日本のデータベースの現状(X 線編)

宇野 伸一郎 / 三浦 昭

文部省宇宙科学研究所 PLAIN センター

I. DARTS 開発の現状

宇野　伸一郎

1　宇宙科学データベース「DARTS」

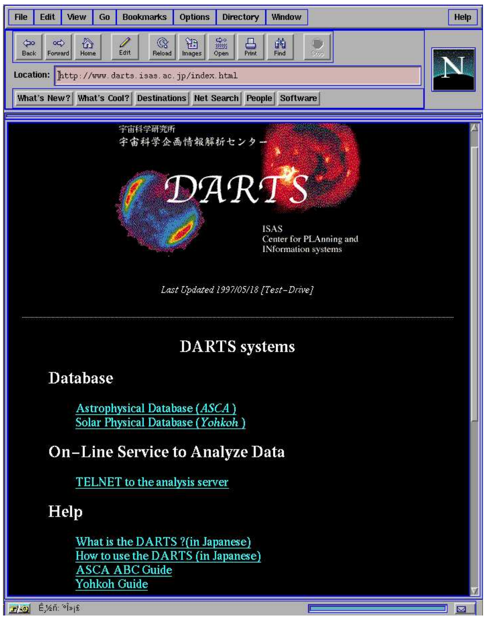
宇宙科学研究所 企画情報解析センターでは、データベースシステム”DARTS” を構築してきた。システムは、1997 年 5 月に試験公開を開始し、現在順調に稼働しつつある。ここでは、「DARTS」開発の現状報告と、今後の開発計画を記す。

データベース「DARTS」の目的は、宇宙研の衛星等により得られた科学データを一般に提供し、それらを使って誰もが自由に研究、解析を行えるようにすることにある。その特徴は、大きく以下のようなものが挙げられる。

1. 各衛星プロジェクトからのデータ提供
2. 複数のユーザーインターフェース
3. 解析ソフト/解析システムの提供 (UNIX 環境)
4. 解析マニュアルの整備
5. (人的，ソフト構成面の) 分業体制
6. 将来の拡張を容易にした構成

2　データベースの内容

データベースの内容は、表 1 のように大きく 2 つに分けられる。まずは、各々の衛星プロジェクトから提供されるデータを格納するもので、これらは各衛星毎にデータベースが作成されている。現在稼働中のものに、X 線天文衛星「あすか」と、太陽観測衛星「ようこう」のものがある。これらのデータ量はそれぞれ数百 GB 程度あり、現在でも増え続けている。この「各衛星のデータベース」では、X 線天文学では「ぎんが」衛星のデータを、また、地球物理学のほうからは「GEOTAIL」のデータをそれぞれ提供できるように現在構築している。また、地球物理学の「あけぼの」のデータや、2000 年度に打ち上げられる予定の「ASTRO-E」など、多くの衛星のデータを提供できるよう、拡張してゆく予定となっている。



http://www.darts.isas.ac.jp/index.html

宇宙科学研究所
宇宙科学企画情報解析センター

DARTS

ISAS
Center for PLAnning and
INformation systems

Last Updated 1997/05/18 [Test-Drive]

DARTS systems

Database

Astrophysical Database (*ASCA*)
Solar Physical Database (*Yohkoh*)

On-Line Service to Analyze Data

TELNET to the analysis server

Help

What is the DARTS ?(in Japanese)
How to use the DARTS (in Japanese)
ASCA ABC Guide
Yohkoh Guide

図 1: DARTS ホームページ

また、これら各衛星のデータとは別に、各衛星データとは独立に作成されているデータベースが存在し、現在は「ネームリゾルバーデータベース」が稼働している。このネームリゾルバーは天体名や座標を分解/解釈して、データベース上の天体と結びつけるものである。このような構成にすることで、各データベースの独立性を確保し、なおかつ、将来への拡張を容易にしてある。

表 1: DARTS データベース システム

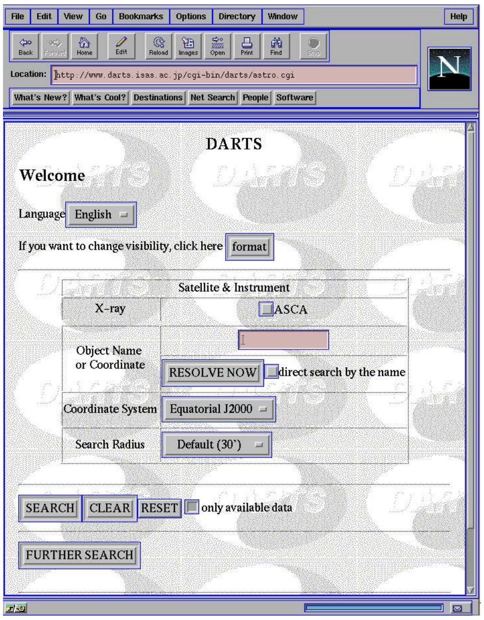
各衛星のデータベース		
	稼働中	構築中
X 線天文学:	あすか	ぎんが、ASTRO-E
太陽物理学:	ようこう	
地球物理学:		GEOTAIL、あけぼの
衛星とは独立のデータベース		
ネームリゾルバー		

