

Title	再帰型神経回路網による単語クラスタリングに関する研究
Author(s)	兵藤, 大輔
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/341
Rights	
Description	Supervisor:林 幸雄, 知識科学研究科, 修士

修 士 論 文

**再帰型神経回路網による
単語クラスタリングに関する研究**

北陸先端科学技術大学院大学
知識科学研究科知識システム基礎学専攻

兵藤 大輔

2002 年 3 月

修 士 論 文

**再帰型神経回路網による
単語クラスタリングに関する研究**

指導教官 林 幸雄 助教授

北陸先端科学技術大学院大学
知識科学研究科知識システム基礎学専攻

050070 兵藤 大輔

審査委員主査	林 幸雄 助教授
審査委員	中森 義輝 教授
審査委員	佐藤 賢二 助教授
審査委員	藤波 努 助教授

提出年月: 2002 年 2 月

目次

第1章	はじめに	1
1.1	まえがき	1
1.2	本研究の目的	1
1.3	本研究の概要	2
第2章	研究背景	3
2.1	SRN の語系列予測学習	3
2.2	Elman の単語クラスター	4
第3章	実験手法	7
3.1	SRN の単語クラスターの獲得	7
3.1.1	RAAM の固定長表現	7
3.1.2	RAAM での隠れ層活性度パターンのクラスタリング	8
3.2	文脈に依存しない単語クラスターの構築	9
3.3	クラスター分析	10
第4章	実験結果	11
4.1	予備実験	11
4.1.1	実験環境	11
4.1.2	実験結果	11
4.2	RAAM による単語クラスター	14
4.2.1	実験環境	14
4.2.2	実験結果	14
4.3	文脈によらない単語クラスターの構築	17
4.3.1	実験環境	17
4.3.2	実験結果	17

4.3.3 文脈依存を除去した隠れ層の活性化パターン	20
第 5 章 考察	21
第 6 章 まとめ	23
謝辞	24
参考文献	26

第1章 はじめに

1.1 まえがき

子供はどのように言語を獲得するのであろうか。言語学、心理学、認知科学、脳科学、計算機科学など多くの分野で、今もそれぞれの立場から、あるいは多面的にこの疑問を解決しようと多くの研究なされている。その中で、言語獲得過程の解明にコネクショニズムのモデルが重要な働きをすると考える研究者も多い。コネクショニズムモデルの中でも、Elman の SRN (simple recurrent neural network) は良く知られている。Elman は SRN に文を構成する単語を系列的な入力データとして順に与え、単語が与えられるごとに、次に与えられる単語を予測させる (語系列予測課題) ことにより、2、3 語の単純な文章の文法構造を学習させられることを示している [2]。さらに関係節を含む複雑な文章の文法構造を学習したとする報告をしている [3][4]。疑問詞 who による疑問文も学習可能であるとされている [12]。加えて、Elman は階層的な単語クラスターを SRN によって表現している。Elman その単語クラスターは文法 (ここでは名詞や動詞といった統語範疇を指す) や単語の意味 (ここでは人や動物といった意味の範疇を指す) のクラスターを持つと指摘し、生得的な知識なしに SRN がそれらを獲得したと考えられるとしている [2]。しかし、そこで表現された単語クラスターが SRN が学習によって文法や意味を獲得した結果によるものかは疑問がある。単に入力されている単語の連なりの前後関係のみによるものである可能性を否定できないからである。

1.2 本研究の目的

本研究は Elman が計算機上での再帰型神経回路網 (simple recurrent neural network, SRN) によって、文中の次の単語を予測する (語系列予測) 課題を実行した結果、文法や単語の意味に関する生得的な知識なしに文法のような (あるいは意味のような) ものを獲得できたとする報告に対して、その妥当性を評価することを目的としている。

1.3 本研究の概要

本研究では Elman の階層的な単語クラスター [2] の追実験を足掛かりに、Elman が示したような単語クラスターが構築できたからといって SRN が学習によって単語クラスター（あたかも意味のような）を獲得したことにならないという可能性を、RAAM(recursive auot-associative memory) による間接的なアプローチによって指摘する。さらに、SRN から得られる単語クラスターは、単純に入力の時系列としての構文的なもののみを反映しているにすぎないことを示すために、文脈に依存しない層を SRN の入力層と隠れ層の間に追加し、その層の活性化パターンから階層的な単語クラスターが得られるか検証する。

本稿では 2 章で研究背景として特に Elman の単語クラスターを詳しく述べる。3 章では本研究の実験手法を述べ、4、5 章で実験結果を提示し考察する。

第2章 研究背景

この章ではコネクショニストが提案してきた手法を概観し、また 2.1 節からは、特に本研究の主背景である Elman の単語クラスタリングについて述べる。

2.1 SRN の語系列予測学習

SRN は図 2.1 のように隠れ層 (hidden units) から文脈層 (context units) への再帰結合を持つ、きわめて単純な再帰型ネットワークである。文脈層は隠れ層の状態を 1 ステップ保持し、次の入力層からの入力とともに、再び隠れ層に伝達する。このフィードバックにより、ネットワークは過去の隠れ層状態を参照することができ、時系列処理ができる。訓練では教師信号を次の時点の入力とすることにより、次の入力を予測させる。Elman は SRN により 0,1 の列で前 2 ビットの XOR が次のビットとなる XOR 時系列入力に対して次のビットの予測をするタスクを行い、また文を構成する単語を系列的な入力データとして順に与え、単語が与えられるごとに、次に与えられる単語を予測させることにより、2、3 語の単純な文章の文法構造を学習させられることを示している [2]。加えて、関係節を含む複雑な文章の文法構造を学習させ [3]、制限したワーキングメモリを増加させていくことにより、最初から複雑な文を正しく学習できることも示している [4]。玉森は疑問文や疑問詞 who による疑問文も学習可能であるとしている [12]。

SRN によって上述の報告のように学習が成功する理由は次のようだと言える [16]。こどもの言語獲得を考える上で、「このような文は文法的でない」という情報を与えられないにも関わらず、過剰に生成される文法的でない文をそぎ落とすのかという問題 (Baker のパラドックス¹) がある [1]。この問題はネットワークによる学習にもつきまとう問題である。語系列予測という課題は文の中の次の単語を予測する学習途中では誤った予測をする。それを次に入力される単語を見て誤りの修正を行うため、実際にはシステム内部で誤った使い方も学習しているのである。課題の選択が良かったと言えるだろう。

¹否定証拠欠如問題、そぎ落とし問題とも言われる

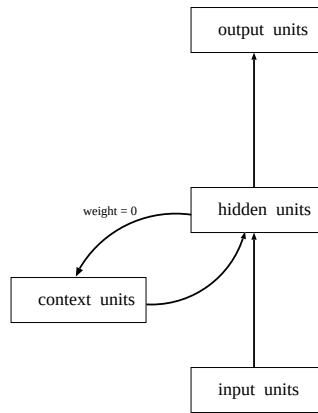


図 2.1: 単純再帰型ネットワーク (SRN)

