数理解析研究所講究録 1205 巻 2001 年 194-199

Turing 機械と等価な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
(Graduate School of Electro-Communications, University of Electro-Communications)
e-mail: {jmoriya,nishino}@ice.uec.ac.jp

1 はじめに

言語学および認知心理学の分野において，人間の自然言語処理，および言語獲得をモデル化するために，J. L. Elman は単純回帰ネットワーク (Simple Recurrent Network; SRN) を提案した．Cleeremans らは，特定の正則文法より生成された記号列における長距離依存関係を，SRN が学習可能であることを実験的に示した [1]．さらに，Elman は，特定の単文や複文における各単語が提示された後，各単語の次の単語を予想するように，SRN をトレーニング可能であることを示した [2, 3]．また，従来提案された他の幾つかのモデルと SRN を，自然言語の学習に関して実験的に比較した結果も示されている [4]．

SRN は有限個のゲートからなるリカレントニューラルネットであり，各ゲートは特定の発火関数を計算する．Siegelmann と Sontag は，部分線形関数を発火関数とするリカレントニューラルネットが，(1) 重みが有理数の場合，Turing 機械と等価であり [5, 7]，(2) 重みが実数の場合，非一様な論理回路と等価であることを示した [6, 7]．部分線形関数を発火関数とする SRN の重みが実数の場合，(2) の結果より，SRN は非一様な論理回路を模倣することが可能になるため，帰納的でない言語を認識することが可能になる．本論では，最初に，発火関数が部分線形関数であり，かつ，重みが実数の場合，帰納的でない言語を認識する SRN が存在することを示す．この結果は既に知られているが，本論では，非一様な論理回路の模倣を含まない，より簡潔な証明を与える．

Siegelmann らの結果を SRN に適用する場合，SRN の辺の重みが実数であることから，SRN の各ゲートの出力は無限の精度を持つことが必要になる．無限精度の出力が可能なゲートを物理的に実現できないので，このようなゲートの存在を仮定することは非現実的である．一方，Kremer は，SRN の各ゲートの出力が固定された精度の場合，SRN が Moore 機械と等価になることを示した [8]．さらに，我々は，SRN の発火関数の値域を有限集合であると仮定した場合に，与えられた SRN と等価な Mealy 機械，および，与えらた Moore 機械と等価な SRN の構成方法を示した [9]．

本論では，SRN の各ゲートが有限集合 $\{0,1\}$ を値域とする閾値関数を計算するという仮定のもとで，十分な計算能力をもつように，SRN を拡張する方法を示す．すなわち，閾値関数を発火関数とする，Turing 機械と等価となる一様なユニット数限定 SRN 族を定義し，その構成法を示す．

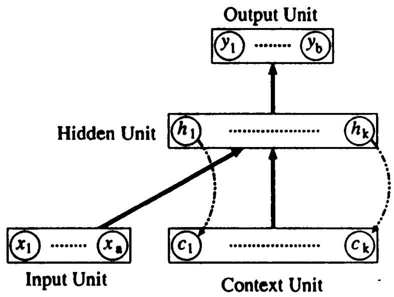

図 1: 単純回帰ネットワーク

2 単純回帰ネットワーク

単純回帰ネットワーク (SRN) とは，入力ユニット，出力ユニット，隠れユニット，文脈ユニットの 4 つのユニットから成る回路であり，各ユニットはゲートの集合である (図 1 参照)．入力ユニットに属するゲートの出力は外部より与えられ，隠れユニットの出力は入力ユニットと文脈ユニットの出力，出力ユニットの出力は隠れユニットの出力にのみ依存する．SRN の主な特徴は，単位時刻前の隠れユニットの出力のコピーを，文脈ユニットが保持できる点にある．SRN 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，以下を満たす 8-項組 $\mathcal{E} = (G, I, O, H, C, A, w, T)$ である．

(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** $I = \{x_1, x_2, \ldots, x_a\} \subseteq V$，**出力ユニット** $O = \{y_1, y_2, \ldots, y_b\} \subseteq V$，**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) $A = (A_1, \ldots, A_k) \in \mathbf{R}^k$ は文脈ユニットの初期出力である．

(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．特に，辺 $(h_i, c_i), h_i \in H, c_i \in C, 1 \le i \le k$ の重みは 1 とする．

(4) $T : V \to \mathbf{R}$ はゲートへの閾値の割り当てとする．ゲート g の閾値を $T(g)$ と表す．

SRN における各ゲートの計算を次のように定義する．

定義 2.2 SRN における任意のゲート gに対して，時刻 $t \in \mathbf{N}$ におけるゲート gの**出力** $g(t)$ を以下のように定義する．

