

概要

近年，機械翻訳の分野で原言語から目的言語に翻訳する統計翻訳が注目されている．統計翻訳は対訳文を用いてフレーズごとの翻訳確率である翻訳モデルや目的言語らしさである言語モデルを学習する．そして，文が入力された時に翻訳モデルと言語モデルの組合せから確率が最大となる文を出力文として得る．統計翻訳において，対訳文数が多ければ多いほど出力文の翻訳精度は高くなることが知られている．しかし，利用できる対訳文数には限りがある．

この問題に対し，対訳コーパスに対訳句コーパスを追加する手法の研究が行われ，自動評価結果が向上したとの報告がある．この研究はセルビア語英語間，スペイン語英語間 [1] そしてブルトン語フランス語 [2] の各翻訳において行われている．

そこで，本研究では同様の手法を日本語英語間の翻訳において行う．辞書のデータから抽出した対訳句コーパスを用いる．日英対訳コーパスを用いて言語モデルの学習する．日英対訳コーパスに対訳句コーパスを加えて翻訳モデルを学習し，統計翻訳を行う方法を提案手法とし，翻訳精度の調査を行う．

日英対訳コーパスは単文コーパスと重文複文コーパス [3] を用いる．対訳句コーパスとして鳥バンク [4] と英辞郎 [5] を用いる．

その結果，自動評価結果と人手評価結果のどちらにおいても提案手法はベースラインよりも翻訳精度が向上した．日本語英語間の統計翻訳において，対訳句コーパスを用いた提案手法は有効であると確認できた．

目次

第1章	はじめに	1
第2章	統計翻訳システム	3
2.1	基本概念	3
2.2	翻訳モデル	4
2.2.1	IBMモデル	4
2.2.2	GIZA++	5
2.2.3	フレーズテーブルの作成方法	5
2.3	言語モデル	8
2.3.1	N -gramモデル	8
2.4	デコーダ	9
2.5	評価方法	9
2.5.1	自動評価法	9
2.5.2	BLEU	10
2.5.3	NIST	10
2.5.4	RIBES	10
2.5.5	METEOR	11
2.5.6	人手評価法	11
第3章	提案手法	13
第4章	実験環境	15
4.1	日英対訳コーパス	15
4.2	翻訳モデルの学習	15
4.3	言語モデルの学習	16
4.4	デコーダのパラメータ	16
4.5	対訳句コーパス	16

4.5.1	鳥バンク	17
4.5.2	英辞郎	18
第5章	翻訳実験	20
5.1	ベースラインと提案手法	21
5.2	評価実験	21
5.3	実験結果	22
5.3.1	自動評価結果	22
5.3.2	人手評価結果	23
5.3.3	出力例	24
第6章	考察	29
6.1	対訳句コーパスごとの翻訳品質の比較	29
6.1.1	出力例	30
6.2	鳥バンクと英辞郎を組み合わせた翻訳実験	33
6.3	鳥バンクと英辞郎の差	34
6.3.1	正しいフレーズテーブルの選択	34
6.3.1.1	例1	34
6.3.1.2	例2	35
6.3.1.3	例3	36
6.3.2	未知語を含む文の割合調査	38
6.3.3	未知語の減少	38
6.3.3.1	例1	39
6.3.3.2	例2	39
6.3.3.3	例3	40
6.3.3.4	例4	40
6.4	自動評価結果と人手評価結果	41
6.4.1	例1	41
6.4.2	例2	42
6.4.3	例3	42
6.4.4	例4	43
6.4.5	例5	44
6.4.6	例6	44

6.4.7	例7	44
6.4.8	自動評価結果と人手評価結果の相違のまとめ	46
6.5	英辞郎の日本語動詞活用	47
6.5.1	実験方法	47
6.5.2	実験結果	48
6.5.2.1	自動評価結果	49
6.5.2.2	人手評価結果	50
6.5.3	英辞郎の日本語動詞活用のまとめ	50
6.6	今後の課題	51
6.6.1	鳥バンクと英辞郎の解析	51
6.6.2	新しい自動評価法の検討	51
第7章 おわりに		52

目 次

2.1	日英統計翻訳の流れ	3
3.1	日英統計翻訳の場合の提案手法の流れ	13
4.1	単語数による対訳句対数の分布 (鳥バンク 英語句)	17
4.2	単語数による対訳句対数の分布 (鳥バンク 日本語句)	18
4.3	単語数による対訳句対数の分布 (英辞郎 英語句)	19
4.4	単語数による対訳句対数の分布 (英辞郎 日本語句)	19

表 目 次

2.1	フレーズテーブルの例	4
2.2	日英方向の単語対応	5
2.3	英日方向の単語対応	5
2.4	積集合 intersection	6
2.5	和集合 union	6
2.6	ヒューリスティック grow-diag-final	6
2.7	作成されたフレーズ対の例	7
2.8	作成されたフレーズテーブルの例	7
4.1	前処理後の単文コーパスの例	15
4.2	前処理後の重文複文コーパスの例	15
4.3	鳥バンク対訳句対の例	17
4.4	クリーニング後の英辞郎の対訳句対の例	18
5.1	翻訳実験の種類	20
5.2	使用した実験データ	20
5.3	対訳句コーパスごとの対訳句対数	21
5.4	言語モデルと翻訳モデルの作成に使用するコーパス	21
5.5	自動評価結果	22
5.6	人手評価結果	23
5.7	提案手法 (鳥バンク) 日英統計翻訳結果 (単文)	24
5.8	提案手法 (鳥バンク) 英日統計翻訳結果 (単文)	25
5.9	提案手法 (鳥バンク) 日英統計翻訳結果 (重文複文)	25
5.10	提案手法 (鳥バンク) 英日統計翻訳結果 (重文複文)	26
5.11	提案手法 (英辞郎) 日英統計翻訳結果 (単文)	26
5.12	提案手法 (英辞郎) 英日統計翻訳結果 (単文)	27
5.13	提案手法 (英辞郎) 日英統計翻訳結果 (重文複文)	27

5.14	提案手法 (英辞郎) 英日統計翻訳結果 (重文複文)	28
6.1	人手評価結果	29
6.2	日英統計翻訳結果 (単文)	30
6.3	英日統計翻訳結果 (単文)	31
6.4	日英統計翻訳結果 (重文複文)	31
6.5	英日統計翻訳結果 (重文複文)	32
6.6	自動評価結果	33
6.7	提案手法 (鳥バンク) ○の例 (日英翻訳 単文)	34
6.8	フレーズテーブルと日本語フレーズと英語フレーズの対応	35
6.9	提案手法 (鳥バンク) ○の例 1(英日翻訳 重文複文)	35
6.10	フレーズテーブルと日本語フレーズと英語フレーズの対応 1	36
6.11	提案手法 (鳥バンク) ○の例 2(英日翻訳 重文複文)	36
6.12	フレーズテーブルと日本語フレーズと英語フレーズの対応 2	37
6.13	出力文 1 万文中に未知語を含む文数	38
6.14	未知語減少例 1	39
6.15	未知語減少例 2	39
6.16	未知語減少例 3	40
6.17	未知語減少例 4	40
6.18	提案手法 (英辞郎) ○と判断した例 (日英翻訳 単文)	42
6.19	提案手法 (鳥バンク) ○と判断した例 (英日翻訳 単文)	42
6.20	提案手法 (鳥バンク) ○と判断した例 (日英翻訳 重文複文)	43
6.21	差無しと判断した例 1(日英翻訳 重文複文)	43
6.22	差無しと判断した例 2(英日翻訳 重文複文)	44
6.23	差無しと判断した例 3(英日翻訳 重文複文)	45
6.24	差無しと判断した例 4(英日翻訳 重文複文)	45
6.25	鳥バンク動詞の例	47
6.26	英辞郎動詞の例	47
6.27	英辞郎の動詞活用例	48
6.28	動詞活用後の対訳句コーパスの数	48
6.29	自動評価結果	49
6.30	人手評価結果	50

第1章 はじめに

近年，機械翻訳の分野で原言語から目的言語に翻訳する統計翻訳が注目されている．統計翻訳は対訳文を用いてフレーズごとの翻訳確率である翻訳モデルや目的言語らしさである言語モデルを学習する．そして，文が入力された時に翻訳モデルと言語モデルの組合せから確率が最大となる文を出力文として得る．統計翻訳において，対訳文数が多ければ多いほど出力文の翻訳精度は高くなることが知られている．しかし，利用できる対訳文数には限りがある．

この問題に対し，対訳コーパスに対訳句コーパスを追加し，自動評価結果が向上したとの報告がある．この研究はセルビア語英語間，スペイン語英語間 [1] そしてブルトン語フランス語 [2] の各翻訳において行われている．

Maja Popovićらはセルビア語英語間，スペイン語英語間において，対訳コーパスに対訳句コーパスを追加し統計翻訳を行った [1]．セルビア語英語間の翻訳で，対訳句コーパスとして 351 対訳句対を使用した．セルビア語英語方向の翻訳において，対訳コーパス 0.2k に対訳句コーパスを追加することによって，BLEU スコアで 0.020 の向上を報告した．また，スペイン語英語間の翻訳で，対訳句コーパスとして 52,566 対訳句対を使用した．スペイン語英語方向の翻訳において，対訳コーパス 1k に対訳句コーパスを追加することによって，BLEU スコアで 0.06 の向上を報告した．

Francis M Tyers はブルトン語フランス語間の翻訳において対訳句コーパスを用いた [2]．その結果，対訳コーパス 27,987 文に語彙変化しない対訳句コーパスとして 7k 対訳句対を追加することによって，BLEU スコアで 0.01 の向上を報告した．また，対訳コーパスに語彙変化させた対訳句コーパスとして 43k 対訳句対を追加することによって BLEU スコアで 0.07 の向上を報告した．

そこで，本研究では日本語英語間の翻訳において，同様の手法を用いる．辞書のデータから抽出した対訳句コーパスを用いる．日英対訳コーパスを用いて言語モデルの学習する．日英対訳コーパスに対訳句コーパスを加えて翻訳モデルを学習し，統計翻訳を行う方法を提案手法とし，翻訳精度の調査を行う．日英対訳コーパスとして単文コーパスと重文複文コーパス [3] を用いる．対訳句コーパスとして鳥バンク [4] と英辞郎 [5] を用

いる。

その結果，自動評価結果と人手評価結果のどちらにおいても提案手法はベースラインよりも翻訳精度が向上した．日本語英語間の統計翻訳において，対訳句コーパスを用いた提案手法は有効であると確認できた．

本論文の構成は以下の通りである．まず，2章で統計翻訳システムの概要を示し，各モデルの学習，評価方法について述べる．3章では，本研究の提案手法について述べる．4章では，実験に用いるデータやツールといった実験環境について述べる．5章では，提案手法の結果を示す．そして，6章で考察を行い，最後に7章で結論を述べ，まとめる．

第2章 統計翻訳システム

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，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．

