Vol.2010-MPS-80 No.9
2010/9/28

同義語情報を用いた確率的単語アライメントモデル

進藤裕之[†1]　藤野昭典[†1]　永田昌明[†1]

二言語間の教師なし単語アライメント問題に対して，単言語リソースである同義語辞書情報を利用して単語対応付けの精度を向上させる手法を提案する．対訳文には同じ意味を表す様々な表現が用いられるため，同義語情報を利用することでデータスパースネスの問題を解消し単語アライメントの精度向上が期待できる．しかし，単語には多義性があり，ある単語ペアが同義語であるかどうかは文脈に大きく依存する．そこで，我々はトピックモデルを利用して，同義語情報を文脈に応じて学習させる同義語の確率モデルを考案する．さらに，同義語モデルを既存の単語アライメントモデルと同時に学習させる枠組みを提案する．対訳コーパスを用いたアライメント実験の結果，同義語情報を用いない場合や，同義語情報を文脈を考慮せずに同義語情報を利用した場合に比べて，提案手法では高い精度が得られることを確認した．

Word Alignment with Synonym Information

Hiroyuki Shindo,[†1] Akinori Fujino[†1]
and Masaaki Nagata[†1]

We present a novel framework for word alignment that incorporates monolingual synonym knowledge to improve word alignment performance. We expect synonym information is helpful to overcome the data sparseness problem of word alignment since there are various lexical forms represent the same meaning in a bilingual corpus. However, synonym relations depend heavily on context or domain since a word in natural language is ambiguous. We design a synonym probabilistic model with a topic model, which uses synonym information according to the context. Moreover, we propose a word alignment framework that jointly trains our synonym model and conventional bilingual model. The experimental results show that our proposed method obtained better results compared to cases where synonym or context information is not used.

1. はじめに

単語アライメント問題は，対訳コーパスが与えられたときに，異なる言語間における単語の対応関係を推定する問題であり，現在のフレーズベースおよび文法ベースの統計的機械翻訳において最も基本的なタスクの一つである．単語アライメントの精度が高ければ，より良い翻訳モデルを構築することができるため，高精度な機械翻訳の実現を期待できる．

これまでに，対訳コーパスの生成モデルに基づく教師なし学習[4),10),13)–15)]や，識別学習に基づく教師あり学習[5),9),12)]など様々な単語アライメント確率モデルが提案されてきた．教師あり学習に基づくアライメント手法は，一定量の人手によりタグづけされた正解単語アライメントが必要となるが，現状では多くの言語対において正解単語アライメントデータを入手することは困難であるため，本論文では教師なし学習に基づく単語アライメント推定手法に焦点を当てる．

統計的機械翻訳で用いられる代表的な教師なし単語アライメントモデルとして，IBM model 1-5[2)]や HMM[13)]がある．また，これらを改良したり付加情報を加えることで単語アライメントの精度を向上させる手法が数多く提案されている．例えば，単言語の知識を利用して原言語と目的言語の単語を機能語か内容語かに分類し，機能語同士または内容語同士を対応させやすくする手法がある[3)]．その他にも，単語間における文法的な依存関係は原言語側と目的言語側の双方で保存されている可能性が高いため，そのような文法的な知識を確率モデルに組み込むことでアライメント精度を向上させるものも存在する[6)]．このような言語の文法的知識はアライメントモデルに対して制約条件として機能し，単語対応の過学習を避けてアライメントの精度を向上させることができる．

一方，現在では自然言語処理に有用な多くの語彙的，意味的言語リソースが利用可能である．例えば WordNet[8)]は 50ヶ国以上の言語で構築されているシソーラスであり，同義語，反意語，上位語や下位語など単語の意味的な関係が記述されている．本論文ではそれらの言語資源のうち，同義語の情報を単語アライメントに利用することを考える．同義語は異なる言語の同じ単語へ対応する傾向にあるため，同義語情報を学習時に活用することで単語アライメントの推定精度の向上が期待できる．例えば，“ 二酸化炭素 ”と“ 炭酸ガス ”は同義語のペアであり，同じ英単語“ carbon dioxide ”に対応することが期待されるため，同義語の

