

特集 「編集委員 2007年の抱負」

知識発見から知識体系発見へ

鷺尾 隆 大阪大学産業科学研究所



知識を発見しっぱなしでよいのか？

人工知能やその周辺でデータマイニングや知識発見の研究が盛んである。この研究に携わる者としては、大変喜ばしい。しかし、人工知能を探求する者としては、これだけでよいのか？ という思いに至っている。

一昔前の人工知能研究では、知識の学習や獲得と並行して、知識でいかに推論するか、知的活動をいかに実現するか、盛んに研究が行われていた。一方、現在は知識の発見や表現方法に研究重点が置かれ、それらを利用した知的活動を実現する枠組みの研究がどこか置き忘れている。確かに、過去のエキスパートシステムや常識推論は、少ない知識をもとに過大な知的な処理を追い求めていたきらいがある。また、推論機構も膨大な知識を扱うには計算量的に非力であった。しかし、これだけ豊富に電子データが入手でき、そこから膨大な知識を高速に導出できる今日、得た知識をもとにいかに計算機に知的処理をさせるか、今一度真剣に探求すべき状況ではないか。

知識をいかに使える形に構成するか？

データマイニングや機械学習は、膨大なデータからその中の事例に対応づけられた、何らかの基準で特徴的パターンを導出する。我々人間の頭脳にも実世界の経験と対応づいた種々膨大なパターンが詰まっていて、計算機が見つけたパターンが人間のもつものと大きく矛盾なく結合できて推論に使用できる場合に、そのパターンは知識となる。つまり今の知識発見手法の多くは、知識であることの判断や利用を人間に委ねている。計算機自身が膨大なデータから見つけた膨大なパターンを知識として知的処理を行うには、まず人間と同様に実データ中の事例との対応を維持しながらバラバラなパターンを相互に関連づけ、まとまりある知識体系を構成しなければならない。知識体系を記述する器はオントロジーやXML、数理関数など種々研究されているが、そのうえで膨大なパターンを結合して推論や計算が可能ないように計算機実装する必要がある。このような知識体系は使用する文脈に応じて、実世界モデルや背景知識、常識知識などと呼ばれる。これにより、計算機自らがデータの事例に根ざし、対象の幅広い性質に関する推論やシミュレーション、予測・推定など知的処理を行えるはずである。

知識体系の構成方法としては、データから見つけた特徴的パターンをたくさんため込み、それらの整合な結合関係を探索する方法が考えられる。しかし、従来の人工

知能研究の経験では、このような探索は膨大なパターンの組合せ爆発に直面する可能性大である。それよりも、個々のパターンがデータの事例にひもづいているので、データをなぞってパターンを結合したり、データからパターンとそれらの相互関係も一緒に発見する方法論が現実的である。またこれにより現実にもしめない形式的結合を避けることもできる。このように、知識発見やデータマイニングを超えて、膨大なデータからの知識体系発見へと研究を進めるならば、人工知能を含む計算機科学はその地平を一步広げるのではないか。

膨大な知識体系をいかに利用するか？

仮に膨大かつまとまりのある知識体系をデータから得たなら、さらにその上で効率的な推論や計算を行う必要がある。しかし、知識体系が膨大なため、従来の推論機構で現実的処理を行うには困難を伴うであろう。これを避ける新たな推論機構の研究も課題となる。

この克服にはいくつかの方向が考えられる。一つは、推論に必要な知識を容易に探索する知識体系の構造化である。近年のデータベースやデータ構造研究において、多様な知識やデータ形式の効率的アクセスを可能にする手法が発展しており、この方向には実現性がある。もう一つは、自己の知識体系の構成を言及する地図を準備する方法である。例えば、検索エンジンのように自己の知識体系の逆引きインデキシングを行い、推論に必要な知識の分野や内容に応じ、関連深い知識やそれらの関係を知識体系から取り出す方法が考えられる。

最後に問題となるのは、膨大な知識体系のどの部分までを、推論の各段階で当該の知的処理に必要な探索範囲とするかである。これは、人工知能において昔から言われているフレーム問題の一種である。しかし、この問題についても、ページランキングに代表されるように、近年の知識や情報の関連性に関する評価方法が役立つのではないだろうか。

過去10年以上にわたって、種々の知識発見、データマイニング、情報検索、オントロジーなどの新しい手法が蓄積されてきた。以上に述べたように、これらを今一度、新たな人工知能研究の枠組み構築に動員する時代が訪れつつあるように思われる。筆者はこのような認識のもとで、新たな研究を展開していきたいと思っている。今年が人工知能研究の新時代への突破口を切り開く元年となることを期待したい。