Modelbase System: A Distributed Model Database on The Internet

M. Hirafuji, K. Tanaka, T. Kiura, A. Otuka

Department of Information Science and Technology

National Agriculture Research Center

Tsukuba 305-8666 Japan

hirafuji@affrc.go.jp

http://model.narc.affrc.go.jp

Introduction

Many kinds of models (simulation programs) and databases are required in agriculture. So far many models and fact databases have been developed. Some groups collected models on their web sites such as

Software database of Japanese agricultural institutes
http://ss.narc.affrc.go.jp/~kiura/meta_index/index.html (Japanese),
WWW-Server for Ecological Modelling
http://dino.wiz.uni-kassel.de/ecobas.html,
CropSyst
http://www.bsyse.wsu.edu/cropsyst/cropsyst.html.
And program libraries such as
Biosystems Analysis Group
http://biosys.bre.orst.edu/src_doc.htm
provides enormous collection of functions and objects. We have collected links to such web sites on http://yummy.narc.affrc.go.jp/Hirafuji/mbres.htm, which includes about two dozens URLs, and the each URL includes also a dozen URLs. So we estimate that there are thousands models at least on the Internet.

It is, however, extremely troublesome to use programs on the web sites. Because users must find candidate models for their purpose by using search engines, and read documents carefully to know specifications of models. Then they need to download a file with several mouse/key operations, and melt it using an archive program. Then they install it and test

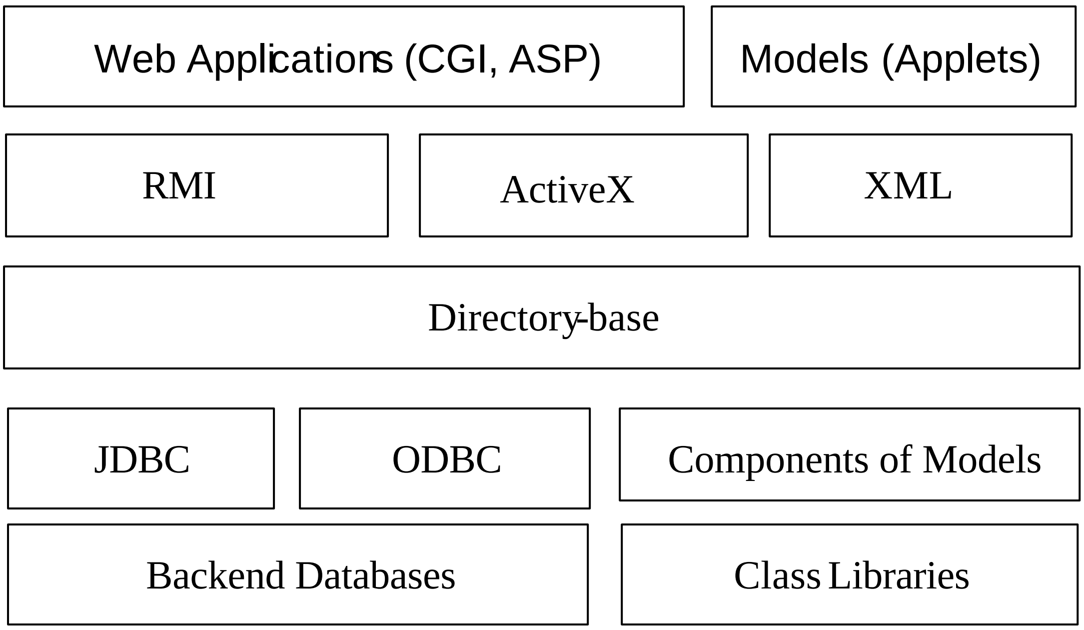


Fig. 1 Layers of the Modelbase system

it. If it does not fit to their purposes, then they must uninstall it. Here, they have a risk that operations of installation/un-installation frequently destroy operating systems such as windows 95/98. Unfortunately some programs require legacy hardware, specific versions of operating systems and utilities. They must prepare all of them before they execute a program.

