

線形分離を可能とする改良型パーセプトロンの使用メモリ改善

渡辺 雄二*

Improvement of Using Memory on Modified Perceptron Enabling Linear Separation

Yuji Watanabe*

要旨： これまで、今では注目されなくなってしまったパーセプトロンまでさかのぼり、パーセプトロンの線形分離不可能問題に対して、二点分離法という手法を用いてこの問題を解決した。しかしながら、入力する数値が大きくなると中間層のニューロン数が増大し処理の使用メモリが膨大になるという問題が残されていた。この理由は、従来手法では中間層ニューロン数が入力する数値に対する指数関数の関係になっていたためである。今回、この中間層ニューロン数は教師データの数だけ用意すれば済むことがわかり、これを試行した。この結果、中間層ニューロン数を削減でき、これにより使用メモリを削減する改善ができた。これによって、ニューラルネットワークの新たな可能性を見出した。

キーワード： パターン認識、ニューラルネットワーク、線形分離不可能問題、二点分離法、トンネル切羽評価問題

1. はじめに

昨今、日刊工業新聞のみならず、各マスメディアから「AI (Artificial Intelligence, 人工知能)」という言葉が発信され続けている。現在、第3次人工知能ブームの真っただ中にあるとあってよく、ニューラルネットワークの発展系であるディープラーニングはその中核的な存在である。2012年、世界的な画像認識のコンペティション ILSVRC (Imagenet Large Scale Visual Recognition Challenge) で、東京大学、オックスフォード大学、イエーナ大学(ドイツ)、ゼロックスなどの名だたる研究機関が開発した人工知能を抑えて、初参加のカナダのトロント大学が開発したスーパー・ヴィジョンが圧倒的な勝利を飾った¹⁾。このコンペでは、ある画像に写っているのがヨットなのか、花なのか、ネコなのかをコンピュータが自動であてるタスクが課され、その正解率の高さ(エラー率の低さ)を競い合う。トロント大学は、他の人工知能を10ポイント以上引き離して、いきなりエラー率15%台をたたき出した。何がトロント大学に勝利をもたらしたのか。それは同大学教授のHintonが中心となって開発した新しい「機械学習」の方法、「ディープラーニング(深層学習)」であった。一方、2016年11月14日の日

本経済新聞(電子版)に「AIで東大合格断念『東ロボくん』偏差値伸びず」という記事が掲載された。内容は、人工知能で東京大合格を目指す「東ロボくん」の開発を進めてきた国立情報学研究所などは、2016年度の大学入試センター試験の模試で偏差値57.1を獲得したが、昨年からはほぼ横ばいで、東大合格圏には達しなかった。今後は東大合格を目標にせず、中高生の読解力を高める研究などに注力する考え、ということである。新井らによれば、最終的に70%以上の大学で合格可能性80%以上を達成した。その一方で、英語では19億文という大量の英文を学習させ、様々なディープラーニング手法を試したにもかかわらず、比較的容易な会話文の穴埋め問題において明確な成果を上げることができなかった²⁾、ということである。

パーセプトロンは、1950年代の後半に、パターン認識を行う機械としてRosenblattによって提案された³⁾。これは、その後約10年間続いたニューラルネットワークブームの中心モデルとなった。しかし、Minskyらは、パーセプトロンは線形分離の可能な学習機械ではあるが、これが不可能なケースもあり、場合によっては中間層のニューロン数が膨大になることや収束速度が遅すぎるということを数学的に証明した⁴⁾。これは、一般に「パーセプトロンの限界説」として知られている。この結果、パーセプトロン研究への関心が急速に薄れて

* 山形県立産業技術短期大学校庄内校
〒998-0102 山形県酒田市京田三丁目57-4
e-mail: watanabe@shonai-cit.ac.jp

* Shonai College of Industry & Technology
3-57-4 Kyoden, Sakata City, Yamagata, 998-0102, Japan
e-mail: watanabe@shonai-cit.ac.jp

