Before getting into the substance of this talk, I will begin by defining exactly what I mean by an intelligent system. Next, I will talk briefly about the applications of this concept to SAS, to human information needs, and the like. The first major section of this talk will describe a basic 3-component intelligent system which could be implemented in SAS. The final section will describe the kinds of information needed in more sophisticated systems, like the human brain or human organizations; it will begin with an intuitive overview aimed at a non-technical audience, and conclude with some of the mathematical and physiological concepts which underly the intuition (with citations).

WHAT IS AN INTELLIGENT SYSTEM?

I define an "intelligent system" as a generalized system able to maximize some measure of success over the long-term in an uncertain, non-linear environment which it must learn to adapt to in an open-minded way. This definition could apply, in principle, to systems made up of equations in a software package or to networks of microchips, nerve cells, ectoplasm, or anything else. The key words in this definition are:

- Generalized -- The system must allow for a wide range of environments, non-linear or stochastic. A real-world system must give greater weight to simpler theories of its environment, but it should not be limited to particular functional forms or specialized structures.

- Adaptive -- The system must be able to learn what it needs to know from direct experience (time-series data).

- Success -- The ultimate measure of success (or "utility") cannot come from the intelligence proper, by definition. It may come from the system designer or user (as with decision tree analysis, which is widely used in industry) or it may come from biological evolution.

Intelligent systems are different from expert systems because:

- Expert systems rely almost entirely on knowledge borrowed from intelligent systems such as humans or SAS runs.

- Expert systems impose domain-dependent structures upon knowledge (e.g., systems for chemists built on atoms/valences/molecules).

- Expert systems use domain-dependent rules (often with some limited intelligence) to respond to very specialized requests.

However, there are also some similarities:

- Initial conditions (prior knowledge) can be inserted into any intelligent system.

- Intelligent systems can learn to communicate, and might even be designed with a capability to exchange knowledge directly.

WHAT APPLICATIONS DOES THE IDEA HAVE?

The theory of intelligent systems can shed light on five important questions:

- To the extent that humans are intelligent systems, what are the basic, unvarying components of human thought and brain dynamics?

- How can groups of humans such as companies or governments become more intelligent, more effective in achieving their own goals?

- If you sell information or knowledge to intelligent systems such as humans or enterprises, what are you selling and what are they buying? What kinds of information do intelligent systems really use?

- How can we build computers or computer-based control systems to be intelligent systems?

- What additional capabilities would SAS need in order to become a more complete support system for intelligent decision-making?

These are all large questions, drawing on different professional disciplines which use different vocabularies. This paper will only provide a quick overview; for more details, see the references.

Note that SAS is a generalized system, unlike most of the more standard systems in artificial intelligence. In this respect, it is more like a true intelligent system than they are. Just as the human baby eventually overtakes the monkey baby, despite the monkey's early advantage from overspecialization, generalized software systems may well overtake more conventional forms of artificial intelligence, if the necessary R&D is carried out.
A BASIC 3-COMPONENT DESIGN FOR INTELLIGENT SYSTEMS

Optimization Over Time: The Starting Point for the Design

Optimization over time is central to the definition of intelligent systems, and central to the difficulty of building them.

Many optimization methods exist for simple, specialized or constrained systems. However, there is only one family of methods which gives the right answer over time in a generalized, uncertain, non-linear environment: dynamic programming.

In dynamic programming, the user supplies a utility function, \( U \), which represents his long-term, basic objectives. In chess, for example, \( U \) might be +1 for winning, -1 for losing, and 0 for anything else. The user also supplies a dynamic model, \( f \), which describes reality at time \( t+1 \) as a function of reality at time \( t \) and actions \( (u(t)) \) at time \( t \). Dynamic programming then translates this long-term optimization problem into a short-term optimization problem, which is more tractable. It calculates the secondary or strategic utility function, \( J \), which has the following property: any operational system which picks its actions, \( u(t) \), so as to maximize \( J \) in the immediate future will automatically do a perfect job of maximizing \( U \) in the long-term future. In chess, for example, the old rule of counting 9 points for a queen, 5 for a castle, etc., is nothing but an attempt to approximate \( J \) for that game; by fighting for points in the short term, players hope to win the game in the long-term. Because this rule gives an imperfect approximation to \( J \), it sometimes makes sense to sacrifice points in the short-term in order to achieve other long-term strategic goals; however, a better approximation to \( J \) should account for these other strategic goals as well. Decision-tree analysis is one form of dynamic programming.

