

概要

機械翻訳には様々な手法がある．機械翻訳手法の例として統計翻訳が挙げられる．統計翻訳では対訳文から言語モデル（目的言語文の統計モデル）と翻訳モデル（原言語と目的言語の関係を表すモデル）を作成する．これらのモデルを利用して原言語文を目的言語文に翻訳する手法である．しかし，統計翻訳では限られた対訳文からモデルを作成する．そのためモデルは不完全であり，不完全なモデルを利用した翻訳は精度が十分ではない．

別の手法として古瀬ら“経験的知識を活用する変換主導型機械翻訳 [6]”を提案した．この手法では，対訳文を解析して得られた経験的知識から変換知識を作成する．この変換知識を利用して原言語文を繰り返し変換することで，目的言語文に翻訳する手法である．対訳文を解析して得られた経験知識を利用することで原言語と目的言語の一般的な現象を捉えることができる．そのため，高い翻訳精度を実現した．しかし，変換知識の作成は人手で行われるため，開発コストは高い．また，古瀬らの研究 [6] では語彙数の少ない対訳文を利用していた．本研究で利用する電子辞書などの例文より抽出した単文 [14] では，入力文に対して翻訳可能な文の割合（以下，カバー率）は低下すると考えられる．

以上のような問題を解決するために，“相対的意味論を利用した変換主導型統計機械翻訳: Transfer Driven Statistical Machine Translation [1]”¹（以下，TDSMT）を提案する．TDSMT では対訳文から変換テーブルを自動作成する．変換テーブルは「 A が B ならば C は D 」の形式である．この変換テーブルと対訳文を利用して翻訳を行う．対訳文を変換し，翻訳を行うため，文法を順守した翻訳が期待できる．さらに，変換テーブルを自動作成することにより“経験的知識を活用する変換主導型機械翻訳 [6]”と比較して低コストで高いカバー率を得られる．

第 I 部では“相対的意味論を利用した変換主導型統計機械翻訳 [1]”の詳細な手法（変換テーブルの自動作成手法と翻訳の手法）を提案する．TDSMT カバー率を調査した．調査のカバー率は 20% 程度であった．このカバー率は統計翻訳と比較してまだまだ低い．カバー率を向上させる手法を考案する必要がある．

そのため，第 II 部では出力文中に未知語を出力する手法を提案する．この手法は，入力文と対訳文を利用して未知語出力用変換テーブルを作成する．未知語出力用変換テーブルは「 A が B ならば C は C 」の形式である．この未知語出力用変換テーブルを変換変換テーブルとあわせて利用し翻訳することで，出力文に未知語を出力する手法である．

¹この名称の“相対的意味論を利用した”は古瀬らの“経験的知識を活用する変換主導型機械翻訳 [6]”と差を明確にするための名称である．

未知語を出力した場合の TDSMT においてカバー率を調査した。調査の結果未知語を出力することでカバー率は 60% 程度まで向上した。さらに、未知語を正しく翻訳できた場合、翻訳精度において、TDSMT は統計翻訳を上回る可能性を示した。

目次

第I部 相対的意味論に基づく 変換主導型統計機械翻訳の提案	1
第1章 はじめに	2
第2章 先行研究	3
2.1 統計翻訳	3
2.1.1 単語に基づく統計翻訳	3
2.1.2 言語モデル	4
2.1.3 翻訳モデル	5
2.1.3.1 IBM 翻訳モデル	5
2.1.3.2 GIZA++	6
2.1.4 デコーダ	6
2.1.5 句に基づく統計翻訳	6
2.1.6 統計翻訳の問題点	7
2.2 古瀬らの研究 [6]	8
2.2.1 変換知識	8
2.2.2 意味距離	9
2.2.3 実験結果	10
2.2.4 問題点	10
2.3 機械翻訳における評価手法	11
2.3.1 自動評価法	11
2.3.2 人手評価法	13
第3章 提案手法：相対的意味論を利用した変換主導型統計機械翻訳 (TDSMT)	14
3.1 概要	14
3.2 相対的意味論 [5]	14

3.2.1	概要	14
3.2.2	機械翻訳における応用	14
3.2.3	機械翻訳における利点	15
3.3	翻訳	17
3.3.1	概要	17
3.3.2	翻訳の手順	17
3.3.2.1	変換テーブルの選択	18
3.3.2.2	変換テーブルの適用	18
3.3.2.3	出力文の決定	19
3.3.3	TDSMT の出力例	19
3.4	学習:変換テーブルの自動作成	21
3.4.1	変換テーブルの自動作成手法の概要	21
3.4.2	変換テーブルの自動作成手法の手順	21
3.4.2.1	対訳単語の作成	22
3.4.2.2	文パターンの作成	23
3.4.2.3	変換テーブルの作成	23
3.4.3	変換テーブルの作成例	24
第4章	実験環境	26
4.1	実験データ	26
4.2	評価方法	27
第5章	実験結果	28
5.1	作成した変換テーブルの数の調査	28
5.2	作成した変換テーブルの精度の調査	28
5.2.1	評価例:変換テーブル(A が B)	29
5.2.2	評価例:変換テーブル(C が D)	30
5.3	カバー率の調査	30
5.4	実験結果のまとめ	31
第6章	考察	32
6.1	TDSMT と Moses[9] との翻訳精度の比較	32
6.1.1	自動評価	32

6.1.2	人手評価	32
6.1.2.1	人手評価例：TDSMT の例	33
6.1.2.2	人手評価例：Moses の例	34
6.1.2.3	人手評価例：差なしの例	35
6.1.2.4	人手評価例：同一出力の例	37
6.2	TDSMT の利点	38
6.3	誤り解析	39
6.3.1	変換テーブルの誤り	39
6.3.2	利用する対訳文の誤り	40
6.3.3	確率値による出力文の選択の誤り	41
6.3.4	手法の改善	43
第7章 おわりに		44
第II部 相対的意味論を利用した 変換主導型統計機械翻訳における未知語の出力		45
第1章 はじめに		46
第2章 従来手法:相対的意味論を利用した変換主導型統計機械翻訳 (TDSMT)		47
2.1	TDSMT の概要	47
2.2	TDSMT の問題	47
第3章 提案手法:未知語を出力可能にした TDSMT		50
3.1	概要	50
3.2	翻訳	51
3.2.1	翻訳の概要	51
3.2.2	翻訳の手順	51
3.2.3	従来手法との変更点	51
3.2.4	提案手法の出力例	52
3.3	学習	52
3.3.1	変換テーブルの自動作成手法	52
3.3.2	未知語出力用変換テーブルの自動作成手法	52

3.3.2.1	未知語出力用変換テーブルの作成	53
3.3.2.2	変換テーブルへの追加	54
3.3.3	未知語出力用変換テーブルの作成例	54
第4章	実験環境	55
4.1	実験データ	55
4.2	評価方法	55
第5章	実験結果	56
5.1	翻訳に利用する変換テーブルの数	56
5.2	カバー率の調査	56
5.3	実験結果のまとめ	57
第6章	考察	58
6.1	提案手法と Moses[9] との翻訳精度の比較	58
6.1.1	自動評価	58
6.1.1.1	自動評価法	58
6.1.1.2	自動評価結果	58
6.1.2	人手評価	59
6.1.2.1	人手評価法	59
6.1.2.2	人手評価結果	59
6.1.2.3	人手評価例：提案手法 の例	59
6.1.2.4	人手評価例：Moses の例	60
6.1.2.5	人手評価例：差なしの例	62
6.1.2.6	同一出力の例	63
6.1.3	提案手法の翻訳精度に関するまとめ	63
6.2	誤り解析	63
6.2.1	変換テーブルの誤り	64
6.2.2	利用する対訳文の誤り	65
6.2.3	確率値による出力文の選択の誤り	66
6.3	未知語の翻訳精度に及ぼす影響	68
6.4	翻訳に利用する確率値について	68

目 次

第 I 部 相対的意味論に基づく

変換主導型統計機械翻訳の提案

2.1 英日統計翻訳の手順	4
2.2 経験的知識を活用する変換主導型機械翻訳のフローチャート	8
3.1 TDSMT による翻訳手法のフローチャート	17
3.2 変換テーブルの自動作成手法のフローチャート	21

第 II 部 相対的意味論を利用した

変換主導型統計機械翻訳における未知語の出力

2.1 カバー率が低下する要因	48
3.1 未知語を出力可能にした TDSTM のフローチャート	50
3.2 未知語出力用変換テーブルの自動作成手法のフローチャート	53

表 目 次

第I部 相対的意味論に基づく 変換主導型統計機械翻訳の提案

2.1	フレーズテーブルの例	7
2.2	変換知識の数と具体例	9
3.1	相対的意味論を利用して自動作成した変換テーブルの例	15
3.2	変換テーブルの選択例	18
3.3	変換テーブルの適用例	19
3.4	変換テーブルの適用例	19
3.5	TDSMT の翻訳例	20
3.6	対訳文から作られる対訳単語	22
3.7	対訳単語と対訳文から作られる文パターン	23
3.8	文パターンと対訳文から作られる変換テーブル	24
3.9	変換テーブルの実際の作成例	25
4.1	実験データ	26
4.2	実験データの例	26
5.1	作成した変換テーブルの数	28
5.2	変換テーブルの評価基準	28
5.3	作成した変換テーブルの精度	29
5.4	変換テーブル (A が B) の評価 の例	29
5.5	変換テーブル (A が B) の評価 の例	29
5.6	変換テーブル (A が B) の評価 \times の例	29
5.7	変換テーブル (C が D) の評価 の例	30
5.8	変換テーブル (C が D) の評価 の例	30
5.9	変換テーブル (C が D) の評価 \times の例	30

5.10	各手法のカバー率	30
6.1	TDSMT と mooses の自動評価結果	32
6.2	対比較評価の評価基準	33
6.3	TDSMT と Moses との対比較評価結果	33
6.4	TDSMT の人手評価例 1	33
6.5	TDSMT の人手評価例 2	34
6.6	TDSMT の人手評価例 3	34
6.7	Moses の人手評価例 1	34
6.8	Moses の人手評価例 2	35
6.9	Moses の人手評価例 3	35
6.10	差なしの人手評価例 1	36
6.11	差なしの人手評価例 2	36
6.12	差なしの人手評価例 3	36
6.13	同一出力の人手評価例	37
6.14	誤りのおおまかな内訳	39
6.15	TDSMT の誤り例 (変換テーブルの誤り)	39
6.16	表 6.15 の変換テーブル 2	40
6.17	TDSMT の誤り例 (対訳文の選択の誤り)	41
6.18	TDSMT の誤り例 (確率値が第一位の翻訳候補文)	41
6.19	TDSMT の誤り例 (確率値が第 2 位の翻訳候補文)	42

第 II 部 相対的意味論を利用した

変換主導型統計機械翻訳における未知語の出力

3.1	提案手法の実際の翻訳例	52
3.2	未知語出力用変換テーブルの作成	54
3.3	未知語出力用変換テーブルの例	54
5.1	翻訳に利用する変換テーブルの平均数	56
5.2	各手法のカバー率	56
6.1	提案手法と Moses の自動評価結果	58
6.2	提案手法と Moses との対比較評価結果	59
6.3	提案手法 の人手評価例 1	59

6.4	提案手法 の人手評価例 2	60
6.5	提案手法 の人手評価例 3	60
6.6	Moses の人手評価例 1	60
6.7	Moses の人手評価例 2	61
6.8	Moses の人手評価例 3	61
6.9	差なしの人手評価例 1	62
6.10	差なしの人手評価例 2	62
6.11	差なしの人手評価例 3	62
6.12	同一出力の人手評価例	63
6.13	誤りのおおまかな内訳	64
6.14	TDSMT の誤り例 (変換テーブルの誤り)	64
6.15	表 6.14 の変換テーブル 1 の詳細	65
6.16	TDSMT の誤り例 (対訳文の選択の誤り)	66
6.17	TDSMT の誤り例 (確率値が第一位の翻訳候補文)	67
6.18	TDSMT の誤り例 (確率値が第三位の翻訳候補文)	67

第I部

相対的意味論に基づく 変換主導型統計機械翻訳の提案

第1章 はじめに

機械翻訳には様々な手法がある。現在，研究活動に広く利用されている手法として統計翻訳の Moses[9] が挙げられる。Moses は対訳文から言語モデルと翻訳モデルを統計的に学習する。この2つのモデルを利用して，原言語文を目的言語に翻訳する。言語モデルは人間には理解に難しい統計数値の集合である。また，この統計数値は膨大な対訳文から作成されている。このため，翻訳に誤りが生じた場合，細かく原因を探索するのは困難である。

また，古瀬らは，“経験的知識を活用する変換主導型機械翻訳 [6]”を提案した。この手法は対訳文から得た経験的知識をもとに人手で作成した変換知識を利用して翻訳を行う。入力文に変換知識に基づいた変換を繰り返して翻訳を行う。このため，翻訳の手順が詳しく解析できる。翻訳に誤りが生じた場合，手順をたどることで誤りの原因を解析することができる。しかし，変換知識は人手で作成されるため，開発コストは高い。

そこで，“相対的意味論に基づく変換主導型統計機械翻訳:Transfer Driven Statistical Machine Translation[1]”（以下，TDSMT）を提案する。TDSMT では，相対的意味論 [5] を利用して，対訳文から「 A が B ならば C は D 」で表現する変換テーブルを自動作成する。入力文を変換テーブルを利用して変換し翻訳を行う。入力文を変換し，翻訳を行うことで文法を順守した翻訳文が期待できる。変換の手順も明確に探索できるため，誤りの原因を解析することも容易である。さらに，変換テーブルは自動で作成されるため，開発コストも低い。

この論文の第2章では，この研究に必要な先行研究を紹介する。第3章では，提案手法である“相対的意味論に基づく変換主導型統計機械翻訳 [1]”の手順や理論を説明する。第4章では，実験データやどのような評価を行ったかを示す。第5章では，第4章で示した評価の結果を示す。第6章では，TDSMT の利点や誤りの解析を紹介する。第7章で，この研究全体をまとめる。

第2章 先行研究

2.1 統計翻訳

統計翻訳システムを原言語 (翻訳の対象となる入力された言語) を日本語文, 目的言語 (翻訳された後に出力される言語) を英語とする日英統計翻訳の場合を例として説明する. なお, 本節は川原らの論文 [4] を引用し, 一部変更を加えて記述している.

2.1.1 単語に基づく統計翻訳

統計翻訳は, 機械翻訳手法の一種である. 日英統計翻訳システムは, 日本語入力文 j が与えられたとき, 全ての組み合わせの中から確率が最大となる英語文 \hat{e} を探索することで翻訳を行う. 以下に基本モデルを示す.

$$\hat{e} = \arg \max_e P(e|j) \quad (2.1)$$

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

$P(j|e)$ は翻訳モデル, $P(e)$ は言語モデルと呼ぶ. \hat{e} を探索する翻訳システムをデコーダと呼ぶ. 図 2.1 に日英統計翻訳の流れを示す.