いった。その後、1986 年にこの解決法として Rumelhart らにより、バックプロパゲーション法が提案された⁵⁾。その後、近年、前述のとおりディープラーニングが脚光を浴びている。ディープラーニングはニューラルネットワークを何段階にも深く（ディープに）積み重ねる手法であり、学習法としてはバックプロパゲーション法などを用いている。一方、パーセプトロンは教師あり学習機能を有する基本モデルとして、ニューラルネットワーク研究の原点に位置するものと考えられる。そこで、パーセプトロンの構成要素であるネットワーク構造や学習アルゴリズムを工夫して線形分離が可能となれば、別の系統からニューラルネットワークの発展に貢献できる可能性がある。筆者らは、これまで各ニューロンの入出力を 1 と 0 の二値に限定し、中間層のニューロンで学習するネットワークモデルを提案し、二点分離法と名づける学習アルゴリズムによって、線形分離が可能となることを示した⁶⁾。しかしながら、この手法では入力する数値が大きくなると中間層ニューロン数が増大し、実用に耐えないことも判明している。今回、この問題を解決するための検討を行った。

2. パーセプトロンの原理

図 1 に脳神経細胞（ニューロン）の概略を示す。各ニューロンの樹状突起は、シナプスを介して他のニューロンからの信号を受け取る。ニューロン間の信号伝達は電気パルスによって行われ、これにより、受け取ったニューロンの状態が変化する。ニューロンは受け取った電気信号の総和が、ある閾値を超えると発火（パルスを発生）し、軸索を通じてそのパルスを伝え、他のニューロンに影響を及ぼす。

このニューロンに対して、1943 年に McCulloch と Pitts は図 2 に示す人工的なモデルニューロン（以下ユニットと呼ぶことにする）を提案した⁷⁾。このユニットを複数個組み合わせたものをニューラルネットワークと呼んでいる。ユニットは、入力 x_i を受け取るがその間に w_i という大きさの結合荷重が係っている。ユニットは、

$$y = w_1x_1 + w_2x_2 + \cdots + w_nx_n = \sum_{i=1}^n w_i x_i \quad (1)$$

の入力を受け取る。ユニットは、 y がある閾値 θ を超えれば $z = 1$ となり、超えなければ $z = 0$ となる。

これを式に表すと、

$$z = 1[y - \theta] \begin{cases} 1 & (y - \theta > 0) \\ 0 & (y - \theta \leq 0) \end{cases} \quad (2)$$

となる⁸⁾。

Hebb は 1949 年に、シナプスの伝達効率が変化するシナプス強化モデルを示した⁹⁾。これは、あるニューロンが発火したとき、それに信号を伝えたニューロンとのシナプス結合は強まるというものである。これを応用し、正しい出力が 1 であるのに 0 と出力した場合は、ユニット間の結合荷重を増加させ、閾値を下げる処理を行う。これは、

$$w \leftarrow w + \epsilon x \quad (3)$$

$$\theta \leftarrow \theta - \epsilon \quad (4)$$

とする（ ϵ は小さな正の定数）。また、これとは逆に、正しい出力が 0 であるのに 1 と出力した場合は、次のように行う。

$$w \leftarrow w - \epsilon x \quad (5)$$

$$\theta \leftarrow \theta + \epsilon \quad (6)$$

式(3)、(4)及び式(5)、(6)の更新処理を繰り返すことによって、所望の入出力関係（教師データ）を実現できることになる¹⁰⁾。これがパーセプトロンの学習則で、誤り訂正学習と呼ばれるものである。本稿では、この 1 個のユニットと学習則を組み合わせたものを基本パーセプトロンと呼ぶことにする⁸⁾。

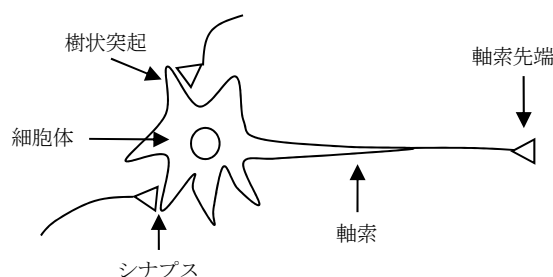


図1 神経細胞（ニューロン）の概略

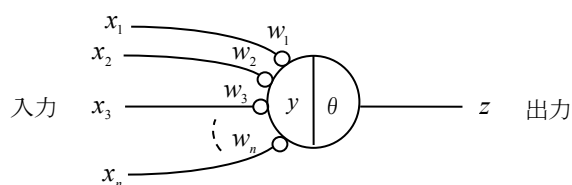


図2 モデルニューロン（ユニット）

3. 二点分離法による解法

各ユニットの入出力を 1 と 0 の二値に限定し、2 入力 1 出力のケースについて検討を進める。式(1)、(2)より、2 入力の基本パーセプトロンは、 $z = 1$ の場合に注目すると、

