# DECISION TREE MODULE WITHIN DECISION SUPPORT SIMULATION SYSTEM

Mohamed Moussa
Dr. Janaka Y. Ruwanpura
Dr. George Jergeas

Project Management Specialization
University of Calgary
2500 University Drive NW
Calgary AB T2N 1N4 CANADA

## ABSTRACT

Decision trees are one of the most easy to use tools in decision analysis. Problems where decision tree branches are based on random variables have not received much attention. This paper introduces a successful application of a Special Purpose Simulation (SPS) program in developing a Decision Tree module that is part of a unified Decision Support System (DSS) template. The DSS template consists of three modules: Decision Tree (DT), shortest and longest path Dynamic Programming (DP) Network, and Cost / Time (CT) Estimate network. The DT module integrates with other modules and allows users to model decision trees with variables that are based on probabilistic random numbers. This paper introduces the DSS-DT module and shows its advantages.

## 1   INTRODUCTION

Decision trees are one of the most attractive and easy to use tools in decision-making. They analyze decision alternatives in a systematic, chronological way and provide an easy to read graphical presentation to decisions under considerations. However, solving DT's using single number estimate has several disadvantages that limit the DT's applications. With the use of computer-based simulation programs, additional features are added to the DT's to increase their usefulness and enhance their use.

SPS programs provide a unique solution to modeling methodologies. They provide tools that can be used by non-expert simulation users. SPS program allows an experienced simulationist to build a template that can model problems of general nature and that can be used by inexperienced user. This paper presents an application of a SPS program in producing a reusable DT template that can be easily used by non-experienced simulation users.

The paper has two objectives:

- to present a successful application of a SPS program in modeling DT's
- to contribute to the body of knowledge by introducing the added features of the DT's built under a SPS platform.

The template is created using Simphony, a SPS computer program developed by the University of Alberta. For more information about Simphony, readers may refer to AbouRizk & Hajjar (1998) and Hajjar & AbouRizk (2002).

The paper starts with highlighting the DT conventional concept including a brief introduction on the utility theory. Then, an introduction to stochastic DT's is provided. Finally DSS-DT Module and its advantages are presented.

## 2   DECISION TREES

A DT is "a graphical presentation of expected value (EV) calculation" (Schuyler 2001) – it is "an attempt to explicitly determine the alternatives available … the expected costs, payoffs and probability of success associated with each alternative" (Mock 1972). The DT's show the sequence of decisions and their consequences with the objective to select the best course of action. By solving the DT the analyst/ decision-maker is seeking the optimum solution based on the maximum (or minimum) Expected Value (EV).

A DT can also be an analysis tool by examining the outcomes under wide range of input variables. The latter is best modeled using stochastic values i.e. inputs are based on probabilistic distribution estimates.

Hespos & Strassmann (1965) distinguish between 1) DT's where single number estimates are used and decisions are made based only on the EV criterion; and 2)

DT's where probabilistic values are used i.e. where the concept of risk/utility is used in selecting decisions. They named the former a "conventional DT" and the latter a "stochastic DT;" same distinguish is used in this paper.

## 2.1 Conventional Decision Trees

Conventional DT's are covered in wide range of publications (Raiffa & Schlaifer 1961; Meredith et al. 1973; Revelle, Whitlatch, and Wright 1997; Taha 1997; Moore & Weatherford 2001; Hillier & Lieberman 2001). A brief explanation on the DT components and computational procedure is provided below for easy reference. A conventional DT consists of five main components:

- Decision Nodes: represented by squares that precede variables or actions that the decision-maker control
- Chance Event Nodes: represented by circles that precede events that cannot be controlled
- Terminal / End Nodes: End points where outcome values (payoff's) are attached
- Decisions: branches out-coming from Decision Nodes; they represent alternative decisions available to decision-makers
- Chance Events: branches out-coming from Chance Event nodes; they represent possible outcomes of decisions; they are assigned probability.

A tree has one root, normally a decision node, drawn chronologically from left to right; branches (decisions & outcomes) are radial lines originated from the nodes. Figure 1 shows a simple DT with two decisions and three chance-events resulting from each decision. To solve a conventional DT, the analyst calculates the expected value (EV) by**:**

1. identifying alternative decisions and their cost.
2. identifying possible outcomes of each decision
3. estimating the probability of the outcomes; if there are subsequent decisions or outcomes steps 1 to 3 are to be repeated
4. drawing the tree chronologically from left to right and calculating the payoffs at end of each branch
5. folding-back the tree to calculate the EV and taking the decision that has the optimum EV.

