

文と画像情報からの相対概念の獲得

中西宏文

(情報科学選修)

A Relative Concept Acquisition from sentence and image information

Hirobumi NAKANISHI

(Department of Information and Computer Sciences)

Abstract

Recent language acquisition systems which efficiently learn the vocabulary and its meanings by using linguistic and non-linguistic information has been studied. Formerly, we made a concept acquisition system for the purpose of formalizing the method to acquire the concept with visual scene and sentence. Our system acquires concepts without a priori knowledge.

The purpose of this study is to acquire the relative concept from sentence and image information. In this study, we assume that each concept about names, locations, colors and size of objects is already known from the method of previous study. We tried to acquire the relative position or relative size between two objects. The algorithm is to store the information about two learning data and to lead the relative relations. Consequently, some relative concepts are acquired by sequentially learning from features of image and its character strings corresponding to sentence. We have realized the first stage of human's concept acquisition process on a computer system.

1. はじめに

人工知能の研究は、人間にのみ可能であった知識の保有・処理等を情報処理技術の立場から計算機処理の対象とし計算機に演繹・推論能力をも持たせようとするもので、将来的には人間の知的機能を模倣し、あるいはその代行をすることも目指している。

人間の場合、いくつかの感覚器を単独であるいは組み合わせて使用し、外部からの刺激を感じて、それらの情報が脳に伝えられ様々な概念を獲得していると考えられる。その中でも多くの場合、中心となるのは視覚と聴覚であり、色や形を始めとして多種多様の物の名前や空間的な位置に関する概念を獲得している。

人工知能の研究においては、学習データに対して種々の解析を行い、それを基に未知入

力に反応するようなシステムが研究されている。しかしながら、生まれて間もない頃から幼児が実際に複雑な処理を行うと考えるには無理があり、それらに関しては認知心理学の側面からの研究もなされている。

筆者らは過去に、実際の音声とその意味に相当する画像情報を与え、音声と画像の情報を用いて計算機に物の名前や位置等の概念を学習させるシステムを作成した[古部 91][中川 93]。つまり、ある物を表現する画像があったとすると、その画像を説明する文を音声によって与えることにより、逐次画像上の形状・色・大きさ・位置といった概念に対応する音声言語を獲得すること、逆に言えば、ある「音」に対応する形状の概念を獲得するシステムである。但し、物の名前や位置等の概念を単語として与えるのではなく、簡単な説明文の音声データ（又は文字列）とそれに対応する画像データとを用いて、形状・大きさ・位置・色等の概念を形成している。

今回の研究では、複数図形が含まれる画像データとその説明文より、図形間の位置関係や相対的な大小関係の概念を獲得するシステムを提案する。なお、従来からの基本方針として、複雑な構文や意味解析は行わず、文と画像という2種類の情報を関連づけ辞書に蓄積するという、比較的シンプルな方法で相対概念の獲得を行った。

2. 本システムの概要

本システムでは、具体的に人間の幼児がどのような情報によって概念形成を行っているのか考察し、これをヒントに工学的に概念形成メカニズムを計算機上で実現することを最終目標にしている。今回は、従来の研究[中川 93]で単一図形を取り扱っていたものを、複数図形に拡張し、今回は2図形間の相対的な概念を研究の対象としている。

本研究では、相対的な概念の獲得に主眼を置いたため、音声と画像データの学習によって単一図形の概念は獲得されていると仮定した。なお、従来の研究では、音声と画像データを用いて実際のシミュレーションを行ったが、音声処理の精度が不十分なために希望する概念が得られない場合があった。それに対処する方法等も検討してきたが、単一図形の概念獲得が不十分であると、複数図形からの概念獲得に大きな支障をきたす。そこで、音声処理の精度の問題を避けるため、単一図形からの獲得概念が完全なもので誤りが無いという仮定の下で、システムを構築した。誤り無しの仮定を行うことで、シミュレーションは実際の音声データを文字列データに置換することができ、音声処理の精度に関係なく相対概念の獲得のみに主眼を置いたシステムとすることができる。

3. 学 習 デ ー タ

実験で扱う画像データは、表1に示すような単一図形を処理して得られるそれぞれの図形に関する特徴量4次元を合わせたものに、相対関係を表す特徴量を付加し、表2に示すような形式の9次元または10次元の画像特徴を表すデータである。ここで表2に示す相対関係は、図形1を基準図形とした時の図形2の状態を表す。

また、画像データに対応する文データは、複数図形の状態を説明したもので、各図形データに対応した表3に示すようなものである。

なお、本研究では相対概念を扱っているため、判断の基準図形を2図形の内どちらにするかで、同一の画像特徴に対し、表3の4,5のように複数の音声データが対応可能であり、