Conventional dynamic programming cannot be applied to problems with many variables, because the calculations are impossibly expensive. Therefore, intelligent systems must be based on generalized methods to approximate the results of dynamic programming. This can be done by having those systems develop a model of strategic utility \( (J) \), based on the same kinds of methods SAS already provides for developing dynamic models \( (f) \): parameter estimation and error minimization. One possible method is called GDHP, Global Dual Heuristic Programming.

Three Core Components of Any Intelligent System \((J, f, u)\):

---

### Examples of Strategic Utility \((J)\)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Basic Utility ((U))</th>
<th>Strategic Utility ((J))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chess</td>
<td>Win/Lose</td>
<td>Queen = 9 points, etc.</td>
</tr>
<tr>
<td>Business</td>
<td>Current Profit</td>
<td>Present Value of Strategic Assets</td>
</tr>
<tr>
<td>Theory</td>
<td>Cash Flow</td>
<td>Performance Measures</td>
</tr>
<tr>
<td>Human</td>
<td>Pleasure/ Pain</td>
<td>Hope/Fear</td>
</tr>
<tr>
<td>Thought</td>
<td>Hunger</td>
<td>Reaction to Job Loss</td>
</tr>
<tr>
<td>Behavioral</td>
<td>Primary</td>
<td>Secondary</td>
</tr>
<tr>
<td>Psychology</td>
<td>Reinforcement</td>
<td>Reinforcement</td>
</tr>
<tr>
<td>Artificial</td>
<td>Artificial Utility Function</td>
<td>Static Position Evaluator (Simon)</td>
</tr>
<tr>
<td>Intelligence</td>
<td>Artificial Intelligence Evaluation Function (Hayes-Roth)</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>Rational Values</td>
<td>Cost/Benefit</td>
</tr>
<tr>
<td>Finance</td>
<td>Long-Term Goals</td>
<td>Measures</td>
</tr>
<tr>
<td>Physics</td>
<td>Lagrangian</td>
<td>Action Function</td>
</tr>
<tr>
<td>Economics</td>
<td>Current Value of Product to You</td>
<td>Market Price or Shadow Price (&quot;Lagrange multipliers&quot;)</td>
</tr>
</tbody>
</table>

Actually, the items in the last row of this table refer to the derivatives of the two kinds of utility. They refer to the change in overall utility which would result from a small change in your level of consumption of a particular good. Likewise, the emotional charge associated with a particular object is really just the derivative of \( J \) with respect to the corresponding variable. (All of these examples could be discussed at greater length; for example, I would contend that human "pleasure" includes parental pleasure in experiencing happy children, but issues like that go beyond the theory of intelligence).
Derivatives are vital to the operation of any intelligent system. For example, when SAS estimates a non-linear model, it starts from the initial guesses which you provide for all the parameters. It adjusts these guesses upwards or downwards, so as to minimize forecasting error. In order to do this efficiently, it has to know whether a small positive adjustment of any parameter would increase error, decrease error, or leave the error unchanged; in other words it has to know the derivative of error with respect to every parameter.

Example of a Minimal Intelligent System

In SAS

SAS already allows users to submit model files, which can then be used in simulation. In theory, SAS could support the use of four-part "intelligence files" or "dynamic decision files" which might look like:

utility function: \( U = (4.1 \times \text{fish} - 10000 \times \text{boattrips})(1.04^{\text{year}-1985}) \);

strategic assessment: \( j = a1*(\text{fish}^{0.5}) \);

model: \( \text{catch}(t+1) = a2*\text{fish}(t) + \text{catch}(t) \);

operations: \( \text{boattrips}(t) = a3*\text{fish}(t) + a4*\text{boattrips}(t) \);

In this example, the user is trying to maximize long-term profit from fishing, assuming that fish will sell for $4.10 each and that boat trips cost $10,000 each; future profits are discounted by 4 percent per year. The strategic assessment here simply reflects the concern that overfishing could cut into future profits. A more realistic example would require more equations and variables in all four parts of the file.

SAS already could let you estimate \( a3 \) and \( a4 \), if you have historical data on the variables in the model. SAS can even estimate the "error" with each equation, which you need in simulating the model. It will give you statistics which show how good the model is. If SAS had a capability to do GDHP, it could go on to pick \( a1, a2, a5, a6, \) and \( a7 \) so as to maximize your expected profit in future years. The operations equation would then tell you how to vary the level of boat trips in response to a changing supply of fish as well as provide you with performance measures, to let you compare this assessment model and operations equation against other alternatives.

