

後続要素の予測を利用する系列予測モデル

A Model of Sequence Prediction with Prediction of subsequent elements

上原 謙吾*

Kengo UEHARA

原田 賢一*

Kennichi HARADA

1 はじめに

Elman(1991, 1993) は、言語の獲得過程のモデルとして単純再帰型ニューラルネットワークモデル (Elman Net) を利用してシミュレーションした結果、学習によって言語が獲得できることを示した [1].

Elman Net は人間の潜在学習の多くの特性を示すことが知られており、現在、この Elman Net を元に、言語にまつわる統語知識や語彙の獲得など、人間の言語獲得のモデルを提案した多くの研究が報告されている [3][5][6].

しかし、Elman のシミュレーションで使用された単語系列 (人工文) は、27 種類の単語 (名詞、動詞、ピリオド) であり、冠詞や前置詞、形容詞などの単語は含まれない。また、対象とする人工文は、関係節を含む平叙文であるが、基本的には SV/SVO 型の構造にすぎず、疑問文や受動態、時制などの文法規則は使用されない。

本研究では、系列要素としてその前接・後接要素との依存度が高い前置詞に焦点を絞り、前置詞を含む文を学習できるように Elman Net の拡張モデルを提案する。前置詞は、後続の名詞によっても規定される要素であるため、1 個先の単語予測課題での処理は困難である。本研究で提案する拡張モデルは 2 個先の単語予測も行う 2 段階予測モジュールを組み込んでおり、後続要素の予測結果を入力情報として同時に利用する。この拡張モデルを利用して、人工文を学習対象とした文法獲得のシミュレーションを行い、Elman Net では検討されていない前置詞などを含む文法規則の獲得を検証する。

2 提案する系列予測モデル

2.1 ニューラルネットワークモデルの概要

Elman Net は、3 層パーセプトロンに文脈層を付加したものである。文脈層は 1 時点前の中間状態層の活性パターンを記憶しており、次の時点の入力の際に中間状態層にフィードバックされる。この特徴により、Elman Net は時系列情報を処理する能力を持つ。

提案する系列予測モデルの概要を図 1 に示す。現在の単語、文脈層からのフィードバックに加えて、現在の単語から 2 個先の単語予測を行った結果も共に入力として使用する。Elman Net の基本形と異なる。

2.2 2 段階予測モジュール

2 段階予測モジュールは、1 個先の単語予測に加えて、その予測結果を元に更に 2 個先の単語予測も行う。2 段階予測モジュールの全体像を図 2 に示す。その基本構造は、Elman Net を直列に 2 つ組み合わせたもので、2 階

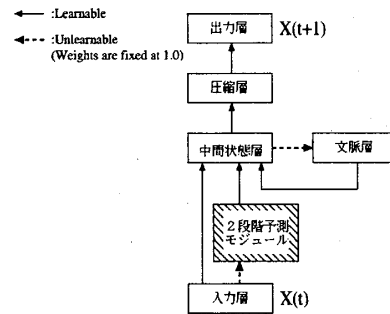


図 1: 提案する系列予測モデルの概要図

部分の入力には 1 階部分の出力値 (1 個先の単語予測値) を利用する。

2 段階予測モジュールとして、(a) 文脈層からのフィードバックを各階だけで使用する over-prop type と、(b) 双方向型の情報処理 [4] のように両階で使用する bi-directional type の 2 種類を実装する。また、誤差を逆伝播させる範囲、その重み付け、1 階部分の出力値から 2 階部分への入力値を計算する方法など、1 段階目と 2 段階目間のデータのコピー方法 (フィルタリング) も比較する。

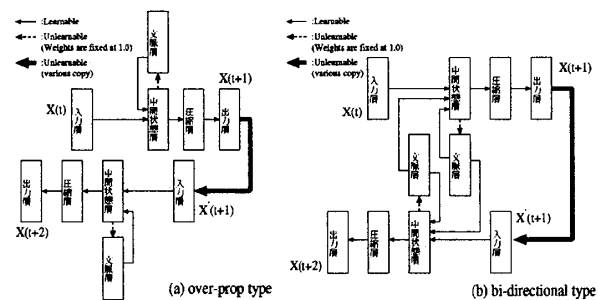


図 2: 2 段階予測モジュール

2.3 学習課題と学習方法

2.3.1 提示する人工文

学習に用いる人工言語は、名詞、動詞に加えて、前置詞、前置詞にのみ接続する名詞などから構成される。学習対象とする人工文は句構造規則 (テンプレート) をランダムに選択して生成される。使用する規則は、語順を制約する統語的規則と、単語間の共起関係を制約する語彙的規則の 2 種類である。本研究では、Elman のシミュレーションとは異なり、前置詞を含む文も句構造規則に含まれている。

