Exystence
European Complex Systems Network of Excellence

COMPLEXITY WORKSHOP
HP Grenoble, Auditorium
November 28th 2005

The Bayesian Aspect of Complex Systems

PIERRE BESSIERE, Senior researcher at the French National Centre for Scientific Research (CNRS).

Introduction
The LAPLACE research group has been working on stochastic models for perception, inference, learning and action since 1992. The central thesis is that a subjectivist approach to probability is a simple mathematical theory of cognition and rationality as an alternative and extension of logic. During this time the group has elaborated the Bayesian Programming methodology and software tools.

The initial problem
15 years ago we were working on artificial intelligence and robotics and were puzzled because we could not seem to get the software right, so we designed the simplest robotics system that we could imagine. We had a dustbin with a light inside and a vertical axle on which was mounted a photoelectric cell. The one control variable was that the axle turned and a single photocell recorded the light level. We started to collect data by simply revolving the axle and recording the angular position against the reading of the photoelectric cell. The results were puzzling because, theoretically there was an obvious revolutionary symmetry to the system and we were expecting to find a corresponding symmetry in the data. We expected to find a symmetrical peak where the photoelectric cell directly faced the light and another symmetrical reflection peak at 180° from this, but we did not get that result.

We checked that the axle was really vertical and centred in the dustbin. We cut out all extraneous light by putting the system under a black cover and we made sure that there was no dust in the dustbin, but we still obtained the same results. What we subsequently found was that the photoelectric cell had a hysteresis and what we were doing, was exploring this memory characteristic of the cell. In other words the hysteresis, or memory, spread the peak beyond the area that it should have occupied and demonstrated, that even in a very simple system, there can be complexity. So a simple theoretical model of the system would not capture the phenomenon because of a hidden variable.

Introduction to Bayesian theory
We subsequently met an American physicist called Edward T. James who

proposed a Bayesian approach to deal with this problem of incompleteness. The approach is quite simple; it is to turn the discrepancy between a model of the phenomenon it is trying to simulate, into a probabilistic model using experimental data and then develop a methodology for building and implementing models. First, a probabilistic description of the phenomenon is made using the collected data and knowledge of physics to pick the variables. Then a joint probability is calculated by making some hypothesis about the dependences between variables.

'Spam application'
One of the success stories of Bayesian techniques has been 'spam' detection in e-mails using Boolean variables. If you have a Boolean measure for each word of the dictionary then spam may be detected via the logic of the sequence. In the case of spam it was a strong hypothesis that the operation of any word is simply conditioned by the fact that it is a spam word. If you neglect any co-occurrence of words the hypothesis works well. The last part of the specification is to give parametrical form for each distribution. The model is validated by simply asking people whether, given the presence of a word in the e-mail it was spam.

Development
The Bayesian programme enabled us to write the most probabilistic models but we still needed a programming language in which to express it, so in 1995 we produced the first version of an inference engine. This was a C++ API which is a direct implementation of a Bayesian programme and is running on most operating systems.

The work then split into two different directions, one more fundamental research and the other more industrial. One of the fundamental European research projects, which ran from 2001 to 2005, was BIBA standing for Bayesian Inspired Brain and Artefact. This was modelling biological processes and cognitive processes using Bayesian inference. There is a big community in life science now using Bayesian models to explain and understand phenomena in neuro-science. The three questions we are trying to answer in the European project are:

- Is there a biological possibility of Bayesian reasoning at a microscopic level and is there any proof that the central nervous system is doing Bayesian or probabilistic calculations?
- Is there any biological possibility of Bayesian reasoning at a macroscopic level, enabling us to build a model of behaviours of humans or animals? There are number of examples of such models existing.
- How can we develop better artefacts using Bayesian reasoning? For example, to derive better algorithms used for driving robots.

Discussion

Questioner 1. By biological possibility of Bayesian reasoning, do you mean that may be how people update their probabilities and learn?

Pierre. I mean more than that. I mean can we model the process of how neurons carry out computation at the microscopic level, and at the macroscopic level can we build better models to explain, for example, psycho-physical experiments such as illusions. There are about equal numbers of people from the life sciences, from the computer sciences and from mathematics working on this. BIBA is both an artefact project and a cognition project.

At the same time however we have discovered a number of different industrial applications of Bayesian inference. We started a company in 2003 with the support of CNRS which now has about 10 people working in it. There is research and development going on with some financial and industrial applications and we are still working to improve the inference engine ProBT.

A French bank in Basle wanted to cover 99% of their risk with as precise an evaluation as possible. There are different kinds of risk and they wanted an evaluation of the loss that could occur with those risks. We are doing Bayesian analysis to see what kind of capital cover they need.

