

# 平成21年度後期自主課題研究

## ニューラルネットワークを用いた英文中の単語の意味判別

情報システム工学科3年 031 田澤遼

指導教員 中山謙二

### 1、 研究課題、背景

ニューラルネットワークを用いて、英文中に存在する、複数の意味を持つ単語を判別させる。今回は、動詞‘have’の意味の判別のみを行う。

ニューラルネットワークを用いることで、ただ単語の意味を訳すのではなく、文脈から判断して単語の意味を読み取れるのではないかと考え、この課題を設定した。

### 2、 研究方法

英文中の単語を、単語一つにつき一つではなく、分類毎に6bitのコードに変換する。

例：apple → 000001, breakfast → 000001, dog → 000010

今回は、最大8単語からなる英文を用いるため、6bit×8=48bitのデータをニューラルネットワークへの入力データとして入力する。

同様に目標値は10種類の意味を10bitのコードとして表す。

例：「～を持っている」 → 0000000001, 「～を飼っている」 → 0000000100

学習には、肯定文、疑問文、否定文を最低一つ用いる。

### 3、 研究成果

学習は全てのデータから正確な数値が出力されたため、学習そのものは上手くいっているものと考えられる。ただ、テスト結果の正答率は、60%とあまり芳しい数字ではなかった。学習させたデータと文型の似通っているものについては、多くの場合で、正しい値が出た。文の意味を決める重要な単語、例えば「～を食べている」という意味を出す場合における「食べ物」という分類の単語が、学習データとテストデータとの間で、入力された位置にズレがある場合では、正しい値が出ることはほとんどなかった。出たとしても、コードの誤認識などの別の要因が絡んできた。

### 4、 まとめ

ニューラルネットワークは、入力するデータの位置が非常に重要となってくる。

そのため、英文のように、データの位置が変わるようなものを扱うのは難しいのではないかと考える。

現在考えられる対策としては、学習の際に、重要単語のコードを、考えられる全ての位置に入力させれば、データの位置のズレという問題は解消できるのではないかと思う。