SANDIA REPORT

SAND2007-5847
Unlimited Release
Printed September 2007

Architectural Considerations for Agent-Based National Scale Policy Models LDRD Final Report

David Strip and George Backus

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2007-5847
Unlimited Release
Printed September 2007

Architectural Considerations for Agent-Based National Scale Policy Models LDRD Final Report

David Strip
Computational Biology

George Backus
Exploratory Simulation Technology

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1316

Abstract

The need to anticipate the consequences of policy decisions becomes ever more important as the magnitude of the potential consequences grows. The multiplicity of connections between the components of society and the economy makes intuitive assessments extremely unreliable. Agent-based modeling has the potential to be a powerful tool in modeling policy impacts. The direct mapping between agents and elements of society and the economy simplify the mapping of real world functions into the world of computation assessment. Our modeling initiative is motivated by the desire to facilitate informed public debate on alternative policies for how we, as a nation, provide healthcare to our population. We explore the implications of this motivation on the design and implementation of a model. We discuss the choice of an agent-based modeling approach and contrast it to micro-simulation and systems dynamics approaches.

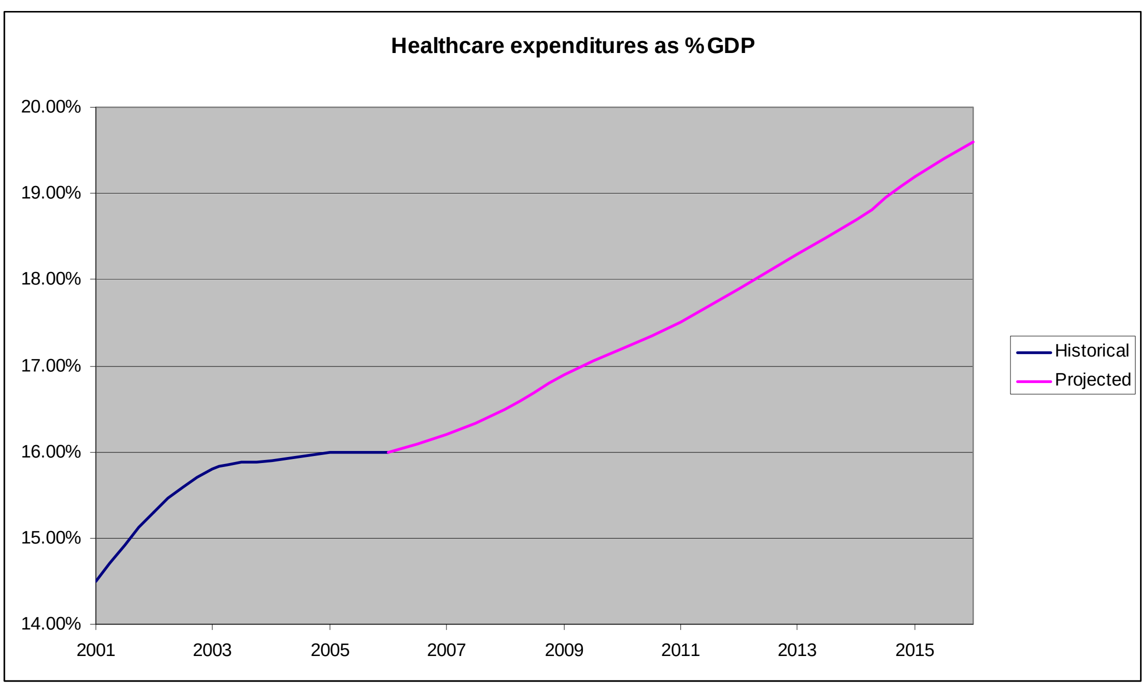
CONTENTS

FIGURES

1. Introduction

The need to anticipate the consequences of policy decisions becomes ever more important as the magnitude of the potential consequences grows. The multiplicity of connections between the components of society and the economy makes intuitive assessments extremely unreliable. Agent-based modeling has the potential to be a powerful tool in modeling policy impacts. The direct mapping between agents and elements of society and the economy simplify the mapping of real world functions into the world of computation assessment. The emergent behavior properties of agent models make them especially well-suited in environments where it is difficult to tease out all the myriad feedback mechanisms. We will study the potential of agent-based systems for policy analysis with a motivating example in health care. Healthcare costs are threatening to swamp the budget, leaving reduced funding for other critical national security needs. Healthcare policy analysis benefits from an abundance of detailed data and a lack of security restrictions which might otherwise accompany a data set of this magnitude and scope.

