

幼児エージェントによるバイアスの形成と 言語獲得過程におけるその効果

篠原 修二[†] 田口 亮[†] 桂田 浩一[†] 新田 恒雄[†]

[†]豊橋技術科学大学大学院工学研究科 〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘1-1
E-mail: †{shinohara,taguchi}@vox.tutkie.tut.ac.jp, ††{katurada,nitta}@tutkie.tut.ac.jp

あらまし 人間の幼児は言葉を学習する際、形状類似バイアス、相互排他性バイアスなどの制約を利用して効率よく学習を行うことが知られている。本研究では、経験を通して言葉を学習する幼児エージェントのモデルを構築し、どのような条件の下で各種バイアスの形成という現象が見出されるのかについて分析した。数値実験の結果、コミュニケーションの観点から導かれるある条件を導入すれば、IA は言葉を学んでいく過程の中で、バイアスを自律的に形成し学習を加速するようになることが示された。

キーワード 幼児エージェント、言語獲得、形状類似性バイアス、相互排他性バイアス

Formation of Learning Bias by Infant Agent and the Effect on Language Acquisition Process

Shuji SHINOHARA[†], Ryo TAGUCHI[†], Koichi KATSURADA[†], and Tsuneo NITTA[†]

[†] Graduate School of Eng., Toyohashi Univ. of Tech. Hibirigaoka 1-1, Tenpaku-cho, Toyohashi-shi, Aichi, 441-8580 Japan

E-mail: †{shinohara,taguchi}@vox.tutkie.tut.ac.jp, ††{katurada,nitta}@tutkie.tut.ac.jp

Abstract It is found that human infants acquire language effectively by utilizing learning biases such as a shape bias and a mutual exclusivity bias, etc.. In this study we explore whether those biases emerge in learning process of artificial infant agent which forms concepts based on its own experiences. We model an intelligent agent by implementing a learning method, which is derived from the assumption that an agent has consistent beliefs in communication, and investigate experiments. Our simulation results show that an infant agent develops learning biases and comes to acquire language more efficiently.

Key words infant agent, language acquisition, shape bias, mutual exclusivity bias

1. はじめに

人間の幼児は1歳半前後の語彙爆発と呼ばれる時期から、驚異的なスピードで新しい言葉を覚えていく。その際幼児は、一事例あるいは数個の事例を示されただけで言葉の意味を正しく推論できることが知られている。これは即時マッピングと呼ばれる。

しかし例えば普通名詞は、無数の対象に適用される。有限個の経験事例から無限の対象に言及する言葉の意味あるいは適用範囲を正しく推論することは、論理的には不可能である。また、あるラベルを初めて聞く幼児にとっては、そもそもそれが普通名詞なのか固有名詞なのかということさえわからない[2],[3]。このような状況において、なぜ人間の幼児は即時マッピングを行うことができるのか。この問題に対して、幼児は全ての可能

性を吟味するのではなく、ある種の制約を利用して、可能性を限定することで効率よく推論を行っている、という考えが提唱されている[4],[5]。

例えば幼児は新奇なラベルを聞く際、それを固有名詞ではなく普通名詞と見なす傾向がある[4]。これは事物カテゴリーバイアスと呼ばれる。普通名詞と見なす場合、どのような対象群にラベルを適用するのかという問題があるが、幼児は形状の類似した対象にラベルの適用範囲を拡張する。これは形状類似バイアスと呼ばれる[5]~[8]。他にも、例えば既知の事物と新奇な事物を目の前にして新奇なラベルを聞いた時、幼児はそのラベルを新奇な事物の名前でであると見なす傾向がある。これは、相互排他性バイアスと呼ばれる[9],[10]。このバイアスは、指差しなどが無い場合に指示対象の同定に役立つ。幼児はこのようなバイアスを持つことで、効率よく言語を学習する。

