

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報理工学研究科 情報・ネットワーク工学専攻 博士前期課程		
氏 名	竹下 虎太郎	学籍番号	2131098
論 文 題 目	品詞情報の獲得を考慮した深層強化学習による 幼児の語彙獲得のモデル化		
要 旨			
<p>本研究では、田口の提案した幼児語彙獲得モデルを参考に、獲得してほしい語彙を名詞から動詞に変化させる親の意図を幼児が理解することで、品詞情報を順に獲得する幼児の語彙獲得モデルを提案した。提案モデルは、オペラント連合学習を模すために強化学習、幼児の心的状態の表現と状態数の大きさに対応するためにニューラルネットワークを用いた。モデルに組み込んだ心理学的要因は、共同注意、親の意図理解であり、特に親の意図理解については親が名詞と動詞のどちらを幼児に学習させたいかという親の意図と、幼児が名詞と動詞のどちらを学習したいかという幼児の意図をそれぞれ用意し、親と幼児の意図が一致した場合には親が喜びか驚きの表情を、一致しなかった場合には、喜び、驚き、悲しみ、怒り、嫌悪、ニュートラルのいずれかの表情を幼児に見せるように設定することで、親の意図を幼児が完全には理解できない事象を再現した。特徴情報は画像や動画等から取得した実データを用いた。</p> <p>名詞の語彙のみを獲得する幼児モデルに対する予備実験では、モデルが名詞語彙を問題なく学習できた。本実験では、1 エピソードの中で名詞の語彙のつぎに動詞の語彙を獲得する、品詞情報の獲得を考慮した幼児モデルに対する検証を行った。予備実験のモデルパラメータを事前学習したモデルと、事前学習なしのモデルで比較を行い、どちらのモデルでも特徴選択で対称性バイアスや名詞・動詞バイアスが確認された。また、事前学習ありのモデルでは正答率が伸び悩んだ結果に対し、事前学習無しのモデルでは正答率がほぼ 1 まで伸びた。この実験により、実世界のように名詞の語彙と動詞の語彙を完全に切り離さずに順に学習することで、徐々に各バイアスの獲得と適用するバイアスのバランスコントロールができ、品詞情報を獲得しつつ語彙も獲得できることがわかった。一方で、名詞語彙と動詞語彙を順ではなく同時に推定するタスクでは、正答率が 0.2 付近を推移し、モデルの学習が困難だった。これは、名詞と動詞の区別をするには動作主である名詞を学習した後に、その関連である動詞を学習するという順序が重要であること、名詞と動詞の結び付けを行う助詞も学習できれば、同時に学習できる可能性があることを示唆している。</p> <p>以上より、本研究で提案した幼児語彙獲得モデルは、名詞のつぎに動詞を学習する順序によって、学習バイアスをコントロールし、名詞と動詞の品詞情報を獲得しつつ語彙も獲得できることが確認できた。</p>			

2022 年度 情報・ネットワーク工学専攻
コンピュータサイエンスプログラム 修士論文

題目 品詞情報の獲得を考慮した深層強化学習による
幼児の語彙獲得のモデル化

2131098 竹下 虎太郎

指導教員

南 泰浩 教授

古賀 久志 准教授

2023 年 1 月 30 日

目次

第 1 章	はじめに	1
第 2 章	研究背景	2
2.1	ロボットを用いた記号設置問題に関する研究	2
2.2	幼児語彙獲得に関する心理学的研究	2
2.2.1	共同注意	2
2.2.2	親の意図理解	3
2.2.3	学習バイアス	3
2.2.4	オペランド連合学習	3
2.3	幼児語彙獲得における品詞情報獲得の重要性	3
第 3 章	研究目的	5
第 4 章	提案手法：品詞情報の獲得を考慮した幼児語彙獲得モデル	6
4.1	強化学習	6
4.1.1	定義	7
4.1.2	DQN	7
4.1.3	DDQN	8
4.1.4	探索アルゴリズム	8
4.2	再帰型ニューラルネットワークモデル	9
4.2.1	RNN	9
4.2.2	LSTM	9
4.3	提案モデルの構造	10
4.4	学習方法	11
4.5	行動空間	12
4.6	物体の特徴情報	12
4.6.1	色特徴量	13
4.6.2	形特徴量	14
4.6.3	味特徴量	14
4.6.4	動き特徴量	15
4.7	親の意図理解	15
4.7.1	親の表情	16

第 5 章	予備実験	18
5.1	実験設定	18
5.2	実験結果と考察	18
第 6 章	本実験	20
6.1	品詞情報の獲得を考慮した幼児の語彙獲得モデルの実験	20
6.1.1	実験設定	20
6.1.2	実験結果と考察	20
6.2	品詞情報を同時に獲得する幼児の語彙獲得モデルの実験	23
6.2.1	実験設定	23
6.2.2	実験結果と考察	24
第 7 章	結論	25
参考文献		27

第 1 章

はじめに

言語処理分野ではコンピュータが語と実態を結び付けるための自律的な特徴選択が困難とされる記号接地問題が大きな課題となっている。記号接地問題を解決するために、人が感覚器官を通して実態の特徴を理解する方法の模倣が着目されている。既存研究として、ロボットがその身体的な特徴を利用して事前の知識なしに語彙獲得する手法 [1] やマルチモーダルな情報を利用した概念・行動・言語を学習可能な統合認知モデルを実現した手法 [2] が提案されている。しかし、これらの手法では人間が語彙を自然に獲得する際に観察される複数の心理学的要素を説明できない。

この問題に対して、心理学的要素を考慮した幼児の語彙獲得機構のモデル化が提案されている [3]。この手法では事物と名称を結びつけるタスクにおいて、幼児が語彙獲得する際に観察される共同注意、親の意図理解、学習バイアスを考慮した上で語彙獲得を行うモデルの構築を行っている。特に幼児の語彙獲得で重要とされる親の意図をモデルに組み込むことで、モデルが親の意図を理解した上で名詞と動詞のどちらも学習可能であることを示した。しかし、この手法では、幼児が語彙獲得を行う品詞の順、特に名詞から動詞の順に学習する様子をモデル化できない。

そこで本研究では田口の提案した幼児語彙獲得モデルを参考に、学習の際に親の意図を変化させることで名詞から動詞の順に語彙獲得を行うモデルを提案する。

本論文は以下のような構成になっている。第 2 章で本研究に取り入れる心理学的な学説について説明する。第 3 章で本研究の目的について詳細に説明する。第 4 章で提案手法である幼児語彙獲得モデルと幼児語彙獲得モデルに用いている手法、品詞情報の獲得を行うための親の意図変化についてそれぞれ説明する。第 5 章で名詞の語彙のみを獲得する幼児モデルを用いた検証実験を行う。第 6 章で第 5 章で行った検証実験を踏まえた提案モデルの検証実験を行う。第 7 章で本研究の結論を述べる。

第2章

研究背景

本章では、ロボットを用いた記号設置問題に関する先行研究、および幼児の語彙獲得を考えるにあたって重要な心理学的要素に関する先行研究について説明する。

2.1 ロボットを用いた記号設置問題に関する研究

ロボットが「りんごを持ってきて」といった人間の指示を受けた場合に、ロボットは「りんご」という単語がどんな物体を指すのかや、「持ってきて」という単語がどんな動作を指すのかを理解していなければ、人間の指示を完遂できない。これは介護用ロボットや案内用ロボット等の人間の生活を支援するロボットの問題点であり、言語処理分野における大きな課題である。これを解決するために、数理的な確率モデルでロボットに言語を理解させる手法を検討した研究 [1] や、マルチモーダルな情報を用いて概念・行動・言語を学習可能な統合認知モデルを実現した研究がある [2]。しかし、これらの研究では幼児が一から言語を学習する上で重要となる共同注意や意図理解等の心理学的要素が組み込まれていない。幼児は何も学習していない真っ新な状態から共同注意や親の意図理解、学習バイアス等の心理学的要素を用いて親の意図を汲み取りながら徐々に言語の学習を行っていく。すなわち、幼児語彙獲得に関する心理学的要素を言語学習モデルに組み込むことは、相手の意図を汲み取りながら言葉の意味を自律的に学習するモデルを実現するために重要である。付け加えるならば、言葉の意味を自律的に学習するモデルは、語彙を一つ一つ明示的に学習する既存のモデルと比較して発展性や応用性に優れると考えられる。したがって、本研究では意図理解によって自律的な言語学習を行うモデルのたたき台となり得る、心理学的要素を考慮した幼児語彙獲得モデルの作成を行う。

2.2 幼児語彙獲得に関する心理学的研究

本節では幼児の語彙獲得において重要な心理学的要素である、共同注意、学習バイアス、親の意図理解、オペラント連合学習について説明する。

2.2.1 共同注意

幼児が語彙を獲得する状況では、周囲の環境の影響、特に母子相互作用の影響を受ける。共同注意とは、親あるいは幼児がある事物に注目するとき、もう一方も同じ対象に注目する行動である [4]。この行動により、幼児が周囲の物に視線を向け、母親は幼児の視線を追って物を共有しながら幼児の生活をサポートすると同時に、幼児は母親の働きかけに合わせるように振る舞う、といった母子の関係世界の構築と維持が行われている。

