

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: 金曜 & グループ 9

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。）

- 155706J 久場翔悟: 担当 Level1
- 155711E 平木宏空: 担当 Level
- 155716F 石塚海斗: 担当 Level3.4
- 155730B 清水隆博: 担当 Level2

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**して**した。具体的には云々。

表 1: OR 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

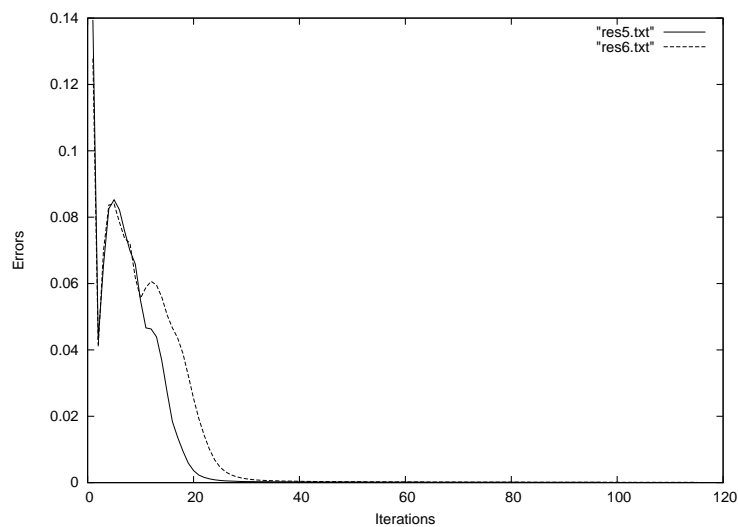


図 1: 重みを更新する様子

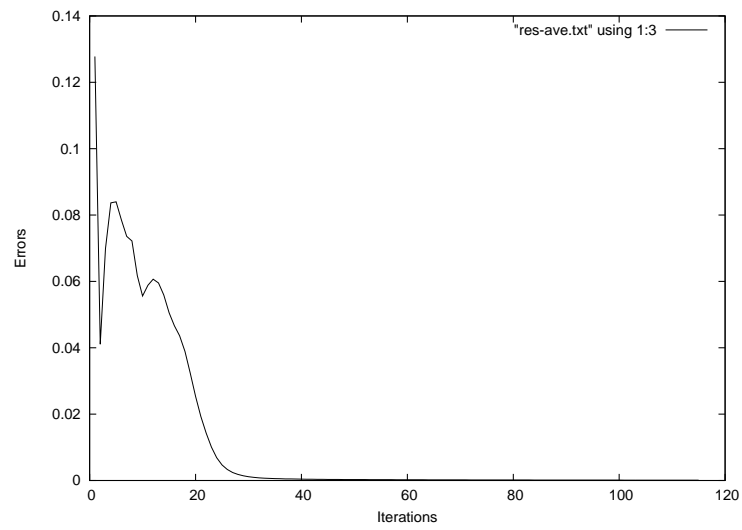


図 2: 重みを更新する様子 (平均値)

1.2.2 考察

(結果から分かったことは?)

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

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)

2.2.2 実行結果

(補足: 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 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
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