†1 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation

Vol.2010-MPS-80 No.9
2010/9/28

情報は正しい単語対応関係を推定するのに有効である．

しかし，一般的に単語の同義関係は単語の表層的な形式ではなく，文脈に大きく左右される．例えば,“ head” と“ forefront” は物理的な位置を表す場合はどちらも「先頭」の意味を表す名詞である．また,“ head ”と“ chief ”は会社などの話題で用いられる場合はどちらも「集団の長」の意味を表す同義語ペアである．しかし,“ forefront” と“ chief” はおそらくどの文脈でも同義語ペアではないであろう．以上のことから，同義語の情報を適切に利用するためには，文脈から単語の語義を推定し，複数の同義語候補の中から正しいものを選択して利用する必要がある．

我々は，教師なし単語アライメント学習のために，語彙的，意味的な言語リソースから収集された同義語情報を利用する新たな確率モデルを提案する．提案法では，文脈に応じて同義語情報を利用するために，トピック変数を導入して同義語ペアの確率モデルを構築し，それを対訳コーパスの生成モデルに基づく単語アライメントモデルと統合する．本論文では，単語アライメントモデルとして，HMM にトピックモデルを取り入れた HM-BiTAM[15] を利用する．本手法を英語とフランス語の単語アライメントタスクへ適用し，アライメント精度が向上することを示す．

2. 単語アライメントモデル

本章では，トピックモデルと HMM に基づく単語アライメントモデルである HM-BiTAM を概説する．

HM-BiTAM は，二言語対訳コーパスの生成モデルであり，潜在変数としてトピック z，アライメント a およびトピック分布ベクトル θ を有する．トピック変数とは，例えば“ 科学 ”,“ ニュース ”，または“ 医療 ”など文の話題を表す変数であり，各対訳文に対して 1 つ割り当てられる．トピック変数によって対訳文の話題が特定されることにより，単語の語義曖昧性解消に有効である．アライメント変数は，目的言語の各単語がどの原言語の単語と対応関係にあるかを表す変数である．トピック分布ベクトルとは，各トピックがどれくらい出現しやすいかという確率をベクトル形式で表現したものである．以下に，HM-BiTAM の生成過程を示す．

(1) $\boldsymbol{\theta} \sim \mathrm{Dirichlet}(\boldsymbol{\alpha})$： ディリクレ分布に従ってトピック分布ベクトルを生成する．

(2) 各対訳文ペア (E_n, F_n) について

(a) $z_n \sim \mathrm{Multinomial}(\boldsymbol{\theta})$： 多項分布に従ってトピック z_n を生成する．

(b) 原言語の各単語位置 $i_n = 1, \ldots, I_n$ について

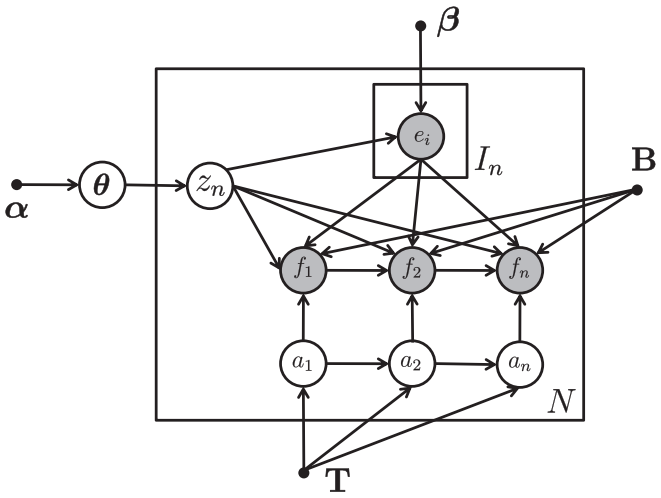

図 1 HM-BiTAM のグラフィカルモデル
Fig. 1 Graphical model of HM-BiTAM

(i) $e_{i_n} \sim p(e_n | z_n; \boldsymbol{\beta})$： トピックに依存した原言語のユニグラムモデルに従って原言語の単語 e_{i_n} を生成する．

