

EM アルゴリズムを用いた絶対概念と相対概念の同時学習

Learning of Absolute Concepts and Relative Concepts Using EM Algorithm

田口 亮^{*1}
Ryo Taguchi

WANG DING^{*1}
Wang Ding

YU QIYUE^{*1}
Yu Qiyue

保黒 政大^{*2}
Masahiro Hoguro

梅崎 太造^{*1}
Taizo Umezaki

^{*1} 名古屋工業大学
Nagoya Institute of Technology

^{*2} 中部大学
Chubu University

This paper proposes a method for the unsupervised learning of absolute concepts and relative concepts from pairs of a symbol representing a word and a feature vector. In a previous work, we have proposed a lexical learning method based on statistical model selection. This method can learn the words that represent a single object, such as proper noun, but cannot learn the adjective-words that represent relative attributes of objects, such as color, position and size. This paper describes a method for learning relative concepts.

1. はじめに

近年、単語の意味をロボットの感覚・運動系の信号と対応付けて学習させる研究が行われ始めた[田口 10, Iwahashi 08]. 従来研究の多くは、単語と共起するセンサ情報を直接的に学習していた。例えば、文献[Iwahashi 08]では画像から抽出された物体の色や形、大きさなどの特徴量と音声の対応関係を学習している。しかし、色や大きさ、位置などの概念は、他の物体との比較や、典型的な概念との比較によって表される。例えば、蟻を指さして「小さい」と言う場合もあれば、同じ蟻でも隣の蟻と比較して「大きい」と言う場合もある。また、「小さい象」は通常、「大きい蟻」よりも大きい。このように言葉には参照点に依存し、相対的な意味を持つ単語が多く存在する。

先の例では「大きい蟻」や「小さい象」などのように、参照点が言語的に明示されていたが、人間同士の対話においては、参照点が明示されない場合がある。例えば、図 1(a)のような場面で、人が「右にあるやつを取って」と発話した場合、取って欲しいものは画像右側のぬいぐるみであることが想像できる。この場合、参照点は画像の中心、または話者の視線の中心となる。一方で図 1(b)のような場面で、「左にあるやつを取って」と発話することも可能である。この場合、取って欲しいものは画像中央の赤い箱であり、参照点は画像右側の白い箱になる。参照点に依存した語意を学習する際も同様に、参照点が明示されない場合でも、複数の候補から適切に参照点を推定しながら語意を学習していく必要がある。本稿では EM アルゴリズム[Dempster 77]を用いることで、参照点の推定と語意の学習とを同時に行う手法を提案する。

2. 語彙学習手法

ユーザがある対象をロボットに提示し、対象が持つ特徴の名前を教示する。本実験では、音声の代わりにテキストのラベルを与える。ここで、教示ラベルを w 、対象の特徴を x_n とする。 n は学習サンプルのインデックス ($1 \leq n \leq N$, N は学習サンプル数) である。また概念の比較対象となる参照点の候補を k とする。参照点 k を基準とした特徴量を x_{nk} と表す。各場面における指示対象と参照点の候補は既知とするが、候補のいずれかが真の参照点なのかは未知である。例えば、特徴量 x_n が指示対象の位置を表す場合、 x_{nk} は参照点 k からの相対位置を表す。各学



図 1: 参照点が明示されない例

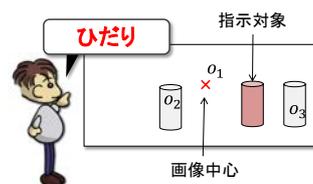


図 2: 教示の例 ($o_1 \sim o_3$ が参照点候補)

習サンプルでは、参照点候補の数に応じて、学習候補となる特徴量 x_{nk} が複数現れる。図 2 の例では、参照点の候補として $o_1 \sim o_3$ が与えられ、各参照点候補と指示対象との距離が学習候補の特徴量 $x_{n1} \sim x_{n3}$ となる。図では o_1, o_2 の右に指示対象があるため、「ひだり」の意味を学習する際には、 o_3 を参照点とし、他の特徴量候補は破棄しなければならない。このように、本タスクでは、学習サンプル毎に正解となる特徴量のみを採用し、学習していくことが望ましい。

