CHANGES IN GLOBAL POLICY ANALYSIS PROCEDURES SUGGESTED BY NEW OPTIMIZATION METHODS AND REQUIRED FOR GLOBAL SURVIVAL

Paul J. Werbos

Dept. GVPT, U. of Maryland
8411 48 Ave., College Park,
Md. 20740, U.S.A.

ABSTRACT

"Conventional" methods of policy analysis, defined as methods which use interest rates to discount future benefits or costs, are based upon classical concepts of microeconomics and static optimization. This paper presents new methods for dynamic optimization, appropriate for complex, nonlinear, stochastic environments. (Application to natural and artificial intelligence are also mentioned.) The paper begins by drawing out the requirements for valid policy analysis in dynamic environments. Then it discusses "crossroads" phenomena, including threats to "global survival," which require drastic changes in substantive policy and policy analysis procedures (e.g. require greater emphasis on programs like Gerard O'Neill's, to expand long-range energy and population horizons). "Foresight," concern for the future without discounting, is also defended as a proposition in ethical philosophy. Finally, the optimization methods proper are summarized.

I. INTRODUCTION AND SUMMARY OF POLICY CONCLUSIONS

In the past few years, experts in many fields have argued that human civilization has finally acquired the ability to wipe itself out totally - through nuclear proliferation, or ecological disaster, or economic collapse due to shortages of energy, food or natural resources, or to the interaction of these problems, etc. Even though these arguments may or may not be correct, their very *possibility* has profound implications for national policy in all fields. Is it "rational" policy to leave it up to chance whether we survive or not? The main theme of this paper is that conventional methods of "rational policy analysis," used throughout industry and government, do in fact leave our survival up to chance. More precisely, the conventional methods depend on the validity of certain mathematical assumptions about the policy environment; threats to global survival constitute a gross violation of these assumptions, enough to insure a gross failure of public policy, but they are actually only one symptom of a more general weakness in the conventional methods. Our goal in this paper is to encourage the application and development of new methods of dynamic optimization which we have proposed, methods which do not share the weaknesses of conventional policy analysis; these methods are appropriate for situations involving many variables, nonlinear interactions, large-scale uncertainty and, above all, an inextricably *dynamic* character to the relations between variables. These methods were originally developed for application to artificial intelligence, but the problem of real-time adaptation to complex environments is central to public policy and natural intelligence as well.

Before we get deep into the details, let us try to summarize the implications of our research for the decision-maker.

Some middle-level decision-makers will ask, "What is the relevance of all this to real decision-making? What do goals and utilities have to offer me in my job?" It is hard to imagine anyone trying to make intelligent decisions without any sort of goal. *Some* criterion of "success" or "failure" must exist, or else the "decisions" being made are irrelevant. *Our concern in this paper is with the procedures used to evolve measures of "success" at all levels of decision-making.* In most decision-making, it is fairly clear that the real decisions boil down to how many dollars must be spent where; quantitative assessments of benefit and performance are clearly the ultimate focus of decision. In some areas, as in foreign policy or military decision-making, the hierarchy of strategy is more complex; dynamic optimization theory is especially valuable in those areas, but we have no space in this paper to discuss many of the details, as we have elsewhere. ([1]; see also DDI [12].) Let us emphasize that we are *not* suggesting that human intuition or freedom should be "replaced" by formal bureaucratic procedures; rather, we are trying to spell out an improvement in the formal mechanisms which exist already, so that it will be easier and more natural for people to translate valid insights from the level of intuition to the level of political reality and vice-versa. Diverse research projects and maximum creative freedom can both be justified very strongly in any decision procedure which "counts" the long-range value of possible but low-probability breakthroughs in any such effort.

In the discussion below, we will note that conventional methods of policy analysis are actually *blocking* the possibility of a comprehensive, rational attack on problems such as threats to the survival of global civilization. By "conventional methods of policy analysis," we mean any existing method which asks us to use interest rates to discount future benefits or costs. Interest rates have become an article of quasi-religious faith in the West, while abstract mathematics have not; therefore, most readers will probably need a concrete example of the failure of interest rates, before they can study our analysis seriously.

S. David Freeman, in *Energy: The New Era* [2], a study sponsored by the Ford Foundation, has sketched out the essence of such an example. Suppose that we "expect" a real level of national consumption of $1 trillion per year, starting in 1978 and continuing for the rest of eternity. Suppose that someone suggests a new program, which would change the situation as follows: (i) consumption will be \$2 trillion per year from now until 2000; (ii) after 2000, consumption will be \$0, and no one will be alive ever after. The "new program" would win out, by the criteria now used in government, so long as the real interest rate used is 4% or larger. Some economists would object that this example is atypical, because it poses alternatives so extreme that real dollars are no longer closely related to marginal utility. However, our point is that extreme alternatives *are in fact at issue* now that global survival is threatened, and also that traditional cost-benefit analysis has no clear prescription for measuring value in units other than "real dollars." Some economists would actually recommend that we *adopt* the new program in this example; in section (III), we will discuss the ethical arguments against this position. In parts of our economy far away from the decision-making center, interest rates probably have some value as an organizational device, to simplify low-cost approximate decision-making; however, an ongoing analysis will be necessary, to determine how best to make this system yield good approximations to the global interest rather than encourage short-sighted decisions throughout industry.

Many public choice theorists have likewise emphasized the impossibility of a "social utility function," as we will be implicitly recommending. Strictly speaking, we are presenting a recommended policy for those actors who accept the concept of "foresight" and the idea that the long-range domain alternatives for the human species are relatively few, as we will discuss in section (II); for such actors, the "social utility function" appropriate is simply the long-range goal of "survival" which they share.

Our conclusions may be summarized as follows:

(1) Virtually all policy decisions by the US government (or other governments) - domestic or international decisions - should be explicitly subordinated to the pursuit of global survival.

In the past, economic organization and structures have been set up to reflect the "microeconomic" ideal of a social computing machine which tries to maximize consumer satisfaction in a kind of steady state. Cost-benefit analysis is merely one application of the microeconomic ideal. But these concepts of "microeconomics" are essentially just an outgrowth and application of Lagrange's method for maximizing a static function of static variables. (For example, see Newman [3] or Samuelson [4].) Our own more generalized methods, to be discussed in sections (IV) and (V), yield the same policy recommendations as Lagrange's method does, *for static policy environments.*

However, when we admit the existence of threats to "global survival," then we admit that we are in a radically different kind of policy environment. We are in a "crossroads" situation. Even though the forks ahead of us are not so simple as "everyone lives" versus "everyone dies," we may still be in a crossroads, as we will discuss in section II. In a crossroads situation, it no longer makes sense to try to maximize the *most likely* level of economic satisfaction. Instead, national policy (which includes *all* decision)should focus on maximizing the *probability* of going down the preferred fork. Strictly speaking, the "benefits" and "costs" of *any* government program should be assessed solely in terms of the impact on the probability of "global survival." This is true even for small, "marginal" programs.

The generalized methods to be discussed below do not force us to choose between the "crossroads" model and the "steady-state" model; we may regard the crossroads model simply as a better approximation than the steady-state model. It is important to have such an approximation, so that human decision-makers can "understand" in intuitive terms what they are supposed to be trying to do. In section II, however, we will argue that the "crossroads" model is an extremely good approximation, and that those few issues which do appear to limit its formal validity still point towards the same changes in policy which the model would recommend.

(2) **Instead of using interest rates to assess future benefits, we should use "shadow prices" ("Lagrange multipliers") which are expressed explicitly as a function of both time and socioeconomic conditions.**

"Dual heuristic programming," the more efficient of the two general optimization methods to be discussed in section IV, is essentially just a method for the calculation of such "shadow prices." The use of shadow prices is already quite common in sophisticated cost-benefit analysis, but rarely if ever are those prices treated as a function of time; furthermore, they are usually calculated by reference to current market conditions, rather than long-term considerations. In principle, shadow prices must be estimated for all variables which enter into our calculations when we predict future outputs or future values of other measures of the viability of society.

When we are called on to evaluate a small-scale project, it is still legitimate to use something which looks a lot like cost-benefit analysis. More precisely, if we can predict *exactly* what the results of the project will be (which is assumed in normal cost-benefit analysis anyway), *or if* the uncertainties involved are statistically independent of major global uncertainties, then we may use the following procedure on all but the largest projects.