1. $g \in H \cup O$の場合．gの入次数を m とし，gに入る辺をそれぞれ $(g_1, g), \ldots, (g_m, g)$ とする．ゲートの**発火関数**を f とすれば，
$$g(t) = f\left(\sum_{i=1}^{m} w(g_i, g) \cdot g_i(t-1) - T(g)\right).$$
2. $g \in C$の場合．$g = c_i, 1 \le i \le k$ならば，時刻 t でのゲート c_iの出力 $c_i(t)$ は，$t = 0$ ならば，初期出力 A より，$c_i(0) = A_i$，また，$t \neq 0$ ならば，
$$c_i(t) = h_i(t-1)$$
である．ただし，$h_i \in H$は c_iに対応する隠れユニット Hのゲートである．
3. $g \in I$の場合．時刻 $t \in \mathbf{N}$ におけるゲート $g \in I$ の出力 $g(t) \in \{0,1\}$ は，外部より与えられるものとする．

発火関数に関しては，次節において説明する．ここで，入力ユニットに属する $a(=|I|)$ 個のゲートに対して，外部より与えられる時刻 t のゲート $x_i \in I$の出力の列 $(x_1(t), x_2(t), \ldots, x_a(t)) \in \{0,1\}^a$ を，SRN への時刻 t での**入力**とする．同様に，出力ユニットに属する $b(=|O|)$ 個のゲートに対して，時刻 t の各ゲートの出力の列 $(y_1(t), y_2(t), \ldots, y_b(t))$ を，SRN の時刻 t での**出力**とする．

SRN による言語の認識を定義する．以下では，SRN を言語の認識機と考える．そのために，以後，SRN の入力ユニットと出力ユニットに属するゲートの個数はそれぞれ 2 個であるものとし，$I = \{x_1, x_2\}, O = \{y_1, y_2\}$ とする．以下，$\{0,1\}$ 上の言語 L を考える．任意の長さ n の語 $w = w_0 w_2 \cdots w_{n-1} \in \{0,1\}^n, w_i \in \{0,1\}$ に対して，w が SRN へ入力されるとは，SRN への入力が以下を満たすことである．$t' \in \mathbf{N}, 0 \le t' \le n-1$ に対して，$x_1(2t') = w_{t'}$かつ $x_2(2t') = 1$ であり，$t' \ge n$ に対して，$x_1(2t') = 0$ かつ $x_2(2t') = 0$. 更に，奇数の t に対して，$x_1(t) = x_2(t) = 0$ とする．

定義 2.3 SRN $\mathcal{E}$が言語 L を**認識する**とは，任意の $w \in \{0,1\}^*$が $\mathcal{E}$へ入力されたとき，(1) $w \in L$ ならば，ある $t_A \in \mathbf{N}$ が存在し，$t = t_A$ならば $y_1(t) = 1$ かつ $y_2(t) = 1$ であり，$t \neq t_A$ならば $y_1(t) = 0$ かつ $y_2(t) = 0$ であり，(2) $w \notin L$ ならば，ある $t_R \in \mathbf{N}$ が存在し，$t = t_R$ならば $y_1(t) = 0$ かつ $y_2(t) = 1$ であり，$t \neq t_R$ならば $y_1(t) = 0$ かつ $y_2(t) = 0$ を満たすことである．

3 帰納的でない言語を認識する SRN

本節では，発火関数の違い，および発火関数に対する制限により生じる SRN の計算能力の違いをみる．

Siegelmann らの結果より，SRN の発火関数が**部分線形関数** (piecewise linear function):

$$f_p(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } x > 1 \end{cases}$$

である場合，つぎのことが導かれる：(1) SRN の重みが**有理数**の場合，SRN は **Turing 機械**と等価である [5, 7]. (2) SRN の重みが**実数**の場合，SRN は**非一様な論理回路**と等価である [6, 7]. SRN の重みが実数の場合，(2) より，SRN は非一様な論理回路を模倣することが可能になるため，SRN は帰納的でない言語を認識することが可能になる．本論では，SRN の発火関数が部分線形関数であり，かつ，重みが実数の場合，非一様な論理回路を模倣することなく，帰納的でない言語を認識する SRN が存在することを示す．

定理 3.1 発火関数が部分線形関数であり，かつ，重みが実数である *SRN* $\mathcal{E}$が存在し，$\mathcal{E}$は帰納的でない言語 L_Uを認識する．

証明の概要：L を任意の帰納的でない言語とする．L に対して，言語 L_Uを次のように定義する．

$$L_U = \{1^n \mid n \text{ の 2 進数表現は } L \text{ の元}\}$$

L_Uも帰納的でない言語である．いま，L_Uに対して，つぎの条件を満たす実数 $r_{L_U} \in \mathbf{R}, 0 \le r_{L_U} \le 1$ を考える：

$$1^n \in L_U \Rightarrow r_{L_U} \text{の小数点以下第 } n \text{ 桁が } 3.$$
$$1^n \notin L_U \Rightarrow r_{L_U} \text{の小数点以下第 } n \text{ 桁が } 1.$$

