Scientific software solutions

DECISION SUPPORT SYSTEMS

Michael Rippen
TESSELLA SUPPORT SERVICES PLC

Issue V2.R1.M0
October 2005

Introduction

A Decision Support System (DSS) is any tool used to improve the process of decision making in complex systems, particularly where information is uncertain or incomplete. There are a number of approaches to DSS systems, each of which assist the process in a different ways. DSS is an old term that now applies collectively to a number of 'new' systems such as OLAP (On Line Analytical Processing), EIS, ESS, expert systems and more. This technical supplement will present DSS as a whole.

A DSS can range from a system to answer simple queries that allow a subsequent decision to be made, to a system that provides detailed querying across a spectrum of related datasets. More complicated systems directly 'answer' questions as opposed to providing static information, in particular high-level 'what-if' scenario modelling.

These latter systems are more important in a DSS and are its prime application area. They encompass systems where knowledge or answers are sought in systems with incomplete or very complex data. A DSS then provides decisions based on rules or algorithms derived from an understanding of the business or application domain.

Who can benefit from Decision Support Systems?

Decision Support Systems were traditionally aimed at managers to assist in high-level decision-making. Of course, their application is much wider than this and they have been applied across the spectrum of business and industry.

Any given DSS can be tailored to provide features across a spectrum of users, firstly to ensure the system is accessible at the right level for different audiences and secondly to provide only the information appropriate for confidentiality and security purposes.   For example, in a drug discovery environment a DSS could have tailored interfaces for Directors, Team Leaders and Scientists.

Development of a Decision Support System

Development of any DSS has a number of generic stages:

- Data gathering and collation
- Data storage, access and mining
- Data analysis
- Data reporting

Each of these is important to ensure that the information provided is both accurate and answers the intended question of the user. To this end, the data must be complete and comprehensive, stored in a scalable and searchable manner, semantic inference must be fully researched and defined, and the system presented to the user in such a way that the necessary questions can be asked of it and the correct answers displayed. This latter point is important – a badly designed user interface may force a certain approach on a user or may not allow the full scope of a problem to be addressed.

What is in a Decision Support System?
All Decision Support Systems have 3 main components:

1. A data store of knowledge.
2. A process by which this knowledge may be systematically interrogated to provide answers to questions.
3. A user interface providing users with an intuitive, accessible tool for gaining the information they require.

It is the second of these that predominantly distinguishes between different Decision Support Systems; some of these are described next.

How does a Decision Support System work?
There are many ways in which Decision Support Systems are implemented. Simple query tools can provide one level of support. The more complex systems employ some form of artificial intelligence. There are a number of approaches to this, three of which are described below:

- ❑ Rule based systems
- ❑ Model based systems
- ❑ Statistical/probabilistic based systems

These systems can be relatively straightforward or very complex. However, both their strength and weakness lies in fully encapsulating the data and the rules and relationships associating it.

Rule based systems
Often called ‘expert systems’ or ‘knowledge based systems’. These are systems containing a database of information with a semantic content, allowing relationships to be defined between the data. A ‘rules’ database is devised which encapsulates these semantic relationships. Data is then analysed using

symbolic logic, allowing a user to iterate through the information, guided by the system, to provide answers to their enquiries. These systems can also explain *why* a particular answer has been given.

Where might such a rule-based system be used?

- Tessella has developed a DSS for Metronet BCV, who are responsible, under a Public Private Partnership, for managing the infrastructure assets of four London Underground lines, such as the Bakerloo line. It is essential that these assets - stations, signalling, tracks and tunnels – comply with the relevant regulatory constraints on safety, both during and at the end of the 30-year period of the partnership. To ensure this, Metronet needed a system to help their managers identify and then justify the most cost-effective strategies for smart and efficient asset maintenance. The new decision support system takes real data about the current state of the assets, and then applies a variety of 'deterioration rules' to forecast the future condition of those assets over time, and how that condition will be affected by future maintenance or replacement. The managers can readily configure these deterioration rules, and examine the effects of any number of factors such as stress, load, or location, or restrict the investment strategies produced by applying caps on expenditure. In addition, dependencies between nearby assets, and the benefits of renewing groups of assets together are all handled. Importantly, the process is completely interactive; this allows the Metronet asset manager to understand, and therefore communicate to senior managers and regulators, *how* the system arrives at the final recommendation and ensures both safety and value for money.

