

マルチメディアインターフェースとしてのAVSの機能評価

北川文夫 池田秀人 中村正規
広島大学総合情報処理センター 東京電力

科学技術計算で処理されるデータは、データ構造が複雑な場合が多いので、可視化や統計処理することで計算結果を分りやすくすることが重要になってきている。しかし、それらのデータ構造は、特定の目的に作られたプログラムに依存した独自なものであり、別のデータ処理ツールで処理するには、いちいちデータ変換する必要があり、利用者にとって障害であるばかりか、同一の対象を処理している認識を失わせてしまう問題がある。この問題を解決するために、同一のデータをいろいろなツールで処理する環境：dataICUを提案する。また、その可視化ツールとしてAVSの評価を行う。

Functional Evaluation for AVS in the DataICU Environment

Fumio Kitagawa Hideto Ikeda Masaki Nakamura

Information Processing Center, Tokyo Electric Power Co.
Hiroshima University

Since the format/structure of scientific data is so complicated, it becomes too important to visualize the data or to analyze the data statistically. But, the format/structure of data depends its analyzing program, then it needs to convert the data before processing them by another processing tools. This causes the barrier of research or lack of recognition of processing data as unique object. To improve such a problem, we propose a new environment for scientific data processing, named dataICU. Additionally, we evaluate the AVS as a visualization tool in dataICU.

1. はじめに

コンピュータ技術の発展に伴い、科学技術計算では、より大規模な解析計算や大規模なシミュレーションが可能になった。それらのプログラムで処理されるデータは、構造が複雑だったり、容量が大きいので、解析結果を分りやすくするための、いろいろなデータ処理／管理ツールが開発されてきている。その中で、科学技術データベースシステムは、科学技術データのデータ管理を容易にして、科学技術計算を支援する目的で開発されている。これは、ビジネス分野でのデータ処理で成功を収めたデータベース技術を応用して、科学技術データ処理を支援しようという発想である。しかし、これらはあまり成功していない。ビジネス分野で成功を収めたデータベース技術が、科学技術計算分野ではなぜ成功しないのであろうか？これは、それぞれのデータの処理の目的を見れば当然のことであることがわかる。すなわち、ビジネス分野では、均質で大量のデータを保存、チェック、検索等を行うことが目的であり、かつ、そのためのアプリケーションの開発が効率的に行われる必要がある。そのために、データの定義やデータの管理、トランザクション制御等の機能をアプリケーションから独立にDBMSに持たせることで、これらの要求を満たすことができた。しかし、科学技術計算で用いるデータは、解析手法やデータモデルの構造が高度な理論に基づく特殊なものが多く、ビジネスデータのように多量のデータを蓄積しておき、それらをまとめて処理するということは非常に少ない。また、データそのものは比較的変化が少ないので、データ構造を規格化してデータベースに蓄積するより、それぞれのデータ特有の構造のまま処理することが求められ、プログラムもそれに適したもののが利用されている。つまり、データ自身の構造も特殊な上に、個々のデータの処理方法が大きく異なるため、ビジネスデータ処理に適したデータベースの手法を適用するには無理があった。したがって、科学技術データを処理するには、個々のデータに対してそのデータに適用できる専用解析／表示ツールが容易に利用できる環境を用意することが必要である。

我々は、このような科学技術データの処理環境のシステムコンセプトを病院の集中治療室に求めた。病院の集中治療室は、各種モニタや治療機器が用意されており、複数の専門医がチー

ムを組んで1人の患者の治療に専念するようできている。病棟の入院患者が均一な管理の元、効率的に治療されるのと対照的である。データベースは、医療でいえば病棟のようなもので、個々のデータの個性は無視され、均質に管理されて効率よく処理される。科学技術計算では、1つのデータを極めて大切に扱い、あらゆる解析を、十分な（計算）資源を使って行い、最大の情報を引きだそうとする。すなわち、集中治療室の患者と似た扱いを受ける。これが、データのための集中治療室：data ICUのシステムコンセプトである。data ICUのデータ構造は、いろいろな解析系が持つもの（ユーザが定義したものも含む）に柔軟に変換できるようになっている。システムはデータ構造辞書を持ち、これを参照して、処理対象のデータが適用できる解析／表示ツールを示したり、自動的にそれらツールに適した構造へ変換する機能を持つ。こうすることで、ユーザは1つの対象（データ）を継続して解析／表示しているという認識で、データをスムーズに処理できる。

