

要約システム ABISYS の精度向上とクレーム文要約への対応

荻本 輝明 佐藤 知里 田代 奈津美 松田 源立 原田 実

青山学院大学 理工学部 情報テクノロジー学科

1. はじめに

昨今の自動要約では、要約対象も幅広くなり、またより人間の要約に近い自然な要約文生成への活発な活動が行われている。ここで単一テキスト内自動要約に関するいくつかの先行研究について紹介する。

畑山ら^[4]は、「キーワード」あるいは「文生成に必要な語句」に着目した重要語句抽出を行い、抽出した語句から文生成をするという要約を行っている。また、上田ら^[5]は、要約結果を“句”の長さで表現する手法を提案し、語間に重要と思われる関係を持つ語の組を列挙することで要約を行っている。jingら^[6]は、人間の自由作成要約の分析から得られた6つの編集操作(文短縮・文の結合・構文的変形・語彙的言い替え・句の抽象化・抽出文の列挙)を用いた「切り貼り」に基づく要約手法を提案している。そして、大竹ら^[7]は、二重修飾に着目した削除を中心とした文内要約と、重要度付与による文選択の二つの部分からなる要約システムYELLOWの開発研究を行っている。これは、文内の二重修飾に関して削除するルールを独自に作成して、要約を行うものである。しかし、削除に関してのルールを作成するには限りがあり、すべての修飾に関して削除ルールを決定するのは不可能である。

以上、全ての研究において問題となっている点は、表層格で述語の格を補うものが多く、語意に依存した適切な格が抽出されていない。さらに意味解析を行っていないので、文中の語意の精度が低く、必須格の抽出誤りがあり、要約文が日本語として不適格な場合が多いことである。

原田研究室ではこれまで、日本語文を意味解析するシステムSAGE^[8]から出力される格フレーム群を使い要約文を生成するシステムABISYS2005^[9]を開発してきた。ABISYS2005は報道記事の要約においては精度の高い要約を行うことができた。しかし、ABISYS2005では要約文を生成するときに重要になる重要語候補を体言のみに限定しているため、動詞節などが重要語候補にならない。そのためクレーム文章によっては、要約文にクレーマーの意図が全く現われないこともあった。また、ABISYS2005では重要語の特徴を記した学習データを作成する作業が煩雑で、学習データへの不適切な語の混入等が生じ、結果として論説文の要約において重要語の選択に誤りがある場合があった。よって本研究では、具体的に以下の3つを行い、ABISYS2007を開発する。

1. **ABISYSのクレーム要約への対応:** 文末節に現れる発話態度を表す用言などの述語節も重要語候補とする。
2. **ABISYSの高精度化:** 述語節の種類自体を判別する文節種類得点、そして発話態度を考慮するモード種類得点を追加する。
3. **学習支援機能の追加:** 学習モデルデータの作成を支援するシステムを開発する。

2. ABISYSのクレーム要約への対応

ABISYS2005では、体言のみを重要語候補としていたが、クレーム文の要約においては適切な要約文章を出力することができなかった。図1は従来の方法でクレーム要約をした結果である。四角で囲まれている部分が要約の際に不要と思われる削除された部分であるが、クレーマーの意見が強くでていると思われる下から1~2行目の「態度が悪すぎる。もっとひどい漏電の時じゃないと通報

してはいけないみたいな印象を持った」の部分が削除されたので、クレームの要約としては不適切である。

少しの漏電だったせいか、修繕に来られた方の態度がとても悪かった。こちらとしても漏電の程度が少なかったので、気を遣いながら通報し、来てもらったのに、受付の人は親切に対応してもらったが、修繕に来られた方は「こんなんでも呼びやがって…」といった態度だった。帰された後、電気を使用したら、まだ漏電しているような気がして怖かったので、再度来てもらった。同じ人が来たが、同じ態度だった。少しの漏電で来てもらったことを気にして、対応していたのに、あまりに態度が悪すぎる。もっとひどい漏電の時じゃないと通報してはいけないみたいな印象を持った。また漏電がしたらいやだなぁと、ビクビクしています。

図1 クレームを従来の方法で要約した結果(40%要約)