2.1 基本概念

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．

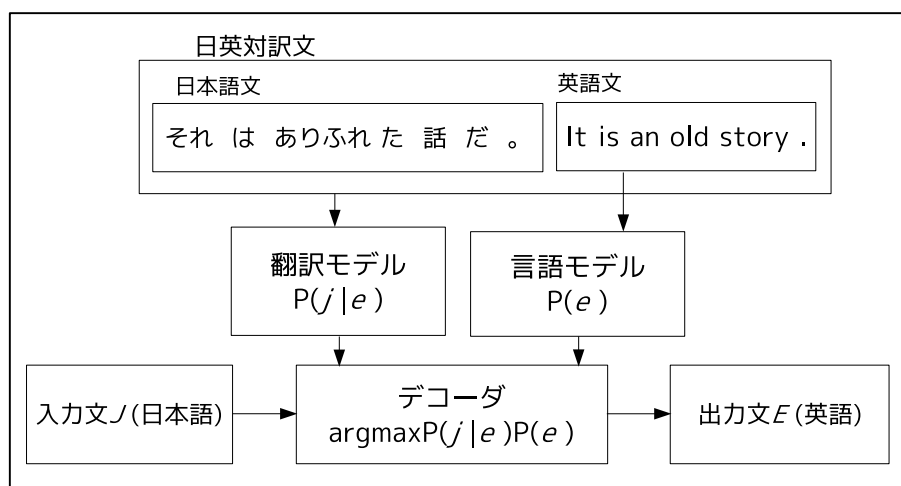


図 2.1: 日英統計翻訳の流れ

2.2 翻訳モデル

翻訳モデルは日本語の単語列から英語の単語列へ確率的に翻訳を行うためのモデルである。翻訳モデルには、大きくわけて語に基づく翻訳モデルと句に基づく翻訳モデル [6] がある。初期の統計翻訳は、語に基づく翻訳モデルを用いていた。しかし、翻訳確率の高さから、現在では句に基づく翻訳モデルが主流となっている。句に基づく翻訳モデルは表 2.1 に示すフレーズテーブルで管理される。

表 2.1: フレーズテーブルの例

日本語 フレーズ	英語 フレーズ	フレーズの 英日翻訳確率	英日方向の単語の 翻訳確率の積	フレーズの 日英翻訳確率	日英方向の単語の 翻訳確率の積	フレーズ ペナルティ
20 秒	twenty seconds	0.5	0.133	1	0.013	2.718
あらし	the storm	0.0625	0.245	0.0476	0.0353	2.718
その 博物館	the museum	0.125	0.037	0.142	0.201	2.718
その 馬	the horse	0.25	0.0575	0.0454	0.182	2.718

左から、日本語フレーズ、英語フレーズ、フレーズの英日翻訳確率 $P(j|e)$ 、英日方向の単語の翻訳確率 (IBM モデル) の積、フレーズの日英翻訳確率 $P(e|j)$ 、日英方向の単語の翻訳確率 (IBM モデル) の積、フレーズペナルティである。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。

2.2.1 IBM モデル

翻訳モデルの代表例として IBM 翻訳モデル [7] がある。IBM モデルは英語文 e 、日本語文 j の翻訳モデル $P(j|e)$ を計算するためにアラインメント a と呼ばれる概念を導入し、以下のような式を考える。なお、アラインメントとは、ある日本語単語 j と英単語 e の対応関係のことを示す。

$$P(j|e) = \sum_a P(j, a|e)$$

IBM モデルでは、日英統計翻訳の場合、英単語は 1 対 n の対応を持ち、日本語の単語は 1 つの英単語のみと対応すると仮定する。また、日本語の単語の対応関係として適切な英単語がなかった場合、英語文の文頭の特異文字 e と対応付けを行う。

2.2.2 GIZA++

GIZA++[8]とは、統計翻訳を用いることを前提に作られた単語対応のアラインメントを行うツールである。IBMモデルを学習し、単語の対応関係の確率値を計算する。

2.2.3 フレーズテーブルの作成方法

GIZA++よりIBMモデルを推定することで最尤な単語アラインメントを得る。これを日英、英日の両方向に対して行う。対訳文，“We ate all the apple pie”“私たちはアップルパイを全部食べてしまった”を学習データとした時の例で日英方向の単語対応の例を表2.2に、英日方向の単語対応の例を表2.3に示す。また、表中の“■”は獲得した最尤な単語アラインメントを表す。

表 2.2: 日英方向の単語対応

	私	たち	は	アップルパイ	を	全て	食べ	て	しまっ	た
We	■	■								
ate							■		■	■
all						■		■		
the					■					
apple										
pie				■						

表 2.3: 英日方向の単語対応

	私	たち	は	アップルパイ	を	全て	食べ	て	しまっ	た
We		■								
ate							■			
all						■				
the					■					
apple				■						
pie				■						

次に、両方向のアラインメントから、両方向に1対多の対応を認めた単語アラインメントの計算を行う。この単語アラインメントは基本的に両方向の単語対応の積集合(intersection)と和集合(union)の中間をヒューリスティックで求める。なお、積集合は両方向ともに単語対応が存在する場合のみ対応を残し、和集合は少なくとも片方向に単

語対応が存在する場合、単語対応を残す。対称な単語対応を求めるヒューリスティック (grow-diag-final など) は、まず積集合から始まり、和集合にしかない単語対応が妥当であるかを判断しながら単語対応を徐々に加える。対称化された単語アラインメントの獲得の例として、表 2.4 に (積集合 intersection) を、表 2.5 に (和集合 union) を、表 2.6 に (ヒューリスティック grow-diag-final) を示す。

表 2.4: 積集合 intersection

	私	たち	は	アップルパイ	を	全て	食べ	て	しまっ	た
We		■								
ate							■			
all						■				
the					■					
apple										
pie				■						

表 2.5: 和集合 union

	私	たち	は	アップルパイ	を	全て	食べ	て	しまっ	た
We	■	■								
ate							■		■	■
all						■		■		
the					■					
apple				■						
pie				■						

表 2.6: ヒューリスティック grow-diag-final

	私	たち	は	アップルパイ	を	全て	食べ	て	しまっ	た
We	■	■								
ate							■			
all						■				
the					■					
apple				■						
pie				■						

対称化された単語アラインメントのうち、矛盾しないすべてのフレーズ対応を得る。抽出したフレーズ対の例を表 2.7 に示す。

表 2.7: 作成されたフレーズ対の例

日本語フレーズ	英語フレーズ
アップルパイ	apple pie
アップルパイを全部食べてしまった	ate all the apple pie
私たち	We
私たちは	We
全て	all
食べてしまった	ate

抽出したフレーズ対応に対して確率付けを行う。日本語フレーズ J_{phrase} と英語フレーズ E_{phrase} からなるフレーズ対応の確率は以下の式で計算される。また、フレーズ対に確率を付与し、作成されたフレーズテーブルの例を表 2.8 に示す。

$$P(J_{phrase}|E_{phrase}) = \frac{\text{学習データ中で } J_{phrase} \text{ と } E_{phrase} \text{ が同時に出現した数}}{\text{学習データ中で } E_{phrase} \text{ が出現した数}} \quad (2.3)$$

$$P(E_{phrase}|J_{phrase}) = \frac{\text{学習データ中で } J_{phrase} \text{ と } E_{phrase} \text{ が同時に出現した数}}{\text{学習データ中で } J_{phrase} \text{ が出現した数}} \quad (2.4)$$

表 2.8: 作成されたフレーズテーブルの例

日本語 フレーズ	英語 フレーズ	フレーズの 英日翻訳確率	英日方向の単語の 翻訳確率の積	フレーズの 日英翻訳確率	日英方向の単語の 翻訳確率の積
アップルパイ	apple pie	1	1	1	1
アップルパイを 全部食べてしまった	ate all the apple pie	1	1	1	1
私たち	We	1	1	1	1
私たちは	We	1	1	1	1
全て	all	1	1	1	1
食べてしまった	ate	1	1	1	1

2.3 言語モデル

言語モデルは翻訳候補の文に対して目的言語の文らしさの指標を与えるモデルである。翻訳モデルでは、訳語の選択や訳語の位置の選択に対する評価を与えることはできるが、作られた翻訳候補が目的言語の文としてふさわしいかどうかを判断する評価を与えることはできない。そのため、言語モデルでは日英統計翻訳の場合、より英語らしい文に対して、高い確率を与えることで、翻訳モデルで翻訳された訳文候補の中から英語として自然な文を選出する。

2.3.1 N -gram モデル

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。また、 $N=1$ のモデルを uni-gram, $N=2$ のモデルを bi-gram, $N=3$ のモデルを tri-gram と特有の呼びかたをする。 $N=4$ 以上は 4-gram など数値を用いて呼ぶ。 N -gram は以下の式で表現される。ここで、 w_i^j は i から j 番目までの単語列を表す。

$$P(w_1^n) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \dots P(w_n|w_1^{n-1}) \quad (2.5)$$

$$\approx P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \dots P(w_n|w_{n-(N-1)}^{n-1}) \quad (2.6)$$

$$= \prod_{i=1}^n P(w_i|w_{i-(N-1)}^{i-1}) \quad (2.7)$$

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。

$$P(w_i|w_{i-(N-1)}^{i-1}) = \frac{C(w_{i-(N-1)}^i)}{C(w_{i-(N-1)}^{i-1})} \quad (2.8)$$

たとえば、“I have dogs .” という単語列に対して $N=2$ とした bi-gram モデルの言語モデルを適応した場合、単語列が生成される確率は以下の式で計算される。

$$P(\text{“I have dogs .”}) \simeq P(I) \times P(\text{have}|I) \times P(\text{dogs|have}) \dots P(\text{.}|dogs) \quad (2.9)$$

tri-gram モデルであれば, $P(\text{dogs}|I \text{ have})$, 4-gram モデルであれば $P(.|I \text{ have dogs})$ となる.

(2.8) 式から信頼性の高い値を推定するためには, 単語列 w_1^n が多く出現している必要がある. しかし, 実際には多くの単語列は出現数が 0 となることが多いため信頼できる値を推定できない場合が多い. 低頻度な語彙の場合, $C(w_{i-(N-1)}^i), C(w_{i-(N-1)}^{i-1})$ の値が小さく, 信頼性が低い. また, 学習データ中に単語列 w_1^i が存在しない場合, この単語列の出現確率は 0 と推定される. そのため, (2.8) 式から信頼できる値を算出するためには, 大規模なコーパスを用いて, 各単語列の出現数を高める必要がある. そこで, 出現頻度の少ない単語列をモデルの学習から削除 (カットオフ) する方法や, 確率が 0 となるのを防ぐために, 大きい確率を小さく, 小さい確率を大きくするスムージング手法が提案されている. スムージングの代表的な手法にバックオフ・スムージングがある. バックオフ・スムージングは学習データに出現しない N -gram の値をより低い次数の $(N-1)$ -gram の値から推定する.

2.4 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. 代表的なデコーダに Moses[9] がある.

2.5 評価方法

本研究では, 翻訳システムによって出力した文の評価に自動評価法と人手評価法を用いる.

2.5.1 自動評価法

機械翻訳システムの翻訳精度を自動的に評価する手法として, あらかじめ用意した正解文と, 翻訳システムで出力した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ある. 本研究では, BLEU[10], NIST[11], RIBES[12], METEOR[13] を用いる.