Even if they could install a program successfully, they must gather current data such as metrological data to execute it. Frequently they must transform the data into required specific format, and they must customize parameters in the program in order to fit it on to their districts and cultivars.

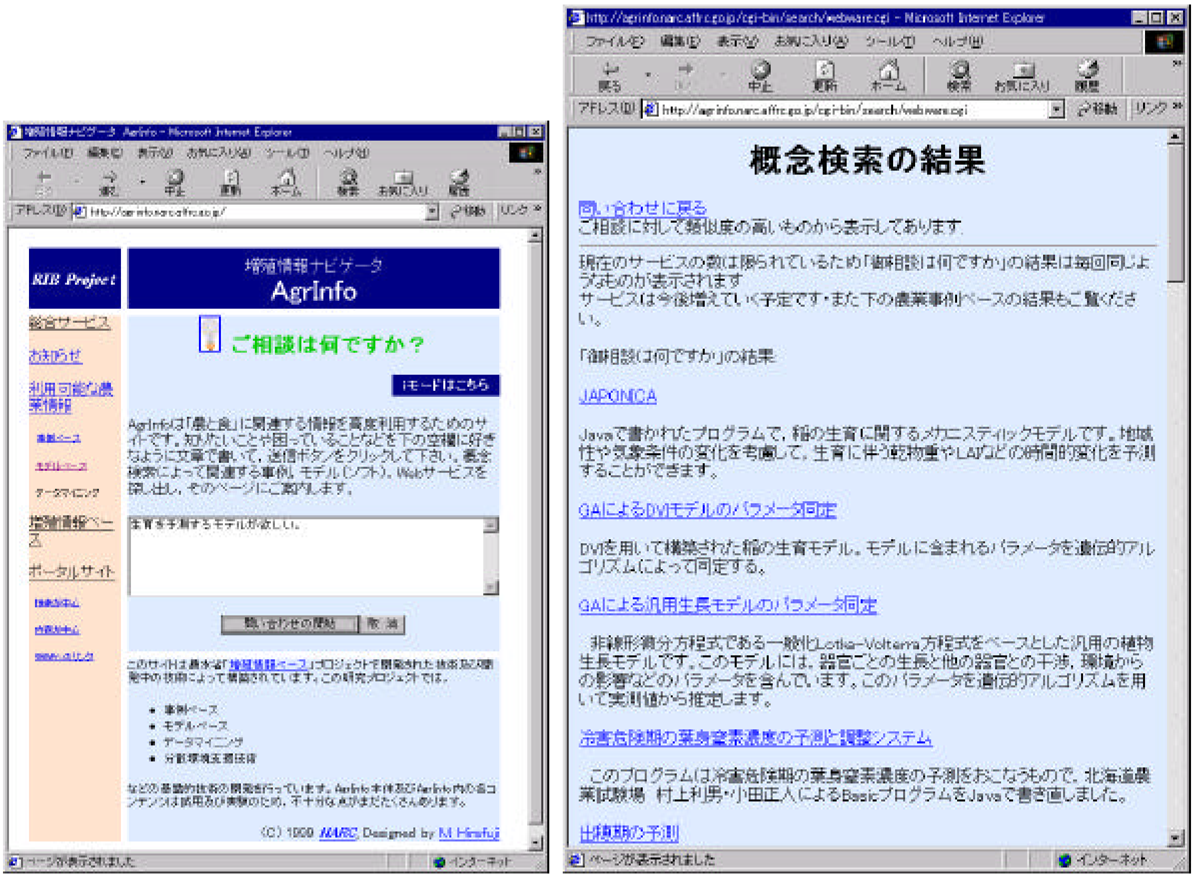
Thus there are hard barriers for users only to try a model. Yet number of models for farming is not sufficient, since enormous number of models should be developed for various cultivars, climates, soils and districts. In addition, combined multidisciplinary models with weed models, soil models, farming system models and management models are indispensable for actual decision support. Where connectable and universal models are needed.

So we propose a concept of modelbase, which is a database of models. Objectives of modelbase are to retrieve models and to develop complex models by combining component models.

