

■宇野毅明 情報学プリンシプル研究系 准教授

【タイトル】

高度な問題を一般のパソコンでも解けるようにする

【本文】

コンピューターに計算させるには、計算の仕方の設計図、つまりアルゴリズムが不可欠です。高性能なコンピューターを使わなければ解けないような問題を、アルゴリズムを書き換えることで、一般に普及しているコンピューターでも十分解けるようにする、それが私の夢です。

アルゴリズムの研究には、理論を極める方向と、現実の応用を広げる方向の2つがあります。理論的な研究では、解こうとする問題について、いちばん厳しい、言い換えれば最悪の条件を想定して、それでも解ける工夫を考えます。しかし、現実の問題ではもっと条件が緩いことが多く、いちばん難しい問題を解くための手法を使うと無用な作業をすることがあります。私が考えているのは、理論的な研究の成果を十分理解しながら、現実の問題をうまく解く方法を見つけ出すことです。

### データマイニングの効率化を図る

データベースの世界では、「データマイニング」という分野があります。例えば、コンビニエンスストアで販売戦略を立てる場合、まず過去の販売データから「弁当を買う人はお茶も買う」というような組み合わせを探ることがよく行われます。ある本を買う人は、別のある本も買う頻度が高い、という組み合わせを探し出して、お薦めの本として表示するのも同じ手法（協調フィルタリング）が基になっています。

このときに、たくさんの組み合わせを調べるのは大変です。また、結果が多すぎても使いにくい。最初から、意味のある組み合わせになりそうなものを調べたほうが効率的です。そのための独自のアルゴリズムをいくつか考案してきました。

### アルゴリズム研究の可能性

アルゴリズムの研究では、その研究成果のユーザーが他の分野の研究者であることが多く、私も積極的に協力しています。ユーザーから問題を提供してもらうことが自分の研究にも役立つからです。

最近、人間の遺伝情報を解析するヒトゲノム解析の計算量を減らすためのアルゴリズム作成にも協力しています。ゲノムを構成する塩基の配列を解析するときは、制限酵素を使って数百塩基単位に切ってから調べるのですが、先頭から順番に切っていく、ということはできません。その酵素の働く部分でゲノムが切り刻まれて、無数の断片が混ざった状態になります。そのため、各断片の配列をまず確定してから、それぞれの端を見て、どの断片同士がつながるかを調べて貼り合せ、元のゲノムを再構成していました。この一連の作業を行うのに、数百台のコンピューターが何日もフル稼働しなければならないほどでした。そこで、私はこの作業に使われているアルゴリズムを見直して、似ている断片をまず分類してから比較するようにしたところ、作業時間をなんと数時間程度に抑えることができました。

このように、アルゴリズムの側面からデータ処理を見直すと、ブレークスルーがあり、今後も

いろいろな問題に取り組みたいと思っています。

(取材・構成 齋藤 淳)