各層を長方形で表している。context units には 1 ステップ前の hidden units の状態がコピーされ、hidden units は input units と context units からの入力を同時にうける。これにより時系列処理が可能になる。

2.2 Elman の単語クラスター

SRN の隠れ層の活性度パターンは分散表現 (distributed representation) であると考えられる。分散表現では一つのユニットが異なる多くの概念の表現に関わり、ある時点でどの概念が表現されているかはユニット活性度の全体的なパターンに反映される。このとき入力の種類度が隠れ層の活性度パターンの間の空間的距離に従い、状態空間において隠れ層の活性度パターンが近いほど、それらに対応する入力は類似していると考えることができる。状態空間内でのパターン間の距離はユークリッド距離で考えることができる。隠れ層活性度パターン間のユークリッド距離が近いほど類似し、離れるほど類似していない。隠れ層活性度パターンを可視的に表せるのであれば、ニューラルネットワークの動作の理解につながるのだが、状態空間が 100 にも 200 にも及ぶような高次元になる場合、それを直接視認し理解するのは困難である。そのため活性度パターンの類似度を考えるのは間接的な手法として一つの良い手立てである。

Elman は、文の次に続く単語を予測する課題を学習させた SRN に、単語を入力した際にあらわれる隠れ層の活性化パターンから図 2.2 の単語クラスターが生じることを示している [2]。入力に用いる単語は 31 ビットで表現され、単語ごとにユニークな 1 ビットだけが、“1”、それ以外は“0”をとる互いに等距離の直交したベクトルである。したがってどの単語間の距離もは等距離かつ直交している。各単語は 31 次元の状態空間において、それぞれ異なる頂点をとっているため、どの単語がどの単語と近いかという生得的な知識は与

えないことになる。

しかし図 2.2 では、私たちが名詞、動詞と分類するような集まり（クラスター）が隠れ層の内部表現で発生していることがわかる。またこのような統語範疇の違いだけでなく、人と動物のような意味的な（意味範疇の）違いのようなものも表現されている。

これら Elman の提示した結果から、前提知識を与えてない SRN の内部表現に、名詞と動詞の違いや、人と動物といった意味的な違いが見られるようである。しかし学習によって SRN が単語の統語範疇、意味範疇での位置づけを獲得したと結論づけるのには疑問が残る。単純に文脈（文中でその単語の前に表れている単語群）のみによるものではないかと考えられるからである。本研究では、SRN の内部表現から得られた単語クラスターが、学習によって単語の統語範疇の違いを獲得したことによるのか、単に文脈により隠れ層に出現しただけの現象なのかを検証する。

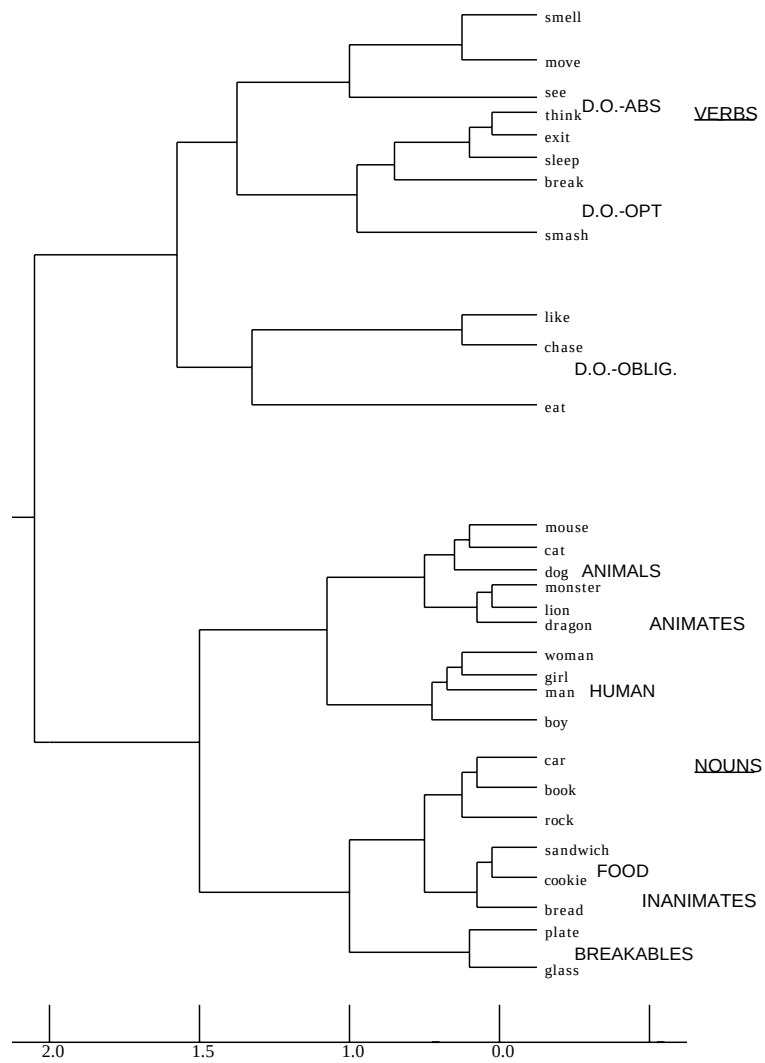


図 2.2: SRN 隠れ層状態の活性化パターンによる単語クラスター

第3章 実験手法

Elman の示した単語クラスターが表現している統語範疇は、SRN の学習による獲得を示すものであるのかについて検証する方法を提案する。

3.1 SRN の単語クラスターの獲得

予備実験により Elman の示した単語クラスターは確かに文中の次単語を予測させる課題を学習させた後の SRN の隠れ層の活性化パターンから確認できた (後節 4.1)。しかし、この結果をもってして、Elman のように SRN が統語範疇を学習によって獲得したと言えるかは疑問が残る。しかしながらその疑問を解消するために、SRN の内部でおこなわれている多次元な動作を解析するのは困難である。例えば、[2] での実験における SRN について解析するとすると、解析に重要なウェイトを占めるであろう隠れ層パターンは 150 次元ものベクトルになる。このような多次元空間の解析は難しい。そこで本論では直接 SRN を解析するのではなく、別のネットワークを用いて間接的に SRN の隠れ層活性化パターンに見られる単語クラスターが学習によるものとは言えないことを示す。この間接的手法として RAAM を用いる。

3.1.1 RAAM の固定長表現

Elman の単語クラスターを検証するために、Elman が使ったデータパターンを利用できれば、結果の比較が容易である。しかし、フィードフォワード型ネットワークを用いたコネクショニスト的な記号処理では可変長のデータがうまく取り扱えないという問題がある。そのため、可変長データを固定長データ表現する必要がある。そこで Pollack の RAAM(recursive auto-associative memory)[7] に着目した。

RAAM は図 3.1 のような再帰型の自己連想記憶ネットワークである。図 3.1 において、TOP、STACK のユニット数をそれぞれ L 、 M とすると、 $M + L$ 次元ベクトルが入力