2.5.2 BLEU

BLEU は語順 (4-gram) が正しい場合に高いスコアを出す。BLEU は以下の式で計算される。

$$BLEU_{score} = BP \times \exp \left(\sum_{n=1}^N \frac{1}{N} \log p_n \right) \quad (2.10)$$

$$p_n = \frac{\sum_i \text{出力文 } i \text{ と正解文 } i \text{ で一致した } N\text{-gram の数}}{\sum_i \text{出力文 } i \text{ の } N\text{-gram の数}} \quad (2.11)$$

ここで、 P_n は出力文と正解文の N -gram の一致率を表している。BLEU はこの一致率を 1-gram から 4-gram まで計算し、その幾何平均をとる。また、出力文が正解文より短い場合、“ $\sum_i \text{出力文 } i \text{ の } N\text{-gram の数}$ ” が小さくなり、不当にスコアが高くなる可能性がある。そこで、正解文より短い文に対するペナルティとして、 BP を用いる。 BP は出力文が正解文より長い場合は 1 となり、出力文が正解文より短い場合は 1 未満の値となる。

2.5.3 NIST

NIST では BLEU と同様に語順の正しさを比較を行うが、5-gram を用いる。NIST は以下の式で計算される。

$$NIST_{score} = BP \times \sum_{N=1}^N \frac{\sum_i \left(\sum_{\substack{\text{出力文 } i \text{ と正解文 } i \text{ に} \\ \text{共通する } w_1 \dots w_n}} Info_i(w_1 \dots w_n) \right)}{\sum_i \text{出力文 } i \text{ と正解文 } i \text{ で一致した } N\text{-gram の数}} \quad (2.12)$$

$$Info_i(w_1 \dots w_n) = \log_2 \frac{\text{評価コーパス中 } (w_1 \dots w_{n-1}) \text{ の数}}{\text{評価コーパス中の } (w_1 \dots w_n) \text{ の数}} \quad (2.13)$$

2.5.4 RIBES

RIBES は、正解文と出力文の間で、共通単語の出現順序を順位相関係数で評価を行う評価法である。RIBES は以下の式で計算される。

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

$$RIBES(K) = NKT \times P^\alpha \quad (2.15)$$

$$P = \frac{n}{h} \quad (2.16)$$

ここで、 NSR はスパイマンの順位相関係数であり、 NKT はケンドールの順位相関係数である。また P はペナルティであり、 n は出力文と正解文との間で共通な単語の数、 h は出力文の単語数で計算する。 α はペナルティに対する主として使用され、 $0 \leq \alpha \leq 1$ の値である。単語の出現順を順位相関係数を用いて評価することで、文全体の語順に着目することができる。

2.5.5 METEOR

METEOR は単語属性 (3 人称単数など) が正しい場合に高いスコアを出す。METEOR は以下の式で計算される。

$$METEOR_{score} = F_{mean} \times (1 - Pen) \quad (2.17)$$

$$F_{mean} = \frac{P \times R}{\alpha \times P + (1 - \alpha) \times R} \quad (2.18)$$

$$Pen = \gamma \times \left(\frac{c}{m}\right)^\beta \quad (2.19)$$

METEOR はまず再現率 R と適合率 P に基づく F 値を求め、次に、単語の非連続性に対するペナルティとして関数 Pen を与える。ペナルティ関数 Pen において、 m は出力文と正解文の単語の一致率を表す。そして、 c は一致した単語を対象に、正解文と語順が同じものを 1 つのまとまりとして統合した場合の、まとまりの数を表す。そのため、出力文と正解文が同じ文であるとき $c=1$ となる。また、一致率の計算において、WordNet による類義語を用いて、似た意味を持つ単語は同一であると判断される。 α , β , γ の値はパラメータである。

BLEU と METEOR と RIBES では 0 から 1 までの間で評価され、NIST では 0 から ∞ までの間で評価される。いずれの評価方法でも、評価方法が高いほど翻訳精度が高いことを表す。なお、本研究では入力文 1 文に対して正解文 1 文を用いて評価を行う。

2.5.6 人手評価法

ベースラインの翻訳結果と提案手法の翻訳結果からそれぞれランダムに 100 文ずつを抽出し、対比較評価を行う。

提案手法の翻訳結果がベースラインの翻訳結果より優れている場合は“提案手法○”とする。一方、提案手法の翻訳結果がベースラインの翻訳結果より劣っている場合は“ベースライン○”とする。また、提案手法の翻訳結果とベースラインの翻訳結果に変化が見

られたが、翻訳品質がどちらも変わらない場合は、“差無し”とし、提案手法の翻訳結果とベースラインの翻訳結果が同じ場合、“同一出力”とする。

第3章 提案手法

本研究では辞書のデータから抽出した対訳句コーパスを日英対訳コーパスに追加し、統計翻訳を行う手法を提案手法とする。

図 3.1 に日英統計翻訳の場合の提案手法の流れを示す。

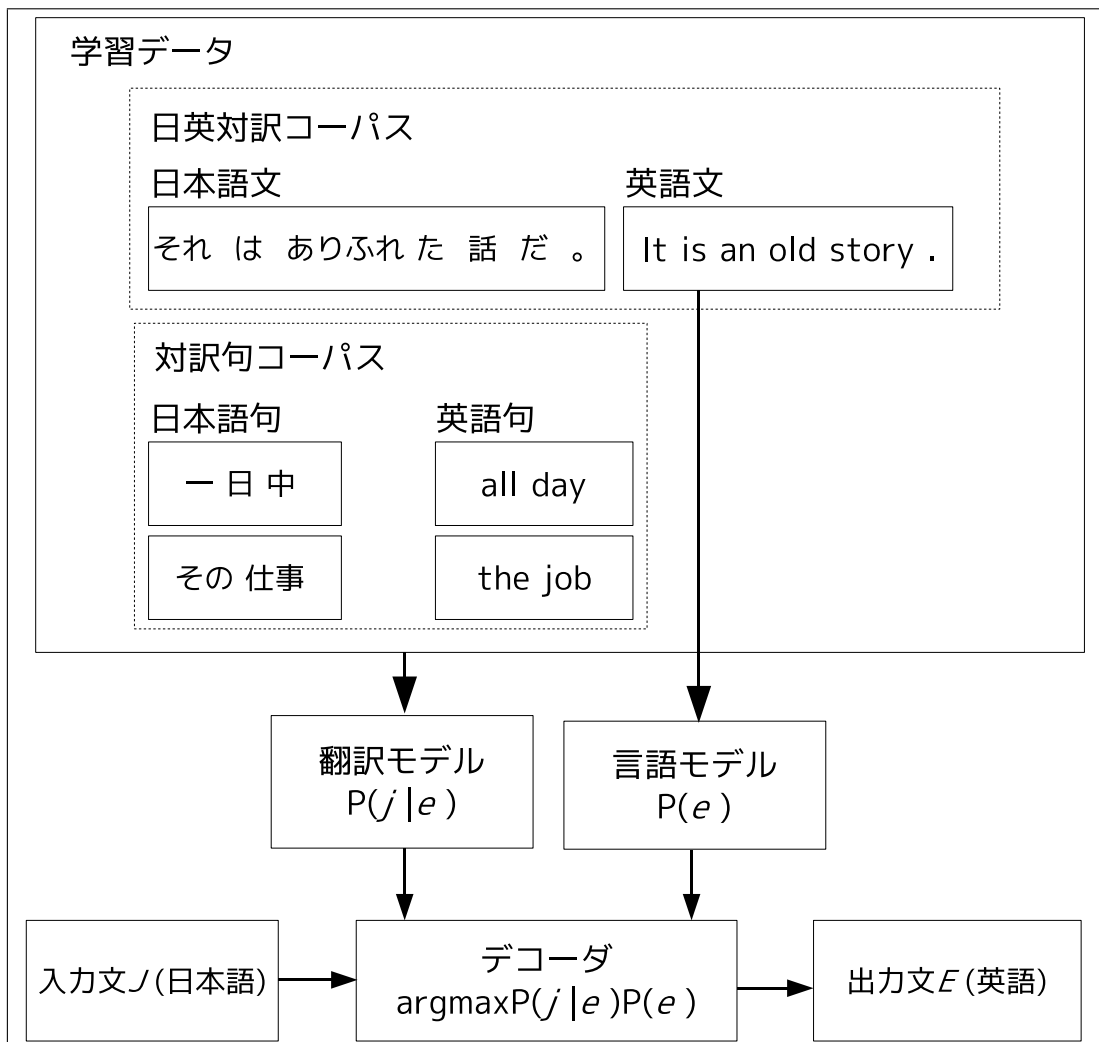


図 3.1: 日英統計翻訳の場合の提案手法の流れ

提案手法の手順を以下に示す.

手順1 日英対訳コーパスを学習データとして言語モデルを作成する

手順2 日英対訳コーパスに対訳句コーパスを追加する

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

第4章 実験環境

4.1 日英対訳コーパス

本研究では日英対訳コーパスとして、単文コーパスと重文複文コーパス [3] を用いる。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[14] を用いて形態素解析を行う。また、英語文に対して “tokenizer.sed [15]” を用いて正規化を行う。前処理後の単文コーパスの例を表 4.1 に、重文複文コーパスの例を表 4.2 にそれぞれ示す。

表 4.1: 前処理後の単文コーパスの例

日本語文	魚がたくさん取れた。
英語文	Many fish were caught .
日本語文	娘は今年中学校に上がった。
英語文	My daughter advanced to middle school this year .
日本語文	彼は踏み台に乗った。
英語文	He got on to a stool .

表 4.2: 前処理後の重文複文コーパスの例

日本語文	勉強をしている間はラジオを切っておきなさい。
英語文	While studying , turn off the radio .
日本語文	たかしは言いつけをよく守る子だ。
英語文	Takashi is an obedient boy .
日本語文	映画が始まるまでにまだ 30 分ある。
英語文	We still have 30 minutes before the movie starts .

4.2 翻訳モデルの学習

翻訳モデルはフレーズテーブルで管理されている。フレーズテーブルの学習には、多くの方法がある。本研究では、Moses の付録である “train-model.perl” [9] を用いる。このプログラムは GIZA++ を利用し、IBM モデルの計算を行う。本研究では、フレーズ

テーブルのヒューリスティックとして “grow-diag-final” を用いる。また、フレーズテーブルを作成する際、フレーズテーブル内の日本語と英語のフレーズ中の単語数の上限として、max-phrase-length が定義されている。たとえば、max-phrase-length の値が7の場合、日本語か英語のいずれかのフレーズ中の単語数が、8以上のフレーズ対は作成されない。本研究では、max-phrase-length の値は20を用いる。

4.3 言語モデルの学習

言語モデルは、 N -gram モデルを用いる。 N -gram モデルの学習には “SRILM[16]” の ngram-count を用いる。 N -gram モデルの次数は、先行研究により 5-gram が有効であることがわかっている。そのため、本研究でも、5-gram の言語モデルを用いる。なお、スムージングに “Kneser-Ney discount” を用いる。

4.4 デコーダのパラメータ

デコーダは “Moses[15]” を使用する。Moses はいくつかのパラメータを設定することができる。パラメータの最適化には、“Minimum Error Rate Training(MERT)[17]” を用いることが一般的である。MERT は目的の評価関数 (通常は BLEU) を最大とするような翻訳結果が選ばれるようにパラメータを調節する。この際、development データと呼ばれる試し翻訳用のデータを与え、各文について、上位 100 文程度の翻訳候補の中で重みを変え、各文について上位にくるようにパラメータを調節する。しかし、パラメータの最適化は development データに依存し、また、多くの時間がかかるため、本研究ではパラメータの最適化は行わない。クロスエントロピーを用いるため、“weight-t” は “0.5 0.0 0.5 0.0 0.0” とする。翻訳時にフレーズの位置変化に柔軟に対応するため、“distortion weight” は 0.2 とする。