本稿では、与えられたラベルと指示対象の対応関係を平均 μ 、分散 σ^2 の正規分布 $N(x_{nk} | \mu, \sigma^2)$ でモデル化する。また、学習サンプル n において参照点候補 k が真の参照点である確率を π_{nk} とおく。そして式(1)に示すモデル尤度を最大化するパラメータ $\theta = (\mu, \sigma^2, \pi_{nk})$ を EM アルゴリズムにより求める。各パラメータの更新式を式(2)~(4)に示す。

$$\theta = \arg \max \sum_{n=1}^N \ln \sum_{k=1}^{M_n} \pi_{nk} N(x_{nk} | \mu, \sigma^2) \quad \dots(1)$$

$$\mu = \frac{1}{N} \sum_{n=1}^N \frac{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2) x_{nk}}{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2)} \quad \dots(2)$$

連絡先: 田口 亮 Email: taguchi.ryo@nitech.ac.jp

TEL&FAX: 052-735-5552

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N \frac{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2) (x_{nk} - \mu)^2}{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2)} \quad \dots(3)$$

$$\pi_{nk} = \frac{\pi_{nk} N(x_{nk} | \mu, \sigma^2)}{\sum_j \pi_{nj} N(x_{nk} | \mu, \sigma^2)} \quad \dots(4)$$

EM アルゴリズムは隠れ変数を含む確率モデルのパラメータを、最尤法に基づいて推定する手法であり、混合正規分布の学習にも用いられる。本手法ではサンプル毎に参照点の確率 π_{nk} を定義しているが、混合正規分布では各サンプルに対して独立な混合重み π_k を用いる。また、混合正規分布の場合には複数の正規分布を用いるが、本手法では単一の正規分布である。

また EM アルゴリズムを用いた動作概念の獲得に関する研究も行われている[Sugiura 11]。この研究では参照点を基準とした軌跡で動作をモデル化している。本研究はこの研究を動作以外の概念に拡張したものである。

3. シミュレーション実験

3.1 実験条件

国立情報学研究所で開発された社会的知能発生学シミュレータの SIGVerse[稲島 12]を用いてシミュレーション実験を行う。実験環境の俯瞰図と教示の様子を図 3 に示す。ユーザは人型のアバターをコントローラで操作し、仮想空間内を移動する。ユーザが移動すると、その後ろをロボットが追いかけるように移動する。ユーザは任意の点で場所名をテキストで教示する。入力されたテキストはロボットの位置座標(x, z)と共に保存される。実験では「本棚」「テレビ」「ソファ」の 3 つのオブジェクトを、場所名を教示する際のランドマークとする。教示では、オブジェクトの名前と、相対的な位置を表す単語(前, 後, 左, 右, 近くのいずれか)を同時にロボットに与える。これは「本棚の前」や「テレビの右」という教示を想定している。ただし、単語の区切りは与えられるものとする。また、参照点もその 3 つのオブジェクトとし、各オブジェクトの重心座標および向き(正面方向)がロボットに与える。実験では、ランドマークと相対位置を表す単語の組み合わせ毎に、少しずつ位置を変えながら 5~6 箇所まで教示した。

3.2 実験結果と考察

左右の学習結果(X 軸)を図 4 に示す。左右の概念を絶対値(仮想環境の中心が原点)として学習すると、左右の 2 つの分布が重なり、正しい意味が得られない。一方で提案手法により学習すると、任意の参照点を原点とした時に、X 軸のプラス側を右、マイナス側を左と呼ぶように学習できた。学習時に推定された参照点を確認した所、ほぼ正しく推定できていた。

