

H-005 単純再帰ネットワークの定量的な評価方法の検討 Quantitative Evaluation of Simple Recurrent Network

竹谷 仁志[†] 岩村 雅一[†] 大町 真一郎[†] 下村 正夫[†] 阿曾 弘具[†]
Hitoshi Takeya Masakazu Iwamura Shinichiro Omachi Masao Shimomura Hiroto Aso

1. はじめに

人の高次の脳機能の一つとして言語機能が挙げられる。言語の獲得過程の研究において、生得的に言語を保持しているという立場と学習によって獲得するという立場がある。後者を支持する研究として、Elmanによるものがある。Elmanは単純再帰ネットワーク(Simple Recurrent Network; 以下SRN)を用いたモデルによる実験を行い、文の系列から文の構造が獲得できることを示した[1]。しかし、Elmanの用いた評価値であるCOSINEでは定性的な評価は可能であるが定量的な評価は困難であると考えられる。そこで本研究では、言語処理モデルの評価方法であるパープレキシティとクロスエントロピー[3]、および確率事象の距離尺度であるダイバージェンスによる定量的な評価を検討する。

2. 言語獲得過程のモデル

本研究では、ElmanのSRNを用いたモデルを使用する。本節ではこのモデルについて説明する。このモデルでは、学習者(SRN)が周囲でなされる会話(文系列)を聞き、次に話される単語を予測しているうちに、結果として言語を獲得するものとする。

本研究では単語を29種類用意し、それぞれ29次元の単位ベクトルで表す。1の位置が単語の種類を表す。文はSVまたはSVOの文型をとる名詞と動詞のみからなり、多数の文を切れ目なく繋げたものを文章とする。語尾変化や句読点等は考えない。

単純再帰ネットワーク(SRN)は三層パーセプトロンに、隠れ層の値を一時保存する文脈層を付加したものである(図1)。入力層と出力層の各ノードを各単語に一対一対応させて入力層と出力層のノード数は29とした。隠れ層と文脈層のノード数は150とした。

ここで文の作成方法について説明する。29種類の単語は人を表す名詞や動物を表す名詞、自動詞や他動詞など12種類に分類され、これをカテゴリと呼ぶ。文は、カテゴリをどのように並べるかが記述されている16種類のテンプレートを用いて作成する。一つのテンプレートをランダムに選択し、さらにテンプレートに記述されているカテゴリの中から単語をランダムに選択して、一つの文を作成する。

3. 評価方法

本節では本研究で用いる4つの評価方法について簡単に説明する。

3.1 COSINEによる評価

単語 x が入力されたときの出力層の各ノードの出力を x の次に入力される単語の生起確率の予測値と考える。 x は入力系列中に複数回現れるので、予測値の平均を求

[†]東北大学大学院工学研究科

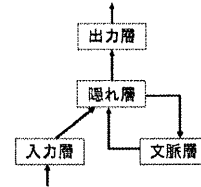


図1: 単純再帰ネットワーク(SRN)

める。これをさらにカテゴリごとに平均したものを12次元ベクトル \mathbf{O}_x とする。本研究では、この \mathbf{O}_x と文の生成規則から算出される予測値の理想値 \mathbf{P}_x の類似度を表す COSINE_x を $\text{COSINE}_x = \frac{\mathbf{O}_x \cdot \mathbf{P}_x}{|\mathbf{O}_x| |\mathbf{P}_x|}$ と定義する。

3.2 パープレキシティによる評価

本研究で新しく用いる3つの評価方法は、情報理論に基づいている。そこで、単語 x の次に続くカテゴリ $C^{(i)}$ の生起確率を確率分布 $P_M(C^{(i)}|x)$ に従うものとして扱う。パープレキシティ $PP(x)$ は、エントロピー $H(x) = -\sum_i P_M(C^{(i)}|x) \log_2 P_M(C^{(i)}|x)$ によって、 $PP(x) = 2^{H(x)}$ と定義される。ここで、 $P_M(C^{(i)}|x)$ としてベクトル \mathbf{O}_x の i 番目の要素を用いる。パープレキシティは単語 x の後に続く平均カテゴリ数を表す。

3.3 クロスエントロピーによる評価

クロスエントロピー $H_{\text{cross}}(x)$ は $H_{\text{cross}}(x) = -\sum_i P(C^{(i)}|x) \log_2 P_M(C^{(i)}|x)$ と定義される。 $P(C^{(i)}|x)$ は単語 x の次にカテゴリ $C^{(i)}$ が続く確率を表し、ベクトル \mathbf{P}_x の i 番目の要素である。クロスエントロピーは、言語モデルがどれだけ正確に言語を近似しているかの評価に用いられる。

3.4 ダイバージェンスによる評価

ダイバージェンス $D(x)$ は $D(x) = \sum_i P(C^{(i)}|x) \log_2 \frac{P(C^{(i)}|x)}{P_M(C^{(i)}|x)} = H_{\text{cross}}(x) - H(x)$ と定義される。ダイバージェンスは二つの確率分布の距離尺度として用いられる。

