数理解析研究所講究録
1148 巻 2000 年 206-211



# 決定性プッシュダウンオートマトンを模倣する拡張単純回帰ネットワークの構成法

守谷 純之介 (Junnosuke Moriya)　　西野 哲朗 (Tetsuro Nishino)

電気通信大学大学院 電気通信学研究科

〒 182-8585 東京都調布市調布ヶ丘 1-5-1
e-mail: {jmoriya,nishino}@ice.uec.ac.jp

## 1 はじめに

人間の言語獲得のメカニズムの解明は，言語学および認知心理学においてもっとも重要な課題である．現在，J. L. Elman によって提案されたモデルである単純回帰ネットワーク (Simple Recurrent Network; SRN) が，言語獲得のモデルとしての妥当性をもつか否かとういう問題が活発に議論されている [8, 11]．これまでの主な結果として，McClelland らは，単純回帰ネットワークが有限オートマトンを学習可能であることを実験的に示した [1]．さらに，Elman は，単文や複文を構成する単語の系列が与えられた単純回帰ネットワークが，各単語の次の単語を予想するように，トレーニング可能であることを示した [2, 3]．本稿では，単純回帰ネットワークの計算能力に着目する．

計算能力の観点から，単純回帰ネットワークは，Siegelmann と Sontag の結果の自然な拡張により，Turing 機械を模倣可能であることがわかる [7]．しかし，そのような模倣を行う場合，単純回帰ネットワークの各ゲートが，出力に関して無限の精度をもつことが必要になる．我々は，これまでに，現実的な仮定として単純回帰ネットワークの各ゲートが計算する関数の値域が有限集合である場合，単純回帰ネットワークの計算能力は出力付き有限オートマトンと一致することを示した [9, 10]．本稿では，単純回帰ネットワークの各ゲートが $\{0,1\}$ を値域とする閾値関数を計算するものとし，単純回帰ネットワークが決定性プッシュダウンオートマトンを模倣可能となるように，単純回帰ネットワークを拡張する．そして，与えられた決定性プッシュダウンオートマトンを模倣する拡張単純回帰ネットワークの構成方法を示す．

## 2 単純回帰ネットワークの拡張

**単純回帰ネットワーク**は，**入力ユニット** (input unit)，**出力ユニット** (output unit)，**隠れユニット** (hidden unit)，**文脈ユニット** (context unit) から成る (図 1参照)．各ユニットは，ゲートの集合であり，各ゲートが計算する関数を**発火関数**と呼ぶ．単純回帰ネットワークの文脈ユニットに属する各ゲートは，そのゲートに対応する 1 単位時間前の隠れユニットのゲートの出力を保持する．そのため，文脈ユニットに属するゲートの個数と隠れユニットのゲートの個数は等しくなければならない．隠れユニットは入力ユニットと文脈ユニットに依存し，出力ユニットは隠れユニットにのみ依存する．単純回帰ネットワークを形式的に定義する．

**定義 2.1 単純回帰ネットワーク (SRN)** とは *8*-項組 $\mathcal{E}=(G,I,O,H,C,A,w,h)$ であり，ここに，

1. $G=(V,E)$ は有限グラフ，ここに $V$ はゲートの有限集合であり，つぎの４つの集合に分割される: **入力ユニット** $I=\{x_1,x_2,\ldots,x_n\}\subseteq V$，**出力ユニット** $O=\{y_1,y_2,\ldots,y_m\}\subseteq V$，**隠れユニット** $H=\{h_1,h_2,\ldots,h_k\}\subseteq V$，および**文脈ユニット** $C=\{c_1,c_2,\ldots,c_k\}\subseteq V$．各ユニットは共通部分をもたないものとする．このとき，$E=\{(v_1,v_2)\mid(v_1,v_2)\in I\times H\cup C\times H\cup H\times C\cup H\times O\}$ は $G$ の辺の集合である．隠れユニットから文脈ユニットのへの辺は $(h_i,c_i),h_i\in H,c_i\in C,1\leq i\leq k$ のみ許される．
2. $A=(A_1,\ldots,A_k)\in\mathbf{R}^k$ は**初期出力**.
3. $w:E\to\mathbf{R}$ は辺への**重み**の割り当てとする．辺 $(v_1,v_2)$ の重みを $w(v_1,v_2)$ と表す．辺 $(h_i,c_i),h_i\in H,c_i\in C,1\leq i\leq k$ の重みは *1* とする.
4. $h:V\to\mathbf{R}$ はゲートへの**閾値**の割り当てとする．ゲート $g$ の閾値を $h(g)$ と表す．文脈ユニットに属する任意のゲート $c\in C$の閾値は，*1* とする.

