Final Report GR/M11370
Materials Algorithms Project (MAP)

ABSTRACT

The Materials Algorithms Project (MAP) website[1], which was started in February 1997 as a centre for the "validation" and distribution of algorithms of use in the modelling of materials, has been developed and expanded. It originally distributed only software written in FORTRAN. This has now been extended to include software written in any standard programming language or programs provided only as executable files. Two new libraries along with many new programs and subroutines have been added to the website, giving access to over 190 programs or routines and 24 materials data-sets. A search engine has now been made available for use on the website, providing the user with the facility either to make a general search of the web pages, or to carry out a more specific search of the keywords which describe the purpose of the software. Maintainance of the website has been semi-automated, with shell script programs making regular updates of the access statistics (daily) and the compressed tar file of the entire library (weekly). Rationalization of the file system on the website, and the formation of standard template files and folders for the creation of new libraries and web-pages will greatly facilitate further expansion of the MAP libraries. A mirror site in North America[2], courtesy of QuesTek Innovations LLC., has now been fully established. Despite its presence, access statistics show that the usage of the MAP website has doubled, both in number of users and in number of requests since the start of this project in February 1999, reaching over 20,000 requests for webpages and 2200 distinct hosts served worldwide during the month of November 1999.

1. INTRODUCTION AND ORIGINAL OBJECTIVES

The Materials Algorithms Project (MAP)[1] was initially created as a perpetual electronic library of FORTRAN programs and routines containing documented algorithms for the modelling of materials. The inclusion of properly documented software in the library by researchers has enabled the subject of materials modelling to grow by avoiding repetition and by avoiding the loss of information when modelling projects end. Although not originally intended, this was extended to include experimental and theoretical data on materials properties. The software and data libraries have been freely accessible to anyone via the world wide web since February 17^{th} 1997. During the first 9 months about 29,000 files were accessed (approximately 3200 per month). More recent access statistics (figure 1) show that during 1998 this had increased to about 8000 per month. Since the start of the current project in February 1999, to develop the MAP website further and increase the amount of software available, there has been a dramatic increase in the number of requests for files made to almost 20,000 a month by November 1999. This increase is also reflected in the number of compressed files downloaded, typically between 5 and 10% of the total number of requests for files made. The number of distinct hosts accessing the MAP website has also doubled from about 1000 in 1998 to over 2000, which shows a growing interest in the software provided by MAP.

MAP was set up originally as a result of funding from the EPSRC (GR/J98073). Additional funding was sought for further development of the project (GR/M11370) with the follow aims:

1. The inclusion of software written in the C programming language in the MAP software libraries.
2. The incorporation of a search engine for the MAP web site.
3. The documentation and inclusion of a large quantity of our own remaining public domain software and data.
4. The automation of the submission of documentation and software. This would reduce the magnitude of the task involved in adding new software to the MAP website and go a long way towards making it self-sustaining.

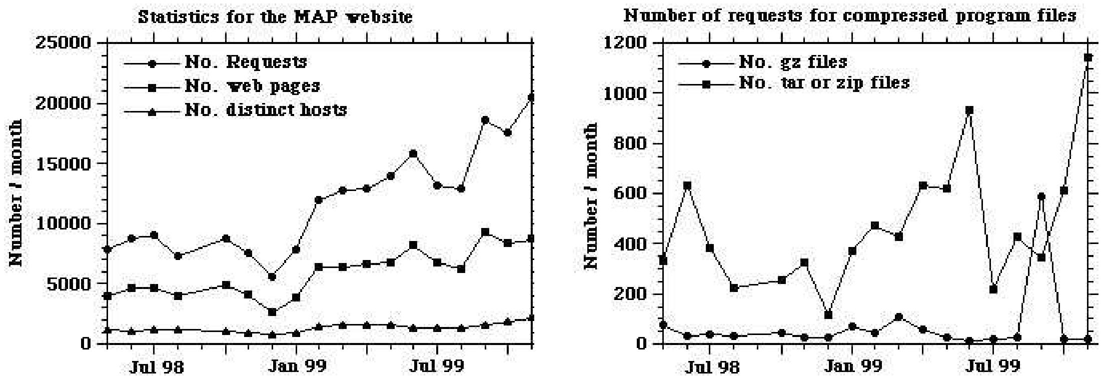


Figure 1. Statistics for the MAP website from May 1998 to November 1999.

