

話者視点の解析を利用した談話構造抽出*

6R-7

永松健司 田中英彦†
 東京大学 工学部

1 はじめに

文章要約などの処理においては、文の接続関係や談話構造などの高次レベルの情報の取得が必要となる。しかし、現在の接続関係や談話構造解析はその精度において充分ではないと思われる。これは一般に、文の内容に関する知識の不足が大いに影響していると考えられるが、理解システム内の知識を、どのような領域において、かつどれだけ増やせばよいのかは明確ではなく、それだけでも困難な問題とされている。

これに対して、文章内に含まれる（と考えられる）筆者の意向の痕跡をできるだけ汲み取れるならば、それを利用することで、談話構造解析などの精度向上を期待できると考えられる。そこで、我々は新聞社説を対象とし、話者（筆者）の視点に注目してその解析を行った上で、視点情報を利用した談話構造抽出を試みている。本稿では、主にその解析の手法について述べる。

2 融合に基づく話者視点の解析

2.1 視点情報の定義

従来、視点に関しては様々な研究がなされてきているが、その定義、意味は種々であり、まさに様々な視点からの研究が行われている状況である。そこで本研究では、次のような情報からなる構造を定義し、視点として取り扱うことにする。

- 話者（筆者）… この視点の保持者
- 中心話題… 文の中心となる話題。この視点構造の対象となる概念を含む節の意味情報で表す
- 比較視点… 中心話題と関連づけられる視点。
- 話者態度… 中心話題に対する話者（筆者）の態度（以下の中から一つを値として取る）
 - + 観測事象 (f) … 中心話題が現在および過去の確認しうる出来事の場合。事実を記述した文に対応する。

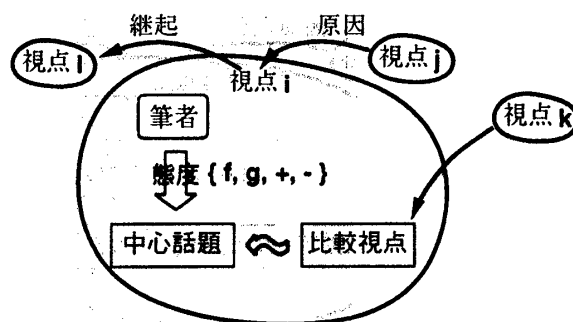


図 1: 視点情報の概念図

- + 推測事象 (g) … 中心話題が原理的に確認できない状況、事態の場合。話者の推測が含まれる。
- + 肯定意見 (+) … 中心話題に対して肯定的意見の場合。話者（筆者）の意見を述べた文に対応する。
- + 否定意見 (-) … 中心話題に対して否定的意見の場合。話者（筆者）の意見を述べた文に対応する。

この視点構造は、話者が、比較視点から中心話題を見て、その話題に対し話者態度に示される態度を取っているという意味を持つ（これを図示したものが図1。大きな丸の中）。

処理において、この構造は、解析対象となる文の中の各節（行為者-行為構造）ごとに解析・生成される。また、比較視点としては、これまでに生成された視点、または予め知識として与えられているプリミティブな視点（空間的、原因-結果、etc.）を取ることができる。

2.2 視点構造の解析

視点構造の中の各情報は次のような方法によって解析し取り出される。

- 中心話題… 中心話題は、その視点構造に対応する節全体の意味情報を指すとは限らない。その下位部分となる視点構造と同じ話題が選択される場合もある。この選択は、構文的に子にあたる節の視点構造の中心話題との比較によって判断する。
- 話者態度… 意見であることを示す副詞句、動詞句が含まれている場合に、その内容から肯定意見、否定意見のいずれかを選択する。

*Extracting Discourse Structures Making use of Analyzing View-point

†Kenji Nagamatsu Hidehiko Tanaka
 {naga,tanaka}@mtl.t.u-tokyo.ac.jp

Faculty of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo, 113, Japan

観測事象か推測事象かの選択の場合は、現在知識ベースに依っている。

- 比較視点… 基本的に、中心話題の中の主な名詞句がどのような格役割を持つかによって、プリミティブな視点構造を与える。但し、構文木で上位にあたる視点構造においては、その中心話題や話者態度などから、より抽象化された視点構造を取る必要もあり、現在は知識ベースでの選択を行う。

