

# 知覚的群化を利用した参照表現の生成

船越 孝太郎<sup>†</sup>

渡辺 聖<sup>†</sup>

徳永 健伸<sup>†</sup>

参照表現とは、特定の物体を他の物体と混同することなく識別する言語表現である。参照表現の生成に関する従来の研究では、対象物体固有の属性と異なる2つの物体間の関係を扱ってきた。しかし外見的特徴の差異が少なく他の物体との関係が対象物体の特定に用を成さない場合、従来の手法では対象物体を特定する自然な参照表現を生成することはできない。この問題に対して我々は知覚的群化を利用した参照表現の生成手法を提案しているが、この手法が扱える状況は強く限定されている。本論文では、我々が提案した手法を拡張し、より一般的な状況に対応できる参照表現の生成手法を提案する。18人の被験者に対する心理実験をおこない、本論文の提案手法を実装したシステムが適切な参照表現を生成できることを確認した。

キーワード: 参照表現, 空間記述, 知覚的群化, 言語生成

## Generating Referring Expressions Using Perceptual Grouping

KOTARO FUNAKOSHI<sup>†</sup> SATORU WATANABE<sup>†</sup> TAKENOBU TOKUNAGA<sup>†</sup>

Referring expressions are the linguistic representations that identify the referent among objects. Past work of generating referring expressions mainly utilized attributes of objects and binary relations between objects. However, such an approach does not work well when there is no distinctive attribute among objects. To overcome this limitation, we proposed a generation method using perceptual grouping of objects. However, this method can deal with very limited situations. This paper proposes an extended method using perceptual grouping that can deal with more general situations. The psychological experiments with 18 subjects showed that the extended method could effectively generate proper referring expressions.

**KeyWords:** *Referring Expression, Spatial Description, Perceptual Grouping, Language Generation*

## 1 はじめに

参照表現の生成は自然言語処理の重要なタスクの1つであり (Byron 2003), 多くの研究者により様々な手法が提案されてきた (Appelt 1985; Dale and Haddock 1991; Dale 1992; Dale and Reiter 1995; Krahmer and Theune 2002; Krahmer, van Erk and Verleg 2003) .

---

<sup>†</sup> 東京工業大学 大学院情報理工学専攻

Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology

参照表現生成に関する従来の研究は、主に対象物体固有の属性と他の物体との関係を扱ってきた。ただし、他の物体との関係は2項関係のみである。そのため従来の手法では、指示すべき物体とその他の物体との間に外見的特徴の差異が少なく、他の物体との2項関係も弁別の用を成さない状況において、適切な参照表現を生成することができない。ここで適切な参照表現とは、自然で過度な冗長性のない表現のことを言う。

例として、図1において対象物体  $c$  を人物  $P$  に示すことを考える。対象物体  $c$  は、外見からは物体  $a$  や物体  $b$  から区別することができない。そこで次の方策として、対象物体  $c$  とテーブルとの間の関係を用いることが考えられる(例えば、「テーブルの右の玉」)。しかし、物体  $a$  も物体  $b$  もテーブルの右にあるため、この状況においては「 $X$  の右の  $Y$ 」という関係に弁別能力はない。テーブルの代わりに物体  $a$  や物体  $b$  を参照物として使うことも意味がない。なぜなら物体  $a$  および物体  $b$  は物体  $c$  が一意に特定できないのと同じ理由によって一意に特定することができないからである。このように、物体の属性と2項関係のみを用いる従来の手法では参照表現の生成に失敗する。手法によっては「玉の前の玉の前の玉」のような論理的には誤りでない表現を生成できるが、適切な参照表現ではない。このような状況は今まで注目されてこなかったが、物体配置の様な状況(例えば(Tanaka, Tokunaga and Shinyama 2004))では頻繁に起こりうる。この場合、「一番手前の玉」という表現が自然かつ簡潔であると考えられる。このような参照表現を生成するためには、話し手は知覚的に特徴のある物体群を認識し、群に含まれる物体の間の  $n$  項関係を用いる必要がある。

この問題に対し、我々は知覚的群化(Thórisson 1994)を用いて物体群を認識し、物体群の間の関係を用いて参照表現を生成する手法を提案した(船越, 渡辺, 栗山, 徳永 2006)。知覚的群化(perceptual grouping)とは外見的に類似した物体や相互に近接した物体を1つの群として認識することである。我々の提案した手法によって物体の  $n$  項関係を利用した参照表現の生成が可能となったが、この手法の想定する状況は同形同色同大の物体を複数配置した2次元空間と

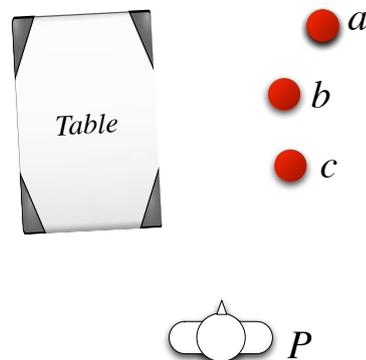


図1: 従来手法で表現生成が困難な例

いう非常に限られたものであったため、一般的な状況には対応できなかった。

本論文では、我々が提案した手法を拡張し、従来より利用されてきた色、形、大きさ等の属性や2項関係も利用できる、知覚的群化に基づく参照表現の生成手法を提案する。(船越他 2006)では、知覚的群化を利用して参照表現を生成するために、参照表現と参照する空間の状況とを結びつける SOG (Sequence of Groups) という中間表現形式を提案した。本論文では、SOG が包含関係以外の関係や物体の属性も表現できるように拡張する。そして拡張した SOG を用いた生成手法を提案し、大学生 18 人に対する心理実験によって実装システムが生成した参照表現を評価する。

本論文の構成は以下の通りである。まず2節では(船越他 2006)で提案した SOG について説明し、その拡張を行なう。3節では拡張した SOG を用いて知覚的群化に基づく参照表現生成手法を提案する。そして提案手法の評価と考察を4節に示す。最後に5節で本論文の結論と今後の課題を述べる。

## 2 SOG

我々は、参照する空間の状況と対象物体を特定する参照表現との間の中間表現として、SOG (Sequence Of Groups) を提案した(船越他 2006)。これは、物体全体の群  $G_0$  から始まり対象物体のみを含む物体群  $G_n$  に至る物体群の列を表現したものである。日本語では、物体全体からより小さな物体群へ参照範囲を絞り込みながら対象物体を特定する。SOG はこの絞り込みの過程を抽象化したものである。SOG 中の物体群の順序は主要部後置型である日本語における群の表現順序に対応する。

(船越他 2006)で想定している状況は同形同色同大の物体を複数配置した2次元空間であったため、物体群の間関係は空間的関係のみであった。したがって群間関係を明示する必要はなく、SOG は以下のように定式化されている。

$$\text{SOG} : [G_0 G_1 \dots G_n]$$

ここで  $G_x$  は物体群であり、 $G_0$  は物体全体の群、 $G_n$  は対象物体のみを含む群である。また群間関係は空間的な絞り込みの関係(内部参照関係)のみを利用している。