次に前後の学習結果(Z 軸)を図 5 に示す。前, 後の相対値を学習した結果、共に Z 軸のプラス側で確率が高くなり、前, 後の意味を正しく学習することができなかった。学習時に推定された参照点を確認した所、教示の際に使用した参照点とは異なる点が推定されていた。今回の実験では、オブジェクトの配置が固定されており、さらに全てのオブジェクトが画面中心を向いているため、多くの相対位置がプラス方向に偏っていた。そのため局所解に陥ったものと考えている。提案手法では、EM アルゴリズムを用いているため、このように局所解に陥ることが考えられる。今後は人やロボットの注意の方向等を利用し、参照点の確率 π_{nk} の初期値を設定することを試みる。これにより、局所解に陥ることを避け、正しい参照点の推定を実現する。

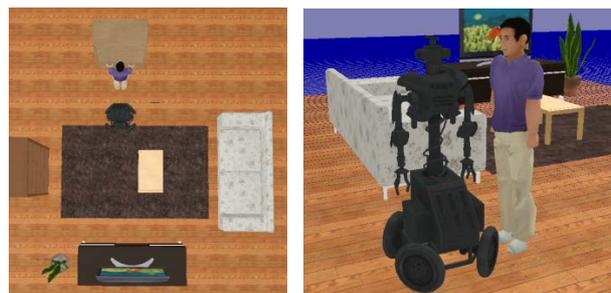


図 3: 実験環境 (左: 俯瞰図, 右: 実験の様子)

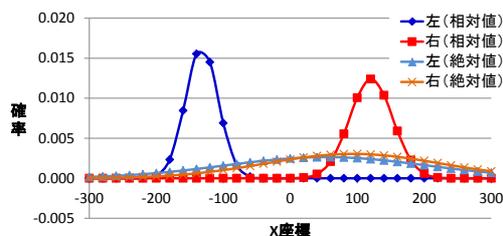


図 4: 左右の学習結果

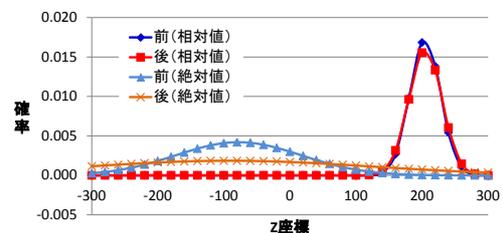


図 5: 前後の学習結果

また、本研究では絶対的概念と相対的概念の判別を学習結果から行うことが実現できていない。提案手法により「本棚」「テレビ」「ソファ」を学習すると、絶対値として学習した場合よりも、狭い分布が学習される。そのため、分布の形状やモデルの尤度では、どちらの概念なのかを判別することはできない。絶対的概念と相対的概念の判別も今後の課題とする。

4. まとめ

参照点に依存した語意の学習手法を提案した。本実験では位置を表す単語を学習したが、今後は色を表す他の単語や、形を表す単語など、異なる属性を対象とする単語を同時に学習する実験を行う。また、参照点推定精度の向上、および絶対的概念と相対的概念の判別方法を検討する。

5. 謝辞

本研究の成果の一部は、科研費若手研究(B)(23700200)および国立情報学研究所共同研究「SIGVerse を用いた自律ロボットの自己位置表象と語彙獲得の研究」の助成を受けた。

参考文献

[田口 10] 田口亮 他: , 人工知能学会論文誌, Vol.25, No.4, pp.5491-5501. (2010)
 [Iwahashi 08] Iwahashi, N., Trans. IEICE, Vol. E91-D, No. 2, pp. 312-321. (2008)
 [Dempster 77] Dempster, A.P., et al., Journal of the Royal Statistical Society. Series B (Methodological) 39 (1): 1-38.(1977)
 [Sugiura 11] Sugiura, K., et al., Advanced Robotics, Vol.25, No.6-7, pp.825-848.(2011)
 [稲島 12] 稲島哲也, 人工知能学会誌, 27(6) pp. 619-624. (2012)