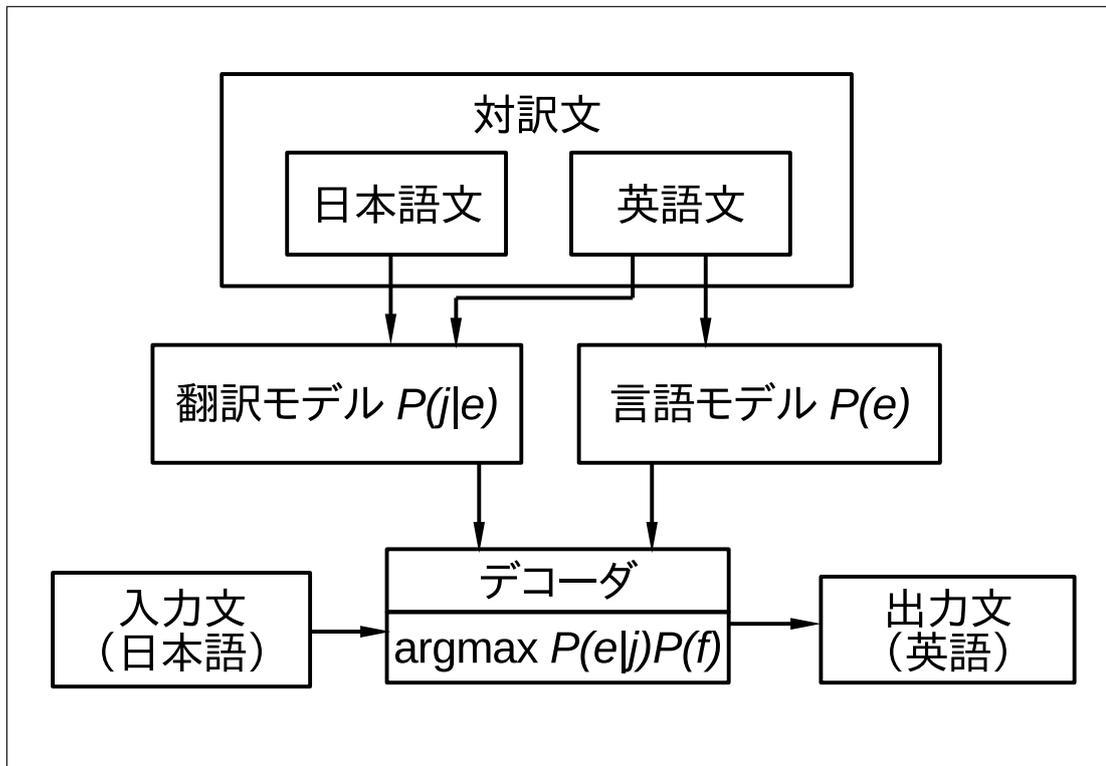


図 2.1 英日統計翻訳の手順

2.1.2 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、 $P(w_i|w_{n-(N-1)}^{i-1})$ 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。

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

2.1.3 翻訳モデル

翻訳モデルは、訳語の尤もらしさを規定する統計モデルである。対訳文 (原言語文と目的言語文) から学習する。翻訳モデルの代表例として、IBM 翻訳モデル [8] がある。

2.1.3.1 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 [8] 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

2.1.3.2 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。

2.1.4 デコーダ

デコーダは、翻訳モデルと言語モデルの全ての組み合わせの中から確率が最大となる出力文を探索して翻訳を行う。翻訳モデルを用いて、原言語と対応する確率が高い単語を探索する。また、言語モデルを用いて、翻訳文が目的言語の文としてふさわしいかの確率を計算する。翻訳モデルの確率と言語モデルの確率の組み合わせでもっともふさわしいものを出力文とする。

2.1.5 句に基づく統計翻訳

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

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。

• Moses

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	

代表的な句に基づく統計翻訳に Moses[9] がある。2.1.3.2 節の GIZA++ によって IBM 翻訳モデルを計算し、言語モデルには SRILM[2] を用いる。Moses はオープンソースの句に基づく統計翻訳である。

2.1.6 統計翻訳の問題点

統計翻訳では、学習文対から統計的な数値を学習し、言語モデルや翻訳モデルを構築する。この言語モデルや翻訳モデルを利用して翻訳を行う。自然言語文は無数に存在するため、限られた学習文対から完全な言語モデルや翻訳モデルを構築するのは不可能である。このため、統計翻訳は精度が低い。

また、言語モデルでは学習文対の種類によって、統計的に偏ったモデルが構築される問題が報告されている。このため、ある種類の学習文対に別の種類の学習文対を追加した場合、翻訳精度は向上しない。または、翻訳精度が低下する。

別の問題として、統計翻訳では学習文対のどの情報が翻訳に利用されたのかが不明確である。このため、統計翻訳に対して、どのような改善を行うのか解析を行うのは容易ではない。このような問題は現在、活発に研究が行われているニューラル機械翻訳にも同じことが言える。

2.2 古瀬らの研究 [6]

古瀬らは“経験的知識を活用する変換主導型機械翻訳 [6]”を提案した。この手法は対訳文に対して、頻度調査や分析を行い4つのレベルの変換知識を人手で作成した。この変換知識と意味距離を利用して、入力文を変換し翻訳を行う。表 2.2 に“経験的知識を活用する変換主導型機械翻訳 [6]”のフローチャートを示す。

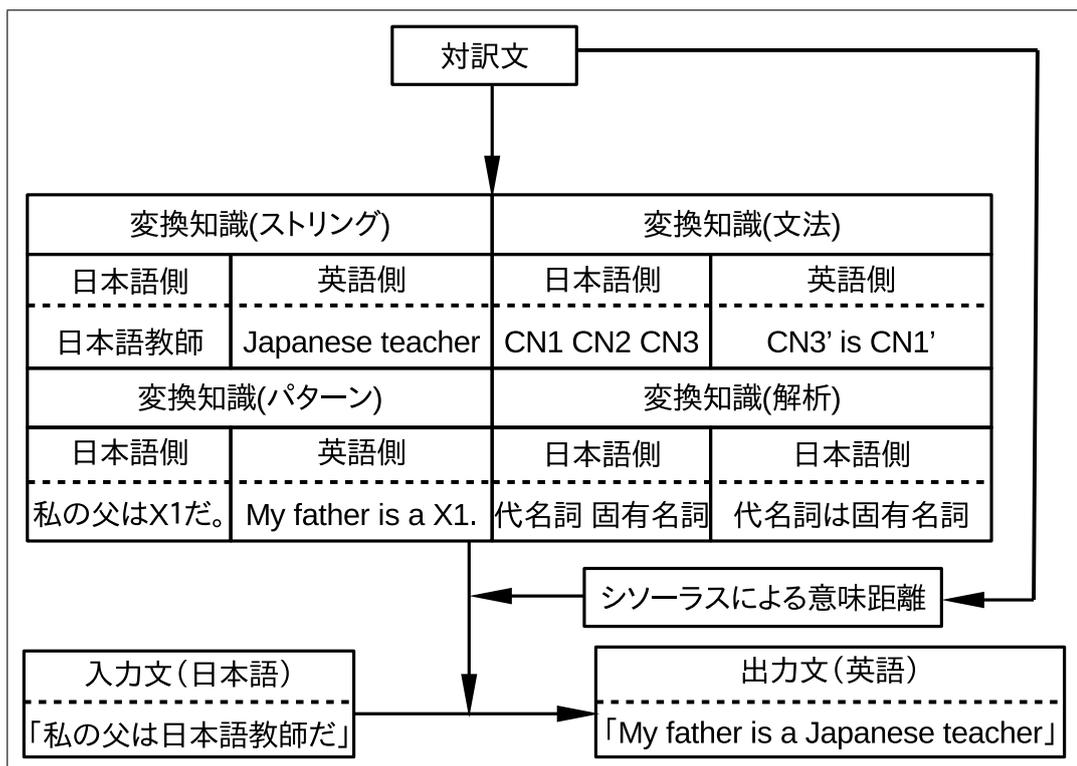


図 2.2 経験的知識を活用する変換主導型機械翻訳のフローチャート

2.2.1 変換知識

変換知識は対訳文を頻度調査や分析することによって得た経験的知識から人手で作成する。対訳文の作成には「国際会議に関する問い合わせ」に関する対訳文を 17,000 文利用した。また変換知識を作成後 17,000 文とは別の訓練データ 825 文を利用して、翻訳実験を行った。この翻訳結果をもとに、変換知識を追加作成した。

変換知識は「ストリングレベル」「パターンレベル」「文法レベル」「解析知識」の4つのレベルが存在する。「ストリングレベル」「パターンレベル」「文法レベル」の変換知識は入

力文（原言語）を出力文（目的言語）に翻訳するために利用する．そのため，原言語を目的言語に変換する知識である．一方，「解析知識」は他3つのレベルの変換知識をうまく利用するため，入力文（原言語）を調整するために利用する．そのため，原言語を原言語に変換する知識である．表2.2に古瀬らが作成した数と具体例を示す．

表 2.2 変換知識の数と具体例

レベル	数	具体例	
ストリングレベル	1783	失礼します	Good-bye
パタンレベル	490	X を Y	Y' X'
文法レベル	45	普通名詞 1 普通名詞 2	普通名詞 1' 普通名詞 2'
解析知識	338	代名詞 固有名詞	代名詞 は 固有名詞

表 2.2 において，「普通名詞 1'」は「普通名詞 1」の翻訳である．

2.2.2 意味距離

変換知識の適用に複数の組み合わせが考えられる入力文があった場合，意味距離が最小となる組み合わせを最終的な出力文とした．古瀬らの利用する意味距離計算では，類語辞典に準拠したシソーラスを利用する．シソーラスの概念階層の意味概念間の位置関係によって，入力文の単語 i と対訳文の単語 e の間の意味距離 $d(i, e)$ を計算する．つまり，意味距離はシソーラスの与え方によって変化する．また，意味距離の値域は 0 から 1 までであり，0 に近いほど両者は類似している．

以上は単語における意味距離の計算法を示した．次に訳し分けの鍵となる入力文の表現と対訳文の表現のいみ距離の計算法を示す．入力文の表現を I ，対訳文の表現を E とする． $d(I, E)$ は I と E を構成する単語の意味距離をもとに算出される．

$$I = (i_1, i_2, \dots, i_t) \quad (2.8)$$

$$E = (e_1, e_2, \dots, e_t) \quad (2.9)$$

$$(2.10)$$

$$d(I, E) = d((i_1, i_2, \dots, i_t), (e_1, e_2, \dots, e_t)) \quad (2.11)$$

$$= \sum -k = 1^t d(i_k, e_k) w_k \quad (2.12)$$

w_k は翻訳における k 番目の要素の重みを表し，値域を 0 から 1 までとする．

2.2.3 実験結果

古瀬らの研究では対訳文に「国際会議に関する問い合わせ」に関する対訳文 [7] を利用した。変換知識の作成には対訳文を 17,000 文利用し，訓練データとして 825 文を利用した。言語対は日本語から英語である。この実験では，テスト文 1056 文に対して，正解率が 60～82%となった。

2.2.4 問題点

古瀬らの研究では，変換知識は人手で作成した。そのため，表 2.2 に示すように変換知識の数は合計 2,656 個と非常に少ない。「国際会議に関する問い合わせ」に関する対訳文 [7] には限られた内容の文が多く存在する。そのため，少ない変換知識の数でも 60%以上の正解率を確保できた。しかし，様々な内容の文に対応するためには大量の変換知識を作成する必要がある。

2.3 機械翻訳における評価手法

機械翻訳における評価手法は自動評価と人手評価がある．自動評価手法の代表的な計算方法を 2.3.1 節で説明する．人手評価手法の中で，今回の実験で利用した対比較評価の方法を 2.3.2 節で説明する．

2.3.1 自動評価法

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．本節では代表的な自動評価手法の説明を示す．なお，本節は川原らの論文 [4] を引用し，一部変更を加えて記述している．

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)[10]，METEOR(Metric for Evaluation of Translation with Explicit ORdering)[11]，RIBES(Rank based Intuitive Bilingual Evaluation Measure)[12] を用いる．

・ BLEU

BLEU[10] 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．

$$\text{BLEU} = \text{BP}_{\text{BLEU}} * \exp\left(\frac{1}{N} \sum_{n=1}^N \log p_n\right) \quad (2.13)$$

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 2.13 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．

$$\text{BP}_{\text{BLEU}} = \begin{cases} 1 & (c > r) \\ e^{1-\frac{r}{c}} & (c \leq r) \end{cases} \quad (2.14)$$

翻訳文の文字数が正解文よりも長い場合には， $\text{BP}_{\text{BLEU}} = 1$ 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，その短さに応じて， 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 2.13 の指数関数 p_n は以下の式で計算される。

$$p_n = \frac{\sum_{\text{全翻訳文}} \sum_{\text{正解文}} \text{正解文と一致する } N\text{-gram 数}}{\sum_{\text{全翻訳文}} \sum_{\text{正解文}} \text{全 } N\text{-gram 数}} \quad (2.15)$$

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。

・ METEOR

METEOR[11] 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。

$$F = \frac{P \times R}{\alpha \times P + (1 - \alpha) \times R} \quad (2.16)$$

$$Pen = \gamma \times (c/m)^\beta \quad (2.17)$$

$$\text{METEOR} = F \times (1 - Pen) \quad (2.18)$$

式 2.17 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。

・ RIBES

RIBES[12] 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 [13] 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．

$$\text{RIBES}(S) = \text{NSR} \times P^\alpha \quad (2.19)$$

$$\text{RIBES}(K) = \text{NKR} \times P^\alpha \quad (2.20)$$

式 2.19 における NSR はスピアマンの順位相関係数である．また，式 2.20 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．

2.3.2 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”，“調査対象手法の出力文”，“ベースラインの出力文”が与えられ，調査対象手法の出力文とベースライン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と考えている．

第3章 提案手法：相対的意味論を利用した変換主導型統計機械翻訳 (TDSMT)

3.1 概要

相対的意味論を利用した変換主導型統計機械翻訳 (以下, TDSMT) は, 学習として, 対訳文から IBM model1 (単語翻訳確率) を利用して変換テーブルを自動作成する. 変換テーブルは相対的意味論に基づいて自動作成する. 変換テーブルの形式は「 A が B ならば C は D 」である. TDSMT は変換テーブルを利用して, 入力文を変換し, 対訳文の日本語側と一致させる. 同じ変換テーブルで対訳文の英語側を変換することによりに翻訳を行う.

まず, 3.2 節で変換テーブルの作成に利用する意味論について論じる. 3.3 節で TDSMT の翻訳の手順を示す. そして, 3.4 節で TDSMT の学習の手順として変換テーブルの自動作成手法を示す.

3.2 相対的意味論 [5]

3.2.1 概要

相対的意味論は「北」の意味を「南の逆の方角」であると意味付ける. このように, ある言葉の意味を別の言葉の意味を利用して定義付ける理論である.

3.2.2 機械翻訳における応用

相対的意味論は, 日本語の言葉の意味を英語の言葉と定義することにより, 翻訳に利用可能である. つまり「日本語句 C の意味は英語句 D である。」として利用する. 本論文では「 A が B ならば C は D 」という変換テーブルの作成の際に利用する. つまり, C の意味 (D) を求めるために A の意味 (B) を利用している.