しかしバイアスの適用が、必ずしも言語学習に有効に作用するわけではない。例えば事物カテゴリーバイアスは、固有名詞を学習する際に妨げになる。また相互排他性バイアスは、「りんご」と「果物」のように、一方が他方を包含する関係にある言葉の学習を妨げる [2], [3]。形状類似バイアスにも問題がある。このバイアスは、幼児の学習すべき言語が形状に関して構造化されていれば学習を加速するが、他の属性、例えば色に関して構造化された言語であれば学習を阻害することになる^{(注1)(注2)}。つまり幼児が効率よく学習を達成するためには、発達段階や環境に応じて適切なバイアスを形成し、またその適用を調整していく必要がある。実際幼児はバイアスをむやみに適用するわけではない。例えば事物カテゴリーバイアスに関して、新奇な対象に新奇なラベルが付けられる時にはそのラベルを普通名詞と解釈するが、既知の対象に新奇なラベルが付けられる場合にはそれを固有名詞と見なす^(注3) [8]。

本稿では、バイアスは言語を学習する過程において形成されるという立場に立ち、幼児エージェント（以下 IA）の単純なモデルを構築することで、バイアスを形成するために必要な能力は如何なるものかに焦点を当てて考察する。

2. 幼児エージェントの学習

本稿では、経験を通して言語を学習していく幼児エージェント（以下 IA と呼ぶ）をモデル化する。本節では、モデルの概要説明、および本稿で用いる用語の定義を行う。

2.1 言語

まず色属性の集合を $C = \{c_1, c_2, c_3, \dots\}$ 、形属性の集合を $F = \{f_1, f_2, f_3, \dots\}$ と表記する。モデルでは簡単のため、対象は色と形という二つの属性のみを持つこととし、色と形の対で表現するものとする。すなわち対象集合は、 $O = \{o \mid o \in C \times F\}$ と定義される。また全ての対象には一つのラベルが付与されるものとする。ここでラベル集合を $L = \{l_1, l_2, l_3, \dots\}$ と表記する。言語とは、対象とラベルの対応関係のことであると、 $Lang: O \rightarrow L$ と定義する。

2.2 信念

IA は、経験を通して

- 「ラベル l は対象 o である」という言明に対する確信度
 - 「対象 o はラベル l である」という言明に対する確信度
- という二つの確信度を形成する。本稿ではこれらの確信度の形成過程を学習と呼ぶ。前者の言明、例えば「いちごは赤くて丸い」という言明は、赤や丸がいちごに共通の属性であるという

(注1): ここで、形状に関して構造化された言語とは、同じ形の対象は他の属性が違っても同じラベルを持つが、他の属性が同じでも形が異なれば異なるラベルを持つような言語である。

(注2): 我々が実際に使用する普通名詞は、形状類似に基づいたカテゴリーと対応している場合が多いと思われる。例えば、「象」という普通名詞に関して、赤い象を見たことがなくても簡単に想像できるが、犬の形をした象は想像することが困難である。このことから、我々が「象」というラベルの適用範囲を決定する際、色よりも形に強く依拠して判断を行っていることが示唆される。もしそうであれば、形状類似バイアスの形成は、なぜ我々の言語が形状に関して構造化されたのかという言語進化の問題と関連付けて考える必要があると思われる。

(注3): ただし、固有名詞と見なすのは対象が動物の時である。

ことを言い表している。一方後者の言明、例えば「赤くて丸いものはいちごである」は赤や丸という属性がいちごに固有であることを言い表している。この意味で、前者を共通度、後者を固有度と呼ぶことにする。また共通度と固有度の組を信念と呼ぶ。IA は経験を通して信念を形成し、その信念に基づいて様々な判断を行う。

2.3 経験

IA は、経験を通して言語を学習していくが、モデルでは学習の際に、例えばいちごを指差しながら「これはいちごです」と明示的に教えることはしない。IA は目の前にあるいくつかの対象を眺めていて、大人はそれらの内のいずれかの対象について話している、という状況を想定する。具体的には、IA は対象集合の中からランダムに選ばれた N 個の対象を見ていて、例えば「いちご」というラベルを耳にする。 N 個の対象の中に必ずいちごは存在するが、それらの内どれがいちごなのかは教えられない^(注4)。このような状況の中で、IA は「いちご」というラベルを耳にした時に見ていた全ての対象をいちごであると理解するものとする。すなわち、見ている対象の内どれがいちごなのかということについての判断は特に行わない。IA はこのような経験を繰り返すものとする。

