Proposal for Development and Implementation of A Global Information and Forecasting Service (GIFTS)

Paul R. Williamson, Ph.D.
Global Vision, Inc., P.O. Box 4394
Ann Arbor, MI 48106-4394
telephone: 734-769-4877
e-mail: paulrw@globechange.org

Myron S. Karasik, MS, MBA
Member, Institute of Electrical and Electronic Engineers
Principal
ColdFrame Inc., 530 Lawrence Expressway, Unit 354
Sunnyvale, CA 94085-4014
telephone: 408-705-6870
e-mail: coldframe@eudoramail.com

February 21, 1994
Updated November 3, 1994 and May 31, 1995; author information updated October 17, 2003.

***Key Terms*: Global- , Computational- , Dynamic- , and Data- Modeling; Social- , Environmental- , and Behavioral- Prediction / Integration; Data Transformation; Data Mining; Data Visualization; Fuzzy Systems.**

ABSTRACT

We propose to form a non-profit research venture to utilize and gradually improve the state of the art of modeling-forecasting technology, and to develop an automated network- service / information-clearinghouse capable of reliably predicting consequences of large scale human activities and interactions with the environment. The exemplar is the weather forecasting service used to predict potential climatological catastrophes. While initially emphasizing the problem of potential destructive conflicts between human populations, this service / clearinghouse will gradually be expanded to cover the full range of interrelated global environmental and social behaviors and changes. A network of data collection ('monitoring') stations, and the central analysis and dissemination of results, will in turn be funded by subscription open to all groups and persons, public and private. These subscribers will receive regular ongoing forecasts of probable political, social, economic, and environmental trends in areas of interest. They can also obtain (for an extra fee) special assessment reports analyzing probable responses to proposed actions of the subscriber by all communities in an area of interest. Subscribers are expected to include regional, national and international public agencies, corporations, non-profit /public interest organizations, and investment groups, as well as individuals. The goals of the service include contributing to the avoidance or mitigation of destructive conflicts or other unforeseen consequences, and the losses they entail, through proper assessment of likely responses to given choices or strategy. The modeling technology is expected to be refined over time and act as a catalyst to better joint decision-making by all parties having influence in a particular domain. Pricing for services will be based on operation costs and R & D efforts as well as the value provided to the subscriber.

THE NEED

Central to the rationale for this Service is the unsatisfied need, on the part of analysts and practitioners, for improved information and forecasting. This need arises because to adequately cope with present-day ecological and social changes will require a much more realistic, than heretofore, picture of the present state of the world and of alternative scenarios concerning the potential future, depending on what actions are taken now.

Everyone who makes choices based on information is, at least implicitly, doing forecasting. However, the quality of management and public policy choices has been greatly impaired by serious limitations to the degree of descriptive and predictive realism currently attainable. For one thing, the information used by analysts and policy makers in their forecasting often has been too specialized or too disjointed to be used effectively for obtaining a realistic picture. Such a picture requires the recognition we are presented with an interconnected complex of problems and pertinent background phenomena. These are directly or indirectly coupled one to another, giving rise to significant mutual interactive effects. Moreover, the standard statistical and other forecasting methods have often given insufficient attention to certain possibilities and fundamental relationships concerning the prediction of social and other global phenomena.

In policy terms, these shortcomings have meant the failure to recognize that poverty, disease, ethnic conflict, pollution, etc., are not solved 'simply' by one or more of promulgation of statutes, expenditure of funds, military force, economic boycotts, or bureaucratic assistance. Sudden shifts can result in substantial losses of investments made by various entities. Further, accidental infringement on cultural taboos can result in eruption of savage conflicts and loss of markets.

We can see how critical these deficiencies are, at this particular moment, by taking a brief look at our world. The recent collapse of the Soviet Union and the momentary or hopefully long term improvement of superpower relationships presents the human community with new opportunities; but, especially if these opportunities are not seized, the accelerated spread of nuclear and other high-technology weapons, the continuing or even increased potential for violent conflict in many areas of the world, economic and environmental stresses, and many other circumstances present great dangers as well. Beside weapons proliferation, we see the advent and worldwide spread of other technological innovations, benevolent and malign.

For example, there is the plethora of high speed communication and transportation links that now connect us; these have made the concepts of 'neighborhood' and 'independence' obsolete and made us interdependent neighbors by default if not by design. At the same time, other more traditional factors remain in effect. Important among them: we continue to belong to multiple communities of interest-- be they geographical, kin, ethnic, occupational, religious, philosophical, political, social, technical, or what not --that influence us in competing ways.

Relationships among these various factors provide important examples of factors being coupled one to another, to which we earlier referred. For example, the technologically induced 'shrinkage' of the world has amplified the impact of multiple interest communities, by bringing the various human cultural traditions into progressively more intimate contact. The latter, in concert with other changes, has exacerbated the potential for conflict because the now shared and radically altered environment cannot uniformly meet everyone's values and behavioral expectations. The ability of the human race to find a creative alternative to this state of affairs is necessary if humans are to prosper or, possibly, even to avoid destroying our planetary home.

The above sustain our key point: To meet these global circumstances and to devise the necessary creative alternative require vast improvements over the present content, form, and availability of information; and they require forecasting that takes account of the relevant possibilities and fundamentals of prediction.

GOALS OF THE SERVICE

The proposed Service is based on a recognition of these requirements. Its ultimate goal would be to aid the efforts of researchers, analysts, planners and policy makers throughout the world, in every pertinent domain, public and private, in acquiring the realistic world picture that we humans must have, in order to cope with ecological and social change. It would pursue this goal by

* developing a comprehensive management and distribution system for the pertinent information

* promoting forecasting based on the recognition, definition and modeling of the couplings, taking account of the other predictive considerations alluded to above.

