

確率伝搬法を用いた評価表現辞書構築

後藤拓馬¹

東京工業大学大学院総合理工学研究科 知能システム科学専攻

近年のインターネット環境の整備やブログなどの普及によって一般の個人が情報を発信する機会がかつてないほど増加してきている。この個人が発信する情報は、ある対象（サービスなど）に関するその人の評価など個人の意見が記述される場合が多い。これらの評価情報を抽出し、整理し、提示することは、評価対象の提供者である企業や、評価対象を利用する立場の一般の人々の双方にとって利点となる。このため、自然言語処理の分野では、評価情報を扱う研究が活発に行われている [1]。

テキストの評価分析に用いられる基盤要素技術に、評価表現辞書の構築が挙げられる [1]。評価表現辞書とは評価表現とその表現がもつ評価極性の対（例：良い-肯定）の集合である。この集合の要素を単語の評価（感情）極性と呼ぶ。

この評価表現辞書を構築する方法の一つとして、語彙ネットワークを利用した手法がある [2, 3]。この手法ではシソーラス²の情報を基にして、語彙をノードとする語彙ネットワークを構築する。そして、評価極性情報をネットワーク上で伝搬させることによって、語彙ネットワーク中のすべてのノード（語彙）の評価極性を求める。

高村ら [2] は語彙ネットワークを一種のスピンシステムとしてモデル化することにより、この評価極性を求める手法を提案した。この手法ではスピンの期待値を計算することで、単語の評価極性を求めている。スピンの期待値の計算にはナイーブ平均場近似が用いられている。

現在、我々は高村らによるスピンモデルを用いた手法の精度的改善を目標として、確率伝搬法を用いた評価表現辞書の構築を試みている。

本講演では、まず高村らによって提案された評価表現辞書構築法を紹介する。その後、我々が試みている確率伝搬法を用いた手法の結果を報告する。

参考文献

- [1] 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”, 自然言語処理学会論文誌, 2006, 13(3), 140-210
- [2] 高村大也, 乾孝司, 奥村学, ” スピンモデルを用いた感情極性抽出”, 情報処理学会論文誌, 2006, 47(2), 627-637
- [3] Jaap Kamps, Maarten Marx, Robert J. Mokken, and Maarten de Rijke, 2004, ”Using wordnet to measure semantic orientation of adjectives”, *In Proceedings of the 4th International Conference on Language Resource and Evaluation (LREC 2004)*, 4, 1115-1118

¹E-mail: takuma@sp.dis.titech.ac.jp

²単語の上位/下位関係, 部分/全体関係, 同義関係, 類義関係などによって単語を分類し, 体系づけた辞書