実際に相対的意味論を利用して自動作成した変換テーブルの例を表 3.1 に示す．表 3.1 内の「()」は対訳文，対訳単語，文パターンの情報から変数に当たる部分が確定される部分を示している．また「{ }」は表の情報だけでは変数に当たる部分が確定できないため，仮で示した変数に当たる部分である．

表 3.1 相対的意味論を利用して自動作成した変換テーブルの例

変換テーブル	A:反省	B:Reflection
	C:嵐	D:The storm
対訳文 1 (日)	(反省) は (知恵) を (増す)	
対訳文 1 (英)	(Reflection) (increases) (wisdom)	
対訳単語 (X00)	増す	increases
対訳単語 (X01)	知恵	wisdom
対訳単語 (X02)	反省	Reflection
文パターン (日)	X02 は X01 を X00	
文パターン (英)	X01 X00 X01	
対訳文 2 (日)	(嵐) は (勢い) を (増す)	
対訳文 2 (英)	(The storm) (grew in } {force)	

表 3.1 の変換テーブルは「“反省”の意味は“Reflection”である」という情報を利用して、「“嵐”の意味は“The storm”である」という情報を対訳文から抽出している．

表 3.1 の変換テーブルでは，一見，A と C は非関連に感じられる．しかし「犬が dog ならば猫は cat」の対偶を考える．この命題の対偶である「猫が cat でないならば犬は dog でない」が成立する．この対偶は表 3.1 の「A が B ならば C は D」が成立する証明となる．

3.2.3 機械翻訳における利点

相対的意味論を機械翻訳に利用することで出力文の文法性が向上すると考えられる．相対的意味論を利用して，変換テーブルを抽出した場合，対訳文中で同じ位置関係にある単語対が変換テーブルとなる．このため，変換テーブルを自動抽出した場合でも，変換テーブル内の単語や句の品詞が揃いやすい．TDSMT は入力文を変換し翻訳する手法である．このような手法では同じ品詞同士で変換した場合，文法構造を変化させないため対訳文の文法構造を維持したまま翻訳を行うことが可能である．逆に，変換テーブル内の品詞が揃っていない場合，文法性が下がる．

また，古瀬らのように経験的知識 [6] を利用した場合と比較して，単純な方法で変換テーブルの作成が可能になる．3.4 節で説明する変換テーブルの自動作成手法は，単純なパターンマッチングに基づいた手法である．

3.3 翻訳

3.3.1 概要

TDSMT における翻訳には、対訳文と変換テーブルを利用する。変換テーブルは「*A*が*B*ならば*C*は*D*」で表現する。変換テーブルを利用することにより、対訳文を入力文と出力文に変換することが可能となる。変換テーブルを利用した変換により、入力文を翻訳するため、文法性の高い出力文が期待できる。図 3.1 に TDSMT による翻訳手法のフローチャートを示す。

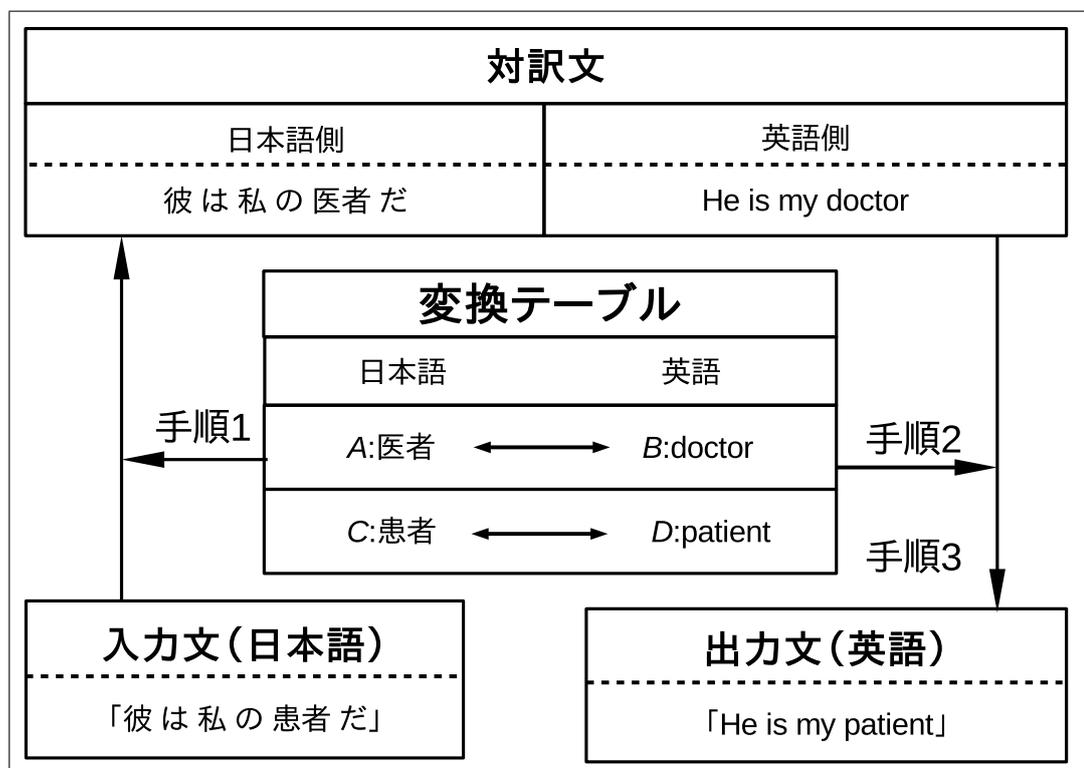


図 3.1 TDSMT による翻訳手法のフローチャート

3.3.2 翻訳の手順

翻訳手法は大きく分けて3つの手順から成る。まず、変換テーブルの日本語側を参照し対訳文の日本語側と入力文を一致させる。そして、同じ変換テーブルの英語側を対訳文の英語側に適用し、出力候補文を作成する。最終的に、言語モデルと変換テーブルに

付与する適用確率を利用して，出力文を出力候補文の中から選択する．日英翻訳の例を利用して手順の詳細を示す．

なお，翻訳は膨大な対訳文と変換テーブルの組み合わせから処理する必要がある．しかし，計算空間や計算時間には限りがあるため，経験的に決定した閾値による枝刈り処理を各手順で利用している．また，翻訳候補文の作成には beam-search による探索技術を利用している．

3.3.2.1 変換テーブルの選択

変換テーブルの日本語側を利用して，入力文を対訳文の日本語側と一致させる変換テーブルを選択する．翻訳で利用する対訳文は入力文と距離の近い 256 文を選択して利用している．この 256 文は入力文ごとに，対訳文全体から選択する．変換テーブルの選択例を表 3.2 に示す．

表 3.2 変換テーブルの選択例

対訳文	
彼は私の <u>医者</u> だ	He is my doctor
変換テーブル	
日本語側	英語側
A:医者	B:doctor
C:患者	D:patient
入力文	
彼は私の <u>患者</u> だ	

3.3.2.2 変換テーブルの適用

選択した変換テーブルの英語側に従い，対訳文の英語側を変換し，出力候補文を作成する．変換テーブルの適用例を表 3.3 に示す．

出力候補文は変換テーブルの選択と適用が成功した文である．出力候補文は大量に作成されることが予測される．このため，経験的に決定した閾値による枝刈りや beam-search を利用し，計算時間の短縮を目指す．

表 3.3 変換テーブルの適用例

対訳文	
彼は私の医者だ	He is my <u>doctor</u>
抽出された変換テーブル	
日本語側	英語側
A:医者	B:doctor
C:患者	D:patient
出力候補文	
He is my <u>patient</u>	

3.3.2.3 出力文の決定

出力候補文の中から，言語モデルと変換テーブルの適用確率を用いて，出力文を決定する．出力文の決定例を表 3.4 に示す．なお，確率は 2 の対数で表記する．よって，負の絶対値の大きい値ほど確率は低い．負の絶対値が小さい値ほど確率は 1 に近づく．

現在，言語モデルには High-order-Joint-Probability[17] を利用している．変換テーブルの適用確率には dice 係数を利用して，計算した確率を参照する．各翻訳候補文に対して，言語モデルと変換テーブルの適用確率の合計値を計算する．確率値の合計値が最も大きい出力候補文を出力文とする．本研究ではこの出力文を評価に利用する．また，上位 4 つの出力候補文を解析用に保存する．

表 3.4 変換テーブルの適用例

出力候補文	言語モデル	変換テーブルの適用確率
He is my <u>patient</u>	-230	-25
He is my <u>patients</u>	-560	-28
最終的な出力文		
He is my patient		

3.3.3 TDSMT の出力例

表 3.5 に TDSMT による実際の翻訳例を示す．入力文の「彼は大学で法律を修めた」に対して，対訳文「彼はテニスで相手を破った：He defeated his opponent at tennis」と 4 つの変換テーブルを利用して，「He completed the law at college」が出力される．表 3.5

の例は出力候補文中で言語モデルと変換テーブル適用確率の合計値が最大の出力候補文である。

表 3.5 TDSMT の翻訳例

入力文	彼は(大学)で(法律)(を)(修めた)	
参照文	He studied law at college	
対訳文(日)	彼は(テニス)で(相手)(を)(破った)	
対訳文(英)	He (defeated) (his) (opponent) at (tennis)	
変換 テーブル 1	A:テニス	B:tennis
	C:大学	D:college
変換 テーブル 2	A:相手	B:opponent
	C:法律	D:law
変換 テーブル 3	A:を	B:his
	C:を	D:the
変換 テーブル 4	A:破った	B:defeated
	C:修めた	D:completed
出力文	He (completed) (the) (law) at (college)	

3.4 学習:変換テーブルの自動作成

3.4.1 変換テーブルの自動作成手法の概要

TDSMT では、学習として、変換テーブルを自動作成する。変換テーブルの形式は「*A*が*B*ならば*C*は*D*」である。また、変換テーブルには適用される確率（変換テーブルの適用確率）を付与する。図 3.2 に変換テーブルの自動作成手法のフローチャートを示す。

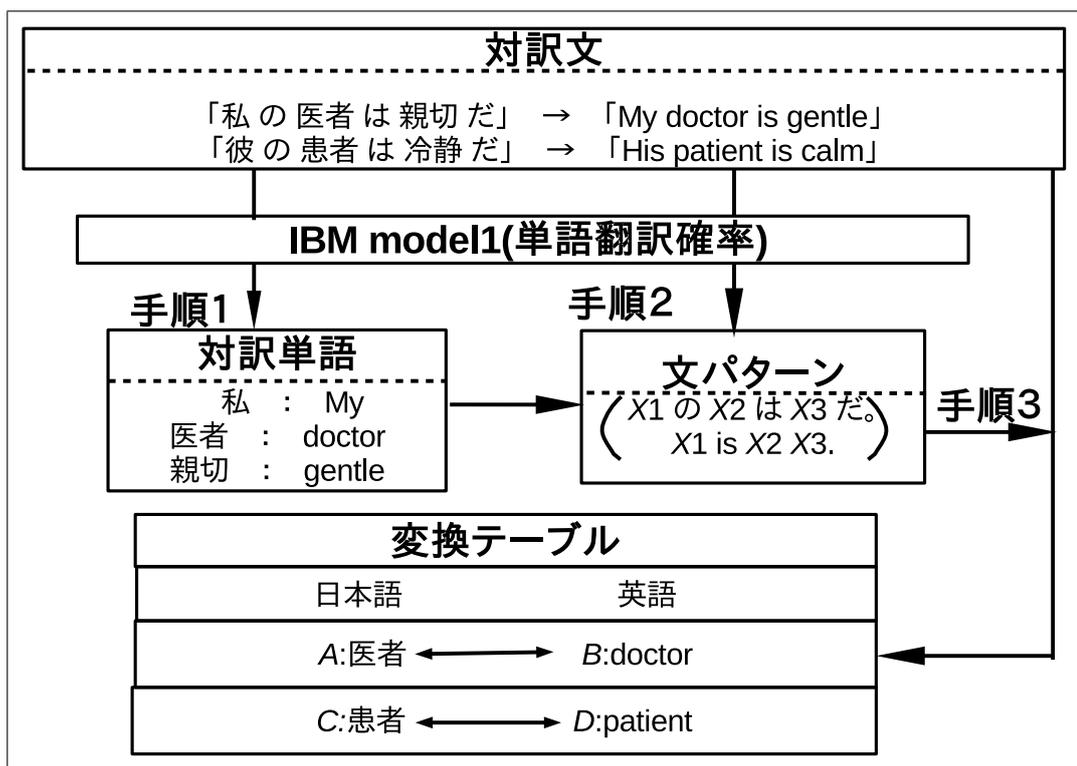


図 3.2 変換テーブルの自動作成手法のフローチャート

3.4.2 変換テーブルの自動作成手法の手順

変換テーブルの自動作成手法は大きく分けて3つの手順から成る。まず、対訳文と IBM model1(単語翻訳確率)を利用して、対訳単語を作成する。次に、対訳単語と対訳文から文パターンを作成する。最後に、文パターンと対訳文を照合し、変換テーブルを作成する。単語日英翻訳による例を利用して手順の詳細を示す。

3.4.2.1 対訳単語の作成

対訳単語とは、対訳文中の日本語単語と英語単語の組み合わせである。対訳単語には日本語単語が英語単語に翻訳される確率と英語単語が日本語単語に翻訳される確率やその順位などの情報が含まれる。

IBM model1(単語翻訳確率)を用いて、対訳文から対訳単語を作成する。単語翻訳確率はGIZA++-v1.0.7[3]を利用して得る。学習回数は4回である。対訳単語の作成例を表3.6に示す。

表 3.6 対訳文から作られる対訳単語

対訳文	
私の 医者は親切だ	My doctor is gentle
対訳単語	
私	My
医者	doctor
親切	gentle
etc...	

