

マルチエージェントコミュニケーションにおける共通の記号接地の創発

荒木 武人[†] 三宅 陽一郎[†]立教大学大学院人工知能科学研究科[†]

1. はじめに

2006年以降、エージェントの学習を用いて言語の創発についての数理シミュレーションを行おうという取り組みが多く行われている[1][2]。しかし、これらのシミュレーションではエージェントそれぞれが用いる記号の接地は一致しておらず[2]、これはそれぞれのエージェントが異なる言語を用いて会話していることを意味する。人間社会の知能においては、共通に接地された言語を用いたコミュニケーションの実現が不可欠であるが、その接地プロセスは明らかにされていない。本稿では複数のエージェントによる共通の言語接地を実現するためのモデルを提案し、シミュレーションによってその有効性を示す。

2. 数字あてゲーム

言語の役割が最も単純かつ明瞭になるのは、聞き手が言語だけを受け取り、かつ聞き手と話し手が固定されている場合である。このような状況の例としては、目隠しをして周囲の人からの声による指示を頼りにスイカを割る挑戦をする「スイカ割り」があげられる。スイカ割りでは、指示をする側は観測した実際の状況を言語化して挑戦者に伝え、挑戦者は逆に受け取った言語から実際の状況を想像し、それを元に行動を取る。これは、指示をする側が実際の状態を言語にエンコードし、挑戦者はその言語から実際の状態をデコードしているような状態となっている。即ち、このゲームを成功させるために言語に期待される役割は、実際の状況を復元できる表現になっていることのみである。

本研究では、上記のスイカ割りを簡単にモデル化したゲームとして、図1に示すような「数字あてゲーム」を提案する。このゲームは二人のプレイヤーが「話し手」と「聞き手」の二つの役割に分かれて行う。まず、話し手がN個の数字のうち一つをランダムに環境から受け取り、それを元に聞き手に言語に見立てたN個の記号のうちから一つを選んで伝える。聞き手はその記号をもとに話し手が受けとった数字を予測し、N個の数字のうちから一つを回答する。回答した数字が正解だった場合は1、不正解だった場合は0の報酬が両方のプレイヤーに与えられる。数字あてゲームの報酬を最大化するためには、話し手がそれぞれの記号を別々の数字に接地し、聞き手がそれを解釈できることが要請される。本研究では、2体のエージェントによる数字あてゲームのシミュレーションを行う。

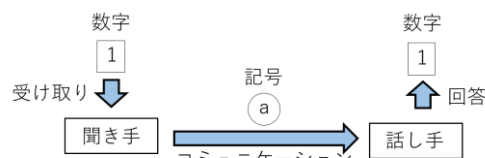


図1: 数字あてゲームの概要図

3. 事前学習

数字あてゲームを行う前に、環境から渡される数字と同じ数字を回答することが、数字あてゲームの成功条件であることをエージェントに学習させる。この学習により、聞き手のエージェントは、話し手が環境から受け取った数字を、聞き手自身が受け取った記号から予測できれば、その数字をそのまま回答できるようになることが期待できる。この事前学習ではエージェントは単独でN個の数字から一つをランダムに環境から受け取り、それを元にN個の数字から一つを回答する。回答した数字が受け取った数字と同じ場合は1、違う場合は0の報酬が与えられる。

4. エージェントアーキテクチャー

エージェントは自身の役割に応じて、アクター、スピーカー、リスナーの三つのモジュールを組み合わせて、図2に示すようなそれぞれの構造を取る。図2(a)は、数字あてゲームの事前学習を行う場合である。環境から受け取った数字はアクターモジュールによって処理され、回答が行われる。図2(b)は、数字あてゲームにおいて聞き手を担う場合である。エージェントは話し手から受け取った記号からリスナーモジュールを用いて話し手が環境から受け取った数字を復元し、それを元にアクターモジュールが回答を行う。図2(c)は、数字あてゲームにおいて話し手を担う場合である。エージェントは環境から受け取った数字からスピーカーモジュールを用いて聞き手に渡す記号を出力する。

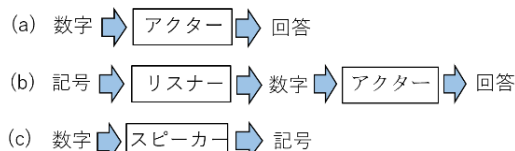


図2: エージェントアーキテクチャー

5. 強化学習

本研究では集中学習や教師なしでの強化学習を用いたエージェント間での言語の創発を目指す。強化学習の手法としてはTD誤差を方策改善の指針としたアクタークリティック法(AC法)を採用した。エージェントを構成するアクター、リスナー、スピーカーの三種類のモジュールはそれぞれAC法のためのActor,

The emergence of a common symbol grounding in multi-agent communication

[†] Taketo Araki and Youichiro Miyake, Graduate School of Artificial Intelligence and Science, Rikkyo University

Critic の二種類のニューラルネットワークを持つ。各モジュールは数字または記号を受け取ると最初に One-hot 表現に変換する。次に、Actor ネットワークを用いて上記の One-hot 表現のベクトルからどの数字または記号を選ぶかの確率分布を出力する。そして、Actor ネットワークの出力した確率分布をサンプリングし、出力する数字または記号を決定する。次に、Critic ネットワークを用いて One-hot 表現のベクトルから、状態価値を計算する。そして、Critic ネットワークが出力した状態価値から損失を計算し、強化学習を行う。

