

Putting the Decision in Decision Support

David A. Swayne

*Computing Research Laboratory for the Environment
Department of Computing and Information Science and
Faculty of Environmental Sciences,
University of Guelph, Guelph, N1G2W1 CANADA
(dswayne@uoguelph.ca)*

Abstract. We consider a novel approach to developing multi-objective environmental decision support applications. We propose using causal probabilistic networks (CPN) to subsume one or more engineering process models, together with CPN implementations of specific decisions and utility measures. So-called decision nodes are set to prior probability assignments of $1/N$ (where the particular decision has N possible values) and utility nodes are based upon the *standard lottery* principle give a value for that decision choice. The decisions can be so-called "one-shot" or sequential (time-dependent or at least seasonal). In the latter case maximization of the utility is calculated by gaming to optimize outcomes over a longer time period. The novelty of this approach stems from the embedding of the models into the CPN, and the potential for increasing the scope of decidability for environmental planners, and the entirely acceptable CPN approach of embedding evidential information in a heterogeneous fashion in addition to the modelling formulation.

Keywords: Graphical probability models, decision support, environmental modelling

1. Introduction

Since 1985, the Computing Research Laboratory for the Environment has been collaborating in the implementation of new information tools and technologies in strategic-level environmental applications. This work has taken on many forms, deriving its primary usefulness by shortening the time between discovery of information technology tools and their application to environmental problem analysis [LA01].

CPN, and more general graphical probability models (GPM), first came to our attention in 1988, as a new paradigm in expert systems implementation that could incorporate uncertainty, dynamic behaviour, and conflicting or inconclusive or partially available evidence [PE86]. GPMs have been extensively used for parameter estimation [VA98], irrigation water management [BA99], fisheries management [VA97], and crop production [GU94]. It is a small step from a so-called *influence diagram* that ecologists use extensively in describing ecosystems to a GP model.

The idea of using a GPM to incorporate the behaviour of an engineering model is new. Its spatial dimension is informative and useful, and its potential for speed is enormous. Incorporation of decisions (using so-called *decision nodes* and

evaluating (through *utility theory*) the inherent "goodness" or "happiness" with the results of the decision is new.

CPN models are based on relatively simple principles. Random variables representing key components of a model are represented as nodes in a graph and causal relationships between variables are represented by arrows or directed edges. If fuel economy E for an automobile is based on engine size S and driving style D , and in turn transportation costs T , a directed graph representing this system would be represented thus (Figure 1). S and D are so-called *causal variables* and T is termed an *evidential variable*. *Basic probability assignment* for the prior states of each node involves the one or more values that the random variable can take, for each value of that variable's parent nodes. Nodes with no parents have unconditional prior probability assignments. *Belief updating* for a variable is accomplished via a recursive process, back through parents, and forward through children [LA88]. The updating process is dependent on the graphical model being representable by a *directed acyclic graph*, principally to avoid indefinite loops in the updating algorithm, although this restriction has been somewhat relaxed. For an exhaustive tutorial, the reader is referred to [RU95].

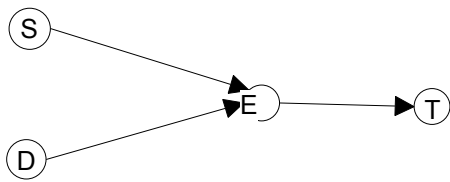


Figure 1. A simple Graphical Probability Model (GPM).

Software should not “make” automatic decisions. It is entirely appropriate to present a value judgement on the quality of decisions about, for example, land use and land use changes, cropping practices, water retention and releases, and other human activities that affect the environment. The decision-maker can then agree or disagree: the character of decision making would change, not the requirement that the decision-maker make the best possible decision given the supporting evidence.

Our laboratory's long-term objectives are: to develop useful decision support systems which evaluate the real or potential impact of decision outcomes; to develop utility measures for environmental activities; use of decision theory to facilitate collaborative work in noisy and untrusted environments; and distributed decision support systems.