(c) 目的言語の各単語位置 $j_n = 1, \ldots, J_n$ について

(i) $a_{j_n} \sim p(a_{j_n} | a_{j_n - 1}; \mathbf{T})$： 一次マルコフモデルに従ってアライメント変数 a_{j_n} を生成する．

(ii) $f_{j_n} \sim p\left(f_{j_n} \middle| e_{a_{j_n}}, z_n; \mathbf{B}\right)$： ポジション j_n に対応する原言語の単語 $e_{a_{j_n}}$ とトピック z_n に依存した単語翻訳モデルに従って目的言語の単語 f_{j_n} を生成する．

ただし，アライメント変数 $a_{j_n} = i_n$ は原言語の単語 e_{i_n} と目的言語の単語 f_{j_n} が対応関係にあることを表す．$\boldsymbol{\alpha}$ はトピック分布ベクトル $\boldsymbol{\theta}$ の確率モデルのパラメータ，$\boldsymbol{\beta} = \{\beta_{k,e}\}$ は原言語の単語出現確率分布である．原言語の単語出現確率分布はトピック k に依存し，$\beta_{k,e}$ は $p(e | z = k)$ に相当する．$\mathbf{B} = \{B_{k,e,f}\}$ は，トピック k の下で原言語の単語 e から目的言語の単語 f への単語翻訳確率を表す．すなわち，$B_{k,e,f}$ は $p(f | e, z = k)$ に相当する．$\mathbf{T} = \{T_{i,i'}\}$ は単語の位置 i から i' への遷移確率である．HM-BiTAM では，従来の HMM 単語アライメントモデルと同様に，アライメント変数が一次マルコフモデルに従うと仮定し

ている．図 1 に HM-BiTAM のグラフィカルモデルを示す．

HM-BiTAM では，対訳文の確率 $p(\mathbf{E},\mathbf{F})$ を以下のようにモデル化する．

$$p(\mathbf{E},\mathbf{F};\mathbf{\Phi}) = \sum_{\mathbf{z}}\sum_{\mathbf{a}}\int p(\{E_n\},\{F_n\},\mathbf{z},\mathbf{a},\boldsymbol{\theta})\,d\boldsymbol{\theta} \tag{1}$$

ただし，$\mathbf{\Phi} = \{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{T},\mathbf{B}\}$ は HM-BiTAM のパラメータセットである．

3. 提案モデル

3.1 同義語データ確率モデル

本章では，提案法で用いる目的言語の同義語ペア $\{f,f'\} = \{(f_m,f'_m)\}_{m=1}^{M}$ の確率モデル $p(\{f,f'\})$ について述べる．自然言語の単語には多義性があるため，言語リソースなどから収集された同義語ペアの同義関係は常に成立するものではなく，文脈に大きく依存する．

ここで，同義語ペアはある共通の“ 意味 ”s という条件の下で独立に生成されると仮定すると，同義語ペアの生成確率を

$$p(f,f') \propto \sum_{m}^{M} p(f|s)\,p(f'|s)\,p(s) \tag{2}$$

とモデル化できる．

本研究では，同義語情報を単語アライメントの学習に利用するために，原言語の単語を利用して目的言語の同義語をモデル化することを考える．例えば，目的言語（日本語）の“ 金星 ”と“ 明星 ”はどちらも原言語（英語）の“ Venus ”に対応するので，原言語の単語を用いて同義語の語義を表現することができる．しかし，Venus には“ 女神 ”の意味もあるように単語には多義性があり，文脈に応じて単語の意味が変わる．そこで，提案手法では，単語の多義性に対処するためにトピック z を導入し，原言語の単語 e とトピック z の組み合わせ (e,z) で目的言語の同義語の語義が定まると考える．そして，その語義に対して同義語のペアが生成されると仮定する．この仮定の下では，同義語のペア集合の生成確率を以下のようにモデル化できる．

$$p(\{f,f'\}) \propto \prod_{(f,f')}\sum_{e,z} p(f|e,z)\,p(f'|e,z)\,p(e,z) \tag{3}$$

