

# 情報工学実験Ⅱレポート（探索アルゴリズム2）

曜日 & グループ番号: 月 & グループ 0

2011 年 12 月 23 日

## 概 要

このレポート（ファイル）は、「情報工学実験Ⅱ・探索アルゴリズムその 2[1]」の実験レポートの骨組みを例示している。あくまでも例示であって、全てをこの通りに従う必要はないが、指示された項目を含めた上で、報告書として他者が読みやすいレポートとなるように工夫する事。

## グループメンバ

（補足：レベル毎に 全員が協力して実施 した上で、レベル毎にレポートをまとめる担当者を決め、全体を一つのレポートとして整理すること。）

- 145763C 仲村 大地: 担当 Level1, 3.3
- 145738B 西銘 明留: 担当 Level2, 3.1, 3.2
- 145717K 泉川真理南: 担当 Level2, 3.1, 3.2
- 145714E 豊 美玲 : 担当 Level3.4

## 提出したレポート一式について

レポート一式は‘‘naha:/home/home/teacher/tnal/jikken1-fri/e945734/’’ にアップロードした。提出したファイルのディレクトリ構成は以下の通りである。

（補足：必ず下記のように整理しろという指定ではない。自分たちでやりやすいように Level 毎に整理しても構わない）

```
./src/      # 作成したプログラム一式
./report/   # レポート関係ファイル、図ファイルを含む、
```

# 1 Level1: 線形分離可能な OR 問題への適用

## 1.1 課題説明

2 入力 1 出力で構成される単純パーセプトロン（ニューラルネットワーク）を用いて、4 つの教師信号を用意した OR 問題へ適用し、重みが適切に学習可能であることを確認する。また、学習が収束する様子をグラフとして示す。

## 1.2 OR 問題を学習させた際の誤差収束度合いについて

### 1.2.1 実験結果

NN では重みを更新する毎に誤差が減るように学習を行うが、その学習の様子は初期の重みをどのように設定したか、学習に用いたパラメータをどのように設定したか、といった対象問題以外の要素に影響して学習の様子が変化する。シード値を変えた際の学習収束回数を表 1 に示す。シード値を 10 回変更して学習させた際の重みを更新する様子を図 1 に、その平均をプロットした平均推移値を図 2 に示す。なお、平均値を求める際には 10 回分の実行データを加工して統合、平均値を算出した。具体的には一行ごとに実行データから抜き取り、10 回分を加算し、加算回数で割って求めている。(src/bp\_mo 配下の integration.sh を参照)

表 1: OR 問題の学習に要した回数

シード値	収束した回数
1000	96
2000	90
3000	111
4000	109
5000	93
6000	99
7000	100
8000	114
9000	113
10000	94
10 試行の平均値	101.9