(STACK と TOP) され、ユニット数 M の隠れ層 (STACK) に伝達され、出力は入力と同じベクトルに復元される (自己連想)。同時に隠れ層状態は STACK として保持され、次の入力に利用される (再帰)。入力層から隠れ層は圧縮、隠れ層から出力層は復元がなされるので、それぞれエンコーダ、デコーダとしての役割を果たしている。1 ステップ毎に TOP と STACK が結合されているものの STACK サイズは固定であるため、TOP に次々入力すれば可変長データを固定長データに変換することができる。

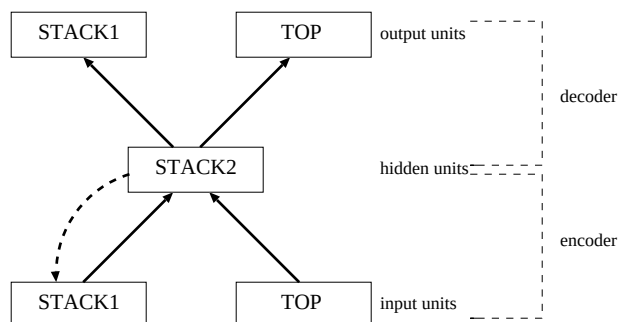


図 3.1: RAAM

各層を長方形で表す。入力データと同一のパターンが出力されるように学習する。したがって、教師信号は入力データである。STACK 1 に 1 ステップ前までの文脈が内部に表現されており、TOP とともに STACK2 に入力される。STACK1 からの 1 ステップ前までの文脈と新しい入力を結合した新しい文脈が STACK2 に表現される (エンコード)。出力は今の入力と 1 ステップ前までの文脈である (デコード)。新しい文脈は STACK1 に保持され次のステップに使われる。

3.1.2 RAAM での隠れ層活性度パターンのクラスタリング

RAAM のエンコーダ部分は新たな入力をそれまで stack されていた入力に結合し、圧縮し、また次の入力を結合するという動きを、STACK がクリアされない限り入力の都度繰り返す。文の中の単語を順に入力していくとすると、単語が入力されたところまでの文の一部が圧縮されて STACK に保持される。本論文ではこのときの RAAM の隠れ層 (STACK) に着目し、単語が入力されて現れる、隠れ層の活性化パターンから階層的なクラスタの構築を試みる。エンコーダ部分での結合荷重の更新は行わない。構築されるクラスタが学習によるものではなく、単に入力単語の系列による隠れ層の活動の結果であることを示す。このことが示されれば、Elman のように単語クラスタの構築によって、

前提知識を与えていない SRN が、文の中の単語を予測する課題を行うだけで単語の意味らしきものを学習している [2] と簡単に言えなくなることになる。

3.2 文脈に依存しない単語クラスターの構築

Elman[2] のように単語のクラスターを SRN 隠れ層での活性度パターンによって構築すると、そのクラスターは学習によって SRN が得た、単に文脈層にスタックされている状態と入力単語を反映しただけであるという可能性が否定できないということは 2.2 節で述べた通りである。そこで図 3.2 のように文脈層との結合のない隠れ層 B(hidden units B) を SRN に追加する。隠れ層 B は入力層状態を圧縮し、隠れ層 A(hidden units A) に伝達するだけで、文脈層状態の影響は受けない。そのため、隠れ層 B に現れる活性度パターン、は文脈層の影響による、単語の前後関係は反映しないはずである。しかし、もし SRN が学習によって単語クラスターを獲得するとすれば、その痕跡が隠れ層 B に観測されても良いはずである。なぜなら、もし隠れ層 B に学習の効果が何も残っていないとすれば、SRN の学習は、次単語予測が行ないやすいように文脈層の表現を変更しただけであるからである。もし Elman が述べたように、SRN が単語の統語範疇や意味のような概念を学習しているのだとするならば、この隠れ層 B の活性化パターンから統語範疇が反映されているとされる階層的な単語クラスターと同様ものが得られると考えられる。

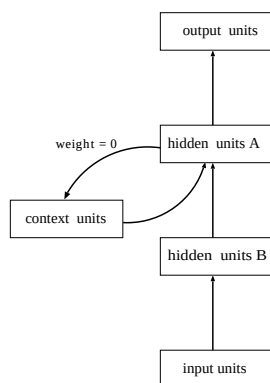


図 3.2: 文脈層との結合のない隠れ層をもつ SRN

各層を長方形であらわす。文脈層と結合のある隠れ層 A と入力層の間に隠れ層 B を加えている。隠れ層 B は文脈層と結合していないため、文脈によらない内部表現を作る。

3.3 クラスター分析

Elman は文中の次単語を予測するように訓練した SRN に、学習済みの単語を入力し、そのときの隠れ層の活性化パターンから階層的なクラスターを作成している [2]。本研究でも単語に対応する隠れユニットの活性化パターンを階層的な木構造によって表現する。用いるクラスタリング手法は以下の通りである。

n 個の単語について、隠れユニット数を p とする。単語に対応する隠れユニットの活性度を $X_{i1}, X_{i2}, \dots, X_{ip} (i = 1, 2, \dots, n)$ とする。

1. 個々の単語に対応する隠れユニットの活性化パターンを決定する。
2. それぞれの単語のパターンについて他の単語すべてのパターンとのユークリッド距離 d_{ij} を測定する。

$$d_{ij} = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2} \quad (i, j = 1, 2, \dots, n) \quad (3.1)$$

3. ユークリッド距離の最も近い単語、もしくはクラスターを結合して、1つのクラスターとする。クラスター a とクラスター b が結合されたクラスターをクラスター c とする。このとき結合後のクラスター c とクラスター x ($x \neq a, b$) との距離 d_{xc} は次の式であらわされる。

$$d_{xc} = \alpha_a d_{xa} + \alpha_b d_{xb} + \beta d_{ab} + \gamma |d_{xa} - d_{xb}| \quad (3.2)$$

また本研究では最近隣法でクラスタリングをするものとし、各パラメータを $\alpha_a = 0.5$ 、 $\alpha_b = 0.5$ 、 $\beta = 0$ 、 $\gamma = -0.5$ で与える。

$$d_{xc} = 0.5d_{xa} + 0.5d_{xb} - 0.5 |d_{xa} - d_{xb}| \quad (3.3)$$

これにより2つのクラスターが1つになるので、クラスターの総数が1つ減少する。

4. 3 の手続きをクラスターが1つになるまで繰り返す。

第4章 実験結果

4.1 予備実験

本研究の実験に先立ち、Elman の構築した単語クラスター [2] の再現を行った。

4.1.1 実験環境

Elman の実験条件（入力データ、各層のユニット数など）と同一の条件を用いた。入力層と出力層のユニット数は 31、隠れ層と文脈層のユニット数は 150 である。単語のカテゴリー分け、例文ルールを表 4.1、4.2 に記載する。使用する単語は 26、例文ルールは 16、それにより作成された例文は 225 文である。また実際の入力データは入力ビットのうち、各単語に対応した 1 ビットのみを 1 とし、他のビットは 0 とすることにより、それぞれの単語が直交し、等距離になるようにした（表 4.3）。学習には BP(back propagation) を用い、例文をランダムに約 4 万文入力した。

