

論文内容の要旨

博士論文題目 統計解析システムにおけるプログラミング言語の研究

申請者 小林 郁典

(論文内容の要旨)

統計解析システムは、実験・観測・調査で得られたデータを様々な視点で眺め、いろいろな種類の要約統計量を計算し、図表を描き、測定した現象の背景にある本質を探るための環境である。それを具体的に行うことは、統計学の成果に基づく統計解析システムの上で実現される。しかし、統計解析は、データに対して定型的にこれらの手法を単純に適用すれば目的を達成できるというものではない。データによって手法を使い分けたり、評価のための判断基準を変更したりしなければならないことがある。このような作業を行うために、統計解析システムには専用のプログラミング言語が要求される。統計解析システムのプログラミング言語は、統計解析システムの可能性を広げるための重要な道具となっている。

本論文は、3つの統計解析システム(重回帰分析支援システム RASS, それを拡張した RASS2, Java 言語ベースの汎用統計解析システム Jasp) のプログラミング言語を、どのように設計・実装したか、これらを使用することでどのような効果があるかを述べている。また、これらを実現するために必要であった関連技術に関する研究成果についても述べている。これらの言語は、利用者が簡単にしかも安定的にプログラミングができるように工夫されている。

1 章は、緒論として、本論文の位置付けと目的を述べている。

2 章では、知的データ解析支援システムに求められるプログラミング言語の条件について考察し、その具体例として実装した知識ベースの重回帰分析支援システム RASS (Regression Analysis Supporting System) のプログラミング言語 (RASS 言語) について説明している。RASS 言語の設計では、RASS が持つ支援機構の枠組みを崩さないこと、「知識処理」に適した論理型プログラミングと「手法を記述する」手続き型プログラミングが混在できること、アルゴリズムを容易に記述できることなどに留意している。

3 章では、RASS を拡張した RASS2 とそのプログラミング言語について述べている。RASS2 は、重回帰分析を知的に支援する環境であるが、その構成や内容を柔軟に変更できるように設計している。それを実現するために、オブジェクト指向言語のオブジェクトに局所的な重みを持つルールを導入して、そのルールに基づいて利用者の解析を自発的に支援することができるモジュールを基本構成要素としている。RASS2 には、このモジュールを記述するための機能が備わっている。

4 章では，対話型プログラミング環境で，容易にオブジェクトに機能を追加する方法を提案し，その実装方法について述べている．それを実現するために，条件部と実行部から成るアクティブルールを，特殊なメソッドとして，オブジェクトの枠組みの中に組み込み，オブジェクトが，アクティブルールに基づいて自発的に振る舞うようにしている．この機能を利用すると，対話型プログラミング上でオブジェクトを操作する利用者が，従来の方法と比べて簡単にオブジェクトを拡張させることができることを，例を用いて示している．この方法は，統計解析システムのプログラミング言語にも適用できる．

5 章では，汎用統計システム Jasp(JAva based Statistical Processor) のプログラミング言語の機能と文法について説明している．Jasp は，Java 言語をベースにして開発されており，先進的なコンピュータ技術を，統計解析において利用できることを目指している．Jasp 言語は，関数の記述を基本とした手続き型のスクリプト言語であるが，オブジェクト指向プログラミングをできるだけ簡便にできる枠組みと柔軟性・拡張性を兼ね備えるように設計されている．Jasp 言語の実装には，Java 言語のスクリプト言語である Pnuts を利用している．

6 章では，本研究で得られた成果を要約して，統計解析システム用のプログラミング言語の開発における指針を述べている．

(論文審査結果の要旨)

日々発生する様々な大量のデータを統計的に解析処理するに当たって、コンピュータを利用することは不可欠なことである。これに対して、多くの統計パッケージや統計解析システムが開発されてきた。しかし、それらを使用する際に、統計解析の詳細な知識が要求されたり、決められた解析処理しかできなかったりして、使用の容易さや柔軟性の面で、必ずしも満足の行くものではなかった。

本論文は、これらの点を改善するために、統計解析システムに必要なプログラミング言語の設計方法と、実装方法について研究し、次の2つの目標に向けて具体的に設計・実装した言語の効果を述べたものである。

(1) 知的支援機構をもつ統計解析システムのプログラミング言語を開発する。

(2) 汎用統計解析システムで拡張性をもつプログラミング言語を開発する。

本論文の主な成果として、3種のプログラミング言語の開発と、新しい特徴として自発性をもつ会話型言語の提案をしており、次のように要約される。

1. RASS 言語：重回帰分析支援システム RASS (Regression Analysis Supporting System) に、「知識処理」に適した論理型プログラミングと「手法を記述する」手続き型プログラミングの機能が混在したプログラミング言語として、RASS 言語を設計し実装している。知的なデータ解析を支援するシステムの実現において、アルゴリズムの記述を容易にして、使いやすさを向上している。
2. RASS2 言語：RASS を拡張して、オブジェクト指向言語のオブジェクトに、局所的なルールを加えたモジュールを基本構成要素とするプログラミング言語 RASS2 を設計し、実装している。RASS2 では、ルールに基づいて、利用者が気付かない状況を検出したり、特別の状況での解析処理を加えたりして、オブジェクトが利用者の解析を自発的に支援することができるようにしている。
3. Jasp 言語：汎用の統計システム Jasp (JAva based Statistical Processor) のプログラミング言語の機能と設計方針を述べ、Java 言語のスクリプト言語である Pnuts を利用して実装している。Jasp 言語は、Java 言語をベースに開発しているので、XML 文書の扱い、C や Fortran プログラムとの結合、その他の先進的なコンピュータ利用技術の利用を可能にして、柔軟性・拡張性を兼ね備えるように設計されている。
4. 自発的オブジェクト：オブジェクトに、条件部と実行部から成るアクティブルールを導入して、対話型プログラミング環境で、オブジェクトに機能を容易に追加でき

るようにしている。これについては、実験環境を整備している段階だが、オブジェクトが、アクティブルールに基づいて自発的に振る舞う機構の有効性を示している。この方法は、統計解析システムのプログラミング言語だけでなく、広い分野の言語に適用できる。

このように、本研究の成果は、使用の容易さと柔軟性を改善した統計解析システムのプログラミング言語を設計・開発して、統計解析システムの可能性を広げていること、その有用性を示していること、新しい言語の設計に示唆を与えていること等、学術上、応用上寄与するところが多い。よって、本論文は博士（工学）の学位論文として価値あるものと認める。