このように体言中心の主題のみを追う ABISYS2005 要約では、クレームの特徴をつかみきれない。そのため、文末節に現れる発話態度を表す形態素なども文の重要性を判断する指標になると考え、本研究 ABISYS2007 では述語節も重要語候補に加えることにした。

3. ABISYS2007の処理の流れ

ABISYSの処理の流れは図2のようにになっている。入力した文章をSAGEを用いて意味解析した結果の格フレームを用いて、下記で説明するような得点を計算し、それらの得点を線形SVMを用いて総合得点とした。その総合得点の大きいものから要約率に応じて抽出し、必須格等を補い要約文章を作成する。

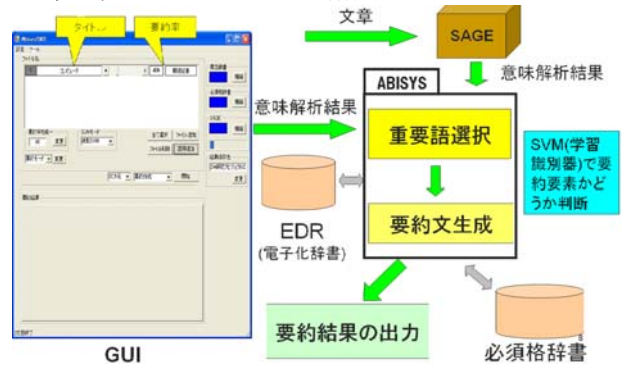


図2 ABISYSの処理の流れ

4. 重要語得点

ABISYS2005で採用していた反復、主題、位置、意見、文脈、モード得点と固有表現得点の7得点では体言を重要語候補とした主題中心の要約には対応できたが、クレーム文等では、体言のみを重要語候補としていたので有効なものを重要語候補としてとれないことが多々あった。そのため、ABISYS2007では、述語節を重要語候補に加えることにした。名詞節、動名詞節、形容名詞節、形容動名詞節、断定名詞節、動詞節、形容詞節、形容動詞節と断定節を重要語候補にするにあたり、他の属性と文節品詞間の共起性などが語の重要性に影響を与えることがあるので、品詞の種類を判別するための文節種類得点を追加した。一方、クレーム文等の要約で述語節における発話態度を表すモードにクレーマーの意図や意志が現われることが多いという事実に着目し、モード種類得点を追加した。その結果、ABISYS2005より適切な要約結果を出力できるようになった。以下に従来から採用している得点と、新しく追加した得点について記述する。

4.1. 従来から採用している得点

4.1.1. 反復得点

佐久間ら⁸⁾は、国立大の文科系学生に日本語文章を与え要約文を作成させ、原文と要約文を比較し、要約文に残りやすい語句の調査を行った。本研究ではこの調査研究を基に、文章中の反復語句に着目し、反復語句に対して加点する。反復語句の抽出では、異なる文に二度以上出現する同一語句、同義語句、同カテゴリ語句を抽出する。ただし、「ある」「こと」「もの」等の反復語句の定義で除外された語は除く。まず反復語句候補の抽出し、反復語句候補の出現数の計算を行う。また、その反復語句候補に対して固有表現がある場合は、0.5回分をさらに回数として追加する。反復回数に応じて、5回以上で55点、4.5回で50点、3~4回で40点、1.5回~2.5で30点、そして1回では25点を加点する。

4.1.2. 主題得点

一般に文章の焦点は新しい情報を表す部分とされているが、ABISYS ではそれとこだわらずに旧情報の部分であっても話者の発話の重点が置かれている語句は焦点とみなすことにした。本研究では、表層格を用いて主題・焦点を決定し、それぞれの語に対して加点した。具体的には、名詞節が名詞A+助詞「は」ならば30点、名詞A+助詞「も」、「が」、「を」、「に」、「で」では20点を名詞Aに主題・焦点得点として加点する。そしてこの名詞Aのうちで他の語句に対して「agent」「object」「a-object」「goal」「o-agent」「time」「place」の深層格をもつ名詞Bを特定しこの名詞Bにも主題・焦点得点として20点加点する。「agent」は有意志動作を引き起こす主体、「o-agent」はSAGE独自のもので、動作・変化の影響を与える主体、「object」は動作・変化の影響を受ける対象、「a-object」は属性をもつ対象、「goal」は要約の主体または対象の最後の位置、「place」は事象の成立する場所であり、「time」は事象のおこる時間をそれぞれ表す。