The EV is calculated so that:

- at chance events the EV equals the sum of the values of the chances multiplied by their probability:

$$EV = \sum_{i=1}^{i=n} X_i P(X_{i)} \qquad (1)$$

n is the number of chance events i's at the Node

- at the decision node the EV equals the optimum EV at the node (minimum or maximum as per the optimization criteria required).

A DT is constructed from the bottom up (from time zero up), analyzed from the top down, and finally implemented from the bottom up (Revelle, Whitlatch, and Wright 1997). Values in the tree should be discounted to the net present value.
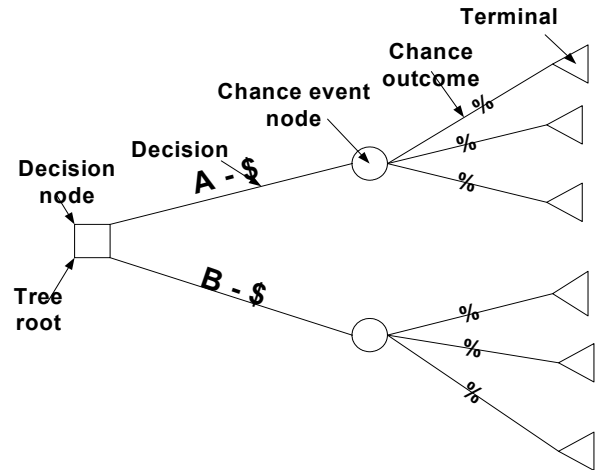


Figure 1: Decision Tree Presentation – Main Components

Taking decisions based on the EV alone and using single values in estimating future events have been subject to several criticism(Schuyler 2001; Hespos & Strassmann 1965; Smith, Hinton, and Lewis 1983; Hartman 2000; Ferrara & Hayya 1970; Durrenberger 1999). A summary of these criticisms are given below.

1. The EV assumes repetitive decisions and does not take into consideration the one-time decision. Decisions based on the EV do not take into consideration the risk attitude of the decision-maker and do not show the risk involved with decisions.
2. In conventional DT's, the number of branches can get larger in complex problems. It assumes no feedback to previous decisions, and hence if feedback is required, the tree gets larger.
3. The use of single number estimate assumes knowledge of future events with certainty and gives no information on the range of possible outcomes or the probability associated with these outcomes.

The conventional DT is unable to reflect the underlying uncertainty in decisions related to future events. To overcome the problem of the EV, analysts may use the utility function of the decision-maker to solve the tree. Utility function measures the relative preference of the decision-maker and the value of payoffs to him/her (Smith, Hinton, and Lewis 1983). By measuring the utility of the decision-

maker, the payoffs are translated to reflect this utility, the DT can then be solved using the utility converted factors (Smith, Hinton, and Lewis 1983). In other words, instead of taking the decision that reveals the optimum Expected Monetary Value (EMV), the decision can be for the maximum Expected Utility (EU). The utility theory considers the risk attitude of the decision-makers, and hence by using the utility function, the tree can be suitable for the one-time decision. For more information on methods of measuring the utility function readers may refer to Hillier & Lieberman (2001), Revelle, Whitlatch, and Wright (1997), and Taha (1997). The Utility concept is considered a theoretical concept more than a practical one (Wideman 1992):

- Measurement of utility is difficult
- In case of several decision-makers – the tree should be solved for each utility function; it can't be used as a communication tool among several decision-makers
- Utility method still doesn't not provide information about the range of possible outcomes and associated probability.

## 2.2 Stochastic Decision Trees

Stochastic trees are "extension of DT's that facilitate the modeling of temporal uncertainties" (Hazen & Pellissier (1996). Stochastic DT's permit the use of probability estimates or frequency distributions for some or all factors affecting decisions (Hespos & Strassmann 1965). Applications of stochastic trees can be found in Hespos & Strassmann (1965) and in Hazen & Pellissier (1996), and Hazen (2000). A stochastic DT is defined in this paper as a tree whose:

- decisions (branches from decision nodes) may be estimated as density distribution function values
- chance events (branches from chance-nodes) may be estimated as density distribution function values
- nodes my be connected to branches with probabilistic or/and constant values.
- chance events problaity may be estimated as a range of the likely probability.

Branches with density distribution values are analogous to large number of branches in a conventional DT (Hespos & Strassmann 1965). Stochastic DT's overweigh the conventional method in that information about decisions and outcomes are estimated in ranges rather than in single numbers reflecting the real-life uncertainty in estimating values of future events. Results from stochastic DT's can be obtained in a probabilistic form; the information can then be analyzed using the concept of utility and risk (Hespos & Strassmann 1965).

## 3 DECISION SUPPORT SYSTEM (DSS) UNDER SIMPHONY PLATFORM

Simphony, a SPS program, has been used in several applications including earth-moving, tunnel construction, range estimating, and PERT time estimate; see Hajjar & AbouRizk (2002) and Ruwanpura et al. (2001).

