社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2006－ＦＩ－ 84 (19)
2006－ＮＬ－175 (19)
2006／9／13

意見文に関する情報付与と意見に焦点を当てた要約の関係の分析

関 洋平[†] 江口浩二[††]
神門典子[††] 青野雅樹[†]

テキストからの意見抽出研究の一つとして，テキスト集合から意見を集約した要約の生成がある．本研究では，このような要約の生成における，文書の意見性（文書ジャンル），文タイプの意見性，意見句，意見保有者，意見の極性，強度などの意見属性の効果について調べた．実験では，DUC2006 の英語の要約データセットから意見に関わる 15 トピックを選択し，評価指標として ROUGE, BE を用いた．まず，提案手法の理論的上限値を調べるために，人手で文書集合に，文タイプの意見性，ならびに，意見保有句，意見の強さなどの文より小さいレベルでの意見属性を付与した．文書ジャンルは NIST によって提供されたものを利用した．また，システムの実現可能性を調査するために，文タイプの意見性ならびに意見の強さについて自動付与を行った．その結果，日本語要約の先行研究と同様に，英語の要約でも文タイプと意見性と文書ジャンルの情報を要約のパラメタとして組み合わせることで，条件によっては精度の向上があることが明らかになった．また，意見の強さが要約のパラメタとして精度向上にわずかながら寄与することがわかった．

An Analysis of Opinion-Focused Summarization using Opinion Annotation

YOHEI SEKI,[†] KOJI EGUCHI,[††] NORIKO KANDO[††] and MASAKI AONO[†]

One important application of sentiment/opinion analysis is to produce summaries by aggregating opinions from multiple texts and provide them to users who request opinions. In this research, we selected opinion-related 15 topics in DUC 2006 and investigated the improvement effect of ROUGE and BE scores to produce opinion-focused summaries by extracting opinionated sentences in opinionated genre of documents and using phrase-level opinion attributes. In order to examine the upper ceiling of the proposed approach, opinionated sentences and opinionated attributes such as opinionated phrases, opinion holders, opinion intensity, and opinion polarity were annotated to source documents by a human assessor. We also utilized document genre information given by NIST. To investigate the feasibility, opinionated sentences and opinion intensity were also automatically annotated. As the past results in Japanese summarization, we found that ROUGE and BE scores improved by combining document genre and sentence type information as summary parameters. In addition, we also found opinion intensity is important as summary parameters.

1. はじめに

近年，テキスト中の意見を分析し，自動抽出する研究が盛んである．このような研究は，評判抽出や世論分析などの社会調査といった応用が期待されている一方で，情報検索[1]や質問応答[2]，さらには自動要約[3)~11)]といったテキストからの情報アクセス技術の向上につながるという点でも強く関心を持たれている．

われわれは，文書単位の意見性の情報（報道記事や社説といったジャンルの区別）と，文単位の意見性の

† 豊橋技術科学大学
Toyohashi University of Technology
†† 国立情報学研究所
National Institute of Informatics

情報の双方を要約生成のパラメータとすることにより，事実だけでなく，意見や知識といった側面に焦点を当てた要約を生成することについて研究を進めてきた[12)~14)]．意見抽出に基づく要約の研究は，他にも行われている[6),7)]．しかし，ここ数年，意見の抽出は，文よりも細かい単位の属性に着目する研究が進んでいる[4),5),15)~19)]．

本研究では，このような背景を受けて，文書集合に意見の属性を付与し，複数文書要約のパラメタとして使用することの効果について調査を行った．データは，要約の評価会 DUC 2006 で使用されたトピックから，意見に関連すると考えられる 15 トピックを選択した．意見情報は，先行研究[12),13)]で，日本語の意見に焦点を当てた要約に効果があることを示した文単位の意見性と文書ジャンルに加えて，意見句，意見の保有者，意見の強さ，意見の極性といった合計 6 つの属性について付与を行った．

本論文の構成は以下の通りである．2 章に関連研究を示す．3 章では実験で使用したデータセットと実験方法について説明する．4 章で結果と考察を示す．5 章でまとめを行う．

