神経修飾を用いたニューラルネットワークによるメタ記憶の進化

Creating metamemory by evolving artificial neural network with neuromodulation

大和 祐介	鈴木 麗璽	有田 隆也
Yusuke Yamato	Reiji Suzuki	Takaya Arita

名古屋大学 大学院情報学研究科

Graduate School of Informatics, Nagoya University

Human have the ability to consciously aware of the presence or absence of their memory. This kind of the ability called metamemory plays important roles in human cognition. We aim to evolve artificial neural networks with neuromodulation, that have a metamemory function. Our constructive approach is based on the repetition of evolutionary experiments, analysis of the evolved networks and refinement of the measure, so as to reducing the gap between the functional properties of behavior and subjective reports of phenomenal experience. This paper reports on the current state of the approach based on the evolutionary experiments with a delayed matching-to-sample task.

1. 序論

本研究の目的は構成論的アプローチによってメタ記憶能力 を持つニューラルネットワークを進化させることである. ヒト は自分の記憶に特定の知識が存在するか,その記憶がどれ程の 確かさであるかを認知できる. このような自らの記憶に対す る主観的な認知能力であるメタ記憶は,ヒト固有の特徴と見 なされており [1],メタ認知の一種として広く研究されてきた. 我々は,この目的のために,Evolutionary Plastic Artificial Neural Networks (EPANN) [2] を用いる. これは,可塑的な ニューラルネットワークを進化によって創発させることによ り,学習システムをデザインし,創り出す計算論的アプローチ であり,生物の知性が進化,発達,学習の相互作用によって創 発したという事実にそのアイディアの源泉がある.

本研究は認知・行動プロセスにおける必須要素として知ら れている神経修飾を使用する.これは修飾ニューロンによって ニューロン間結合の可塑性を動的に変更するものである[3].さ らに、ある種の進化戦略[4]を用いて、神経修飾を用いたニュー ラルネットワークのトポロジーと結合重みを進化させることに より、ターゲットとする認知機能、メタ記憶を自動的に生成す ることを狙う.すでに我々は予備的実験により、そのようなア プローチによって進化させたネットワークがサルを対象に開発 された遅延見合わせ課題(DMTS)に基づくメタ記憶の行動 主義的な基準を満たすことを示している[5].

しかし、メタ記憶はそれほど単純なものではない [6]. なぜ ならば、動物の行動を観察するだけで自分の記憶を監視できる と結論づけることは非常に難しいからである.たとえ、ヒトを 用いた被験者実験でさえも、最新の技術を用いて脳活動を計測 するだけで主観的な認知能力を同定することは簡単ではない. そもそも、その難しさはメタ記憶の定義自体にも依存する.た だし、我々のアプローチのように人工的なニューラルネット ワークの進化を用いる場合は、生物を扱う場合とは違い、原理 的にはニューラルネットワークで起こるすべてのプロセスのメ カニズムを理解することができるという強みがある.

本研究は, 先行研究 [5] を出発点とし, 進化させたネットワー

クを批判的に分析し、メタ記憶とは異なるようなメカニズムに 基づいているネットワークが進化することを排除するようにメ タ記憶の基準を改訂する.このような進化実験、進化したネッ トワークの分析、メタ記憶の基準の改訂を繰り返していくこと により、真のメタ記憶の実現に近づけていく.

2. 方法論

2.1 メタ記憶の尺度

Hamptonは、アカゲザルを対象とした遅延見合わせ課題を 考案した.動物に対するメタ記憶実験には回避応答パラダイ ムと情報探索パラダイムがあり、彼の実験は前者に属する.ま ず、サルはあるイメージを記憶する.その後、設定した遅延時 間を経て、テスト(4つのイメージから記憶した1つを選択す る)を受けるか回避するかを選択する.ただし、一定の確率で 選択させずに強制的にテストを受けさせる.テストでは記憶し たものと同じイメージを選択した場合、高品質の報酬を受け取 るが、それ以外の場合は何も受け取れない.回避した場合には 中程度の報酬を受け取る.メタ記憶を持つならば、選択して受 けた場合と強制されて受けた場合の正解率を比較すると前者の 方が高く、また、遅延時間の増加に従ってその差は大きくなる はずである.Hampton はサル1匹がこの基準を満たすことを 示した [7].