SRN における各ゲートの発火関数の計算を次のように定義する．時刻 $t$ におけるゲート $g$の出力を $g(t)$ で表すものとする．いま，SRN のあるゲート $g$に対して，$g$の入次数を $k$とし，$g$に入る辺をそれぞれ $(g_1,g),\ldots,(g_k,g)$ とする．ゲート $g$の発火関

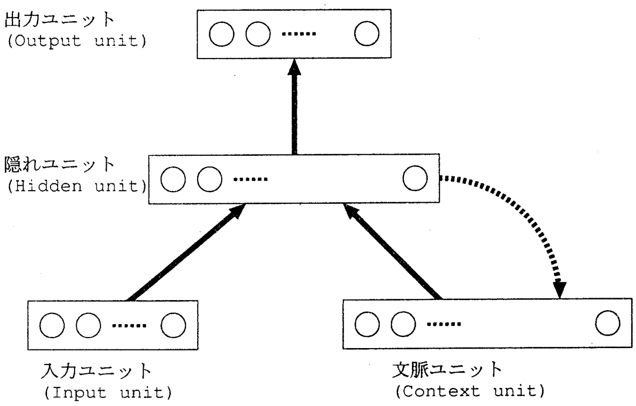


図 1: 単純回帰ネットーワーク

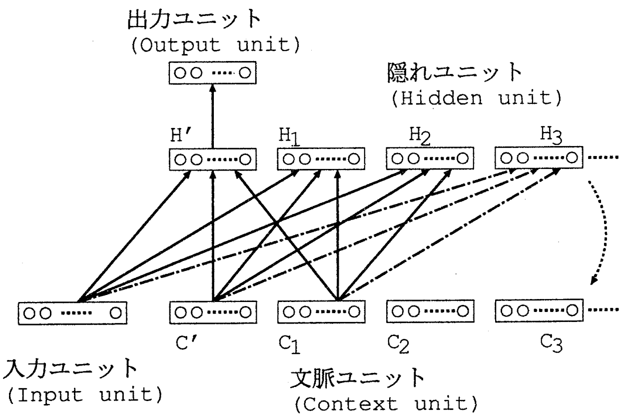


図 2: 拡張単純回帰ネットーワーク

数を $f_g$ とするとき，時刻 $t$ のゲート $g$ の出力を発火関数 $f_g$ によって

$$g(t) = f_g\left(\sum_{i=1}^{k} w(g_i, g)\cdot g_i(t-1) - h(g)\right)$$

と定義する．また，文脈ユニットに属するゲート $c_i, 1 \leq i \leq k$ の時刻 $t$ での出力を $c_i(t)$ とするとき，時刻 $t = 0$ での出力 $c_i(0)$ は，初期出力 $A$ より，$c_i(0) = A_i$ と定義する．

ニューラルネットの研究分野において，各ゲートが計算する発火関数として，つぎのような発火関数が用いられる (例えば，[6] 参照)．**閾値関数** (threshold function):

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0. \end{cases}$$

**ロジスティック関数** (logistic function):

$$f(x) = \frac{1}{1+e^{-x}}.$$

ここに，$e$ は自然対数である．**部分線形関数** (piecewise linear function):

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1. \end{cases}$$

SRN の各ゲートが計算する発火関数が，ロジスティック関数，もしくは部分線形関数である場合，SRN の各辺の重みが実数であることから，各ゲートの出力は実数になることに注意する．各ゲートが計算する発火関数が部分線形関数である場合，Siegelmann らの結果の自然な拡張により，SRN は Turing 機械を模倣することが可能になる [6, 7]．しかしこの場合，先に述べたように，SRN の各辺の重みが実数であることから，各ゲートの出力は実数であり，ゲートの出力に関して無限の精度が必要とされる．

我々は，これまでに，SRN の各ゲートが計算する発火関数の値域が有限集合である場合，SRN の計算能力は出力付き有限オートマトンと等しくなることを示した [9, 10]．本論では，各ゲートが計算する発火関数を閾値関数として，SRN の計算能力の拡張を試みる．