る。共同注意を通じて、幼児は語と実態を結び付けて語彙獲得を行う。

2.2.2 親の意図理解

幼児は親との共同注意によって親の意図理解を行う。親の意図を理解しようとすることは、語彙獲得において重要でありコミュニケーション能力の向上につながると考えられる。親と幼児が異なる事物を注視するとき、親が注視する事物の名称を幼児に伝えたときの幼児の語彙学習を調べた実験 [5, 6] がある。実験の結果、幼児は自身が注視していた事物ではなく親が注視していた事物の名称を学習していた。このことから、幼児は親の意図理解を積極的に行っており、後述する学習バイアスを喚起するものが親の意図理解であると考えられている。

2.2.3 学習バイアス

新たな語を学習する際に、心理学では、その語が対象の全体を指すのか部分を指すのか、または対象が持つ属性を指すのかなど、語に対する指示の範囲を同定する必要があるとされている。この同定のためバイアスと呼ぶ暗黙の知識を仮定する仮説がある。このバイアスは、名詞バイアス、事物全体バイアス、相互排他性バイアス、形状類似バイアス、対称性バイアスなどがある。例えば名詞バイアスや形状類似バイアス、対称性バイアスは、幼児がカテゴリーの形成を行う大きな要因とみなされている。例えば、取っ手のあるコップと取っ手のないコップを、幼児はどちらもコップだと認識できる。これは幼児がそれぞれのコップを固有の事物とみなさず、抽象的なコップの形状カテゴリー分類しているためである。以上のバイアスは、個々の品詞の語彙獲得には役立つが、日本語語彙全体を獲得する上では、バイアスが干渉しあう。この現象を避けるには、どのバイアスを適用するかをコントロールする必要がある。本研究では品詞情報を順に獲得していく過程で幼児が親の意図を理解し、バイアスを適応的に学習していく手法を提案する。

2.2.4 オペラント連合学習

幼児は親の態度や表情から親の意図を理解して親が喜ぶ良い行動と、親が喜ばない悪い行動を学習しており、良い行動についてはその行動の頻度を増やし (正の強化)、悪い行動については別の行動を試す (負の強化) ことで試行錯誤しながら行動の学習を行っていると考えられている。このような様々な刺激による能動的な学習をオペラント連合学習と呼ぶ。オペラント学習は「パブロフの犬」として広く知られているパブロフ型条件付けと混同されることがあるが、パブロフ型条件付けがある刺激に対して受動的な学習によって対象が無自覚に学習するのに対し、オペラント学習は対象が自発的に行動することで試行錯誤を繰り返し、能動的に学習するという点で異なっている。

2.3 幼児語彙獲得における品詞情報獲得の重要性

心理学の既存研究で、語彙獲得において名詞は一目で知覚可能な具体物を指すために他の品詞に比べ優位性があること、名詞のつぎに動詞を獲得することが示されている [7]。名詞と述部の区別は、「人や物のような具体的概念」と「活動、状態の変化、因果関係のような叙述的概念」とされる。事物の指示物を知覚的に利用の方が学習として容易であり、名詞に対応したカテゴリは、動詞や他の述部に対応したカテゴリより、単純で基本的であるとされる。また、幼児の意図理解は名詞を学習するときと動詞を学習するときでは異なって機能

することが示されている [8]。これらから、幼児の語彙獲得において品詞情報の獲得に順序があることがわかる。幼児に獲得してほしい語彙を名詞から動詞に徐々に変化させる親の意図を幼児が理解することで品詞情報を順に獲得できると考えられる。本研究では幼児が名詞から動詞の順に語彙獲得を行う様子に着目する。

第3章

研究目的

本研究では、田口の提案した幼児語彙獲得モデルを参考に、獲得してほしい語彙を名詞から動詞に変化させる親の意図を幼児が理解することで、品詞情報を順に獲得する幼児の語彙獲得モデルを提案する。提案モデルは深層強化学習を用いて構築する。幼児の語彙獲得機構において名詞のつぎに動詞を学んでいく品詞情報の獲得は重要な事項である。品詞情報の獲得を考慮することで、より実際の幼児に近い語彙獲得過程のモデル化を行う。また、提案するモデルは明示的なラベル付けによる学習とは異なり、親の意図という暗黙的な情報を理解する学習を行うため、他者とのやり取りから言語を自律的に学習するモデルの土台になる。

第 4 章

提案手法：品詞情報の獲得を考慮した幼児語彙獲得モデル

本章では、本論文で提案する幼児の語彙獲得モデルの構造と、使用する深層強化学習を構成する強化学習と再帰型ニューラルネットワークモデルの理論、品詞情報の獲得を考慮するための親の意図変化について説明する。

4.1 強化学習

強化学習における「強化」とは、心理学用語で刺激と反応の結び付けを強めることを指し、強化学習とは人工的なエージェントを用意し、エージェントの行動選択に応じて報酬を与えることで行動と報酬の結び付けを行う機械学習の手法である。報酬や嫌悪的刺激に適応して自発的な行動学習を行うオペランド条件付けのモデル化にも利用できる [12]。本研究では幼児が (幼児にとっては) 未知の物体の特徴を取得した後に未知の物体の名称を答え、親がその正否に応じて報酬や嫌悪的刺激を与えるという強化学習の手法を用いることで、幼児に物体の名称を語彙として獲得させる。以下では、強化学習の詳細な理論について述べる。図 4.1 に、強化学習のモデルを記述する。本研究における幼児語彙獲得モデルには強化学習の手法の一つであり、ゲームやロボットの AI に用いられている Deep-Q Networks(DQN) を改良した Double Deep-Q Networks(DDQN) を用いた。

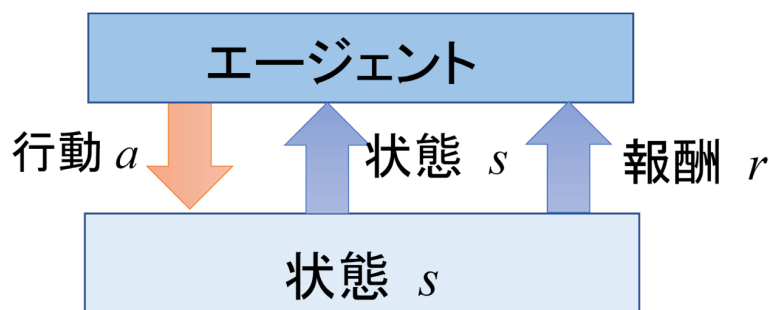


図 4.1 強化学習の基本原則

4.1.1 定義

ある環境下において何らかのタスクを実行するために行動を起こすエージェントを定義する．そのエージェントが置かれている環境を状態 $s \in S$ ，状態 s においてエージェントがとる行動を行動 $a \in A(s)$ とする．また，状態 s において行動 a をとって状態 s' となった際に得られる報酬を報酬 $r(s, a, s')$ とする．

このとき，強化学習では得られる報酬が最も大きくなる行動を選択するように，エージェントのパラメータが学習される [10]．この学習機構を用いると，前述した心理学のオペラント連合学習をモデル化できる．しかし，オペラント連合学習は，これまで，エージェント（動物や幼児）の心的状態を反映していないという批判が存在する．これに対する工学モデルとして，POMDP という手法が考案されている [11]．この手法では，ここでの状態をエージェントの心的状態と考え，外部の情報によってこの心的状態が変化する．これにより，エージェントの心的状態を考慮したモデルとなり，オペラント連合学習のモデルより，複雑なモデルを実現できる．本論文のモデルでは，このような心的状態を仮定するモデルをニューラルネットワークにより実現する．

4.1.2 DQN

状態 s において行動 a が選択される確率を方策 $\pi(s, a)$ とする．現在の時刻より未来の報酬に割引率 γ を順次かけることで，未来の不確実性を表現した累積の報酬を割引報酬和 $G_t = \sum_{\tau=0}^{\infty} \gamma^{\tau} r_{t+1+\tau}$ ($0 \leq \gamma \leq 1$) を定義する．方策 π に従って状態 s において行動 a を実行したとき得られる収益の期待値を行動価値関数 $Q^{\pi}(s, a) = E^{\pi}(G_{t+1} | S_t = s, A_t = a)$ とする．

DQN は各状態において最も割引報酬和が大きくなる最適行動価値関数 $Q^*(s, a) = \max_{\pi} Q^{\pi}(s, a)$ をニューラルネットワークを用いて近似する手法である [10]．DQN には学習を安定させるための代表的な三つの手法が存在する．

一つ目が，clipping である．ニューラルネットワークはモデルの出力と教師データの値の差を損失関数と呼び，損失関数を最小化するように重みの更新を行う．Clipping は損失関数を -1 から 1 の範囲にクリップすることでパラメータ更新の値を制限し，学習を安定させる手法である．

二つ目が，Experience Replay(ER) である．ER は行動を試行した結果 (s, a, r, s') をメモリに保管しておき，学習の際にバッチサイズ分だけランダムに取り出して使用することで，データの順番に対する相関の学習を防ぐ．

最後が，neural fitted Q である．実世界のタスクを扱う際には，最適行動価値関数は自明でない場合が多い．そのような場合において，DQN では正解データ y_t に見立てたターゲット y_t^{DQN} を 1 ステップ前のパラメータ θ^{i-1} を用いて式 4.1 のように定める．

$$y_t^{DQN} \equiv r(s, a, s') + \gamma \max_{a'} Q(s', a'; \theta^{i-1}) \quad (4.1)$$

