

場所細胞 — その生成モデルと強化学習への適用 —

東京工業大学 ○青田 佳人 三宅 美博 鶴飼 正二

Place cells — Its modeling and application to the Q-learning. —

Tokyo Inst. Tech. ○Yosito Aota Yoshihiro Miyake Seiji Ukai

Abstract— Hippocampal place cells in rats show location-specific activity which is influenced by the visual input and self-locomotor experience. We present a model to demonstrate how place cells could be generated from the views that the rat receives as visual input. The temporal pattern learning as the hippocampal circuit associates the views. Behavior of the model rat determines whether place cells become functional or not. Place cells are generated when the model rat looks repeatedly from left to right while moving within the model space. However, if the model rat looks around too frequently, place cells are generated but become unstable. Then, we apply the model to the Q-learning.

Key Words: Rat, Hippocampus, Place cell, Generation, Neural network, Q-learning

1. まえがき

移動ロボットにおいて、ある地点から目標地点まで常に最短経路を選択しつつ移動できることは重要である。新規な環境においてこれを自律的に行うには、その環境の学習が必要であるが、この場合いかに効率よく学習を遂行するかが問題となる。

ここでカメラ付き移動ロボットを考えた場合、ロボットからの「見え」（視覚情報）と次に移動すべき「方向」の対応関係を直接学習させた場合と、見えからロボットの今いる「位置」を割り出すことができたとして、その「位置」と「方向」の対応を学習させた場合とでは、当然後者のほうがはるかに学習の効率がよいであろう。問題は、新規な環境において「見え」と「位置」の対応関係をいかにロボットに学習させるか、である。

ところで、ラットの海馬には「場所細胞」と呼ばれる神経細胞が存在する[1]。場所細胞とは、ラットがある特定の位置にきた時に反応する細胞のことで、その活動は主としてラット自身の行動（動機や運動情報など）と視覚情報に左右されるが、視覚や行動と直接対応しているわけではない。場所細胞は、あたかも自分を外から眺めているように反応するのである。このような、他者中心的（アロセントリック）な視点をコードしている細胞は、まさに前述の「見え」から「位置」を割り出す、我々の求める細胞と言えよう。

そのため本研究では、場所細胞の生成機構をモデル化することを目的とした。そして、構築したモデルの出力を各状態として次の移動方向を決める強化学習に当てはめ、「見え」を直接各状態にした場合との比較を行った。なお本研究では、ある空間内のある限られた領域において、ラットがどの方向に向いているとも発火する細胞を場所細胞と定義する。

さて、場所細胞の生成を問題にしてその数理モデル化を試みた研究は、現在までに多数報告されている（例えば、[2][3][4][5][6]）。しかしながらそれらはいずれも、視覚に関する情報のみから場所細胞の生成過程の説明を試みており、行動に関する影響が考慮されていなかった。さらにいずれのモデルも、視覚入力として複数の方向からの視覚情報を同時に入れたり、手掛かり刺激との間の角度を入力したりなど、実質的に既に入力の段階で位置の情報となっていた。視覚と行動の両方を考慮したマクノートンら[7][8]のモデルも、位置と等価の情報を視覚情報として入力しており、かつラットは海馬において場所細胞をあらかじめ持っているということが前提となっていた。

場所細胞の生成を視覚情報のみから説明しようとする上記

のような不自然な入力になると判断し、我々は見えと行動の両方が組みとなり、それと海馬特有の構造が場所細胞を発現させると考える。

本研究ではまず、見えと行動の両方の影響を受け、それに海馬特有の構造から発現される場所細胞の数理モデル化を目指した。特に、海馬に見られる巡回回路型の構造[9]に時系列学習則をあてはめ、それとラット自身の行動が組み合わさって、場所細胞の生成に必要な見えの情報を選択し連合すると考えてモデル化を行った。そしてそのようにして生成された場所細胞が、学習にどう寄与しているかを調べた。第2章ではモデルの構成を説明し、第3章においてそのシミュレーション結果と考察を、第4章においては通常の強化学習と本モデルを強化学習へ適用したものととの比較を、第5章においてまとめと今後の課題を述べる。

2. モデル

2.1 モデルの概要

本モデルは、4層のニューラルネットワークから成り、その結合様式は生理学的・解剖学的に知られている海馬の結合様式に矛盾しないように設定した。その概略をFig.1に示す。

R層は感覚入力層であり、大脳新皮質の連合野を想定している。視覚情報（モデルラットが外界を見たときの見えに対応する情報）がここに入力される。

E層はR層の入力パターンを自己組織化マップ[10][11]によってクラスタリングする層である。連合野の情報の集約する所として、内嗅野、海馬台皮質または海馬の歯状回を想定している。