As a practical matter, SAS would have to give you options (like linear regression) to estimate each part of your model as cheaply as possible. To handle large, multiequation systems, it would have to improve convergence methods and use low-cost methods to calculate the derivatives required with such methods. Many users would be more interested in the derivatives (for sensitivity analysis or shadow prices) than in the operations equation itself; in other words, the lack of valid long-term shadow prices is a fundamental problem in government finance, central to the concerns of both futurists and environmentalists among others. SAS could also allow for more advanced methods, user options, and a generalization of stepwise regression. Some users have advocated a capability to handle linked systems of models, as in the system used by Decision Focus, Inc. Ideally, a version of SAS could even be developed for the new "supercomputers", massively parallel computers analogous to the human brain. All of these things turn out to be feasible, but they require major R&D support which does not now exist.

Calculations Required to Implement a Basic Intelligent System

My previous diagram of an intelligent system simply assumed, on faith, that you know what non-linear regression is. This diagram describes the required calculations in more detail. Note the two circles labeled "FIT"; these represent the two places where the predictions of a system component are compared against actuality. The lower "FIT" circle simply calculates the sum of squared error, across time and variables (properly weighted). The upper "FIT" circle refers to the similar loss function used in GDHP (although much of the error calculation is actually done inside the strategic assessment system). The FITT circles can also be interpreted as reality testing or quality control nodes.

The solid black lines refer to flows of information for the variables in the system. The other lines refer to additional flows of information required to calculate all the derivatives needed in estimating the system. This diagram assumes that you will use "dynamic feedback", a method which allows you to calculate all the required derivatives, even for a very large model, at a cost comparable to that of exercising the model itself. The dotted lines represent the derivatives of \( j \) with respect to all the variables in the system, including the
parameters which govern action; these derivatives are used in estimating the parameters of the tactical operations system. The dashed lines represent the derivatives of FIT with respect to the variables in the strategic assessment system \( J \) and the dynamic model \( f \). Methods do exist for estimating thousands of parameters, so as to minimize error or maximize success, when the derivatives are known.

Dynamic feedback and its cousins have been used a number of times at the Department of Energy, mainly for sensitivity analysis with large-scale simultaneous-equation models.

The arrangement in this diagram has an interesting parallel to Freud's theory of ego development, in which "psychic energy" (the dotted line) drives psychological growth. The solid lines correspond to gross, electro-chemical flows in the brain, while the other lines would have to correspond to chemical flows backwards along the same network. Extensive flows of feedback also help to explain the success of Japanese corporations (though they also lead to a slower decision process).

ADVANCED INTELLIGENT SYSTEMS

Humans and other mammals have capabilities far beyond the minimal system which I just talked about. If you are selling information to human beings, you are selling a more complex product than that discussion suggests.

Before exploring these complexities on a mathematical or scientific level, I would like to motivate them by describing, in common-sense terms, what kinds of information exist in a more advanced system. This discussion will be highly intuitive, but it will be followed by a technical explanation of the underlying mathematics.

Most of you have heard people talk about the difference between "data" and "information". Managers or users of information systems are usually very concerned about the wide gap between raw data, on the one hand, and the kind of information their clients really need, on the other. They often develop taxonomies (like "data" versus "information" versus "solutions") to express these concerns. High-level government policy makers tend to be very serious about concepts like "vision", even though lower-level people may treat such words like a disease. The discussion which follows provides a similar taxonomy based upon the more detailed mathematics of advanced intelligent systems. It represents the underlying categories of information which, I would claim, are hard-wired into the brain.
the far future. Wisdom is the ability to place the proper value on those aspects of the world which the brain is already aware of. (It is a function of the "limbic system" in the brain, although it draws on other parts of the brain for support). Market economics tries to imitate Wisdom by putting the proper prices (values) on physical goods.

Vision is the ability to perceive those aspects of the world which are crucial to the validity of Wisdom. It is the ability to perceive variables or patterns which organize one's values, hopes and fears into an integrated and clear picture, so that Wisdom becomes less like groping in the dark and more like carpentry. Vision typically grows from the interaction with Wisdom, Insight, and creative imagination.