Developers use the Simphony editor to develop a SPS template. A template is a collection of modeling elements that are targeted for a single domain (Hajjar & AbouRizk 2002). The user builds a simulation template in Simphony by creating modeling elements that resemble real components of a model/system and linking them together. Simphony provides a hierarchical modeling feature. A system can be represented by an abstracted model at a level; at a lower level, each of the model elements can have its own child model, which represents the sub-system inside that element (Hajjar & AbouRizk 2002 and Simphony User's Guide 2000).

The DSS template includes the elements required to model the three modules. The template consists of 16 modeling elements, which are colour coded so that elements with similar colours are used by one module. The DT module uses 10 modeling elements (see Table 1) which are required to model a stochastic DT.

Table 1: DT Modeling Elements

| Element Name | Function |
|---|---|
| Root Element # 1 | The higher level element of a network |
| Start Node # 2 | The first element in a DT or any other network |
| DT Decision Node # 3 | Decision node in a DT; precedes decisions alternatives |
| DT Chance Node # 4 | Chance node in a DT; precedes chance events |
| DT Terminal # 5 | Terminal of a DT branch |
| DT Branches Sum # 6 | Sums values of DT Terminals. Combines risk values of any decision or combination of decisions. |
| Resource # 13 | Represents resources / costs associated to a decision. |
| Probability Relation Parent # 14 | The parent element for chance events (Probability Relation Child # 15). |
| Probability Relation Child # 15 | An event. Child of a Probability Relation Parent. |
| Link/Branch # 16 | A decision. |

Figure 2 shows the DSS template with the DT elements marked.

Figure 3 shows a DT using the DSS notations. As seen in Figure 3, the DSS has additional elements not in the conventional DT's (see Figure 1). The "DT Branches
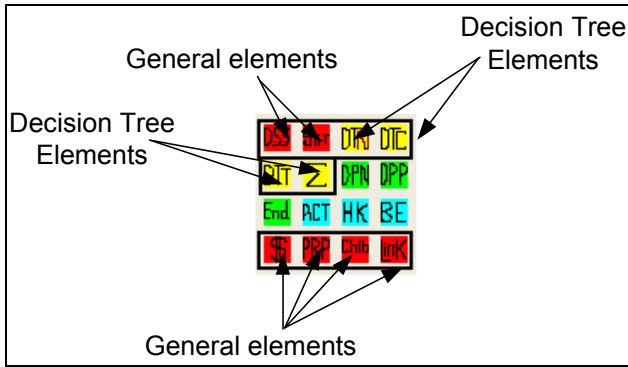
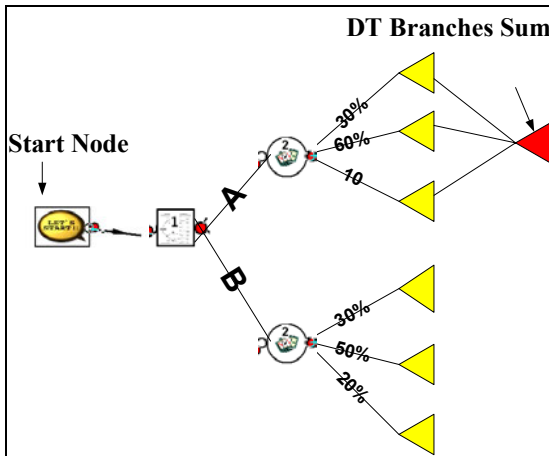Figure 2: DSS Template - Modeling Elements Icons



Figure 3: DT Using DSS Notations

Sum" collects the payoffs of realized branches and produces the risk/utility profile of any decision or any combination of decisions. The "Start Node" creates the required entities for the simulation (initiates the simulation).

## 4 DSS TEMPLATE - DECISION TREE MODULE

A DT using DSS should be constructed in the lower level of the "Root Element # 1" (See table 1). Figure 4 shows a conceptual graph showing the relationship between the "Root Element" and the DT. This section describes the features of the DT-DSS and shows how these features enhance DT applications.

### 4.1 DT - DSS Template Features

The DT- DSS template has the following features:

- Each decision and chance-event can be estimated using the cost and/or benefit values; cost indicates spending (e.g. cash-out); benefit indicates income (e.g. cash-in).

- For each cost/benefit of a decision or of a chance-event, users have the option to input the values either as constant value (single number) or as a density distribution function (Uniform, Beta, Exponential, or Normal distribution).

- A node in DT-DSS (a decision node or a chance-node) may be connected to a branch with probabilistic (density distribution) value, a constant value, or both.