First, spell out the expected "costs" and "benefits" by year in terms of tangible output variables. ("Inputs," here, are simply negative outputs.) Then, for each type of output and year, find the "expected shadow price"; the expected shadow price for a variable-year is simply the shadow price for that variable in that year, *averaged* over the conditions which might occur in that year according to their probability of occurrence. To evaluate the total net benefit in a given future year, simply add up benefits times price over all the different types of benefit. To evaluate the total net benefit of the project, simply add up the net benefits in all years. No discounting is required, because the shadow prices should already account for tradeoffs between years.

This procedure is fairly simple as it stands, but there is one hard part: where do we get the "expected shadow prices"? These are not identical with expected market prices. They require a great deal of careful calculation, by people with a great deal of resources. From an institutional point of view, therefore, it is unrealistic to ask that each little group doing cost-benefit analysis should calculate its own estimates of these shadow prices; special departments are needed, dedicated to this task, at the highest levels of government and intergovernmental organization. Such departments would also be needed to help cope with decisions too large to fit into the classical format; for example, O'Neill's plan [5] for the exploitation of lunar materials and orbital space has implications which are too large to be regarded as "marginal."

In an economy in "equilibrium," these shadow prices would correspond closely to market prices. However, in a crossroads situation, the correct shadow price of a variable is simply equal to the marginal impact of that variable on the *probability of global survival.* Thus the difference between careless estimates of the shadow prices and careful estimates of them may be the difference between drifting along versus coping directly with the threats to global survival.

Even when we use the best methods available, there will be a lot of uncertainty and imperfection in our estimates of the shadow prices; however, it is better to approximate these shadow prices as best we can, albeit imperfectly, rather than accept a "simplification" so drastic that it glosses over the essence of the policy problems we face today (i.e. it ignores what we are actually trying to approximate.).

(3) **"Planning," the attempt to choose policies on the assumption that one can thereby "plan" the future in a deterministic manner, misses the requirements of global survival even more drastically than do market mechanisms.**

One might hope to avoid the complexities of "dual heuristic programming," to be discussed below, by using *deterministic* models of the policy environment. "Planning" in the classical sense requires that we can predict the future which will result from a given set of policies and conditions in the present; it does not allow for models with uncertainty built into them. Using such models, as with the classical "calculus of variations," one might avoid the need to specify shadow prices as function of "conditions," since social conditions would be a known, fixed function of time; specifying the time would implicitly specify the conditions as well, so that we could consider the shadow prices to be a function of time alone. We need only consider shadow prices along the *planned trajectory* for society.

While some simplifications will be needed in the use of heuristic dynamic programming, the assumption of determinism goes too far; it leads to two drastic mathematical problems. By considering shadow prices only on the "expected trajectory" for the present strategy of action, we lose those features of the general mathematical method which "pull" us out of local minima; or, in simpler words, when we don't make room for visualizing alternative social directions, far from the current trends, then we can easily get stuck in a "rut," in a self-reinforcing trend towards decay or mediocrity. (Formally, the information about qualitatively different possibilities lies in the boundary conditions, not on the current trajectory.) Furthermore, in a "crossroads" situation, the assumption of absolute certainty leads to severe discontinuities; in effect, it urges every decision maker to totally *disregard* the issues of global survival, right up to the point where a slight continuation of this activity would be precisely the "straw that breaks the camel's back." Clearly, the pursuit of such a hair-trigger optimum is a matter of "flirting with disaster"; real uncertainties are what determine how far one can rationally risk going in such a direction, and therefore uncertainty is inextricably part of the essence of this decision problem. Again, in a crossroads situation, the current shadow price equals the impact on the probability of global survival; estimating a *probability*, a *level* of uncertainty, is the main focus of decision, and should not be swept under the rug.

When this kind of decision method is implemented, there will inevitably be pressures towards "simplifications," like determinism, which are both unnecessary and dangerous. In particular, all approximations should be checked to make sure they do *not* weaken either the individuals or the system as a whole, by limiting them to easily predictable channels of expression or by indirectly discounting the value of "imaginative solutions." Indeed, we would not have been led to write this paper in the first place if we had not noted the tendency of *present* decision-making systems to miss the logical importance of key creative opportunities such as those discussed by Gerard K. O'Neill [5].

(4) **The development of an open set of explicit designs and contingency plans is absolutely essential both to accuracy in estimating the shadow prices and to "global survival."**

The need for explicit designs is best seen by example.

Suppose that we tried to design a television set by the use of cost-benefit analysis and zero-based budgeting. Suppose that we define "benefit" as television quality, and "cost" as dollar cost. *No matter how much* weight we place on the benefits, our analysis will fail. As per zero-based budgeting, we begin with zero spending on all components. Then we evaluate the probability of adding one transistor. Does it improve reception? By itself, not at all. So it has zero benefit, positive cost: we reject it. Likewise we reject adding a capacitor, a resistor, etc. We end up with nothing.

Designing a world energy system is just as complex as designing a television set. It certainly is plausible that the Soviet approach, based on presidium full of production engineers trying to *design* an energy system without regard to fine-tuning, will do better than the American status quo, based on hordes of lawyers and economists picking apart little pieces of an energy plan as if they represented independent movements from the status quo. Classical cost-benefit analysis, like the mathematical method of steepest ascent, leads to an optimal policy *so long as* the initial policy is "good enough." In the case of an energy system or a television, "good enough" means a *design that works*, and that works by principles the same as those with the optimal design.

Dual heuristic programming is not so limited as cost-benefit analysis. It allows fine-tuning *without* requiring such ideal "starting values." But it doesn't offer something for nothing; in order to arrive at a working energy system, or a working television, we still need some human being (or other intelligence) somewhere to select out, from the trillions of conceivable "designs," a design which works.

Low-level designs can be teated as "technological options," as in cost-benefit analysis; for example, the details of the design of a television set would not enter into our global policy analysis. Instead, we would modify our global model to indicate that certain inputs of resources will yield a certain output, now that this design exists, if the design is used. However, high-level "designs," such as a proposed global energy system, have impacts across a wide range of global variables; they are best handled as *scenarios*, as *configurations* of global variables which (hopefully) are already present in our global model.

The accuracy of dual heuristic programming depends on two key factors: (i) the *quality* of the system we use to predict the probabilities of future outcomes as a function of present policy; (ii) the quality of *scenario generation*. Even if a central agency in the government handles this analysis, there is no reason why scenario generation should be carried out in an authoritarian manner; in principle, *more* scenarios and *more diversity* in the scenarios means *greater accuracy*, so that a central agency should be assigned to *add* to its list of scenarios anything suggested by anyone else, if at all possible. Likewise, the development of probabilistic forecasting systems can involve a network of people, coordinated at a central point but not limited to one agency. For purposes of fine-tuning, it is especially important to work over scenarios in which a design almost works or almost fails; in theory, if the units of time are small enough, if the borderline scenarios receive enough attention, and if we use dual heuristic programming, then a "marginal" or "incremental" approach to policy-making is still quite adequate even for complex, "lumpy" decision problems. By the way, if different types of scenarios consistently lead to different policies, in a noticeable manner, then a hybrid policy, which accounts for the existence of two such "regions", would give us a universal optimum.

Finally, let us note that Congress (and the public) will agree to budgetary decisions based on this approach only if it has some "feeling" for the substantive rationale behind the decisions generated; likewise, the competence and morale of verbal operatives, like those in the State Department, will depend heavily on their substantive understanding of what is going on. Thus translation *from* the final framework, and *vice-versa* with questions, is very important as part of the formal system, even if such translation cannot be all-pervasive. Explicit "designs" for "alternative futures" should be very prominent as a device for communicating the substantive rationale behind policy. To strengthen these developments, without eroding democracy, we would recommend strongly that the prime responsibility for this analysis go to a new department of the Congressional Budget Office; we would recommend that the new methods be phased in "smoothly," so that changes in policy are developed *after* the reasons behind existing policies have been thoroughly probed, *except* in areas such as energy and space which require urgent change.

Furthermore, we would recommend the development of international groups to estimate the shadow prices on a global level. With classical monetary procedures, it has been possible to set up groups like the IMF, despite the conflicts of interest; here, where we are talking about common global interests, and methods which are specific to a given ideology, joint analysis could be both possible and a major step towards world cooperation. Also, from a scientific point of view, global analyses are capable of validity at a much higher level than analyses which treat only one nation as a rational actor.

II. CROSSROADS PHENOMENA IN DYNAMIC DECISION PROBLEMS

We assume that the reader is already familiar with the arguments by Raiffa [6] or others that any rational decision maker should be able to accept the concept of a cardinal utility function.

