

修士論文

スパイキングニューラルネットワーク
におけるスパイク時刻と数の学習法
ー 不要な要素の除去による学習性能の向上 ー

平成27年度修了

三重大学大学院工学研究科

博士前期課程 電気電子工学専攻

松本 崇

目次

第1章	はじめに	1
1.1	はじめに	1
第2章	SpikeProp	4
2.0.1	SpikeProp のネットワーク構造と学習法	4
2.0.2	SpikeProp の問題点	6
第3章	荷重減衰による余分なスパイクの抑制	10
3.1	提案法—調整 WD 法—	10
3.2	実験	13
3.2.1	時間版 XOR の学習	13
3.2.2	Iris データセットの学習	17
第4章	初期ネットワーク構造の余分な要素除去	22
4.1	提案法—不要中間層ユニット除去—	22
4.1.1	中間層ユニット数調整の方針	22
4.1.2	ユニット除去のタイミング	23
4.1.3	除去対象のユニットの決定法	23
4.1.4	実験	24
4.1.5	検討	28
4.2	提案法—不要時間遅れ除去—	28
4.2.1	基本アイデア	28

4.2.2	判定法	29
4.2.3	不要時間遅れ除去法 of 具体案	29
4.2.4	実験	30
4.3	提案法—不要な副結合除去—	36
4.3.1	基本アイデア	36
4.3.2	判定法	36
4.3.3	抑制法	37
4.3.4	実験	38
第 5 章 まとめ		43
謝辞		45
参考文献		46
発表		48

表 目 次

3.1	時間版 XOR 問題	14
3.2	学習に用いたネットワークの構成 (XOR 問題)	14
3.3	調整 WD 法の各種パラメータ (XOR 問題)	15
3.4	学習に用いたネットワークの構成 (Iris 問題)	18
3.5	調整 WD 法の各種パラメータ (Iris 問題)	18
3.6	初期中間層ユニット数, 及び初期副結合数を変化させた結果 (学習 成功率) (XOR 問題)	19
3.7	初期中間層ユニット数, 及び初期副結合数を変化させた結果 (平均 学習回数) (XOR 問題)	19
3.8	時間遅れを変化させた結果 (XOR 問題)	20
3.9	時間遅れを変化させた結果 (Iris 問題)	20
3.10	初期副結合数を変化させた結果 (Iris 問題)	20
4.1	学習に用いたネットワークの構成	25
4.2	中間層ユニット除去法による学習成功率	26
4.3	中間層ユニット除去法による平均学習回数	27
4.4	中間層ユニット除去法による最終中間層ユニット数	27
4.5	学習に用いたネットワークの構成 (XOR 問題)	33
4.6	不要時間遅れ除去による学習成功率 (XOR 問題)	33
4.7	不要時間遅れ除去による平均学習回数 (XOR 問題)	34
4.8	学習に用いたネットワークの構成 (Iris 問題)	34

4.9 不要時間遅れ除去による学習成功率 (Iris 問題)	35
4.10 不要時間遅れ除去による平均学習回数 (Iris 問題)	35
4.11 学習に用いたネットワークの構成 (XOR 問題)	39
4.12 不要な副結合除去による学習成功率 (XOR 問題)	39
4.13 不要な副結合除去による平均学習回数 (XOR 問題)	40
4.14 学習に用いたネットワークの構成 (Iris 問題)	40
4.15 不要な副結合除去による学習成功率 (Iris 問題)	41
4.16 不要な副結合除去による平均学習回数 (Iris 問題)	41

目 次

2.1	SpikeProp のネットワーク構造	5
2.2	スパイク時刻の調整に成功した割合	7
2.3	スパイクの時刻とその数の調整に成功した割合	8
2.4	出力ユニットの発火数	9
3.1	学習率 η の切り替え	11
3.2	係数 ρ の選択	12
3.3	学習成功率 (XOR 問題)	15
3.4	平均学習回数 (XOR 問題)	16
4.1	ユニット除去のタイミング	24
4.2	除去対象のユニットの決定法	25
4.3	不要時間遅れの除去 (入力層-中間層間)	30
4.4	不要時間遅れの除去 (中間層-出力層間)	30
4.5	時間遅れ除去の流れ	31
4.6	時間遅れ除去の流れ (詳細)	32
4.7	結合毎の要・不要の判断	37
4.8	提案法を加えたフローチャート	38

第1章 はじめに

1.1 はじめに

今日，計算機が担う役割は大きなものとなっており，多くの情報を処理する技術の発展が大きな課題となってきた。その中でも近年，情報処理の分野で機械学習が注目を集めている。これは，多数のデータを計算機に与え学習させることで，複雑な動作を実現できるようにするものである。例えば，機械学習をコンピュータ将棋に適用することで，プロ棋士に匹敵する計算機を作り出すことができた [1]。機械学習と一口に言っても k-mean 法，クラスタリング，決定木やニューラルネットワーク (NN) 等，さまざま存在する。この中でも NN は人の脳をモデル化したものであり，計算機が不得意とする曖昧な判断も可能と考えられている。

ニューラルネットワークは 1958 年，Rosenblatt によって提案された [2]。当初のニューラルネットワークは入力層と出力層の二層から成るモデル（単純パーセプトロン）であり，簡単なパターン認識であれば解くことができ，注目を集めた。しかし，Rosenblatt によって提案されたモデルでは非線形分離問題（例えば，XOR 問題）を解くことができず，限界があることがわかった。後の 1986 年，Rumelhart らが多層パーセプトロンを提案した [3]。これは Rosenblatt によって提案された二層から成るものではなく入力層，中間層（隠れ層），出力層の多層から成るモデルである。このモデルは誤差逆伝播法（BP 法）で学習を行うことで非線形分離問題を解くことができるようになった。後にニューラルネットワークの研究は広く行われ，現在までに様々なモデルが提案され，応用されてきた [4-6]。

中でもユニット間のやりとりをスパイク（パルス）で行うスパイクニューラ

ルネットワーク (SNN) が注目を集めている [7, 8]. 生体のニューロンは, スパイク状の電気信号のやりとりを通じて信号処理をしていることが知られている. SNN はこれを工学的にモデル化したものであり, ニューラルネットワークの中では生体に近いモデルであるといえる. そして, SNN はスパイク列により情報を表現するため, 時系列信号の取り扱いが可能となったモデルである. また, 信号伝達がスパイクにより行われるため, ノイズに強くハードウェア実装時に有利であるといった利点もある.

SNN は, その入出力における情報表現手法により, 大きく二種に分けることができる. 一つはスパイクの密度を用い情報を表現するモデル, もう一つはスパイクのタイミングによるより情報を表現モデルである. 前者はスパイクの密度を観測するための時間がかかってしまうため, 応答の速さの観点から後者のほうが好ましいとされている [9].

スパイクのタイミングにより情報表現を行う SNN の学習法としてさまざまなものが提案されている [10–14]. その一つに Booiij らによって提案された SpikeProp がある [13]. これは, Bohte らにより提案された SpikeProp[14] を拡張したモデルである.

Bohte らのモデルでは, 出力ユニットの発火時刻を学習できるものの, 各ユニットは1つのスパイクしか出力できなかった. Booiij らは各ユニットが複数スパイクを出力できるよう, また複数の出力スパイクの時刻を学習できるように拡張した. これにより, スパイク列の学習が可能となり, Booiij らの SpikeProp はさまざまな時系列情報処理システムへの応用が期待されている. しかし, Booiij らの学習法では出力スパイクの数を陽に調整できず, 余分な出力スパイクをも出力してしまう. これは, SpikeProp の機能をスパイク列からスパイク列への変換であるとみなしたときに, 好ましくない動作である.

そこで, 本研究では, Booiij らのモデルにおいて, 出力スパイクの発火時刻だけでなく, それらの数も同時に学習できるようにすることを目標とする. 本稿では

Booiij らの SpikeProp ネットワークにおいて、不要な結合の結合荷重抑制法を提案・適用し、この目標を実現する。また、この方法のみでは任意のパラメータである初期のネットワーク構造（中間層ユニット数，時間遅れ，副結合数）に学習性能が依存してしまう。そのため，本稿では依存性を解消し，学習性能の向上をするため，不要なネットワーク要素を学習毎に除去する方法も併せて提案・適用する。

以下に本論文の構成を示す。

2 章では SpikeProp とその学習法について述べる。3 章では不要な結合の結合荷重を抑制することでスパイクの数，またその時刻を学習させる方法の提案・検討を行う。その検討結果を踏まえ，4 章では SpikeProp の初期ネットワーク構造の余分な要素の除去をすることによるネットワークの自動構築手法を提案する。5 章で本研究についてまとめる。

第2章 SpikeProp

この章では Booiij らによって提案された SpikeProp モデルの概略と問題点について説明する。

2.0.1 SpikeProp のネットワーク構造と学習法

SpikeProp は入出力関係にスパイク（パルス）を用いるフィードフォワード・ニューラルネットワークのモデルである。例えば、図 2.1 に示すような階層型のネットワーク構造を持つ。また、ユニット間の結合は 1 本ではなく複数の副結合からなっており、それぞれが結合荷重と時間遅れという二種のパラメータを持つ。時間遅れの値は、副結合番号（1 から m ）により決まる定数として、 d^1, d^2, \dots, d^m のように表す。本稿では、入力層と中間層の間、中間層と出力層の間に関わらず、層間の結合は全結合とし、各結合は同様の副結合を持つものとした。また、ユニット i からユニット j への第 k 副結合の結合荷重は w_{ij}^k のように表す。

各ユニットは、その内部状態があらかじめ決められた発火しきい値 θ を超えた時、発火しスパイクを出力する。ここでユニット i の内部状態 x_i は、式 (1) に従い変化する。式 (2.2) はスパイクに対する応答を示すスパイク応答関数 ϵ 、式 (2.3) は

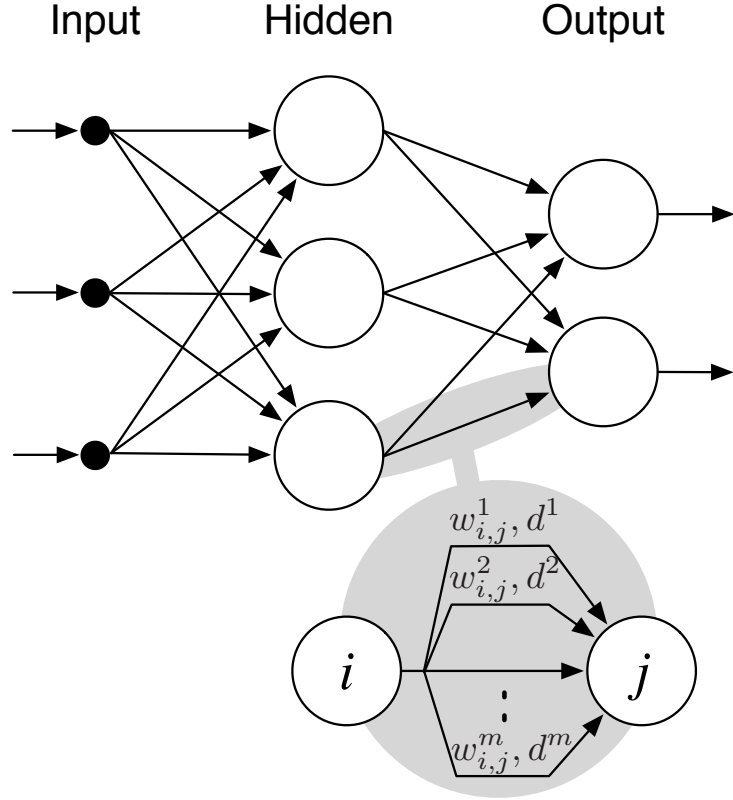


図 2.1: SpikeProp のネットワーク構造

自身の発火による不応期を表す不応期関数 κ である.

$$\begin{aligned}
 x(t) = & \sum_{t_j^f \in F_j} \kappa(t - t_j^f) \\
 & + \sum_{i \in \Gamma_j} \sum_{t_i^f \in F_i} \sum_{k=1}^m w_{ij}^k \epsilon(t - t_i^f - d^k)
 \end{aligned} \tag{2.1}$$

$$\epsilon(t) = \begin{cases} 0 & (t < 0) \\ \exp\left(-\frac{t}{\tau_m}\right) - \exp\left(-\frac{t}{\tau_s}\right) & (t \geq 0) \end{cases} \tag{2.2}$$

$$\kappa(t) = \begin{cases} 0 & (t < 0) \\ -\theta \exp\left(-\frac{t}{\tau_r}\right) & (t \geq 0) \end{cases} \tag{2.3}$$

ここでここで Γ_j はユニット j に接続するユニットの番号の集合, F_i はユニット i の発火時刻の集合であり, m は各ユニット間の副結合数, τ_m, τ_s, τ_r はスパイク応

答関数・不応期関数の時定数, t_i^f はユニット i の f 番目のスパイクの時刻である.

