

CB Summarizer の'言語的'オプション機能

野村直之

(株)ジャストシステム
自然言語製品開発本部

文書要約・構造化製品 CB Summarizer は、複数文書を入力として、一つなぎの要約結果を、3 フレーム HTML 文書の形で生成する。フレーム間は省略箇所の境界でリンク付けし、indicative な要約と informative な要約との融合をはかっている。

http://www.justsystem.co.jp/cb/product/index_sum.html

文書レイアウト解析、言語判定、形態素解析、日本語構文解析、陳述と接続関係の制約を抽出する談話解析、さらに複合語の出現頻度や談話キュー等に基づくフレーズ単位の重要度判定、そして指示語の先行詞への関連付け制約等を適用しつつ、総合的に重要度の判定を行い、要約結果を生成する。フレーズ単位の重要度判定の際に、陳述要素の与える正負の重み係数を加減することにより、事実の記述を優先して要約するか意見の記述を優先するかの度合いを制御する。この度合いは、詳細設定画面上のスライダーを模した GUI で指定することができる。

A 'Linguistic' Option Control for CB Summarizer

Naoyuki NOMURA

NLP Products Development Department

Justsystem Corporation

nomuran@justsystem.co.jp

CB Summarizer is a text and HTML document summarizer, which automatically generates a documents index HTML frame, a unified summary frame, a combined original documents frame, and thousands of hyperlinks between these frames. Thus, it gives the unified functionality of 'indicative' and 'informative' summaries. This demonstration highlights CB Summarizer's ability to control the output summary by a GUI 'slide bar,' which lets the user specify the weight between 'opinion description biased summary' and 'fact description biased summary.'

【事実優先で要約】 v s 【意見優先で要約】 の対比例：

パソコンで日本語を自由に使えるようにした日本語変換システム「ATOK」は、日本語に強いジャストシステムの礎を築きました。その「ATOK」を搭載した「一太郎」は日本語ワードプロセッサのスタンダードとして多くのお客様にご活用いただいております。「一太郎」以外にも、「花子」、「三四郎」といったアプリケーションソフトにより、オフィス業務一般をデジタルで表現する手段をご提供してまいりました。こうして、ジャストシステムは、パソコンの歴史と共に歩んでまいりました。今日、オフィスの道具は、かまびらかなパソコンがコミュニケーション・メディアとして進化しつつあります。この進化は、ネットワーク時代、グループウェア時代に入った今日、未来へ向かって、さらに力強くなるようとしているのであります。ジャストシステムも、OM、JOSS、CB Server と、コミュニケーションを強力に支援するアプリケーション、インフラを提供しております。