4.1.3. 位置得点

畑山らは、日本経済新聞において重要語が含まれる文の出現位置を調査し分析を行った。その研究結果を元に本研究では、文の最初と最後の部分には重要な語が現れると考え、報道記事では、文頭から第1文から第3文まで順に50点、25点、10点を加点する。論説文ではそれに加え文末の第1文から第3文まで順に30点、20点、5点を加点する。クレーム文では、文のどの位置に重要語があるか判断できないため、位置得点は考慮していない。

4.1.4. 意見得点

論説文、クレーム文においては、筆者の主張・意見、希望が重視されるものである。そこで意見語を特定するための意見語句(「思う」など)を調査し、その語句と深層関係がある語、さらにその語とも深層関係がある語句に対して意見得点として20点を加点する。これは意見語句から距離が2以内の深層関係において筆者の意見が集中しているためである。例えば、図3では、「思う」という見解語句に対して、2つ先の深層関係語句である、「中退は」と「今後も」に対して意見得点として20点を加点する。

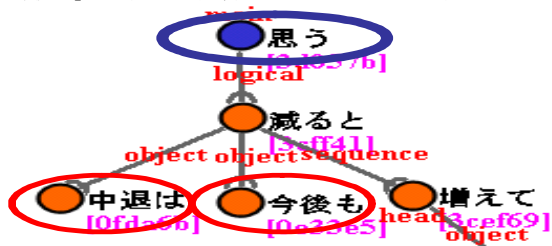


図3 意見得点の計算

4.1.5. 文脈得点

隣接する2文の間に存在する意味的な関係、つまり文間深層格を定義し、この文間深層格と要約の関係性を調査し、要約文への残りやすさに応じて、その文中の体言に文脈得点として表1のように加点した。加点する文間深層格の種類は、理由、原因、目的、詳細化、例提示、例説明、質問-応答、対比、逆接、条件、仮定、時間的順接、転換、並列と補足である。

表1 文間深層格の得点表

文間深層格	前文格		後文格	
	前文	現在文	現在文	後文
理由	5	20	20	5
原因	0	20	20	0
目的	0	20	20	0
詳細化	-5	10	10	-5
例提示	-10	30	30	-10
例説明	-10	0	0	-10
質問-応答	20	0	0	20
対比	10	0	0	10
逆接	20	3	3	20
条件	10	0	0	10
仮定	10	0	0	10
時間的順接	5	0	0	5
転換	0	5	5	0
並列	0	10	10	0
補足	0	10	10	0

4.1.6. ムード得点

ムードとは文の属性の一種であり、事態や相手に対する筆者の判断・態度を表す文法形式のことであり、SAGEによって述語節に付与される。筆者の判断・態度というものは、要約文において重要な役割をはたす。特に、「～すべきだ」(当為)や「～である」(断定)といったものは文中の要点を示す場合が多い。よって、本システムでは、SAGEの意味グラフにおいて、ムードを含む文節をルートとする部分グラフ内の体言に加点する。具体的には、「断定」に20点、「確信」に20点、「当為」に30点、「意志」に15点、「願望」に20点、「命令」に30点、「勧誘」に15点、「否定」に10点、「依頼」に15点そして「禁止」に30点を加点する。なお、このムード得点は、主題がなにかを捉える為に用いられ、後述の本年度追加したムード種類得点とは役割が異なる。

4.1.7. 固有表現得点

固有表現とは、人名、地名、組織名などの事物に特有の表現をもつ単語を表すもので、主に、固有名称が相当する。要約において、固有表現は、その特有の事物を表していることから、欠かせない情報であり、話題の中心になることが多い。よって、本研究において、重要語を選択するにあたり、固有表現の影響を及ぼすことができるように、固有表現得点として実現した。具体的には、SAGEが出力する文節域のreferentが存在する場合、固有表現であるので固有表現得点として50点を加点する。

4.2. 本年度追加した得点

4.2.1. 文節種類得点

クレーム文等では、体言だけを重要語候補としていては有効なものを重要語候補としてとれないことが多々あり、述語節を重要語候補に加えることにした。そこで語の品詞情報を得点化するための9得点を導入した。各得点は各文節が名詞節、動名詞節、形容名詞節、形容動名詞節、断定名詞節、動詞節、形容詞節、形容動詞節、断定節である場合に1、そうでなければ0として与えられる。

