リカレントネットワーク

浅川伸一 <asakawa@twcu.ac.jp>

一般に時系列情報や文脈依存の情報を処理させるために考えられたネットワークとして単純回帰型ネットワーク (simple recurrent network, SRN) あるいはリカレントネットと呼ばれる回路があります[1]。 これらの例は、提案者

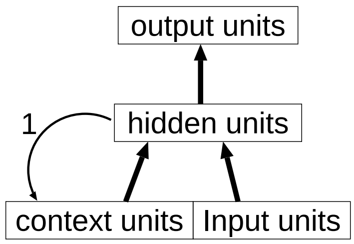

Figure 1: エルマンネット

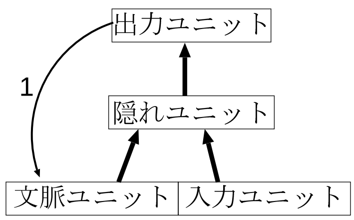

Figure 2: ジョーダンネット

の名前で エルマンネット (Elman net, 図 1), ジョーダンネット (Jordan net, 図 2) と呼ばれます。図では、ある層の一部が帰還信号 (recurrent feedback) を受け取るようになっています。このような帰還信号を受け取る素子を文脈層あるいは関連層と呼ぶ。

エルマンネットでは、入力層は入力信号を処理する入力ユニットと、直前の中間層の状態を入力とする文脈ユニットとで構成されています。文脈ユニットは以前の中間層をコピーするためだけ (すなわち中間層から文脈ユニット

[1]時間的な順序関係を表現しようとしたものに 時間遅れネット (Time Delay Network) とよばれるものも存在します。

への結合強度は 1.0) です。結合強度の学習は順方向の結合についてだけ行われるので、通常の誤差逆伝播法がそのまま適用できます。

ある時刻 t で処理される内容は、その時点での入力信号と、それ以前の時刻 $t-1$ までで処理された回路の状態を表す信号とを同時に処理することになります。すなわち、文脈層は過去の状態を記憶していることを意味します。この結果、ある時刻 t でのネットワークの状態は現在の入力と過去の入力履歴の集合によって決まることになります。例えば、図 3 において過去の影響を考えれば時刻 t における中間層の状態 $h(t)$ は

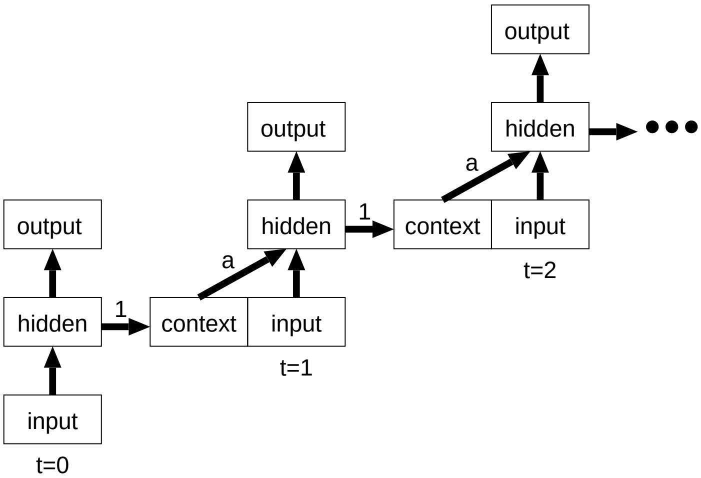

Figure 3: エルマンネットの時間発展

$$
\begin{aligned}
h(t) &= I(t) + ah(t-1) \\
&= I(t) + a(I(t-1) + h(t-2)) \\
&= I(t) + aI(t-1) + a^2 I(t-2) + a^2 h(t-3) \\
&= \sum_{\tau=0}^{T} a^{\tau} I(t-\tau)
\end{aligned} \qquad (1)
$$

と表すことができます。ここで $I(t)$ は時刻 t における入力を表します。文脈層からの影響 a が 1 より小さければ過去の入力からの影響が指数関数に従って小さくなることを表しています (1 より小さければ過去の影響は急速に小さくなる)。このことは、エルマンネットが一つ前の状態を保存しておくという単純な構造にもかかわらず、過去履歴に依存した出力を生じることを示しています。

次にもう一つ別の簡単な例として図 4 のような中間層の素子数が 2 個の場合を考えてみましょう。時刻 t において入力層からの情報が途切れたとき、2 つの中間層素子の状態が $h_1^2 = h_2^2 = 1$ を満たし、かつ 文脈層から中間層