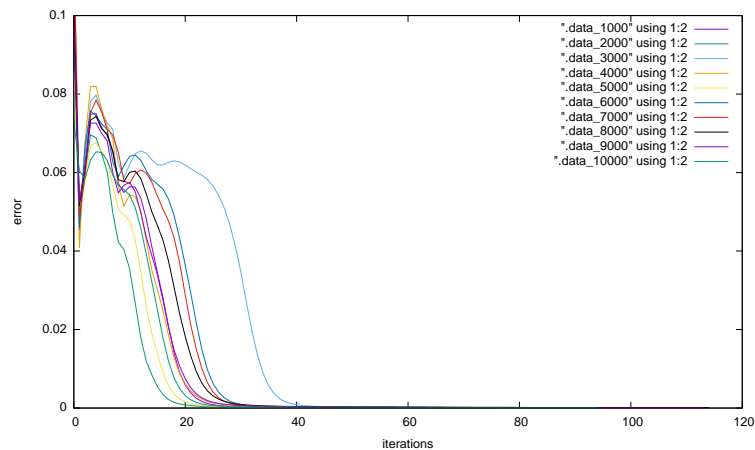


図 1: 重みを更新する様子

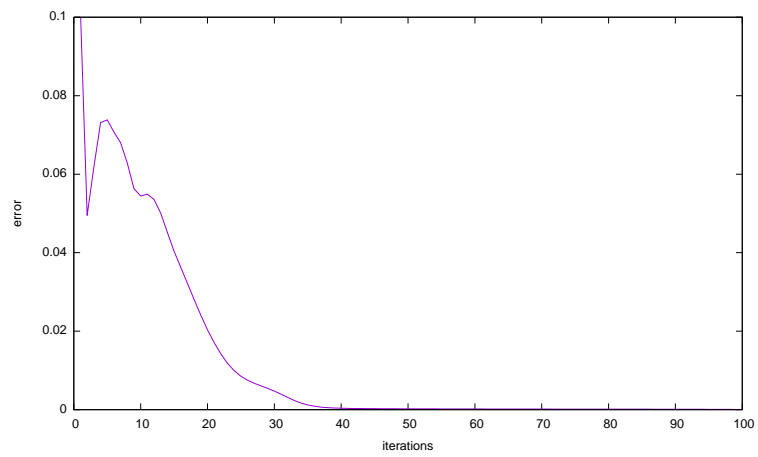


図 2: 重みを更新する様子 (平均値)

### 1.2.2 考察

実験よりわかったことを以下に箇条書きする．

- すべて FINISH2 で終わったため，誤差が最小誤差より小さくなっている．
- 実験が収束したのは 100 前後で，seed 値をこれ以上大きくしてみても同じように 100 前後で収束すると推測する．
- 学習曲線は，error が 0.05 くらいまで急激に下がり，その後若干上がり，そのあと最小誤差へと遷移している．

## 2 Level2: 線形分離不可能な ExOR 問題への適用

### 2.1 課題説明

階層型ニューラルネットワークを ExOR 問題へ適用し、線形分離できない問題においても学習可能であることを確認する。特に Level2 では、この問題を解決するために中間層を導入することで拡張した階層型ニューラルネットワークにより学習可能であることを確認する。

### 2.2 階層型 NN による学習

#### 2.2.1 最適なパラメータを探すためのアプローチ

まず最初に seed 値を 1 に固定し適当なパラメータの値を決め実行、を繰り返し、その実行結果の中から学習回数が少ないもののパラメータの傾向を探り出した。HIDDEN を 10 にすると安定して学習回数が減っているように見えたため、HIDDEN を 10 に固定し、その後今まで試行したものの中から一番学習回数が少なかった ETA と alpha の値を元に微調整を繰り返した。

上記の方法により以下のパラメータを求めた。

ETA=1.65

ALPHA=0.97

HIDDEN=10

#### 2.2.2 実行結果

シー ド値を変えた際の学習収束回数を以下に示す。

表 2: 階層型 NN による ExOR 問題の学習に要した回数

シード値	収束した回数
1000	29
2000	39
3000	89
4000	99
5000	57
6000	81
7000	28
8000	49
9000	43
10000	63
10 試行の平均値	57.7

以下に 10 試行分の学習曲線と、その平均学習曲線を示す。

#### 2.2.3 考察

グラフは、x 軸が収束回数、y 軸が誤差に対応している。seed 値 1000 10000 まで実行したときの平均のグラフから、平均して誤差が 0 に限りなく近づくのは、試行回数 80 90 前後ということがわかる。