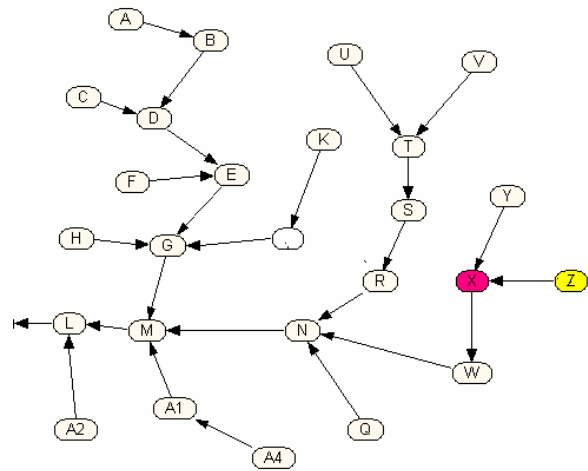
Examples of difficult decisions in the environment abound. Water consumers, for example farmers using irrigation, will compete with barge transporters and hydroelectric generators. All will impinge on forest stands that require flooding and also on fish that are triggered into reproductive mode by a Spring flood and are made uncomfortable by lowering of water quantity and quality. Mining operations that require the entrainment of receiving waters may individually meet water quality requirements. Collectively they may fail to meet the expectations of controlling ministries. Who is at fault?

Suppose, for example, we are modelling erosion and sediment transport in an agricultural drainage network. We must break the watershed into individual plots, with identical slope, land use, soil type, owner (decision-maker). For some models such as AGNPS [YO86], [PA98], these plots are broken into rectangular tiles or other discretization. We will use a model after Dorner [DO2001] et al. called GAMES [RU86]. Within each homogeneous plot, a model like GAMES can predict the generation of sediment within the plot, and a spatial network diagrams the pathway of sediment transported from above and to the plot below in elevation (Figure 2).

2. Agricultural Non Point Source Pollution

Watersheds in agricultural areas are composed of one or more farms, plus pieces of land with non-farm use. These farms are themselves sometimes broken up into individual pieces of land with different land uses. We would call specific acts of different land uses of farm land a *decision*. The cumulative effect of several decisions requires the application of a process model for the environmental effect of activities on the watershed, basin or river system.

Figure 2. A tree structure generated from a set of



plots. The primary downslope direction is indicated by arrows. The receiving waters for this hypothetical example occur "below" node L.

Within a plot or cell, a *decision* can be made about land use or cropping practice that will have an effect on this ratio. Parallel to the physical network is a set of *decision nodes*, one per plot, whose aggregated effect constitutes a *management practice*, possibly for every plot in the whole watershed. Impermeable plots such as roads or buildings, tile-drained fields and point sources are also elements that could be included in the whole watershed model. These other objects would have to incorporate differing models. For example, a housing subdivision might use a standard urban stormwater model. The decision nodes for each plot have, for each of N decision choices, prior probability 1/N. When a particular decision choice for a single plot is selected, setting its value to 1 and the remainder to 0, its effect on the sediment or pollutant is propagated throughout the CPN, to determine the effect of the decision.

For this next discussion, we ignore the model mechanics of GAMES, and concentrate on the example of sediment transport. Then the sediment

S(D) delivered downstream (down-slope) through a particular plot is related to the sediment delivered from the immediate upstream plots $\Sigma S(U)$ and sediment generated internally within the current plot, S(G) by $S(D) = \alpha(\Sigma S(U) + S(G))$. The factor α is called the sediment *cell delivery ratio* (CDR). The CDR is generated by the factors considered in the development of the GAMES model: slope, aspect ratio with respect to steepest descent, roughness, cropping practices, soil type. (Figure 3). Deterministic relationships within the model are calculated directly.