4.5 対訳句コーパス

本研究では対訳句コーパスとして鳥バンク [4] と英辞郎 [5] を用いる。

4.5.1 鳥バンク

鳥バンクは自然言語処理のための言語知識ベースを収録したデータバンクであり，日本語の重文と複文を対象とする「意味類型パターン辞書」が収録されている．本研究では鳥バンクから抽出した 698,472 対訳句対 [18] を用いる．対訳句対の例を表 4.3 に示す．

表 4.3: 鳥バンク対訳句対の例

日本語句	発酵槽に移し
英語句	transferred to the fermenter
日本語句	摩擦熱
英語句	Friction accounts
日本語句	コート の すそ
英語句	the edge of my coat

また，単語数による対訳句対数の分布を調査した．鳥バンクの英語句の結果を図 4.1 に，日本語句の結果を図 4.2 に示す．

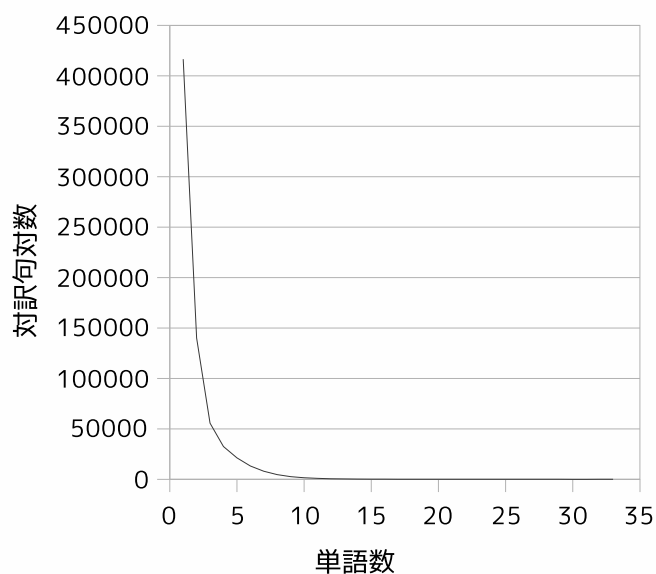


図 4.1: 単語数による対訳句対数の分布 (鳥バンク 英語句)

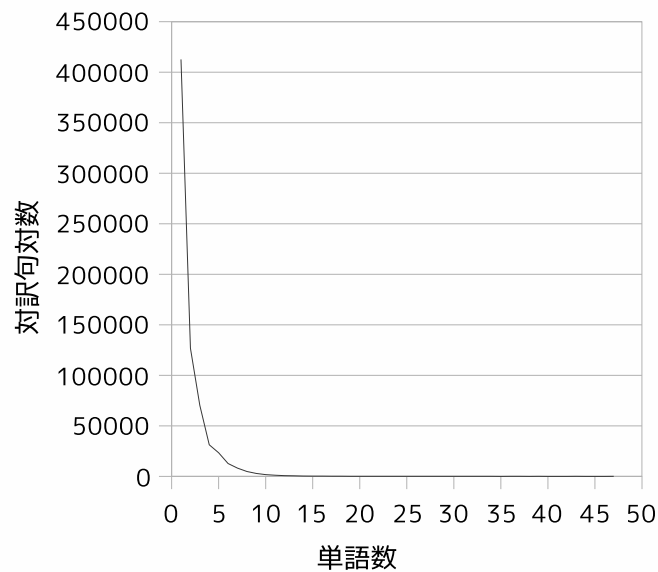


図 4.2: 単語数による対訳句対数の分布 (鳥バンク 日本語句)

4.5.2 英辞郎

英辞郎は、EDP(Electronic Dictionary Project) がアップデートし続けている英和・和英辞書である。英辞郎のデータには対訳句対の他に翻訳例や注釈や読み仮名、本来の文に出てこない“～”等の記号が含まれる。

本研究では英辞郎のクリーニングを行い、必要な英語と日本語の対訳句対のみの形にした 1,350,299 対訳句対を用いる。クリーニング後の英辞郎の対訳句対の例を表 4.4 に示す。

表 4.4: クリーニング後の英辞郎の対訳句対の例

日本語句	が どのように 貧乏 に せよ
英語句	be someone ever so poor
日本語句	から 出て くる
英語句	come out from
日本語句	の 結果 として 生じる
英語句	come out from

また、単語数による対訳句対数の分布を調査した。英辞郎の英語句の結果を図 4.3 に、日本語句の結果を図 4.4 に示す。

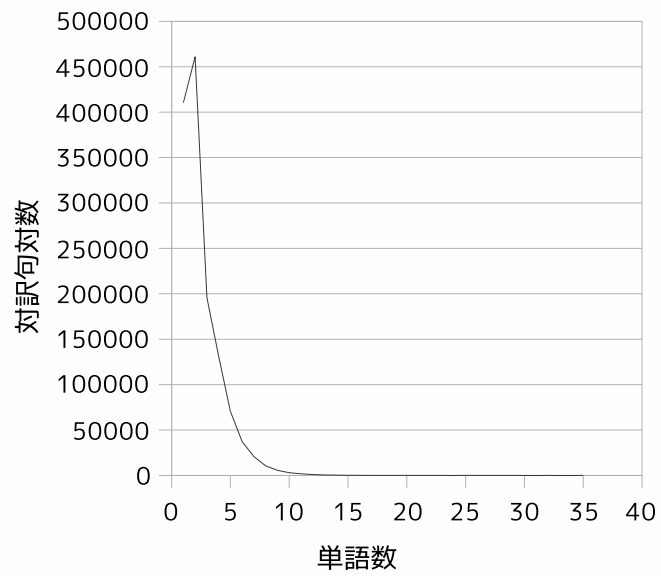


図 4.3: 単語数による対訳句対数の分布 (英辞郎 英語句)

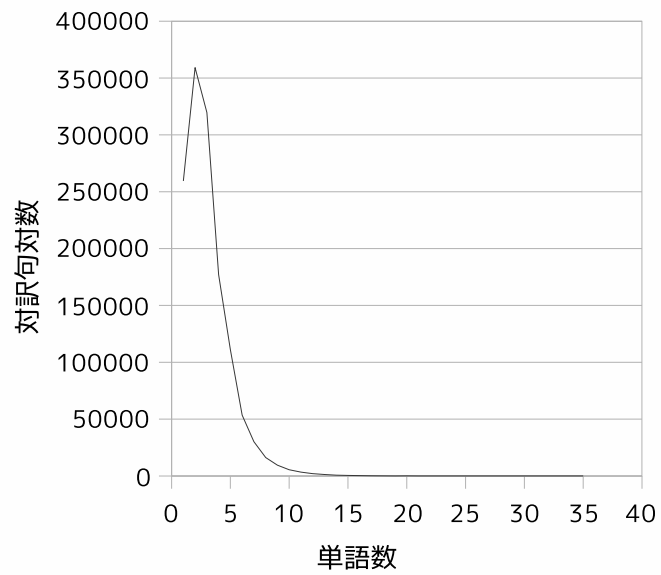


図 4.4: 単語数による対訳句対数の分布 (英辞郎 日本語句)

第5章 翻訳実験

本研究では日英対訳コーパスとして、単文コーパスと重文複文コーパス [3] を用いた2種類の翻訳実験を行う。また、対訳句コーパスとして鳥バンク [4] と英辞郎 [5] の2種類を用いる。翻訳実験は日英統計翻訳と英日統計翻訳を行う。したがって、合計8種類の翻訳実験を行う。

翻訳実験の種類を表 5.1 に示す。

表 5.1: 翻訳実験の種類

日英対訳コーパス	対訳句コーパス	統計翻訳
単文コーパス	鳥バンク	日英
単文コーパス	鳥バンク	英日
単文コーパス	英辞郎	日英
単文コーパス	英辞郎	英日
重文複文コーパス	鳥バンク	日英
重文複文コーパス	鳥バンク	英日
重文複文コーパス	英辞郎	日英
重文複文コーパス	英辞郎	英日

単文コーパスの実験は単文コーパス 182,988 文から、重文複文コーパスの実験では重文複文コーパス 102,712 文から表 5.2 の内訳で用いる。また、対訳句対数を表 5.3 に示す。

表 5.2: 使用した実験データ

	単文	重文複文
日英対訳コーパス	100,000	91,712
テストデータ	10,000	10,000

表 5.3: 対訳句コーパスごとの対訳句対数

対訳句コーパス名	対訳句対数
鳥バンク	698,472
英辞郎	1,350,299

5.1 ベースラインと提案手法

ベースラインでは言語モデルと翻訳モデルの作成に学習データを用いる。提案手法では言語モデルの作成に学習データを，翻訳モデルの作成に学習データと対訳句コーパスを用いる。本研究において対訳句コーパスとして，鳥バンクを用いる翻訳実験を提案手法(鳥バンク)と呼び，英辞郎を用いる翻訳実験を提案手法(英辞郎)と呼ぶ。言語モデルと翻訳モデルの作成に使用するコーパスについて表 5.4 にまとめる。

表 5.4: 言語モデルと翻訳モデルの作成に使用するコーパス

翻訳実験	学習データ	
	翻訳モデル	言語モデル
ベースライン	日英対訳コーパス	日英対訳コーパス
提案手法(鳥バンク)	日英対訳コーパス+鳥バンク	日英対訳コーパス
提案手法(英辞郎)	日英対訳コーパス+英辞郎	日英対訳コーパス

5.2 評価実験

本研究では，出力文の評価として自動評価法と人手評価法を行う。自動評価法は“BLEU[10]”，“NIST[11]”，“RIBES[12]”，“METEOR[13]”を用いる。人手評価法はベースラインと提案手法の出力文からランダムに 100 文ずつを抽出し，対比較評価を行う。

5.3 実験結果

5.3.1 自動評価結果

自動評価結果を表 5.5 に示す.

表 5.5: 自動評価結果

(1) 日英統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1195	4.521	0.6999	0.4923
提案手法 (鳥バンク)	0.1401	4.964	0.7144	0.5189
提案手法 (英辞郎)	0.1379	5.017	0.7173	0.5199

(2) 英日統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1620	4.384	0.6241	-
提案手法 (鳥バンク)	0.1701	4.759	0.6371	-
提案手法 (英辞郎)	0.1813	4.795	0.6503	-

(3) 日英統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.0852	3.743	0.6420	0.4310
提案手法 (鳥バンク)	0.2193	6.086	0.7101	0.5533
提案手法 (英辞郎)	0.1068	4.318	0.6627	0.4609

(4) 英日統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1239	3.953	0.5618	-
提案手法 (鳥バンク)	0.2275	5.895	0.6459	-
提案手法 (英辞郎)	0.1383	4.276	0.5782	-

表 5.5 より, すべての自動評価法において, 提案手法はベースラインより翻訳精度が向上した. また, 重文複文コーパスにおいて提案手法 (鳥バンク) は提案手法 (英辞郎) より大幅に翻訳精度が向上した.