2. 関連研究

本節では，意見の属性判定に関する最近の研究についてまとめる．なお，評判抽出タスクなどでも 3 つ組の抽出といった関連研究が行われているが，既に他のサーベイ論文[20)]で紹介されており，紙数の制限もあるため，詳細は割愛する．また，文書ジャンルや文タイプを対象とした先行研究も多いが，これについては Lin ら[21)]をご参照されたい．

2.1 意見保有者 (Opinion Holder) の抽出

ここ 1 年ほど，意見保有者 (Opinion Holder) ☆の自動抽出に関する研究[4),5),15)~17)]が盛んである．Choi ら[17)]は，意見保有者が名詞であること，意見を表現できる意味クラスに属すること，意見句と依存関係があることに着目し，条件付確率場を用いることで，6 割以上の F 値で意見保有者が抽出できることを示した．この研究では，1 文中に複数の意見保有者が存在する場合などが問題となっていた．Kim ら[15)]は，この点に着目し，1 文中の複数の意見を区別し，最大エントロピーモデル[15)]や FrameNet[16)]に基づきそれぞれに対応する意見の保有者やトピックを推定する手法を提案した．

一方 Stoyanov ら[4),5)]は，意見指向要約の実現に向けて，意見保有者間の共参照関係の解決 (coreference resolution) を試みている．彼らは機械学習アプローチにおいて，訓練データとして意見情報を付与したコーパスでは，すべての名詞の共参照関係が含まれないという問題を解決するために，部分教師ありクラスタリング (Partially Supervised Clustering) という手法を提案している．

2.2 意見の極性 (Opinion Polarity) の判定

意見の極性（肯定，否定，中立）の判定は，主に文書や文を単位として研究が行われてきたが，最近では，句や節といったより細かい単位を対象として判定を行う研究が進んでいる．基本的な手法は，句または節に含まれている意見を表す語が肯定語であるか否定語であるかといった方法[15)]に基づいているが，問題点として，推測，信念などの意見性は中立的な極性を保つことが多い点と，周囲の文脈により意見を表す語が持つ極性が反転することも多い点が上げられる．Wilson ら[18)]は，極性を持つか中立かを分類してから，極性を持つ句について肯定・否定の分類を行う手法を提案した．

2.3 意見の強さ (Opinion Strength) の判定

意見の強さの研究は，上記 2 つの属性と比べると少ない．Cesarano ら[23)]は，文書単位の意見の強さを判定するために，意見を表す語の強さを計算する方法を示した．彼らの提案では，まず，被験者に文書ごとの意見の強さを 0 から 1 の間の数値で判定してもらい，文書中の意見を表す語の相対出現頻度を考慮し，意見を表す単語（形容詞）の強さを決める．

Wilson ら[19)]は，MPQA コーパス[22)]における意見句に付与された neutral, low, middle, high, extreme の 5 値の意見の強さを利用し，手がかり句が，5 値に対応して出現する確率が閾値を超えたものだけを特徴素として利用することにより，BoosTexter や RIPPER などの機械学習アルゴリズムで節を単位とした意見の強さが分類できることを示している．

2.4 意見要約への応用

意見属性を利用して意見要約技術への応用を試みた研究としては，初期の研究として Cardie ら[3)]がある．この研究の現状は，既に説明したように，Stoyanov ら[4),5)]において，意見保有者をパラメタとして利用することで，複数の見方を区別した意見を要約として生成することを目指している．さらに，意見を問う質問に対する応答技術について，Stoyanov ら[2)]は，従来の質問応答技術と意見を問う質問応答技術の違いや，意見保有者を利用したフィルタリングについて論

☆ MPQA コーパス[22)]を利用した研究では，意見の出所 (Opinion Source) という言葉が使用されることが多い．

じている．

そのほか，意見文を抽出することで，要約生成に応用する研究としては，映画レビューを対象としたもの[7]や，TREC Novelty Track のデータを利用したもの[6]がある．また，抽出した意見を可視化することでユーザに提示する研究も行われている[8),9)]．また，佐々木ら[10]は，論点と代表的な意見を抽出し可視化することを試みている．