図 2 に同義語データモデルのグラフィカルモデルを示す．

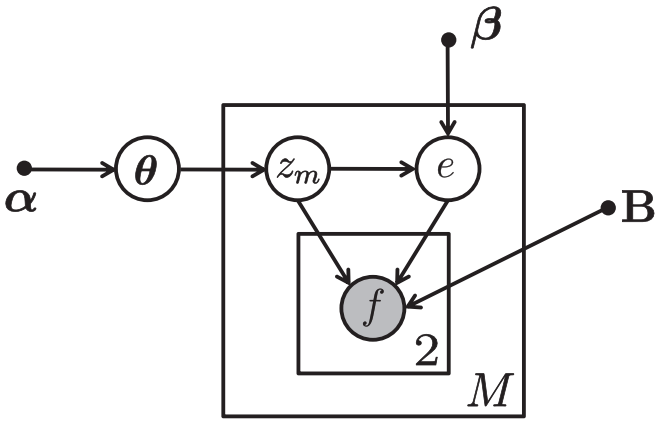

図 2 同義語モデルのグラフィカルモデル
Fig. 2 Graphical model of synonym pair model

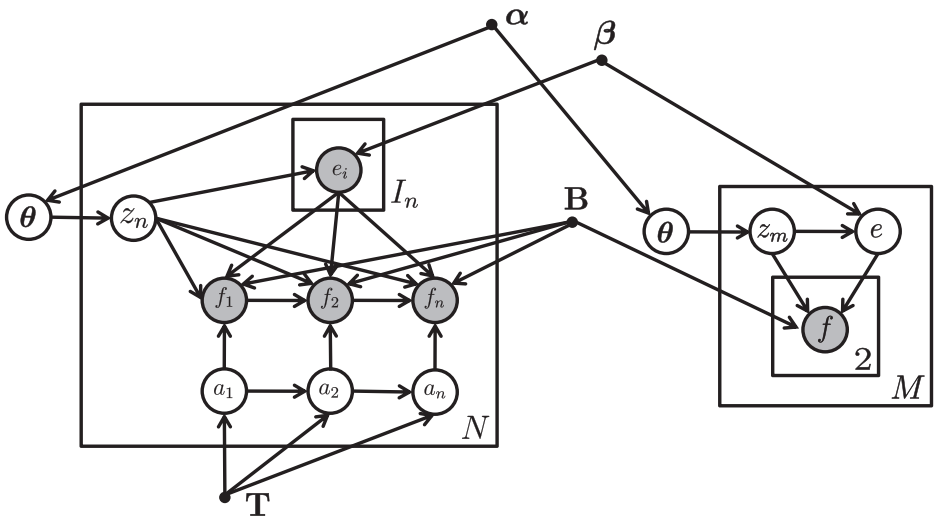

図 3 パラメータを共有した提案手法のグラフィカルモデル
Fig. 3 Graphical model of proposed method

3.2 同義語モデルと統合された単語アライメントモデル

提案手法では，上述の HM-BiTAM と同義語モデルを統合したモデルを用いて単語アライメントを学習させる．この統合は，HM-BiTAM と同義語モデルに含まれる目的言語と原言語の確率モデルおよびトピックの確率モデル $p(f|e,z)$，$p(e|z)$，$p(z)$ のパラメータを互いに共有することにより実現する．具体的には，式 3 の同義語モデル $p(\{f,f'\})$ を，HM-BiTAM と同じパラメータセット $\mathbf{\Phi}=\{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{T},\mathbf{B}\}$ を用いて以下のようにパラメータ化すればよい．

$$p(f|e,z=k) \equiv p(f|e,z=k;\mathbf{B}) = B_{k,e,f}, \tag{4}$$

$$\begin{aligned} p(e,z=k) &\equiv p(e|z=k;\boldsymbol{\beta})\,p(z=k;\boldsymbol{\alpha}) \\ &= \beta_{k,e}\int \theta_k p(\boldsymbol{\theta};\boldsymbol{\alpha})\,d\boldsymbol{\theta}. \end{aligned} \tag{5}$$

