Vol.2013-NL-214 No.6
2013/11/14

係り受け関係の階層化とその共起に基づいた構文木モデルを利用した構文解析手法の提案

大野 一樹[1,a)]　波多野 賢治[2,b)]

概要：本稿では係り受け関係の階層化とその共起頻度を素性とした構文木モデルを生成し，これを利用した構文解析手法を提案する．我々は以前に文末の文節を根とし，文末の文節から他の文節へとコンテキストを辿る係り受け関係の階層化に基づいた n-gram ベースの構文木モデルを生成した上で日本語の構文解析を行ってきた．しかし，この手法は文節を構成する形態素の品詞を素性としており，単一の係り受け木の生起頻度を考慮してとしているだけのため，係り受け関係の生起に考慮されるべきコンテキストが不足しており，全体の精度としては従来手法と比較して優れた結果を得ることができなかった．そのため，係り受け関係の階層化だけでなくその共起頻度に基づいた構文木モデルを利用して構文解析を行うことで解析精度の向上を行う．

1. はじめに

日本語や中国語のような格文法を基本文法とした言語における構文解析は，文節間の修飾関係や意味的役割を表現する係り受け関係の取得である．そのため，統計的手法に基づいた既存の日本語構文解析手法は文節間に係り受け関係が成立するかどうかを SVM を用いて構文木モデルを構築しており，そうして生成された構文木モデルを用いて構文解析を行っている [1]．この工藤らの手法はすべての文節に着目し，直後の文節に対して係り受け関係を持つかどうかというアルゴリズムを繰り返すことで構文解析を行っている．そのため，ある文節が直後の文節に対して強い係り受け関係を持つとすると，その文節に対して係り受け関係を生じやすく，結果として同じ格を持つ文節が連続して出現するような複雑な構造を持つ文に対して誤解析が生じる．ゆえに，充分な構文解析精度を得られないといった課題がある．

英語に対する構文解析においては木置換文法 (TSG: Tree Substitution Grammar) [2], [3] を利用した構文解析手法が注目されている．構文木モデル生成の際に， TSG を利用することで，構文木の深さをコンテキストとして考慮した構文木モデルを構築することができ，これにより弱文脈依存性を取り入れた構文解析を行うことができる．そのため， TSG を利用した構文解析手法は現存する英語の構文解析手法の中でも最高精度の構文解析精度を実現している [4], [5]．

このことから，日本語構文木モデルにおいても同様に係り受け関係を持つ文節間の周辺にある文節をコンテキストとして考慮した構文木モデルを構築し，これを構文解析に利用することで，工藤らの構文解析手法における誤解析の防止を図ることが可能であると考えられる．

そこで，我々は日本語の語順が自由であるという特徴を保ちながら構文木を係り受け関係の階層化に基づいて構築し，そのモデルに基づいた構文解析手法の提案を行った [6]．提案した構文木モデルは係り受け関係を持つ任意の長さの文節を単位とした構文木モデルであり，係り受け関係を持つ文節間の前後の文節をコンテキストとして構文木モデルに取り込んでいる．これにより，工藤らの構文解析手法の問題点であった複雑な構造を持つ文に対する構文解析の問題点を解決することができた．しかし，構文木モデル構築の際の素性の不足から京都大学テキストコーパスに基づいた構文解析精度は工藤らの手法に比べると，劣る方法となっている．

そのため，本稿ではコンテキストの共起を新たな素性として構文木モデルに取り込み，これを利用した構文解析手法を提案し構文解析精度の向上を図る．

2. 基本的事項

本節では， Pitman-Yor 過程と 3 節で述べる木置換文法を利用した構文木モデルの生成および提案する構文木モデル構築の際に利用する階層 Pitman-Yor 過程について述

1 同志社大学大学院文化情報学研究科
2 同志社大学文化情報学部
a) ohno@ilab.doshisha.ac.jp
b) khatano@mail.doshisha.ac.jp