本稿では、2節でdata ICUの設計理念とシステム設計を示し、3節でその実現方法を紹介する。その中で、data ICUの可視化ツールとしてAVSの機能の評価を行う。

2. データのための集中治療室： data ICU

科学技術計算では、その処理目的や解析プログラムにより、独自の構造を持つデータを入力し、解析結果も解析プログラム独自の構造を持っている。科学技術計算で代表的な利用分野とプログラムには、次のものがある。

- ・流体力学 PLOT 3D
- ・構造解析 NASTRAN
- ・分子軌道計算 Gaussian

また、シミュレーションでは、各種独自の形式があり、中でも有限要素法ではUCD(Unstructured Cell Data)が代表的である。この他、リモートセンシングデータ、NCSAがコンピュータ間の科学技術データ移行のために定めた形式：HDF(Hierarchical Data Format)や、各研究者が独自に開発したものなど、科学技術計算では多種のデータ形式が使われている。

さて、このように多種あるデータ形式は、それぞれ専用の解析プログラムへの入力データやその解析結果のものであるから、1つの解析

プログラムを使うだけの場合は、効率よく扱われる。しかし、一旦この処理手順から一步踏み込んで、他の解析プログラムや可視化ツール、統計ツール等を用いて別の視点からデータを見ようすると、それが非常に困難なことが分る。科学技術計算で得られたデータは、構造が複雑だったりデータ容量が大きかったりするので、可視化や統計処理又は他のツール等を使わないと本当に必要な結果が得られたかどうかが分りにくい。そのため、読み込めるデータ形式が汎用的な可視化や統計システムがいくつか開発されている。AVSやS-PLUS等である。しかし、これらのシステムを使って、ユーザが自分のデータを処理する場合、独自のデータをこれらのツールに読ませるための変換が必要なため、簡単には利用できない場合が多い。

以上のような特徴を持つ科学技術データの処理（特に可視化）支援システムが、データベース技術を用いて開発されてきている。最近のものでは、Tigra[1]がある。これは、ODBの上に、検索とグラフィックビューアをビジュアル環境で組み合わせてアプリケーションを構築できる環境を提供し、イメージデータの階層もオブジェクトの階層構造と連動してビューできる等の機能を持つ。また、イメージデータ等の大容量データの処理を高速化するためにバッファリングの最適化等も行い、大量のイメージデータを処理する場合に有効なシステムと考えられる。USD[2]も同様にDBMS中心の科学技術データベースである。しかし、これらシステムは均質で大量のデータを効率よく処理するという、データベースの発想であるため、イメージデータのように、解像度別に階層構造を構成するといった比較的簡単な構造をしたデータには適するかもしれないが、より複雑な構造をしたデータには適さないように見える。すなわち、オブジェクト指向でのモデリングで、複雑なデータ構造が格納できたとしても、下層のオブジェクトを独立して処理するということは、あまりないからである。例えば、自動車の持主は自動車全体を1つの対象と捉えるわけで、エンジンの調子とか、タイヤの減りぐあいを気にかけるのは、それが部品のことではあるが自動車本来の機能を發揮するかどうかという点から捉えているからである。単体の部品だけに注目する必要は、ほとんどないであろう。自動車会社から見ればこのようにはならず、

個々の部品まで注目する必要があろう。同様に、科学技術データも複雑な構造を持つが、その構造をOOモデルで再構築するメリットは少ないよう見える。むしろ、科学技術データは、いろいろな解析プログラムや可視化ツールが直接処理できるそれぞれのデータ構造に、柔軟に変換できる共通なデータ構造で管理すべきであろう。

以上のことから、我々は、科学技術計算で処理されるデータを、その意味を変えずに構造を変換できる柔軟な形式で持ち、いろいろなアプリケーションのもつデータ構造を辞書に記述しておくことにより、それらのデータ構造に変換して実行させるのが、最も自然で効率が良いと考え、対象データそのものを多角的に処理する支援環境が必要であるという結論に達した。しかも、個々のデータ処理ツールは既存の研究成果を利用し、支援環境はそれらのツール間でデータを円滑に渡して処理しやすくすることのみを主な機能とする。これは、「データベース研究で何をするべきか」という議論[3,4]に指摘されるような、既存の成果を有效地に利用し、後追い開発することを避ける立場である。