C3層は、E層からの入力を時間的系列にしたがって連合する層である。C3層は海馬のCA3を想定している。重松ら[12]は時系列学習が可能な、Hebb則を拡張した学習則を提案したが、E層→C3層間の結合ではそれを簡略化した学習則を用いて、E層で起こった活動の時間的な前後関係を結ぶようにした。これは、歯状回→CA3→CA1→海馬台皮質・内嗅野→歯状回という海馬体における巡回回路[12]を模した形となっている。

C1層は、C3層からの入力を競合学習によって分類する層であり、海馬のCA1を想定している。場所に対応した発火パターンは、この層とC3層において見られる。

以上のようにモデルを構築し、モデルラットの行動を通して複数の見えを選択し、それらの見えを時系列学習によって連合させ場所細胞としてコードさせようと試みた。

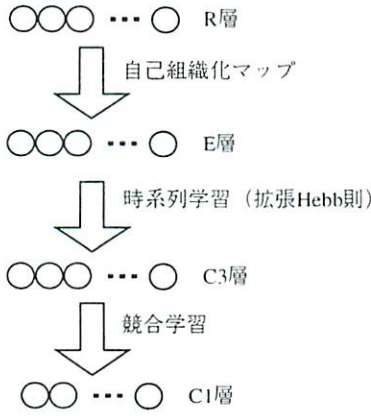


Fig. 1 Structure of the model.

2.2 モデル空間

今回のシミュレーションでは、Fig.2に示したような円形のフィールドを設定した。これは、場所細胞の実験にしばしば用いられているミルクタンク[13]を模したものである。ミルクタンクとは、円形の桶にラットの足が底に届かない程度に水を入れ白く濁らせたもので、通常は桶の一部の領域にラットの足場となる台を置いてラットを遊泳させる。本研究では 21×21 の格子図を円形に切り取り、その中をモデルのラットが1マスずつ動くようにした。円の外には、目印となる8つのパターンを置いた。

2.3 R層

R層は感覚入力層で、0~1の値をとる15個のニューロンから成っている。モデルラットがフィールド内を移動するたびに、モデルラットのいる場所と向いている方向に応じた視覚情報が入力される。紙数の都合上詳細は省くが、それぞれの視覚情報はモデル空間内の位置と方向が近ければそれだけお互いに似たパターンとなっており、実際の見えの移り変わりに近い形にしている。後で見るように、一つの見え(とそれに近い見えの集まり)とE層の一つの細胞が1対1対応になってさえいれば、R層の設定はあまり問題にならない。

2.4 E層

E層は、自己組織化マップによってR層のパターンの特徴抽出を行う層である。E層は700個のニューロンから成り、時刻 t におけるE層 i 番目のニューロンの出力を y_{Ei}^t 、膜電位を q_{Ei}^t 、R層 i 番目のニューロンの出力を y_{Ri}^t とおくと以下の式に従う。

$$q_{Ei}^t = \sum_{j=1}^{15} |y_{Rj}^t - w_{REji}^t| \quad (1)$$

$$y_{Ei}^t = \begin{cases} 1, & i = i_0, \quad q_{Ei_0}^t = \min_i(q_{Ei}^t) \\ 0, & \text{others} \end{cases} \quad (2)$$

w_{REji}^t はR層からE層への結合度で、自己組織化マップの学習則[10][11]にしたがって変化する。本研究ではR層→E層の学習を、以降の層間の学習に先立って前学習として行った。そして80万ステップ計算した結果Fig.3のようになった。図中の番号は、モデル空間においてモデルラットが→方向を向いているとき、E層で発火した細胞の番号を示す。図中261番の細胞のように複数の場所に対応したのもも出現したが、おおむね似た

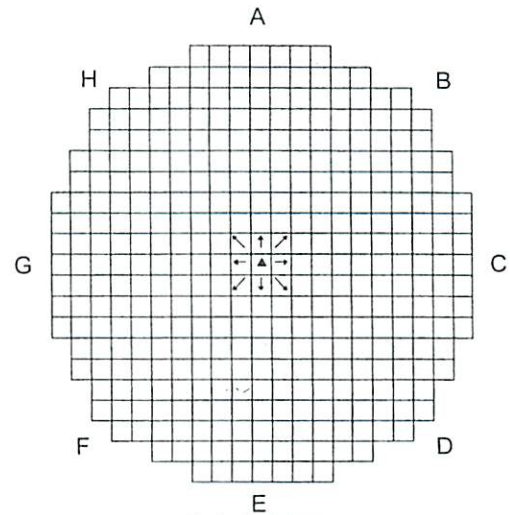


Fig.2 Model space.