To elaborate, the proposed Service would foster a more scientific methodology for forecasting. Recognizing that adequate empirical knowledge is indispensable to this process, the Service would combine or network already accumulated information bases and add to them, to form a more complete and usable information base. This information or knowledge base would represent the distillation of facts, relationships and mathematically valid computational techniques. As described below, these informational and forecasting elements would include a new approach to developing and refining knowledge and information concerning the cultural 'lenses' through which we humans filter our experiences and define our potential behaviors.

In carrying out these elements, heavy reliance would be placed on developing and using automated (computerized) modeling tools. While the very complexity, of the global system, that must be addressed has been a barrier to implementing such tools in the past, the Service would make use of recent developments that allow this barrier to be circumvented or lowered. Similarly, the Service would take advantage of computer developments that now greatly improve the ability to develop and access information of requisite quality.

In addition to the above goals, we have two other normative considerations in mind. To reinforce democratic practices and institutions, we would promote the Service, not as a central manager or prime mover of decision making, but instead, as an aid to non-centrally directed decision making by a great many public and private users. There will be no exclusions as to who may subscribe or contribute information. Further, to advance the science of forecasting, we would make the Service available and readily accessible, at appropriately determined moderate fees, to all research parties to promote extension, refinement and elaboration of the technology.

The second consideration is that the information and forecasting system should be "multi-cultural." This means that the system should provide a culturally neutral method by which the values and other cultural 'facts' of each human community can be described and made into predictively usable information. Of course the Service is informed by ethical norms, referred or alluded to throughout this proposal, concerning our desire for a viable and just future. Ultimately, we seek to create a credible simulation and forecasting tool that transcends individual cultures, to aid in promoting a peaceful equilibrium between differing cultural-normative systems and in harmonizing human actions within the full ecological context of the world that we inhabit. And the methods we intend to follow are based on the premises of Western science.

Beyond those normative and methodological commitments, however, the Service would seek to avoid making judgments about the moral correctness of values or behaviors. Rather, it would seek to describe the cultural facts in forms that are computationally tractable. Such descriptions would then be combined with such 'hard' facts as geologic, climatic, health and economic conditions. These data would then be used to projectively model various important trends and probabilities-- e.g. to produce the equivalent of 'behavioral weather maps' showing trends and predicting possible 'stormy' conditions --as well as the affects of possible interventions and strategies. Again, the end purpose would be to determine which of them might best prevent, contain, ameliorate, accentuate, or otherwise appropriately respond to impending social and environmental changes.

The rest of the proposal describes what is necessary to achieve an operational service.

TECHNICAL REQUIREMENTS

The GENIUS system would begin as a global electronic network supporting multimedia (voice, alphanumeric, image, video) information types with researchers throughout the world, each having access to a workstation capable of accepting their data and analyses. This network of database and analytical activities would be supported by a set of interrelated electronic processors and software tools.

The key problem is to define a workable, operational model that captures the complexity of the interrelationships in a tractable form. No doubt, many models are possible, and any given one capable of improvement. A long term function of the Service will be to encourage a variety of alternative approaches, by many investigators. The following elements therefore merely define how we will approach the problem of the first such model.

While analytic and creative activity is essential to the aim of better forecasting, an adequate data base is a parallel and equally essential aspect. Often, the modeling task will involve an interplay between choosing and developing data, and selecting and testing dynamic (predictive) concepts. The choice of concept determines what kind of information shall be regarded as data and the predictive efforts tell us, by their success or failure, whether the information accumulation efforts are on the right track; conversely, the availability of data constrain what concepts can be tested and used, and relationships that the data reveal constrain what is plausible to be tested. In sum, data and model development are so closely related that it would be most unwise to consider either one alone. Thus it is hardly surprising that in the following, we will find it necessary to move back and forth among various connected ideas about data and prediction.

Basic modeling approach

The first consideration is that the model will embody a "distributed network" principle, meaning:

1) the world is represented as a system of interactions (flows or movements of information, energy, commodities, materials, persons, and other living things) among distinct "actors" (regions, nations, provinces, localities, or groups); and

2) each such actor is represented as a separate set of computational sub-routines corresponding to a distinct sub-model within the overall model.

In addition to representing social entities, the model will represent the environment in which they function. Many environmental aspects correspond to more "physical" aspects of the modeling problem, in which variables are more operationally defined and relationships (the couplings) among them are understood deterministically. For instance, clearly climate and geological forces and conditions are mutually affected. Each has different time frames of processes, one has to do with atmospheric and cosmological sources (sunspots) which can shift radically in days, the other has long-term forces which define the mineralogical wealth and soil characteristics. Both drive biological forces by defining ecological possibilities. Climate and geology interact with economics and technology when it comes to conversion of physical resources to goods and waste products. Each effects the other; consider acid rain, water pollution, over-farming/desertification, etc. as examples of such cross-linking.

The representation of these comparatively more deterministic factors will provide part of the environmental background (context) within which the societal components of the model interact. In addition, for many social and other processes not understood deterministically, nevertheless there are 'hard' statistics that can provide probability or population distribution functions for per capita wealth, education, health, etc. These comparatively tractable deterministic and probabilistic factors will provide the context within which the modeling activity will address, in ways we discuss below, the 'softer', more complex aspects of human and other global system behavior.

Together, these elements will be used to drive a 'Monte Carlo' simulation engine (a well-proven, automated means generating many scenarios to determine likelihood of outcomes) to effectively explore the range of possible states and their likelihood. Thus the overall model will integrate probabilistic and deterministic aspects to explore the range of possible states within, and connections among, elements. One of the uses of these capabilities will be to simulate the interactions of the actors under the perturbations of various events.

Representing memory in the system - the problem of history

The former naive belief, in the ability of simple cybernetic feedback systems and statistical forecasting techniques to adequately capture, by themselves, the full range of interactions of all elements subject to policy decisions on the large scale, has been lost. We have stressed that the interconnections among various problems and factors is an important complication to adequate forecasting and planning. Moreover every action creates in its wake a largely unforeseen range of results largely due to the fact that every 'problem' represents the culmination of a long history of events. This historical aspect is an important part of the complexity problem.