The other example is stock picking and we are opening a web site in 2006 which will chart the evolution of stocks. It involves a number of relative indicators for each company about whether the stock will go up or down. These consist of technical ratios involving expert evaluations and data and we did a fusion of all the information to try to get better understanding from the past six months of analysis

Questioner 1. How does that compare with standard portfolio theory which also allows you to calculate risk and construct rational portfolios of stocks? It is a standard procedure which is used today.

Pierre. We used the same method for systematic fund management two years ago for another big French bank and, basically, married the approach with all the other tools that they had. The idea was not to predict whether the stock would go up down, but whether one stock will perform better than another. That is a simpler problem because you do not need to do any prediction about how international events could influence the stock as a whole. The data is the differential between one stock and all the others.

We also have worked on stock management for shops selling tools and hardware. They carry a huge number of items, so there are several tens of thousands of references and as some of the items are very expensive they want to reduce the quantity of stock that they carry. With a huge database of sales during the last 10 years, it is possible to do a detailed analysis of what they are likely to sell during the next year which is very important

information for them.

Other industrial applications are in the car industry where we have several projects on automatic drivers for cars to help prevent accidents. These drivers analyse a video image for moving parts, resulting in a probabilistic map of what is likely to occupy the space in front of the car from moment to moment. The system has a memory so that, when a possible moving target disappears behind something else, the probability of its presence does not fall to zero straightaway which means the automatic driver takes its probable reappearance into account.

Questioner 1. Is probabilistic reasoning acceptable in this kind of situation?

Pierre. Yes, this is an aid, not a replacement of the human driver. It is especially important for bus drivers because most of the accidents occur when buses are arriving or departing from a stop and because the driver has a blind spot. If the system detects that the driver is going to make a move which is dangerous, an alarm can be fitted to alert the driver to the situation.

Questioner 2. Is the tool you are using the right one to use for all problems given to you by industry, or do you consider other approaches that would also solve the problem? Do you try to formulate the problem generally and then solve it using a Bayesian approach?

Pierre. We are confident and competent using Bayesian inference, but not in using other methods, so if we cannot use the approach then we will recommend something else. However, we do have an automatic driver for an electric car which follows the road, avoids obstacles and will overtake another car and make emergency stops.

Questioner 1. In that case if the system detects that something or someone is about to appear in front of the car you have a simple instruction that stops the car. Is the probability programme updated for that to take place?

Pierre. Yes, the system will just stop the car, but it also has to do several things at the same time such as follow the road in a given direction, avoid a car which it is passing, take account of a moving pedestrian and stop if necessary. All this is modelled with different probabilistic models and what is nice is that it is quite easy to build hierarchical models to make more complicated decisions. We also work on prevention models which try to predict failure and prevent it. Probabilistic inference is very good for diagnosis.

Questioner 3. Can you give us some idea where Bayesian inference is going in biological models?

Pierre. If I could prove that the central nervous system is doing Bayesian inference I would be very happy. There is some very interesting work being

done by several teams around the world analysing the ways in which the central nervous system could be working with Bayesian models, analysing how a single neuron is working-or assemblies of neurons.

Questioner 3. How do you update the changing dynamics with probabilities? Driving is one thing where, perhaps, the dynamics information does not change significantly, but in informational terms updating a 'spam' filtering system is interesting. Spam is constantly being updated to get around filters and as the filters update so the spam updates as well. How do you determine the appropriate dynamics?

Pierre. There are Bayes filter models with hidden updating. I think that, in general, Bayes filtering is a very effective tool to analyse dynamics, especially a dynamic training model. It means the reference to the past analysis is very important. Using Bayes filtering means you never decide exactly what the state of the system is, because it is always a probability distribution. That probability distribution constitutes the memory of all that has happened in the past. It is not an exact memory of course, but it is a summary of all that is important to the evolution of the system. One very interesting way to analyse the central nervous system is not to consider it as storing states, but storing probabilities about states. That is a different approach from what has been done in the past with modelling the central nervous system.

Questioner 4. How do you resolve the legal aspect of your involvement with banks? I mean who accepts responsibility for actions carried out on advice from the model?

Pierre. Well, first of all, decisions taken concern the bank and not clients of the bank. For example, one thing that they wanted to demonstrate to the banks regulatory body was that they had sufficient cover for capital risk. So it was a matter of convincing the regulators. For operational risk it was a rough assessment, I think an outlay of some 15% of the amount of money that the bank had to deal with. The model was merely to help their argument on risk assessment. We are providing the tools but we also have to use the expert knowledge given to us by the bank in order to build the model. So we are working together with them.

Questioner 4. So it is not a black box model?
Pierre. Absolutely not. It is very different from that because the model is built with two kinds of expertise. First of all there is the construction of the model, in terms of the differences between the variables, and secondly, in operational risk management there is a huge amount of data. For example, in the case of fraudulent use of credit cards there is a huge amount of statistical data. There are cases where evaluation of the risk simply comes down to expert judgment and that has to be given weight along with that given to the statistical data.