

## 特集 「編集委員今年の抱負 2011」

# 知識は増えた？



市瀬 龍太郎 国立情報学研究所

最近、ある研究会で、情報は爆発的に増えたが、その割に、知識は爆発的に増えていないのではないかという話を聞いた。近年の情報通信技術の発展に伴い、人々が利用できる情報が飛躍的に増えたというのは、多くの人々に異存がないであろう。さて、知識はどうであろうか。

知識の捉え方には、さまざまな考え方がある。一つの考え方として、情報を一段階抽象化させて、利用できるような形式で整理したものと捉えることができるであろう。そのように捉えると、知識の基盤となる情報が増えれば、当然、知識が増えるための潜在的な環境が整ったといえる。しかし、情報がいくら増えても情報を一段階抽象化させる処理を行わなければならないため、単純に情報と知識が同じ割合で増えるとまではいいきれない。では、情報を一段階抽象化させるという処理能力は、情報の増加とともに増えたのであろうか。現状では、情報を整理して、知識に変換するという主体は、人間と機械の2通りが考えられるであろう。

人間が情報を知識に変換する場合には、正確性、信頼性は非常に高い。例えば、知識の集積として、辞書を考えてみよう。辞書の作成では、雑多な言葉の山から必要な情報を抜き出し、編集者、執筆者の多くの知見を背景としてなされる。人間が情報処理をした知識の集積である辞書は、正確性、信頼性などが高い。しかし、知識の量という観点で見ると大きな問題点が生じる。個々の人間がもつ、情報を知識に変換するという情報処理能力には限界がある。そのため、情報がいくらあったとしても、専門家が多くの情報を知識に変換することは難しい。しかし、この問題も情報技術の発展に伴い、状況が大きく変化している。確かに、個々の人間の情報処理能力は限られているが、たくさんの人間の情報処理能力を統合して扱えば、個々の人間の能力の限界を打ち破ることが可能となる。インターネット上で多くの人々によりつくられるウィキペディアは、その代表とされており、情報を知識に変換するという作業は、個々の人々がやっていた時代に比べて非常に大きな変化を遂げている。

一方、機械が情報を知識に変換する場合には、人間の場合とは逆に、正確性、信頼性が低いが大量の情報の処理に向くと一般的にいえるであろう。私が近年、積極的に行ってきた研究テーマは、ここに含まれているといえる。大量の情報から知識を獲得する「データマイニング」、さまざまな情報を統合することで知識に変換する「意味的情報統合」、情報を整理して概念を生成したり概念間

の対応関係を同定したりする「オントロジー学習」や「オントロジーマッピング」、過去の情報から未来を予測する知識を生成する「機械学習」などの技術である。これらの技術は、近年になり、大きな進歩を遂げ、正確性、信頼性が従来に比べて大幅に向上している。しかも、これらの技術は、大量の情報が得られることで、正確性や信頼性が上がるという性質をもっているため、人間が主体の場合と同様に、近年になりますます多くの情報が知識へと変換できるような環境が整ってきているといえるであろう。

このように考えると、情報が爆発的に増え、情報が知識へと変換される処理手法も急速に整ってきているため、環境的には、知識が爆発的に増えてもおかしくない状況になっているといえるであろう。ここで、改めて周囲の状況を見回してみると、実は、知識の爆発はすでに起こっているのではないかと思われる。

セマンティック Web の世界では、結合データ (Linked Data) と呼ばれるデータの Web が急速に広まっている。結合データでは、ある情報と別の情報の間を関係で示し、それらの情報を関係でつなげていくことで、さまざまな情報を関係とともに提示していく枠組みである。例えば、ある映画という情報があった場合には、その監督が誰であるかの情報、そして、その監督は別に何の映画をつくっているかの情報などがそれぞれ関連づけられることによって、インターネット上で提供されている。この枠組みは、人工知能において、知識表現として長年用いられてきた意味ネットワークとほとんど同じである。そして、重要なことは、これらの結合データは、上記で述べたようなたくさんの人間や進歩した機械の力を借り、急速に拡大をしていることである。

人工知能の歴史をたどると知識は、必ず言及される大きな論点である。しかし、これまでは、有用で大量の知識をいかにしてつくるかということが問題とされ、知識工学や機械学習、データマイニングの研究が多くなされてきた。しかし、上記で述べてきたように、これからは膨大な知識が存在したうえで、いかに人工知能の研究を行っていくのかというのが試されるようになる。しかも、これまでの人間が読む Web の知識とは違い、機械で処理可能な Web の知識として。

このような知識基盤を活かし、社会の期待に応えるような人工知能研究に、今年は取り組んでいきたいと思う。