べる．

2.1 Pitman-Yor 過程

ノンパラメトリックベイズモデルの一つである Pitman-Yor 過程 [7] は，自然言語処理の n -gram 分布を扱う際の利用がその一つの例として存在する．観測された単語集合 W に含まれる単語 w が出現する n-gram 分布を生成する確率過程 G を Pitman-Yor 過程 $PY(d,\theta,G_0)$ を用いて式 (1) に表すことができる．

$$G \sim PY(d,\theta,G_0) \tag{1}$$

このとき， Pitman-Yor 過程の三つのパラメータ d,θ,G_0 について， d は $0 \leq d \leq 1$ の範囲を取り，観測度数を実際の度数よりも低く見積もるために用いられるディスカウント項と呼ばれるパラメータ，θ は Pitman-Yor 過程によって生成される確率過程の基底分布 $G_0 = [G_0(w)]_{w \in W}$ への依存の強さを示すパラメータ，G_o は基底分布であり，一般的に一様分布が用いられる．

Pitman-Yor 過程はディリクレ過程にディスカウント項 d のパラメータを加えたディリクレ過程の拡張である．そのため， Pitman-Yor 過程も ディリクレ過程と同様，無限次元のディリクレ分布を生成することのできる確率過程，すなわち加算無限性をサポートした確率過程である．Pitman-Yor 過程において $d = 0$ のとき， Pitman-Yor 過程は 式 (2) のディリクレ過程 $DP(\theta,G_0)$ と等価である．

$$G \sim DP(\theta,G_0) = \theta G_0 \tag{2}$$

式 (2) におけるディリクレ過程 G は，観測された出現単語の事象空間である r の大きさの離散空間の任意の分割に対して式 (3) で表されるディリクレ分布 $Dir(\theta G_0(w_1),\cdots,\theta G_0(w_r))$ に近似する．

$$(G(w_1),\cdots,G(w_r)) \sim Dir(\theta G_0(w_1),\cdots,\theta G_0(w_r)) \tag{3}$$

一般的にこのような確率過程を用いて n-gram 分布 G を生成するための処理として，中華料理店過程 [8] が利用される．n-gram 分布 G に存在する単語を客とし，その客が順に無限の客が座ることのできる無限のテーブル (G_0 に相当) のある店に入店してくる．1 番目の客は 1 番目のテーブルに着席する．そして，続く客は 1 番目の客と同じ単語であればすでに人が座っているテーブルに座り， G_0 において新しく観測された単語であれば新しいテーブルに座る．なお，新たに客が入店してきてすでにテーブルが存在する場合，$\theta + d \cdot (\text{テーブルの総数}) * G_0(w_k)$ の確率で新しいテーブルに座る．この中華料理店過程に基づいて観測した単語のテーブル数について t とし，単語 w_k の出現頻度を c_k， Pitman-Yor 過程 G によって生成される分布の全単語数を $c.$ として，式 (4) に Pitman-Yor 過程の導出を表すことができる．

$$PY(d,\theta,G_0) = \frac{c_k - d}{\theta + c_.} + \frac{\theta + dt}{\theta + c_.} G_0(w_k) \tag{4}$$

なお， Pitman-Yor 過程 に対して付与されるパラメータ d,θ はノンパラメトリックであり，G_0 に近似するような分布をマルコフ連鎖モンテカルロ法の一種であるギブスサンプリング を用いてパラメータを収束させて推定を行うものである．

2.2 階層 Pitman-Yor 過程

階層 Pitman-Yor 過程 [9] は Pitman-Yor 過程を階層化した確率過程である．観測された単語の n-gram 分布を基底分布とする Pitman-Yor 過程を再帰的に計算することにより階層化を行う．階層 Pitman-Yor 過程では $n-1$ の単語長によって構成されるコンテキスト $\mathbf{u}$ の n-gram 分布を事前分布とした n-gram 分布を生成することが可能となる．n-gram 分布 G_u はコンテキスト u の元で出現する単語の n-gram 分布である $G_{\pi(\mathbf{u})}$ を基底分布とした Pitman-Yor 過程によって下記のように生成される．

$$G_u \sim PY(d_{|u|},\theta_{|u|},G_{\pi(u)}) \tag{5}$$

ここで， 依存パラメータ $\theta_{|u|}$ とディスカウントパラメータ $d_{|u|}$ はコンテキスト $\mathbf{u}$ の長さ $|\mathbf{u}|$ に基づくパラメータである．$\pi(\mathbf{u})$ はコンテキスト u が生起するコンテキストを表しており， $G_{\pi|\mathbf{u}|}$ はこのコンテキストに基づいた確率分布を表している．

