体の翻訳精度が向上し、提案手法の有効性が確認できた。一方で、提案手法の大きな問題として、未知語処理前の文の翻訳精度が低い場合には、未知語が正しく翻訳できたとしても文全体の翻訳精度は向上しないことが分かった。また、未知語の翻訳精度低下につながる原因として、頻度1単語が未知語の多くを占めているという問題と、日本語1単語に対して複数の英単語を必要とするような複合語の未知語を正しく翻訳できないという問題が存在することが分かった。今後は、未知語処理前の文の翻訳精度を向上させる手法を検討したい。

謝辞

最後に、三年間に渡り、本研究のご指導を賜りました鳥取大学工学部知能情報工学科自然言語処理研究室の村上仁一准教授、村田真樹教授に深く感謝すると共に、厚く御礼申し上げます。また、ご多忙中にも関わらず、有益なるご助言を賜りました鳥取大学岩井儀雄教授と鳥取大学木村周平教授に厚く御礼申し上げます。そして、日常の議論を通じて多くの知識や示唆を頂いた同研究室の皆様には深謝いたします。また、参考にさせていただいた論文の著者の方々に対して、深く感謝申し上げます。

参考文献

- [1] 日野聡子, 村上仁一, 徳久雅人: “日英統計翻訳における対訳句コーパスの効果”, 言語処理学会第 18 回年次大会発表論文集, pp.491-494, 2012.
- [2] 藤原勇, 村上仁一, 徳久雅人: “句に基づく統計翻訳における未知語処理の 1 手法”, 言語処理学会第 19 回年次大会発表論文集, pp.580-583, 2013.
- [3] Franz Josef Och, Hermann Ney: “A Systematic Comparison of Various Statistical Alignment Models”, Computational Linguistics, Volume 29, Number 1, pp.19-51, 2003.
- [4] 江木孝史, 村上仁一, 徳久雅人: “句に基づく対訳句パターンの自動作成と統計的手法を用いた英日パターン翻訳”, 言語処理学会第 20 回年次大会, A6-2, pp.951-954, 2014.
- [5] 村上仁一: “パターンに基づく統計機械翻訳の概要と問題点について”, 電子情報通信学会技術研究報告, 言語理解とコミュニケーション, NLC2017-3, pp.13-18, 2017.
- [6] Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, Robert L. Mercer: “The mathematics of statistical machine translation: Parameter Estimation”, Computational Linguistics, Volume 19, Number 2, pp.263-311, 1993.
- [7] Moses: Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation”, Proceedings of the ACL 2007 Demo and Poster Sessions, pp.177-180, 2007.
- [8] BLEU: Papineni Kishore, Salim Roukos, Todd Ward, Wei-Jing Zhu, “BLEU: a method for automatic evaluation of machine translation”, 40th Annual meeting of the Association for Computational Linguistics, pp.311-318, 2002.

- [9] METEOR: Banerjee Satanjeev , Lavie Alon , “METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments” , Proceedings of Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at the 43th Annual Meeting of the Association of Computational Linguistics (ACL-2005) , pp.65-72 , 2005.
- [10] RIBES: 平尾努, 磯崎秀樹, Kevin Duh, 須藤克仁, 塚田元, 永田昌明, “RIBES: 順位相関に基づく翻訳の自動評価法” , 言語処理学会第 17 年次大会発表論文集 , pp.1111-1114 , 2011.
- [11] 藤原勇: “パターン翻訳を用いた学習データ増加手法の検討” , 平成 24 年度鳥取大学修士論文, 2013.
- [12] 松本拓也, 村上仁一, 徳久雅人: “機械翻訳における人手評価と自動評価の考察” , 言語処理学会第 18 回年次大会発表論文集 , pp.505-508 , 2012.
- [13] Popović Maja, and Ney Hermann “Statistical Machine Translation with a small amount of bilingual training data” , 5th LREC SALT MIL Workshop on Minority Languages. 2006.
- [14] 村上仁一, 藤波進: “日本語と英語の対訳文対の収集と著作権の考察” , 第一回コーパス日本語学ワークショップ , pp.119-130 , 2012 .
- [15] Mecab : mecab-0.97.tar.gz , mecab-ipadic-2.7.0-20070801.tar.gz
<http://mecab.sourceforge.net/>.
- [16] Moses: Open Source Toolkit for Statistical Machine Translation , Proceedings of the ACL 2007 Demo and Poster Sessions , pp.177-180 , Prague , 2007.
- [17] 中村友哉: “対訳単語辞書の精度調査” , 平成 28 年度鳥取大学卒業論文, 2017.
- [18] 川原宰, 村上仁一: “パターンに基づく日英統計翻訳とその未知語処理” , 言語処理学会第 23 回年次大会発表論文集 , pp.624-627 , 2017.

研究業績

1. 川原宰, 村上仁一：“文字ベース翻訳による未知語処理”，言語処理学会第 22 回年次大会発表論文集，pp.207-210，2016.
2. 川原宰, 村上仁一：“パターンに基づく日英統計翻訳とその未知語処理”，言語処理学会第 23 回年次大会発表論文集，pp.624-627，2017.
3. 川原宰, 村上仁一：“日英翻訳において IBM Model 1 を用いた未知語処理”，言語処理学会第 24 回年次大会発表論文集 (3 月掲載予定)，2018.