The Modelbase System

We assumed that legacy models are written by various languages such as Basic (NEC N88BASIC, MS Visual Basic), Fortran, C, C++, Lisp, Prolog, spread sheets (Excel, Lotus 1-2-3) and DBMS (Access, Oracle, Informix), and new component models will be written only by Java. Here, DBMS and conventional databases such as metrological databases are also treated as legacy models. Modelbase must be able to treat any models, and all models registered in modelbase must run anywhere by using only popular browsers. And all models should be reused efficiently to develop new models. We propose a multiplayer architecture to realize this concept as shown in Fig. 1.

Legacy models are implemented as web applications, which provide interactive services to execute models by CGI or ASP (Active Server Pages). New models are implemented as applets or Java Beans, which can



(a) Entrance page of the modelbase system. (b) Results of conceptual retrieval

Fig. 2. An example runs of the modelbase system. In the query box, Japanese natural language text was entered, “I want to run plant growth models”. Users can immediately execute candidate models by only clinking an underlined name.

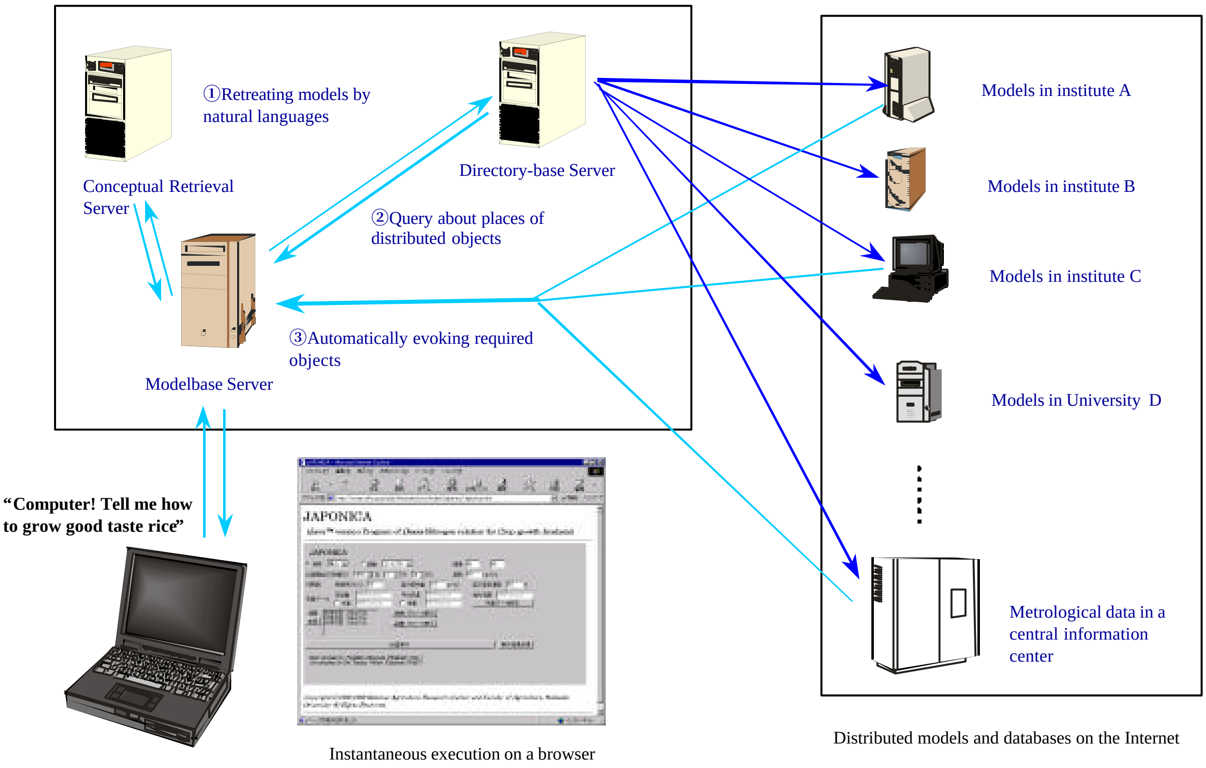


Fig. 3 Structure of the modelbase system

run immediately on any PCs.
Remote models can communicate to other models through RMI. For example, a plant growth model can get automatically metrological data from a web application through an universal metrological database broker. The directory-base in Fig. 1 is a mediator, which informs logical places of all registered models like as DNS.
Input/output of web applications are HTML, so models need to decode/encode intrinsic data into/from HTML files. Instead of RMI and HTML, we propose also XML for standard input/output format of models for advanced modelbase systems.

