TITLE
*ASP Conference Series, Vol. **VOLUME**, **PUBLICATION YEAR***
EDITORS

Computational Astrophysics Tools for the GRID

C. Simon Jeffery

Armagh Observatory, College Hill, Armagh BT61 9DG, Northern Ireland

Abstract. The newest generation of telescopes and detectors and facilities like the Virtual Observatory (VO) will deliver enormous volumes of astronomical data. The analysis of these data relies heavily on computer-generated models of growing sophistication and realism. To carry out simulations at increasingly high spatial and temporal resolution and physical dimension and in a growing parameter space will require a new approach to data modelling. A proposal to develop toolkits to faciliate such modelling and to integrate them within a heterogeneous network (such as the VO) is described.

1. Inroduction

The proposals outlined in this poster have been developed by a team comprising the author, Dr Robert J. Allan (CLRC), Dr Martin A. Barstow (Univ. of Leicester), Prof Tom J. Millar (UMIST), Dr Juliet C. Pickering (ICSTM) and Dr Jeremy Yates (UCL), together with other associates, involved in a range of science projects:

- measuring the properties and dynamics of plasma, fields and particles in the near-Sun heliosphere,
- investigating the fine-scale structure and dynamics of the Sun's magnetised atmosphere,
- the structure and evolution of stellar atmospheres other than the Sun,
- probing the atmospheres of stellar remnants including white dwarfs, hot subdwarfs, helium stars and R CrB stars,
- the structure and chemistry of hot molecular cores,
- shock chemistry associated with protostellar jets,
- the evolution of asymptotic giant branch stars and planetary nebula,
- photochemistry in the envelopes of AGB stars and proto-planetary disks,
- the processing of dust from AGB stars through the ISM and protostellar cores to planets,
- the hydrodynamics of planetary atmospheres,
- and the excitation of cosmic masers.

Our goal is to develop software tools and a computational framework that will facilitate modelling astronomical data in large parameter spaces using widely distributed and heterogeneous computer systems. It builds on experience and resources acquired during the lifetime of the Collaborative Computational Project No. 7 for the Analysis of Astronomical Spectra (Jeffery 1996).

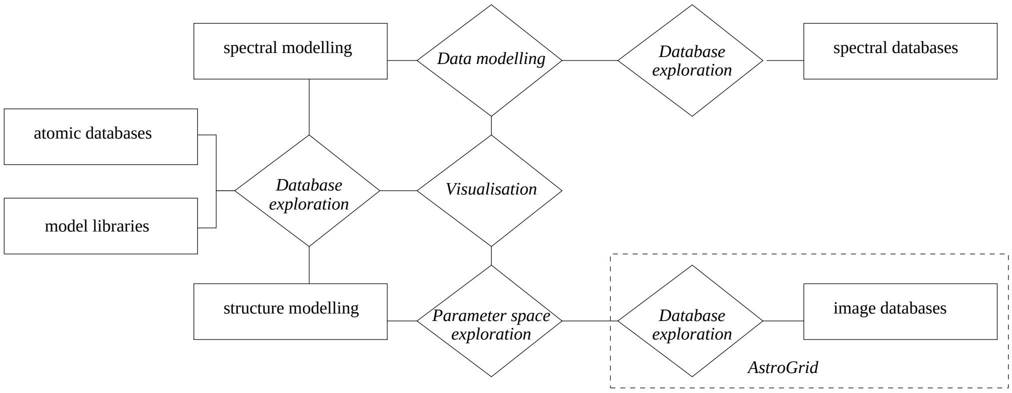

Figure 1. Schematic view of relation between theoretical models, atomic data and spectroscopic and imaging data in a computational grid.

2. Interpreting Data

The interpretation of nearly all astronomical phenomena involves fitting a predictive model to quantitative data by some process – χ^2 minimization or data inversion, for example. This could involve combining new observations with data from established archives or new databases. A goal-seeking algorithm uses this data to constrain a sequence of theoretical calculations. In turn these simulations may need to retrieve data from other sources, atomic data, pre-computed model atmospheres and so on. Already, data resources are distributed and computations are frequently done remotely from the scientist.

The principal elements of the problem to be solved are the fitting of multi-parameter theoretical models to multi-element observational data. A key application is the modelling of astronomical spectra. These may contain several thousand elements (in frequency, time and space). Spectra are produced in media in which the temperature, density, velocity and chemical structure must be determined empirically – these are the 'free' parameters in the model. Other examples include the structure of astrophysical jets, or modelling the light curves of active galactic nuclei.

The general modelling problem will have many of the following properties. The parameter space may be very large, but only specific volumes will be astrophysically interesting. The models may be linear in some parameters, but non-linear in others. Theoretical models require ever increasing cpu, RAM and data resources at geographically remote locations. The physical sizes of individual models or grids of models are increasing, exceeding tens of gigabytes in many cases. Distributed computer systems on a cluster or intranet allow different regions of parameter space to be searched at the same time, but monitoring and guidance require overall control systems and, possibly, human intervention.