したがって，同義語ペアの確率は以下のように書ける．

$$\begin{aligned} p(f,f';\mathbf{\Phi}) &\propto \sum_{e,z}\int p(f|e,z)\,p(f'|e,z)\,p(e|z)\,p(z|\boldsymbol{\theta})\,p(\boldsymbol{\theta})\,d\boldsymbol{\theta} \\ &= \sum_{k,e}\beta_{k,e}B_{k,e,f}B_{k,e,f'}\int p(z=k|\boldsymbol{\theta})\,p(\boldsymbol{\theta})\,d\boldsymbol{\theta} \\ &= \frac{1}{\sum_{k'}\alpha_{k'}}\sum_{k,e}\alpha_k\beta_{k,e}B_{k,e,f}B_{k,e,f'} \end{aligned} \tag{6}$$

図 3 に HM-BiTAM と統合した同義語モデルのグラフィカルモデルを示す．提案手法では，HM-BiTAM の対数周辺尤度と同義語モデルの対数周辺尤度を同時に最大化させる $\mathbf{\Phi}$ を推定値 $\hat{\mathbf{\Phi}}$ とする．

$$\hat{\mathbf{\Phi}} = \operatorname*{argmax}_{\mathbf{\Phi}} \left\{ \log p(\mathbf{E},\mathbf{F};\mathbf{\Phi}) + \zeta \log p\left(\left\{f,f'\right\};\mathbf{\Phi}\right)\right\} \tag{7}$$

ただし，ζ は同義語モデルの重みを調節するハイパーパラメータである．推定値 $\hat{\mathbf{\Phi}}$ の学習法は，変分アルゴリズム[1] を用いる．具体的な更新式は，スペースの都合上省略する．

4. 実　　験

4.1 実験設定

提案手法の評価を行うため，Hansards データセット[7] を用いて単語アライメントの実験を行った．Hansards データセットは，英語とフランス語の二言語対訳コーパスで，規模は１００万文以上である．本データセットのうち，４４７文は人手による正解単語アライメント情報が付与されている．我々は，この４４７文の中からランダムに選択した１００文を開発用データセット，残りの３３７文を評価用テストデータとした．開発用データセットは重み ζ を最適化するために用いた．トレーニングデータは以下のように構成した．まず，評価用テストデータから正解単語アライメント情報を削除した３３７文の対訳文と，人手で正解の付与されていない残りの対訳文の中からランダムに 10k，50k，100k の文を選択し，これを混合したものをトレーニングデータとした．したがって，トレーニングデータには必ず評価用テストデータの対訳文が含まれている．このトレーニングデータを用いて単語アライメントの教師なし学習を行い，３３７文の評価用対訳文の推定結果と正解を比較して単語アライメントの精度を評価した．したがって，トレーニングに正解単語アライメントの情報は一切使用していない．

英語およびフランス語の同義語辞書は，それぞれ WordNet 2.1[8] および Wolf 0.1.4[11] から収集した．WordNet は英語の意味的な概念を扱う言語リソースで，単語が synset と呼ばれる同義語のグループに分類されており．同義語ペアのデータを得ることができる．WOLF は WordNet やその他の各種言語リソースから構築されたフランス語の WordNet である．我々は，これらのリソースから得られた同義語ペアのうち，いずれの単語もトレーニングデータ中に含まれる場合のみ学習に使用した．

我々は，GIZA++ 1.0.3[10]，HM-BiTAM および提案手法で単語アライメント精度の比較を行った．ただし，HM-BiTAM は我々が独自に実装したものを用いている．GIZA++は，IBM model-4 による単語アライメントであり，HM-BiTAM は式 7 で $\zeta=0$ に相当する．

IBM model，HM-BiTAM や提案手法のような雑音のある通信路モデルに基づく単語アライメントモデルでは，原言語と目的言語を入れ替えることにより二方向の単語アライメント結果が得られる．本実験では，英語を原言語，フランス語を目的言語とした場合と，原言語をフランス語，目的言語を英語とした場合の二方向の結果を“ GROW ”ヒューリスティクス[10),14)] を用いて統合し，一方向よりも高精度かつロバストな予測単語アライメントを得た．

本データセットには，S（sure）または P（probable）の二種類の正解単語アライメントのラベルが人手によって付与されている．S アライメントは，確実に対応関係である単語ペアに対して付与されたアライメント情報であり，P アライメントはそれ以外の（不確実な）アライメントである．予測した単語アライメントの精度は，Precision，Recall，F-measure，AER を用いて評価した．これらの尺度は，単語アライメント問題で標準的に用いられる評