The relationship of the sediment to the so-called probabilities is a formal, rather than an actual one. The total sediment incoming to a cell can be normalized with respect to the maximum that can be generated, with $\alpha=1$ and the maximum sedimentation from above and within being the normalizing constant, to make the S(.) behave like a probability density function. This normalization is itself automatically generated by the *probabilistic inference engine* (typically NETICA™ or HUGIN™).

All models that might fit in this formulation would be "compartment" models, whether they be river reaches, fields, or segments of a manufactured drainage system. They will have a compartment (analogous to the "plot"), immediate upstream compartments, and one or more downstream compartments. The amount of a quantity (excess water, sediment, pesticide) to be delivered downstream is the sum of what is delivered from upstream and what is generated locally, a fraction of which is retained locally and the remainder transmitted (or *delivered*) downstream.

The actual generation of the densities for the probability tables (more precisely, the potential tables, since they do not need to be normalized) is a Monte Carlo process, over ranges of the acceptable parameter values for the model and the particular characteristics of the watershed. Most of the exploratory work of CPN modelling sediment transport in a watershed was completed by Dorner [DO01], but the methodology for development and normalization of overall watershed effects is largely untouched. All plots are not equal in their effect on receiving waters. It is necessary to formulate a correct normalization, under the control of the decision application. Impact estimation is central to conversion of a *localized decision process*, cell-by-cell, of transport modelling past a single cell to a *global decision process* for which a normalized utility measure can be developed over many cells in a network. Our modelling framework is not restricted to a single process model. In fact, several can co-exist, and even vote on an outcome. As well, independent bits of evidence can also contribute to the decision-making process. Causal

networks allow us to do all sorts of mixed inferencing. Targeting, as in attempting to achieve loadings less than or equal a maximum value for, say, sediment and mixed inferencing (including intercausal inferencing) are other useful purposes for graphical probability models.

3. A More Complex Model.

A similar problem is found in the modelling of prediction and amelioration of floods in a river course such as the Saar River in Germany. As with the nonpoint source example, there is incoming water, a retention capacity and an outlet for each reach or sub-catchment. Production would consist mainly of precipitation. A local decision to prevent flooding by passing water downstream works up to a threshold, a global (or wider-area) decision to permit local flooding may improve the larger outcome at the expense of a locally "bad" outcome. The normalization and the predicted downstream effect are not at all clear from the global perspective, even if the locally optimal solutions are at hand.

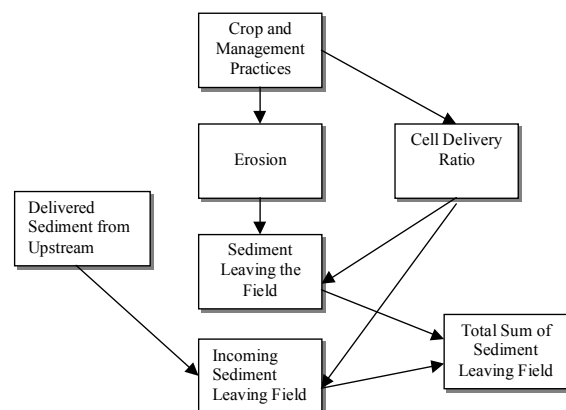


Figure 3. GAMES sediment transport as a CPN, and as an example of a compartment model. (See also Figure 4).

GAMES is a "lumped model" and does not have a significant "event driven" component. The process of developing and sustaining the mechanics of the flood prediction is definitely more complex. As a first attempt, a sequence of states with a *state transition function* can be employed. This transition from state to state is called the *situation calculus* and may prove sufficient, provided the segments have a well-defined duration in a particular state. The computation problem then revolves around maintaining sufficient past and future copies of each sector to propagate the river behaviour accurately.

Whether “situation calculus” is adequate for this type of decision process is an unanswered question. If it is not, CPN models can in fact behave dynamically. Through API calls, one subnet can alter another subnet according to a message received. We have attempted this process in a CPN simulation of hostage takings as part of exploratory work of MSc student Balachandran, in 1994, where a time-lapse changes a node representing the mental state of a hostage-taker. This is a representation of event models, and is a tool for animating analysis. A rainfall node can “play” a storm event, plot by plot to evaluate particular decisions.