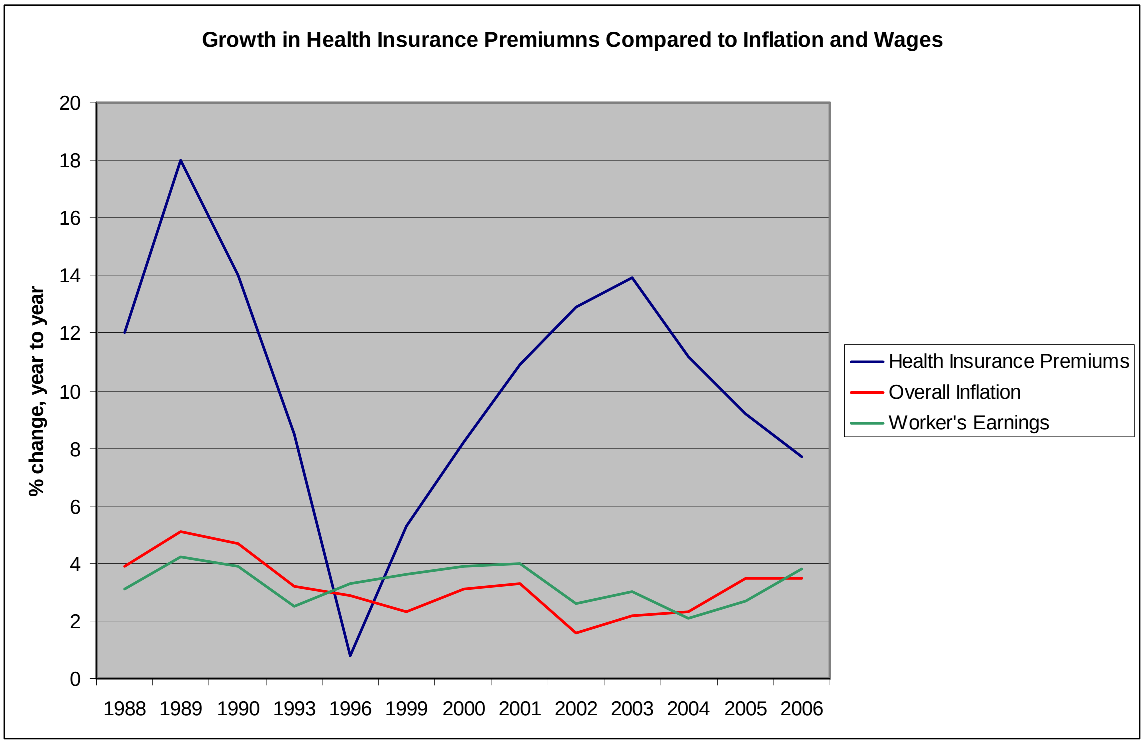
Crisis is perhaps the most common descriptor for the state of the US healthcare system among analysts and commentators. Rapidly increasing costs are placing a financial burden on federal and local governments, businesses, and individual citizens. Healthcare costs consume 16% of GDP in the United States and because these costs are growing faster than the economy, they account for an increasing fraction every year. The Office of the Actuary of the CMS forecasts these costs to grow to almost 20% by 2016.