- Pharmaceutical Industry – Process Design

An example of this in the pharmaceutical industry might be in the area of process design in the drug development stage. Given the known characteristics of the drug to be manufactured, the Decision Support System would be set up to address the question of how best to design and configure the details of a production set-up.

Model Based Systems

These systems derive answers to questions by taking input from a user, combining that with information in the data store and then running an analytical model on the data. This model can be very simple, such as plotting a series of curves in Microsoft Excel and choosing the one with a fitted parameter above a particular threshold, or can be very complicated, such as solving differential equations in one, two or three dimensions, perhaps as a function of time. The scope of any given system is relatively tightly defined – an analytical model has its own well defined application area – but the range of questions that can be asked may still be very complex as the models may have several degrees of freedom and several possible constraints. These systems are useful where empirical information is either not available or not appropriate to a particular application area.

With the extensive scientific background of our staff, Tessella has a great deal of experience in Model Based Decision Support Systems, some examples of which are described below:

- Pharmaceuticals – High Throughput Screening

This approach is being employed in the pharmaceutical industry through the use of genetic algorithms. Firstly, partitioning existing chemical and biological results data and then selecting the best sample set for new high throughput screens – this approach runs ‘simulated screens’ on the computer to determine those samples most likely to be ‘active’ in a new screen.

- Rail Industry – Asset Management

Tessella has been involved in a number of asset inspection, monitoring and management systems for the rail industry. A decision support tool to assist in the management of rail and weld defects was developed by Tessella. This tool is now used in several areas in Network Rail’s Southern and Eastern regions. Tessella also successfully tested an integrated asset management process, which combines newly developed mechanized inspection techniques and software based decision support models, to provide a powerful track management tool. The trial has achieved significant improvements in condition and reliability of track. On another project, a suite of products for Railtrack and its maintenance contractors is currently being rolled out in a turnkey contract on the West Coast Main Line to give prior warning of potential asset failures, so that corrective action can be taken to avoid the failure occurring.

- NIREX - TDPSA

Time Dependent Probabilistic Safety Assessment (TDPSA) is a system designed to model the possible environmental impact of the disposal of radioactive waste in underground repositories, over long periods of time. The parameter space is huge and a system of algebraic and differential equations models the system as a network of interconnecting sub-models. This allows a series of what-ifs to be assessed, based on the outcome of the model simulations.

- TRL - Whole Life Cost Model

'Whole Life Cost Model' has been developed, and is used to determine the best maintenance strategy for large road networks (typically hundreds of kilometres) over their lifetime. This predicts when treatments would be required and weighs up the financial advantage of delaying them or choosing cheaper treatments, against extra costs that may incur as a result of more extensive, future treatments. The model also assesses the impact on road users.

- HRI - Crop-Rotation Decision Support System

An EU funded project is currently underway, co-ordinated by the Horticultural Research International (HRI) with partners in six European countries, to develop a Decision Support System to plan crop rotations across Europe. This has the dual aims of maximising economic yield and minimising environmental pollution through nitrogen leaching. At the core of the system are a number of sophisticated mathematical models simulating nitrogen movement in plant and soil environments. Tessella is involved in developing a framework for the system as a whole, which will allows user at all levels, such as farmer, scientist or EU policy maker, to plan design and test the effectiveness of different crop-rotation practices through a series of what-if scenarios.

Statistic Based Systems

The most popular of these are systems based on 'Bayesian Statistics'. This approach involves the development of 'belief networks' to model the system and its properties. Each of these properties are assigned likelihood probabilities and are also linked via conditional probabilities, i.e. what are the probabilities that property A will take particular values, given the value of property B. The final network is a complex web of influences and conditional probabilities. The values of these probabilities are initially 'guessed', either based on historical data or professional and technical 'gut-feel'. Bayesian theory then allows refinement of these probabilities and the network as a whole, as more information about the system is obtained. The 'answers' to questions are then based on a 'most probable' or 'least (probably) costly' (and so on) basis.

Bayesian belief networks are some of the most flexible Decision Support Systems, with a wide application area supporting everything from top-level business and marketing decision making, to technical diagnostic systems. Of all the approaches to DSS, this is the most natural solution to the system with 'incomplete knowledge' – by definition, as knowledge is gained; the system updates its internal relationships.

