

参照表現の生成における属性選択

黒沢 崇浩

Philipp Spanger

徳永 健伸

東京工業大学

東京工業大学

東京工業大学

工学部 情報工学科

大学院情報理工学研究科

大学院情報理工学研究科

計算工学専攻

計算工学専攻

kurosawa.t.aa@m.titech.ac.jp

philipp@cl.cs.titech.ac.jp

take@cl.cs.titech.ac.jp

1 はじめに

参照表現とは、「赤いコップ」「右を向いた緑の椅子」などのように、特定の物体を他の物体と区別して指し示す表現である。

人間とコンピュータとの自然なコミュニケーションを目指す上で、参照表現の処理は重要な課題の一つである。この分野では人間が生成した参照表現の理解や、指示対象を指し示す参照表現の生成などが研究されているが、本研究では人間に近い自然な参照表現を生成することを目的とする。これまでの参照表現生成の研究では、物体の特徴を使って対象を特定するための最小な表現の生成 [3, 4]、空間的な配置からの自然な表現の生成 [1, 2] などが行われてきた。本研究では人間の属性選択の傾向を利用して、人間に近い参照表現の生成を試みる。

まず、人間の生成した参照表現のコーパスを用いて、人間が参照表現を生成する際にどのような属性の組合せを用いるかを分析した (4 節)。またその結果より、周辺の物体の状況から適切な属性を選択するアルゴリズムを作成した。最後にアルゴリズムが生成した属性の組を、属性選択に使用したコーパスを用いて評価した (5 節)。

2 コーパス

本研究で用いるコーパスは、Aberdeen 大学の TUNA project^{*1}において作成された TUNA コーパスを簡略化したものである。これは、TUNA project の一環である GRE Challenge [5]^{*2}で用いられたものである。

^{*1} <http://www.csd.abdn.ac.uk/research/tuna/>

^{*2} <http://www.csd.abdn.ac.uk/research/evaluation/>

2.1 コーパスの構成

コーパスは、平面に配置された複数の物体と、その中の 1 つを指示するために人間が生成した参照表現の組から構成されている。平面は 5 × 3 マスの二次元平面であり、その中には指示する目標物 (ターゲット) と他の物体 (ディストラクタ) が配置されている。この物体が配置された平面のそれぞれを**事例**と呼ぶ。各事例には家具のみが配置されている家具ドメインと、人間のみが配置されている人間ドメインがある。配置されている各物体にはそれぞれ表 1 のような属性が定義されている。

表 1: 各事例で定義されている属性

家具		人間	
属性名	属性値 (例)	属性名	属性値 (例)
type	chair	type	person
colour	red	haircolour	light
orientation	right	orientation	left
size	large	age	old
x-dimension	2	x-dimension	1
y-dimension	4	y-dimension	2
		has suit	true
		has shirt	false
		has tie	true
		has beard	false
		has hair	true
		has glasses	true

コーパスの一部の事例に、人間が生成した参照表現が対象物体を指し示していないなどの問題があった。これらについては人間の属性選択の規則の参考にならないと判断し、削除した。この結果、家具ドメインには 278、人間ドメインには 230 の事例がコーパスに残った。また、対象を一意に特定できない参照表現の事例もあったが、これらは人間が選ぶ規則の参考になると

判断し、使用した。これについての分析を4.5節で行っている。

2.2 コーパスの問題点

現在扱っているコーパスは、元の言語表現からなるTUNAコーパスを簡略化し、物体の持っている属性名と値の組のみを抽出したものである。そのため、元のコーパスより情報量が少なくなっている。例えば「左上にある椅子」と「机の左上にある椅子」という表現はどちらもx-dimension、y-dimensionの属性値のみで表現されており、区別できない。より深い分析のためには、x-dimension、y-dimensionなどではなく、元の言語表現が必要である。

3 基本アルゴリズムと評価方法

3.1 基本アルゴリズム

表1の各属性について、事例内にターゲットと異なる属性値を持っているディストラクタがいくつ存在するかを調べ、これを属性の弁別能力とする。弁別能力の高い順に属性を採用して、事例内でターゲットを一意に特定できた段階でその属性の組を出力とする。この方針で、以下の2つの基本アルゴリズムを作成した。

静的アルゴリズム 事例内で最初に1度だけ弁別能力を計算し、その順で属性を選択する。

動的アルゴリズム 属性を1つ選択する毎に、その時点で採用された属性の組で弁別できないディストラクタのみ用いて新たに弁別能力を計算する。

3.2 評価方法

2つの基本アルゴリズムを基にプログラムを作成し、プログラムの出力した属性の組を、人間の選択した属性の組とのdice係数を求めて評価した。dice係数は、人間とプログラムで共通した属性の数をC、プログラムの選択した属性の数をP、人間の選択した数をHとして、式(1)で求める。

$$dice = \frac{2C}{P+H} \quad (0 \leq dice \leq 1) \quad (1)$$

3.3 基本アルゴリズムの評価

表2がdice係数の平均である。基本アルゴリズムでは、対象をユニークに特定は出来るが、4.1節の選択率や4.2節の属性間の相性といった人間が自然に生成する参照表現に存在する傾向が無視されている。これについて、4節で分析し、5節で実装を行う。