2.2.3 考察

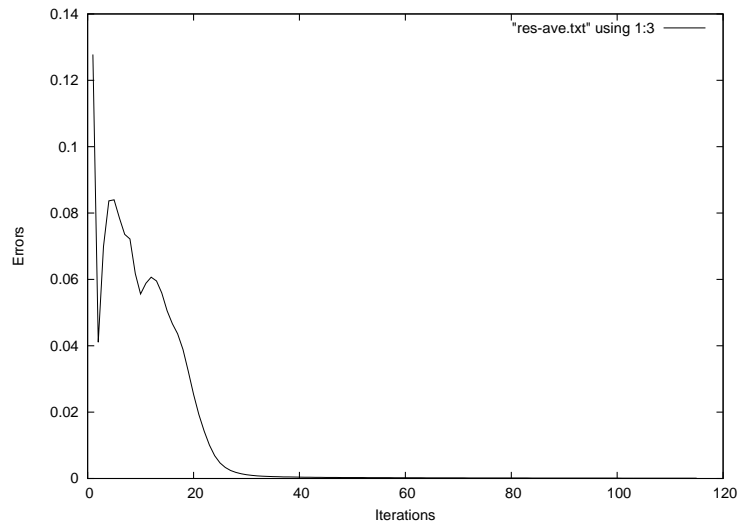


図 3: 重みを更新する様子（平均値）

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、3つのパラメータ ETA, ALPHA, HIDDEN を一つずつ変動させ、総当たりすることで最適なパラメータを探索した。具体的には ETA を 0.00 - 1.98 までの 0.01 刻み, ALPHA を 0.00 - 1.98 までの 0.01 刻み, HIDDEN を 1 - 15 までの 1 刻みの計 594015 回 で変動させていき、探索を行った。

探索を行った各パラメータによる実行結果より、最も iteration が低いパラメータ群を最適な組み合わせとして決定するという方法をとった。// また,seed 値については、各 seed 値についても総当たりを行いたいが、それでは、実行時間がレポート提出期間内に終わらないため、今回の総当たりは seed 値を 1000 に固定して行った。

3.2.2 実行結果

seed 値 1000 において パラメータを総当たりで探索した結果,iteration が最低となったパラメータは ETA 1.83 ALPHA 0.73 HIDDEN 15 であった。

このパラメータを用いて各 seed 値で実行した際の学習収束回数を以下の図 3 に示す。

以下の図 4 に求めたパラメータを用いた際の学習曲線を示す。

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	173
200	258
300	213
400	250
500	369
600	196
700	121
800	204
900	1325
1000	69
10 試行の平均値	317.8