SpikeProp では, 各入出力をスパイクの時刻により表現している. 出力スパイクの時刻を調節するために, 式 (2.4) により定められた誤差 E を減らすように, 誤差逆伝搬法に基づいた方法により, 全ての副結合の結合荷重を調整することで学習する. これにより, 各スパイクの時刻を詳細に調整できる.

$$E = \sum_{p \in P} E_p, \quad E_p = \frac{1}{2} \sum_{j \in J} \sum_l \left(t_j^{p,l} - \hat{t}_j^{p,l} \right)^2 \quad (2.4)$$

ここで, J は出力層ユニットの集合を, P は教師パターンの集合を表す. また, $t_j^{p,l}$, $\hat{t}_j^{p,l}$ はパターン p におけるユニット j の l 番目のスパイクの出力時刻および希望する出力時刻を表す. 学習は, 式 (2.5) にしたがって, 結合荷重を繰り返し更新することで行う.

$$\Delta w_{ij}^k = -\eta \frac{\partial E}{\partial w_{ij}^k} \quad (2.5)$$

ここで, η は学習率であり正の実数である.

なお, 学習を繰り返すと, ユニットの発火回数が不足することがある. 学習中に一つでもそのようなユニットが生じた場合には, 誤差の逆伝搬量を求めることができなくなる. そのため, 再びそのユニットが発火するように, そのユニットへのすべての結合の結合荷重を大きくし, 学習を継続する. 本稿では, 対象となる結合荷重に一定値を加算することを, 当該ユニットが発火するまで繰り返す.

2.0.2 SpikeProp の問題点

SpikeProp を, 時系列情報処理に用いる際の問題点は, 出力のスパイク数を意図的に調整できないことである. 出力スパイクが不足する場合は, 前節の最後に述べたように結合荷重を補正するため問題ない. しかし, 出力スパイクが過剰な場合に問題が生じる. これは, ネットワークが出力する最初の何個かのスパイクについての発火時刻のずれだけを式 (2.4) で評価していることに起因する. 過剰に (期

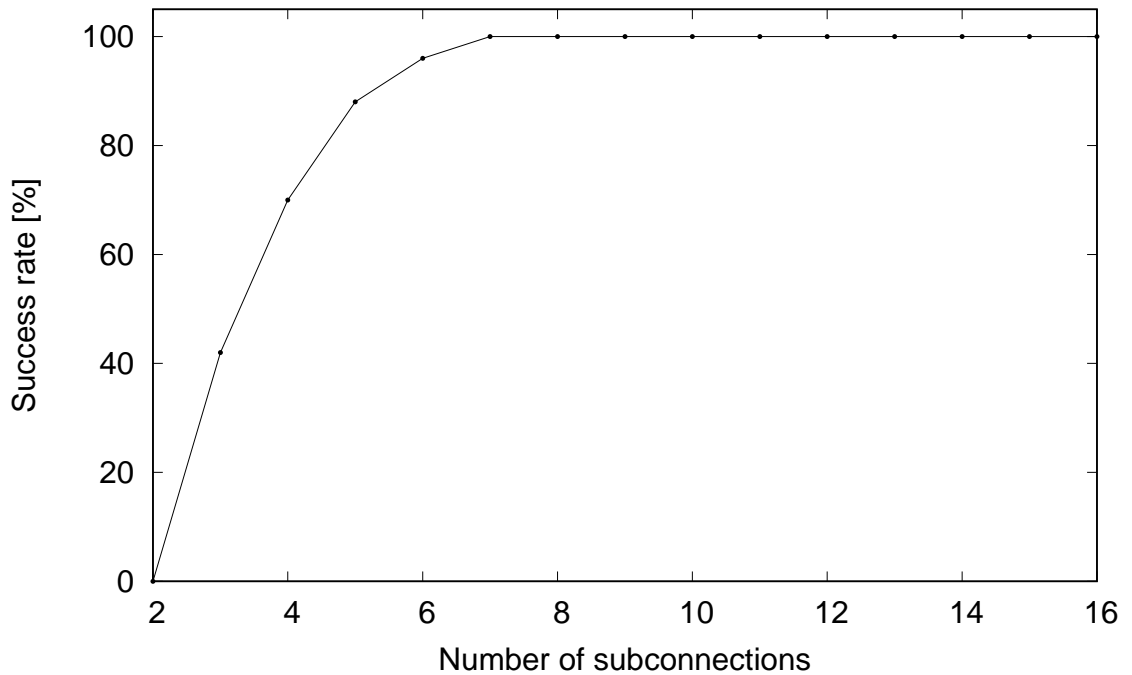


図 2.2: スパイク時刻の調整に成功した割合

待した以上に) 出力されたスパイクは、誤差に何ら影響を与えない。そのため、誤差を減少させるだけでは、これらのスパイクは学習により抑制されない。これは、期待したスパイク列だけを出力するネットワークを確実に得ることはできないことを意味する。SpikeProp の機能をスパイク系列からスパイク系列への変換であると考えたとき、これは好ましい動作ではない。

実際に、Booij らの論文で使用されている時間版 XOR 問題を、副結合数 m を 2 から 16 本まで変化させ Booij らの手法を用いて学習させた (詳細な学習条件は 3.2.1 節を参照)。このときの学習成功率と副結合数 m の関係を図 2.2 に示す。ここで学習成功とは、出力スパイク数に関わらず、スパイク時刻の調整に成功した (式 (2.4) による誤差が所定の学習回数以内に十分に減少した) ことを意味する。副結合が 3 本を超えると学習に成功し、7 本を超えると学習に 100% 成功した。時間版 XOR 問題では、全てのパターンで出力スパイク一つあたり一つの教師出力しか期待していない。そこで、スパイクの時刻だけでなくスパイクの数も合っていると

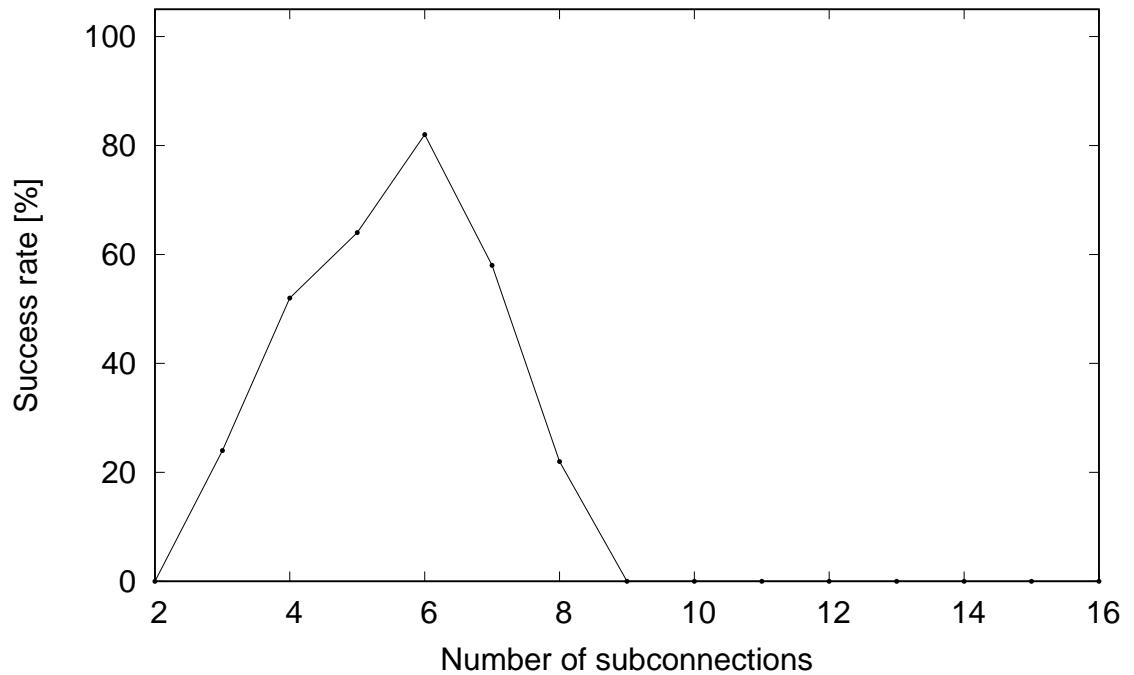


図 2.3: スパイクの時刻とその数の調整に成功した割合

きに学習成功と見なすと、学習成功率は図 2.3 のようになった。先ほどと異なり、副結合が 3 本から 8 本までの場合にのみ学習に成功した。また、その成功率は最大で 82% であった。図 2.4 に、副結合が 9 本以上の場合について、出力層ユニットが出力したスパイク数を示す。副結合が増加するのに伴い、出力スパイクが増加している。出力スパイクとしては 1 個だけを期待しているので、いずれの場合でも複数の余分なスパイクが出力されている。

以上より、Booij らの学習法ではスパイク数を調整することはできないが、副結合数を調整することで多少は解決できることが分かる。各ユニットは、内部電気が発火しきい値を超えたときに、発火しスパイクを出力する。式 (2.1) から、副結合が少ないときには内部電位が大きくなりにくい。すなわち、副結合数を少なくすることで、各ユニットへの入力（刺激）が弱くなり、そのユニットが発火しにくくなる。

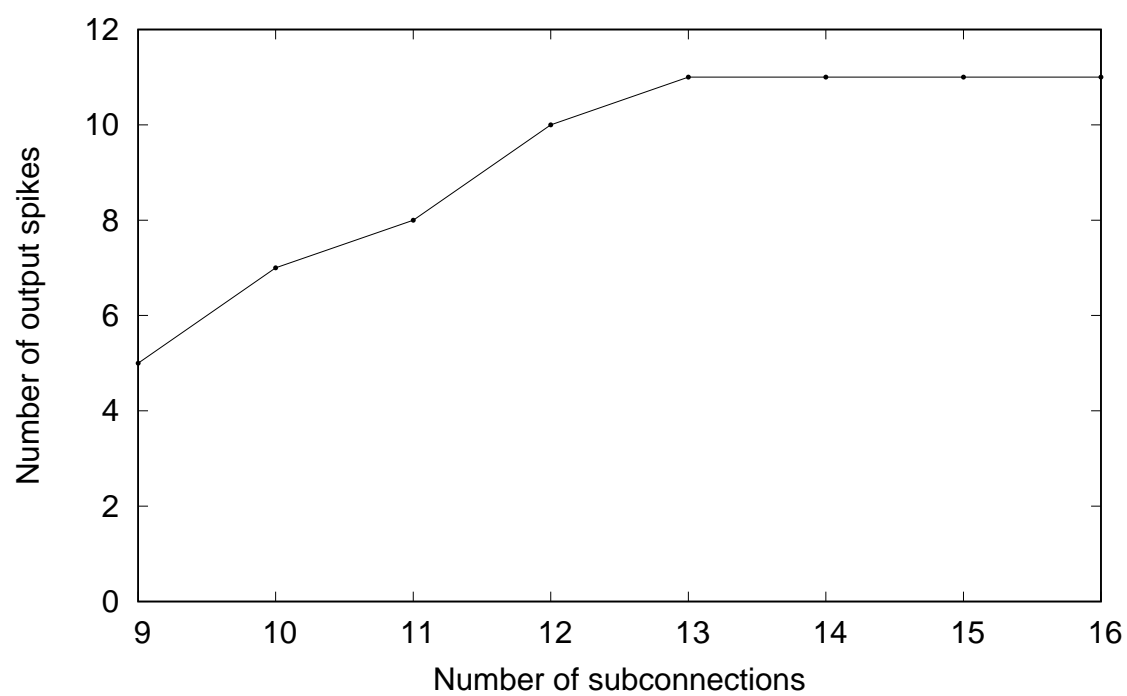


図 2.4: 出力ユニットの発火数

第3章 荷重減衰による余分なスパイクの抑制

前節の議論より，学習中に余分な副結合を除くことで，スパイク時刻とスパイク数をともに学習できると考えた．この考えに基づいた学習法をこの章で提案する．

3.1 提案法—調整 WD 法—

スパイキングニューラルネットワークに限らなければ，学習中に余分な結合を除く学習法が多数提案されてきた [15, 16]．本稿ではこの中から，BP モデル向けに提案された荷重減衰（WD: Weight Decay）法 [16] に着目する．この手法は，各結合の要・不要を陽に判定することなく，余分な結合を除くことができる．そのため，異なるモデルである SpikeProp にも容易に適用できると考えた．