例として，$L_U = \{1, 111, 1111, \ldots\}$ ならば，$r_{L_U} = 0.3133\cdots$ となる．

L_uを受理する SRN $\mathcal{E}$は実数 r_{L_U} を辺の重みとしてもつ．そして，$\mathcal{E}$は次のように動作する．長さ n の入力に対して，

(1) 入力が 1^nであるか否かを判定する．すなわち，任意の $t \in \mathbf{N}, 0 \le 2t \le 2n$ において，$x_1(2t) = 1, x_2(2t) = 1$ であるか否かを判定する．入力が 1^nでないならば，$y_1 = 0, y_2 = 1$ となる．

(2) 入力より，値 $1/10 + \cdots + 1/10^n$を計算する．値 $1/10 + \cdots + 1/10^n$は，入力の長さ n を保存するためのカウンタの値である．

(3) カウンタの値 $1/10 + \cdots + 1/10^n$を利用し，r_{L_U} の小数点以下第 n 桁 r_nが 3 であるか否かを判定する．$r_n = 3$ ならば，$y_1 = 1, y_2 = 1$ となり，$r_n = 1$ ならば，$y_1 = 0, y_2 = 1$ となる．□

SRN の発火関数が部分線形関数であり，重みが実数である場合，SRN は帰納的でない言語を認識することから，SRN は計算論的観点から妥当な計算モデルとはいえない．

次節において，定義域が有限集合 $\{0,1\}$ である閾値関数を発火関数とし，計算能力が Turing 機械と等価となる SRN の拡張を考える．

4 一様なユニット数限定 SRN 族

本節より，SRN の各ゲートの発火関数を閾値関数 (threshold function):

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

とする．よって，各ゲートの出力は $\{0,1\}$ の要素であり，SRN の各ゲートが計算する発火関数の値域は有限集合である．また，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，[10] 参照)．よって，以後 SRN の辺の重みとゲートの閾値を整数とする．本論では，SRN を拡張し，新たに**ユニット数限定単純回帰ネットワーク**を以下のように定義する（図 2参照）．

定義 4.1 $s \in \mathbf{N}$ とする．s **ユニット限定単純回帰ネットワーク**（s ユニット限定 *SRN*）とは，以下を満たす 8-項組 $\mathcal{E}(s) = (G, I, O, H, C, A, w, T)$ である．

(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** $I = \{x_1, x_2\} \subseteq V$，**出力ユニット** $O = \{y_1, y_2\} \subseteq V$，**隠れユニット** $H = \{h_1, h_2, \ldots, h_{6s+k}\} \subseteq V$，および**文脈ユニット** $C = \{c_1, c_2, \ldots, c_{6s+k}\} \subseteq V$．各ユニットは共通部分をもたないものとする．隠れユニットと文脈ユニットはそれぞれ**部分隠れユニット** $H_0, H_1, H_{-1}, \ldots, H_{\lceil s/2 \rceil}, H_{-\lfloor s/2 \rfloor}$ と**部分文脈ユニット** $C_0, C_1, C_{-1}, \ldots, C_{\lceil s/2 \rceil}, C_{-\lfloor s/2 \rfloor}$ に分割される．ただし，各 C_i，および各 H_iはそれぞれ共通部分をもたないものする．定数 $k \in \mathbf{N}$ は $|H_0| = |C_0| = k$, $k \geq 4$. を満たすものとする．また，

$$|H_1| = |C_1| = |H_{-1}| = |C_{-1}| = \cdots = 6,$$

とする．C_0, H_0には，それぞれに 4 個の特別なゲートが含まれ，それらを $c^{\mathrm{I}}, c^{\mathrm{II}}, c^A, c^B \in C_0$, $h^{\mathrm{I}}, h^{\mathrm{II}}, h^A, h^B \in H_0$ とする．また，任意の $i \in \mathbf{Z}, i \neq 0, -\lfloor s/2 \rfloor \leq i \leq \lceil s/2 \rceil$ に対して，C_i, H_iに属する 6 個のゲートを

$$C_i = \{c_i^{A,1}, c_i^{A,2}, c_i^{A,3}, c_i^{B,1}, c_i^{B,2}, c_i^{B,3}\}$$

$$H_i = \{h_i^{A,1}, h_i^{A,2}, h_i^{A,3}, h_i^{B,1}, h_i^{B,2}, h_i^{B,3}\}$$

とする．

$$C_i^A = \{c_i^{A,1}, c_i^{A,2}, c_i^{A,3}\}, C_i^B = \{c_i^{B,1}, c_i^{B,2}, c_i^{B,3}\}$$