$$w_1x_1 + w_2x_2 - \theta > 0$$

となり、これは一次不等式である。係数 w_1 、 w_2 及び定数項 θ の値によっても変わってくるが、例えば図 3 の斜線で示した部分がこの一次不等式の表す領域である。2 入力の基本パーセプトロンは、平面を直線で分離する機能を持つ。そして、基本パーセプトロン 1 個だけでは、排他的論理和 (XOR) と論理一致 (XNOR) の 2 ケースについて、これを実現できない。図 3 で、白丸を 0、黒丸を 1 とすれば、これは XOR を表しており、この分離が不可能なことは明らかである。

筆者らは、これまでの研究において、中間層の各ユニットが二点ずつに対応する、二点分離法と名づけた改良型パーセプトロンを用いて、この問題を解決している⁶⁾。これは、図 4 に示すような 3 層パーセプトロンを用い、出力層のユニットをあらかじめ論理和 (OR) に固定し、中間層ユニットで学習する方法である。なお、入力層ユニットは 1 か 0 かの値を中間層ユニットに伝える役割を果たす。学習法のポイントは、ユニット C の出力が 1 となるケースにおいては、ユニット A とユニット B の両方が 1 となるような学習は行わず、ユニット A は(0,0)、(0,1)の 2 点について、またユニット B は(1,0)、(1,1)の 2 点についてのみの学習を行うことである。なお、ユニット C の出力が 0 となるケースにおいては、ユニット A、B とも 0 となる学習を行う。これによって、XOR と XNOR への対応が可能となり、全ての組み合わせ論理回路を実現できるようになった。

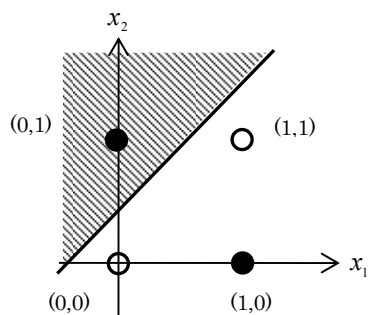


図 3 線形分離性

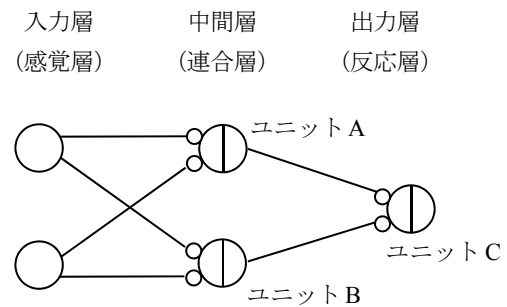


図 4 3 層パーセプトロン

4. 二点分離法の問題点

二点分離法では、入力層ユニットと出力層ユニットに教師データを与えて学習させる場合、1 と 0 の二値に限定する。このため、例えば画像データなどを扱う場合は、データを 1 ビットずつに分解しこれをユニットに対応させる操作が必要となってくる。このため、多入力多出力の検討が必要となってくる。

まず、3 入力 1 出力については図 5 に示すように立方体となり、入力データ点の数は最大 8 点となる。これに対する二点分離法の適用を考えると、二点ずつペアにして線形分離をするので、中間層のユニット数は 4 個となる。このことから、入力層のユニット数を n とすると、中間層のユニット数は、

$$\frac{2^n}{2} = 2^{n-1} \quad (7)$$

となり、入力層のユニット数に対して指数関数の関係にある。次に、3 入力 2 出力について検討する。この場合のネットワーク構造は、図 6 のようになる。二点分離法では、出力層ユニット 1 個毎に、それぞれ独立した中間層ユニット群が必要となってくる。したがって、多入力多出力の中間層のユニット数 m は、入力層のユニット数を n 、出力層のユニット数を k としたとき、