3. モデル

本節では、IA の学習過程をモデル化する。ただし本稿では、IA0, IA2, IA3 という学習能力の異なる三種類のモデルを構築する。

3.1 IA0

IA0 は、「様々な対象を眺めている状況の中であるラベルを聞く」という経験を繰り返す。IA0 はそれらの経験を通して、対象とラベルの各組に対して、 l_k を聞いた時に o_i を見た頻度 $V(o_i|l_k)$ と o_i を見た時に l_k を聞いた頻度 $V(l_k|o_i)$ を記憶していく。ただしモデルでは、遠い過去の記憶ほど忘れやすいと仮定し、IA0 が記憶している内的な頻度分布を以下の式を用いて更新する [11]。ただし初期値は全て 0 とする。

$$V(o_i|l_k) \leftarrow V(o_i|l_k) + \alpha[r - V(o_i|l_k)] \quad (1)$$

$$\text{ただし、} r = \begin{cases} 1, & l_k \text{ を聞き } o_i \text{ を見た時} \\ 0, & l_k \text{ を聞き } o_i \text{ を見なかった時。} \end{cases}$$

$$V(l_k|o_i) \leftarrow V(l_k|o_i) + \alpha[r - V(l_k|o_i)] \quad (2)$$

$$\text{ただし、} r = \begin{cases} 1, & o_i \text{ を見て } l_k \text{ を聞いた時} \\ 0, & o_i \text{ を見て } l_k \text{ を聞かなかった時。} \end{cases}$$

ここで α ($0 \leq \alpha \leq 1$) は学習率と呼ばれるパラメータである。

次に、形成された頻度を正規化することで共通度および固有度を計算する。ただし、信念は経験に依存して徐々に形成されると仮定し、ボルツマン分布を用いて正規化する。ここで、共通度と固有度を各々 $B(o_i|l_k)$ 、 $B(l_k|o_i)$ と表記することにし、

$$B(o_i|l_k) = \frac{\exp\{V(o_i|l_k)/\tau\}}{\sum_{o \in O} \exp\{V(o|l_k)/\tau\}} \quad (3)$$

(注4): ここで $N=1$ とすれば、明示的な教示と考えられる。

$$B(l_k|o_i) = \frac{\exp\{V(l_k|o_i)/\tau\}}{\sum_{l \in L} \exp\{V(l|o_i)/\tau\}} \quad (4)$$

と定義する。ここで τ は温度と呼ばれる正定数であり、温度が高い場合には全ての共通度が同程度になるように設定される。一方温度が低い場合には、 V の差がより増幅される。

3.2 IA1

IA0において、頻度 V を記憶するためには、 $2|O||L| = 2|C||F||L|$ のメモリー空間が必要とされる。IA1では、メモリー空間を節約するため、色と形の頻度を別々に記憶することにする。この時必要とされるメモリー空間は、 $2(|C| + |F|)|L|$ になる。IA1は、属性毎に記憶された各頻度に対して、共通度と固有度を別々に計算する。対象の共通度および固有度は、属性毎に求めた値を掛け合わせることで計算する。

3.3 「何？」と「どれ？」

IA2の定義を行う前に、「 o は何？」と「 l はどれ？」という質問に対するIAの回答法について述べる。はじめに「これ何？」という質問に対する回答法を述べる。IAがこの質問に答えるためには、属性値の入力に対しラベルを出力しなければならないが、その際に必要となるのは共通度ではない。例えば赤い対象を示され「これは何？」と質問されて「これはいちごである」と答える際に必要とされるのは、「赤いものはいちごである」という固有度に基づく言明の正当性である。「いちごは赤い」という共通度に関する言明が真であったとしても、「赤いものはいちごである」という言明が真であるとは限らない。つまり「これは何？」という質問に対しIAは、提示された対象 o_i に対して最も固有度の高いラベルを選択するものとする。すなわち、この質問に対するIAの解答 $What(o_i)$ は、

$$What(o_i) = l_k \in L, s.t. B(l_k|o_i) = \max_{l \in L} B(l|o_i) \quad (5)$$

である。

次に、あるラベル l_k を用いて「 l_k はどれ？」と聞かれた場合の回答法について述べる。「何？」の場合とは対照的に「どれ？」という質問の場合、回答方法としていくつかの候補が考えられる。第一の候補として、式(6)に示すような方法が考えられる。つまり、まずIAは各対象について、 $What$ を用いてそれが何かを判断する。次に、もしそれらの内で l_k と判断した対象あればそれを選ぶ。一方提示された全ての対象が l_k ではないと判断した場合、「どれも l_k ではない(null)」と回答する。

$$Which1(l_k) = \begin{cases} o_i, & \exists o_i \in O, s.t. What(o_i) = l_k \\ null, & \forall o \in O, What(o) \neq l_k \end{cases} \quad (6)$$

