

Infant Agents 間の対話による概念獲得と対話戦略

Acquisition of Concept and Dialogue Strategies through the Interaction among Infant Agents

田口 亮 山本 航 桂田 浩一 新田 恒雄
Ryo Taguchi Wataru Yamamoto Kouichi Katsurada Tsuneo Nitta

豊橋技術科学大学 大学院工学研究科
Graduate School of Engineering, Toyohashi University of Technology

This paper described acquisition of concept and dialogue strategies through the interaction among infant agents (IAs). In experiments, an IA who has learned all the concepts teaches an uninformed IA along with a dialogue strategy. Two types of concept acquisition experiments are executed: (A) the dialogue strategy is given beforehand and (B) the strategy is automatically given by using reinforcement learning. The experimental results show the importance of strategies in dialogue exchange among IAs and the same dialogue strategy as human designs is given by the reinforced learning.

1. はじめに

マルチモーダル対話 (MMI) システムでは、モダリティを解釈・統合する中で、利用者意図に則した動作を効率良く行うことが望まれる。対話を効率良く進めるには、人間同士の対話を利用する、共通の価値原理をコンピュータに持たせることが必要になる。我々は、複数の人間が個別に育てたコンピュータ内の幼児エージェント (以下 Infant Agents: IAs と呼ぶ) 間で、マルチモーダル対話を行わせる実験から、共通の価値原理を自動獲得する研究を行っている。本報告では前回 [新田 02] に続き、価値原理の一つである対話戦略の獲得を目指す。前半では、全ての概念を獲得した IA (教示 IA) が、他の IAs の質問に答える形で対話を進めた際に、予め付与した対話戦略を変えた場合の比較実験結果を報告する。続いて後半では、教示側の対話戦略に的を絞る、強化学習による自動獲得実験結果を報告する。

2. Infant Agent

IAs は同一空間内にある画像オブジェクトの特徴を視覚情報として、人間や他の IAs が発する音声の特徴を聴覚情報として受け取り、両者の対応関係を概念として獲得する。前報では、聴覚情報に音声特徴ベクトルを用いたが、今回の実験では、IA の音声認識部が出力するモーラ単位のシンボル列を用い、かつ実験の便宜上シンボル列に誤りはないものとした。

(1) 感情と表情

IA は平常・快・不快の三つの感情を持ち、これらの感情は顔の表情によって他の IA に伝達されるとする。通常、IA は平常の状態にあるが、概念を獲得した場合、教示発話に易しい未知語 (4モーラ未満) が含まれている場合に快になり、難しい未知語 (4モーラ以上) が含まれている場合は不快になる。また快・不快は持続せず、その都度平常に戻るとする。教示 IA は、質問 IA の表情を手掛かりに、対話戦略を獲得することになる。

(2) 記憶システム

IA は視覚情報を一時的に格納する感覚記憶と、獲得概念を辞書として保持する長期記憶を持つ。また、相手の表情 (上記感覚記憶の一部) が不快でなければ、理解のシグナルとして概念とペアでこれを記憶する。これは対話中だけ保持する短期記憶であるが、後述の実験ではこの機構を利用する場合と、利

用しない場合を比較する。この他、発話の際には内容を格納する発話メモリを別に持つ。

(3) 基本行動

IA が意識的に取る行動は、① なにもしない、② 話題とするオブジェクトをランダムに選択する、③ オブジェクトを指差す、④ オブジェクトを移動する、⑤ 概念の抽出元を概念辞書にする、⑥ 抽出元を一つ前の発話メモリにする、⑦ 抽出元から発話メモリに概念を追加する、⑧ 発話メモリの内容を発話する、の八つである。なお、表情を変える行動は、ここでは無意識に現出するものとした。

(4) 対話学習

IA は他の IAs と対話することで概念を共有化する。本報では、全ての概念を獲得した IA (教示 IA) が、知識を持たない IA (質問 IA) に答える形で対話学習を行う。役割を固定するため、教示 IA の行動は上記の①⑤⑥⑦⑧に、また質問 IA の行動は②③④に制限した。