また，奥村ら[11]は，意見要約コーパスを作成しており，要約の評価会 TSC-4 のデータとして共同研究者らと研究を進めている．意見の属性としては，意見の下位分類タグ 8 種類を節単位で付与している．

以上のように，意見の属性の自動判定ならびに要約生成技術への応用についても研究が進みつつある．本稿では，意見の属性を積極的に利用した要約の生成について，分析する．

3. 実　　験

本節では，前節の背景を受けて，意見の属性と意見に焦点を当てた要約生成との関連を調査するために行った実験について説明する．まず，実験データについて，次に文書に付与した意見情報について，最後に実験方法について説明する．

3.1 実験データ

英語の要約データとしては，意見に焦点を当てた要約データは著者の知る限り存在しない．そこで，本研究では，要約の評価会である DUC 2006* のデータから，意見に関連した内容を問う質問を含んでいる 15 トピックを選択し，分析を行った．DUC は，2000 年から**毎年米国標準技術院 (NIST) が開催している．2005 年からは，実世界の人間の複雑な質問，すなわち，名前，日付，量など断片的な事実（Factoid）では回答できない質問の回答を提供することをモデル化した，複雑な質問に焦点を当てた要約作成タスク (complex question-focused summarization) が行われている．なお，質問は 1 トピックにつき複数（最大で 4 つ程度）あるものも多く，そのうち 1 つだけが意見を問う質問であるトピックも含まれている．要約作成の対象となる元の記事は AP 通信，ニューヨークタイムズ，新華社通信から選択され，1 トピックあたり 25 記事である．報道記事 (news story)，その他のジャンルが両方含まれている．また，250 語のモデル要約が，4 名の判定者により作成されている．

表 1　実験データ：DUC2006 から選択した 15 トピック

ID	トピック
D0601	アメリカ先住民族居留地システム-賛否両論-
D0603	湿地の価値と保存
D0604	スター・ウォーズ エピソード I に対する期待と反応
D0606	世界の気候変化の影響
D0609	イスラエルによるヨルダン川西岸への入植
D0610	自宅学習-賛否両論-
D0615	進化論/創造論に関する議論
D0619	同性愛者と共和党
D0623	禁煙法
D0624	ステファン・ローレンス
D0628	注意欠陥（多動性）障害-診断と治療
D0635	ブッシュ知事在任中のテキサス州における死刑
D0636	自動車労働組合とアメリカ自動車製造業者の問題
D0641	地球温暖化
D0642	ウゴ・チャベス

3.2 意見属性の付与

本研究では，3.1 節で説明した実験データの記事集合について，以下の意見の属性を付与した．判定者は，米国人の翻訳家 1 名が担当した．判定者は，文書集合がどのように要約されたかの知識は与えられていない．

(1) それぞれの文が意見を表しているか否か
(2) 文中に含まれる意見句を，Wiebe ら[24]を参考に，以下の 3 つのタイプに分類して抽出
  (a) 心的状態の明示的な記述 (explicit mentions of private states)
  (b) 発話/記述イベント (speaking/writing events)
  (c) 主観を表現する要素 (expressive subjective elements)
(3) 意見句に対応する意見保有者
  (a) 心的状態を表明した人または組織など
  (b) 発話/記述イベントの主体
  (c) 主観を表現する主体（書き手/話し手など）
(4) 意見句の強さ
  確実 (Extreme)，強い (High)，中くらい (Middle)，弱い (Low)
(5) 意見句の極性
  肯定 (Positive)，否定 (Negative)，その他

また，システムの実現可能性を調査するために，文タイプの意見性について，SVM^{light}***（多項式カーネル，コストは 1.2）を用いて，MPQA コーパス[22]を訓練データとして，Hatzivassiloglou ら[25]の形容詞

* http://duc.nist.gov
** 初年度は 5 チームのみの参加で，ロードマップの作成に重点が置かれた．2001 年以降は毎年 20 チーム以上が参加しており，われわれは 2005 年と 2006 年に参加した．
*** http://svmlight.joachims.org/

辞書と General Inquirer[26] 中の極性語をタイプ別に9つの特徴素として用いることで，自動判定を行った．MPQA コーパスについて 5-fold cross validation をした結果と，DUC の 15 トピックについての人間の手動付与に対する正確さ (accuracy)，精度 (precision)，再現率 (recall) を **表 2** に示す。