4. The Nature of Utility, and Its Representation in a CPN

Utility theory is well described in many advanced textbooks such as Russell's and Norvig's [RU95]. We will not repeat in detail the derivation, but we will summarize briefly. Utility is a measure of the quality of an outcome. A decision can be thought of as an action, with one outcome if taken and a different outcome if not taken. If all circumstances surrounding the decision are known, the particular value of an outcome will occur with probability p . This *utility measure* should have certain characteristics. One of the most important is monotonicity: a better outcome means higher utility. Money is a good form of utility, if an imperfect one for extreme values.

A standard lottery L with outcomes A and B is given by $L = [p, A ; (1-p), B]$, where p is the probability of outcome A and $(1-p)$ is the probability of outcome B . If an agent's preferences obey the axioms of utility, then there exists a real valued function U that operates on states such that $U(A) > U(B)$ iff A is preferred to B , and $U(A) = U(B)$ iff the agent is indifferent between A and B . If A is the utility value of the best possible outcome, and B is the worst, setting A to 1 and B to 0, replaces L by a probability p .

The validity of the normalization transformation depends on several axioms, but does permit utility nodes being inserted into a CPN under the correct conditions. The actualization of utility in a CPN model is not, in general a trivial process, and requires many other mathematical constraints. The enumeration of the outcomes from each combination of *chance nodes* (those nodes set by modelling and other physical variables) and decision nodes generates a utility value for each decision taken or not taken. It is important to understand that the organization and separation of decision, chance and utility nodes are a prerequisite to the task of CPN model building being a doable task. The utility value depends on the values of the

chance nodes, whose values depend in turn on physical conditions and decisions taken.

5. Optimizing Outcomes

Decisions are made in the absence of perfect information. As each chance node is resolved by a new piece of evidence, the belief can be updated. Likewise, as each decision node is fixed, the updated utility is calculated. We are concerned with calculating, given the available evidence, *MEU*, the *maximum expected utility* of an action A given evidence E , where the expected utility EU is given by the following equation for *Expected Utility* (EU),

$$EU(A|E) = \sum_i P(result_i(A)|E, Do(A)) \times U(result_i(A))$$

where $result_i(A)$ represents, possible outcome states of action A , i ranges over the different outcomes. and $Do(A)$: the proposition that action A is executed in the current state.

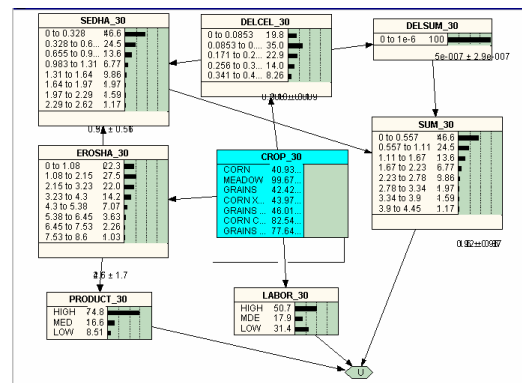


Figure 4. One spatial cell within a GAMES-based decision network.

6. Problems

Populating a CPN with numbers is a complex problem. For each state of the causal or parent nodes of a particular node, a value for the node must be given. The complexity of this calculation presents problems for the physical nodes in the physical models. Models of sufficient size to be realistic will be extremely large. Fortunately, the interconnection (Figure 2) in the geographical domain is sparse, being the connection between current node and its parents. Decisions are made that affect generation of quantity (eg. sediment, Figure 3), and the resulting sparse system updates without generating significant problems for the CPN algorithm. The main issue, after computational complexity, is developing user interfaces for CPN construction and data management. A typical display of the CPN should be via a GIS GUI (Figure 6), not a graphical representation of the CPN itself (Figure 5).