3　検索システム

検索システムは、様々なユーザーの要求に合うよう、現在のところ2通り、近い将来3通りのユーザインターフェースを用意する。

1. GUIによる基本的な検索
2. より詳しい条件での検索
3. UNIX コマンドラインによる検索

「1.GUIによる基本的な検索」は、初めてデータベースを使う人や、目的の観測対象が明確に一つもしくは2-3個決まっている、といった方に主に利用していただくことを想定している。図 2 のように、選択項目は必要最小限とし、名前もしくは座標を入れて検索させればよい、というユーザーが特に何も考えずに使える設計になっている。



Location: http://www.darts.isas.ac.jp/cgi-bin/darts/astro.cgi

DARTS

Welcome

Language English

If you want to change visibility, click here

Satellite & Instrument

X-ray ☐ASCA

Object Name or Coordinate

☐direct search by the name

Coordinate System Equatorial J2000

Search Radius Default (30')

☐only available data

図 2: ASCA 基本検索画面

次に、「2.より詳しい条件での検索」がある。これは、観測日、観測対象などによる分類、検索など、ユーザーの細かい要求に対応して、特定のデータを検索する、といったものである。

そして現在、「3.UNIX コマンドラインによる検索」ができるよう、構築を行っている。このコマンドライン検索はスクリプトなどによる連続処理などをするパワーユーザー向けに対応しようとするものである。

データベースというものの最大の利点は、大量のデータに一度にアクセスできるという事だ。そして、その構築目的は、1.膨大なデータから目的の情報を簡単に取り出すこと、だけでなく、2.大量のデータを連続的に検索/処理し、大局的な情報の把握を可能とすることにもあると私は考える。残念ながら世にあるデータベースは、2の視点に欠けているものが多い。全体の情報を見渡し、統合的にひとつのカテゴリー/ターゲットを解析してゆくことのできるデータベースは、サイエンスに新たな視点を提供するものであると確信している。

現状の GUI は、ひとつのデータを探り当てるといったことはいいが、パワーユースには限界がある。そこで DARTS では、次の節で述べるユーザー解析システムと、データ検索システムを統合してゆくことを考えている。これにより、ある程度解析に慣れたユーザーであれば、ブラウザを使わずに、telnet などでサーバーに入り、データの検索、選択、解析、そしてそれらの連続処理、といったものを行なっていけるようになる。

4　ユーザー解析システム

DARTS では、UNIX による解析環境を提供している。これは、他にあまり見られない大きな特徴といえるだろう。このユーザー解析システムには、現在稼働している各衛星データの解析ツール群が用意されており、ユーザーは簡単な登録手続きだけで、これら解析ツール群を利用することができる。これよって、データを探して持っていく、というだけではなく、データをその場で見て、解析する、ということができるようになっている。つまり、自分のワークステーションに、「ようこう」「あすか」などの解析ツールをインストールすることなく、データ解析が実現できることになっている。

さらに、この解析システムが、先に言ったコマンドライン検索システムとあわさる事によって、非常に強力なデータベースとなることができる。例えば、観測されたすべての活動銀河中心核に関して、スペクトルを作成し、その特徴を総合的に議論する、といったことができるようになる。現状の GUI でこれをやっていたら非常に手間がかかってしまうことは容易に想像できるであろう。また、たとえば、「あすか」がこれまで観測した全視野のイメージを天球に投影する、といったこともできるようになっている。このように、大量のデータを処理し、解析することができる、というのが「データベース」といったものの最大の利点のひとつだと

我々は考えている。

5 マニュアル

解析システムを提供する上では、マニュアルの整備は不可欠だ。そのため、DARTS では現在マニュアル整備が進められている。ここでは「あすか」を例に挙げる。

「あすか」は、X 線 CCD カメラと、蛍光比例係数管を積んだ衛星で、各検出器で、イメージ、光度曲線、スペクトルがそれぞれ得られる、ハイテク衛星である。取得されるデータは X 線望遠鏡、各検出器の応答関数を反映しており、このため、データベースから、スペクトルファイルがすぐに出てくるというわけにはゆかない。スペクトルひとつ導出するにせよ、生データの選択や応答関数の選択など数段階のプロセスを経ないと得られない。