しかしながら，学習を終えるたびにターゲットの更新を行う場合，正解データが毎回変化するのと同義であるため，収束が難しくなる．そのため，ターゲットを一定期間固定し学習を安定させる．また，一定期間毎に固定していたターゲットのパラメータ θ^- を更新し，最適行動価値関数の近似を図る手法が neural fitted Q である．

$$y_t^{DQN} \equiv r(s, a, s') + \gamma \max_{a'} Q(s', a'; \theta^-) \quad (4.2)$$

4.1.3 DDQN

式 4.2 のように、DQN はターゲットの更新に $\max_{a'} Q(s', a'; \theta^-)$ を用いている。しかし、この更新方法では選択した行動の行動価値を過大評価する可能性がある。

したがって、DDQN ではターゲット y_t^{DDQN} を式 4.3 のように決定する（ただし、 θ は学習中の Q 関数のパラメータ、 θ^- はターゲットの Q 関数のパラメータ）。

$$y_t^{DDQN} \equiv r(s, a, s') + \gamma Q_{\theta^-}(s', \arg\max_a Q_{\theta}(s', a)) \quad (4.3)$$

式 4.3 は、次状態 s' の下で $\arg\max_a Q_{\theta}(s', a)$ によって行動 a を決定し、行動 a の期待値 $Q_{\theta^-}(s', a)$ によって Q_{θ} を更新する構成となる。このように、DDQN では行動を試行する Q 関数と行動の価値を推定する Q 関数を分け、過大評価を防ぐ。

4.1.4 探索アルゴリズム

強化学習では報酬を多く得るために最適方策を可能な限り早く見つけ出し、可能な限り多く最適方策に基づいて行動を実行することが重要となる。しかしながら、最適方策を見つけ出す為には局所解にトラップされないように数多くの探索をする必要があり、数多くの探索をすると最適方策で実行できる行動回数が減るという問題がある。これを「探索」と「活用」のトレードオフと呼ぶ。この探索回数をなるべく小さくするために強化学習では多くの探索アルゴリズムが提案されている。

最も単純な探索アルゴリズムは Greedy アルゴリズムと Random アルゴリズムである。前者は常にその時点で最適な方策を取り続けるアルゴリズムであり、Random アルゴリズムは常に無作為な行動をとり続けるアルゴリズムである。前者は最適な方策を早い段階で見つけ出せれば得られる報酬が最大になるという利点があるが、一方で未探索の領域が最も大きくなり、局所解にトラップされやすい問題がある。後者は十分な試行を行えば未探索領域が無くなるという利点があるが、一方で最適な方策を見つけ出した後もランダムに行動を実行するため、得られる報酬が小さくなるという問題が挙げられる。これらの単純なアルゴリズムに対しても「探索」と「活用」のトレードオフの問題が見受けられる。

ϵ -Greedy アルゴリズムは Greedy アルゴリズムと Random アルゴリズムを組み合わせたアルゴリズムであり、 ϵ という小さい定数を設定し、 ϵ の確率で Random アルゴリズムを、 $1-\epsilon$ の確率で Greedy アルゴリズムを適用するアルゴリズムである。この組み合わせによって Greedy アルゴリズムの未探索領域が大きくなり、局所解にトラップされやすい問題や Random アルゴリズムの得られる報酬が小さくなる問題の解決を図っている。しかし、 ϵ -Greedy アルゴリズムは十分に探索が終わった後も ϵ の確率で Random アルゴリズムを適用するため、その分報酬が小さくなる問題がある。この解決策として、学習の初期段階では ϵ を大きくし、学習が進むにつれて ϵ を徐々に小さくする焼きなまし ϵ -Greedy アルゴリズムが提案されている。焼きなまし ϵ -Greedy アルゴリズムはニューラルネットワーク強化学習の分野で広く用いられる手法である。強化学習の非常に単純な問題設定の一つに、多腕バンディット問題がある。多腕バンディット問題とはそれぞれ異なる確率で報酬を返すように設定されているいくつかのスロットマシンに対して、エージェントがスロットを回す（アームを選択する）という単純な行動のみで報酬を最大化する問題である。当然、それぞれのスロットマシンが報酬を返す確率は不明であり、エージェントは探索を行ってより高い確率で報酬を返すスロットマシンを見つけて出し、そのスロットマシンをなるべく多く回して報酬最大化を目指す流れとなっている。この多腕バンディット問題に対する有名な解法として UCB アルゴリズムが存在する。UCB アルゴリズムは、未選択の

アームを必ず選択し、未選択のアームが存在しなければこれまでのアームの選択回数と得られた報酬を考慮して探索と活用のバランスを取るよう設計されたアルゴリズムである。

この UCB アルゴリズムを派生させたアルゴリズムの一つに KL-UCB が存在する。KL-UCB は KL (Kullback Leibler) divergence という 2 つの確率分布の差異を計る尺度を UCB アルゴリズムに適用したアルゴリズムである [12]。KL-UCB は理論限界を達成しており、理論上最も性能が良い強化学習の探索アルゴリズムの一つである。

本研究では、田口の比較実験 [3] より、幼児語彙獲得モデルの検証にて精度がよかった焼きなまし ϵ -Greedy 法を採用する。

4.2 再帰型ニューラルネットワークモデル

4.2.1 RNN

Recurrent Neural Network(RNN) はニューラルネットワークの一種であり、入力情報と共に 1 ステップ前の中間層の出力結果を入力として与え、時系列データを取り扱う。しかし、RNN には勾配消失問題があり、長期依存の時系列データを取り扱えない。

このモデルは時刻 t における入力情報 s_t 、RNN の前ステップの出力情報 h_{t-1} を用いて、式 4.4 のように表される。

$$h_t = RNN(s_t, h_{t-1}) \quad (4.4)$$

4.2.2 LSTM

LSTM は RNN と同様に時系列データを取り扱える再帰型ニューラルネットワーク [13] であり、勾配消失問題の存在する RNN を拡張し、内部情報を保持する媒介変数 cell を新たに導入することで、長期依存した過去の情報の保持を可能とするモデルである。このモデルは、時刻 t における入力情報 s_t 、LSTM の前ステップの出力情報 h_{t-1} 、及び LSTM の前ステップの内部情報 c_{t-1} を用いて、式 4.5 のように表される。

$$h_t, c_t = LSTM(s_t, h_{t-1}, c_{t-1}) \quad (4.5)$$

この式 4.5 をさらに 1 ステップ前の入力情報 s_{t-1} 、出力情報 h_{t-2} 、および内部情報 c_{t-2} を用いて式変形すると、式 4.6 のようになる。

$$\begin{aligned} h_t, c_t &= LSTM(s_t, h_{t-1}, c_{t-1}) \\ &= LSTM(s_t, LSTM(s_{t-1}, h_{t-2}, c_{t-2})) \end{aligned} \quad (4.6)$$

このような LSTM モデルの概要図を図 4.2 に示す。

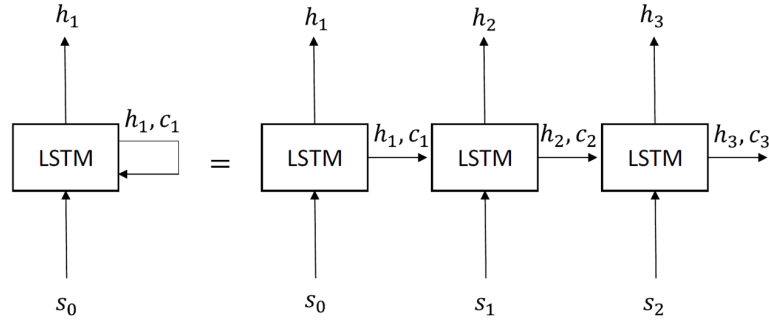


図 4.2 LSTM モデルの概要図

4.3 提案モデルの構造

以上の二つのモデルを用いた幼児の語彙獲得モデルの構造を図 4.3 に示す．また，図 4.4 は LSTM が幼児語彙獲得のモデルにおいてどのような役割を果たすのかを示した図である．

図 4.3 は，まず，初期状態 s_0 を図 4.4 のように心的状態を表す LSTM に入力し，その出力結果を DQN1 に入力として与える．これにより，DQN1 はどの特徴を選択するかを出力した後，LSTM は状態 s_1 の特徴情報を更新し，この状態をつぎの LSTM に入力し，その出力結果を DQN2 の入力として与えることで，DQN2 は名称の選択を行う．この処理を名称が一致するまで繰り返す．なお，図 4.3 における a_{DQN} はモデルのテスト時には $a_{DQN} = \underset{a}{\operatorname{argmax}} Q(s, a)$ となる．本研究ではこのモデルを使用する．