各ゲートが発火関数として閾値関数を計算する場合，重みと閾値を整数に制限しても，ゲートの計算能力は変わらないことが知られている (例えば，[5] 参照)．よって，以下では，SRN の重みと閾値として整数のみ考える．本論では，単純回帰ネットワークを以下のように拡張する (図 2 参照)．

**定義 2.2 拡張単純回帰ネットワーク**とは 8-項組 $\mathcal{E} = (G, I, O, H, C, A, w, h)$ であり，ここに，

*1'.* $G = (V, E)$ は有向グラフ，ここに $V$ はゲートの集合であり，つぎの 4 つの集合に分割される：**入力ユニット** $I = \{x_1, x_2, \ldots, x_n\} \subseteq V$, **出力ユニット** $O = \{y_1, y_2, \ldots, y_m\} \subseteq V$, **隠れユニット** $H = \{h_1, h_2, \ldots\} \subseteq V$, および**文脈ユニット** $C = \{c_1, c_2, \ldots\} \subseteq V$．各ユニットは共通部分をもたないものとする．また，隠れユニットと文脈ユニットはつぎのように分割される．

$$H = H' \cup H_1 \cup H_2 \cup \cdots,$$

$$C = C' \cup C_1 \cup C_2 \cup \cdots.$$

$H', H_1, H_2, \ldots$ および $C', C_1, C_2, \ldots$ は，それぞれ共通部分をもたないものとする．また，$k, k' \in \mathbf{N}$ を定数とし，$|H'| = |C'| = k$, $|H_1| = |C_1| = |H_2| = |C_2| = \cdots = k'$ を満たすものとする．このとき，

$$\begin{aligned} E \;=\; & \{(v_1, v_2) \mid (v_1, v_2) \in I \times H' \cup I \times H_1 \\ & \cup C' \times H \cup H \times C \cup C_1 \times H' \cup \\ & C_1 \times H_1 \cup C_1 \times H_2 \cup H' \times O, \end{aligned}$$

または，任意の $i \geq 2$ に対して，
$(v_1, v_2) \in I \times H_i \cup\ C' \times H_i \cup$
$C_1 \times H_i \cup C_i \times H_{i-1} \cup C_i \times H_i$
$\cup\, C_i \times H_{i+1}\}$

は $G$ の辺の集合である．隠れユニットから文脈ユニットのへの辺は，任意の $i$ に対して，$(h_i, c_i), h_i \in H, c_i \in C$のみ許される．

*2'.* $A = (A_1, \ldots, A_{k+k'}) \in \mathbf{B}^{k+k'}$は初期出力．

*3'.* $w : E \rightarrow \mathbf{Z}$ は辺への**重み**の割り当てとする．任意の $i$ に対して，辺 $(h_i, c_i), h_i \in H, c_i \in C$ の重みは *1* とする．いま，$H_i, H_{i'}, i \neq i', i \geq 3, i' \geq 3$に属するゲートを$h_{i,j}, h_{i',j}\ 1 \leq j \leq k'$ とする．このとき，任意の $i, i'$に対して，

$$\begin{array}{rcll} w(x, h_{i,j}) & = & w(x, h_{i',j}), & x \in I, \\ w(c', h_{i,j}) & = & w(c', h_{i',j}), & c \in C', \\ w(c, h_{i,j}) & = & w(c, h_{i',j}), & c \in C_1 \end{array}$$

を満たすものとする．さらに，$C_i$に属するゲートを $c_{i,j}, 1 \leq j \leq k$としたとき，任意の $1 \leq j, j' \leq k'$に対して，

$$\begin{array}{rcl} w(c_{i,j}, h_{i,j'}) & = & w(c_{i',j}, h_{i',j'}), \\ w(c_{i,j}, h_{i+1,j'}) & = & w(c_{i',j}, h_{i'+1,j'}), \\ w(c_{i,j}, h_{i-1,j'}) & = & w(c_{i',j}, h_{i'-1,j'}) \end{array}$$

を満たすものとする．

*4'.* $h : V \rightarrow \mathbf{Z}$ はゲートへの**閾値**の割り当てとする．文脈ユニットに属する任意のゲート $c \in C$ の閾値は，*1* とする．

