

比較記述テキスト中の比較表現の分析

飯田 諒† 飯田 龍‡ 徳永 健伸†

†東京工業大学 大学院情報理工学研究科 ‡情報通信研究機構

iida.r.ab@m.titech.ac.jp, ryu.iida@nict.go.jp, take@cl.cs.titech.ac.jp

1 はじめに

参照表現の生成はテキスト生成の中心的な課題として研究されてきた [4]. 既存の参照表現の生成に関する研究の主な関心は、指示対象を他の対象と区別するために、指示対象が持つ属性の中からどの属性を選び、言語化するかという点にある。また、そのような属性を選択する上で曖昧性なく簡潔な表現で他の対象と区別することが中心的な研究課題であり、そのために種々の属性選択のためのアルゴリズムが提案されている [1, 2]. これに対して我々は、複数の対象が様々な特徴を与えられた際に、聞き手が対象を比較する上で必要な情報を提供するテキスト (以下、**比較記述テキスト**) の生成を目指している。対象の特徴を記述する点は既存の参照表現生成と同じであるが、目的が他の対象からの「弁別」ではなく、他の対象との「比較」であるという点が異なる。

比較記述テキスト生成の利用例として、インターネット上の商品の評判サイトなどに見られる商品の情報のうち、明示的に比較すべき商品・属性を特定し、それらに基づいてある商品群に対して特に言及すべき内容を文章に要約して提示することが考えられる。商品比較は表 1 に示す各対象 (商品) の各属性について属性値が与えられた形式で提示されることが多い。人間がこの表から各対象の情報を読み取る際には、表の内容を解釈し、各対象が他の対象に比べてどのような特徴があるかを把握する必要があり、表の規模・複雑さ次第では多大なコストを必要とする。そこで、このような表の内容のうち特筆すべき内容のみを文章として生成することで、ユーザが表を解釈する負担を軽減することができると考えられる。

ある対象に関する数値データなどの非言語情報から対象に関するテキストを生成する研究はこれまでも行われてきた。例えば、気象情報データから天気予報士が読み上げる天気予報の原稿を生成するシステム [9] や株価の変動データをテキスト化するシステム [5] などが開発されている。しかし、これらの既存研究は個別の場所の天気や株価について独立にテキストを生成しており、対象間の関係については生成を行っていないため、その関係についてはユーザが個別に読み取る必要がある。一方で、本研究では、対象横断的に中立的なテキストを生成するため、その点が大きく異なる。

対象の特徴を記述する際に、類似した対象との相違点を説明することでわかりやすい説明ができると言われていた [3, 6]. このように、他の対象と対比することで、言及したい対象を説明する方略は初期のテキスト生成でも使われてきた [7]. 例えば、直接比較 [8] という比較手法では、2つの対象を比較して、類似点と相違点を述べる方略をとる。この際に、2つの対象をそれぞれ偏りなく

記述するという点では本研究の比較記述テキスト生成の目的と類似しているが、直接比較は対象が2つに限定されているため、3つ以上の対象を比較する場合の生成の方略については言及されていない。本稿でもこの種の局所的な比較方略は個々の対象の特徴を記述する際には用いるが、比較記述テキストで中心的に扱うのは3つ以上の対象を横断的に比較する大域的な比較である。

比較記述テキストの自動生成を実現するために、我々はこれまでに表 1 に示すような表形式で提示された複数の商品とその属性値を対象にして人間に比較記述テキストを生成させ、それらの分析を行った [10]. その分析では、表の中の商品と属性の組み合わせで特定される項目 (以下、**セル**) がどの程度言及されやすいかを調査し、さらに回帰分析の問題として各セルが言及される割合を特定する手法を提案した。これは比較記述テキスト生成における属性選択の問題を解いていると考えることができる。この延長として本稿では、選択された属性を記述する際にどのような表現を使うかについて調査する。我々が収集した比較記述テキストには、ある対象の属性について述べる際に他の対象の同じ属性と比較した記述が多く含まれている。これは上述の局所的な比較表現を使った属性の記述であると考えられる。そこで本稿では、特に選択された属性を表現する上で他の対象との比較表現を使用する傾向について調査し、それをモデル化することを試みる。以下、「比較表現」は局所的な比較を行う表現という意味で使用する。

本稿では、まず 2 節でこれまでに構築した比較記述テキストコーパスの概要について述べる。3 節で、比較表現に対するアノテーション結果について報告する。さらに、4 節で各セルの比較表現の用いられやすさに関する分析と、比較対象の選択の傾向に関する分析結果について報告する。最後に 5 節でまとめと今後の課題について述べる。