$Which1$ を採用すると、「どれも l_k ではない」と判断する場合があり得る。この場合、 l_k は何も指示しない無意味なラベルになってしまう。またこの方法を採用すると、選択の際に全ての固有度を参照しなければならなくなるため、計算量が多くなるという難点がある。

上記の問題を解消する方法として、 l_k に関する固有度のみを比較し、それらの内最も値の高いものを選択するという方法が考えられる。

$$Which2(l_k) = o_i \in O, s.t. B(l_k|o_i) = \max_{o \in O} B(l_k|o) \quad (7)$$

他にもこの質問に回答するには、ラベルを入力とし対象を出力する点に着目するならば、共通度を使う、つまり l_k に関する共通度を比較し、それらの内最も値の高いものを選択するという方法が考えられる。

$$Which3(l_k) = o_i \in O, s.t. B(o_i|l_k) = \max_{o \in O} B(o|l_k) \quad (8)$$

これら二つの方法を用いる時、IAは任意のラベル入力に対して何らかの回答を出力できる。また l_k に関する固有度あるいは共通度のみを比較すればよいから、 $Which1$ を用いる場合に比べて計算量が大幅に減少する。

しかしこれらの方法を用いると、 o_i を見せられ「これ何？」と質問されれば l_k と回答するにもかかわらず、「 l_k はどれ？」と聞かれれば $o_j (i \neq j)$ と回答する場合が起こり得る。ある対象に対して、例えば「これはいちごである」と「いちごはこれではない」という信念を同時に持つことは、論理的には矛盾ではないが、このような信念を持つと、コミュニケーションに不都合が生じる場合がある。

例を述べよう。AさんがBさんに「 l_k を取って」と依頼し、Bに l_k を取ってもらう場面を想定する。ただしAとBは全く同じ信念を持つと仮定する。ここでAが使用するラベル l_k とは、彼が取ってもらいたい対象 o_i に対して、彼が「 o_i は何？」と質問された時に回答するラベルである。すなわち $l_k = What(o_i)$ である。一方「 l_k を取って」と依頼されたBが選択する対象 o_j とは、「 l_k はどれ？」という質問をされた際に彼が返答する対象である。仮に彼らが判断に $Which3$ を用いているとすれば、 $o_j = Which3(l_k)$ である。つまり彼らのコミュニケーションが成立する、すなわち $o_i = o_j$ であるためには、 $Which3(What(o_i)) = o_i$ という関係が成立していなければならない^(注5)。しかし l_k が普通名詞の場合、Aは対象 o_i そのものでなくとも彼が l_k だと考える対象、すなわち $What(o_i) = What(o_j)$ を満たす対象 o_j を取ってもらえれば良いだろう。このように条件を弱めるならば、AとBの間でコミュニケーションが成立するためには、 $What(Which3(What(o_i))) = What(o_i)$ という関係が成立すればよい。逆に言えば、この関係を満たさない時、AとBが全く同じ信念を持っていたとしても、二人の間でコミュニケーションが成立しない場合が生じる。

以上に「どれ？」に対する三つの回答法を示したが、どの方法にも問題がある。我々人間が、三つの回答法(あるいは他の方法もあるかもしれないが)の内のどの方法を採用しているのかを詮議することも重要ではあるが、どの方法を採用したとしても重大な問題が残る。どの方法を使っても結果が一致する、あるいは結果に不整合が生じないような信念を持つことができようか? 次のIA2では、どの方法を採用しても結果に不都合が生じないように信念を修正する機構を与える。

