

DRQNによる 幼児の語彙獲得のモデル化

野口 輝

南 泰浩

電気通信大学大学院 情報理工学研究科

noguchi@sd.is.uec.ac.jp minami@is.uec.ac.jp

1 はじめに

言語処理の分野において、現実の「実体」と「記号」が結びつける記号設置問題がある。本研究ではこの問題を解決するための一歩として人間の、特に幼児の語彙獲得方法を参考にして語の意味を獲得するモデルを構築する。

2 研究背景

自然言語の分野での、人の語彙獲得と工学的な記号接地問題との違いは以下の2点である。

1. 経験と「記号」を単位とする学習
2. 入力と教師信号の不確定・確定な学習

2.1 経験と記号を基本とする学習

人間は「リンゴ」という単語を覚える際に、実世界に存在する「実体」のリンゴを感覚器官を通して経験する。そして経験した情報を「記号」のリンゴと重ね合わせることで「リンゴ」という単語を理解している。しかしながら言語処理における記号処理では「リンゴ」という単語を学習する際に、「甘い、丸い」という記号を単位として単語と結びつけようとしているためその意味を十分には把握できない。

2.2 入力と教師信号の確定・不確定な学習

言語処理における記号処理と人間とでは学習方法が異なる。前者では何かを学習する際に、完全な正解・不正解データを教師として与えることが多い。しかしながら、人間の学習、特に幼児の学習では完全な正解・不正解データが与えられるわけではない。母親が教師となり、幼児が取得した特徴を見て正解を与えるわけではなく、幼児の行動を見て、その結果に対して「ほめる・頭をなでる」といった報酬を与えることにより幼児に語の意味を獲得させる。

3 研究目的

工学的な側面からの幼児の語彙獲得の研究は近年多く行われているが、心理学的な複数の現象を考慮した上での幼児の語彙獲得モデルは少ない。そこで、本稿では心理学的要素を考慮した上で、幼児の語彙獲得モデルを構築していく。これにより心理学的な面では語彙学習現象の解明を、工学的な面では経験による記号接地のモデル化の有効性を検証していく。

4 幼児の語彙獲得現象

言語処理において、人の語彙学習と記号接地問題との違いを記載したが、本研究で扱う心理学的要素である幼児の語彙獲得現象は、以下に示す共同注視・学習におけるバイアス・意図の理解である。

4.1 共同注視

共同注視とは母親と子供が注意を共有する行為である。Butterworth[1] は子供が注意しているものに対して、そのあとから親が注視し、それに言及するほうが幼児の語彙学習が進むということを示した。

4.2 語彙学習におけるバイアス

幼児の語彙学習において、少ない刺激でも確実に語彙を獲得できる理由として、様々な先験的な知識であるバイアスを仮定する説がある。特に、名詞カテゴリーバイアスや形状類似バイアスは、幼児がカテゴリー形成する際に適用される原理と考えられている。例えば取手のついてないコップと取手のついているコップを幼児に見せても、どちらもコップであると判断することができる。これは幼児が各々のコップを対象ではなく、形のあり方としてカテゴリー区別しているためである。

4.3 意図の理解

親の意図を読み取ろうとすることは、語彙の意味獲得のためには重要であり、コミュニケーション能力を促進させることに繋がる。幼児は親の意図を理解しようと努めるため、親の目線・身振りなどにも注視し、結果として親が「何」をしようとしていたかという意図を学習している。

5 従来手法

知識獲得モデルの従来手法として、曾傑らによる対象物の特徴や機能に関する知識を獲得する対話システムがある [2]。これは大規模な対話のデータから、対象物の特徴や機能に関する知識を獲得し、その知識を物体を表す属性として、スロットに埋め込むことで、扱う対象物の知識獲得を行うものである。

6 提案手法

本研究では幼児の学習に注目し、未知の環境の中で実際に体験する一連の手続きから物体と記号を結びつける語彙獲得システムの構築を行う。初期状態として物体と名称を結びつけるためのルールなどは与えず、自ら試行錯誤することでそれらの知識を獲得できるかを検証していく。またこの際に親の意図の認識率を可変にすることで、その認識率による語彙獲得の様相の比較を行い、心理学的な面での要素による違いなども見ていく。これら表現するために本研究では深層強化学習という手法を用いた。