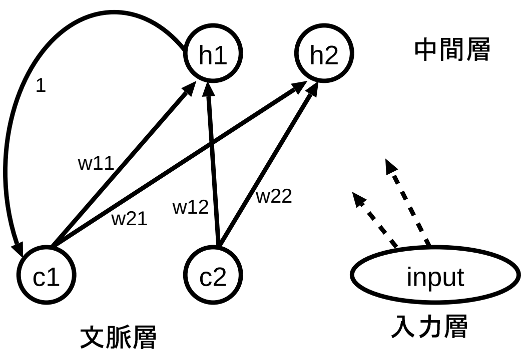

Figure 4: 2 素子だけの単純なエルマンネット

への結合係数が

$$\boldsymbol{W} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} = \begin{pmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{pmatrix} \tag{2}$$

のようになっていたとします。このとき $\boldsymbol{W}$ は回転行列であるから、例えば $\theta = \frac{\pi}{2}$ なら、単位円上の４点を循環します。この４点に対応する状態と特定の出力が結び付いているのなら、入力が 0 であるかぎり 4 状態を永遠に循環する出力が得られます。$n\theta = 2\pi$ を満たすかぎり n 個の状態を循環する出力を得ることができますが、θ の取り方は任意であるので、エルマンネットは十分に長い系列を記憶する能力があることがわかります。しきい値の存在により生体では $\theta \pm \epsilon$ は同じ状態として扱われると考えれば、短期記憶あるいは作業記憶のモデルとしても有効でしょう。直前の入力系列を再生する能力がしきい値の存在により $n = 7 \pm 2$ 程度 (すなわち $2\pi/7$ だけ回転する行列 $\boldsymbol{W}$ を用いる) の状態しか保持できないと仮定すれば、単純な短期記憶のモデルの一つになりうるのです。

エルマンら (**?**) は

Manyyearsagoaboyandgirllivedbytheseatheyplayedhappily.

のような文章の区切りを見つけることをネットワークに要求しました。ここでの入力は文中の 1 音素で、出力は次の 1 音素を予測するように訓練されました。

図からエラーは語頭で高く、語末まで減少している様子が分かります。エラー曲線を確信度と解釈すると、単語内の次にくる音素をかなり確信を持って予測していると解釈できます。一方、入力が単語の終りに達すると次にどの語がくるか予想できないので、結果としてノコギリ状のエラー曲線になるわけです。エラーの特徴については、実際観察される幼児の言葉の誤りと類似していて a boy を aboy と切り出してしまうオーバーシューティング (overshooting)

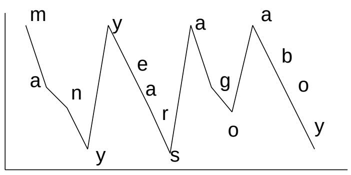

Figure 5: エルマンによる音素予測課題のエラー曲線

や、they を the y とするアンダーシューティング (undershooting) のエラーが観察されています。

このような単純回帰型ネットワークの強力な性質を利用して、エルマンネットでは言語情報処理などへの応用が試みられています。一方、ジョーダンネットでは運動制御への応用が試みられています。エルマンネットとジョーダンネットの違いは、後の処理で利用する形式が出力層で表現される形がよいのか、中間層の形式の方がよいのかという違いです。例えば、入力信号を視覚情報、ジョーダンネットの出力を現在の手や腕の位置だとすると、現在の手や腕の位置と視覚情報とから、手や腕を操作して目標物を掴むための運動を制御する問題を解くことができるでしょう。

さらにアトラクタネット (attractor net) と呼ばれるネットワークも提案されています。以下に英単語の意味を学習するためにヒントンとシャリス (**?**) によって提案されたモデルを示します。

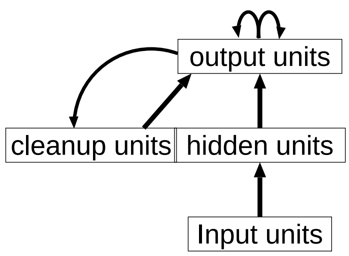

Figure 6: ヒントンとシャリスのアトラクタネット

一般にリカレントネットを数学的に解析するためには、出力ベクトル $\boldsymbol{y}$ に対して、文脈層からの結合係数行列を $\boldsymbol{W}$ とし時刻 t における入力を $\boldsymbol{I}(t)$ と表すことにすれば

$$\boldsymbol{y}(t) = \boldsymbol{W}\boldsymbol{y}(t-1) + \boldsymbol{I}(t), \tag{3}$$

と表すことができます。各素子への入力は互いに独立であると考えれば、$\boldsymbol{W}$ の対角成分にに $\boldsymbol{I}(t)$ の対応する値を加えることで $\boldsymbol{W}$ と $\boldsymbol{I}(t)$ をまとめて $\boldsymbol{W}(t)$ と表すことができる。これにより、(3) 式は

$$\boldsymbol{y}(t) = \boldsymbol{W}(t)\boldsymbol{y}(t-1), \tag{4}$$

と表現できる。(4) 式からリカレントネットの挙動は行列 $\boldsymbol{W}(t)$ の固有値によって定まると言えます。例えば、固有値が 1 より大きければ、対応する固有ベクトルの方向へ大きくなる。固有方程式が複素解を持つならば先に見たような回転する解が得られることになります。