これらの考え方を元に、科学技術計算支援環境として設計したのがデータのための集中治療室：data ICUである。data ICUでは、文字通りデータがその中心にあり、それを取り巻くように、解析プログラムや統計プログラム、可視化プログラム等が用意され、データを多角的に処理できる環境を提供する。data ICUのシステム要件は、いくつか考えられるが、最も重要なのは、データ変換の透明性である。これは、ユーザから見れば、1つのデータは、いろいろな処理をしていても同じ解析対象（オブジェクト）と映るし、アプリケーションから見れば、接続性の良さと見ることができる。図1にdata ICUのシステム概念を示す。

・ストレッチャー：ユーザファイルからのデータ（解析対象）の読み込み、又は、オペレーションを行ったデータのユーザファイルへの書き出しを行う。読み込みでは、bedに格納する際、パーサ又はデータチェッカーがデータ構造の特定又はデータの誤りを検出する。データに誤りがなければ内部データ構造に変換し格納する。病院で、患者を集中治療室へ運ぶ、ストレッチャーに由来する。

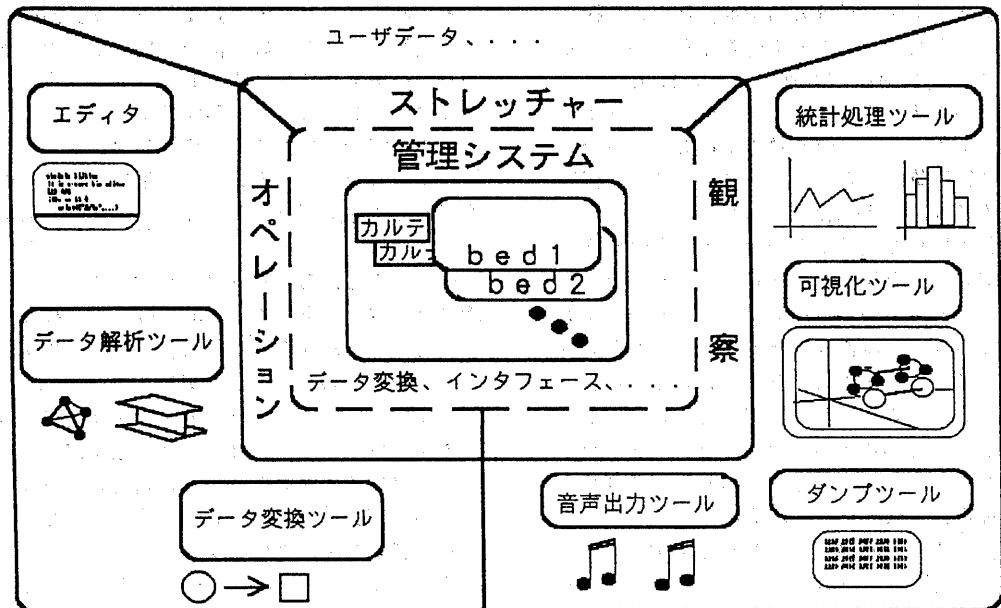


図1 data ICU概念図

・ b e d : データが data ICU に読み込まれたときのデータの格納場所。システム内部で使われるデータ用 DB のこと。DBMS は、柔軟な構造をしており、どんなデータタイプも格納できるものを用いる。データ名称、データフォーマット名称や、それぞれのデータの意味（座標位置、電位、流量、とか単位など）、データの処理状態を示す管理情報等が記録されたカルテと共に、内部形式に変換されたデータが管理される。ここに格納されたデータは、オペレーションや観察の操作が行われるときに、データ構造をそれぞれのツールに適した形式に変換してから渡される。オペレーションで変更されたデータは、元のデータ名等が書かれたカルテと共に、新しい b e d に乗せられる。名前は集中治療室で患者をのせる b e d に由来する。

・ オレーション：ベッドに乗せられた data に手を加え、 data を変更する操作。具体的には、(1) 科学技術計算を行うプログラムによるデータ解析やシミュレーション等を行う、(2) エディタによるテキスト、グラフィック、図面等のデータ編集、(3) 明示的にデータフォーマットを変換するなどの操作が行われる。集中治療室の手術（オレーション）に由来する。