とし，H_i^Aと H_i^Bも同様に定義する．

Eは G の辺の集合であり，$E = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6$と分割される．

$$E_1 = \{(v_1, v_2) \mid (v_1, v_2) \in I \times H_0 \cup H_0 \times O \ \cup\ H \times C \ \cup\ C_0 \times H_0\}$$

とし，隠れユニット H から文脈ユニット C 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．$E_2$は C_1, C_{-1}から H_0への辺であり，

$$E_2 = \{(v_1, v_2) \mid (v_1, v_2) \in C_1 \times H_0 \ \cup\ C_{-1} \times H_0\}$$

とする．E_3は C_0から H_1, H_{-1}への辺であり，

$$E_3 = \{(v_1, v_2) \mid (v_1, v_2) \in \{c^j\} \times \{h_1^{j,1}\} \cup \{c^j\} \times \{h_{-1}^{j,3}\},\ j \in \{A, B\}\}$$

とする．E_4は，$i \in \mathbf{Z}, i \neq 0, \lfloor s/2 \rfloor \leq i \leq \lceil s/2 \rceil$ に対して，C_0から H_iへの辺からなり，

$$E_4 = \{(v_1, v_2) \mid (v_1, v_2) \in \{c^{\mathrm{I}}, c^{\mathrm{II}}\} \times H_i\}$$

とする．E_5は，任意の $i \in \mathbf{Z}, i \neq 0, 1, -1, -\lfloor s/2 \rfloor + 1 \leq i \leq \lceil s/2 \rceil - 1$ に対して，C_{i-1}, C_i, C_{i+1}から H_iへの辺であり，

$$E_5 = \{(v_1, v_2) \mid (v_1, v_2) \in C_{i-1}^j \times \{h_i^{j,1}\} \cup C_i^j \times \{h_i^{j,2}\} \ \cup\ C_{i+1}^j \times \{h_i^{j,3}\},\ j \in \{A, B\}\}$$

とする．E_6は，$i' = \lceil s/2 \rceil, -1$ と $i'' = -\lfloor s/2 \rfloor, 1$ に対して，

$$E_6 = \{(v_1, v_2) \mid (v_1, v_2) \in C_{i'-1}^j \times \{h_{i'}^{j,3}\} \cup C_{i'}^j \times \{h_{i'}^{j,2}\} \ \cup\ C_{i''}^j \times \{h_{i''}^{j,2}\} \ \cup\ C_{i''+1}^j \times \{h_{i''}^{j,1}\},\ j \in \{A, B\}\}$$

とする．

(2) $A = (A_1, \ldots, A_k) \in \{0,1\}^k$ は初期出力．

(3) $w : E \to \mathbf{Z}$ 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．また，次の辺の重みは固定される．

- 任意の i に対して，辺 $(h_i, c_i), h_i \in H, c_i \in C$の重みを 1 とする．
- E_3, E_5, E_6に属する全ての辺の重みを 1 とする．
- E_4に属する辺の重みは，任意の i と $j \in \{A, B\}$ に対して，$w(c^{\mathrm{I}}, h_i^{j,1}) = 1$, $w(c^{\mathrm{I}}, h_i^{j,2}) = 0$, $w(c^{\mathrm{I}}, h_i^{j,3}) = -1$, $w(c^{\mathrm{II}}, h_i^{j,1}) = 1$, $w(c^{\mathrm{II}}, h_i^{j,2}) = -1$, $w(c^{\mathrm{II}}, h_i^{j,3}) = 1$ とする．

(4) $T : V \to \mathbf{Z}$ はゲートへの閾値の割り当てとする．ゲート vの閾値を $T(v)$ と表す．また，次のゲートの閾値は固定される．文脈ユニットに属する任意のゲートの閾値は 1 とする．任意の i と $j \in \{A, B\}$ に対して，$T(h_i^{j,1}) = 3$, $T(h_i^{j,2}) = 1$, $T(h_i^{j,3}) = 2$ とする．

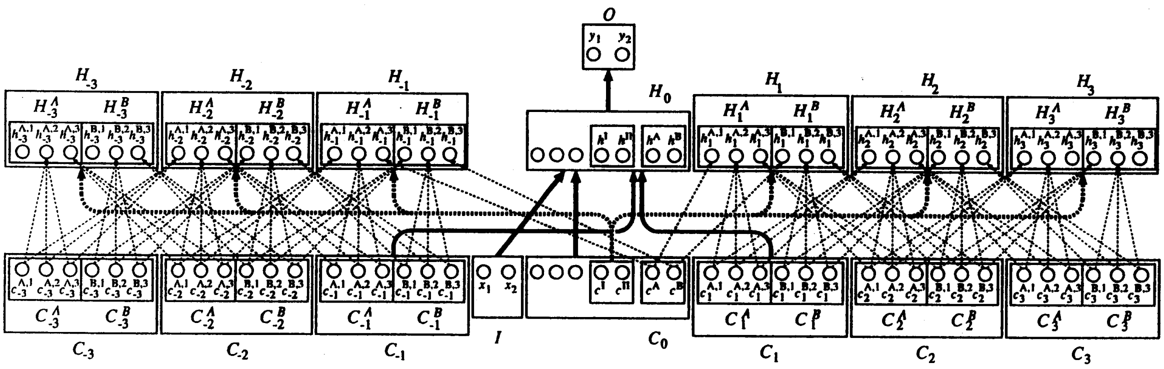

図 2: 6ユニット限定単純回帰ネットワーク

sユニット限定SRNにおける各ゲートの出力は，定義2.2に従うものとする．ただし，$C_i, i \neq 0$に属する任意のゲートcの時刻$t=0$での出力は，全て0であるものとする．

