

物体の集合関係を含む参照表現の解析

渡辺聖[†] 船越孝太郎[†] 栗山直子^{††} 徳永健伸[†]

[†] 東京工業大学 大学院情報理工学研究科

152-8552 東京都目黒区大岡山 2-12-1

^{††} 東京工業大学 大学院社会理工学研究科

152-8552 東京都目黒区大岡山 2-12-1

E-mail: †{satoru_w,koh,take}@cl.cs.titech.ac.jp, ††kuriyama@hum.titech.ac.jp

あらまし 本論文では物体の集合関係を含む参照表現の解析を扱う。実/仮想世界における対話ロボットの重要なタスクの1つとして参照表現の利用が挙げられる。そこで我々は集合関係を含む参照表現を採集するため、被験者実験を行なった。その実験データに基づいて、参照表現を解析し対象物体を特定するための手法を提案する。提案手法の実装システムからは、79%の正解率が得られている。

キーワード 参照表現, 知覚的集合化, 対話ロボット, 参照解決, 対象特定

Understanding Referring Expressions Involving Perceptual Grouping

Satoru WATANABE[†], Kotaro FUNAKOSHI[†], Naoko KURIYAMA^{††}, and Takenobu
TOKUNAGA[†]

[†] Graduate School of Information Science and Engineering, Tokyo Institute of Technology
Ookayama 2-12-1, Meguro-ku, Tokyo, 152-8552 Japan

^{††} Graduate School of Decision Science and Technology, Tokyo Institute of Technology
Ookayama 2-12-1, Meguro-ku, Tokyo, 152-8552 Japan

E-mail: †{satoru_w,koh,take}@cl.cs.titech.ac.jp, ††kuriyama@hum.titech.ac.jp

Abstract This paper deals with understanding referring expressions involving perceptual grouping. Using referring expressions is an important ability of conversational robots aimed at real/virtual world interaction. We conducted a psychological experiment to collect referring expressions involving perceptual grouping. Based on the collected data, a set of methods to identify referents are presented. Our method could correctly identify the referents of the collected expressions at accuracy of 79%.

Key words referring expression, perceptual grouping, conversational robots, referential resolution, object identification

1. 序 論

参照表現とは対象物体を他の物体と混同せずに特定する言語表現である。参照表現の利用は実/仮想世界における対話ロボットの重要なタスクの1つである [1]。参照表現には談話文脈 (例:「さっきの〜」) やパラ言語情報 (話し手の視線や指差し動作) などを利用する表現も含まれるが、本論文においては対象物体固有の属性や他の物体との関係のみを用いる表現を研究対象とする。(例:「ドアの前にある赤い椅子」)

我々は参照表現の理解と生成における知覚的集合化 (Perceptual Grouping) の重要性を指摘した [11]。知覚的集合化 [9] と

は外見的特徴の差異の少ない物体や相互に近接した物体を1つの集合として認識することである。従来の参照表現の生成手法 [2-4, 6, 7] は、対象物体固有の属性や他の物体との2項関係のみを用いることが多かった。そのため、指示すべき物体とその他の物体との間に外見的特徴の有意な差がなく、他の物体との位置関係も弁別用の用を成さないような状況では、適切な参照表現を生成することができない。そのような状況においては、知覚的集合化に基づく表現生成が必要になる。

例として図1の状況を考える。属性や2項関係のみを利用する従来のアルゴリズムでは、物体cを示すために「玉の手前の玉で玉の奥の玉」といった不自然な表現を生成してしまうか、

参照表現の生成は失敗となっていた。例1のような自然な表現を生成するためには、外見的に類似した物体もしくは近接した物体を集約化し、その集合の中でのn項間関係を利用する必要がある。

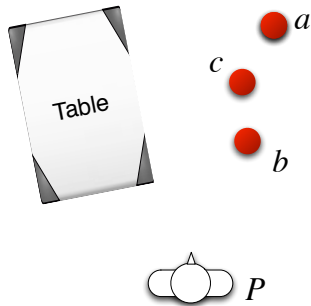


図1 従来の手法が適用できない例

図1のような状況においては、物体 a, b, c からなる集合を認識することで例1のような表現の理解や生成が可能になる。

(例1) 真ん中の玉

本論文では、物体の集合関係を含む参照表現の解析手法を提案する。参照表現の解析は以下の2段階からなる。(1) 意味解析では、意味情報を抽出するための言語解析を行なう。(2) 対象特定では、得られた意味情報を利用して対象物体の選別を行なう。