近年,動物の認知に関する研究が盛んであり,様々な議論が なされている.特にメタ記憶に関する動物実験に対する批判は 記憶しているかどうかの主観的な認知をヒトを使った実験での アンケート調査等によらずに,動物の行動のみから推測してい る点にある.生物の主観的な体験ではなく,メタ記憶を持って いるならばするはずの行動を調べている点である.Hampton の実験においても記憶の有無に関わらず難しいイメージを固定 的に回避して報酬を最大化することでクリア可能と言える[6] [8].我々は,行動の機能と主観的な体験の間に存在するギャッ プを埋めるため,メタ記憶の定義,進化実験,進化したネット ワークの分析,基準の改訂,…という反復に基づく構成論的ア プローチにより真のメタ記憶を持つニューラルネットワークの 進化を目指す.

行動基準

Hampton の実験パラダイムで、選択条件と強制条件のテ

連絡先: 大和祐介, 名古屋大学 大学院情報学研究科 複雑系科 学専攻, 〒 464-8601 名古屋市千種区不老町, 080-2712-4721, yamato@alife.cs.is.nagoya-u.ac.jp

ストの正答率に差が生じる.

ニューラル基準1
行動基準を満たすが、これが入力のみに基づくものではない。

以降,この定義に基づき進化実験と分析を行う.

2.2 タスク

進化実験に用いるタスクは, Hampton の実験 [7] に基づく. まず、個体は学習フェーズで5桁のビットからなるパターン (00001,00010,00100,01000,10000)のうち1つを正解とし て受け取る.次に,遅延フェーズでノイズとして複数回 00000 のパターンを受け取る.その後,選択フェーズへと移行する. この時, 確率 2/3 で選択条件, 確率 1/3 で強制条件となる. 選 択条件では,個体はテストの選択を指示する入力を受けて,受 けるか回避するか選択する.回避を選択すると得点 0.3 を受け 取る.一方,強制条件の場合,テストフェーズへ移行する.テ ストフェーズでは、全パターンがランダムな順に入力される. 正解と同じパターンが入力された時に解答を選択した場合,個 体は得点 1.0 を得る. それ以外の場合は得点を得られない. な お、学習フェーズでは確率的に正答不能パターン(00000)を 含める.これは、回避選択が最適行動である状況を作り出して 回避選択能力の進化を促すためである. 遅延量 N_{delay} の大き さはべき分布に従う乱数は以下の式(1)で定める.

$$N_{delay} = \lfloor \frac{-1}{\lambda \times \ln\left(R\right)} \rfloor + 1, \qquad (1)$$

ここで,λは分布の形状に関するパラメータ,*R*は [0,1] の範 囲の一様乱数である.以上のタスクを複数回行い,得点の合計 値を個体の適応度とする.



図 1: Hampton の実験パラダイムに基づくタスクの流れ.

2.3 モデル

個体は神経修飾を導入したニューラルネットワークを持つ. ニューロンとして、7つの入力ニューロン以外は、可変個数の 標準ニューロン(内、3つが出力ニューロン)と修飾ニューロ ンである.入力は、パターンの各ビット、選択条件の指示、バ イアスであり、それぞれガウシアンノイズ(σ,μ)を加算した ものが入力される. 修飾ニューロンは, 標準ニューロンとは異 なり, ニューロンの出力値をネットワークの結合重みの更新率 とすることにより, 結合重みの更新の可塑性(学習率)を動的 に変更する. 標準ニューロン *i* の活性 *a_i* は式 (2), 修飾ニュー ロン *i* の活性 *m_i* は式 (3) によって計算する.



図 2: ニューラルネットワークの入力と出力.

$$a_i = \sum_{j \in Std} w_{ji} \cdot o_j \tag{2}$$

$$m_i = \sum_{j \in Mod} w_{ji} \cdot o_j, \tag{3}$$

ここで, w はシナプス前ニューロン j とシナプス後ニューロン i の結合重み, Std および Mod はニューロン i に結合している 標準ニューロンと修飾ニューロンの集合である. o_j はニューロ ン j の出力であり, $o_j = \tanh(a_j)$ として計算される. ニュー ロン $j \ge i$ 間の結合重みは以下の式によって更新する.

$$\Delta w_{ji} = \tanh m_i \cdot \eta \cdot (Ao_j o_i + Bo_j + Co_i + D), \quad (4)$$

o_j と o_i はシナプス前ニューロン j とシナプス後ニューロン i の出力である. η, A, B, C, D は進化可能なパラメータで あり,進化によって多様な更新規則を表現することができる. ネットワークの構造や結合重みは進化戦略により進化させ る.アルゴリズムは基本的に Soltoggio らの手法 [3] と同じで ある.遺伝子座として,ニューロン数,ニューロンの種類, 2 次元の結合重み行列,結合重みの更新規則に関する 5 パラメー タを持つ.結合重みとパラメータ A, B, C, D は [-1, 1], η は [-100, 100] の範囲で進化する.遺伝子型 GhV の結合重み は,式 (5) によって初期表現型の結合重み PhV に変換される.