ここで，sユニット限定SRNの**符号化**を考える．sユニット限定SRNの符号化は，通常の重み付き有向グラフと同様に，符号化するものとする．任意のsユニット限定SRN $\mathcal{E}(s)$に対して，定義より，つぎのパラメータを定めることにより，$\mathcal{E}(s)$の符号化を定義することができる．(A) C_0とH_0に属するゲートの個数，(B) E_1とE_2の辺の重み，(C) H_0とOに属するゲートの閾値，(D) 初期出力$A \in \{0,1\}^k$，(E) sの値．C_0, H_0, I, Oに属するゲートの個数は定数個であることから，(A)から(D)のパラメータは，定数長さの記号列で符号化できることに注意する．以上より，$\mathcal{E}(s)$の符号化は，(A)から(D)のパラメータの符号化とsのtally表現により，実現することが可能になる．つぎに，**一様なsユニット限定SRN族**を定義する．

定義 4.2 いま，関数$s : \mathbf{N} \to \mathbf{N}$に対し，$E = \{\mathcal{E}(s(n)) \mid n \in \mathbf{N}\}$を$s(n)$ユニット限定SRN族とする．$E$が**一様な$s(n)$ユニット限定SRN族**であるとは，入力$1^n$に対して，$s(n)$ユニット限定SRN $\mathcal{E}(s(n))$の符号化を出力するTuring機械が存在することである．

定義より，関数$s(n)$は領域構成可能であることに注意する．つぎに，一様なsユニット限定SRN族による言語の認識を定義する．

定義 4.3 一様な$s(n)$ユニット限定SRN族Eが言語Lを**認識する**とは，長さnの任意の$w \in \{0,1\}^n$が$\mathcal{E}(s(n))$へ入力されたとき，定義2.3に従い，$w \in L$ならば$\mathcal{E}(s(n))$はwを受理し，$w \notin L$ならば$\mathcal{E}(s(n))$はwを拒否することである．

ユニット数限定SRNの性質

ユニット数限定SRNの定義において，部分隠れユニットH_iの各ゲートへの辺とその辺の重み，各ゲートの閾値はそれぞれ定数であった．これら定義により与えられた定数より，部分隠れユニットH_iはつぎの補題を満たす．

補題 4.1 任意の時刻tにおいて，$j \in \{A, B\}$に対し，H_i^jに属する3個のゲートのうち，出力が1であるゲートは高々1個しか存在しない．

部分文脈ユニットC_iと部分隠れユニットH_iは，それぞれ6個のゲートからなる．いま，それら6個のゲートの出力列に対して，長さ2の2進数列を以下のように定義する．

定義 4.4 ある時刻tのH_iの出力列に対して，$\alpha_{H_i}(t) = \alpha_{H_i^A}(t)\alpha_{H_i^B}(t) \in \{0,1\}^2$は，$j \in \{A, B\}$に対し，(1) H_i^jの3つのゲートの中に出力1のゲートが存在するならば，$\alpha_{H_i^j}(t) = 1$, (2) H_i^jの3つのゲートの中に出力1ゲートが存在しないならば，$\alpha_{H_i^j}(t) = 0$. C_iに対しても同様に定義される．

以上のように定義された2進数列を用いることにより，次の補題が示される．

補題 4.2 いま，ある$\alpha_{C_i}(t) \in \{0,1\}^2$に対して，

(1) $c^{\mathrm{I}}(t) = 0$かつ$c^{\mathrm{II}}(t) = 0$，または，$c^{\mathrm{I}}(t) = 1$かつ$c^{\mathrm{II}}(t) = 0 \Rightarrow \alpha_{C_i}(t) = \alpha_{H_i}(t+1)$,

(2) $c^{\mathrm{I}}(t) = 0$かつ$c^{\mathrm{II}}(t) = 1 \Rightarrow \alpha_{C_i}(t) = \alpha_{H_{i-1}}(t+1)$,

(3) $c^{\mathrm{I}}(t) = 1$かつ$c^{\mathrm{II}}(t) = 1 \Rightarrow \alpha_{C_i}(t) = \alpha_{H_{i+1}}(t+1)$.

