

クレーム内容の自動分類

原田 実* 川又 真綱**

青山学院大学 理工学部 情報テクノロジー学科*

青山学院大学大学院 理工学研究科 理工学専攻 知能情報コース**

1. 序論

1.1. 研究背景

近年企業価値を表す指標の 1 つに CS (顧客満足度) が重要視され始めている。CS 向上の方策として、CC (コンタクトセンタ) を充実させることが、特に B to C でビジネスを展開するような企業にとっては重要な要素となっている。しかし、近年増加傾向のメール問合せに対して素早い応答を行うには、問合せ対応人員の増大など、膨大な人件費が必要である。メールによるクレームに対して人手で行われている知的作業を、自動化・支援することができれば多くの企業にとって有益である。

1.2. 先行研究の問題点

CC における作業自動化・作業支援に関する研究はすでにいくつか行われている。

岩崎ら(2005)[1]は、コールセンタでの対話データをもとに、営業日報を自動生成するための重要文抽出手法を研究している。この研究において各文は「挨拶」・「応答」・「質問」・「結論」・「返事」・「その他」の 6 クラスに分類されると定義している。しかし、分類対象は電話の対話データであり、このクラス構成をそのまま本研究のクレームクラスとすることは望ましくない。また分類は SVM を利用し、素性として文章を形態素解析した結果と語の品詞を利用しているため、語の意味や語間の関係を無視していると言う問題もある。

田村ら(2006)[2]は、複数文からなる質問の依存関係を解析してそれぞれの文の従属関係を決定する手法を提案している。ただ、問合せ文に核となる質問文を発見できないと何の処理も行われず、意見文・要求文のみで構成されるような文章に対応できない。

1.3. 研究課題

CC におけるメール問合せ処理は自動化がそれほど進んでいない領域で、クレーム文に対する分類支援や、正確な分類を自動的に行うような研究はこれまでなされていない。そこで本研究では、クレーム文の分類支援を目的としたクレーム分類システム CLASS (CLAIM Classification System) の試作を行う。

表 1 クレームクラスの定義

クラス名	定義
要求	文末節に対象への要求が表れている文
意見	文末節とその近辺に話者の主張が表れている文
あいさつ	あいさつ・感謝を表している文
質問	文末節に疑問が表れている文
評価の記述	人・物・サービスへの評価が述語節に表されている文
タイトル	名詞節・断定節で構成される文
状況の記述	話者が置かれた状況が述語節に表われている文
その他	上記7クラスに分類できない文

2. クレーム文の分類と各クラスの定義

2.1. クレーム文の特徴

クレーム文に関わらず、日本語文は客観的な事柄である命題と、書き手の主観的表現によって構成される。クレーム文の分類にお

いて、書き手の主観的表現は非常に重要な分類の手がかりとなる。また、日本語文において主観的表現は文末付近に表れる。よって、文末近辺の部分意味グラフを手がかりとして、クレーム文を後述する 8 つのクレームクラスへ分類した。

2.2. クレームクラス

各クレームクラスを表 1 のように定義する。

3. クレーム分類システム CLASS

3.1. 入力形式

本研究において試作するクレーム分類システム CLASS は、青山学院大学原田研究室で研究開発が行われている日本語意味解析システム SAGE[3]の出力である意味グラフを入力として扱う。

3.2. 出力形式

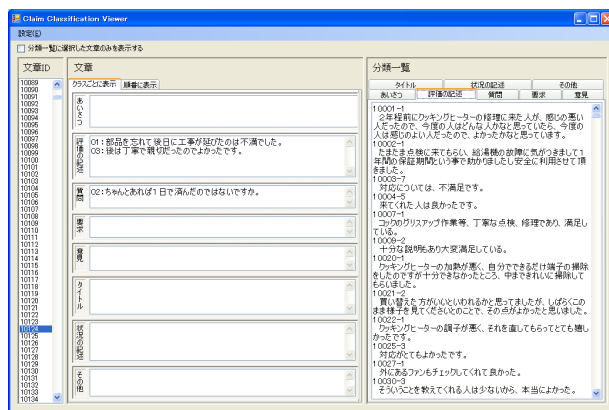


図 1 CLASS の出力

CLASS の出力は、各クレーム文を前述の 8 つのクレームクラスに分類した結果を付与した XML 形式のデータである。CLASS の出力データを、表示ツールを用いて表したのが図 1 である。左側のクレーム文章の ID を選択すると、中央部分にどのようにクレーム文が分類されたかが表示される。