4.1.2 実験結果

作成した例文の単語を入力し、次に続く単語を予測する課題を学習させた後、学習に用いた例文を入力し、入力される単語毎に現れる活性化パターンの平均をとり、その単語の活性化パターンを決定した。3.3 節で示したクラスタリングアルゴリズムを用いた結果、図 4.1 のような結果となった。Elman の示したクラスター [2] のように明確に名詞、動詞のクラスターは構築できなかったが、大きな 3 つのクラスターにおいて、文中、1 番目（主語）、3 番目（目的語）の両方に出現している単語のクラスター、2 番目（動詞）にのみ出現しているクラスター、3 番目（目的語）にのみ出現しているクラスターという様になっている。小さなクラスターに着目すると、完全ではないとはいえ、出現パターンが似ているものほど、近い位置に集まっていることが確認できた。この結果から確かに Elman が示したような単語クラスターが SRN によって構築できることが分かった。

表 4.1: 入力単語とカテゴリー分け

No	Word	Category
1	man	N-HUM
2	woman	N-HUM
3	boy	N-HUM
4	girl	N-HUM
5	mouse	N-ANIM
6	cat	N-ANIM
7	dog	N-ANIM
8	lion	N-ANIM,N-AGRESS
9	dragon	N-AGRESS
10	monster	N-AGRESS
11	book	N-INANIM
12	rock	N-INANIM
13	sandwich	N-FOOD
14	cookie	N-FOOD
15	bread	N-FOOD
16	plate	N-FRAG
17	glass	N-FRAG
18	see	V-PERCEPT,V-TRAN
19	break	V-AGPAT,V-DESTROY
20	smash	V-DESTROY
21	eat	V-EAT
22	chase	V-TRAN
23	move	V-AGPAT
24	smell	V-PECEPT
25	think	V-INTRAN
26	sleep	V-INTRAN

左端の No は後節のクラスター図の要素番号に対応している。

表 4.2: 例文パターン

WORD1	WORD2	WORD3
N-HUM	V-EAT	N-FOOD
N-HUM	V-ERCEPT	N-INANIM
N-HUM	V-DESTROY	N-FRAG
N-HUM	V-INTRAN	
N-HUM	V-TRAN	N-HUM
N-HUM	V-AGPAT	N-INANIM
N-HUM	V-AGPAT	
N-ANIM	V-EAT	N-FOOD
N-ANIM	V-TRAN	N-ANIM
N-ANIM	V-AGPAT	N-INANIM
N-ANIM	V-AGPAT	
N-INANIM	V-AGPAT	
N-AGRESS	V-DESTROY	N-FRAG
N-AGRESS	V-EAT	N-HUM
N-AGRESS	V-EAT	N-ANIM
N-AGRESS	V-EAT	N-FOOD

表 4.3: 単語に対応付けるビットパターン

man	01000000000000000000000000000000
book	00000000000001000000000000000000
see	00000000000000000001000000000000

各単語にはこのようにビット列のうち1個所だけビットを立てたパターンを対応づける（ローカルコーディング）。

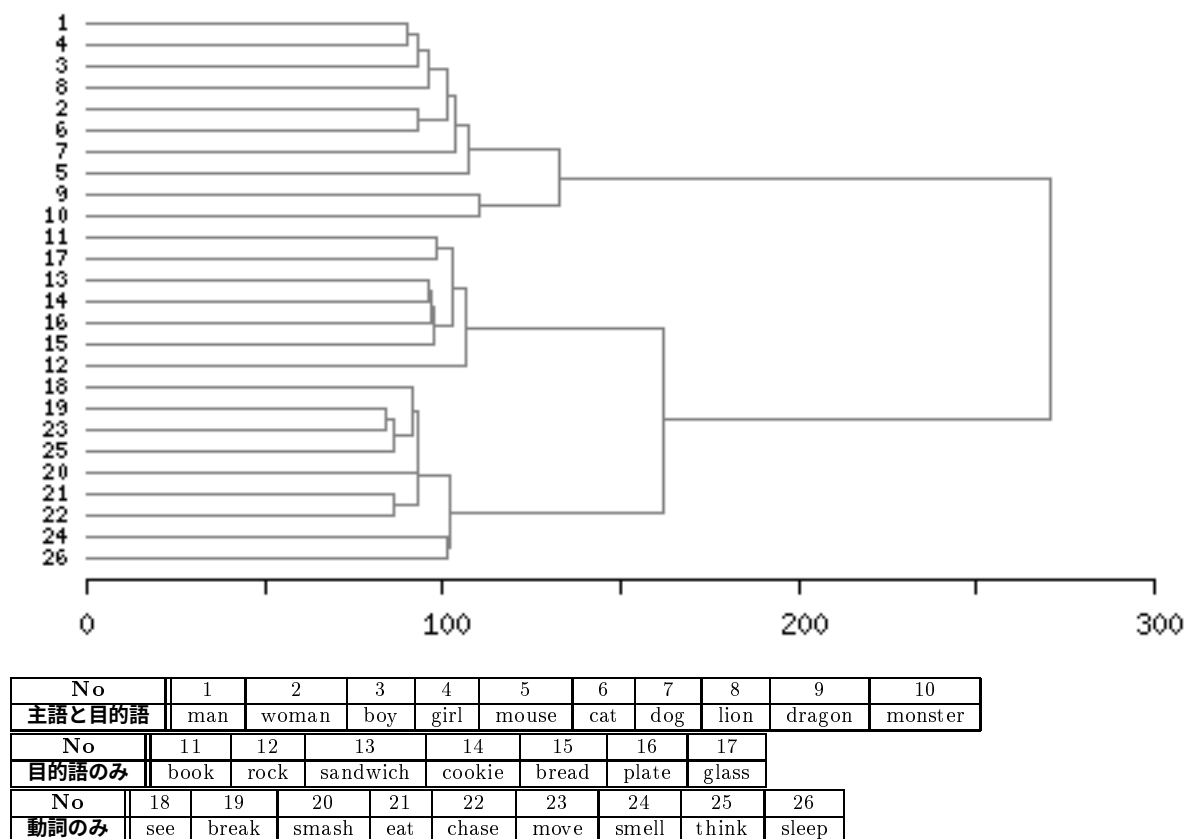


図 4.1: 追実験で構築した単語クラスター

横軸はクラスター間のユークリッド平方距離。文中で主語と目的語に相当する位置両方にあらわれる単語のクラスター（1-10）、目的語の位置にのみある単語のクラスター（11-17）、動詞の位置にある単語のクラスター（18-26）の大きな3つのクラスターを持っている。