式 (5) において再帰的にコンテキストを遡る操作を繰り返すことにより， n-gram 分布を生成する確率過程を生成することができる．この操作を深さ $n-1$ のコンテキストからコンテキストが存在しない状態，すなわち基底分布に G_ϕ を獲得するまで行う．

上述のようにして生成される階層 Pitman-Yor 過程は深さ n の Suffix-Array で表現される．このとき，それぞれのノードは $n-1$ のフレーズによって構成されるコンテキストのもとで，生起するとされており，また，そのノードは他のノードのコンテキストになっている．つまり，階層 Pitman-Yor 過程は観測された単語の出現確率によって長さ $n-1$ のコンテキストを考慮して対象の単語の生成確率を求めることができる．これにより， Kneser–and–Ney スムージング [10] を適用した n-gram 分布と同様の再現性の高い言語モデルを生成することができる．

3. 木置換文法

木置換文法 (TSG)[11] は文脈自由文法 (CFG: Context Free Grammar) の拡張である．TSG, CFG はともにそれぞれの書き換えルールによって構文木の非終端記号を書き換えていくことにより，入力文に対してトップダウンに構文木を構成する．CFG が特定のノードに対して，深さ 1 の部分木である書き換え規則を利用して書き換えるのに対

して， TSG は任意の深さの部分木でノードの書き換えを行っていく．そのため， TSG は任意のコンテキストに基づいた弱文脈依存性のある言語モデルを形成する．

TSG は $G = (T, N, S, R)$ の四つの要素によって記述される．T, N はそれぞれ非終端記号，終端記号の集合である．また， S はルートを表し， $S \in N$ であり， R は部分木の結合規則を表している．一般的に英語の構文木は句構造規則によって木構造に表されるため， TSG もまた木構造によって構文木モデルを表現している．

このとき，ルートノードは非終端記号としてラベル付けされ，葉ノードは終端記号あるいは非終端記号としてラベル付けされている．構文木の非終端記号を，任意の深さの構文木である部分木で再帰的に書き換えていくことにより，入力として与えられた文に対して対応する構文木を生成する．

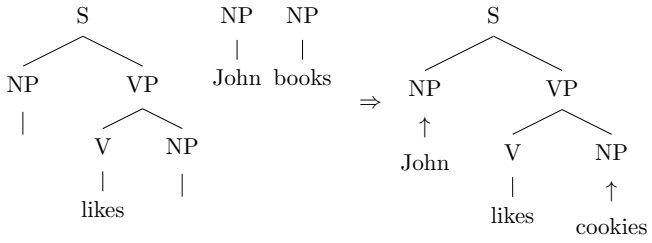


図 **1** TSG による構文解析

たとえば， 図 **1** では S → NP (VP (V like) NP) という書き換え規則により，非終端記号 S が部分木 (S NP (VP (V likes) NP)) に書き換えられている．この書き換え規則は二つの名詞あるいは名詞句を非終端記号としている．そして，これらの非終端記号である名詞は (NP John) と (NP cookies) の部分木によって書き換えられる．つまり，木置換文法を言語モデルとして利用して構文解析を行う際は解析対象の文に対して最適な部分木を設定し，その部分木の非終端シンボルに対して他の部分木の置換規則を用いて最適な置換操作を行うことで，文について構文木の生成を行う．