以降、2節では被験者実験によって収集したデータと集合関係を含む参照表現の構造について示す。3節と4節ではそれぞれ(1) 意味解析と(2) 対象特定について説明する。そして5節で提案手法の実装結果について考察し、6節でまとめとする。

2. 収集データ

我々は集合関係を含む参照表現を集めるために、大学生42人を対象に実験を行なった。以降、実験手法と収集データから得られた参照表現の構造について示す。

2.1 実験手法

同形・同大・同色の物体を複数個配置した2次元画像を被験者に示し、破線で囲まれた物体を画像中に描かれた第3者に伝える場合、どのように表現するかを尋ねた。物体の配置は、3~9個の物体を人為的に非整列に配置したもの12種類を用意した。図2に被験者に提示した画像の一例を示す。ただし画像中のラベル a, b, \dots, f, x は説明のために付与したものであり、被験者には示していない。各被験者の具体的なタスクは、画像内の人物に対して破線で囲まれた物体を取らせるような命令を生成することである。ただし適切な表現を生成することが不可能であると判断した場合には放棄することも許し、次に進ませた。対象物体を示す参照表現はこれら被験者の生成した命令から採集した。

2.2 参照表現の構造

上記の実験では42人の学生にそれぞれ12個の配置を示したので合計504表現が収集される予定であったが、放棄が15個と対象物体特定が明らかに不可能であると判断される表現が

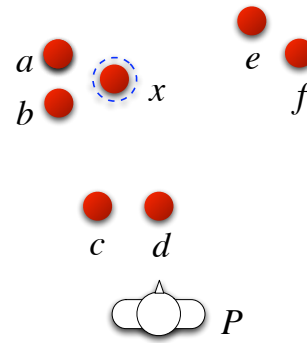


図2 表現採集に用いた視覚刺激の例

13個存在したため、収集された有効な参照表現は476個であった。得られた参照表現の分析から、参照表現は物体全体の表現から始まり徐々に集合を絞り込みながら対象物体に至るというような表現であることがわかった。よって参照表現は被験者の絞り込みの処理過程に即した集合の列(SOG, Sequence Of Groups)として定式化できる。被験者の絞り込みの過程では複数の集合が参照される状況が多く見られた。このような場合、それらの集合間の関係は内部包含関係と外部参照関係の2種類に分類できることがわかった。関係は2種類存在するが、収集した参照表現全体の約73%は包含関係のみを含む表現であった。以下に示す例は図2において対象物体 x を特定する表現で、実際に収集されたものである。参照表現に対応するSOGをその下に示してある。

(例2) 左奥にある3つの玉のうち一番右の玉

SOG: $\{[a, b, c, d, e, f, x], [a, b, x], \{x\}\}$ ^(注1)

ここで

- $\{a, b, c, d, e, f, x\}$: 画像中の全物体 (全体集合)
- $\{a, b, x\}$: 左奥の3つの玉
- $\{x\}$: 対象物体

絞り込みは必ず全体集合から始まるのでSOGもまた全体集合から始まり対象物体のみを含む集合で終わるという形となる。以上のように、収集された参照表現をSOGに変換することによって表現を抽象化し分類することができる。得られた表現は1配置につき40個程であるが、平均8.4個のSOGに分類することができた。このデータを表1に示す。

3. 意味解析

本論文では言語表現から必要な情報を抽出するために構文解析ではなく単純なパターンマッチを用いる。これは収集した表現があまりに口語的でありそのバリエーションが多岐に渡るため、それらに完全に対応する文法を作成することが困難であるからである。当然ながらパターンマッチによる処理には限界があり、構文解析による処理を今後検討する必要がある。しかし、本論文で用いた手法は当座の目標を果たすには十分に機能している。

(注1): SOGは大括弧で囲んで表現する。

表 1 布置ごとの表現数

布置番号	1	2	3	4	5	6	7	8	9	10	11	12	平均
有効表現数	41	40	41	41	42	37	42	32	42	41	41	36	39.7
SOG のことなり数	5	6	8	8	6	12	4	15	4	11	5	17	8.4