- Cost may be estimated using detailed estimate rather than based on one lump sum estimate (whether it is constant or probabilistic estimate). With the "Resource" modeling element, users can assign any number of resources to a decision. Resources may represent labour, materials, money allocation, etc. The cost of each resource is a function of the quantity of the resource and the unit rate. Both resource quantity and unit-rate can be either constant or density distribution function. The number of resources assigned to a decision is limited by the computer capacity. Assigning resources to decisions provides the modelers a flexibility in estimating the cost based on different combination of density distribution estimates.

Decisions' cost may also be estimated as a function of the cost of another network. The DSS template supports a hierarchal structure that allows networks to be modeled in a lower level of some modeling elements. The hierarchical characteristic facilitates integration among the modules i.e. the output of a network may be the input of another network. For example, the cost of a decision in a DT problem may be estimated as the cost of a project resulting from a cost range estimate modeled by the CT network module.

The first element in any network is the "Root Element", the DT is constructed at a lower level of the "Root Element". The "Root Element" collects the network information. The "Root Element" may be constructed in a lower level of another element. In such case, the information collected by the root element is used as input to the element in the higher level.

Figure 5 shows a conceptual diagram that illustrates the hierarchal relationship in the DT–DSS modules. The first element in the hierarchy is the "Root Element." It includes the DT network in its lower level. A decision in the DT network has another "Root Element" in its lower level. The latter "Root Element" has a project network in its lower level. Similarly more/other networks may be linked to other elements. The tree is calculated from down-up i.e. networks at the lower levels are calculated first.

This feature allows the user to estimate the cost of a decision based on the cost of another stochastic network. Such network may be another DT, a DP or a CT network with or without probabilistic realization of nodes, probabilistic values, and/or resources and networks in a lower level.
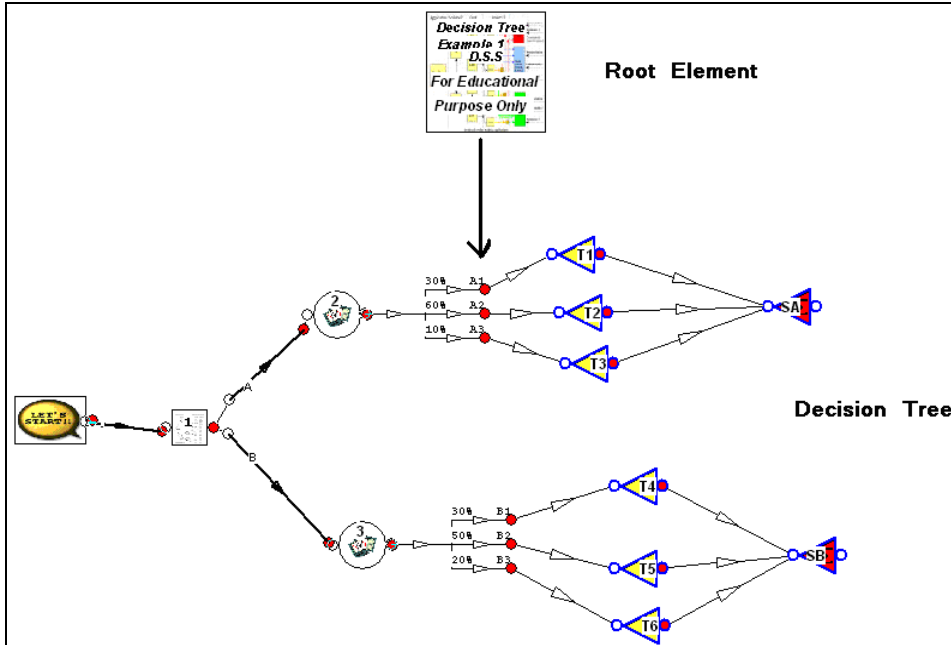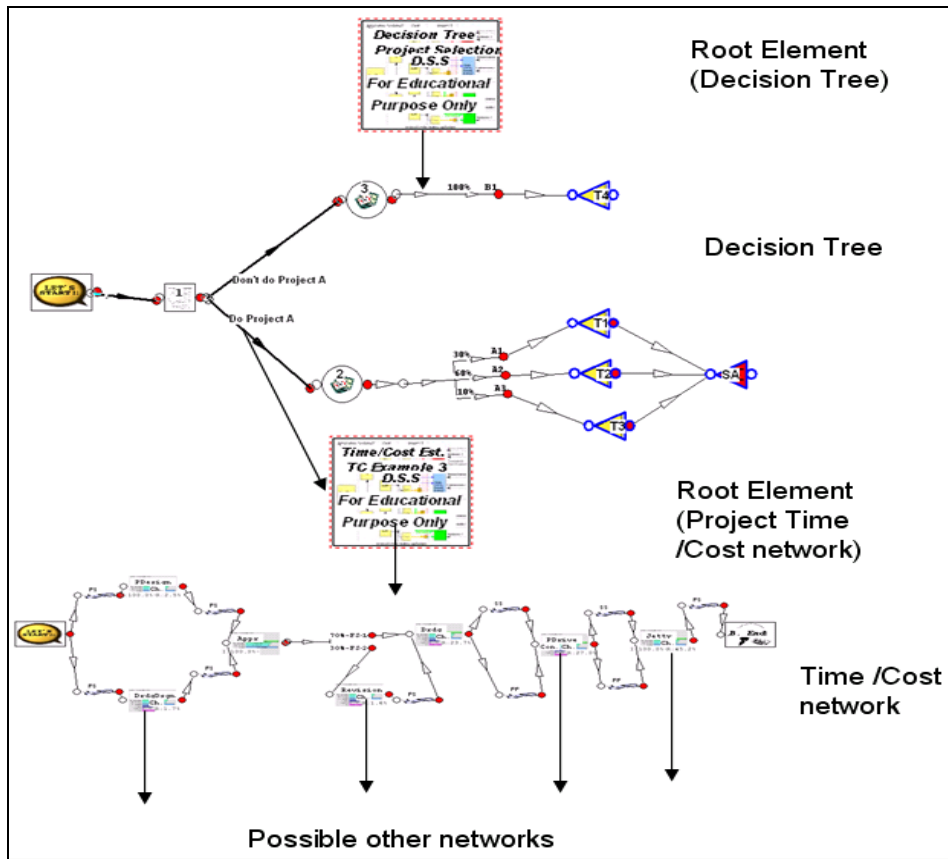
Figure 4: Decision Tree – DSS Screen