One aspect of this complexity is that many factors (concerning, for instance, changes in social structures, definitive crisis-response situations, and many other salient economic-political events) vary over time in an apparently irregular manner which seems to escape any simple predictive formula. One way of dealing with this irregularity relies on the approach of adaptive learning by examining sufficiently many historical changes, to discern patterns. (For instance, present-day world political changes may be found to possess characteristics in common with previous changes.) This possibility calls for data that are historically deep enough to include many instances of the changes in question. However, in many cases this possibility may be foiled by another aspect of historical complexity to which we turn next.

In reality, the nominally independent 'trajectories' or histories commingle and interact, changing each other in various ways. In many cases, to treat the symptoms dealt with by the customary policies does not adequately address this interactive reality, thus it misses the true forces that are in play. Proper perspectives require attending to the couplings or "cross-fertilization" effects as having this longitudinal (over time) character, rather than simply existing of a given moment.

Our way of meeting this requirement is that the sub-models of our distributed network will represent the actors as "intelligent," meaning that they possess, in some degree, a memory of past events and are capable of acting based on the contents of their memory plus information on the current state of the system. Through the latter feature, we will seek to emulate present attitudes and behavioral traits (including dysfunctional traits such as ethnic prejudices in Bosnia and resistance to change in General Motors or IBM) as the product of evolution from past conditions and experiences.

Only in very recent years has the mathematical modeling community been able to develop a mechanism to describe such complex systems. Similarly, only recently have computational technologies advanced to support large, heterogeneous databases and visual modeling techniques capable of depicting alternative complex scenarios within a manageable period of the investigator's time. Therefore, this aspect of the Service is only now a very real possibility.

Data requirements

We have chosen to organize the huge variety of human and environmental information necessary to support the above modeling approach, into nine major types. These concern:

* Geological forces and conditions

*Climatological, oceanic, and cosmological conditions

*Biological factors - non human species (physical and behavioral traits)

*Biological - human factors (physical and behavioral traits, including health, epidemiological, reproductive, and metabolic)

*Social structures that define the actors of the model (cultural, social, and ethnic groups; and provincial, national, international, and transnational political and economic units)

*Culture (language and value systems)

*Human artifacts, technologies, and wealth creation potential (including capital and capital infrastructure)

*Economic-political organization and behavior (including distribution of wealth and other resources)

*Interactions (material, energy, and information flows among relevant structures and elements)

Each of these nine categories span a variety of information which, to be adequately tracked, requires a multiplicity of indicators. In meeting these requirements, our priorities will emphasize reproducible information and comparability over broad spatial and temporal domains. To the extent possible, we will downplay reliance on anecdotal and impressionistic information and emphasize the development of scientific data.

Corresponding to the key role of history discussed above, there is an essential historical dimension to many of these types of information. To meet the corresponding data requirements, there are written and archaeological evidence from which to infer human history, the geologic record for climatic and mineralogical data, and genetic and paleontologic information for species data-- all of which we will utilize.

Recalling our earlier distinction between comparatively 'hard' versus 'soft' aspects of modeling, one can anticipate that this emphasis on scientific data will meet with greatly varying success, depending on the type of information. As we noted in the opening section, one of least tractable kinds of information concerns human culture, to which we turn next.

Representing the effects of language

The problem of finding a computationally tractable way to represent human culture is important to the tasks of describing human values and perceptions (i.e. the cultural 'lenses' we mentioned); these, in turn, may be essential to understanding the learning process central to the fact of history. They also relate to the problem of achieving normatively neutral cultural description. Of course, in posing these matters we seek to deal with a deeper 'meaning' of culture than simply its appearance as gross physical artifacts. The salient characteristics of a cultural define normative behavioral patterns for its members based on social structures, historically catalytic events and traditional teachings filtered through authoritative figures. Information of this type is describable primarily in linguistic terms and can be measured effectively by statistical measures (i.e., percentage of population reporting certain indicators), Periodic opinion surveys (assays) will then allow us measure changes in these salient characteristics. The representation problem is developing a consistent 'linguistic yardstick'. Once this is done, a new class of hard data, 'cultural' in nature, is created, to be modeled along with other data. The initially chosen means of linguistic quantification is the method of 'semantic networks'. This method characterizes words into linguistic categories-- for instance, a lion is a carnivore, which in turn is a mammal, also which is a life-form living in a Savannah, etc. The 'is-a' connector can then be weighted by a degree of membership where appropriate. The purpose of this tool is to be able to analyze specific linguistic statements into these taxonomic categories to determine commonalties and relationships.

Integrating linguistic data (i.e. relative frequencies) with other 'hard' data can be accomplished through 'neural network - fuzzy set' techniques, used very effectively in modern control systems and pattern recognition. The term 'fuzzy' deals with terms such as 'more', 'less', 'strongly agree', 'strongly disagree', etc. When assessing preferences (the human utility function) these terms will make sense to a person rather than an arbitrary numerical scheme. Ultimately we will map them to a numerical scheme to create a computable function for a computer program.

These methods would be applied in a fashion something like the following: Every individual is the member of one or more communities; each of the latter share a common history, culture (language, prescribed behaviors, taboos), and an accepted political structure to determine authority. At least one community will have economic components and provide a means of livelihood; and at least one will provide the primary social matrix (usually family, kin system, clan, tribe or people). The methods and data are used to infer the relationships and relative weights between such communities in a manner congruent with human perceptions and values. This would be accomplished by generating a membership vector (list of elements) whose elements represent the degree of membership corresponding with the perceived cultural identities and preferences. This vector will determine the contribution to overall behavior by the values expressed by the community in question. On a global basis, we can predict probable behaviors by population segments and relative contribution to total. These 'memberships' define the clustering of actor-elements of our distributed network model.