2 比較記述テキストコーパス

我々はこれまでに比較記述テキストの自動生成を目的として、複数人の記述者に比較記述テキストを記述させた。また、属性選択の傾向やテキスト中の比較表現の利用傾向を分析するための準備として、収集したテキストに対して、主題として記述される対象と比較される対象に関するアノテーションを行った。

2.1 比較記述テキストの収集実験

収集実験では、表 1 に示すような 10 種類の表 (以下、**評価表**) を作成し、20 名の記述者に各評価表について各製品の特徴を比較して述べるテキストを記述させ、合計 200 の比較記述テキストを収集した。この比較の対象と

表 1: 記述者に与えられる評価表の例

属性	テレビ A	テレビ B	テレビ C	テレビ D	テレビ E	属性の説明
デザイン	★★★★★	★★	★★	★	★★★★	見た目のよさ、質感
サイズ	★★	★★	★	★★	★	省スペース性・コンパクトさ
機能性	★	★★★★★	★★★★	★★	★★	機能が充実しているか
応答性能	★★★★	★	★★★★★	★★★★	★★	動画などで残像が残らないか
音質	★★	★★★★	★★★★	★★★★★	★	音質の良さ
画質	★★★★	★	★★	★★	★★★★★	映像の鮮明さなど
操作性	★★	★★★★	★★★★★	★★★★	★★	リモコンや本体の操作のしやすさ
価格	75,000	70,000	80,000	75,000	75,000	価格
満足度	★★★★	★★★★	★★★★	★★★★	★★	総合的に評価した値

しては、家電製品を選択し、各家電製品に対する評価項目（属性）は価格.com¹のレビューに利用されている項目を用いた。比較する各製品の属性値には、星で記載された5段階評価、もしくは実数値が記入されている。属性の項目は価格.comで使用されているものをそのまま利用したが、属性値は異なる表の特性を調べるために、5種類の要因を考慮して恣意的に決定した。5種類の要因の詳細と各評価表と要因の関係については文献[10]を参照されたい。価格.comに掲載された家電製品のうち、テレビ、掃除機、電子レンジ、(ブルーレイ・DVD)レコーダ、炊飯器の5種類を記述対象とする製品クラスとして選択し、各製品クラスに対してそれぞれ2種類の評価表を作成し、合計10種の評価表を用意した。また、各評価表に対して400字程度の比較記述テキストを生成するように指示を与えた。

2.2 比較表現のアノテーション

2.1で説明した実験で収集した比較記述テキストに対し、各製品の記述に使われた比較表現のアノテーションをする前に先行研究[10]で実施したように、テキスト中のどの部分がどの対象のどの属性に言及しているかをアノテーションする。まず評価表のセルに対する言及は、「(Aのデザインが)良い」等の評価表現を伴うため、この評価表現を言及部分として抽出する。次に、その抽出した評価表現に対して、どの対象のどの属性が言及されているかをアノテーションし、さらにその評価の極性をアノテーションする。例えば、「テレビAは画質が非常に優れている」のような記述では、評価表現として「優れ」を抽出し、評価対象が「A」、評価の属性が「画質」、評価の極性が「肯定」の3つの情報をアノテーションする。そして、それらに追加する形で、比較表現のアノテーションとして、以下の3種類のタグ情報を付与する。

比較対象

記述者は、言及すべき評価対象に加えて、必要に応じて比較対象を適切な比較表現とともに記述する。そこで、タグ「比較対象」を利用して、このような比較対象として選択された製品名を評価対象と関連付ける。例えば、「テレビAはテレビBに比べてデザインが良いですが」という記述では、評価対象として特徴を記述されているテレビAに対し、その比較の対象として明示的に記述されているテレビBが比較対象となる。

比較の種類

比較表現は「Aに対してBは～」のような対比的な記述と「(A, B, Cの中で)Aが最も～」のような

評価対象を含むグループの中で評価対象がどのようなかという記述に大別できる。本研究では前者を「対比」、後者を「包含」と呼び、比較対象がどちらの比較表現で言及されるかをアノテーションする。例えば、「AはBより属性Xで優れている」という記述では、図1に示すように、評価対象Aが比較対象Bと対比的に比較されているため「比較の種類」は「対比」となる。一方、「Aは属性Xで最も良い評価を得ている」という記述では、図1に示すように、評価対象Aを含めた全対象のグループを考え、その中で属性Xにおいて、Aが最も優れていることを述べているため、「比較の種類」は「包含」となる。