意味解析は以下の 3 段階からなる。

(1) 分割

参照表現を集合に対応する部分表現に分割する。

(2) 情報抽出

分割された各部分表現から意味情報を抽出する。

(3) 関係特定

部分表現の示す物体集合の間の関係を特定する。

3.1 分割

第一段階として、参照表現を部分表現に分割する。この処理は SOG($[G_0, G_1, \dots, G_k]$) の形式を生成することに対応する。ここで集合の数が決まるが、各集合の要素は未定である。 G_0 は全体集合で、 G_k は対象物体である。ここではパターンマッチにより部分表現の境界を探し出している。境界は、例えば「のうちの」というような手がかり語により検出される。手がかり語は 2 節の収集データから抽出した。

日本語は主要部後置型であるので、部分表現の順序は SOG の集合の順序と同一となる。ただし全体集合は言及されないことが多い。

例えば、2.2 節の例 2 は以下の 2 つの部分表現に分割される。

(例 3) 左奥にある 3 つの玉のうちの

(例 4) 一番右の玉

これらの境界は「のうちの」で表現されている。例 3 は図 2 の中の集合 $G_1 : \{a, b, x\}$ に対応しており、例 4 は対象物体の集合 $G_k : x$ に対応している。

3.2 情報抽出

第 2 節で示したドメイン設定では、物体や集合を特定するために利用される情報の型の種類は以下に示されるものに限定される。情報は各部分表現から正規表現によって抽出する。ただし同種の情報は同一部分表現内には 1 つしか現れないと仮定する。

- 物体の種類
- 集合の要素数
- 位置情報
- 序数

例 3 からは以下の情報が抽出できる。

- 物体の種類は「玉」である。
- 集合の要素数は「3」である。
- 集合は「左奥」にある。

また例 4 からは以下の情報が抽出できる。

- 物体の種類は「玉」である。
- 集合の要素数は「1」である。
(最後尾の部分集合は単体を示すはずである。)
- 集合は「最も右」にある。

3.3 関係の特定

上述のように、各部分表現は特定の集合を 1 つ示している。(物体が 1 つのときも集合として扱う。) このステップでは前のステップで抽出された情報から、連続する 2 つの集合の間の関係が内部包含関係と外部参照関係のどちらであるかを決定する。

内部包含関係の場合は、さらに以下の 3 つの下位関係に分類される。

- 有向幾何関係
例：3 つのうちの 右 の玉
- 無向幾何関係
例：真ん中 の玉
- 順序関係
例：右から 3 番目 の玉

また外部参照関係は以下の 2 つの関係に分類される。

- 有向幾何関係
例：3 つの玉の 右 の玉
- 無向幾何関係
例：3 つの玉の 最も近くにある 玉

関係の特定は優先度付きの規則を適用して行なう。規則は収集データを参考に人手で作成した。

4. 対象物体の特定

前段階で意味解析を行なったので、この節では抽出された意味情報を利用して対象物体を特定するアルゴリズムを説明する。本論文では各布置の参照枠 [8] は適切に解決されているものとする。

対象物体を特定する処理は、左から右へ、すなわち G_1 (注2) から G_k (G_i については節 2.2 と節 3.1 を参照) へ対象物体特定アルゴリズムを順次適用して集合の要素を決定していくことによって行なう。 G_i に適用するアルゴリズムは G_i と G_{i-1} の関係によって決定する。要素の決定した G_i は G_{i+1} の要素決定に利用する。この処理は、たとえ参照表現が曖昧であっても意味解析処理と同様に決定的に行なう。

節 3.3 の各関係に対する 5 つの対象物体特定アルゴリズムを以下に示す。各アルゴリズムは、引数として (1) 物体の集合と (2) 対象を特定する情報をとる関数として与えられる。

4.1 内部包含関係

4.1.1 有向幾何関係

この関係は「手前の」や「右の」のような表現に対応し、 $Id_{intra}(C, n, \vec{d})$ という関数として与えられる。引数 C は物体の集合であり、直交座標形で表現された座標を持つ。この C は全体集合 G_0 であるか 1 ステップ前に特定された集合 G_{i-1} である。引数 n は特定すべき集合 G_i の要素数である。また \vec{d} は位

(注2) : G_0 は常に全体集合であり、処理を必要としない。

置情報に対応する方向ベクトルであり、参照枠によって変化する。もし参照軸が上向き（y 軸方向）であるならば「右」に対応するベクトルは (1,0) となる。実装では右・左・手前・奥の方向とそれぞれの中間方向をとって、8 方向を利用している。関数 $Id_{intra}(C, n, \vec{d})$ の出力は要素数 n の集合となる。