すなわち，補題4.2は，ある時刻tにおけるα_{C_i}が，時刻$t+1$においてH_{i-1}, H_i, H_{i+1}のいずれかにコピーされることを示している．

5 Turing機械と等価な一様ユニット数限定SRN族の構成法

一様なユニット数限定SRN族はTuring機械により容易に模倣可能であることが解る．これより，Turing機械を模倣する一様なユニット数限定SRN族が存在することを示す．以下では，単にTuring機械と言った場合，決定性片方向無限1テープTuring

機械を指すものとする．模倣を行う Turing 機械を $M=(Q,\Sigma,\Gamma,\delta,q_0,F_A,F_R)$ とし，ここに，(1) Q を状態の有限集合，(2) $\Sigma=\{0,1\}$ を入力アルファベット，(3) $\Gamma=\{0,1,\sharp\}$ をテープアルファベット，$\sharp\in\Sigma-\Gamma$ を空白記号，(4) $q_0\in Q$ を初期状態，(5) $F_A\subset Q$ を受理状態，$F_R\subset Q$ を拒否状態，$F_A\cap F_R=\emptyset$，$F=F_A\cup F_R$，(6) δを $(Q-F)\times\Gamma$から$Q\times\Gamma\times\{L,S,R\}$ への関数とし，遷移関数とする．

いま，Mは $S(n)$ 領域限定であり，関数 $S(n)$ は領域構成可能かつ，$S(n)\geq n$ を満たすものと仮定する．このとき，Mを模倣する一様な $2S(n)$ ユニット限定 SRN 族を構成する．関数 $S(n)$ は領域構成可能であることから，$2S(n)$ ユニット限定 SRN の族は一様である．すでに $2S(n)$ の値が既知であることから，$2S(n)$ ユニット限定SRN を決定するには，(1) C_0と H_0に属するゲートの個数，(2) E_1と E_2の辺の重み，(3) H_0 と Oに属するゲートの閾値，(4) 初期出力 $A\in\{0,1\}^k$ を決定すればよい．以上の各パラメータを Mより決定する．このとき，$2S(n)$ ユニット限定SRN が，つぎのような動作を行うように，各パラメータを決定する．

(1) 入力ユニットより与えられる入力を H_{-n}から H_1 にコピーする．

(2) H_{-n}から H_1の内容を h^A,h^Bと H_1から H_{n-1}にコピーする．

(3) Mの模倣を開始する．

$2S(n)$ ユニット限定 SRN $\mathcal{E}(2S(n))=(G,I,O,H,C,A,w,T)$ を以下のように構成する．ただし，(1) と (2) を実現する I，H_0，C_0に属するゲートの閾値，及びゲート間の重みは，容易に構成可能であることから，以下の構成では言及せず，(3) を実現する構成法のみを示す．

1. 部分文脈ユニット C_0と部分隠れユニット H_0は，それぞれ $2|Q|+|Q|\times|\Gamma|+4$ 個のゲートを含むように構成する．各ゲートは，次のように対応する．$2|Q|$ 個の各ゲートは，Mの状態 $q\in Q$ に対応している．状態 $q\in Q$ に対応する 2 個ゲートを c_q^1,c_q^2とする．また，$|Q|\times|\Gamma|$ 個の各ゲートは，Mの状態とテープ記号の 2 項組 $(q,X)\in Q\times\Gamma$ に対応している．2 項組 $(q,X)\in Q\times\Gamma$に対応するゲートを $c_{(q,X)}$とする．残りの 4 個のゲートは，$c_0^{A,1},c_0^{A,3},c_0^{B,1},c_0^{A,3}$ とする．H_0に対しても同様とする．

2. C_0から H_0への辺の重みを次のように定義する．

   $\delta(q,X)=(p,Y,\beta)$ を Mの遷移関数とする．このとき，C_0に属する状態 qに対応しているゲートの 1 つ c_q^1から H_0に属する 2 項組 (q,X) に対応しているゲート $h_{(q,X)}$ への辺の重みを

   $$w(c_q^1,h_{(q,X)})=1$$

   とする．C_0に属する 2 項組 (q,X) に対応しているゲート $c_{(q,X)}$から H_0に属する状態 p に対応しているゲート h_p^2 への辺の重みを

   $$w(c_{(q,X)},h_p^2)=1$$

   とする．つぎに，

   - $Y=1$ ならば，ゲート $c_{(q,X)}$から h^Aへの辺の重みを 1 とし，h^Bへの辺の重みを-2 とする．
   - $Y=0$ ならば，ゲート $c_{(q,X)}$から h^Aへの辺の重みを-2 とし，h^Bへの辺の重みを-1 とする．
   - $Y=\sharp$ ならば，ゲート $c_{(q,X)}$から h^A,h^Bへの辺の重みをそれぞれ-2 とする．

