

# 名詞概念獲得システム SINCA における ラベル獲得規則の性能評価

内田 ゆず 荒木 健治  
Yuzu Uchida Kenji Araki  
{yuzu,araki}@media.eng.hokudai.ac.jp

北海道大学大学院 情報科学研究科  
Graduate School of Information Science and Technology, Hokkaido University

## 1. はじめに

近年, さまざまな用途に合わせたロボットが盛んに開発されており, 人と相互作用するようなものも登場している[1][2]. その中でも人と共同生活を営むことを前提としたロボットには高いコミュニケーション能力が必要であると考えられる. さらに, 日々変化する環境に対応するために, あらかじめ与えられた知識だけでなく, 動的に獲得した知識を実生活に応用する能力も必要であろう.

我々は, 人間の幼児の言語獲得能力を手がかりに, 語彙や文法の知識を持たない状態から名詞概念を獲得するシステム SINCA (System for Noun Concepts Acquisition from utterances for Image) の構築を行い, 性能評価を行ってきた[3][4]. SINCA は画像に関する内容の発話(書き起こしテキスト)を入力として得ることで, 名詞概念(画像に対するラベル)を獲得することが可能である. また, SINCA は名詞概念を獲得するにつれて, 名詞概念が含まれやすいフレーズを”ラベル獲得規則”として学習していく. 我々は, ラベル獲得規則で表されるフレーズがどの程度一般的で, 自然な表現なのかについて Web を用いた評価を行った[5]. その結果, Web を用いて規則の一般性を計り, 一般性の低い規則を淘汰することが可能であることを明らかにした.

しかし, 各々の規則が持つ名詞概念を抽出する能力の評価は行ってこなかった. 高い精度で名詞概念を抽出することの出来る規則が存在すれば, その規則を用いることで大規模な文書の中から未知の名詞概念を抽出することが可能となり, ブートストラップ型の名詞概念獲得システムの実現が可能になると考えられる.

本稿では, Web を用いた名詞概念獲得規則の性能評価実験と, その結果について述べる.

## 2. SINCAの概要

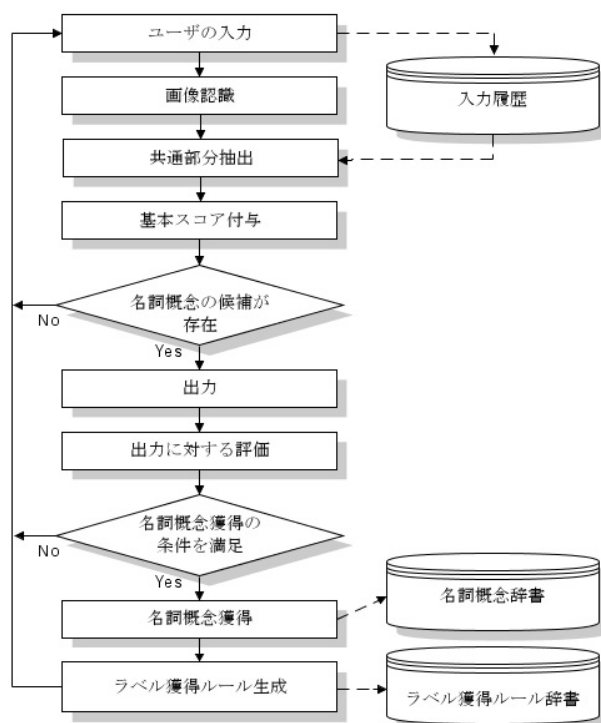


図 1 SINCA の処理の流れ

SINCA の処理の流れを図 1 に示す. また, 本章では, それぞれの処理について簡単に述べる.

### 2.1 入力

入力は画像と文の対である. 入力画像は Web カメラ (USB-CAMCHAT2/アイ・オー・データ機器. 有効画素数: 30 万画素) からキャプチャされた画像 (以降画像 P と呼ぶ), 入力文は画像 P を見せながらユーザが幼児に話かける発話 1 文 (以降文 S と呼ぶ) である. 入力画像は, ユーザが自由に被写体を選び撮影するものである.

