

What's on the CD?

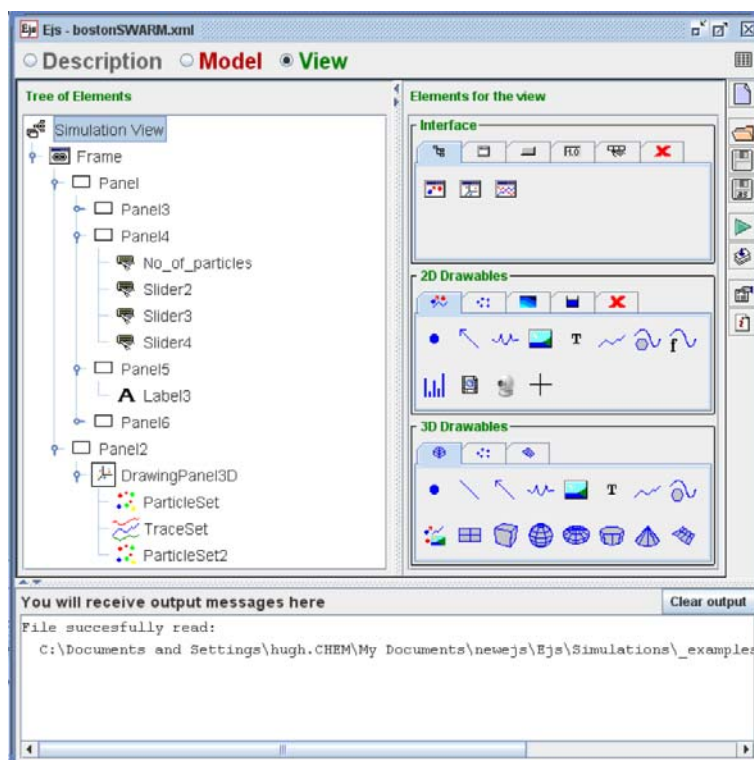
There are several types of information on this CD:

1 Easy Java Simulations

Easy Java Simulations (Ejs) is a software package designed to make the use of Java simple. Although it has been written with the preparation of scientific simulations particularly in mind, it is also potentially valuable for any user of Java who would like most of the details of preparing an attractive interface to be handled automatically.

Ejs has a variety of useful features built in, such as two- and three-dimensional display tools, solution of differential equations, plotting of fields and surfaces, display of images and film clips, easy preparation of frames and windows, widgets built in and so on.

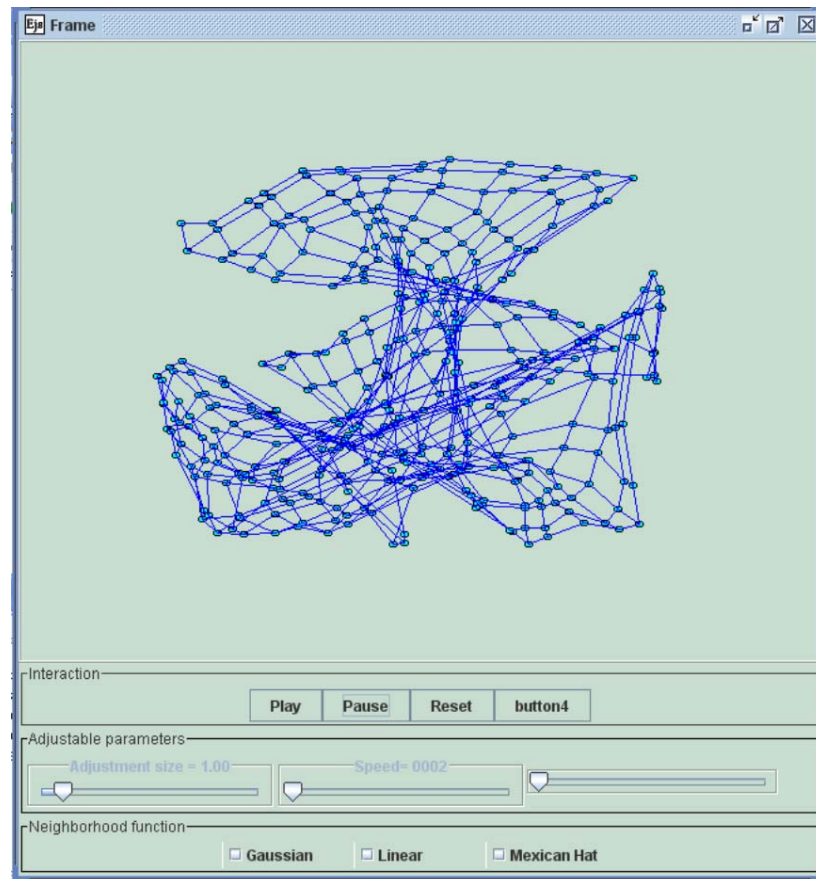
Ejs is free, but if you use it in your work, the developer, Prof. Francisco Esquembre of the University of Murcia in Spain, and collaborators should be acknowledged in any published material (<http://www.um.es/fem/>).