ような見え(R層の入力パターン)はE層の同じ一つの細胞に対応するようになった。このようにR層→E層の学習を行い、E層の一つの細胞を1かたまりの見えに対応させた上でE層以降の学習を行った。

2.5 C3層

C3層は、E層から入力されたパターンを時間的系列にしたがって連合する層である。C3層は700個のニューロンから成り、 i 番目のニューロンの出力を y_{C3i}^t 、膜電位を q_{C3i}^t とおくと以下の式に従う。

$$q_{C3i}^t = \sum_{j=1}^{700} w_{ECji}^t y_{Ej}^t + y_{Ei}^t \quad (3)$$

$$y_{C3i}^t = \begin{cases} 1, & q_{C3i}^t > 0 \wedge i = i_0, i_1, \dots, i_M, \\ q_{C3i_0}^t = \max(q_{C3i}^t), \\ q_{C3i_1}^t = \max_{i \neq i_0}(q_{C3i}^t), \\ \dots, \\ q_{C3i_M}^t = \max_{i \neq i_0, i_1, \dots, i_{M-1}}(q_{C3i}^t) \\ 0, & \text{others} \end{cases} \quad (4)$$

すなわちC3層では、 q_{C3i}^t が0より上でかつ上から M (本研究では4)番目まで大きい値の q_{C3i}^t を持つ細胞が出力1を取り、残りは0を取る。ここで w_{ECji}^t はE層からC3層への結合度で、次式に従う。

$$w_{ECji}^{t+1} = \begin{cases} w_{ECji}^t + \alpha m_j^t (1 - w_{ECji}^t), \\ m_j^t y_{C3i}^t > 0 \\ w_{ECji}^t, & \text{others} \end{cases} \quad (5)$$

$$m_j^t = \begin{cases} M_{max}, & m_j^{t-1} = 0 \wedge y_{Ej}^t = 1 \\ m_j^{t-1} - 1, & m_j^{t-1} > 0 \\ 0, & \text{others} \end{cases} \quad (6)$$

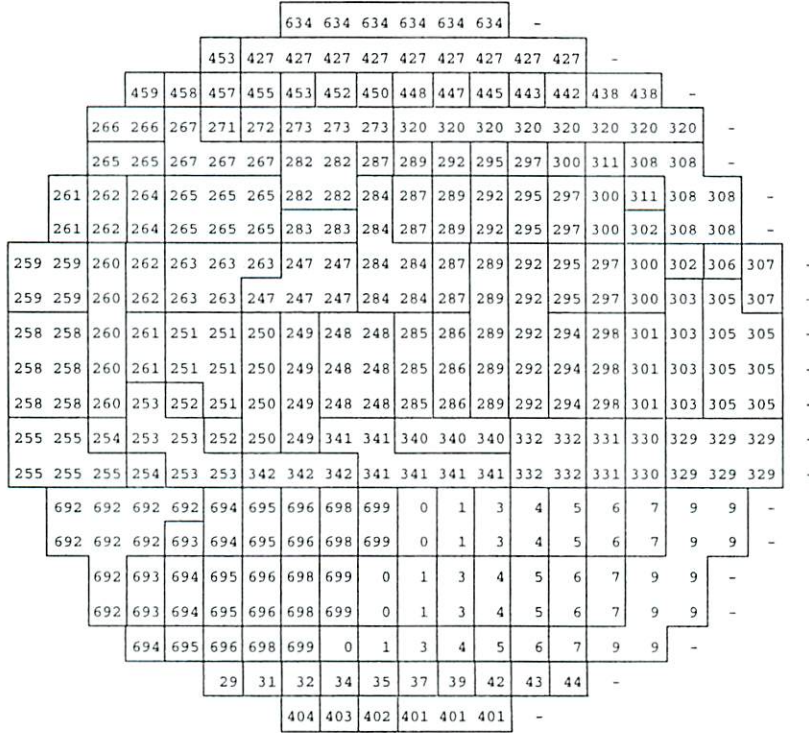


Fig.3 No. of the E-cell which reacted to the view of \rightarrow direction at each positions.

m_j^t はE層j番目の細胞の出力の履歴値で、 M_{mx} は正の整数値(本研究では6)である。これは重松ら[12]の提案した学習則を簡略化したものであり、例えば時刻 $t-1$ でE層j番目の細胞の出力の履歴値 m_j^{t-1} が0でかつ時刻 t で y_{Ej}^t が1になったときに m_j^t が M_{mx} になり、その後1ステップ毎に1ずつ引かれて最後には0に戻る。この m_j^t が0より上のものをHebb則の対象にすることにより時系列学習が可能となり、巡回回路による時系列学習[14]と同等の効果を有することになる。