(注5): $Which3(What(o_i)) = o_i$ を満たす対象 o_i は、 o_i の属する概念を代表する典型的事例と見なすことができる。

3.4 IA2

IA2はIA1と同様に、共通度と固有度を属性毎に別々に計算する。加えてIA2は、次の三つの条件式を満足するように信念を修正する機構を持つ。

$$\forall o \in O, \forall l \in L, B(o|l) = \frac{|L|^2}{|O|} B(l|o) \quad (9)$$

$$\forall l \in L', \text{What}(\text{Which2}(l)) = l \quad (10)$$

$$\forall l \in L', \text{What}(\text{Which3}(l)) = l \quad (11)$$

ただし、 $L' = \{l | l \in L, o \in O, \text{What}(o) = l\}$ である。IA2の信念が以上の条件式を満足する時、前項に述べた問題は全て解消される。

ここに述べた機構を実装する一方法として、モデルでは以下に示すように、属性毎に共通度(固有度)を固有度(共通度)によって相互規定することにする。すなわち、 $\forall c_i \in C, \forall f_j \in F, \forall l_k \in L$ に対して

$$B_C(l_k|c_i) = \frac{B_C(c_i|l_k)}{\sum_{l \in L} B_C(c_i|l)} \quad (12)$$

$$B_C(c_i|l_k) = \frac{B_C(l_k|c_i)}{\sum_{c \in C} B_C(l_k|c)} \quad (13)$$

$$B_F(l_k|f_j) = \frac{B_F(f_j|l_k)}{\sum_{l \in L} B_F(f_j|l)} \quad (14)$$

$$B_F(f_j|l_k) = \frac{B_F(l_k|f_j)}{\sum_{f \in F} B_F(l_k|f)} \quad (15)$$

という条件式を満足するように信念を修正する。

4. 数値実験

本節では各IAを用いて行った数値実験結果を示す。パラメータの値は、 $\alpha = 0.01, \tau = 0.01, |C| = 16, |F| = |L| = 22, |O| = |C||F| = 352$ に設定した。IAが一度の経験で見ると対象数は5に固定した。またIAに教示するのは、形に関して構造化された言語である。IAは毎ステップ、対象集合からランダム選ばれた5個の対象を見ると同時に、あるラベルを聞くという経験を繰り返す。

4.1 実験1：正解率

この実験では各IAの学習能力を調べるために正解率を測定した(注6)。図1に、各IAの正解率の時間経過を示す。図から各IAは正解率1を達成するが、学習速度は大きく異なることがわかる。特にIA0の学習速度が極めて遅い。

4.2 実験2：事物カテゴリーバイアス

この実験では、IAが見ることができる対象を一部に限定する。つまりIAはいくらか経験を繰り返しても、見たことのない未知の対象が数多く存在する。具体的には c_1 あるいは f_1 の属性を持つ対象のみを経験できるように設定した。この時、全ての対象の内では経験できる対象の割合は約10%になる。

図2に各ステップにおける各IAの正解率、およびIAがそのステップまでに見た対象の割合の時間経過を示す。見たこと

(注6)：ここで正解率とは、全ての対象に対して「これは何？」と質問し、IAの回答が教示言語と同じならば正解、異なるならば間違えとして正解の割合を表したものである。

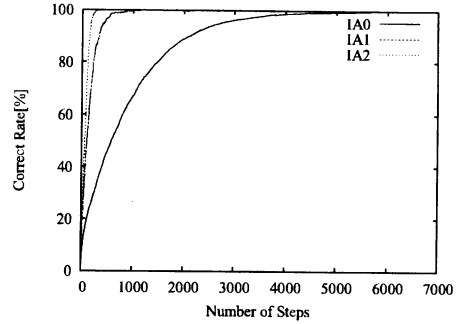


図1 各IAの正解率比較

Fig. 1 Comparison of correct rate among IAs

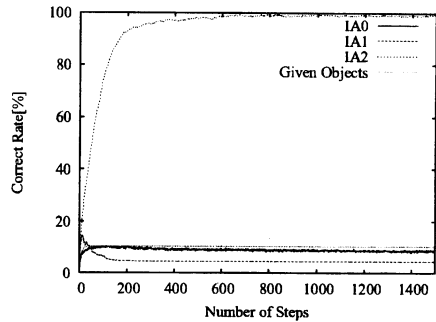


図2 各IAの汎化能力の差異

Fig. 2 Difference of generalization ability among IAs.