$$m = k2^{n-1} \quad (8)$$

となる。

具体的には、入力層のユニット数が 32、出力層のユニット数が 1 の場合は、

$$m = k2^{n-1} = 1 \times 2^{31} = 2,147,483,648$$

となる。このようなケースでは、計算機を用いて処理する場合、メモリの確保が困難となってくる。

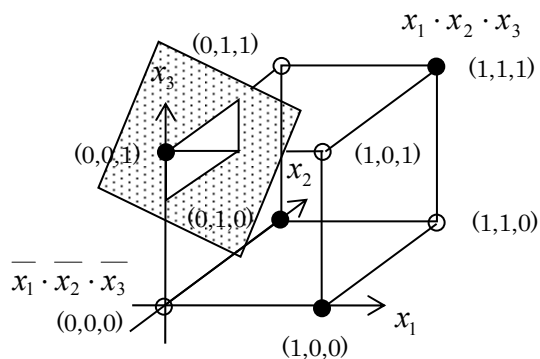


図5 3入力1出力の場合

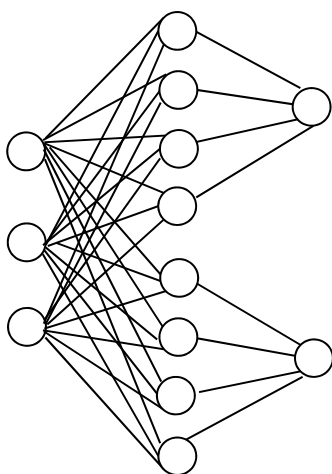


図6 3入力2出力のネットワーク構造

5. 使用メモリ改善の検討

中間層のユニット数を削減するための検討を行った。

二点分離法では、出力が0となる学習の場合は、全ての中間層ユニットの出力が0になる学習を行う。出力が1となる学習の場合は、1個の中間層の担当ユニットのみが1になる学習を行っている。このことは、何を意味しているのか。出力が1になるケースは、どれか一つの中間層ユニットが1になる学習を行えばよいということを示している。このように考えると、あらかじめ式(8)のような膨大なユニットを用意する必要はなく、最大でも教師データの数だけ用意すれば十分である。そして、出力が1となるケースは教師データの並びに対して中間層ユニットを対応させ、対応する一つのユニットのみが学習するようにすればよい。

具体的な検討を進める。まず、表1の学習を行うケースを考える。教師データ数は4であるので、中間層ユニットを4個用意する。そして、教師データ1から4の学習を担当する中間層ユニットを図7のように一つずつ対応させる。このようにすると、中間層ユニットを削減するという目標とは一見矛盾するようにみえる。しかしながら、教師データ数はそれほど多くない場合が多い。例えば、ニューラルネットワークなどのベンチマークデータとして用いられるMNIST（エムニスト）という手書きの数字画像セットの場合、計70,000枚の画像が含まれており、そのうち60,000枚はトレーニングセット（教師データ）で10,000枚はテストセット（推定データ）である¹⁾。教師データが60,000枚であるならば、問題なく対応可能である。

中間層ユニット数を教師データと同数にする学習アルゴリズムを一般化してまとめる。次の三つから構成される。なお、出力層のユニットは、あらかじめORに固定しておく。

(i) 出力層ユニットの出力が1となるべきところ0であれば、教師データに対応する一つの中間層ユニットのみが1となる学習を行う。

(ii) 出力層ユニットの出力が0となるべきところ1であれば、中間層ユニットからの出力が1となった全てのものについて0となる学習を行う。

(iii) 全ての教師データが満足されるまで学習を繰り返す。

表1 XORの真理値表

x_1	x_2	z
0	0	0
0	1	1
1	0	1
1	1	0

… 教師データ1
… 教師データ2
… 教師データ3
… 教師データ4

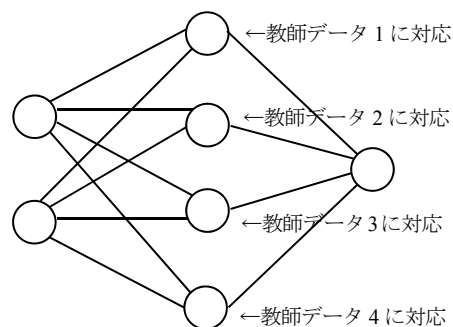


図7 使用メモリ改善案

6. 試行

前章で検討した学習アルゴリズムで、表 1 を学習し、再現（推定）できることを確認した。

この後、付表 1 と付表 2 に対して試行した。これらはトンネル切羽評価用のデータであり⁹⁾、トンネル掘削時、岩盤の圧縮強度、風化変質、割目間隔、割目状態、湧水量、劣化に対して評価点を与えるというデータ構造になっている。この評価点は、トンネルを支える工法（様々なバリエーションがある）を選択する際の指標となり、100 点に近くなるほど岩盤強度が高く安価な工法で済む、ということになってくる。この評価点は、専門化が決定したものであるが、これまでこれを改良型パーセプトロンで学習して推定し、これによって専門家のノウハウを自動で評価算出できないかという研究を行ってきた⁹⁾。付表 1 と付表 2 は、この研究時と同じデータとなっている。