Figure 5: Decision Tree Integration Example

The user may use any combination of the three methods to input values in the model i.e. cost of a decision may be a result of resource costs, a network cost, or/and a direct estimate. Such flexibility provides the modeler the chance to model DT's whose costs do not necessarily fit into a standard probability distribution. (For the importance of using a correct distribution refer to AbouRizk, Halpin, and Wilson (1994)).

- probability at the chance events may be estimated as single numbers (i.e. the total estimate of chance events at each chance node should be 100%) or as a range of the likely probabilities. The range of the likely probability allows using range of estimates. For example, in a three chance-events-node, the user may specify the probability of outcomes as 30-40% for the first event, 60-75% for the second event, and 5-10% for the third event. Such flexibility provides users a more realistic approach in estimating the likelihood of events' occurrence.
- The DSS permits feedback to previous decisions or event nodes and accepts loops (cycles) based on probabilistic estimate. it also accepts
- Optimization of the EV can be with either minimum or maximum EV as specified by the user .

## 4.2 DT- DSS Template Calculation and Results

During each simulation run, the DSS assigns independent random values to each of the tree variables based on the specified distribution and parameters. Variables include the costs, the benefits, and the probability ranges at chance event nodes. If there are resources or networks at the lower level of any decision, the DSS calculates these resources and networks first; then adds such costs to any other costs input directly into the decision. The tree is then calculated through:

- forward path calculation: the DSS calculates the probability of realization at each node and the payoffs at the terminals. At each chance-node the DSS generates an independent random number based on which one branch is realized. The utility/risk at each realized terminal is calculated as the payoff at this terminal. Utility/risk of unrealized terminals is zero. The program collects the statistics of utility/risk at each terminal. The statistical results provide the probability and the range of payoff's resulting from each decision or a combination of decisions. Assessment of the risk/utility is later explained in section 4.3.
- backward calculation: the program calculates the EV at each node using the payoff values calculated in the previous step; the DSS also calculates the optimality index for each decision (the prob-

ability for a decision to fall on the path that yields the optimum EV); the optimality index = the number of times a decision falls on the tree solution path / number of iterations. (tree solution path is the combination of decisions that yield the tree optimum EV)

Once all simulation iterations are is complete, the program collects the statistical results:

- for input values so that the user may review the estimates provided – this is particularly important if the input values are delivered from another network or from the costs of allocated resources
- for calculated results: 1) EV 's at any node, 2) Payoff's at terminals, and 3) Risk/utility for any decision and any combination of decisions at the "DT Terminal" and "DT Branches Sum" nodes

## 4.3 Assessment of Utility and Risk Using the DSS-DT Module

The DT-DSS may be used for any decision tree application. We provide a hypothetical example to explain how the DT-DSS may be used in supporting the decision process. Assume that a company must select between two projects: Project A and Project B (see Figures 1 and 3). Inputs (estimated costs of projects and estimated outcomes) are estimated as density distribution values. After simulating the model for a large number of iterations (here 1000 runs), the outputs show the range of possible outcomes from each project.