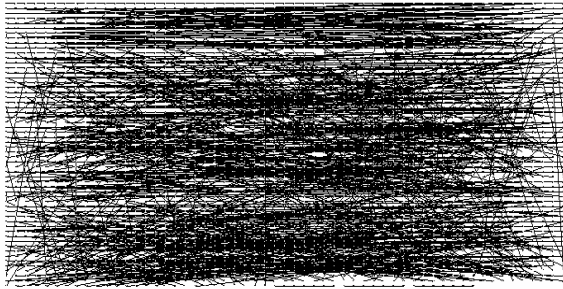


Figure 5: The Graphical Probability Model for the Stratford Avon Watershed [DO00]. There are more than 7000 nodes in this figure.

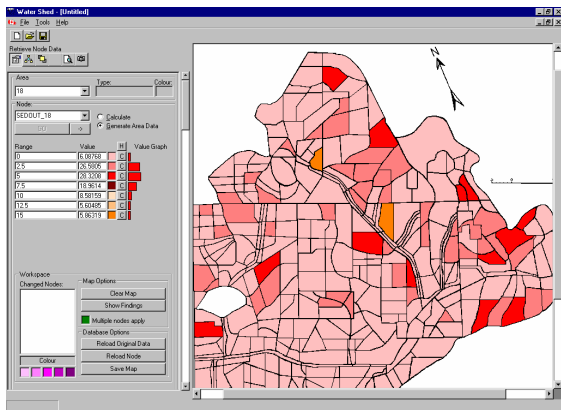


Figure 6. A slightly better user interface [DO00]. The left frame has functionality to interact with individual nodes and parameters. Several system applications are used to structure and populate the networks.

The first problem arises when continuous distributions are approximated by discrete ones. Most commercial tools, such as Netica™ or

Hugin™, work with discrete states, finite in number. There results a tradeoff between accuracy and speed. Typical parameter distributions are lognormal, and there results a long tail with little information content, but a tremendous number of excess states when the continuous distribution curve is replaced by a discrete one. There is typically a "pinching in" of results, as compared to actual model running. (Figure 7). The modal probability of the distribution of parameters on output nodes does not change, but the extremal values are optimistically presented. This problem is remedied with the aid of GUI tools for handling a large number of states in the discretization, and in the population of the network with sufficient numbers of states, at a cost in size and speed of updating numerical values.

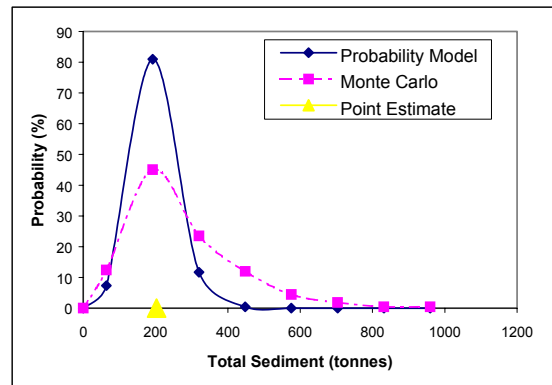


Figure 7: Distortion of output, a result of discretization of lognormal distribution of inputs.