表 2 意見文自動付与の正確さ，精度，再現率

	正確さ	精度	再現率
k-fold	0.602	0.610	0.657
DUC	0.527	0.419	0.540

また，意見の強さとして，SVM^{light} が各文に対して出力する実数値のうち，0 以上の値を用いることで，実現可能性について検証した．詳細は，次の節の評価の項で説明する．

3.3 実験方法

われわれは，先行研究 [12),13)] で，日本語について，30 トピックに関連した新聞記事の集合から，事実，意見，知識を区別した要約データセット ViewSumm30 を開発し，文を単位とした意見性と文書ジャンルを重要文抽出のパラメタとして組み合わせることで，モデル要約に対するカバレッジが，意見情報を用いないベースラインシステムに対して有意に向上することを明らかにした．

本稿でも，文を抽出する際の重み付けのパラメータとして，文を単位とした意見性，文書ジャンル，意見の属性を重要文抽出のパラメタとすることで，意見情報を用いないベースラインシステムと比べて要約評価精度の向上に寄与するかどうかについて調査を行った．ベースラインシステムについては，日本語の研究 [12),13)] と同様に，クラスタリングに基づく複数文書要約システムを実現して用いた．このシステムでは，文書集合を段落単位で区切り，各段落を内容語の類似度を用いてクラスタリングして，それぞれのクラスタから，重要文を抽出して組み合わせて要約を作成する．DUC 2006 に参加した結果の詳細は，14) を参照のこと☆．

質問文の意見性判定

先行研究と異なる点として，本稿で対象とした要約データは，複数の質問に焦点を当てた要約を作成している点がある．本稿では，Stoyanov らの仮説 [2)] に基づき，意見を問う質問が文書中の意見に関連するものとした．実験データ中の質問は，必ずしも意見を問うものではないため，意見に焦点を当てた質問とそうでない質問を区別する必要がある．本研究では，3.2 節で説明したように SVM^{light} を用いて，Hatzivassiloglou ら [25)] の形容詞辞書と General Inquirer[26)] 中の極性語をタイプ別に 9 つの特徴素として用いることで，文書集合中の意見文を自動判定しているが，この手法を質問文の意見性の自動判定についても利用した．ただし，質問文の意見性判定の際には，WordNet ☆☆ を用いることで，同義語と上位概念が前記 2 つの極性語辞書に登録されている語を展開した上で，辞書との照合を行った．また，6 個のキーワード（commentary, arugument, discuss, react, reaction, concern）が質問中に含まれる場合には，無条件に意見を問う質問とした．

さらに，肯定意見・否定意見を問う質問について，質問文中の内容語の WordNet における上位概念が good や bad の概念を含むかどうかにより，自動判定を行った．

意見に焦点を当てた質問を考慮した要約生成

DUC では，要約文字数は 250 語と設定されている．また，要約生成に際して焦点を当てる質問の数はトピックごとに異なり，1〜4 文程度である．意見を問う質問であるかどうかは，質問ごとに異なるため，本研究では，250 語を，文単位で分割した質問の数で割り，各質問文ごとに対応する要約を割った文字数内で生成し，最終的に出版時系列順に並べなおすかたちで要約を生成した．質問が意見に焦点を当てた質問である場合には，意見情報に関連するパラメタを利用して重み付けを行う．なお，要約ごとに余剰した文字数は調整するようなアルゴリズムを実現している．

各クラスタからの重要文抽出は，以下の式 1 に基づきそれぞれの文の重みを計算することで実現した．ただし，L(s) は，文の文書内位置に基づいた重み，Q(s) は文は，質問中の内容語が文に含まれる数，H(s) は，文の見出しの内容語が文に含まれる数，T(s) は，文中の内容語の tf.idf 値の合計数（idf は DUC 2006 で用いた総文書数を基準として計算），N(s) は質問中の固有表現のタイプに基づいた固有表現数（固有表現の抽出には OAK[27)] を使用）を示している．各パラメタの重みは，$a_1 = 0.8$, $a_2 = 1/$見出し中の内容語の数, $a_3 = 1$, $a_4 = 0.2$ とした．

$$W(s) = L(s) \times \underline{G(s)} \times (a_1 \times Q(s) + a_2 \times H(s) + a_3 \times T(s) + a_4 \times N(s) + a_5 \times \underline{S(s)} + a_6 \times \underline{Pos(s)} + a_7 \times \underline{Neg(s)}). \quad (1)$$