2. THE MAP LIBRARIES

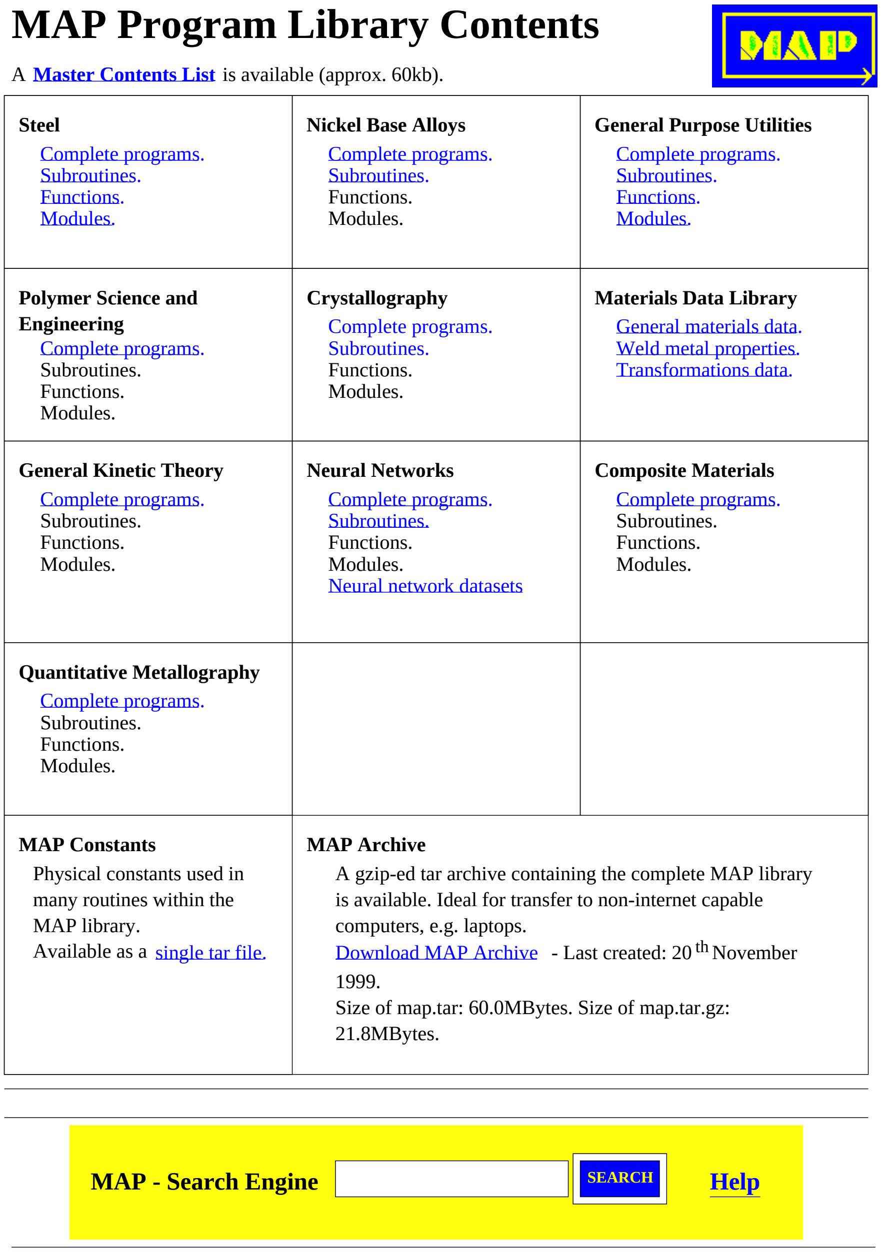
MAP incorporates a number of distinct libraries which contain software or data which are subject-related. Access to these libraries is through the MAP Program Library Contents page shown in figure 2. Each library contains sub-libraries for complete programs, subroutines, functions and modules. The links (blue) on the Library Contents page access the index pages of the sub-libraries. These index pages list the names of the routines/data-sets available, along with a short description of their purpose, and provide a link to the individual web pages containing detailed information about their function, downloading and use. Only those sub-libraries with active links on the Program Library Contents page currently contain software. During the last year a number of general changes and improvements have been made to the MAP website:

2.1 The incorporation of software in programming languages other than FORTRAN

Although the initial intention had been to create separate libraries for FORTRAN and C code, it was decided that it would be much simpler to maintain the subject-related libraries, already present in MAP, and open them up to computer code written not only in C but in *any* standard computer programming language. This involved only a minor change to the titles on the web pages - the phrase FORTRAN Library was changed to Program Library. It then became necessary, however, to state explicitly for each entry in the MAP libraries the programming language of the software. This was done by including a clear statement both on the index page of each sub-library and in the specification section of the web page for each program or routine. The library was also opened up to software which is available only as an executable file (provided it contains no commercial software). In this case details of the hardware and operating system required to run the program are also given in the specifications section.