4.2.2. ムード種類得点

語のムードを得点化するため15得点を導入した。各得点は各文節が断定、兆候、確信、当為、伝聞、疑問、意志、願望、命令、推量、勧誘、否定、二重否定、依頼、禁止である場合に1、そうでなければ0として与えられる。前述のムード得点は主題の特徴をつかむのには有効であるが、クレーム文の要約において完全に話者の伝えたいことを捉えられないことがある。よってはじめから述

語節自体にムードを得点として付与することでより正確に発話態度を捉えることが出来るようになった。

5. 学習支援機能

ABISYS では、要約における重要語を選択するために用いる様々な得点を総合化する手段として、学習識別器サポートベクターマシン (SVM) をもちいている。そのため、文の種類ごとにどのような得点が重要であるか半断するための学習モデルデータファイルを用意する必要がある。ABISYS2005 ではこのモデルデータファイルを完成させるには手作業で重要語と非重要語を入力しなければならず、特に大量に重要語の特徴を学習させる場合には大変な作業であった。また今後、さらに要約分野を増やしていくことを考慮すると、学習支援機能が不可欠であるため、新たに学習支援ツールを作成した。ただしこの場合でもあらかじめ人間が理想的な要約を作成しておく必要がある。

本ツールでは、学習させたい文章の原文と人手による重要文抽出型の要約文の CSV ファイルを自動的に比較し、重要候補後の各得点を保持するスコアファイルや重要・非重要を示す分類済みファイルを生成することができる。つまり要約文に現われる文節が原文中のどの文節に対応するかを検出し、それらを正データに分類し、他を負データに分類する。また、各文節に対して、各得点を計算する。

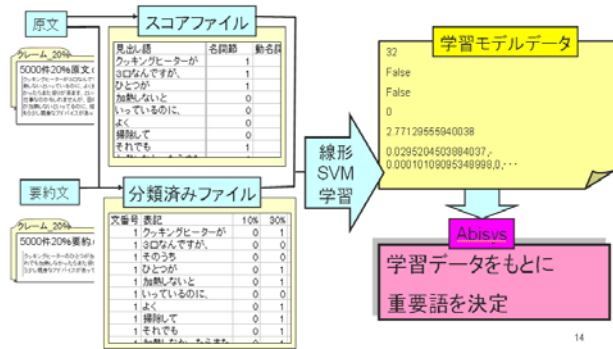


図4 学習支援ツールの処理の流れ

6. 重要語の学習

6.1. 重要語句得点の総合化

ABISYS ではパターン判別方法であるSVMを採用して重要語文節を検出する。具体的には4章で述べた9変数で定義された空間、つまり「反復得点」、「意見得点」、「位置得点」、「主題・焦点得点」、「文脈得点」、「ムード得点」、「固有表現得点」、「文節種類得点」(名詞節、動名詞節、形容名詞節、形容動名詞節、断定名詞節、動詞節、形容詞節、形容動詞節と断定節の計9次元)、「ムード種類得点」(断定、兆候、確信、当為、伝聞、疑問、意志、願望、命令、推量、勧誘、否定、二重否定、依頼と禁止の計15次元)の9得点の計31次元空間において、重要・非重要語を分離する超平面を学習する。ただし、反復得点、意見得点、位置得点、主題・焦点得点、文脈得点、ムード得点と固有表現得点は50点等の得点がついているが、文節種類得点及びムード種類得点は1か0かのみなので、あらかじめ反復得点等ABISYS2005で採用していた7得点を平均0.5分散1に正規化した。また、この7得点は体言に対して与えられているものなので、7得点計算後、対応する述語節に得点を付与することにした。例えばある文章をSAGEにて意味解析した結果の一部が図5のようになったとき、文節は、「某一流メーカー」であり、体言は「某」、「一流」、「メーカー」である。このとき、「某」、「一流」、「メーカー」の体言がムード得点としてそれぞれ20点をもつとき、その得点を文節の「某一流

メーカー」に与える。よって、この場合の「某一流メーカー」のムード得点は、60点となる。

f. 234.某一流メーカーです。..DA.237.235.□□□□.[断定]
s: 235.某,ボウ,..1090d7,FTM,JN1,..
s: 236.一流,イチリュウ,..0e49d3,FTM,JN1,..
s: 237.メーカー,..10b05c,FTM,JN1,..
s: 238.です,デス,だ,2621ba,HAN,JJD,判定詞,デス列基本形
s: 239. ..0ee33b,KUT,JSY,..