6.1 深層強化学習

強化学習とは未知の環境下で実施に試行しながら報酬が最大になるような最適な行動を学習する手法である。状態 s が与えられ、エージェントはそれに応じて行動 a を選択する。この結果、その行動 a に応じて状態が遷移し、目的を達成した場合に報酬 r が得られるという仕組みになっている。今回は強化学習の中でも Q 学習という手法を使用する。この更新式は以下のとおりである。

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \max_{a'} Q(s', a') - Q(s, a)) \quad (1)$$

また、近年では強化学習と入力の不確定性や状態の共有化を行うため深層学習を組み合わせた深層強化学習が、これまでにない性能を発揮することに成功している [3]。その理由として多くのパラメータを持った高次元の非線形関数である多層ニューラルネットワークにより状態を表現することで、扱える状態空間数が見かけ上増えたためと考えられる。

6.2 提案モデル

本研究では、深層学習と Q 学習を組み合わせた DQN (Deep Q-Network) を用いることで、言語獲得モデルの構築を行った。提案モデルと具体的な処理の流れとしては図 1 のとおりである。未知の物体からどの特徴を取得するか選択し、選択された特徴からさらに名称を選択するようなモデルとなっている。また、この処理の例では、未知の物体から「本」を当てる問題となっている。

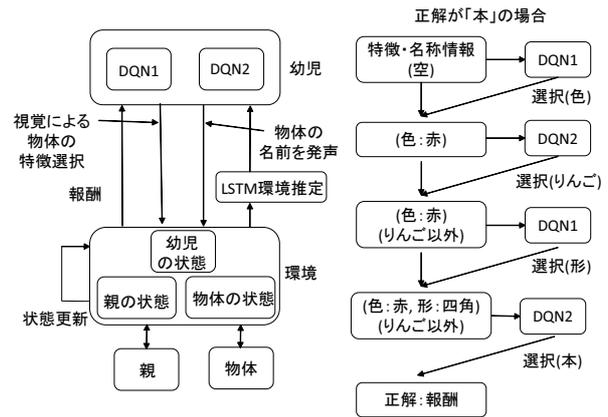


図 1: 提案モデル (左) と処理の例 (右)

しかし、幼児の語彙獲得では、必ずしも特徴などを正確に把握できるわけではない。そのため取得した特徴情報をそのまま使用せず、LSTM という過去情報を保持することができる深層モデルを通すことで状態を確率的に表す操作を行った。LSTM の式は以下のとおりである。

$$\begin{aligned} h_t, c_t &= LSTM(s_t, h_{t-1}, c_{t-1}) \\ &= LSTM(s_t, LSTM(s_{t-1}, h_{t-2}, c_{t-2})) \end{aligned} \quad (2)$$

LSTM は過去情報を保持することができる深層モデルであり、式 (2) のような再帰的な計算を行う。現在時刻 t における入力を表現する s_t 、内部パラメータを表す h_t 、過去の入力情報や予測結果を表現している c_t で表される。つまり今回のモデルは DQN と LSTM を組み合わせた DRQN (Deep Recurrent Q-Network) がベースとなっている。提案モデルの詳細を図 2 で示す。LSTM を共有しているのはそちらの方が性能が良いというが示されているためである [4]。

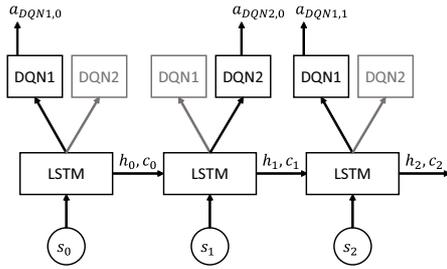


図 2: 提案モデルの詳細

7 実験

7.1 実験準備

提案モデルである語彙獲得モデルの構築と、それを使用した意図理解の重要性の検証を行った。検証方法としては用意した 8 つの物体名称・動作 (りんご・オレンジ・ボール・本・ブロック・早い・遅い・回る) と、6 つの特徴 (色 (RGB)・形・大きさ・味 (甘味, 酸味)・固さ・動き) を用いることで、物体が持つ特徴集合と名称を結びつけられることを示す。特徴の表現には RGB や単純な数字評価 (2 値) などを使用することで記号が単位となることを避る。200 episode 毎の訓練の後に 30 episode のテストを行い、名称一致回数で評価する。

