

よって、ストラテジ・マップをたどることができ、最終的に解析目的に適した統計手法を用いて解析を行うことが可能となる。

4. インプリメンテーション

系統立った知識を計算機に組み込んで利用することのできるツールとして、「HyperLink (発売: ウェブトレイン)」というソフトウェアを用いた。つまり、作成したストラテジ・マップを基に、階層的な統計解析手順を互いに関係付け、参照できるようにした。また、統計解析プログラムについては我々が開発したものを修正して組み込んだ。これらの作業の結果、統計解析コンサルテーションシステムを構築することができた。

5. ま と め

統計解析の知識体系は、階層的な関係が定義されており、また追加・修正が行われる頻度も高い。ストラテジ・マップは視覚的に捕らえる事ができ、また柔軟に構築できる点からも、今回の問題に適していると思われる。

今後の課題としては、システムが提示した手順に対して、なぜその様な決定に至ったかを説明する機能や、統計プログラムが計算した統計量に対する解釈方法に対して、どの様にシステムに取り入れていくかが残されている。

いずれにしても、試作と試用を繰り返して、本システムを鍛えていく必要があると思われる。

参 考 文 献

- Affi, A.A. and Clark, Virginia (1990). *Computer-aided Multivariate Analysis*, 2nd ed., Van Nostrand, Princeton, New Jersey.
- Chabris, Christopher F. (1989). *Artificial Intelligence and Turbo C*, Multiscience Press, New York (岩谷 宏 訳, 『Turbo Cによる人工知能』, 工学社, 東京).
- Gale, W.A. (1986). *Artificial Intelligence and Statistics*, Addison Wesley, Reading, Massachusetts.
- Gebhardt, F. (1990). *An Expert System Strategy for Selecting Interesting Results*, Proceedings of Compstat 90, Physica-Verlag, Heidelberg.
- 林 篤裕, 垂水共之 (1990). データ解析における手法選択のコンサルテーションシステム, 統計エキスパートシステム研究会配布資料.
- HyperLink ユーザーズマニュアル (1990). ウェブトレイン, 東京.
- Quillian, M.R. (1969). The teachable language comprehender, *Comm. ACM*, **12**, 459-476.
- 垂水共之, 林 篤裕 (1988). 『統計解析ソフトウェア Seto/B』, 共立出版, 東京.
- 上野春樹, 石塚 満 (1987). 『知識の表現と利用』, オーム社, 東京.
- Van den Berg, G.M. and Visser, R.A. (1990). *Knowledge Modelling for Statistical Consultation System; Two Empirical Studies*, Proceedings of Compstat 90, Physica-Verlag, Heidelberg.

データ解析チューターシステムの試作を終えて

北海道教育大学 函館分校 宇田川 拓 雄

1. はじめに

コンピュータを用いたデータ解析は、以前はコンピュータ言語によってデータ処理プログラムを自作して行っていた。従って統計学とコンピュータの知識を持った少数の人だけがそれを

行うことが可能であった。ところが、近年、コンピュータ性能が向上し、統計パッケージが普及したために、データ解析は随分と手軽に行えるようになった。また、研究成果を数値データの形で発表することが一般的になったために、コンピュータを用いてデータ解析を行いたい人の数は確実に増えている。

データ解析法は便利な技術であり、うまく使えば大きな効果が期待できる。その意味で、データ解析法が普及するのは大学の研究者にとっても、また社会的にも好ましいことである。しかし、データ解析を実際に研究に使用するにも、また学生に教育するにも、解決しなければならない幾つかの問題がある。この研究はそれを解決しようとする試みである。

2. データ解析実践上の問題点

データ解析を行うには色々な分野の知識が必要である。例えば、コンピュータのハードウェアとソフトウェアの知識、統計学の知識、統計パッケージの知識、当該分野の専門知識などである。それらが有機的に組合わされて初めて解析を成功させることができる。

問題は、現在のところ、データ解析を必要としている各分野の専門家の中にはコンピュータやデータ解析の訓練をほとんど受けていない人が少なくないことである。これらの人々はその分野での専門家であり、第一線の研究者であるが、ゆっくり統計学の基礎やデータ解析法を学ぶ時間や機会に恵まれない。

このような人々が自分で学ぼうとしても、データ解析の実際を丁寧に解説した良い書物は少ない。データ解析は実際にデータを解析しつつ経験を通じて学ばなければならないが、そのための参考書もない。統計パッケージのマニュアルはそのソフトの使い方の説明だけであって、実際にデータを解析した時に生ずる様々な具体的な疑問に答えてくれるものではない。