6. 共通の言語接地のための教師あり学習

本研究では、エージェント内のリスナーとスピーカークの二つのモジュールを教師あり学習を用いて同期することでエージェント間に共通の記号接地の実現を試みる。図3に示すように環境から入力を受け取ったリスナー、スピーカークいずれかのモジュールの出力を用いて、その出力からもう片方のモジュールが元の入力を復元できるように交差エントロピー誤差を用いて学習を行う。これにより、記号と数字の対応が相互的になることが期待される。ただし、リスナー、スピーカークどちらのモジュール出力が確率的にサンプリングされるため、学習が不安定になりやすい。この学習の不安定化を避けるため、サンプリングを行う前の Actor ネットワークが出力する確率をそのままう片方のモジュールに渡すように実装を行った。また、教師あり学習の影響が強すぎると、強化学習が進まなくなる。そこで、強化学習での損失にこの教師あり学習での損失を 0.1 倍してから加え、この損失を用いてエージェントを学習させることで、エージェント間に共通かつ、ゲームの報酬最大化ができる記号接地が形成されることを目指す。

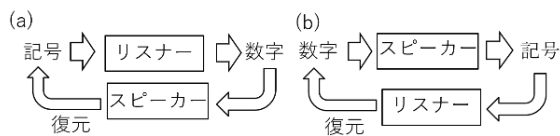


図3: 教師あり学習の概要図

7. 実験

実証実験のため、数字として 1 から 8、記号として A から H を用いて、N=8 でのエージェントによる数字あてゲームおよびその事前学習を以下の手順で行った。最初に、二体のエージェントを用意し、それぞれをエージェント 1、エージェント 2 とする。そして、それぞれのエージェントに数字あてゲームの事前学習を行い、アクターモジュールを学習させた。事前学習では、学習率 0.001 で 10000 回の学習を行った。次に、事前学習を行った二体のエージェントを用いて数字あてゲームを行い、リスナーとスピーカークの二つのモジュールを学習させた。数字あてゲームではそれぞれのエージェントが交互に話し手と聞き手となって、学習率 0.001 で 30000 回のゲームを行った。

8. 結果

事前学習の評価のため、報酬期待値を評価指標とする。実験の結果、事前学習での報酬期待値はエージェント 1 で 0.9826、エージェント 2 で 0.9836 であり、数字あてゲームを行うにあたり問題のない高い精度であることが確認できる。

次に、数字あてゲームで報酬を最大化できるような記号接地ができていないかを評価するために、数字あてゲームでの報酬期待値を指標とする。図 4(a)に報酬期待値の遷移を示す。実験終了時の報酬期待値はエージェント 1 で 0.9737、エージェント 2 で 0.9705 と高く、図から学習が安定して進んだことも確認できる。数字あてゲームでの記号接地の様子を確認するために、実験で用いた二体のエージェントが話し手として、それぞれの数字に対して出力する記号の確率分布の自然対数を図5に示す。図から二つのエージェントがそれぞれの記号を別々の数字に同じように接地させていることが確認できる。また、定量的に評価するために、それぞれの数字に対してこの二つの確率分布の内積をとり、平均したものを言語接地の類似度とし、シミュレーション中の遷移を図 4(b)に示した。実験終了時の類似度は 0.9845 と高く、この図から共通の言語接地が安定して行われていることが確認できる。

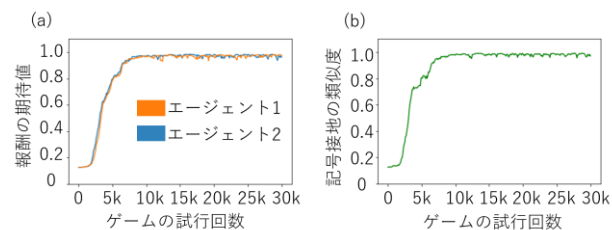


図4: 数字あてゲームにおける(a)報酬の期待値と(b)記号接地の類似度の遷移

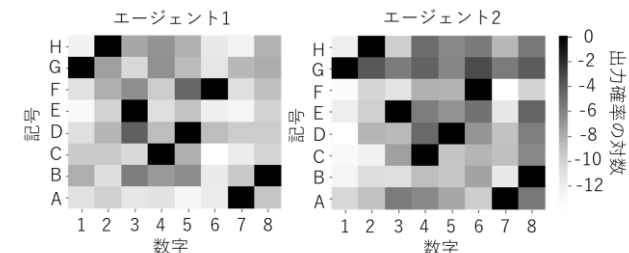


図5: それぞれの数字に対して出力される記号の確率分布のヒートマップ

9. まとめ

本研究では、教師あり学習を用いてエージェント内の記号を解釈するネットワークと、記号を発するネットワークを同期させることで、記号と数字の対応が相互的になり、エージェント間に共通の記号接地の形成を確認することができた。これらの結果は、記号接地プロセスにおいて言語と実際の事物が相互に結びついていることが、同じ集団における共通の言語の形成を促進することを示している。

参考文献

[1] Mazurowski, M. A., and Zurada, J. M. Emergence of communication in multi-agent systems using reinforcement learning. *IEEE International Conference on Computational Cybernetics*, (2006).
 [2] Lin, T. et al. Learning to Ground Multi-Agent Communication with Autoencoders. *Advances in Neural Information Processing Systems*, (2021).