This technology could find application in the pharmaceutical industry, for example, in clinical trials and in the determination of subject population groups based on the accumulation of data continually gathered.
Another example of this approach in the pharmaceutical industry is in the portfolio management of drug discovery and development pipelines to best plan a strategy for candidate drug development given the uncertain global climate and the high-risk nature of the industry as a whole.

Tessella has been involved in a number of projects employing Bayesian statistics and many projects specifically utilising Bayesian Belief Networks (BBN). One such project, for a large consumer goods manufacturer, developed a bespoke BBN that helps in the decision making process by showing what factors positively or negatively influence a desired outcome, through a process of 'what-if' scenario planning.

Tessella has developed a BBN system for a large petrochemical company which assists in the decision making process for a range of issues, from where to drill a well to where to invest spare capital reserves.

Hybrid Systems

In large multi-national companies, such as those in the pharmaceutical and engineering industries, it is likely that a full DSS will be an amalgam of various approaches, with each component tailored to optimize the decision making process in its scope. A model-based system could feed in to a rules based system or they could work side by side in building up a solution to the overall problem.

IT Requirements for Decision Support Systems

Decision Support Systems are user interactive computer systems, which can vary in scale enormously depending on their application. The IT requirements are thus heavily dependent on the scope. These requirements span not only the development of the DSS itself, but also the architectural infrastructure under-pinning it, along with the hardware requirements to host and support it. For a desktop DSS the demands are small, but for an enterprise DSS the development of the system is a significant and complex undertaking, with all the commensurate issues surrounding distributed application development. An enterprise system will typically be web based with a commercial strength database, such as Oracle, for data storage, application server and web server technology, and a distributed web application as a front end.

Irrespective of the scale and nature of deployment, the DSS itself will be useless unless the requirements defining its application area are very carefully established – the software engineering of such a system is critical to its success, even more so than with traditional applications which provide a defined amount of fixed functionality – a badly specified DSS may 'work' perfectly as a software system, but will not provide the necessary benefits unless the application domain and purpose is clearly defined, the data gathering is complete and the rules and modelling methodologies clearly set up.

Requirements must be developed through coordination of personnel with a balance of IT and business domain knowledge, to ensure that the system defined and the data gathered and presented, match the business need.

Data Gathering

As has previously been mentioned, data classification and gathering is the first essential stage in developing a DSS. It simply cannot be stressed strongly enough: a DSS is only as good as the data and rules under-pinning it. Depending on the specific nature of the project, the approach to data gathering may vary but it will always be the case that data gathering is a major component, requiring some or all of the following:

- A clear definition of the problem to be solved, its scope and bounds.
- A solid understanding of the business domain, facilitating identification of the data to be gathered.
- A detailed programme of work collecting data, which may well be an ongoing process subsequent to delivery, with new data being fed into a system to either extend its range of analysis or to update 'knowledge' about a system providing updated statistical information for a BBN, for example.
- A process to identify and define the rules relating the components of the data.

The elapsed timescale for the data gathering process is likely to be greater than that for IT development and it is a component that should not be underestimated.

For example, data collection often involves a labour intensive collation of physical data and measurements on a national or international level. This can be necessary both for the initial data store and as part of the business process that the DSS supports. As an example of the first, the previously mentioned

HRI project has an initial two year period of gathering empirical data across Europe for physical soil characteristics, weather patterns and crop characteristics. This will be augmented continually as new regions and crop types are characterized. An example of the second is the railway asset monitoring system described earlier - this will work out the most 'appropriate' assets to test at any given point in time based on the information previously stored. The tests and measurements made will then feedback into the system to update the information, along with information about reported failures and problems elsewhere and so on.

Data Acquisition – Process Integration

A DSS is often only one component in some of the more dynamic and complex systems required in industry and business. Integration of business processes and the appropriate choice of approach, is a complex problem in itself. One such area of integration is precisely that of data gathering, both initially and as part of the business problem. Data acquisition is a subject that merits its own technical article, but as alluded to earlier, it is itself often a component in some of the systems of which a DSS is a part. Data acquisition can be used in two contexts with DSS:

- As part of the system process 'pipeline', continually feeding into the data gathering process for a DSS.
- As a part of the working of a DSS providing direct input into the models, for example, as described for asset monitoring.