Figure 6 shows the simulation results for projects A and B. For project A, the optimality index (chances that the project provides the maximum EV) is 59.5% while that for project B is 40.6%. The maximum, the minimum, the average and the standard deviation of the EV can be retrieved and the users can examine the cumulative density distribution (CDF) graph for the EV of each decision (Figure 7) (x-axis=$, y-axis=probability).



Project A



Project B
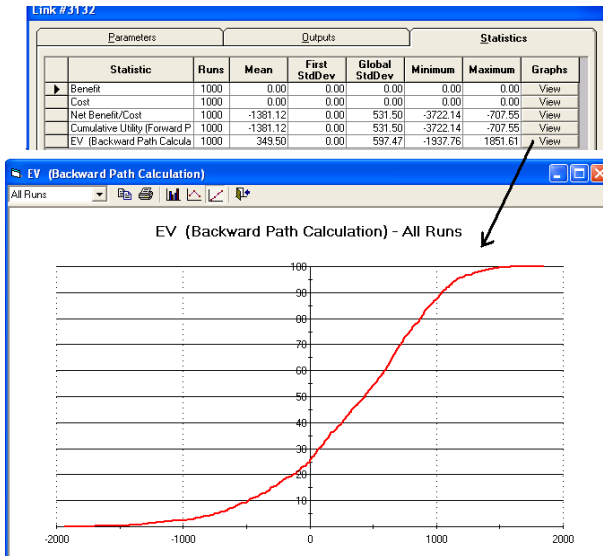
Figure 6: Output Screen Decision A and B

Figure 7: Project A Statistics Screen and the EV – CDF

The risk/utility of the projects A and B is provided through statistics of the "DT Branches Sum" (Figure 8). Figures 9 and 10 show the CDF graph for each decision. The "DT Branches Sum" collects the value of the payoffs for the realized chance-event branches; hence, the statistical results reflect the possible range of outcomes for each decision.
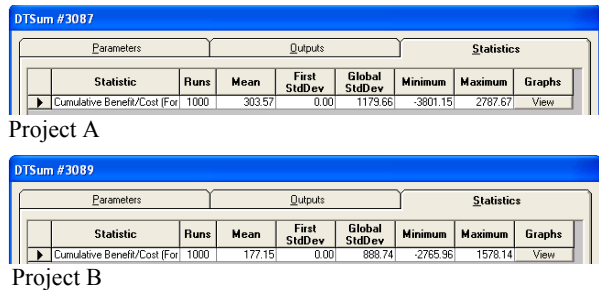


Project A



Project B

Figure 8: Outputs Results – Utility/Risk at Branches Sum Example
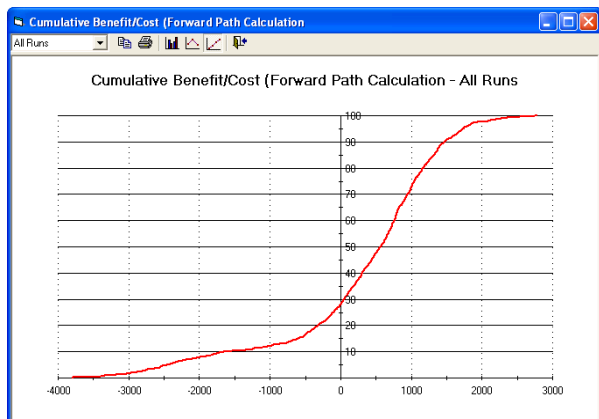


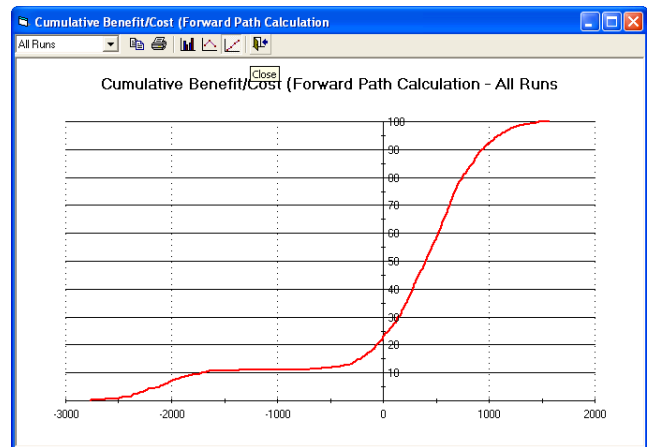Figure 9: CDF Graph Decision A



Figure 10: CDF Graph Decision B

By analyzing the outputs of the utility/risk shown in figure 8, the decision maker can take the proper decision for his/her problem. Recognizing the larger loss of project "A" (-$3,801 for "A" versus -$2,765 for "B") and the small difference of EV optimality (59.5% versus 40.6%), a risk-averse decision-maker may select decision B. Also, from Figures 9 and 10 (CDF graphs of the information shown in figure 8), the decision-maker may consider that the chance of loss in "A" is about 29% while in "B" is about 23%. Another decision-maker may look at the possible outcome at a specific probability – for example at 70% chance of success. At such probability, project "A" yields $1,000 or less while "B" yields $700 or less. A risk-taker may take decision "A" for the higher income.