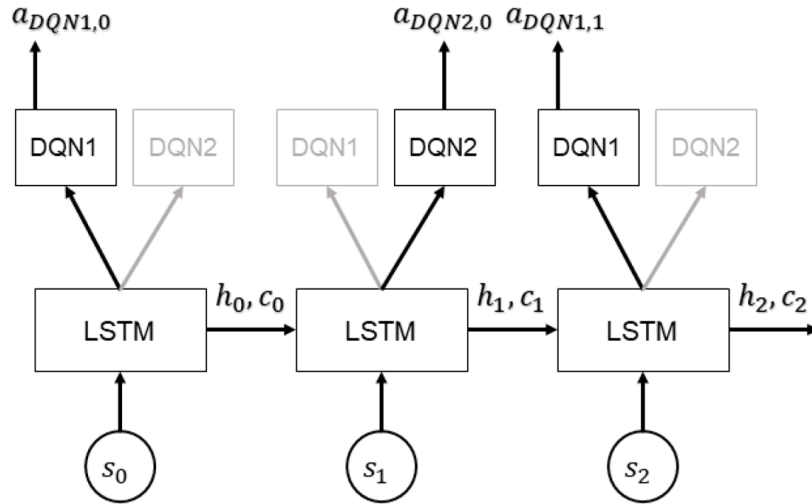


図 4.3 幼児語彙獲得モデルの構造

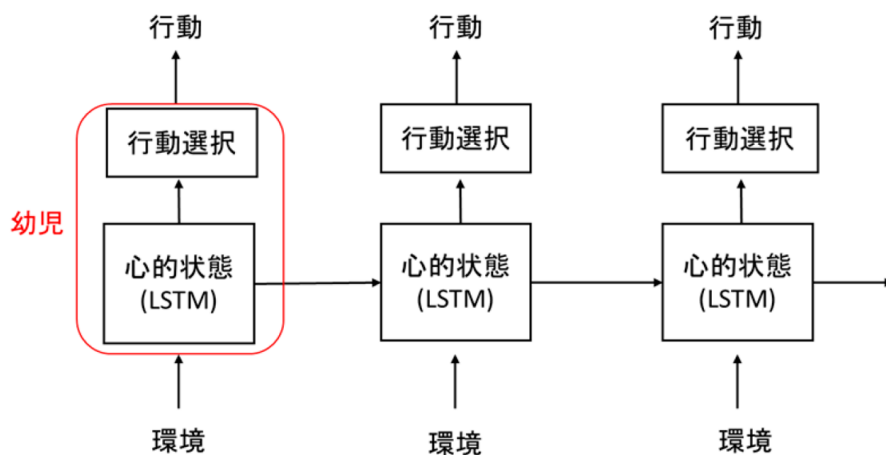


図 4.4 心的状態を表す LSTM モデル

4.4 学習方法

品詞情報を順に獲得する幼児の語彙獲得モデルにおける 1 エピソードは、以下の通りである。また、この一連の流れをまとめたものを図 4.5 に示す。1 エピソードの中でモデルは名詞のつぎに関連する動詞のラベル予測を行う。ラベル予測のためにそれぞれの品詞の特徴情報を取得する。

1. 学習する名詞の語彙を無作為に決定して、その語彙に関連する動詞の語彙を決定する。
2. モデルは、選ばれた名詞の語彙の特徴情報を取得する。
3. 選ばれた語彙が名詞であることを表す親の意図情報を取得し、幼児の意図との照らし合わせを行う。
4. モデルは語彙の特徴情報と親の意図情報から名詞語彙のラベルを予測する。
 - (a) 正解の場合、正の報酬を与える。
 - (b) 不正解の場合、負の報酬を与える。
5. モデルは、選ばれた動詞の語彙の特徴情報を取得する。
6. 選ばれた語彙が動詞であることを表す親の意図情報を取得し、幼児の意図との照らし合わせを行う。
7. モデルは語彙の特徴情報と親の意図情報から動詞語彙のラベルを予測する。
 - (a) 名詞語彙のラベル予測結果 (4.a) と動詞語彙のラベル予測結果 (7.a) の両方が正解の場合、正の報酬を与えて、エピソードを終える。
 - (b) 名詞語彙のラベル予測結果 (4.a) と動詞語彙のラベル予測結果 (7.a) の片方でも不正解の場合、負の報酬を与えて、2 に戻る。
8. 試行回数が打ち切りの回数に到達した場合、エピソードを終える。

上記のような未知の物体の特徴情報を学習し、物体の名称を予測する仕組みを構築するため、強化学習と再帰型ニューラルネットワークを用いた。

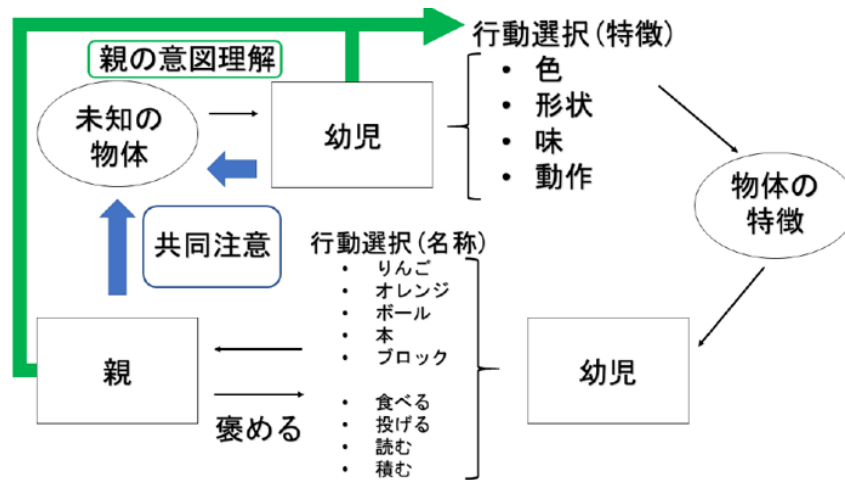


図 4.5 幼児語彙獲得モデルの一連の流れ

4.5 行動空間

提案手法である幼児語彙獲得モデルでは未知の物体に対してどの特徴量を選択して取得するかという特徴量の選択と選択した特徴量から未知の物体の名称を推定する 2 種類の行動について最適化が行われる。前者の特徴量選択については色、形、味、動きの 4 種類の特徴情報を用意した。後者の名称選択については名詞については“りんご”、“ボール”、“本”、“ブロック”の 4 種類を用意し、動詞については“食べる”、“投げる”、“読む”、“積む”の 4 種類を用意した。これらの 8 種類の名称に加えて、未知の物体に対する確信度が高くない時を想定した選択肢として“分からない”を用意し、計 9 種類の行動が選択できる。なお、“分からない”という行動選択に対しても他の名称と同様に扱うと“分からない”が唯一正解出来ない選択肢として確信度の高さに関わらず選択されなくなることが予期されたため、“分からない”が選択された場合に小さい正の報酬を与えるようにした。加えて、“分からない”以外の名称選択時に不正解だった場合にも小さい負の報酬を与えるようにした。これにより、確信度が低い場合には不正解になって負の報酬が与えられるのを避けるために“分からない”を選択し、確信度が高い場合にはより大きな報酬を手に入れるために名称を当てにいくような方策の学習が期待される。

4.6 物体の特徴情報

本研究では画像や動画から取得した実データで構成された特徴情報（色、形、味、動き）を用いて物体の特徴リストを作成した。リストにはニューラルネットワークによって抽出した特徴量ではなく、人間が感じ取っている物体の特徴を実物体から抽出し、それを特徴量として用いた。ここでの仮定は、幼児がこれらの特徴量の全部を一度に調べられず、一つずつ順番に特徴を取得するというものである。これは、人間の特徴量抽出の遅さをモデル化したものである。それぞれの特徴量抽出に用いた手法を以下に示す。なお、りんご一つをとっても日本人にとって一般的である赤いりんごだけでなく緑色のりんごも存在する。このような個体差を考慮するため、色と形特徴量の抽出で用いる画像を物体毎に 3 種類ずつ用意した。

4.6.1 色特徴量

本研究では、物体の色の特徴量としてカラーヒストグラムを用いる。カラーヒストグラムは画像認識で広く用いられる特徴量であり、ピクセル毎の RGB 値の頻度のヒストグラムである。このヒストグラムを特徴量とする機械学習により、類似した色合いの画像に分類できる。すなわち、カラーヒストグラムを用いることで赤や青等のような単純な色だけでなく、言葉で形容することが困難な色や複数の色を色特徴量として表現できる。

なお、本研究では R 値, G 値, B 値がそれぞれ 256 色のままヒストグラムを作成すると、色数が非常に大きくなるため、64 色に圧縮する。また、図 4.6 から分かる通り、画像の中心以外の余分な情報を取得するとカラーヒストグラムが上手く形成されない場合があるため、図 4.7 のようにマスクをかけて画像中心部の特徴のみを取得するようにする。