表2: 基本アルゴリズムのdice係数

	家具	人間
静的	0.286	0.257
動的	0.308	0.257

4 参照表現の分析

4.1 属性の選択率

属性名の選択率への影響 利用されやすい属性と利用されにくい属性の存在を調べるため、式2で定義される選択率の属性名による違いを調べた。式中の「属性Aを含む事例数」とは、人間の生成した参照表現に属性Aが含まれていた事例数である。表3は家具ドメインの場合だが、属性によって選択率にばらつきがある事が分かる。

$$\text{属性Aの選択率} = \frac{\text{属性Aを含む事例数}}{\text{全事例数}} \quad (2)$$

colourとtypeの選択率が非常に高く、特にtypeに関しては、殆ど無条件で選ばれていると言える。typeの選択率は人間ドメインでも高く、ターゲットとディストラクタが全て人間であるにもかかわらず9割近い値だった。これは、参照表現を言語化する際にtypeが表現の主辞となるため、typeについて言及する事が自然だからだと考えられる。更に人間ドメインでは、眼鏡や髭などの顔に関する属性の選択率が高く、服装に関しての選択率は低い値であった。ただ服装についてはこのコーパスでは属性が非常に少なく、多様な服装をしている人間に対して参照表現を生成すると別の結果となる可能性も考えられる。

表3: 属性名の選択率への影響(家具ドメイン)

属性名	出現数	選択率(%)
colour	253	91.00
orientation	91	32.73
size	101	36.33
type	269	96.76
x-dimension	50	17.98
y-dimension	73	26.25

弁別能力の選択率への影響 属性の弁別能力の違いでの選択率の変化を調べたところ、弁別能力と選択率に正の相関を持つ属性と、弁別能力と選択率に殆ど関係の見られない属性が存在する事が分かった。

弁別能力に依存しない属性には、物体の位置を示す dimension などが挙げられる。特に dimension に関して言えば、同じ値を持つかどうかより、どのように配置されているかが重要であるためと考えられる。

属性値の選択率への影響 同じ属性名の中で属性値によって利用される傾向に偏りがあるかどうかを、それぞれの属性値での選択率から調査した。その結果、家具のドメインでは orientation が後ろ向きの場合が他の向きと比べて採用される割合が高く、約2倍であった。

人間のドメインの特徴としては、hasGlasses・hasBeard は殆どが“true”の時のみ採用されている。これらは“true”の場合のみ物体の特徴となり、“false”の場合にわざわざ言及する事を人間が好まないため採用しないと考えられる。

だが hasHair については、“true”の場合と“false”の場合で同じ程度に採用されていた。またこれは 4.2 節でも後述するが、hasHair が“true”の場合の多くは hairColour と共起していた。これから、「hasHair の“true” + hairColour」と「hasHair の“false”」は、人間の頭髮に関する情報として同じくらい採用されていると考えられる。

4.2 属性の共起

属性の中には、別の属性と共に選択されやすいものや、反対にある属性が選ばれると選ばれにくくなるものがある。この共起の度合いを測るために、属性間依存度 (式 3) を定義した。属性間依存度が 1 に近いと出現に依存関係はなく、1 より大きくなればなるほど同時に現れやすい。反対に 1 より小さいほど同時には現れにくくなっている。式中の $P(A, B)$ は A と B が同時に選択される確率で、 $P(X)$ は X が選択される確率である。家具のドメインの属性間依存度の表 4 からは、orientation と dimension が共起しにくい事などが読み取れる。

$$\text{属性間依存度}_{AB} = \frac{P(A, B)}{P(A) \times P(B)} \quad (3)$$

また表は割愛するが、hairColour は hasHair か hasBeard と必ず共起し、hasHair も 8 割は hairColour と共起している。

4.3 位置情報の選択

4.1 節より、位置情報は弁別能力と関係なしに選択されている。位置情報の採用に関して弁別能力以外の

表 4: 家具の属性間依存度 (type はほぼ 1 なので割愛)

	colour	orientation	size	x-dim	y-dim
colour					
orientation	1.06				
size	1.08	1.27			
x-dim	0.81	0.18	0.33		
y-dim	0.87	0.16	0.37	3.19	