ここで，文脈ユニット $C'$と $C_1$に属するゲート $c_i, 1 \leq i \leq k + k'$の時刻 $t = 0$ での出力 $c_i(0)$ は，初期出力 $A$ より，$c_i(0) = A_i$と定義する．その他のゲートの時刻 $t = 0$ での出力は，すべて 0 であるとする．

## 3 決定性プッシュダウンオートマトン

決定性プッシュダウンオートマトン (DPDA) を以下のように定義する (例えば [4] 参照)．

**定義 3.1** DPDA $M$とは 7 項組 $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ である．ただし，

*1.* $Q$ は**状態**の有限集合，

*2.* $\Sigma$ は**入力アルファベット**と呼ばれるアルファベット，

*3.* $\Gamma$ は**スタック記号**と呼ばれるアルファベット，

*4.* $q_0 \in Q$ は**初期状態**，

*5.* $Z_0 \in \Gamma$ は**初期スタック記号**と呼ばれる特定のスタック記号，

*6.* $F \subseteq Q$ は**最終状態**の集合，

*7.* $\delta : Q \times (\Sigma \cup \{\epsilon\} \times \Gamma \rightarrow Q \times \Gamma^*$の型をした遷移関数．ただし，次の条件を満たす．

   *a.* ある $a \in \Sigma$ に対して，$\delta(q, a, X) \neq \emptyset$ ならば，$\delta(q, \epsilon, X) = \emptyset$.

   *b.* 任意の $q \in Q, a \in \Sigma \cup \{\epsilon\}, X \in \Gamma$ に対して，$|\delta(q, a, X)| \leq 1$.

とする．

定義より，DPDA の動作としてつぎの 2 種類が考えられる．

I. 入力記号を 1 つ使用する動作．

II. 入力記号を使用しない動作 ($\epsilon$-動作と呼ぶ)．

DPDA によって受理される言語を**決定性文脈自由言語** (deterministic context-free language，決定性 CFL) と呼ぶ．DPDA の性質として，つぎのことが知られている (例えば [4] 参照)．

**補題 3.1** 任意の *DPDA* $M$に対して，$M$と等価な *DPDA* $M' = (Q', \Sigma', \Gamma', \delta', q'_0, Z'_0, F')$ が存在し，$\delta'(q, a, X) = (p, \gamma)$ ならば，$|\gamma| \leq 2$ を満たす．

**補題 3.2** 任意の *DPDA* $M$に対して，$M$と等価な *DPDA* $M' = (Q', \Sigma', \Gamma', \delta', q'_0, Z'_0, F')$ が存在し，$\delta'(q, a, X) = (p, \gamma)$ ならば，$\gamma$は$\epsilon$，$X$，$YX$のいずれかの形である．

以上の補題より，任意の DPDA に対して，スタック操作に関し，

1. トップ記号を pop するか，

2. スタック記号を 1 つ push するか，

3. トップ記号を変更しないか，

のいずれかの場合のみ考えればよい．

次に，DPDA に与えられる入力と SRN に与えられる入力の対応について考える．DPDA に関して，次のことが知られている．

**補題 3.3** 任意の *DPDA* $M$に対して，$M$と等価な *DPDA* $M'$が存在し，$M'$は任意の入力に対して入力列全体を読み込む．

**補題 3.4** 任意の *DPDA* $M$に対して，$M$と等価な *DPDA* $M'$が存在し，$M'$が受理状態では$\epsilon$-動作をもたない．

補題 3.3より，ある DPDA $M$が入力を途中まで読み込んだ時点で，$\epsilon$-動作によって入力を最後まで読み込まないようなことは起こりえないと仮定してよい．

いま，DPDA $M$に受理される言語を $L$ とする．$w \in L$ の長さを $n$ とし，$w = w_1 w_2 \cdots w_n, w_i \in \Sigma, 1 \leq i \leq n$ と仮定する．$w$が $M$によって受理される際，ある時点で入力記号 $w_i$を使用する動作が行われ，つぎの時点では入力記号を使用しない動作，すなわち$\epsilon$- 動作が行われたとする．このようなDPDA の$\epsilon$-動作に対応して，SRN では，文脈ユニットのある特別なゲート $B$の出力が 1 である間 (busy 状態である) は，入力を受けつけないものとする．すなわち，$w_i$に対応する入力が SRN に与えられた後，SRN の文脈ユニットのゲート $B$の出力が 1 であるならば，$w_i$のつぎの入力である $w_{i+1}$に対応する入力は SRN に与えられず，ゲート $B$の出力が 0 になった時点で，$w_{i+1}$に対応する入力が SRN に与えらるものとする．