It would also be possible to deal with the circumstance of ambiguous identity. Every individual attempts to integrate the various communities into a personally congruent and fulfilling life. Sometimes this is easy, more often these days, not, because in this shrinking world, we find ourselves more likely to belong to more groups having less in common. For example, the Bosnian Muslim Biophysicist and Musician might have more in common with a Pakistani Muslim Poet or a Hindu Biophysicist than a Bosnian Serb Musician (even though the last and the first may have two things in common rather than one). This sort of non additive combination is the stuff of 'fuzzy' systems where degrees of membership and relative weights can be accurately modeled. Furthermore, fuzzy systems techniques are 'mappable' to (i.e. capable of expression within) neural network computing models.

Such methods integrate individual elements operating in parallel and cross-linked with one another. As such, they provide a reasonably close computing analog to an environment-- such as the global environment --that consists of many mutually interconnected entities, each with its own evolving histories. This would correspond to the distributed network and learning concepts that we spoke of earlier. The result would also constitute, or lead to, a description of the organization of local communities consonant with the subjective viewpoint of the members of the culture.

The organization of the data would be first done in a taxonomic fashion, distinguishing core from peripheral values, ranking authoritative sources, mapping their domains, coding the evaluative criteria and ranking preferred behaviors. Next, we would cross-reference each cultural description to a 'normalized' standard by which to define a metric of consonance or dissonance between goals, criteria and behaviors. We then would measure the degree of interconnectedness among the cultural entities, and the degree to which the members of a given population group 'belong' to a community.

The linguistic part of the data would be inferred from assays performed on populations, to determine rankings of cultures and their relative impacts on individual behavior. A variant would use research experts who are local to the various social units and who specialize in the description of their cultural communities, as sources. (One use of the GENIUS global workstation network would be to extract these opinions.) Thus expert opinion would be treated as itself a kind of data. No doubt, each contributor might use his or her own terminology to describe social relationships, history, and cultural values; but the analytic technique would create a non-specific, general category structure, and map these disparate individual descriptions into it. In this manner, the 'raw' cultural information would be processed by a 'learning' system that would classify and model data describing the various communities and the relationships among them.

The above analyses would be done in a manner that was reproducible and consistent over extended time-frames and spatial domains. To the extent afforded by such reproducibility and consistency, compared to idiosyncratic human judgment, these analyses would be unbiased with respect to the values. Use of these techniques in such actor delineation-organization tasks is further described in Appendix I, which we offer as a discussion vehicle for clarifying and refining a suitable approach to this task.

Summary of Technical Requirements

The modeling of nonlinear systems and large scale, interconnected systems is still in development, but the basic conceptual frameworks are there. The technology to gather the indicated data in a cost-effective fashion and disseminate it is currently in existence and is becoming more powerful and flexible every day, as is the framework for structuring the multi-faceted distributed database. The ability to accommodate those data that are linguistic in nature and/or for which degree of membership is an important valuation, is also in hand. For such data, the framework for analysis and evaluation would follow the concepts of semantic networks modeling and fuzzy system controllers. In conjunction with these elements, the neural network modeling tools can provide an excellent 'normalized' ranking mechanism and a systems interconnection matrix, whose components are 'learned weights' that give relative strengths of component signals.

Thus the major issue is the integration of the appropriate tools, concepts and people into a shared project to achieve the goal of scientifically validated and practical global knowledge that we have set forth in the previous paragraphs.

BRINGING THE SERVICE TO OPERATIONAL STATUS

What are the resources, the sequence of steps by which we would use them, and the costs, to achieve a functioning prototype of the Service? We now consider these issues under sub-headings for project organization, personnel requirements, other resource requirements, budget, and for four project working groups.

Project organization

Overall responsibility for the GENIUS project and for policy matters affecting it will be vested in a parent non-profit corporation established expressly to promote the global information and forecast dissemination goal. This corporation will hold full ownership rights, possibly shared with outside sponsors/collaborators, to the results and assets of the project. The corporation will be governed by a Board of Directors, which will elect the officers of the corporation; act as primary liaison with outside persons, organizations, and funding sources; and exercise full powers of oversight, including final review and approval of all management actions, of funding and expenditure decisions, and of personnel actions.

Project tasks will be divided among a Project Management Group, a Strategic Planning Group, an Applications Development Group, and a Data Acquisition /Development Group. Each of these groups will be headed by an area Manager, who together will act as the primary co-investigators /developers of the project. These Managers will be chosen by the Board of Directors for fixed, renewable terms, during which they may also serve as corporate officers and Board members. The head of the Project Management Group will be Project Manager and will be responsible for overall operational management of the project. The functions of each of these Groups are further discussed in separate sections, below.

Choosing the key personnel will be particularly important to success. The history of human technical breakthroughs is replete with good examples of success and failure. Large scale projects are successful (Apollo, Manhattan) when there is a core group of participants whose complementary skills are brought together under a shared goal. Managers and Directors must have the ability to interact on a collegial basis and to follow a 'management-team' approach. The leader of the organization must have the vision, the tenacity, and the respect of others, to provide the necessary cohesion and to keep the project on course.

We recommend breaking the development process into two phases. Phase I, lasting approximately 12 months, will consist of: assembling initial personnel; start-up; and addressing the fundamental issues of project management and integration, strategic planning (service /operations planning, marketing, funding), applications development (choice of modeling tools/methods, Service design plan), and data acquisition /development. In its second half, it will also include a start on certain long lead-time data development tasks.

The Phase I deliverable will be a full development plan for the Service, including specification of full team, techniques and technologies to employ, and project schedule and budget.

Phase II, lasting approximately 18 - 24 months, will consist of building the validated prototype system for the domain specified.

The actual prototype for GENIUS can be built using the existing Internet data communications network. This can be supplemented by NREN's multimedia capabilities as they are brought into implementation. Several institutions have excellent programs and resources for socio-cultural-economic-political data development. Various research institutes, governmental agencies, such as (in the United States) the Meteorological and Geological surveys, and various United Nations and other international agencies will also be able to provide substantial contributions in developing the databases. International relations data development and data-based research activities taking place at a number of universities (e.g. Colorado at Boulder, SUNY at Binghamton, Maryland at College Park, Michigan) will be able to act as important resources in the data development and modeling tasks. The neural modeling technologies are strongest in the Southern California region, where there is a concentration of experts who will be able to build the first working prototype.