「対比」の関係

「包含」の関係

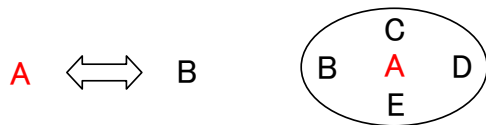


図 1: 「対比」の関係と「包含」の関係

評価対象と比較対象の評価の差

評価対象と比較対象が比較される際に、ある属性に関して評価対象の評価が等価であるのか、相違であるのかを区別することが可能であり、これらの違いによって生成の実現方法が異なる可能性がある。そこで、比較表現のアノテーション時には、この属性値の等価性を「相違」もしくは「等価」というラベルでアノテーションする。例えば、「AはBより属性Xで優れています」では、Aの属性Xによる評価はBの属性Xによる評価より優れており、属性値間に違いがあるため「相違」をアノテーションする。一方、「AはBと同様に属性Xで良い評価です」の場合は、AとBの属性Xによる評価に差はないため、「等価」をアノテーションする。

これらのタグ情報を用いてアノテーションを実施し、その結果について分析を行った。先行研究[10]で行った属性選択の分析は、各対象の特徴を記述する上でどの属性を用いるかを特定するために、評価表から算出される特徴量から各セルの言及されやすさを特定するモデルを導入した。しかし、このモデルだけでは、どのセルを比較記述テキスト中で言及するかを特定できても、その特徴を記述するのに比較表現を使うかどうかは決定できない。比較表現を含むテキストを生成するためには、あるセルについて言及した上で、比較表現を用いるかの決定と、比較対象とする対象を選択することが必要である。そこで本研究では比較表現の利用傾向と比較対象の選

¹http://kakaku.com

択における傾向を明らかにするために、比較表現のアノテーション結果を分析した。

3 比較表現使用の分析

比較表現のアノテーションをしたコーパスを対象にして、まず比較表現の利用傾向について調査した。ここでは比較記述テキスト中の「AはBに対して音質が優れた」とような1つの評価表現を単位として比較表現を計上する場合と、評価表のセルを単位として比較表現が使われたと計上する場合の2通りで調査した。セル単位の場合は、例えば「AはBに比べて音質と画質が優れている」というような記述では、評価対象Aの音質と画質の2つで比較が行われていると計上する。実施した比較表現のアノテーションは、比較の種類が「対比」と「包含」の2種類、評価対象と比較対象の評価の差が「等価」と「相違」の2種類あるため、それぞれの組み合わせとして4種類に分類できる。表2に各単位で計上した各分類の利用数を示す。表2が示すように「包含・相違（包含かつ相違の分類を指す）」の比較が最も多い。

表 2: 比較表現の分布

	評価表現単位	セル単位
対比・相違	261	303
対比・等価	66	83
包含・相違	577	692
包含・等価	0	0

次に比較記述テキストを生成した記述者ごとに、比較表現の利用傾向にどのような差があるかを調査した。図2に各記述者が各分類の比較表現をどの程度利用したかを示す。図2から記述者によって比較表現の利用傾向が異なることがわかる。縦軸は評価表現単位で計上した比較表現の利用数を表しており、記述者によって比較表現の利用数には差があることがわかる。各記述者の比較表現の利用数を平均すると45.0で標準偏差は35.4であった。また分類の観点では、「包含・相違」の比較を割合多く用いる記述者がいる一方で、「対比・相違」の比較を割合多く用いる記述者もいることがわかる。

4 比較表現の利用と比較対象の選択

3節の調査結果を元に、比較記述テキスト生成で比較表現を適切に使用するための分析を行った。本研究では、特に利用数の多い「包含・相違」と「対比・相違」の比較に対して、異なる手法で分析した。