The decision-maker is now more informed about the range of possible outcomes of each decision alternative and the probability associated with any value and is in a position better than that of using the EV alone. In case of repetitive decisions, where the EV is argued to be an adequate decision criterion, the decision-maker is able to get the possible range of EV's for each decision (Figure 7) and the value at any specific probability.

Similarly any node can be examined. Also the results of each run can be obtained in a database file for more analysis if required. In case of sequential decisions, the "DT Branches Sum" may connect to any terminal representing any decision or to several terminals representing a combination of decisions. Figure 11 shows a sequential decision tree where the "DT Branches Sum" (marked with arrows) were used to collect information about the risk/utility of several decisions.

## 5    CONCLUSION

Decision Trees (DT's) provide powerful methodology to visualize and analyze decisions. Conventional methods of solving DT's do not provide information about the risk involved in decisions and are not suitable for one time decisions. We introduced an application to DT's by using a
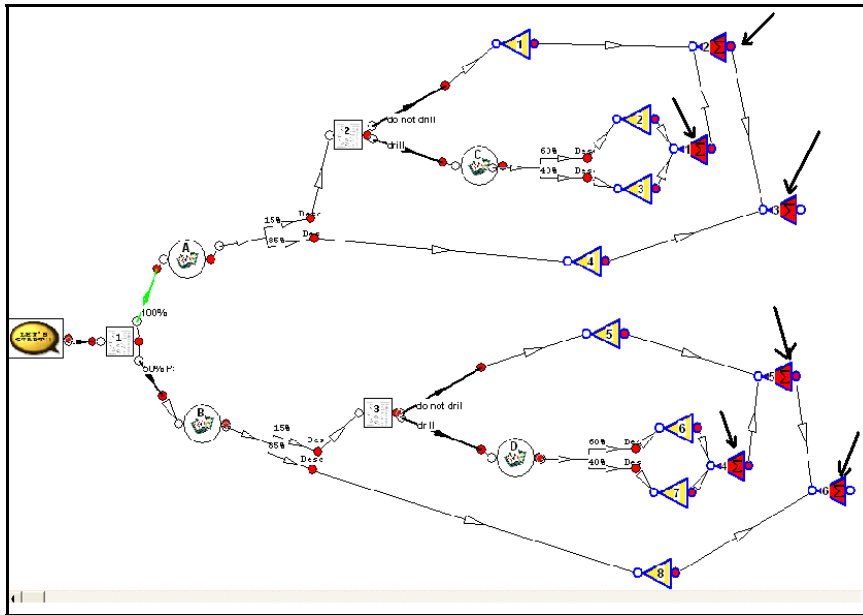
Figure 11:  Sequential DT - "DT Branches Sum"

Decision Support System (DSS) template that is developed under Simphony – a Special Purpose Simulation (SPS) program. The DSS-DT module accepts constant and density distribution values and integrates with other modules in the DSS template. Decisions' costs can be estimated directly by inputting values in the elements or by detailed estimate. The DSS-DT Module accepts loops (feedback to previous decisions), allows estimating probability of events based on range estimate, provides the optimality index for each decision, and provides the utility/risk profile for any decision and any combination of decisions.

## ACKNOWLEDGMENTS

## REFERENCES

AbouRizk, S. M. & Hajjar, D. (1998). "A framework for applying simulation in construction." Canadian Journal of Civil Engineering, pp. 604-617. Vol. 25, No. 3.

AbouRizk, S. M., Halpin, D.W. & Wilson J.R. (1994). "Fitting Beta distribution based on sample data." Journal of Construction Engineering and Management, pp. 288-305, Vol. 120, No. 2.

Durrenberger, M. R. (1999). "True estimate reduces project risk." PM Network, pp. 45-48, May 1999.

Ferrara, W. L & Hayya, J. C. (1970). "Toward probabilistic profit budgets."  Management Accounting, pp. 23-28, October 1970.

Hajjar, D. & AbouRizk S.M. (2002). "Unified modeling methodology for construction simulation." Journal of Construction Engineering and Management, pp. 174 – 185. Vol. 128,  No. 2.

Hartman, F. T. (2000). Don't park your brain outside: A practical guide to improving shareholder value with SMART management. Project Management Institute, pp. 159-190 & pp. 103-109..