関数 $Id_{intra}(C, n, \vec{d})$ はまず、全ての組み合わせの要素数 n の集合 $C_{S_1}, C_{S_2}, \dots, C_{S_m}$ を C から生成する。ここで各集合 C_{S_i} から式 (1) で表現されるスコア $is(C_{S_i})$ を計算する。 $is(C_{S_i})$ は C_{S_i} の各要素の近接性を示している。 a は定数である。 $dist(v1, v2)$ は C_{S_i} の要素である $v1$ と $v2$ の間の距離である。

$$is(C_{S_i}) = \begin{cases} \prod_{v1, v2 \in C_{S_i}} \exp(-a \cdot dist(v1, v2)) & |C_{S_i}| > 1 \\ 1 & |C_{S_i}| = 1 \end{cases} \quad (1)$$

C_{S_i} の最終的なスコア $s(C_{S_i}, \vec{d})$ は式 (2) から得られる。 $c(\vec{v}, \vec{d})$ は \vec{v} の \vec{d} 方向の成分であり、 \vec{v} は C_{S_i} の要素である v の位置ベクトルである。

$$s(C_{S_i}, \vec{d}) = is(C_{S_i}) \prod_{v \in C_{S_i}} \frac{c(\vec{v}, \vec{d})}{C_{S_i, \vec{d}}^{max}} \quad (2)$$

ただし、

$$c(\vec{v}, \vec{d}) = \frac{\vec{v} \cdot \vec{d}}{|\vec{d}|}$$

$$C_{S_i, \vec{d}}^{max} = \max_{v \in C_{S_i}} c(\vec{v}, \vec{d})$$

関数 $Id_{intra}(C, n, \vec{d})$ は式 (3) で与えられる。

$$Id_{intra}(C, n, \vec{d}) = \operatorname{argmax}_{C_{S_i} \in C_{S^n}} s(C_{S_i}, \vec{d}) \quad (3)$$

図 3 は関数 Id_{intra} の説明図である。実線で囲まれた 2 つの玉が出力となっている。 Id_{intra} は、より \vec{d} の方向に位置する物体を選択している。

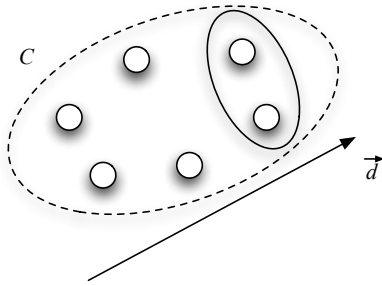


図 3 Id_{intra} の説明図 ($n = 2$)

ここで例として、図 2 に対して例 3 の部分表現「左奥にある 3 つの玉のうちの」が参照している物体を考える。3.1 節で示したように、例 3 の部分表現は G_1 (要素は未定) に対応している。ここで引数 C は G_0 、つまり全体集合 $\{a, b, c, d, e, f, x\}$ (2.2 節) である。要素数 n は 3 であり、 \vec{d} は図 2 の人物 P の視点からの「左奥」の方向を表すベクトルである。以上の入力から、関数 $Id_{intra}(C, n, \vec{d})$ の計算結果は $\{a, b, x\}$ となる。

4.1.2 無向幾何関係

この関係は「中心にある」というような表現に対応し、この関数は式 (4) の $Ind_{intra}(C, n)$ として与えられる。引数の C と n は前節と同様である。

$$Ind_{intra}(C, n) = \operatorname{argmax}_{C_{S_i} \in C_{S^n}} is(C_{S_i}) \cdot js(C_{S_i}, c) \quad (4)$$

式 (4) の c は集合 C のバウンディングボックスの中心である。関数 $js(C_{S_i}, c)$ は式 (5) で与えられ、 c と集合 C_{S_i} の各要素との近接性を示している。

$$js(C_{S_i}, c) = \prod_{v \in C_{S_i}} \exp(-a \cdot dist(v, c)) \quad (5)$$

4.1.3 順序関係

この関係は「左から 3 番目」というような表現に対応し、式 (6) の $Io(C, j, \vec{d})$ という関数として与えられる。引数 C は物体の集合であり、直交座標形の座標を持っている。 j は序数であり、 \vec{d} は位置的順列^(注3)の方向を表すベクトルである。順序を用いる場合、その部分表現はただ 1 つの物体を特定しているということが収集データより類推される。ゆえに関数 Io の出力は単数の物体となる。

$$Io(C, j, \vec{d}) = \operatorname{arg}_{v \in C}^j c(\vec{v}, \vec{d}) \quad (6)$$