このことに対する対処が研究のポイントとなる。

また、画像データに関しても、どちらを基準図形にするかにより、

(四角形：基準図形) (三角形) 上側

(三角形：基準図形) (四角形) 下側

のように2通りの表現が可能である。

表1. 単一図形の特徴量

(y 1, y 2, y 3, y 4)
y 1・・・形状 (屈曲点数)
y 2・・・大きさ (1:中, 2:大, 3:小)
y 3・・・色 (1:白, 2:黒, 3:灰色)
y 4・・・位置 (1:中心, 2:右, 3:左, 4:上, 5:下)

表2. 相対関係の概念獲得に用いる画像特徴

(図形1の特徴量)	(図形2の特徴量)	(相対関係)
相対関係: 2:右側, 3:左側, 4:上側, 5:下側		
または 2:より大きい, 3:より小さい		

表3. 相対概念の獲得に用いる音声データ例

1. 丸が三角形の上方にあります
2. 三角形の上方に丸があります
3. 三角形の左方に四角形があります
4. 四角形の右方に三角形があります
5. 三角形の左方に四角形があります
6. 丸は三角形より大である

4. 概念の獲得アルゴリズム

今回提案する方法は、単一図形からの概念獲得の際に用いた方法を発展させたもので、画像と文の対応づけができたものを辞書部として登録し、逐次学習を行うものである。

2つの学習データをマッチングし差異と共通部分を抽出した場合、その差異のあった部分を特徴部、共通部分を共通部と呼ぶ。ここでは、単一図形概念学習より、[丸], [三角形], [四角形], [黒い], [白い], [大きい], [小さい], [が], [あります]に関する概念(但し, [が], [あります]に関しては対応する画像特徴はないので無意味な概念)として獲得済みと仮定している。また、学習を進めながらそれぞれの概念の出現頻度をカウントする。アルゴリズムは、2つの入力文の間、ならびに入力と登録済みの辞書内容との間で音声間および画像間の共通部を抽出することに基づいた方法である。

以下に、図1に示す簡単な学習データ例を用いて、アルゴリズムを示す。(－はその次元を使用しないことを示す)

1	丸が三角形の上方にあります (3 1 1 - 0 1 1 - 4)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>○</td></tr><tr><td>△</td></tr></table>	○	△
○				
△				
2	三角形の上方に丸があります (3 1 1 - 0 1 1 - 4)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>○</td></tr><tr><td>△</td></tr></table>	○	△
○				
△				
3	丸が三角形の上方にあります (0 1 1 - 3 1 1 - 5)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>○</td></tr><tr><td>△</td></tr></table>	○	△
○				
△				
4	三角形の上方に丸があります (0 1 1 - 3 1 1 - 5)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>○</td></tr><tr><td>△</td></tr></table>	○	△
○				
△				
5	三角形の下方に丸があります (3 1 1 - 0 1 1 - 5)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>△</td></tr><tr><td>○</td></tr></table>	△	○
△				
○				
6	丸が三角形の下方にあります (0 1 1 - 3 1 1 - 4)	<table border="1" style="text-align: center; width: 40px; height: 40px;"><tr><td>△</td></tr><tr><td>○</td></tr></table>	△	○
△				
○				
.....				

図1. 学習データ例

1° 最初の学習データを記憶する暗記学習。

分類番号 : 1
 共通画像特徴 : 3 1 1 - 0 1 1 -
 相対関係 : 4
 画像順序 : 基準
 共通文字概念 : 丸が三角形の上方にあります
 入力文字列 : 丸が三角形の上方にあります
 共通概念対 : 丸が三角形の上方にあります

2° 最初の2データを比較し、共通部があれば、共通文字・画像等を辞書に新規登録。

辞書への登録内容

(画像順序欄は、2データ間の画像特徴が同一順序で照合できるか逆順で照合できるかを示す)

分類番号 : 1
 共通画像特徴 : 3 1 1 - 0 1 1 -
 相対関係 : 4
 画像順序 : 同一
 共通文字概念 : 丸が三角形の上方にあります
 入力文字列 : 三角形の上方に丸があります
 共通概念対 : 丸が 三角形の上方に

- 3° 次のデータを入力。全ての辞書と画像特徴を比較し、一番近い辞書を検索。画像特徴が全て同じ場合、同一番号で分類し登録、それ以外の場合は、文字・画像それぞれ共通部を新辞書として新分類番号で登録。

分類番号	: 1
共通画像特徴	: 3 1 1 - 0 1 1 -
相対関係	: 5
画像順序	: 逆順
共通文字概念	: 丸が三角形の上方にあります
入力文字列	: 丸が三角形の上方にあります
共通概念対	: 丸が三角形の上方にあります

- 4° 現入力データとn個前までの入力と比較し、共通部を抽出し、それが辞書にあるかどうかを検索し、同じものがなければ、新規登録。
- 5° 入力データがなくなるまで、2°へ。
- 6° 概念の出現頻度をもとに、どの概念がキーとなるかを決定する。

この場合、[が] [あります]の、2概念が必ず現れるので、この2つがキーとなる。但し、学習文に [あります] を含まないものを許したり、[が] を [は] に置換した場合は、閾値により選択する必要があるが、少なくとも特に偏りのない学習データを使用する限り、他の概念より多い出現頻度は保証される。