入力文は全てひらがなで表記され、入力文に形態素解析などの前処理は一切施されない。ひらがなで表記するのは、ユーザによって表記に揺れが生じることと、入力された文字列自体に意味が含まれてしまうことを避けるためであり、形態素解析などを行わないのは、幼児が正確な品詞分割などの能力を持っていないと考えるためである。

## 2.2 画像認識

過去に同じ被写体が写った画像が入力されたかどうかを判断する。ここでは、エボリューション・ロボティクス社の ERSP3.1 (Evolution Robotics Software Platform) [6]に含まれる”ERSP ビジョン”を用いた。ERSP 3.1 は、ロボット製品の作成を目的とした総合開発プラットフォームで、照明や物体の位置が管理されていない現実的な環境の中でも、ロボットや装置が2次元と3次元の物体を認識することができる画像認識ツールである。

## 2.3 共通部分抽出

システムは入力を得ると、過去に画像 P とともに入力された文と文 S を比較して、字面が一致する文字列を切り出す。この切り出された文字列を共通部分と呼ぶ。これ以降の処理で共通部分は、画像 P に対応するラベルの候補として扱われる。

## 2.4 基本スコア付与

抽出された共通部分には基本スコアが付与される。基本スコアとは、その共通部分のラベルとしての確からしさを表した値であり、出現頻度が高く、文字数が多く、他の画像と共に出現することのない共通部分ほど高いスコアを与えられる。基本スコアの計算式は式(1)のようになる。

$$SCORE = \alpha \times \frac{F}{PN} \times \sqrt{L} \dots \dots \dots (1)$$

(1)式で、 $\alpha$  は共通部分が他の画像とともに出現している場合スコアを減少させるようにはたらく係数、 $F$  は共通部分が同一画像と共に出現した頻度、 $PN$  は画像の出現回数、 $L$  は共通部分の文字数である。

## 2.5 出力

2.4 で述べた方法で求めた基本スコアが閾値  $\theta$  を超えた共通部分は、画像 P に含まれる事物のラベルに適している可能性が高いと判断され、テキストで出力される。予備実験の結果、閾値  $\theta$  は 8.0 とした[7]。

## 2.6 出力に対する評価

システムの出力に対してユーザは次の 3 つのキーワードのうち、最も相応しいものを選び、入力する。

- ・「じょうず」：ラベルとして適切である
- ・「おいしい」：ラベルとしては適切でないが意味はわかる
- ・「ちがうよ」：意味がわからない

幼児がこれらのキーワードを完全に理解するとは考えられないが、実際には、大人の表情や声の調子で感じ取ることのできる情報は多い。本手法ではそれらの情報の代わりにキーワードを用いることとする。

ユーザの反応によってその共通部分のスコアは再計算される。具体的には、基本スコアに係数  $\beta$  を乗ずる。係数  $\beta$  は、予備実験から、ユーザの評価が「じょうず」の場合は 1.5、「おいしい」の場合は 0.8、「ちがうよ」の場合は 0.2 とした[6]。

## 2.7 名詞概念獲得

「入力」から「ユーザの評価」の処理を繰り返した結果、再計算されたスコアが閾値  $I$  を超え、さらに「じょうず」という評価を得たことがある共通部分は画像 P のラベルとして獲得される。予備実験の結果、閾値  $I$  は 30.0 とした[6]。

## 2.8 ラベル獲得ルール生成

ラベル獲得ルールとは、再帰的な名詞獲得を行うためのルールである。人間は過去に得た知識を活用し、より効率的に学習を進めていく。本手法ではそのような再帰的な学習を次のようにして実現している。

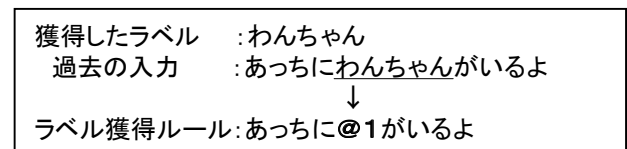


図 2 ラベル獲得ルールの生成例

システムが文字列 S をある事物に関する正しいラベルとして獲得すると、その事物に関する過去の入力文のうち、文字列 S を含む文から、ラベル獲得ルールを生成する。ラベル獲得ルールとは、図 2 のようにラベルの部分を変数とすることで、入力文を抽象化したものである。次に、生成したラベル獲得ルールに合致する入力文があった場合、変数部@1 に相当する部分を切り出し、スコアを上昇させる。

これは、人間は様々な表現を聞いているうちに、どのような表現がラベルを示すものなのかを学習して、より効率的に学習を進めていると考え、その様子をモデル化したものである。実際に、語彙爆発期の幼児は、一度大人が事物を指して言葉を発するのを聞いただけで、正しくその言葉を使うことができ

る（即時マッピング）ことが知られている[8]。ラベル獲得ルールを生成することで、本システムでもこれと似た現象を再現することができると考えられる。

### 3. 実験

#### 3.1 実験の目的

Web を日本語文書のコーパスとみなし、SINCA が生成するラベル獲得ルールの性能を評価する。ここでの性能とは、あるラベル獲得ルールによって名詞概念が切り出される可能性を表す。