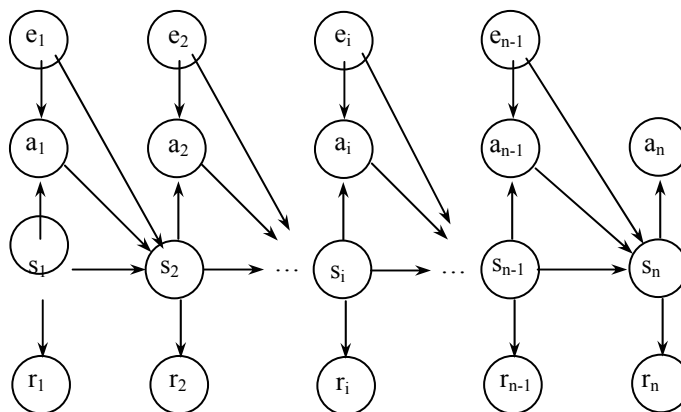


Figure 8. A Markov decision process (MDP) for a single plot of land, over time

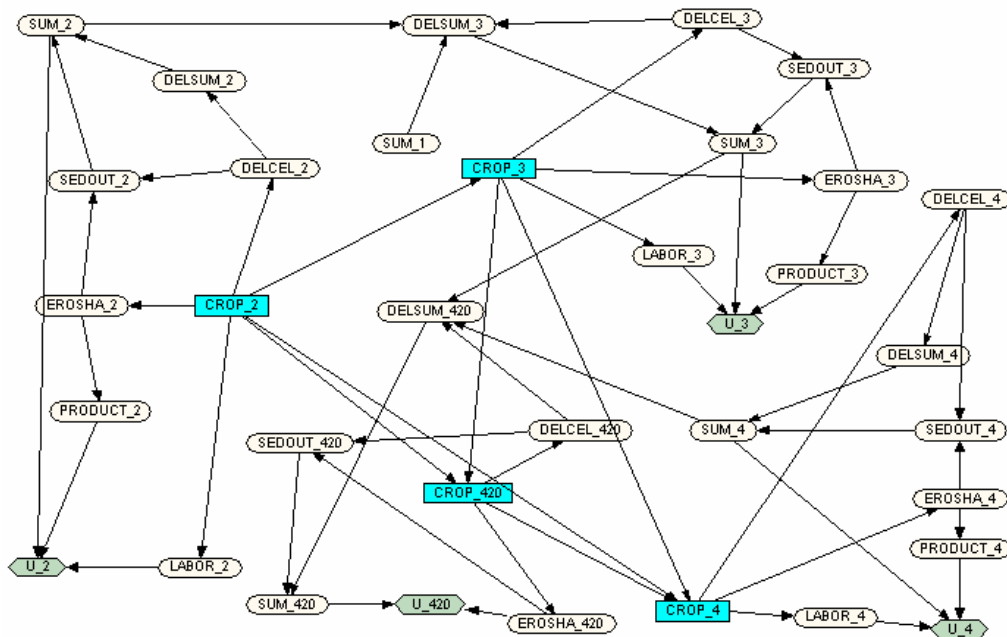


Figure 9: A sample four-cell decision network. CROP 1-4 are decision nodes and the hexagonal nodes are the utility nodes. Overall utility is not represented.

The generation of the input parameter distributions also requires considerable research. Since the CPN model is in fact a transformation of all potential model runs, it cannot generate an output for a parameter that has not been used in the Monte Carlo generation that was used to populate the CPN in the first place. Also, when a parameter is unknown at prior time, a distribution of output values is dependent upon a realistic distribution of input parameters [HA98]. In other words, generation of a typical model is a significant work, rather than a simple input exercise.

When the decisions occur seasonally, over time, the consequences of the sequence of decisions must be calculated over all possible (or at least plausible) decision sequences. The calculation proceeds, making the decisions one at a time and investigating the set of outcomes given each decision at time j , investigating the utility of all possible future decisions, making the best choice and advancing to the next state $j+1$ (Figure 8). This algorithm, including an understanding of the *discount value* of future decisions in our environmental context is a target for future investigation. Recognizing the complexity of the search space, we hypothesize that much of the speed gained by CPN modelling would be lost, if many nodes are examined simultaneously. Solutions would be for single decisions over time, with the inherent loss of flexibility. Hence we

envision the need to build such models and test approximate solutions.