なお、以下の条件に当てはまる対訳単語は削除し、以降の処理には使用しない。これは、学習速度、学習の安定性を目的とした枝刈り処理である。

- 日本語 → 英語方向の単語翻訳確率が 0.01 以下
- 英語 → 日本語方向の単語翻訳確率が 0.01 以下
- 日本語 → 英語方向の単語翻訳確率の順位が 8 位以降
- 英語 → 日本語方向の単語翻訳確率の順位が 8 位以降
- 対訳単語の対訳文中の共起頻度が 1 回以下

3.4.2.2 文パターンの作成

文パターンとは、対訳文中の対訳単語に当たる部分を変数化した文対である。また、文パターンには文パターンの作成に利用した対訳文（文パターン原文）や単語翻訳確率を利用して得た文パターンの確率を情報として含む。

対訳文と対訳単語を照合する。対訳文中で対訳単語にあたる箇所を変数化し、文パターンを作成する。文パターンの作成例を表 3.7 に示す。

表 3.7 対訳単語と対訳文から作られる文パターン

対訳単語	
私	My
医者	doctor
親切	gentle
対訳文（文パターン原文）	
私の医者は親切だ	My doctor is gentle
文パターン	
X1 の X2 は X3 だ	X1 X2 is X3

3.4.2.3 変換テーブルの作成

変換テーブルは「A が B ならば C は D」という形式である。「A が B」には文パターンの作成の際に利用した対訳単語が当たる。そして、「C は D」には文パターン原文とは別の対訳文に文パターンを当てることによって抽出した日本語句と英語句の対が当たる。また、変換テーブルには IBM model1 を利用して計算した変換テーブルの適用確率、dice 係数を利用して計算した変換テーブルの適用確率、またその順位などを情報として含む。

文パターンと文パターン原文とは別の対訳文を照合する。変数化した対訳単語と変数に当たる句を変換テーブルとする。変換テーブルの作成例を表 3.8 に示す。

なお、以下の条件に当てはまる変換テーブルは削除し、以降の処理には使用しない。これは、学習速度、学習の安定性を目的とした枝刈り処理である。

- 「C は D」日本語 → 英語方向の翻訳確率の順位が 16 位以降 (IBM model1 による確率を参照)

表 3.8 文パターンと対訳文から作られる変換テーブル

文パターン	
X1 の X2(医者) は X3 だ	X1 X2(doctor) is X3
対訳文 (文パターン原文)	
私の <u>医者</u> は <u>親切</u> だ	My <u>doctor</u> is <u>gentle</u>
対訳文	
彼の <u>患者</u> は <u>冷静</u> だ	His <u>patient</u> is <u>calm</u>
X2 における変換テーブル	
日本語側	英語側
A:医者	B:doctor
C:患者	D:patient

- 「C は D」英語 → 日本語方向の翻訳確率の順位が 64 位以降 (IBM model1 による確率を参照)
- 「C は D」の対訳文中の共起頻度が 1 回以下
- 「A は B」日本語 → 英語方向の単語翻訳確率の順位が 4 位以降 (IBM model1 による確率を参照)
- 「A は B」英語 → 日本語方向の単語翻訳確率の順位が 4 位以降 (IBM model1 による確率を参照)
- 「A は B」の対訳文中の共起頻度が 1 回以下

3.4.3 変換テーブルの作成例

変換テーブルの実際の作成例を表 3.9 に示す。表 3.9 内の「()」は対訳文，対訳単語，文パターンの情報から変数に当たる部分が確定される部分を示している。また「{}」は

表の情報だけでは変数に当たる部分が確定できないため，仮で示した変数に当たる部分である．

表 3.9 変換テーブルの実際の作成例

変換テーブル	A:反省	B:Reflection
	C:嵐	D:The storm
対訳文 1 (日)	(反省) は (知恵) を (増す)	
対訳文 1 (英)	(Reflection) (increases) (wisdom)	
対訳単語 (X00)	増す	increases
対訳単語 (X01)	知恵	wisdom
対訳単語 (X02)	反省	Reflection
文パターン (日)	X02 は X01 を X00	
文パターン (英)	X01 X00 X01	
対訳文 2 (日)	(嵐) は (勢い) を (増す)	
対訳文 2 (英)	(The storm) (grew in } {force)	

表 3.9 では対訳文 1 「反省は知恵を増す : Reflection increases wisdom」から 3 つの対訳単語を作成した．そして，対訳単語と対訳文 1 から文パターン「X02 は X01 を X00 : X01 X00 X01」を作成した．この文パターンと対訳文 2 「嵐は勢いを増す : The storm grew in force」を照合することにより変換テーブルが作成された．

第4章 実験環境

4.1 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 4.1 に示す。

表 4.1 実験データ

対訳文	160,000 文
入力文	1,000 文

本実験ではオープンデータによる実験を行い、クローズデータによる実験は行わない。つまり、入力文を利用した学習は行わない。表 4.2 に対訳文と入力文の例を示す。

表 4.2 実験データの例

対訳文	
対訳文 (日)	滓は底に沈澱する。
対訳文 (英)	The sediment settles at the bottom.
対訳文 (日)	彼女はいたずらに引っかかるような子ではない。
対訳文 (英)	She is not a girl to trick with.
対訳文 (日)	それには料金がかかります。
対訳文 (英)	There is a fee for that.
入力文	
入力文 (日)	勝手に電話を使ってください。
参照文 (英)	You are welcome to our telephone.
入力文 (日)	彼は颯と軍を引き揚げた。
参照文 (英)	All of a sudden he withdrew his army.
入力文 (日)	仕事は速やかに進んでいる。
参照文 (英)	The work is making rapid progress.

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 4.2 の例はこの処理を行った後の状態である。

入力文はコーパスから抽出しているため、英語側が存在する。しかし、この入力文の英語側を利用した学習は行わない。入力文の英語側は参照文として、自動評価や人手評価の際の参考にする。

4.2 評価方法

本研究では TDSMT を 4 つの基準で評価する。本研究では主に、TDSMT の翻訳可能な入力文の割合（カバー率）に焦点を当てて評価を行う。これは、TDSMT は手法が厳格であるため、統計翻訳と比較して、カバー率が低くなることが予測されるためである。

カバー率の原因を観測するため、変換テーブルの数や精度、実際に翻訳に利用する変換テーブルの数を調査した。報告は出力が得られた順に行う。

1) 作成した変換テーブルの数

3.4 節の方法で作成した変換テーブルの数を調査する。

2) 作成した変換テーブルの精度

3.4 節の方法で作成した変換テーブルの中からランダムで 100 個抜き出し、人手評価を行い精度を調査する。

3) 翻訳に利用する変換テーブルの数

従来手法と提案手法で翻訳に利用する変換テーブルの数を比較調査する。なお、変換テーブルの数は入力文一文あたりの平均数で示す。

4) カバー率の調査

TDSMT で翻訳実験を行う。1,000 文を翻訳し、カバー率を調査する。なお、この調査においては、翻訳の正誤は問わない。

第5章 実験結果

5.1 作成した変換テーブルの数の調査

変換テーブルの自動作成手法で作成した変換テーブルの数を以下の表 5.1 に示す。

表 5.1 作成した変換テーブルの数

変換テーブルの数	5,744,175(個)
----------	--------------

表 5.1 より，変換テーブルの大量作成に成功した．人手で表 5.1 に示す数の変換テーブルを作成，管理することは現実的ではない．このような大量の変換テーブルを作成，利用できるため，TDSMT は高い翻訳精度を期待できる．

5.2 作成した変換テーブルの精度の調査

表 5.1 の変換テーブルの数から，ランダムに 100 個を抜き出し，その精度を人手評価した．なお，「 A が B ならば C は D 」の「 A が B 」の部分と「 C は D 」の部分分割し，それぞれの対応を評価した．評価基準を表 5.2 に示す．

表 5.2 変換テーブルの評価基準

	日本語側と英語側が正しく対応している
	日英のどちらかに余分な単語を含む
×	日本語側と英語側が間違っただ対応をしている

評価結果を表 5.3 に示す．表 5.3 より変換テーブルの精度は高いことがわかる．

表 5.3 より変換テーブルの精度は高いことがわかる．精度の高い変換テーブルを利用して翻訳を行うことで，翻訳精度も高いことが予測される．

「 A が B 」の部分は単語の対応であり，「 C は D 」の部分は句の対応であるため，「 A が B 」の部分と「 C は D 」の部分の評価には差が生じていると考えられる．また，「 A が B 」

表 5.3 作成した変換テーブルの精度

			×
A が B	95	2	3
C は D	75	16	9

の部分と「C は D」の部分で作成方法は異なる。「A が B」の部分は IBM model1 の単語翻訳確率から作成する。一方、「C は D」の部分は単純なパターンマッチングによって作成する。この作成方法の違いも評価の差に影響していると考えられる。

5.2.1 評価例：変換テーブル(A が B)

変換テーブルの A が B の部分に対する評価例を表 5.4, 表 5.5 表 5.6 に示す。

表 5.4 変換テーブル(A が B) の評価 の例

評価	変換テーブル(A が B)	
	台風	typhoon
	はためいた	fluttered
	新	new

表 5.5 変換テーブル(A が B) の評価 の例

評価	変換テーブル(A が B)	
	ぱっと	burst
	し	have

表 5.6 変換テーブル(A が B) の評価 × の例

評価	変換テーブル(A が B)	
×	ずたずた	pieces
×	した	They
×	を	a

5.2.2 評価例：変換テーブル (C が D)

変換テーブルの A が B の部分に対する評価例を表 5.7, 表 5.8 表 5.9 に示す .

表 5.7 変換テーブル (C が D) の評価 の例

評価	変換テーブル (C が D)	
	逃亡	escape
	都市 の 中心	city center
	を 務めた	acted as

表 5.8 変換テーブル (C が D) の評価 の例

評価	変換テーブル (C が D)	
	その 試験	test
	肩	his sholders
	大声 で 助け を	for help

表 5.9 変換テーブル (C が D) の評価 × の例

評価	変換テーブル (C が D)	
×	歌っ	,please
×	待ち なさい	your fork in
×	彼女 の 指輪	diamond

5.3 カバー率の調査

TDSMT において入力文 1,000 文に対して翻訳実験を行った . カバー率を表 5.10 に示す .

表 5.10 各手法のカバー率

カバー率	19.6 % (196 文/1,000 文)
------	------------------------

表 5.10 より , TDSMT ではカバー率は低かった . Moses や NMT では , ほぼ 100% のカバー率を有している .

5.4 実験結果のまとめ

今回、行った TDSMT の翻訳実験では、変換テーブルが大量に作成できた。しかし、カバー率はまだまだ低かった。カバー率は Moses や NMT ではほぼ 100% のカバー率を有しているため、現在主流となっている手法と公正に比較することは難しい。

カバー率が低い問題の原因として、beam-search のとり方などがある。しかし、最も大きい問題点は入力文中に対訳文に存在しない単語や対訳文中に出現頻度が少なく学習が難しい単語が出現した場合、その入力文は翻訳不可になってしまう。学習の工程で、変換テーブルを作成する際、学習の難しい単語や句は枝刈り処理を行う。このため、学習した変換テーブルと対訳文の組み合わせでは対応できない。このような現象が、カバー率の低下の原因となっている。

第6章 考察

6.1 TDSMT と Moses[9] との翻訳精度の比較

この節では、TDSMT の翻訳精度について考察する。一般的な統計翻訳の手法である moses と比較調査を行うことによって、TDSMT の翻訳精度がどの程度有効であるかを調査した。

5.4 節に示すように、TDSMT と moses ではカバー率に差がある。そのため、TDSMT で出力可能な入力文を対象にして翻訳精度の調査を行う。

6.1.1 自動評価

TDSMT で出力可能であった入力文 196 文で翻訳実験を行った。TDSMT と moses の出力文に対して、自動評価を行った。評価手法は 2.3.1 に示した、BLEU、METEOR、RIBES を採用した。結果を表 6.1 に示す。

表 6.1 TDSMT と moses の自動評価結果

手法	BLEU[10]	METEOR[11]	RIBES[12]
moses	0.263	0.565	0.791
TDSMT	0.232	0.536	0.805

表 6.1 より、BLEU 値と METEOR 値では TDSMT は moses より低かった。RIBES では TDSMT と moses は同等程度という結果となった。

6.1.2 人手評価

TDSMT で出力可能であった入力文 196 文からランダムに抽出した 100 文に対して、翻訳実験を行って、人手評価を行った。人手評価は 2.3.2 節に示した対比較調査を行った。調査対象手法を TDSMT、ベースラインを Moses とした。どちらの出力が TDSMT の出

力文で、どちらが Moses の出力文かが評価者がわかる場合、評価が TDSMT に有利になる可能性がある。そのため、出力文をランダムに入れ替えて評価を行った。なお評価者は、本論文の著者一名で行った。

評価基準を表 6.2 に示す。

表 6.2 対比較評価の評価基準

従来手法	TDSMT の出力文の方が優れている
Moses	Moses の出力文の方が優れている
差なし	二つの出力文の優劣がつけにくい
同一出力	二つの出力文が一致している

結果を表 6.3 に示す。

表 6.3 TDSMT と Moses との対比較評価結果

TDSMT	Moses	差なし	同一出力
16	15	47	22

表 6.3 より、TDSMT は翻訳文の出力に成功した場合、Moses と同等の翻訳精度を示した。今回の TDSMT はまだまだ開発段階である。今後改善を行うことにより、moses に勝る翻訳精度が期待できる。

6.1.2.1 人手評価例：TDSMT の例

TDSMT の評価例を表 6.4、表 6.5 及び表 6.6 に示す。

表 6.4 TDSMT の人手評価例 1

入力文	彼女はろうそくを吹き消した
参照文	She puffed the candle out
TDSMT	She blew out the candle
Moses	She is a candle

表 6.4 の例では、TDSMT では入力文と同じ意味の翻訳文が出力された。しかし、Moses の出力では動詞が be 動詞のみであり、「吹き消した」の意味が読み取れない。そのため、評価を TDSMT とした。

表 6.5 TDSMT の人手評価例 2

入力文	潮が差している
参照文	The tide is rising
TDSMT	The tide is rose
Moses	The tide is in

表 6.5 の例では，TDSMT の出力文には「潮」と「差す」という意味が「tide」と「rose」という英単語として含まれている．しかし，Moses の出力文には「差す」を表すような単語は出現していないため，TDSMT とした．TDSMT の出力文の時制を考慮すると，完全に入力文の意味を表しているは言い難い．しかし，対比較評価では，どちらがより正解に近いかを評価するため，この文の評価を TDSMT とした．

表 6.6 TDSMT の人手評価例 3

入力文	田畑 は 火山灰 に 埋まった
参照文	The fiels was overwhelmed by an accumulation of volcanic ashes
TDSMT	The fields were buried in volcanic ashes
Moses	The volcanic ash was buried in the field

表 6.6 の例では，TDSMT の出力文では入力文の意味が正しく読み取ることができる．しかし，Moses の出力文では，文構造は受動態のままで，動作主と目的格が入れ替わっている．このため，入力文の意味を読み取るのは難しい．よって，この文の評価を TDSMT とした．

6.1.2.2 人手評価例：Moses の例

Moses の評価例を表 6.7，表 6.8 及び表 6.9 に示す．

表 6.7 Moses の人手評価例 1

入力文	日程 を よく 調べて ください
参照文	Please check the dates carefully
TDSMT	Please investigating a good schedule
Moses	Please check carefully the schedule

表 6.7 の例では、TDSMT では入力文の「よく」が「日程」を修飾する文法で出力された。Moses の出力では「carefully」を利用し動詞の「check」を修飾する文法で出力した。意識するとどちらの出力文も入力文の意味を読み取れる。Moses の出力文の方が入力文の修飾関係に近いと評価を Moses とした。

表 6.8 Moses の人手評価例 2

入力文	この犬は私のそばを離れない
参照文	This dog won't leave my side
TDSMT	The dog is chained to to me
Moses	This dog is chained to me

表 6.8 の例では、両方の出力文で「そばを離れない」の訳出に「is chained to」が使われている。意識をすれば入力文の意味と読み取ることが可能である。しかし、TDSMT の出力文では「chained」の後が「to to」と重複して出力されている。このため、この文の評価を Moses とした。TDSMT では変換テーブルに過不足がある場合や、助詞「の」や「を」などに変換テーブルが適用され場合、このような問題が多く観測されている。

表 6.9 Moses の人手評価例 3

入力文	日本人は箸でものを食べる
参照文	The Japanese eat their food with hashi
TDSMT	He eats Japanese things with chopsticks
Moses	The Japanese eat anything with chopsticks