2.2 The inclusion of new libraries and sub-libraries

In the past, the files, directories and web pages required to install a new library or sub-library had been created only as required. This was generally a somewhat timing-consuming task. We have, however, now set in place the entire file structure for the MAP web-site, with web pages for all the index files of the libraries and sub-libraries shown on the Library Contents page. A *template* directory, containing all the necessary files and sub-directories required for adding a new library to the website, has also been created; this will reduce considerably the amount of work required for future extensions of the website. Two new libraries have been added to MAP for programs relating to quantitative metallography (Quantitative Metallography Library), and for the modelling and analysis of properties and phenomena associated with composites, laminates and coatings (Composite Materials Library).



MAP Program Library Contents

A Master Contents List is available (approx. 60kb).

Steel Complete programs. Subroutines. Functions. Modules.	**Nickel Base Alloys** Complete programs. Subroutines. Functions. Modules.	**General Purpose Utilities** Complete programs. Subroutines. Functions. Modules.
Polymer Science and Engineering Complete programs. Subroutines. Functions. Modules.	**Crystallography** Complete programs. Subroutines. Functions. Modules.	**Materials Data Library** General materials data. Weld metal properties. Transformations data.
General Kinetic Theory Complete programs. Subroutines. Functions. Modules.	**Neural Networks** Complete programs. Subroutines. Functions. Modules. Neural network datasets	**Composite Materials** Complete programs. Subroutines. Functions. Modules.
Quantitative Metallography Complete programs. Subroutines. Functions. Modules.		
MAP Constants Physical constants used in many routines within the MAP library. Available as a single tar file.	**MAP Archive** A gzip-ed tar archive containing the complete MAP library is available. Ideal for transfer to non-internet capable computers, e.g. laptops. Download MAP Archive - Last created: 20 th November 1999. Size of map.tar: 60.0MBytes. Size of map.tar.gz: 21.8MBytes.	

MAP - Search Engine SEARCH Help

Figure 2. MAP Program Library Contents page.

2.3 The data library

Each entry in the MAP library is given a unique identifying name of the form *MAP_LIB_NAME*, where *NAME* is the authors own name for the program or data-set and *LIB* is the name of the library. The nine software libraries have the names COMP, CRYSTAL, KINETIC, METALL, NEURAL, NICKEL, POLY, STEEL and UTIL. The names of the data-sets in the data library, however, had been of the form MAP_NAME. In order to standardize the naming system and distinguish more clearly between data-sets and software, the names of the entries in the Data Library have been changed to the form MAP_DATA_NAME.

2.4 Standard HTML programming language

The programming language used for creating the web pages on the MAP website is HTML 3.2 (Hypertext Markup Language version 3.2). The W3C HTML Validation Service (at http://validator.w3.org/) was used to check all the web pages on the MAP website for conformity with this standard. Any programming errors and non-conformity with the HTML 3.2 standard were removed. All new pages added to the site were also checked in this way.

2.5 Other minor changes

Several minor changes were made to the format of the documentation given on the web pages for each library entry. Namely, the addition of the date of submission of the code to MAP, a contact e-mail address to the authors of the code (optional), and a *Keywords* section for entries in the Data Library (already present in the software libraries). The first two of these minor alterations, however, could not be introduced retrospectively for software and data-sets already present on the MAP website, but were implemented for most of the additions carried out during 1999.

3. THE INCORPORATION OF A SEARCH ENGINE

A search engine has now been successfully implemented for carrying out a search of the files in the MAP webspace. In the interests of maintaining security from network hackers, it was deemed unsafe to provide MAP with its own search engine; instead we have made use of the search engine run by the University of Cambridge. Apart from carrying out a simple search of the files on the MAP website for the words listed by the user, it also has the capability of restricting its search to the titles of each web pages, or the contents of other labels (*meta tags*) which are placed in the header section of each web pages. In order to make full use of this facility, this section of all the web pages on MAP was expanded and standardized. An example of this header section is as follows:

```
<head>
 <title>MAP Subroutine MAP_STEEL_AVOLF</title>
 <meta name="author"   content="Bhadeshia, Parker">
 <meta name="keywords" content="allotriomorphic ferrite, heterogeneous
            nucleation, volume fraction, area fraction, continuous cooling">
 <meta name="description" content="To calculate the volume and area fractions
            of allotriomorphic ferrite formed during continuous cooling
            transformation from austenite, for heterogeneous nucleation.">
 <meta name="reply-to" content="map@msm.cam.ac.uk">
 <meta http-equiv="Last-Modified" content="Fri, 3 Sept 1999 9:45 GMT">
</head>
```