6. Promise

There are presentation methods and tools to assist in consensus building in untrusted and distributed environments. We wish to investigate these. The most important contribution of a graphical model comes from its effect on the model construction phase. The most exciting part of expert system implementation occurs when the techniques are employed in politically difficult situations. Our experience with an older technology, in acid rain negotiations tells us that if parties to a dispute can work on the structure of a model, and agree on that structure, it is difficult to argue with results that are derived from well-founded data and methods. Our agricultural example has the ability to quickly become both untrusted and distributed. Watersheds have many landowners, public and private. Compensation plans for taking river and lake margins out of agricultural production would quickly lead to dispute over the amount of compensation. It is important for designers, implementers and users of DSS software to remember the "shoot the messenger" axiom in such circumstances.

Another simpler approach to design, rather than stakeholders collaborating on CPN development, is to treat the model design as being only partially

observable. Our interest in this area is a research one, and far from any practical implementation. We hypothesize that, in a confrontational situation, a protagonist can propose a strategy for either his/her cause or for the antagonist. The speed of GPM updates allows the consequences to be determined quickly and a decision to accept or defer can be given with some confidence. The effect on the decision-making of an alternative process model, say for fisheries' management, could be quickly determined.

This untrusted environment need not be antagonistic. We will eventually develop and test secure distributed model for the decision-making process, incorporating a messaging architecture to enable "stakeholders" to be physically separated, communicating over an information carrier (via World-Wide Web). Each stakeholder will see his or her sectors as the central view of the environmental problem, together with the global perspective to the outer edge of the relevant problem horizon (the limits of the watershed or basin, or the length of the river). Provision for continuance of the model in the face of "outages" or change in participants, verification of predictions, communication and recommendations for security in decision-making all fall within our interest in approximations to fully observable CPN models.

We have implemented a test version of this approach for a slight variation of the "noisy and untrusted" environment, namely for a hierarchical model of atmospheric transport of acid precipitation. A CPN representing larger scale transport is superimposed on a number of smaller scale models, representing local generation and local deposition was developed by the explorations of a Guelph MSc student, Huang to demonstrate simple replacement and overlap of physically based CPN models. The partitioning of the observed deposition into local and long-range can only work with observations. A range of possible partitionings is the best that can be done for future prediction.

7. Discussion

There will be considerable work before practical implementations can succeed. These algorithms grow very quickly in size. Maintenance is an issue, even for simple problems. On the plus side, there are economies to be made for some domains. In a watershed model, for example, the most important pieces of the network are those adjacent to the rivers, lakes and ponds. As with any model, complexity can be doled out parsimoniously to where it is most needed.

For complex (time-dependent) decisions, the search space for optimum solutions grows exponentially. Commonsense application of decisions, based on

groupings of land units based on same crop type or land use can keep investigations out of parts of the solution space that would be unreasonable.

Searches in sequential problem spaces (Figure 7) that do not involve utility are more complex implementations of *targeting*, previously described. Simple targeting is classically solved this backward problem by forward shooting. For one-shot decisions, the probabilistic shell produces the desired result to fit the target (or demonstrates its infeasibility) in a single update.

For time dependent, sequential decisions, policy iteration is one way to implement targeting over more than one time-step. It involves beginning at the end-point (in time) and working backward from the ideal outcome, until it is clear which choices are best, then stepping backward in time for an iteration. This avenue of search is unlikely to work, except when it quickly converges to the obvious. Search in utility space might be best accomplished via a genetic algorithm, since the potential for growth of possible solutions is high.

Use of utility theory holds promise for comparing "apples and oranges". Economic trade-offs can be calculated on the basis of risk analysis, with the outputs of the GPM used directly to calculate economic gains or losses resulting from changes to agricultural practice. In our current model, marginally productive land that requires a large energy input (fertilizer, pesticides) might generate a low utility value if the price of the commodity being produced were hypothesized to fall below a certain threshold value.

We are starting to use the EU equation to incorporate competing utilities. We are currently attempting to model environmental degradation together with labour and energy inputs and commodity prices. We hope to gain an understanding of how we might interpret the real cost of taking marginal land, or land along the river or lakeshores, or land where erosion is problematic. Perhaps, a simple iteration to calculate the net loss of income for input to a subsidy program is the more useful approach.