$\operatorname{arg}_{x \in S}^j f(x)$ は、集合 S の要素の中で関数 f により算出される値が j 番目に大きいものを表している。

例として (例 4) の部分表現「一番右の玉」が参照している物体を考える。引数 C は前ステップで (例 3) の部分表現の参照物として特定された集合 $G_1: \{a, b, x\}$ である。序数 j は 1 であり、 \vec{d} は図 2 の人物 P の視点からみた「右」を表すベクトルである。以上の入力から、 $Io(C, j, \vec{d})$ の計算結果は $\{x\}$ となる。

4.2 外部参照関係

4.2.1 有向幾何関係

この関係は「～の右の」というような表現に対応し、式 (7) の関数 $Id_{inter}(T, r, n, \vec{d})$ として与えられる。引数 T はこのドメインにおける全体集合である。 r は外部参照関係の参照物であるが、これは集合 G_{i-1} となる。 n と \vec{d} は内部包含関係 (4.1 節) のものと同じである。

関数 $Id_{inter}(T, r, n, \vec{d})$ はまず全体集合 T の要素から、全ての組み合わせの要素数 n の集合 $T_{S_1}, T_{S_2}, \dots, T_{S_m}$ を生成する。生成された集合 $T_{S_1}, T_{S_2}, \dots, T_{S_m}$ の集合を T_{S^n} とする。

$$Id_{inter}(T, r, n, \vec{d}) = \operatorname{argmax}_{T_{S_i} \in T_{S^n}} t(T_{S_i}, r, \vec{d}) \quad (7)$$

ただし

$$t(T_{S_i}, r, \vec{d}) = is(T_{S_i}) \cdot js(T_{S_i}, r) \cdot ks(T_{S_i}, r, \vec{d})$$

関数 $is(T_{S_i})$ は式 (1) のものと同じである。 T_{S_i} は T の要素数が

(注3)：集合内の関係には位置関係以外にも、大きさや色などの固有属性の程度の差を表す関係 (例：「2 番目に大きな玉」「最も明るい色の玉」) も存在するが、今回は位置関係による順序のみを対象とする。

n である部分集合である。 T_{s_i} の要素 v は $c(\vec{v}-\vec{r}, \vec{d}) > 0$ ^(注4) を満たさなければならない。 関数 $js(T_{s_i}, r)$ は式 (5) のものと同じである。 関数 $ks(T_{s_i}, r, \vec{d})$ は式 (8) で与えられる。 $\theta^{\vec{d}, \vec{v}-\vec{r}}$ は \vec{d} と $\vec{v}-\vec{r}$ の挟角である。

$$ks(T_{s_i}, r, \vec{d}) = \prod_{v \in T_{s_i}} \exp(-\tan^2 \theta^{\vec{d}, \vec{v}-\vec{r}}) \quad (8)$$

$$= \prod_{v \in T_{s_i}} \exp\left(-\left(\frac{|(\vec{v}-\vec{r})||\vec{d}|}{(\vec{v}-\vec{r}) \cdot \vec{d}}\right)^2 + 1\right)$$

図 4 は Id_{inter} の説明図である。 実線で囲まれた 2 つの物体が選択された物体である。 Id_{inter} は参照物 r との距離が短く、より \vec{d} に近い方向にある物体を選択する。

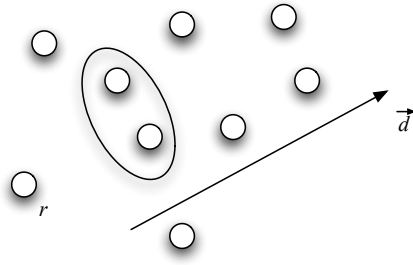


図 4 Id_{inter} の説明図 ($n = 2$)

関数 is や関数 js , 関数 ks は参考文献 [10] で定義されるポテンシャル関数である。ポテンシャル関数は位置の座標の尤もらしさを表している。ここで、 a の値は 0.001 としている。

