3B2-R-2-1

CMA-TWEANNを用いたニューラルネットの構造進化および結合強度最適化に基づく政策探索

Policy Search with CMA-TWEANN: Efficient Topological and Parametric Evolution of Artificial Neural Networks

森口博貴[*1] 本位田真一[*1*2]
Hirotaka Moriguchi Shinichi Honiden

[*1] 東京大学 [*2] 国立情報学研究所
The University of Tokyo National Institute of Informatics

Neuroevolutionary algorithms are successful methods for optimizing neural networks, especially for learning a neural policy (controller) in reinforcement learning tasks. Their significant advantage over gradient-based algorithms is the capability to search network topology as well as connection weights. CMA-TWEANN, a state-of-the-art topology evolving method, is known to be efficient compared to other existing methods. However, its efficiency was only validated on a simplistic benchmark task. This paper tries to validate its scalability on a more challenging pendulum swing-up task.

1. 序論

進化型ニューラルネットは適応度関数を最大化（最小化）するように，進化計算を用いてニューラルネットのトポロジや結合強度を最適化する手法である [Yao 99, Floreano 08]．また見方を変えれば，適応度関数を通じて最適関数の近似を行なっていると考えることも出来る．たとえばニューラルネットを学習エージェントの政策として用い，その政策の評価値を適応度関数とすれば，強化学習における最適政策の近似，すなわち政策の直接探索が行える [Whiteson 06, Nolfi 04]．人工知能分野における多くの課題が関数近似問題に帰着されるため，進化型ニューラルネットは幅広い応用範囲を持つと考えられる．

我々が提案したCMA-TWEANN [Moriguchi 12] は，ニューラルネットの結合強度だけでなくトポロジも同時に最適化するような，トポロジ進化型ニューラルネットの一つである．倒立振子の安定化ベンチマークによる実験では，CMA-TWEANNが既存の進化型ニューラルネットの中で最も少ない試行回数で倒立振子を安定化させることが示されており，その探索の効率性に期待が持たれている．しかし倒立振子の安定化は，極めて小型のトポロジを持つニューラルネットでも，適切な結合強度最適化が為されれば可能であり，トポロジ進化型ニューラルネットが必須なタスク—より複雑なトポロジを要するタスク—へのスケーラビリティを有するか否かは明らかでない．

本研究では，倒立振子タスクより解くことが困難な二重振子の振り上げ・安定化タスクを用いて，CMA-TWEANNのスケーラビリティを検証する．関節トルクが小さい二重振子の振り上げタスクにおいては，初めに振子を振ることで振り上げに必要な運動量を蓄積させなければならない．このようなタスクは従来のロボット制御手法では，系の運動方程式などといったドメイン知識無しで解くことが困難であることが知られている [Atkeson 08]．そのため，ドメイン知識が無い場合には，進化型ニューラルネットを用いた直接的な政策探索が有用であると考えられる．

実験から，CMA-TWEANNが大きなトポロジを必要とするタスクにおいても，効果的に機能することを示す．

連絡先: 森口博貴，東京大学情報理工学系研究科コンピュータ科学専攻，hmori@nii.ac.jp

2. CMA-TWEANNの概要

本節では，CMA-TWEANN の動作について概説する．CMA-TWEANN は CMA-ES (Covariance Matrix Adaptation Evolution Strategy [Hansen 01]) と呼ばれる関数最適化手法を用いて，結合強度最適化を効率的に行う．それと同時に，トポロジについても探索が行われるよう設計されている．CMA-TWEANN はリカレント型ニューラルネットを対象として進化させる．

以下ではまず初めに，CMA-ESを用いて固定トポロジのネットワークの結合強度を最適化する方法について説明し，その後CMA-TWEANN の概要を述べる．

2.1 CMA-ESを用いた結合強度最適化

最大化すべき適応度関数 $f:\Psi\to\mathbb{R}$（ここで Ψ は全てのニューラルネットの集合）が与えられたとき，進化型ニューラルネットは以下のように定義される最適なニューラルネット ν^* を探索する，

$$\nu^* = \operatorname*{argmax}_{\nu\in\Psi} f(\nu). \tag{1}$$

固定トポロジのネットワークの結合強度を最適化する場合は，式 (1) の最適化は以下の関数最適化に帰着できる：

$$\boldsymbol{x}^* = \operatorname*{argmax}_{\boldsymbol{x}\in\mathbb{R}^N} f(g(\boldsymbol{x})), \tag{2}$$

