

概念体系作成支援システムのデータ構造

2M-4

安達久博 橋本秀樹 熊野 明 泉 佑二 天野真家

株式会社日本電子化辞書研究所 株式会社東芝ソフトウェアエンジニアリング

1. はじめに

株式会社日本電子化辞書研究所(EDR)で開発している電子化辞書は、言葉に依存する単語辞書と言葉に隠れた常識を知識ベースの形で表した概念辞書から成る。単語辞書では、語が持つ語義を細分化し、細分化された語義を概念の構成単位である概念素に対応させる。概念辞書の中で概念体系は、単語辞書で設定された概念素間を上位下位関係に従って構造化する。この上位下位関係は、概念の持つ属性の継承関係を保証する。また、概念記述のコンパクト化のために用いられる。[1]

本稿では、上述の概念体系の作成を効率良く支援するシステムのデータ構造について、実現例を提示し、考察を行う。[2],[3]

2. データ構造とその設計方針

概念とその相互関係は、概念体系を形成しこの概念体系は色々な形の図式(例えば、分類表、系統樹等)をもって表現することができる。また概念には、論理的关系と存在的関係があり、双方とも色々な場合同じ対象について使われるため、同時に概念レコードに論理的と存在的概念の記述がoccurする。[4] これが多重継承の問題の要因である。

このため、概念体系をネットワークで表現するか、木構造で表現するか問題である。我々は、概念に対する見方の違いによる上位概念の多重リンクを明確に観点で区別し、多重リンクの発生した時点で概念ノードを分化させ木構造で体系データを表現する。(図1参照) これにより、データ構造の単純化と概念の相互関係の観点の明確化が実現できる。

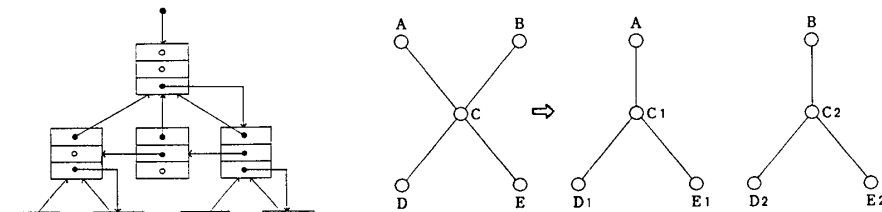


図1. 木構造への変換

4. おわりに

概念体系のデータ構造を木構造で表現することにより概念ノードのリンクの付替え等の編集機能制御の簡単化と、概念体系上でのノードと属性情報の管理が効率良く行える等の利点がある。今後は、高速化のためのデータ構造について検討を行う予定である。

【参考文献】

[1] 内田他：自然言語処理のための電子化辞書の構成法，第35回情報処理学会全国大会
 [2] 泉 他：UNIX環境を利用した概念体系作成支援ツール，第36回情報処理学会全国大会
 [3] 井原他：概念の体系化に関する考察，第36回情報処理学会全国大会
 [4] 尾関周二，Christian GALINSKI：ターミノロジー学，文理閣 1987.6