

5X-1

画像・音声メディアからの言語概念獲得における 対話を用いた学習の効率性評価*

伊東 洋一郎 黒田 和史 荒井秀一 宮内新†
武蔵工業大学‡

1 はじめに

近年、言語概念の獲得を試みる研究が数多く行われている。これらの多くは予め言語シンボルを与え、その上でより高度な概念の獲得を目指している。しかし、言語シンボルの概念を獲得しようとした場合、言語シンボルは概念を内包しているため、予め与えることは、獲得すべき概念を与えることになる。そのため、非シンボリックな情報のみからの概念獲得が必要であると考えられる。そこで、我々は既に言語シンボル概念の獲得モデルを提案している [1][2]。この獲得モデルにおいて、画像・音声信号を用いた対話から概念の獲得が行えることが確認されている。しかし、対話からでは獲得された概念を定量的に評価することができない。

対話により表象される刺激は、概念を具象化した情報であるため、その情報は概念の一部でしかない。そこで、計算機内部に構築された概念を直接評価することとした。さらに、構築される概念は、与える画像・音声信号による刺激に依存する。そこで、刺激の提示方法による獲得概念の違いを評価する。

2 概念獲得の枠組

2.1 概念の表現

提案している言語シンボル概念の獲得モデルにおいて、概念は画像・音声刺激により構築され、入力刺激の特徴量の分布の形で表される。画像概念は、画像領域の面積や、周囲長などの特徴量ベクトルの分布と、概形を表す HMM(Hidden Markov Model) を用いて保持し、音声概念は、単語の HMM により保持している。

また、同時に受容した画像刺激・音声刺激は同じ概念を表すとする提示同時性に基づき、同時に受容した刺激の概念同士にリンクを張る。この画像概念、音声概

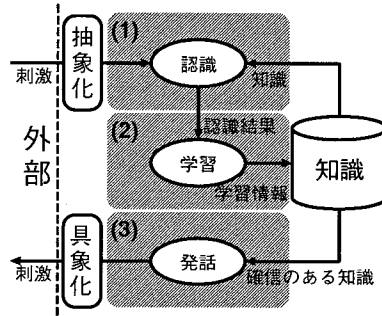


図 1: 対話を用いた概念獲得における情報の流れ

念、リンクを含めて統合的な言語シンボル概念を表すものとする。

2.2 概念の獲得

提案している概念獲得モデルにおける情報の流れを図 1 に示す。このモデルは、刺激を受容することで駆動するモデルとなっている。

認識 受容刺激は特徴量を抽出し認識処理を行う。受容刺激がいずれの概念に該当するのか、あるいは、未知の概念であるのか判断をする。

学習 認識結果より入力刺激が未知の概念であれば、新規概念を生成する。既知の概念であれば、該当する概念の分布に、入力刺激を加えることで概念の学習を行う

発話 発話の動機が発生した時、保持する概念を具象化することで発話を行う。

2.3 確信度

受容刺激の認識は確信度を用いて行われる。確信度は、各概念において受容刺激がその概念であるとする事後確率で表される [3]。

概念 (k) として保持する分布の平均と、入力刺激 (x) の特徴量とのマハラノビス距離を D_k とし、分布が正規

*Evaluation of Learning Efficiency for Acquiring the Concept of Linguistic Method through dialogue Using Image and Voice Media

†Yoichiro ITO, Kazufumi KURODA, Shuichi ARAI, Arata MIYAUCHI

‡Musashi Institute of Technology

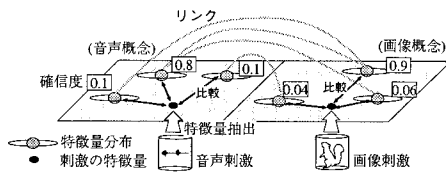


図 2: 入力刺激と確信度

分布に従うと仮定すると、確信度は式 1 で定義される。

$$P(k|x) = \frac{\exp(-\frac{D_k^2}{2})}{\sum_{j=1}^m \exp(-\frac{D_j^2}{2})} \quad (1)$$

入力刺激と、各概念の確信度の関係を図 2 に示す。確信度が 1 に近い概念は明確に認識できていることを表す。

3 獲得概念の評価

概念の獲得が行われることは、対話を通じ確認することができる。しかし、概念を具象化するということは、概念の一部を表象するのみで概念全てを表しているわけではない。そこで、計算機内部に構築された概念を直接評価することとした。

獲得したい概念は、バリエーションのある刺激に対しても表現可能な概念、または、詳細情報が学習された概念である。

ゾウの概念であれば、足の曲がったゾウや、ハナの向きが違うゾウに関しても表現可能であることが望ましい。つまり、人間がゾウであると判断する刺激に対し同じようにゾウであると判断できる概念が獲得したい概念である。このような概念を評価するためには、バリエーションのある刺激を与え、確信度の計算を行うことで概念の評価を行うものとする。

また、詳細情報が学習された概念とは、他の概念と明確に異なる概念である。概念の部分的な情報、構造情報が学習された概念は、他の刺激を与えても、その時の確信度が高くなる概念である。ゾウの概念が、ゾウのハナや、ミミ、シッポなどを学習していれば、キリンやリスといった刺激を受容しても高い確信度は計算されない。この様にして評価を行うことにする。

4 実験

前述した評価方法を用いて、獲得される概念が入力刺激に依存し、学習方法により獲得される概念が異なることを評価する。

まず始めに対話を用いない 2 通りの学習方法により概念の獲得を行う。

刺激提示方法 a 1 つの物体について十分に理解が行われるように学習を行い、その後、次の物体について学習を行うといった方法である。しかし、刺激を与える人間には十分に理解が行われたかは分からないため、多くの刺激の入力を必要とする。この時に学習させる概念としては、ゾウ、キリン、ウサギ、カンガルー、リスなどの物体概念、および、物体の部分構造を表す、ミミ、クビ、シッポ、短い、細いなどである。

刺激提示方法 2 各物体をランダムな順に教える方法である。この方法についても同様な概念の学習を行わせる。

つぎに、対話を用いた概念の獲得を行う。この方法は、対話を行うことで計算機が受容刺激が明確に認識できたか否かを人間(教師者)に示すため、不明確であった概念を表す刺激が多く、また、明確に認識された概念については与える刺激を少なくする。このような対話を用いて、刺激提示方法 a、及び、刺激提示方法 b の学習を行う。これにより獲得された概念を評価する。

5 まとめ

提案している言語概念獲得モデルについて、獲得概念の定量的評価を行った。概念の獲得において、ランダムに刺激を与えるのではなく、系統立てて学習を行うことで、明確な概念を獲得できることが確認できた。また、対話を行うことで、効率的な学習を行うことが確認できる。これにより、提案している対話による言語概念獲得モデルが学習効率の観点から効果的であるということが言える。

参考文献

- [1] 黒田, 荒井, 石川: “画像・音声メディアを用いた人間-エージェント間対話による言語シンボル概念の獲得”, 電子情報通信学会, ソフトウェアエージェントとその応用 特集ワークショップ (SAA2000) 講演論文集, pp.297-304, 2000.11
- [2] 伊東, 黒田, 荒井: “画像・音声メディアを用いた言語概念獲得モデルにおける学習効率の評価”, 信学全大, D-8-15, 2001.3
- [3] 小松, 荒井: “エージェントの確信度を用いた画像・音声情報からの物体概念獲得”, 信学技報, AI96-38