のある対象の割合(図ではGiven Objectsと表示)は、時間経過とともに上昇するが、実験の設定上すぐさま10%で飽和する。IA0の正解率は、最終的に10%を少し下回る程度で収束する。すなわちIA0は、経験した対象については確実に学習できるが、未知の対象については全く対処できない。IA1の正解率は、最終的に5%程度になる。すなわち経験した対象についても全て学習できるわけではない。一方IA2は、未知の対象が数多くあるにもかかわらず、100%の正解率を達成した。このようにIA2のみが極めて高い汎化能力を持つ。

4.3 実験3：形状類似バイアス

この実験では、各ステップの学習後、IAが今までに見たことのないような新奇な色と形を持つ対象 o_{novel} を提示し、それに対して初めて聞く新奇なラベル l_{novel} を付与する。この教示を一度だけ行った後、 o_{novel} と色は同じだが形は異なるような対象群 O_C と、形は o_{novel} と同じだが色は異なるような対象群 O_F に含まれる各対象について、「これは何？」と質問する。

図3, 4, 5は各IAが、 O_C と O_F 各々の対象群の内、 l_{novel} と判断した対象の割合を示す。IA0は O_C と O_F のどちらにも含まれる対象に対して、 l_{novel} と判断する割合は、10%程度と非常に低い。またその割合は学習の進行具合とは無関係である。

IA1は学習初期の段階では、 O_C と O_F どちらにも含まれる対象も60%以上の比較的高い割合で l_{novel} と判断する。しかし、

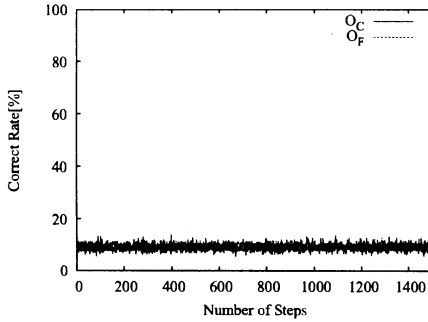


図3 形状類似バイアス: IA0
Fig.3 Shape bias: IA0

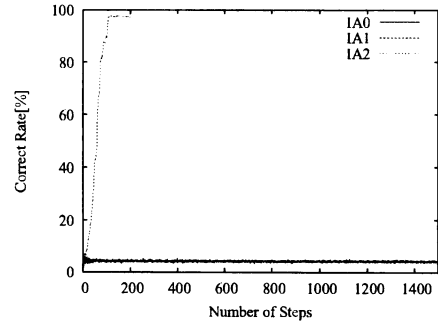


図6 相互排他性バイアス
Fig.6 Mutual exclusivity bias

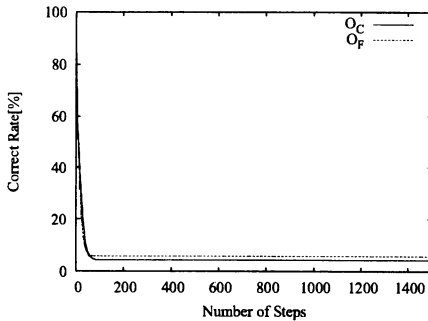


図4 形状類似バイアス: IA1
Fig.4 Shape bias: IA1

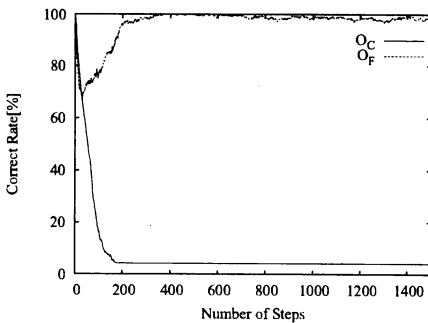


図5 形状類似バイアス: IA2
Fig.5 Shape bias: IA2

学習の進行とともに、その割合は10%以下にまで減少する。このように IA0 および IA1 は、学習初期において違いが見られるものの、最終的には O_C と O_F に含まれる対象を l_{novel} と判断することはほとんどない。以上の結果は、彼らは教示された対象以外の対象を l_{novel} と見なさない、つまり l_{novel} を固有名として解釈するというを表す。