4.2.2 無向幾何関係

この関係は「～の近く」というような表現に対応し、式 (9) の関数 $Ind_{inter}(T, r, n)$ として与えられる。全体集合 T や参照物 r , 要素数 n は上述のものと同様である。

$$Ind_{inter}(T, r, n) = \operatorname{argmax}_{T_{s_i} \in T_s^n} is(T_{s_i}) \cdot js(T_{s_i}, r) \quad (9)$$

関数 $is(T_{s_i})$ と関数 $js(T_{s_i}, r)$ はそれぞれ式 (1) と式 (5) のものと同様である。

式 (9) と式 (4) はほとんど同じであるが、引数に違いがある。

この関係の特別な例として、否定の関係がある。「～を除いて」や「～を省いて」というような表現に対応していて、その関数 $Ie(T, r)$ は式 (10) で与えられる。引数 T, r は上述のものと同様である。

$$Ie(T, r) = T - r \quad (10)$$

上式 $T - r$ は集合 T と集合 r の差集合を表している。

(注4): 「～の隣の」という表現の場合、左右双方向に適用できるのでこの制約は必要ない。

5. 評価結果と考察

5.1 評価結果

3 節と 4 節の手法を Perl で実装し 2 節の収集データに適用した。

表 2 は全ての関係 (内部包含関係と外部参照関係) を含む表現に対しての解析結果である。しかし、収集データの中には提案手法が適用不可能な表現も存在する。以下にこのような表現を分類する。

- 複数の集合を並列的に参照する表現
例: 「右と左にそれぞれ玉が固まっていて～」
- 複数の文から成る表現
- 幾何学的配置を利用する表現
例: 「2 列目の～」 「～を平行四辺形として」
- etc.

これらの表現は原理的に提案手法では扱えない。表 2 の上段は全表現についてのもので、下段は提案手法が適用可能な表現についての解析結果である。各分類に対して、表現数と意味解析の成功率、対象物体特定の成功率を示している。上段の対象物体特定の成功率が参照表現解析の最終的なパフォーマンスである。

441 個の適用可能な表現のうち、分割の境界は 259 個存在した。実装システムの検出したものは 257 個であるが、そのうち 254 個が正しい境界であった。よって分割の精度は 98.9% (254/257) となり、再現率は 98.1% (254/259) となる。

提案手法が適用不可能な 35 表現のうち 8 表現で対象物体の特定が成功している。これにより表 2 の Ident. は 1.7% 上昇している。提案手法がこれらの表現の意味情報を十分に利用できなかったということを考えると、これらの成功は偶然の産物である。しかしこれは本手法の頑健さとすることができ、実用性の観点からは重要であると考えられる。

表 2 実験結果: 全表現

参照表現	表現数	意味解析	対象特定
全表現	476	87.0%	79.0%
対象表現	441	93.9%	83.4%

表 3 は内部包含関係のみを含む表現についての結果である。また表 4 は内部包含関係のみでなく、外部参照関係をも含む表現に対しての解析結果である。

表 3 実験結果: 内部包含関係のみ

参照表現	表現数	意味解析	対象特定
全表現	349	91.4%	86.8%
対象表現	332	96.1%	89.2%

表 4 実験結果: 外部参照関係を含む

参照表現	表現数	意味解析	対象特定
全表現	127	74.8%	57.5%
対象表現	109	87.2%	66.1%

次に我々は44人の被験者に2節の収集データ476表現のうちの117表現の参照物体特定作業をさせるという実験を行なった。被験者は2節とは全く異なっている。この実験から、各表現につき12~13個のデータが得られた。表5は117表現に対する人と本手法の対象物体特定における正解率の比較を示している。この結果から、本手法のパフォーマンスは人のパフォーマンスに近いことがわかる。3行目ではシステムの正解率が人に勝っているが、実験の母数が少ないため有意な差ではないと考える。

表5 正解率：人 vs. 提案手法

参照表現	表現数	人	提案手法
全表現	117	87.3%	84.6%
包含関係のみ	78	94.4%	88.5%
参照関係含む	39	75.1%	76.9%

5.2 誤り解析

5.2.1 意味解析段階

5.1節に示したように、表現の分割(3.1節)については精度、再現率共に十分な成果が得られている。また集合間の関係の特定(3.3節)でも失敗はほとんどみられなかった。