本論文では色、形、大きさ等の属性情報も利用した参照表現の生成を目的とするため物体群間関係が多様化する。そこで我々は SOG の群間に関係を挿入する。拡張した SOG は以下のように定式化できる。

$$\text{SOG} : [G_0 R_0 G_1 R_1 \dots G_i R_i \dots R_{n-1} G_n]$$

ここで、 $R_x$  は群間関係を示している。以後、断りなしに SOG という場合は、この拡張した SOG を指す。

## 2.1 群間の関係

$R_i$  は群  $G_i$  と群  $G_{i+1}$  を結ぶ関係を表す．関係には内部参照関係と外部参照関係の 2 種類がある．

- 内部参照関係

群  $G_i$  から  $G_{i+1}$  への絞り込みの関係であり,  $G_i \supset G_{i+1}$  となる．内部参照関係は, 絞り込みに利用する素性の種類に応じて下位範疇に細分類できる．これらの下位範疇を以下の記号で表す．

$\xrightarrow{type}$  : 物体のタイプ  
 $\xrightarrow{space}$  : 位置関係  
 $\xrightarrow{shape}$  : 物体の形  
 $\xrightarrow{color}$  : 物体の色  
 $\xrightarrow{size}$  : 物体の大きさ

- 外部参照関係

$G_i \cap G_{i+1} = \phi$  となり, この関係は空間的な関係に限られる．外部参照関係は記号  $\Rightarrow$  によって表す．

以上より,  $R_i$  は以下のように定式化できる．

$$R_i \in \{ \xrightarrow{space}, \xrightarrow{type}, \xrightarrow{shape}, \xrightarrow{color}, \xrightarrow{size}, \Rightarrow \}$$

図 2 に示す状況で, 物体  $b_4$  を指示する参照表現とその参照表現に対応する SOG の例を示す．全体群は  $\{all\}$  と略記する．下線は言語表現と群・関係の間の対応関係を表す．

参照表現: 「手前の<sub>(1)</sub>机の<sub>(2)</sub>上の<sub>(3)</sub>黒い<sub>(4)</sub>玉<sub>(5)</sub>」  
 SOG:  $[\{all\} \xrightarrow{type} \{t1, t2, t3\} \xrightarrow{space}_{(1)} \{t2\}_{(2)} \Rightarrow_{(3)} \{b3, b4\} \xrightarrow{color}_{(4)} \{b4\}_{(5)}]$

## 3 参照表現の生成

本論文が提案する参照表現の生成アルゴリズムは以下の 4 ステップから成る．

- Step 1 知覚的群化
- Step 2 SOG の生成
- Step 3 言語表現の付与
- Step 4 順位付け

以降, これらのステップをそれぞれ説明する．また, 例として図 2 中の状況を利用する．提案手法は, 物体の重なりを許すことで「~の上の~」という位置関係にも対応する．

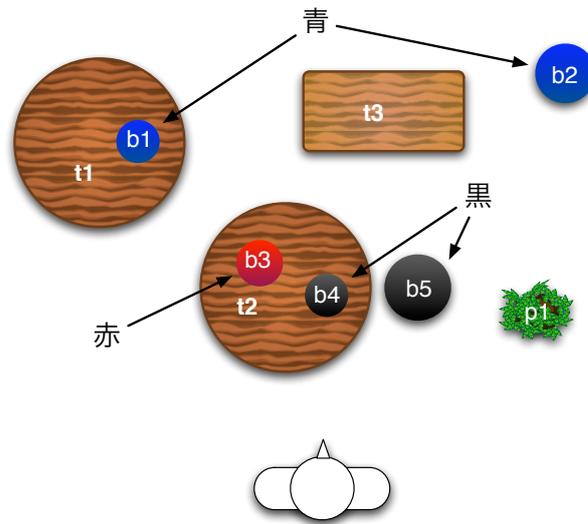


図 2: 参照表現を生成する状況の例

### Step 1: 知覚的群化

知覚的群化は以下の 5 つの素性に対して行う。

- (1) 物体のタイプ
- (2) 物体の形
- (3) 物体の色
- (4) 物体の大きさ
- (5) 物体の位置

ただし、「(5) 物体の位置」に関しては以下の 2 種類の知覚的群化を行なう。

- (5.1) 物体間の近接性：近接した物体を群化する。
- (5.2) 閉包：他の物体または特定の領域に囲まれている物体を群化する。

図 2 の例の場合、それぞれの机の上に乗っている玉を群化する。

物体間の近接性 (5.1) については、(Thórisson 1994) の手法を用いて知覚的群化を行なう。同手法は色や大きさや形などの近接性以外の素性にも素性毎の「距離」を定義することによって適用可能であるが、本論文ではこれらの素性については物体毎に予め定めたカテゴリに分類した。タイプには「机、植木、玉」、形には「四角、丸」、色には「赤、青、緑、黒」、大きさには「大、中、小」のカテゴリを用意し、Step 3 で付与する言語表現もカテゴリ毎に用意した。

知覚的群化に際しては「(1) 物体のタイプ」を特に重要視する。なぜなら、人は一般的に異なるタイプの物体を同一の物体群として捕らえることは少なく、また「タイプ」という素性はその物体を最も単純に表現できるものだからである。そこで本手法では、まず物体のタイプを

利用した群化を行ない, それぞれのタイプの物体群に対して (2) ~ (5) の素性を利用した群化を行なう.

Thórisson は下の 3 通りの素性の組み合わせを知覚的群化の異なる方略として認めている.

- 形と近接性
- 色と近接性
- 大きさと近接性

つまり, Thórisson の知覚的群化の手法では, 形, 色, 大きさが類似しているだけでなく, 各物体が相互に近接している場合のみそれらを群化する. しかし, 視覚情報から物体の群化を行なうだけの場合, この方略は有効だが, 物体群を利用して参照表現を生成しようとする場合 (例えば図 2 の状況で「青い 2 つの玉のうち」の様に青い 2 つの玉を群化した表現を生成したい場合), 距離の離れた物体同士を色や形, 大きさ等の素性から 1 つの物体群として扱うことが必要となる. よって本手法では, 素性ごとに単独で知覚的群化を行なう.

生成した各物体群には, 知覚的群化の際に利用した素性に応じてラベルを付与する. このラベルは, 次のステップの SOG 生成において使用する. 本手法が対象とする状況では,  $\{type, shape, color, size, space\}$  の 5 つのラベルを定義する. また特別な群として, 全体群と各物体単体の群も生成する. 単体群に対しては *space* ラベルを与える. 全体群にはラベルは必要ない. 下の例が示すように, 一つの群は複数のラベルを持ちうる.

図 2 の状況に対して知覚的群化を行なった結果生成される物体群を以下に示す.

素性	ラベル	認識された群
全体群	なし	$\{t1, t2, t3, p1, b1, b2, b3, b4, b5\}$
単体群	<i>space</i>	$\{t1\}, \{t2\}, \{t3\}, \{p1\}, \{b1\}, \{b2\}, \{b3\}, \{b4\}, \{b5\}$
タイプ	<i>type</i>	$\{t1, t2, t3\}, \{p1\}, \{b1, b2, b3, b4, b5\}$
形	<i>shape</i>	$\{t1, t2\}, \{t3\}$
色	<i>color</i>	$\{b1, b2\}, \{b3\}, \{b4, b5\}$
大きさ	<i>size</i>	$\{b1, b3, b4\}, \{b2, b5\}$
近接性	<i>space</i>	$\{t2, t3\}, \{b1, b3, b4, b5\}, \{b3, b4, b5\}$
閉包	<i>space</i>	$\{b1\}, \{b3, b4\}$

## Step 2: SOG の生成

Step 1 で生成した群をもとに, SOG を生成する. SOG 生成は, 自然言語生成のいわゆるコンテンツ・プランニングの段階に相当する. 生成アルゴリズムを図 3 ~ 図 4 に示す.