4. 評価

COSINE、パープレキシティ、クロスエントロピー、ダイバージェンスの4つによって学習後のSRNを評価する。文を連結した27476語の単語系列を入力系列として使用し、SRNの結合強度、各ノードの閾値の初期値をランダムに与えてBP法による学習を行った。入力系列の提示回数(学習回数)を1回から11回まで変えて、各評価を行った。4つの評価尺度の学習回数に対する推

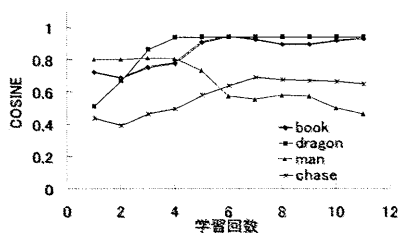


図 2: COSINE の推移

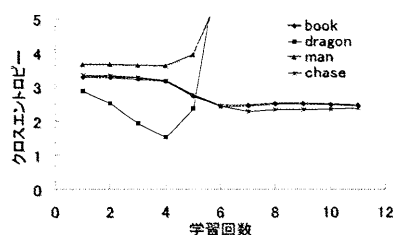


図 4: クロスエントロピーの推移

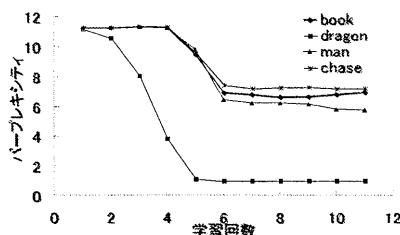


図 3: パープレキシティの推移

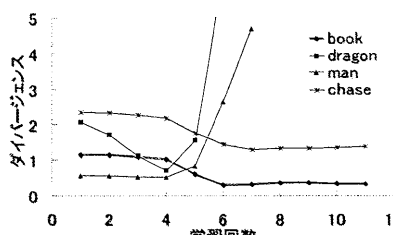


図 5: ダイバージェンスの推移

移を示す。今回は 29 種類の単語の内、4 つの単語 book、chase、dragon、man について注目する。SRN の学習結果を人間が評価した結果、book は理想的な出力 chase はある程度理想的な出力、dragon と man は理想とはかけ離れた出力が得られた。

まず、COSINE の推移 (図 2) から man を除いた 3 つの単語は学習によって理想の生起確率に近い出力をするようになると評価される。しかし、パープレキシティの推移 (図 3) に注目すると、dragon は学習回数が 5 回を越えと値が 1 になっている。これは、カテゴリを一つしか予測していないことを示しており、実際に SRN の出力はそのカテゴリに属する単語に対応するノードのみ大きな値をとっていた。文の生成規則では、dragon の次には 2 つのカテゴリが続くことになっており、そのうち生成される確率の高い方のカテゴリのみ予測していた。これらのことから、COSINE による評価では良い結果であると評価された dragon は、パープレキシティに注目することで予測すべきカテゴリを予測できていないことがわかった。つまり、COSINE による評価で大きな値をとる場合でも、予測すべきカテゴリを正しく予測できていない場合があり、必ずしも有効な評価基準であるとは言えない。

次に、man について注目するとパープレキシティの推移から学習によって次に続くカテゴリの数を減らしていると評価された。しかし、COSINE に注目すると man については理想の生起確率との類似度が下がってしまっている。つまり、次に続くカテゴリを絞り込んではいないが、そのカテゴリが誤っているものと考えられる。パープレキシティによる評価では大体いくつくらいのカテゴリに絞り込んだかは評価できるが、そのカテゴリが正しいかどうかまでは評価できない。

次にクロスエントロピーの推移 (図 4) を見ると、dragon、man について、正しいカテゴリを絞り込ん

でいないという評価が出来る。ダイバージェンスも同様 (図 5) である。しかし、どの単語もダイバージェンスの理想は 0 となるのに対し、クロスエントロピーはそうではない。つまり、クロスエントロピーの推移を見た時に、最小値に収束しているようであってもそれが確かに最小値であるとは言えない。実際、book と chase について注目すると、クロスエントロピーはほぼ同様に収束しているのに対し、ダイバージェンスは book と chase で収束した時の値が違い、後どれだけ値が減り得るのかが分かる。この点からダイバージェンスによる評価の方が優れた評価であると考えられる。

5. まとめ

本研究では、Elman による SRN を用いたモデルの定量的な評価方法として、言語処理モデルの評価方法に使用されるパープレキシティ、クロスエントロピー、確率事象の距離の評価に使用されるダイバージェンスの 4 つを用い、評価を行った。パープレキシティによって学習の結果どの程度カテゴリを絞り込めたかをカテゴリの数で評価できた。また、クロスエントロピー、ダイバージェンスによる評価方法によって COSINE では誤って良い結果と評価されてしまう悪い結果も正しく評価出来た。今後は、クロスエントロピーとダイバージェンスの値の物理的な意味付けを検討する必要がある。

参考文献

- [1] J. L. Elman. "Finding Structure in Time," Cognitive Science, vol.14, pp.179-211, 1990.
- [2] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett, "Re-thinking Innateness," MIT Press, Cambridge, 1996.
- [3] 北研二, "確率的言語モデル," 東京大学出版会, 1999.