CPN examples have been built for a few test watersheds, for a few parameters. For these watersheds, utility estimation and the partner technique, policy iteration (a less constrained and simpler iteration that uses utility to develop qualitative recommendations) have been used only on small subsets of the whole watersheds that have been investigated (Figure 9). This tool is looking for a chance to prove itself in a complex situation.

8. Acknowledgement

Students Sarah Dorner, Chris Pal, Jie Shi, Chunsen Pan and Chris Newald have generated the images in this paper as part of their study. Chris developed the user interface for the watershed application. Deepa Balachandran and Tony Huang, respectively, developed versions of the software for hostage-taking simulation and analysis, and acid rain deposition prediction. Environment Canada, Natural Sciences and Engineering Research Council have provided partial financial assistance. Ralf Denzer, students and other colleagues from the Environmental Informatics Group at Saarbrücken have also provided invaluable advice and support.

9. References

- [BA99] Batchelor, C., J. Cain, 1999. Application of belief networks to water management studies. *Agric. Water Manage.* 40:51-57.
- [DO00] Dorner, Sarah, 2000, *Graphical Probability Models for Evaluating Best Management Practices for Agricultural Watersheds*, MSc Thesis, School of Engineering, University of Guelph.
- [DO01] Dorner, Sarah, David A. Swayne, R.P. Rudra, and Chris Newald. 2001. Integrating parametric uncertainty and modelling results into an advisory system for watershed management. *Advances in Environmental Research*, 5(4): 445-451.
- [GU94] Gu, Y., D.R. Peiris, J.W. Crawford, J.W. McNicol, B. Marshall, R.A. Jefferies, 1994. An application of belief networks to future crop production. *IEEE*: 305-309.
- [HA98] Haan, C.T., D.E. Storm, T. Al-Issa, S. Prabhu, G.J. Sabbagh, D.R. Edwards., 1998. Effect of parameter distributions on uncertainty analysis of hydrologic models. *Trans. ASAE*. 41(1): 65-70.
- [LA01] Lam, David and David Swayne. 2001. Issues of EIS software design: some lessons learned in the past decade. *Environmental Modelling and Software*. Elsevier Science, 16(5): 419-426.
- [LA98] Lauritzen, S.L., D.J. Spiegelhalter., 1988. Local computations with probabilities on graphical structures and their applications to expert systems. *J. R. Statist. Soc.* 50: 157-224.
- [PA98] Parson, S.C., J.M.Hamlett, P.D. Robillard, M.A. Foster, 1998. Determining the decision making risk from AGNPS simulations. *Trans. ASAE* 41(6): 1679-1689.
- [PE86] Pearl, J. 1986, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman.
- [RU86] Rudra, R.P., W.T. Dickinson, D.J. Clark and G.J. Wall. 1986. GAMES - A screening model of soil erosion and fluvial sedimentation in agricultural watersheds. *Canadian Water Resources Journal*, 11(4): 58-71.
- [RU95] Russell, Stuart and Peter Norvig, 1995, *Artificial Intelligence: A Modern Approach*. Prentice-Hall.
- [VA95] Varis, O., 1995. Belief networks for modeling and assessment of environmental change. *Environmetrics* 6: 439-444.
- [VA97] Varis, O., S. Kuikka., 1997. Joint use of multiple environmental assessment models by a Bayesian meta-model: the Baltic salmon case. *Ecol. Modeling* 102: 341-351.
- [VA98] Varis, O., 1998. A belief network approach to optimization and parameter estimation: application to resource and environmental management. *Artif. Intelligence* 101: 135-163.
- [YO86] Young R., C. Onstad, D. Bosch, W. Anderson, 1986. *Agricultural Nonpoint Source Pollution Model: A Watershed Analysis Tool*, Model Documentation, Agricultural Research Service, U.S. Department of Agriculture, Morris, MN.