図 3 に定義した 3 つの変数, *Target*, *AllGroups*, *SOGList* は大域変数である. *Target* は参照表現によって指示する対象物, *AllGroups* は Step 1 で生成されたすべての群の集合, *SOGList*

は生成された SOG のリストである。Target と AllGroups が与えられると、関数 makeSOG はすべての可能な SOG を深さ優先で生成し、SOGList に追加する。

**makeSOG(図 3)** makeSOG は最初に SOG の第 1 要素として全体群を追加する。次に空間内に存在する物体のタイプの中から対象物体より顕現性が高いか等しいタイプを選択し、SOG を拡張する。顕現性の高い物体を優先的に手がかりとして使うことにより、聞き手の理解が容易になると期待できる。ここでは、大きさに着目して、机、植木、ボールの順に顕現性が高いと仮定する。これを繰り返し、選択する物体のタイプがなくなったら終了する。

**search(図 4)** この関数は生成途中の SOG を引数とする。SOG の最後の要素を LastGroup とし、以下の場合に分けて処理を行なう。

1. LastGroup が対象物体のみからなる単体群の場合 (05-06 行)  
この場合 SOG は完成しているので、SOGList に SOG を追加して終了。
2. LastGroup が対象物体以外の単体群からなる場合 (08-14 行)  
この場合は、LastGroup の単体群と外部参照関係によって関係付けられる対象物体を含む群を探し、この群によって SOG を拡張する。まず、対象物体を含み、LastGroup に含まれる物体 (参照物体) から適切な位置関係にある物体群を探し GroupList に代入する。適切な位置関係とは、以下の条件 (a), (b), (c) を満たすものである。ここで実装上の効率化のため、条件 (a) に該当する方向を記録しておく。
  - (a) 対象物体を含む群の全要素が参照物体から「奥、右奥、右、右手前、手前、左手前、左、左奥、上」<sup>1</sup>のいずれかの表現で表せる同一の方向にある。
  - (b) (a) と同一の方向で、なおかつ対象物体を含む群のいずれの要素よりも参照物体から近い位置に対象物体と同じタイプの他の物体が存在しない。
  - (c) 対象物体を含む群はいずれかのタイプの全体群ではない。
 次に外部参照関係 ⇒ と GroupList 中のそれぞれの物体群を追加した SOG を生成し、関数 search(図 4) を再帰的に呼び出す。GroupList 中の物体群を全て適用したら終了。
3. LastGroup が対象物体を含む複数の物体を含む場合 (17-26 行)  
この場合は、新しい素性 (内部参照関係) を使って LastGroup をさらに絞り込む。AllGroups 中の各物体群と LastGroup との積集合をとった群 NewG を生成し、付与されているラベルもコピーする。NewG が対象物体を含むならば、その時点の SOG と NewG を引数として関数 extend(図 5) を呼び出す。ただし、NewG が重複しないように GroupList を利用してチェックする。AllGroups 中の物体群を全て適用したら終了。
4. LastGroup が対象物体以外の複数の物体を含む場合 (28-32 行)  
この場合は、当面の目標として対象物体以外の単体群を作る。AllGroups の中から、LastGroup に包含される物体群 Group を選択し、それぞれの Group に対してその時点の

<sup>1</sup> 参照物体の重心を原点とし、空間内の人物が向いている方向を「奥」の方向として方向を 8 等分し、この順に方向を割り当てる。

SOG と Group を引数として関数 extend(図 5) を呼び出す。LastGroup に包含される物体群 Group を全て適用したら終了。

extend(図 5) この関数は生成途中の SOG (SOG) と次に追加する群 (Group) を引数とする。Group に付与されているラベルのリストを LabelList に取り出す (ただし *type* ラベルは除外する)。それぞれのラベルから SOG の末尾の物体群と Group を結び関係、および Group を SOG のコピー SOGcopy に追加する。そして関数 search(図 4) を SOGcopy に対して呼び出す。

図 2 の状況において対象物体を  $b_1$  としたときに生成される SOG を以下に示す。全体群は  $\{all\}$  と略記する。

1.  $[\{all\} \xrightarrow{type} \{t_1, t_2, t_3\} \xrightarrow{space} \{t_1\} \Rightarrow \{b_1\}]$
2.  $[\{all\} \xrightarrow{type} \{t_1, t_2, t_3\} \xrightarrow{shape} \{t_1, t_2\} \xrightarrow{space} \{t_1\} \Rightarrow \{b_1\}]$
3.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{space} \{b_1\}]$
4.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{color} \{b_1, b_2\} \xrightarrow{space} \{b_1\}]$
5.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{color} \{b_1, b_2\} \xrightarrow{size} \{b_1\}]$
6.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{size} \{b_1, b_4, b_3\} \xrightarrow{space} \{b_1\}]$
7.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{size} \{b_1, b_4, b_3\} \xrightarrow{color} \{b_1\}]$
8.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{space} \{b_1, b_3, b_4, b_5\} \xrightarrow{space} \{b_1\}]$
9.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{space} \{b_1, b_3, b_4, b_5\} \xrightarrow{color} \{b_1\}]$
10.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{space} \{b_1, b_3, b_4, b_5\} \xrightarrow{size} \{b_1, b_4, b_3\} \xrightarrow{space} \{b_1\}]$
11.  $[\{all\} \xrightarrow{type} \{b_1, b_2, b_3, b_4, b_5\} \xrightarrow{space} \{b_3, b_4, b_5, b_1\} \xrightarrow{size} \{b_1, b_4, b_3\} \xrightarrow{color} \{b_1\}]$

### Step 3: 言語表現の付与

SOG の各要素に表現を付与することで、参照表現を生成する。表現の付与には以下の規則を用いる。規則 1 は物体群の表現を、規則 2 は群間の関係の表現を生成する規則である。それぞれの規則の中では、各下位規則は表記順に優先度が高いものとする。

[規則 1]: 物体群に対する表現付与

規則 1.1 全体群 ( $\{all\}$ ) は言語化しない

(船越他 2006) で述べられているとおり、全体群が明示的に言語化されることがまれであるという被験者実験の結果を参考にした。

規則 1.2 各単体群には「タイプ名」もしくは「タイプ名 + 『の』」を付与

単体群が SOG の最後の要素でない場合は、外部参照関係 ( $\Rightarrow$ ) が後続するため「の」を付与する。規則 2.1 によりタイプの絞り込みは言語化せず、単体群に対して「タイプ名」を付与することで必要な情報を言語化する。

```

Target      # 対象物体
AllGroups   # 知覚的群化によって生成された全ての群のリスト
SOGList     # SOG のリスト

01:makeSOG()
02: SOG = [ ];                                # 物体群と関係を要素とするリスト
03: All = getAll();                            # 物体の全体群を得る
04: add(All, SOG);                             # SOG に追加
05: TypeList = getAllTypes(All);              # 物体のタイプの集合を得る
06: TypeOfTarget = getType(Target);           # 対象物体のタイプを得る
07: TargetSaliency = saliency(TypeOfTarget);  # 対象物体のタイプの顕現性
08: for each Type in TypeList do              # Type∈{Table, Plant, Ball}
09:   if saliency(Type)≥TargetSaliency then  # 顕現性を比較 (Table>Plant>Ball を仮定)
10:     Group = get(AllGroups, Type);          # そのタイプの全体群を得る
11:     SOGcopy = copy(SOG);                  # SOG のコピーを作る
12:     add(type, SOGcopy);                    # type の関係を末尾に追加
13:     add(Group, SOGcopy);                  # Group を末尾に追加
14:     search(SOGcopy);
15:   end if
16: end for
17:return
    
```

