

特集・データベース天文学 その(1)

水 本 好 彦

〈文部科学省国立天文台光学赤外線天文学・観測システム研究系／天文学データ解析計算センター

〒181-8588 東京都三鷹市大沢 2-21-1〉

e-mail: mizumoto.y@nao.ac.jp

データベース天文学と聞いてイメージが沸く人はどのくらいいるだろうか。“☆☆天文学”というと電波天文学とかX線天文学の様に観測波長を冠したものと理論天文学のように研究方法を冠したものがある。データベースというのは観測波長ではないから研究方法を冠したものであることは想像がつくが、次に「データベースってなんだろう」とかえて疑問が増す。実際、天文学の研究者のなかでも共通の認識はまだないのではないだろうか。

今月号から始まる特集ではデータベース天文学の現状を紹介するとともに、データベース天文学が今後5年間に急速に発展して新しい天文学を開くだろうという予感と期待をまとめた。

1. アーカイブデータの利用

天文学の基本は観測である。まず空を見る。観測から洞察を得、モデルを作る。そのモデルがどのくらい正しいか観測によって確かめる。物理学や化学だとモデルは実験によって確かめられる。一方、天文学は相手が宇宙だから実験はとてもできない。だから観測に徹してきた。最近では計算機を使った数値実験が大規模に行えるようになったが、その結果の正否はやはり観測で確かめなければならない。もう一つの特徴は同じ観測は二度とできないこと。実験室での実験は条件を整えればいつでも再実験ができるが、天文観測は天体現象が時間変化している可能性がある。従って、観測データは極めて重要である。そのため(かどうかは怪しいが)観測データをアーカイブデータとして大切に保管する習慣がある。

実際、アーカイブデータはいろいろな形で利用されてきた。例えば、シューメーカー・レヴィ第9彗星が1995年に木星に衝突し、その衝突痕跡は1年近くも残っていた。小望遠鏡による眼視観測でもはっきりと確認できたことから、過去の観測例が残っているかもしれない。探してみたらやはり見つか

った(「ついに見つかった衝突痕跡」, 天文月報1997年6月号)。

もう一つの例、おうし座のかに星雲は約1000年前に起こった超新星爆発の残骸である。今いくら詳しく観測してもその正確な爆発の時期を決めることは難しい。そこで過去の文献を調べてみると、中国の歴史書や藤原定家の名月記に1054年の客星として記録されていることは有名である(<http://centaurs.mtk.nao.ac.jp/~avell/history.html>)。このように過去の文献や観測記録を調べる天文研究を、貝塚や遺跡を発掘する考古学に習って天文考古学と呼ぼう。

観測データはそれを観測した天文台がアーカイブするが、20世紀半ばから始まった天文衛星による観測も同様で、データは衛星毎にアーカイブされ公開されている。一度ある目的のために研究されたデータも別の観点から解析することにより新しい結果をもたらすことがある。さらに、複数のアーカイブデータを比較する事により一つだけでは分からなかった天体を検出できることもある。例えば、おとめ座のBrown dwarfなどはPalomar Digitized Sky Surveyの可視光データ(DSS)と2 Micron All Sky Surveyの赤外データ(2MASS)を比較することに

より発見された (Adam J. Burgasser, et al; *Astrophys. J.*, 522, L65-L68, 1999). このように観測波長の異なる複数のアーカイブデータを利用すると、単独では見損なっていたものが浮かび上がってくるという大きな利点がある。このような研究は最近では普通に行われるようになり、アーカイブ天文学と呼ばれている。このようなアーカイブ天文学の誕生は天文観測データの爆発的増大とその公開に因るところが大きい。

2. 観測データの洪水

20 世紀には科学技術の進展に伴い、電波天文学や X 線天文学など新しい天文学が次々と誕生した。一般に新たな観測波長で宇宙を観測すると、何が見えるか予想がつかない。そこで、天球の広い領域にわたってべたっと観測するサーベイ観測が行われる。特に初期の高エネルギー天文衛星では、観測装置の角度分解能は悪いが比較的広視野という特徴を生かし全天サーベイ観測が行われ、それにより新しい天文学が生まれてきた。電波望遠鏡と天文衛星の観測はビッグバン宇宙論を導き、人類の宇宙観を一変させた。さらに、詳細な光学観測により宇宙の大規模構造が発見された。広範囲にわたる観測データの積み重ねによって宇宙論の発展がもたらされた。このように、サーベイ観測は天文学の王道である。また、ハッブル宇宙望遠鏡によるハッブルディープフィールドというサーベイ観測はその高い角度分解能と観測データの迅速な公開により天文学の世界に大きなインパクトを与えた。