3. データ解析教育上の問題点

実社会や研究者の間でデータ解析が普及してきたために、大学で学生や院生にデータ解析法を教育する必要が生じている。データ解析を教育するには少なくとも(1)コンピュータシステム、(2)統計パッケージ、(3)データ解析法のテキスト、(4)実習用の例示データ、それに(5)当該分野におけるデータ解析の経験的知識、の5つがそろっていなければならない。

教育の現場での問題は、これら5つを取りそろえるのが困難なことである。コンピュータシステムの維持管理を専門に行う人員のいない大学、学部などではコンピュータやソフトを常時使用できるように維持しておくことは容易なことではない。教育を始める前の準備に大変な手間がかかる。このうちコンピュータそのものは大学の予算措置に関わることであって、いかんともしがたい。しかし、あとの4つについては知的な努力である程度解決可能であるように思われる。

4. ハイパーテキスト+シミュレーター

そこで、1つの解決策として、統計処理プログラム、テキスト、例示データ、経験的知識を組合わせて1つのソフトウェアにしてしまうことを思い立った。1989年から3カ年の文部省科学研究費(試験研究)を得て、「統計データ解析チューターシステム」の試作を行った。実際のシステムを作成する前に、実験的なデータ解析セミナーを数回にわたって行い、データ解析教育に関する経験的知識を収集した。例示データは研究グループのメンバーが、その研究テーマや背景について十分に理解できているものを選んだ。

このシステムは、一種のハイパーテキストである。基本設計は次のようになっている。それはまず、データ解析の知識をテキスト(文章)として持っている。テキストの内部から、テキス

トの中で用いられている専門用語の説明を見ることができる。また図、写真、音声や音楽、アニメーションなどの特殊効果がセットされており、テキストや例示データの理解に役立てることができる。さらに必要に応じてテキストの任意の箇所を参照することができる。

ユーザーはこのシステムによって、データ解析に関するほとんど全部の知識をディスプレイ上で自由にわたり歩くことができる。このことを「ナビゲート」(航海)という。また、データ解析実習の段階で、テキスト外部の統計プログラムを呼出し、用意されている例示データを用いて、データ解析を実際に体験できる。解析結果について、システム中に用意してあるアドバイスや専門知識、経験的ノウハウ、間接的に関係する付随的な知識を参照して学ぶことができる。

試作システムの内容は「データ解析の基礎」、「因子分析」、「回帰分析」からなる。コンピュータはMac II (5MB, 40MBHD)、ハイパーテキストはHyperCard、統計プログラムはSPSSを用いた。例示データは小学校児童の体力データ、学生の身体測定データ、アイヌの衣服の色のデータ、幼稚園児の栄養のデータなどである。因子分析と回帰分析についてはデータの背景や分析テーマの詳しい説明をつけた。特殊効果としては簡単なメロディー、アニメーション、イラスト、スキャナーで取り込んだ写真を用いた。

5. チューターシステムによる教育

1991年の秋からの学期で、本学の「データ解析法」の授業で使用した。授業は、講義のあと、演習の課題を与え、それに答えさせる形で進めた。演習は「内容の理解」と「データ解析実習」の2つの部分からできている。学生は課題の指示に従ってチューターシステムを操作して、設問に答え、レポートを作成する。設問の答えやヒントはチューターシステムの中に用意してある。例示データの解析結果をもとに出力結果の解釈を行う際にも、アドバイスやヒントを参照することができる。

チューターシステムによる教育では、学生は一人で知識や経験の「海」をナビゲートしなければならない。このことによって学生は自主的に学習に取り組むようになった。その結果、学生の興味や関心、学力に応じた教育が可能になる。理解度も深まったし、学生の学習意欲も高くなった。データ解析は上手にできるならもともと面白いものなのである。

6. 今後の展開

この研究の特徴は、すでに良く知られているハイパーテキストの概念を、データ解析に適用して、実際に教育に使えるシステムに作り上げたことである。

開発システムはMacで動くが、それをもっと広く普及しているパソコン(PC-9801)やEWSの上で動くシステムに移植する計画である。また広く一般に使ってもらうには統計手法と例示データを増やすことが必要であり、解説内容をもっとやさしく分かりやすいものにしなければならない。

さらに、われわれはチューターシステムはエキスパートシステム構築のための前段階であると考えており、データ解析教育のAI化へ向けての具体的な作業を検討している。

参 考 文 献

- 宇田川拓雄(1992). 統計解析チューターシステムの開発, 平成3年度文部省科学研究費補助金(試験研究B)研究成果報告書.