表 6.9 の例では、この入力文では TDSMT の出力文と Moses の出力文で正しい翻訳文が出力されていない。TDSMT の出力文では主語が彼になっている。Moses の出力文では「anything」ではなく「something」を用いるのが正しい。両方の出力文が誤っている場合、どちらの出力文がより、入力文の意味を読み取りやすいかを基準とするため、この文の評価を Moses とした。

6.1.2.3 人手評価例：差なしの例

Moses の評価例を表 6.10、表 6.11 及び表 6.12 に示す。

表 6.10 では、主語は入力文において省略されているため「He(彼)」と「I(私)」で差はない。また「誘惑に負ける」に対する表現も「lost temptation」と「gave in to temptation」

表 6.10 差なしの人手評価例 1

入力文	一時の誘惑に負けた
参照文	He succumbed to a momentary temptation
TDSMT	He lost temptation one time
Moses	I gave in to temptation of one o'clock

では表現が違うが意味は同じであるため、差はない。「一時の」に対して、TDSMT では「one time(前の, かつての)」, Moses では「of one o'clock(時間の 1:00 の)」と出力された。どちらも、入力文全体の意味から不適切であるため、評価を差なしとした。

表 6.11 差なしの人手評価例 2

入力文	子供達は公園に遊びに行った
参照文	The children went to play in the park
TDSMT	The children went to see in the park
Moses	The children went to the park

表 6.11 では、TDSMT の出力文と Moses の出力文の差は「see in」があるかどうかである。しかし、この句があるかどうかにかかわらず、入力文の「遊びに」の意味が読み取れない。そのため、この文の評価を差なしとした。

表 6.12 差なしの人手評価例 3

入力文	その試みは失敗に終わった
参照文	That try ended in failure
TDSMT	The attempt was abortive
Moses	The attempt ended in failure

表 6.12 では、TDSMT の出力文と Moses の出力文の差は「失敗に終わった」の訳出である。TDSMT の出力文では「was abortive」と訳し、Moses の出力文では「ended in failure」と訳された。Moses の出力文は参照文と同じだが、TDSMT の出力文では異なっている。しかし、「abortive」には「失敗した」という意味があり、入力文の意味を正しく読み取ることが可能であるため、この文の評価を差なしとした。なお「attempt」には「試み」という名詞の意味が存在し、入力文の翻訳に成功している。

6.1.2.4 人手評価例：同一出力の例

Moses の評価例を表 6.13 に示す .

表 6.13 同一出力の人手評価例

入力文	電車が混んでいる
参照文	The train is crowded
TDSMT	The train is crowded
Moses	The train is crowded

表 6.13 では TDSMT の出力と Moses の出力が完全に一致している . このような場合 , 評価を同一出力とした .

6.2 TDSMT の利点

TDSMT は翻訳の手順を詳しく解析することが可能である。TDSMT では、翻訳に利用した対訳文、変換テーブルを調べることが可能である。また、利用された変換テーブルの作成に使用した対訳文、対訳単語も調べることが可能である。このように、TDSMT は確率計算以外のほぼすべての翻訳手順が明示されている手法である。このため、翻訳の誤りの追求が可能となる。つまり、改善点の探索が非常に容易である。

一方、現在主流となっているニューラル機械翻訳や Moses などの句に基づく統計翻訳では、このような詳しい解析は困難である。これらの手法は統計的な計算を中心とした手法である。計算結果は示されるが、どの情報が翻訳結果に影響したのか調べることは困難である。データベースの追加以外の方法で統計的な確率を中心とした手法を発展させるのは困難である。

現在行われているニューラル機械翻訳に関する研究では、複数のニューラル機械翻訳の手法から性能の高いモジュールを取り出し、組み合わせるような研究が多い。つまり、実際の翻訳結果を解析し、誤りに基づいた改善を目指す研究は少ない。

TDSMT では誤りに基づいた改善が行えるため、確実な翻訳精度の向上が期待できる。今回出力されたカバー率や、翻訳精度は発展途上の結果である。改善をすすめることで十分な翻訳性能を持つ手法になり得ることが TDSMT の利点である。

6.3 誤り解析

6.2 節では，TDSMT で翻訳の誤りの原因を解析することが可能であるとした．実際の翻訳例を利用して，TDSMT による翻訳手順の誤り解析の例を示す．なお，6.1.2 節の対比較評価で，moses と評価された入力文において解析した．

Moses と評価された入力文 14 文中の誤りのおおまかな内訳を表 6.14 に示す．

表 6.14 誤りのおおまかな内訳

変換テーブルの誤り	8/15
利用する対訳文の誤り	4/15
確率値による出力文の選択の誤り	3/15

表 6.14 より，変換テーブルの誤りにより，翻訳が失敗するケースが多かった．6.3.1 節，6.3.2 節，6.3.3 節に誤りの例を示す．

6.3.1 変換テーブルの誤り

表 6.15 TDSMT の誤り例（変換テーブルの誤り）

入力文	(授業) に (間に合った)	
参照文	I was in time for the class	
対訳文 (日)	(ハイキング) に (行った)	
対訳文 (英)	I (went) (hiking)	
変換 テーブル 1	A:ハイキング	B:hiking
	C:授業	D:class
変換 テーブル 2	A:行った	B:went
	C:間に合った	D:just
出力文	I (just) (class)	

表 6.15 の翻訳例では，対訳文の選択は適切であると言えるが，変換テーブル 2 の「C は D」の部分不適切と言える。「間に合った」の英語訳は「was in time for」などがある．この変換テーブル 2 の作成に利用した情報を表 6.16 に示す．

表 6.16 では対訳文 2 では「間に合った : just barely met」に文パターンが照合することで正しい変換テーブルが得られる．しかし，このようにした場合，他の変数には対応

表 6.16 表 6.15 の変換テーブル 2

対訳文 1 (日)	(ハイキング) (に) (行った)	
対訳文 1 (英)	(I) (went) (hiking)	
対訳単語 (X00)	に	I
対訳単語 (X01)	ハイキング	hiking
対訳単語 (X02)	行った	went
文パターン (日)	X01 X00 X02	
文パターン (英)	X00 X02 X01	
対訳文 2 (日)	(辛くも 締め切り) { { に) (間に合った)	
対訳文 2 (英)	(I) (just) (barely met the deadline)	
変換テーブル	A:行った	B:went
	C:間に合った	D:just

が誤った部分が当たる。「辛くも締め切り：the deadline」そのため、この対訳文 2 と文パターンの組み合わせは不適切である。

表 6.16 では対訳文 1 と対訳文 2 には文法の差が少しある。対訳文 2 では文頭に「辛くも」という形容詞がある。これは英語側では「just barely」にあたる。このため、文パターンを照合した際誤った変換テーブルが作成された。

また、仮に「C は D」が「間に合った：just barely met」の変換テーブルを得られたとしても表 6.15 の入力文には一般的には使われない。このように、目的語や主語に対応した動詞を利用しなくてはならない問題がある。このような問題点は言語モデルで補うのが現在のきかい 翻訳の主流である。TDSMT でもこの流れに従う。

また、対訳文には「間に合った」の表現に関して「電車に間に合った：was in time, just made」「開演時間に間に合う：was just in time for」など様々な表現がある。これらの対訳文を利用して、正しい表現を抽出必要がある。

6.3.2 利用する対訳文の誤り

表 6.17 により、たくさんの誤りの原因が考えられる。一つは翻訳に利用した対訳文が不適切である。入力文は主語が存在するが、翻訳に利用した対訳文では主語が省略されている。このような場合、変換により翻訳文を導くのは困難である。具体的には、変換テーブル 1 を利用して、入力文の主語である「わたし」と対訳文の目的語である「彼女」が変換されている。このため、出力文で不適切な位置に「my」という単語が出力されて

表 6.17 TDSMT の誤り例 (対訳文の選択の誤り)

入力文	(わたし) (は 彼) (に) 郵便 で (小包) (を) 送った	
参照文	I mailed him a parcel	
対訳文 (日)	(彼女) (に) (小包) 郵便 で (本) (を) 送った	
対訳文 (英)	I sent (a) (book) (to) (her) by (parcel) post	
変換 テーブル 1	A:彼女	B:her
	C:わたし	D:my
変換 テーブル 2	A:に	B:to
	C:は 彼	D:him
変換 テーブル 3	A:小包	B:parcel
	C:に	D:in
変換 テーブル 4	A:本	B:book
	C:小包	D:parcel
変換 テーブル 5	A:を	B:a
	C:を	D:the
出力文	I sent (the) parcel (him) (my) by (in) post	

しまった。

翻訳確率 (言語モデル + 変換テーブルの適用確率) の順位で 4 位までを解析しても , 表 6.17 に示す対訳文以外の対訳文を利用する出力候補文は存在しなかった。

6.3.3 確率値による出力文の選択の誤り

表 6.18 TDSMT の誤り例 (確率値が第一位の翻訳候補文)

入力文	(あいつ) はいつも不平ばかり言っている	
参照文	That fellow is grumbling all the time	
対訳文 (日)	(彼) はいつも不平ばかり言っている	
対訳文 (英)	(He) is always complaining	
変換 テーブル 1	A:彼	B:he
	C:あいつ	D:He's
出力文	(He's) always complaining	
言語モデル	high order joint probability	-250.3029
変換テーブルの適用確率		-16.8469

表 6.19 TDSMT の誤り例 (確率値が第 2 位の翻訳候補文)

入力文	(あいつ) はいつも不平ばかり言っている	
参照文	That fellow is grumbling all the time	
対訳文 (日)	(彼) はいつも不平ばかり言っている	
対訳文 (英)	(He) is always complaining	
変換	A:彼	B:he
テーブル 1	C:あいつ	D:That guy
出力文	(That guy) always complaining	
言語モデル	high order joint probability	-264.7922
変換テーブルの適用確率	-19.5901	

表 6.18 と表 6.19 に同一の入力文における第一位の翻訳候補文と第二位の翻訳候補文を示す。なお、確率値は対数をとっているため、0 に近い値が大きい確率を示す。一方、マイナスの絶対値が大きくなるほど低い確率を示す。

入力文「あいつはいつも不平ばかり言っている」に対して、表 6.18 の翻訳候補文が出力された。しかし、表 6.18 では「あいつ」が「he's」と翻訳され、英語の文法にあわない。一方、表 6.19 の第二位の翻訳候補文では「あいつ」が「That guy」に翻訳され、文法構造をくずしておらず、入力文の意味を読み取りやすい翻訳となっている。このため、表 6.19 を出力文とすることが好ましい。しかし、言語モデルの確率、変換テーブルの適用確率ともに表 6.18 の翻訳候補文に劣っている。そのため、表 6.18 の翻訳候補文が出力文に選択された。

言語モデルは現在 “high order joint probability” を採用している。この言語モデルは目的言語の言語モデルの構築に原言語文と目的言語文を利用する。そのため、出力候補文の言語モデルを計算するために入力文を利用する。そのため、「He's」のような対訳文中に大量に存在する単語はモデル構築の上で、確率値が高くなりやすい。このため、表 6.19 と比較して、表 6.18 の言語モデルの方が言語モデルの値が高くなったとかがえられる。

変換テーブルの適用確率は IBM model1 の単語翻訳確率をもとに対訳文中の単語の出現頻度や共起頻度をもとに推定される。そのため、言語モデルの場合と同様に、人称代名詞のような対訳文中の頻度の高い単語を含む変換テーブルは適用確率が大きくなりやすい。このため、表 6.19 と比較して、表 6.18 の変換テーブルの適用確率の値が高くなったとかがえられる。

言語モデルの計算方法、変換テーブルの適用確率の計算方法をより最適なものにする

必要がある．また，2つの確率値をうまく利用して，出力文を決定する方法を考案する必要がある．

6.3.4 手法の改善

TDSMT の誤りの原因の一つに変換テーブルの「 C は D 」の部分が誤っていることが挙げられる．TDSMT による翻訳では，変換テーブルの誤りが出力文に出力される．人手評価で Moses よりも翻訳精度が低い出力文ではこの問題が多く観測された．今後は，変換テーブルにおいて，表 5.3 に見られる，評価 \square や評価 \times を削減する方法を考案する．改善法について，いくつか列挙する．

- 1) 対訳単語のみから変換テーブルを作成
- 2) 未知語を翻訳した変換テーブルの作成
- 3) 句対句の変換テーブルの作成
- 4) 一変数の文パターンのみで変換テーブルの作成

第7章 おわりに

これまで多くの翻訳手法が提案されてきた．代表的な翻訳手法として，統計翻訳やニューラル機械翻訳が挙げられる．しかし，これらの手法では，翻訳精度は十分ではない．また，これらの手法は翻訳手順の解析は困難であり，効率的な改善方の提案は難しい．

そこで，“相対的意味論に基づく変換主導型統計機械翻訳（以下，TDSMT）”を提案した．TDSMTでは翻訳可能な入力文は少なかった．しかし，出力結果は Moses と同等程度の翻訳精度を持っている．さらに，翻訳の手順や，変換テーブルの作成の手順を詳しく解析することが可能である．このため，多くの改善点を発見し，今後，カバー率と翻訳精度の両面で発展する可能性が示された．

今回の実験では，対訳文に存在しない語や，対訳文中の出現頻度が少なく学習が難しい語が入力文中に存在した場合，その入力文は翻訳不可能となっている．今後，このような学習が難しい語を未知語として，原言語のまま出力文に出力する手法を考案し，大幅なカバー率の向上を目指す．

第II部

相対的意味論を利用した 変換主導型統計機械翻訳における未知語 の出力

第1章 はじめに

現在，機械翻訳の分野では統計翻訳やニューラル機械翻訳など様々な手法が研究されている．しかし，どの手法も十分な翻訳精度にたどり着いていない．そこで第I部では“相対的意味論を利用した変換主導型統計機械翻訳 [1](以下，TDSMT)”を提案した．

TDSMT は対訳文と「 A (日本語単語) が B (英語単語) ならば C (日本語句) は D (英語句)」で表現する変換テーブルを利用する．学習として，対訳文から変換テーブルの自動作成を行う．その後，対訳文と変換テーブルを利用して，翻訳を行う．TDSMT は入力文の変換によって出力文を作成するため，文法性の高い翻訳が期待できる．第I部の実験では，変換テーブルが大量に作成され，その精度も高かった．しかし，翻訳可能な入力文の割合（以下，カバー率）は約 20% 程度と低かった．カバー率が低い原因として，対訳文中に存在しない語や，対訳文中の頻度が低く学習が難しい語（以下，未知語）が入力文中に一つでも存在すれば，翻訳不可能になることが挙げられる．また，TDSMT は翻訳文の出力に成功した場合 Moses と同等の翻訳精度がある．

そこで，第II部では，TDSMT において，出力文に未知語を出力する手法を提案する．未知語を出力するために未知語出力用変換テーブルを作成し，変換テーブルに追加する．未知語出力用変換テーブルは「 A (日本語単語) が B (英語単語) ならば C (日本語句) は C (日本語句)」で表現される．また未知語出力用変換テーブルは入力文と対訳文から作成する．

