

要約支援システムCOGITOの文脈表現形式

安原 宏 小松 英二 加藤 安彦
(沖電気工業株式会社 総合システム研究所)

1. はじめに

我々は、知的対話機能実現のための一環として、文脈処理及び要約処理の研究を目的として要約支援システムCOGITOの開発を行っており、特に文脈処理に重点を置いて進めている。そこでは、テキストは、Prolog形式の中間言語に展開されたものを用いる。中間言語のテキストは、文脈処理を行うことにより、照応付けられ、一貫性を持ち、構造化されていく。従来、文に対して、文法カテゴリが定義されているが、我々はテキストに対して構造化表現を与えることにより、文脈処理のゴールを、より明確なものに設定した。以下ハイパー・ロジックと名付けるものは、この文脈表現形式を表わす。図1に示すようにハイパー・テキストとは、構造化、ブラウジング機能等において類似の概念であるが、ロジックプログラミングの枠組みでテキスト表現を行うことが特徴である。本稿では、ハイパー・ロジックの定義及び利用方法について述べる。

2. ハイパー・ロジック

文脈解析後の論理表現は様々な文脈的關係を含んでおり、構造化がおこなわれている。ハイパー・ロジックは、文脈処理の結果得られた意味表現をテキストの構造化の立場から整理し強化した意味表現である。図2にハイパー・ロジックの表現を示す。また図3にハイパー・ロジックの表現要素を示す。

COGITOにおいては、各關係が入力データになっており、それを用いて照応などの文脈処理により、照応關係、文屬性、一貫性關係を抽出する。さらにそれらを用いて、パラグラフ關係及び要約の抽出を行う。述語論理で表現しているため、テキストの検索、合成などが同じロジックプログラムで処理できるメリットがある。テキスト管理手法として話題になっているハイパー・テキストの特徴であるブラウジングに対して、ハイパー・ロジックではキーワードレベルの検索から、パラグラフレベルの主題に至るまで、多様な機能を持たせており、推論を伴う知的な検索が可能である。

本研究は第5世代コンピュータプロジェクトの一環としてICOTからの委託で行われたものである。

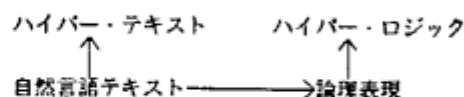


図1 テキストの構造化

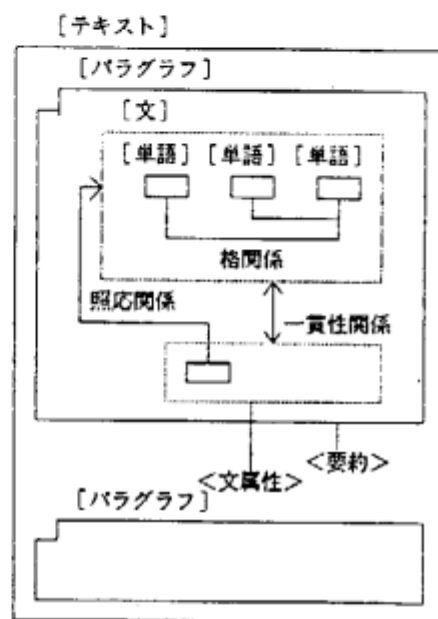


図2 ハイパー・ロジックの表現

格關係	: 中間言語として生成される。
照応關係	: 指示語、省略、名詞同士の關係。
文屬性	: 文の機能的な分類など。
一貫性關係	: 文同士の關係。
パラグラフ關係	: パラグラフ同士の關係。
要約	: パラグラフ毎の要約、パラグラフの中間言語のサブセットとして保持する。

図3 ハイパー・ロジックの表現要素

3. おわりに

テキストの構造化手法としてハイパー・ロジックを導入した。テキスト処理をロジックプログラムの枠内で実現することにより、テキスト全体が、宣言的かつ手続き的に扱えるようになり、その結果、テキスト処理の応用範囲が広がった。

また文脈処理の観点から見ると、ハイパー・ロジックは、論理文法の自然な拡張を示唆しており、認知科学的にも興味ある研究課題である。