・ 観察：元の data を変更せずに、ユーザが望む出力形態に data を加工をして、出力装

置に出力する。次の観察方法がある。(1) 統計処理プログラムにより、データの統計処理を行い、クロス表やグラフ表示を行う。(2) データの可視化ツールにより、データを2次元や3次元で表示し、色やレンダリング手法、視点等を変えることにより最適な表示を行う。(3) 楽譜の演奏や、テキストの読み上げを行うプログラムにより、音声出力をを行う。(4) ダンプツールにより、データのダンプ表示を行う。(5) データのカルテを見る。

また、 data ICU の機能を実現させるための内部プログラムは、データを入力するもの、各ツールを起動する際の起動パラメータやスクリプト生成をするもの、ツールを起動するもの、データを変換するもの、処理の状態を管理するもの、ユーザインタフェースを行うものがある。

次に、 data ICU をユーザが利用する場合のユーザインタフェースを図2に示す。これは、1つのウインドウであり、ストレッチャー、オペレーション、観察、登録の4個のボタンとオブジェクト、データタイプ、状態を示す3個のテキストボックスからなる。ユーザは、次の手順で処理を行う。(1) ストレッチャーボタンを押すと、データタイプ選択リストが表示されるので、1つ選択する。すると、ユーザが持つデータのうち該当データタイプのみのオブジェクト（データ）一覧がアイコン表示される。

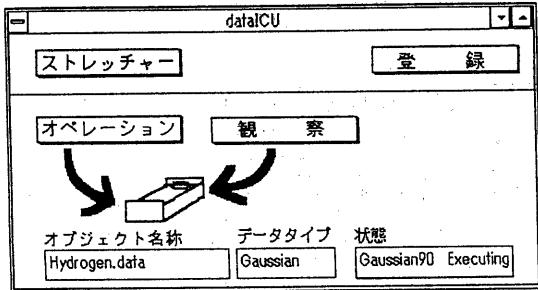


図2 ユーザインターフェース

目的のオブジェクトを指定すると、データが内部形式に変換されDBに読み込まれ、オブジェクト名称、データタイプにそれぞれの値が表示され、bedアイコンが表示される。状態は、データの種類により適切に表示される。(2) 次に、オペレーションを選択すると、オブジェクトのデータタイプに適用可能なツールアイコンが表示されるので、1つを選択し、bedへドラッグする。すると、内部データはツールのデータ構造に変換されてからツールに渡され、実行される。その際、実行ウィンドウが開きツールの提供する各種の操作ができる。解析ツールの場合は、バッチ型で実行されることが多く、その場合、実行ウィンドウには「オペレーション中」のメッセージが表示されるだけである。このとき、dataICUを終了させ、ジョブが終了した時点で再起動すれば、継続した処理が行える。オペレーションの場合は、データが変更されるので、状態表示にそのことが表示され、新たなbedが作成される。(3) 観察のボタンを押すと、同様にデータタイプに適用可能なツールアイコンが表示されるので、1つ選択しbedへドラッグする。観察用のツールが起動するので、パラメータ等を変更して、ユーザが望む表示を得ることができる。また、カルテの表示ではアクティブオブジェクトのいろいろな属性情報が得られる。アクティブオブジェクトを変更して、オペレーションを行う前のカルテを見ることもできる。(4) 登録ボタンを押すと、データタイプとツールの登録が行える。それぞれ登録用のウィンドウが開き、必要項目を入力することで登録ができる。

以上の様に、今まで科学技術計算でのデータ処理でツール間を移動するには、ユーザが意識して各ツールに適合するようにデータ又は、パラメータを変更するなどが必要であったが、dataICUを用いることで、この煩わしさか

ら解放され、1つのオブジェクトを継続して処理することができるので、本来の目的を達成しやすくなると考えられる。

3. 実現方法

ここでは、dataICUの実現方法を、広島大学総合情報処理センターのConvergentシステムを例にとり、説明する。dataICUシステム構成例を図3に示す。それぞれの要素の機能と特徴を次に示す。

- ・DBMSは池田ら[5,6]が開発したB5構造をしたもの用いる。B5構造は、格納できるデータタイプが自由に設定でき、構造や容量が制限されないので、dataICUには適したものである。格納されるデータは、内部形式に変換された科学技術計算データとそのカルテ、データ構造辞書、各ツール起動パラメータリストなどである。個々のソフトウェアの概要を示す。