Let us begin by dealing with the problem of optimal social decision-making, assuming the existence of an (ultimate) social utility function; at the end of our discussion, we will come back to this assumption, and point out that it does not affect our conclusions. We will not actually recommend that the government construct such a function; rather, we will conclude that we should pursue "survival" as an operational goal, regardless of what the ultimate function is and regardless of whether it exists. Likewise, at the end of our discussion, we will come back to the further assumption we will make, that "long-term" stability equals infinite stability.

How can we decide what actions to take, in order to maximize the expected value of our utility function across future time? The answer to this questions depends on what we assume about the environment. If we assume that the environment obeys know deterministic laws, then we can use the classical optimization technique, called the "calculus of variations." The "calculus of variations" is based on the use of shadow-prices, "Lagrange multipliers," which can also be used as a mechanism for decentralizing decisions. This method has been generalized somewhat by Kuhn and Tucker, to allow for constraints, but in any case it is the mathematical technique behind virtually all of the economists' approaches to optimal decision-making. (Again, see Newman [3] or Samuelson [4].) These approaches, in turn, are the foundation of ordinary cost-benefit techniques.

Still, the assumption of determinism is very disturbing. It is very disturbing, for example, that the theorems of ordinary "capital theory" used to justify the use of interest rates require, as an assumption, that all investors have a perfect long-range ability to see the future. Certainly this assumption is *not applicable* to situations of uncertainty. Is it valid as a simplifying assumption, as an approximation? When uncertainty is so massive that we must worry about the probability of survival itself, "perfect knowledge and certainty" would seem to be a very poor approximation. The issues discussed in section (I) cannot be assessed analytically, except with the help of a method which allows us to ask *how big* are the implications of uncertainty. Furthermore, most of the classical work in economics, based on the calculus of variations, assumes that we *discount the future* in our ultimate assessment of utility; foresight, as we will discuss it is section (III), is ruled out because *the theorems don't work if it is allowed for.*

Many elaborate and variegated rationalizations have been conjured up, after the fact, to justify the practice of discounting. The most important motive, however, is that discounting allows certain decision procedures to "work."

From a theoretical point of view, however, cost-benefit methods need not be as limited as they may seem. There exists an optimization technique which is valid for virtually *any* decision process, even with massive amounts of uncertainty in it, *so long as* there is future discounting (or a finite time horizon) to prevent mathematical divergences. This technique is called *dynamic programming*, a familiar mainstay of industry and engineering. It is based upon the Bellman equation, which may be written:

$$J(\underset{\sim}{x}(t),t) = U(\underset{\sim}{x}(t),t) + \underset{\underset{\sim}{M}(t)}{\mathrm{Max}} \; E(J(\underset{\sim}{x}(t+1),t+1))$$

The purpose of this equation is to allow us to convert a complex, *long-range* optimization problem into a *short-range* optimization problem which we can solve in a straight forward way. In the long-term, what we want to maximize is "U", our true utility function; more exactly, we want to maximize the *sum* of U overall future periods of time. "U" means the utility function; "$\underset{\sim}{x}(t)$" means the situation at time "t"; thus "$U(\underset{\sim}{x}(t),t)$" means "the utility of the situation we are in at time t." (Notice that we put an extra "t" here to emphasize that the utility might look *smaller* to us today, even for a *given* $\underset{\sim}{x}(t)$, if it is far in the future.) U is a measure of the *intrinsic value* to us of being in situation "$\underset{\sim}{x}$"; we would call it the "intrinsic utility function."

The purpose of the Bellman equation is to allow us to compute a different function, "$J(\underset{\sim}{x}(t),t)$." $J(\underset{\sim}{x}(t),t)$ may be called the "strategic utility function." It tells us how good it is for us to be in situation "$\underset{\sim}{x}$", by accounting for more than just the intrinsic value of being in situation "x"; it also accounts for our ability to *move* from the present situation to future situations of greater utility.

In order to maximize the total utility of U, over all future time, we *need only* try to maximize J in the *immediate future*. In our equation, $\underset{\sim}{M}(t)$ means "the actions taken at time t"; at any time t, the equation asks us to pick those actions which will maximize the expected value of J, of strategic utility, in the immediate future (t+1).

At first glance, this mathematics may seem a bit obscure and alien to most practitioners. Still, any intelligent system - whether a political system, a brain, a "soul," or an ambitious artificial intelligence - must come to grips with making decisions under uncertainty in a dynamic situation. Long-range optimization *must* be translated into manageable short-term decision problems, even if approximations must be used to make this feasible. We would argue, therefore, that *any* such system must have at least four basic capabilities within it: (i) a subsystem to define the instrinsic utility, U, for any situation "$\underset{\sim}{x}$", or at least for situations as they are encountered; in other words, there must be a built-in utility function, $U(\underset{\sim}{x})$; (ii) a subsystem to create models of the external environment; this subsystem must allow one to estimate the *probability* of possible results of alternative actions, so that we can compute the "expected level of J" as demanded by the Bellman equation; (iii) a subsystem to use these models, to translate the intrinsic utility function ($U(\underset{\sim}{x})$) into a strategic utility function("J"), a system to measure the *strategic* value of being in any situation (or at least the derivatives of strategic value); (iv) an implementation system, which picks actual *actions* designed to maximize expected utility in the long-term.

Since a capability of this sort must already be built into the human brain, we should not be surprised to find ourselves already using the kind of approach, unconsciously, without even knowing that we are doing it. Let us look at a few examples. In chess, for example, our ultimate goal is to win; we may assign an intrinsic utility of +1 to a victory, -1 to a defeat, and 0 to a stalemate. *But in order to win,* we need a near-term criterion of how good our strategic position is. It is common to count nine points for a queen, five points for a rook, and so on, as a crude measure of strategic utility. In certain sophisticated circles, it is common to go on and count points for position, under certain rules. (This gets to be as complex as the Goren systems, etc., for counting poimts in bridge.) Studies of chess playing have shown that the very best players do not really look ahead very many moves, explicitly; they are distinguished, more than anything else, by their ability to assess the strategic utility of various near-term outcomes in a realistic manner.

In formulating political strategies, we try to assess strategic utility - "J" - on the basis of our sense of the long-term dynamics of history. It is those dynamics which tell us what to expect in the next period of time and the times which follow. For example, it is a common myth that American foreign policy making is oriented totally towards the short-term future. Yet top American officials (Colby, CIA) have explained why they would regard a pro-American dictatorship as more desirable than a pro-Soviet dictatorship, as the outcome of a five-year struggle beginning at the time of initial decision. They have emphasized that they regard *both* types of dictatorship as equally undesirable, in terms of their *intrinsic value*, but that the pro-American dictatorship would be more likely to change into a democracy in the *still more distant future*. The focus is on possibilities for change *long after* the end of the struggle five years in the future. In effect, they are assessing the "J" of two alternative outcomes five years in the future; they are assessing it, not on the basis of intrinsic utility, but on the basis of assumptions about the long-range dynamics of history which may or may not be justified in fact.

The essence of our proposal is that these kinds of assumptions should be based on careful, explicit study, in a context which also allows the possibility of international negotiations; they should not be based on ad hoc, unnoticed decisions by government contractors or field commanders. It is very important, for example, *how big* we rate the difference between the two outcomes; it is vital that we put more energy into the problems where the outcome is really important to us, in the long term, instead of dissipating it on traditional kinds of conflict which are easy for us to understand. In other words, we must make an *explicit* effort to estimate the true "J" function, and to apply these estimates to high-level operational decisions: this is the crux of our whole proposal.

Note, by the way, that many of these international decision making problems are being dealt with now by Bayesian decision-tree analysis. (See DDI [12].) Decision-tree analysis is a method based upon dynamic programming in a more intuitive form. (See Raiffa [6], footnote on page 23.) It might be used, for example, to construct a tree of possible decisions and contingencies, to handle US strategy in a five-year struggle such as the one mentioned above. Nevertheless, this method depends *very critically* on estimating the "utility" of the various possible *outcomes*, at the end of the tree; in this case, we must evaluate possible outcomes five years in the future. As in this example, the "utility of outcomes" does not normally refer to the *intrinsic utility;* rather, it refers to the *strategic utility,* as estimated by accounting for one's understanding of strategic considerations in the more distant future. If we plug in the wrong strategic utility estimates, then Bayesian decision analysis can act like a badly driven race-car: it will take us systematically, efficiently and at high speed over a cliff (e.g. towards a world of "pro-American dictatorships.") *The same exact danger exists with all verbal decision-making and analysis,* perhaps to a worse degree; in verbal policy analysis, the assessment of eventual outcomes may not even be carried out explicitly, and may therefore be done in an unconscious manner. In either case, with mathematical *or* verbal analysis, we recommend that the evaluation of strategic utilities be treated as a separate, vital, substantive task, based on explicit studies of long-range social dynamics.