入出力層のユニット数は、数値を二進数で表した場合のビット数だけあればよい。入力層ユニットは、湧水量に対しては 2 ビット、これ以外は 3 ビットずつ、合計 17 ビットであるので 17 個用意する。出力層ユニットは、評価点の数値は符号ビットを含めて 8 ビットあればよいので 8 個とした。従来手法での中間層ユニット数は、 $8 \times 2^{17-1} = 524,288$ であった。今回の提案手法では、 $8 \times 50 = 400$ （50 は教師データ数）で済む。従来手法と提案手法で付表 1 を学習後、付表 2 の未知データを入力し推定させた。結合荷重と閾値の初期値は、両手法とも全て 0 とした。中間層ユニット数と相関係数を表 2 にまとめ、付表 2 の最右例に提案手法による推定値を記載した。相関係数は、専門家による評価点と改良型パーセプトロンが推定した値がどの程度相関しているのかを求めたものであり、1 に近くなればなるほど相関が高い。つまり、汎化能力が高いことを表している。提案手法は、従来手法に対して遜色のない相関係数を示している。なお、学習と推定を合わせた処理時間は、提案手法では従来手法の約 1,000 分の 1 となった。

表 2 トンネル切羽評価の試行結果

	中間層ユニット数	相関係数
従来手法	524,288	0.788
提案手法	400	0.771

7. おわりに

これまで、パーセプトロンの線形分離不可能問題に対して、二点分離法という手法を用いてこの問題を解決してきた。しかしながら、入力する数値が大きくなると用意すべき入力層ユニットが多くなり、このため中間層のユニット数が増大して処理の使用メモリが膨大になるという問題が残されていた。この理由は、従来手法では中間層ユニット数が入力する数値に対する指数関数の関係になっていたためである。今回、中間層ユニット数は教師データ分だけで済むことがわかり、これを試行した。この結果、中間層ユニット数を削減でき、使用メモリを削減する改善ができることがわかった。これによって、ニューラルネットワークの新たな可能性を見出した。

使用メモリをさらに削減する方法を考えてみると、次のようなアイデアがある。出力層ユニットの出力が 1 となるケースの学習は、対応する中間層ユニットが 1 となる学習を行えばよいのであるから、教師データの中で 1 となるケース分だけの中間層ユニットを用意すれば十分である。これを行えば、さらに半分程度にすることができる。しかしながら、これについてはこれまでの処理との整合性が取れるかを検討する必要があるので、今後の課題としたい。

文 献

- 1) 松尾豊：人工知能は人間を超えるか、KADOKAWA、東京、263p. (2015)
- 2) 新井紀子他：小特集「ロボットは東大に入れるか」という企て、情報処理、Vol.58, No.7, pp.598-617 (2017)
- 3) Rosenblatt, F. : *Principles of neurodynamics*, Spartan, Washington, D.C, 616p. (1962)
- 4) Minsky, M. L. and Papert, S. A. : *Perceptrons*, The MIT Press, Cambridge, Massachusetts (1969) (中野他訳：パーセプトロン (改訂版), パーソナルメディア、東京、340p. (1993))
- 5) Rumelhart, D. E., Hinton, G. E. and Williams, R. J. : Learning representations by back-propagating errors, *Nature*, Vol.323, No.9, pp.533-536 (1986)
- 6) Watanabe, Y., Kitajima, T., Kimura, M. and Harada, K. : Solution with modified perceptron to tunnel cutting face evaluation problems, *Geoinformatics*, Vol.17, No.2, pp.61-70 (2006)
- 7) McCulloch, W.S. and Pitts, W. : A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, Vol.5, pp.115-133 (1943)
- 8) 中野馨編：ニューロコンピュータの基礎、コロナ社、東京、248p. (1990)
- 9) Hebb, D. O. : *The organization of behavior*, John Wiley and Sons, New York (1949) (白井常訳：行動の機構、岩波書店、東京、413p. (1957))
- 10) 中野馨編：C でつくる脳の情報システム、近代科学社、東京、259p. (1995)