3.3. CLASS における処理の流れ

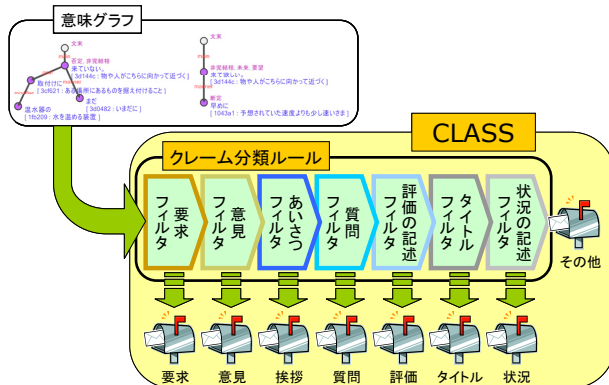


図 2 CLASS の処理の流れ

CLASS の処理の流れを図 2 に示す。CLASS は入力として SAGE の出力である意味グラフを受け取ると、各クラスへのフィルタを

通過させ、8つのクレームクラスのどれか1つに必ず分類する。分類結果は図1のように表示される。

3.4. フィルタ



図3 意味グラフ: もっと親切な対応をするべき。

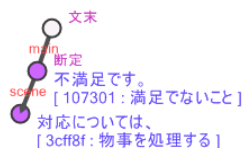


図4 意味グラフ: 対応については不満足です。

文末付近の部分意味グラフを手がかりとして、各クレームクラスに分類するフィルタを構築した。

フィルタの例を示す。例えば意見フィルタでは、図3に示するような意味グラフの場合、文末節にモード「当為」を含むので、書き手が読み手に対して意見を述べていると判断できる。よってこの文を意見クラスに分類する。

また評価の記述フィルタでは、図4に示するような意味グラフを受け取ると文末節の語の概念を参照し、上位概念に評価を表す概念を持つので、この文を評価の記述クラスに分類する。

このようなフィルタを各クレームクラスに対して設定し、クレーム文の分類を行う。本研究では8つのクレームクラスのフィルタに対し、全53の条件を設定した。

3.5. フィルタ処理による効果

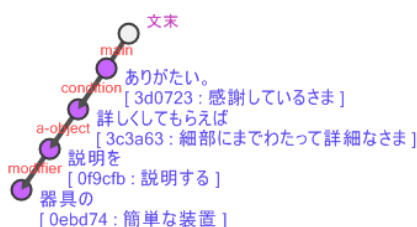


図5 意味グラフ: 器具の説明を詳しくしてもらえればありがたい。

提案手法は、SAGEを用いて意味解析を行うことと、各フィルタを通過させる順番を固定することに大きな特徴がある。

岩崎ら(2005)は文を形態素解析した結果を素性として、SVMを用いて各クラスへの分類を実現している。しかし分類精度は不十分であると言える。これは文が複数のクラスに分類されることを許容していること、文の意味情報を利用していないことが原因として考えられる。

そこで本研究では分類されるクレームクラスを1つに制限した。またクレーム文を意味解析することで、文内の語の意味やモード、語間の関係を考慮した各クレームクラスへの分類を可能にした。さらに、クレームクラスへの分類を図2に示すように順序だてて行うことで、各クレームクラスへの分類の正確性と速度を向上させることに成功した。

例えばあいさつフィルタと意見フィルタを考える。あいさつフィルタは、文末にあいさつや感謝を表す語があるとあいさつクラスに分類する、というルールを含んでいる。ここで図5のような文があるとする。もしあいさつフィルタを先に通過すると、あいさつフィルタの分類ルールは「器具の説明を詳しくしてもらいたい」という意見を暗に文が含んでいるのに、文末の「ありがたい」に反応してしまい、あいさつクラスに分類してしまう。しかしその前に意見フィルタを通過させることで、文末があいさつや評価の記述をあらわす概念であっても、条件をあらわす深層格であるcondition格のアークが文末ノードから出ていると、意見クラスに分類することが可能になる。

4. 評価実験

実際に企業に送られたクレームデータ 5272 文章の中から無作為に210文章(583文)を抽出し、人手でクレームクラスへの正解分類データを作成した。その後、CLASSが分類したデータと照合を行い、各クレームクラスへの分類精度を測定した結果、表2のようになった。

表2 各クレームクラスへの分類精度

クレームクラス	再現率	適合率	F値
要求クラス	92.7%	90.5%	91.6%
意見クラス	87.2%	90.4%	88.8%
あいさつクラス	100.0%	83.3%	90.9%
質問クラス	90.6%	100.0%	95.1%
評価の記述クラス	87.8%	91.3%	89.5%
タイトルクラス	100.0%	100.0%	100.0%
状況の記述クラス	96.6%	85.4%	90.7%
その他クラス	27.3%	57.1%	36.9%

5. 考察

フィルタを作成したその他以外の各クレームクラスへの分類精度はF値で88%を超え、本研究において作成したフィルタが妥当であったといえる。

クレーム文の分類の際、文を意味解析したことで、3.5節に示すように語の概念・文節のモード・語と語の関係などを考慮することができた。また、フィルタの通過順序を固定し、フィルタ間に依存関係を持たせたことで、ソースコードの冗長性が軽減され、実行速度が向上した。

本研究では分類のみに焦点をあてたが、今後これらの各クレームクラスへ分類された結果を用いて、自動応答などの領域への応用が可能であると思われる。また、各クラスへの分類精度を更に高め、さまざまなパターンのクレームに対しても正確な分類ができるように改良する必要がある。

6. 参考文献

- [1] 岩崎礼次郎, 荒木健治. “コールセンターの対話データを対象とした営業日報自動生成のための重要文抽出手法”, 第19回人工知能学会全国大会論文集, 2005.
- [2] 田村晃裕, 高村大也, 奥村学. “質問事項の抽出とその依存関係の特定”, 言語処理学会第12回年次大会 B2-4, 2006.
- [3] 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.