#### 3.2 実験方法

SINCA が 30 枚の画像に対するラベルを獲得するまで入力を行う。この際に使われた入力文はアンケートによって 31 名の回答者から得た 396 文で、全ての文にラベルを示す単語が含まれている。この過程で生成された各ラベル獲得ルールについて、次のような処理を行う。

ラベル獲得ルールの変数部を、ワイルドカードを示す記号"\*"で置換し、ルール全体をダブルクォーテーションで囲む。このようにしてフレーズ検索が可能な形で検索クエリを生成する。Google でこれらのクエリを検索した際に出力される最大 50 件分のスニペットを対象として、ルールの変数部にあたる単語の品詞を調査する。名詞概念獲得に対するルールの有効性を次の条件により定める。

- 変数部が文字列に挟まれる形態のルール（例：これは@っていうんだよ）の場合、変数部に 1 つの名詞が対応する
- 変数部が文字列に挟まれない形態のルール（例：@がいるよ、かわいい@）の場合、変数部に対応する直近の 1 単語が名詞である

各ルールについてスニペットから得られた全フレーズについて、上記の条件を満足しているかどうかの分析を行った。上記の条件を満足しているフレーズの割合を有効率と呼ぶ。

スニペット中にはさまざまな単語が含まれ、形態素解析器では対応しきれないことが予想され、また、今回は品詞の獲得は研究の対象としていないため、品詞の決定は第一著者が人手で行った。

#### 3.3 実験結果

30 枚の画像に対するラベルを獲得するまでに、85 種類のラベル獲得ルールが生成された。そのうち、Google による検索ヒット件数が 1 件以上存在したルールは 61 種類であった。各ラベル獲得ルールと一致する表現のフレーズは 2,916 文抽出された。有効率が 100%であったルールは 9 種類存在した。80%以上であったルールは 22 種類存在し、これは全体の約 26%を占めている。また、検索ヒット数と有効率の間の相関係数は 0.04 であった。

各ルールが得た検索ヒット件数、有効率と変数部に対応する名詞の具体例を抜粋して表 1 にまとめる。

表 1 実験結果の抜粋

| ルール          | 名詞率 (%)        | 検索件数      | 名詞の例 |      |       |
|--------------|----------------|-----------|------|------|-------|
| かわいいお@だね     | 100.00 (49/49) | 109,000   | 人形   | 弁当   | 店     |
| @がこっちをみてるよ   | 100.00 (19/19) | 19        | カエル  | ウサギ  | お姉さん  |
| これは@っていうんだよ  | 98.04 (50/51)  | 329,000   | ベーグル | 手袋   | すごろく  |
| ここに@があります    | 97.56 (40/41)  | 573,000   | 花    | かご   | 病院    |
| @はおおきいね      | 84.21 (16/19)  | 653       | 影響   | 問題   | 魚     |
| だいすきな@だよ     | 81.40 (35/43)  | 71,500    | ケーキ  | 小説   | オムライス |
| @はちょっとこわいね   | 77.27 (17/22)  | 538       | 動物   | 火山   | 夜     |
| @ってかわいいでしょう  | 76.92 (30/39)  | 11,300    | 子犬   | 子供   | 男の子   |
| おかあさんの@だね    | 62.07 (18/29)  | 32        | 宝物   | 味方   | 誕生日   |
| @はわかるかな      | 60.78 (31/51)  | 79,300    | 理由   | 雰囲気  | 形     |
| これは@っていうの    | 55.32 (26/47)  | 57,000    | 亭主関白 | デコボン | シイタケ  |
| @だねえ         | 50.72 (35/69)  | 1,970,000 | 雨    | 電車   | 浪花節   |
| あかい@だ        | 49.21 (31/63)  | 472       | あくま  | お部屋  | 葉っぱ   |
| ちっちゃい@だねー    | 36.11 (13/36)  | 37        | 爪    | つらら  | 成功    |
| @っていえるようになった | 33.33 (17/51)  | 7,320     | ピザ   | ママ   | ゾウさん  |
| あれはね@        | 22.45 (11/49)  | 150,000   | 僕    | 丸の内口 | お茶漬け  |
| こ@さんかな       | 3.33 (1/30)    | 358,000   | ぶた   | -    | -     |
| おいしい@よー      | 0.00 (0/50)    | 55,500    | -    | -    | -     |