4.2 RAAMによる単語クラスター

Elman が構築した単語クラスター同様のクラスターが RAAM によっても構築されるかどうかを調べる。

4.2.1 実験環境

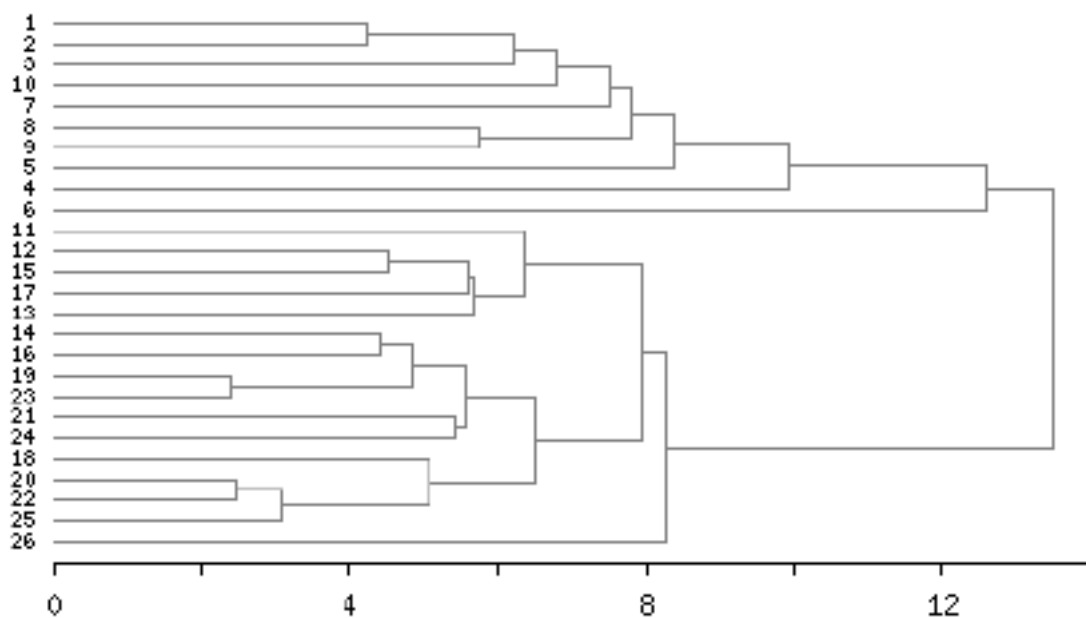
RAAM に入力するデータは Elman の追実験で使用了もの（表 4.1、4.2、4.3 をもとに作成したデータ）を使う。入力層、出力層のユニット数は Elman[2] 同様 31。隠れ層のユニット数は 10 とするが、確認のため Elman[2] と同じ 150 でも行った。結合荷重は初期値を-1 から 1 の間でランダムに設定し、重みの更新は行わない。単語データを入力する毎に隠れ層の活性化パターンをとり、各単語の文脈毎の活性度パターンの平均を計算し、その値を単語のパターンとする。この様に決められた単語のパターンでクラスタリングする（3.3 節）。

4.2.2 実験結果

隠れ層の活性度パターンから構築した単語クラスターは隠れ層のユニット数 10 の場合、図 4.2 の様になった。図 4.1 とクラスタ間の距離のスケールこそ異なりはするが、良く似た階層的な単語クラスターになったと言える。また隠れ層のユニット数を 150 とした場合は、図 4.3 のようになり、図 4.1 同様、名詞や動詞的用法のような大きな 2 つのクラスターとなり、小さなクラスターも比較的、用法の近いもので構成されている。

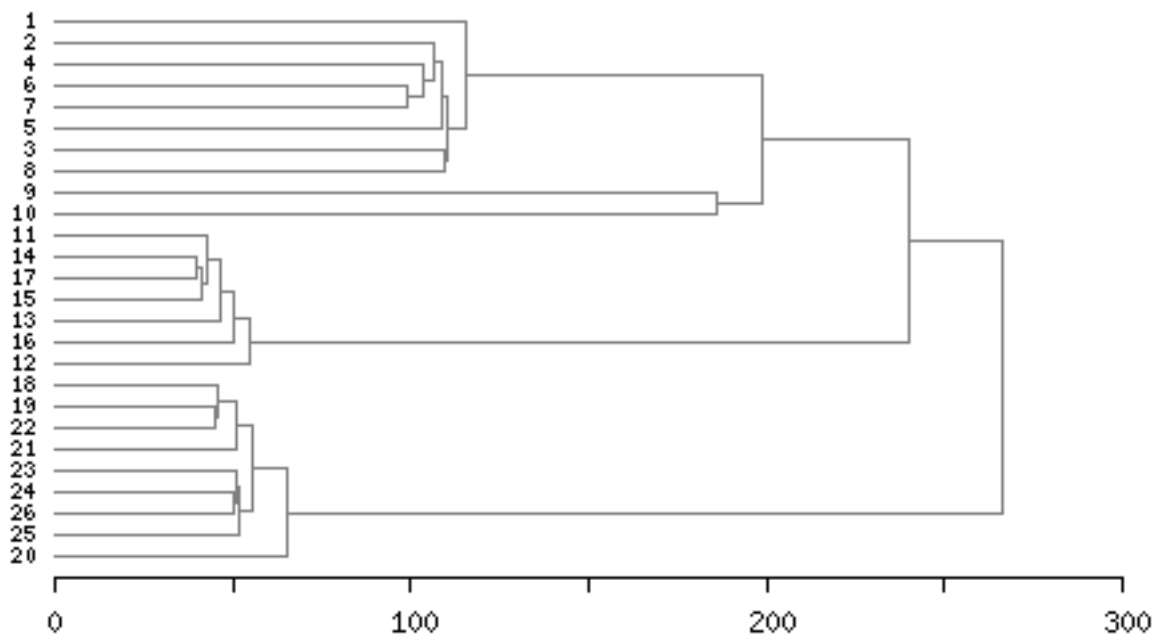
ここで重要な点は学習のしていない RAAM によって、Elman[2] の単語クラスターが構築できたことである。

これらの結果から文中の次単語を予測させる課題を行わせた SRN の隠れ層の活性化パターンにより構築される階層的な単語のクラスターは、特に SRN が単語の統語範疇を学習したことによるものではないと言える。



No	1	2	3	4	5	6	7	8	9	10
主語と目的語	man	woman	boy	girl	mouse	cat	dog	lion	dragon	monster
No	11	12	13	14	15	16	17			
目的語のみ	book	rock	sandwich	cookie	bread	plate	glass			
No	18	19	20	21	22	23	24	25	26	
動詞のみ	see	break	smash	eat	chase	move	smell	think	sleep	

図 4.2: RAAM 隠れ層 (10 ユニット) 活性化パターン単語クラスター
 横軸はクラスター間のユークリッド平方距離。文中で主語と目的語に相当する位置両方にあらわれる単語のクラスター (1-10) は大別できるが、目的語の位置にのみある単語のクラスター (11-17) と動詞の位置にある単語のクラスター (18-26) は隠れ層のユニット数が多い図 4.3 に比べクラスターの分類感度が良くない。