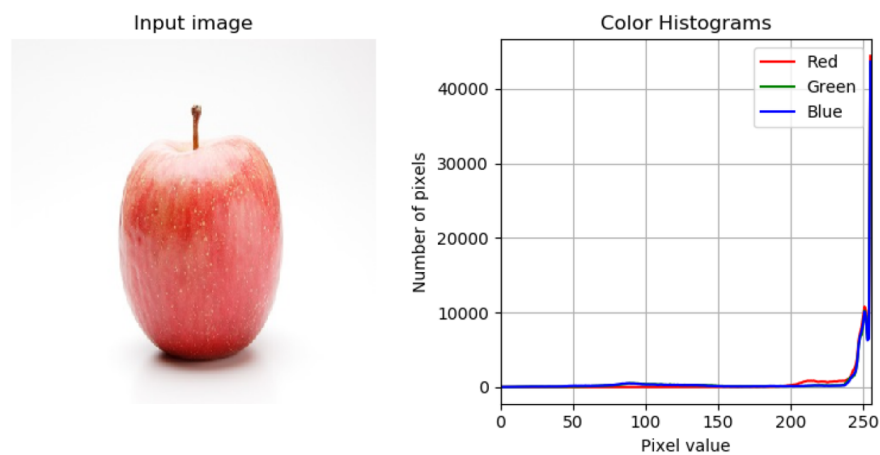


図 4.6 入力画像とカラーヒストグラム

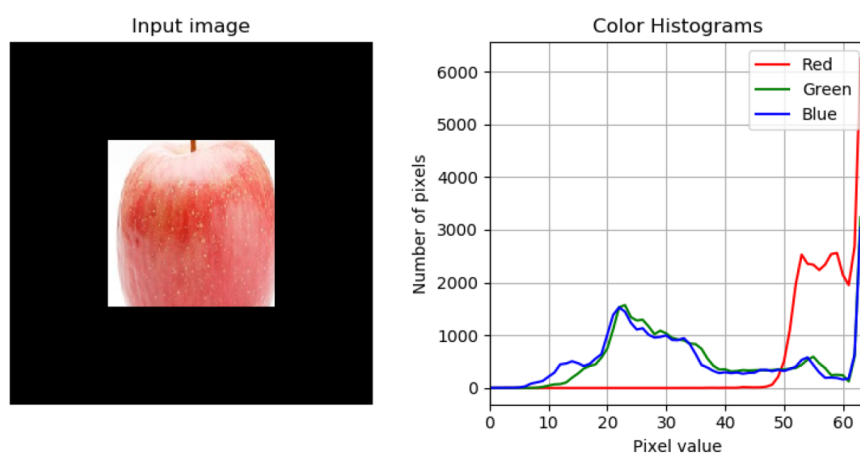


図 4.7 マスクした入力画像とカラーヒストグラム

4.6.2 形特徴量

Histogram of Oriented Gradient(HOG) は局所領域の輝度勾配の方向からヒストグラムを作成し、画像の特徴情報を抽出する手法で、図 4.8 のように物体の大まかな形状を表せる。すなわち、HOG を特徴情報として用いることで人間が視覚的に捉えている物体の形状を機械が学習できる。また、HOG は輝度勾配の情報をを用いることから、画像をリサイズしてサイズを一律に変更してからヒストグラムを作成することで、特徴量の次元を統一できる。

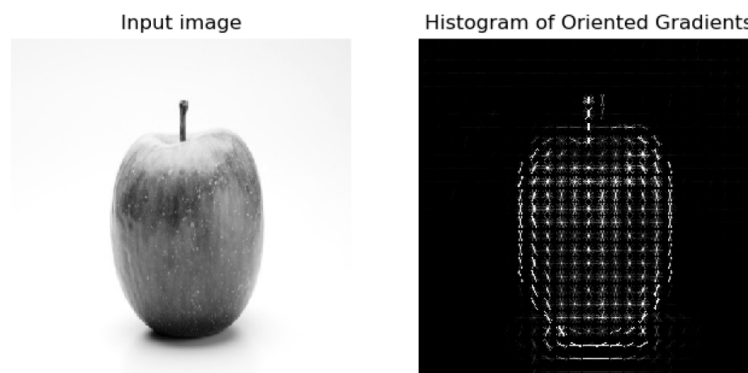


図 4.8 入力画像と HOG

4.6.3 味特徴量

人間は食べ物に含まれる化学物質を舌で感知することで味を感じる。味の種類は甘味、酸味、塩味、苦味、旨味の五味が基本味だと考えられている。そこで、本研究では文部科学省によって作成された日本食品標準成分表 [14] より、甘味、酸味、塩味、苦味、旨味のそれぞれの味を構成する化学物質の値を取得し、特徴情報とする。すなわち、物体が食品である場合には味特徴情報の取得によってその物体がどのような味であるかを機械が学習できる。それぞれの味に対応する化学物質は以下の通りである [15, 16, 17, 18, 19]。

- 甘味・・・フルクトース、グルコース、ガラクトース、スクロース、マルトース、ラクトース、トレハロース、ソルビトール、マンニトール、グリシン、アラニン、スレオニン、プロリン、セリン
- 酸味・・・有機酸、アスパラギン酸
- 塩味・・・食塩相当量
- 苦味・・・カフェイン、フェニルアラニン、チロシン、アルギニン、イソロイシン、ロイシン、バリン、メチオニン、ヒスチジン
- 旨味・・・グルタミン酸

4.6.4 動き特徴量

Dense Trajectories

動き特徴量については Dense Trajectories を用いる [20]. Dense Trajectories は時系列方向に特徴点追跡を行い、その軌跡上で Histogram of Oriented Gradients(HOG), Histogram of OpticalFlow(HOF), Motion Boundary Histograms(MBH) の局所特徴量を計算し、ベクトル化する手法である. すなわち, Dense Trajectories を用いることで人や動物などの動きを特徴量ベクトルとして取得できる.

HOF は画素の移動を表すオプティカルフローを前後の 2 フレームから求めた後、局所領域のフローの方向と強度からヒストグラムを作成し、動画の特徴量を抽出する. 一方で, MBH は HOF を求めるのに用いられたオプティカルフローを微分して動画の特徴量を抽出する. Dense Trajectories では x 方向微分の MBHx と y 方向微分の MBHy を用いる.

Moments in Time

Dense Trajectories に渡す動作の動画には Moments in Time Dataset を用いる [21]. Moments in Time Dataset は三秒の動作の動画に行動のラベルが付与されている行動認識用のデータセットである.

このデータセットから物体に関係のある動作の動画 (リンゴ-eating, 本-reading など) を DenseTrajectories に渡し、その結果求められたベクトルを特徴情報として用いる. つまり, DenseTrajectories と Moments in Time Dataset の組み合わせによって取得した動き特徴量は物体をどのようにして用いるかを表現している.

4.7 親の意図理解

ここまでは親の意図を考慮しない幼児語彙獲得モデルの基本的な枠組みについて説明した. ここでは親の意図理解機構の構成について説明する. 親の意図理解はモデルの名称候補に動詞を加え、親が名詞と動詞のどちらを幼児に教えたかを幼児に入力として与える. すなわち, 正解ラベルが名詞と動詞のどちらであるかという情報がモデルに与えられる. しかし, 幼児は常に共同注意によって親と同一の物体を注視している訳ではないこと, 物体の特徴を 100% 親と同じように捉えているわけではないこと (例えば, 親は本が読み物であることに注目してほしいにもかかわらず, 幼児は本の色や絵柄に注目してしまうような事態が起きてしまうこと) などから, 親の意図を完全に理解できない. この現象を再現するために, 親だけでなく幼児に対しても幼児の意図が存在すると考え, 幼児と親の意図が一致した場合に親が幼児に対して肯定的なアクションを返す機構とした. 具体的には, 幼児がランダムに名詞と動詞のどちらかを選択し (幼児の意図), 親は自身の意図と照らし合わせる. 幼児の意図と親の意図が一致していた場合には, 親は喜びと驚きのどちらかの表情をランダムに幼児に見せる. 一方で, 幼児の意図と親の意図が一致していなかった場合には親は喜び, 驚き, 悲しみ, 怒り, 嫌悪, ニュートラルのいずれかの表情をランダムに幼児に見せる. 幼児の意図と親の意図が一致しなかった場合にも喜びや驚きの表情を見せるようにしている理由としては, 親は幼児が間違った行動を行った場合にも行動自体に価値を見出して喜んでしまう可能性があることに加えて, 前述したように幼児が親の意図を 100% 理解できないことをモデル化するためである. 図 4.9 はこのような親の意図理解機構を加えたモデルの構造を表している. なお, 図 4.9 における親の意図は幼児が捉えた親の意図であり, 実際の親の意図と一致しているとは限らない.