The title specifies the name and type of routine or data; it's format was standardized. The authors specified in the *author* meta tag are the authors of the code or data and not, as previously, the authors of the web page. The *keywords* and *description* meta tags had not previously existed and were introduced into the header section of all the web pages on MAP. The contents of the *keywords* meta tag are descriptive of the contents of the web page; in the case of library documentation web pages, the keywords given here are generally the same as those given in the keywords section of the documentation itself. Similarly, the *description* meta tag is generally the same as that given in the section *Purpose of code* on the documentation web pages. The remaining two meta tags remain unchanged and give the e-mail address of the MAP Administrator and the date the page had last been modified.

The search engine can make a restricted search of the first four of these meta tags in the header section. For example a search of the keywords for the word 'weld' can be made by typing *keywords: weld.* This use of the search engine is explained with examples in a small help page, accessible from a link next to the search engine itself (see figure 2). The inclusion of these meta tags also makes the results of any search much more helpful; they now show the title and description for each web page found.
e.g. the seach string *title: COOLCU* gives the result:

MAP Subroutine MAP_STEEL_COOLCU
Calculates the cooling curve for the fusion zone of a steel weld deposit.
http://www.msm.cam.ac.uk/map/steel/subs/coolcu-b.html - size 10.2K

4. AUTOMATION OF THE MAP WEBSITE

One of the aims of this project, was to make the web-site as self-sustaining as possible, thereby reducing the amount of manpower required to maintain and make new additions to the website. This was achieved in the following ways:

4.1 Automation of the access statistics

A shell script program was created, which is run automatically each night to produce a file showing the latest access statistics. The program also updates the MAP web page `/map/access.html`, which contains a link to these statistics. The file produced on the last day of each month is kept as a record of the access statistics for that month. The need for someone to run the access statistics program and save the results to disc on the last day of each month has now been completely removed.

4.2 Creation of a compressed file containing the entire MAP library

A zipped tar file containing the entire library had already been available. It had been updated only spasmodically. This process of updating has now been automated by running a shell program once a week. Each week the old file is deleted, a new one created and the Program Library Contents page (figure 2) updated to show the size and creation date of the new zip file.

4.3 Submission of code/data and documentation

The original proposal suggested that the submission of new code or data and its accompanying documentation, for inclusion in MAP, could be automated. Firstly it should be emphasised that, for security reasons, the downloading of files onto the website by external users cannot be comtemplated. It could have been possible to produce a web page containing a form which would need to be completed with the appropriate information for creating the required MAP documentation. This information could then have been submitted by sending it via e-mail to the MAP Administrator. It was decided, however, not to do this for the following reasons:

(a) Experience has shown that the submission of forms via e-mails frequently fails, often due to an incorrectly configured computer/web browser. Consequently much time can be wasted filling in the form, only to lose all the information after attempting to e-mail it and finding that the web-browser has failed to do so.

(b) The text of the web pages and other documentation generally contain mathematical expressions or equations. It is highly desirable that contributors do continue to give mathematical details related to their software. It would be difficult, however, to submit mathematical expressions or equations simply by filling in boxes on a form using plain text.

It was therefore decided to provide well commented, standard template files which could be downloaded and edited by the contributor to produce the required MAP web pages and documentation. (Anyone capable of producing software for inclusion in MAP would certainly be computer-literate enough to edit a text file and insert the required information about their software!) Standard template files were produced in both HTML and TEXfor creating documentation for

(a) complete programs
(b) subroutines
(c) functions
(d) modules
(e) data

The templates differ slightly, depending on the type of code (program/subroutine *etc.*) or data in question. A comment line was placed at each point in the file where text needs to be inserted by the contributor. Every effort was made to minimise the number of HTML or TEXcommands the contributor would have to insert himself into the file. A set of comprehensive notes describing what information should be inserted at what point in the files was created and can be viewed on-line on the MAP website at `/map/notesmain.html`. This includes a short list of useful HTML and TEXcommands. Whilst a knowledge of HTML and TEXwould be helpful, the template files were designed specifically to enable even those with no knowledge of either to edit them and include the specified information. These files, along with the code or data, could be sent to the MAP Administrator, who would then make any necessary minor corrections, 'validate' the software, create the tar file and include it in the library. The advantage of this procedure is that the contributors have greater flexibility in the writing of their documentation and would be able to change and edit what they write much more easily.