$$PhV = \begin{cases} 0 & (|GhV^{3}| < 0.01) \\ 10 \cdot GhV^{3} & (otherwise). \end{cases}$$
(5)

親個体は、空間的な局所性を考慮したトーナメント選択によっ て選ばれる.各個体は配列に格納され、ランダムなオフセット を伴ったサイズ5の連続的なセグメントに区切られる.その 後、各セグメントの最大適応度個体が親として選択され、0.1 の確率で交叉、それ以外は自身のコピーを子とする.交叉は、 親集団からランダムに選択される.親同士の結合重みの部分 行列と対応するニューロンを交換して子を生成する(交叉点 は結合重み行列の各軸上からランダムに選択).また、更新規 則パラメータに対しては一様交叉を適用する.突然変異率は 0.1 とし,結合重みと更新規則パラメータに対して $\mu = 0.0$, $\sigma = 0.3$, (η のみ $\mu = 0.0$, $\sigma = 3.0$)のガウシアンノイズを加 算する.ニューロンの挿入,複製,削除はそれぞれ 0.04, 0.06, 0.02の確率で適用される.追加されたニューロンの種類は 0.5 ずつの確率で決定され,結合重みは [-1, 1]の範囲で初期化さ れる.このようにして N 個体が 1 世代として生成され,G世 代繰り返される.

3. 実験結果

個体数 N = 300, 世代数 G = 1000, 課題数 T = 300, 正 答不能課題数 U = 20, N_{delay} のパラメータ $\lambda = 0.7$, 最大 ニューロン数を 16 とする. 各試行は 300 回のタスクのうち, 正当不可能課題を 20 回含む. 初期世代の各エージェントのパ ラメータはタスクの開始時に可能な範囲内でランダムな値に設 定される.

この設定で進化実験を 20 試行行った結果,17 試行で行動基 準を満たす個体が進化した.分析の結果,典型的な 2 種類(タ イプ A とタイプ B)の個体を発見した.タイプ B 個体は 17 試行全てで進化したが,タイプ A 個体は 17 試行中 8 試行のみ 進化した.図 3,4 はそれぞれタイプ A,タイプ B の個体の遅 延変化による得点の変化を表す.両タイプとも強制条件より選 択条件での正答率が高い.つまり行動基準を満たす.しかし, タイプ A は遅延時間の増加に従い回避選択率が増加したのに 対し,タイプ B はほとんど変化がない.



図 3: タイプ A の典型的な振る舞い (reward: 217.7).



図 4: タイプ B の典型的な振る舞い (reward: 220.4).

次に、タイプAとタイプBの振舞いの違いを調べた.表1, 2は、各パターン別の強制条件、選択条件、及び回避タスクの 平均正答率を表す.回避タスクの正答率は回避したテストを 仮に解答していたとして算出した.遅延は各パターン 40 とし た.タイプA は受け取るパターンに関わらず,強制条件より も選択条件の方が正答率が高く,回避タスクの正答率も低い. これは,正答できないテストを回避していることを表してお り,ニューラル基準1を満たす.一方,タイプB はパターン 01000 が入力された場合のみ回避行動を多用し,その他のパ ターンでは全く回避を行わなかった.これは,特定の刺激に対 して回避を関連付けるような回避応答パラダイムにおいて指摘 された振舞いであり,行動基準は満たすがニューラル基準1は 満たさない個体に典型的な振る舞いである.ただし一方で,パ ターン 01000 についてはタイプA と同様に記憶を忘却した課 題に対して回避を選択する傾向があった.つまり個体はニュー ラル基準1を満たすものであるが,特定パターンのみに対し てのみの振る舞いで満たしていると言える.

表 1: 各パターンに対するの正解率に関するデータ (タイプ A).

Bit patterns	Accuracy in forced tests	Accuracy in chosen tests (probability of taking)	Accuracy in declined tests (probability of declining)
00001	0.671	0.825(0.756)	0.291 (0.244)
00010	0.667	0.825(0.800)	0.172 (0.200)
00100	0.663	0.788(0.753)	0.302(0.247)
01000	0.689	0.809(0.799)	0.287(0.201)
10000	0.714	0.814(0.787)	0.203(0.213)

表 2: 各パターンに対するの正解率に関するデータ (タイプ B).