荷重減衰では，学習評価関数を誤差そのものではなく結合荷重に関する L2-正則化項を付加したものに変更することで，学習により不要な結合を減衰させる．SpikeProp に適用した場合，式 (3.1) に示した E' を学習評価関数として，これを学習により最小化する．第一項は SpikeProp の誤差 E ，第二項は荷重減衰による正則化項である．第二項により誤差を減少させることに寄与しない結合荷重は徐々に減衰する．

$$E' = E + \frac{\rho}{\eta} \sum_{w \in W} w^2 \quad (3.1)$$

ここで， W はネットワーク内の全結合荷重の集合， ρ は荷重減衰の影響の強さを調整する定数である．この評価関数を最小化するための各結合荷重の更新式は式

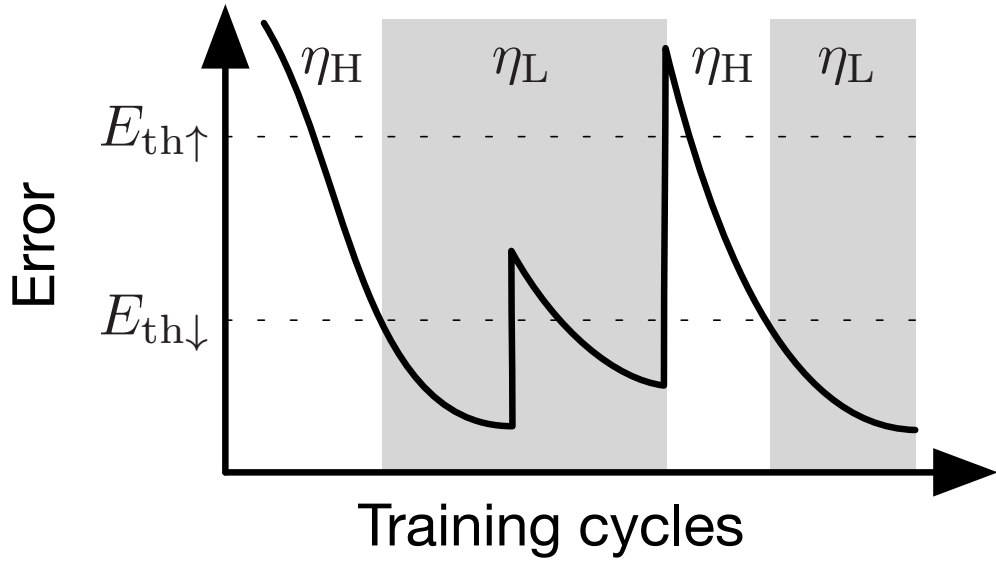


図 3.1: 学習率 η の切り替え

(3.2) のようになる．

$$\Delta w_{ij}^k = -\eta \frac{\partial E}{\partial w_{ij}^k} - 2\rho w_{ij}^k \quad (3.2)$$

過度な荷重減衰の適用は，学習性能の低下を引き起こす．SpikeProp においては，過度に結合荷重が減衰すると，内部電位が低下しすぎてユニットが発火しにくくなる．ユニットが発火なくなると，そのユニットに繋がる結合の結合荷重が補正される．この補正は式 (3.2) とは関係なく行われる．そのため，過度に荷重減衰を適用することは，学習過程に悪影響をおよぼすだろう．

そこで，学習の進行に応じて，学習率 η と係数 ρ を調整する．具体的には， η は誤差が十分に減少したら小さくし， ρ は余分なスパイクが出力されなくなったら小さくする．このようにすることで，学習終盤で発火時刻・スパイク数の一方のみが調整しきれていない場合は，それに集中して調整するようになり，学習が効果的に進むと考えられる．なお簡単のため，いずれのパラメータも値を 2 段階で切り替える．

学習率 η は， η_H, η_L ($\eta_H > \eta_L$) を以下の基準で切り替える．学習開始時から誤差

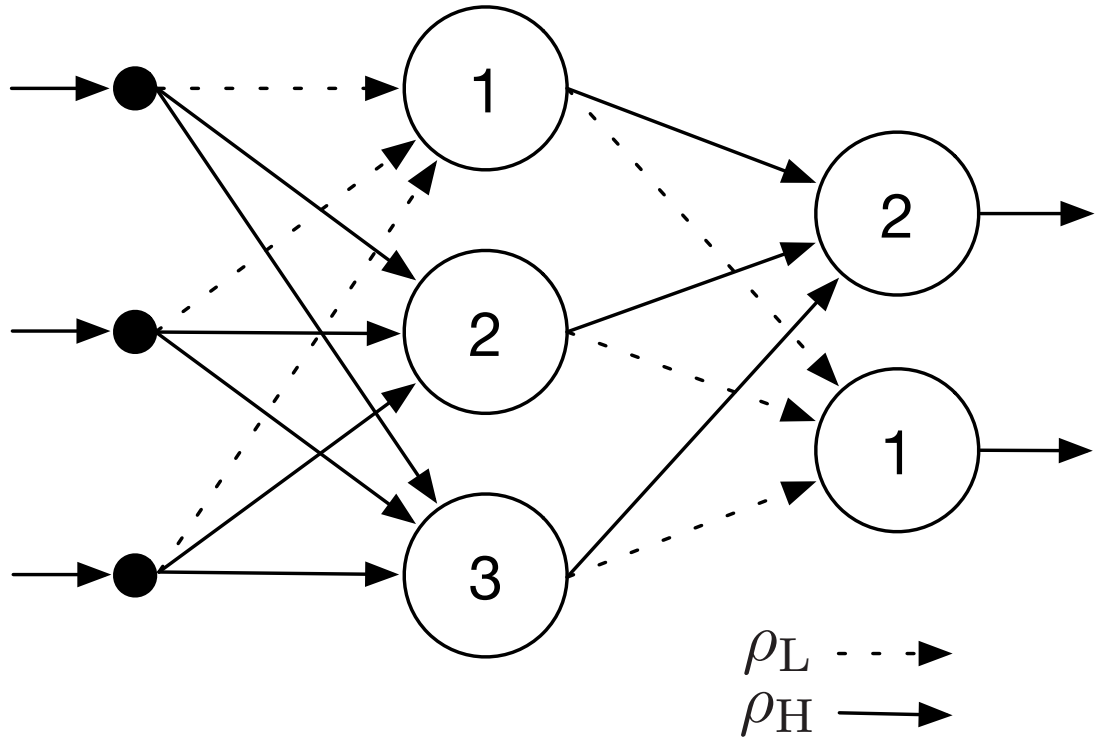


図 3.2: 係数 ρ の選択

がある値 ($E_{th\downarrow}$) まで減少するまでは, η_H を用いることで, 誤差を減少させる効果を強くする. 誤差が $E_{th\downarrow}$ を下回った後は, η_L を用い余分なスパイクを抑える学習に切り替える. その後, 誤差が上昇し誤差が $E_{th\uparrow}$ ($E_{th\uparrow} > E_{th\downarrow}$) を上回った後は, 学習率として η_H を再び用い, 以下同様に学習率を調整する. 図 3.1 は, 学習の進行に伴う η の切り替えの様子を表している. このように, ヒステリシスな特性を持たせることで η が頻繁に切り替わることを防いでいる.

係数 ρ は, ρ_H, ρ_L ($\rho_H > \rho_L$) を以下の基準でユニット毎に切り替える. 余分な発火をしているユニットへの副結合については ρ_H を用いることで, 発火を抑制する. そうでないユニットについては ρ_L を用いることで, 学習を阻害しないようにする. なお, 出力層ユニットについては, 余分な発火をしているかどうかを教師パターンより判定できる. 中間層ユニットについては, 必要な発火数は不明であるため, あらかじめに定めたしきい値 (θ_{num}) と比較することで判定する. なおこ

のしきい値は，出力層ユニットが十分に発火できるように設定しなければならない．図 3.2 に， ρ の選択例を示す．各ユニット内部の数字は，その出力スパイク数を表すとする．教師出力で期待しているスパイク数および θ_{num} を 1 とすると，発火回数が 2 回以上のユニットは余分なスパイクを出力している．図では 3 個のユニットが該当しているので，それらのユニットへの結合には ρ_H を適用し，他の結合には ρ_L を適用する．

3.2 実験

この節では，簡単な実験をとおり，調整 WD 法（提案法）により実際に余分なスパイクの抑制ができるのかを検証する．

3.2.1 時間版 XOR の学習

この節では，Bohte ら，Booij らの論文で学習対象として用いられていた時間版 XOR 問題の学習をつうじて，提案法の学習性能（学習成功率と平均学習回数）を確認する．比較は，従来法（Booij らの学習法），WD 法（パラメータ (η, ρ) 固定の荷重減衰法），および調整 WD 法（提案法）の 3 手法の間で行う．

実験条件

学習対象の時間版 XOR 問題は，Booij らの論文でベンチマーク問題として取り上げられていた教師パターンセットである．これは，論理値をスパイクの発火時刻によりエンコード（早い時刻が論理値 1，遅い時刻が論理値 0）し，排他的論理和を表したものである（表 3.1）．本来これは 2 入力 1 出力の問題だが，文献 [14] での実験にならい，バイアス入力として時刻 0 に発火するスパイクをパターンにかかわらず第 3 の入力 3 として与えた．

表 3.1: 時間版 XOR 問題

Input			Output
1	2	3	
0	0	0	16
0	6	0	10
6	0	0	10
6	6	0	16

[ms]

表 3.2: 学習に用いたネットワークの構成 (XOR 問題)

Number of input units	3
Number of hidden units	5
Number of output units	1
Number of sub-connections m	$2, 3, \dots, 16$
Delay d^k ($k = 1, 2, \dots, m$)	k
Threshold for activity θ	7
Time constant τ_m, τ_s, τ_r	10.0, 5.0, 10.0

実験に用いたネットワークの各パラメータは、Booij らの実験に準じて定めた (表 3.2)。中間層ユニットについても Booij らの実験に準じ、5 個のうち 4 個は興奮性ユニット、1 個は抑制性ユニットとした。初期荷重は、中間層ユニットへの結合については $[-1, 5]$ 、興奮性の中間層ユニットからの結合は $[0, 5]$ 、抑制性の中間層ユニットからの結合は $[-1, 0]$ の一様乱数とした。従来法では学習率 η を 0.2 とした。WD 法では、学習率 η は 0.2、係数 ρ は 0.002 とした。これらパラメータは学習成功率が高くなるように、試行錯誤により決定した。

また、調整 WD 法のパラメータを表 3.3 に示す。これらパラメータについても学習成功率が高くなるように、試行錯誤により決定した。学習終了条件は、誤差

表 3.3: 調整 WD 法の各種パラメータ (XOR 問題)