3. 対話戦略を予め付与した概念獲得実験

これまでの、初出の概念は未知語のため一単語発話からしか獲得できない、という条件下で実験を行ってきた。しかし、自然な対話を想定すると、既知語と未知語が混在する発話中から未知語を検出し記憶する必要がある。そこで本報では、発話内の未知語を切り出し、新しい概念として登録できるように変更した上で、以下の教示戦略比較実験を行った。

3.1 教示戦略

(1) ターン数に合わせて概念数を漸増

教示 IA はターン数に合わせて (ここでは 10 ターン毎)、発話概念数を一つずつ増加させる (以下、時間戦略と呼ぶ)。

(2) 教示履歴を用いて概念数を漸増

教示 IA は履歴から、教示予定の発話内容に未教示の概念が何個含まれているかを知ることができる。そこでこの情報を利用し、未教示概念が一つ現れるか、もしくは発話可能な単語数の上限 (今回は 3 単語) になるまで発話メモリに概念を追加し、その後発話する (以下、履歴戦略と呼ぶ)。

3.2 実験条件

今回は、形状 (丸、三角、四角)、色 (赤、青、白) の物体概念、

連絡先: Ryo Taguchi [taguchi@vox.tutkie.tut.ac.jp]

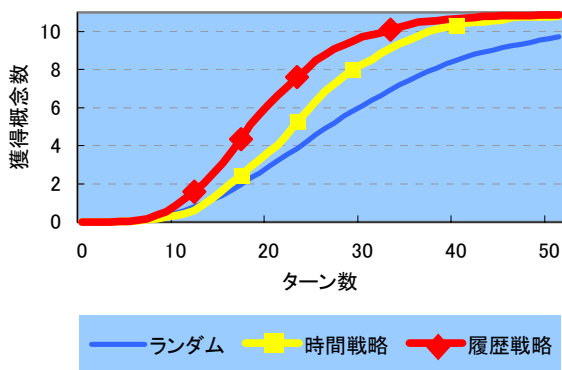


図1: 付与した戦略の概念獲得過程

空間概念としての位置(上, 下, 左, 右), および動作概念(移動)の計11個の概念を使用した. 教示IAは11個全ての概念を持ち, 質問IAは概念を持たない. 学習は質問IAが, 対象のオブジェクトを指差し, または移動することから始まり, 教示IAはそれに対して適切な概念を教示する. 教示により質問IAが不快になった場合はもう一度質問を繰り返す, それ以外の場合は他のオブジェクトに対して質問する. これを繰り返す, 質問IAの獲得概念数が11個になったら終了する. また概念は, 以下の条件を全て満たした時に獲得したと仮定した.

- ① 3回以上学習している
- ② ある属性内(形, 色, 位置, 動作)で, 同じ特徴(赤など)が現れた頻度が90%以上ある
- ③ ②を満たす属性が一つだけである

比較のため, 前述した基本行動⑦(概念追加)と⑧(発話)をランダムに行った場合の実験も行った. また, 教示IAが発話する際の概念の抽出元は概念辞書とし, 前の発話は利用しない.

3.3 実験結果

各戦略それぞれについて1000回の対話実験を行い, それらを平均した結果を図1に示す. インタクションの中で相手の状況を考慮し, 相互信念を作ろうとする履歴戦略が効率良いという結果が得られた.

4. 対話戦略の自動獲得

ここでは, 強化学習の一手法である Q-learning [Watkins 92] を用いて, 環境との相互作用の中でIA自身に対話戦略を獲得させるを試みる.

4.1 実験条件

状態を構成する属性の組み合わせには, 以下の3種類を用意し実験A~Cを行った.