☆ ただし，この結果は文書ジャンルを利用していない

☆☆ http://wordnet.princeton.edu/

下線部は，意見に焦点を当てた場合に使用するパラメタである．G(s) は，文書ジャンルのパラメタであり，文書ジャンルが News Story である場合に，0 とした．S(s) は，文を単位とした意見性または，それに関連した句の水準の意見属性を利用して，文が条件を満たすときに重みを加える．また，肯定・否定を問う質問である場合には，極性語辞書を利用して文中の肯定語・否定語の頻度 (Pos(s), Neg(s)) を重みとして加える．実験では $a_5 = a_6 = a_7 = 1$ とした．

評　　価

評価は，DUC でも公式に採用されている自動評価ツールである ROUGE[28] と BE[29] を採用した．また，意見の属性を用いる本提案手法の理論的上限値を調べるために，人手で意見属性を付与し，以下の重み付け戦略について比較を行った．

(1) 意見属性を重みとして利用しない（ベースライン）
(2) 文書ジャンルを利用（News Story 中の文を要約から排除）
(3) 文単位の意見性を利用（意見文を重み付けして抽出）
  (a) 意見文をすべて重み付け
  (b) 意見句の強さが L 以外のものである意見文を重み付け
  (c) 質問中に含まれる意見保有者を含む意見文を重み付け
  (d) 肯定・否定を問う質問である場合に，意見句の極性が対応する文を重み付け
(4) (2) と (3) のうち，効果があるものを組み合わせて重み付け

また，システムの実現可能性について，3.2 節で説明したように，文書集合に SVM^{light} を用いて自動付与した文単位の意見性を用いて，前項の (3)(a) の効果について検証した．また，意見の強さとして，SVM^{light} が各文に対して出力する実数値のうち，0 以上の値を用いて，式 1 の $a_5 \times \underline{S(s)}$ の項に乗することで，(3)(b) の実現可能性について検証した．

(3)(c)，(3)(d) については，人手で付与した情報について，目立った向上は見られなかったため，結果からは割愛する．

4. 結果と考察

4.1 結　　果

実験結果を **表 3** に示す．ベースラインは，前節の実験方法の評価の (1)，組み合わせは (4)，文タイプは (3)(a)，意見の強さは (3)(b)，文書ジャンルは (2) に対応する．文書ジャンルは入力文書にあらかじめ付与されており，自動付与は行っていない．また，理論的上限値のうち，BE の評価結果で意見の属性を組み合わせたものについては，Wilcoxon 検定（両側検定，有意確率 0.05 以下）でベースラインからの有意差があることがわかった☆．

表 3　意見情報を用いた要約精度の向上

タイプ		ROUGE -2	ROUGE -SU4	BE
ベースライン		0.07244	0.12838	0.02708
理論的上限 (人手付与)	組み合わせ	0.07758	0.13286	0.03543*
	文タイプ	0.07484	0.13066	0.02932
	意見の強さ	0.07504	0.13083	0.02932
	文書ジャンル	0.07514	0.13166	0.0314
システム (自動付与)	組み合わせ	0.07521	0.13087	0.03065
	文タイプ	0.07152	0.12658	0.02677
	意見の強さ	0.07252	0.1294	0.02646