ρ_H	0.002	ρ_L	0.0005
η_H	0.2	η_L	0.05
$E_{th\uparrow}$	1.5	$E_{th\downarrow}$	0.7
θ_{num}	1		

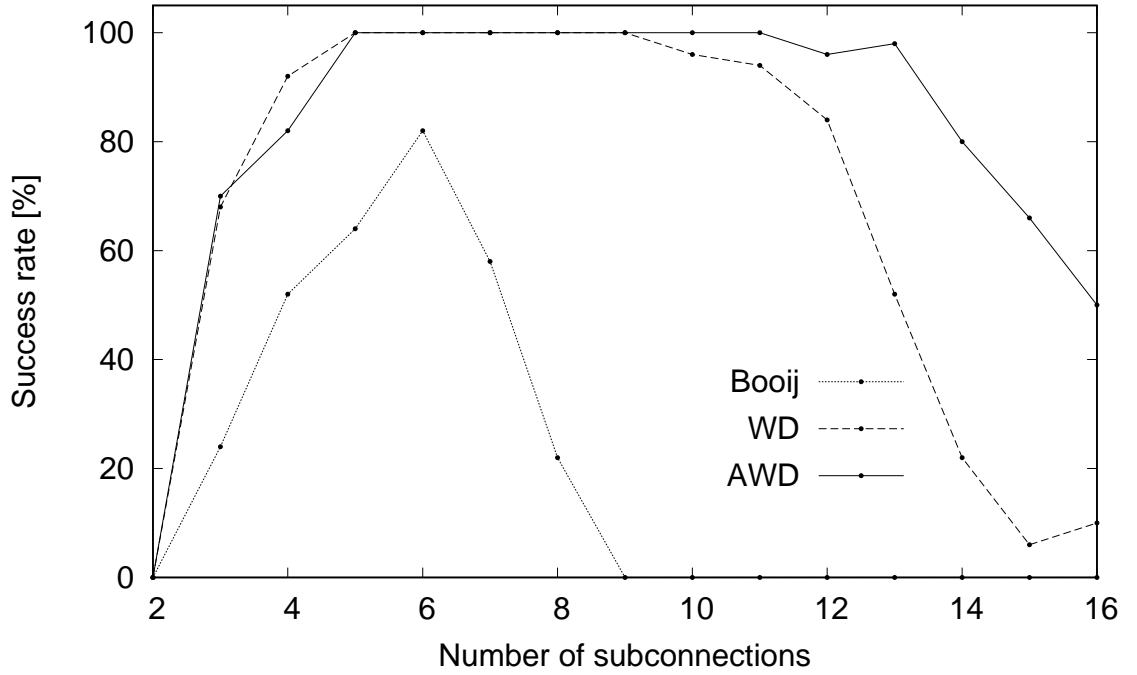


図 3.3: 学習成功率 (XOR 問題)

が十分に小さくなり ($E < 1$) かつ出力層ユニットが余分なスパイクを出力していないこととした。なお、結合荷重が更新された回数を学習回数とし、学習回数が 1 万回を超えた場合は学習失敗とみなした。実験は、異なる 50 種の初期荷重から行った。

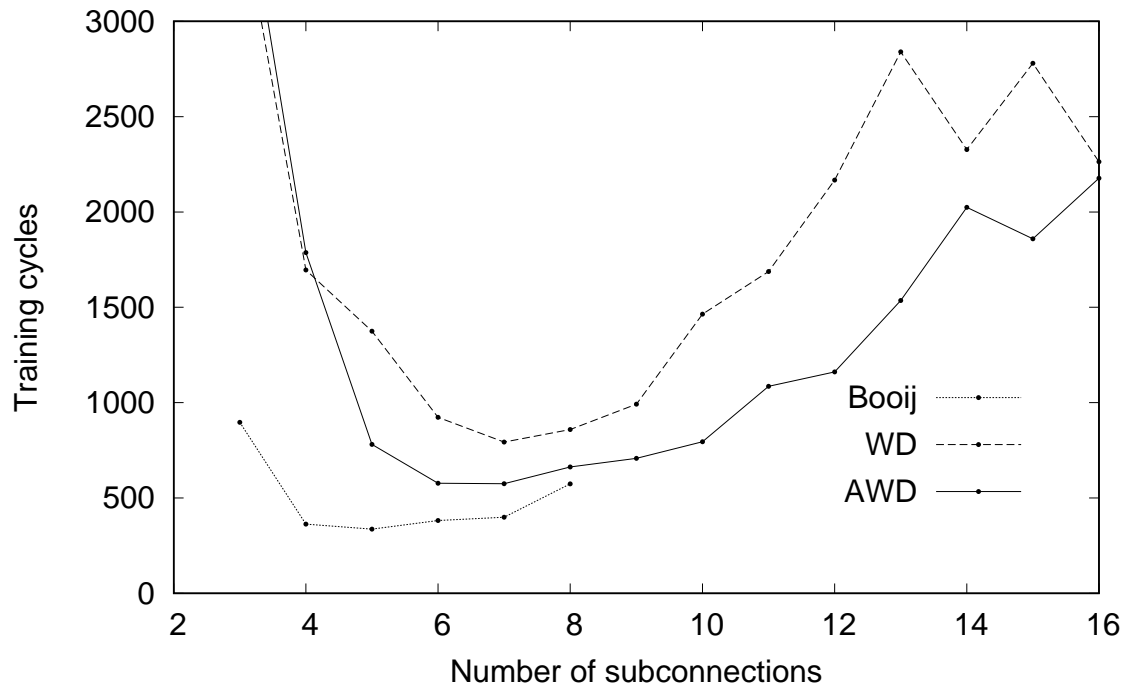


図 3.4: 平均学習回数 (XOR 問題)

結果

図 3.3 に、副結合数 m と学習成功率の関係を示す。実線が調整 WD 法 (AWD と表記)、破線が WD 法 (WD と表記)、点線が従来法 (Booij と表記) による結果である。従来法では学習成功率は最大で 82% だったが、荷重減衰を適用することで副結合数が 5 から 9 の場合に学習が 100% 成功するようになった。また、調整 WD 法と WD 法を比べると、前者の方が副結合数が多いときに学習成功率がより向上した。

図 3.4 に、副結合数 m と学習成功時の平均学習回数の関係を示す。実線が調整 WD 法、破線が WD 法、点線が従来法による結果である。いずれの結果も、学習回数は、副結合が 3 本の時が最も多く、副結合数の増加に伴いいったん減少し、その後、増加する傾向を示した。荷重減衰を適用することで従来法と比べ学習回数は増加したが、増加量は WD 法に比べ調整 WD 法の方が少なかった。

以上より、荷重減衰を適用することで、スパイクの時刻だけでなくスパイク数

も確実に学習できるようになったといえる。その反面、学習回数は従来法と比べ増加した。しかし、適切な副結合で学習したときは従来法での最悪の平均学習回数(約900回)を下回っており、学習に100%成功するようになったことをふまえると、荷重減衰を適用することは有効であったと考える。また、調整WD法により η と ρ を適応的に切り替えることで、単純に荷重減衰を適用した場合と比べ学習性能が向上した。

3.2.2 Iris データセットの学習

この節では、Iris データセット [17] の学習に対しても、調整WD法(提案法)が効果があることを、従来法との比較をつうじ確認する。

実験条件

学習対象の Iris データセットは、パターン認識の分野で古くからベンチマーク問題として用いられてきたデータセットである。Iris データセットは、4つの特徴量から品種(3種類)を特定するものである。各品種毎に、50件のデータがデータセットに含まれている。ここでは、4つの特徴量をそれぞれ時刻0msから6msの区間に正規化して入力した。なおXOR問題の時と同じく、バイアス信号も入力として与えた。出力には各品種に対応したユニットを用意し、入力されたパターンに該当する品種に対応するユニットに対しては10ms、そうでないユニットに対しては15msのスパイクを教師信号として与えた。各品種からランダムに5件ずつ選んだデータを基にした計15組のパターンを、教師パターンとして用いた。

実験に用いたネットワークの各パラメータを表3.4に示す。中間層ユニットは、興奮性・抑制性は区別はつけない。初期荷重は、全ての結合について $[-1,5]$ の1様乱数とした。従来法では学習率 η を0.2とした。

表 3.4: 学習に用いたネットワークの構成 (Iris 問題)

Number of input units	5
Number of hidden units	10
Number of output units	3
Number of sub-connections m	8
Delay d^k ($k = 1, 2, \dots, m$)	k
Threshold for activity θ	7
Time constant τ_m, τ_s, τ_r	10.0, 5.0, 10.0

表 3.5: 調整 WD 法の各種パラメータ (Iris 問題)

ρ_H	0.002	ρ_L	0
η_H	0.05	η_L	0.0125
$E_{th\uparrow}$	45	$E_{th\downarrow}$	21
θ_{num}	1		

また，調整 WD 法のパラメータを表 3.5 に示す．これらのパラメータについても学習成功率が高くなるように，試行錯誤により決定した．

学習終了条件は，誤差が十分に小さくなり（すべての教師パターン p について $E_p < 2$ ）かつ出力層ユニットが余分なスパイクを出力していないこととした．なお，結合荷重が更新された回数を学習回数とし，学習回数が 1 万回を超えた場合は学習失敗とみなした．実験は，異なる 20 種の初期荷重から行った．

結果

従来法では，全ての試行において学習に失敗した．いずれの場合も，誤差は十分に減少したが，発火回数が十分に減少しなかった．すべての試行で，誤差は学習終了条件を満たすことはできて，その場合の平均学習回数は 1,973 回であった．また 1 万回学習した時点で，各出力ユニットは約 10 個の余分なスパイクを出力し

表 3.6: 初期中間層ユニット数, 及び初期副結合数を変化させた結果 (学習成功率)
(XOR 問題)

Number of hidden units	Number of sub-connections		
	8	12	16
5	100	96	50
10	100	100	64
15	100	98	56

表 3.7: 初期中間層ユニット数, 及び初期副結合数を変化させた結果 (平均学習回数)
(XOR 問題)

Number of hidden units	Number of sub-connections		
	8	12	16
5	663	1,161	2,177
10	615	890	2,186
15	785	1,343	3,065

ていた. なお, 他のネットワーク構造・パラメータの値を用いた場合でも, 学習に成功することはなかった.

調整 WD 法では, 全ての試行において, スパイク時刻・スパイク数の両方の学習終了条件を満たすことができた. この場合の平均学習回数は 1,168 回であった.

調整 WD 法を用いることで, 時間版 XOR 問題の学習の時と同様にスパイク時刻だけでなくスパイク数の調整に成功しただけでなく, 従来法と比べ学習回数が大きく減少した.

表 3.8: 時間遅れを変化させた結果 (XOR 問題)

	Delay (min - Max)		
	1-8	1-12	1-16
Success rate[%]	96	94	50
Average of learning cycles	1,442	1,080	2,177

表 3.9: 時間遅れを変化させた結果 (Iris 問題)

	Delay(min-Max)		
	1-8	1-12	1-16
Success rate[%]	85	80	45
Average of learning cycles	1,980	3,967	3,110

検討

調整 WD 法を用いることで SpikeProp において出力のスパイクの数, またその時刻を学習することができるようになった. しかし, 例えば XOR 問題であれば図 3.3 の副結合数が 14 から 16 本の場合のように, 副結合数が多くなるにつれて学習成功率の低下が見られる.