Hazen, G. (2000). "Preference factoring for stochastic trees." Management Sciences, pp. 389-403. Vol. 46, No. 3.

Hazen, G. B. & Pellissier J.M. (1996). "Recursive utility for stochastic trees." Operation Research, pp. 788-809. Vol. 44, No. 5.

Hespos, R. F. & Strassmann P.A. (1965). "Stochastic decision trees for the analysis of investment decisions." Management Science. pp. B244 - B259. Vol. 11.

Hillier, F.S. & Lieberman G.J. (2001). Introduction to Operation Research (7th ed.). The McGraw-Hill Companies, Inc. New York.

Meredith, D. D., Wong, K.W., Woodhead, R.W. & Wortman R.H. (1973) Design and planning of engineering systems. Prentice-hall. New Jersey.

Mock, T. J. (1972). "Decision tree approach to the methodological decision process." Accounting Review, pp. 826. Vol. 47, No. 4.

Moore, J. H. & Weatherford L.R. (2001). Decision modeling with Microsoft Excel (6th ed.). Prentice Hall. New Jersey.

Raiffa, Howard & Schlaifer, Robert. (1961). Studies in Managerial Economics: Applied Statistical Decision Theory. Division of Research: Graduate School of Business Administration. Harvard University. Boston.

Revelle, C. S., Whitlatch E.E, &Wright J.R. (1997). Civil and environmental systems engineering. Prentice Hall,

Inc. Upper Saddle River, New Jersey. Reprinted with correction June 1999.

Ruwanpura, J. Y., AbouRizk S.M., Er K.C., & Fernando S. (2001). "Special purpose simulation templates for tunnel construction operations." Canadian Journal of Civil Engineering pp. 22-237. Vol. 28.

Schuyler, J. (2001). Risk and decision analysis in projects (2nd ed.). Project Management Institute. pp 41 –74.

Simphony User's Guide. (2000). NSERC/Alberta Construction Industry Research Chair. University of Alberta.

Smith, A. A., Hinton E. & Lewis R.W. (1983). Civil engineering systems analysis and design. John Wiley & Sons.

Taha, H. A. (1997). Operations research: An Introduction (6th ed.). Prentice Hall. Upper Saddle River, New Jersey.

Wideman, M.R. (ed). (1992) Project & program risk management: A guide to managing project risks & opportunities. Project Management Institute.

## AUTHOR BIOGRAPHIES

**MOHAMED MOUSSA** is a Ph.D. candidate in the Department of Civil Engineering at the University of Calgary. He holds a Bachelor of Science in Architecture Engineering – Ain Shams University / Egypt, a Post graduate Diploma in Management Information System and Computer Science, and a Master of Science in Civil Engineering-Project Management specialization from University of Calgary. He has more than 15 years of construction management experience in heavy civil and marine construction project.

**DR. JANAKA Y. RUWANPURA**, a former US Fulbright Scholar, is an Assistant Professor in the Project Management Specialization at the University of Calgary. He is also an Associate Chair in Design Engineering. He earned his B.Sc. (Honours) from the University of Moratuwa, Sri Lanka and his M.S. in Construction Management from Arizona State University, and Ph.D. in Construction Engineering and Management from University of Alberta. Dr. Ruwanpura has experience working as a lecturer, professional quantity surveyor, project manager, and program developer for public and private sector firms in Asia and Canada. He has initiated various project management and simulation based research projects with industry support and has published over 40 papers in refereed journals and conference proceedings. His current expertise includes simulation and modeling, risk management, cost forecasting, construction productivity improvement, sustainable development, and project management. Dr. Ruwanpura is an active consultant in PM including training and guidance for projects in international countries such as Mexico, China, Portugal, and Sri Lanka.

**DR. GEORGE JERGEAS** is an Associate Professor in the Project Management Specialization at the University of Calgary. He earned his B.Sc. in Civil Engineering from the University of Baghdad, Iraq, and his M.Sc and Ph.D. in Construction Management from Loughborough University, UK. He has over 20 years of industry experience. Dr. Jergeas joined the university as a full-time academic coming from the industry in 1994. His career in industry was in the construction of international infrastructure projects. During his employment with Revay and Associates, he investigated over 30 construction projects, both in Canada and the USA, experiencing cost overrun and delay. Dr. Jergeas' experience in industry and subsequent research has demonstrated a strong interest in improving the efficiency of construction projects. Dr. Jergeas' expertise includes construction team building and partnering, claims and disputes, delay analysis, alternative dispute resolution, risk allocation and management, project controls and contract administration. Dr. Jergeas is also an active project and construction management consultant for both public and private sector organizations in Canada and has provided extensive training to many overseas industry participants.