・オペレーション：

(1) データ解析ツール

NASTRAN 構造解析

Gaussian90 . . . 分子起動計算

MOPAC 分子起動計算

ユーザ開発プログラム . . . ユーザが開発した
個々のプログラム

(2) エディタ

Emacs, xed . . . テキストエディタ

イメージエディタ、図面エディタ（計画）

(3) データ変換ツール

バイナリASCII変換、イメージデータや、いろいろなソフトウェアの持つフォーマット変換を明示的に行う。

・観察：

(1) 統計処理ツール

S-PLUS . . . 豊富な統計手法と、グラフ化機能を持つ汎用統計プログラム

(2) 可視化ツール

AVS . . . 汎用可視化ツール

(3) 音声出力

MIDIや音声合成ボード出力（計画）

(4) ダンプツール

ダンプコマンドやXのイメージダンプ

(5) カルテ

データの属性値の一覧表の表示。

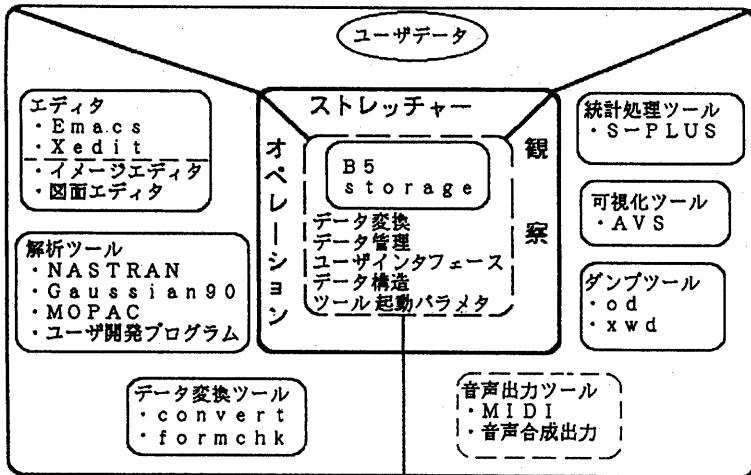


図3 data ICU実現構成例

システムプログラムのそれぞれの実現方法は次のように行う。

データ変換：内部形式データは、標準化されたデータ形式とその意味を合わせて持つており、データ変換は、データ構造辞書に記述された個々のアプリケーションのデータ構造へのマッピングとして実現する。

各ツールの実行：ツールの起動パラメータを適切に設定したり、実行マクロを生成してプログラムを起動する。

データ管理：データ管理情報はB5構造に格納されるので、B5構造へのアクセス法を使い、実現する。

ユーザインターフェース：X-Window上のwindowとして実現する。各ツールの実行は、それぞれのツールが持つウインドウで実行される。

これらの、機能をスムーズに実行させるには、次のような技術が必要である。

- ・ファイル又はデータの属性値の拡大。UNIXのファイル属性は、オーナー、日付、リードライトフラグがあるが、データの中身を示す属性が無いので、ファイル名の拡張子を明示してつけているのが、現状である。data ICUの考え方では、データはそれぞれの内容を示す属性が附属すべきであり、名前に連結させるべきではない。

- ・データ変換を効率よくするための、パイプ機能と同期制御が必要である。構造化したデータは、いくつかの異なるタイプの基本データタイプの組み合わせがあり、それぞれのデータタイプに対して、目的のデータタイプに変換する必

要がある。そのために、UNIXのパイプの機能に加え、効率よく変換するには同期をとることも必要である。

この他、実現に向けての技術的課題は多くあると思われるが、今後の研究や議論を経て明らかにしていきたい。

さて、科学技術計算では、データモデルが高度な理論に基づいたり、データ量が大きいなどの理由から、可視化することが計算結果を理解する上で重要になっている。このようなことから、次に、data ICUの可視化ツールとしてAVSの機能評価を行う。

AVSの表示機能は、基本的に次の3種に分類できる。

- (1) 座標表示。位置を示す座標データに、その点の値（温度や速度）が与えられるデータで、値に対応して色や、矢印の長さ等を変えて表示する。

- (2) イメージ表示。ピクセルデータを表示。

- (3) グラフ表示。簡単な棒グラフ、折れ線グラフ、スキャッタグラム、コンター等の表示。

このうち、座標表示がAVSの特徴的なものである。特に3次元图形の表示では、入力データの加工機能(FILTER MODULES)や座標データを球で表示するかポリゴンで表示するかという图形のマッピング機能(MAPPER MODULES)、シェーディング方法や光源などの表示機能(DATA OUTPUT MODULES)が豊富に用意されていて、自由に組み合わせて使うことができ、ユーザが最も望む表示をさせることができる。また、これらの機能は、通常画面上で機能モジュールアイコンを接続して一連の処理(Networkと呼