図 3: makeSOG

規則 1.3 各タイプの全体群は言語化しない

理由は規則 1.1 に同じ。

規則 1.4 後続が内部参照関係  $\xrightarrow{space}$  の場合、「個数 + タイプ名 + 『のうち』」を付与、その他の関係が後続する場合は言語化しない

後続する関係が空間的な関係 ( $\xrightarrow{space}$ ) 以外の場合、その関係は名詞 (タイプ名) の前の関係の列として言語化できる。例えば、「大きい玉のうちの赤い玉」は「大きい赤い玉」と表現できる。外部参照関係については、単体群以外には後続しないので規則 1.2 で全て処理される。

[規則 2]: 群間の関係に対する表現付与

規則 2.1 タイプの関係 ( $\xrightarrow{type}$ ) は言語化しない

規則 1.2 を参照。

規則 2.2 各関係に関係を示す表現を付与

$\xrightarrow{shape}$ ,  $\xrightarrow{color}$ ,  $\xrightarrow{size}$  にはそれぞれの属性の値の表現を付与する。⇒,  $\xrightarrow{space}$  に対する表現の付与は以下の場合分けに従って行なう。ここで、 $|G_i|$  は  $G_i$  の要素数を表す。

内部参照関係 ( $G_i \xrightarrow{space} G_{i+1}$ )

- $|G_i| = 2$  のとき

絞り込みの前後で要素数は必ず少なくなるため、 $|G_{i+1}| = 1$  である。各物体の座標

```

01:search(SOG)
02: LastGroup = getLastElement(SOG);      # SOG の最後に位置する物体群を得る
03: Card = getCardinality(LastGroup);     # 物体群の要素数を得る
04: if Card == 1 then
05:   if containsTarget(LastGroup) then   # 群が対象物体を含んでいるかを調べる
06:     add(SOG, SOGList);
07:   else
08:     GroupList = searchTargetGroups(LastGroup);
           # LastGroup から適切な位置にあり, 対象物体を含んでいる物体群の集合を得る
09:     for each Group in GroupList do
10:       SOGcopy = copy(SOG);
11:       add( $\Rightarrow$ , SOGcopy);
12:       add(Group, SOGcopy);
13:       search(SOGcopy);
14:     end for
15:   end if
16: elseif containsTarget(LastGroup) then
17:   Checked = [ ];
18:   for each Group in AllGroups do
19:     NewG = Intersect(Group, LastGroup); # 交差をとった群を生成
20:     Labels = getLabels(Group);
21:     setLabels(Labels, NewG);           # ラベルを Group から NewG へコピー
22:     if containsTarget(NewG) & !contains(Checked, NewG) then
23:       add(NewG, Checked);
24:       extend(SOG, Group);
25:     end if
26:   end for
27: else
28:   for each Group of AllGroups do
29:     if contains(LastGroup, Group) then
30:       extend(SOG, Group);
31:     end if
32:   end for
33: end if
34:return

```

図 4: search

```

01:extend(SOG, Group)
02: LabelList = getLabels(Group);      # Group に付与されている属性ラベルのリストを得る
03: for each Label in LabelList do     # Label ∈ {space, shape, color, size}
04:   SOGcopy = copy(SOG);             # SOG のコピーを作る
05:   add( $\xrightarrow{\text{Label}}$ , SOGcopy);      # 関係を末尾に追加
06:   add(Group, SOGcopy);             # Group を末尾に追加
07:   search(SOGcopy);
08: end for
09:return

```

図 5: extend

から、4つの表現「[右/左/手前/奥]の」のいずれかを付与する。

- $|G_i| \geq 3$  かつ  $|G_{i+1}| = 1$  のとき  
各物体の座標を参考にし、以下の表現の妥当性をこの順に調べ、妥当な表現を付与する。
  - (1) 「一番[手前/奥/右/左/右手前/左手前/右奥/左奥]の」
  - (2) 「真ん中の」
  - (3) 「[左/右/手前/奥]から  $j$  番目の」
- $|G_{i+1}| \geq 2$  のとき  
各物体群の座標を参考にし、「[右/左/手前/奥/真ん中/右手前/左手前/右奥/左奥]の」の中から妥当な表現を付与する。

外部参照関係 ( $G_i \Rightarrow G_{i+1}$ )

- search の性質からこの場合は  $|G_i| = 1$  となる。ここでは Step 2 の SOG 生成のときに記録した方向に対応する表現（「[奥/右奥/右/右手前/手前/左手前/左/左奥/上]の」）を付与する。

図 2 の状況において Step 2 で生成した SOG に表現を付与すると以下ようになる。

1.  $[\{all\} \xrightarrow{type} \{t1, t2, t3\} \xrightarrow{space}_{(1)} \{t1\}_{(2)} \Rightarrow_{(3)} \{b1\}_{(4)}]$   
「一番左の<sub>(1)</sub> 机の<sub>(2)</sub> 上の<sub>(3)</sub> 玉<sub>(4)</sub>」
2.  $[\{all\} \xrightarrow{type} \{t1, t2, t3\} \xrightarrow{shape}_{(1)} \{t1, t2\}_{(2)} \xrightarrow{space}_{(3)} \{t1\}_{(4)} \Rightarrow_{(5)} \{b1\}_{(6)}]$   
「丸い<sub>(1)</sub> 2つの机のうち<sub>(2)</sub> 左の<sub>(3)</sub> 机の<sub>(4)</sub> 上の<sub>(5)</sub> 玉<sub>(6)</sub>」
3.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{space}_{(1)} \{b1\}_{(2)}]$   
「一番左の<sub>(1)</sub> 玉<sub>(2)</sub>」
4.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{color}_{(1)} \{b1, b2\}_{(2)} \xrightarrow{space}_{(3)} \{b1\}_{(4)}]$   
「青い<sub>(1)</sub> 2つの玉のうち<sub>(2)</sub> 左の<sub>(3)</sub> 玉<sub>(4)</sub>」
5.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{color}_{(1)} \{b1, b2\} \xrightarrow{size}_{(2)} \{b1\}_{(3)}]$   
「青い<sub>(1)</sub> 小さい<sub>(2)</sub> 玉<sub>(3)</sub>」
6.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{size}_{(1)} \{b1, b3, b4\}_{(2)} \xrightarrow{space}_{(3)} \{b1\}_{(4)}]$   
「小さい<sub>(1)</sub> 3つの玉のうち<sub>(2)</sub> 一番左の<sub>(3)</sub> 玉<sub>(4)</sub>」
7.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{size}_{(1)} \{b1, b3, b4\} \xrightarrow{color}_{(2)} \{b1\}_{(3)}]$   
「小さい<sub>(1)</sub> 青い<sub>(2)</sub> 玉<sub>(3)</sub>」
8.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{space}_{(1)} \{b1, b3, b4, b5\}_{(2)} \xrightarrow{space}_{(3)} \{b1\}_{(4)}]$   
「左の<sub>(1)</sub> 4つの玉のうち<sub>(2)</sub> 一番左の<sub>(3)</sub> 玉<sub>(4)</sub>」
9.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{space}_{(1)} \{b1, b3, b4, b5\} \xrightarrow{color}_{(2)} \{b1\}_{(3)}]$   
「左の<sub>(1)</sub> 青い<sub>(2)</sub> 玉<sub>(3)</sub>」