いま，$\mathcal{E}$の出力ユニットに属するゲートを $O = \{o\}$ とする．DPDA が最後の入力記号を読み終った後で，$\epsilon$-動作を行う場合，補題 3.4より，入力が受理されるならば，最終状態に到達した後で，DPDA が$\epsilon$-動作を行うことはない．このような DPDA の動作に対応して，SRN では，SRN に入力が全て与えられた後，出力ゲートが 1 を出力するならば，SRN は入力を受理するものとする．

いま，DPDA $M$が受理する言語を $L(M)$ とし，拡張 SRN $\mathcal{E}$が受理する言語を $L(\mathcal{E})$ する．$L(M) = L(\mathcal{E})$ であるならば，DPDA $M$と SRN $\mathcal{E}$は**等価**であるという．

## 4 単純回帰ネットワークの構成法

いま，DPDA $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ が与えられたとする．このとき，$M$を模倣する SRN $\mathcal{E}$を次のように構成する．

1. 入力ユニットに属するゲートの個数を $|\Sigma|$ とする．入力ユニットの各ゲートは，$M$の各入力記号に対応するものとする．

2. 出力ユニットを $O = \{o\}$ とする．

3. 文脈ユニット $C'$に属するゲートの個数を $|Q| \cdot (|\Sigma|+1) \cdot |\Gamma|$ とする．SRN の定義より，隠れユニット $H'$に属するゲートの個数も $|Q| \cdot (|\Sigma|+1) \cdot |\Gamma|$ となる．ここで，文脈ユニット $C'$，および隠れユニット $H'$の各ゲートは，$M$の状態，入力記号，スタック記号の 3 項組に対応するものと考える．すなわち，$M$の状態 $q$，入力記号 $a$，スタック記号 $X$の 3 項組 $(q, a, X)$ に対し，$(q, a, X)$ に対応ゲートが必ず 1 つ存在するものとする．さらに，全ての状態とスタック記号に対して，3 項組 $(q, \epsilon, X)$ に対応するゲートも存在するものとする．$C'$および $H'$に属するゲートの個数を $|\Sigma| \cdot (|\Gamma|+1) \cdot |Q|$ としたことから，各 3 項組と各ゲートの間には一対一対応が存在する．$C'$および $H'$に属する各ゲートは，各ゲートに対応する 3 項組の第 1 項の要素である状態を表現するゲートと考える．

4. 文脈ユニット $C_i, i \geq 1$，および隠れユニット $H_i, i \geq 1$ それぞれに属する定数個のゲートを，以下のように 3 つの部分に分けて考える．3 つの部分は，それぞれ，**1.** スタックを push する場合，**2.** トップ記号を変更しない場合，**3.** スタックのトップ記号を pop する場合の 3 つの各場合に対応している．$C_i$および $H_i$の 3 つの部分をそれぞれ

$$C_i^1, C_i^2, C_i^3 \subseteq C_i, \quad H_i^1, H_i^2, H_i^3 \subseteq H_i$$

とし，$C_i^1, C_i^2, C_i^3$は共通部分を持たないものとし，$H_i^1, H_i^2, H_i^3$も共通部分を持たないものとする．各部分は，$M$の状態 $q$，入力記号 $a$，2 つのスタック記号 $X, Y$からなる 4 項組 $(q, \epsilon, X, Y)$ に対応するゲートから構成され，第 4 項の要素であるスタック記号 $Y$を表現するゲートと考える．いま，$|\Gamma| = \gamma$とし，$\Gamma = \{X_1, X_2, \ldots, X_\gamma\}, X_1 = Z_0$とする．構成は，$M$の遷移関数に対して，以下のように行われる．$x \in \Sigma$ または $x = \epsilon$とする．