2 Instructions for Installing and Running Ejs

3 Sample Programs in Artificial Intelligence

Several programs that illustrate the use of AI techniques are contained in the zipped Ejs file in a sub-directory within “Simulations”. There are programs illustrating Genetic Algorithms, Self-organizing maps, Artificial Neural Networks and Growing Cell Structures, each of which can be run using Ejs.



These programs are not intended to provide a platform that can be used, without modification, to tackle the solution of scientific problems using AI methods – there are several commercial packages that can be used for the major AI methods, such as Artificial Neural Networks, Genetic Algorithms and Expert Systems.

Instead, the aim of the book is to encourage you to investigate how the methods work, and then write your own programs. The sample programs therefore are provided so you can “play around” with some simple examples of AI in science. They are neither particularly sophisticated in what they can do, not especially clever in the way they are written. However, they may help to give you a feel for the way that a GA, for example, behaves.

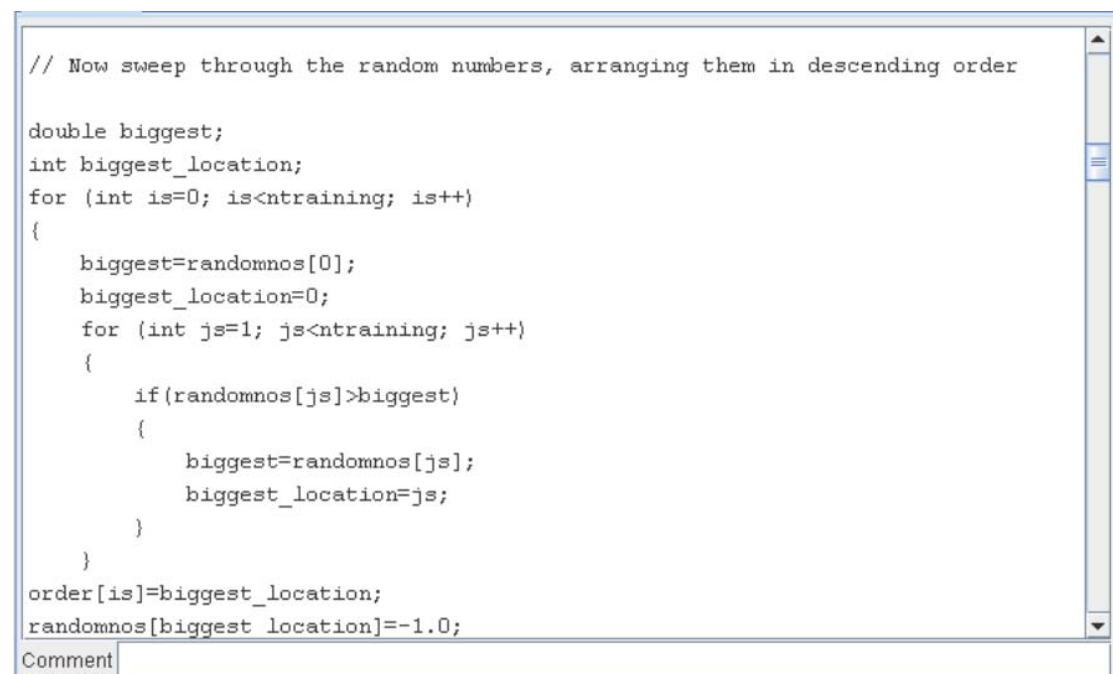
4 An Introduction to Running the Sample Programs

An introduction in pdf format is provided for each AI program provided with Ejs. This is designed to help you get the program running and understand how to use it. The introduction also includes suggestions for short investigations using each program.

5 Program listings

Not everyone wants to use Java; many scientists prefer other languages such as C++ or FORTRAN. Consequently, there are also ascii text listings for each program. Although the listings are taken directly from the Java source code, they are fairly well commented, so most people who know how to program in a different computer language should be able to understand what each section of code does. The listings, therefore, should be helpful to anyone who does not want to learn Java but still wishes to use the programs to help understand how simple AI programs work.

Note that you are welcome to use parts of the Java code itself in an application you write (though use of a large piece of code as opposed to snippets should be acknowledged in any published work). However, the Java code will need some modification if you want to use it outside Ejs, since both the declaration of variables and a suitable interface display will need to be added.