後の追実験より, 調整 WD 法では初期のネットワーク構造 (初期中間層ユニッ

表 3.10: 初期副結合数を変化させた結果 (Iris 問題)

	Number of Sub-connections		
	8	12	16
Success rate[%]	100	85	45
Average of learning cycle	1,204	3,200	3,110

ト数，時間遅れ，初期副結合数）にその学習性能が依存する傾向にあることがわかった．その様子を表 3.6, 3.7, 3.8 に示す．この結果は時間版 XOR 問題を 50 試行学習させた結果である．表 3.6, 3.7 については副結合数を 8, 12, 16 と変化させた場合（時間遅れ幅固定）中間層ユニット数を 5, 10, 15 と増やした場合，表 3.8 は中間層ユニット数 5 個，副結合 16 本に固定し，時間遅れの最大値を変化させた結果である．表 3.6, 3.7 では，中間層ユニットが多くなるにつれ，平均学習回数が増加傾向にあり，副結合数が多くなるにつれて学習成功率が悪く，平均学習回数が増加傾向にあることがわかる．また，表 3.8 では，最大時間遅れが大きくなるにつれ学習成功率が悪く，平均学習回数が増加傾向にあることがわかる．また，Iris 問題も同様に副結合数を多くした場合（時間遅れ幅固定）（表 3.9），最大時間遅れを大きくした場合（副結合数 16 本固定）（表 3.10）ではその学習性能が悪くなる傾向にある．

これらより，調整 WD 法を用いた SpikeProp はその学習性能（学習成功率及び平均学習回数）が初期のネットワーク構造に依存してしまう．

本稿では，この依存性を解消し，学習性能の向上を目指すため 4 章でそれぞれの不要なネットワーク要素の除去法について検討する．4.1 節では中間層ユニットについて，4.2 節では時間遅れについて，4.3 節では副結合についてそれぞれ実験・検討を行う．

第4章 初期ネットワーク構造の余分な要素除去

この章では SpikeProp の初期のネットワーク構造（初期中間層ユニット数，時間遅れ，初期副結合数）の不要な要素を除去する方法を提案する．そして，それぞれの提案法により学習性能の向上をめざす．

4.1 提案法——不要中間層ユニット除去——

この節では，SpikeProp において，スパイクの発火時刻だけでなくスパイク数も確実に学習できるように，中間層ユニット数を自動的に調節しつつ調整 WD 法にて学習する手法を提案する．4.1.1 節で基本方針について検討し，冗長な中間層ユニットを学習中に除去する方針を定める．この方針では，どのタイミングで・どのユニットを除去するのかが重要である．そこで 4.1.2 節でタイミングについて，4.1.3 節でユニットの選択について述べる．

4.1.1 中間層ユニット数調整の方針

中間層ユニット数を学習中に調整する方針としては，中間層ユニットが不足する場合に追加してゆくものと，冗長な中間層ユニットを除去してゆくものが考えられる．本稿では，後者について検討する．これは，ユニットを追加する場合は，適切な中間層ユニットを作成することが困難であると考えたためである．冗長な

中間層ユニットを除去する場合は、適切なユニットを選択できれば、除去すること自体は容易である。

この手法により、学習開始時に十分な数の中間層ユニットを用意すれば、適切な中間層ユニット数を持つネットワークを自動で構築できるようになる。その結果、学習を確実に行うことができるようになる。

4.1.2 ユニット除去のタイミング

ユニットを除去する際には、その影響をゼロにすることは難しい。したがって、影響の少ないタイミングを狙いユニットを除去する必要がある。そのため本稿では、ユニット除去のタイミングを誤差が十分に減ったときとした。これはこの時点では学習は終盤に達している。そのため、不要なユニットを判断しやすい。逆に、誤差が十分に減っていない時点で不要なユニットを判定しても、学習の進行により要・不要が変わってしまう可能性がある。このタイミングは、学習中の誤差の値によって判断できる。ここでは、図 4.1 のように誤差 E が E_d 以下になった時点で、誤差が十分に減ったと判断する。

4.1.3 除去対象のユニットの決定法

除去するユニットの候補は学習に不要でなければならない。不要なユニットの判断基準としては除去しても出力に与える影響が少ないというものである。SpikeProp のネットワークは 3 層の階層型ネットワークであり、除去できるユニットは中間層のネットワークに限られる。本稿では、図 4.2 のように不要な中間層ユニットは発火しなくなったユニットとした。このユニットは除去しても、出力に影響はないと考えたためである。これは、発火していないということは、その先に信号が送られていないということであり、そのユニットが不要である可能性が高い。また、前節で述べたように、ユニットが学習中に発火しなくなった場合、そのユニット

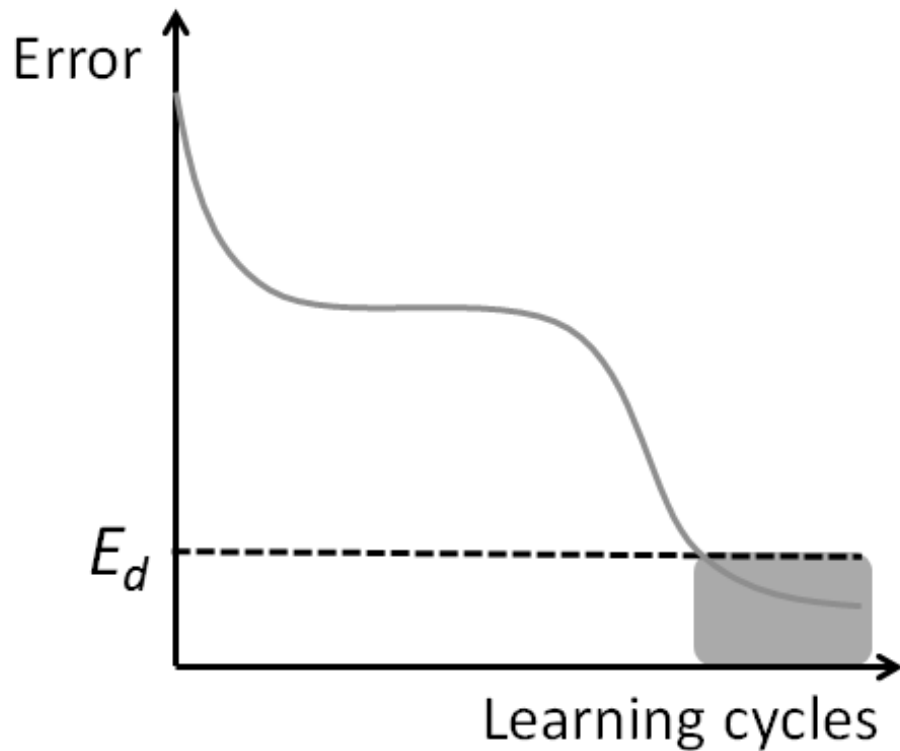


図 4.1: ユニット除去のタイミング

への結合の結合荷重を操作することで、再び発火するようにしている。そのような操作をすることなく、そのままユニットを除去すれば、容易にそのユニットを除去できる。ここで発火しなくなったユニットとはある全てパターンで発火しなかった場合ではなく、1つのパターンでも発火しなかった場合のユニットとした。また興奮性ユニットの方が多く存在おり、除去しても学習の障害に繋がりにくいと考えたため、除去するユニットの対象は興奮性のユニットのみとした。

実験によってこの方法の有効性について検証する。

4.1.4 実験

この節では、簡単な実験を通し、中間層ユニット除去法の有効性について検討する。実験では調整 WD 法と中間層ユニット除去法の学習性能で比較・検討する。

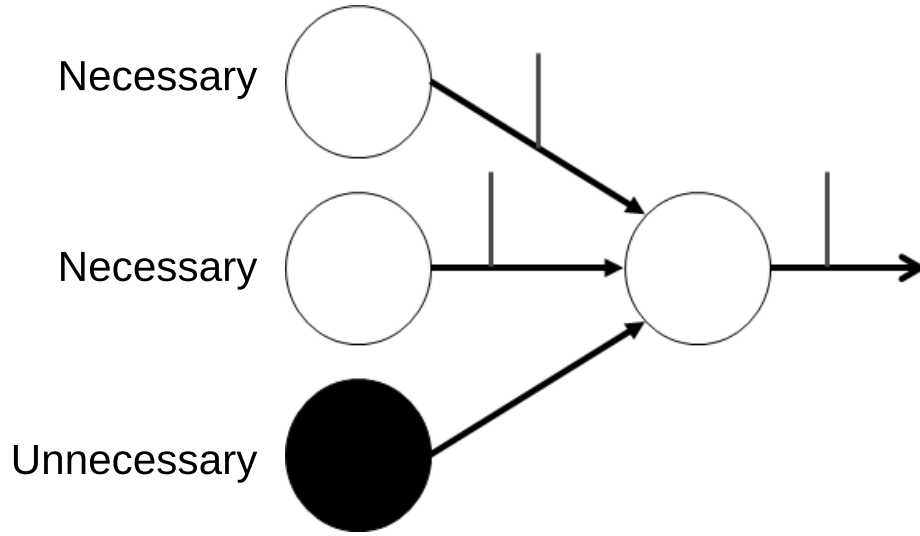


図 4.2: 除去対象のユニットの決定法

表 4.1: 学習に用いたネットワークの構成

Number of input units	3
Number of hidden units	5, 10, 15
Number of output units	1
Number of sub-connections m	8, 12, 16
Delay d^k ($k = 1, 2, \dots, m$)	k
Threshold for activity	7

実験条件

学習対象は，時間版 XOR 問題を対象とし，条件は 3.2.1 節と同様であり，Spike-Prop のネットワークを表 4.1 に示す構成で学習する．また，学習パラメータは，表 3.3 を用いた．また，中間層ユニットの削除法のパラメータ E_d は 1.0 とした． E_d については学習終盤と考えられるかつ，学習の障害に繋がらない程度に中間層ユニットを除去するものを試行錯誤により決定した．削除対象のユニット選択時に，複数のユニットが同時に選択された場合は，それらの中からランダムに 1 つ選択し，そのユニットを削除した．

表 4.2: 中間層ユニット除去法による学習成功率

Number of hidden units	Number of sub-connections		
	8	12	16
5	100	100	44
10	100	98	72
15	100	100	64

結果

中間層ユニット除去法とおよび調整 WD 法の学習性能を比較する。

まず，学習成功率について検証する．表 4.2 に中間層ユニット除去法による学習成功率を示す．調整 WD 単体（表 3.6）と中間層ユニット除去法（表 4.2）と比較すると除去による学習成功率の低下は見られず，副結合数 16 本に対しては中間層ユニットが多い場合，学習成功率が少し良くなった。

つぎに学習回数について検証する．表 4.3 に中間層ユニット除去法を用いた場合の平均学習回数を示す．調整 WD（表 3.7 では中間ユニットが増えるほど多くの学習回数を要した．しかし，中間層ユニット除去法を用いた場合（表 4.3），特に 16 本の場合の増加傾向が抑制することができた．たとえば調整 WD 法では中間層ユニット数が 5 個から 15 個へ増えた場合約 900 回の増加であったが，不要な中間層ユニット除去により約 100 回程度に抑えることができた。

最後に最終的に残った中間層ユニット数を示す．表 4.4 に中間層ユニット除去法により得られたネットワークの最終的な中間ユニット数を示す．中間層ユニット数がいずれの場合でも副結合数が多くなるにつれ，除去がより効果的に働き，副結合数が少ない場合はあまり働かない結果となった．そして，いずれの場合でも学習の阻害になるほどの除去は行われておらず，過剰にユニットを削除することなくネットワークを構築できたといえる。

表 4.3: 中間層ユニット除去法による平均学習回数

Number of hidden units	Number of sub-connrctions		
	8	12	16
5	665	1,150	1,735
10	618	765	1,664
15	785	1,079	1,847

表 4.4: 中間層ユニット除去法による最終中間層ユニット数

Number of hidden units	Number of sub-connrctions		
	8	12	16
5	4.9	4.5	4.7
10	9.9	9.2	8.1
15	14.8	12.1	10.1

4.1.5 検討

調整 WD 法と中間層ユニット除去法の学習成功率を比較した場合、特定の場合のみ中間層ユニット除去法の方が学習性能が良いことが分かった。また、学習回数でも特定の場合、増加傾向の抑制につながった。傾向として、副結合数が多い場合については中間層ユニットを除去をより行う傾向にあり、副結合数が少ない場合についてはあまり中間層ユニット除去が行われない傾向にあることが分かった。そのため、この方法では初期ネットワーク依存性は改善しきれなかったといえる。