また、61種のラベル獲得ルールにおける有効率を図3にまとめる。

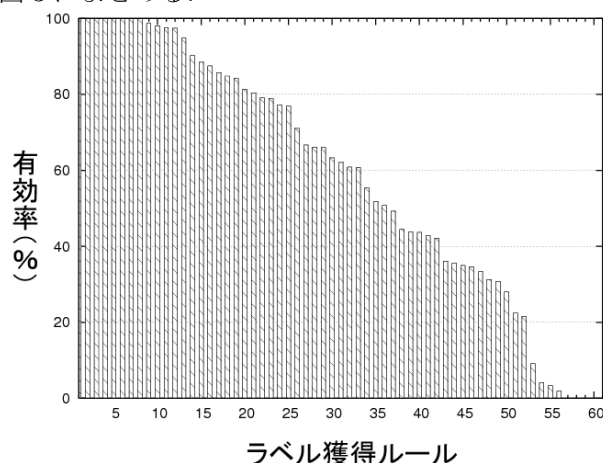


図3 各ラベル獲得ルールの有効率

#### 4. 考察

24種類（生成されたルール全体の28.2%）のラベル獲得ルールはGoogleによる検索ヒット件数が0件であった。これは、ルールが全てひらがなで表記されていることに原因の一端がある。SINCAの入力は全てひらがなで行われるため、ルールもひらがな表記になることは不可避である。しかし、全体の7割以上のルールは何らかの検索結果を得ることができているので大きな影響はないと考えられる。

名詞概念獲得ルールによって抽出された名詞に着目すると、Webの文書から得られた名詞は、具体名詞、抽象名詞、固有名詞など多岐にわたっている（表1参照）。ルールによって動物を表す名詞が抽出されやすい、抽象名詞が抽出されやすい、などの特徴も見られる。

検索ヒット数と有効率の間の相関は弱く、ルールの有効性は頻度（頻繁に使われる表現であるかどうか）に依存しないと言える。

次に、有効率の観点から考察する。全体の26%を占める22種類のルールが有効率80%以上を達成しており、高い精度で名詞の抽出が可能であることが明らかになった。このようなルールを有効に活用することで、SINCAは新たな名詞概念をブートストラップ式に学習することが可能であると考えられる。また、名詞だけでなく、形容詞、副詞、動詞などを抽出することが出来るルールを獲得することで、品詞の獲得の実現が期待できる。

#### 5. まとめ

名詞概念獲得システムSINCAにおける、名詞概念獲得ルールの有効性の評価実験を行った。その結果、生成された名詞概念獲得ルールの約26%は80%以上の割合で正確に名詞概念を抽出可能であることが明らかになった。この結果は、名詞概念獲得

ルールを用いることでSINCAがブートストラップ型の学習を実現させることが可能となることを示唆している。同様にして、他の品詞を抽出できるルールを生成することができれば、幅広い言語を獲得するシステムの構築も可能であると考えられる。

今後は、今回の実験結果をふまえて、名詞概念獲得ルールを有効活用することによってSINCAの性能向上を目指す予定である。また、我々はすでに実際の幼児と大人の対話データの収集を行っている。このデータを使ってSINCAと幼児の名詞概念獲得過程を比較する予定である。

#### 参考文献

- [1] T. Shibata, T. Mitsui, K. Wada, and K. Tanie : Subjective Evaluation of Seal Robot: Paro - Tabulation and Analysis of Questionnaire Results, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 13-19, 2002
- [2] 神田崇行, 石黒浩, 小野哲雄, 今井倫太, 中津良平 : 人間と相互作用する自律ロボット Robovie の評価, 『日本ロボット学会誌』, Vol. 20, No. 3, pp. 315-323, 2002
- [3] 内田 ゆず, 荒木 健治 : 画像に対する発話からの名詞概念の獲得, 情報処理学会研究報告, 2006-NL-176 , pp.81-86, 2006.
- [4] 内田ゆず, 荒木健治:"画像に対する発話からの名詞概念獲得システムにおける音素認識の導入について", 言語処理学会第13回年次大会発表論文集, pp.994-997, 2007.
- [5] Yuzu Uchida and Kenji Araki : A System for Acquisition of Noun Concepts from Utterances for Images Using the Label Acquisition Rules, Lecture Notes in Artificial Intelligence (LNAI) 4830, pp. 798-802.
- [6] エボリューション・ロボティクス社 : <http://www.evolution.com/products/ersp/>
- [7] 内田ゆず, 荒木健治 : 言語獲得システムにおける類似度に基づくラベル拡張手法の提案, 平成17年電気・情報関係学会北海道支部連合大会講演論文集, 180, 2005.
- [8] 今井むつみ : ことばの学習のパラドックス, 共立出版, 1997.