

この様に、用語は多重階層で構造化され、更に同義語表現を多く持ち、それ故、多義性も多く見受けられることが分かる。しかしながら、SS-KWEIC法の仕様から抽出される階層関係は階層関係の中でも包含関係に位置することが分かる。

4 実際の文からの意味関係の解析

実際の文から意味関係を人手で解析し、それを概念構造として表現することを試みた。また、文は複数に渡るものでなく、「。」で区切られる一つの単文を使用し、解析を行った。

例文1: SEがプログラムを実装する。

この文の意味の核となるものは、「SE」、「プログラム」、「実装する」である。これらを概念構造として表現すると、以下の図.2の様になる。楕円で示される物が用語で、四角で示される物が関係に当たる。

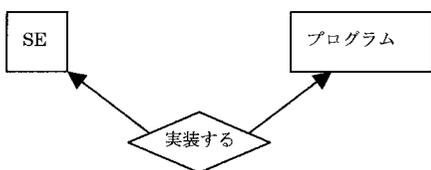


図.2 文からの意味関係1

例文2: 知識情報処理システムに「常識による推論」や「不完全な知識からの推論」などのより高度な情報処理能力を与えようとする試みがある。

この文の意味の核となるものは、「知識情報処理システム」、「情報処理能力」、「与える」であり、更に「情報処理能力」とは「常識による推論」、「不完全な知識からの推論」の事である。よって、概念構造を表現すると、以下の図.3の様になる。

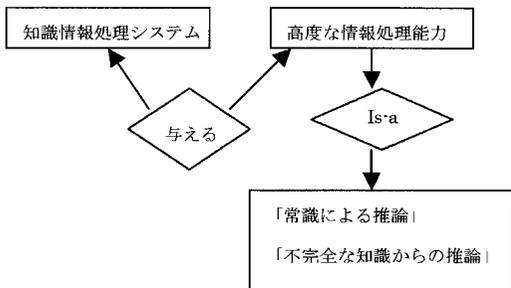


図.3 文からの意味関係2

5 考察

今回、概念間に存在する意味関係を忠実に捉え、それを構造として表現することの利点や構造化のための概念の表現形式について述べた。

各種意味関係抽出手法により用語を概念と見立て階層関係、同値関係などの用語の特性を利用して、情報の意味を概念間の構造として表現することが有用であると示すことができた。

実際の文章を参考に意味関係を見てきたが、単純な二項関係となるような文章だけではなく、図.3に示されるように、「高度な情報処理能力」には入れ子構造として is-a 関係で結ばれる「常識による推論」、「不完全な知識からの推論」などがあり、これらは更に多項関係と言える。この様に情報には様々な意味関係があり、これらを表現するには多項関係や入れ子構造を支持することが可能な情報構造が必要となることがわかった。

7 まとめ

今回実際の文章から見た情報の意味内容の特性から、これらを全て支持できる概念構造を作成することが可能となれば、類推、帰納推論、仮説生成などの高度な処理が可能となると考えられ、更に情報構造はグラフの二項関係を多項関係に拡張したハイパーグラフをさらに関係の相対性、入れ子関係などに対応できるように拡張した、均質化二部グラフモデルの実装中である。

謝辞

本研究を行うにあたって、国立情報学研究所で作成されたNII-NACSISコレクションのNTCIR1を使用いたしました。深く感謝いたします。

参考文献

- [1] 日本語形態素解析システム JUMAN:
<http://pine.kuee.kyoto-u.ac.jp/nl-resource/juman.html>
- [2] NACSIS テストコレクション:
<http://research.nii.ac.jp/ntcir/index-ja.html>
- [3] 藤原譲, 情報学基礎論の現状と展望-学習・思考機構と超脳計算機への応用-, 情報知識学会誌, Vol.1, 9, No. 1, pp13-29, 1998.
- [4] 藤原譲, 安江虹: 意味理解のための情報構造モデルとその応用, 情報知識学会第4回研究報告会, 1966
- [5] 平野朋成, 藤原譲: 概念構造のためのモデル, 情報知識学会, 第3回研究報告会, 1995