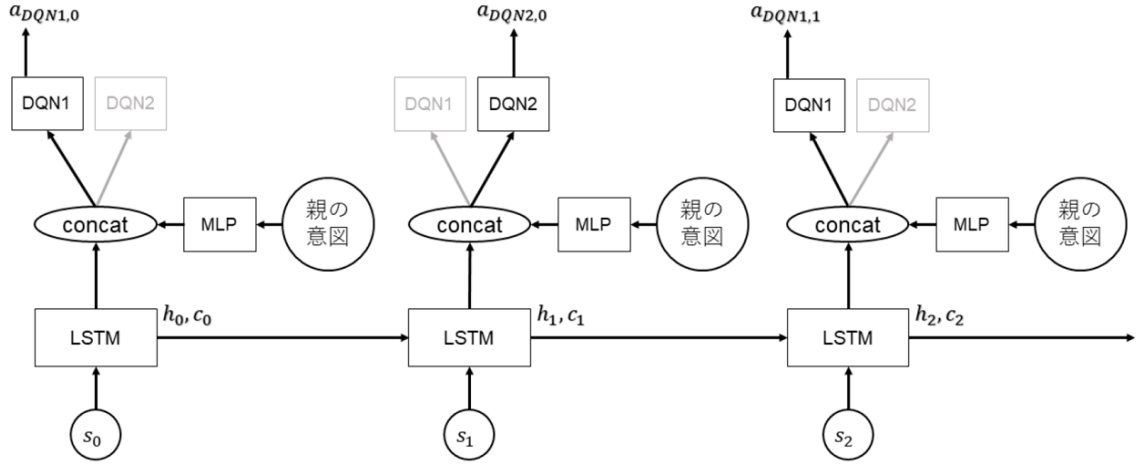


図 4.9 親の意図理解機構を加えた幼児の語彙獲得モデル

子供が親が選択した品詞と同じ品詞に注目出来なかった際に嫌悪的刺激を与える。また、親が名詞を選択した際に幼児が後述する色、形状、味、動作の物体の特徴から動作の特徴に注目した場合、または親が動詞を選択した際に動作以外の色、形状、味の特徴に注目した場合に嫌悪的刺激を与えている。これらは、親と幼児の双方が同一の品詞に注目した場合に学習がより進む共同注視機構が実現している。具体的には、幼児が物体自体の特徴である色、形状、味に注目した場合には名詞の学習を幼児が行うと仮定し、その時に親が選択した品詞が名詞でなかった場合や、逆に幼児が物体自体ではなく物体を用いる動作の特徴に注目した場合には動詞の学習を幼児が行うと仮定し、その時に親が選択した品詞が動詞でなかった場合には嫌悪的刺激を与えるようにした。なお、親と幼児の品詞選択が一致した場合に正の報酬を与える機構にする場合、幼児が最適な特徴選択を行うと誤認してしまうため、報酬の与え方として嫌悪的刺激を利用している。

品詞情報を順に獲得する様子をモデル化するため、親の意図は 1 エピソードの中で名詞から動詞に変化させる。モデルは親の意図の変化を親の表情から理解し、名詞のつぎに関連する動詞を出力することが正の報酬につながることを学習する。第 6 章 本実験では、名詞のつぎに関連する動詞を出力するモデルの学習結果を示すことで、モデルが品詞情報を順に獲得することを示す。

4.7.1 親の表情

親の表情については JAFFE(The Japanese Female Facial Expression) データセット [22] を用いた。JAFFE は 10 人の日本人女性について 7 つの表情 (喜び、驚き、悲しみ、怒り、嫌悪、恐怖、ニュートラル) を収録したデータセットである。図 4.10 は JAFFE データベースの一例である。本モデルでは 10 人の女性の中の 1 人の女性をターゲットに、7 つの表情の内、恐怖を除いた喜び、驚き、悲しみ、怒り、嫌悪、ニュートラルの 6 種類の表情を使用した。親の意図理解機構については、この親の表情を入力として畳み込みニューラルネットワークを用いることで幼児が捉えた親の意図を出力している。



図 4.10 使用する親の表情例

第 5 章

予備実験

本章では、名詞の語彙を獲得する幼児モデルに対する実験結果を示す。

前述したように、名詞は他の品詞に比べ単純で基本的であることから、幼児が一番最初に学ぶ品詞であることが多い。本章では、名詞の語彙を獲得する幼児モデルを構築し結果を確認することで、幼児語彙獲得モデルの構造や学習機構の検証を行う。

5.1 実験設定

この実験では、動詞の語彙を取り除き名詞の語彙のみをモデルに与えた。したがって、語彙の名称選択は名詞の 4 つと「分からない」の計 5 つである。実験の条件として、以下の流れを 1 エピソードとした。また、打ち切りの回数を 5 回 (5 ステップ)、エピソード数を 10000 エピソードとし、テストを 50 エピソード毎の訓練後に 50 エピソード行った。

1. 学習する名詞の語彙を無作為に決定する。
2. モデルは、選ばれた語彙の特徴情報を取得する。
3. 選ばれた語彙が名詞であることを示す親の意図情報を取得し、幼児の意図との照らし合わせを行う。
4. モデルは語彙の特徴情報と親の意図情報から名詞語彙のラベルを予測する。
 - (a) 正解の場合、正の報酬を与えて、エピソードを終える。
 - (b) 不正解の場合、負の報酬を与えて、2. に戻る。
5. 試行回数が打ち切りの回数に到達した場合、エピソードを終える。

5.2 実験結果と考察

図 5.1 に実験結果を示す。

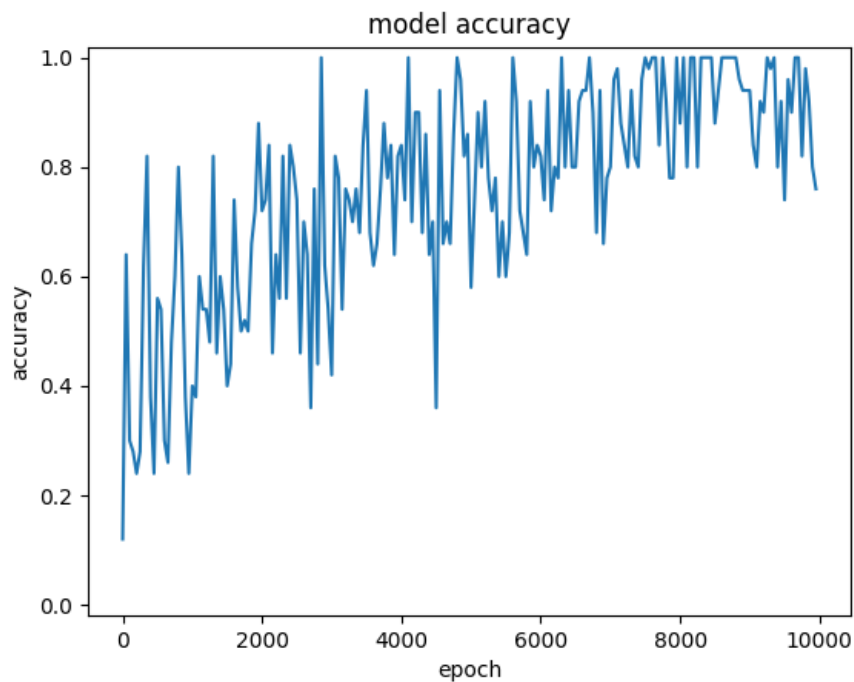


図 5.1 名詞の語彙を獲得する幼児モデルの正答率の遷移

縦軸が正答率 (accuracy)、横軸がエピソード数である。実験の結果、名詞のみを学ぶ幼児の語彙獲得モデルの正答率は 0.8 (80%) 程度となっており、名詞の語彙を学習できていることがわかる。その一方で 4000 エピソード付近で正答率が大きく下がってしまう場合がある。本研究では強化学習の探索アルゴリズムに焼きなまし ϵ -Greedy アルゴリズムを用いているため、一定確率でランダムな行動をとって学習を行う。その結果、学習によって方策が変わったことが原因と考えられる。焼きなまし ϵ -Greedy アルゴリズムは学習終盤にはほぼランダムな行動をとらず、Greedy アルゴリズムをとるため、正答率は 0.8 付近を保つことができている。

第 6 章

本実験

本章では，提案する品詞情報の獲得を考慮した幼児の語彙獲得モデルの実験結果を示す．

6.1 品詞情報の獲得を考慮した幼児の語彙獲得モデルの実験

6.1.1 実験設定

本実験では，学習方法を 4.4 学習方法で述べた内容として，名詞語彙のつぎに動詞語彙のラベル予測を行った．打ち切り回数を 5 回（5 ステップ），エピソード数を 10000 エピソードとし，テストを 50 エピソード毎の訓練後に 50 エピソード行った．また，与える特徴情報，語彙を 4.5 行動空間で述べた内容とした．また，親の意図は同エピソード内で名詞から動詞に変化させた．

本実験では，第 5 章 予備実験結果のモデルパラメータを事前学習したモデルと，事前学習なしのモデルの正答率を比較する．事前学習ありのモデルは，名詞語彙の学習の際に獲得した名詞バイアスやその他バイアスをあらかじめ獲得した状態で学習を始めて，各バイアスを干渉させずにうまくコントロールして名詞と動詞の語彙を区別して学習できるのか実験する．事前学習なしのモデルは，名詞と動詞の両方の品詞バイアスやその他バイアスを徐々に獲得していく中で，バイアスのコントロールにより品詞情報を獲得し，名詞と動詞の語彙を区別して学習できるのか実験する．