   (a) **$\delta(q, x, X) = (q', YX)$ が $M$の遷移関数である場合．** すなわち，スタック記号を 1 つ push する場合である．このとき，スタックにトップ記号をのせる場合に対応する部分 $C_i^1$および $H_i^1$に，$\gamma$個の 4 項組 $(q, x, X, X_1), (q, x, X, X_2), \ldots, (q, x, X, X_\gamma)$ に対応する$\gamma$個のゲートを加える．

   (b) **$\delta(q, x, X) = (q', X)$ が $M$の遷移関数である場合．** すなわち，トップ記号を変更しない場合である．このとき，トップ記号を変更しない場合に対応する部分 $C_i^2$および $H_i^2$に，$\gamma$個の 4 項組 $(q, x, X, X_1), (q, x, X, X_2), \ldots, (q, x, X, X_\gamma)$ に対応する$\gamma$個のゲートを加える．

   (c) **$\delta(q, x, X) = (q', \epsilon)$ が $M$の遷移関数である場合．** すなわち，スタックのトップ記号を pop する場合である．このとき，スタックのトップ記号を消す場合に対応する部分 $C_i^3$および $H_i^3$に，$\gamma$個の 4 項組 $(q, x, X, X_1), (q, x, X, X_2), \ldots, (q, x, X, X_\gamma)$ に対応する$\gamma$個のゲートを加える．

5. 以上のようにして構成された各ユニット間において，辺の重みをつぎのように決定する．

   (a) **入力ユニット $I$から隠れユニット $H'$への辺の重み．** $M$の任意の入力記号 $a$ に対して，$a$ に対応する入力ユニットに属するゲートを $x_a$とする．このとき，隠れユニッ

ト $H'$に属する各ゲートに対し，各ゲートが対応する $M$の状態，入力記号，スタック記号の 3 項組において，入力記号が $a$ に対応するならば，$x_a$からそのゲートへの辺の重みは1．そうでないならば，0 とする．

(b) **入力ユニット $I$から隠れユニット $H_i, i \geq 1$ への辺の重み．** $M$の任意の入力記号 $a$ に対応する入力ユニットに属するゲートを $x_a$とする．このとき，隠れユニット $H_i$ に属する各ゲートに対し，各ゲートが対応する 4 項組において，入力記号が $a$ に対応するならば，$x_a$からそのゲートへの辺の重みを 1，そうではないゲートへの辺の重みを 0 とする．

(c) **文脈ユニット $C'$から隠れユニット $H'$への辺の重み．** いま，$X, Y \in \Gamma$とし，$\xi$ を $\epsilon, X, XY$ のいずれかであるとする．

i. **$\delta(q, a, X) = (q', \xi), a \in \Sigma$が $M$の遷移関数である場合．** 文脈ユニット $C'$ の各ゲートに対し，各ゲートと対応する 3 項組において，状態が $q$である全てのゲートから，3 項組 $(q'a, X)$ と対応する隠れユニット $H'$のゲートへの辺の重みを 1 とする．

ii. **$\delta(q, \epsilon, X) = (q', \xi)$ が $M$の遷移関数である場合．** 文脈ユニット $C'$の各ゲートに対し，各ゲートと対応する 3 項組において，状態が $q$である全てのゲートから，3 項組 $(q'\epsilon, X)$ と対応する隠れユニット $H'$のゲートへの辺の重みを 2 とする．

iii. その他の $C'$のゲートから $H'$へのゲートの辺の重みを 0 とする．

(d) **文脈ユニット $C_1$から隠れユニット $H'$への辺の重み．** すでに，文脈ユニット $C_1$ に属するゲートは，$M$の状態，入力記号，2 つのスタック記号からなる 4 項組 $(q, a, X, Y)$ に対応することを述べた．$M$ のスタック記号$Y$に対して，文脈ユニット $C_1$に属し，各ゲートが対応する 4 項組の第 4 項の要素が$Y$に対応する任意のゲートを $c_Y$とする．このとき，隠れユニット $H'$に属する各ゲートに対し，各ゲートに対応する $M$の状態，入力記号，スタック記号の 3 項組において，スタック記号が $Y$に対応するならば，$c_Y$からそのゲートへの辺の重みは 1．そうでないならば，0 とする．

