Information Handling in Astronomy - Historical Vistas
Ed. A. Heck, (c) Kluwer 2003

FOREWORD

This book is dedicated to the memory of Gisèle Mersch whose life ended prematurely in June 2002. Back in the 1970s, when few people were using them, Gisèle introduced me to the arcane secrets of then advanced multivariate statistical methodologies.

I was already involved in more classical statistical studies undertaken at Paris Observatory with Jean Jung: developing and applying maximum-likelihood algorithms to stellar photometric and kinematic data in order to derive absolute luminosities, distances and velocities in the solar neighborhood.

But what could be envisaged with those methodologies was something of another dimension: for the first time, I could really see how to extract information from massive amounts of data without calling for elaborated physical or mechanical theories.

Several pioneering applications were developed under Gisèle's guidance and with her collaboration to study the delicate interface between spectroscopic and photometric data. Thus errors in spectral classifications were investigated as well as predictions of spectral classifications from photometric indices (see Heck 1976, Heck *et al.* 1977, Heck & Mersch 1980 and Mersch & Heck 1980), with very interesting results for the time. Gisèle also took part in studies of period determination algorithms (see Mersch & Heck 1981, Manfroid *et al.* 1983 and Heck *et al.* 1985).

Gisèle's generosity, patience and dedication were impressive. She had set up a statistical consultancy service for the other departments at the University of Liège, Belgium. She would often tell the following anecdote which is full of lessons worthy of considerations by students.

One day, she was approached by someone from the human sciences. That gentleman, who obviously knew little of the elementary mathematical problematics, brought her a case study with n observations and m unknown 'parameters' to be determined, with $n < m$. Gisèle kindly explained him that, in such a situation – less observations than variables – she could not do anything. He had to collect a bigger sample of observations if he wanted

the case to be solved. How could she dare! He started threatening to file a complaint with her boss and even higher up in the University if she was to persist in such a non-cooperative attitude.

Shared between offence, compassion, and a strong need to laugh, Gisèle kept however her best face and said that, in such conditions, she had indeed no choice. She invited the arrogant gentleman to come back a couple of days later. After he left, it took her five minutes to write a short Fortran program printing in huge characters on one of those large pages of the computer printout in usage at that time:

"The case has too many unkown parameters
for the number of observations.
It cannot be solved."

You have certainly guessed the end of the story. When he came back, the gentleman had no difficulty to accept the verdict of the machine. It was pure truth since the computer had said it.

Also for students, we often took as an example the paper by Heck *et al.* (1977) where four mathematicians working in different disciplines (astrophysics, medicine, psychology and statistics) collaborated efficiently on a single project: once agreement on the vocabulary used had been reached (for instance, the term 'parameter' did not mean initially the same thing for everybody), the intellectual processes and statistical procedures were the same whether the individuals dealt with were stars, cancer patients or laboratory rats.

Those investigations were expanded later on and other methodologies were investigated with other partners (see *e.g.* Murtagh & Heck 1987, Heck & Murtagh 1989 and Heck & Murtagh 1993), always with the same fascination Gisèle had lit up. Such studies were the forerunners of today's data mining and knowledge building methodologies.

It should be kept in mind that these were never intended to replace physical analysis. They should be seen as complementary, useful to run rough preliminary investigations, to sort out ideas, to put a new ('objective' or 'independent') light on a problem, or to point out aspects which would not come out in a classical approach. Physical analysis is necessary to subsequently refine and interpret the results, and to take care of the details. Nowadays, with many 'virtual observatory' projects dealing with huge amounts of data, those intellectual investments of the past are more than ever justified.

*****

This book completes, with emphasis on history, an earlier volume entitled *Information Handling in Astronomy* and published in the same series (Heck 2000).

After a few considerations by the Editor on the evolution of astronomical data and information handling methodologies in the second half of the last century, E. Biémont reviews how the measurements of time, a fundamental parameter for our science, evolved over ... time.

Several chapters are then devoted to astronomical data processing, starting with a personal account by R. Albrecht followed by contributions centered on specific systems: IHAP (P. Grosbøl & P. Biereichel), FITS (E. Greisen), MIDAS (K. Banse) and AIPS (E. Greisen).