この論文の第2章では，従来手法である“相対的意味論に基づく変換主導型統計機械翻訳 [1]”を紹介する．第3章では，提案手法である“未知語を出力可能にした TDSMT[1]”の手順を説明する．第4章では，実験データやどのような評価を行ったかを示す．第5章では，第4章で示した評価の結果を示す．第6章では，提案手法の利点や誤りの解析を紹介する．第7章で，この研究全体をまとめる．

第2章 従来手法:相対的意味論を利用した変換主導型統計機械翻訳 (TDSMT)

2.1 TDSMT の概要

TDSMT の概要は第 I 部 3 節に示す。なお、この手法の入力文に対する翻訳可能な文の割合（以下、カバー率）は約 20% 程度であった。また、翻訳文の出力に成功した場合は Moses と同等程度の翻訳精度を示した。

2.2 TDSMT の問題

TDSMT の最も大きな問題点はカバー率が低いことである。TDSMT では変換テーブルを利用して、入力文を変換し、対訳文の日本語側と一致しなければ翻訳できない。さらに、同じ変換テーブルの「 B 」の部分が対訳文の英語側に存在する必要がある。このように、翻訳可能な条件を非常に厳格にすることで文法性を確保しているが、一方で、カバー率を低下させる結果となっている。

図 2.1 に具体的な例を示す。

図 2.1 において、変換テーブル 1 では対訳文の日本語側の「医者」と入力文の「患者」を変換できないため、適用できない。また、変換テーブル 2 では、対訳文の「医者」と入力文の「患者」を変換することが可能である。しかし、変換テーブルの「 A が B 」の部分が「医者 が surgeon」となっている。しかし、対訳文の日本語側に「医者」という単語は存在するが、対訳文の英語側では「surgeon」という単語は存在しない。このため、変換テーブル 2 は適用できない。

このように、変換テーブルを適用するには、対訳文の日本語側と入力文において、変換テーブルの A と C が存在し、対訳文の日本語側と英語側において、変換テーブルの A と B が存在する必要がある。このような条件で入力文を変換して、対訳文と一致させる

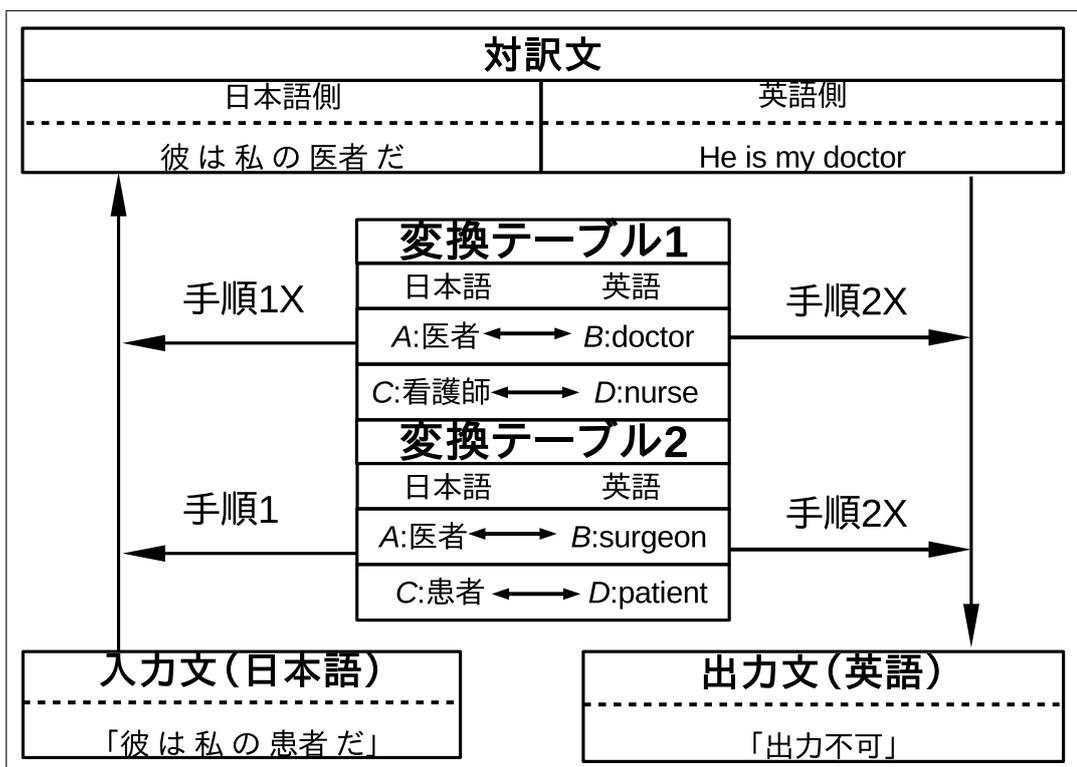


図 2.1 カバー率が低下する要因

には大量の変換テーブルを用意する必要がある。

また、対訳文に存在しない語や対訳文中の出現頻度が少なく学習が困難な語(以下、未知語)が入力文に一つでも存在した場合、この入力文は翻訳不可能となる。

第3章 提案手法:未知語を出力可能にしたTDSMT

3.1 概要

本論文では、2.2節の問題を解決するために、未知語を出力可能にしたTDSMTを提案する。未知語を出力するために、入力文の単語と対訳文を利用して未知語出力用変換テーブルを作成する。未知語出力用変換テーブルの形式は「AがBならばCはC」である。この未知語出力用変換テーブルを第I部3.4節で作成した変換テーブルに追加し、翻訳を行う。翻訳は第I部3.3節と同様の手順で行う。図3.1に未知語を出力可能にしたTDSMTのフローチャートを示す。

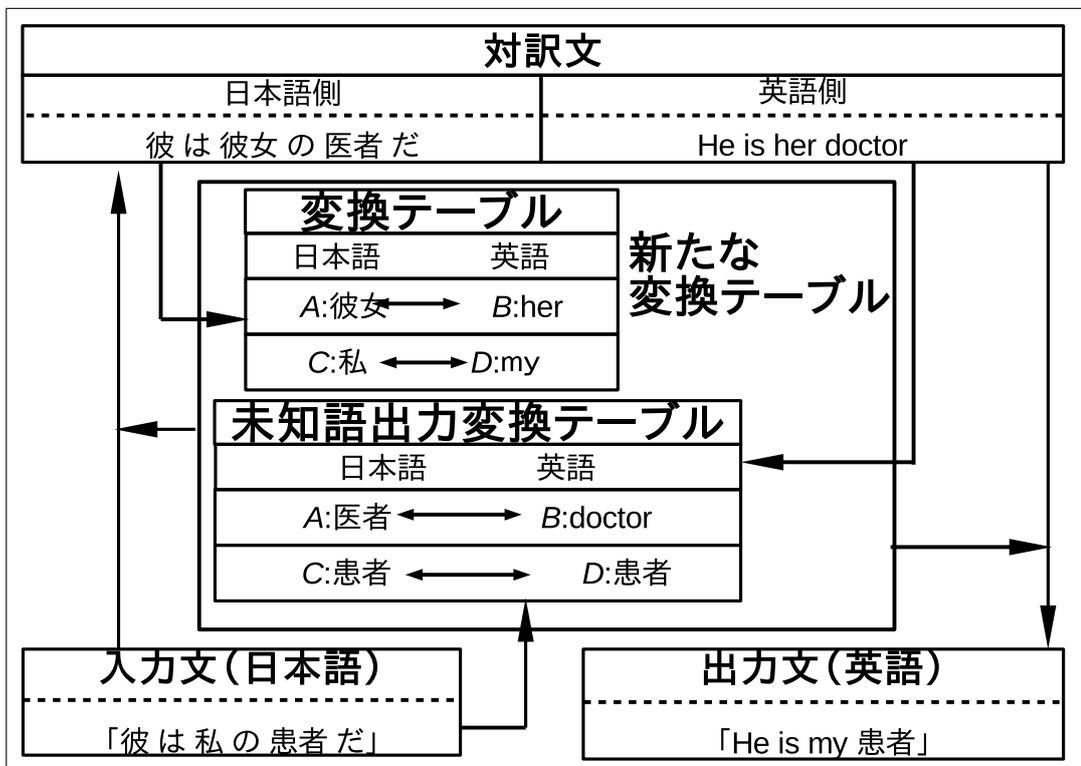


図 3.1 未知語を出力可能にした TDSMT のフローチャート

3.2 翻訳

3.2.1 翻訳の概要

提案手法における翻訳には対訳文と未知語出力用変換テーブルを追加した変換テーブルを利用する。未知語出力用変換テーブルを利用することで、従来手法と比較してカバー率の向上が期待できる。

3.2.2 翻訳の手順

提案手法における翻訳の手順は従来手法の手順と同様に行う。詳細は第 I 部 3.3 節に示す。

3.2.3 従来手法との変更点

提案手法では、未知語出力用変換テーブルを変換テーブルに追加した。未知語出力用変換テーブルと変換テーブルを効率良く扱うために、翻訳の設定を変更した。以下に変更点を示す。

- 経験的に決定した重みの利用
- 言語モデルの計算方式の変更

未知語出力用変換テーブルと変換テーブルでは、同じ適用確率の計算方法を利用することができない。そのため、未知語出力用変換テーブルには経験的に決定した重みを利用して、変換テーブルの適用確率との出現率を調整した。この重みは、経験的に決定しているため、必ずしも最適な値であるとは限らない。

従来手法では、出力文に出現する単語は目的言語単語のみで構成されていた。しかし、提案手法では目的言語単語に加え、未知語として出力した原言語単語が含まれる。このため、同じ計算方法を利用すると正しい言語モデルが計算できない可能性がある。このため、今回の実験では、出力文中に含まれる原言語単語は省いた状態で出力文の言語モデルの計算を行った。

3.2.4 提案手法の出力例

表 3.1 に提案手法による実際の翻訳例を示す．入力文の「国内事情がよくなった」に対して，対訳文「外が暗くなった：It has gotten dark outside」と2つの変換テーブルを利用して，「It has gotten dark 国内事情」が出力される．表 3.1 の例は出力候補文中で言語モデルと変換テーブル適用確率の合計値が最大の出力候補文である．

表 3.1 提案手法の実際の翻訳例

入力文	(国内事情)が(よく)なった	
参照文	The domestic conditions have improved	
対訳文(日)	(外)が(暗く)なった	
対訳文(英)	It has gotten (dark) (outside)	
変換 テーブル1	A:外	B:outside
	C:国内事情	D:国内事情
変換 テーブル2	A:暗く	B:dark
	C:よく	D:better
出力文	It has gotten (dark) (国内事情)	

3.3 学習

提案手法では3.2節の翻訳を実現するために，学習として，変換テーブルの自動作成と未知語出力用変換テーブルの自動作成を行う．変換テーブルの自動作成手法は従来手法と同様の手順で行う．また未知語出力用変換テーブルは対訳文より得た対訳単語と入力文から自動作成する．

3.3.1 変換テーブルの自動作成手法

提案手法では学習の一つとして，変換テーブルの自動作成を行う．詳しい手法は第I部3.4節と同様である．

3.3.2 未知語出力用変換テーブルの自動作成手法

従来手法の学習は変換テーブルの自動作成手法のみである．提案手法はさらに，未知語出力用変換テーブルの自動作成手法という学習を加える．未知語出力用変換テーブル

は入力文と、第I部 3.4.2.1 節で作成した対訳単語から作成する。図 3.2 に未知語出力用変換テーブルの自動作成手法のフローチャートを示す。

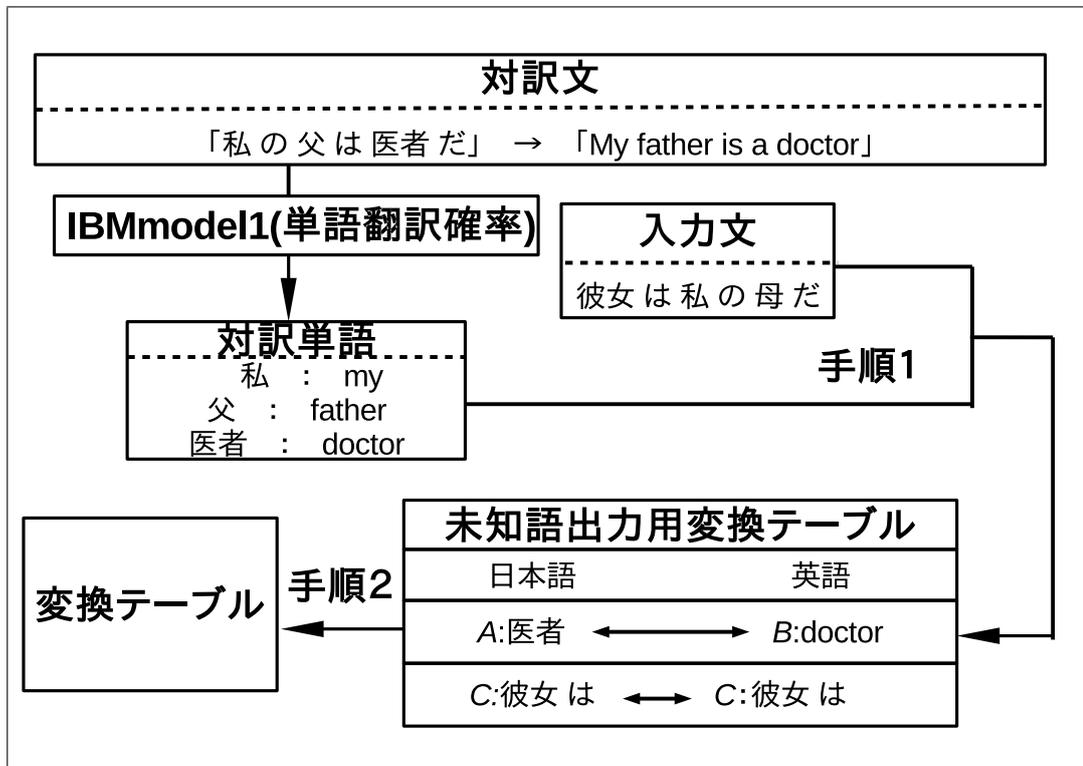


図 3.2 未知語出力用変換テーブルの自動作成手法のフローチャート

未知語出力用変換テーブルの自動作成手法は大きく分けて2つの手順から成る。まず、対訳単語と入力文を利用して、未知語出力用変換テーブルの作成を行う。そして、未知語出力用変換テーブルに適用確率の重みを加え、変換テーブルに追加する。日英翻訳による例を利用して詳細な手順を示す。

3.3.2.1 未知語出力用変換テーブルの作成

入力文と第I部 3.4.2.1 節で作成された対訳単語から未知語出力用変換テーブルを作成する。未知語出力用変換テーブルの形式は「AがBならばCはC」である。未知語出力用変換テーブルの作成例を表 3.2 に示す。なお、今回の実験では翻訳における計算時間の短縮のため、「CならばC」の部分にあたる単語は2単語以下に制限する。

表 3.2 未知語出力用変換テーブルの作成

対訳単語	
医者	doctor
入力文	
彼女 は 私 の 母 だ	
未知語出力用変換テーブル	
日本語側	英語側
A:医者	B:doctor
C:彼女	C:彼女
A:医者	B:doctor
C:彼女 は	C:彼女 は
etc...	