(e) **文脈ユニット $C_1$から隠れユニット $H_1$への辺の重み．** $x \in \Sigma$または $x = \epsilon$とする．

i. **$\delta(q, x, X) = (q', YX)$ が $M$の遷移関数である場合．** このとき，$C_1$に属するゲートの内，各ゲートが対応する 4 項組の第 4 項の要素が $X$である全てのゲートから，$H_1^1$に属するゲートの内，対応する 4 項組の第 1 項の要素が $q$，第 2 項の要素が $x$，第 3 項の要素が $X$，第 4 項の要素が $Y$であるゲートへの辺の重みを 2 とする．そうではないゲートへの辺の重みを 0 とする．

ii. **$\delta(q, x, X) = (q', X)$ が $M$の遷移関数である場合．** このとき，$C_1$に属するゲートの内，各ゲートが対応する 4 項組の第 4 項の要素が $X$である全てのゲートから，$H_1^2$に属するゲートの内，対応する 4 項組の第 1 項の要素が $q$，第 2 項の要素が $x$，第 3 項の要素が $X$，第 4 項の要素が $X$であるゲートへの辺の重みを 2 とする．そうではないゲートへの辺の重みを 0 とする．

iii. **$\delta(q, x, X) = (q', \epsilon)$ が $M$の遷移関数である場合．** このとき，$C_1$に属するゲートの内，各ゲートが対応する 4 項組の第 4 項の要素が $X$である全てのゲートから，$H_1^3$に属するゲートの内，対応する 4 項組の第 1 項の要素が $q$，第 2 項の要素が $x$，第 3 項の要素が $X$である全てのゲートへの辺の重みを 1 とする．そうではないゲートへの辺の重みを 0 とする．

(f) **文脈ユニット $C_1$から隠れユニット $H_2$への辺の重み．** $C_1$に属するゲートの内，各ゲートが対応する 4 項組の第 4 項の要素が $X$である全てのゲートから，$H_2$に属するゲートの内，対応する 4 項組の第 3 項の要素が $X$である全てのゲートへの辺の重みを 1 とする．更に，$H_2^1$に属するゲートの内，対応する 4 項組の第 4 項の要素が $X$である全てのゲートへの辺の重みを 1 とする．このとき，$H_2^1$に属するゲートの内，対応する 4 項組の第 3 項の要素が $X$，かつ，第 4 項の要素が $X$である全てのゲートへの辺の重みは 2 とする．そうではないゲートへの辺の重みを 0 とする．

(g) **文脈ユニット $C_1$から隠れユニット $H_i, i \geq 3$ への辺の重み．** $C_1$に属する各ゲートが対応する 4 項組において，第 4 項の要素がスタック記号が $X$に対応するゲートから，$H_i$に属する各ゲートが対応する 4 項組において，第 3 項が $X$に対応するゲートへの辺の重みを 1 とする．そうではないゲートへの辺の重みを 0 とする．

(h) **文脈ユニット $C'$から隠れユニット $H_i, i \geq 1$ への辺の重み．** $M$の任意の状態 $q$について場合分けする．

i. $\delta(q,a,X), a \in \Sigma$**が** $M$**の遷移関数に含まれる場合**．$C'$に属する各ゲートが対応する3項組において，状態が$q$に対応するゲートから，$H_i$に属する各ゲートが対応する4項組において，状態が$q$に対応するゲートへの辺の重みを1とする．そうではないゲートへの辺の重みを0とする．

ii. $\delta(q,\epsilon,X)$ **が** $M$**の遷移関数に含まれる場合**．$C'$に属する各ゲートが対応する3項組において，状態が$q$に対応するゲートから，$H_i$に属する各ゲートが対応する4項組において，状態が$q$に対応するゲートへの辺の重みを2とする．そうではないゲートへの辺の重みを0とする．

(i) **文脈ユニット** $C_i, i \geq 2$ **から隠れユニット間** $H_j$**への辺の重み**．定義より，$C_i$に属するゲートから出る辺は，$H_{i-1}, H_i, H_{i+1}$のいずれかに属するゲートへの辺のみ許されている．文脈ユニット$H_i$から隠れユニット$H_j$への辺の重みは，次のように構成される．いま，文脈ユニット$C_i$のゲート$c_X$は，対応する4項組の第4項がスタック記号$X$に対応している任意のゲートとする．

i. **文脈ユニット** $C_i$**から隠れユニット** $H_{i-1}$**への辺の重み**．$c_X$から，$H^3_{i-1}$に属し，対応する4項組の第4項がスタック記号$X$に対応しているゲートへの辺の重みを1とする．そうではないゲートへの辺の重みを0とする．