近年の天文観測の急速な発展の原因はデジタル信号処理技術の進歩にある。写真乾板に代わる CCD と計算機の登場は光学観測に一大革命をもたらした。この 30 年あまりの半導体技術の進歩を支配するムーアの法則 (半導体の集積度や速度は 18 ヶ月ごとに 2 倍になる) は CCD の進歩についても成立する。例えば、すばる望遠鏡の主焦点カメラは 2K × 4K の CCD チップ 10 個を並べた 8 千万画素の性能を持ち、一回の撮像で視野 24' × 30', サ

イズ 160MB の画像が得られる。このような巨大 CCD の登場により、観測データは爆発的に増加している。それに比べて研究者のデータ処理能力は一定であるため、たとえ高速の計算機を用いたとしても、一画像毎に処理結果を目で見て確認していたのではデータ解析処理が追いつかない状況になっている^{*脚注1}。一方、口径 10 m クラスの大型光学赤外線望遠鏡が次々と登場する中で、数年前まで最高性能を誇っていた 4 m クラスの望遠鏡をはじめ中小望遠鏡によるサーベイ (SuperCOSMOS, Sloan Digital Sky Survey, 2Micron All Sky Survey, DPOSS, UKIDSS (3.8 m UKIRT IR wide-field camera), VISTA (4 m wide field telescope) によるサーベイなど) や次世代天文観測衛星 (XMM-Newton, PRIME, MAP, FIRST など) では大規模サーベイ観測の時代を迎え、更なるデータの洪水が予想される。次世代大規模サーベイの特徴は、数 TB という規模の大きさ以外に、データアーカイブの公開が計画の一部として含まれていることにある。ムーアの法則に従えば 1 年半でアーカイブデータの量は 2 倍になる。従って、サーベイ観測のデータが 1 年半後に公開されるとすると、研究者は世界中に蓄積された観測データの約半分をアーカイブデータとして常にご利用することができるようになるのである。

3. 数値宇宙のパラダイム

CCD の誕生が天文観測データの洪水をもたらしたように、計算機の進歩が天文観測の手法そのものを変える可能性がある。アーカイブデータは観測された宇宙の数値データの集合である。これを多次元メモリー空間に射影・記憶された「数値宇宙」

* 脚注 1 天文観測は、肉眼 → 筒を覗く → 筒の両端にレンズを付けた望遠鏡 (17 世紀はじめから) を覗く → 望遠鏡に写真乾板をつけて写真を撮り (20 世紀前半) それを見る → 望遠鏡に CCD カメラをつけてデジタル写真を撮り (20 世紀後半から) それを計算機で処理して見る、というように変わってきた。次はどうなるのだろうか？

だと考えよう。すると望遠鏡や観測装置は実宇宙を数値宇宙に射影する装置といえる。観測装置によって得られた歪んだ画像を較正し物理パラメータを求めるというデータ解析作業は、数値宇宙から求める天体データを探し出し、歪んだ数値宇宙データを歪みのない数値宇宙データに変換する「一般化望遠鏡」と、歪みのなくなった数値宇宙データから物理パラメータを計測する「一般化観測装置」と定義し直せる。数値宇宙に含まれる観測データはサーベイデータのように均質で質の保証されたデータであることが必須条件であるが、電波からガンマ線まであらゆる種類の観測データが含まれ得る。

この数値宇宙パラダイムでは、天文観測は数値宇宙に対して行われる。もしも数値宇宙の目的の領域が空(=「から」：観測データがまだない)なら、実望遠鏡が実宇宙を観測し、観測値を数値宇宙に加える。一流の観測天文学者になるためには望遠鏡による観測経験とデータ解析の訓練が必要であるといわれているが、数値宇宙観測では一般化望遠鏡によるデータベース検索が観測であり、特別な経験や訓練は不要となる。もちろん実望遠鏡による実際の観測は経験を積んだ少数の専門家がデータの質を保証して行う必要はある。天文衛星による観測ははじめからこのようになっている。現実には、すばる望遠鏡のような大望遠鏡や天文観測衛星で目的の天体観測を行うための観測時間獲得の競争は厳しい。