   更に，

   - $\beta=L$ ならば，ゲート $c_{(q,X)}$から $h^{\mathrm{I}},h^{\mathrm{II}}$への辺の重みをそれぞれ 1 とし，
   - $\beta=R$ならば，ゲート $c_{(q,X)}$から h^{II}への辺の重みを 1 とする．

   つぎに，任意の $q\in Q$，$X\in\Gamma$に対して，$c^A,c_0^{A,1},c_0^{A,3}$から 2 項組 (q,X) の Xが 1 であるすべてのゲート $c_{(q,1)}$への辺の重みを 1 にし，Xが$\sharp$ であるすべてのゲート $c_{(q,\sharp)}$への辺の重みを-2 にする．同様に，$c^B,c_0^{B,1},c_0^{B,3}$から 2 項組 (q,X) の Xが 0 であるすべてのゲート $c_{(q,0)}$への辺の重みを 1 にし，Xが$\sharp$ であるすべてのゲート $c_{(q,\sharp)}$への辺の重みを-2 にする．

   残りの辺の重みは次のようになる．

   - c_q^2から h_q^1への辺の重みを 1 とする．
   - c^A，$c_0^{A,1}$，$c_0^{A,3}$から h^Aへの辺の重みと，c^B，$c^{B,1}$，$c^{B,3}$から h^Bへの辺の重みをそれぞれ 1 とする．
   - c^{I}からの辺の重みは，$h_0^{A,1}$，$h_0^{B,1}$への重みをそれぞれ 1，$h_0^{A,3}$，$h_0^{B,3}$への重みをぞれぞれ -1 とする．
   - c^{II}からの辺の重みは，$h_0^{A,1}$，$h_0^{B,1}$への重みをそれぞれ 1，h^A,h^Bへの重みをぞれぞれ-1，$h_0^{A,3}$，$h_0^{B,3}$への重みをそれぞれ 1 とする．

3. H_0に属するゲートから出力ユニット Oに属する 2 個のゲート y_1と y_2への辺の重みを次のように定義する．H_0に属する状態 $q\in Q$ に対応しているゲート h_qより，$q\in F_A$ならば，h_qから y_1と y_2 への辺の重みをそれぞれ 1,1 とし，$q\in F_R$ならば，h_qから y_2への辺の重みを 1 とする．

4. C_1,C_{-1}に属するゲートから H_0に属するゲートへの辺の重みは，$j\in\{A,B\}$ に対して，C_1^jに属する 3 個のゲートから $h_0^{j,3}$への辺の重みと，C_{-1}^jに属する 3 個のゲートから $h_0^{j,1}$への辺の重みを全て 1 とする．

5. H_0に属するゲートと出力ユニットOに属するゲートの閾値を次のように定義する．

   H_0に属するゲートの閾値は，任意の状態 qに対して，q対応するゲート 2 個のゲート h_q^1, h_q^2の閾値を 1，任意の 2 項組 (q, X) に対して，$X \in \{0,1\}$ ならば (q, X) に対応するゲート $h_{(q,X)}$の閾値を 2，$X = \sharp$ ならば $h_{(q,X)}$の閾値を-1 とする．また，$h^{\mathrm{I}}, h^{\mathrm{II}}, h^A, h^B$の各ゲートの閾値を 1 とし，$h_0^{A,1}, h_0^{A,1}$の閾値を 3，$h_0^{A,3}, h_0^{A,3}$の閾値を 2 とする．

   Oに属する 2 個のゲート y_1と y_2の閾値は，ともに 1 とする．

以上の構成法より得られるユニット数限定 SRN と Mとの対応関係は，次のようになる．

(1) Mのテープ上に書かれた記号は，定義 4.4 において，部分文脈ユニットおよび部分隠れユニットの出力により定義された 2 進数列に対応する．テープ記号 $0, 1, \sharp$ は，それぞれ$\alpha_{C_i} = 10$, $\alpha_{C_i} = 01$, $\alpha_{C_i} = 00$ に対応する．

(2) Mが状態 qでテープ上の記号 Xをよみ，テープ記号を Yに書き換え，現在の状態 qから次の状態 p へ遷移する動作は，次のように模倣される．

C_0に属するゲート $c_0^{A,1}, c^A, c_0^{A,3}, c_0^{B,1}, c^B, c_0^{B,3}$の出力が，$M$のヘッドが見ている記号 Xに対応した出力であるとする．すなわち，$X = 0$ ならば，$c_0^{A,1}, c^A, c_0^{A,3}$, の内 1 個のゲートが出力 1 であり，$c_0^{B,1}, c^B, c_0^{B,3}$ の出力はすべて 0 である．同様に，$X = 1$ ならば，$c_0^{A,1}, c^A, c_0^{A,3}$, の出力はすべて 0 であり，$c_0^{B,1}, c^B, c_0^{B,3}$の内 1 個のゲートが出力 1 である．$X = \sharp$ ならば，$c_0^{A,1}, c^A, c_0^{A,3}$, $c_0^{B,1}, c^B, c_0^{B,3}$の出力は全て 0 である．