TSG は書き換え規則によって非終端記号を部分木に書き換えることで，構文木を生成するため，この構文木が書き換えられる確率を計算する必要がある．確率的木置換文法 (PTSG:Probabilistic Tree Substitution Grammar) は，TSG に部分木 e によって非終端記号 c を書き換える確率 $P(e|c)$ を付与して拡張したものである．この書き換え規則の確率は $P(e|c)$ は学習データに基づいて統計的に計算される．PTSG の書き換え規則の分布 G_c は Pitman-Yor 過程 [9] を用いて以下の式 (6) のようにして生成される．

$$G_c \sim \mathrm{PY}(d_c, \theta_c, G_{\pi(c)}) \tag{6}$$

ここで， d_c と θ_c は非終端記号 c がコンテキストが与えられたときの Pitman-Yor 過程のパラメータである．$G_{\pi(c)}$ は c を終端記号として持つ部分木の生起する無限次元の分布である．構文木 e を生成するとき，G_ϕ 非終端記号 $c_1, \cdots, c_m$ を葉ノードとして持つ部分木 e_1 を得る．同様に $e_2, \cdots, e_m$ を基底分布 $G_{\pi(c_1)}, \cdots, G_{\pi(c_m)}$ から生成する．この処理を構文木を生成し終えるまで，繰り返すことで，入力文に対する構文木を得る．

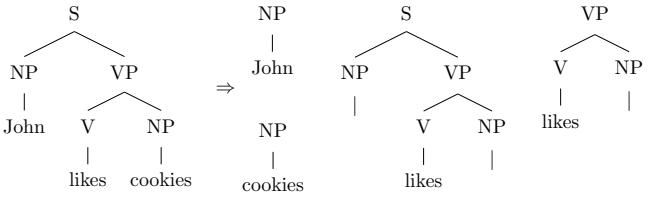


図 **2** TSG に基づく構文木の獲得

しかしながら， c のもとで e が生起する確率という PTSG の導出においては部分木の分割の問題がある．部分木の分割については Gibbs サンプラーを繰り返し用いることで，再帰的に最適な単位の構文木を決定することができる．図 **2** では構文木をランダムに部分木に分割することにより，特定の部分木のパタンを獲得している．

以上の操作に処理によって， TSG を利用した構文解析手法は弱文脈依存性に基づいた構文解析を行うことが可能でき，この手法は現在において英語の構文解析において最高精度を誇る手法 [5] のベースとなっている．しかしながら， TSG は構文木を木構造で表現するため，日本語において係り受け関係や語順の自由性をサポートできない．よって，我々は係り受け関係を階層化することで，弱文脈依存性と語順の自由性を考慮した TSG の日本語への拡張を行った日本語構文解析手法を提案している．

4. 係り受け関係の階層化とその共起を考慮した構文解析手法

本節では，我々が以前に提案した弱文脈依存性を考慮した係り受け関係の階層化による構文木モデルとこれを利用した日本語構文解析手法とこの問題点について述べ，この問題点の解決を図るために生成した構文木に含まれるコンテキストの共起を考慮する．

4.1 係り受け関係の階層化に基づいた構文木モデル

係り受けの生起するコンテキストを考慮するため，n-gram ベースな係り受け関係の生起確率に基づいた構文木モデルを構築する．係り受け関係の生起を n-gram のマルコフ過程として考えると，係り先の文節をコンテキストとしたとき，係り受け関係を持つ係り元の文節が生起する確率を計算することができる．そのため，文末の文節を開始状態とした n-gram によっての文節間の係り受け関係を状態の遷移として，構文木モデルを構成する．