ii. **文脈ユニット** $C_i$**から隠れユニット** $H_i$**への辺の重み**．$c_X$から，$H^2_i$に属し，対応する4項組の第4項がスタック記号$X$に対応しているゲートへの辺の重みを1とする．そうではないゲートへの辺の重みを0とする．

iii. **文脈ユニット** $C_i$**から隠れユニット** $H_{i+1}$**への辺の重み**．$c_X$から，$H^3_{i+1}$に属し，対応する4項組の第4項がスタック記号$X$に対応しているゲートへの辺の重みを1とする．そうではないゲートへの辺の重みを0とする．

(j) **隠れユニット** $H'$**から出力ユニット** $O$**への辺の重み**．構成より，出力ユニット$O$に属するゲートは1つであり，そのゲートを$o$とした．隠れユニット$H'$に属する各ゲートに対して，各ゲートが対応する4項組の第1項が状態$q \in F$に対応しているならば，それらのゲートから$o$への辺の重みを1とする．それ以外のゲートからの辺の重みを0とする．

6. 各ゲートの閾値をつぎのように決定する．

(a) **文脈ユニット** $H'$**に属するゲートの閾値**：各ゲートの閾値を3とする．

(b) **文脈ユニット** $H_i, i \geq 1$ **に属するゲートの閾値**：各ゲートの閾値を4とする．

(c) **出力ユニット** $O$**に属するゲートの閾値**：ゲート$o$の閾値を1とする．

以上のように構成されたSRNに関して，任意の時刻に対し，隠れユニット$H', H_1, H_2, \ldots$，文脈ユニット$C', C_1, C_2, \ldots$の各々に属するゲートの内，高々1つのゲートが出力1であることに注意する．

以上の構成法より，以下の定理を得る．

**定理 4.1** 任意の*DPDA* $M$に対し，$M$と等価な拡張*SRN* $\mathcal{E}$が存在する．

## 5 まとめ

今後の課題として，DPDAを模倣可能である拡張SRNが，DPDAの計算能力を越えないことを示す必要がある．また，拡張SRNが自然言語を獲得する計算モデルとして妥当性をもつことが最も重要な課題であった．そのために，再帰的遷移ネットワークおよび，拡張遷移ネットワークを模倣する拡張SRNの構成方法を示す必要がある．

## 参考文献


[1] Axel Cleeremans, David Servan-Schreiber, and James L. McClelland. Finite state automata and simple recurrent networks. *Neural Computation*, Vol. 1, No. 3, pp. 372–381, 1989.

[2] Jeffrey L. Elman. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, Vol. 7, pp. 195–225, 1991.

[3] Jeffrey L. Elman. Learning and development in neural networks: the importance of starting small. *Cognition*, Vol. 48, pp. 71–99, 1993.

[4] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to Automata Theory, Language and Computation.* Addison-Wesley, 1979.

[5] Ian Parberry. *Circuit Copmlexity and Neural Networks.* The MIT Press, 1994.

[6] Have T. Siegelmann. *Neural Networks and Analog Computation:Beyound the Turing limits.* Birkhäuser, 1999.

[7] Have T. Siegelmann and Eduardo D. Sontag. On the computational power of neural nets. In *Proceedings of the Fifth ACM Workshop on Computational Learning Theory*, July 1992.

[8] 橋田浩一, 大津由紀雄, 今西典子, Yosef Grodzinsky, 萩原裕子, 錦見美貴子. 言語の獲得と喪失, 岩波講座 言語の科学, 第10巻, 第1,4章. 岩波書店, 1999.

[9] 守谷純之介, 西野哲朗. 単純回帰ネットワークの計算能力について. 京都大学数理解析研究所講究録「計算モデルとアルゴリズム」, Vol. 1093, pp. 188–193, 1999.

[10] 守谷純之介, 西野哲朗. 単純回帰ネットワークを模倣するmealy機械の構成法. 電子情報通信学会コンピュテーション研究会資料, COMP98-27, pp. 51–58, 1998.

[11] 大津由紀雄, 坂本勉, 乾敏郎, 西光義弘, 岡田伸夫. 言語科学と関連領域, 岩波講座 言語の科学, 第11巻, 第2章. 岩波書店, 1998.