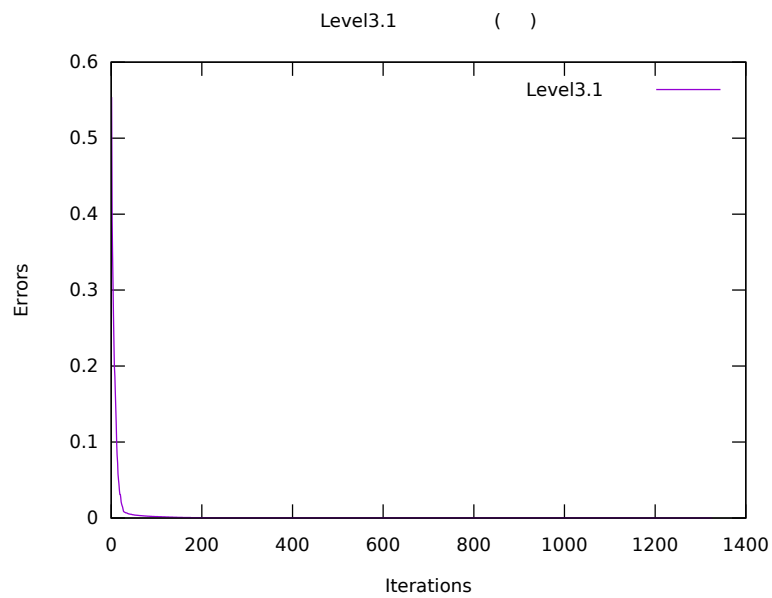


図 4: 重みを更新する様子（平均値）

上記の図 4 では, 変化が見えづらいため変化の著しい iteration 50 までの図を下図 5 に示す。

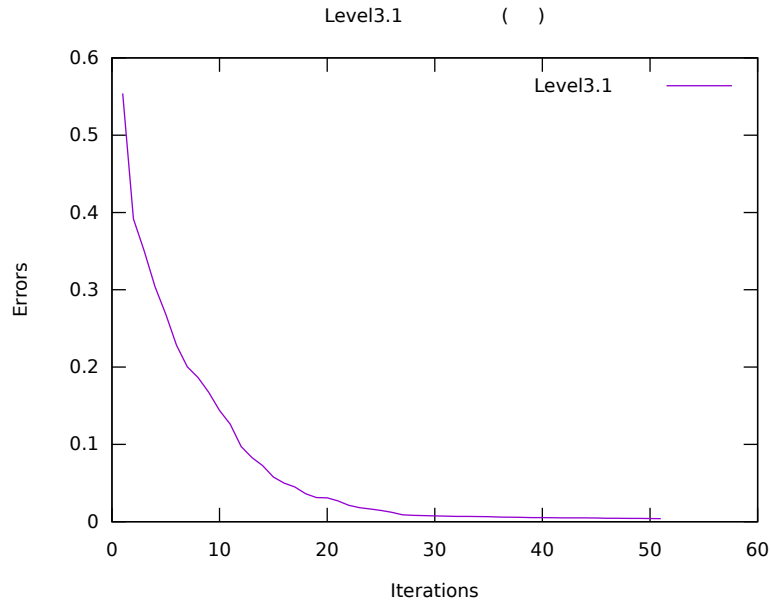


図 5: 重みを更新する様子（平均値）拡大図

3.2.3 考察

今回 seed 値を固定してパラメータの組み合わせの総当たりを調べたのには, seed 値を可変にして総当たりを行うと膨大な時間がかかるという理由と, seed 値はニューロンへの入力的重要度の初期値を決定するために用いられるだけであるから, 重要度の初期値がどうであれ, ETA, ALPHA, HIDDEN のパラメータが最適な組み合わせであれば, 学習に影響はないであろうという大雑把な予測からであった。しかし, 今回の実験結果 (図 3) を見ると, seed 値 1000 と seed 値 900 の学習収束回数間には 1200 回以上の差がある。このことから, ニューロンへの入力信号の重要度も学習の収束回数へ影響を及ぼし, "最適なパラメータの組み合わせ"には ETA, ALPHA, HIDDEN の 3 つの他に seed 値 も含まれたものであったと考察する。

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, 3 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

level3.3 では文字認識プログラム nn_num において, 評価用データが教師用データと異なる場合, 認識結果に違いが見られるかを実験する。

3.4.1 アプローチ

実験を行うにあたり、学習用データと評価用データがずれていくにあたり、認識率がどう変動するか仮説を検討した。文字認識の場合、計算機に学習させるアルゴリズムによって最終的な認識に至るまでのプロセスが異なると考える。しかし、あくまで認識は学習用データを利用し行う。教師あり学習の場合、明確な答えが定義されている学習用データそのものが入力された場合の認識率は最大であると考え。また、今回の実験で使用した文字認識プログラムは0か1の数値データに基づいて学習を行う。人間では誤差の範囲であるが、計算機の場合1bitの変化で計算した内容が大幅に変化していくのではないかと推測した。従って、学習用データからずれていくに連れて文字を誤認識する確率が上昇するのではないかと仮説を立てた。

3.4.2 結果

3.4.3 考察

3.5 Level3.4: 認識率を高める工夫

3.5.1 対象とする問題点

今回対象として考えた問題点は、入力と比べサイズが異なる場合、及び、位置がずれている場合である。

3.5.2 改善方法の提案

上記問題に対して提案された改善方法は、入力データを細かくいくつかのブロックに分け、"1"が入力されているブロックだけに対してのみ、認識をかけるという方法である。ブロックに分け、入力のあるブロック(今回のテストデータにおける"1"が入力されているブロック)に対してのみ相対的な位置やその傾向を用いて認識することで、サイズが異なる場合や位置のずれに対応できると考えた。また、今回の改善方法における正確性の向上について、分けるブロック数を増やす、ブロック化したデータをさらにブロック化する(ブロック化を多重に行う)等の提案もされた。

3.5.3 考察

今回、対象とした問題は入力データ(今回のテストデータにおける"1"が入力されている部分)の全体的な位置は変化するものの、"1"が入力されているデータ間の相対的な位置が変化しないものであるため、相対的な位置を学習することで認識率を向上させるという方法をとった。しかし、この方法では入力データにノイズが混入する等の、テストデータ内の"1"が入力されてるデータ間の相対的な位置が変化するような問題に対応できていない。このことから、グループ内の討論では、各問題に対しての改善方法を複数実装した方が入力データを正しく認識できる確率が向上するという結果に至った。

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（当間）

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