一方で，n-gram モデルでは n が小さいと学習データの

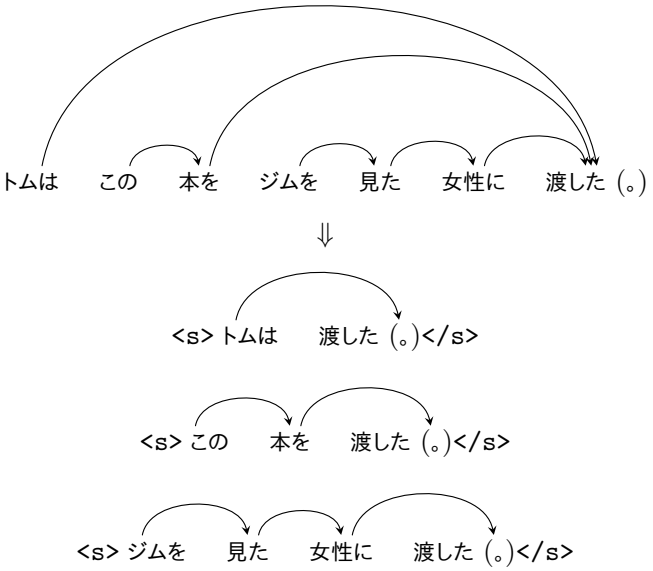


図 **3** 係り受け関係の階層化に基づいた構文木モデルの構築

パープレキシティが大きくなるという問題がある．逆に大きいと状態数が爆発的に増加し，モデルのサイズが大きくなってしまう．そこで，本手法では階層 Pitman-Yor 過程の拡張であり，n を任意の変数として扱うことのできる可変長階層 Pitman-Yor 過程 [12] を利用する．これにより，係り受け関係を持つ任意の長さの文節を構文木の単位とした構文木モデルとして扱うことができる．

これにより，文末の文節を初期のコンテキスト，つまり，根としてこれに対する係り受け関係を持つ文節が生起する確率を計算するとき，任意の係り先の文節数を考慮した構文木モデルを生成することができる．

例えば，図 **3** では，文末の文節に基づいた三つの係り受け木を観測する．図中で用いられている <s> はそれ以上，文節が係り元の文節から係り受け関係を受け取らないということを表しており，逆に </s> はそれ以上，係り受け関係を持たないつまり，それが文末の文節であるということを表している．これにより，文末の文節を根として階層化を行い，文末の文節「渡した」をコンテキストとして任意の文節「·」から，係り受け関係が発生する係り受け関係の生起確率 $P_D(\cdot|\text{渡した})$ について計算を行う．そのため，任意の長さの係り受け関係を持つ文節コンテキストを構文木モデルに取り込んだ再現率の高い構文木モデルを生成することが可能となる．

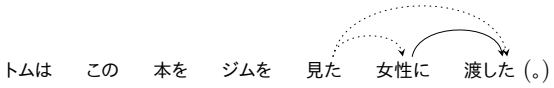
4.2 構文解析アルゴリズム

4.1 節で構築した構文木モデルは文末の文節を根として階層化されているため，構文解析の際には文末の文節から前方の文節へと入力文に対して CYK アルゴリズム [13] を用いてボトムアップに解析を行うことで，構文解析を行うことができる [6].

日本語構文解析において日本語の係り受け関係は以下の制約を持つ．

- 日本語は主要部終端型言語である．そのため，文節からの係り受け関係は右側の文節に対して発生し，すべての文節はその係り受け関係の係り先の文節を一つ持つ．
- 係り受け関係は交差しない．

(a)



(b)

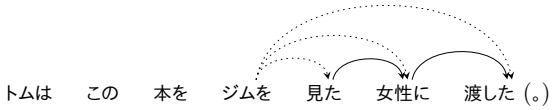


図 **4** 係り受け関係の階層化に基づく構文解析の例

これらの制約のもとで構文解析は行われるが，この際，第一に文末の文節はその直前の文節から係り受け関係を得るため，文末の文節とその直前の文節をコンテキストとして，文末の文節の直前の文節に対して係り受け関係を持つ文節の探索を行う．例えば，図 **4**(a) では，「渡した」という文節は文末の文節であるため，その直前の文節である「女性に」という文節から係り受け関係を得る．

そして，「渡した」という文節が「女性に」という文節から係り受け関係を得る確率を $P_D(\text{女性に}\,|\,\text{渡した})$ とし，この確率に基づいて，他の係り受け関係を探索する．「渡した」という文節が「女性に」という文節から係り受け関係を得るという条件のもとで，「女性に」という文節が「見た」という文節から係り受け関係を得る確率を $P_D(\text{見た}\,|\,\text{女性に渡した})$ と表す．また，「渡した」という文節が「見た」という文節から係り受け関係を得る確率を $P_D(\text{見た}\,|\,\text{渡した})$ と表し，図 4(a) において点線で描かれている矢印に相当するこれらの確率を算出する．