図5 SAGEでの意味解析結果

6.2. 要約文の生成

学習モデルデータを用いて線形の半判別関数を作成し、各重要語候補ごとにその値を求め、昇順にソートし、その上位のものから順に与えられた要約率を超えるまで以下の処理を行う。

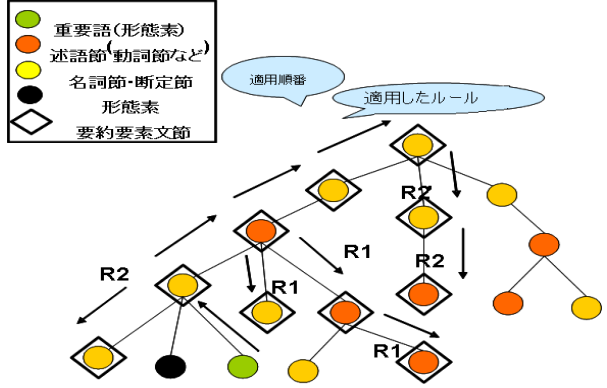


図6 要約文の生成

R1:重要語句(述語節)が要約要素語になれば、必須格を抽出する。
R2:名詞節、断定節、助詞節が要約要素語になれば、一定のルールで隣接格を出力する。

R2':名詞節が要約要素語であり、主述語であれば、体言止めの可能性があるため、その名詞をサ変名詞と考えた動詞の必須格である名詞節と連体詞節を抽出する。

補完処理:その他括弧の対応など出力に不足している部分を補完する。その後、述語節における必須各の抽出を行う。述語節における必須格の抽出とは、R1に当たる部分である。ここでいう、述語節とは、動詞節、形容詞節、形容動詞節、断定節、動名詞節、形容名詞節、形容動名詞節、断定名詞節をさす。述語節の内、断定節と断定名詞節を除く節の主辞が要約要素語になった場合、必須格の計算を行って求めた格とその他の必須格でリンクされた文節を要約要素語として抽出する。

7. 結果

7.1. 学習データの作成

学習データとして用意した事例文章を人間が作成した理想的な要約文をもとに重要・非重要語の2つに判別しておく。この学習データを分野(報道記事、論説文、クレーム文)ごと要約率(10%, 20%, 30%, 90%)ごとに生成した。これらのデータに対して線形SVMを適用し、判別式 $g(x) = wx + b > 0$ の w と b を学習する。ここで x は得点ベクトル、 w は各得点の重み、 b はthresholdである。要約率は、分野と要約率に対応して、総合得点 $g(x)$ の高い順に重要語句とする。

7.2. 本年度のABISYSの出力

ABISYS2005の主題を捉えた要約では図1のようにクレームの要約として残すべきであると思われる「態度が悪すぎる。ひどい漏電の時じゃないと通報してほしくないみたいなの印象を持った」の部分が生成されないのに対し、ABISYS2007の発話態度を捉えた要約では図6のようにクレーム要約において必要な部分を生成することができた。

少しの漏電だったせいか、修繕に来られた方の態度がとても悪かった。こちらとしても漏電の程度が少なかったため、気を遣いながら通報、来てもらったのに、受付の人は親切に対応してもらったが、修繕に来られた方は「こんなんで呼びやがって…」といった態度だった。帰られた後、電気を使用したら、まだ漏電しているような気がして怖かったので、再度来てもらった。同じ人が来たが、同じ態度だった。少しの漏電で来てもらったことを気にして、応対していたのに、あまりに態度が悪すぎる。もっとひどい漏電の時じゃないと通報してはいけないみたいな印象を持った。また漏電がしたらいやだなあとビクビクしています。

図7 クレームを本年度の方法で要約した結果(40%要約)

