

特 集

◇自然言語処理 (情報検索・抽出・分類)

Affect Analysis of Textual Input Utterance in Japanese and Its Application in Human-Computer Interaction

Michal Ptaszynski (ミハウ プタシンスキ)

ptaszynski@ieee.org

北海道大学大学院情報科学研究科

指導教員: 荒木 健治

博士学位 (情報科学), 2010 年 9 月 24 日

取得



キーワード: affect analysis, emotions, affective computing, affect sensing from text, emotional intelligence, emoticons, contextual appropriateness of emotion, conversational agents.

概要: In this dissertation I present my research on enhancing machines with Emotional Intelligence. I develop a set of affect analysis tools and propose methods for efficient utilization of emotive information obtained by these tools. The first system developed is ML-Ask, a system for affect analysis of textual input utterance in Japanese. ML-Ask first separates emotive utterances from non-emotive and in the emotive utterances seeks for expressions of specific emotion types. The second system, CAO, is a system for analysis of emoticons. CAO extracts emoticons from input and determines the specific emotion types they express. Firstly, it matches the extracted emoticons to a predetermined emoticon database containing over ten thousand emoticon samples extracted from the Web and annotated automatically. The emoticons difficult to deal with this way are automatically divided into semantic areas representing "mouths" or "eyes" and matched to a semantic area database created and annotated automatically. The above systems are then utilized in two methods for enhancing of Human-Computer Interaction. The first is a method for automatic evaluation of conversational agents. In this method, the affect analysis systems are used to analyze users' emotional engagement during conversation. This data is reinterpreted to specify general attitudes to the conversational agent and its performance. The method revealed similar tendencies to a standard usability questionnaire, proving its applicability in automatic evaluation of Japanese-speaking conversational agents. Finally, I present a method for determining whether emotions expressed by speaker are appropriate for the context of the conversation. In this method, affect analysis systems estimate the speaker's affective states and a Web mining technique gathers from the Internet emotive associations consisting of a list of emotions that should be expressed at the moment. Implementing the above methods to a conversational agent could allow it choose appropriate conversational procedures and therefore enhance human-computer interaction. I conclude the dissertation with a discussion on possible applications for the proposed systems and methods, and further work needed for implementation of the full scope of Emotional Intelligence in machines.

公表論文: Ptaszynski, M., Maciejewski, J., Dybala, P., Rzepka, R. and Araki, K.: CAO: A fully automatic emoticon analysis system, *Proc. 24th AAAI Conf. on Artificial Intelligence (AAAI-10)*, pp. 1026-1032, Atlanta, Georgia, USA (July 11 - 15, 2010)

現職: 日本学術振興会外国人特別研究員@北海学園大学

論文入手先: http://arakilab.media.eng.hokudai.ac.jp/~ptaszynski/data/2010_Doctoral_dissertation_Ptaszynski.pdf

抱負: 近頃注目を浴び始めている感情処理研究を出発点としてユーザ感情を理解し適宜反応ができる対話エージェントを開発したいと思います。そのためには大規模データを基にした会話中感情モデルなど、これまで挑戦されなかった研究が必要となります。研究を長い目で見てそれを試みたいと思います。

◇自然言語処理 (要約)

携帯端末への情報提示を指向したテキスト要約に関する研究

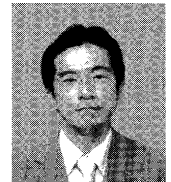
長谷川 隆明

hasegawa.takaaki@lab.ntt.co.jp

東京工業大学大学院総合理工学研究科

指導教員: 奥村 学

博士 (工学), 2010 年 3 月 26 日 取得



キーワード: テキスト要約, 重要文抽出, 文短縮, 関係抽出.

概要: インターネットに接続可能な携帯端末の普及により、いつでもどこでもネットワーク上にある情報にアクセスすることが可能になった。しかしながら、携帯端末の画面は小さいために一度に表示できる文字数も限られるため、アクセスの対象となる情報について大意を保持しつつ短く要約する必要がある。本論文では、要約の対象、要約の利用目的、要約の長さという観点から、携帯端末向けの情報提示を指向したテキスト要約について論じる。電子メールを対象とし携帯端末への着信通知として利用するときの 500 Byte 以内の要約を作成する方法、および Web ページを対象とし携帯端末における検索結果の提示として利用するときの 70 Byte 以内の概要文を生成する方法を提案し、いずれも従来の方法よりも良い精度が得られることを示す。また、携帯端末向けの短い要約を生成するために情報抽出の結果を用いてアブストラクトを生成するアプローチがある。情報抽出のためのテンプレートを獲得するには高いコストがかかるが、これを低コストで獲得する方法を提案する。大規模なテキストから教師なし学習により、テンプレートに必要な情報の一つである関係を抽出するタスクで評価実験を行い、提案手法の有効性を示す。

公表論文: 長谷川隆明, 西川 仁, 今村賢治, 菊井玄一郎, 奥村 学: 携帯端末のための Web ページからの概要文生成, *人工知能学会論文誌*, Vol. 25, No. 1, pp. 133-143 (2010)

Hasegawa, T., Sekine, S. and Grishman, R.: Discovering relations among named entities from large corpora, *Proc. 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004)*, pp. 415-422 (2004)

現職: 日本電信電話株式会社 NTT サイバースペース研究所
論文入手先: 東京工業大学附属図書館, <http://www.libra.titech.ac.jp/welcome.php>

抱負: 社会に役に立つアプリケーションを提供するために、テキスト要約や情報抽出の研究開発を行ってきました。今後は学位取得で得られた知見を活かして、実用的かつ本質的な自然言語処理の研究開発に取り組んでいきます。