4.2 提案法—不要時間遅れ除去—

この節では、SpikeProp において、スパイクの発火時刻だけでなくスパイク数も学習できるように、不要な時間遅れを除去する方法について検討する。4.2.1 節で基本アイデア、4.2.2 節で不要な時間遅れの判定法、4.2.3 節で提案法の具体案を行う。

4.2.1 基本アイデア

不要な時間遅れを除去する方法として、①時間遅れを学習する、②不要な時間遅れを持つ結合を除去するの 2 つの方針が考えられる。しかし、時間遅れを学習することは計算コストが高くなることにつながる。そのため、本提案では不要な時間遅れを除去することで余分な出力スパイクの抑制を試みる。ここではネットワーク内部で教師出力パターンを出力することに貢献していない結合を検出し、それらに共通した時間遅れを不要な時間遅れと判定する。

4.2.2 判定法

この節では前節のアイデアを実現するために必要となる不要な結合を判定する方法を検討する。

調整 WD では式 (3.2) により各結合荷重を調節している。ここで、式 (3.2) の第一項はネットワークの出力誤差を減少させるために、第二項では不要な結合荷重の大きさを 0 となるように働く。ここで第一項の値が 0 となるような結合を不要な結合とみなすことにする。この値が 0 であれば、第二項の働きでその結合は徐々に 0 に近づく。しかし、完全に 0 になるには時間がかかり、ネットワークの規模が大きい（結合数が多い）場合に学習性能の低下を引き起こしていると考えられる。そのため、第二項が 0 となった結合に別途対策を行うことで学習性能の低下を防ぐことができる。

ただし、この考えに基づいて結合そのものを除去すると除去しすぎ、学習の阻害に繋がる可能性がある。また、式 (3.2) の第一項が 0 となるのはその時点で不要な結合であり、最終的な不要な結合とは必ずしも一致しない。これらをふまえ、同じ時間遅れを持つ結合が大域的に不要であればその時間遅れが不要であると判断する。これにより過剰に結合を除去しないようになる。

次節で提案法の具体的な手順を示す。

4.2.3 不要時間遅れ除去法の実案

除去の具体的な流れを図 4.5 に示す。除去は結合荷重の更新ごとに行う。不要な時間遅れの判断は図 4.3, 4.4 のように入力層-中間層の結合、もしくは中間層-出力層の結合に分けて判定を行う。(図 4.6) それぞれに含まれるすべての結合で同じ時間遅れを持つ結合すべてが同時に不要と判断されたとき、その時間遅れを持つ結合をすべてその結合荷重を 0 にする。これにより、不要な時間遅れに対処する。

次節で不要時間遅れ除去法の有効性の確認を検証する。

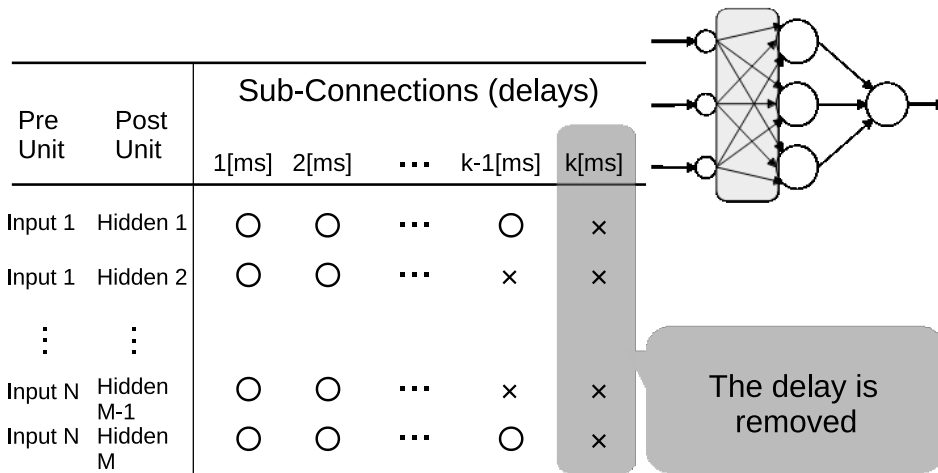


図 4.3: 不要時間遅れの除去（入力層-中間層間）

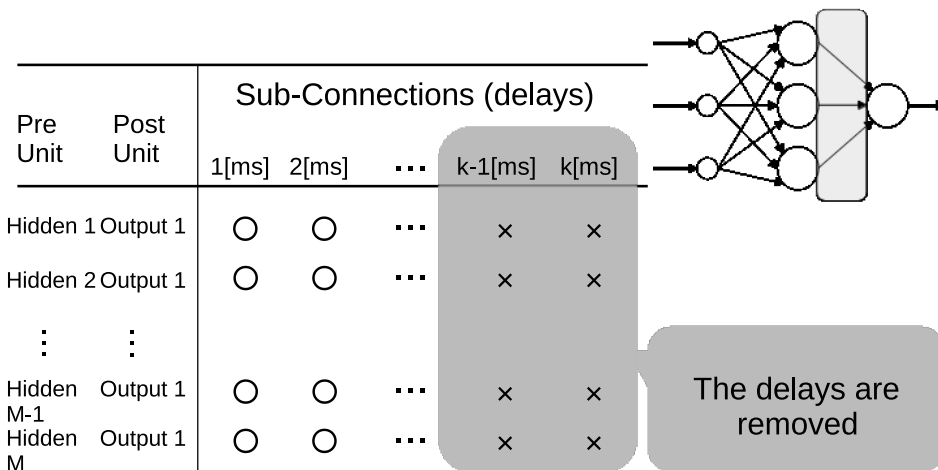


図 4.4: 不要時間遅れの除去（中間層-出力層間）

4.2.4 実験

この節では，時間版 XOR 問題，Iris 問題を用いた実験をとおり，不要な時間遅れ除去法の有効性検討をする．

実験条件（時間版 XOR 問題）

ここでは調整 WD 法と不要な時間遅れ除去法について，学習成功率と平均学習回数と比較する．また，ネットワーク全体（表では RD(IHO) と表記），入力層-中

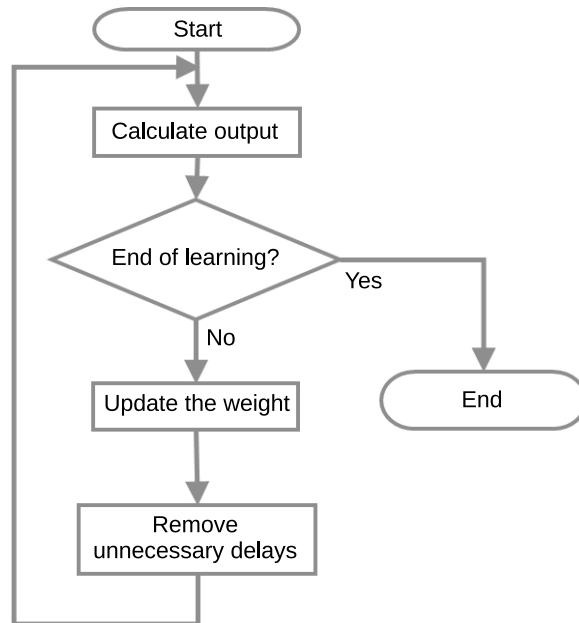


図 4.5: 時間遅れ除去の流れ

間層間のみ（表では RD(IH) と表記），中間層-出力層間のみ（表では RD(HO) と表記）に対しての適用の効果の違いも検討する．

学習対象は，時間版 XOR 問題を対象とし，条件は 3.2.1 節と同様であり，Spike-Prop のネットワークを表 4.5 に示す構成で学習する．また，学習パラメータは，表 3.3 を用いた．

結果（時間版 XOR 問題）

はじめに学習成功率について示す．表 4.6 に各手法による学習成功率を検討する．調整 WD 単体では，最も良い場合の学習成功率は 96%を達成しているが，最大時間遅れが大きくなる毎に成功率が低下し，最大時間遅れが 16[ms] の場合には 50%にまで低下した．それに対し，不要時間遅れ除去法をネットワーク全体，入力層-中間層間に適用した場合にはいずれの場合も 98%以上の学習成功率を得た．しかし，不要時間遅れ除去法を中間層-出力層間に適用した場合は調整 WD 法単体と大差なかった．

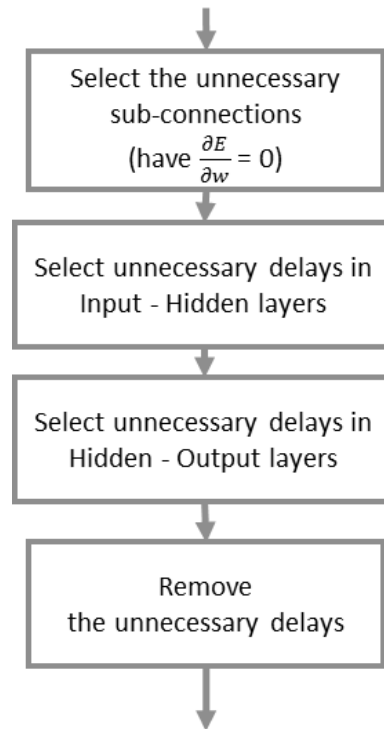


図 4.6: 時間遅れ除去の流れ (詳細)

次に学習回数について検討する．表 4.7 に各手法の学習に成功した場合の平均学習回数を示す．調整 WD では最大時間遅れが大きくなるとともに学習回数も増加していた．不要時間遅れ除去法をネットワーク全体，入力層-中間層間に適用した場合ではむしろ減少傾向にあった．しかし，平均学習回数についても不要時間遅れ除去法を中間層-出力層間に適用した場合は調整 WD 法と大差なかった．

検討

この節では不要な時間遅れ除去法を入力層-中間層間のみ，中間層-出力層間のみ，ネットワーク全体に対しての適用した場合の効果の違いを検討する．不要な時間遅れ除去法を入力層-中間層間のみ，ネットワーク全体に対しての適用した場合は効果的に働いたが，中間層-出力層間のみに適用させた場合は効果的に働かなかった．これは中間層-出力層間は調整 WD 法のみで不要な時間遅れを除去できること

表 4.5: 学習に用いたネットワークの構成 (XOR 問題)

Number of input units	3
Number of hidden units	5
Number of output units	1
Number of sub-connections m	16
Delay (min)	1
Delay (Max)	8, 12, 16
Threshold for activity	7

表 4.6: 不要時間遅れ除去による学習成功率 (XOR 問題)

	時間遅れ (最小-最大)		
	1-8	1-12	1-16
AWD	96	94	50
RD(IHO)	100	100	100
RD(IH)	98	100	100
RD(HO)	100	96	42

が考えられる。そのため、不要な時間遅れ除去法を入力層-中間層間のみ、ネットワーク全体に対しての適用した場合は効果的に働いたと考えた。また、不要な時間遅れ除去法を入力層-中間層間のみ、ネットワーク全体に対しての適用させた場合のそれぞれを比較した場合、学習成功率については大きな差がなかった。しかし、平均学習回数については入力-中間層間のみ適用させた方が少ない回数で学習が終了している。そのため、除去を行う場合、入力層-中間層間のみ適用し、中間層-出力層間については調整 WD 法のみで良いと考えることができる。

表 4.7: 不要時間遅れ除去による平均学習回数 (XOR 問題)

	Delay(min-Max)		
	1-8	1-12	1-16
AWD	1,442	1,080	2,177
RD(IHO)	1,486	1,219	962
RD(IH)	1,470	780	838
RD(HO)	1581	1,572	3,222

表 4.8: 学習に用いたネットワークの構成 (Iris 問題)

Number of input units	5
Number of hidden units	10
Number of output units	3
Number of sub-connections m	16
Delay (min)	1
Delay (Max)	8, 12, 16
Threshold for activity	7

実験条件 (Iris 問題)

