

## 進化的深層強化学習を用いた言語創発における 遺伝と学習の相互作用の解明

### Elucidating the Interaction between Genetic Inheritance and Learning in Language Emergence using Evolutionary Deep Reinforcement Learning

若林 啓

筑波大学 図書館情報メディア系

#### 1. 研究目的

ヒト言語の創造性と世代を超えた進化の起源を解明するには、遺伝と学習の相互作用を理解することが不可欠である (Dor & Jablonka 2001). 言語進化の過程は直接観察が難しく、化石証拠や生物実験による実証的証拠にも限界があるため、深層強化学習を用いた計算機シミュレーションによって言語の創発と進化を分析する研究が近年注目されている。しかし、計算コストの制約により、既存研究は主に一世代内で創発する言語の分析に留まっており、遺伝的進化と創発言語の関係は未解明の課題として残されている。

報告者らはこれまでに、進化計算と深層強化学習を組み合わせた進化的深層強化学習アプローチを提案し、遺伝と学習の相互作用メカニズムの理解を進めるための枠組みを構築してきた。本研究では、この枠組みを用いて、各個体に与えられる学習機会の量が、遺伝的進化と創発言語の性質に与える影響を分析する。ヒトの言語は、言語表現の部分の意味の組み合わせによって文全体の意味を表現する「構成性」を持つが、深層強化学習エージェントの創発言語に必ずしも自然に現れるとは限らない (Chaabouni ら 2020)。一方、構成的な言語は新規個体に学習されやすい (Cheng ら 2023) ため、進化過程で定着しやすい可能性がある。本研究では、ニューラルネットワークの初期重みを遺伝子として扱い、学習機会が多いほど構成的言語を学習しやすい遺伝子が広まるという遺伝的同化仮説を、シミュレーションによって検証する。

#### 2. 研究成果の内容

本研究では、20 個体のニューラルネットワークエージェントからなる集団を、遺伝的アルゴリズムに従って 1,000 世代にわたり進化させる進化的深層強化学習シミュレーションを、複数の設定で実施した。各世代では、個体が集団内の他個体と繰り返しペアを組み、「再構成ゲーム」と呼ばれるコミュニケーションゲームを通じて学習する。送信者は観測した状態に基づいてメッセージを出力し、受信者はそのメッセージから状態を再構成する。この過程で、集団内にメッセージと状態の対応関係が形成され、創発言語が成立する。遺伝子はニューラルネットワークの初期重みパラメータを表すベクトルとし、コミュニケーション成功率の高い個体の遺伝子ほど次世代に引き

継がれやすくなった。複数条件での実験から、以下の知見が得られた。

- (1) 各世代で学習機会がある環境では、学習機会のない環境に比べて、短期間の学習でコミュニケーション成功率を高められる遺伝子が広まりやすいことが示された。学習機会のない設定では、出生時から成功率の高い遺伝子が選好される一方、学習機会のある設定では、出生時の成功率が低くても、学習によって成功率を高められる遺伝子が選好される。興味深いことに、学習機会のある環境で進化した遺伝子は、ある程度の進化の過程を経た後には出生時の成功率も向上し、学習機会のない環境と同程度に達した。これは、前世代の集団が学習によって獲得した創発言語を生得的に理解しやすい遺伝子が有利になる、遺伝的同化の一種と考えられる。
- (2) 1 世代あたりの学習機会が多いほど、集団内に定着する創発言語の構成性が高まる傾向が示された。これは、構成的な言語を選好する初期重みパラメータが存在し、それが短期間での学習による成功率向上においても有利であることを示している。この結果は、構成的な言語が新規個体にとって学習しやすいという既存の知見 (Cheng ら 2023) を発展させ、学習機会の量が構成的な創発言語の定着を促す進化的圧力となる可能性を示唆するものである。

さらに本研究では、1 回の相互作用で報酬が得られる再構成ゲームに加え、報酬獲得まで複数回の相互作用を要する環境へ枠組みを拡張し、報酬遅延がコミュニケーションの創発に与える影響も分析した。その結果、報酬遅延のある環境ではコミュニケーションの創発自体が難しくなり、遺伝的進化がより重要な役割を果たす可能性が示された。

### 3. 学際共同利用プログラムが果たした役割と意義

学際共同利用プログラムで提供された計算機を用いて、ニューラルネットワークエージェント集団の学習を分散並列で処理することによって、1,000 世代にわたる進化的深層強化学習シミュレーションを現実的な実行時間で行うことが可能になった。本プログラムによる支援は、言語創発における進化と学習の相互作用に関する基礎研究の促進において大きな意義を果たしたと言える。

### 4. 今後の展望

本研究では、ニューラルネットワークの構造として多層パーセプトロンを用いたが、再帰型ニューラルネットワークなど、メッセージの系列構造をより適切に処理できる構造を用いることで、より複雑な創発言語を伴う場合への知見の一般化が期待される。この場合には、初期重みパラメータの空間が大規模化し、単純な遺伝的アルゴリズムでは十分な進化が難しくなるため、より効率的な進化アルゴリズムの開発が必要になると考えられる。

5. 成果発表

(1) 学術論文

・ Naoki Inoue, Kei Wakabayashi : Communication Emergence under Reward Delay: The Role of Direct Utility in Temporally Extended Communication Games. *Artificial Life and Robotics*, vol.31, Issue:1, 11pages, 2026.2

(2) 学会発表

・ 外谷 直人, 若林 啓 : ノイズロバストな学習手法を用いた創発言語における構成性の改善. 第 40 回人工知能学会全国大会論文集 (JSAI), 2026.5

・ Naoto Toya, Kei Wakabayashi : Investigating the Baldwin Effect in Emergent Collective Communication via Evolutionary Deep Reinforcement Learning. *ALIFE2025 workshop Emergence and evolution of language and communication*, 2025.10

・ Naoki Inoue, Kei Wakabayashi : A Computational Analysis Method for the Emergence of Communication with Temporally Extended Communication Games. *ALIFE2025 workshop Emergence and evolution of language and communication*, 2025.10

使用計算機	使用計算機に○	配分リソース※		
		当初配分	移行*	一般利用による追加
Pegasus	○	16000		
Miyabi-G				
Miyabi-C				
※配分リソースについてはノード時間積をご記入ください。 *バジェット移行を行った場合、「+2000」「-1000」のように記入				