そして，それぞれの係り受け関係が生起する確率 $P_D(\text{見た}\,|\,\text{女性に渡した})$ と $P_D(\text{見た}\,|\,\text{渡した})$ とを比較し，確率が大きい係り受け関係を採用する．$P_D(\text{見た}\,|\,\text{女性に渡した})$ が $P_D(\text{見た}\,|\,\text{渡した})$ よりも高い確率を示す場合，「見た」という文節から「女性」にという文節に係り受け関係を持つとして，図 4(b) において実線として描かれているような係り受け木を得る．このプロセスを文頭の文節に到達するまで繰り返すことにより，入力文に対して係り受け関係を持つ任意の長さの文節をコンテキストとして考慮した構文解析を行うことが可能となっている．これにより， 従来の日本語構文解析手法 [1] における問題点を解決することができている [6].

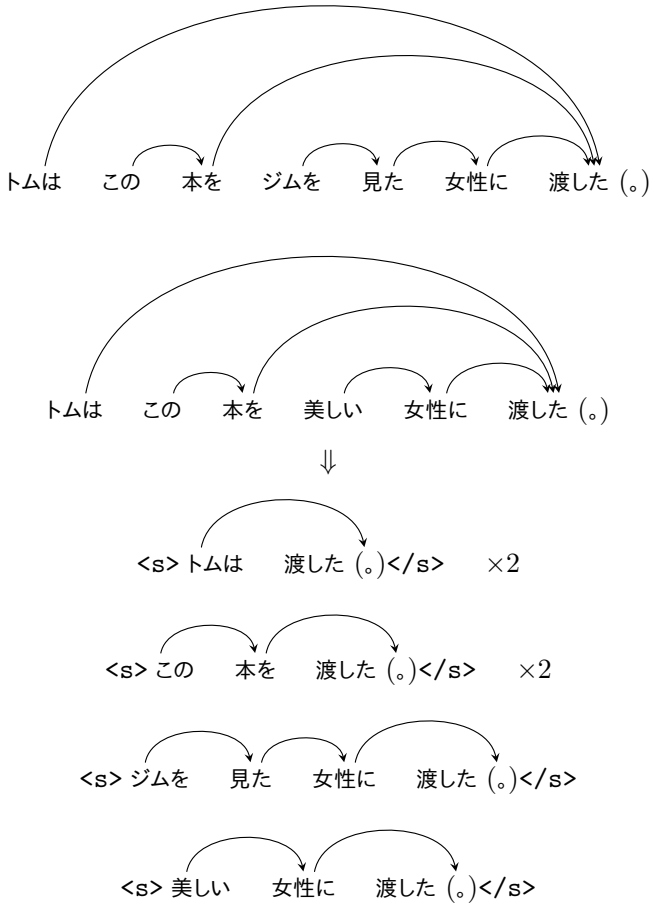


図 **5** 可変長 Pitman-Yor 過程を利用した構文木モデルの構築

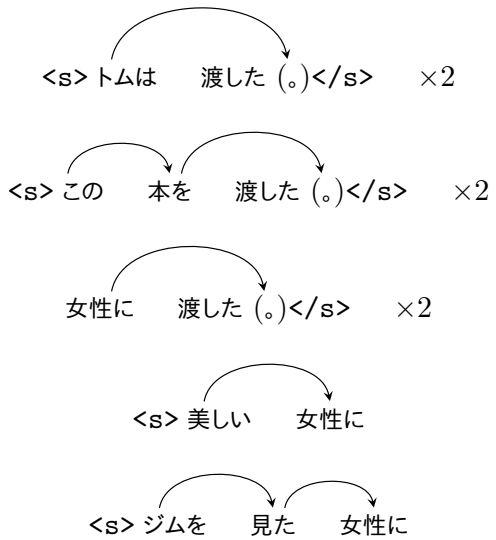


図 **6** 可変長 Pitman-Yor 過程による係り受け木の分割

4.3 コンテキストの共起を考慮した構文解析手法

係り受け関係の階層化に基づいた構文解析手法 [6] では，適切な単位の係り受け関係を持つ文節を構文木として獲得するために，可変長階層 Pitman-Yor 過程を利用していた．しかし，可変長階層 Pitman-Yor 過程は無限長の n-gram を構成した後にそれを頻度を基準に適切な長さの n-gram に分割するため， n-gram のコンテキストが途中で失われることがあった．