In the classical approach, a single task may combine model construction, data organisation and goal-seeking algorithms. The increasing sophistication of the model construction process and the need to outsource data organisation (e.g. atomic data, satellite archives) suggests that this will not be sufficient.

The overall procedure may need to be broken down into several tasks (Fig. 1), and distributed across a computational grid.

3. Automated Analysis

A decade ago, classical methods for the fine analysis of a high-resolution stellar spectrum would take several months of effort; even the process of spectral classification could be time consuming. Multi-object, 2D and echelle spectrographs, high-speed detectors, high spectral resolutions, heterodyne receivers, and long-running survey programmes are generating ever-increasing volumes of high-quality astronomical spectra. Gains in S/N provided by large-aperture ground-based telescopes are complemented in the UV and X-ray by a variety of space missions (e.g. HST, FUSE, Chandra, XMM-Newton). Recent technological developments promise spectral resolutions at X-ray and EUV wavelengths to rival those routinely provided by UV and optical observatories. Array detectors and interferometers being developed for the submillimeter will enable the 3-D exploration of interstellar clouds. Datasets may contain up to tens of thousands of frequency points, a few hundred positions in 1 or 2 spatial dimensions, tens to thousands of snapshots in time, or spectra of a few thousand objects from a single night's observing. Many scientific goals (e.g. identifying solar analogues in the Galaxy, measuring lithium abundances in young low-mass stars, distinguishing brown dwarfs from free-floating planets) require the analysis of large samples. Establishing precise astronomical dimensions from the study of time-dependent phenomena (e.g. planets, active galactic nuclei, pulsating stars, binaries) may require the analysis of long series of spectra with wavelength coverage from the infrared to the X-ray. Under these circumstances, extracting anything but trivial information on a competitive timescale now demands a revolution in the methods of spectral modelling.

We have piloted a scheme to automatically extract effective temperatures, surface gravity, helium abundance, radial velocity and interstellar reddening from a series of moderate-resolution spectra of a variable star (Jeffery, Woolf & Pollacco 2001). Extending this to extract chemical information for more than one spectrum was prohibitive with the resources available. The classical solution would be to move the problem onto a faster computer. That offered by grid technology would be to distribute specific tasks across a network, allocating tasks to optimized architectures and available resources. In the face of imminent data inundation, the scientific motivation for automating the data modelling process is compelling.

4. Dynamic Modelling

In addition to observational data, theoretical simulations regularly use libraries of atomic data and of pre-calculated grids of models. Such data libraries are characteristically different to observational data and lie outside the interest of most contemporary projects. Their significance is that in a computational grid, applications may be designed to access or save data in remote libraries as required, thereby automating some of the most laborious steps in a scientific investigation.

4.1. Atomic and Molecular Databases.

Most astrophysical models require atomic and molecular data. There is a confusing array of experimental and theoretical data of varying quality and completeness. This costs astronomers dearly in terms of wasted time and resources and may often result in a flawed analysis of an astronomical spectrum which may have been obtained at great expense. The solution is to develop methods using metadata to allow an astronomer or, better, an application to automatically access the most appropriate databases around the world in real time. These would select the best available data for a particular problem and take advantage of the very latest, state-of-the-art measurements or calculations. New tools would be required to access databases of varying formats and to handle large quantities of data efficiently. Standards for nomenclature, units and file formats would be required. Where search tools fail to find specific atomic data, an automatic request (e.g. from an R-matrix calculation) could be generated and used to update relevant databases. Thus astrophysical requirements could lead to directly to the generation of required atomic data 'on demand'.

4.2. Model libraries.

In general, astrophysical models have a much shorter lifetime than astronomical observations; obsolescence is an inevitable result of progress. However, the computational cost of producing many models remains large compared with that of data storage and remote retrieval. When computed for a sufficiently general volume of parameter space, precomputed grids of models provide a vital resource for interpreting data. A simple example is the Kurucz database of model structures, flux distributions and synthetic colours for stellar atmospheres (Kurucz 1991 and subsequent), which have found applications throughout astronomy. These might be supplemented by databases of high-resolution flux and intensity spectra for normal stars. Similar databases might subsequently be generated with non-standard compositions or improved atomic data.

Applications in a computational grid need to know about such model libraries and have tools to explore them and to introduce new data, especially where these extend the volume of parameter space represented by existing models.

5. Project Realisation

The ideas outlined in this poster are challenging, but they are widely recognised as important to the future of astronomical research. The project team are exploring means and resources by which their goals may be realised.

References

Jeffery C.S. 1996, QJRAS, 37, 39

Jeffery C.S., Woolf V.M., & Pollacco D. 2001, A&A 376, 497

Kurucz R.L. 1991, Precision Photometry: Astrophysics of the Galaxy, ed. A.G.Davis Philip, A.R.Upgren, & K.A.Janes (L.Davis Press, Schenectady)