10.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{space}_{(1)} \{b1, b3, b4, b5\} \xrightarrow{size}_{(2)} \{b1, b3, b4\}_{(3)} \xrightarrow{space}_{(4)} \{b1\}_{(5)}]$   
 「左の<sub>(1)</sub> 小さい<sub>(2)</sub> 3つの玉のうち<sub>(3)</sub> 一番左の<sub>(4)</sub> 玉<sub>(5)</sub>」
11.  $[\{all\} \xrightarrow{type} \{b1, b2, b3, b4, b5\} \xrightarrow{space}_{(1)} \{b1, b3, b4, b5\} \xrightarrow{size}_{(2)} \{b1, b3, b4\} \xrightarrow{color}_{(3)} \{b1\}_{(4)}]$   
 「左の<sub>(1)</sub> 小さい<sub>(2)</sub> 青い<sub>(3)</sub> 玉<sub>(4)</sub>」

#### Step 4: 順位付け

出力表現を決定するために、表現中で使用された関係と表現の長さを考慮して、各表現にスコアを与える。

最初に SOG 内の各関係に対して  $[0, 1]$  の範囲でコストを与える。関係のコストは以下のよう  
 に決定する。これらのコストは (Dale and Reiter 1995) で述べられている素性の優先順位に  
 従う。

---

$\xrightarrow{type}$	: (無視)
$\xrightarrow{shape}$	: 0.2
$\xrightarrow{color}$	: 0.4
$\xrightarrow{size}$	: 「大きい」: 0.6, 「小さい」: 0.8, 「中くらいの」: 1.0
$\xrightarrow{space}$ , $\Rightarrow$	: コスト関数を (Tokunaga, Koyama and Saito 2005) で提案されたポ テンシャル関数に従って定義した。ただし、関係「~の上にある ~」のコストは 0 とする。

---

次に、関係のコスト  $cost\_rel$  を各関係のコストの平均値として求める。そして表現長のコス  
 ト  $cost\_len$  を以下の式で求める。表現長は文字数で測る。

$$cost\_len = \frac{\text{length}(\text{expression})}{\max_i \text{length}(\text{expression}_i)}$$

コスト  $cost\_rel$  及び  $cost\_len$  を用い、表現のスコア  $score$  を以下のように定める。

$$score = \frac{1}{\alpha \times cost\_rel + (1 - \alpha) \times cost\_len}$$

$\alpha$  の値は、次節の実験では 0.5 に固定した。

図 2 の状況において生成した表現にコストを付与し、スコアの高い順に示す。各行の表現の  
 右の数値はその表現のスコア ( $score, \alpha = 0.5$ )、括弧内の数値は左から関係のコスト ( $cost\_rel$ )、  
 表現長のコスト ( $cost\_len$ ) である。

1. 「一番左の玉」	3.66 (0.251, 0.294)
2. 「左の青い玉」	2.62 (0.468, 0.294)
3. 「一番左の机の上の玉」	2.44 (0.289, 0.529)
4. 「青い2つの玉のうち左の玉」	2.20 (0.204, 0.706)
5. 「青い小さい玉」	2.10 (0.600, 0.353)
6. 「小さい青い玉」	2.10 (0.600, 0.353)
7. 「左の小さい青い玉」	1.90 (0.578, 0.471)
8. 「丸い2つの机のうち左の机の上の玉」	1.67 (0.259, 0.941)
9. 「左の4つの玉のうち一番左の玉」	1.64 (0.393, 0.824)
10. 「小さい3つの玉のうち一番左の玉」	1.42 (0.526, 0.882)
11. 「左の小さい3つの玉のうち一番左の玉」	1.31 (0.529, 1.000)

## 4 評価と考察

提案手法を評価するため、上記のアルゴリズムを Java で実装し、大学生 18 人を対象に心理実験を行なった。

### 4.1 実験

実装システムが生成した参照表現を評価する被験者実験を実験 1 と実験 2 に分けて行なった。被験者は実験 1 の後に実験 2 を行なったが、それらの関連については被験者には知らせていない。

実験に使用した 20 布置と、各布置に対して実装システムが生成した上位 5 つの参照表現を付録に示す。布置は物体の個数も位置もランダムに決定したもので、対象物体もランダムに選んだ。ただし布置は、参照表現が 5 つ以上生成され、かつ対象物体と同じタイプの物体が 2 つ以上存在するものに限定した。これは「机」のような 1 単語の表現では対象物体を特定できないようにするためである。

**実験 1** この実験は、システムによって生成した参照表現を人間が解釈した時に、どの程度正確に対象物体を同定できるかを評価するために行なった。被験者に、布置とその布置中の対象物体 1 つに対して生成した参照表現のうち最もスコアが高かったものを示し、その参照表現が指し示す物体を選ばせた。図 6 は実験に使用した視覚刺激の例である。

**実験 2** この実験はシステムに実装したスコア付けがどの程度人間の直感に合致しているかを評価するために行なった。対象物体を明示した布置の画像と、その布置中の対象物体 1 つに対して生成した参照表現をスコアが高い順に 5 つ被験者に示し、その中で被験者が最もよいと思う表現を選ばせた。判断基準は明確に設定せず各自の判断に任せた。図 7 は実験に使用した視

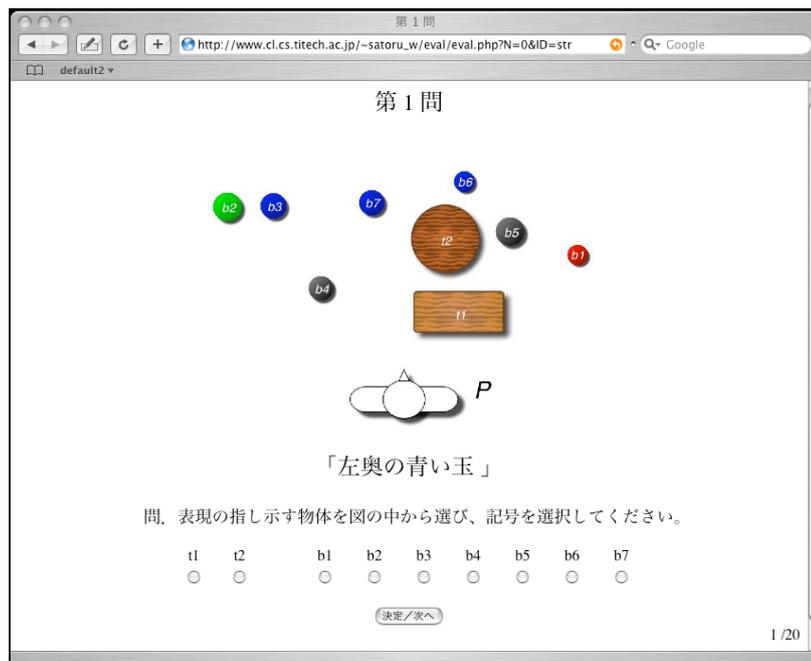


図 6: 実験 1 に使用した視覚刺激の例

覚刺激の例である．布置は課題 1 と同じものの 20 種を使用した．被験者に示した 5 つの参照表現は実験 1 で評価した参照表現も含む．

## 4.2 結果

表 1 に実験 1 の結果を示す．対象物体特定の正解率は全体で 95.0%であった．このことから，実装システムが生成した参照表現は高い対象物体特定能力をもつと言える．他の布置に比べ，布置 20 (付録参照) は正解率が低い，これは位置の計算を直交座標系に準じて行なったためと考えられる．布置 20 中の対象物体に対して実装システムは，対象物体が垂直方向において人物  $P$  に最も近い，「一番手前の玉」という表現を生成している．この表現に対してほとんどの被験者は図中の一番右側の玉を選択した．このことから「一番手前」という表現では，人物  $P$  と物体間のユークリッド距離が最も重要な要素となっていると考えられる．