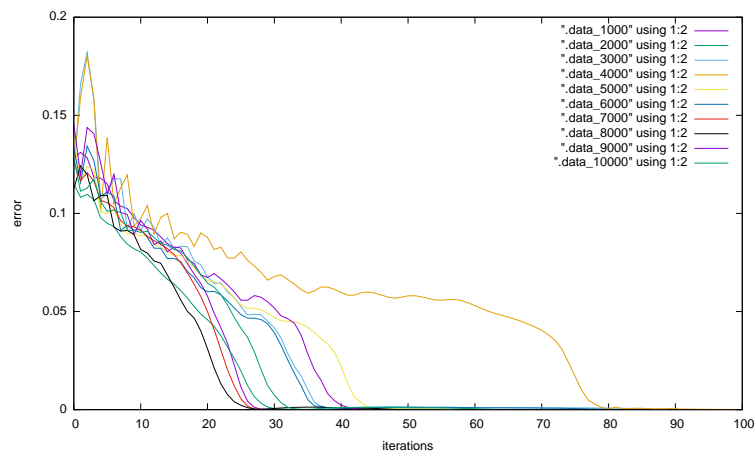


図 3: seed 値 1000 10000 まで実行したときの結果

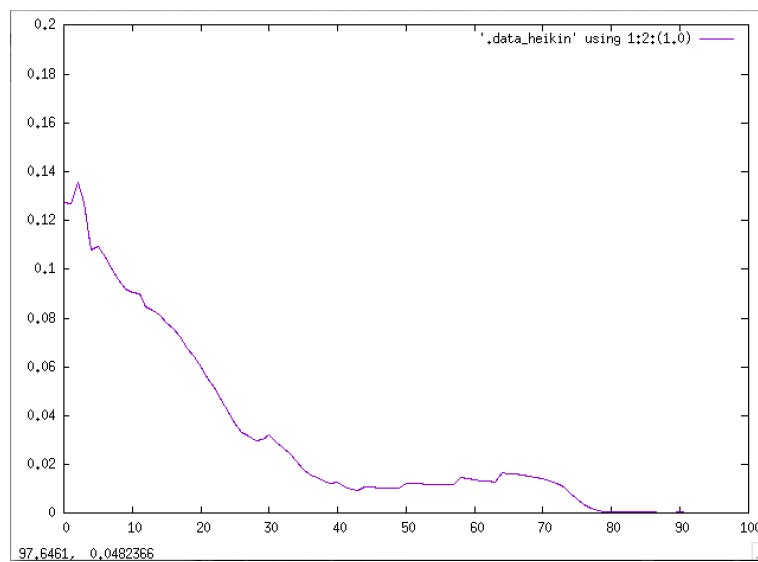


図 4: seed 値 1000 10000 まで実行したときの平均

### 3 Level3: 応用事例：文字認識問題への適用

#### 3.1 課題説明

階層型 NN を文字認識に適用し、考察する。特に、用意された教師データと認識のしやすさに関する関係性や、学習最適化のためのパラメータのチューニングおよび、より柔軟性の高い認識方法に関する検討を行う。

#### 3.2 Level3.1: パラメータのチューニング

##### 3.2.1 最適なパラメータを探すためのアプローチ

指定された条件下において学習が効率良く行われるパラメータの組み合わせを探すため、\*\*して\*\*することでパラメータを調整した。

(補足：全パターンを調べても良いし、いくつかのパターンを調べても良いが、どのような方法で調整したら良いかを考えよう。その上で、自分たちがどのように取り組んだのか(=アプローチ)を説明しよう。)

3.2.2 実行結果

（補足：ボーナスポイントの確認がありますので、シード値 10 パターンで試した際の収束に要した学習回数と、その平均回数が分かるように明示してください。）

表 3: 階層型 NN による文字認識問題の学習に要した回数

シード値	収束した回数
100	hoge
200	hoge
300	hoge
400	hoge
500	hoge
600	hoge
700	hoge
800	hoge
900	hoge
1000	hoge
10 試行の平均値	hoge

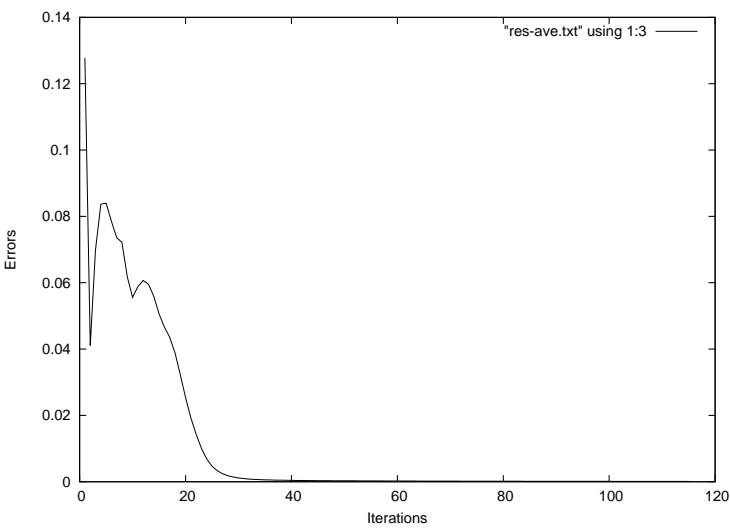


図 5: 重みを更新する様子（平均値）

3.2.3 考察

3.3 Level3.2: パラメータと収束能力の関連性について

3.3.1 関係性を確認するためのアプローチ

3つのパラメータがどのような関係にあるかを検証するため、ケース 1,2,,, を設定し、学習曲線からその関係性について考察する。

### 3.3.2 結果

### 3.3.3 考察

## 3.4 Level3.3: 任意の評価用データを用いた評価

### 3.4.1 アプローチ

(仮説 1)

