

視覚化を利用した知識ベース管理システムの設計と実装

重松伴典[†] 平石広典[†] 溝口文雄[†]

東京理科大学 理工学部 [†]

1. はじめに

エキスパートシステムは専門家の知識を格納した知識ベースを元に推論し、一般人にも専門家のような判断が可能となるシステムである。

エキスパートシステムは知識ベースと推論エンジンから構成され、知識ベースは専門家の知識を格納したデータベースである。知識の表現にルール形式を用いた知識ベースでは、知識ベースは質問と仮説と「if (条件) then (適用する内容)」という形式のルールとで構築される。質問は専門的判断を下す根拠を得るための要素である。仮説は仮説名と確信度から成る。ルールは専門家の経験的な知識を表現する。推論エンジンは知識ベースを用いて推論し、推論結果の仮説と確信度を結論として提示する。

このような知識ベースでは、個々の要素は理解しやすい。しかし、知識ベースの各要素が推論の中でどう用いられるかを、データを直接読んで理解することは困難である。

そこで本研究では、知識ベースのある結論に対して、確信度がどの質問によりどう推移するかを三次元視覚化することにより解決する。これにより、ある結論の確信度は、推論の中でどう推移するかを視覚的に理解できる。そして、知識ベースの確信度の推移に関する特徴を把握することが可能となる。したがって本研究では、このような知識ベースの視覚化システムの設計と実装を行う。

2. 関連研究

本研究に関連する研究には SemNet[1]がある。SemNet は知識ベースの検索を行うために、知識ベースを三次元で視覚化したもので、知識ベースの要素をノードとして配置している。しかし、視覚化した図からノードの座標値以外の意味を読み取れないという問題点がある。

3. 設計

本研究で三次元視覚化をする際に解決すべき問題点は、知識ベースの要素をただ視覚化しただけでは、任意のノードの集合がどんな意味を持つのかを理解できない点である。つまり、ノードの集合が何か有益な意味を持つように、座標系を設計する必要がある。

そのために、本研究では、ある結論に関して知識ベースを視覚化し、質問の順番や確信度の推移を読み取れるような座標系を設計する。

最初に、確信度は任意とした、ある結論を導く全質問パターンを網羅的推論により求める。

次に、座標軸について考える。求めた全質問パターンから、ある質問の「質問頻度の低さ」を x 軸、ある質問が平均何番目かという「平均ステップ数の高さ」を y 軸、ある質問をされる時点の「確信度の高さ」を z 軸で表示する。

そして、ノードについて考える。ノードについては図 1 のように、まず質問をノードとして表示する。また、質問の選択肢が確信度をどの位変化させるかという影響力をノードの大きさとする。そして、質問順序に基づいてノードをリンクする。図 1 ではノード A の質問の次にはノード B の質問が聞かれることを意味している。最後の結論を導く質問ノードに関しては、その座標で確信度を表すノードをリンクする。

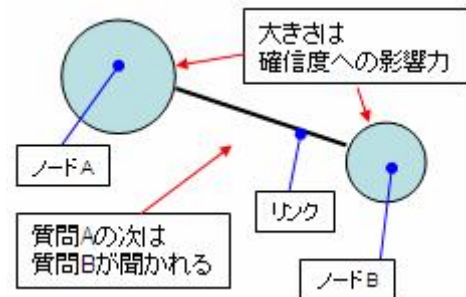


図 1: ノードの表示とその関係

このように設定した座標系にノードを表示すると、まず座標から座標値がわかる。次にノード自体の大きさから、確信度への影響力がわかる。そして、最初のノードから、結論を導くノードへの経路を見ると質問毎の確信度の推移が

Design and mounting of knowledge base management system using visualization

[†]Tomonori Shigematsu, Hironori Hiraishi, Fumio Mizoguchi, Faculty of Sci. and Tech., Tokyo University of Science

わかり、その特徴を把握できる。最後に、全体のノード分布を見ると、座標空間におけるノード密度の粗密がわかる。これにより、確信度を細かく計算している箇所、いない箇所がわかる。つまり、全体のノード分布から知識ベース毎に異なる、確信度計算の特徴が把握できる。