*：Wilcoxon の符号付き順位検定（両側，有意確率 $\leq$0.05）で有意差あり

4.2 考　　察

表 4 に，理論的上限とシステムの実現可能性を調査した結果のうち，パラメタを組み合わせものについてのトピック別の向上の状況を示す．

表 4　トピック別の向上の状況

Topic	ROUGE -2		ROUGE -SU4		BE	
	理論値	システム	理論値	システム	理論値	システム
D0601	○	○	○	○	○	○
D0603	○	×	○	×	○	×
D0604	○	○	○	○	○	○
D0606	○	×	○	×	○	○
D0609	○	○	×	×	×	×
D0610	○	×	×	×	○	○
D0615	○	○	○	○	○	○
D0619	×	×	×	×	○	○
D0623	○	×	○	×	×	×
D0624	○	○	○	○	○	○
D0628	○	○	○	○	○	○
D0635	×	×	×	×	○	○
D0636	○	○	○	○	×	×
D0641	×	×	×	×	×	×
D0642	○	○	○	○	○	○

○:向上，×：悪化

- D0601, D0603, D0604, D0606, D0615, D0624, D0628, D0642 の 8 つのトピックは理論的上限値は，すべての評価尺度について値が向上している．このうち，D0603, D0606 については自動評価の尺度が悪化しており，意見文の自動付与に改善の余地があることがわかる．
- 残り 7 つのトピックのうち，D0619 については，意見に焦点を当てた質問の判定が誤っていたこと

☆ 補足：トピック数が 15 トピックで，有意差検定をするにはやや厳しい条件である．

がわかった．D0635 については，1 文中に複数の質問が含まれており，意見に焦点を当てた質問がそのうちの 1 つであるが区別できない点が問題である．この問題には，Harabagiu ら [30] が提唱しているような構文情報やキーワードを利用した質問分割が必要と考えられる．その他のトピックは，文書ジャンルが News Story である文を排除している点が悪化の原因となっている．

- 意見保有者が質問文に照合した場合の重み付けがうまくいかなかったのはいくつかの理由があるが，質問文中の内容語と照合する数が少ないか，逆に多すぎる結果になってしまうことが原因として挙げられる．
- 肯定・否定を問う質問については，辞書を基にした極性語頻度が精度向上に寄与しているにもかかわらず，句を単位として手動で付与した極性は向上に寄与しなかった．原因については，継続して分析を行う必要がある．

5. おわりに

近年の意見抽出，意見要約に関する研究の高まりを受け，意見抽出で用いられる属性について，意見要約への効果について調査した．英語の要約データ DUC 2006 から意見に関連するトピックとして選択した 15 トピックについて調査した結果，先行研究 [12),13)] で日本語の要約について得られた結果と同様に，文書ジャンルと意見文の情報を要約のパラメタとして利用することにより，条件によっては，要約の評価が向上することがわかった．また，付与した句の水準の属性のうち，特に意見の強さが要約の精度向上に寄与することがわかった．ただし，要約データは質問に焦点を当てて要約を作成することに主眼がある点に注意する必要がある．

また，われわれは，日本語についても，毎日新聞，読売新聞の 1998 年〜2001 年の記事からトピックごとに 10 記事程度を選択し，**表 5** に示す要約データ OpinionSumm15 を作成しており，文書集合には本稿で示した意見に関する属性を付与している．今後，評価手法を検討したうえで，このデータについての調査結果もあらためて報告を行う予定である．

表 5　日本語要約データ：OpinionSumm15

ID	トピック	記事数	記事文字数合計
S310	防空識別圏	8	9007
S320	政党助成金	7	13087
S330	小選挙区比例代表並立制	7	12296
S340	ビール業界	10	10017
S350	海運 合併	8	9232
S360	経営健全化 大手銀行	8	12916
S370	マイライン 開始	9	15357
S380	日米外相会談 田中 パウエル	11	13148
S390	臨時国会 PKO 改正	11	11939
S400	ギリシャユーロ加盟	9	8477
S410	医療保険制度改革関連法案	11	11155
S420	官房機密費 松尾室長逮捕	9	11078
S430	教科書検定	10	13518
S440	高祖氏 選挙違反	9	14964
S450	京都議定書 COP7	7	12578

補　足

本稿の研究とは一部の属性やデータ，研究目的などが異なるが，NTCIR-6 において，パイロットタスクとして Opinion Analysis の評価 ☆ を行う予定である．詳細は，脚注にあるホームページをご参照されたい．

謝　辞

この研究の一部は，文部科学省科学研究費補助金若手研究 (B) (課題番号 18700241)，若手研究 (A)（課題番号 17680011），萌芽研究（課題番号 16650053）を受けて遂行された．