数値宇宙にアクセスする計算機ネットワークが天文台で、データベースがいつでも使える望遠鏡であるという世界が現実になろうとしている。すでに述べたように、新しい観測データが続々と公開されている。しかし、データアーカイブの量が多くなるとそこから自分の必要とするデータを探すのが一苦勞で、さらに、それを処理するにはデータ量が多過ぎてそのデータの保存されているところに行かなければできないということもある。この問題を最新の計算機・通信・情報

技術を使って解決し数値宇宙と一般化望遠鏡を実現しようというのが、仮想天文台(VO: Virtual Observatory)である。米欧では仮想天文台の建設がすでに始まっている(アメリカのNVO(<http://www.srl.caltech.edu/nvo/> 参照)、ヨーロッパ南天文台のAVO(<http://www.eso.org/projects/avo/> 参照)など)。日本では国立天文台が中心となり、数値宇宙パラダイムの実現を目指して検討が行われている。

4. データベース天文学の世界

一般化望遠鏡と一般化観測装置は観測波長を限らないし、数値宇宙もあらゆる波長の観測データを含む。そこで、数値宇宙を観測すれば、赤外線天文学とか電波天文学とかはもちろんのこと「光学赤外線電波天文学」とか「電波 X 線天文学」とか複数の波長に渡った天文学の研究が容易にできるようになる。こうなると多波長のデータを使った研究方法が当たり前になり、それをただの「天文学」と呼び、狭い波長領域のデータだけを使った研究を例えば「可視天文学」とか呼ぶ時代が来るかもしれない。その中間の時期にあるのが「データベース天文学」だろう。

はじめに述べたように、新しい波長で宇宙を観測すると何が見えるか予想がつかない。データベース天文学も同じで、どんな発見があるのか分からないから「わくわく」する。データベース天文学の研究は

異なった波長で、

異なった装置で、

異なった時間に

観測された異なったデータの膨大な塊(数値宇宙)の中からあるパターンを探すという方法を採用。これは金や石油の鉱脈探し(探鉱)になぞらえて「データマイニング」と呼ばれる。その一つの方法は、星の分類に用いられる HR 図(星の色(温度)と明るさの 2 パラメータの分布図)を拡張したものである。多数の天体について様々なパラメータ(n 種類)を測定し、その n 次元パラメータ空間の中

での天体の分布を調べる。分布図^{*脚注2}の中でデータ点が密集した塊（クラスター）の形や種類、塊にならないデータ点などを調べ、天体の分類を行う。

例えば、同じ領域の長期間にわたる観測データの中から明るさや位置の時間変化を調べることで、高赤方偏移の Hypernova やガンマ線バースト天体の探索、クエーサーの時間変化の研究、マイクロレンジングや系内惑星、太陽系近傍天体の探索などができる。天体の2つの色の分布図からは高赤方偏移のクエーサーや、冷たい白色矮星、褐色矮星など稀な天体の探索ができる。複数の色での天体位置の相関から cosmic string（宇宙紐）探索などもできる。

このような研究では大量データの統計処理を駆使するとともに、 n 次元パラメータ空間内のデータ点分布の特徴を分析するための道具が必要となる。人間の高度な空間認識能力を生かすための、 n 次元パラメータ空間を2次元または3次元に投影して見せる可視化技術が威力を発揮するだろう。実際に n 次元パラメータ空間の3次元投影の中に入

り込んでデータ点の分布を詳細に観察することができ。子供の頃に屋根の上に寝ころんで夜空を見上げると星空の中に自分が浮かんでいる様な感じを経験された方も多いただろう。データベース天文学者は数値宇宙という仮想空間に入り込んで観察するとき、ふと懐かしい想いが甦るのではないだろうか。データベース天文学によって、宇宙年齢のような宇宙論パラメータを膨大なデータの統計処理で正確に決定するような天文学の精密化^{*脚注3}と、未知天体の発見ラッシュが期待される。

Database Astronomy

Yoshihiko MIZUMOTO

National Astronomical Observatory, Mitaka 181-8588, Tokyo, Japan.

Abstract: A database astronomy era will come soon. Astronomers of the near future will use large astronomical databases as virtual telescopes.

The huge astronomical databases could be a digital universe, capable of looking at the whole sky at once, over a wide range of electromagnetic wavelength.

*脚注2 分布図といっても n 次元だから平面図ではない。

*脚注3 「大雑把」という意味での「天文学的」という言葉が死語になる？