IA2 は、何も学習していない初期の段階においては、 O_C と O_F どちらに含まれる対象も l_{novel} と判断する。つまり o_{novel} と形か色のどちらかが同じであれば、 l_{novel} と判断する。しか

し学習が進行するにつれて、 l_{novel} と判断する対象の割合が減少していく。 O_C 内の対象に関しては、その割合は5%程度になるまで減少し続ける。一方 O_F 内の対象に関しては、その割合は一旦減少するが、30 ステップあたりから上昇に転じ、400 ステップあたりではほぼ100%に回復する。これは形が o_{novel} と同じであれば色が何であっても l_{novel} と判断するが、色が o_{novel} と同じであっても形が異なれば l_{novel} とは判断しないということを示している。つまり、IA2 は学習が進行するにつれて、ある対象が l_{novel} であるか否かを、形状のみを基準に判断するようになる。このように IA2 には、形状類似バイアスが見出される。

また対比実験として、色に関して構造化された言語を IA2 に学習させた場合、どのようなバイアスが形成されるかを分析した^(注7)。紙面の都合上結果は省略するが、IA2 は成長するにつれてある対象が l_{novel} であるか否かを、色のみを基準に判断するようになる。

4.4 実験4: 相互排他性バイアス

ここでは、各ステップの学習後、IA に既知の対象と新奇な対象 o_{novel} の二つを提示し、新奇なラベル l_{novel} を用いて「 l_{novel} はどちら？」と質問する^(注8)。この質問を全ての既知の対象と o_{novel} の間で行う。

図6 は各 IA が新奇な対象を選択した割合を示す。IA2 において、学習の進行とともに、新奇な対象を選択する割合が増加し、200 ステップあたりでその割合は100%に達する。一方 IA0 および IA1 は、ほとんどの場合において既知の対象の方を l_{novel} であると判断する。すなわち IA2 のみに相互排他性バイアスが見出された。

4.5 実験5: バイアスの適用制限

この実験では、各ステップの学習後、既知のある対象 o_{known} に新奇なラベル l_{novel} を強制的に付与し、それを一度だけ教示した後、その対象と色あるいは形が同じ対象群 (O_{CF} と表記する) に含まれる対象に対して「これは何？」と質問する。

(注7): ここで色に関して構造化された言語とは、色が同じ対象には同じラベルを付与し、かつ色が異なる対象には、全て異なるラベルを付与した言語である。

(注8): 回答法には *Which1* を用いた。ただし、この場合選択対象は二つに限定されているので、評価は全ての対象ではなく二つの対象に対してのみ行う。

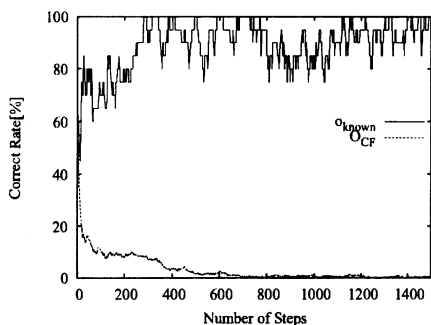


図7 バイアスの適用制限: IA2
Fig. 7 Restriction on application to bias

図7は o_{known} と O_{CF} 内の対象について、IA2が l_{novel} と判断した割合を示す。IA2は o_{known} を、80%程度の高い割合で l_{novel} と判断する。一方 O_{CF} 内の対象については、学習初期には50%程度の割合で l_{novel} と判断するが、学習が進行するにつれてその割合が減少し、700ステップあたりでほぼ0%に収束する。つまりIA2は成長するにつれて、既知の対象に付けられた新奇なラベルを固有名と解釈するようになる。

5. まとめと今後の課題

IA0は、対象とラベルの対応関係を対象毎に個別に記憶することができる。このため、どのような言語を与えられようとも学習可能である。しかしその優れた記憶力のため、逆に汎化能力を持つことができず、教示されない未知の対象については全く対処できない。つまりIA0は、ラベルを固有名と見なして学習を行うエージェントであると解釈できる。

