

中国語感性処理における感性辞書の構築

張 羽[†] 任 福繼[†] 黒岩 眞吾[†]

本稿では、中国語感性処理における感性辞書の構築について述べる。中国語情報処理の領域（情報に従って検索して、テキストのフィルタリング、自動の要約、ホームページの評価）などの技術が継続的な発展し、研究者はテキストの中の語句の表した感情属性の情報を利用することを試みて実用的なシステムの知能化のレベルを高めることを始める。そのため語句の表すプラス面、マイナス面の感情を評価する属性の特徴がますます学者達の関心を集めている。

キーワード：感情評価、イメージ値、貶義語

Construction of Emotion Dictionary in Chinese Emotion Processing

YU ZHANG,[†] FUJI REN[†] and SHINGO KUROIWA[†]

Abstract: With the constant development of different technology of the Chinese information processing field, such as text filtering, information retrieve and web pages' evaluation and so on. the researcher begins to try to make use of the words in the text that show emotion attribute information to improve the intelligent level of the practical system. So the character that the word shows negative or affirmative evaluation of emotion is paid more and more attention to by scholars. This paper describes the construction of emotion dictionary in Chinese emotion processing using the character of evaluation of words emotions.

Keywords : evaluation of word emotions, image value, commendatory word, derogatory word

1. はじめに

感情は人間の発展と日常生活で非常に重要な効果を生む。感性の情報は多くの領域ですでに応用されている。例えば、人工の心理の理論を採用した商品販売専門家システムが作られた。これは人性化の商品の設計と消費者の心理を考慮するシステムとして開発された。SONYのペット型ロボットAIBO, ロボットSDR3-Xのような感情を持つ機械は人間と機械の距離をより近くにした。

感性の機能とは何ですか。認知心理学者Frijdaによると、「感性は、外界や体内の様々な事象がその人にとって重要な目標の達成にどのように影響するかを評価し、適切な行動を準備する過程である」とされる。すなわち、感性の機能は「状況（内部状態を含む）を現在の目標（関心）との関わりという観点から評価し、適切な行動を準備すること」ということになる。

自然言語対話に関していくつかの研究が行われている。しかしながら 対話システムに対してユ

[†] 徳島大学知能情報工学研究科, Graduate School of Engineering, Department of information science and Intelligent Systems, The University of Tokushima

ユーザとの親和性の高い対話を求めるとき従来のタスク指向の対話処理の他にユーザの感情の推定やシステムの感情の表出といった感情情報処理が必要となる。中国語情報処理の領域（情報に従って検索して、テキストのフィルタリング、自動要約、ホームページの評価）などいくつかの技術が継続的な発展をし、研究者はテキストの中の語句の表す感情属性の情報を利用することを試みて実用的なシステムの知能化のレベルを高めることを始める。

2. 中国語感性情報処理の現状

感性情報処理は新しい研究分野で、画像、音声、心理学、医学など各領域で、研究が行われている。しかし、中国において、音声、画像の方面の感性について研究が多いが、感情情報処理に関する研究は萌芽段階にあり、感情に関する知識やその運用手段は明らかにされていない。自然言語処理の方面の感性についての研究が始まったばかりなので、米国と、日本の各国より行った研究はあまり多くなく、不完全な状態である。

中国語感情処理に関する研究はちょうど歩きだして、公開の資源は少ないし、利用できる資源は更に少ないである。感性情報の抽出を行うには情報が不十分である。感性辞書は感性を分析する一つの重要な要素であると考えられている。データベースの構築や中国語感性処理システムの構築することにより、感情の“快/不快”というパラメータのみでは推定できない感情についてこの辞書の構築が必要がある。そして、その状況として、中国語の感性処理の研究を行う。

まず、これらの課題の研究プロセスの中で、どのように感情を分類するのが最も重要な問題で、感情を代表する言葉がある。中国の伝統の文化を結び付けること、言語の習慣及び感情的な評価を考慮して、研究が始まっている。

本研究では考慮する感情の種類を表1に示す13種類とした。この12種類の感情は、Plutchikの多因子分析論「」による8つの基本情緒に加え、中国語の感情システムに基づいた24種類からなる感情のうちの4種類を選んだものである。しかし、感情を全面的に表さないとき、1つのセンテンスはどんな感情も持たないという問題があったので、もう1つの定義を必要とする。そして、“平静”を加えた。

快乐	悲伤	恐惧	憎恶	愤怒	惊讶	+	平
爱	期待	不安	后悔	尊敬	羞疚		静

表1：本研究で定義する感情の種類

快乐=喜び 悲伤=悲しみ 恐惧=恐れ 憎恶=嫌悪 愤怒=怒り 惊讶=驚き 爱=愛
期待=期待 不安=不安 后悔=後悔 尊敬=尊敬 羞疚=恥 平静=平静

3. 感性辞書の構築

研究によって、語句が含む感情の属性はそのセンテンスの中の共起する語句に対してとても大きい制限があって、ある文中で共起する語句はよく同一の感情の傾向を持つことを求める。例えば、“溃逃”[敗走]は中国語の貶義語である。それがセンテンスの中に含まれる時に、共起する主語は多くの場合「悪人」の属性を持つ語である。

