# 感情を理解する？テキストマイニング

津田和彦
ビジネス科学研究科教授

「バカ」 いきなりですが、これがあなたに向けられた言葉なら、あなたはどんな感情を発しますか？　大半の方は“不快”または“怒り”の感情を発すると思います。しかし、「今日は綺麗だよ」と言われた彼女は頬を染めながら言った。「バカ」 の「バカ」はどうでしょう？

言葉は、人間の持つコミュニケーション手段です。コミュニケーションにおいては、ただ単に内容を理解するだけではなく、その言葉の奥底にある感情を読み取ることは、非常に重要な要因です。

私の研究室では、近年実用化が進みつつあるテキストマイニングツールにおいて、言葉の持つ感情をコンピュータに理解させる研究を実施しています。ここでは、この取り組みについて紹介します。

## どう役立つの？

この研究は、上記のような恋愛小説の一説を、コンピュータに理解させることが目的ではありません。最近、インターネットの発達と共に、『お客様の声』を Web. や電子メールで受け付ける企業が多くなっています。この収集した“お客様の声”を分析することを目的として研究を実施しています。

最近、既存のテキストマイニングツールを用いて『お客様の声』分析しても、役立つ情報が得られない、という声を聞きます。ある意味、役立つ情報が得られないのは、当然なのです。

なぜなら、『お客様の声』は必ずと言っていいほど“何”を“どうする”という２つの要素で構成されています。ところが、既存のテキストマイニングツールでは、主に“何”について分析しています。けれども、分析者が第1に欲しているのは“どうする”なのです。

表１に『お客様の声』の一例を示します。

例えば、表1の2行目の声では、「ポイント」に対する意見であることが解っても何ら問題は解決しません。「増えません」という言葉から、お客様が［不満］の感情を抱いていることが解らねばならないのです。もし、お客様が[不満]の感情を抱いていることが理解できれば、素早い対応を取ることができ、顧客の不満を解消すると共に顧客満足度を向上させることができます。

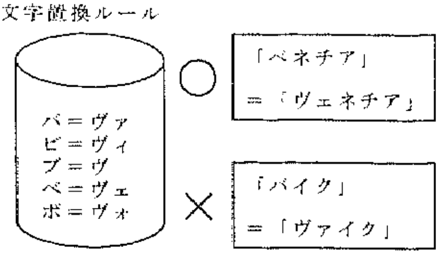


**図1　文字置換ルールの適用**

どうしてできなかったの？

これまで書いたことは、ごくごく当然のことを再確認しただけで、人間なら簡単に理解できることです。では、既存のテキストマイニングツールは、どうして、こんな当然のことが出来ないのでしょうか？理由は、“何”を理解するのに対して“どうする”を理解するのは、非常に複雑な処理が必要、ということです。“何”は、主に名詞で構成されますが、“どうする”は、動詞で構成されます。

名詞の処理は、

1.「ベネチア」と「ヴェネチア」のような文字表記の揺れへの対処
2.「コンパクトディスク」と「CD」のような同義語や類義語への対処
3.「トラック」や「バス」は「自動車」の仲間という意味的な階層関係への対処

を考慮するだけで実現できます。

といっても、最も簡単な1.への対処だけでも、図1に示すように、単純な文字置換ルールでは「ベネチア」と「ヴェネチア」には対処できますが、「バイク」を「ヴァイク」と表記すると誤りとなるため、対処できないなど、非常に困難な問題を解決しなければなりません。2.や3.への対処は、同義語・類義語辞書やシソーラス辞典などに相当する知識辞書を構築しなければならないので、非常に多大かつ面倒な作業ですが・・・

**表1　コールセンターへ寄せられた「お客様の声」のサンプル**

| お客様の声 | 何 | どうする［感情］ |
|---|---|---|
| セキュリティが心配です。 | セキュリティ | 心配［不安］ |
| 買ったのにポイントが増えません。 | ポイント | 増えない［不満］ |
| 音楽を聴くことができないのですが… | 音楽 | できない［不満］ |
| IDはどのように使うのですか？ | ID | 使うのか［疑問］ |
| 音質は満足できないこともない。 | 音質 | 満足の2重否定［？］ |

これに対して、動詞には、名詞にはない「か、き、く、く、け、け」などのような活用が存在します。この処理が複雑な上、活用語尾のあとに続く助動詞や助述表現と呼ばれる語尾の処理も非常に複雑です。例えば、表1の3行目のように文末の表記は、「が･･･」など文法に従わないものや、5行目のような2重否定表現も多々あり、非常にバラエティに富んだ表現がなされます。

この助述表現には、発話者の感情が含まれます。例えば、表1の1行目であれば、「心配です」という表現から［不安］の感情を持っていることが読み取れます。しかし、3行目や5行目では、容易に感情を理解することはできません。

図2に示すように、抽出した表現と感情のリンクを構築するのは容易な作業ではありません。このときに文頭に書いた「バカ」のようにシュチエーションによる差異も考慮すると莫大な知識ベースの構築が必要となります。

更に図2右部に示すように、テキストマイニングツールで取得した直接感情にたいして対処を行うか否かで、顧客満足度がどう変化するのか？ということまでを知識として構築できれば、誰からも「良い」と言われるテキストマイニングツールが出来るでしょう。

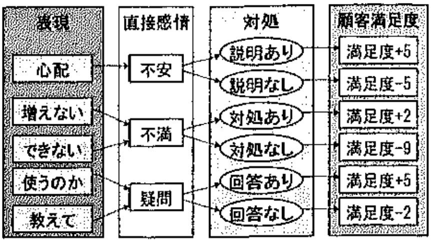

図2　動詞句および対処と感情の関係

## おわりに

言葉は文化であり、1年間に万単位で単語や表現が生まれているという説もあります。また、人間は声の強弱や高低、速度でも感情を伝えています。言葉をコンピュータに理解させることは、果てしない道程ですが、ゴールに向かって苦闘の日々をおくっています。

（つだ　かずひこ／自然言語処理、情報検索）