参考文献

1) Eguchi, K. and Lavrenko, V.: Sentiment Retrieval using Generative Models, *Proc. of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, Sydney, Austraria, pp. 345–354 (2006).
2) Stoyanov, V., Cardie, C. and Wiebe, J.: Multi-Perspective Question Answering Using the OpQA Corpus, *Proc. of the Human Language Technology Conf. / Conf. on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, Vancouver, Canada, pp. 923–930 (2005).
3) Cardie, C., Wiebe, J., Wilson, T. and Litman, D.: Combining Low-Level and Summary Representations of Opinions for Multi-Perspective Question Answering, *Proc. of AAAI Spring Sympo. on New Directions in Question Answering*, Stanford, CA, pp. 20–27 (2003).
4) Stoyanov, V. and Cardie, C.: Toward Opinion Summarization: Linking the Sources, *Proc. of Wksp. on Sentiment and Subjectivity in Text at Proc. of the 21th Int'l Conf. on Computational Linguistics / the 44th Ann. Meeting of the Assoc. for Computational Linguistics (COLING/ACL 2006)*, Sydney, Austraria, pp.

☆ http://research.nii.ac.jp/ntcir/ntcir-ws6/opinion/index-en.html

9–14 (2006).
5) Stoyanov, V. and Cardie, C.: Partially Supervised Coreference Resolution for Opinion Sumarization through Structured Rule Learning, *Proc. of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, Sydney, Austraria, pp. 336–344 (2006).
6) Ku, L., Lee, L.-Y., Wu, T.-H. and Chen, H.-H.: Major Topic Detection and its Application to Opinion Summarization, *Proc. of the 28th ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2005)*, Salvador, Brazil, pp. 627–628 (2005).
7) Pang, B. and Lee, L.: A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts, *Proc. of the 42nd Ann. Meeting of the Assoc. for Computational Linguistics (ACL2004)*, Barcelona, Spain, pp. 271–278 (2004).
8) Liu, B., Hu, M. and Cheng, J.: Opinion Observer: Analyzing and Comparing Opinions on the Web, *Proc. of the 14th Int'l World Wide Web Conf. (WWW 2005)*, Chiba, Japan (2005). [cited 2004-10-26]. Portable Document Format. Available from: <http://www.cs.uic.edu/~liub/publications/www05-p536.pdf>.
9) 立石健二, 福島俊一, 小林のぞみ, 高橋哲朗, 藤田篤, 乾健太郎, 松本裕治: Web 文書集合からの意見情報抽出と着眼点に基づく要約生成, 情報処理学会自然言語処理研究会 (NL-163-1), pp. 1–9 (2004).
10) 佐々木千晴, 藤井敦, 石川徹也: 意思決定支援のための主観情報マイニング, 言語処理学会第 11 回年次大会 (NLP2006) (2006).
11) 奥村学, 平尾努, 難波英嗣: TSC4: 意見要約コーパスとそれを用いたワークショップ, 言語処理学会第 11 回年次大会 (NLP2005) (2005).
12) 関洋平, 江口浩二, 神門典子: 利用者の情報要求を考慮した観点に基づく複数文書要約とその評価, 情処学論:データベース, Vol. 46, No. SIG8 (TOD26), pp. 106–119 (2005).
13) Seki, Y., Eguchi, K. and Kando, N.: Multi-Document Viewpoint Summarization Focused on Facts, Opinion and Knowledge, *Computing Attitude and Affect in Text: Theory and Applications* (Shanahan, J. G., Qu, Y. and Wiebe, J.(eds.)), The Information Retrieval Series, Vol. 20, Springer, Dordrecht, The Netherlands, chapter 24, pp. 317–336 (2005).
14) Seki, Y., Eguchi, K., Kando, N. and Aono, M.: Opinion-focused Summarization and its Analysis at DUC 2006, *Proc. of the Document Understanding Conf. Wksp. 2006 (DUC 2006) at the Human Language Technology Conf. - Noarth American chapter of the Association for Computational Linguistics (HLT-NAACL 2006)*, New York Marriott, pp. 122–130 (2006).
15) Kim, S.-M. and Hovy, E.: Identifying and Analyzing Judgment Opinions, *Proc. of the Human Language Technology Conf. of the North American Chapter of the Association of Computational Linguistics (HLT-NAACL 2006)*, New York City, USA, pp. 200–207 (2006).
16) Kim, S.-M. and Hovy, E.: Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text, *Proc. of Wksp. on Sentiment and Subjectivity in Text at Proc. of the 21th Int'l Conf. on Computational Linguistics / the 44th Ann. Meeting of the Assoc. for Computational Linguistics (COLING/ACL 2006)*, Sydney, Austraria, pp. 1–8 (2006).
17) Choi, Y., Cardie, C., Riloff, E. and Patwardhan, S.: Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns, *Proc. of the 2005 Human Language Technology Conf. and Conf. on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, Vancouver, B. C. (2005).
18) Wilson, T., Wiebe, J. and Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis, *Proc. of the 2005 Human Language Technology Conf. and Conf. on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, Vancouver, B. C. (2005).
19) Wilson, T., Wiebe, J. and Hwa, R.: Recognizing Strong and Weak Opinion Clauses, *Computational Intelligence*, Vol. 22, No. 2, pp. 73–99 (2006).
20) 乾孝司, 奥村学: テキストを対象とした評価情報の分析に関する研究動向, 自然言語処理, Vol. 13, No. 3, pp. 201–241 (2006).
21) Lin, W.-H., Wilson, T., Wiebe, J. and Hauptmann, A.: Which Side are You on? Identifying Perspectives at the Document and Sentence Levels, *Proc. of the 10th Conference on Natural Language Learning (CONLL 2006)*, New York Marriott (2006).
22) Wiebe, J. M., Breck, E., Buckley, C., Cardie, C., Davis, P., Fraser, B., Litman, D., Pierce, D., Riloff, E. and Wilson, T.: MPQA: Multi-Perspective Question Answering Opinion Corpus Version 1.1 (2005). [cited 2005-8-26]. Available from: <http://nrrc.mitre.org/NRRC/02_results/mpqa.html>.