学習時のデータ(教師データ)との違いが少ない程認識率が高く、逆に教師データとの違いが多い程認識率が低くなる  
との仮定の下、如何に示す評価データを用意した。

- a\_deter1.txt
- a\_deter2.txt
- a\_deter3.txt
- a\_deter4.txt
- a\_deter5.txt
- a\_deter6.txt

(仮説 2)

学習時のデータ(教師データ)との違いが多い程認識率が高く、逆に教師データとの違いが少ない程認識率が低くなる  
との仮定の下、評価データは仮説 1 と同様のものを用いる。

(仮説 3)

学習データの一部分だけの場合、一部分が似ているものの認識率が高いという仮定の下、「あ」のデータを上下左右の半  
分だけのデータを用意した。また他の文字の左半分だけのデータも用意した。

- a\_up.txt
- a\_down.txt
- a\_left.txt
- a\_half.txt
- ka\_half.txt
- sa\_half.txt
- ta\_half.txt
- na\_half.txt
- ha\_half.txt
- ma\_half.txt
- ya\_half.txt
- ra\_half.txt
- wa\_half.txt

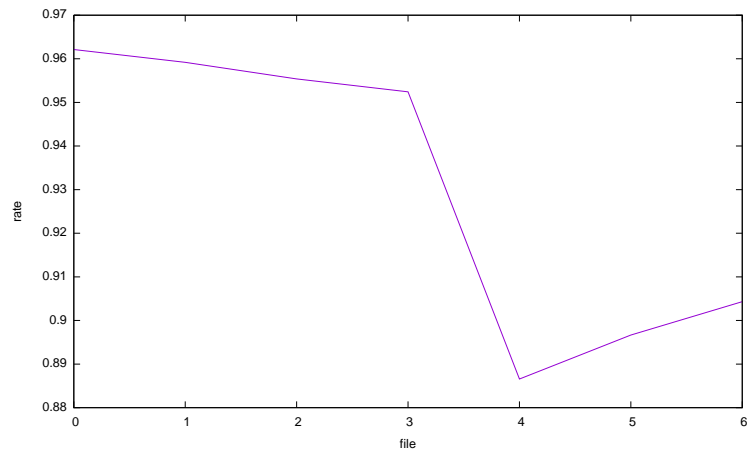


図 6: 仮説 1: 認証率 (rate) と違い

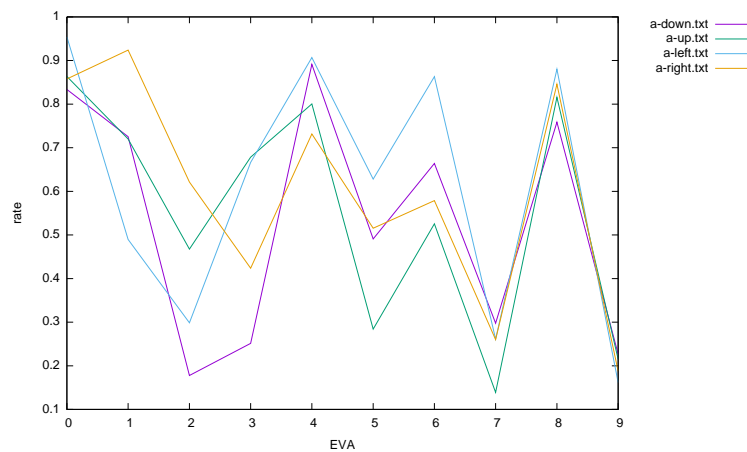


図 7: 仮説 3: 「あ」のデータ

### 3.4.2 結果

上記の X 軸は, 0 が学習データと同じデータ, 1 以降は a\_deter.txt の番号となっている. またこの番号が上がるにつれて, 元データよりデータを劣化 (1->0) にしている.