5.3.2 人手評価結果

日英統計翻訳，英日統計翻訳においてベースラインと提案手法の人手評価を行った。ベースライン○は提案手法がベースラインより翻訳品質が劣っていることを示し，提案手法○は提案手法がベースラインより翻訳品質が優れていることを示す。また，差無しは翻訳品質に差が無いことを示し，同一出力は出力文が完全に同一であることを示す。人手評価結果を表 5.6 に示す。

表 5.6: 人手評価結果

(1) 日英統計翻訳 単文コーパス

	提案手法○	ベースライン○	差無し	同一出力
提案手法 (鳥バンク)	14	3	74	9
提案手法 (英辞郎)	13	2	77	8

(2) 英日統計翻訳 単文コーパス

	提案手法○	ベースライン○	差無し	同一出力
提案手法 (鳥バンク)	13	3	63	11
提案手法 (英辞郎)	14	4	71	11

(3) 日英統計翻訳 重文複文コーパス

	提案手法○	ベースライン○	差無し	同一出力
提案手法 (鳥バンク)	11	2	84	3
提案手法 (英辞郎)	10	0	90	0

(4) 英日統計翻訳 重文複文コーパス

	提案手法○	ベースライン○	差無し	同一出力
提案手法 (鳥バンク)	21	6	67	6
提案手法 (英辞郎)	14	4	76	6

表 5.6 より，すべての人手評価結果において，提案手法はベースラインよりも優れている。

5.3.3 出力例

ベースラインと提案手法(鳥バンク)の対比較の出力例を表 5.7 から表 5.10 に示す。また、ベースラインと提案手法(英辞郎)の対比較の出力例を表 5.11 から表 5.14 に示す。

表 5.7: 提案手法(鳥バンク)日英統計翻訳結果(単文)

ベースライン○の出力例	
入力文	全 国民 の 関 心 が 選 挙 に 集 ま っ て い る 。
正解文	The interest of every citizen is focussed on the elections .
ベースライン	The whole nation is concentrated in the election .
提案手法(鳥バンク)	The whole nation attention focused on the election .
提案手法(鳥バンク)○の出力例	
入力文	朝 顔 の 手 が 伸 び て き た 。
正解文	The vines of the morning glory are growing .
ベースライン	My hands came in the morning .
提案手法(鳥バンク)	My morning glories has grown up .
差無し of 出力例	
入力文	わ た し は 一 音 楽 家 に す ぎ な い 。
正解文	I am only a musician .
ベースライン	I only a musician .
提案手法(鳥バンク)	I have no more than a musician .

表 5.8: 提案手法 (鳥バンク) 英日統計翻訳結果 (単文)

ベースライン○の出力例	
入力文	I want you to be very careful .
正解文	十分な注意を望みます。
ベースライン	非常に気をつけて欲しい。
提案手法 (鳥バンク)	に注意してほしい。
提案手法 (鳥バンク) ○の出力例	
入力文	A siren pierced the morning quiet .
正解文	サイレンが朝の静けさをつんざいた。
ベースライン	朝のサイレンが静かに突き刺さった。
提案手法 (鳥バンク)	サイレンが静かな朝の静けさを破った。
差無し of 出力例	
入力文	He cracked a bone .
正解文	彼の骨にひびが入った。
ベースライン	彼は骨と鳴った。
提案手法 (鳥バンク)	骨を割った。

表 5.9: 提案手法 (鳥バンク) 日英統計翻訳結果 (重文複文)

ベースライン○の出力例	
入力文	商売が上手く行って金持ちになった。
正解文	Business went well and they became wealthy .
ベースライン	He became a rich man business has been well .
提案手法 (鳥バンク)	Business went well with growing rich .
提案手法 (鳥バンク) ○の出力例	
入力文	新陳代謝を変えるのは難しい。
正解文	It is hard to change one 's metabolism .
ベースライン	It is difficult to alter 新陳代謝 .
提案手法 (鳥バンク)	It is hard to change metabolism .
差無し of 出力例	
入力文	傘を持って来ればよかった。
正解文	I wish I had brought an umbrella .
ベースライン	How I had an umbrella with you .
提案手法 (鳥バンク)	Take an umbrella with should come .

表 5.10: 提案手法 (鳥バンク) 英日統計翻訳結果 (重文複文)

ベースライン○の出力例	
入力文	He is too intent on his studies to think of anything else .
正解文	学問に一心不乱だから余事を省る暇が無い。
ベースライン	彼は学問に一心で他事を顧みる暇が無い。
提案手法 (鳥バンク)	彼は学問に一心に何を考えている。
提案手法 (鳥バンク) ○の出力例	
入力文	He said awful things while under the influence of alcohol .
正解文	彼は酒気を帯びてひどいことを言った。
ベースライン	彼の言ったことをひどくアルコールの間に影響を与える。
提案手法 (鳥バンク)	アルコールの影響を受けている間にひどいことを言った。
差無しの出力例	
入力文	Say it , do not spray it .
正解文	つばを飛ばさずに話せよ。
ベースライン	それ Say しぶきがない。
提案手法 (鳥バンク)	それはしぶきを言ってしまった。

表 5.11: 提案手法 (英辞郎) 日英統計翻訳結果 (単文)

ベースライン○の出力例	
入力文	その件はすっかり決まった。
正解文	The matter is quite decided .
ベースライン	The matter has been decided .
提案手法 (英辞郎)	The matter can all set .
提案手法 (英辞郎) ○の出力例	
入力文	世界の人口は年ごとに膨張を続けている。
正解文	The population of the world goes on increasing year by year .
ベースライン	The population of the world are continuing their goes on every year .
提案手法 (英辞郎)	The population of the world have been expansion a year by year .
差無しの出力例	
入力文	古代エジプトの文化はきわめて高度なものであった。
正解文	The culture of ancient Egypt was highly advanced .
ベースライン	The culture of a high on what was very ancient Egypt .
提案手法 (英辞郎)	The very highly of Egyptian civilization was very .

表 5.12: 提案手法 (英辞郎) 英日統計翻訳結果 (単文)

ベースライン○の出力例	
入力文	He ca n't even write his own name properly .
正解文	あの人は自分の名前も満足に書けない。
ベースライン	彼は正しくは自分の名前も書けない。
提案手法 (英辞郎)	彼は自分の名前満足に書けない。
提案手法 (英辞郎) ○の出力例	
入力文	He drove in a pile there .
正解文	彼はそこに杭を打った。
ベースライン	彼はその場で山になった。
提案手法 (英辞郎)	彼はそこに杭を打ち込んだ。
差無しの出力例	
入力文	We can do nothing but follow the march of events .
正解文	事件の流れに身をまかせる以外ない。
ベースライン	私たちは行進の出来事に従っ 以外は何もできない。
提案手法 (英辞郎)	私たちはできるだけにの行進した。

表 5.13: 提案手法 (英辞郎) 日英統計翻訳結果 (重文複文)

提案手法 (英辞郎) ○の出力例	
入力文	名古屋 から 金沢 へ 行く には いくつか の ルート が ある 。
正解文	There are several ways to go from Nagoya to Kanazawa .
ベースライン	There is some channels 金沢 from Nagoya to go .
提案手法 (英辞郎)	There is some route to go to Kanazawa from Nagoya .
差無しの出力例	
入力文	彼は事を軽蔑してかかったから失敗したのだ。
正解文	He made light of the matter , and failed .
ベースライン	He took the matter with contempt , I failed .
提案手法 (英辞郎)	He took it down to the failure .

表 5.14: 提案手法 (英辞郎) 英日統計翻訳結果 (重文複文)

ベースライン○の出力例	
入力文	I lent him a hand in doing it .
正解文	それを する について 彼に 手を 貸した。
ベースライン	私は 彼に 手を 貸して それを した。
提案手法 (英辞郎)	私は 彼に 手を 貸して くれた そう する。
提案手法 (英辞郎) ○の出力例	
入力文	I did not feel any pain in my foot at the time I was injured .
正解文	足に けがを した 直後 には 痛みを 感じ なかった。
ベースライン	その ときは けがを した 足が 痛く ても なかった。
提案手法 (英辞郎)	私の 足に 怪我を した 時 には 苦痛 とは 思わ なかった。
差無し of 出力例	
入力文	He had a clear recollection of having witnessed the event .
正解文	その 出来事 を 目撃 した のを はっきり と 覚えて いた。
ベースライン	彼の 顔が はっきり した こと を 催した。
提案手法 (英辞郎)	証人 として 記憶 している こと を 明らかに した。

第6章 考察

6.1 対訳句コーパスごとの翻訳品質の比較

提案手法 (鳥バンク) と提案手法 (英辞郎) の翻訳品質を比較するために、人手評価を行う。提案手法 (鳥バンク) と提案手法 (英辞郎) の出力文からランダムに 100 文ずつを抽出し、対比較評価を行った。

提案手法 (鳥バンク) ○は提案手法 (鳥バンク) が提案手法 (英辞郎) より翻訳品質が優れていることを示し、提案手法 (英辞郎) ○は提案手法 (鳥バンク) が提案手法 (英辞郎) より翻訳品質が劣っていることを示す。また、差無しは翻訳品質に差が無いことを示し、同一出力は出力文が完全に同一であることを示す。

人手評価結果を表 6.1 に示す。

表 6.1: 人手評価結果

(1) 日英統計翻訳 単文コーパス

提案手法 (鳥バンク) ○	提案手法 (英辞郎) ○	差無し	同一出力
12	6	77	5

(2) 英日統計翻訳 単文コーパス

提案手法 (鳥バンク) ○	提案手法 (英辞郎) ○	差無し	同一出力
12	7	67	14

(3) 日英統計翻訳 重文複文コーパス

提案手法 (鳥バンク) ○	提案手法 (英辞郎) ○	差無し	同一出力
12	4	83	1

(4) 英日統計翻訳 重文複文コーパス

提案手法 (鳥バンク) ○	提案手法 (英辞郎) ○	差無し	同一出力
12	4	78	6

表 6.1 より, 提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。

6.1.1 出力例

提案手法 (鳥バンク) と提案手法 (英辞郎) の対比較の出力例を表 6.2 から表 6.5 に示す。

表 6.2: 日英統計翻訳結果 (単文)

提案手法 (鳥バンク) ○の出力例	
入力文	彼は涙を流した。
正解文	He shed tears .
提案手法 (鳥バンク)	He was moved to tears .
提案手法 (英辞郎)	He drew tears .
提案手法 (英辞郎) ○の出力例	
入力文	明日 此処で会議が行われます。
正解文	The meeting will be held here tomorrow .
提案手法 (鳥バンク)	The meeting was held here tomorrow .
提案手法 (英辞郎)	The meeting will be held here tomorrow .
差無しの出力例	
入力文	金の指輪を買いたい。
正解文	I would like to buy a ring of gold .
提案手法 (鳥バンク)	I 'd like to buy a gold ring .
提案手法 (英辞郎)	I want to buy a gold ring .

表 6.3: 英日統計翻訳結果 (単文)

提案手法 (鳥バンク) ○の出力例	
入力文	Please let me join this group .
正解文	わたしも このグループに加えてください。
提案手法 (鳥バンク)	この会に入れてください。
提案手法 (英辞郎)	この会にてください。
提案手法 (英辞郎) ○の出力例	
入力文	The chairman pounded his desk with a gavel .
正解文	議長が槌で机を叩いた。
提案手法 (鳥バンク)	彼の机の上に gavel 議長をたたいている。
提案手法 (英辞郎)	議長は小槌で机をたたいていた。
差無しの出力例	
入力文	I do not take breakfast from to time .
正解文	私はときどき朝食を抜かす。
提案手法 (鳥バンク)	朝食に時間を持っていない。
提案手法 (英辞郎)	ない時からに朝食を取る。

表 6.4: 日英統計翻訳結果 (重文複文)

提案手法 (鳥バンク) ○の出力例	
入力文	少女は男の甘いことばに惑わされた。
正解文	The girl was seduced by his honeyed words .
提案手法 (鳥バンク)	The girl was deluded by a man 's sweet words .
提案手法 (英辞郎)	The girl 's honeyed words deluded by man .
提案手法 (英辞郎) ○の出力例	
入力文	知識は自然を支配する力となる。
正解文	Knowledge means power over nature .
提案手法 (鳥バンク)	Knowledge is control nature is a power .
提案手法 (英辞郎)	Knowledge is a power over nature .
差無しの出力例	
入力文	彼が居なくても何ともない。
正解文	I do not miss him at all .
提案手法 (鳥バンク)	He can not be quite .
提案手法 (英辞郎)	He could not have to nothing .