ここでは時間版 XOR 問題以外の問題に対しても不要な時間遅れ除去法が有効に働くかを Iris 問題を用いて確認する。また、前節より、不要な時間遅れ除去法は入力層-中間層のみに適用させる。

学習対象は、Iris 問題を対象とし、条件は 3.2.2 節と同様であり、SpikeProp のネットワークを表 4.8 に示す構成で学習する。また、学習パラメータは、表 3.5 を用いた。

表 4.9: 不要時間遅れ除去による学習成功率 (Iris 問題)

	Delay(min-Max)		
	1-8	1-12	1-16
AWD	85	80	45
RD(IH)	85	100	100

表 4.10: 不要時間遅れ除去による平均学習回数 (Iris 問題)

	Delay(min-Max)		
	1-8	1-12	1-16
AWD	1,980	3,967	3,110
RD(IH)	1,980	2,962	2,872

結果 (Iris 問題)

はじめに学習成功率について示す．表 4.9 に各手法による学習成功率を検討する．調整 WD 単体では，最も良い場合の学習成功率は 100%を達成しているが，最大時間遅れが大きい場合には 55%にまで低下した．それに対し，不要時間遅れ除去法を入力層-中間層間に適用させた場合ではいずれの場合も 85%以上の学習成功率を得た．

次に学習回数について検討する．表 4.10 に各手法の学習に成功した場合の平均学習回数を示す．調整 WD と不要時間遅れ除去法と比較すると，最大時間遅れが小さい場合は同じ結果となったが，最大時間遅れが大きくなるにつれて調整 WD 法のみで起こっていた学習回数の増加傾向が不要な時間遅れ除去法により，抑制することができ，少ない学習回数で学習が終了することが分かった．

以上より，不要時間遅れ除去法は学習させる問題に寄らず，また，過剰に除去することなく効果的に働くことを確認した．

4.3 提案法—不要な副結合除去—

この章では結合毎で不要な結合を見つけ、その結合を除去する方法について検討する。そして、初期ネットワーク構造の要素の一つである副結合に対処する。

4.3.1 節で基本アイデア、4.3.2 節で不要な副結合の判定法について、4.3.3 節で抑制法について説明する。

4.3.1 基本アイデア

不要な副結合を抑制する方法として本稿では学習毎に要・不要を判定し、不要と判定された副結合の結合荷重をすばやく抑制することを考えた。

次節で具体的な要・不要の判定法、抑制法について説明する。

4.3.2 判定法

この節では、前節のアイデアを実現する方法を検討する。この方法は不要な副結合を判定する部分と不要と判定された副結合の結合荷重を抑制する部分からなる。

不要な副結合を判定する方法は各ユニットの入力・出力スパイクのタイミングに着目する。具体的には、ユニットの出力スパイクの後に入力されたスパイクは不要であると考え、この概念図を図 4.7 に示す。ユニットの出力スパイク（最下段）より前に副結合 1, 2 を通じて入力されたスパイクはユニットの発火に貢献している。そのため、必要な結合であることがわかる。しかし、ユニットの発火後に入力された副結合 3, 4 は発火に貢献していない。そのため、本提案法では副結合 3, 4 を不要な結合であると見なすこととした。

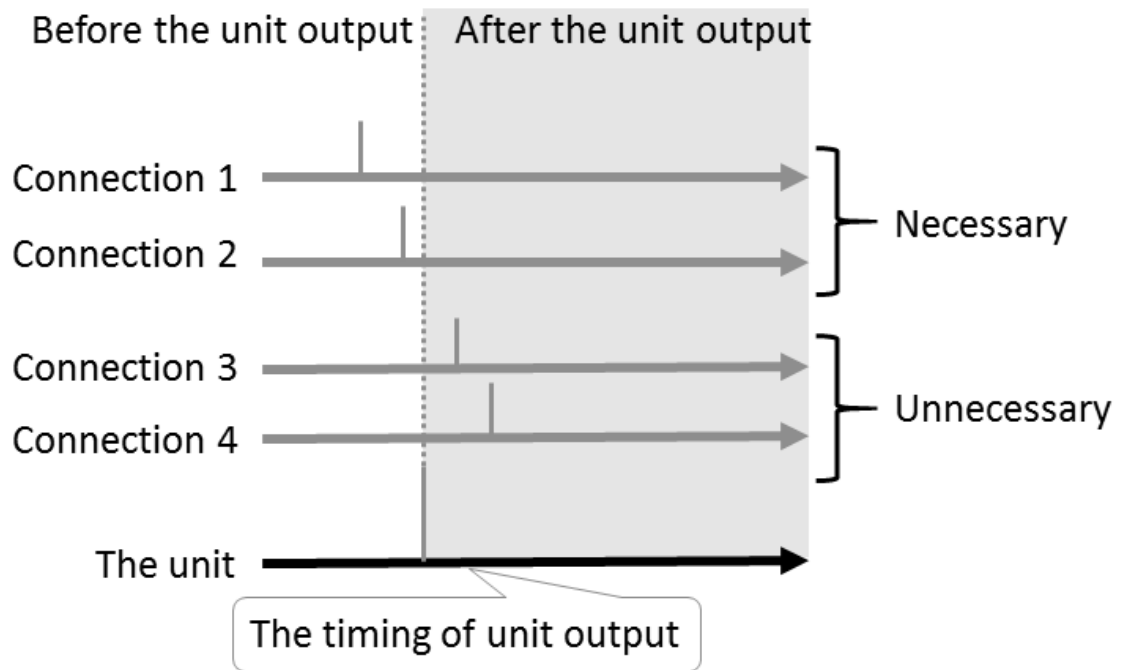


図 4.7: 結合毎の要・不要の判断

4.3.3 抑制法

素早く抑制する方法として、WD 法の効果を強める、除去（結合荷重を 0 にする）の二つの方針が考えられる。しかし、前者では時間がかかる可能性、また、効果を強めすぎた場合除去と変わらない状況になることが考えられる。そのため、本提案法では除去を採用した。

以上をまとめると、提案手法を適用した学習の手順は図 4.8 に示すように学習毎に行う。なお、前節にしたがい、副結合除去法は入力層-中間層間の結合に対してのみ適用する。また、必要な発火は調整 WD 法で用いた θ_{num} を用い、決定することとした。

次節で不要な副結合除去法の有効性の検討をする。

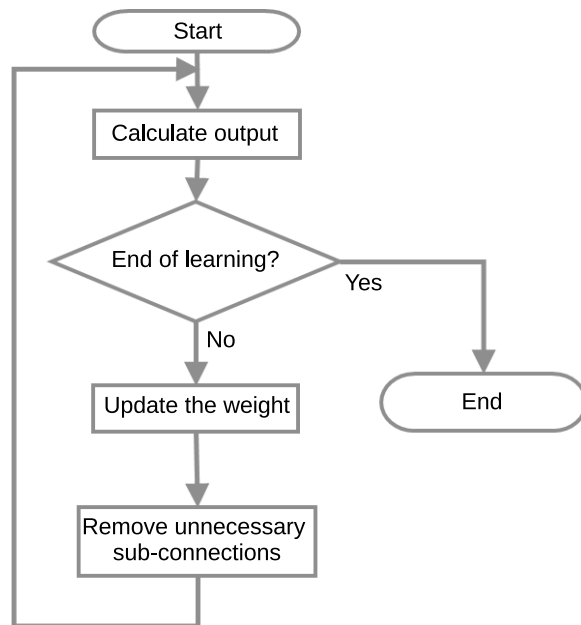


図 4.8: 提案法を加えたフローチャート

4.3.4 実験

この節では、時間版 XOR 問題、Iris 問題を用いた実験をとおり、不要な時間遅れ除去法の有効性検討をする。

実験条件（時間版 XOR 問題）

ここでは調整 WD 法と不要な時間遅れ除去法（表では RC と表記）について、学習成功率と平均学習回数と比較する。

学習対象は、時間版 XOR 問題を対象とし、条件は 3.2.1 節と同様であり、Spike-Prop のネットワークを表 4.11 に示す構成で学習する。また、学習パラメータは、表 3.3 を用いた。

結果（時間版 XOR 問題）

はじめに各手法による学習成功率について検討する。表 4.12 に各手法による学習成功率を示す。調整 WD 単体では、副結合が 8,12 本であれば 100%であったが、

表 4.11: 学習に用いたネットワークの構成 (XOR 問題)

Number of input units	3
Number of hidden units	5, 10, 15
Number of output units	1
Number of sub-connections m	8, 12, 16
Delay d^k ($k = 1, 2, \dots, m$)	k
Threshold for activity	7

表 4.12: 不要な副結合除去による学習成功率 (XOR 問題)

	Number of sub-connections		
	8	12	16
AWD	100	92	50
SC	100	100	100

[%]

16 本に増えた場合は 50% に低下し、初期の副結合数に対する依存があった。それに対し、不要な副結合除去法はいずれも 100% の学習成功率を得た。これは副結合数が少ない場合もその性能を損なう事なく、多い場合にも対応できたことを表している。

次に学習回数について検討する。表 4.13 に各手法の学習に成功した場合の平均学習回数を示す。調整 WD では副結合数が大きくなるとともに学習回数も増加し、副結合数が 16 本のときには 8 本の時と比べ 3 倍を超えた。不要な副結合除去法も最大時間遅れとともに学習回数は増加しているが、いずれの場合でも調整 WD 法を下回る結果となった。特に 8 本と 16 本と比較すると、2 倍程度に抑えることができた。

表 4.13: 不要な副結合除去による平均学習回数 (XOR 問題)

	Number of sub-connections		
	8	12	16
AWD	663	1,161	2,177
SC	519	625	1,092

表 4.14: 学習に用いたネットワークの構成 (Iris 問題)

Number of input units	5
Number of hidden units	10
Number of output units	3
Number of sub-connections m	8, 12, 16
Delay d^k ($k = 1, 2, \dots, m$)	k
Threshold for activity θ	7

実験条件 (Iris 問題)

ここでは時間版 XOR 問題以外の問題に対しても不要な時間遅れ除去法が有効に働くかを Iris 問題を用いて確認する。また、前節より、不要な時間遅れ除去法は入力層-中間層のみに適用させる。

学習対象は、Iris 問題を対象とし、条件は 3.2.2 節と同様であり、SpikeProp のネットワークを表 4.14 に示す構成で学習する。また、学習パラメータは、表 3.5 を用いた。

結果 (Iris 問題)

はじめに学習成功率について示す。表 4.15 に各手法による学習成功率を検討する。調整 WD 単体では、最も良い場合の学習成功率は 100%を達成しているが、副

表 4.15: 不要な副結合除去による学習成功率 (Iris 問題)

	Number of Sub-connections		
	8	12	16
AWD	100	85	45
SC	100	100	100

表 4.16: 不要な副結合除去による平均学習回数 (Iris 問題)

	Number of Sub-connections		
	8	12	16
AWD	1,204	3,200	3,110
SC	985	1,121	1,390

結合数が多くなるにつれて 55%にまで低下した。それに対し、不要な副結合除去法を入力層-中間層間に適用させた場合ではいずれの場合も 100%の学習成功率を得ることができた。

次に学習回数について検討する。表 4.10 に各手法の学習に成功した場合の平均学習回数を示す。調整 WD と不要な副結合除去と比較すると、調整 WD 法では平均学習回数は増加傾向にあったが、不要な副結合の除去により、その増加傾向は大きく抑制することができた。特に調整 WD 法のみでは 16 本は 8 本のときの約 2.6 倍学習に時間を要していたが、不要な副結合除去法を適用した場合、約 1.4 倍程度に抑えることができた。

これらにより、不要な副結合除去法についても学習させる問題に寄らず効果的に働くことを確認した。

以上より、初期のネットワーク構造の学習性能依存性を解消し、学習性能向上をするため、不要な中間層ユニット除去、不要な時間遅れ除去、不要な副結合除

去のそれぞれの方法を提案・検討した.

不要な中間層ユニット除去についてはユニットの除去ができたものの, 初期ネットワーク依存性は改善しきれなかった. また, 不要な時間遅れ, 不要な副結合の除去により, 初期の最大時間遅れが大きい場合や副結合が多い場合に対しても学習性能の向上が見られた.