No	1	2	3	4	5	6	7	8	9	10
主語と目的語	man	woman	boy	girl	mouse	cat	dog	lion	dragon	monster
No	11	12	13	14	15	16	17			
目的語のみ	book	rock	sandwich	cookie	bread	plate	glass			
No	18	19	20	21	22	23	24	25	26	
動詞のみ	see	break	smash	eat	chase	move	smell	think	sleep	

図 4.3: RAAM 隠れ層 (150 ユニット) 活性化パターン単語クラスター

横軸はクラスター間のユークリッド平方距離。図の要素は 12 頁の表 4.1 の番号に対応する。文中で主語と目的語に相当する位置両方にあらわれる単語のクラスター (1-10)、目的語の位置にのみある単語のクラスター (11-17)、動詞の位置にある単語のクラスター (18-26) の大きな 3 つのクラスターがはっきり出ている。

4.3 文脈によらない単語クラスターの構築

4.3.1 実験環境

SRN に文脈層と結合のない隠れ層を追加し、例文の中の次の単語を予測するタスクを与える。入力データは 4.1、4.2 節と同じものを使用する。各層のユニット数は入力層、出力層では 31、隠れ層 A では 150、隠れ層 B では 5 とする（図 4.4）。文脈層と結合のない隠れ層 B について、階層的な単語クラスターの有無を調べる。

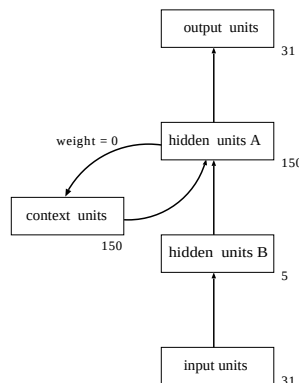


図 4.4: 文脈層と結合のない隠れ層を加えた SRN
各層は長方形であらわされる

4.3.2 実験結果

文脈層と結合のない隠れ層 B の活性化パターンによる階層的な単語クラスタリングを試みた（図 4.5）。この図には用法の近いものが近くの配置になるというような特徴（例えば図 4.1 のような）は見られない。この時、文脈層と直接結合している隠れ層 A で、同じようにクラスタリングを行うと図 4.6 のように、用法の近い単語が近くに配置される階層的なクラスターができているのがわかる。

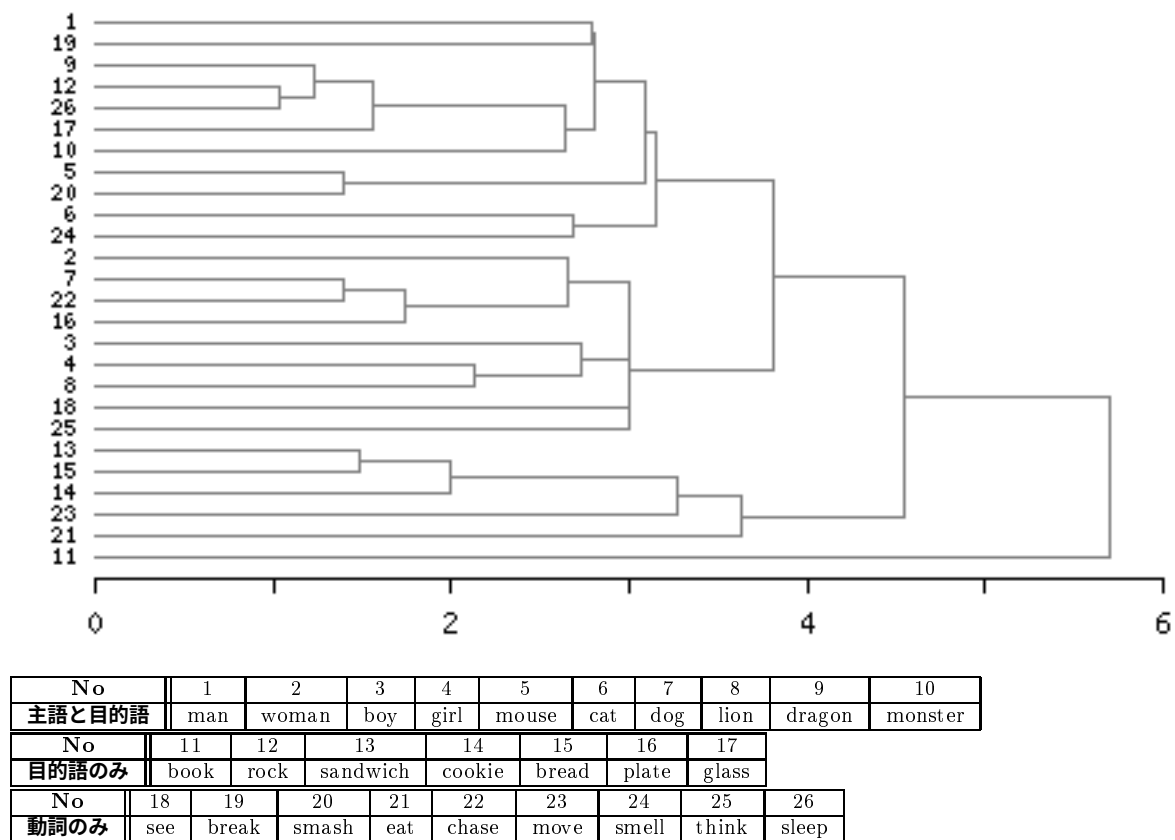
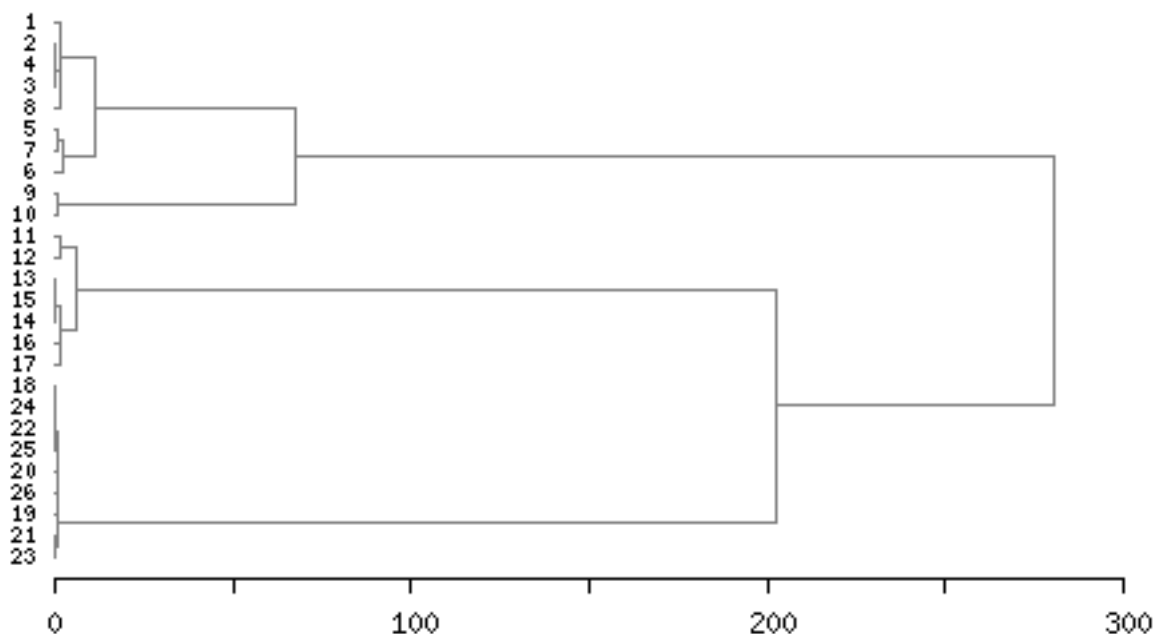