Furthermore, when we go from ordinary dynamic programming to a consideration of "crossroads" phenomena, the classical approaches turn out to be invalid. The first step in carrying out this program is to consider Howard's [7] generalization of Bellman's equation:

$$J(\underset{\sim}{x}(t)) = U(\underset{\sim}{x}(t)) + \underset{\underset{\sim}{M}(t)}{\mathrm{Max}} \; E(J(\underset{\sim}{x}(t+1))) - \overline{U},$$

where $\overline{U}$ represents the long-term expected value of utility under the optimal strategy. Howard's formula allows us to deal with the possibility of "foresight," the possibility that we do *not* discount the future. Under these conditions, the sum of utility over all future times may turn out to be infinite; therefore, to aim at an equivalent sort of goal, we try to maximize the long-term expected value of utility. (One could add a "t", here too, to the utility function, but with "foresight" assumed, this is unnecessary; formally speaking, our version of Howard's formula is a slight simplification of the original.) The term "U" is normally just a *calibration factor,* to make sure that J does not go to infinity as our planning horizon goes to infinity. The elegant "turnpike theorems" of modern economics represent a successful attempt to deal with this kind of foresight; they demonstrate that conflicts in short-term utilities become irrelevant if one assumes foresight, under the conditions required by the theorem.

Unfortunately, even Howard's procedure is not always valid, under conditions of foresight. Consider the system shown in Figure 1. Howard's generalization of dynamic programming would be a valid procedure to apply to decisions at any of the states *within* the double lines. Each of the sets of states within double lines represents a "mixing process" or a "sink." (This notion corresponds roughly to the notion of "basin"

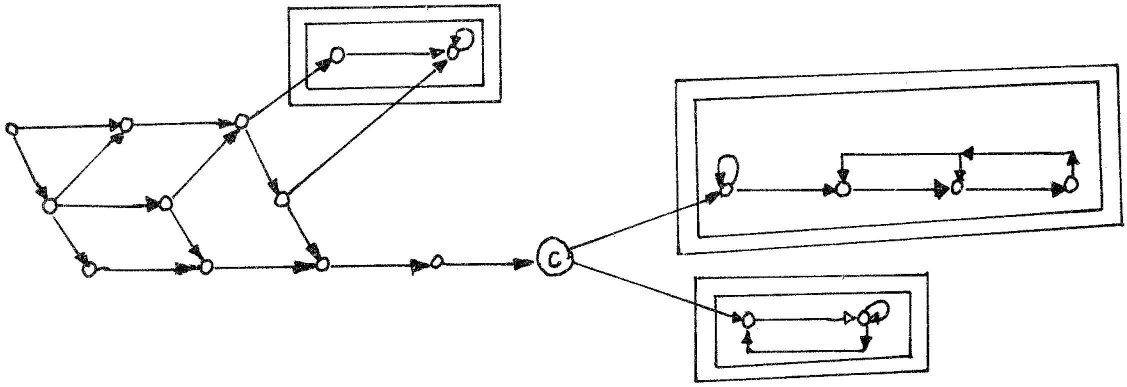

Figure 1: Example of the State Diagram of a Crossroads Situation. Circles are states. Arrows are transitions possible with any strategy, but probabilities may vary. Double lines enclose "sinks", groups of states leading to and containing distinct "mixing processes.."

in catastrophe theory, but the Markhov theory formulation, which we are using, has existed for well over a century.) In a mixing process, the equilibrium probability distribution will depend only on the strategy of action thosen, *not* on the initial state; thus "$\overline{U}$" is well-defined, for each strategy, as a constant, as a number which is not a function of the initial state. In such a situation, the effect of *any* given action, at a given time, will be finite; it can change the state at the next period of time, and the utility experienced in the next period of time, but, in the long-term, $\overline{U}$ will be the same for *any* of these states, regardless of what action is taken.

On the other hand, the states outside of the double lines do not have this property. For example, we have marked a "c" on one of these states, on the state where we decide which of two mixing processes we will spend the rest of eternity in. It is clear that $\overline{U}$ may be different in the two mixing processes; at any rate, there is no reason at all, apriori, to assume that it must be the same, in our diagram. It is clear that $\overline{U}$, the expected value of utility under the optimal strategy, may be *different in different states;* in particular, it may have one value for all the states in one mixing process, and another for all the states in another. Thus we have $\overline{U}(\underset{\sim}{x})$, a function of the state. We may still calculate $J(\underset{\sim}{x})$, by Howard's formula, under these conditions; it is still valid to maximize J *within* a mixing process, where Howard's formula is valid in its original form. However, it is clearly *not* valid to maximize J in a state such as "c". In order to maximize the long-term expected value of utility, in a state such as "c", we clearly want to maximize $\overline{U}(\underset{\sim}{x}(t+1))$; in other words, we do our best to enter into a mixing process with the highest possible $\overline{U}$. When a finite action can affect the infinite future, these infinite effects dominate the analysis. The state "c" represents the crossroads of history.

In general, it is clear that we must try to maximize $\overline{U}(\underset{\sim}{x}(t+1))$, whenever we can have some effect upon $\overline{U}(\underset{\sim}{x}(t+1)$; then, if there is no unique set of actions which maximizes this, we narrow down our choice further by picking the actions which maximize $J(\underset{\sim}{x}(t+1))$. In other words, we use a "lexicographic preference order," as if $\overline{U}$ were the "first letter in the word we are looking up in the dictionary," and J the "second letter." In practice, of course, this means that $\overline{U}$ is infinitely more important than J whenever we are near a "crossroads situation," and that J has no practical impact on our decisions.

Now let us translate this back into its practical implications. Maximizing $\overline{U}$ amounts to "maximizing the probability that we wind up in the right alternative future (alternative mixing process) in the long-term." If one of these futures involves the survival of global civilization in some form, while the others represent chaos or death, then "maximizing the probability of global survival" is the appropriate criterion for making decisions.

Actaully, however, one may question whether "death" and "survival" are the real mixing-processes/sinks in this situation. While "death" is clearly a sink, one may question whether biological survival of the species, by itself, would be enough to insure that the species would be able to reach certain desirable and conceivable possibilities. We have argued at length [1] that there are actually *three* terminal situations worth considering here: (i) death; (ii) zero growth; (iii) growth for at least centuries, based on full exploitation of the solar system and a multidue of diverse, dispersed "habitats." A prolonged state of zero growth, along with a high concentration of agricultural populations on earth, produces economic and biological pressures which are self-reinforcing, we have argued, and which generate competitive forms of social behavior ("K selection," in the language of E.O. Wilson, Chapter 4, [8].) which eventually limit all possibility of further growth or even of continued prosperity. Unfortunately, there is no space in this paper to reproduce that substantive argument, which is admittedly quite critical to the substantive value of our analysis. For now, let us merely note that we are developing the policy implications of the hypothesis that there is a small number of "mixing processes" in the human future, and that "global survival" is our shorthand terminology for "reaching the preferred sink;" these policy implications remain valid, even for those who disagree with us in their substantive analysis of what the "sinks" actually are. Note, by the way, that "M" represents the set of *all* decisions taken by society; the mathematics ask us to pick *all* the components of $\underset{\sim}{M}$, *all* decisions, in order to maximize $\overline{U}(\underset{\sim}{x}(t+1))$. In other words, *all* decisions must be based upon this single criterion. The classical economic concepts of present value may be a reasonable approximation to "J"; they are *not* an approximation to $\overline{U}$. (Though they would still be of some use, of course, in estimating $\overline{U}$.)

Now let us come back to the two fine points, which we promised to discuss at the end of this section.

First, there is the question of whether it is valid to talk about "social utility functions" as we have. Without "foresight," we admit that aggregation is not possible. Paradoxically, we find ourselves driven by the same motive as the economic discounters, but in reverse: at a higher level of mathematics, the theory works *without* discounting rather than *with* discounting. With foresight, one can go back over the mathematics above, with the understanding that "$\underset{\sim}{M}$" represents all the actions of an *individual* and U the utilities of an *individual;* since it is rational for any political actor to tailor *his* actions to maximize the probability of the "best" long-range social alternative (mixing process), it is rational for rational actors to unite on the basis of identical operational objectives, and to resolve objective disagreements by joint objective analysis.