In the latter case, the system as a whole will be 'open'. This system takes data input (for example physical measurements), runs them through a DSS analysis tool, perhaps a Bayesian diagnostic system, and reports to an operator who can take action based on the data presented. An example of this is given by a project Tessella worked on for Powergen, which developed a system for power station condition monitoring – this system attempted to balance two conflicting objectives (maximum energy output, minimum polluting emissions) by automating the operating process and applying complex data acquisition techniques to make several key physical measurements. All of the information was fed into a DSS, processed, and fed to an operator providing information on the best courses of action to take to keep the system running optimally. A closed system is one that takes the central decision analysis system and automates the feedback response, according to rules devised during the 'manual' operation.

Commercial Systems

There are Commercial Off The Shelf (COTS) packages available, particularly for the development of Bayesian Belief Networks. The development of this Bayesian approach is relatively generic and such COTS systems can be used as a base ‘engine’ for a bespoke system. However, in practice individual business needs, particularly larger scale systems, are rarely catered for fully with COTS systems and a great deal of customisation or data interfacing needs to be developed anyway – they also often provide a great deal of functionality not relevant to a business. When considering a COTS package, a detailed evaluation is essential, not just of its functionality and features, but of how it can provide the functionality *required* by the business.

Conclusion

Decision Support Systems are used extensively in business and industry to assist in decision-making across a wide spectrum of problem areas. There are a number of approaches and techniques employed, from simple data reporting tools, to sophisticated AI systems using Bayesian statistics or genetic algorithms. With the increase in their use must come a word of caution – a decision support system will not eliminate ‘bad’ decisions, or badly drawn conclusions derived from asking the ‘wrong’ questions. The output of the DSS must be examined critically and used together with existing understanding of the wider business or application areas. This being said, systems are now being developed which are providing enormous benefits, both in time and cost savings, and this is only set to improve as the technology advances.

Tessella - providing innovative software solutions to scientific, technical and engineering problems

Tessella specializes in the application of innovative software solutions to scientific, technical and engineering problems. Our services cover software design and development, IT consultancy, infrastructure support and project management.

Other Technical Supplements published by Tessella include:

- ❑ Active Server Pages
- ❑ Archiving of Electronic Information
- ❑ Asset Monitoring
- ❑ Automated GUI Testing
- ❑ Bayesian Statistics
- ❑ Beowulf Clusters
- ❑ Beyond LIMS
- ❑ C++
- ❑ Choosing and Using a LIMS
- ❑ COM
- ❑ Computational Fluid Dynamics
- ❑ Computer Image Processing
- ❑ Decision Support Systems
- ❑ Development for the Mobile Platform
- ❑ Digital Preservation Practical Experiences
- ❑ e-GIF
- ❑ Electronic Data Capture
- ❑ Electronic Lab Notebooks
- ❑ Evolutionary Computing
- ❑ Excel
- ❑ Extending the Life of Software
- ❑ FDA 21 CFR Part 11
- ❑ Formulation
- ❑ FORTRAN 90
- ❑ Grid Computing
- ❑ High Throughput Experimentation
- ❑ High Throughput Screening
- ❑ Instrumentation
- ❑ Integrated Lab Systems
- ❑ J2EE
- ❑ Java
- ❑ Linux
- ❑ Microsoft .NET
- ❑ n-tier Architecture
- ❑ Object Oriented Programming
- ❑ Open Source and Free Software
- ❑ Pocket PC
- ❑ Portable GUI Development
- ❑ Real Time Systems
- ❑ Regression Testing
- ❑ Security and the Internet
- ❑ Simulation
- ❑ Soft Computing
- ❑ Software Development Cycle
- ❑ Software Documentation
- ❑ Software Portability
- ❑ Software Re-engineering
- ❑ Software Specification
- ❑ SQL
- ❑ UNIX Inter-Process Comms
- ❑ UNIX System Performance
- ❑ Web Services
- ❑ Web Scripting
- ❑ Windows 2000 Services
- ❑ Workflow Systems

Certificate No. FM 22778

INVESTOR IN PEOPLE

Tessella Support Services plc
3 Vineyard Chambers, Abingdon, Oxon, OX14 3PX, England
Tel: (+44) (0) 1235 555511 Fax: (+44) (0) 1235 553301
info@tessella.com www.tessella.com