SIRMM: Searchable Information Repository of Mathematical Models

F. Giannino¹, M. R. Guarracino², V. Monetti¹, N. Romano¹, and G. Toraldo¹

¹ Department of Agricultural Engineering, University of Naples, Italy
Via Università 100, 80055, Portici (I)
² High Performance Computing and Networking Institute
Italian National Research Council
Via P. Castellino, 111, 80131 Napoli (I)

ABSTRACT
Mathematics programs in high school, graduate and undergraduate university classes, confront students with real-world problems to be solved with software tools, through a modeling process which requires mathematical concepts. We propose an e-learning application for an educational mathematics environment, whose aim is to build a bridge between mathematics world and application world. It is going to offer to the students a friendly and easy to use tool, to deal with mathematics from the applications and modeling viewpoints, and, on the other way around, to recover the mathematical issues they need for a fully understanding of technical disciplines such as chemistry, engineering or economics.

1. INTRODUCTION
Mathematics plays a major role in science education and understanding: it allows the optimal management of physical systems, the prediction of results, the transparent representation of cause and effect relationships. The global market and the technologically advanced world ask more and more qualified professional skills to use modelling tools, and therefore mathematics in a broad sense, to efficiently support sustainable decisions on a quantitative level, to test conjectures, and to develop scenarios of complex systems.

In such a perspective, mathematics programs, in high schools as in graduate and undergraduate university classes, confront students with real-world problems to be solved with software tools, through a modeling process which requires mathematical concepts, that become more and more advanced moving from high school to undergraduate and graduate programs. On the other hand, the wide spread of computational tools, in conjunction with internet technologies, which enable both distance learning and the access to a huge amount of data and information, makes the "problem solving stage" (which was generally, up to ten years ago, very limited because of less developed computational power) much more realistic, exciting and challenging. For example, a linear optimisation problem (emerging from an equilibrium problem or from some economical market planning strategy) with some thousands of variables today is potentially solvable by a high school student; that is made possible not just because a cheap computer can do it, but mainly because user friendly software tools (interactive computational environments) are available to make the computation, and no difficult programming tasks must be performed! Students are required to incorporate an applied mathematical and computational approach in their classes of physics, chemistry, biology, economy, business, finance, and all branches of engineering, as well as social sciences. Several components are then involved in making the educational process to work in a proper way:

1. Mathematical tools available in textbooks of different kind: ranging from the more theoretical and methodological ones (sometimes very specialist and devoted to specific branches), to the more "numerical" ones, devoted to computational and modeling issues. Books of different kind share a large quantity of information, but their approaches (and languages!) are often quite heterogeneous.
2. Computational tools available in books and on the net are also very heterogeneous. Moreover, technical descriptions very often are unsuitable for users who are only interested in the use of such a tool at a first elementary level.
3. Problems description tools, i.e. books devoted to different application fields, that give the general ideas, describe problems, bring to theoretical models, sometimes to computational issue, but in general need support from the two categories above.

Merging all the ingredients is a hard task, usually more difficult than solving the problem itself. Learning Technology Systems (LTS) [1] provide knowledge through Internet and WWW technologies, offering a powerful tool to integrate knowledge and to build interdisciplinary educational paths [10]. They often provide facilities to adapt to the context, allowing users to change certain system parameters and adapt their behavior accordingly [11].

In the present work we propose an LTS for an educational applied mathematics virtual environment, whose aim is to build a bridge between mathematics and application world, offering to the students a friendly and easy to use tool to deal with mathematics from the applications and modeling viewpoints, and, on the other way around, to recover the mathematical issues they need for a fully understanding of technical disciplines such as chemistry, engineering or economics. It is adaptable in the sense it enables a dynamic content fruition, with respect to user needs, and knowledge discovery. To reach this goal, as it will be detailed in the following, we want to integrate existing resources in an application that can be searched and navigated homogeneously and consistently.

General goals we want to achieved are:
• to make an innovative model for distant teaching-learning for pre-graduate and postgraduate students capable to integrate different disciplines and themes;
• to stimulate the development of analysis capabilities in dynamic, complex, multidimensional scenarios, overcoming present knowledge separation and obstacles in
multidisciplinary teams from apportionment and lack of communication;  
• to offer new forms of constructive interaction among different typologies of cognitive agents, stimulating a permanent flow of knowledge between producers and consumers to create a constant feedback of innovative ideas on education contents and methodology.