図 4.5: 文脈層と結合のない隠れ層 B での単語クラスター
 横軸はクラスター間のユークリッド平方距離。特に統語範疇や意味範疇のような
 概念が表現されている構造を持たない。



No	1	2	3	4	5	6	7	8	9	10
主語と目的語	man	woman	boy	girl	mouse	cat	dog	lion	dragon	monster
No	11	12	13	14	15	16	17			
目的語のみ	book	rock	sandwich	cookie	bread	plate	glass			
No	18	19	20	21	22	23	24	25	26	
動詞のみ	see	break	smash	eat	chase	move	smell	think	sleep	

図 4.6: 文脈層と結合のある隠れ層 A での単語クラスター

横軸はクラスター間のユークリッド平方距離。学習していないにも関わらず、文中主語と目的語に相当する位置に両方あらわれる単語のクラスター (1-10)、目的語の位置にのみある単語のクラスター (11-17)、動詞の位置にある単語のクラスター (18-26) の大きな3つのクラスターがはっきり出ている。また小さいクラスターを見ても意味的 (人や動物のような概念) に近いものが配置されている。8が1-4 とクラスターを作っているが、8は2つのカテゴリー (N-ANIM,N-AGRESS) に分類されており例文生成の上で、N-HUM(1-4) とほぼ同様になるためである

4.3.3 文脈依存を除去した隠れ層の活性化パターン

それでは文脈自体が存在しない場合はどうであろうか。語系列予測課題を与えて学習させた後、文ではなく、独立した単語を入力した場合隠れ層 A での活性化パターンはどのようなのであろうか。ここでの実験は図 4.4 の SRN に 4.3.1 節と同じタスクを行わせた後、文脈層を除いて（単語一つ入力するごとに文脈層のクリアを行う）、隠れ層 A の活性化パターンで階層的クラスターを構築している（図 4.7）。これも用法の近いものが近くに配置されるとは言えない。しかし図 4.5 とクラスター間の距離のオーダーの違いこそあれ、よく似たクラスターとなっている。隠れ層 B で圧縮された入力が入力層 A に伝達されただけあり、隠れ層 B から隠れ層 A への結合によっても、統語範疇を反映した単語クラスターが構築できなかったと結論することができる。

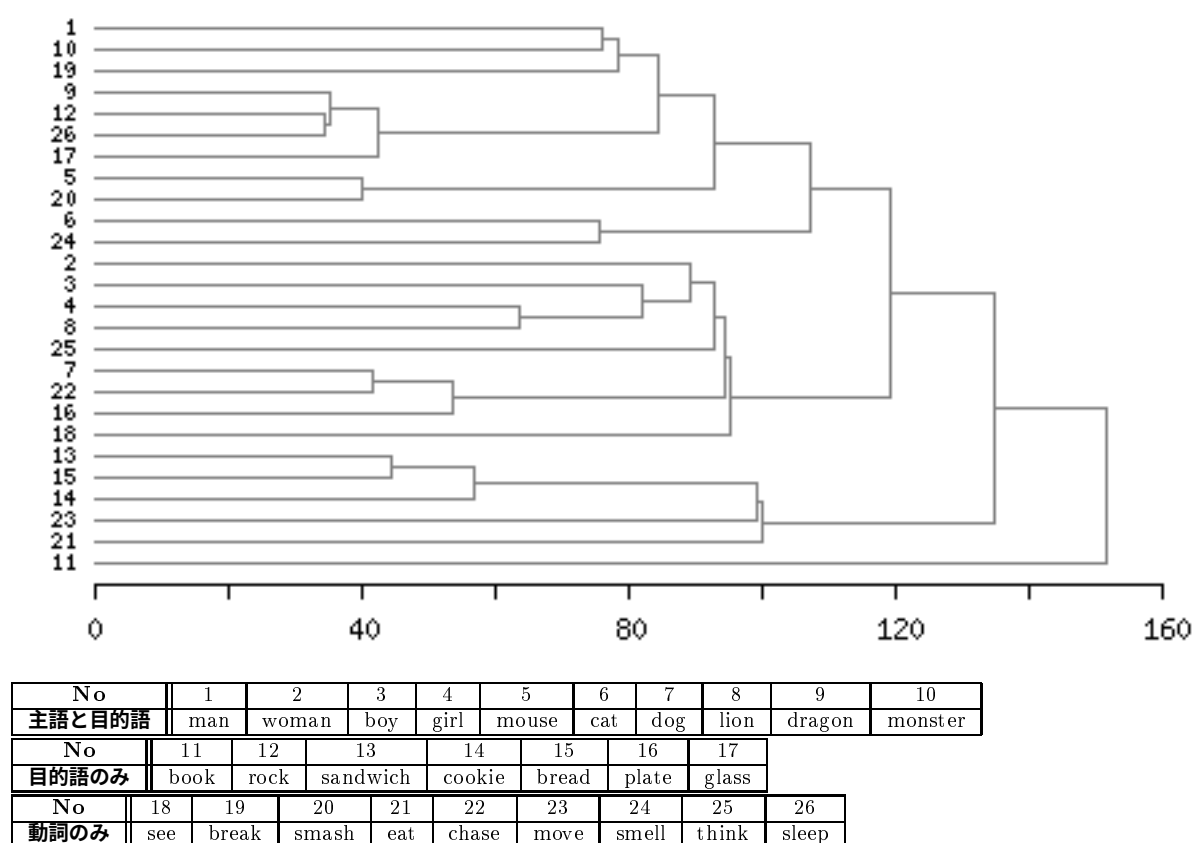


図 4.7: 学習後に文脈依存を取り除いた隠れ層 A での単語クラスター
横軸はクラスター間のユークリッド平方距離。文としてではなく、単語を一つずつ入力して文脈に依存のない入力とした。特に統語範疇、意味範疇のような概念が表現されている構造がない。

第5章 考察

RAAM の学習をさせないエンコーダ動作によって、図 4.2、4.3 によって、文法的、意味的と解釈されうるクラスターをもつ階層的な単語クラスターが構築できた。この結果が示すものは、前提知識のないネットワークが単に単語列（文）の時系列処理を行うだけでも統語範疇的な単語クラスターができるということである。つまり統語範疇や単語の意味の近さを示すようなクラスターを持つ階層的な単語クラスターが表現できるからといって、そのネットワークが統語範疇や単語の意味を学習し、獲得したとは言い切れないということである。それは学習させていない SRN でも同様の単語クラスターができたことから明らかである（図??）。