これによりC3層においては、E層で発火した細胞を順次連合して行く。E層の各細胞は個々の見えに対応していたから、C3層は複数の見えを順に連合することになる。

場所細胞とは、ある空間内のある限られた領域において、ラットがどの方向に向いていようと発火する細胞のことであった。そのため本モデルにおいて、モデルラットの行動を通して空間内のある領域の全方向分の見えを選択し、選択した見えをC3層の細胞によって連合させれば、以降その細胞は場所細胞としての機能を持つであろう。本研究ではその点を狙って後の行動様式を設定している。

2.6 C1層

C1層は、競合学習によってC3層のパターンを分類する層である。場所細胞はすでにC3層においてみられると期待されるが、この層はC3層において発現されるパターンを直交化して結果を見易くしようというものである。

C1層は20個のニューロンから成り、時刻 t における i 番目のニューロンの出力を y_{C1i}^t 、膜電位を q_{C1i}^t とおくと以下の式に従う。

$$q_{C1i}^t = \sum_{j=1}^{700} w_{CCji}^t y_{C3j}^t \quad (7)$$

$$y_{C1i}^t = \begin{cases} 1, & i = i_0, \quad q_{C1i_0}^t = \max(q_{C1i}^t) \\ 0, & \text{others} \end{cases} \quad (8)$$

w_{CCji}^t はC3層からC1層への結合度で、競合学習を行い次々に従って変化する。

$$w_{CCji}^{t+1} = \begin{cases} w_{CCji}^t + \beta (y_{C3j}^t - w_{CCji}^t), & y_{C1i}^t = 1 \\ w_{CCji}^t, & \text{others} \end{cases} \quad (9)$$

以上のようにモデルを構築し、次に述べる行動様式にしたがってシミュレーションを行った。なお本研究では、パラメータは $\alpha = 0.002$, $\beta = 0.02$ として全ての計算を行った。

2.7 行動様式

本研究では行動が見えの時系列学習にどのような影響を与えるのかを調べる上で、その第一歩としてまず環境を固定して、実験者がモデルラットの移動のさせ方(行動様式)を何種類か用意し、その行動様式の違いによる場所細胞の働き方を調べた。行動様式は次の5種類を用意した(Fig.4)。

A. Random

モデルラットがフィールド内をランダムに1マスずつ移動する

B. Random+8

モデルラットがランダムに移動する度に周りを見渡す(8方向分の視覚情報を順次入力する)

C. Random+5S

5マス直進した後、ランダムに方向を選び、再び5マス進むということをくり返す

D. Octagon

一辺の長さが1~8マスの八角形をランダムに選び、1周するか壁にあたった時点で八角形を選びなおし、ランダムな方向から始める

E. Octagon+4

Dと同様に移動するが、一辺終わるたびに進行方向+左右2方向ずつを眺めてから次の一辺に進む

Aは、モデルラットがランダムに移動しても場所細胞が生成されるかどうかを調べるために設定した。Bは、ラットが新規の環境に置かれた場合に見られる行動を模したものである。Cは、Aで場所細胞の生成が見られた場合に、行動のどのような点が場所細胞生成の要因となったかを調べるための一つとして、直進的な行動の重要性を見るために設定した。DとEは実際のラットの動きを模したものであり、周りをとときとき見ているか見ていないかの違いを調べるために設定した。

以上の行動様式をそれぞれ別々にモデルに適用した。

3. シミュレーション結果

この章ではAからEの各行動様式を本モデルに適用した結果を示す。全体としては、まずモデルラットの行動様式の違いが場所細胞生成の有無を分けたことを示して、行動と時系列学習の関係を考察し、次に場所細胞が生成されても行動様式によってはその活動が不安定であったことを示してその考察を行う。

Fig.5は、C1層の細胞において見られる典型的な反応結果を示している。色の濃い領域ほど活動が高く、より多くの方向の見えと反応を有することを示す。計算量は、各行動毎に10万回である。Fig.5からわかるように、場所細胞としての反応を現すようになったのはRandom, Random+8, Octagon+4の3つの行動様式で、Random+5SとOctagonの行動様式に関しては場所細胞の生成は見られなかった。

ここで場所細胞の定義を本モデルにあわせてもう少し具体的にしよう。本モデルでは、モデル空間内において7 or 8方向の見えに反応したマス目を1つ以上40個以下まで持つC1層の細胞を場所細胞と定義する。このように定義して、各行動様式