The remaining of the work is organized as follows: next section describes related ongoing projects; then the application is detailed. Section 3 outlines the underlying models database schema and provides an example. Section 4 shows the reference software architecture. Finally the work is discussed and conclusions are gathered.

2. RELATED WORK

Many crucial problems in science, technology, environment, medical sciences and engineering cannot be often elaborated with traditional methods but require innovative approaches to be solved in which mathematical analysis is a basic tool. Therefore, an innovative approach for dealing with applied mathematics and, at the same time, for learning is evolving and new technologies enable different kinds of interactions between students and knowledge.

Several web sites have been implemented in the last few years, whose purpose is mainly to supply a new context for learning and teaching mathematics, where it is possible to alleviate the constraints of time and place, and to provide an excellent degree of flexibility of learning and advanced interactivity between tutors and learners [3] [4]; granting maintenance and reusability of resources (“…this website literally shows how to turn ideas and skills once or still thought to be un-teachable natural talents into directly teachable ideas and skills, readily transferred with words, simple stories and images” – from: Appetizers and Lessons For Mathematics and Reason, Alan Selby,aselby@whyslopes.com).

ZIB (http://elib.zib.de/) is an example of an Electronic Library for mathematical Software that operates in the field of information technology. Galaxy (http://www.einet.net/galaxy/Science/Mathematics.html) is a web site of miscellaneous branches of mathematics; it provides an index of pointers to resources which are gathered into categories. See also (http://archives.math.utk.edu/) for a collection of mathematics-related resources. Those projects differ from our in that resources are only catalogued, without any effort to provide a single view through the navigation of various resources. In some web sites (http://euclid.math. fsu.edu/Science/math.html, http://netlib.org) mathematics virtual libraries are available and maintained by universities. CSC, Center for Scientific Computing of Finland (http://www.csc.fi/math_topics/General.html), provides modeling, computing, and information services for universities, research institutions and industrial companies. GAMS (http://gams.nist.gov/), is a Guide to Available Mathematical Software, a cross-index and virtual repository of mathematical and statistical software components of use in computational science and engineering. Those usually provide tools that need to be installed on specific platforms and operating system to be used. This approach works for specialists, but gives no help to people whose aim is to better understand a discipline solving a case study in a controlled environment.

Moreover, recent trends in scientific literature show an increasing application of generic modeling environments (see [5] and [6] for review), originally developed for educational purposes, whose more recent versions are today used for advanced and data intensive modeling simulation [7]. Some of these are already able “to create state-of-the-art web-based interactive simulations”, like STELLA (www.hs-inc.com) which is a tool for helping to build “systems citizens” [8] [9] for the world community, or to implement their own model catalogue. The same happens for Simulistics (www.simulistics.com), that develops and distributes SIMILE, a software tool for computer simulation of complex dynamic systems in the earth, environmental and life sciences. Again, Powersim’s mission (www.powersim.com) is to provide the leading simulation software and consulting services in the area of business simulation. The MathWorks (www.mathworks.com) is a leading developer and supplier of technical computing software in the world. Matlab is an integrated technical computing environment that combines numeric computation, advanced graphics and visualization, and a high-level programming; Simulink is an interactive tool to model, simulate and analyze dynamic systems. Mathematica seamlessly integrates a numeric and symbolic computational engine, graphics system, programming language, documentation system, and advanced connectivity to other applications. All those modelling environments can help in the process of learning but, if used stand-alone, they do not provide a tool that can create a path to follow to understand the whole process from the definition of a problem to the determination of specific software tool for its solution.

Finally, many databases are available on internet to promote the communication and distribution of applied models in different scientific area: REM (http://dino.wiz.uni-kassel.de/ecobas.html) is a meta-database for existing mathematical models and StatLib (http://lib.stat.cmu.edu) is a system for distributing statistical software, datasets, and information by e-mail, FTP and HTTP. Although a huge amount of information is spread around in the net, still a large gap exists among the different components of modern science, and scientific and technological education.