Bit patterns	Accuracy in forced tests	Accuracy in chosen tests (probability of taking)	Accuracy in declined tests (probability of declining)
00001	0.843	0.823 (1.0)	-(0.0)
00010	0.840	0.835 (1.0)	-(0.0)
00100	0.794	0.789(1.0)	-(0.0)
01000	0.448	0.729(0.436)	0.294(0.564)
10000	0.758	0.770(1.0)	-(0.0)

次に,タイプAのネットワーク構造と振る舞いを分析する. 図5で典型的なタイプAのネットワーク構造を示す.ネット



図 5: タイプ A のニューラルネットワーク.

ワークの挙動の観察により,図6のようにパターンの各ビッ ト入力ニューロンと特定のニューロン間の結合重みでターゲッ トパターンの構造が保持されていることがわかった.具体的に は、ターゲットパターンの1であるビットに対応する結合重 みのみが正となっており、テストフェーズにおいてこの結合重 みが入力に対するマスクとして働く.その結果、ターゲットパ ターンが入力された場合に解答を選択する.



図 6: ターゲットパターンの構造を反映する結合重み.

また, 忘却は次のようなメカニズムによることがわかった. 記憶が保たれている状態では, テスト回答ニューロンが受け取 る修飾ニューロンからの出力の総和は0で, 記憶を反映する 結合構造は変化しない. ノイズにより修飾ニューロンの活性に 変化が生じて, テスト回答ニューロンが受け取る修飾ニューロ ンからの出力の総和が0以外に変化すると結合重みが更新さ れる. この更新量が大きい場合に記憶の忘却が生じる.

図7は、ネットワークの発展による時間結合重みの変化を 表す. 同図より、忘却が発生すると同時に周辺の結合重みも変 化することで回避を選択していることがわかる. これは、修飾 ニューロンによる修飾先の同時更新によるものであり、記憶の 忘却原因を引き金とした受動的な性質だと考えられる. つま り、記憶のモニタのような能動的な機能として、自身の記憶を 自己参照しているものではない.

4. 結論

本稿では神経修飾モデルを用いた進化実験を行ってメタ記憶 の進化を調べた.我々は、メタ認知を扱う回避応答パラダイム に基づく実験の議論において指摘を受けた振る舞いを排除する ために2つの基準(行動基準とニューラル基準1)を定義した. また、進化実験を行い2種類の代表的な個体を発見し、どち らもニューラル基準1を満たすことを確認した.さらに、ネッ トワークの構造とダイナミクスを解析し、メタ記憶の機能を獲 得したネットワークの記憶の保持と忘却、記憶の忘却時にテス トの回避を選択するメカニズムを明確にすることで、ニューラ ル基準1を満たす個体の振舞は自身の記憶の参照に基づくも のではないとわかった.

参考文献

- Endel Tulving and Stephen A Madigan. Memory and verbal learning. *Annual review of psychology*, Vol. 21, No. 1, pp. 437–484, 1970.
- [2] Andrea Soltoggio, Kenneth O Stanley, and Sebastian Risi. Born to learn: The inspiration, progress, and future of evolved plastic artificial neural networks. *Neural Networks*, Vol. 108, pp. 48–67, 2018.



図 7: ネットワークの結合重みの時間発展. ノイズにより修 飾ニューロンの活性が変化しする. これにより, ビット入力 ニューロンと標準ニューロンの間の記憶を保持する結合重みが 正から負へと変化 (A). これは記憶の忘却に相当. それと同時 に, その他の結合重みも反転 (B). その結果, 選択フェーズで 回避選択を決定する標準ニューロンが活性化し回避を選択.

- [3] Andrea Soltoggio, John A Bullinaria, Claudio Mattiussi, Peter Dürr, and Dario Floreano. Evolutionary advantages of neuromodulated plasticity in dynamic, rewardbased scenarios. In *Proceedings of the 11th international conference on artificial life (Alife XI)*, No. LIS-CONF-2008-012, pp. 569–576. MIT Press, 2008.
- [4] Thomas Bäck, David B Fogel, and Zbigniew Michalewicz. Handbook of evolutionary computation. CRC Press, 1997.
- [5] Masaru Sudo, Reiji Suzuki, and Takaya Arita. Can agents with neuromodulation know when they remember? Proc. of the 19th International Symposium on Artificial Life and Robotics, pp. 330–334, 2014.
- [6] Josep Call. Do apes know that they could be wrong? Animal cognition, Vol. 13, No. 5, pp. 689–700, 2010.
- [7] Robert R Hampton. Rhesus monkeys know when they remember. Proceedings of the National Academy of Sciences, Vol. 98, No. 9, pp. 5359–5362, 2001.
- [8] Jonathon D Crystal and Allison L Foote. Metacognition in animals. *Comparative cognition & behavior reviews*, Vol. 4, pp. 1–16, 2009.