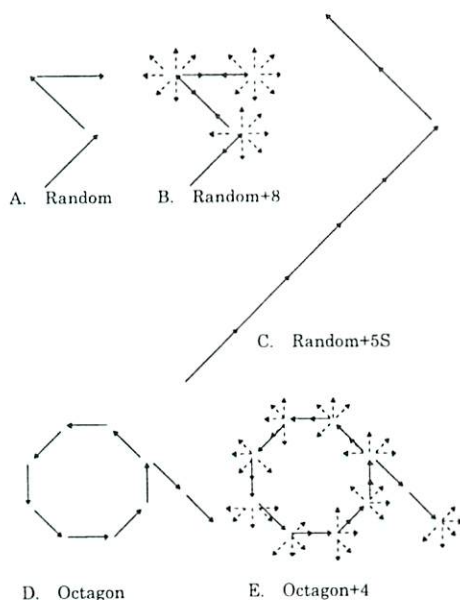


Fig.4 Movement patterns.

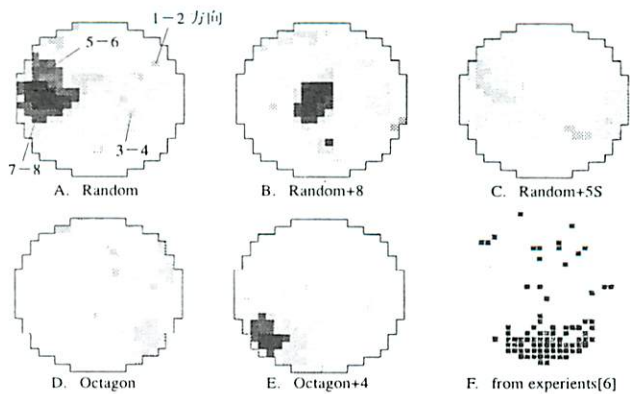


Fig.5 Examples of C1-cell reaction for each movement pattern.

毎に場所細胞の合計をプロットしたのがFig.6である。

Fig.6を見ると、行動様式によって場所細胞生成の有無がハッキリと別れたのがよく分かる。特にOctagonとOctagon+4は、進行方向に対して左右の見えをモデルに時々入力しているか否かの違いだけでも関わらず、それが場所細胞生成の有無を分けた。また、Random+5SとOctagonのように直進的な行動が多い場合には、モデルの細胞が場所情報を持つことは少ないが、仮に直進的な行動が多くても、Octagon+4のように適当な間隔で「周りを眺める」という行動があれば、場所細胞は十分に生成され得るということが分かった。

以上の結果を理解するために、ここで本モデルにおいて場所細胞が生成されるための要件を考えてみよう。C3層は時間的に近い順序で入力された見えの情報を連合するから、場所細胞の生成には、モデルラットはその短い時間の間に同じ位置で異なる方向をなるべく多く見るのが望ましい。そのように行動することによって、モデルラットが次に同じ位置に来たときに同じ細胞が発火する確率が高くなり、その位置でその細胞がまだ連合していない方向の見えも連合しやすくなる。こうして、ある位置の全方向分の見えが同じ一つの細胞に連合されて場所細胞になるのである。

そのために、同じ場所で異なる方向を見る機会が多いRandom, Random+8, Octagon+4のような行動では場所細胞が生成され、直線的に進んで周りを見ることのないRandom+5SやOctagonのような行動では場所細胞の生成が見られなかったと考えられる。

行動を通して見えを選択し、選択された見えをC3層で連合することによって場所細胞が生成されるという関係は、C1層において場所細胞ができるまでの過程を調べた結果からもうかがえる。Fig.7は、モデルラットがRandomとして行動したときの典型的な例を示している。一番上の段はモデルラットがモデ

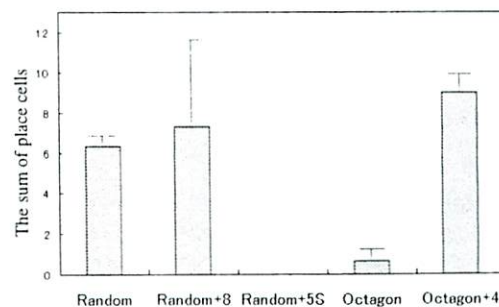


Fig.6 Movement patterns decide whether place cells are produced or not

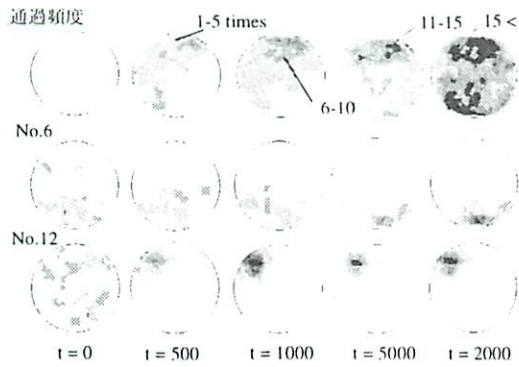


Fig.7 Simulation results of the place cell production process.

