

# 言語認知工学 最終レポート

2210259 米谷 祐希

2025 年 2 月 14 日

## 問 1

- (1) × 単語 apple の意味を外界との対応付けで推測・学習することは、記号接地の一例である。
- (2) × プロトタイプとは、カテゴリーの最も典型的な仮想上のメンバのことである。
- (3) × 「数学」と「物理学」は、同位関係にある。
- (4) × BoWとは、自然言語処理において文書や言語表現を単語の集まりで表現する方法のことである。
- (5) ○
- (6) ○

## 問 2

単語 doctor に対する、単語 nurse によるプライミング効果は意味プライミングであるので、その効果を検証するための比較対象としては、プライミングを行わずに単に doctor と示したり、全く関係ない語を示して doctor の反応時間を測定し、比較すれば良い。

## 問 3

まず最初に

$$d_1^m = \frac{\partial E}{\partial z_1^m} = \frac{\partial E}{\partial a} \frac{\partial a}{\partial z_1^m} \quad (1)$$

ここで、シグモイド関数の微分を考えると、

$$\frac{\partial a}{\partial z_1^m} = f_0'(z) = f_0(z)(1 - f_0(z)) = a(1 - a) \quad (\because a = a_1^m = f_0(z_1^m)) \quad (2)$$

であり、

$$\frac{\partial E}{\partial a} = -\frac{y}{a} + \frac{1-y}{1-a} \quad (3)$$

であるので、

$$\begin{aligned} d_1^m &= \left(-\frac{y}{a} + \frac{1-y}{1-a}\right)a(1-a) \\ &= -y(1-a) + (1-y)a \\ &= a - y = a_1^m - y \end{aligned}$$

と導出できる。

## 問 4

リカレントネットワークは隠れ層の情報を保持して、次の処理へとつなげることができ、文章などの時間軸データとなったものに対して利点となっている。一方で、フィードフォワードネットワークは、実装がシンプルであり学習もリカレントネットワークでの勾配消失問題に対して軽減することができるという利点がある。さらに文章分類において、特徴抽出により固定長のベクトルにすることができる場合には、時系列データとして扱う必要がなくなるという点でフィードフォワードネットワークの効率に利点がある。

## 問 5

意味空間モデルにおいて単語というのはベクトルで表される。そのベクトルというのは、その他の単語・文脈からどのように使われるか、再起頻度などによって計算されるものであり、Glenberg and Robertson が言うように、人間の実世界での意味の認識・獲得の過程と対応付けられていないという点が問題である。例えば犬というものを想像したときに単語ベクトルの「犬」という単語は他の語とどんな関係性にあるという情報を持っているのみであり、実際に人間がふわふわして・小さい・動くもの、のような視覚的であったり感覚的な特徴情報を保持しているわけではないということである。

## 問 6

### (1)

$v_{34}$  は、文章  $D_4$  における、単語  $w_3$  の出現頻度を表している。

### (2)

$v_{34}$  は、段落、文、前後数単語のある範囲での、 $w_3$  と  $w_4$  が同時に出現する、共起頻度を表している。

(3)

$w_1$  とコサイン類似度が高い単語は  $\frac{w_1 \cdot w_i}{|w_1| \cdot |w_i|}$  が最大となる単語  $w_i$  を選択すれば良い。  
計算の結果そうなるものは、単語  $w_5$  であり、コサイン類似度は、

$$\frac{w_1 \cdot w_i}{|w_1| \cdot |w_i|} = \frac{37}{\sqrt{78} \cdot \sqrt{26}} = \frac{37}{26\sqrt{3}} \quad (4)$$

となる。

(4)

この問題では td-idf は  $l_{ij}g_i$  と表される。これに値を当てはめて計算すると、

$$\begin{aligned} l_{ij}g_i &= \log_2(1 + f_{ij}) \cdot \log_2(n/n_i) \\ &= \log_2(1 + 6) \cdot \log_2(5/3) \\ &= \log_2 7(\log_2 5 - \log_2 3) \end{aligned}$$

となる。

## 問 7

(1)

ノードの次数とは、そのノードの接合しているリンク数なので、次数が 2 次のは a,b,e,f である。

(2)

それぞれのノード  $i$  における  $C_i$  を計算した。

表 1: それぞれのノードにおける  $C_i$

ノード $i$	$C_i$
a	1
b	1
c	1/3
d	1/3
e	1
f	1
g	2/3

よって

$$C = \frac{\sum C_i}{|V|} = \frac{16}{21} \quad (5)$$

**(3)**

それぞれのペアについての最短経路長をすべて足し、 ${}_7C_2$  で割ると平均最短経路長になるので、

$$L = \frac{34}{21} \quad (6)$$

となる。

**問 8**

第 6 章で扱っている評価方法の中で、これは順位付きの評価であるが、一度 6 つのすべてを対象として順位なしとして比較すると、どちらのプログラムも 6 個の出力の中に人間が連想した 3 つの連想語は含まれている。つまり再現率は 100% である。そして 3 つの連想後に対して 6 個の出力をしているので、適合率は 50% である。よって順位なしとして扱った場合の評価はどちらも同じである。

次に順位付きとしてみて、平均適合率を計算する。

**プログラム A**

第 2,3,4 位にあるので、

$$(\text{平均適合率}) = \frac{\frac{1}{2} + \frac{2}{3} + \frac{3}{4}}{3} = \frac{23}{36} \quad (7)$$

**プログラム B**

第 1,4,6 位にあるので、

$$(\text{平均適合率}) = \frac{\frac{1}{1} + \frac{2}{4} + \frac{3}{6}}{3} = \frac{24}{36} \quad (8)$$

と言う結果になったので、プログラム B の優れているという評価ができる。