### 3.4.3 考察

結果として

- 仮説 1,2: 違いが増えるほど認証率は低くなるが, ほとんどデータを劣化させた場合少し高くなる.
- 仮説 3: 「あ」のデータを移動させた場合, 「あ」としての認証率が高く, 他にも「な」や「ら」, 「か」などの文字としての認証率も高かった.
- 仮説 3: 学習データ全ての文字を半分にしてみた場合, 「あ」, 「か」, 「な」, 「ら」の認証率が高かった.

というデータを得た.

図 6 のデータの中で, ほとんどデータを劣化させた a\_deter5.txt と a\_deter6.txt の認証率が若干高くなったのは, ほとんどデータが少なく判断できなかったことで全体的に認証率が上がったのではないかと推測する.

図 7,8 より, 似た部分の多い「あ」, 「か」, 「な」, 「ら」の 4 つの認証率が高いことがわかった.

よって, 仮説 1 と 3 が正しいという結果になった.



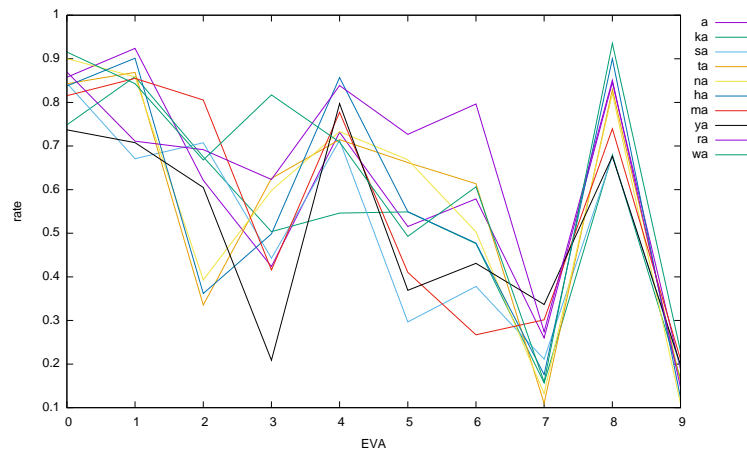


図 8: 仮説 3:学習データすべての半分

### 3.5 Level3.4: 認識率を高める工夫

#### 3.5.1 対象とする問題点

入力されたデータが想定していた入力と比べてサイズが異なったり、位置がずれている等、文字の一部が欠けている以外にも多様な要因によるデータ（情報）の劣化が考えられる。認識率を高めるにはどのような点を工夫すれば良いか？どのような方法でも構わないので、検討せよ。

#### 3.5.2 改善方法の提案

認識率を高めるために改善すべき方法として、学習データの数を増やすということがあげられる。

#### 3.5.3 考察

今回の実験では、学習サンプルデータがそれぞれ1種類ずつしかなかった。そのため、いろいろなパターンの文字の形を学習させる際に少々学習に手間取ったように見える。この問題を解決するための方法として学習データの増加を行う必要がある。

人の脳をモデルとしているのならば、人が行うような学習をさせればいいのだから、いろいろなデータを蓄えておいて、それを元に分析を行うことはごく自然なことである。一つの事象にとらわれることなく学習を行うために、データの増加は必要だと考えた。1つの文字に対して複数のパターンを用意し、それに基づいて学習をさせれば、与えたデータをより細かく学習させることができるだろう。

また、複数のパターンを用意して、それに共通する部分や同じような形を記憶させておけば、さらに学習パターンは広がるはずだ、これを実現させるためにはさらにコードを書かないといけなくなるので、容易にできるのはデータ数を増やすことだと考えた。

## 4 その他: 実験の内容・進め方に関するコメント等

(補足: 今後の為に参考にしたいので、情報工学実験 2・探索アルゴリズム 1,2 で扱った内容、実験の進め方等について意見があれば書いてください(当然、どのような意見であってもレポートの評価を下げる事はしません。)。 「授業評価アンケート」の際に書いてもらっても構いません。)

(補足: 参考文献は thebibliography 環境を使って列挙し、本文中で適切な箇所で引用するようにしましょう。例えば下記文献は、アブストラクト中で引用しています)

## 参考文献

[1] 情報工学実験 2: 探索アルゴリズムその 2 ( 當間 )

<http://www.eva.ie.u-ryukyu.ac.jp/~tnal/2011/info2/search2/>