ここで，$\boldsymbol{x}$ は結合強度パラメタのベクトル，$g(\boldsymbol{x}):\mathbb{R}^N\to\Psi$ は結合強度パラメタから，そのパラメタで特徴付けられたニューラルネットへの写像である．$f\circ g:\mathbb{R}^N\to\mathbb{R}$ を，パラメタベクトル $\boldsymbol{x}$ 上に定義された適応度関数と考えれば，式 (2) の最適化は実関数上の最適化となり，Nelder-Mead 法 [Nelder 65] や CMA-ES などの関数最適化アルゴリズムを用いて解くことができる．

CMA-ESによる関数最適化では，探索点の近傍にランダムに子個体を生成し，探索点を移動させていく．CMA-ESの特徴は，近傍探索に用いられる探索分布 $\sigma\mathcal{N}(\boldsymbol{0},\boldsymbol{C})$ のステップサイズ σ および共分散行列 $\boldsymbol{C}$ を，自己適応させる点にある．特に共分散行列の自己適応によって CMA-ES はパラメタ間に相関がある場合に効率的な探索ができる．式 (2) で定義されるニューラルネットの結合強度最適化では変数間に相関があるため，CMA-ES はニューラルネットの結合強度を非常に効率的に探索で

きることが知られている．CMA-ES を用いた結合強度最適化は CMA-NeuroES と呼ばれる [Igel 03, HeidrichMeisner 09]．CMA-NeuroES は効率的な探索が可能である一方，固定のトポロジを仮定しているためトポロジの探索は出来ない．そのため，ユーザは事前に適切なトポロジを知り，アルゴリズムに与える必要がある．CMA-ES の詳細な説明は本研究の範囲を超えるため，アルゴリズムの詳細などについては文献 [Hansen 01] を参照されたい．

2.2 CMA-TWEANN

CMA-NeuroES は，固定トポロジを前提としているため，トポロジの進化・探索ができないという課題を持つ．CMA-TWEANN は，この課題を解決するためトポロジの拡張方法を導入し，結合強度とトポロジの同時探索を可能にする．CMA-TWEANN の擬似コードを Algorithm 1 に示す．アルゴリズムは最小のトポロジ，すなわち隠れ層のない完全結合リカレント型ニューラルネットから探索を開始する．以下では CMA-TWEANN の動作を概説するが，詳細については文献 [Moriguchi 12] を参照されたい．

Algorithm 1 CMA-TWEANN

最小のトポロジを設定
結合強度ベクトル：$\langle \boldsymbol{x} \rangle_{\mathrm{w}} \leftarrow \mathbf{0}$
共分散行列：$\boldsymbol{C} \leftarrow \boldsymbol{I}$
ステップサイズ：$\sigma \leftarrow \sigma_{init}$
while 終了条件が満たされない **do**
  for $i = 1 \rightarrow \lambda$ **do**
    $\boldsymbol{x}_i \leftarrow \langle \boldsymbol{x} \rangle_{\mathrm{w}} + \sigma \boldsymbol{z}_i (\sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}))$
  end for
  $\boldsymbol{x}_{i..\lambda}$ の評価値に基づいて共分散行列 $\boldsymbol{C}$ を更新
  $\boldsymbol{x}_{i..\lambda}$ の評価値に基づいてステップサイズ σ を更新
  $\langle \boldsymbol{x} \rangle_{\mathrm{w}} \leftarrow \sum_{i=0}^{\mu} w_i \boldsymbol{x}_i$
  if AddNode あるいは AddEdge オペレータの適用 **then**
    トポロジを拡張
    探索点 $\boldsymbol{x}$ を新たな探索空間に写像
    共分散行列 $\boldsymbol{C}$ を拡張
    探索次元数に依存するハイパーパラメタを更新
  end if
  $\sigma < \sigma_{min}$ の場合 $\sigma = \sigma_{min}$ に調整
end while

CMA-ES の一世代は共分散行列とステップサイズパラメタの更新で完了するため，CMA-TWEANN ではこの更新の完了後に二種類のトポロジ突然変異オペレータを適用する．これらのオペレータの動作を図 1 に示す．

トポロジ変異の発生は確率的に決定される．AddNode オペレータは $P(\alpha_n)$ の確率で，AddEdge オペレータは $P(\alpha_e)$ の確率でそれぞれ適用され，$P(\alpha_n \cap \alpha_e) = 0$ である．$P(\alpha_n)$ および $P(\alpha_e)$ をそれぞれ零に設定すれば CMA-NeuroES と同等のアルゴリズムとなることに注意されたい．追加される辺の数 ΔN は，AddNode では $\Delta N = 2$ に，AddEdge では $\Delta N = 1$ になる．

トポロジが拡張される際に，現在の探索点も N 次元の探索空間から $N + \Delta N$ 次元空間に以下のように写像される必要がある：

$$\langle \boldsymbol{x} \rangle^{(g)} = \boldsymbol{G} \langle \hat{\boldsymbol{x}} \rangle^{(g)},$$