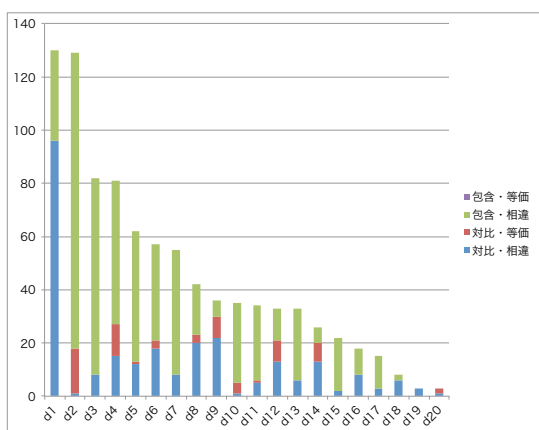


図 2: 各記述者の比較表現の利用傾向

4.1 「包含・相違」の比較を対象とした回帰モデル

「包含・相違」の比較表現数は今回構築したコーパスでは最も多く存在し、それらの比較はほとんどが最上級の表現を用いた比較であった。また、この種の比較表現の数は0から12の分布であったため、各セルに対して最上級の表現の用いられやすさを推定する回帰モデルの導出を行った。また、比較表現を用いるかどうかは言及することが前提となるため、全記述者の半数である10名以上が言及したセル244件を分析対象とした。回帰分析では、目的変数は各セルの「包含・相違」の比較表現が用いられた数とし、説明変数には次の5種類の素性を利用した。

属性値の唯一性 (F1)

各セルの属性値が属性内の観点で唯一の属性値かどうかを表す二値素性である。例えば、表1のデザインにおけるAとDとEはそれぞれ★★★★★, ★, ★★★で唯一の属性値であるため1を取り、BとCは★★で唯一の属性値ではないため0を取る。

属性内で最大もしくは最小の属性値 (F2)

各セルの属性値が属性内の観点で最大もしくは最小の属性値かどうかを表す二値素性である。例えば、表1のデザインにおけるAとDは1を取り、B, C, Eは0を取る。

対象の属性値と属性内の属性値の平均の差 (F3)

属性内で属性値の平均を計算し、各セルの属性値と差を取った値である。例えば、表1の(A, デザイン)のセルの場合、デザインの属性値の平均を計算し、(A, デザイン)の属性値との差を計算する。

対象の属性値と製品内の属性値の平均の差 (F4)

製品内で属性値の平均を計算し、各セルの属性値と差を取った値である。例えば、表1の(A, デザイン)のセルの場合、Aの属性値の平均を計算し、(A, デザイン)の属性値との差を計算する。

対象の属性値と表全体の属性値の平均の差 (F5)

表内で属性値の平均を計算し、各セルの属性値と差を取った値である。例えば、表1の(A, デザイン)のセルの場合、表1全体の属性値の平均を計算し、(A, デザイン)の属性値との差を計算する。

素性の計算では、文献[10]で述べられている正規化を各属性値に対して適用したものをを用いた。このような設定で各セルの最上級の比較の用いられやすさを数値化する回帰モデルを導入した。複数の素性を組み合わせる重回帰分析も行ったところ、自由度調整済み決定係数が最も高いのは、素性F1とF5の2つを説明変数とした回帰分析であり、その際の決定係数は、0.512であった。

4.2 「対比・相違」の比較における比較対象の選択

3節で既に述べた通り、303件観測された「対比・相違」の比較では、評価対象を除いた他の全ての対象を「その他の製品」とような表現を伴って、比較対象として選択するものも含まれている。しかし、本研究では特定の比較対象の選択を扱うこととして、これらの比較は分析対象から除いた。記述者の1人は、他の記述者に比べて「対比・相違」比較の利用数が極端に多いため、分析するデータに偏りを与えないために、今回の分析対象から除外した。このような基準によって限定された「対比・相違」の比較表現に対して比較対象の選択における傾向を分析する。ここで「対比・相違」の比較が用いら

れたセルは102種類あり、今回はこれらのセルに言及した際に、どの比較対象と対比するかという問題を扱う。例えば、「Cは使いやすさと機能・メニューでAより高い評価を受けています」のような記述を生成するために、評価対象Cの属性「使いやすさ」と「機能・メニュー」(セル)に言及する際に、「対比」の比較表現を使うとしたら比較対象としてAを選ぶという問題を解くことになる。対象 x の属性 p の値を $v(x, p)$ で表わすとし、属性 a について言及する際の対象 x の比較対象として対象 y の選択されやすさ $C(x, y)$ を以下の式で定義する。

$$C(x, y) = |v(x, a) - v(y, a)| - \text{ave}_{p \neq a} |v(x, p) - v(y, p)|$$

つまり、着目する属性値の両者の差が大きく、その他の属性値の差の平均が小さい程、比較されやすいと考える。コーパス中で対比が用いられた102件のセルに対して、全ての比較対象との組み合わせのスコアを計算し、スコア基準で上位1件の比較対象が実際に対比されていた比較対象となっている場合を正解として表3に示す結果を得た。