7.3. SVMの精度評価

表2 SVM_Classifyの精度

	報道記事		論説文		クレーム文	
	精度	再現率	精度	再現率	精度	再現率
ABISYS 2005	75%	37%	60%	12%	56%	96%
ABISYS 2007	77%	17%	64%	5%	61%	99%

本研究では、SVMモデルデータの作成のためにTinySVM⁸⁾を用いた。TinySVMには、クローズドテストを行うSVM_Classifyが用意されており、それを用いて、重要語が正しく重要ごととして選別しているか、精度と再現率を計算した結果、表2のような結果になった。報道記事と論説文では精度向上、再現率低下。クレーム文では顕著な改善が見られた。

7.4. 人間による要約結果の評価

被験者11名により、人間の作成した要約、本年度のシステムの要約、従来のシステムの要約とWordでの要約の4つの要約および原文を提示し、原文の重要な内容をどの程度要約がカバーしているか、要約の読み易さの2つの評価基準で、要約の順序付けを行った。そして、順位の平均を計算した結果が以下の表である。評価は、値が低いほうがよい要約となる。

表3 主観評価の結果

		Abisys2007	Abisys2005	Word要約	人手要約
報道記事	読みやすさ	1.99	2.38	3.99	2.25
	内容カバー	2.35	2.47	3.36	1.82
	総合	2.35	2.47	3.36	1.82
論説文	読みやすさ	2.45	1.91	3.79	1.84
	内容カバー	2.43	2.36	3.35	1.86
	総合	2.49	2.12	3.68	1.71
クレーム文	読みやすさ	2.60	3.09	2.60	1.71
	内容カバー	2.33	3.44	2.71	1.47
	総合	2.40	3.39	2.70	1.51

上記の結果より報道記事、クレーム文では読みやすさ、内容のカバー、総合のいずれの点においても文末節の発話態度を考慮したほうが体言中心の要約よりも主観評価においては平均順位が高いことがわかった。この結果よりクレーム処理におけるABISYS2007の有効性が示された。一方、論説文では、名詞節中心の主題を捉えた要約のほうがやや順位が高いことがわかった。

8. 終わりに

本研究により、ABISYSは述語節も重要語候補とすることができるようになり、また文末節に現れる発話態度も考慮できるようになった。結果としてクレーム文等の従来のABISYSでは要約が困難な文章も適切に要約できるようになった。今後は、論説文では、なぜ主題を中心とした要約がよいのかを考慮しつつ、場合によっては、ABISYS2005とABISYS2007の手法を組み合わせ、論説文では、主題中心、報道記事やクレーム文では、発話態度中心の要約を自

動的にできるようにしていくことで、より人間が使いやすい要約システムを開発することができると思われる。また本年度手を加えていない要約文の生成方法を本年度採用した手法用に見直すことで、より様々な分野において自然な要約文を出力できるようになるとと思われる。

9. 参考文献

- 1) 川口純一,青木洋,松田源立,原田実:"意味解析システムSAGEの精度向上"情報処理学会第69回全国大会論文集,1C-04,第2分冊 pp.77-78.(2007.3).
- 2) 田中信彰, 面来道彦, 野口貴, 矢後友和, 韓東力, 原田実: "意味解析を踏まえた自動要約システム ABISYS", 言語処理学会論文誌, Vol.13, No.1, pp.143-164, (2006.1).
- 3) 佐久間まゆみ: 文章構造と要約文の諸相,第3刷,くろしお出版(2000)
- 4) 畑山満美子, 松尾義博, 白井諭: 重要語句抽出による新聞記事自動要約, 自然言語処理, Vol.9, No.4, pp.55-70(2002)
- 5) 上田良寛, 岡満美子, 小山剛弘, 宮内忠信: 区表現要約手法に基づく要約システムの開発と評価, 自然言語処理, Vol.9, No.4, pp.75-96(2002).
- 6) Jing,H. and McKeown,K: Cut and Paste Based Text Summarization,In Prec. of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics, pp.178-185(2000).
- 7) 大竹清敬, 岡本大吾, 児玉充, 増山繁: 自由作成要約に対応した新聞記事要約システム YELLOW, 情報処理学会論文誌, Vol.43, No.SIG2, pp.37-47(2002).
- 8) 奈良先端科学技術大学院大学情報科学研究科松本研究室, TinySVM, <http://chasen.org/~taku/software/TinySVM/>.