Insight. Insight is an accurate, unified sense of how the world works, including the world of people, feelings, and symbolic reasoning. Insight may be expressed in words or mathematics, or it may remain subjective. In all cases, it is a sense of how things can change, with what probability. It is closely related to what Freud calls "ego", although psychologists often use that term in other ways. It is a function of the major cells in the cerebral cortex, including both the "left side" and "right side" of the brain. Formal mathematical models used to make forecasts or scenarios are a mechanical imitation of (or stimulus to) Insight.

Memory. Memory is a way station between insight and awareness. The brain records the past in many ways, and all of these are sometimes called "memory". True memory is a record of past events which directly changes our expectations of the future, and perturbs the results of our Insight. For example, a child who has been attacked by a man in a blue shirt may cringe when he sees a different person wearing a blue shirt, even if his knowledge and ego tell him that this new person is trustworthy. Statesmen may cringe from anything which reminds them of the Vietnam. True memory can operate with or without the ability to consciously recall past events.

Freud and others have identified true memory as the source of many neuroses and mental blockes; however, memory provides an essential raw material for the creation of Insight, just as pet bogs and swamps are an essential raw material for nature's production of coal. All mammals possess true memory. More primitive animals, like reptiles or fish, may possess a faculty of memory but not of Memory; as a result, their insight never becomes very deep. Some institutions seem to use reptilian thinking in making decisions. As Memory is absorbed into Insight, it loses its ability to perturb the ego in irrational ways; this process of absorption is crucial to growth of the mind. (True memory is a function of medium-sized cells in the cerebral cortex, but it only happens when the events to be remembered seem to be important as defined by the faculty of Wisdom).

Imputed Reality. Imputed Reality refers to one's picture of the world. Knowledge of the present. The human cerebral cortex builds up an image of the world, based on the combination of insight, experience, and awareness. Imputed Reality includes those beliefs about the state of the world which we arrive at after considering the limitations (and logical implications) of the data which enters our conscious awareness. Imputed Reality includes an awareness of those patterns which tends to hold together and form a basis for Insight (and memory). Our Imputed Reality may waver at times, when we are uncertain of what to believe, but the wavering can be limited if we include uncertainty itself as part of our reality.

Awareness. Awareness refers to the regular flow of information from the world (internal and external) into full conscious attention. This information is true only in the sense that it truly reaches consciousness; it is up to the conscious mind (cerebral cortex) to decide what relation it has to truth or falsehood in the real world. Nevertheless, it is an important job for the thalamus and the lower centers of the brain to sift through all the raw data which enter the nervous system and send up the most important information, in a regular and condensed form so that the conscious mind will find the information useful after having enough experience with trying to relate it to other such information. Awareness provides regular information, like the screen of a television set or the weekly polls on Presidential popularity, unlike Imputed Reality, which focuses on the richness and diversity of patterns which come and go.

Crude Inputs. Crude Inputs include all the information that reaches the system, misleading or not, which may or may not ever reach awareness. Many psychological problems in human beings relate to important information which never reaches the awareness level for one reason or another. Some of the data may affect the actions of the system in an important way, through peripheral action sites which are not part of the conscious mind.

Action. There are major parts of the nervous system which are not oriented towards knowledge or policy, but towards carrying through on short-term aspects of the values which trickle down from the conscious mind. One of these, the cerebellum, has a limited degree of foresight in extrapolating the priorities which it receives from above, but operates in a real-time mode which is too fast for the higher centers to keep up with in detail. These lower systems also control the muscles of the eye and thereby help to
Mathematics of Advanced Dynamic Models

The previous discussion was simply an intuitive translation of concepts taken from mathematics and neurophysiology. The possibilities for upgrading the dynamic model component do mirror the previous discussion fairly closely.

Insight. The forecasting equations proper can be estimated by methods more robust than regression. In technical terms, this requires robustness against the independence assumption; this is known to be an important issue, but has received inadequate attention because the research requires a difficult combination of sophisticated epistemology and pragmatic experiments. Regression and other maximum likelihood methods have severe problems in handling lagged endogenous variables and near-singular multivariate error distributions; methods have been found which reduce these problems, but more research is needed. In intelligent systems, it may be enough simply to weight the error in forecasting each variable in the "reality" vector by some function of its marginal strategic utility or the like. In using SAS to develop forecasting equations for the Department of Energy, these problems are serious, but I have minimized them by avoiding lagged endogenous variables, by avoiding multiequation estimation methods, and by using functions of time in various estimation experiments.