* 慶應義塾大学 理工学部 情報工学科

2.3.2 学習課題

提案する系列予測モデルを用いたシミュレーションでは、Elmanのシミュレーションと同じく、ある単語系列が与えられたときの1個先の単語予測を学習課題とする。ただし、2段階予測モジュールの学習フェイズでは、2個先の単語予測課題も同時に行う。

2.3.3 学習方法

提案モデルを使用して系列予測を学習するには、あらかじめ2段階予測モジュールを学習させておく必要がある。つまり、以下の手順で学習を行う。

1. 2段階予測モジュールの学習
 まず、2段階予測モジュール(図2)の学習を行う。2段階予測モジュールの学習は、1階部分に対しては1個先の単語を、2階部分に対しては2個先の単語を、それぞれ教師信号として与える。
2. 提案する系列予測モデルの学習
 2段階予測モジュールの学習が終了した後、その学習済みのネットワークを組み込んだものが提案する系列予測モデル(図1)である。提案モデルの学習は、Elmanのシミュレーションと同様に、1個先の単語を教師信号として与える。

3 結果

提案した系列予測モデルのシミュレーションでは、与えられた単語系列(人工文)から1個先の単語を予測することが学習課題となる。ただし、Elmanのシミュレーションと同じく、このような予測課題の結果を評価する場合、予測モデルの出力値(予測値)は1個先の単語の生起確率となるので不確定である。本研究では、出力値を元に正解単語の順位を平均して、予測精度を評価する。

単語数43(含まれる前置詞はinとto)の場合、Elman Netの基本型と提案モデルで正解単語の平均予測順位を比較した結果を図3に示す。2段階予測モジュールの組み合わせによっては、提案モデルが予測精度を上げることが確認された。

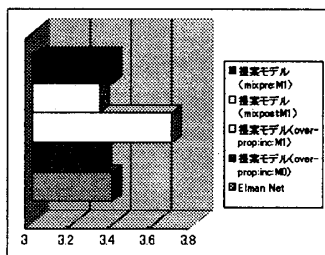


図3: 正解単語の平均予測順位

また、単語数54(含まれる前置詞はin,to,from,on)の場合、bi-directional(図2)を組み込んだ提案モデルに対して単語系列の予測学習を行い、学習終了後にその中間状態層の活性パターンを階層的クラスタリングを用いて分類した結果を図1に示す。前置詞が完全に分類されていないが、前置詞を含む人工文からも単語の系列特徴が得られることが確認された。

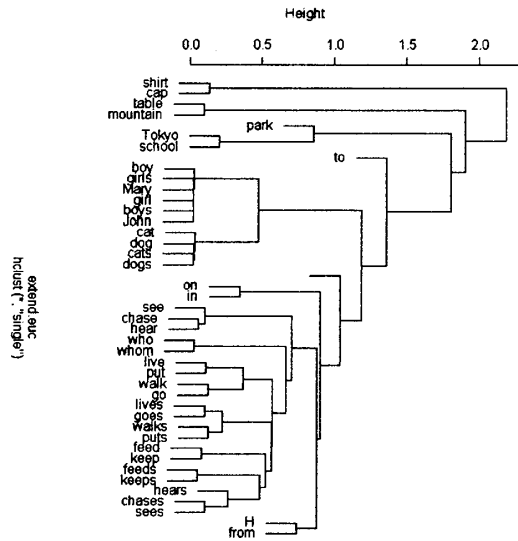


図4: 中間状態層に抽出された単語系列特徴

4 まとめ

本研究では、後続要素からも束縛を受ける単語系列を処理するために、Elman Netに2段階予測モジュールを組み込んだ系列予測モデルを提案した。提案モデルは、1個先の単語予測に後続要素(2個先の単語)の予測結果を利用できるように、2段階予測モジュールを組み込んだ。この結果、提案モデルを利用して前置詞を含む人工文に対しても文法規則が獲得できることを示した。

参考文献

- [1] Elman, J. L.: Distributed representations, simple recurrent networks, and grammatical structure, Machine Learning, 7, 195-225, 1991.
- [2] 乾: 文理解過程のネットワークモデル, 心理学評論, 40(3), 303-316, 1997.
- [3] 玉森, 乾: Elman ネットによる統語範疇の配列と格関係の学習, 認知科学, 6(3), 359-368, 1999.
- [4] 和久屋, 信太: “双方向型計算機に基いた神経回路モデルによる時系列予測: 計算機生成時系列データ” Data Set D”を用いた汎化能力の検討”, 電子情報通信学会論文誌 D-II, Vol.J85-D-II, No.1, 101-111, 2002.
- [5] 下斗米, 遠山, 大森: 文法メタ知識による語彙学習加速のコネクショニストモデル Cognitive Studies, Vol.10, No.1, 104-111, 2003.
- [6] 森藤, 乾: 自己組織化マップを含む Elman 型ネットワークを用いた文法知識の獲得モデル, 日本認知心理学会第1回大会論文集, 2003.