表3: 比較対象の選択

評価表	分析セル数	正解数	正解率
テレビ1	9	9	1.00
テレビ2	9	3	0.33
掃除機1	10	4	0.40
掃除機2	11	11	1.00
電子レンジ1	13	8	0.62
電子レンジ2	11	5	0.45
レコーダ1	6	3	0.50
レコーダ2	10	6	0.60
炊飯器1	10	6	0.60
炊飯器2	13	12	0.92
平均			0.64

表3が示すように、評価表ごとによって正解率が異なっている。正解率の平均は0.64であり、全体の6割以上のセルで正しく比較対象を選択できる。しかし、評価表によっては正解率が低いものもあるため、これらの差異が生じた原因を分析し、より正確に比較対象が選択できるようなスコアの導出が今後の課題として考えられる。

5 おわりに

本稿では、比較記述テキスト中の比較表現を生成するために、収集した比較記述テキストに対して局所的比較表現のアノテーションを行い、特に「相違」、つまり、比較される対象間で優劣がある場合に生成される明示的な比較表現に対して、分析を行った。「包含・相違」の比較に対しては、どのセルで最上級の表現を用いるかという問題として、5種類の素性を用い、比較の用いられやすさを数値化する回帰モデルを作成し、そのモデルの有効性を調査した。その結果、「属性の唯一性」と「対象の属性値と表全体の属性値の平均の差」の2つの素性を用いた場合が最も性能が良く、決定係数として0.512という結果を得た。「対比・相違」の比較では、対比的な比較を用いる際に評価対象に対してどの対象を比較対象として選択するかについて分析を行った。対象間のスコアを属性値の差異を利用して算出し、そのスコアがどの程度正しく比較対象を選択できるかを分析した。スコアの順位付けを行い、上位1件が記述者のうち少なくとも1名が

対比した比較対象であるか調査したところ、正解率0.64で比較対象を選択できることがわかった。

本研究では、分類した比較表現が各セルでどのように利用されるかを分析したが、各セルで比較表現の用いられやすさを予測する問題の分析対象は、「包含・相違」の比較に限られている。「対比・相違」の比較に関しては、比較表現が用いられるという前提のもとで、比較対象の選択に関する分析を行ったが、各セルで「対比・相違」の比較が用いられるかどうかに関しては分析できていない。そのため、本研究で設定した素性の洗練もしくは新たな素性を追加することで、「包含・相違」以外の比較表現の利用に関する予測モデルを開発することが今後の課題として挙げられる。また、属性選択のモデルと比較表現のモデル、さらに比較対象選択のモデルを開発して組み合わせることで、実際に比較記述テキストを表層化することが、重要な今後の課題として考えられる。

参考文献

- [1] Robert Dale. Cooking up referring expressions. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 68–75, 1989.
- [2] Robert Dale and Ehud Reiter. Computational interpretation of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, Vol. 19, No. 2, pp. 233–263, 1995.
- [3] Athanasios Karasimos and Amy Isard. Multilingual evaluation of a natural language generation system. In *Proceedings of 4th International Conference on Language Resources and Evaluation (LREC 2004)*, pp. 829–832, 2004.
- [4] Emiel Krahmer and Kees van Deemter. Computational generation of referring expressions: A survey. *Computational Linguistics*, Vol. 38, No. 1, pp. 173–218, 2012.
- [5] Karen Kukich. Design of a knowledge-based report generator. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics (ACL 1983)*, pp. 145–150, 1983.
- [6] Matthew Marge, Amy Isard, and Johanna Moore. Creation of a new domain and evaluation of comparison generation in a natural language generation system. In *Proceedings of the 5th International Natural Language Generation Conference (INLG 2008)*, pp. 169–172, 2008.
- [7] Kathleen McKeown. *Text Generation*. Cambridge University Press, 1985.
- [8] Maria Milosavljevic and Robert Dale. Strategies for comparison in encyclopedia descriptions. In *Proceedings of 8th International Natural Language Generation Workshop (INLG 1996)*, pp. 161–170, 1996.
- [9] Somayajulu G. Sripada, Ehud Reiter, Jim Hunter, and Jin Yu. Exploiting a parallel TEXT-DATA corpus. In *Proceedings of the Corpus Linguistics 2003 conference*, pp. 734–743, 2003.
- [10] 飯田諒, 飯田龍, 徳永健伸. 比較記述テキスト生成のための属性選択. 研究報告自然言語処理 (NL), Vol. 2014-NL-218, pp. 1–7, 2014.