ここで $\langle \boldsymbol{x} \rangle^{(g)}$ は写像先の点，$\langle \hat{\boldsymbol{x}} \rangle^{(g)}$ は写像元の点，そして $\boldsymbol{G}$ は (i, i) 要素が 1 で，それ以外の要素が零であるような $(N + \Delta N) \times N$ 行列である．

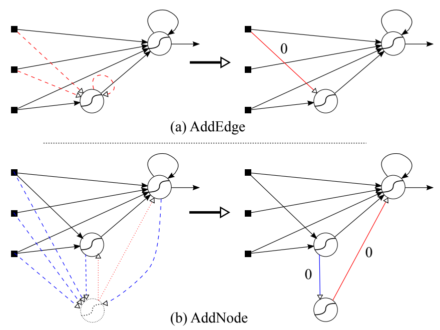

図 1: 二種のトポロジ変異オペレータ（文献 [Moriguchi 12] より再掲）(a) AddEdge オペレータ：赤い波線が導入されうる辺を表す．それらのうち一つが選ばれ実際に導入される．導入された辺は右図で赤い実線によって表されている．(b) AddNode オペレータ：新しく点線で示された隠れノードが導入される．またそのノードに対する入力として追加されうる辺が青い破線で，ノードからの出力として追加されうる辺が緑の点線で示されている．選択された辺は右図で実践によって示される．追加される辺には全て零の結合強度が与えられる．

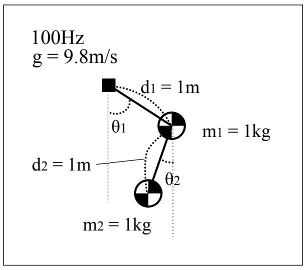

図 2: 二重振子の構成．

共分散行列も同様に拡張される必要がある．拡張は以下のように行われる：

$$\boldsymbol{C}^{(g)} = \left(\begin{array}{c|c} \hat{\boldsymbol{C}}^{(g)} & 0 \\ \hline 0 & \boldsymbol{E} \end{array} \right) \tag{3}$$

ここで $\boldsymbol{C}^{(g)}$ は拡張後の $(N + \Delta N) \times (N + \Delta N)$ 共分散行列，$\hat{\boldsymbol{C}}^{(g)}$ は元の $N \times N$ 共分散行列，$\boldsymbol{E}$ は各対角要素が $\delta_{add} = (\sigma_{init}/\sigma^{(g)})^2$ であるような $\Delta N \times \Delta N$ 対角行列である．$\boldsymbol{E}$ は新たに加えられたパラメタ軸方向の探索ステップ分散を σ_{init}^2 と設定する．この更新ルールは既存の辺と新たに加えられる辺の結合強度については無相関であると仮定したものである．また既存の辺同士の相関は，それまでと同様に扱われる．一方で，新規の辺と既存の辺の間の相関は，共分散行列 $\boldsymbol{C}$ が適応するにつれて徐々に学習される．

3. 実験

CMA-TWEANN のスケーラビリティを検証するため，本研究では二重振子の振り上げ・安定化タスクを用いて性能評価を行い，CMA-NeuroES と比較する．

3.1 実験設定

実験に用いる二重振子を図 2 に図示する．タスクの目的は，吊り下げられて均衡状態にある振子を，各関節にトルクを加え

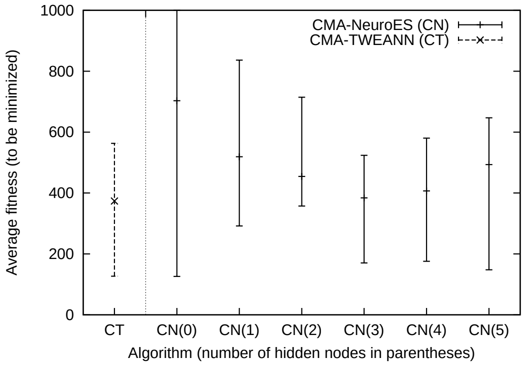

図 3: 実験結果．CT は CMA-TWEANN を，CN は CMA-NeuroES を表す．各アルゴリズムを用いて 10 回タスクを解き，最終的にベストな評価値について，平均値・最小値・最大値を示している．CN(0) については最大値が大きく表示範囲を超えているが，最大値は 3557.58 であった．

ることで直立状態に遷移させ，そのまま安定させることである．今回は初めの 500 ステップ (=5 秒間) で振子を振り上げ，その後 300 ステップ (=3 秒間) 直立状態を維持することを目的として，以下の評価関数を用意した，

$$f(x) = \sum_{t=500}^{800} (\cos(\theta_1(t)) + \cos(\theta_2(t)) - 2)^2,$$

ここで $\theta_1(t), \theta_2(t)$ は各関節の時刻 t における角度を表す．この関数を最小化する政策を発見することが，CMA-TWEANN および CMA-NeuroES の目標となる．