例えば，図 **5** のような文が学習データとして与えられた場合，「女性に」という文節が「渡した」にかかるような係り受け関係が重複して観測される．そのため，可変長階層 Pitman-Yor 過程を利用して係り受け関係を持つ任意の文節の長さを決定するときに，「女性に」という文節で構文木が分割された形で図 **6** のように構文木モデルに取り込まれる可能性がある．これらは独立したコンテキストとなるため，係り受け関係を得る文節を推定する際のコンテキストが不足してしまう．

そこで我々は可変長階層 Pitman-Yor 過程を階層 Pitman-Yor 過程のノードとした入れ子構造である Nested Hierarchical Pitman-Yor 過程 [14] を利用することで，コンテキストの共起確率を考慮することでこの問題点の解決を図った．この効果ついて次節となる 6 節で評価実験を行い，精度の向上について確認する．

5. 評価実験

1995 年度の毎日新聞の一部のデータに対して様々な言語情報が人手で付与された京都大学テキストコーパスは形態素，文節間の係り受け関係等が示されている日本語コーパスの一つであり，形態素解析や構文解析といった自然言語処理の基礎的なタスクに利用される．本節ではこのコーパスを用いて評価実験を行った結果を示す．

5.1 比較実験

係り受け関係の階層化に基づいた構文木モデルによる構文解析手法を従来手法とし，この従来手法によって構築された構文木モデルに含まれる構文木のコンテキストの共起確率を考慮し，提案した構文解析手法とを比較する．評価実験のモデル生成の際には京都大学テキストコーパスの 1995 年 1 月 1 日分のデータを学習データとして利用した．構文木モデルの生成では各文の係り受け関係を文末の文節に基づいた形態素および各文節の品詞体系からなる係り受け関係の構造を抽出する．そして，各係り受け関係に対して文末の文節を根として階層化し Nested Hierarchical Pitman-Yor 過程によって構文木モデルを生成する．Nested Hierarchical Pitman-Yor 過程のパラメータ推定には Gibbs イテレーションを 50 回繰り返し，構文木モデルを生成した．

Nested Hierarchical Pitman-Yor 過程によって生成された構文木モデルにおける uni-gram モデルは可変長階層 Pitman-Yor 過程に基づいた構文木モデルと等価であるので，従来手法の Nested Hierarchical Pitman-Yor 過程の uni-gram モデルを構文木モデルとして構文解析を行った．一方で，コンテキストの共起頻度を考慮した構文解析手法では，可変長階層 Pitman-Yor 過程による構文木の bi-gram モデルを用いることでコンテキストの共起確率を考慮した

構文解析を行う．これらの構文解析手法の評価には，1995 年 1 月 3 日のデータから無作為に 200 文を抽出したものをテストデータとして使用する．

5.2 結果

文節間の係り受け関係は一般的に係り受け解析手法の評価に利用される式 (7)[1] を利用して係り受け解析の精度を測定した．

$$\begin{aligned} X &= \text{各手法によって得られた文節間の係り受け関係と正解データの係り受け関係の一致数} \\ Y &= \text{テストデータの文節間の係り受け総数} \\ \text{正解率} &= \frac{X}{Y} \end{aligned} \tag{7}$$

表 **1** の実験結果より，構文木モデルに含まれる係り受け関係を持つ任意の長さ文節を構文木におけるコンテキストの共起を考慮することにより，従来手法と比較して，わずかであるが精度の向上を確認することができた．精度の向上がわずかであったことの要因の一つとして， uni-gram に基づいた構文木モデルにおいてコンテキストが途中で途切れる頻度が少なかったためであるといえる．コンテキストが途中で途切れる頻度が少ないことの理由としては，文節の素性を形態素によって表現することで文節が細分化されていることにある．そのため，文節の素性の粒度について検討することが課題として残った．