“敌人已经大败溃逃” -----中国語

“敵は大敗して総崩れになって逃げた” -----日本語

しかし、一致しない感情の傾向を表す語が共起することもある。例えば、“摆脱”[脱却]は

貶義語ではない。しかし、通常は消極的な感情の傾向を表す語句と共に現れる。例えば、

“摆脱 危险” -----中国語
 $\frac{a}{b}$

“危険から脱却する” -----日本語

“b” 語句は消極的なマイナス面を表しているのだが、全体のセンテンスは表現するのが積極的で、プラス面の傾向を表現する。この例から“摆脱”も積極的な、肯定的なニュアンスを持つことになる。もしこれらの規則を実際のコーパスの中から抽出できれば、これらの規則に対して定量分析を行って、知識コーパスの中で融合され、モデル化すると中国語情報処理に対しても外国人の中国語教育に対しても有効であると考えられる。

本研究では、「現代汉语词典」と「常用褒贬义词语详解词典」収録されている「感情」（名詞）、「感情状態」（形容詞）「感情動作」（動詞）に分類される語に対して人手により感情パラメータを設定することにした。まず、感情パラメータの設定について説明する。「現代汉语词典」から抽出した感情属性に属している語が、本研究の12種類の感情の内どの感情に当てはまるかを調べ、最も当てはまる感情については1、全く当てはまらない感情については0と表す。これは属性が「嫌悪」となっている場合は「嫌悪」、属性が「憎悪」や「悪意」である場合は「嫌悪」とするなど、作業者が判断する。

例えば、「自責」という語については「後悔」が最も適当なので、「後悔」という感情に対しての値を与える。その他の感情に対しては全て「0」という値を与える。また、0か1の値だけではなく、各感情属性語に対してイメージ値というものを付与する。イメージ値とは「常用褒贬义词语详解词典」で定義されている文脈に依存しない感情的な評価であり、「褒義」から[貶義]の5階段に区分されたものである。表2で表す。

褒義	积极	中性	消极	貶義
++	+	0	-	--

表2：イメージ値

ここで、「褒義」=褒める意味 「积极」=積極的「中性」=中性「消极」=消極的「貶義」=けなす意味

番号	単語	褒義	积极	中性	消极	貶義	付注
		++	+	0	-	--	
1	けち					●	けちなやつ
2	まるい	●		●			円満である/かたがちまるい
3	小暗い			●		●	小暗い部屋
4	むごい					●	いたわしい/残酷な
5	ものうい					●	

表3：日本語の例

中国語の感情辞書は感情語、品詞、読み、感情種類、イメージ値、及び2ヶ月間中国の人民日報に出て来る感情語の使用頻度を記述する。

感情語の数量

喜び	112	悲しみ	102	恐れ	50
愛	230	期待	29	不安	108

嫌悪	238	怒り	249	驚き	102
後悔	38	尊敬	198	恥	26

	品詞	感情語	使用頻度	読み	静态评价(褒贬义)	感情種類	中性	积极, 消极度
	a	可恶	1	ke3wu4	贬义(品性词)	憎恨		0
	a	狞笑		ning2xiao4	贬义	憎恶		0
	a	伪善		wei3shan4	贬义	憎恶		0
	a	卑劣	2	bei1lie4	贬义(品性词)	憎恶		0
	a	疯狂	6	feng1kuang2	贬义	憎恶		0
	a	万恶		wang4e4	贬义	憎恶		0
	a	怯弱	866	qie4ruo4	贬义	憎恶		0
	a	滔天		tao1tian1	贬义	憎恶		0
	a	狡猾	498	jiao3hua2	否定, 贬义(态度词)	憎恶		0
	a	阴险	1305	yin1xian3	贬义	憎恶		0
	a	阴森	1304	yin1sen1	贬义	憎恶		0
	a	狠毒		hen3du2	否定, 贬义(态度词)	憎恶		0

表3：感情語統計の一部

4. おわりに

本研究では、自然言語処理における感情処理について述べた。特に、感情および自然言語処理における感性情報処理を紹介し、中国語感性情報処理の現状を紹介した。中国語について感情類別を分類し、中国語の感情辞書(一部のみ)を構築し感情語にイメージ値を付与した。今後の課題は感性辞書を整理し、中国語の特徴を考慮し言語コーパスを構築する。

謝辞 本研究は科学研究費補助金基盤研究B (No. 14380166)によって実施された。

参考文献

- [1].王治敏,朱学锋,俞士汶“基于现代汉语语法信息词典的词语感情评价研究”2003
- [2].许小颖,陶建华“汉语情感系统中情感划分的研究”第一届中国情感计算及智能交互学术会议论文集, p199-p205, 2003
- [3].Plutchik R., “The Multifactor-Analytic Theory of Emotion” the journal of Psychology, vol.50, pp.153-171, Jan.1960
- [4].辻 三郎“感性の科学” 1999
- [5].目良 和也,市村 匠,相沢 輝昭,山下 利之,“語の好感度に基づく自然言語発話からの情緒生起手法”,人工知能学会論文誌,17 卷,3 号,p.186-p.1952002
- [6].松本 和幸,任福継,黒岩 眞吾“感情計測システムについて”電子情報通信学会論文誌 p.55- p.60 2003
- [7].郑怀德,孟庆海“汉语形容词用法词典” 2003.12 商务印书馆
- [8].郭先珍,张伟,刘缙,王玲玲“常用褒贬义词语详解词典” 1996 商务印书馆
- [9].飛田良文・浅田秀子“現代形容詞用法辞典” 1991 東京堂出版