6.1.2 実験結果と考察

事前学習ありのモデルの実験結果を図 6.1，事前学習なしのモデルの実験結果を図 6.2 に示す．

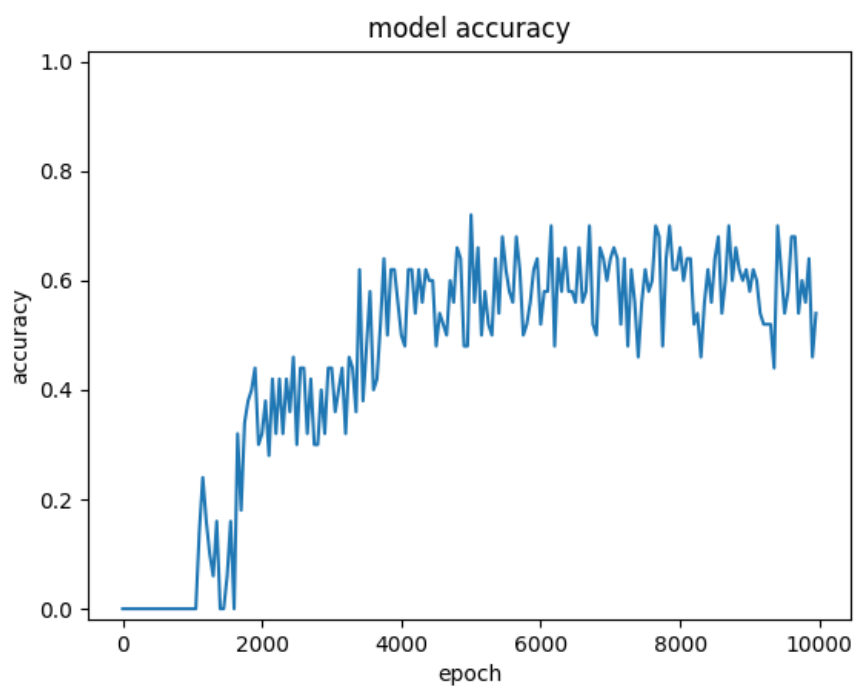


図 6.1 品詞情報の獲得を考慮した幼児モデルの正答率の推移（事前学習あり）

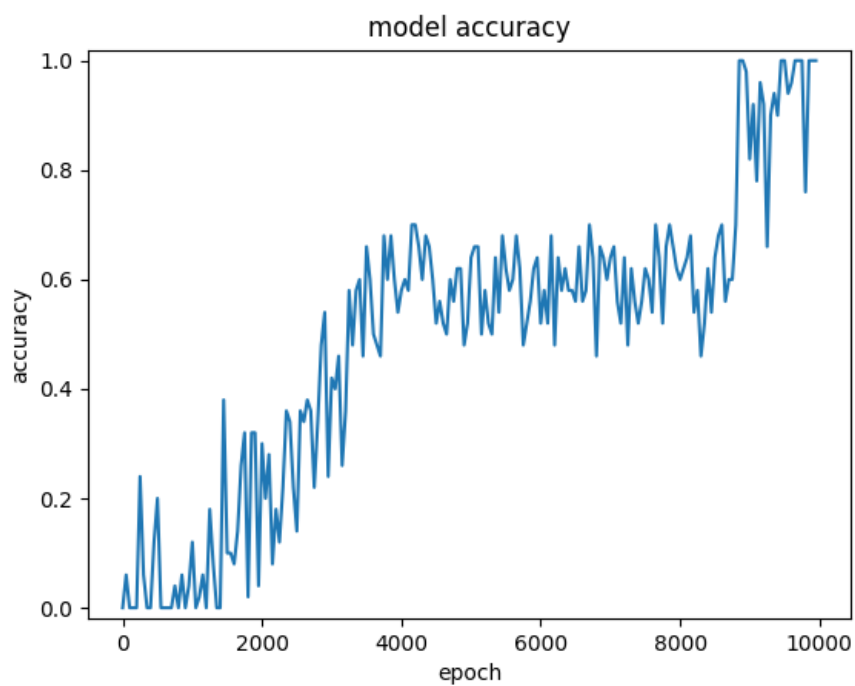


図 6.2 品詞情報の獲得を考慮した幼児モデルの正答率の推移（事前学習なし）

縦軸が正答率（accuracy）、横軸がエピソード数である。事前学習ありのモデルは 2000 エピソード前後から 4000 エピソード前後にかけて徐々に正答率を 0.6（60%）前後に伸ばしていき、9000 エピソード前後では正答率を 1（%）前後まで伸ばしている。一方で事前学習なしのモデルは、同様に 2000 エピソード前後から 4000 エピソード前後にかけて徐々に正答率を 0.6（60%）前後に伸ばしているが、学習終了まで正答率が大きく変わらなかった。事前学習ありのモデルの正答率が伸び悩んだ理由として、名詞語彙が「apple」、動詞語彙が「eat」の場合がほぼ正答できず、打ち切り回数上限の 5 ステップ目では「分からない」を出力していたことが挙げられる。これは、名詞語彙を推定するタスクでの事前学習で獲得した名詞バイアスと本実験での名詞語彙のつぎに動詞語彙を推定するタスクで獲得した動詞バイアスが干渉し合い、特徴選択や名称予測がうまくできなかったと考えられる。

事前学習ありのモデルの特徴選択の推移を、名詞語彙の場合は表 6.1、動詞語彙の場合は表 6.2 にそれぞれ示す。

表 6.1 事前学習ありモデル 特徴選択 名詞語彙の場合

エピソード	特徴選択 1 ステップ目	特徴選択 2 ステップ目	特徴選択 3 ステップ目
0～2000	色・味	動き・味	動き
2000～4000	色	動き・形	動き・形
4000～10000	色・形	動き・味	なし

表 6.2 事前学習ありモデル 特徴選択 動詞語彙の場合

エピソード	特徴選択 1 ステップ目	特徴選択 2 ステップ目	特徴選択 3 ステップ目
0～2000	動き・色・味	動き・色・味	動き・色・形
2000～4000	動き・色・味	動き・味・色	動き
4000～10000	動き	動き・形	なし

事前学習なしのモデルの特徴選択の推移を、名詞語彙の場合は表 6.3、動詞語彙の場合は表 6.4 にそれぞれ示す。

表 6.3 事前学習なしモデル 特徴選択 名詞語彙の場合

エピソード	特徴選択 1 ステップ目	特徴選択 2 ステップ目	特徴選択 3 ステップ目
0～2000	色・味	動き・味	動き
2000～4000	色	動き・形	動き・形
4000～9000	色・味	動き・形	なし
9000～10000	味	動き・色	なし

表 6.4 事前学習なしモデル 特徴選択 名詞語彙の場合

エピソード	特徴選択 1 ステップ目	特徴選択 2 ステップ目	特徴選択 3 ステップ目
0~2000	動き・色・味	動き・色・味	動き・色・形
2000~4000	動き・色・味	動き・味・色	動き
4000~9000	動き	動き・形	なし
9000~10000	動き	動き	なし

特徴選択が「なし」の場合は、それ以前のステップで名称を正しく推定できエピソードが終了していることを示す。事前学習の有無に関わらず、エピソードが進んでいくうちに特徴選択に規則性が生まれ、それに伴い正答率も上がっている。本モデルは、選択した特徴を根拠に未知の物体の名称を推定するタスクを通じて、語彙が示す物体や動作がどのような特徴を持つかを学習している。これは、特徴と名称の対称性バイアスを形成しているといえる。また、名詞語彙における特徴選択では動きを選んでいく場合が多い。これは、名詞語彙のつぎに動詞語彙のラベル予測を行う本タスクを学習していくうちに、動きから名詞を予測する対称性バイアスも形成されていることがわかる。

以上より、名詞語彙と動詞語彙を順に学習することで、品詞を区別しつつ関連する語彙の獲得が行えることがわかった。なお、名詞語彙と動詞語彙の学習は完全に切り離して行くと、本実験で示した事前学習ありのモデルのようにバイアス同士が干渉し合うことで語彙獲得がうまくいかない。実世界のように名詞語彙と動詞語彙の学習を完全に切り離さず、徐々に名詞と動詞の両方の品詞バイアスや対称性バイアスを獲得していくことで、本実験で示した事前学習無しのモデルのように各バイアスをコントロールしつつ語彙獲得できる。

6.2 品詞情報を同時に獲得する幼児の語彙獲得モデルの実験

この実験では、前節で述べた実験とは異なり、名詞語彙と動詞語彙を同時にラベル予測を行った。同時に学習しても品詞の区別は行えるのか、心理学の研究で示されている品詞情報の獲得順序の重要性を確認する。

6.2.1 実験設定

実験の条件として、以下の流れを 1 エピソードとした。また、打ち切りの回数を 5 回（5 ステップ）、エピソード数を 10000 エピソードとし、テストを 50 エピソード毎の訓練後に 50 エピソード行った。動作の元である物体を親の指示対象の主とするため親の意図を名詞とした。モデルの出力の全結合層を 1 つから 2 つに増やし、名詞語彙の予測出力と、動詞語彙の予測出力の 2 つを行えるように変更した。

1. 学習する名詞の語彙を無作為に決定して、その語彙に関連する動詞の語彙を決定する。
2. モデルは、選ばれた語彙の特徴情報を取得する。
3. 選ばれた語彙が名詞であることを示す親の意図情報を取得し、幼児の意図との照らし合わせを行う。
4. モデルは語彙の特徴情報と親の意図情報から名詞語彙のラベルと関連する動詞語彙のラベルを予測する。
 - (a) 正解の場合、正の報酬を与えて、エピソードを終える。
 - (b) 不正解の場合、負の報酬を与えて、2. に戻る。
5. 試行回数が打ち切りの回数に到達した場合、エピソードを終える。