各ステップにおいて，ニューラルネットには関節の角度 θ_1, θ_2，角速度 $\dot{\theta}_1, \dot{\theta}_2$，およびバイアス項として 1 が入力として与えられる．それぞれ角度，角速度は値がおよそ $[-1:1]$ の範囲に収まるようスケールされる．ニューラルネットは二つの出力ノードを持ち，その出力は $[-10:10]\mathrm{kg}\cdot\mathrm{m}$ の範囲にスケールされ各関節のトルクとして利用される．シグモイド関数には $\varphi(x) = \frac{x}{1+|x|}$ を用いた．

CMA-NeuroES には初期の隠れノード数が異なるトポロジを与え，トポロジの違いが探索性能の違いを産むことを明らかにする．ネットワークは全て完全結合リカレントネットワークである．それに対し，CMA-TWEANN には入力・出力が直接つながった最小のトポロジを与える．

CMA-NeuroES および CMA-TWEANN に共通するパラメタは，初期のステップサイズ σ_{init} およびステップサイズの下限 σ_{min} である．今回の実験ではこれらのパラメタを $\sigma_{init} = 0.5, \sigma_{min} = 0.001$ と設定した．また CMA-TWEANN のパラメタである，AddEdge, AddNode オペレータの適用確率はそれぞれ $P(\alpha_n) = 0.005$ および $P(\alpha_e) = 0.02$ と設定した．

3.2 実験結果

各アルゴリズムを用いて振子の振り上げ・安定タスクを解いた実験結果を図 3 に示す．

この結果からは以下のことが分かる．二重振子の振り上げ・安定タスクにおいては，適切なトポロジを事前に与えなければ CMA-NeuroES は効果的な探索ができない．具体的には，隠れノード数が 0 のトポロジでは最も性能が悪く，それが 3 に増加するまで徐々に性能が向上した．そして，それより隠れノード数が増加するとふただび性能が徐々に低下した．一方で，CMA-TWEANN は適切なトポロジを与えなくとも自動的にトポロジを探索するため，適切なトポロジに関する事前知識を要さないことも分かる．CMA-NeuroES は隠れノード数が 3 の時に最も良い性能を示すが，CMA-TWEANN は適切な $P(\alpha_n), P(\alpha_e)$ が与えられればそれに匹敵する性能を示した．

4. まとめ

本研究では，われわれが文献 [Moriguchi 12] で提案した CMA-TWEANN について，適切なトポロジの発見が必須なタスクにおいても効果的に機能するか検証した．

二重振子の振り上げ・安定タスクは，倒立振子とは異なり単純なトポロジのみを有するニューラルネットでは効果的に解くことができず，適切なトポロジの発見が必須である．そのようなタスクにおいても，CMA-TWEANN は事前知識なしに適切なトポロジの発見と結合強度の最適化を自動的に行い，効果的な政策探索が実現できることが示された．

参考文献

[Atkeson 08] Atkeson, C. Stephens,B.: Random sampling of states in dynamic programming, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 38(4):924–929 (2008)

[Floreano 08] Floreano, D. Dürr, P. and Mattiussi, C: Neuroevolution: from architectures to learning, Evolutionary Intelligence, 1(1):47–62 (2008)

[Hansen 01] Hansen, N. Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies, Evolutionary computation, 9(2):159–195 (2001)

[HeidrichMeisner 09] Heidrich-Meisner, V. Igel, C.: Neuroevolution strategies for episodic reinforcement learning, Journal of Algorithms-Cognition Informatics and Logic, 64(4):152–168 (2009)

[Igel 03] Igel, C.: Neuroevolution for reinforcement learning using evolution strategies, In proceedings of CEC, pp. 2588–2595 (2003)

[Moriguchi 12] Moriguchi, H. Honiden, S.: CMA-TWEANN: Efficient Optimization of Neural Networks via Self-Adaptation and Seamless Augmentation , In proceedings of GECCO, to appear (2012)

[Nelder 65] Nelder, J. Mead, R.: A simplex method for function minimization, The Computer Journal, 7(4):308 (1965)

[Nolfi 04] Nolfi, S. Floreano, D.: Evolutionary Robotics, The MIT Press, Cambridge, MA (2004)

[Whiteson 06] Whiteson, S. Stone, P.: Evolutionary Function Approximation for Reinforcement Learning, The Journal of Machine Learning Research, 7:877–917 (2006)

[Yao 99] Yao, X: Evolving artificial neural networks, Proceedings of the IEEE, 87(9):1423–1447 (1999)