This logic might be undermined, in theory, by the issue of *whose* descendants get to survive. After enough generations, however, natural selection should actually decide all the actual characteristics of who survives; the theoretical ancestors will be a purely academic issue. At any rate, the long-term survival of the human race is best insured by maintaining a large, diverse, adaptive gene pool, especially if things get tight, and this appears to be strongly preferred as the "tacit solution" to the game of coping with any residual conflicts. (See Schelling [9] regarding such games.)

Second, there is the difficulty that the impact of human actions may never truly be eternal. If we stay within the solar system, and depend on finite energy sources such as the sun, then sooner or later these energy sources will die and we will all die with them. If we can travel beyond this solar system, or exploit some energy source which frees us from dependence on the solar system, then we must realize that other species in the universe can do the same; our impact on universal history may ultimately be irrelevant. In a greater context, then, we are still maximizing a "J", not a "$\bar{U}$". Nevertheless, a mere billion years of history will still outweigh the next thirty, if each year is given equal importance. Our crossroads model is a local approximation, but adequate for the purpose. Within the greater universal system, survival of the fittest will still favor behavior oriented towards the *final outcome* in terms of survival; thus the arguments in the next section can be extended to the greater context, for those who are concerned about such issues.

III. FORESIGHT: THE RATIONALITY OF CARING FOR THE FUTURE

The central, driving assumption behind this paper is that we "should" care about the distant future; all of the rest of our analysis has been the result of a slow crystallization of thought about that core of axiom. There are two obvious objections to this kind of axiom: (i) it is a normative assumption, not a part of objective political science; (ii) it is the kind of issue which will be decided by politics in the end, by politics of one variety or another, not by academia. Nevertheless, the whole business of public policy is normative; it is impossible to have a "post-behavioral revolution," emphasizing public policy issues, without bringing back ethics as a subject for intelligent discourse. At any rate, it may be more appropriate for academics to spell out their ethical assumptions, and discuss them explicitly, rather than hide them as implicit assumptions. Our discussion of ethical issues may not have power over the political process, but at least it may be one input to that process; even the "objective" aspects of political science are bound by the same limitation.

Our claim is that "rationality" implies "foresight."

Ethical philosophy in the US does not provide a very strong basis for evaluating this kind of action-oriented proposition. Years ago, students of the later Wittgenstein [10] noted that a large part of philosophical disputes revolve about empty questions of semantics; therefore, they decided that the best way to approach such disputes is to clarify the common usage of the words involved, and answer the various questions accordingly. This approach has come to dominate most of the universities of Britain and the United States. In the extreme, one might take a poll of Americans, on what kinds of overt behavior they associate with the word "rationality"; this would decide whether rationality implies foresight.

The Anglo-American analytic tradition has produced one relatively strong and positive theory of ethics - the theory of Rawls. Rawls has used such concepts as parsimony, borrowed from scientific philosophy, to justify a definition of the word "justice." His definition requires equal concern for all human beings. This kind of reasoning would seem to imply equal concern for each person-year in the future, as well; after all, why should we care less for our own descendants in the future than for people we will never see on distant continents? Why should time be different from space as a delimiter of concern? (Indeed, relativity theory tells us that the distinction between the two is artificial anyway.)

Nevertheless, one can analyze the issue somewhat better, even within the Anglo-American tradition. In using a word like "rationality," we can remember that the *popular usage* of the word does not refer to overt behavior. It refers to maximum usage of the capacity of "reason," of the capacity of "intelligence" built into the human brain. In asking what is "rational," we are asking what a human brain would decide to do if it made use of the capacity of of "intelligence" to the fullest possible degree.

This analysis leads to an approach more consistent with the Continental schools of philosophy. A statement of the form "I *should* do this" is simply not related to a unique well-defined proposition (i.e. a calculable logical formula) about the state of the external universe; it can be verified neither by deduction nor by any other form of objective, scientific reasoning. Of course, people like Bertrand Russell have been saying this for years. On the other hand, there *is* objective content to a statement such as "I *would* do this, if I had enough personal experience and did enough thinking to resolve my present logical uncertainties, conflicts and lacks of conviction, without forgetting either what bothers me now or the objective uncertainties in my environment." Or in brief: "I *would* do this if I were wise." This statement is a statement about the "I", about psychology, about the nature of human intelligence. It can be dealt with by a combination of self-analysis, along the lines of the existentialists, and of scientific psychology. Only with *self-analysis*, which connects our direct perception of our feelings with our analyses of our feelings, can we learn to deal with the "objective" visual and auditory in as effective and scientific a manner as we ordinarily deal with the "objective" visual and auditory inputs. Only in that way, we would claim, is the human brain able to arrive at a real sense of conviction about what to do, a sense of conviction which extends beyond verbal statements to the will to act in a definite direction, with all the power of one's mind united. (We would agree with the analytic philosophers that confusions involving the use of words or of symbolic reasoning in general is a central problem in "philosophy," perhaps *the* central problem, but we claim that these confusions cannot be resolved on a purely symbolic, analytic level.)

In order to establish a bridge between our feelings and our theories about our feelings, it helps to look at well-constructed theories.

The question of foresight has been studied in great detail by psychoanalysts. Freud [11] has even defined the "reality principle" as the ability to defer immediate gratification. Similarly, "impulse control" - the ability to resist present pleasures for the sake of long-term goals - has received an enormous amount of attention from psychiatrists. People may *appear* to care less for the future than for the present, when they feel unable to make a definite connection between present actions and future states of being; this inability may be based on three factors: (i) a weakness of their predictive capability; (ii) an inability to visualize subjectively what certain logical possibilities would really be like; or (iii) it may be based on a fully competent assessment of limited knowledge which leads to uncertain, probabilistic conclusions. As people overcome their neuroses, it is quite clear that the first two problems begin to disappear, and the person grows more and more future-oriented. The third possibility - rational uncertainty - is something accounted for automatically in the decision-making process we have recommended. (A few pathological cases, such as many classical nation-states, do not even account for rational uncertainty; here, we have argued strongly against decision-making based on such a position.)

Many economists would be very uncomfortable with the argument above. They often find it hard to believe that the human brain does not have a built-in interest rate calculator, like a bank. It might be conceivable that there is a special organ, somewhere in the brain, which turns on every twelve months, totes up a 10% interest rate, and disappears. However, this does not fit very well with the idea of evolution. Assuming that the human mind is the product of some kind of process of natural selection, we know that the criterion for survival has no discount factor in it; what survives, in evolution, is what survives a billion-year process, not a ten-year process. Concern for the next generation certainly seems built into the psychology of healthy mammals. Uncertainties exist, of course, which limit the practical amount of attention to the future. However, the analytic faculties of the human mind seem quite capable of letting uncertainty reduce the attention paid to the future; there is no reason why uncertainties should be double-counted, by means of an interest rate applied to the primary reinforcement input to the analytic faculties.

In a growing economy, we can observe individuals with enormous amounts of money, and a willingness to put all of it into irrevocable trusts for future generations, still trading stocks and bonds on the basis of a 6% real discount factor; the interest rate reflect, not a lack of concern for future lives, but a variety of *near-term strategic opportunities* (the "J" function, not the "U" function) *for cash.* Similarly, many individuals borrow money on the assumption of future growth or for the sake of investment. Concretely, the usual criteria of compound interest and dollars (i.e. present value in dollars at 6% real interest rate) tell us that it would be "beneficial" to obliterate the planet earth, thirty years from now, if we could double the level of consumption in the meantime; *whatever* the theory of the matter, would you as an individual support such a program with your own energy? Would you consider it ethically "good"?

Certain mystical groups have opposed the idea of foresight, on the theory that spiritual development requires that we pay attention only to the problems immediately at hand. Taken to the extreme, this attitude would suggest that people quit graduate school, college or their jobs; problems on the job are "problems immediately at hand" only because one has to stick with the job for the sake of future income or other future benefits. The problem of long-range survival is a problem immediately at hand, just as much as these other problems are. There is no evidence that forces outside of man are encouraging us towards complacency in this matter; on the contrary, we have been very lucky to get as many reminders as we have that the problems of energy, food and ecology are very real. Perhaps a fearful obsession with self preservation in a negative way might inhibit the growth of many, as envisioned by people like Teilhard de Chardin [13]; however, our own emphasis on a "preferred mixing process" which is based on continued growth points towards a more positive approach.