表 2 に実験 2 の結果を示す．表現 No.1～5 の順にスコアが高かった参照表現を表している．上位 2 位の参照表現の得票率が全体の 72%を占めていることから，本論文で提案した参照表現に対する順位付けの手法が有効であるといえる．



図 7: 実験 2 に使用した視覚刺激の例

表 1: 実験結果: 正解率

布置	1	2	3	4	5	6	7	8	9	10	
正解率	0.89	1	1	1	1	1	1	0.94	1	1	
布置	11	12	13	14	15	16	17	18	19	20	平均
正解率	1	0.94	1	1	1	1	1	1	1	0.17	0.95

表 2: 実験結果: 得票率

表現 No.	1	2	3	4	5	計
度数	134	125	59	22	20	360
得票率	0.37	0.35	0.16	0.06	0.06	1

### 4.3 生成能力の限界

提案手法で生成可能な参照表現には原理的な限界がある．その代表的な 3 つを以下に示す．

並列表現 SOG は直列的に焦点を遷移していくため，例えば「机の前で，木の左の玉」という様な並列的に他の物体を参照する表現は生成できない．また「～と～の間の」という関係を利

用する表現も生成できない．この様な並列表現を生成するためには SOG の更なる拡張が必要となる．

多重に他の物体を参照する表現 例えば「四角い机の左の木の奥にある玉」という表現では「机 木 玉」という順に 3 タイプの物体を参照するが，提案手法では対象物体と異なるタイプの物体を 2 つ以上参照するこのような表現は生成できない．しかしこの制限は，参照表現の簡潔さの観点から意図的に設けたものであり，Step 2 の SOG 生成を拡張すれば多重に他の物体を参照する表現は生成できる．

複数の他の物体を参照する表現 外部参照関係における焦点の遷移には，単体から群への遷移のみを定義している．そのため提案手法では，例えば図 2 の状況において物体  $b_1$  に対して，

参照表現: 「丸い机の上の青い玉」

SOG:  $\{[all]\} \xrightarrow{type} \{t_1, t_2, t_3\} \xrightarrow{shape} \{t_1, t_2\} \Rightarrow \{b_1, b_3, b_4\} \xrightarrow{color} \{b_1\}$

の様に，SOG 中で複数の机 ( $\{t_1, t_2\}$ ) を外部参照する表現を生成できない．これを解決するためには，SOG 生成手法と表現付与の規則を工夫する必要があり，今後の課題である．

#### 4.4 その他の課題

位置計算 現在のシステムでは，妥当な位置表現を選択するために，直交座標系を 8 方向に分割して位置を計算するという単純な手法を使っている．適切な位置表現を付与するための座標系や参照枠・視点の選択方法について，今後心理実験などをもとに解明する必要がある．

簡潔性と曖昧性 SOG への表現付与の規則 1.4 は，物体群の言語化を省略することで表現の簡潔さと自然さを得ることを目的としている．例えば「黒い 3 つの玉のうち小さい玉」という表現は規則 1.4 より「黒い小さい玉」という簡潔で人間がより自然と感じる表現となる．しかし，「丸い 2 つの机のうち右の机」という表現は，一般的に「丸い右の机」とした方が自然であるが，この場合「2 つの机のうち」という部分表現が表す物体群の後には「右の」という部分表現が表す群間の空間的内部参照関係  $\xrightarrow{space}$  が続くため規則 1.4 による省略は行なわれない．この例から規則 1.4 が一見不十分だと感じるかもしれないが，空間的内部参照関係の前の物体群を省略してしまうと対象物体特定に曖昧性を生じる場合がある．例えば図 8 の布置で，物体  $b_7$  を指し示す「青い 3 つの玉のうち左から 2 番目の玉」という表現を省略して「青い左から 2 番目の玉」とすると「青い玉」で「左から 2 番目の玉」である物体  $b_3$  との間に曖昧性が生じる．このように，対象物体を特定する情報を適切に言語化する方法にも，さらに改良の余地がある．

人間が生成する表現との比較 提案手法のスコア付けの手法は Incremental Algorithm (Dale and Reiter 1995) を参考に人手で作成したものである．実験 2 の結果からスコアの高いものがよ

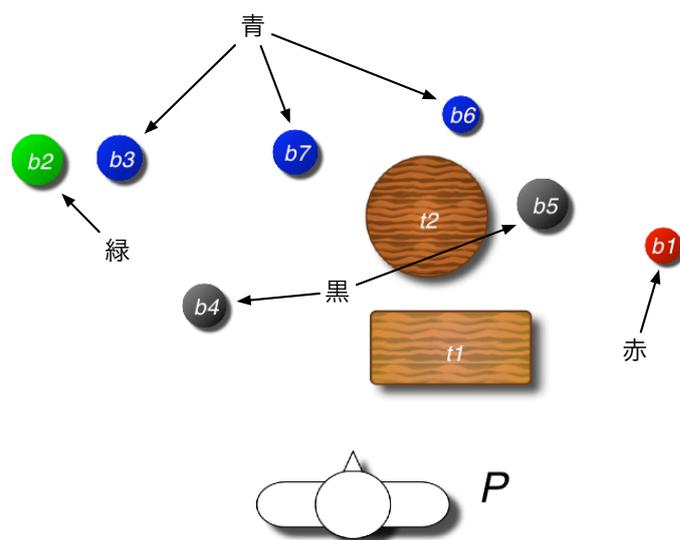


図 8: 表現を省略できない布置の例

りよい表現として選ばれていることより、提案手法のスコア付けの手法が有効であるといえる。しかしながら、この結果は、実装システムが生成した参照表現が人間の生成する参照表現にどこまで近づけたかを示すものではない。人間が生成する参照表現との比較は今後の課題である。

## 5 まとめ

本論文では知覚的群化を利用した参照表現の生成手法を提案した。参照表現と参照される空間の状況との間の中間表現である SOG (船越他 2006) を拡張し、空間的包含関係でない位置関係や物体の属性も扱えるようにした。

提案手法を実装したシステムが生成した参照表現の対象物体特定の精度は 95%であった。またスコア付けによる参照表現の順位付けは被験者による高い評価を得た。提案手法の生成能力には制限があり、ある種の表現は原理的に生成できない。しかし、提案手法は対象を特定する表現を生成する十分な能力を備えている。

## 参考文献

- Appelt, D. E. (1985). "Planning English Referring Expressions." *Artificial Intelligence*, **26**, 1-33.
- Byron, D. K. (2003). "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*.
- Dale, R. (1992). “Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes.” MIT Press, Cambridge.
- Dale, R. and Haddock, N. (1991). “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.
- Dale, R. and Reiter, E. (1995). “Computational interpretations of the Gricean maxims in the generation of referring expressions.” *Cognitive Science*, **19** (2), 233–263.
- Krahmer, E. and Theune, M. (2002). “Efficient context-sensitive generation of descriptions.” In Kees van Deemter and Rodger Kibble, editors, *Information Sharing: Givenness and Newness in Language Processing*. CSLI Publications, Stanford, California.
- Krahmer, E., van Erk, S., and Verleg, A. (2003). “Graph-Based Generation of Referring Expressions.” *Computational Linguistics*, **29** (1), 53–72.
- Tanaka, H., Tokunaga, T., and Shinyama, Y. (2004). “Animated Agents Capable of Understanding Natural Language and Performing Actions.” In Prendinger, H. and Ishizuka, M. (Eds.), *Life-Like Characters*, pp. 429–444. Springer.
- Thórisson, K. R. (1994). “Simulated Perceptual Grouping: An Application to Human-Computer Interaction.” In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 876–881.
- Tokunaga, T., Koyama, T., and Saito, S. (2005). “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.
- 船越孝太郎, 渡辺聖, 栗山直子, 徳永健伸 (2006). “知覚的群化に基づく参照表現の生成.” *自然言語処理*, **13** (2), 79–97.