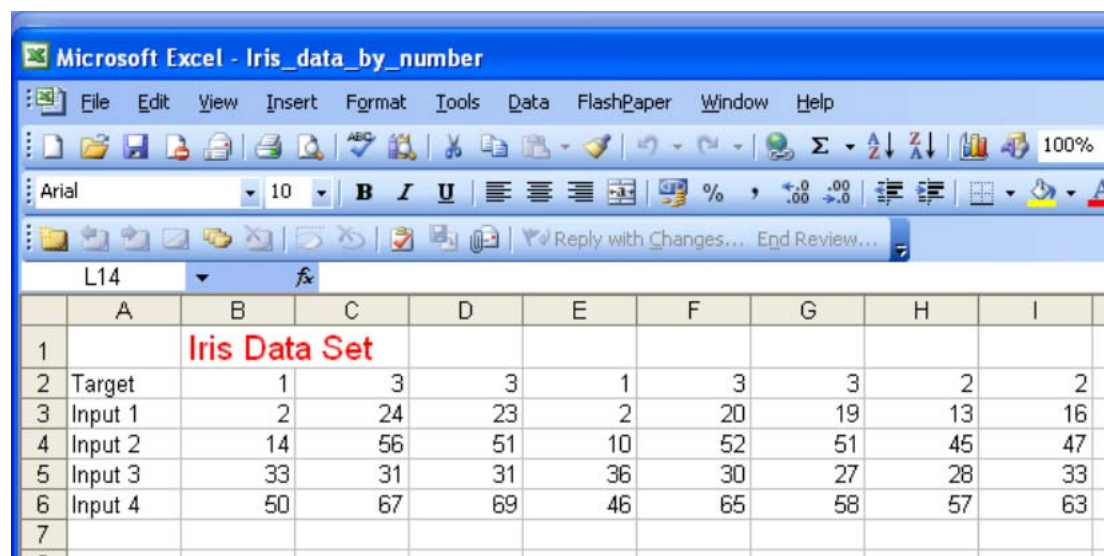
```
// Now sweep through the random numbers, arranging them in descending order

double biggest;
int biggest_location;
for (int is=0; is<ntraining; is++)
{
    biggest=randomnos[0];
    biggest_location=0;
    for (int js=1; js<ntraining; js++)
    {
        if(randomnos[js]>biggest)
        {
            biggest=randomnos[js];
            biggest_location=js;
        }
    }
    order[is]=biggest_location;
    randomnos[biggest_location]=-1.0;
}
```

Comment

6 Iris data

The Iris dataset is a well-known example of data that can be used to test some types of AI programs; in particular, it is a nice set on which to test Artificial Neural Networks. The data set is available on the Internet, but for convenience it is also provided in two forms on the CD



	A	B	C	D	E	F	G	H	I
1		Iris Data Set							
2	Target	1	3	3	1	3	3	2	2
3	Input 1	2	24	23	2	20	19	13	16
4	Input 2	14	56	51	10	52	51	45	47
5	Input 3	33	31	31	36	30	27	28	33
6	Input 4	50	67	69	46	65	58	57	63
7									
8									

7 References

The number of papers in the scientific literature that describe the use of Artificial Intelligence in science and engineering is growing by several hundred per month. Even if it were feasible to provide on this CD a complete current list of all papers that relate to the topics discussed in this book, such a list would be out of date within a matter of weeks.

Djurdjevic, D.P., and Biggs, N.J., *Ab initio* protein fold prediction using evolutionary algorithms: influence of design and control parameters on performance, *J. Comp. Chem.*, 27, 1177, 2006.

Fernandez M., Caballero J., Fernandez L., et al., Protein radial distribution function (P-RDF) and Bayesian-Regularized Genetic Neural Networks for modeling protein conformational stability: Chymotrypsin inhibitor 2 mutants. *J. Mol. Graph. and Modelling*. 26, 748-759, 2007.

Ghasemi J., Saaidpour S., Quantitative structure-property relationship study of n-octanol-water partition coefficients of some of diverse drugs using multiple linear regression, *Analytica Chimica Acta*. 604, 99-106, 2007.

Gobin O.C., Joaristi A.M., Schuth F., Multi-objective optimization in combinatorial chemistry applied to the selective catalytic reduction of NO with C₃H₈, *J. Catalysis*, 252, 205-214, 2007.

Hartfield, R.J., Interpretation of spectroscopic data from the iodine molecule using a genetic algorithm, *Appl. Math. Comp.*, 177, 597, 2006.

A list of representative references is provided instead. These are chosen to illustrate how the kinds of methods on which this books concentrates are used to tackle real scientific problems. The presence on the list of a particular paper is not meant to indicate that it is of more importance than those not listed, but there are many interesting applications included. It is hoped that these will provide a fruitful starting point for further research into the work being carried out by many different groups around the world.

A final note...

The material on this disk will be updated periodically, both through the addition of further programs and data and the correction of any errors or points of ambiguity. If you wish to be provided with a (free) updated disk when available, email Hugh.Cartwright@chem.ox.ac.uk. It is expected that revised CDs will be available roughly every 8-10 months.

Hugh Cartwright
Oxford, February 2008.