2.3 視点の融合

以上のようにして各節に対し生成された視点構造の内、同じ視点だと思われるものについては一つにまとめる必要がある。つまり、ある二つの視点構造において、話者、話者態度がそれぞれ同じ、かつ一方の中心話題がもう一方の中心話題の上位概念となる場合、これら二つの視点構造は同一の視点として融合される。

上位概念かどうかの判断は、現在のところ、あらかじめ作成しておいた名詞・動詞概念の上位-下位関係データによって、中心話題中の主な名詞・動詞間の概念関係を調べている。

3 視点構造を利用した談話構造抽出

3.1 視点間関係の解析

前節のようにして生成された視点構造は、それだけでは十分な機能を果たせない。つまり、これらの各視点構造がどのような関係で結び付いているかを決定する必要がある。

- これに対して、文の接続関係(主に接続詞)を利用した関係付けを行うことで対応している。すなわち、
- 視点構造が対応している節どうしの意味関係(原因、継起、etc.)。動詞の格情報などから解析する。
- 接続詞(接続助詞)による接続関係。文(節)間を結ぶ接続詞の意味から解析する。

を解析し、それを各視点間関係として判断する。

3.2 談話構造の抽出

すべての視点構造の解析およびその間の関係付けが完了した時点で、最終的な談話構造(接続関係)の抽出処理が行われる。

これには、視点構造間の接続関係、話者態度および中心話題の組から接続関係への写像をテーブルとして持っており、それを参照することで結果を得る。その時、中心話題に対しては話題間の距離を考えることで柔軟に対応する。

4 システムの構成

現在、前述のような手法に基づいた、談話構造解析システムを実装しているところである。その概念的な構成を図2に示す。処理の手順としては、

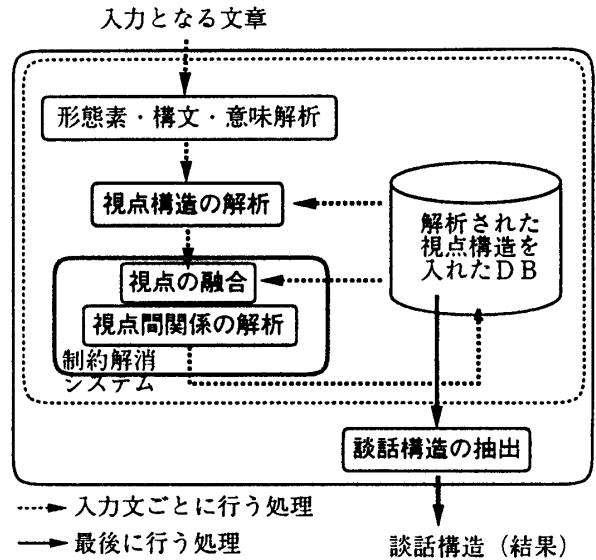


図2: システム構成の概念図

1. 入力となる文章(新聞社説)を順に解析し、
2. 各文に対して視点解析(視点構造の解析、視点の融合、視点間関係の解析)が行われる。この時、それ以前の解析で得られている視点構造および知識として予め与えられている視点情報を蓄えたデータベースへのアクセスを行う。
3. 視点解析の結果得られた視点構造(およびこの過程で変更を受けた以前の視点構造)をデータベースへと追加する。
4. 1.から3.までを入力文すべてに対して繰り返す。
5. 得られた視点構造(データベース中にある)すべてから、談話構造の抽出を行う。

のようになる。

ここで、2.の視点の融合および視点間関係の解析部分は、以前に得られた視点構造の内容をも変更しうるものであるため、全体の意味的な一貫性を保持するために、制約解消を行う機構を入れて矛盾を検出するようにする。

5 おわりに

視点情報を利用して談話構造の抽出を行う手法について述べた。現在、以上で述べた手法を基に、実装を行っているところである。実装依存となる詳細も含めて、この結果を報告できることを目指している。

参考文献

[1] J.M. Wiebe. Tracking point of view in narrative. *Computational Linguistics*, Vol. 20, No. 2, pp. 233-287, 1994.
 [2] 田村俊哉, 田村直良. 文章の表現形式に基づいた要約文章の生成について. 情報処理学会自然言語処理研究報告, 1992. 92-1.