IV. NEW METHODS FOR OPTIMIZATION

We have recommended the creation of special agencies to estimate shadow prices and evaluate large-scale programs crucial to global survival. These agencies should not limit themselves to formalized mathematical methods, but should use an eclectic combination of verbal, mathematical and intuitive techniques, in order to get full value from the human resources available. From a logical point of view, however, it is very critical that we do have optimization methods available which are fully adequate for this task.

Long-range planning, under conditions of uncertainty, can be done, in theory, by the use of dynamic programming, as discussed in section (III). However, as a *numerical technique,* dynamic programming requires us to continually update an estimate of "J" *for every possible contingency which might occur.* In a complex world, there are simply too many contingencies; for this reason, dynamic programming is not feasible for systems involving more than a handful of continuous variables. Bayesian decision analysis, a variant of dynamic programming, suffers from the same liability; contingencies multiply quickly, when one tries to be realistic, and, as the Bayesians say, "the tree becomes a bushy mess." (See Raiffa [6] and DDI [12].) This does not totally invalidate the use of Bayesian analysis in long-range planning. One can still draw up trees which represent different facets of the long-range strategic situation. One can hope that human intelligence will allow us to string together a set of such trees into a coherent analysis; after all, this is not too different from the common belief that we can arrive at a coherent analysis by stringing together a series of English sentences.

Nevertheless, it would be reassuring if we could generalize dynamic programming somehow, so that it can be applied to complex multivariate problems. This we have done. Formally speaking, our talk about social "machinery" to maximize something in the long-term makes sense only now, when we can point to a mathematical system which allows such maximization in the general case.

The chief difficulty in dynamic programming is the representation of the "J" function. If we assume, apriori, that "J" can be *any function,* we must treat its value for each set of values for the arguments as a unique quantity. We allow an infinte number of degrees of freedom, if the arguments are continuous. This is reminiscent of one-way analysis of variance in statistics, where we have separate, unrelated predictions of our dependent variables for *each possible state* of the independent variables(s). This sort of approach has turned out to be very inappropriate, in predicting complex political systems; instead, statisticians and econometricians prefer to formulate *explicit algebraic models,* models which contain a moderate number of parameters in them but which yield predictions for what might happen over a very broad range of contingencies. This requires human effort, in formulating, comparing and upgrading alternative models; still, it is much better than treating each contingency as unique. It is possible, in principle, to allow computers to scan a wide variety of algebraic forms, as possible models; this may be worse than human quality control, but it is probably still a lot better than treating each contingency as unique.

We can follow a similar procedure in dealing with dynamic programming. We can construct a "model" of J. More precisely, we can set up computer programs, designed to input a model of J as an algebraic expression with certain parameters in it identified as requiring estimation. As in nonlinear regression, we can allow the user to put in his or her own initial guesses, as a matter of choice. Also, we can print out the final "degree of fit," for the final model, as in regression. (Here, the "degree of fit" is measured as expected value of U across future time which would result from trying to maximize the user's version of "J".) As with regression analysis, our final model is only an *approximation*; however, since models from different human users can be compared and synthesized, the final strategic model should be better than the best ideas of the humans, *for coping with the policy problem as it has been described.* (i.e. for a given utility function and a given model of the objective probabilities.) Here, the "degrees of freedom" in the model affect the number of calculations required, not the quantity of data required, for accurate estimation.

In Howard's version of dynamic programming, we generate J by successive approximations; in each step, we reset $J(\underset{\sim}{x}(t))$ to equal $U(\underset{\sim}{x}(t))+\text{Max}E(J(\underset{\sim}{x}(t+1)))-\overline{U}$, where the latter value of J is determined by the old estimate of the function J. In each step, we also pick a new set of actions, to maximize $J(\underset{\sim}{x}(t+1))$. This approach can be adapted easily to our purpose. We can *fit* the parameters of $J(\underset{\sim}{x}(t))$, as in statistics, to be as close as possible to our previous estimate of $U(\underset{\sim}{x}(t))+\text{Max}E(J(\underset{\sim}{x}(t+1)))-\overline{U}$. In theory, we could work to find the optimal set of actions in each step for each $\underset{\sim}{x}(t)$ tried out or simulated; in practice, we could derive the actions from an "action model," whose parameters may also be estimated as part of this process. With the latter strategy, we can afford to use simple simulation to give us the equivalent of a carefully-computed expectation value. (Note that $\overline{U}$ is not too critical here; in each iteration, we can store scale factors for J, both additive and multiplicative, to make sure that nothing diverges. This will handle crossroads problems.)

In theory, this system can look ahead only one extra unit of time per iteration. However, if we estimate these parameters by computing the gradient, in each iteration, and plug it into a fast quasiNewtonian convergence method, convergence will be much faster in practice, yet still practical with many parameters in one problem. To compute this gradient inexpensively, with a complex network model, we recommend the use of the "dynamic feedback" algorithm, discussed in our previous work. (Chapter II of [14].)

The method above we have called "heuristic dynamic programming." [15]

Let us suggest another method, which is more efficient for complex situations. It is similar in some respects to differential dynamic programming, developed by Jacobson and Mayne [16]. Instead of estimating $J(\underset{\sim}{x})$, for a raw input vector x, we estimate the functions i(x)," for each component, x_i, of x. "$\lambda_i(\underset{\sim}{x})$" represents the *derivative* of J with respect to x_i. In effect, it represents the "shadow price" of x_i. Since there are many x_i, this would mean many functions to estimate, and it *could* mean many parameters. However, since these λ_i are actually interdependent, we could formulate "network" models in which different λ_i share many parameters and terms.

In each time cycle, our method proceeds as follows. From a given situation $\underset{\sim}{x}(t)$, we carry out a simulation of $\underset{\sim}{x}(t+1)$ by first simulating the set of random numbers, $\underset{\sim}{w}(t)$, required by our stochastic model of reality; then, for those values of $\underset{\sim}{w}(t)$, we get a sample value of the gradient of likelihood with respect to the parameters of $\lambda_i(t)$ by trying to fit each function $\lambda_i(t)$ to match our estimate of:

$$\sum_j \frac{\partial^+ x_j(t+1)}{\partial x_i(t)} \lambda_j(t+1) \text{(plus 1 if "}x_i\text{" refers to "U")}$$

This computation can be done inexpensively, with complex network models, by the dynamic feedback method mentioned above. Note that we added a plus sign in the derivative, to indicate that we wish to measure influence *forwards* in time, as formalized by our concept of "ordered derivative," the mathematical basis of our dynamic feedback method. (See section (II-xii) of [14].) Once again, we can update or optimize action strategies, and use a quasiNewtonian convergence method to estimate all the parameters. Note that the formula just above comes simply from differentiating Howard's formula.

The method described above is the simple form of what we have called "dual heuristic programming," or simply "DHP." [15] We would expect that DHP is considerably more efficient than heuristic dynamic programming. Heuristic dynamic programming exploits our knowledge that $J(\underset{\sim}{x}(t+1))$ has a particular value, but DHP, in effect, *explains* the goodness or badness of J by noting how much any *changes* in $x_i(t+1)$ would have changed J at time t+1; it evaluates the components of $\underset{\sim}{x}(t)$, not blindly, but by accounting for their *known* causal effects on x(t+1), as explained by our model of the world; in effect, it *exploits* more knowledge, and thereby arrives at more complete feedback. The difference is comparable to the difference between punishing all students in a class when there is a disturbance, as opposed to punishing only those who are known to be responsible. (In the example, admittedly, there are complicating interactions, but with DHP we are looking at derivatives, at local linear relations.)

Unfortunately, DHP is very vulnerable to crossroads effects. In biological organisms, however, natural selection would point towards an intrinsic utility function, U, set to "B(t+1) - B(t)", where "B" is a measure of biological success; thus $\overline{U}$ has always been zero, and we have no crossroads to fear. A similar procedure would work in political planning; "B" could be measured in many way - as active population, as real GNP or even as the probability of "survival" - without changing the recommended policy, if we postulate only two real sinks in the long-term. (With three sinks, however, we would have to rate their comparative utility.)