一方IA1とIA2は、共通度と固有度を属性毎に分解して記憶し、使用する場面では属性毎の値を掛け合わせることで統合する。分解と統合という操作によって情報が失われるため、どのような言語でも学習できるわけではない。情報損失の程度は、IA1においては、教示された対象さえ学習することを困難にする。しかし一方、IA2においては、汎化能力という肯定的意味合いを持ち、事物カテゴリーバイアスを形成する要因となる。

IA2は、ラベルを初めて聞く場合そのラベルを普通名詞と判断するが、その際ラベルの適用範囲を形状が同じ対象に拡張する。すなわち、IA2は形状類似バイアスを持つ。ただしIA2はこれらのバイアスを無制限に適用するわけではない。新奇なラベルを付与された対象が既知の場合には、そのラベルを固有名詞と見なす。また、IA2には相互排他性バイアスも見出される。IA2は、構造化された言語を学習していく中で、これらのバイアスを自立的に形成あるいは調整し、効率よく学習を行うようになる。

以上に述べたように、IA2ではバイアスが学習によって形成される。ただし、IA0およびIA1においてはバイアスの形成が見られないという意味で、バイアスの形成能力自体が生得的であると言える。IA1とIA2の差異は、固有度と共通度の相互作用を認めるか否かという点にある。IA1においては、新奇なラ

ベルを学習する際、そのラベルに関する信念は、既に持っている信念とは無関係に形成される。一方IA2においては、新奇なラベルに対する信念形成は、固有度と共通度の相互作用のために、既に獲得している信念からの影響を強く受ける。このため、学習が進行するにつれて、初めて聞くラベルであってもその信念形成にはバイアスがかかるようになる。

最後に今後の課題について述べる。数値実験の結果から、バイアスの形成には、固有度と共通度が相互に作用する機構を持つ必要のあることが示唆された。IA2においてこの機構は所与であり、この機構が生物進化の過程の中でどのように形成されてきたのかという問題については、モデルを進化シミュレーションに拡張して分析する必要がある。

本稿のモデルでは、各対象に対して一つのラベルを付与するという条件を設定した。このため、例えば「りんご」と「果物」のような階層関係を明示的に扱うことができない。階層関係を扱うには、一つの対象に複数のラベルが付与されることを認める必要がある。またモデルでは、対象は色と形の二つの属性のみを持つと仮定した。今後は他の属性や物体と物質の違いなども扱えるようにモデルを拡張したい。

文 献

- [1] S. A. Kripke: "Wittgenstein on Rules and Private Language -An Elementary Exposition-", Basil Blackwell, Oxford (1982).
- [2] 針生, 今井: "語意学習メカニズムにおける制約の役割とその生得性", 心の生得性—言語・概念獲得に生得的制約は必要か— (今井 (編)), 共立出版, 東京, pp. 131-171 (2000).
- [3] 今井: "概念発達と言語発達における類似性の役割", 類似から見た心 (鈴木 宏昭 (編)), 共立出版, 東京, pp. 148-178 (2001).
- [4] E. M. Markman: "Categorization and naming in children: Problems of induction", MIT Press, Cambridge (1989).
- [5] M. Imai, D. Gentner, N. Uchida: "Children's theory of word meanings: The role of shape similarity in early acquisition", *Cognitive Development*, **9**, pp. 45-75 (1994).
- [6] B. Landau, L. B. Smith, S. S. Jones: "The importance of shape in early lexical learning", *Cognitive Development*, **3**, pp. 299-321 (1988).
- [7] M. Imai, D. Gentner: "A crosslinguistic study of early word meaning: Universal ontology and linguistic influence", *Cognition*, **62**, pp. 169-200 (1997).
- [8] M. Imai, E. Haryu: "Learning proper nouns and common nouns without clues from syntax", *Child Development*, **72**, pp. 787-802 (2001).
- [9] E. M. Markman, G. F. Wachtel: "Children's use of mutual exclusivity to constrain the meaning of words", *Cognitive Psychology*, **20**, pp. 121-157 (1988).
- [10] 針生: "幼児における事物名解釈方略の発達検討—相互排他性と文脈の利用をめぐる—", *教育心理学研究*, pp. 11-20 (1991).
- [11] R. S. Sutton, A. G. Barto: "Reinforcement Learning -An Introduction-", MIT Press, Cambridge (1998).