23) Cesarano, C., Picariello, A., Reforgiato, D., Sagoff, A., Subrahmanian, V. S. and Dorr, B.: Opinion Analysis in Document Databases, *Proc. of AAAI Spring Sympo. on Computational Approaches to Analysing Weblogs*, Stanford, CA, pp. 21–26 (2006).
24) Wiebe, J., Wilson, T. and Cardie, C.: Annotating Expressions of Opinions and Emotions in Language, *Language Resources and Evaluation*, Vol. 39, No. 2-3, pp. 165–210 (2005).
25) Hatzivassiloglou, V. and Wiebe, J. M.: Lists of manually and automatically identified gradable, polar, and dynamic adjectives, gzipped tar file (2000). [cited 2005-8-26]. Available from: <http://www.cs.pitt.edu/wiebe/pubs/coling00/coling00adjs.tar.gz>.
26) Stone, P. J.: The General-Inquirer [online] (2000). [cited 2005-8-26]. Available from: <http://www.wjh.harvard.edu/~inquirer/spreadsheet_guide.htm>.
27) Sekine, S.: OAK System (English Sentence Analyzer) Version 0.1 [online] (2002). [cited 2005-8-26]. Available from: <http://nlp.cs.nyu.edu/oak/>.
28) Lin, C.-Y.: ROUGE - Recall-Oriented Understudy for Gisting Evaluation - Version 1.5.5 [online] (2005). [cited 2005-8-26]. Available from: <http://www.isi.edu/~cyl/ROUGE/>.
29) Hovy, E., Lin, C.-Y., Fukumoto, J., McKeown, K. and Nenkova, A.: Basic Elements (BE) Version 1.1 [online] (2005). [cited 2005-8-26]. Available from: <http://www.isi.edu/~cyl/BE/>.
30) Harabagiu, S., Locatusu, F. and Hickl, A.: Answering Complex Questions with Random Walk Models, *Proc. of the 29th ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2006)*, Seattle, WA, pp. 220–227 (2006).