6.2.2 実験結果と考察

図 6.3 に実験結果を示す．縦軸が正答率（accuracy），横軸がエピソード数である．

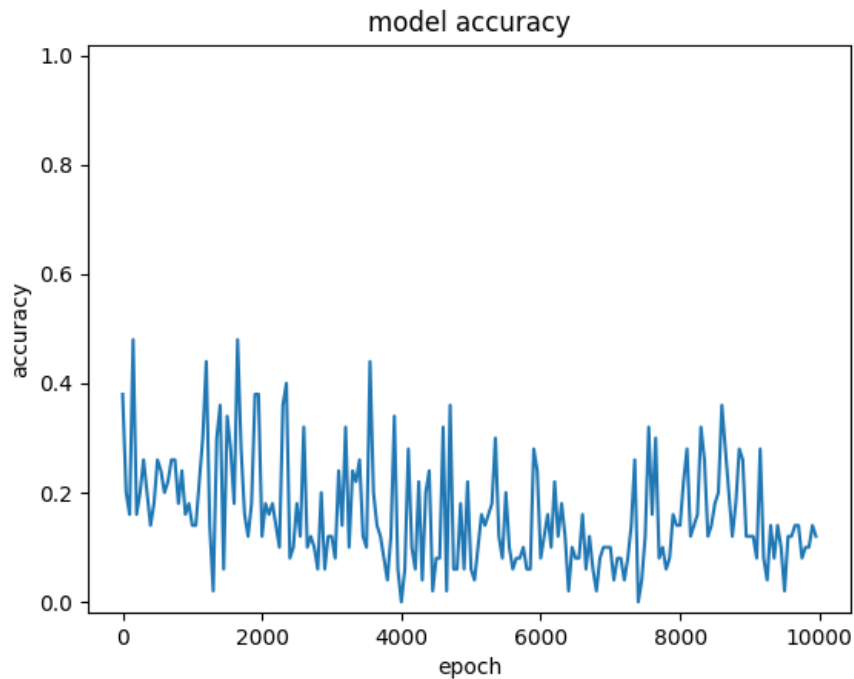


図 6.3 品詞情報を同時に獲得する幼児モデルの正答率の推移

正答率は 0.2（20％）以下になってしまい，学習できないことが確認できた．これは，名詞と動詞の区別をするには動作主である名詞を学習し，その関連として動詞を学習するという順序が重要であり，同時に学習するのは難しいといえる．名詞と動詞だけでなく，名詞と動詞の結び付けを行う助詞も学習できれば，名詞と動詞の関連を同時に学習することができることが考えられる．

第7章

結論

本研究では、田口の提案した幼児語彙獲得モデルを参考に、獲得してほしい語彙を名詞から動詞に変化させる親の意図を幼児が理解することで、品詞情報を順に獲得する幼児の語彙獲得モデルを提案した。提案モデルは、オペランド連合学習を模すために強化学習、幼児の心的状態の表現と状態数の大きさに対応するためにニューラルネットワークを用いた。モデルに組み込んだ心理学的要因は、共同注意、親の意図理解であり、特に親の意図理解については親が名詞と動詞のどちらを幼児に学習させたいかという親の意図と、幼児が名詞と動詞のどちらを学習したいかという幼児の意図をそれぞれ用意し、親と幼児の意図が一致した場合には親が喜びか驚きの表情を、一致しなかった場合には、喜び、驚き、悲しみ、怒り、嫌悪、ニュートラルのいずれかの表情を幼児に見せるように設定することで、親の意図を幼児が完全には理解できない事象を再現した。特徴情報は画像や動画等から取得した実データを用いた。

名詞の語彙のみを獲得する幼児モデルに対する予備実験では、モデルが名詞語彙を問題なく学習できた。本実験では、1 エピソードの中で名詞の語彙のつぎに動詞の語彙を獲得する、品詞情報の獲得を考慮した幼児モデルに対する検証を行った。予備実験のモデルパラメータを事前学習したモデルと、事前学習なしのモデルで比較を行い、どちらのモデルでも特徴選択で対称性バイアスや名詞・動詞バイアスが確認された。また、事前学習ありのモデルでは正答率が伸び悩んだ結果に対し、事前学習無しのモデルでは正答率がほぼ1まで伸びた。この実験により、実世界のように名詞の語彙と動詞の語彙を完全に切り離さずに順に学習することで、徐々に各バイアスの獲得と適用するバイアスのバランスコントロールができ、品詞情報を獲得しつつ語彙も獲得できることがわかった。一方で、名詞と動詞を順ではなく同時に予測するタスクでは、正答率が0.2付近を推移し、モデルの学習が困難だった。これは、名詞と動詞の区別をするには動作主である名詞を学習した後に、その関連である動詞を学習するという順序が重要であること、名詞と動詞の結び付けを行う助詞も学習できれば、同時に学習できる可能性があることを示唆している。

以上より、本研究で提案した幼児語彙獲得モデルは、名詞のつぎに動詞を学習する順序によって、学習バイアスをコントロールし、名詞と動詞の品詞情報を獲得しつつ語彙も獲得できることが確認された。

本研究での実験で扱った語彙数は、実世界の幼児が学習している語彙数と比べて非常に少ない。語彙数やその特徴を増やした場合、どのように各バイアスがコントロールされ学習が進んでいき、品詞情報の獲得に影響するのか検証する必要がある。また、本研究で行われた実験によって、名詞と動詞を同時に学習すると区別ができず品詞情報の獲得が困難であること、名詞のつぎに動詞を学ぶことで品詞情報を順に獲得できることがわかった。しかし、実世界の幼児は本研究のように単語だけを聞いて語彙獲得を行っているのではなく、文章から語彙獲得を行っていると考えられる。名詞と動詞を関連付ける助詞を学習させることで、名詞と動詞を同時に、かつ文章を学習するモデルが構築できると考えられる。

謝辞

本研究を進めるにあたり，ご指導いただいた南泰浩教授，古賀久志准教授には感謝申し上げます．また，研究に関する多くのご助言，ご助力をいただいた南研究室の皆様には感謝いたします．

参考文献

- [1] 長井隆行, 青木達哉, 中村友昭, ” 言語を理解するロボット実現への確率ロボティクスのアプローチ, ” システム／制御／情報, 2016.
- [2] 宮澤和貴, 青木達哉, 堀井隆斗, 長井隆行, ” 統合認知モデルによるロボットの概念・行動・言語の同時学習, ” 人工知能学会全国大会論文集, 2019.
- [3] 田口真輝, ” 心理学的要素を考慮したニューラルネットワーク強化学習による幼児語彙獲得のモデル化, ” 2020 年度情報・ネットワーク工学専攻コンピュータサイエンスプログラム修士論文, 2021.
- [4] 大藪泰, ” 共同注意という子育て環境, ” 早稲田大学総合人文科学研究センター研究誌, 2019.
- [5] Baldwin, D.A., ” Infants’ s contribution to the achievement of joint reference, ” ChildDevelopment, 1991.
- [6] Baldwin, D.A., ” Infants’ s ability to consult the speaker for clues to word reference, ” journal of Child Language, 1993.
- [7] Gentner, D. and Boroditsky, L., ” Individuation, relativity, and early word learning, ” Language acquisition and conceptual development, 2001.
- [8] Tomasello, M., ” ことばをつくる: 言語習得の認知言語学的アプローチ, ” 辻幸夫 (訳), 野村益 (訳), 出原健一 (訳), 慶應義塾大学出版会, 2008.
- [9] 片平健太郎, ” 学習の理論から強化学習, 計算論モデリングへ, ” 特集古くて新しい学習心理学, 2017.
- [10] 牧野貴樹, ほか, ” これからの強化学習, ” 森北出版株式会社, 2016.
- [11] 木村元, Kaelbling, L.P., ” 部分観測マルコフ決定過程下での強化学習, ” 人工知能学会誌, 1997.
- [12] Garivier, A. and Capp e, O., ” The KL-UCB algorithm for boundedstochastic banditsand beyond, ” InProceedings of the 24th AnnualConference on Learning Theory (COLT), JMLR Workshop and Conference Proceedings, 2011.
- [13] Hochreiter and Schmidhuber, J., ” Long short-term memory, ” Neural computation, 1997.
- [14] 科学技術・学術政策局政策課資源室, ” 日本食品標準成分表 2015 年版 (七訂), ” <http://www.mext.go.jp/amenu/syokuhinseibun/1365297.htm>, (accessed 2023-01-20).
- [15] 二宮恒彦, ” アミノ酸の呈味に関する研究, ” 調理科学, 1968.
- [16] 前橋健二, ” 甘味の基礎知識, ” 日本醸造協会誌, 2011.
- [17] ” 食とアミノ酸, ” AJINOMOTO CO.,INC, <https://www.ajinomoto.co.jp/amino/chikara/shoku.html>, (accessed 2023-01-20).
- [18] 中西載慶, ” アミノ酸はおいしい旨い・甘い・苦い・酸っぱい, ” <https://www.nodai.ac.jp/research/teacher-column/0011/>, (accessed 2023-01-20).
- [19] ” 有機酸とはどんなもの? 種類と使用されている食品, ” ANDY INC., <https://taberugo.net/242>,

(accessed 2023-01-20).

- [20] Wang, H., Klaser, A., Schmid, C. and Liu, C., “Action Recognition by Dense Trajectories, ” In IEEE Conference on Computer Vision & Pattern Recognition, 2011.
- [21] Monfort, M., Andonian, A., et al. “Moments in time dataset: one million videos for event understanding, ” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
- [22] Lyons, M.J., Akamatsu, S., Kamachi, M. and Gyoba, J. “Coding Facial Expressions with Gabor Wavelets,” 3rd IEEE International Conference on Automatic Face and Gesture Recognition, 1998.