In the mammalism brain, we would expect that simple DHP is used in the cerebellum, where a *quick* evaluation of overt action is essential. In the forebrain, we would expect that *globalized* DHP (to be discussed below) is implemented by the following steps: (i) at time t+1, compute $J(\underset{\sim}{x}(t+1))$ by using a model of strategic utility, as in heuristic dynamic programming; (ii) use dynamic feedback to compute the $\lambda_j(t+1)$ as the *derivatives* of $J(\underset{\sim}{x}(t+1))$; (iii) feed back these estimated $\lambda_j(t+1)$, by our procedure above, to compute the "true" $\underset{\sim}{\lambda}_i(t)$; (iv) after comparing true and actual $\lambda_i(t)$, correct the J network by computing the derivatives of error with respect to the parameters of the J network; this requires "second order dynamic feedback," using dynamic feedback to calculate the derivatives of a weighted sum of quantities which were themselves computed by dynamic feedback; however the procedures in Werbos [14], Chapter II, are quite explicit on how to do this. In the mammalian brain, we would claim that the first-order dynamic feedback calculations are performed by the well-known "retrograde" chemical transmission, flowing back from cell to cell along small tubes inside of the brain cells; the second-order calculations must be performed by a similar "double retrograde" network of flows, which contains forwards and backwards flows both. On a computer, GDHP is much easier to implement, as we will see below. GDHP has two advantages over simple DHP: (i) it insures logical consistency between the shadow price formulas for different variables; (ii) it gives us a measure of global strategic utility, which is useful in evaluating very large programs or in intuitive understanding of policy.

The practical use of these methods requires the prior availability of stochastic predictive models of the global environment. Such models could come either from statistics, or from "judgemental models" on Bayesian lines. However, the development of good judgemental *models* will require a development of Bayesian techniques to a higher level than has been considered in the past. This will require careful studies of the effectiveness of variants of these techniques, in cases where the measures of performance have a high variance.

V. OPERATIONAL ASPECTS OF DHP

Dual heuristic programming was developed initially for use in computer simulation of natural intelligence. Strictly speaking, it is a method for estimating two types of policy parameter *together* - "value parameters" (used to compute "shadow prices" or "values", or, in GDHP, strategic utility) and action parameters (appearing in the "model strategy). Indeed, it allows us to put a given parameter on both lists, so that we may allow the parameters and equations for these two functions to overlap each other; we need not worry about restricting a given parameter to one category or the other.

We do *not* mean to suggest that national decisions should all be carried out on a computer; however, it is easier to understand the goals and concepts of the method by seeing how it *would* be implemented on a conventional computer system. It took many decades to translate Lagrange's method for optimization into coherent recommendations for social organization; a complete adaptation of DHP will not be less complicated.

The goals of DHP, in simple form, are Shown in Figures 2 and 3. As in conventional optimization, we start with a utility function, which we want to maximize in the long term, and a completely specified model of *objective* reality (of how the external world works). We have described elsewhere advanced statistical procedures which are appropriate for estimating the "model_proper", the objective model. (See [17], [18], [19], and the section on "syncretism" in [15].) Our goal here is to find the strategy of action which maximizes the utility function in the long-term, for a model_proper which has already been fully estimated. Instead of asking the computer to consider *all possible* strategies, the human user must give the computer the functional form of both the

```
available_data: productive_capacity, investment, population,malnutrition;
policy_parameters: k1, k2, k3, k4, k5;
policy_variables: consumption;
model_proper:
        productive_capacity(t+1) = (1+g1)*productive_capacity(t)
                                   + g2*investment(t);
investment(t+1) = productive_capacity(t+1) - consumption(t+1);
population(t+1) = g3*population(t);
malnutrition(t+1) = g4*((consumption(t+1)/population(t+1))**g5);
intrinsic_utility_function: log(consumption/population)-malnutrition;
evaluation_model:
        value(productive_capacity) = k1*productive_capacity
                                   + k2*log(productive_capacity);
model_strategy:
        consumption = k3*productive_capacity + k4*log(productive_capacity)
                    + k5*malnutrition*productive_capacity;
```

Figure 2: Sample "policy model" As It Might Appear As a Computer File With Simple DHP. The policy problem is to decide on the level of consumption to be chosen in each year.

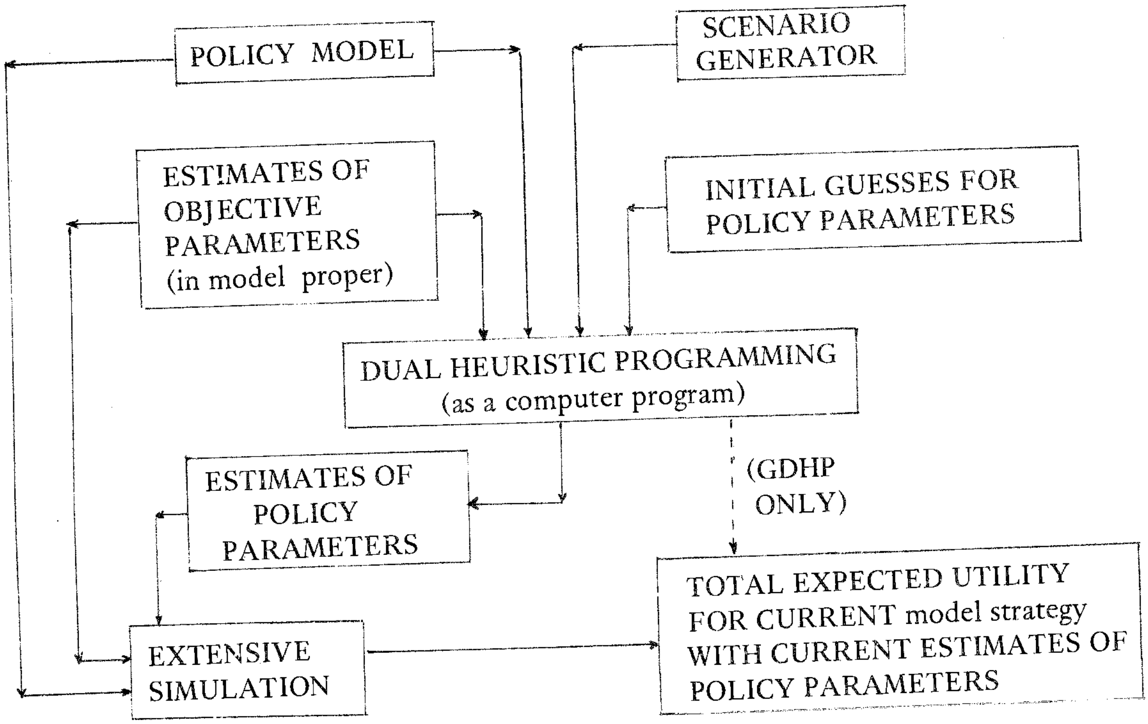

Figure 3. Inputs and Outputs of Dual Heuristic Programming.

strategy and the *shadow prices;* we ask the computer to pick out the best possible combination which *fits* the form supplied by the user. Thus the goal of the computer is to estimate the "policy parameters," the parameters which appear in the "evaluation_model" and the "model_strategy."

Note that our "evaluation_model" refers only to the "shadow price" or "marginal value" of a given variable. Conventional computer convergence routines require a calculation *both* of the function to be maximized (e.g. strategic utility) *and* of its derivatives with respect to the parameters to be estimated. However, for some quasiNewtonian methods, the derivatives alone are enough. We would hypothesize that the human brain employs such a method; with a complex nonlinear network, we doubt the value of performing global "point checks" which evaluate *all* the parameter updates of a given cycle as if these updates were closely related to each other. Likewise, in DHP and GDHP we anticipate the use of such a convergence method. Note that the success of the human brain in adapting billions of such parameters in parallel strongly implies the existence of adequate "sparse" quasiNewtonian methods for megaparameter problems; for our empirical work in model estimation, we have successfully tested a few possibilities in that direction [20].

To make global comparisons of one strategy against another, with simple DHP, one can choose between only two approaches: (i) synthesize the two, and see if some of the policy parameters are close enough to zero that they may be safely deleted; (ii) carry out a simulation test of the two strategies, based on a much wider set of scenarios, etc., than DHP itself requires. With "globalized dual heuristic programming," however, we can ask the user to submit *one equation,* giving a model of strategic utility, instead of a multiequation evaluation model. The *computer* can apply dynamic feedback to differentiate this equation, and generate a set of evaluation equations which are adjusted *exactly as if the user had specified them separately;* the only difference is that the consistency of the equations is guaranteed, and we can then use the resulting J to *evaluate our present strategic position with the current strategy.* This allows some simple, direct comparison of strategies, although the safe use of this comparison remains to be spelled out.