実験A・ターン数(0~5, 50ターンまでの10ターン単位の値)

- ・発話メモリ概念数(0~3)

実験B・相手の表情(0: 平常, 1: 快, 2: 不快)

- ・前の発話メモリ中の概念数(0~3)
- ・概念の抽出元メモリ
(0: 未定義, 1: 概念辞書, 2: 前の発話メモリ)

- ・発話メモリ概念数(0~3)

実験C・発話メモリ概念数(0~3)

- ・発話メモリに含まれる未教示概念数(0~3).

実験Bでは, 教示の前に概念の抽出元を選択(⑤か⑥)しなければならないが, 実験AとCでは, 前の発話を教示に利用しないため, 常に⑤が選択されるものとして省略した.

表1: 獲得された戦略の例

獲得戦略 A		獲得戦略 B		獲得戦略 C	
状態	行動	状態	行動	状態	行動
00	⑦概念追加	0000	⑤概念辞書	00	⑦概念追加
01	⑧発話	0010	⑦概念追加	10	⑦概念追加
10	⑦概念追加	0011	⑦概念追加	11	⑧発話
11	⑧発話	0013	⑧発話	20	⑦概念追加
~	~	~	~	21	⑧発話
51	⑦概念追加	2300	⑥前発話	30	⑧発話
52	⑦概念追加	2320	⑦概念追加	31	⑧発話
53	⑧発話	2321	⑧発話		

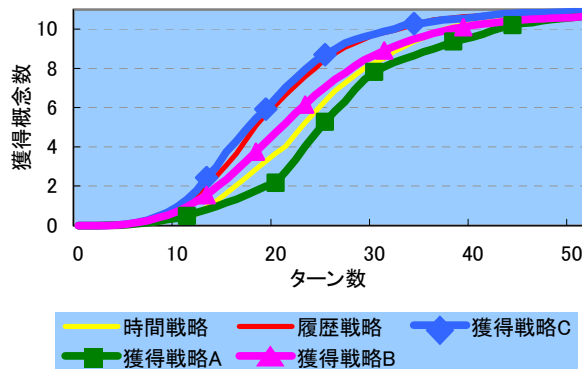


図2: 獲得された戦略の概念獲得過程

報酬は質問IAが快の表情をした時に+2, 不快・平常の場合に-1, 行動⑤⑥(抽出元の選択)及び行動⑦(概念追加)による内部遷移の場合は0とした. なお, Q-learning の学習率 α は0.001, 割引率 γ は0.9と設定した. 学習時の行動選択には ϵ グリーディ手法($\epsilon = 0.5$)を用いている.

4.2 実験結果

3で説明した対話学習を一万回(実験Bの場合は十万回)繰り返し, 戦略を学習させた. 実験A~Cで獲得された戦略(獲得戦略A~C)の一部を表1に, 各戦略を用いた概念獲得の結果を図2に示す. 獲得戦略Aは20ターンまで1語, 30~40が2語, それ以外は3語での教示を行っている. 実験Aでは, 得られる報酬の期待値が過去の状態遷移に強く依存するため, ϵ グリーディによる探索行動が学習結果に悪影響を与える. 獲得戦略Bは, 質問IAが不快な表情をして質問を繰り返した時に前の発話内容から1語発話し, それ以外は3語での教示を行っている. 相手が不快な表情をしたら発話を訂正するという相互対話に基づく戦略を獲得した. 獲得戦略Cは履歴戦略と同じ戦略を示した. 短期記憶を利用し, 発話中の未教示概念の数を制御することで, 結果的に一番効率の良い戦略となった.

5. まとめ

本報告では, 教示に用いる対話戦略を, 対話の相互作用を通して自動獲得を試み, 人間が付与した戦略と同等の戦略が強化学習を通して得られることを示した. 今後は, 質問と教示双方を含む対話戦略の自動獲得を目指したい.

参考文献

- [新田 02] 新田ほか: Infant Agents 間での対話による概念知識獲得, 人工知能学会全国大会, 2002 1A1-07.
- [Watkins 92] C.J.C.H. Watkins et al: Q-learning, Machine Learning 8, pp.279-292, 1992.