3. THE APPLICATION

The basic idea of the SIRMM is to provide a flexible, tool enabling the content and pedagogical knowledge of physics, biology, and other scientific disciplines to be developed by using computer models and modeling. This tool will support the students in:

• understanding scientific phenomena,
• building models and simulation,
• making experimental data analysis,
• using numerical and existing computational analysis tools,
• building new computational tools,
• getting useful links to improve their knowledge of specific topics.

SIRMM should be a searchable interdisciplinary database, offering a large variety of application, giving, for each problem:

1. a clear description of the phenomena,
2. a modeling definition through a careful analysis of the data problem and the relationship among them,
3. a description and analysis of the mathematical issues,
4. a numerical-computational approach,
5. case studies with analysis and interpretation of the results,
6. link to additional source of information, for any of the above points.
SIRMM is not supposed to be a simple collection of problems, rather an adaptable educational tool to be used in a wide range of courses to enhance scientific education, enlightening on the role of computer modeling in science. The flexibility mainly stands in the possibility for the proposed LTS to be used at different levels of complexity, so that not all steps from 1 to 6 need to be considered; moreover for each step, also different levels of complexity can be considered. For example, in problems of dynamic modeling of populations, different formulations can be proposed to represent the same physical phenomenon, such as a simple one based on finite difference and a more complex one which brings to the solution of a system of differential equation. Similarly, in the computational perspective, black box solvers can be used as well as more sophisticated approaches which ask the student to make his own coding. Obviously, the system cannot result in an exhaustive database. The goal will be mainly to build a general framework, with a large sample of problems which be well representative of the widest possible range of applications.

4. THE MODELS DATABASE

We now present the framework of the repository and the functions defined on it. Figure 4 describes the entity-relation flow, thus outlining the relations existing among the data. Since problems can belong to different scientific areas, the first entity in the database is Area, which is uniquely identified by the name of the discipline itself. In every discipline Problems arise and they can be described in natural language. An instance of such problems is called a Case Study. References can be given for each Problem, and they can regard every aspect of it ranging in its relevance, description, solution algorithms or methods and so on. The description of a Model, as defined by Von Neumann, is composed of the equations and the description in natural language of what they represent in that context. To solve a specific Problem, one should express it in terms of equations, and then find a suitable method of solution. For each Model, different algorithms exists, which can be a better choice for a particular instance of a Problem and related data. The algorithm is implemented in a Software that takes input Data in a certain format and produces output Data (results) that can be suitable for other elaboration.

Figure 2 provides an example of population of the repository and its use. In the field of Ecology a well known problem is the dynamic of a population. A particular instance of this problem can be found when we decide to study how Roe (Capreolus capreolus) population behaves in a particular context, such as North Italy National Park. For that particular problem, there will be references to describe details concerning previous studies and how results have been obtained. One or more models can be available for the description of the phenomena in mathematical terms. Depending on the model, a set of software tools can be available, and each of them implements a different method, accepts different data formats, and is suitable, for example, if the data set is small and dense and not if it is huge and sparse.

The above considerations lead us to consider the problem from a user perspective. The users can be divided into producers and consumers of the information contained in SIRMM. Given a certain case study, the producer will provide information, filling each and all the sections of the repository with data. This will create the kernel of knowledge of the system and will initiate the process of population of the system. The consumer will access the system to understand how to solve a problem.

This is a step by step process, in which she or he first identifies if her or his problem is already described, or at least similar to another present, in SIRMM. It is then possible to analyze input data to guide the user through model definition and the selection of useful algorithms, methods and tools and additional information available. If the problem was not in the repository, and the user was able to solve it, the user problem can be added to the system. Also, if a new method is conceived by a user to solve a known problem, it can be added by the user. The consumer becomes therefore a producer of knowledge that can be acquired by the system, thus realizing an economy of thought and starting a process that is fed by the use of the system itself.

This is a shift of paradigm in how scientific repositories are used: there is no static knowledge that is pushed to users, but rather a dynamic structured information pulled by users.
5. REFERENCE SOFTWARE ARCHITECTURE

In this section a description of all components of the software architecture will be given. This is a path already followed by many open systems, see for example [b]. All components are logically divided in four layers:

**Application specific:** the layer encloses all software components that will be specifically implemented for the application.