Choosing the boundaries of this prototype is by no means a trivial exercise. We suggest that a pilot, in the form of a complete model of the Caribbean region (island and boundary countries) plus Canada, would be a worthwhile starting point. Indeed the area is a locus of an issue of vital current interest, the North American Free Trade Agreement; and the area has a great many characteristics of a multicultural environment. Parts of the model will represent the effects of other regions (remainder of South America, Europe, Asia / Oceana, Africa) which can in the future be modeled in greater detail. In addition to the human societal components of the area, factors such as climate, geology, non-human populations and human biology will be modeled in detail.

The 'marketing' of the effort to properly develop the system and subsequent promotion of the forecasting, diagnostic, and prescriptive facilities to political, economic and other authorities, in communities everywhere in the world, will be crucial steps. All this effort is greatly diminished in value unless the 'policy-makers' come to feel they can rely on the Service as a more effective means of achieving the goals for their communities of interest, and can feel confident that loose ends won't come back to surprise them. The dissemination of the intermediate results of the effort and establishing the validity of the approach in a scientific and culturally neutral way are key. Conferences for experts, policy-makers, and interested public, as appropriate, will need to occur from time to time. How results are portrayed and communicated will be an important marketing issue. Know your audience!

Personnel requirements

Project participants will consist of Directors and officers, Group managers, consultants, project assistants, and outside contractors and organization representatives. Phase I personnel will primarily involve the Group managers, plus a few consultants and assistants as necessary; full staffing will occur in Phase II. The technical skills and areas of expertise we need to utilize are clear:

Climatology - geology
Biology-- population, ecological, developmental
Human health-- epidemiology, population dynamics
Ethnography - social systems
Economics
Political science
Historiography
Demography
Mathematics - physics-- complex, nonlinear systems
Scientific computation
Neural computing
Simulation - modeling-- Monte Carlo techniques
Expert systems development-- linguistic nets
Marketing research - statistical survey technique.

Further, we will need technical support for developing, installing and operating the automated systems. This will require programmers and network specialists.

In addition to the important criteria that we mentioned earlier, selection considerations will stress special expertise and ability to think in global terms, mutual complementarity of backgrounds, interest in the issues presented herein, and long-term ability to work together. It is also important that participants span a complimentary range of expertise and interests from mathematical, physical, biological, and social inquiries, in systems modeling, and in visualization and network technologies. These criteria will enable the project to make use of known and emergent complementarities among the different areas, to effectively compare and combine alternative forecasting techniques, and to address issues such as the possibility and usefulness of representing human and ecological phenomena as a single system. Careful choice of participants will enable the GENIUS project to effectively address such matters without the narrowness of perspective that has severely and needlessly handicapped previous forecasting efforts. As specific pieces of work are identified, the co-investigators will seek out appropriate support personnel and consultants to complete the data definition, gathering and integration into the total project. We expect to utilize as project assistants, persons who are thoroughly experienced and skilled in appropriate management and research functions, including data development, but who also are already familiar with the global-systemic perspective expressed in this proposal.

Other Resource Requirements

The following equipment and communication services will be necessary:

Communications services processor
Database services processor (with optical disk storage)
Analytical and expert application services
Communications network interfaces
Database application
Expert shell
Programming development environment
Workstation and workstation software.

These elements are needed primarily for Phase II, although basic communications, text processing and text and, to some extent, numerical database retrieval applications will be needed in Phase I for a few workstations. It is hoped vendors will contribute equipment and software at below market pricing as a means of participating in the project.

Budget

(Herein numbers within parentheses refer to notes at the end of each sub-section.)

Phase I:
$192,000 project team salary and benefits
$100,000 grants/contracts to outside sources
$ 25,000 equipment, software, and services
$ 65,000 data acquisition/development (1)

$382,000 TOTAL Phase I

Phase II (2):

Equipment:
$200,000 servers (3)
$150,000 application development tools / database software
$ 90,000 communication services (4)

Data acquisition/development (5):
$500,000

Application development:
$720,000 in-house work (6)
$ 60,000 outside consulting

Marketing:
$ 50,000 pre-development
$250,000 post-development

Project management (7):
$200,000

TOTAL Phase II:
$2,220,000 for 18 - 24 month prototype

.

GRAND TOTAL:
$2,602,000 Phases I and II.

Notes to budget:
(1) Outside contracting for long lead-time data acquisition/development to be initiated during the second half of the Phase I period.
(2) All entries are rough estimates, to be refined during the Phase I efforts.
(3) Including operating systems. In current parlance, 'servers' refer to central computer processing resources.
(4) For 18 months @ $5,000/month.
(5) For area manager and approximately 15 development teams (both in-house and grants to outside researchers) for 18 months to build prototype database.

(6) For area manager and approximately 10 full-time programmers and development experts.
(7) For project planning/oversight meetings, project manager, assistant, facility and expenses.

As mentioned before, project tasks will be divided among several working Groups with differing responsibilities, to which we now turn.

Project Management Group

This group, consisting primarily of the Project Manager and assistants, will be responsible to do the following:

1. Provide the necessary cohesion and leadership to keep the project on course (see above);

2. Report to the Directors on the status of the project, recommend actions to them, and oversee the implementing of their decisions;

3. Monitor, coordinate, receive reports and recommendations from, and convey information to, the other working Groups, outside collaborators, and contractors;

4. Manage financial transactions and record keeping;

5. Arrange for logistical and other operational matters;

6. Act, as needed, to assist the other working Groups in timely completion of tasks.

Strategic Planning Group

This group will be responsible for the non-computational aspects of planning and developing the Service, including the following:

1. Recommend where to locate the Service and its branches;

2. Determine staffing requirements for administrative support;

3. Draft and implement a business plan, including planned fees and subscriptions for actual use of the system;

4. Plan and carry out appropriate conferences and forums for Service development (as appropriate, in conjunction with marketing: see below);

5. Determine the specific kinds of networking, intervention, and support to be offered to research and action programs, to institutions, and to the public (as appropriate, in conjunction with marketing: see below);

6. Determine the identity of relevant institutions, and political/social action groups and networks; contact these groups where appropriate (in conjunction with marketing: see below);

7. Plan and carry out marketing activities (1), development of clientele (2), information dissemination, conferences, networking, and other steps to obtain outside funding, acceptance, technical assistance, and other forms of support for the Service.