Retrieval engine for modelbase

We implemented two ways to access models. First way is a simple hieratical menu of models, which is available on the modelbase servers. Users can survey all models with categories at a glance. Second way is a query text window for conceptual retrieval
However, conventional retrieval methods using keyword search cannot find objective models quickly, since terminology in agriculture is multidisciplinary and users are not always experts. Keyword search requires skill as database searchers. In fact, we retrieve too many pages or too few pages by using conventional search engines with keywords. For example, 1,752 pages were found by "Yahoo! Japan" and 174,481 pages were found by "Excite Japan" with keywords, 'model and growth'.

For the modelbase system, we employed concept retrieval engine [1]. This retrieval engine was developed for agricultural casebase reasoning systems, so it can find similar conceptual texts based on LSI (Latent Semantic Indexing) [2] algorithm. The conceptual retrieval engine can treat texts written by natural language (Japanese). That is, the conceptual retrieval engine processes explanation texts of models and a query text by morphological analysis using 200,000 basic words and 28,000 agricultural words, and find the most similar models that fit to the query.
Entrance of the modelbase system is http://agrinfo.narc.affrc.go.jp/.

Implementation

The modelbase system consists of several servers, a modelbase server, a conceptual retrieval server, a directory-base server and other application service provider that serve web applications and Java applets of models. Legacy models such as database services and applications provided by CGI or ASP are conventional ordinal web servers. Communication procedure was standardized among model applets and databases in order to construct new models and complex applications using models simply in the framework of modelbase. This complicated structure is hidden from end users (Fig. 3), so end users can access all models with seamless operations.
The modelbase system includes, currently, 18 models and 30 web applications such as rice growth models,