意味解析の段階での失敗の多くは情報抽出(3.2節)の失敗によるものであった。その表現の多くは5.1節に示したように提案手法が適用不可能な表現である。2.2節に示したようにSOGは直列的に物体集合を参照しているため、並列処理を必要とするような表現からは十分な情報を抽出することができない。このような表現は18表現存在した。

また複数の文章から成り立つ表現が10表現存在するが、このような表現もSOGで適切に扱うことができない。

5.2.2 対象特定段階

「右の方にあるうちの、真ん中の玉」というような、その要素数に言及せずに集合を示す表現が存在した。「右の方にあるうちの」というような要素数の情報が利用できない表現から物体集合を特定するのは困難である。しかし、人は高い集合認識能力によってこのような表現の特定する物体を理解することが可能である。[9]に示される集合認識の手法がこのような状況に利用できるかもしれない。

10個の表現は幾何学的配置を利用していた。そのうちの7個は列を利用したものであり、残りは他の図形を利用したものであった。これは[9]に示される集合認識手法を利用しても扱うことができない。このような幾何学的配置を認識できる手法を開発する必要がある。

6. まとめ

本論文では、集合の関係を含む参照表現の解析について取り組んだ。それは2つの段階で成り立っており、それぞれ意味解析と対象物体の特定である。1段階目では被験者実験によって収集された参照表現について、単純なパターンマッチを用いて解析を行なった。これは87.0%の成功率を得た。次に意味解析によって表現から抽出された情報を用いて対象物体の特定を行

ない、その成功率は79.0%であった。

本論文で利用した物体の属性に関する情報は限られているが、提案手法は色や大きさのような物体の異なる属性への対応、拡張も容易であると考えられる。

単文の参照表現に対しての解析としては本論文の結果は十分であり、改良を行なっても飛躍的なパフォーマンスの上昇は得られないであろう。今後の課題としては提案手法を利用して対話的に対象物体を特定することが考えられる。そのような状況ではユーザとシステムは談話を通して複数の発話によって漸進的に対象物体を特定していく(例5)。このような機能により対話ロボットの利便性は大幅に増えるであろう。[5]に示される参照表現生成モデルはよい基盤となるだろう。

(例5) 談話による対象物体特定の例

U: 丸い机が見えますか?

S: はい。

U: その手前に玉がありますね?

S: 3つの赤い玉があります。

U: 左のやつを取ってください。

S: はい。

文 献

- [1] Donna K. Byron. Understanding referring expressions in situated language: Some challenges for real-world agents. In *the First International Workshop on Language Understanding and Agents for the Real World*, 2003.
- [2] Robert Dale. Generating referring expressions: Constructing descriptions in a domain of objects and processes, 1992. MIT Press, Cambridge.
- [3] Robert Dale and Nicholas Haddock. Generating referring expressions involving relations. In *Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics(EACL'91)*, pp. 161-166, 1991.
- [4] Robert Dale and Ehud Reiter. Computational interpretations of the gricean maxims in the generation of referring expressions. *Cognitive Science*, Vol. 19, No. 2, pp. 233-263, 1995.
- [5] Peter Heeman and Graem Hirst. Collaborating referring expressions. *Computational Linguistics*, Vol. 21, No. 3, pp. 351-382, 1995.
- [6] Emiel Krahmer and Mariët Theune. Efficient context-sensitive generation of descriptions, 2002. In Kees van Deemter and Rodger Kibble, editors, *Information Sharing: Givenness and Newness in Language Processing*. CSLI Publications, Stanford, California.
- [7] Emiel Krahmer, Sebastiaan van Erk, and André Verleg. Graph-based generation of referring expressions. *Computational Linguistics*, Vol. 29, No. 1, pp. 53-72, 2003.
- [8] Stephen C. Levinson, editor. *Space in Language and Cognition*. Cambridge University Press, 2003.
- [9] K. R. Thórisson. Simulated perceptual grouping: An application to human-computer interaction. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 876-881, 1994.
- [10] Takenobu Tokunaga, Tomofumi Koyama, and Suguru Saito. Meaning of Japanese spatial nouns. In *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics: Formalisms and Applications*, pp. 93-100, 2005.
- [11] 渡辺聖, 船越孝太郎, 栗山直子, 徳永健伸. 集合内の関係を用いた参照表現の生成. 言語処理学会第10回年次大会, pp. 516-519, Mar 2004.