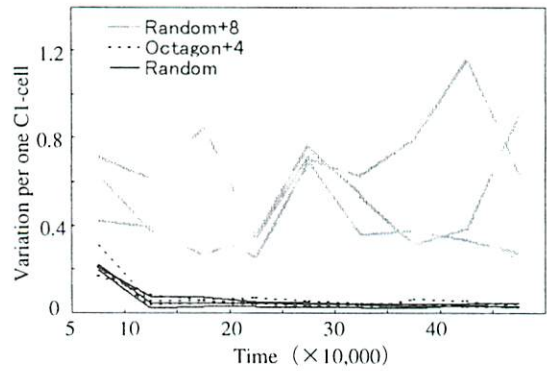


Fig.10 The process of variation per one C1-cell.

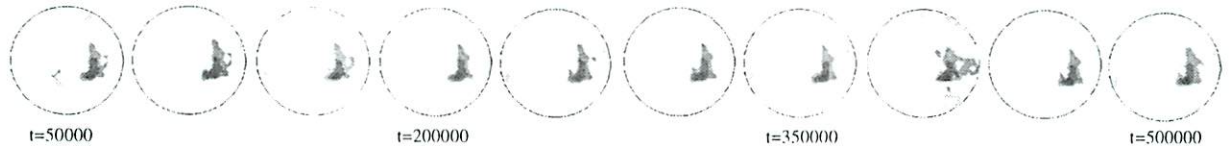


Fig.8 The time course of reaction on the C1-cell (Octagon+4).

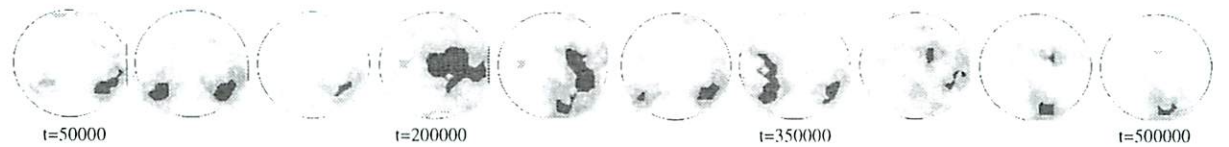


Fig.9 The time course of reaction on the C1-cell (Random+8).

ル空間内の各位置を通過した頻度の経時変化を示し、まん中と一番下の段はそれぞれC1層6番目と12番目の細胞の活動の経時変化を示している。

Fig.7に示したように、はじめの500ステップまでに個々の細胞のおおまかな場所フィールドが決められ、モデルラットがその領域を通過することによって正確に場所フィールドが確定されるというパターンが多く見られた。なお、この結果は実際のラットにおける場所細胞の生成過程によく似ている(例えば実験結果として[15]を参照)。

次に、Random、Random+8、Octagon+4において生成された場所細胞の安定性を調べた。その結果、RandomとOctagon+4で生成された場所細胞については計算中ずっと安定に存在し続けたが、Random+8については常に変化がみられ、場所細胞の生成消滅と、場所領域の変形等を繰り返した。その典型的な例をFig.8、Fig.9に示す。RandomとOctagon+4で生成された場所細胞は、Fig.8のようにその形をほとんど変えずに存在し続けた。しかしRandom+8では、Fig.9のように不規則に変化し続け、ときには場所フィールドが消滅したり、逆にモデル空間全体の8割以上を一つの細胞で占めたりした。

以上のような、モデルラットの取る行動様式の違いによる安定性の差は、C1層の細胞1個あたりの活動の変動量を調べると顕著であった。初期値をランダムに換えて各行動様式あたり3回計算をした結果をFig.10に示す。Fig.10を見るとOctagon+4とRandomの変動量は計算の初期を除いて低くとどまっていたが、Random+8の変動量は常に高く、その幅も大きかった。

Random+8で生成された場所細胞はなぜ不安定なのだろうか。結論から言うと、C1層の細胞同士が競合を行うためである。先の議論で、モデルラットが周りをよく見ればそれだけ場所細胞が生成されやすくなることが分かった。ここでRandom+8の場