Notes to Strategic Planning Group:
(1) Also see above, our earlier remarks on marketing.
(2) The term refers to political authorities and other parties who rely on the system. Clientele development is to be accomplished, inter alia, by demonstrating successful simulation of actual events and showing the impact of strategies actually followed in comparison with alternative strategies.

Applications Development Group

The task of this group is to make the GENIUS system operationally functional. As difficult as all the other tasks may be, this task is the hardest. Forecasting entails identifying and accounting for the connections among the many problems, factors, and entities, and discovering how to predict, to the extent possible, global system behavior and the consequences of human actions. For reasons explained more fully in Appendix II, we have chosen to break this task into two parts: the first concerning modeling, programming, validating, etc., the second concerning indicator development.

Modeling.

This aspect will involve the following responsibilities:

1. Identify appropriate models and research approaches for interpreting information and assessing trends (e.g., in the econometric, economic-political, and environmental areas, models such as SARU, GLOBUS, Project Link, FUGI, US Environmental Assessment Model), to serve as candidate sub-systems for integration into the model);

2. Refine the general conceptual model into detailed operational elements;

3. Decide how best to represent elemental behaviors, parameters, and conditions (whether deterministic or statistical methods apply, whether 'fuzzy' methods apply, etc.) and/or the appropriate emphasis among alternative measurement modalities; in particular, decide the relative emphasis and respective uses of 'quantitative' versus 'qualitative' or linguistic modalities;

4. Decide the scope and relative emphasis in representing various types of social and environmental phenomena;

5. Develop appropriate neural model elements for processing input variables, and determine weights;

6. Develop a semantic net model for linguistic analysis;

7. Develop a generative system to simulate behaviors over time, driven by exogenous events ('factual occurrences') (1);

8. Integrate sets of elements into coupled subsystems and test behavioral characteristics;

9. Implement overall model using appropriate development tools;

10. Develop and implement a test plan to verify precision by modeling past histories versus predicted trajectories;

11. Develop a demonstration version of the model for conferences, forums and marketing, including its use to project results of alternative policies (see exogenous events generative system, above, and Note 1).

Notes to Modeling, etc.:
(1) We expect that the model will predict some historical events of specified characteristics from the antecedent conditions; or it will produce a class of historical incidents, having common antecedent conditions, with a frequency of occurrence that fits the historically observed frequency for those antecedents. If such events had been historically preceded by a particular strategy or set of choices by one of the actors, then we could 'perturb' the model by introducing alternative strategies or choices (the exogenous events) which might have been taken (but were not), and observing the predicted result. In this way we would use the benefit of hindsight to evaluate the actual strategy or choices.

Indicator development.

1. Working with the modelers, develop or select appropriate operational, reproducible indicators and procedures for quantification of the "objective" part of the database.

2. Working with the modelers, develop the taxonomic part (1) of the database;

3. Working with the modelers and the Data Development Group, determine the match or mismatch between needed and available indicators, and seek remedies for the latter;

4. Survey existing programs of fundamental and applied research for suggestions of possible approaches to, and collaborations in, indicator selection-definition (e.g. the Correlates of War project at University of Michigan, the National Center for Atmospheric Research, the Santa Fe Institute); make appropriate contacts.

Notes to Indicator development:
(1) The taxonomic part is based on the semantic network category structure and techniques for obtaining standardized descriptors and normalized rankings.

Data Acquisition/Development Group

This group will have the following responsibilities:

1. Locate other data development activities and information repositories among educational institutions, research groups, government bodies, the United Nations, other international organizations, public interest groups, and private parties;

2. Contact and negotiate with other organizations to obtain or gain access to necessary information and data, or to collaborate in developing it;

3. Examine political and other institutional barriers to information and data availability and consider possible solutions and alternatives;

4. Develop an appropriate global network and procedures for gathering-generating data, including both "objective" and cultural (opinion-based) data;

5. Identify and address problems of data reliability;

6. Devise appropriate labeling, tagging, and coding schemes;

7. Develop or choose management schema and tools for routine, easy data entry, update, retrieval, use; and for database structural changes;

8. Develop interfaces to other independent subsystems to accomplish routine, scheduled update of externally-sourced database elements;

9. Create and test the database, using the results of the previous steps.

SUMMARY

A working prototype system able to effectively predict the behavior of multiple communities of interest to common stimuli-- is a marketer's dream and a politician's salvation. The ability to sell analytical services much the way econometric modeling companies do is clearly in the cards and obviously the significance of this modeling is much greater and worth much more. Once the prototype is in place, expansion and refinement could then be funded through subsequent research project subscriptions (for expanding into different domains) or even selling equity to subscribers (nations, corporations, other organizations, individuals). Fees for services based on scope of issue, depth of strategic modeling, and ability to pay, could be developed as well.

It is expected that a Caribbean model by itself could generate 200 subscriptions at $1,000 per month, sufficient to cover the routine operations and ongoing model refinement. Expansion to other areas (South America, Europe, Africa) would be funded through cash flow and obtaining further capital infusion from interested parties. Once a demonstrable product proves the technology, the expansion costs will be relatively easy to get. The biggest effort will then be in data acquisition for development and monitoring.

In essence we are seeking to create a major aid to the human capacity to think as a species and in global terms before we act, and to build and reinforce functional behaviors among human social units and individuals.