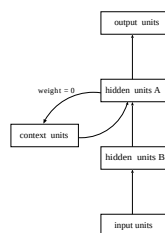


図 5.1: 文脈層と結合のない隠れ層を加えた SRN

文脈層と直結している隠れ層において単語の内部表現から統語範疇を反映した単語クラスターが構築できたからといって、それはネットワークが学習によって、単語の統語範疇を獲得したことを正当化するものではないことがわかった。しかし、これだけでは単語クラスターが例文の構文的なものに引き摺られている可能性を指摘できただけで、完全に SRN が文法や単語の意味を獲得していないと断言することはできない。そこで図 5.1 のような文脈層と結合のない隠れ層 B を SRN に加え、その隠れ層 B で同じように単語クラスターを構築を試みた。結果として、文脈層と直結する隠れ層 A では、文法的、意味的なものを反映しているようなクラスターが構築できた（図 4.6）にも関わらず、隠れ層 B ではそのような単語クラスター（図 4.5）はできなかった。各単語に対する活性度パターンは文脈によって異なるパターンも持つので、ここまでの実験では、Elman が行なった

ように、学習後のネットワークに多くの例文を与えて、そうした異なる文脈のベクトルを単語毎に平均したパターンを用いて単語クラスターを構築してきた。隠れ層 B から隠れ層 A の結合に、統語範疇を反映したクラスターを作成する能力があるか否かを検証することにした。そこで学習後に文ではなく単語を 1 つずつ入力したときの隠れ層 A の活性度パターンから単語クラスターを構築した (図 4.7)。しかし期待するような単語クラスターは構築できなかった。どちらも文脈依存のない状態での単語クラスターの構築であったと考えると、統語範疇や単語の意味の近さを反映するクラスターをもつ階層的な単語クラスターは、学習後であっても文脈に依存した形でしかないかのようである。

本研究での実験結果をまとめると、Elman の示した単語クラスターは以下の様に言える。

- 学習の有無に特に左右されず、文脈情報を直接使うだけで構築できる。
- 学習後でも入力層側 (隠れ層 B) には構築できない。

ということになる。つまり、Elman が提示した単語クラスター [2] は、SRN が文法 (名詞や動詞) や意味 (人や動物) というような概念の獲得を反映した結果ではないということが分かった。

第6章 まとめ

本論文では、単語クラスターが名詞や動詞のような概念を示すかのようなクラスターを”文法のようなもの”、また、人や動物という概念を示すかのようなクラスターを”意味のようなもの”であるという曖昧ともいえる定義のもとで検証してきた。確かに Elman の示したクラスターの本構造は、構文的情報（名詞や動詞）ではない情報に従って、構造化されているようにも見える。しかし与えられた情報が例文のみであり、実世界にあるものや動作と結び付くような”意味”が反映されているわけではないのである。Elman の目論見は、見かけ上、構文情報しかない例文から「学習」によって裏にある単語間の意味的な（意味そのものではなく、意味が似ているといった）関係が抽出できるのではないかということであったのだろう。本論文では Elman の言い分である、「学習」によるものであったという点で否定的な結論となったが、ネットワークが文法カテゴリやそれに反映している意味カテゴリを獲得できないということを示すものではない。

本論文では文例から学習できる「意味」とは何であるのかを考えている。さらにそれを考えていくために本研究の今後の課題を付す。本論文では 2、3 語の単純な文でのみの検証であるが、さらに長い文において直接文脈情報を使うだけで単語クラスターができるのかを調べる。また文法規則のどこまでが文脈情報のみで表現できるのかについても検証する。つまり文例だけでできることの切り分けを行なうのである。

謝辞

本論文の執筆にあたり多くの方々にお世話になりました。

櫻井彰人教授には JAIST 退官後も何度も東京から足を運んでいただき、また、メールなどを通して数多くご指導いただきました。大変感謝しています。林幸雄助教授には的確なアドバイスをいただき、特に櫻井教授退官後は路頭に迷うところを拾い上げ面倒を見ていただき感謝いたします。荒木修助手には優しく、適切なアドバイスをいただき、感謝いたします。研究室の先輩方、そして同期の皆さんとは、研究だけでなく多く時間を共に過ごしてきたことは JAIST の山ごもり生活をとても楽しいものとしてくれました。ありがとうございました。

最後にこの年になるまで働く苦勞を知らない自分を、温かく見守り、支援してくれた両親に最大の感謝を捧げます。

参考文献

- [1] C. L. Baker. Syntactic theory and the projection problem. *Linguistic Inquiry*, 10:532–581, 1979.
- [2] Jeffrey L. Elman. Finding Structure in Time. *Cognitive Science*, 14:179–211, 1990.
- [3] Jeffrey L. Elman. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7:195–225, 1991.
- [4] Jeffrey L. Elman. Learning and development in neural networks:The importance of starting small. *Cognition*, 48:71–99, 1993.
- [5] Jeffrey L. Elman, Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. *Rethinking Innateness: A connectionist perspective on development*. Cambridge,MA:MIT Press, July 1996.
- [6] J. Feldman, M. Fnaty, and N Goddard. Computing with Structured Neural Networks. *IEEE Computer*, 21:91–104, 1988.
- [7] Jordan B. Pollack. Recursive Distributed Representations. *Artificial Intelligence*, 46:77–105, 1990.
- [8] Frank Rosenblatt. The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 65:368–408, 1958.
- [9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning Internal Representations by Error Propagation. *Parallel Distributed Processing*, 1986.
- [10] 今井 むつみ, editor. **心の生得性**. 共立出版, August 2000.
- [11] 橋田 浩一, 大津 由紀雄, 今西 典子, Yosef Grodzinsky, and 錦見 美貴子. **言語の獲得と喪失（言語の科学 10）**. 岩波書店, February 1999.

- [12] 玉森 彩弥香 and 乾 敏郎. Elman ネットによる統語範疇の配列と格関係の学習. **認知科学**, 6(3), 1999.
- [13] 小林 春美 and 佐々木 正人, editors. **子供たちの言語獲得**. 大修館書店, October 1997.
- [14] 原田 哲治. 再帰型神経回路網による文法の獲得. Master's thesis, 北陸先端科学技術大学院大学, March 2001.
- [15] 金水 敏 and 今仁 生美. **意味と文脈**. 岩波書店, March 2000.
- [16] 大津 由紀雄, 坂本 勉, 乾 敏郎, 西光 義弘, and 岡田 伸夫. **言語科学と関連領域（言語の科学 11）**. 岩波書店, September 1998.
- [17] 郡司 隆男, 安部 泰明, 白井 賢一郎, 坂原 茂, and 松本 祐治. **意味（言語の科学 4）**. 岩波書店, March 1998.