付 録

付表1 トンネル切羽評価用教師データ

圧縮強度	風化変質	割目間隔	割目状態	湧水量	劣化	評価点
4	3	4	4	1	1	33
4	2	2	4	2	1	45
4	2	5	4	2	1	36
5	3	4	4	1	1	25
6	4	5	5	2	2	-5
5	3	4	4	2	2	20
5	3	5	4	1	1	22
5	3	4	4	1	2	25
5	4	4	5	1	1	11
6	3	5	5	2	2	2
6	4	5	5	1	2	0
6	4	5	5	1	1	0
4	2	4	4	2	2	34
4	2	4	3	1	1	46
4	2	5	5	1	1	29
6	4	4	5	2	2	-2
4	2	2	4	2	2	40
5	4	5	5	2	3	1
5	3	5	5	1	2	15
4	2	5	4	1	2	36
5	2	5	4	1	2	28
5	3	5	4	2	2	17
4	2	4	4	1	1	39
5	4	5	5	1	1	8
5	4	5	5	2	2	3
6	3	4	5	3	4	-5
4	2	4	4	2	1	39
4	3	3	3	1	1	43
4	3	5	4	1	1	30
5	4	4	5	1	2	11
4	2	2	3	2	1	52
4	2	4	5	2	2	27
5	3	4	5	1	1	18
5	3	5	5	1	1	15
3	2	3	4	2	1	50
6	4	4	5	1	2	3
6	3	4	4	1	1	17
4	2	4	4	1	2	39
4	3	5	4	1	2	30
4	4	5	4	2	2	18
4	2	3	4	2	2	37
5	2	5	5	2	2	16
1	2	2	4	2	1	68
5	3	4	5	1	2	18
4	3	3	4	1	1	36
4	1	3	2	1	1	64
4	2	3	4	2	1	42
5	3	4	4	2	1	25
6	3	4	5	1	1	10
6	4	5	4	2	2	2

付表2 トンネル切羽評価用未知データと提案手法による推定値データ

圧縮強度	風化変質	割目間隔	割目状態	湧水量	劣化	評価点	推定値
6	4	5	5	1	3	-5	0
5	4	4	4	1	1	18	24
5	2	5	5	1	2	21	19
4	2	5	4	2	2	31	32
4	4	4	3	1	1	33	45
2	1	2	3	2	1	74	75
3	2	3	4	1	1	50	43
5	3	5	5	2	2	10	15
5	3	5	4	1	2	22	30
5	4	5	5	1	2	8	8
5	2	4	5	1	1	24	0
4	2	4	3	2	1	46	52
4	2	4	5	2	1	32	4
3	2	2	2	2	1	68	50
4	2	5	4	1	1	36	32
5	4	4	5	2	2	6	3
6	4	5	4	1	2	7	30
5	2	4	4	1	1	31	31
4	3	5	4	2	2	25	30
4	2	3	4	1	1	42	36
3	2	3	5	1	1	43	47
6	4	5	4	1	3	2	30
4	1	4	5	1	1	39	24
4	3	3	4	2	2	31	36
3	2	2	3	2	2	55	52
4	3	4	3	1	1	40	46
4	2	4	5	1	1	32	31
4	2	3	3	1	2	49	37
5	2	5	4	1	1	28	0
5	3	5	3	2	1	29	13
6	4	4	5	1	1	3	1
4	3	4	4	2	2	28	35
6	3	5	4	1	2	14	30
5	4	5	3	1	1	22	24
3	1	2	1	1	1	82	85
5	4	3	4	2	3	14	21
4	3	4	3	2	2	35	45
4	4	5	4	1	1	23	29
6	4	4	5	1	3	-2	3
5	4	5	4	1	2	15	24
4	2	4	3	1	2	46	40
5	4	4	4	2	2	13	31
2	2	3	4	1	1	57	44
5	2	3	4	2	1	34	58
3	2	3	1	1	1	72	44
6	3	4	5	2	2	5	5
6	3	5	5	2	4	-3	-5
4	2	3	5	2	2	30	59
5	2	4	4	2	1	31	6
5	3	4	3	1	1	32	22