This use of template files will help maintain the standard file format created for the website, and this is particularly important for the correct functioning of the search engine. It will considerably reduce the amount of effort required by potential contributors and/or the MAP Administrator to produce suitable documentation for new additions to the library. The provision of a template-directory (section 2.2) and the fact that the entire file structure for the MAP website is now in place, will also facilitate the addition of new code or data to the libraries.

5. NEW ADDITIONS TO MAP AND USAGE OF THE WEBSITE

The process of incorporating our own public domain software on MAP was also continued, along with the addition of software from external contributors. A total of 26 programs, 30 subroutines, 3 functions and 1 data-set have been added to the MAP website since February 1999. Eight of these programs came from external contributors.

In order to improve the access time for people in other parts of the world, a mirror site has been set up in North America[2] by QuesTek Innovations LLC. Links between *Master* and *Mirror* sites have been added to MAP's home page, giving users a choice of site. The *Mirror* site first started up in May 1999 and continues to be updated on a daily basis. The slight reduction in the usage of the *Master* website in Cambridge during July and August (figure 1) may well be due to the presence of the *Mirror* site. (It should be emphasised that any requests from the *Mirror* site have been excluded from the statistics shown in figure 1.) Neverthless, by the end of Novemeber 1999 usage of the site, both in terms of the number of requests made for web pages or software and in terms of the number distinct hosts accessing the site, had approximately doubled since the start of the project.

The MAP website is used by people worldwide (see reference 3 for further details). Usage of more than 9% of the total traffic has, during one month or more since May 1998, been incurred by hosts in Spain, Argentina, Indonesia, Poland, Japan, France, South Korea and Germany. Links to the MAP website can also be found on many websites thoughout the world, including:

Facility for the Analysis of Chemical Thermodynamics at CRCT, École Polytechnique de Montréal in Canada.[4]
Modeling of Weld Microstructure Development at the Oak Ridge National Laboratory in U.S.A.[5]
Data-Free-Way for Nuclear Materials in Japan.[6]
Chemistry Link Korea: Softwares in Chemistry at the Department of Chemistry, Hanyang University in South Korea.[7]
VTT Information Service in Finland.[8]

References

1. MAP website at http://www.msm.cam.ac.uk/map/.
2. MAP*Mirror* site at http://www.questek.com/map/.
3. *Materials Algorithms Project* - Final report for EPSRC Grant GR/J98073.
4. http://www.crct.polymtl.ca/FACT/websites.htm
5. http://engm01.ms.ornl.gov/
6. http://inaba.nrim.go.jp/netnavi/link_F0831.htm
7. http://www.chem.hanyang.ac.kr/Link/software.html
8. http://www.vtt.fi/inf/inflinks/aihelu.htm

SUMMARY

Introduction

The Materials Algorithms Project (MAP) website, which was started in February 1997 as a centre for the "validation" and distribution of algorithms of use in the modelling of materials, has been developed and expanded. It originally distributed only software written in FORTRAN. This has now been extended to include software written in any standard programming language or programs provided only as executable files. Two new libraries along with many new programs and subroutines have been added to the website, giving access to over 190 programs or routines and 24 materials data-sets. A search engine has now been made available for use on the website, providing the user with the facility either to make a general search of the web pages, or to carry out a more specific search of the keywords which describe the purpose of the software. Maintainance of the website has been semi-automated, with shell script programs making regular updates of the access statistics (daily) and the compressed tar file of the entire library (weekly). Rationalization of the file system on the website, and the formation of standard template files and folders for the creation of new libraries and web-pages will greatly facilitate further expansion of the MAP libraries. A mirror site in North America, courtesy of QuesTek Innovations LLC., has now been fully established. Despite its presence, access statistics show that the usage of the MAP website has doubled, both in number of users and in number of requests since the start of this project in February 1999, reaching over 20,000 requests for webpages and 2200 distinct hosts served worldwide during the month of November 1999.

Achievements

1. MAP has become a central resource, recognised internationally, of freely available, documented materials algorithms and data.
2. There have been several reported cases of research publications which have exploited and developed MAP software, and this is a growing trend.
3. The system is now sufficiently automated for it to continue growing without EPSRC support.