表 6.5: 英日統計翻訳結果 (重文複文)

提案手法 (鳥バンク) ○の出力例	
入力文	Many people are discontented with their salaries .
正解文	給料に不満をもっている人は多い。
提案手法 (鳥バンク)	多くの人は給料に不満をもっている。
提案手法 (英辞郎)	給料は満足している者がたくさんある。
提案手法 (英辞郎) ○の出力例	
入力文	He wearied himself by labor .
正解文	彼は労働して疲れた。
提案手法 (鳥バンク)	彼は疲れて仕事だった。
提案手法 (英辞郎)	彼は自分の仕事にうんざりした。
差無し of 出力例	
入力文	When I entered the room , everyone had been waiting for me .
正解文	座敷へ通ると、もうみんな待っていた。
提案手法 (鳥バンク)	私が部屋へ入ると、みんな待った。
提案手法 (英辞郎)	私が部屋に入ったときには、みんながあつたのを待った。

6.2 鳥バンクと英辞郎を組み合わせた翻訳実験

対訳句コーパスとして鳥バンクと英辞郎の両方を用いた翻訳実験を“鳥バンク+英辞郎”と呼ぶ。“鳥バンク+英辞郎”の自動評価結果を表 6.6 に示す。

表 6.6: 自動評価結果

(1) 日英統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1195	4.521	0.6999	0.4923
提案手法 (鳥バンク)	0.1401	4.964	0.7144	0.5189
提案手法 (英辞郎)	0.1379	5.017	0.7173	0.5199
鳥バンク+英辞郎	0.1385	4.973	0.7134	0.5133

(2) 英日統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1620	4.384	0.6241	-
提案手法 (鳥バンク)	0.1701	4.759	0.6371	-
提案手法 (英辞郎)	0.1813	4.795	0.6503	-
鳥バンク+英辞郎	0.1752	4.827	0.6405	-

(3) 日英統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.0852	3.743	0.6420	0.4310
提案手法 (鳥バンク)	0.2193	6.086	0.7101	0.5533
提案手法 (英辞郎)	0.1068	4.318	0.6627	0.4609
鳥バンク+英辞郎	0.2179	6.055	0.7095	0.5517

(4) 英日統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1239	3.953	0.5618	-
提案手法 (鳥バンク)	0.2275	5.895	0.6459	-
提案手法 (英辞郎)	0.1383	4.276	0.5782	-
鳥バンク+英辞郎	0.2307	5.940	0.6483	-

表 6.6 より，単文コーパスの翻訳実験において“鳥バンク+英辞郎”は提案手法 (英辞郎) よりも劣っている。重文複文コーパスの日英翻訳実験において，“鳥バンク+英辞郎”は提案手法 (鳥バンク) よりも劣っている。また，重文複文コーパスの英日翻訳実験において，“鳥バンク+英辞郎”は提案手法 (鳥バンク) よりも優れている結果になった。この結果から，ただ学習データに対訳句コーパスを加えるのではなく，効果のあるフレーズのみを厳選して加える必要があると考える。

6.3 鳥バンクと英辞郎の差

提案手法 (鳥バンク) が提案手法 (英辞郎) よりも翻訳精度が良い原因として，“正しいフレーズテーブルの選択”と“未知語の減少”が考えられる。

6.3.1 正しいフレーズテーブルの選択

正しいフレーズテーブルの選択により，提案手法 (鳥バンク) が提案手法 (英辞郎) よりも翻訳精度が向上した例を示す。

6.3.1.1 例 1

表 6.1 の単文コーパスの日英統計翻訳において，提案手法 (鳥バンク) ○と判断した文を表 6.7 に示す。

表 6.7: 提案手法 (鳥バンク) ○の例 (日英翻訳 単文)

入力文	その映画俳優は婚約を破棄した。
正解文	The movie star called off his engagement .
提案手法 (鳥バンク)	The film star canceled the engagement .
提案手法 (英辞郎)	The movie is young actor broke off .

表 6.7 において，提案手法 (鳥バンク) と提案手法 (英辞郎) の出力文に用いたフレーズテーブルと日本語フレーズと英語フレーズの対応を表 6.8 に示す。

表 6.8: フレーズテーブルと日本語フレーズと英語フレーズの対応

(1) 提案手法 (鳥バンク) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
その 映画俳優 は婚約を破棄した。	The film star canceled the engagement .

(2) 提案手法 (英辞郎) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
その映画 は 婚約 俳優 を破棄した 。	The movie is young actor broke off .

表 6.8 において、提案手法 (鳥バンク) の出力文は鳥バンクに “は 婚約 を 破棄 した 。 canceled the engagement .” の対訳句対が存在し、フレーズテーブルが作成された。

6.3.1.2 例 2

表 6.1 の重文複文コーパスの英日統計翻訳において、提案手法 (鳥バンク) ○と判断した文を表 6.9 に示す。

表 6.9: 提案手法 (鳥バンク) ○の例 1(英日翻訳 重文複文)

入力文	I must be trotting home .
正解文	急いでうちへ帰らないといけない。
提案手法 (鳥バンク)	急いでうちへ帰らなければならない。
提案手法 (英辞郎)	繋駕速歩帰らなければなりません。

表 6.9 において、提案手法 (鳥バンク) と提案手法 (英辞郎) 出力文に用いたフレーズテーブルと日本語フレーズと英語フレーズの対応を表 6.10 に示す。

表 6.10: フレーズテーブルと日本語フレーズと英語フレーズの対応 1

(1) 提案手法 (鳥バンク) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
I must be trotting home .	なければなら ない 急いで うちへ 帰ら 。

(2) 提案手法 (英辞郎) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
I must be trotting home .	なければ なり ませ ん 繫 駕 速 歩 帰 ら 。

表 6.10 において、提案手法 (鳥バンク) の出力文は鳥バンクに “be trotting home 急いで うちへ 帰ら” の対訳句対が存在し、フレーズテーブルが作成された。

6.3.1.3 例 3

表 6.1 の重文複文コーパスの英日統計翻訳において、提案手法 (鳥バンク) ○と判断した文を表 6.11 に示す。

表 6.11: 提案手法 (鳥バンク) ○の例 2(英日翻訳 重文複文)

入力文	He cried so much that his tears went dry .
正解文	あまり 泣いて 涙が 涸れて しまった。
提案手法 (鳥バンク)	彼は あまり 泣き 過ぎて 涙が 涸れて しまった。
提案手法 (英辞郎)	彼は 涙を 流して くらい 泣いた。

表 6.11 において、提案手法 (鳥バンク) と提案手法 (英辞郎) 出力文に用いたフレーズテーブルと日本語フレーズと英語フレーズの対応を表 6.12 に示す。

表 6.12: フレーズテーブルと日本語フレーズと英語フレーズの対応 2

(1) 提案手法 (鳥バンク) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
He cried so much that his tears went dry .	彼は あまり 泣き 過ぎて 涙 が 涸れて しまっ た。

(2) 提案手法 (英辞郎) の出力文に用いられたフレーズテーブル

日本語フレーズ	英語フレーズ
He cried so much that his tears went dry .	彼は 泣い くらい 涙 を 流し た て 。

表 6.12 において, 提案手法 (鳥バンク) の出力文は鳥バンクに “cried so much that あまり 泣き 過ぎて ”と “his tears went dry 涙 が 涸れて しまっ ”の対訳句対が存在し, フレーズテーブルが作成された.

このように, 対訳句コーパスに入力文に対して有効な対訳句対が存在したとき, 翻訳品質は良くなる傾向がある.

6.3.2 未知語を含む文の割合調査

ベースラインよりも提案手法の翻訳精度が向上した原因として，出力文における未知語の減少が考えられる．そこで，各翻訳実験の出力文1万文中に未知語を含む文数を調査した．調査結果を表6.13に示す．

表 6.13: 出力文1万文中に未知語を含む文数

	単文		重文複文	
	日英翻訳	英日翻訳	日英翻訳	英日翻訳
ベースライン	4,676	3,868	5,568	4,455
提案手法 (鳥バンク)	1,516	1,541	563	497
提案手法 (英辞郎)	1,200	1,210	1,819	1,274

表6.13の結果から，どちらの提案手法もベースラインの出力文と比べて未知語を含む文数の減少が確認できた．さらに，単文コーパスの翻訳実験において，提案手法(英辞郎)は最も未知語が減少した．また，重文複文コーパスの翻訳実験において，提案手法(鳥バンク)は最も未知語が減少した．この結果は自動評価結果における単文コーパスを用いた翻訳実験での“ベースライン<提案手法(鳥バンク)<提案手法(英辞郎)”の傾向と，重文複文コーパスを用いた翻訳実験での“ベースライン<提案手法(英辞郎)<提案手法(鳥バンク)”という傾向と同じである．未知語の減少が提案手法の翻訳精度の向上の原因であると考えられる．

6.3.3 未知語の減少

表6.1の重文複文コーパスの英日統計翻訳の提案手法(鳥バンク)○において，未知語の減少によって翻訳品質が向上した文は12文中4文であった．以下に例を示す．

6.3.3.1 例 1

表 6.1 の重文複文コーパスの英日統計翻訳の提案手法 (鳥バンク) ○の出力文を表 6.14 に示す。

表 6.14: 未知語減少例 1

入力文	He volunteered , irrespective of the consequences .
正解文	彼は結果は問題にせず志願した。
提案手法 (鳥バンク)	後先にかまわず自発的に申し出た。
提案手法 (英辞郎)	彼は結果を考えずに volunteered だ。

表 6.14 において、日英対訳コーパスには存在しない “volunteered 自発的に申し出” の対訳句対が鳥バンクに含まれていたため、出力文のフレーズテーブルに用いられた。英辞郎には “volunteered ” の原形である “volunteer ” を含む対訳句対が存在したが、“volunteered” は存在しなかったため、未知語を含む文として出力された。

6.3.3.2 例 2

表 6.1 の重文複文コーパスの英日統計翻訳の提案手法 (鳥バンク) ○の出力文を表 6.15 に示す。

表 6.15: 未知語減少例 2

入力文	At present our M Department Store is handling two world-famous Scotches .
正解文	現在我が M デパートは世界的に有名な 2 種類のスコッチを取り扱っています。
提案手法 (鳥バンク)	私たちの現在の M デパートは世界的に有名なスコッチを扱う。
提案手法 (英辞郎)	M デパートは現在、当社の世界的に有名な 2 Scotches 。

表 6.15 において、日英対訳コーパスには存在しない “Scotches スコッチ” の対訳句対が鳥バンクに含まれていたため、出力文のフレーズテーブルに用いられた。英辞郎には “Scotches” の単数形である “Scotch” を含む対訳句対が存在したが、“Scotches” は存在しなかったため、未知語を含む文として出力された。

6.3.3.3 例3

表 6.1 の重文複文コーパスの英日統計翻訳の提案手法 (鳥バンク) ○の出力文を表 6.16 に示す.