いま，C_0に属する状態 qに対応するゲート c_q^1の出力は 1 であるとする．また，他の状態 $q' \in Q, q' \neq q$に対応するゲート $c_{q'}^1$の出力は全て 0 であるとする．これは，Mが状態 qであることに対応する．

このとき，ゲート c_q^1の出力と，$c_0^{A,1}$, c^A, $c_0^{A,3}$, $c_0^{B,1}$, c^B, $c_0^{B,3}$の出力より，2 項組 (q, X) に対応するゲート $h_{(q,X)}$の出力が 1 となる．$h_{(q,X)}$の出力が 1 となることは，Mが状態 qであり，ヘッドが見ている記号が X であることに対応する．ゲート $h_{(q,X)}$の出力がゲート $c_{(q,X)}$にコピーされ，$c_{(q,X)}$の出力により，c^Aと c^Bの出力を Yに対応する出力に変更し，ゲート h_p^2の出力を 1 とする．そして，ゲート c_p^2の出力が 1 であることから，ゲート h_p^1の出力が 1 となる．以上のように，Mのテープ記号の書き換えと状態の遷移は動作を模倣する．

(3) Mのヘッドの動作は，C_iの出力に対応する定義 4.4 で定義した 2 進数列α_{C_i}を，(i) ヘッドが右へ動く場合は H_{i-1}に，(ii) ヘッドが左へ動く場合は H_{i+1}に，(iii) 動かさない場合は H_iにコピーすることに，それぞれ対応している．これは，次のように模倣される．$\delta(q, X) = (p, Y, \beta)$ を Mの遷移関数とする．そして，(2) の対応に従い，2 項組 (q, X) に対応するゲート $C_{(q,X)}$ の出力が 1 であるとする．$\beta = L$ ならば，ゲート $C_{(q,X)}$の出力が 1 であることから，ゲート $h^{\mathrm{I}}, h^{\mathrm{II}}$の出力が 1,1 となり，補題 4.2 より，$\alpha_{C_i}$が H_{i+1}にコピーされる．同様に，$\beta = R$ならば，ゲート $C_{(q,X)}$ の出力が 1 であることから，ゲート $h^{\mathrm{I}}, h^{\mathrm{II}}$の出力が 0,1 となり，$\alpha_{C_i}$が H_{i-1}にコピーされる．$\beta = \sharp$ ならば，ゲート $h^{\mathrm{I}}, h^{\mathrm{II}}$の出力は 0,0 となり，$\alpha_{C_i}$が H_iにコピーされる．以上のように，Mのヘッドの動作は模倣される．

以上のように Mの動作を模倣するユニット数限定 SRN が得られるので，Mが使用する領域量を考慮することにより，以下の定理を得る．

定理 5.1 L を $S(n)$ 領域限定 *Turing* 機械 Mにより認識される言語とする．ただし，関数 $S(n)$ は領域構成可能であり，$S(n) \geq n$ を満たすものとする．このとき，L を認識する一様な $2S(n)$ ユニット限定 *SRN* 族が存在する．

参考文献

[1] A. Cleeremans, D. Servan-Schreiber and J. L. McClelland, "Finite state automata and simple recurrent networks," Neural Computation, vol. 1, no. 3, pp. 372–381, 1989.

[2] J. L. Elman, "Distributed representations, simple recurrent networks, and grammatical structure," Machine Learning, vol. 7, pp. 195–225, 1991.

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

[4] S. Lawrence, C. L. Giles and S. Fong, "Natural language grammatical inference with recurrent neural networks," IEEE Trans. on Knowledge and Data Engineering, vol. 12, no. 1, pp. 126–140, 2000.

[5] H. T. Siegelmann and E. D. Sontag, "On the computational power of neural nets," Journal of Computer and System Science, vol. 50, no. 2, pp. 132–150, 1995.

[6] H. T. Siegelmann and E. D. Sontag, "Analog computation via neural networks," Theoretical Computer Science, vol. 131, pp. 331–360, 1994.

[7] H. T. Siegelmann, "Neural Networks and Analog Computation : Beyound the Turing limits," Birkhäuser, 1999.

[8] S. C. Kremer, "On the computational power of elman-style recurrent networks," IEEE Trans. Neural Networks, vol. 6, no. 4, pp. 1000–1004, 1995.

[9] J. Moriya and T. Nishino, "Relationships between the Computational Capabilities of Simple Recurrent Networks and Finite Automata," IEICE Trans. Fundamentals, (to appear).

[10] I. Parberry, "Circuit Copmlexity and Neural Networks," The MIT Press, 1994.