Data mining promises to dig up new drugs

February 2 2009

(PhysOrg.com) -- A robot scientist that can make informed guesses about how effective different chemical compounds will be at fighting different diseases could revolutionise the pharmaceutical industry by developing more effective treatments more cheaply and quickly than current methods.
The robot, known as Eve, uses advanced artificial intelligence combined with innovative data mining and knowledge discovery techniques to analyse the results of pharmacological experiments it conducts itself.

By relating the chemical structure of different compounds to their pharmacological activity, Eve is able to learn which chemical compounds should be tested next, bringing a degree of predictability to drug screening procedures that, until now, have tended to be a bit hit and miss.

“Over time, Eve will learn to pick out the chemical compounds that are likely to be most effective against a certain target by analysing data from past experiments and comparing chemical structures to their pharmacological properties,” explains Saso Dzeroski, a researcher at the Jozef Stefan Institute in Ljubljana who helped develop Eve’s data mining capabilities.

“That should help scientists and pharmaceutical companies identify more effective compounds to treat different diseases, allowing them to find drug leads in a fraction of the time and at a fraction of the cost of current methods.”

Eve could minimise the need for random testing of chemical compounds, Dzeroski says, noting that the robot scientist is the first computer system capable of originating its own experiments, physically performing them, interpreting the results and then repeating the cycle.

Currently, when a new drug is sought pharmacological researchers conduct a blind study of tens or hundreds of thousands of chemical compounds, applying them to an assay for a disease. The results of those tests determine the so-called Quantitative Structure-Activity Relationships (QSARs) that relate the structure of a chemical compound to its pharmacological activity.

Exhaustive testing like this is time-consuming, costly and generally has to be repeated each time a new drug is sought.

More “intelligent” approach to drug discovery

Eve offers a more “intelligent” approach, says Ross King, a computer science researcher at the University of Wales, Aberystwyth where Eve is to be installed.

The robot conducts the QSAR testing in assays itself, analyses the results and stores the data for future use. Over the course of numerous experiments, Eve learns which chemical structures are likely to be effective in specific assays. So, instead of choosing compounds to test at random, it can pick ones that are more likely to be effective.

“We have carried out some preliminary trials and the compounds picked by Eve show more promise than those selected randomly,” Dzeroski says.

New data mining techniques developed by a team of researchers led by Dzeroski lie at the heart of Eve’s groundbreaking drug discovery capabilities. Working in the EU-funded IQ project, the team developed new methods to analyse complex data, including chemical structures, from databases such as that in which Eve stores the results of its experiments.

Unlike most data mining approaches, in which an individual analysis is carried out on a single dataset, such as a spreadsheet, the techniques developed in the IQ project allow knowledge discovery processes, consisting of several analysis steps, to be carried out across multiple sets of complex data.

The techniques rely on the use of so-called inductive databases that contain not only raw data but also information about patterns and models valid in the data. In the case of drug discovery, the structures of the chemical compounds tested and their effectiveness would be the raw data, while molecular structures that appear commonly in effective compounds would be patterns, and the equations that predict a compound’s effectiveness would be models.

From experimental data collected by Eve, patterns would emerge that can then be used to make informed guesses about which compounds

should be effective and which probably will not be. The same data mining techniques are also being applied by the IQ project partners in other fields, including genomics, systems biology and environmental sciences.

“Because much more than raw data is being analysed, the same process for identifying different patterns can be reused, regardless of whether you are trying to develop a drug to treat AIDS or tuberculosis,” Dzeroski explains.

Eve will initially be put to work at the University of Wales to search for compounds that could be effective in treating malaria and schistosomiasis, so-called Third World diseases that are the focus of only limited research by commercial drug companies.

King says their mission is both to demonstrate that the data mining technology works and to find new leads that could result in new drugs being developed in the future.

Dzeroski foresees more robots like Eve being put to use in research labs and drug companies over the coming years. And although it will take 10 to 15 years for new drugs, based on compounds picked out by Eve, to start being used in treatments, the work done now “could have a major impact on the pharmaceutical industry and on healthcare in general in the future,” he says.

The IQ project received funding under the ICT strand of the EU’s Sixth Framework Programme for research.

Provided by ICT Results