The approach that we have outlined would provide this aid by greatly increasing the ability to differentiate between what can, and cannot, reliably be foreseen about our global village and its environment; and by allowing one to understand more fully than ever before, in each specific problem or situation, the differences among one well-intentioned policy intervention, another, or none at all. This aid would also take the form of revealing alternative, culturally congruent courses of action arising from credible sources within the interacting social units-- the creative alternatives of which we spoke at the beginning of this proposal. Such knowledge would allow valuable new counsel to policy makers, other practitioners, and all those who care about the future.

APPENDIX I
NEURAL NETWORKS AND 'FUZZY MODELING'

Neural modeling uses the concept of a simple 'neuron' processor which takes many inputs and has a single output (on or off). The 'on' signal is achieved if the linearly weighted sum of the input signals exceeds a certain threshold value. The output in turn can be fed into any number of subsequent stage neurons to be factored together to produce next stage outputs, etc., until we get a final level which generates an overt behavioral change. Weights can change over time as the model spontaneously 'learns' from past mistakes. In our case we have mutually interacting individuals and groups. At the individual level we can model the various strengths (or weights) of direct environmental signals and signals from other neurons that reinforce or inhibit certain responses. The degree of contribution from the other neurons depends on prior experiences. These can be modeled using the concepts of proximity and credibility. Proximity would indicate a stronger weight, while credibility would indicate the sign or direction of influence. Thus, messages from nearby friends are important, as well as those from nearby enemies; casual strangers have little influence; while culturally authoritative sources (pope, imam, etc.) could have a significant contribution as well. The NPN model (negative, positive, neutral) allows us to rank the signals in a standardized fashion and indicate the direction of the signal in terms of reinforcement (+) or inhibition (-). The strengths would be normalized to the set [-1, 1] for simplicity and ability to use Markovian methods. The thresholds can be modeled using a vector whose individual component values are in the set [-1, 1]. The individual component represents the degree of dissonance allowed for a given environmental variable of concern. Thus each neuron acts as if it seeks to maintain some acceptable dissonance between its desired state and actual perceived state. The measures of each of these objectives can then be ranked.

We can treat individual neurons as cohesive groups depending on their degree of membership in the various communities of interest. This is similar to renormalization methods in physics to handle large volumes of elements having the same 'physical behaviors'. Thus our little example, of a Muslim Bosnian biophysicist versus a Serb Bosnian biophysicist being closely or distantly related in terms of degree of membership in Serb / Muslimness; or being Bosnian or being a biophysicist. We can then determine the resulting "objective vectors" of an individual which reflects the individual's degree of membership in various communities. Each community has an objective vector consisting of the set of desired values or norms. This technique allows us to aggregate both values assessed by all groups together with those applying only to some. If being Serb is overriding, the Bosnian and biophysicist valuations are overridden. Aggregation methods can determine the behavior of closely interconnected entities, mimicking crowd and other group behaviors. In essence, each community in which we have membership contributes something to the overall result, with competing values 'netting' out, leading to the overall result.

The development of a standardized taxonomy, to create a common meta language to describe values, behaviors and evaluative filters for facts, can be developed through the tool of semantic networks. The semantic network uses the concepts of categories or classes as a sieve to group linguistic terms that refer to common aspects. Thus a characteristic of a person or entity can be thought of as an instance of a certain category and, furthermore, a degree of membership in that category can be defined. The tools of fuzzy logic allow us to define membership functions which then are well defined set-theoretic entities that can be combined in computationally precise ways. All this allows us to translate back into the cultural milieu's own terms the facts and strategies prescribed and to measure the potential responses according to a single uniform measure. This will actually be the hardest part since it is where our own biases tend to occur. The assay rankings will be important as clues to connotations of the semantics we use, since the responses, in terms of strengths of agreement or disagreement with the premises, define the membership classes.

The links between the neurons are incorporated in an 'interconnection matrix'. Proximity and credibility of source signals are factored in the weights of the signals. The opinion-making process is then structurally described by this matrix. The elements of the weights can change over time as new sources arise and perceived credibility is reduced or enhanced based on history and evaluation of state.

When we look at the model of the individual agent (person or grouping of like persons), we see the following: What we have here is the representation of the fact that every individual has a personal history (memory) recording the results of previous encounters with reality (dissonance is a cumulative function), and is equipped by nature (through biological and physical resources and limitations) to perform a set of actions subject to culture (group history, values, behavioral paradigms). The agent has the ability to sense and filter events and messages from various sources (other agents, etc.) to determine an appropriate action which will affect its domain. An action taken will impact both the agent, through learning how its domain behaves when acted upon, as well as other agents linked through the common domain.

The common domain can 'expand' depending on time frames of delays to incorporate larger numbers of agents who are indirectly affected or loosely coupled to the domain. For example, a town which has a plant that generates wastes into a nearby river can affect many other communities downstream and also, through ground water, even communities not on the river. The plant may generate important economic goods that impact the livelihood of many other communities and play a significant role in international trade, engineering research, etc. The recognition that domains cannot be narrowly construed is crucial to obtain the requisite linkages to make forecasting accurate.

Within this model, conflict between agents is a result of the computed difference between the normalized "objective" vectors of the agents. The goals of each agent are described as components of a set of all goals. The strengths within the range [-1,1] indicate the relative desirability or non-desirability of a particular state of affairs in one's environment. Values of zero or nearly so indicate a 'do not care' situation. Further, each agent can be described as being a member of a number of communities (kin, tribe, clan, people, nation-state, religion, economic class, political party, etc.) There is a combination of goals from each of these source group memberships leading to an overall goal set of the agent. In some cases, the ranking of certain goals from one source will override that of another. Thus a strong desire from one group will override weak desires. On the other hand, opposing tendencies might either cancel as a way of avoiding internal conflict or might go with one or the other.

Amalgamations of agents can be modeled through statistical distribution functions indicating the likelihood of having a certain level of education, longevity, economic resources, etc. Therefore we can create equivalence classes of agents that partition the state space, having a certain 'deterministic' modeled behavioral response to events and other agents and having a statistical measure of resources and capabilities.