We then move to publications-oriented chapters, by H.A. Abt (Editor) and B. Corbin (Librarian) while G. Eichhorn recalls the development of the Astronomy Digital Library. Next, A. Heck reviews the evolution from early 20^{th} century directories to current online yellow-page services.

Two chapters then deal with education, first by J.E. Bishop on pre-college astronomy education in the US, then by C.C. Petersen on the rôle of planetariums.

Then A. Batten and D. McNally discuss the changing rôle of the *International Astronomical Union* in providing and organizing information, followed by D.H.P. Jones discussing a sometimes controversial matter: the impact of the *Carte du Ciel* project on the development of astrophysics in Europe, and thus on the collection of related data on that continent. The book concludes with a review by S. Dunlop of amateur data and discoveries in the 20^{th} century.

It has been a privilege and a great honor to be given the opportunity of compiling this book and interacting with the various contributors. The quality of the authors, the scope of experiences they cover, the messages they convey make of this book the natural complement of the first volume.

The reader will certainly enjoy as much as I did going through such a variety of well-inspired chapters from so many different horizons, be it also because the contributors have done their best to write in a way understandable to readers not necessarily hyperspecialized in astronomy while providing specific detailed information, as well as plenty of pointers and bibliographical elements. Especially enlightening are those 'lessons learned' sections where authors make a critical review of the experience gained.

It is also a very pleasant duty to pay tribute here to the various people at *Kluwer Academic Publishers* who quickly understood the interest of such a volume and enthusiastically agreed to produce it.

Special thanks are due to Artist C. Gerling whose 'Emergence of Knowledge' (2002) illustrates the cover of this volume.

André Heck
Picos de Europa
November 2002

References

1. Heck, A. 1976, An Application of Multivariate Analysis to a Photometric Catalogue, *Astron. Astrophys.* **47**, 129-135.
2. Heck, A. (Ed.) 2000, Information Handling in Astronomy, Kluwer Acad. Publ., Dordrecht, x + 242 pp. (ISBN 0-7923-6494-5).
3. Heck, A., Albert, A., Defays, D. & Mersch, G. 1977, Detection of Errors in Spectral Classification by Cluster Analysis, *Astron. Astrophys.* **61**, 563-566.
4. Heck, A., Manfroid, J. & Mersch, G. 1985, On Period Determination Methods, *Astron. Astrophys. Suppl.* **59**, 63-72.
5. Heck, A. & Mersch, G. 1980, Prediction of Spectral Classification from Photometric Observations – Application to the uvbyβ Photometry and the MK Spectral Classification. I. Prediction Assuming a Luminosity Class, *Astron. Astrophys.* **83**, 287-296.
6. Heck, A. & Murtagh, F. 1989, Knowledge-Based Systems in Astronomy, Springer-Verlag, Heidelberg, ii + 280 pp. (ISBN 3-540-51044-3)
7. Heck, A. & Murtagh, F. 1993, Intelligent Information Retrieval: The Case of Astronomy and Related Space Sciences, Kluwer Acad. Publ., Dordrecht, iv + 214 pp. (ISBN 0-7923-2295-9)
8. Manfroid, J., Heck, A. & Mersch, G. 1983, Comparative Study of Period Determination Methods, in *Statistical Methods in Astronomy*, **ESA SP-201**, 117-121.
9. Mersch, G. & Heck, A. 1980, Prediction of Spectral Classification from Photometric Observations – Application to the uvbyβ Photometry and the MK Spectral Classification. II. General Case, *Astron. Astrophys.* **85**, 93-100.
10. Mersch, G. & Heck, A. 1981, Preliminary Results of a Statistical Study of Some Period Determination Methods, in *Upper Main Sequence CP Stars*, 23rd Liège Astrophys. Coll., 299-305.
11. Murtagh, F. & Heck, A. 1987, Multivariate Data Analysis with Astronomical Applications, Kluwer Acad. Publ., Dordrecht, xvi + 210 pp. (ISBN 90-277-2425-3).