表 6.16: 未知語減少例 3

入力文	Promises are cheap .
正解文	約束 を する だけ なら お 安い 御用 だ 。
提案手法 (鳥バンク)	約束 は 安い 。
提案手法 (英辞郎)	Promises は 安い 。

表 6.16 において, 日英対訳コーパスには存在しない “Promises 約束” の対訳句対が鳥バンクに含まれていたため, 出力文のフレーズテーブルに用いられた. 英辞郎には “Promises” の単数形である “Promise” を含む対訳句対が存在したが, “Promises” は存在しなかったため, 未知語を含む文として出力された.

6.3.3.4 例4

表 6.1 の重文複文コーパスの英日統計翻訳の提案手法 (鳥バンク) ○の出力文を表 6.17 に示す.

表 6.17: 未知語減少例 4

入力文	Hirano fought with the robber at the risk of his life .
正解文	平野 さん は 身 の 危険 を 顧み ず 強盗 と 戦った 。
提案手法 (鳥バンク)	彼は 命 を かけて 強盗 と 戦った 。
提案手法 (英辞郎)	Hirano に 泥棒 を 賭して 戦った 。

表 6.17 において, 日英対訳コーパスには存在しない “Hirano 彼 は” の対訳句対が鳥バンクに含まれていたため, 出力文のフレーズテーブルに用いられた. 英辞郎には固有名詞である “Hirano” が存在しなかったため, 未知語を含む文として出力された.

このように, 英辞郎には多くの動詞や名詞の対訳句対が存在するが, 動詞の活用が原型であったり, 名詞が単数形である場合が多い.

英辞郎の英語フレーズにおいて、動詞を過去形や現在進行形に変化させた対訳句対を追加すれば、翻訳精度は向上すると考える。

また、日本語フレーズにおいても同様に、動詞を活用させることによって未知語が減少し、翻訳精度が向上すると考える。

6.4 自動評価結果と人手評価結果

表 5.5 の自動評価結果と表 6.1 の人手評価結果をまとめる。

- 1 自動評価結果では、単文コーパスを用いた翻訳実験において、提案手法 (英辞郎) は提案手法 (鳥バンク) よりも優れている。重文複文コーパスを用いた翻訳実験において、提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。
- 2 人手評価結果では、単文コーパスを用いた翻訳実験および重文複文コーパスを用いた翻訳実験において、提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。

つまり、自動評価結果と人手評価結果が相違していた。

また、自動評価結果において、重文複文コーパスを用いた翻訳実験において、提案手法 (鳥バンク) は提案手法 (英辞郎) よりも大幅に翻訳精度が向上している。

その原因について考察していく。

自動評価結果と人手評価結果が相違している文の例を 6.4.1 節から 6.4.3 節に示す。重文複文コーパスを用いた翻訳実験の自動評価結果において、提案手法 (鳥バンク) が提案手法 (英辞郎) よりも大幅に翻訳精度している例を 6.4.4 節から 6.4.7 節に示す。

6.4.1 例 1

表 6.1 の単文コーパスの日英統計翻訳において、提案手法 (英辞郎) ○と判断した文と、その文の自動評価結果を表 6.18 に示す。

表 6.18 において、METEOR 以外の自動評価法において、提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。しかし、対比較評価では提案手法 (英辞郎) は提案手法 (鳥バンク) よりも優れている。

BLEU, NIST は正解文と出力文との語順が正しい場合にスコアが高くなる。RIBES は正解文と出力文の間の共通単語の出現順を評価する。正解文と提案手法 (鳥バンク) の出力文には “proposed to her .” が共通している。そのため、BLEU, NIST, RIBES では提案手法 (鳥バンク) が高いスコアを出力してしまった。

表 6.18: 提案手法 (英辞郎) ○と判断した例 (日英翻訳 単文)

		BLEU	NIST	RIBES	METEOR
入力文	彼は彼女にプロポーズした。				
正解文	He proposed to her .				
提案手法 (鳥バンク)	I proposed to her .	0.6687	1.858	0.9457	0.7292
提案手法 (英辞郎)	He made a proposal to her .	0.0000	1.327	0.8694	0.7695

6.4.2 例2

表 6.1 の単文コーパスの英日統計翻訳において、提案手法 (鳥バンク) ○と判断した文と、その文の自動評価結果を表 6.19 に示す。

表 6.19: 提案手法 (鳥バンク) ○と判断した例 (英日翻訳 単文)

		BLEU	NIST	RIBES
入力文	The physician inquired about his condition .			
正解文	医者は彼の容態を尋ねた。			
提案手法 (鳥バンク)	医者は彼の病状はどうかと尋ねた。	0.3066	1.849	0.8409
提案手法 (英辞郎)	医者は彼の条件について尋ねた。	0.4317	2.466	0.9036

表 6.19 において、すべての自動評価法において、提案手法 (英辞郎) は提案手法 (鳥バンク) よりも優れている。しかし、対比較評価では提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。

6.4.3 例3

表 6.1 の重文複文コーパスの日英統計翻訳において、提案手法 (鳥バンク) ○と判断した文と、その文の自動評価結果を表 6.20 に示す。

表 6.20 において、すべての自動評価法において、提案手法 (英辞郎) は提案手法 (鳥バンク) よりも優れている。しかし、対比較評価では提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。

表 6.20: 提案手法 (鳥バンク) ○と判断した例 (日英翻訳 重文複文)

		BLEU	NIST	RIBES	METEOR
入力文	もっとぴったりした機会を待ちましょう。				
正解文	Let 's wait for a more fitting occasion .				
提案手法 (鳥バンク)	I will wait for a more fitting opportunity .	0.4671	2.113	0.9036	0.6772
提案手法 (英辞郎)	Let 's wait for a more tight .	0.6240	2.616	0.9552	0.7383

郎) よりも優れている。

6.4.4 例4

自動評価結果における提案手法 (鳥バンク) の大幅な向上の例を示す。

表 6.1 の日英翻訳の重文複文コーパスで“差無し”と判断した文と，その文の自動評価結果を表 6.21 に示す。

表 6.21: 差無しと判断した例 1(日英翻訳 重文複文)

		BLEU	NIST	RIBES	METEOR
入力文	私は人の物笑いになるのは好まない。				
正解文	I do not like to subject myself to ridicule .				
提案手法 (鳥バンク)	I do not like to be a man of .	0.4111	1.893	0.8801	0.6433
提案手法 (英辞郎)	He does not like to be a man of I scoff .	0.0000	1.301	0.7598	0.5308

表 6.21 において，すべての自動評価法において，提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。特に，BLEU スコアでは提案手法 (鳥バンク) と提案手法 (英辞郎) の差は “0.4111” である。

6.4.5 例5

自動評価結果における提案手法(鳥バンク)の大幅な向上の例を示す。

表 6.1 の日英翻訳の重文複文コーパスで“差無し”と判断した文と，その文の自動評価結果を表 6.22 に示す。

表 6.22: 差無しと判断した例 2(英日翻訳 重文複文)

		BLEU	NIST	RIBES	METEOR
入力文	試験日が迫って、学生達は、懸命に勉強をし始めた。				
正解文	As the exam is coming near , students started studying hard .				
提案手法(鳥バンク)	The day of the students , and started studying hard .	0.2464	2.210	0.8429	0.6226
提案手法(英辞郎)	The days on the test , and began to study hard , the students .	0.0000	1.195	0.4458	0.6242

表 6.22 において，METEOR 以外のすべての自動評価法において，提案手法(鳥バンク)は提案手法(英辞郎)よりも優れている。特に，BLEU スコアでは提案手法(鳥バンク)と提案手法(英辞郎)の差は“0.2464”である。

6.4.6 例6

自動評価結果における提案手法(鳥バンク)の大幅な向上の例を示す。

表 6.1 の英日翻訳の重文複文コーパスで“差無し”と判断した文と，その文の自動評価結果を表 6.23 に示す。

表 6.23 において，すべての自動評価法において，提案手法(鳥バンク)は提案手法(英辞郎)よりも優れている。特に，BLEU スコアでは提案手法(鳥バンク)と提案手法(英辞郎)の差は“0.3586”である。

6.4.7 例7

自動評価結果における提案手法(鳥バンク)の大幅な向上の例を示す。

表 6.23: 差無しと判断した例 3(英日翻訳 重文複文)

		BLEU	NIST	RIBES
入力文	When there was at last a break in the flow of traffic , I crossed the street .			
正解文	やっと車の流れがとぎれたので道路を横断した。			
提案手法 (鳥バンク)	車の流れをしているときには、道路を横断した。	0.3586	2.019	0.9306
提案手法 (英辞郎)	最後のチャンスをしていたときには、車の流れて道を渡った。	0.0000	1.405	0.0000

表 6.1 の英日翻訳の重文複文コーパスで“差無し”と判断した文と、その文の自動評価結果を表 6.24 に示す。

表 6.24: 差無しと判断した例 4(英日翻訳 重文複文)

		BLEU	NIST	RIBES
入力文	Apprehensive of failure , no one will make the attempt .			
正解文	失敗を恐れて指を染める者が無い。			
提案手法 (鳥バンク)	失敗を恐れてだれもした。	0.2527	1.634	0.9036
提案手法 (英辞郎)	Apprehensive の試みは失敗をするだろう。	0.0000	0.9026	0.0000

表 6.24 において、すべての自動評価法において、提案手法 (鳥バンク) は提案手法 (英辞郎) よりも優れている。特に、BLEU スコアでは提案手法 (鳥バンク) と提案手法 (英辞郎) の差は “0.2527” である。

対比較評価ではどちらも差が無いと判断できるが、出力文によっては自動評価結果が過大評価される場合がある。これが重文複文コーパスの翻訳実験の自動評価結果において、提案手法 (鳥バンク) が提案手法 (英辞郎) よりも大幅に翻訳精度が向上した原因ではないかと考えられる。

6.4.8 自動評価結果と人手評価結果の相違のまとめ

文によって6.4.1節から6.4.3節のように、自動評価結果と人手評価結果が相違してしまう。6.4.4節から6.4.7節のように人が見て翻訳品質に差がない文対の一方が過大評価されてしまうことがある。原因として、現在の自動評価法は単語対応で評価してしまうため、出力文と正解文の一部の単語が一致していればスコアが高くなるからではないかと考える。この問題の解決策として、評価方法として比較的信頼性の高い人手での対比較評価を行うようにするか、単語対応での自動評価法だけでなく、人手評価法に近い自動評価法を考案することであると考える。

6.5 英辞郎の日本語動詞活用

自動評価結果と人手評価結果において、提案手法(鳥バンク)と提案手法(英辞郎)の翻訳精度に差があることがわかった。

鳥バンクは重文複文コーパスから抽出した対訳句対であるので、重文複文コーパスとの親和性が高く、翻訳精度が向上しやすいのではないかと考える。一方、英辞郎は辞書のデータであるので、重文複文コーパスとの親和性が低く、翻訳精度の向上の幅が小さいのではないかと考える。