We can therefore project physical scenarios using climatic, geological and biological information, to forecast agricultural and economic occurrences and then model the results of these perturbations on the social/cultural/political matrix of human societies. Similarly we can model the human impacts on our physical environment and project its ramifications as well. Most of all, we can model the likelihood that given strategies will be pursued by differing communities under various economic, political and social stresses, based on the historical/cultural continuities and our ability to detect dissonance between actual domain states and desired ones for each community. Dissonance is the metric that measures the difference in value of the existing state of affairs versus the desired state of affairs. For a Bosnian Serb whose goal of being a member of a Serbian political entity is ranked at +1 (most important), living in a non-Serbian political entity is ranked as (-1) and thus the dissonance is 2. The value, indicating maximum difference between the vectors, will indicate that actions to redress the state of affairs is highly likely: when combined with other members of the same political entity who are adverse to a Serbian dominated political entity, one has then the makings of a serious conflict. Developing strategies that allow multiple communities to reduce dissonance simultaneously is the art which the Service will attempt to validate before they become policies inflicted on the communities in question.

APPENDIX II
RELATIONSHIP BETWEEN THE TASK OF CHOOSING INDICATORS, AND OTHER ASPECTS OF APPLICATIONS DEVELOPMENT

Roughly speaking, there are several parts to the task of applications development: specifying the predictive model (which likely includes creating or selecting several sub-models, then specifying how they shall be synthesized into one); programming the computational algorithms that will execute the model; developing the empirically operational indicators that will represent the 'real world' in the computations; calibrating-testing-refining the algorithm; and altering the model where appropriate. Closely related is the task of obtaining suitable computer hardware and software, and getting it into running order. Realistically, many iterations of this process are to be expected before obtaining a satisfactory result.

While all of these elements are intricately bound together, it is useful to regard empirical indicator development as a separate activity; albeit one requiring close consultation with the others. The reason for this is that getting a match, between variables fitting the model-builders' logic and variables corresponding to information that the world actually provides, is a problem by itself. To address this problem, we would suggest that 'indicator development' be constituted as a separate sub-group that consults closely with both the rest of the Applications Group and with the Data Acquisition/Development Group and that acts to mediate between them.

BIBLIOGRAPHY

Casti, John (1992) *Reality Rules*, Volumes 1 and 2, Wiley Interscience

Gelernter, David (1991) *Mirror Worlds*. New York, Oxford:, Oxford Univ. Press .

Goldstein, Joshua S. (1988) *Long Cycles: Prosperity and War in the Modern Age*. New Haven, CT: Yale Univ. Press.

Hammerstrom, D. (1993) "Neural Networks at Work", *IEEE Spectrum* Vol. 30, No. 6, June p. 26 ff.

Iberall, Arthur S. (1975) "On Nature, Man, and Society: a Basis for Scientific Modeling." *Annals of Biomedical Engineering* 3: 344-385.

Kaiser, David (1990), *Politics and War* Harvard University Press.

Karasik, Myron (1983) "Conservation Laws and the Dynamics of Economic Systems", IEEE Systems, Man and Cybernetics Conference *Proceedings* 1983 p 756 - 762

Karasik, Myron (1984) "Conflict, Cooperation and Differentiation", IEEE Systems, Man and Cybernetics Conference *Proceedings* 1984, p 12 - 17.

Karasik, Myron (1986) "Dynamics of Ensembles of Autonomous Systems", Pergamon Press (*Proceedings of Fifth Annual Conference on Mathematical Modeling*)

Kauffman, Stuart (1993), *Origins of Order*, Oxford University Press, Ch. 9 and 10.

Kennedy, Paul (1988) *The Rise and Fall of the Great Powers*. New York: Random House.

Kosko, Bart (1991), *Neural Networks and Fuzzy Systems*, Prentice-Hall

Kosko, Bart and Isaka, Satoru (1993) "Fuzzy Logic", *Scientific American*, June

Lewin, Roger (1992), *Complexity*, Macmillan

Mamdani, E. and Gaines, B. (1981) *Fuzzy Reasoning and its Applications.* Academic Press

Negoita C., and Ralescu, D. (1987) *Simulation, Knowledge-Based Computing and Fuzzy Statistics*. Van Nostrand

Prigogine, Ilya (1980) *From Being to Becoming.* Freeman

Rashevsky, Nicholas (1968) *Looking at History Through Mathematics.* Cambridge, MA, M.I.T. Press.

Rich, E. and Knight , K. (1991) *Artificial Intelligence.* McGraw Hill, Ch. 16-21.

Richardson, Lewis F. (1960) *Arms and Insecurity.* Pittsburgh: The Boxwood Press and Chicago: Quadrangle Books, pp. 12-36, 95-97, 145-148.

Singer, J. David (1980) "Accounting for International War: The State of the Discipline." *Annual Review of Sociology*, 6:249-367.

Small, Melvin and J. David Singer (1982) *Resort to Arms: International and Civil Wars, 1816 -1980*. Beverly Hills, CA: Sage.

Taagepera, Rein (1968) "Growth Curves of Empires." *General Systems* XIII: 171-175.

Wilkinson, David (1980) *Deadly Quarrels*. Berkeley, CA: Univ. of California Press.

Wilkinson, David (1987) "Central Civilization." *Comparative Civilizations Review* 17 (Fall): 31-59.

Williamson, Paul: (1985) "A 'Physical' Model of World Politics," *Systems Research*, 2 (October): 221-235. Abridgment in *Proceedings*, Society for General Systems Research: 1110-1119.

Williamson, Paul (1989) "Modeling International Linkages," in Manfred Kochen (ed.) *The Small World*. Ablex Pub., Norwood, NJ: 108-127.

Williamson , Paul (1992) "The Need for an Integrated Global System Approach," *Proceedings*, International Federation of Automatic Control/Scientists for Peace Workshop on Supplementary Ways of Improving International Stability," Bolton, Ontario.