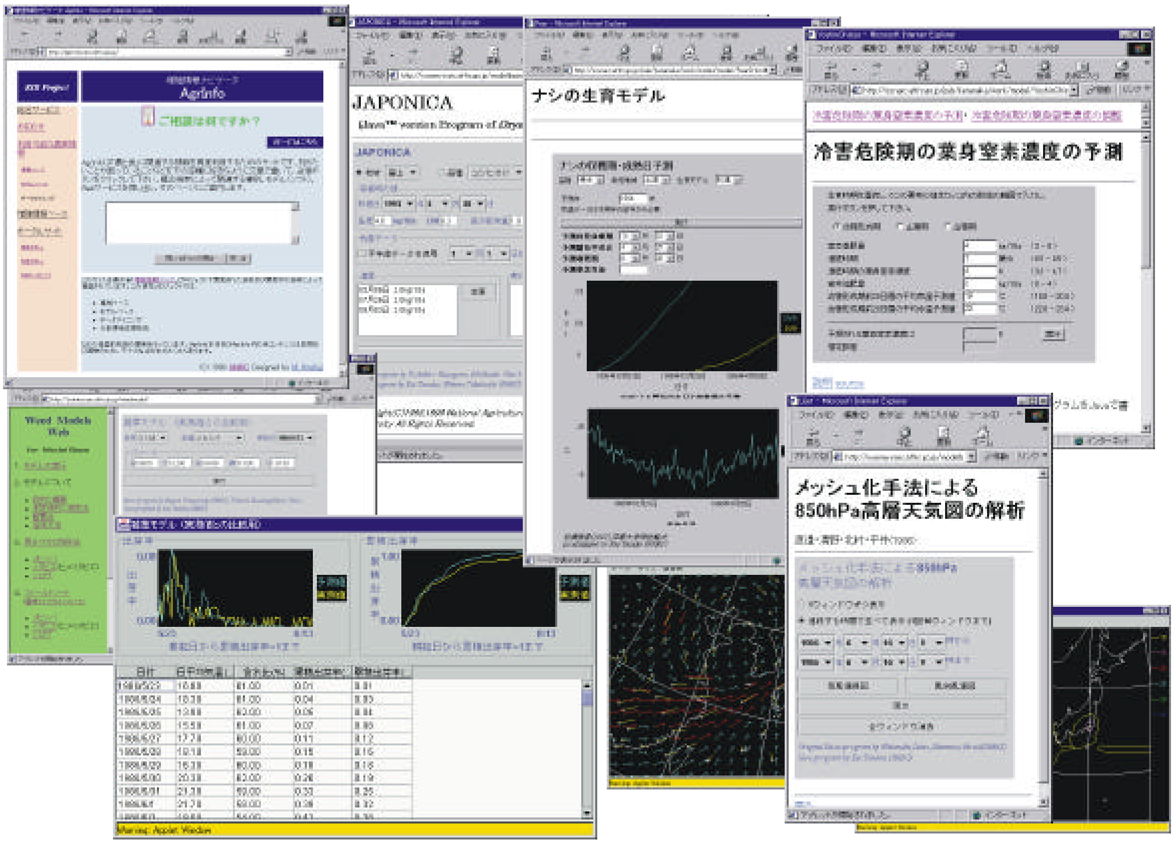


Fig.4 Sample models registered in the modelbase system. All these models are written by Java.

weed models and pear growth models (Fig. 4).

The conceptual retrieval engine is also implemented as a web application (CGI) in the modelbase system, so other kinds of retrieval engines such as amazing 3D visual retrieval engine will be available.

Cyfars with modelbase

This modelbase system will be really useful, when farmers can use both the wireless Internet service and a wearable PC in actual fields. Modelbase may be one of killer applications for farmers. Farmers with cutting edge wearable technologies will be able to enjoy their advanced cool works in ubiquitous computing environment.

So we propose a new concept of advanced mobile frames as 'CYFERS' (CYber FARmerS) to create novel lifestyles in comfortable farming area.

Cyfars will enjoy online voice/video chat services while working, and they will be able to connect modelbase systems anywhere when they find problems. It will be possible that Cyfars work as part-time specialists of gardening and greenery with staying in their district.

Cyfars with a small CCD camera can automatically upload image data of fields and plants in real time while working. Already we have developed an automatic image processing system[3,4], which will be applied to automatic data entry of plant growth models. Using this program, we want to gather enormous raw data about growing plants such as LAI (Leaf Area Index). Then Cyfars will be able to develop their original plant growth models as professional modelers.

Additionally we are developing utility programs to extract environmental numerical raw data such as temperature, relative humidity and light intensity only from image data. Then Cyfars will acquire a function of mobile weather robot stations.

Acknowledgement

This study was funded by the Japan's national research project of MAFF (Ministry of Agriculture, Forestry and Fisheries), "Fundamental Research of Reproductive Information-Base to Develop Agricultural Production Assistance Systems" (http://model.narc.affrc.go.jp/RIP.htm).

References

1) Otuka, A., 1999: Case Retrieval and Management System for Agriculture Case Base, Proc. of EFITA/99 Perspectives of Modern Information

and Communication Systems in Agriculture, Food Production and Environmental Control (Volume B), pp.679-686.

2) Deerwester, S. Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R., 1990: Indexing by Latent Semantic Analysis, Journal of American Society of Information Science, 41 (6), pp.391-407.

3) Otobe, K. Tanaka K., Hirafuji, M., 1998: Image Processing and Interactive Selection with Java based on Genetic Algorithms, 3rd IFAC/CIGR Workshop on Artificial Intelligence in Agriculture, pp. 88-93, 1998.

4) M. Hirafuji, M., Hagan, S., Otobe, K., Tanaka, K., 1998: An Alternative Strategy for Evolutionary Computation based on Research of Consciousness, Proc. of the 3rd Asia-Pacific conference on control & measurement.