合を考えると、モデルラットがある軌跡を辿ってモデル空間内のある位置に来たとき、Random+8の場合そこでひと回りし、結果としてC1層のいずれかの細胞がその位置で場所細胞を形成する。そしてモデルラットがその位置から1マス動いて再びひと回りしたとき、C1層の同じ細胞が発火する可能性も高いであろうが、別の細胞が発火して場所細胞を形成する可能性もある。こうして、次にモデルラットが同じ位置に来たときに、辿ってきた軌跡によって異なる場所細胞が発火するようになり、C1層の細胞同士で活動の競合が起きるのである。

逆にRandomやOctagon+4では、周りを見る機会がRandom+8ほど近接で起こることは少ないために、C1層で形成された場所細胞はうまく「住みわけ」し競合が起らないのであろう。

4. 強化学習への適用

ラットにおいて場所細胞は近道選択などに有効に利用されているに違いないが、実際にどのように役立てられているのかは分かっていない。ここでは本モデルの出力(C1層の出力)を状態として強化学習の一つであるQ-learning[16]に単純に当てはめ、見えを状態とする通常のQ-learningと比べて全体の計算量が少なくて済むことの実証を試みる。

本研究では1ステップQ-学習アルゴリズムを用いた。詳細は浅田ら[17]を参照いただきたい。なお行動選択は、現在の状態を s 、行動を a 、そのときの行動価値観数の値を $Q(s,a)$ とすると、 $Q(s,a)+\gamma$ の高い値を取る a を次の行動として選択することにした。 γ は乱数なので、あまり学習の進んでいない内は行動をランダムに選択することになる。

タスクとしては 5×5 の2次元空間を xy 座標軸上で考えた位置(1,1)を出発点として、(5,5)のゴールに9ステップ以内に辿り着くことを考えた。計算条件は次の2種類を用意した。

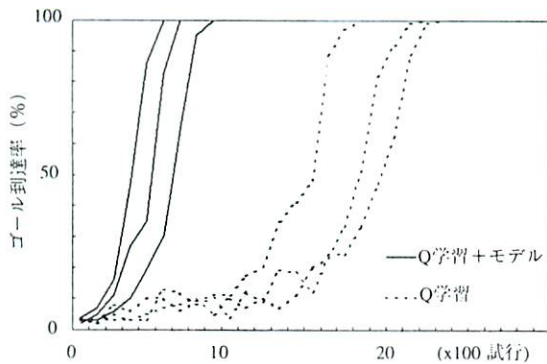


Fig.11 Learning curves.

(1) 通常のQ学習

フィールド上の各位置の各方向に対応した見えをそれぞれ一つの状態とみなし、絶対座標による(他者中心的な)8方向を行動の種類として持つQ学習を行う。

(2) Q学習+本モデル

フィールド上の各位置の各方向に対応した見えをそれぞれ本モデルE層の一つの細胞に対応するとし、C1層の各細胞をそれぞれ一つの状態とみなしてQ学習を行う。

Q学習のパラメータはどちらも同じ値を用いた。

(1)と(2)それぞれ3回ずつ計算した結果をFig.11に示す。Fig.11を見ると明らかなように、モデルを適用した方のQ学習は通常のQ学習に比べて2分の1から3分の1の試行数で収束した。またこのとき、C1層では場所細胞が形成されていたが(Fig.12)、1回の計算当たりC1層の1つか2つの細胞が場所を表現するようになる程度であり、場所細胞でモデル空間全てを埋めるほどの数は生成されていなかった。

以上の結果を解釈すると、Q学習+本モデルでは、 $Q(s,a)$ の更新と場所細胞の生成が同時に進む。言い換えれば空間内を探索しながらモデル内で空間の表現を場所細胞として縮約していくので、学習が進むにつれてスタート地点とゴールの間の距離がモデル内で「縮む」のである。そのため更新すべき $Q(s,a)$ の数が少なくなり、より速く学習が収束するのであろう。

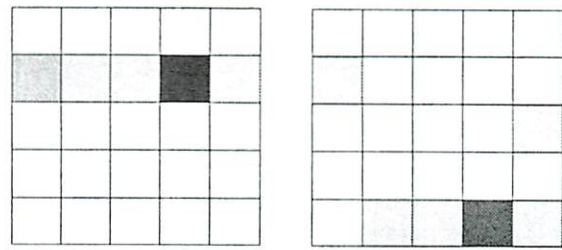
しかし形成された場所細胞の数が少なかったことから分かる通り、状態空間の縮小が最適な形で行われたわけではなかった。実際の環境ではゴールの位置が他の位置に移動するなどの変化があることを考えると、場所細胞が空間全体に均一に分布するのが適切と想像されるが、本モデルをQ学習に単純に当てはめただけではそのような分布が生じることはなかった。