3.3.2.2 変換テーブルへの追加

未知語出力用変換テーブルを 3.3.1 節で作成した変換テーブルに追加する。この際、未知語出力用変換テーブルの適用確率に重みをつける。重みは確率値で 2^{-64} である。

3.3.3 未知語出力用変換テーブルの作成例

未知語出力用変換テーブルの例を表 3.3 に示す。

表 3.3 未知語出力用変換テーブルの例

未知語出力用	A:戦った	B:fought
変換テーブル	C:持ち込まれた	C:持ち込まれた
対訳文(日)	彼らは <u>勇敢に戦った</u>	
対訳文(英)	They <u>fought</u> bravely	
入力文	その事件は調停に <u>持ち込まれた</u>	

表 3.3 では入力文「その事件は調停に持ち込まれた」と対訳文「彼らは勇敢に戦った：They fought bravely」から作成された対訳単語「戦った：fought」から未知語出力用変換テーブルが作成された。

第4章 実験環境

4.1 実験データ

本実験では第I部4.1節と同様の実験データを利用する．第I部と同様にオープンデータによる実験を行い，クローズデータによる実験を行わない．実験データの前処理や後処理の条件も第Iと同様である．

4.2 評価方法

本研究では提案手法を2つの基準で評価する．

1) 翻訳に利用する変換テーブルの数

従来手法と提案手法で翻訳に利用する変換テーブルの数を比較調査する．なお，変換テーブルの数は入力文一文あたりの平均数で示す．

2) カバー率の調査

TDSMTで翻訳実験を行う．1,000文を翻訳し，カバー率を調査する．なお，この調査においては，翻訳の正誤は問わない．

本研究では主に，提案手法は従来手法と比較したカバー率に焦点を当てて評価を行う．これは，問題点である従来手法の低いカバー率を向上させる手段として，提案手法は適切であるかを評価するためである．

また，提案手法では変換テーブルを追加した．そのため，計算空間と計算時間は従来手法と比較して増加すると考えられる．カバー率とこの計算空間と計算時間の関係性を評価するため，一つ目の基準を設定した．

第5章 実験結果

5.1 翻訳に利用する変換テーブルの数

従来手法と提案手法で翻訳に利用する変換テーブルの数を表 5.1 に示す．なお，変換テーブルの数は入力文一文あたりの平均数で示す．

表 5.1 翻訳に利用する変換テーブルの平均数

従来手法	13,503(個)
提案手法	173,004(個)

表 5.1 より提案手法では従来手法と比較し，翻訳に利用する変換テーブルの数は大幅に増加している．このため，提案手法では翻訳に必要な計算空間と計算時間も大幅に増加していると考えられる．

変換テーブルは文パターンと対訳文が照合した場合のみ作成され，更に枝刈りを行なっているため，数は少ない．しかし，未知語出力用変換テーブルは入力文の単語と対訳単語を全展開し作成するため，数は多くなる．未知語を出力するために用意する変換テーブルであるため，枝刈りは難しい．

5.2 カバー率の調査

従来手法と提案手法において入力文 1,000 文での翻訳実験を行った．各手法のカバー率を表 5.2 に示す．

表 5.2 各手法のカバー率

従来手法	19.6 % (196 文/1,000 文)
提案手法	60.6 % (606 文/1,000 文)

表 5.2 より，従来手法ではカバー率は低かった．しかし，未知語出力用変換テーブルを利用することでカバー率は大きく上昇した．しかし，Moses やニューラル機械翻訳のカバー率にはまだまだ及ばない．

5.3 実験結果のまとめ

今回の実験において，カバー率は約 60% まで上昇した．このため，未知語を出力する手法は TDSMT においてカバー率を向上させるのに有効であることがわかる．しかし，Moses やニューラル機械翻訳ではカバー率は約 100% であるためまだまだ及ばない．

今回の実験では，未知語出力用変換テーブルの作成において制限をかけている．3.3.2 節に示すように，未知語出力用変換テーブルの「 C が C 」の部分は 2 単語までである．この制限を取り除くことにより，カバー率は約 100% まで上昇する．しかし，TDSMT の利点である，文法性の維持に影響を及ぼす可能性が考えられる．また，入力文をそのまま出力文として出力する可能性がある．このため「 C が C 」の部分の単語を増やす方法はおこなっていない．

また，カバー率の向上を実現するために，計算空間と計算時間の増加を伴った．これは 5.1 節により示された．

第6章 考察

6.1 提案手法と Moses[9] との翻訳精度の比較

この節では，提案手法の翻訳精度について考察する．一般的な統計翻訳の手法である mose と比較調査を行うことによって，TDSMT の翻訳精度がどの程度有効であるかを調査した．

表 5.2 に示すように，提案手法と moses ではカバー率に差がある．そのため，提案手法で出力可能な入力文を対象にして翻訳精度の超差を行う．

6.1.1 自動評価

6.1.1.1 自動評価法

自動評価法は第 I 部の実験と同じ手法を利用する．詳細は第 I 部 2.3.1 節に示す．

6.1.1.2 自動評価結果

提案手法で出力可能であった入力文 606 文で翻訳実験を行った．提案手法と moses の出力文に対して，自動評価を行った．結果を 6.1 に示す．

表 6.1 提案手法と Moses の自動評価結果

手法	BLEU[10]	METEOR[11]	RIBES[12]
moses	0.191	0.477	0.739
提案手法	0.105	0.354	0.666

6.1 により，すべての評価値で提案手法は moses より翻訳精度が低かった．moses と比較して，提案手法は未知語として，出力文に原言語単語を出力する割合が高い．自動評価では参照文と一致度合いのみで評価を行うため，未知語が多い提案手法の方が評価値が下がったと考えている．

6.1.2 人手評価

6.1.2.1 人手評価法

人手評価法は第 I 部の実験と同じ手法を利用する．詳細は第 I 部 2.3.2 節に示す．人手評価基準を第 I 部 6.2 節に示す．

6.1.2.2 人手評価結果

提案手法で出力可能であった入力文 606 文からランダムに抽出した 100 文に対して，翻訳実験を行った．提案手法と Moses の出力文を対象に対比較評価を行った．結果を表 6.2 に示す．なお，未知語が出現した場合はその未知語が正しく翻訳されると想定して評価を行った．

表 6.2 提案手法と Moses との対比較評価結果

提案手法	Moses	差なし	同一出力
25	14	52	9

表 6.2 より，提案手法の場合では Moses に勝る翻訳精度を示した．未知語を正しく翻訳することで，TDSMT は Moses よりも精度の高い翻訳結果を得ることが期待できる．

6.1.2.3 人手評価例：提案手法 の例

提案手法 の評価例を表 6.3，表 6.4 及び表 6.5 に示す．

表 6.3 提案手法 の人手評価例 1

入力文	彼は国立劇場に出演した
参照文	He appeared on the stage of the National Theater
提案手法	He appeared on 国立劇場
Moses	He played a part in the National

表 6.3 では，未知語が正しく翻訳された場合，TDSMT の出力文は「He appeared on National Theater」となり，入力文の意味がおおよそ読み取れる．Moses の出力文について，動詞の「played a part」は「役割を果たした」や「芝居をした」などがあり，入力

表 6.4 提案手法 の人手評価例 2

入力文	スナップ 写真 を 撮った
参照文	He snapped a picture
提案手法	I took a スナップ 写真
Moses	I took a photograph was away

文の意味を読み取るには意識が必要である．また，入力文の「劇場」の意味が不足している．そのため，提案手法 と評価した．

表6.4では，TDSMTの出力文では「スナップ 写真」が「snapshot」「candid photograph」などと翻訳された場合，入力文の意味を正しく読み取ることができる．一方，Mosesの出力文では「was away」が不必要に訳出されている．このため，Mosesの出力文は入力文の意味を読み取り難いため，評価を提案手法 とした．

表 6.5 提案手法 の人手評価例 3

入力文	その 病気 は 腎臓 を 冒す
参照文	The disease affects the kidneys
提案手法	The disease affects 腎臓
Moses	That affects the kidneys don't be sick

表 6.5 では，TDSMTの出力文では「腎臓」が「the kidneys」などと翻訳された場合，入力文の意味を正しく読み取ることができる．一方，Mosesの出力文では主語が「That」と不明確に訳出されている．また，文の終わりに「don't be sick」という単語列が生成され，入力文の意味を読み取るのは難しい．このため，この文の評価を提案手法 とした．

6.1.2.4 人手評価例：Moses の例

moses の評価例を表 6.6，表 6.7 及び表 6.8 に示す．

表 6.6 Moses の人手評価例 1

入力文	いっさいを 白紙 に 返そう
参照文	Let's forget all about it
提案手法	He 返そう いっさい in in a blank exam paper
Moses	Let's forget everything about his all

表 6.6 では、Moses の出力文は入力文の意味を読み取ることができる。入力文では目的となる事柄「白紙」にする内容が明確に記されていないため、「his」が入ったとしても入力文の意味と変わらないと考えられる。提案手法の出力文では、「返そう っさい」の部分が正しく翻訳されたとしても、その後続く「in in」や文全体の「Let's ~」のような文構造がないため、入力文の意味を読み取るのは難しい。このため、評価を Moses とした。

表 6.7 Moses の人手評価例 2

入力文	授業 に 間に合った
参照文	I was in time for the class
提案手法	I just class
Moses	I was just in time for the class

表 6.7 では、提案手法の出力文では動詞がかけているため、入力文の意味を読み取るのは難しい。Moses の出力文では出力文に「just」が含まれている。参照文と比較して、「なんとか間に合った」という表現になるが、入力文の意味はる程度読み取れる。そのため、提案手法の出力文より Moses の出力文の方が入力文の意味を読み取れるため、評価を Moses とした。

表 6.8 Moses の人手評価例 3

入力文	この論文にはタイプミスがたくさんある
参照文	There are a lot of typos in this paper
提案手法	This paper lot タイプミス
Moses	This paper has a lot of errors in type

表 6.8 では、提案手法の出力文では動詞がない。また、「タイプミス」が「タイプミスをした：made typo」のように動詞に訳出されたとしても、主語が「This paper」では入力文の意味を読み取るには難しい。一方で、Moses の出力文では「この論文にはミスがたくさんある」という意味は読み取ることができる。このミスがタイプミスであることはその後の「in type¹という単語列で意識が可能である。そのため、提案手法の出力文と比較して、Moses の出力文の方が入力文の意味が読み取りやすいとして、この文の評価を Moses とした。

¹「in type」の辞書的な意味は「活字に組まれて」などがあたる

6.1.2.5 人手評価例：差なしの例

差なしの評価例を表 6.9，表 6.10 及び表 6.11 に示す．

表 6.9 差なしの人手評価例 1

入力文	雨が止みそうだ
参照文	The rain promises to hold up
提案手法	The rain 止みそう
Moses	The rain is likely to 止み

表 6.9 の提案手法の出力文では「止みそう」が「is stopping」と正しく翻訳された場合入力文の意味が読み取れる．また，Moses の出力文では「止み」が「stop」と正しく翻訳された場合，入力文の意味が読み取れる．どちらの出力文も未知語が正しく翻訳された場合入力文の意味が正しく読み取れるため，評価を差なしとした．

表 6.10 差なしの人手評価例 2

入力文	十分な駐車スペースがある
参照文	There is ample parking
提案手法	There is enough 駐車スペース
Moses	There is plenty of parking space

表 6.10 では，提案手法の出力文において「駐車スペース」が「parking」や「berth」等と翻訳された場合，入力文の意味を正しく読み取ることができる．また，Moses の出力文においても「plenty of」で「たくさんの」という意味になり，「十分な」とほぼ同義の意味が読み取れる．このため，この文の評価を差なしとした．

表 6.11 差なしの人手評価例 3

入力文	名選手だってときにはエラーはする
参照文	Even the best player sometimes makes an error
提案手法	Sometimes だって players 名エラーは
Moses	The error, sometimes even the runners in the name

表 6.11 では，どちらの出力文も入力文の意味を読み取ることは難しい．また，どちらの出力文も英語の文法に則っていないと判断した．このため，この文の評価を差なしとした．

6.1.2.6 同一出力の例

同一出力の評価例を表 6.12 に示す。

表 6.12 同一出力の人手評価例

入力文	彼の生涯は多事であった
参照文	He has had an eventful life
提案手法	His life was 多事
Moses	His life was 多事

表 6.12 のように、提案手法と Moses の出力が一致した場合、同一出力と評価した。

6.1.3 提案手法の翻訳精度に関するまとめ

今回の実験では、自動評価では提案手法は Moses より翻訳精度が低かった。人手評価では提案手法の方が Moses より翻訳精度が高かった。自動評価は参照文の一致度合いを利用して、翻訳精度を計算する。そのため、評価値が低下した原因は翻訳精度が低下したためではない可能性がある。出力文中に目的言語以外の単語を出力したため評価値が低下したとも考えられる。自然言語では表現の曖昧性や多様性が存在するため、一つの参照文に基づいた自動評価結果は信頼性が低い。一方、人手評価は評価者が表現の曖昧性や多様性を考慮するため、この問題が少ない。

このような点も考慮した。第 I 部で、未知語を出力しない TDSMT は moses と同等の翻訳精度であった。しかし、未知語を出力する TDSMT では翻訳精度は向上し、moses を上回る翻訳精度となったとかがえられる。

6.2 誤り解析

第 I 部 6.2 節で、TDSMT の利点として翻訳の誤り原因を解析することが可能であると示した。この利点は提案手法においても同じである。実際の翻訳例を利用して、提案手法の誤り解析を行った。なお、6.1.2 節の対比較評価において、Moses と評価された入力文において解析を行った。

Moses と評価された入力文 14 文中の誤りのおおまかな内訳を表 6.13 に示す。

表 6.13 誤りのおおまかな内訳

変換テーブルの誤り	7/14
利用する対訳文の誤り	4/14
確率値による出力文の選択の誤り	3/14

第 I 部 6.3 節の表 6.14 と比較くすると，誤りの内訳は TDSMT とほぼ同じであった．TDSMT と同様に，変換テーブルの誤りが誤り原因として多かった．6.2.1 節，6.2.2 節，6.2.3 節に誤りの例を示す．

6.2.1 変換テーブルの誤り

表 6.14 TDSMT の誤り例（変換テーブルの誤り）

入力文	(この国)の(所得税法)は(全面的)に(改正)された	
参照文	The nation's income-tax laws have undergone sweeping revision	
対訳文(日)	(彼)の(抗議)は(完全)に(無視)された	
対訳文(英)	(His) (protest) was (completely) (ignored)	
変換 テーブル 1	A:彼	B:His
	C:この国	D:This country
変換 テーブル 2	A:抗議	B:protest
	C:所得税法	D:所得税法
変換 テーブル 3	A:完全	B:completely
	C:全面的	D:全面的
変換 テーブル 4	A:無視	B:ignored
	C:改正	D:改正
出力文	(This country) (所得税法) was (全面的) (改正)	

表 6.14 の翻訳例では，入力文と同じ構文を持つ対訳文を利用して翻訳が行われている．しかし，変換テーブル 1 のおいて「B」と「D」の部分に注目すると「B」は所有格であるが「D」は所有格ではない．そのため，対訳文の構文を維持できていない．なお，Moses の対訳文は「They made a sweeping revision of the income tax law of this country」となっている．Moses の出力文では英単語「of」を利用して「この国の」という表現を正しく翻訳している．このため，評価は Moses となった．この変換テーブル 1 の作成に利用した情報を表 6.15 に示す．