手段を見つけるため、下記の方法で周囲の状況を分析した。

1. ターゲットから、それと colour や type などが類似したデストラクタまでの距離
2. デストラクタの重心と、ターゲットの間の距離

残念ながら、この分析結果は dimension の採用に関する有効なものではなかった。

4.4 参照表現の冗長性

人間が生成した参照表現の中には、その表現に使われている属性の一部で対象物体の特定が可能であるにもかかわらず、わざわざ冗長な属性を付加して表現を生成している場合がある。これは、冗長性の付加により表現の頑健性を得るためと考えられる。コーパス中の事例について人間が冗長な表現を用いているものとそうでないもので分けると、家具ドメインでは 220/278 で、人間ドメインでは 213/230 であった。冗長な表現がこれだけ選択されているからには、冗長な表現はコンピュータの表現を人間のものに近づけるために必要なものだと言えよう。

4.5 対象物体を特定できない事例

人間の参照表現の中には、冗長なものとは逆に不十分なもの—対象を一意に特定できないもの—も存在している。プログラムには実装しないが、人間の認知的な傾向を読み取るためこれらの特徴を分析する。また、これらの殆どは家具の事例である。

デストラクタとの距離 対象物にフォーカスしていると、事例の全体を認識していない可能性が考えられる。この仮説を、参照表現で除外できないデストラクタとターゲットとの間の距離で測定してみたところ、平均距離が 1.8 であった。事例の半分以上は周囲 8 マスに弁別できないデストラクタが存在したので、遠くにあるものを見逃しているとは考えにくい。これから、dimension の影響はほぼないものと考えられる。

選択されていない属性 上の結果を受け、dimension 以外の属性に着目してみる。ターゲットとディストラクタで属性値が違ってもかかわらず人間が選択しなかったものを、各事例にて算出する。結果の表 5 より、殆どに orientation あるいは size が含まれている。これらは視点からの対象の形状に違いを及ぼすものであり、形状の変わらない colour のみ違う事例は殆ど存在しない。これは、色彩より形状の認知的負荷が高いためにおこっていると考えられる。

表 5: 各事例で異なる属性名

属性名	個数
colour	3
orientation	19
size	20
colour,ori,size	6

5 プログラムの実装と評価

5.1 基本アルゴリズムの改良

3.1 節のアルゴリズムを、4 節の調査結果を基に改良した。

選択率 4.1 節より各属性の弁別能力にそれぞれの選択率をかけあわせてスコアとし、スコア順に属性を選択する。特に type については、最初に無条件で選択する。また一部の属性では、属性値によりスコアの変更や非選択といった変更を加える。

共起 4.2 よりある属性が選択された時に、その属性と共起しやすい属性の選択率を上げて逆の属性は選択率を下げる。

冗長 4.4 よりターゲットを一意に特定できる属性の組を選択した後に、冗長性として次点の候補だった属性を付加する。

組合せ 上記 3 つの手法を併用する。またここでは、冗長な属性の選択の際に一意に特定できる属性の組に type が含まれていない時は、type を冗長な属性としている。

それぞれ静的・動的アルゴリズムにて実装し、評価した結果が表 6 である。共起のみの場合に改善率が思わしくないのは、正解率が低い場合に共起を利用すると間違った答えに関連した答えを選択してしまう事から妥当だと考えられる。

表 6: 各プログラムでの平均 dice 係数

		家具	人間
選択率	静的	0.702	0.669
	動的	0.784	0.669
共起	静的	0.286	0.282
	動的	0.247	0.282
冗長	静的	0.382	0.340
	動的	0.404	0.340
組合せ	静的	0.775	0.703
	動的	0.798	0.703

6 まとめ

属性の選択について、本研究の手法で人間との一致率が改善することができる。だが、今回は冗長性など人間の持つ一部の傾向に関する分析が十分とは言えない。また、属性選択だけでは位置情報 (4.3 節) などの対処に限界がある。更に一致率を改善するためには、もう一步踏み込んだ選択アルゴリズムや空間的な関係からの参照表現の生成 [1, 2] が必要となる。

参考文献

- [1] 船越孝太郎, 渡辺聖, 栗山直子, 徳永健伸. 知覚的群化に基づく参照表現の生成. 自然言語処理. Vol.13. No.2. pp.79 – 97. 2006. Apr
- [2] 船越孝太郎, 渡辺聖, 徳永健伸. 知覚的群化を利用した参照表現の生成. 自然言語処理. Vol.14. No.1. pp.87 – 110. 2007. Jan
- [3] Dale, R., and Reiter, E. (1995). Computational interpretation of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 19(2): 233-263.
- [4] Reiter, E. (1990a). The computational complexity of avoiding false implicatures. *Proceedings of the 28th Annual meeting of the Association for Computational Linguistics, ACL-90*.
- [5] Philipp Spanger, Kurosawa Takahiro, Tokunaga Takenobu. Anja Belz and Sebastian Vargas. TITCH: Attribute selection based on discrimination power and frequency. *Proceedings of the MT Summit XI Workshop: Using Corpora for Natural Language Generation*. pp.98 – 100. 2007. Sep