5. まとめと今後の課題

以上により、視覚情報としての見えの情報のみを入力として、それにモデルラット自身の行動と海馬体サーキットを近似した時系列学習を組み合わせることにより、ラット海馬に見られる場所細胞の生成過程を従来までよりも自然な形で説明することができた。また、本モデルを強化学習に単純に適用した例から、空間学習における場所細胞の有効性が示された。

場所細胞は自身の行動パターンにあわせて視覚情報を適度に統合するため、本モデルのような場所細胞生成機構を利用すれば、移動ロボット等における実環境の学習が従来までよりも格段に速く収束すると考えられる。

しかしながら、強化学習への適用において場所細胞の有効性



a) C1層66番目の細胞の活動 b) C1層128番目の細胞の活動

Fig.12 Examples of C1-cell reaction (Q-learning+model).

が示されたとはいえ、生成された場所細胞の分布パターンが環境に対して最適なパターンとは言えなかった。本モデルは自身の行動に応じて場所細胞の安定性を左右できたことから、移動ロボットへの適用を考えた場合、今後は、変化して止まない実環境にあわせて場所細胞が柔軟に分布するような行動システムを構築する必要がある。また、単純に強化学習に適用しただけでは、ゴール地点が移動するたびに学習をやり直さなければならぬ。この点もあわせて新しい行動・学習システムを考えなければならないであろう。

文献

[1] J. O'Keef and J. Dostrovsky, "The hippocampus as a spatial map preliminary evidence from unit activity in the freely moving rat," *Brain Res.*, vol.34, pp.171-175, 1971.

[2] D. Zisler, "A computational model of hippocampal place fields," *Behav. Neurosci.*, vol.99, pp.1006-1018, 1985.

[3] P.E. Sharp, "Computer simulation of hippocampal place cells," *Psychobiology*, vol.19, pp.103-115, 1991.

[4] N.A. Scmajuk, and H.T. Blair, "Stimulus configuration, spatial learning, and hippocampal function," *Behav. Brain Res.*, vol.59, pp.103-117, 1993.

[5] M.L. Shapiro, and P.A. Hetherington, "A simple Network model simulates hippocampal place fields: parametric analysis and physiological prediction," *Behav. Neurosci.*, vol.107, pp.34-50, 1993.

[6] N. Burgess, M. Recce, and J. O'Keef, "A model of hippocampal function," *Neural Networks*, vol.7, pp.1065-1081, 1994.

[7] B.L. McNaughton, C.A. Barnes, J.L. Gerrard, K. Gothard, M.W. Jung, J.J. Knierim, H. Kudrimoti, Y. Qin, W.E. Skaggs, M. Suster, and K.L. Weaver, "Deciphering the hippocampal polyglot: the hippocampus as a path integration system," *J. Exp. Biol.*, vol.199, pp.173-185, 1996.

[8] A. Samsonovich, and B.L. McNaughton, "Path integration and cognitive mapping in a continuous attractor neural network model," *J. Neurosci.*, vol.17, pp.5900-5920, 1997.

[9] 加藤 宏司, 伊藤 志, 上藤 佳久, "可塑性-海馬のシナプス長期増強を中心にして," in *神経細胞が行う情報処理とそのメカニズム-脳とコンピュータ3*, 松本 元, 大津 展之 編., pp.89-138, 培風館, 東京, 1991.

[10] T. Kohonen, "Self-organized formation of topographically correct feature maps," *Biological Cybernetics*, vol.43, pp.59-69, 1982.

[11] T. Kohonen, "Self-organization and associative memory," 3rd ed., Springer-Verlag, New York, 1989.

[12] 重松 征史, "神経回路素子と学習・記憶," in *脳・心・コンピュータ*, 日本物理学会 編., pp.173-193, 丸善, 東京, 1996.

[13] R. Morris, "Developments of a water-maze procedure for studying spatial learning in the rat," *J. Neurosci. Meth.*, vol.11, pp.47-60, 1984.

[14] 森田 昌彦, "連想記憶の海馬モデル," *信学論(D-II)*, vol.J72-D-II, pp.279-288, 1989.

[15] A.J. Hill, "First occurrence of hippocampal spatial firing in a new environment," *Exp. Neurology*, vol.62, pp.282-297, 1978.

[16] C.J.C.H. Watkins, "Learning from delayed rewards," PhD thesis, King's College, University of Cambridge, May, 1989.

[17] 浅田 稔, 野田 彰一, 依積田 健, 細田 耕, "視覚に基づく強化学習によるロボットの行動獲得," *日本ロボット学会誌*, vol.13, pp.68-74, 1995.