今後は効果的なユニット除去法についてさらに検討するとともに, それぞれの方法がそれぞれ独立的に適用されているため, これらの方法を組み合わせた方法を検討する必要がある.

第5章 まとめ

本研究では，Booiij らの SpikeProp においてスパイクの時刻とその数を学習させる方法について検討した．

Booiij らの SpikeProp は複数本の出力スパイクの時刻の学習が行える一方で学習を行った本数以降の発生するスパイクについては，有無，時刻を調整できない．このため，実際に時系列情報処理などに応用を考えた際に不具合に繋がる可能性がある．

そこで，古典的なシグモイド型ニューラルネットワークモデルにおいて耐ノイズ性の向上の手法として使われる Weight Decay に注目し，Weight Decay の効果によって，不要な結合荷重を 0 に近づけることができる．これを適用することで，学習に必要なではない入力信号の変化は小さくなると予想される．しかし単純に適用するのではなく，荷重減衰の働きの強さを学習の進行に応じて切り替えることで，これがより効果的に働くような方法を提案した（調整 WD 法）．調整 WD 法を Booiij らの学習法に適用した結果，スパイクの時刻とその数を学習させることができ，また，単純に WD 法を適用する場合と比較し，学習性能（学習成功率・学習回数）の向上に効果があることを，複数のベンチマーク問題の学習をつうじ確認した．しかし，この方法は初期のネットワーク構造（初期中間層ユニット数，時間遅れ，初期副結合数）に学習性能が依存することが分かった．

そこで，十分に多い中間ユニットを持つネットワークから，不要な中間ユニットを除去することで自動で適切なユニット数にする不要中間層ユニット除去について検討した．除去するタイミングについては十分に誤差が小さくなったとき，除去対象のユニットは発火しないユニットとし，除去を行った．この方法を適用し

た結果、初期ネットワーク依存性は改善しきれなかった。

また、時間遅れについては不要な時間遅れを除去する方法を検討した。除去するタイミングは学習毎、除去対象は学習で不要と判断された時間遅れを持つ結合を除去した。不要と判断する基準は結合荷重更新式の第一項を参考にし、その値が0になり、かつそれが大域的に共通した時間遅れである場合、その時間遅れを不要な時間遅れとし、除去を行った。この方法を適用した結果、最大時間遅れの大きさに依存することなく、学習性能の向上が見られた。

また、初期の副結合数が多い場合については不要な副結合を学習毎に判断し、除去する方法について検討した。除去する対象は信号の伝播先のユニットの発火タイミングを基準とし、そのユニットの発火に貢献する結合は必要な結合、貢献しない結合は不要な結合と判断し、不要な結合の除去を行った。この方法を適用した結果、初期副結合数に依らず、学習性能の向上が見られた。これらの方法により、Booij らの SpikeProp は時系列情報を随時処理できる可能性が示された。

調整 WD 法の学習率 η 及び WD のパラメータ ρ は問題に合わせて学習性能の最も良いものを試行錯誤により決定している。そのため今後はそれらを自動調整する方法の検討する必要がある。また今後は効果的な中間層ユニット除去法の検討・適用、そして、今回提案した方法がそれぞれ独立的に適用されているため、これらの方法を組み合わせた方法を検討する必要がある。これらの方法を組み合わせることでより学習性能の向上やより小さいネットワークを自動構成することができ、実装しやすいことが考えられる。また、Iris 問題の学習が可能となったため、汎化能力の検討や、より複雑で一般的な問題の学習を行い、SpikeProp の時系列情報の処理能力を検討していく必要がある。

謝辞

本論文は、著者が三重大大学にて行った研究をまとめたものである。本論文を進めるにあたり、懇切丁寧なご指導と御督励を賜った三重大大学の鶴岡信治教授，高瀬治彦准教授，川中普晴助教，野呂雄一准教授に感謝いたします。また，日頃熱心に討論していただいた情報処理研究室の皆様方に厚く御礼申し上げます。

最後に，本論文をまとめるにあたり，助言，討論，その他お世話になったすべての方々に感謝いたします。

参考文献

- [1] 伊藤英紀:コンピュータ将棋におけるクラスタ並列探索, 電子情報通信学会誌 96(2), pp.117-123 (2013).
- [2] Rosenblatt, Frank : The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Psychological Review 65 (6), pp.386–408 (1958).
- [3] David E. Rumelhart; James L. McClelland : Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Cambridge, MA: MIT Press, (1986).
- [4] 郭 帥, 榎本 誠, 川村 春美, 大谷 淳 : ニューラルネットワークを利用する顔画像からの似顔絵自動生成の検討 : 主観評価に基づく比較検討, 第 14 回情報科学技術フォーラム講演論文集, pp.231–232, (2015).
- [5] 渡辺 栄治, 東 拓矢, 尾関 孝史, 小濱 剛 : 講義における講師の動作の分析, 映像情報メディア学会技術報告, vol. 38, No.46, pp.61–64, (2014).
- [6] 原田 達也 : 機械学習による画像理解, 映像情報メディア学会誌, vol.69, pp. 117–123,(2015).
- [7] 池田 英彬, 山内 ゆかり : 追学習の機構を備えた Spiking Hopfield Network の提案, 電子情報通信学会技術研究報告 (NC), vol. 112, pp.49–54, (2013).
- [8] 大保 武慶, 久保田 直行 : スパイキングニューラルネットワークを用いた人間状態計測システム, 26th Fuzzy System Symposium, pp.234–239, (2010).

- [9] Sander M. Bohte: The evidence for neural information processing with precise spike-times : A survey, *Natural Computing*, Vol.3, pp.195–206 (2004).
- [10] 黒江康明 : スパイキングニューラルネットワーク — 学習法を中心として, システム／制御／情報, Vol.48, No.2, pp.57–62, (2004).
- [11] A. Kasiński and F. Ponulak : Comparison of Supervised Learning Methods for Spike Time Coding in Spiking Neural Networks,’ *International Journal of Applied Mathematics and Computer Science*, Vol.16, No.1, pp.101–113, (2006).
- [12] A. Mohemmed and S. Schliebs, : SPAN Spike Pattern Association Neuron for Learning Spatio-Temporal Spike Patterns, *International Journal Neural Systems*, Vol.22, No.4, (2012).
- [13] Olaf Booij, Hieu tat Nguyen: A gradient descent rule for spiking neurons emitting multiple spikes, *Information Processing Letters*, Vol.95, pp.552–558 (2005).
- [14] S. M. Bohte, J. N. Kok, and J. A. La Poutre: Error-backpropagation in temporally encoded networks of spiking neurons, *Neurocomputing*, Vol.48, pp.17–37 (2002).
- [15] R. Reed : Pruning Algorithms — A Survey, *IEEE Trans. Neural Networks*, Vol.4, No.5, pp.740–747, (1993).
- [16] D.C. Plaut, S.J. Nowlan and G.E. Hinton : Experiments on learning by back propagation, Technical Report CMU-CS-86-126, Carnegie-Mellon University, (1986).
- [17] M. Lichman : UCI Machine Learning Repository, University of California, <http://archive.ics.uci.edu/ml/>, 参照 Dec. 10, 2015.

発表

- [A1] Takashi Matsumoto, Yuta Shin, Haruhiko Takase, Hiroharu Kawanaka and Shinji Tsuruoka: Suppress Redundant Spikes in Extended SpikeProp — Effect of Adjustment of Weight Decay Parameters—, Proceedings of the Sixth International Workshop on Regional Innovation Studies, pp. 59–62, (2014).
- [A2] Takashi Matsumoto, Yuta Shin, Haruhiko Takase, Hiroharu Kawanaka and Shinji Tsuruoka: A Learning Method for Extended SpikeProp without Redundant Spikes— Automatic Adjustment of Hidden Units—, Proc. of the 7th Intl. Conf. on Soft Computing and Intelligent Systems, and the 15th Intl Symposium on Advanced Intelligent Systems, pp. 1465–1469, (2014).
- [A3] Takashi Matsumoto, Yuta Shin, Haruhiko Takase, Hiroharu Kawanaka and Shinji Tsuruoka: A Learning Method for SpikeProp without Redundant Spikes — Suppress Redundant Weights According to Spike Time—, Proceedings of the Seventh International Workshop on Regional Innovation Studies, pp. 76–79, (2015).
Outstanding Paper Award
- [A4] Takashi Matsumoto, Haruhiko Takase, Hiroharu Kawanaka and Shinji Tsuruoka: A Learning Method for SpikeProp without Redundant Spike — Effective Removal of Unnecessary Elements—, International Symposium for Sustainability by Engineering at MIU, (2015).
(Poster Session)

- [A5] Takashi Matsumoto, Haruhiko Takase, Hiroharu Kawanaka and Shinji Tsuruoka: A Learning Method for SpikeProp without Redundant Spikes — Automatic Adjusting Delay of Connections —, IEEE 8th International Workshop on Computational Intelligence and Applications, pp. 15–19, (2015).
- [A6] Takutoshi Nakayama, Haruhiko Takase, Hiroharu Kawanaka, Shinji Tsuruoka and Takashi Matsumoto: Suppress Redundant Spikes in SpikeProp —Locally Application of Removing Redundant Delays—, 7th International Conference on Emerging Trends in Engineering & Technology, pp.184–187, (2015).
- Best Paper Award
- [A7] 松本崇：余分なスパイクを出さない拡張 SpikeProp の学習-不要な結合の除去の効果の検討；平成 25 年度電子情報通信学会 東海支部 卒業研究発表会 論文集, p.1, (2014).
- [A8] 松本崇，新友太，高瀬治彦，川中普晴，鶴岡信治：スパイクニューラルネットワークにおける複数スパイクの発火時刻の学習法—ユニット削除の効果— 第 37 回東海ファジィ研究会予稿集, pp. 10-1–10-4, (2014).
- [A9] 松本崇，新友太，高瀬治彦，川中普晴，鶴岡信治：拡張 SpikeProp における余分な出力スパイクの抑制-調整 WD 法による学習性能の向上-平成 26 年度電気・電子・情報関係学会東海支部連合大会講演論文集, J5-1, (2014).
- [A10] 松本崇，新友太，高瀬治彦，川中普晴，鶴岡信治：スパイクニューラルネットワークにおける複数スパイクの発火時刻の学習法—ユニット削減による学習成功率の向上—, 第 30 回ファジィシステムシンポジウム講演論文集, pp. 190～195, (2014).

- [A11] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp において余分なスパイクを抑制する学習法—時間遅れに着目した一検討—, 電子情報通信学会技術研究報告 (NC), Vol.114, No. 362, pp. 49–53, (2014).
- [A12] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp において余分なスパイクを抑制する学習法—適用部位による効果の違い—, 第38回東海ファジィ研究会予稿集, pp. 8-1–8-5, (2015).
- [A13] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: 時系列情報処理を行うニューラルネットワーク, 三重大学アカデミックフェア 2015, (2015).
- [A14] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp において余分なスパイクを抑制する学習法 —ネットワーク構造の自動調整による構造依存性の軽減—, 第14回情報科学技術フォーラム講演論文集, pp. 317–322, (2015).
- [A15] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: 拡張 SpikeProp における余分な出力スパイクの抑制 — 発火タイミングに着目した一検討—, 平成27年度電気・電子・情報関係学会東海支部連合大会講演論文集, N3-4, (2015).
- [A16] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp において余分なスパイクを抑制する学習法—不要な副結合抑制に関する検討—, 電子情報通信学会技術研究報告 (NC), Vol.115, No. 384, pp. 43–48, (2015).
- [A17] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: 時系列情報処理を行うニューラルネットワーク - 不要な要素除去による性能向上 -, 三重大学アカデミックフェア 2016, (2016).
- [A18] 松本崇, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp における余分な出力スパイクの抑制— 不要な結合除去に関する検討 —, 2016年電子情報通信学会総合大会 (3/16 発表予定).