10k		Precision	Recall	F-measure	AER
GIZA++	standard	0.856	0.718	0.781	0.207
	with SRH	0.874	0.720	0.789	0.198
HM-BiTAM	standard	0.869	0.788	0.826	0.169
	with SRH	0.884	0.790	0.834	0.160
Proposed		**0.941**	**0.808**	**0.870**	**0.123**

(a)

50k		Precision	Recall	F-measure	AER
GIZA++	standard	0.905	0.770	0.832	0.156
	with SRH	0.903	0.759	0.825	0.164
HM-BiTAM	standard	0.901	0.814	0.855	0.140
	with SRH	0.899	0.808	0.853	0.145
Proposed		**0.947**	**0.824**	**0.881**	**0.112**

(b)

100k		Precision	Recall	F-measure	AER
GIZA++	standard	0.925	0.791	0.853	0.136
	with SRH	**0.934**	0.803	0.864	0.126
HM-BiTAM	standard	0.898	0.851	0.874	0.124
	with SRH	0.909	0.860	0.879	0.114
Proposed		0.927	**0.862**	**0.893**	**0.103**

(c)

表 1 単語アライメント精度の比較．トレーニングデータのサイズはそれぞれ (a) 10k，(b) 50k，(c) 100k.
Table 1 Comparison of word alignment accuracy. The best results are indicated in bold type. The training data set sizes are (a) 10k, (b) 50k, (c) 100k.

価基準である[10]．

4.2 結　　果

表１は，10k，50k，100k のトレーニングデータで学習された単語アライメントの精度を示している．提案手法がＦ値とＡＥＲで最も良い性能を示している．この結果から，我々の同義語データを利用した単語アライメント手法は，精度を向上させることに効果的であることがわかる．

# vocabularies		10k	50k	100k
English	standard	8578	16924	22817
	with SRH	5435	7235	13978
French	standard	10791	21872	30294
	with SRH	9737	20077	27970

表 2 10k，50k，100k のトレーニングデータの語彙数
Table 2 The number of vocabularies in the 10k, 50k and 100k data sets.

# synonyms	10k	50k	100k
English	7756	17273	23187
French	1677	2524	2980

表 3 10k，50k，100k のトレーニングデータの同義語ペア数
Table 3 The number of synonym pairs in the 10k, 50k and 100k data sets.

前述のように，我々の主なアイディアは，トピック変数と，異なる言語の単語を用いて同義語ペアの精密なモデル化を行うというものである．言語リソースから収集された同義語ペアが同義語であるかは，文脈に大きく依存する．この問題に対処するため，我々はトピック変数を導入した同義語ペアのモデル化を行い，単語の語義曖昧性を解消しつつ対訳文の文脈に応じた同義語ペアを単語アライメント学習に利用可能となった．本モデルの効果を検証するために，我々は同義語データを単語アライメント学習に利用する単純なヒューリスティクス（SRH: Synonym Replacement Heuristics）を用いてテストを行った．SRH は，同義語データ中に含まれる同義語ペアを利用し，トレーニングデータ中の単語を片方の同義語に置き換えるというヒューリスティクスである．例えば，単語 A と単語 B が同義語ペアである場合，トレーニングデータ中の全ての単語 B は単語 A に置換される．SRH では，“ head ”のように複数の同義語ペアをもつ単語は，どの同義語ペアに置き換えられるかはランダムに決定される．したがって，文脈に応じて正しく単語が同義語に置換された場合，単語アライメントの精度向上が期待できるが，そうでない場合は逆に精度を悪化させる恐れがある．表２に示すように，SRH によりトレーニングセットにおける英語とフランス語の語彙数は期待通り大きく減少した．

SRH を実行した後，GIZA++および HM-BiTAM の単語アライメント精度を検証した．SRH は，10k と 100k のデータセットでは精度が若干向上したが，50k のデータセットでは