The workings of DHP may be described, very roughly, as a 5-step process, which we will outline below. In this, the "scenario generator" is very critical, as we noted in section (I). In theory, the scenario generator is simply the basis for a kind of Monte Carlo integration. As with normal Monte Carlo integration, we know that any "adequate" generator (i.e. diverse enough scenarios) leads to convergence to the true J function, if possible, but there is no formal algorithm to give us the "best" generator. As with the parameters of quasiNewtonian convergence methods, we simply have to experiment with a variety of alternatives which are equally good apriori. Theoretical analyses are possible, but will have to be relatively loose, as per operations research, and rooted heavily in experimentation. By the way, note that DHP was originally designed for a real-time, parallel processing system and can therefore be implemented more efficiently in such a form.

In DHP, the core subroutine generates "derivatives of strategic utility with respect to policy parameters," which then get plugged into quasiNewtonian convergence routines as described above. This subroutine goes through 5 steps, which we will outline crudely:

(1) GENERATE SCENARIO: $\underset{\sim}{x}(t)$.

(2) USE MODEL STRATEGY TO CALCULATE POLICY VARIABLES: $\underset{\sim}{m}(t)$.

(3) LOCATE n PRESELECTED SETS OF RANDOM NUMBERS $\underset{\sim}{w}(t)$.

(4) DO FOR EACH SET OF RANDOM NUMBERS, SET GOES FROM 1 TO n:

(4a) SIMULATE $\underset{\sim}{x}(t+1)$, BY INSERTING INTO THE model_proper THE SCENARIO $\underset{\sim}{x}(t)$ FROM STEP (1), THE $\underset{\sim}{m}(t)$ FROM STEP (2), AND THE CURRENT SET OF RANDOM NUMBERS. THE LATTER GO INTO THE IMPLICIT TERM "+Noise(t)" ADDED TO EACH model_proper EQUATION, MULTIPLIED BY THE NOISE STANDARD DEVIATION, WHICH SHOULD BE STORED SOMEWHERE AS A PARAMETER OF THE model_proper.

(4b) CALCULATE $\underset{\sim}{\lambda}(t+1)$, THE ARRAY OF "VALUES", BY INSERTING $\underset{\sim}{x}(t+1)$ FROM STEP (4a) INTO THE evaluation_model.

(4c) CALCULATE $\underset{\sim}{\lambda}(t)$ BY THE FORMULA:

$$\lambda_i(t) = \Big(\sum_j \frac{\partial^+ x_j(t+1)}{\partial x_i(t)} \lambda_j(t+1)\Big) + \frac{\partial U(\underset{\sim}{x}(t))}{\partial x_i(t)} \quad ,$$

WHERE "U" IS THE INTRINSIC UTILITY FUNCTION, AND THE DERIVATIVE OF "$x_j(t+1)$" IS THE DERIVATIVE OF THE OUTPUT OF THE model_proper, CONTROLLING FOR THE INPUTS IN (4a).

(4d) CALCULATE $\underset{\sim}{\lambda}(t)$ DIRECTLY FROM $\underset{\sim}{x}(t)$, AS IN (4b).

(4e) DEFINE THE EVALUATION ERROR, $e\lambda_i$, AS $\lambda_i(t)$ FROM STEP (4c) MINUS $\lambda_i(t)$ FROM STEP (4c), MULTIPLIED BY A WEIGHTING FACTOR. IN PARALLEL WITH ROBUST ESTIMATION [19], WE RECOMMEND A WEIGHTING FACTOR EQUAL TO THE STANDARD DEVIATION OF x_i DIVIDED BY THAT OF λ_i.

(4f) CALCULATED $e1(k_j)$, FOR EACH POLICY VARIABLE, k_j, AS THE SUM OVER "i" OF $e\lambda_i$ MULTIPLIED BY THE DERIVATIVE OF λ_i (AS CALCULATED IN STEP (4c)) WITH RESPECT TO k_j. OPERATIONALLY, INSTEAD OF DOING THIS AS A SUM OVER "i", WE MAY USE DYNAMIC FEEDBACK TO CALCULATE IT IN ONE PASS. FOR POLICY VARIABLES WHICH APPEAR IN THE model_strategy ONLY, THIS WILL BE ZERO, UNLESS THE model_strategy AND evaluation_model ARE MIXED TOGETHER.

(4g) CALCULATE $e2(k_j)$, FOR EACH POLICY VARIABLE k_j AS:

$$e2(k_j) = \sum_i \lambda_i(t+1) \sum_p \frac{\partial^+ x_i(t+1)}{\partial m_p(t)} \frac{\partial m_p(t)}{\partial k_j} ,$$

WHERE $\lambda_i(t+1)$ IS AS CALCULATED IN STEP (4b), WHERE THE LEFTMOST DERIVATIVE REFERS TO "m_p" AS AN INPUT IN STEP (4a), AND WHERE THE RIGHTMOST DERIVATIVE REFERS TO "k_j" AS AN INPUT IN STEP (2).

(5) DEFINE THE CURRENT "DERIVATIVE OF UTILITY" WITH RESPECT TO EACH POLICY PARAMETER k_j AS THE AVERAGE OVER ALL SIMULATIONS AND SCENARIOS OF $e2(k_j)$-$e1(k_j)$.

REFERENCES

1. Werbos, P.J. "Strategic Planning for Global Survival", Report 4 of the Advanced Forecasting Group of the Maryland Crisis Warning and Management Project, Feb. 1977, 83 p. Please address requests to Dr. Stephen Andriole, Director, Cybernetics Technology Office, Defense Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, Va. 22209.
2. Freeman, S.D. *Energy: The New Era.* Vintage Books, New York, 1974.
3. Newman, P., ed. *Readings in Mathematical Economics, Volume II: Capital and Growth.* John Hopkins, Baltimore, 1968.
4. Samuelson, P.A. The *Foundations of Economic Analysis.* Atheneum, New York, 1965.
5. O'Neill, G.K. *The High Frontier.* Bantam, New York, 1978.
6. Raiffa, H. *Decision Analysis: Introductory Lectures on Making Choices Under Uncertainty.* Addison-Wesley, Reading Mass., 1968.
7. Howard, R. *Dynamic Programming and Markhov Processes.* MIT Press, Cambridge Mass., 1960.
8. Wilson, E.O. *Sociobiology: The New Synthesis.* Belknap Press of Harvard U. Press, Cambridge Mass., 1975.
9. Schelling, T.C. *The Strategy of Conflict.* Oxford Press, New York, 1960.
10. Wittgenstein, L. *Philosophical Investigations.* Third Edition. MacMillan, New York, 1973.
11. Freud, S. *Beyond the Pleasure Principle.* Hogarth, London, 1955.
12. Barclay, S., Brown, R.V., Kelly, C.W., Peterson, C.R., Phillips, C.R., and Selvidge, J. *Handbook for Decision Analysis.* Technical Report TR-77-6-30, Sept. 1977, 276 p. Available from Decisions and Designs, Inc., Suite 600, 8400 Westpark Drive, MacLean, Va. 22101.
13. de Chardin, T. *The Activation of Human Energy.*
14. Werbos, P.J. *Beyond Regression: New Tool for Prediction and Analysis in the Behavioral Sciences.* Ph.D. thesis under the Committee on Applied Mathematics, Harvard U., November 1974. Microfiche copies available from Archives Dept., Widener Library, Harvard U., Cambridge, Mass. 02138. Text copy available for perusal in Gordon-McKay Library, Harvard U. Copies in text may be requested from the author.
15. Werbos, P.J. "Advanced Forecasting Methods for Global Crisis Warning and Models of Intelligence." General Systems Yearbook 22 (1977), 25-38.
16. Jaconson, D.H., and Mayne, D.Q. *Differential Dynamic Programming.* Elsevier, New York, 1970.
17. Werbos, P.J., and Titus, J. "An Empirical Test of New Forecasting Methods Derived From a Theory of Intelligence: The Prediction of Conflict in Latin America." IEEE Trans. on Systems, Man and Cybernetics (September 1978).
18. Werbos, P.J. and Titus, J. "The First Full Test of Werbos's Robust Method: Predicting Conflict in Latin America", Report 10 of the Advanced Forecasting Group, Jan. 1978, 150p. Available from same source as reference [1].
19. Werbos, P.J. "Bias in Favor of the Future and Bias in Favor of Critical Variables . . . ", submitted to IEEE Transactions on Systems Man and Cybernetics, July 1978, 44p. Until publication, available from the author.
20. Werbos, P.J. "New Convergence Methods for Multisectoral Optimization and Estimation", Report 6 of the Advanced Forecasting Group, April 1977, 25 p. Available on same basis as reference [1].