Blog Text information extraction Natural language processing Opinion search

Sentiment Analyzer (HYOBAN Cloud)

NTT Cyber Space Laboratories

On the Internet nowadays, word-of-mouth information, so-called "opinion information", written on CGM*¹ (such as blogs and bulletin boards) is attracting much attention as a factor that strongly influences consumer behavior. As for conventional blog searches and web searches, however, it is difficult to retrieve opinion information such as "favorite", "convenient", and "pretty" from a huge number of search results. This opinion-information indexing technology—referred to as "Sentiment Analyzer"—was developed for easily and precisely accessing such opinion information posted on CGM. By performing analysis of text and extraction of opinion information by referencing a large-scale semantic dictionary developed by NTT Laboratories, this technology can speedily and precisely extract and display opinion information related to arbitrary keywords (such as product and shop names).

"HYOBAN Cloud" trial using this opinion-information indexing technology has been performed since October 2006 on NTT Group's comprehensive portal site "goo". "HYOBAN Cloud" provides a fresh awareness to users by connecting opinions expressed on the latest blogs to web-search results and presenting the resulting connected information. Furthermore, presenting the information in the form of a "tag cloud*2" enables the "voices" of many people to be heard in a consolidated form and opinion information to be efficiently found.

From now onwards, through feedback of practicality-evaluation results obtained with "HYOBAN Cloud", we will continue our investigations on commercialization of this opinion-information indexing technology as an "opinion-search service" on "goo" and on its expansion into the corporate marketing business.

- *1 CGM: Consumer Generated Media
- *2 tag cloud: Tags concerning various items are listed on web sites using tags in the form of "social bookmarks". In response to the utilization frequency of tags, etc. of tags, tag characters are often displayed in different sizes.

Opinion search trial (HYOBAN Cloud)