ところで、対訳句コーパスに含まれる動詞句に着目すると、鳥バンクの日本語句は終止形で終わるだけでなく、活用しているものも含まれる。例を表 6.25 に示す。

表 6.25: 鳥バンク動詞の例

日本語句	やっとそこに着い
英語句	I finally arrived
日本語句	とても緊張し
英語句	am very nervous
日本語句	酒を飲ん
英語句	drinking

一方、英辞郎の日本語句は終止形で終わる動詞が多い。例を表 6.26 に示す。

表 6.26: 英辞郎動詞の例

日本語句	セカンド・オピニオンを求める
英語句	ask for second opinions
日本語句	カーテンを引く
英語句	draw a curtain
日本語句	外へ伸ばす
英語句	run out

そのため、翻訳精度に差が出た可能性がある。そこで、英辞郎の動詞句を活用させ、翻訳実験を行う。

6.5.1 実験方法

英辞郎の動詞句を活用させ、翻訳実験を行う。活用させた動詞句を用いた翻訳実験を英辞郎(動詞活用)と呼ぶ。

手順を以下に示す。

- 手順1 日英対訳コーパスを学習データとして言語モデルを作成する
- 手順2 英辞郎の日本語フレーズを MeCab を用いて形態素解析をする
- 手順3 手順2の結果から，フレーズの最後の単語が動詞である対訳句を抽出する
- 手順4 動詞を語幹のみ（なし），未然1，未然2，連用1，連用2，連用3，終止，連体，仮定，命令，可能の11種類にそれぞれに活用させる
- 手順5 日英対訳コーパスに英辞郎と手順4で作成した動詞フレーズを追加する
- 手順6 手順5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7 手順1と手順6で作成したモデルを用いて統計翻訳を行う

手順3で抽出できた動詞フレーズは286,828句であった。手順4で動詞の活用例を表6.27に示し，手順4で作成した対訳句対数を表6.28示す。

表 6.27: 英辞郎の動詞活用例

活用形	動詞フレーズ
基本形	外へ伸ばす
なし	外へ伸ば
未然1	外へ伸ばそ
未然2	外へ伸ばさ
連用1	外へ伸ばし
連用2	外へ伸ばし
連用3	外へ伸ば
終止	外へ伸ばす
連体	外へ伸ばす
仮定	外へ伸ばせ
命令	外へ伸ばせ
可能	外へ伸ばせる

表 6.28: 動詞活用後の対訳句コーパスの数

対訳句コーパス名	対訳句対数
英辞郎	1,350,299
英辞郎 (動詞活用)	3,155,108

6.5.2 実験結果

6.5節で作成した英辞郎 (動詞活用) を用いて翻訳実験を行った。

6.5.2.1 自動評価結果

自動評価結果を表 6.29 に示す.

表 6.29: 自動評価結果

(1) 日英統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1195	4.521	0.6999	0.4923
提案手法 (鳥バンク)	0.1401	4.964	0.7144	0.5189
提案手法 (英辞郎)	0.1379	5.017	0.7173	0.5199
英辞郎 (動詞活用)	0.1280	4.856	0.6986	0.5004

(2) 英日統計翻訳 単文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1620	4.384	0.6241	-
提案手法 (鳥バンク)	0.1701	4.759	0.6371	-
提案手法 (英辞郎)	0.1813	4.795	0.6503	-
英辞郎 (動詞活用)	0.1686	4.551	0.6347	-

(3) 日英統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.0852	3.743	0.6420	0.4310
提案手法 (鳥バンク)	0.2193	6.086	0.7101	0.5533
提案手法 (英辞郎)	0.1068	4.318	0.6627	0.4609
英辞郎 (動詞活用)	0.1038	4.282	0.6560	0.4527

(4) 英日統計翻訳 重文複文コーパス

	BLEU	NIST	RIBES	METEOR
ベースライン	0.1239	3.953	0.5618	-
提案手法 (鳥バンク)	0.2275	5.895	0.6459	-
提案手法 (英辞郎)	0.1383	4.276	0.5782	-
英辞郎 (動詞活用)	0.1317	4.150	0.5654	-

表 6.29 より, 全ての翻訳実験において英辞郎 (動詞活用) は提案手法 (英辞郎) よりも劣っていることがわかる.

6.5.2.2 人手評価結果

提案手法 (英辞郎) と英辞郎 (動詞活用) の出力文からランダムに 100 文ずつを抽出し、対比較評価を行った。人手評価結果を表 6.30 に示す。

表 6.30: 人手評価結果

(1) 日英統計翻訳 単文コーパス

提案手法 (英辞郎) ○	英辞郎 (動詞活用) ○	差無し	同一出力
3	4	91	2

(2) 英日統計翻訳 単文コーパス

提案手法 (英辞郎) ○	英辞郎 (動詞活用) ○	差無し	同一出力
5	10	77	8

(3) 日英統計翻訳 重文複文コーパス

提案手法 (英辞郎) ○	英辞郎 (動詞活用) ○	差無し	同一出力
2	3	90	5

(4) 英日統計翻訳 重文複文コーパス

提案手法 (英辞郎) ○	英辞郎 (動詞活用) ○	差無し	同一出力
8	6	84	2

表 6.30 の結果より、英辞郎 (動詞活用) と提案手法 (英辞郎) の翻訳品質にはほとんど差がないことがわかった。

6.5.3 英辞郎の日本語動詞活用のまとめ

鳥バンクを用いた提案手法の翻訳精度向上の原因として、英辞郎には無い動詞の活用が考えられた。そこで英辞郎において、末尾単語が動詞である日本語句を 11 種類の活用形に活用させ、学習データに追加し翻訳実験を行った。しかし、自動評価結果と人手評価結果から翻訳精度の向上は見られなかった。他の手法として、英辞郎の英語句において、名詞を複数形にする。動詞を過去形、過去完了形、進行形、三人称単数現在形にする。日本語句において、動詞を過去形、進行形にするなどの方法が考えられる。

6.6 今後の課題

今後の研究として以下の2点が挙げられる。

- 鳥バンクと英辞郎の解析
- 新しい自動評価法の検討

6.6.1 鳥バンクと英辞郎の解析

本研究では対訳句コーパスとして鳥バンクと英辞郎の2種類を用いた。単文コーパスを用いた翻訳実験では、対訳句コーパスとして鳥バンクよりも英辞郎を用いた提案手法が自動評価結果において翻訳精度が向上した。一方、重文複文コーパスを用いた翻訳実験では、対訳句コーパスとして英辞郎よりも鳥バンクを用いた提案手法が自動評価結果において翻訳精度が大幅に向上した。

統計翻訳において、翻訳精度は対訳コーパスの多さに依存することから、対訳句対数が多い英辞郎を用いた提案手法の翻訳精度向上に期待したが、重文複文コーパスの翻訳実験においては英辞郎の半数しか対訳句対数を持たない鳥バンクが翻訳精度向上に貢献した。

対訳句対をただ無造作に対訳コーパスに加えるのでは翻訳精度の向上の見込みはなく、入力文に対し、効果的な対訳句対を厳選して加える必要があると考える。

特に重文複文コーパスの実験において、鳥バンクは効果的であったので、英辞郎の対訳句対を鳥バンクの対訳句対の形式に近づければ翻訳精度は向上すると考える。

6.6.2 新しい自動評価法の検討

現在の自動評価方法では、自動評価結果と人手評価結果が相違してしまうことがある。また、人が見て翻訳品質に差がない対比較文の一方が過大評価されてしまうという問題点がある。

原因として、自動評価法は単語対応で評価してしまうため、出力文と正解文の一部の単語が一致していればスコアが高くなるからではないかと考える。この問題の解決策として、評価方法として比較的信頼性の高い人手での対比較評価を行うようにするか、単語対応での自動評価法だけではなく、人手評価法に近い自動評価法を考案することであると考える。

第7章 おわりに

本研究では、提案手法として対訳句コーパスを日英対訳コーパスに追加したコーパスを学習データとして使用し、統計翻訳を行った。対訳句コーパスとして鳥バンクと英辞郎を用い、単文コーパスと重文複文コーパスに対して日英統計翻訳と英日統計翻訳をそれぞれ行った。したがって合計8種類の翻訳実験を行った。

その結果、全ての自動評価結果において、提案手法はベースラインよりも翻訳精度が向上した。また、人手評価法においてもベースラインより提案手法の翻訳精度が良いことを示すことができ、提案手法の有効性を示すことができた。特に重文複文コーパスの翻訳実験において、鳥バンクを用いた提案手法はベースラインや英辞郎を用いた提案手法と比較して大幅に翻訳精度が向上した。

鳥バンクを用いた提案手法の翻訳精度向上の原因として、英辞郎には無い動詞の活用が考えられた。そこで英辞郎の日本語フレーズにおいて、末尾単語が動詞であるフレーズを11種類の活用形に活用させ、学習データに追加し翻訳実験を行った。しかし、自動評価結果と人手評価結果から翻訳精度の向上は見られなかった。

また、自動評価結果と人手評価結果が異なっている翻訳実験について、その出力文を調査した。その結果、新たな自動評価法を考案しなければならないことがわかった。

謝辞

最後に、3年間に渡りご指導いただきました鳥取大学工学部知能情報工学科計算機工学講座C研究室の村田真樹教授，村上仁一准教授，徳久雅人講師そして計算機工学講座C研究室の方々に厚く御礼申し上げます。

本論文をまとめるにあたって，ご指導頂きました松村幸輝教授にお礼申し上げます。

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] Popović Maja, and Ney Hermann “Statistical Machine Translation with a small amount of bilingual training data”, 5th LREC SALTML Workshop on Minority Languages. 2006.
- [2] Francis M Tyers “Rule-based augmentation of training data in Breton-French statistical machine translation”, 13th Annual Conference of the European Association for Machine Translation, pp.213-217. 2009.
- [3] 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.
- [4] 鳥バンク : <http://unicorn.ike.tottori-u.ac.jp/toribank/>
- [5] 英辞郎 : <http://www.alc.co.jp/>
- [6] Franz Josef Och, Hermann Ney, “A Systematic Comparison of Various Statistical Alignment Models”, Computational Linguistics, volume 29, number 1, pp.19-51. 2003.
- [7] Peter F.Brown, Stephen A.Della Pietra, Vincent J.Della Pietra, Robert L.Mercer, “The mathematics of statistical machine translation:Parameter Estimation”, Computational Linguistics, 1993.
- [8] GIZA++ : <http://www.fjoch.com/GIZA++>
- [9] Moses: Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Ondřej 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] NIST: “Automatic Evaluation of Machine Translation Quality Using n-gram Co-Occurrence Statistics” Proceedings of the Human Language Technology Conference (HLT), pp.128-132. 2002.
- [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] 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.
- [14] MeCab: Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis”, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237, 2004.
- [15] tokenizer.sed
<http://www.cis.upenn.edu/treebank/tokenizer.sed>
- [16] SRILM: Andreas Stolcke, “SRILM - an Extensible Language Modeling Toolkit”, 7th International Conference on Spoken Language Processing, pp.901-904, 2002.
- [17] Mert: Franz Josef Och: “Minimum Error Rate Training in Statistical Machine Translation”, In Proceeding of the 41st Annual Meeting of the Association for Computational Linguistics, pp.160-167, 2003.
- [18] 鏡味良太, 村上仁一, 徳久雅人, 池原悟, “統計翻訳における人手で作成された大規模フレーズテーブルの効果”, 言語処理学会第 14 回年次大会, pp.224-227, 2008.