7.2 名詞バイアス獲得の検証

ここでは、バイアスとい知識が自然に獲得できるかを調べるため、親が、幼児に名詞だけを教え続けることで、幼児が提示した名詞の特徴を理解するかを検証した。名詞だけに着目するため単語の名称のうち「早い・遅い・回る」の動作を抜いた。特徴としては、「色 (RGB)・形・大きさ・味・固さ」を用意した。この実験では幼児は親の意図理解を行っていない。特徴の選択から名称を当てる手続きを検証した。その結果を表 1 に示す。

実験結果から学習が進むにつれ、規則性のある特徴の選択方法になることが分かる。これに伴い、名称の一致率が上がっていく。学習の初期 200 回目では、特徴選択では大きさのみを選択している、また、不一致回数も高い。しかし、学習 2000 回目では特徴選択に規則性が生まれ、記号一致率も高くなった。このことから名詞のみを与える環境から一定の規則性が発見できることが分かる。特徴量選択の順序は、意味の木構造を表現しており、物体のカテゴリ表現を実現している。このような特徴量の選択の順番は、ある種のバイアスであると考えられ、今回のモデルを用いることで、形状バイアスが幼児の学習過程から構成できることが分かる。

表 1: 学習回数による名称一致率と特徴量選択順番の差

学習回数	一致率	特徴選択 1 回目	2 回目	3 回目
学習回数 200	30%	特徴: 色・大きさ	大きさ	大きさ
学習回数 1000	35%	味	味・大きさ	味・大きさ
学習回数 2000	77%	味	色・固さ	...

7.3 幼児の語彙獲得モデルの検証

意図理解の有用性を検証するために、提案モデルの LSTM の入力に親の意図も追加した。ここでの親の意図は、幼児に対してどの特徴を取得して欲しいかが表現されている。ただし、幼児は初めから親の意図を完全に理解することは出来ない。これをモデル化するため、今回は親の意図 (真の意図) の認識率をコンピュータ上で操作し、親の意図として幼児モデルの LSTM の入力に与える。実験では親の意図を上手く認識できる場合・できない場合とでは学習の速度や精度に何か違いが生まれるかを調べ、意図理解の重要性を検証する。その他の実験環境としては、図 2 の LSTM 層は 100 次元、DQN の層 (tanh 層) は 50 次元である。学習を行うために最初に物体のリストを作成した。リストに載せる値の範囲を表 2 に示す。この範囲に従って、実際の値を付与した。今回の実験環境では物体の名称 (名詞) を一致させる学習の割合が 80% であり、動作 (形容詞) を一致させる学習が 20% である。真の意図の認識率を 20% (ほぼ認識できていない) と 80% (ほぼ認識できている) に設定して二つの結果の比較を行った。この結果を図 3 に示す。

表 2: 用意したデータの特徴

物体	色 (RGB)	大きさ	形	味	固さ
りんご	赤	小	丸	甘 (2)	柔 (3)
オレンジ	オレンジ	小	丸	酸っぱい (2)	柔 (3)
ボール	白	大	丸	なし	柔 (3)
本	赤・オレンジ	小・大	四角	なし	柔・固
ブロック	赤・白	大	四角	なし	柔・固

() 内の数字は種類数

甘い (2) は甘い 1 と甘い 2 で分類可能

実験の結果、認識率が高い場合は、認識率が低い場合と比較して早期に名称の一致回数が高くなり、また安定性も優れていた。ただし、認識率が高い場合でも動作の名称 (形容詞) は一致させることができなかった。これは実験環境では物体の名称 (名詞) を学習する機会が多い (名詞出現率 80%) ため、名詞を獲得するためのバイアスが生まれ、形容詞を上手く学習できないためだと考えられる。