精度が悪化した．これは，同義語ペアの情報を誤った文脈で利用してしまったことにより，単語アライメントの精度が悪化してしまった影響であると考えられる．

5. 結　　論

我々は，同義語の情報を教師なし単語アライメントモデルへ利用する枠組みを提案した．ある単語ペアが同義語であるかどうかは文脈に大きく依存するため，我々はトピックモデルを利用して文脈を特定し，単語の語義曖昧性を解消しながら同義語ペアの確率モデルを考案した．また，同義語モデルのパラメータを二言語の単語アライメントモデルと同時に学習することで，同義語の情報を単語アライメントへ利用する枠組みを提案した．我々の手法は，二言語の対訳情報と単言語の同義語情報を効率的に利用し，教師なし単語アライメントの精度を向上させた．今後は，本手法を異なる言語間での単語アライメント問題へ適用することや，統計的機械翻訳へ応用することが考えられる．

参 考 文 献

1) Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A. F. M. and West, M.: The variational Bayesian EM algorithm for incomplete data: with application to scoring graphical model structures, *Bayesian Statistics 7: Proceedings of the 7th Valencia International Meeting, June 2-6, 2002*, Oxford University Press, USA, p.453 (2003).
2) Brown, P.F., DellaPietra, V.J., DellaPietra, S.A. and Mercer, R.L.: The mathematics of statistical machine translation: Parameter estimation, *Computational linguistics*, Vol.19, No.2, pp.263–311 (1993).
3) Deng, Y. and Gao, Y.: Guiding Statistical Word Alignment Models With Prior Knowledge, *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Prague, Czech Republic, Association for Computational Linguistics, pp.1–8 (2007).
4) Fraser, A. and Marcu, D.: Getting the structure right for word alignment: LEAF, *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, Prague, Czech Republic, Association for Computational Linguistics, pp. 51–60 (2007).
5) Haghighi, A., Blitzer, J., DeNero, J. and Klein, D.: Better Word Alignments with Supervised ITG Models, *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, Suntec, Singapore, Association for Computational Linguistics, pp.923–931 (2009).
6) Ma, Y., Ozdowska, S., Sun, Y. and Way, A.: Improving Word Alignment Using Syntactic Dependencies, *Proceedings of the ACL-08:HLT Second Workshop on Syntax and Structure in Statistical Translation (SSST-2)*, Columbus, Ohio, Association for Computational Linguistics, pp.69–77 (2008).
7) Mihalcea, R. and Pedersen, T.: An evaluation exercise for word alignment, *Proceedings of the HLT-NAACL 2003 Workshop on building and using parallel texts: data driven machine translation and beyond-Volume 3*, Association for Computational Linguistics, p.10 (2003).
8) Miller, G. A.: WordNet: a lexical database for English, *Communications of the ACM*, Vol.38, No.11, p.41 (1995).
9) Moore, R.C., Yih, W.-t. and Bode, A.: Improved Discriminative Bilingual Word Alignment, *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, Association for Computational Linguistics, pp.513–520 (2006).
10) Och, F.J. and Ney, H.: A systematic comparison of various statistical alignment models, *Computational Linguistics*, Vol.29, No.1, pp.19–51 (2003).
11) Sagot, B. and Fiser, D.: Building a free French wordnet from multilingual resources, *Proceedings of Ontolex* (2008).
12) Taskar, B., Simon, L.-J. and Dan, K.: A Discriminative Matching Approach to Word Alignment, *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, Vancouver, British Columbia, Canada, Association for Computational Linguistics, pp.73–80 (2005).
13) Vogel, S., Ney, H. and Tillmann, C.: HMM-based word alignment in statistical translation, *Proceedings of the 16th Conference on Computational Linguistics-Volume 2*, Association for Computational Linguistics Morristown, NJ, USA, pp. 836–841 (1996).
14) Zhao, B. and Xing, E. P.: BiTAM: Bilingual topic admixture models for word alignment, *Proceedings of the COLING/ACL on Main Conference Poster Sessions*, Association for Computational Linguistics, p.976 (2006).
15) Zhao, B. and Xing, E.P.: HM-BiTAM: Bilingual Topic Exploration, Word Alignment, and Translation, *Advances in Neural Information Processing Systems 20*, Cambridge, MA, MIT Press, pp.1689–1696 (2008).