Memory. Practical engineers often make forecasts by direct use of historical scatter-plots, without any use of regression or other tools which require a choice of functional form. Likewise, some government agencies have made use of "precedent-based forecasting", in which new situations are projected to evolve just like the closest historical analogy in the database. "Nearest-neighbor" methods in artificial intelligence involve the same concept. In mammals, this capability appears to be integrated into the forecasting system. (In real-time systems, like the brain, this faculty may also stabilize long-term aspects of imputed reality).

Imputed Reality. Dynamic programming assumes a Markov Chain model, a model which projects all of reality at time t+1 as a function of reality and actions at time t. To achieve this, in principle, one needs an (approximate) non-linear filtering capability, integrated into the modelling process. This could require a generalization of Kalman filtering, which is now widely used, to the non-linear case. Alternatively, it could require a generalization of methods by Box and Jenkins for addressing "errors in data". In psychological terms, this is the primary source of "object permanence" and "short-term memory". Likewise, one needs to express reality as a vector whose components have minimal cross-correlations in error, in order to get valid simulations. (This is similar in spirit to factor analysis, but may provide a new approach to "feature generation" in pattern recognition).

Mathematics of Advanced Operations Systems

The tactical operations system can be upgraded by adding a high-speed optimization component. This is less relevant to SAS, and more relevant to real-time systems, which need to act quickly on information before it can go through the entire cycle of strategic assessment. GDHP, by itself, can provide the value weights ("Wisdom") needed for optimal decisions. For annual models in a system like SAS, it may be unnecessary to provide a separate faculty of "Vision". In fact, it may even be good enough to simulate the future a hundred years out, and use simpler methods by Jacobson and Wayne to do the optimization. (Actually, with hundred year time horizons, GDHP may have much greater numerical efficiency). However, some systems, like the human brain or factory controllers, need to use dynamic models which forecast a tenth of a second into the future rather than a year; it is not feasible to simulate such models out a hundred years, in normal simulations. For systems of this sort, there can be serious numerical problems in trying to achieve a long period of effective foresight while minimizing the chance of instability. Even the human brain may not represent a perfect solution to this dilemma; it may simply represent a working tradeoff.

At any rate, there is a need for further research into convergence methods for use with GDHP. There may also be a need for a separate faculty of "Vision", designed to develop independent variables which...
straighten out the long-term relations and help GDHP converge better. For example, one could install a subsystem which tries to predict the derivatives of J as a function of current reality; the strategic assessment proper could take current reality and the predictions from this subsystem as its inputs. This might allow a more rapid, explicit translation of its inputs. This might allow a more rapid, explicit translation from primary goals to the secondary goals required to implement them. This possibility is fully consistent with the complexities of the human limbic system.

Comparison with the Human Brain

All of the lines in this wiring diagram represent direct electrochemical pathways. I have used Grossman's book, Essentials of Physiology Psychology, as my basic check, but have used other sources to verify other aspects of this theory, far beyond the scope of this talk. For readers with little background in neurophysiology, the works of Paul MacLean introduce the important facts.

The dotted lines represent basic features beyond those of a minimal intelligent system. Two upgrades appear to explain these features (in addition to those discussed above):

- Restriction of body movement during sleep (simulations of distant possibilities to enhance GDHP numerical efficiency).
- Ability of each component to input calculations from other components as information (as a source of independent variables), and even to reinforce these inputs when useful.

Dynamic feedback, a backwards chemical flow from the limbic system to the cortex, helps explain why the limbic system has strong functional impact on the rest of the brain despite the scarcity of direct, electrical lines from the limbic system to the rest. Neurologists have shown that large-scale backwards chemical flows can occur, but more research is needed in living organisms.

FOOTNOTE AND REFERENCES

*The views in this paper are the author's, not those of the Energy Department or any of its components.

SAS is the registered trademark of the SAS Institute, Cary, North Carolina, USA.

More information may be obtained from the author, at 202-252-1197, or at Forrestal E1-621, Washington, D.C. 20585; some of the information presented here is based on the following papers by the author:


A Statistical Analysis of What Drives Industrial Energy Demand, DOE/EIA-0420/3, Chapter 4. (Discusses robust estimation in a practical context). Available free from National Energy Information Center at 202-252-8800, or check with me.