以上のことから、本設計で三次元視覚化を行うと、ある結論について、確信度の推移を視覚的に把握できる。よって知識ベースの確信度の推移に関する特徴を把握できる。そして、全体のノード分布の粗密から、知識ベースの確信度計算における特徴を把握することが可能となる。

4. 実装

本研究で用いるエキスパートシステムでは、知識ベースは XML 形式で記述され、ネイティブ XML データベース eXist[2]で管理している。推論エンジンは XMLQuery を用いて必要に応じて情報を取り出すことにより推論を行う。

システム構成は図 2 の通りである。本システムは知識ベースに対し網羅的な推論を行う部分と、結果を視覚化する部分とで構成されている。

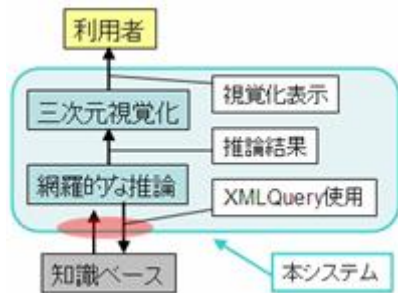


図 2: 本システムの構成

網羅的な推論部分は java で実装を行った。深さ優先の推論を行い、ある質問のある選択肢からは結論が導かれ、次の質問が存在しない場合には、別の選択肢で解答する。このような再帰的な推論を繰り返し、網羅的な推論を行う。

視覚化表示部分は java3D で実装を行った。まず、表示するノードの形状の決定と色分けを行った。ノードは球で表現し、配色については、結論を導く質問ノードを赤色、結論の確信度を示すノードを黄色、それ以外の質問ノードを白色とした。よって質問ノードは白と赤で、結論の確信度ノードは黄色で表現される。次に、質問の確信度への影響力を表すノードの大きさは半透明な球により表現した。そして、色分けした質問ノードの球と半透明な球を三次元空間内に配置し、質問順序通りにリンクした。

5. 結果と考察

本システムにより、知識ベースの結論一つに関して視覚化を行った結果が図 3 である。

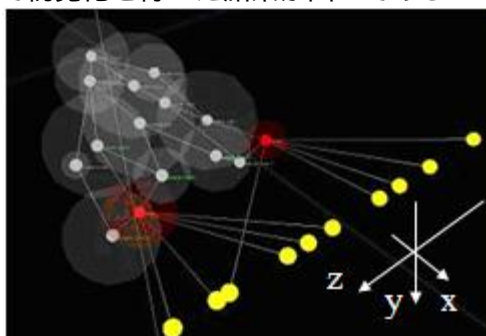


図 3: 視覚化表示図

まずノードの座標からその座標値がわかる。

次に、ノードの経路を見ると、確信度の推移がわかる。この例では結論を出す赤ノードは二つある。その先の z 軸と平行に並んだ黄ノードの配置を見ると、赤ノードの片方は比較的高い確信度の結論を導き、もう片方は比較的低い確信度の結論を導いていることが読み取れる。

そして、全体の質問ノードの分布を見ると、白と赤のノードはほぼ一部分にあり、黄のノードは z 軸方向に大きなムラ無く分布している事から、ノード密度の極端な偏りは無いと言える。

よって、この知識ベースは、この結論を導く推論過程に関しては、極端な質問や確信度の偏りが無いという特徴を持つと言える。

6. おわりに

本研究では、エキスパートシステムの知識ベースに関して、ある結論を導く質問の流れを視覚化表示するシステムの設計と実装を行った。

本システムを用いた視覚化によって、ある結論を導く質問の流れを視覚的に把握できる。また、全体のノードの粗密から、知識ベースの確信度計算の特徴を把握することが可能である。

これらの情報は、構築済みの知識ベースを更新する際に利用できると考えられる。

参考文献

- [SemNet, 1988]K. M. Fairchild, S. E. Poltrock, and G. W. Furnas, "SemNet: Threedimensional graphic representation of large knowledge bases", In R. Guindon, pp. 201-233. Lawrence Erlbaum Associates, 1988
- [eXist, 2005] eXist, Open Source Native XML Database, <http://exist.sourceforge.net/api/>