	CaboCha-0.66	従来手法	提案手法
正解率 (%)	88.1	77.5	78.3

表 **1** 実験結果

6. おわりに

本稿では日本語における構文解析手法において係り受け関係を階層化し，構文木モデルを構築したときに構文木に含まれるコンテキストの共起を考慮した構文解析手法に拡張することで，精度の向上を図った．

評価実験の結果，わずかな構文解析精度の向上を観測したが，これは文節の素性を形態素の並びとして構文木を構築した場合に，文節が細分化され構文木モデルが分割されにくいという要因が挙げられる．これにより，構文木モデルも肥大化してしまうことが考えられるため，今後の課題として文節の素性の粒度について検討する必要がある．また，係り受け関係がアノテーションされた京都大学テキストコーパスの全データを用いた評価実験についても評価の観点から行うべきであり，さらなる精度の向上を目指す際に必要事項である．

謝辞　本研究の一部は，日本学術振興会科学研究費補助金挑戦的萌芽研究 (課題番号: 25540150) の支援による．ここに記して謝意を表す．

参考文献

[1] Kudo, T. and Matsumoto, Y.: Japanese Dependency Analysis using Cascaded Chunking, *CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pp. 63–69 (2002).
[2] Cohn, T., Blunsom, P. and Goldwater, S.: Inducing Tree-Substitution Grammars, *Journal of Machine Learning Research*, Vol. 11, pp. 3053–3096 (2010).
[3] Post, M. and Gildea, D.: Weight Pushing and Binarization for Fixed-Grammar Parsing, *Proceedings of the 11th International Conference on Parsing Technologies (IWPT'09)*, Association for Computational Linguistics, pp. 89–98 (2009).
[4] Blunsom, P. and Cohn, T.: Unsupervised induction of tree substitution grammars for dependency parsing, *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, (EMNLP '10), Association for Computational Linguistics, pp. 1204–1213 (2010).
[5] Shindo, H., Miyao, Y., Fujino, A. and Nagata, M.: Bayesian symbol-refined tree substitution grammars for syntactic parsing, *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*, ACL '12, Association for Computational Linguistics, pp. 440–448 (2012).
[6] 大野一樹，波多野賢治：係り受け関係の階層化に基づいた構文木モデルによる構文解析手法の提案，2013 年度 情報処理学会関西支部 支部大会 (2013).
[7] Pitman, J. and Yor, M.: The Two-Parameter Poisson-Dirichlet Distribution Derived from a Stable Subordinator, *The Annals of Probability*, Vol. 25, No. 2, pp. 855–900 (1997).
[8] Pitman, J.: Exchangeable and partially exchangeable random partitions, *Probability Theory and Related Fields*, Vol. 102, No. 2, pp. 145–158 (1995).
[9] Teh, Y. W.: A hierarchical Bayesian language model based on Pitman-Yor processes, *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, ACL-44, Association for Computational Linguistics, pp. 985–992 (2006).
[10] Kneser, R. and Ney, H.: Improved backing-off for M-gram language modeling, *Acoustics, Speech and Signal Processing*, Vol. 1, pp. 181–184 (1995).
[11] Cohn, T. and Lapata, M.: Sentence Compression as Tree Transduction, *Journal of Artificial Intelligence Research (JAIR)*, Vol. 34, pp. 637–674 (2009).
[12] Mochihashi, D. and Sumita, E.: The Infinite Markov Model, *Advances in Neural Information Processing Systems 20 (NIPS 2007)*, pp. 1017–1024 (2007).
[13] Jurafsky, D. and Martin, J. H.: *Speech and Language Processing (2nd Edition) (Prentice Hall Series in Artificial Intelligence)*, Prentice Hall, 2nd edition (2008).
[14] Mochihashi, D., Yamada, T. and Ueda, N.: Bayesian unsupervised word segmentation with nested Pitman-Yor language modeling, *In Proc. of ACL* (2009).