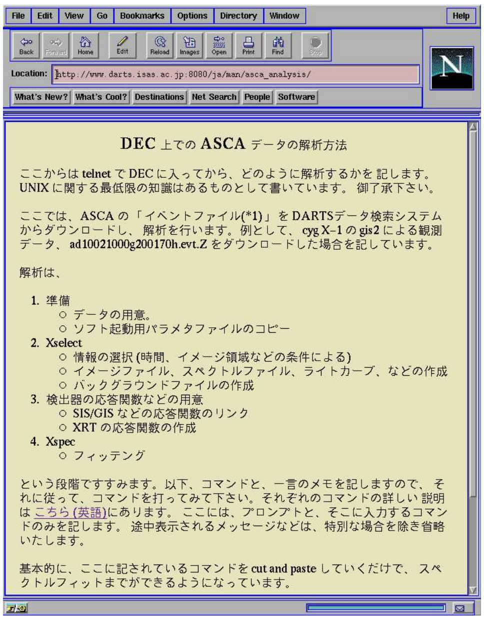
この「あすか」の解析に関しては、「ASCA abc」という英語のマニュアルが存在する。これを最初から読んで解析する、というのが本来の姿であるが、それだけでは、専門家以外にとっては使いやすいとはいえない。そこで、DARTS では 2 段階のマニュアルを用意している。

まずひとつは、各種ツールの動作を最低限のレベルまで解説したものだ。これは「ASCA abc」まで入らなくても、ある程度、基本的なツールが使いこなせるようになる、というコンセプトのものである。また、スペクトル導出の他にも、基本的な時系列解析やイメージ解析も可能となるようなものを現在準備中である。

もうひとつのマニュアルは、図 3 に示すように、「ほんの、最初の一歩」という部分を記したものだ。このマニュアルにしたがって、記されているコマンドを、一字一句打ち込んでいけば、(マウスでカットアンドペーストしてゆけば) スペクトルの導出ができてしまう、というコンセプトのものだ。この「最初の一歩マニュアル」には DARTS システムの「基本的な検索」から、データを解析サーバに落してから、初めてスペクトルを導出する方法が 3 段階で書かれており、何も知らなくてもスペクトルの導出まではできるようになっている。

このように、ユーザーは、ほぼ何も考えずに、X 線のスペクトルを得られるような状況に来ている。しかし、これと同時に、データベースを作る側としては、「データ処理プロセスをブラックボックスにしない」という点を明確にしようとしている。データ解析は、検出器の下ろしてくる生のデータを自分で、どのような基準で選択し、ひとつひとつ検討してゆくところから始まる。このプロセスに目をつぶり、処理されたデータだけを持ってゆくというのは「解析」などとは言えないだろう。そういった意味で、データベース側は以下のようなところを基本としている。

1. 最も生な形でのデータが、必ず提供されている
2. プロセスされたデータのログが参照可能である
3. データ処理プロセスが再現可能である



http://www.darts.isas.ac.jp:8080/ja/man/asca_analysis/

DEC 上での ASCA データの解析方法

ここからは telnet で DEC に入ってから、どのように解析するかを 記します。UNIX に関する最低限の知識はあるものとして書いています。 御了承下さい。

ここでは、ASCA の「イベントファイル(*1)」を DARTSデータ検索システムからダウンロードし、 解析を行います。例として、 cyg X–1 の gis2 による観測データ、 ad10021000g200170h.evt.Z をダウンロードした場合を記しています。

解析は、

1. 準備
   - ○ データの用意。
   - ○ ソフト起動用パラメタファイルのコピー
2. Xselect
   - ○ 情報の選択 (時間、イメージ領域などの条件による)
   - ○ イメージファイル、スペクトルファイル、ライトカーブ、などの作成
   - ○ バックグラウンドファイルの作成
3. 検出器の応答関数などの用意
   - ○ SIS/GIS などの応答関数のリンク
   - ○ XRT の応答関数の作成
4. Xspec
   - ○ フィッテング

という段階ですすみます。以下、コマンドと、一言のメモを記しますので、 それに従って、コマンドを打ってみて下さい。それぞれのコマンドの詳しい 説明は こちら (英語)にあります。 ここには、プロンプトと、そこに入力するコマンドのみを記します。 途中表示されるメッセージなどは、特別な場合を除き省略いたします。

基本的に、ここに記されているコマンドを cut and paste していくだけで、 スペクトルフィットまでができるようになっています。

図 3: ASCA のマニュアルページ

6 今後の開発

今後の開発としては、以下のようなものがある。

1. X 線天文衛星「ぎんが」データベースの作成
2. 「GEOTAIL」データベースの作成
3. マニュアルの整備
4. コマンドラインによるデータベースアクセス

また、この他に、現在ある「あすか」「ようこう」の改良などもある。特に、「あすか」に関しては、ソースファインディングなどを用いた「rev2」と呼ばれるデータを提供開始する予定である。