表 6.15 表 6.14 の変換テーブル 1 の詳細

対訳文 1 (日)	(彼) の (動機) は (複雑) で (あった)	
対訳文 1 (英)	(His) (motives) (were) (complex)	
対訳単語 (X00)	あった	were
対訳単語 (X01)	動機	motives
対訳単語 (X02)	彼	His
対訳単語 (X03)	複雑	complex
文パターン (日)	X02 の X01 は X03 で X00	
文パターン (英)	X02 X01 X00 X03	
対訳文 2 (日)	(この国) の (経済) は (健全) で (ある)	
対訳文 2 (英)	(This country) (has) { a healthy } { economy}	
変換テーブル	A:彼	B:his
	C:この国	D:This country

表 6.15 の対訳文 1 では「彼の動機: His motives」が主語となっている。しかし、対訳文 2 では「この国: This country」が主語となっている。このような構文の違いが「B」と「D」で品詞が一致しない変換テーブルを作成する原因となっている。

なお、「この国の: This country's」を含む対訳文は存在したが、「CはD」の部分が「この国は This country's」という変換テーブルはなかった。変換テーブルの自動作成の際、様々な枝刈りを行なっている。この枝刈りの工程で必要な変換テーブルが削除されてしまった。今後、枝刈りの方法や閾値を見直す必要がある。

6.2.2 利用する対訳文の誤り

表 6.16 の翻訳例では、たくさんの誤りの原因が考えられる。一つは翻訳に利用した対訳文が不適切である。入力文と対訳文の日本語側は構文に近いが、対訳文の英語側と同じ構文で入力文の意味を表すことは困難である。入力文と対訳文で「出た」が共通しているが、対訳文の「appear」を利用して入力文を表現するには構文を変える必要がある。主語の部分に「勉強に張り」という表現を入れる必要がある。このため、どの対訳文を利用して翻訳を行うかの選択方法を考案する必要がある。

翻訳確率(言語モデル+変換テーブルの適用確率)の順位で4位までを解析しても、表 6.16 に示す対訳文以外の対訳文を利用する出力候補文は存在しなかった。

なお、変換テーブルの問題も同時に出現している。「昨夜」は「Last night」が正しい

表 6.16 TDSMT の誤り例 (対訳文の選択の誤り)

入力文	(最近) 彼は (勉強) (に 張り) (が) 出た	
参照文	He has recently felt a strong inducement to his study	
対訳文 (日)	(昨夜) 彼は (テレビ) (番組) (に) 出た	
対訳文 (英)	Last (night) he appeared (on) a (TV) (program)	
変換 テーブル 1	A:昨夜 C:最近	B:night D:最近
変換 テーブル 2	A:にテレビ C:勉強	B:TV D:working
変換 テーブル 3	A:番組 C:に 張り	B:program D:に 張り
変換 テーブル 4	A:に C:が	B:on D:in
出力文	Last (最近) he appeared (in) a (working) (に 張り)	

が、「night」のみと対応した変換テーブルが利用している。これは現在、「AがBならばCはD」の「AがB」の部分は単語対応に限っている。このため、変換テーブル1のような変換テーブルが作成されてしまう。「AのB」の部分も句にした変換テーブルを利用することで翻訳精度が向上する可能性が考えられる。

6.2.3 確率値による出力文の選択の誤り

表 6.17 と表 6.18 に同一の入力文における第 1 位の翻訳候補文と第 3 位の翻訳候補文を示す。なお、確率値は対数をとっているため、0 に近い値が大きい確率を示す。一方、マイナスの絶対値が大きくなるほど低い確率を示す。

入力文「今年は穀物がどっさりとれた」に対して、表 6.17 の翻訳候補文が出力された。しかし、表 6.17 では未知語が正しく翻訳されたとしても完全な翻訳になるのは難しい。一方、表 6.18 の第三位の翻訳候補文では「どっさり」が「lot(「a lot of」の一部)」と、「穀物」が「grain」と未知語処理された場合、正しい翻訳文となる。表 6.18 の翻訳候補文は文法構造をくずしておらず、入力文の意味を読み取りやすい翻訳となっている。このため、表 6.18 を出力文とすることが好ましい。変換テーブルの適用確率は表 6.18 の翻訳候補文が優っている。しかし、言語モデルの確率は表 6.17 の翻訳候補文に劣っている。2 つの確率の合計値は表 6.17 の方が優っているため、表 6.17 の翻訳候補文が出力文に選択された。

表 6.17 TDSMT の誤り例 (確率値が第一位の翻訳候補文)

入力文	(今年) は (穀物) が (どっさり とれた)	
参照文	There was a heavy yield of grain this year	
対訳文 (日)	(それ) は (話) が (違う)	
対訳文 (英)	That's a different story	
変換 テーブル 1	A:それ	B:That's
	C:今年	D:This year
変換 テーブル 2	A:話	B:story
	C:穀物	D:穀物
変換 テーブル 3	A:違う	B:different
	C:どっさり とれた	D:どっさり とれた
出力文	(This year) a (どっさり とれた) (穀物)	
言語モデル	high order joint probability	-326.956671
変換テーブルの適用確率	-222.625002	

表 6.18 TDSMT の誤り例 (確率値が第三位の翻訳候補文)

入力文	今年 は (穀物) が (どっさり) とれた	
参照文	There was a heavy yield of grain this year	
対訳文 (日)	今年 は (米) が (たくさん) とれた	
対訳文 (英)	This year a (lot) of (rice) was produced	
変換 テーブル 1	A:米	B:rice
	C:穀物	D:穀物
変換 テーブル 2	A:たくさん	B:lot
	C:どっさり	D:どっさり
出力文	This year a (どっさり) of (穀物) was produced	
言語モデル	high order joint probability	-373.710876
変換テーブルの適用確率	-208.304391	

現在，言語モデルは出力文に出現した未知語を除いて計算される．つまり，表 6.17 の出力候補文では「This year a」，表 6.18 の出力候補文では「This year a of was produced」で計算される．そのため，たくさんの未知語を含む出力文は言語モデルの値が高くなる傾向にある．このため，未知語を多く含む文は最終的な出力文に選ばれやすい傾向にある．また，言語モデルは長い文では，その値が小さくなる傾向にある．そのため，文長に依存しない言語モデルを考案する必要がある．

また，今回の例では，変換テーブルの確率は正しい翻訳に近い表 6.18 の翻訳候補文の方が高かった．このため，2つの確率値をうまく利用して，出力文を決定する方法を考案する必要がある．

6.3 未知語の翻訳精度に及ぼす影響

提案手法では，未知語を出力することで，カバー率の向上を目指した．しかし，翻訳精度にも影響があると考えられる．Moses では学習が難しい単語も対訳文中の少ない情報を利用して，翻訳する傾向にある．一方，提案手法では，対訳文中の頻度の少ない句（出現頻度が一回）はほとんど未知語として出力されている．こうすることにより，対訳文中の文法が出力文に保存されたため，翻訳精度が向上したと考えている．

今回の実験では，この未知語の翻訳精度に及ぼす影響がどの程度あるか観測することができなかった．そのため，文法性に着目した評価を行う必要がある．また，統計翻訳やニューラル機械翻訳に関しても，学習の難しい語は未知語として出力することで翻訳精度が向上する可能性も示された．

6.4 翻訳に利用する確率値について

本研究では，変換テーブルの自動作成や最終的な出力文の決定に確率値を利用している．変換テーブルの自動作成には，IBM model1(単語翻訳確率)を利用して，対訳単語を作成する．そして，IBM model1(単語翻訳確率)から変換テーブルの適用確率を計算する．また，出力文の決定には High order Joint Probability(言語モデル)と IBM model1(単語翻訳確率)から計算した変換テーブルの適用確率を利用する．

今回の実験では，TDSMT の手法のおおまかな理論，考え方を示すため，確率値の詳しい計算には触れていない．しかし，変換テーブルの適用確率の計算方法や，利用する言語モデルの選択など今後詳しく調査する必要がある．

変換テーブルの適用確率は単純に「*A*が*B*」の部分の単語翻訳確率と「*C*が*D*」の部分の単語翻訳確率を掛けあわしている。「*A*が*B*」と「*C*が*D*」の関連性を示す指標は利用していない。そのため、変換テーブルの適用確率が適切でない問題がある。今後、「*A*が*B*」と「*C*が*D*」の関連性を示す確率値を考案したい。

出力文の決定の際、翻訳候補文の言語モデルの確率値と変換テーブルの適用確率値を単純に掛けあわせた確率をもとに順位をつけている。しかし、言語モデルの確率値は変換テーブルの適用確率値と比較して、非常に小さい。この場合、2つの確率値を出力文の決定に利用しているが、言語モデルの確率値によってほぼ決定してしまう。このように、出力文の決定に2つの確率値をうまく利用する手法を考案する必要がある。

また、単純に順位づけするだけでは、高順位のものがほぼ同じ翻訳候補文があつまり、考察がうまくできない。出力文の決定方法にはまだまだ改善点があると考えられる。

第7章 おわりに

第I部では“相対的意味論に基づく変換主導型統計機械翻訳(以下, TDSMT)”を提案した。TDSMTでは, Mosesと同等程度の翻訳精度を示した。しかし, 翻訳可能な入力文の割合(以下, カバー率)は20%程度と低かった。TDSMTは翻訳の手順を詳しく追跡することが可能である。解析の結果, 対訳文中に存在しない単語や対訳文中に出現頻度の少ない語(以下, 未知語)が入力文中に存在した場合, その入力文は翻訳不可能になってしまう。このことが, TDSMTでカバー率の低い最も大きな原因だと考えられる。

そこで本実験では, TDSMTにおいて未知語を出力する手法を提案した。提案手法では, 従来の変換テーブルに加え, 入力文と対訳文から作成した未知語出力用変換テーブルを作成する。未知語出力用変換テーブルを加えた変換テーブルを利用し, 翻訳を行うことで, 出力文に未知語を出力することが可能となった。提案手法によりカバー率は60%程度まで向上した。しかし, 翻訳に利用される変換テーブルの数は10倍程度に増加した。そのため, 従来手法と比較して, 翻訳に必要な計算空間と計算時間は増加した。

さらに, 提案手法と一般的な統計翻訳の手法であるmosesとの翻訳精度の比較を行った結果, 未知語が正しく翻訳できた場合, 提案手法は翻訳精度が高かった。今後, 未知語を正しく翻訳する手法や新たな変換テーブルの作成手法を考案することで高い翻訳精度を持つ翻訳手法になる可能性が示された。

まとめ

第I部では、“相対的意味論に基づく変換主導型統計機械翻訳（以下，TDSMT）”を提案した．現在主流となっている翻訳手法では，翻訳精度が十分ではなく，また翻訳手順の解析が難しい．この問題を解決するために新しい翻訳手法として，TDSMT を提案した．TDSMT の学習である変換テーブルの自動作成手法により，約 570 万個の変換テーブルの作成に成功した．また，変換テーブルの精度は高かった．しかし，TDSMT の翻訳実験により，入力文に対する翻訳可能な入力文の割合（以下，カバー率）は 20% 程度と低かった．TDSMT の翻訳精度を Moses と比較した．この結果，TDSMT は翻訳文の出力に成功した場合，Moses と同等程度の翻訳精度をもつことがわかった．また，TDSMT は開発段階であり，手法の改善によりさらに，翻訳精度の高い翻訳手法になることが期待できる．

TDSMT は翻訳の手順を詳しく解析することが可能である．翻訳の手順を詳しく解析することによって，誤りの原因をすることが出来る．誤りの原因をふまえた改善法を提案することにより，効率の良い発展が期待できる．この発展性が TDSMT の利点である．

第II部では，TDSMT において，未知語を出力する手法を提案した．低いカバー率の原因は対訳文中に存在しない語や，対訳文中に出現する頻度が低く，学習が困難である語（未知語）が入力文中に出現した場合，翻訳不可能になってしまうことである．この問題を解決するために，未知語を出力文中に出力する手法を提案した．未知語出力用変換テーブルの自動作成手法を利用することにより，未知語出力用変換テーブルは大量に作成された．この結果，翻訳に利用する変換テーブルの数は 10 倍程度まで増加した．このため，提案手法は従来の TDSMT よりも多くの計算空間と計算時間を要する．しかし，未知語を出力することによって，カバー率は 60% 程度まで増加した．

また，Moses と翻訳精度の比較を行った．人手評価では，提案手法は Moses よりも翻訳精度が高い．これは，学習の難しい語を少ない情報から翻訳するよりも，未知語として出力することで，文法を維持した翻訳文を生成することができると考えられる．

本実験において，新しい手法である“相対的意味論に基づく変換主導型統計機械翻訳”を提案し，未知語を出力可能にして，カバー率と翻訳精度を向上させた．しかし，カバー

率と翻訳精度はまだまだ低い。このため、出力文の解析と改善法の提案を行う必要がある。

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徳久雅人講師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] 安場裕人, 村上仁一. “変換主導型統計機械翻訳の提案”, 自然言語処理学会第24回年次大会, ポスター(3), P7-9, pp.789-792, Mar.2018.
- [2] SRILM(The SRI Language Modeling Toolkit) :<http://www.speech.sri.com/projects/srilm/.srilm1.5.12.tar.gz>
- [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] 川原宰: “統計翻訳における自動作成した対訳単語辞書を用いた未知語処理”, 平成29年度鳥取大学修士論文, 2018.
- [5] ナシーム・ニコラス・タレブ: “ブラック・スワン [上] 不確実性とリスクの本質”, Jun.2009.
- [6] 古瀬蔵, 隅田英一郎, 飯田仁. “経験的知識を活用する変換主導型機械翻訳”, 情報処理学会論文誌, Vol.35 No.3 pp.414-425. Mar.1994.
- [7] 江原暉将, 小倉健太郎, 篠崎直子, 森本逞, 樽松明: 電話またはキーボードを介した対話に基づくデータベース ADDの構築, 情報処理学会論文誌, Vol. 28, No.4, pp. 448-456, 1992.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [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] 松本大輝 : “翻訳における言語モデルとしての High order Joint Probability の分野依存性の分析” , 平成 30 年度鳥取大学卒業論文 , 2019.

研究業績

1. 安場裕人, 村上仁一：“人手対訳句を利用したパターンベース統計翻訳”，言語処理学会第 23 回年次大会発表論文集，pp.737-740，Mar.2017.
2. 安場裕人, 村上仁一：“パターンに基づく統計翻訳における人手対訳句の利用”，言語理解とコミュニケーション研究会，電子情報通信学会技術研究報告，vol.117,no.82, NLC2017-1,pp.1-6，Jun.2017.
3. 安場裕人, 村上仁一：“変換主導型統計機械翻訳の提案”，言語処理学会第 24 回年次大会発表論文集，pp.789-792，Mar.2018.
4. 安場裕人, 村上仁一：“相対的意味論に基づく変換主導型統計機械翻訳～未知語の出力～”，言語処理学会第 25 回年次大会発表論文集 (掲載予定)，Mar.2019.