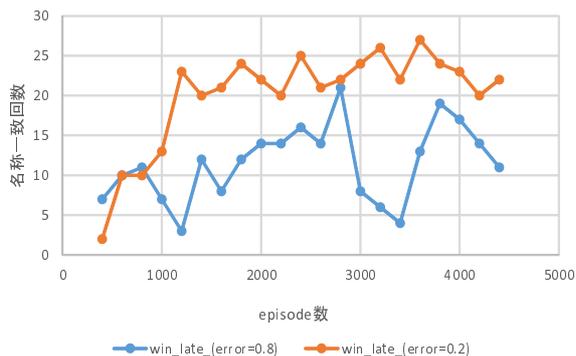


図 3: 実験結果 (名称の一致回数の比較)

また提案モデルは名称選択の不一致が起きてても何度か名称を選びなおすことを許している。このため、特徴や意図を考慮せず、名称を全て選択することで名称が一致する場合もある。しかしながら、表 3 と 4 から、意図認識率が高い場合は、認識率が低い場合と比べて、規則性が見られ特徴や意図を考慮しながら名称を選択していると考えられる。

このことから、提案モデルでは物体と名称を一致させるために必要な特徴選択の規則性を学習していることがうかがえた。つまり物体の名称 (名詞) をあてる問題が多い実験環境から、提案モデルは親の意図を考慮しながら物体と名称を一致させるために必要な最適な特徴を選択するようになったと考えられ、子供の意図理解の重要性がうかがえた。

表 3: 学習回数による特徴量選択順番の差 (認識率が低い場合)

学習回数	特徴選択 1 回目	2 回目	3 回目
学習回数 200	特徴: 固さ	固さ	固さ・形
学習回数 1000	固さ or 色 or 味	色・大きさ	色・固さ
学習回数 2000	動き・味	大きさ・固さ・味	形・味
学習回数 3000	動き・大きさ	固さ・形	固さ・大きさ

表 4: 学習回数による特徴量選択順番の差 (認識率が高い場合)

学習回数	特徴選択 1 回目	2 回目	3 回目
学習回数 200	特徴: 大きさ	大きさ	大きさ
学習回数 1000	固さ	形	大きさ
学習回数 2000	味 or 固さ or 大きさ	形	形
学習回数 3000	大きさ	形 or 味	...

8 考察

本研究では強化学習によって学習するニューラルネットワークを用いて、自律的に語彙を獲得するモデルを構築し、心理学的・工学的な面を考慮したうえでモデル

の検証を行った。結果として特徴・名称を選択する木構造のような概念モデルが自動的に構築されることが分かった。このような木構造は、木構造の段階を途中で打ち切ることにより、物体の抽象化ができ一種のバイアスとなると考えられる。この実験を通して、形状バイアスや名詞バイアスといったものが、学習の過程から獲得できることが初めて分かった。またこの検証により早期に安定した学習を行うためには親の意図が不可欠であることも分かった。この場合の意図とは例えば、名詞を多く教えるなどといった親の行動を指す。これらのことから提案モデルは経験による記号接地のモデル化において有効であることを示せたと考えられる。

今後の課題として、より現実に近いデータでの検証や、親の意図やバイアス以外の幼児の語彙獲得現象の重要性を示すような実験をしていく。また今回の実験では DRQN を使用しており、実際に試行した経験を保管し、学習の際に保管した経験からランダムに選択し学習している。しかし今回の実験では正解に近づかない経験を多く選択し学習することがあり、学習速度や精度に影響を及ぼした場合は存在した。そこで経験をランダムに選択するのではなく、重要度に応じた正規化を行いこの上で一様に選択することをやりたい。

参考文献

- [1] Butterworth, G., & Jarrett, N. L. M. 1991 "What minds have in common is space : Spatial mechanism serving joint visual attention in infancy." *British journal of Developmental Psychology*, 9, 55-72
- [2] 曾 傑, 高瀬裕, 中野有紀子. "対象物の特徴や機能に関する知識を獲得する対話システム" *The 31st Annual Conference of the Japanese Society for Artificial Intelligence*, 2017.
- [3] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra & Martin Riedmiller (2013) "Plating Atari with Deep Reinforcement Learning" *NIPS Deep Learning Workshop 2013*
- [4] Tiancheng Zhao, Maxine Eskenazi (2016) "Towards End-to-End Learning for Dialog State Tracking and Management using Deep Reinforcement Learning" *SIGDIAL 2016*