**Application general:** the layer comprises all components that are not specific to present application and that can be re-used by future application with similar demands.

**Middleware:** the layer contains all services that are not readily available in the operating system of the server hosting the repository and delivering it on internet.

**System software:** the layer offers components managing basic services provided by operating system, which are not in its kernel.

**Figure 3** depicts the proposed architecture: it identifies all components, which are then organized into layers, and it shows dependencies among them. The list of components contained in the diagram, although not exhaustive, highlights the most important of these subsystems and permit to understand which is the impact of adding new components to the system. In the Application specific layer we find Application authoring component, which contains the design templates common to all parts of the application. Consumer and Producer management components comprise functions used by those who consume and produce information stored in the repository. Content searching enables searches in the repository. System management gathers all procedures for back up, security, systems operation check, and resource monitoring. Application delivery will focus on user specific needs such as the use of specialized devices for content fruition.

In Application general layer can be found all management tools that specifically operate on data, metadata, files and content; moreover a Web delivery component handles differences in web clients used to access.

Middleware offers components to build secure tunnels for data transfer and authentication, a Java virtual machine for server side execution, a library of application programming interfaces (API) and applets for Java and API to access and manage the database.

The System software layer contains the software that implements network protocols, relational database management system and computing facilities. All components need to perform efficiently, scale linearly with the dimension of the repository and users, permit modular integration of new features, and be portable with respect to operating systems and middleware.

The application will provide software packages in which users can interact without any concern about the underlying fabric and software infrastructure. Tools needed will be provided by the open source community and all software will be developed with the idea that will be incorporated into SIRMM if it has characteristics that make it easy to distribute, install and ensure use by the community. The system will be publicly available and will cross the frontiers of universities and other institutions, to create a virtual organization in which it is natural to share knowledge.

### AREAS

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ecology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Science</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### REQUIREMENTS

<table>
<thead>
<tr>
<th>Problem</th>
<th>Description</th>
<th>Resource</th>
<th>Model</th>
<th>Data</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### REFERENCE SOFTWARE ARCHITECTURE

**Figure 2 Example of information in database**

<table>
<thead>
<tr>
<th>Software</th>
<th>Name</th>
<th>Description</th>
<th>Web link</th>
</tr>
</thead>
</table>

### CASE STUDY

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Description</th>
<th>Resource</th>
<th>Model</th>
<th>Data</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### CASE STUDY

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Description</th>
<th>Resource</th>
<th>Model</th>
<th>Data</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### CONCLUSIONS

Main features of LTSS are strong interactivity allowed between tutors and learners, flexibility and modularity, creation and distribution of on-line material, reusability of resources. LTSS can play different roles, from supplementing traditional teaching, to supplying a complete virtual classroom environment. They represent a new technology-based instructional system able to supply the students (in principle) with the classical knowledge in a friendly, easy to use, efficient and pedagogically correct manner, new, alternative or complementary with respect to the classical one. LTSS are definitely going to affect Teaching and learning processes in next years.

Much less clear is the answer to the question about which kind of knowledge the LTSS are going to support, although they can offer much more in terms of redefinition and reorganization of disciplines and contents. Of course that would require in addition to computer science expertise, a very strong interaction between different expertise. SIRMM is an attempt to do that, in
a field of “Mathematics and Applications” that we prefer to “Applied Mathematics”, since we are thinking about a sort of two-dimensional field, having a dramatic growth rate in both directions (Mathematics, Applications).

The SIRMM virtual environment does not reproduce any classical discipline, since it attempts to bring in one comprehensive “mathematical frame” items coming from all sort of fields, becoming a powerful tool for classical disciplines, but also becoming an Educational Virtual Classroom environment by itself, possibly supported by on-line learning material from other EVC (specific for the application fields). Unlike many other EVC, SIRMM can not be reproduced in a classical classroom environment, since it simulate a many to many learning relationship, where the teaching entity is something dynamically changing: this is the difficult task of the whole project, but also the challenging part of it. This is the difficult task of the whole application, but also the challenging part of it.

7. ACKNOWLEDGMENTS

Present work has been partially founded by Italian Ministry for Education, University and Research within research project “Distant learning and knowledge integration in life sciences” (PON 2003-2005).

8. REFERENCES