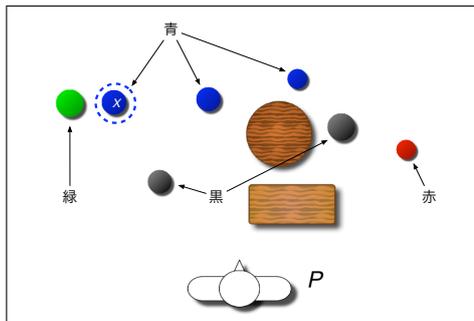
## 付録：実験に用いた布置と生成された参照表現

4 節で使用した布置と生成された参照表現をスコアの高い順に以下に示す。机の形や玉の色，物体の個数はランダムに決定した。物体の個数は以下の範囲に制限した。

- 机：0 ~ 3 個
- 木：0 ~ 2 個
- 玉：3 ~ 9 個

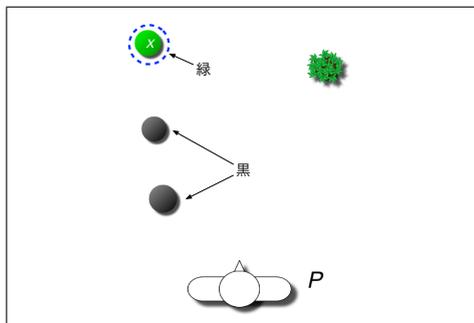
対象物体もランダムに選択し，図中では X を付けて，破線で囲んである。

布置 1



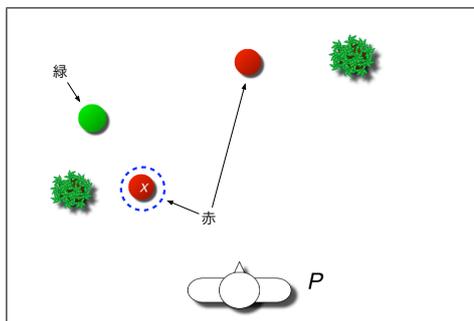
1. 左奥の青い玉
2. 左から 2 番目の玉
3. 青い 3 つの玉のうち一番左の玉
4. 左奥の小さい玉
5. 青い大きい 2 つの玉のうち左の玉

布置 2



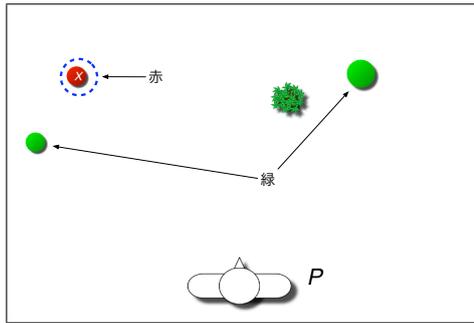
1. 一番奥の玉
2. 緑の玉
3. 小さい緑の玉
4. 木の左の玉
5. 小さい 2 つの玉のうち奥の玉

布置 3



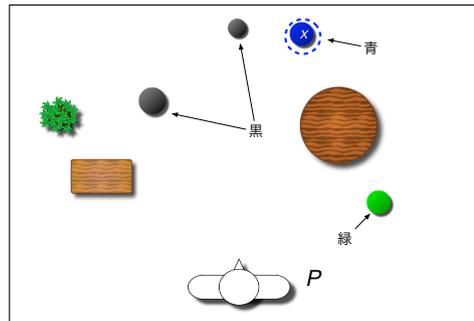
1. 左の木の右の玉
2. 一番手前の玉
3. 左の赤い玉
4. 左の木の右奥の赤い玉
5. 左の小さい玉

布置 4



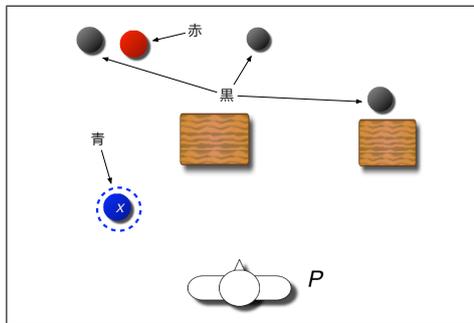
1. 左の赤い玉
2. 赤い玉
3. 木の左の赤い玉
4. 木の左の玉
5. 左の2つの玉のうち奥の玉

布置 5



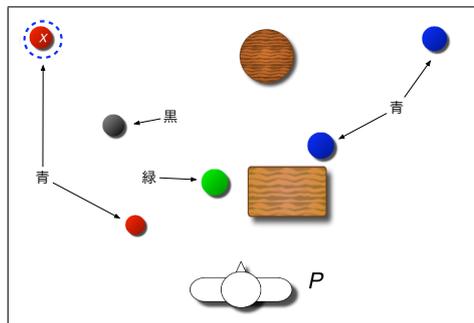
1. 青い玉
2. 奥の青い玉
3. 丸い机の奥の玉
4. 丸い机の左奥の青い玉
5. 右の机の奥の玉

布置 6



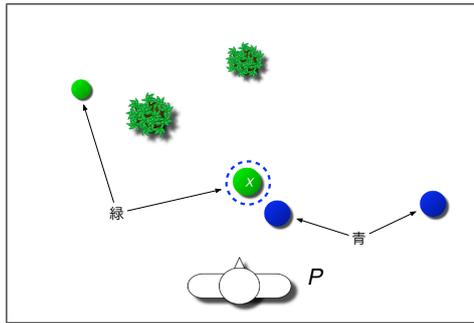
1. 一番手前の玉
2. 青い玉
3. 左の机の左の玉
4. 大きい青い玉
5. 大きい青い玉

布置 7



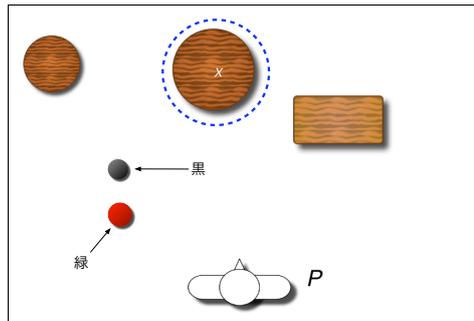
1. 一番左奥の玉
2. 赤い大きい玉
3. 赤い2つの玉のうち奥の玉
4. 左の赤い大きい玉
5. 左の赤い2つの玉のうち奥の玉

布置 8



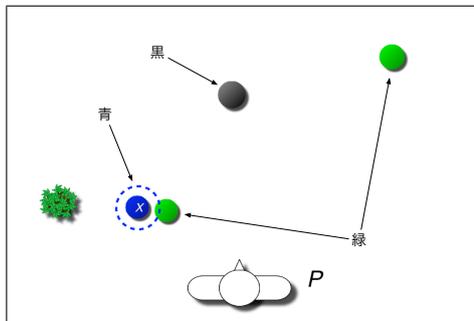
1. 右の緑の玉
2. 真ん中の緑の玉
3. 緑の大きい玉
4. 大きい緑の玉
5. 緑の2つの玉のうち右の玉