- 7° 次に相対関係に関する概念獲得のため登録された同一番号に分類された辞書中で、キーとなる概念の前後関係の共通なもの（前後両方またはどちらか）をグループ分けし、それぞれのグループ内の全ての学習文の文字概念・画像特徴について共通部と特徴部を抽出し、特徴部に関しては同じ概念同士にグループ分けする。

・丸が

共通画像特徴	: 3 1 1 - 0 1 1 -
相対関係	: 4
画像順序	: 基準
入力文字列	: 丸が三角形の上方にあります

共通画像特徴	: 3 1 1 - 0 1 1 -
相対関係	: 4
画像順序	: 同一
入力文字列	: 三角形の上方に丸があります

共通画像特徴：3 1 1 - 0 1 1 -
 相対関係 : 5
 画像順序 : 逆順
 入力文字列 : 丸が三角形の上方にあります

共通画像特徴：3 1 1 - 0 1 1 -
 相対関係 : 5
 画像順序 : 同一
 入力文字列 : 三角形の下方に丸があります

共通画像特徴：3 1 1 - 0 1 1 -
 相対関係 : 4
 画像順序 : 逆順
 入力文字列 : 丸が三角形の下方にあります

共通部

丸が三角形 あります 3 1 1 - 0 1 1 - - -

特徴部

の上方に - - - - - 4 同一 (入力2に対応)

の上方に - - - - - 5 逆順 (入力3に対応)

の下方に - - - - - 5 同一 (入力4に対応)

の下方に - - - - - 4 逆順 (入力5に対応)

8° 次に、同一分類番号の辞書内で画像特徴の全く同一のものをグループ化し、その中で
 の共通文字概念対を導出する。

丸が三角形の上方にあります

3 1 1 - 0 1 1 - 4 基準

三角形の上方に丸があります

3 1 1 - 0 1 1 - 4 同一

三角形の上方に丸があります

3 1 1 - 0 1 1 - 5 逆順

共通概念対 丸が 三角形の上方

5. 獲得される概念

4章で述べたアルゴリズムに従って逐次学習を行うことにより、文字情報と画像情報の
 内で、相対関係を表す概念がそれぞれ抽出できる。これらから得られる結果を、実際の文
 字・画像情報に対応づけてまとめると以下のような結論が得られる。

・7° から、[の上方に]、[の下方に]について、基準図形を逆順にすると相対関係の画像

パラメータが逆関係のものに変化することが獲得できる。さらに、それぞれの関係から[の上方に]、[の下方に] が逆の相対位置関係をもつことが導出できる。

・上記の8° から、[が]、[の上方に] の付随する図形が同一の場合：

画像順序同一→相対関係を表す画像パラメータ不変

画像順序逆順→相対関係の画像パラメータが逆関係のものに変化

といった結論が得られる。

つまり、[が] は、[の上方に]、[の下方に] といった相対関係（視点の方向）を表す概念と相互に関連し、文法上の主格に相当する概念として捉えることができる。また、左右の相対位置関係が混在しても同様に概念が獲得できる。

6. ま と め

本研究では、視覚と聴覚という2つの外的刺激を結びつけて未知の入力に対する相対概念の獲得方法について検討し実験を行った。これは計算機が環境から長期の自動学習を行う初期の段階をシミュレートする実験である。

その結果、逐次、画像と音声に対応づけながら学習を行い、蓄積結果を後処理することで、構文・意味解析等の処理を行うことなく、位置や大きさ等に関する相対概念が獲得でき、人間の相対概念獲得手法の初期段階が計算機上で実現できた。

今回は、かなり画像に対する説明文の表現が画一的な言い回しに限定されているが、今後表現方法を多様化した際の学習データ数と獲得概念数の関係、アルゴリズムの効率化等、また学習データに誤りが含まれる場合の性能評価等、さらには、同一の役割をする複数の概念や2種類以上の相対概念を同時に扱う際のアルゴリズムについても検討を進めていきたい。また、相対概念に関する結論を自動抽出するための手法、誤りを含んだデータに対する頑健なアルゴリズムの開発等も進める必要がある。

謝辞：本研究を行うにあたっては、財団法人堀情報科学振興財団より、第2回堀情報科学振興財団研究助成として、多大な援助をいただいた。ここに感謝の意を表します。

(平成6年9月12日受理)

参 考 文 献

- [古部 91] 古部好計, 中西宏文, 辰巳昭治, 中川聖一: 音声と画像の対応付けに基づく概念の獲得, 人工知能学会全国大会論文集, 2-10 (1991)
- [中西 92] 中西宏文: 視聴覚情報を用いた言語・概念の獲得, 愛知教育大学研究報告, No. 41 (自然科学編), pp. 49-60 (1992)
- [中川 93] 中川聖一, 中西宏文, 古部好計, 板橋光義: 視聴覚情報の統合化に基づく概念の獲得, 人工知能学会誌, Vol. 8, No. 4, pp. 109-118 (1993)