ぶ) を構築するが、コマンド記述ファイルとしても可能になっている。

A V S の入力機能は、A V S 形式の読込の他、代表的な解析プログラムの実行結果の読込モジュールや、Fieldデータという汎用のデータ定義形式に従ったField属性を記述することによる、汎用データ入力がある。また、CやFORTRANでデータをA V S に読み込むモジュールを作成するためのライブラリも提供されている。

次に、d a t a I C U で A V S を使って可視化する手順を、Gau us ianを例に説明する。| | の中は、システムの動作。

(1) 初期データをストレッチャーで b e d に運び、{内部データに変換する} (2) オペレーションから、Gau us ianを選択し実行する。{データをGau us ianフォーマットに変換し、実行パラメータを設定して実行させる。}| (3) この結果に対し、観察から A V S を選択する。{データをA V S で読めるようにコンバートし、A V S のコマンド言語で、(i) Gau us ianデータを読込、(ii) 等電位面を生成するモジュールを呼び、(iii) 原子は球形で表示するパラメターをつけてレンダラーを起動し、(iv) ピットマップに表示するコマンドスクリプトを生成し、このスクリプトを A V S に実行させる。}| (4) ディスプレイ上に別のウインドウが開き、原子の配置とそれを取り巻く電子雲の状態が表示される。

A V S では、表示オブジェクトに対して、色や視点、光源の位置、断面の表示等の変更を簡単に行えるので、上記の処理結果を見てユーザが表示上で不満な点を変更するのに、とまどうことは少ないであろう。

4. 終わりに

本稿では、科学技術データ処理支援環境システム：d a t a I C U を提案した。d a t a I C U は、科学技術データの特徴である個々のデータの特殊性を考慮し、それらのデータ構造に柔軟に変換できる構造を内部データに採用した。これは、科学技術データを処理するとき、継続して同一対象を多角的に見ているという認識をユーザに持たせる効果があると同時に、いろいろな分析ツールを利用するの、にいちいちユーザがデータ変換しなければならないという、障害を取り除くものである。d a t a I C U の

基本要件や実現方法、データ構造辞書など、更に調査や議論をして明かにしていかなければならない事いくつかあるが、それは今後の課題としたい。

d a t a I C U の考え方は、今までにないものであり、科学技術データばかりでなく、マルチメディア等の他の分野へも広く応用できるもので、ユーザから見たデータ処理がきわめて自然に行えるようになることが、期待できる。

参考文献

- [1] Stonebraker, C. et. al., "Tioga: Providing Data Management Support for Scientific Visualization Applications," Proceedings of the 19th VLDB Conference, Ireland, 1993.
- [2] Johnson, R.R. et. al., "USD - A Database Management System for Scientific Research," Proc. SIGMOD International Conference on Management of Data, 1992.
- [3] Stonebraker, C., Agrawal, R., Dayal, U., Neuhold, E. and Reuter, A., "DBMS Research at a Crossroads: The Vienna Update," Proc. of the 19th VLDB Conference, Ireland, 1993.
- [4] グループ討論：データベース研究で何を行なうべきか、情報処理学会データベースシステム研究会研究報告 No. 94, 1993.
- [5] Ikeda, H. et. al., "INTERVISION: A New Hypermedia System Focusing on Dynamic Media," Proceedings of the 2nd DASFAA, Japan, 1991.
- [6] 中村正規, 池田秀人, 北川文夫, 基本 5 列表を応用了した付加価値ストレージシステム (V A S S), 情報処理学会データベースシステム研究会研究報告 No. 96-10, 1993.
- [7] Springmeyer, R. et. al., "A Characterization of the Scientific Data Analysis," Process, Proc. of IEEE Visualization Conf., USA, 1992
- [8] French, J. et. al., "Summary of the Final Report of the NSF Workshop on Scientific Database Management," ACM SIGMOD Record, Vol. 19, No. 4, 1990.