布置 9



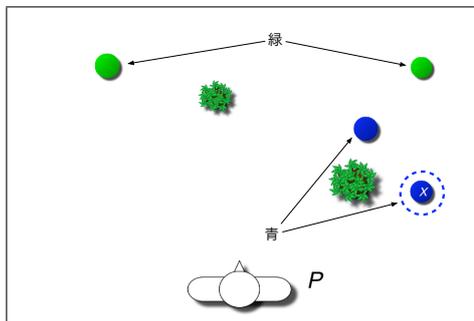
1. 真ん中の机
2. 丸い大きい机
3. 大きい丸い机
4. 丸い2つの机のうち右の机
5. 大きい2つの机のうち左の机

布置 10



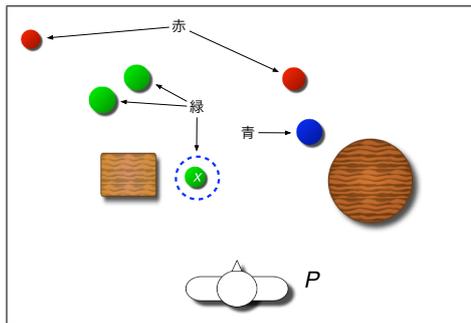
1. 青い玉
2. 手前の青い玉
3. 左の青い玉
4. 木の右の青い玉
5. 木の右の青い玉

布置 11



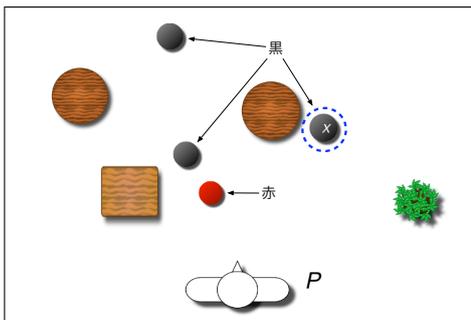
1. 右の木の右の玉
2. 一番手前の玉
3. 右の青い小さい玉
4. 右の小さい青い玉
5. 青い小さい玉

布置 12



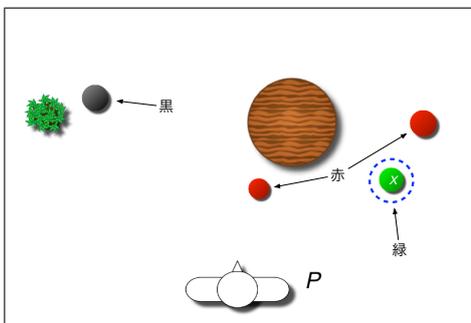
1. 左の机の右の玉
2. 四角い机の右の玉
3. 右から 3 番目の玉
4. 小さい机の右の玉
5. 緑の 3 つの玉のうち一番右手前の玉

布置 13



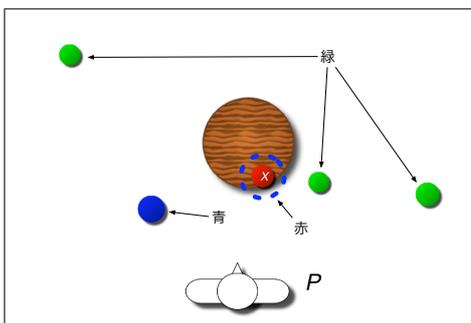
1. 一番右の玉
2. 一番右の机の右の玉
3. 木の左の玉
4. 黒い 3 つの玉のうち一番右の玉
5. 丸い 2 つの机のうち右の机の右の玉

布置 14



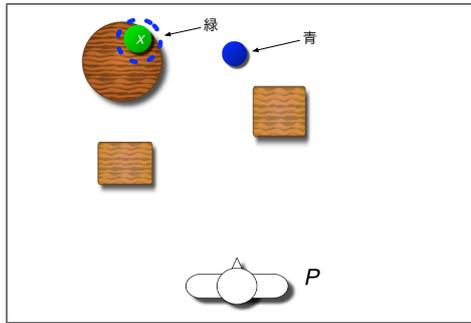
1. 机の右の緑の玉
2. 緑の玉
3. 右の緑の玉
4. 右の緑の玉
5. 大きい緑の玉

布置 15



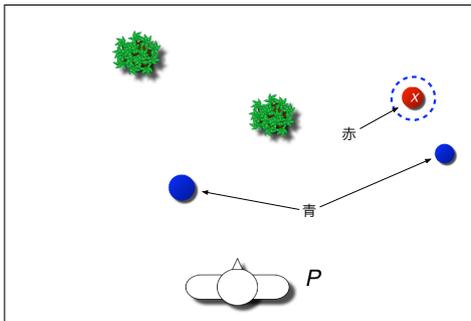
1. 机の上の玉
2. 赤い玉
3. 手前の赤い玉
4. 真ん中の玉
5. 真ん中の赤い玉

布置 16



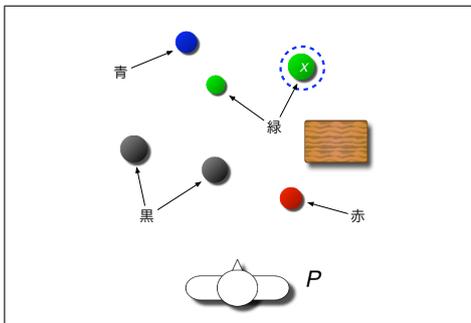
1. 左の玉
2. 丸い机の上の玉
3. 緑の玉
4. 左の丸い机の上の玉
5. 一番左奥の机の上の玉

布置 17



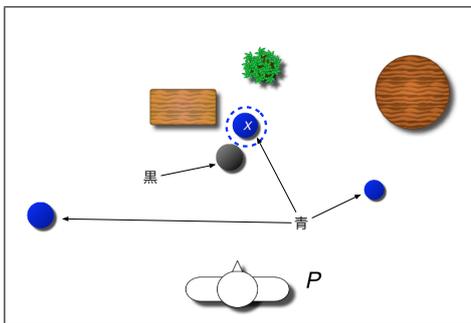
1. 右の赤い玉
2. 赤い玉
3. 一番奥の玉
4. 右の木の右の赤い玉
5. 右の木の右の玉

布置 18



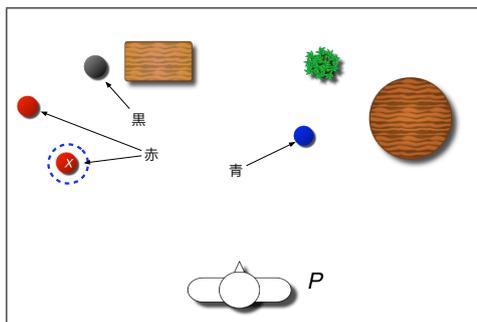
1. 机の奥の玉
2. 一番右奥の玉
3. 緑の大きい玉
4. 大きい緑の玉
5. 緑の2つの玉のうち右の玉

布置 19



1. 真ん中の青い玉
2. 左の机の右の玉
3. 四角い机の右の玉
4. 左の机の右手前の青い玉
5. 木の手前の青い玉

布置 20



1. 一番手前の玉
2. 赤い2つの玉のうち手前の玉
3. 左の3つの玉のうち一番手前の玉
4. 左の赤い2つの玉のうち手前の玉
5. 左の机の左手前の3つの玉のうち一番手前の玉