Source: www.cms.hhs.gov/NationalHealthExpendData/downloads/proj2006.pdf

Figure 1: Healthcare as %GDP

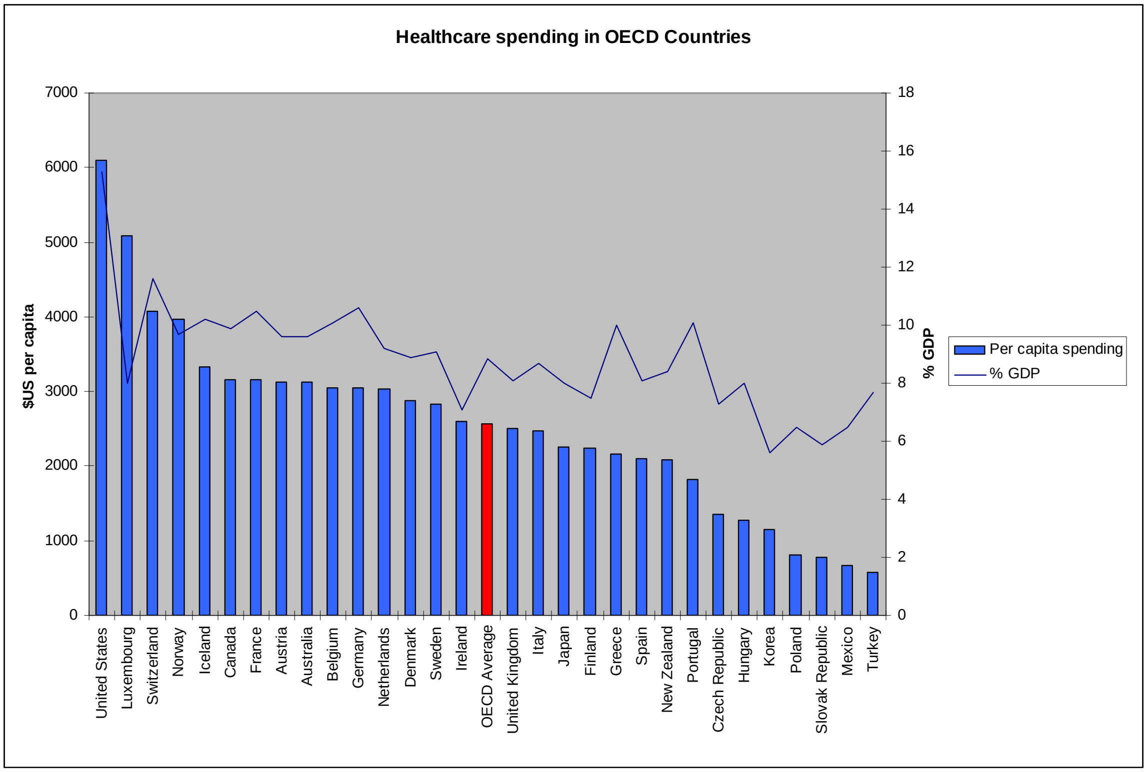
Health insurance premiums continue to grow, further challenging the competitiveness of American firms. While down from a peak annual growth rate of 13.9% in 2003, the 2006 growth rate is still 7.7%, well above inflation. (See Figure 2.)



Source: Kaiser Family Foundation - www.kff.org/insurance/7527/sections/ehbs06-1-1.cfm

Figure 2: Growth in Insurance Premiums

In spite of the fact that the United States leads the industrialized world in terms of its expenditure on healthcare as a fraction of GDP and in per capita spending (see Figure 3 below), it lags the industrialized world in key health indicators such as infant mortality and life expectancy [1]. Japan, with one of the lowest health care costs among industrialized nations, also has one of the longest life expectancies. For example, the life expectancy of a Japanese woman is 86 years while that of her US counterpart is 80 years [1]. The US ties for 35th place worldwide for the probability that an infant will live to see his or her fifth birthday, lagging behind Cuba, Slovenia, Cyprus, and Croatia, as well as most of the leading economies such as Japan and the countries of Western Europe. US averages, computed over a highly diverse society's population, may in fact paint a deceptively rosy picture, hiding disparities that are much larger than those in countries dominated by national healthcare systems.



Source: www.oecd.org/document/16/0,3343,en_2649_37407_2085200_1_1_1_37407,00.html

Figure 3 International Health Care Costs

Current growth patterns cannot be sustained. A significant increase in demand can be expected as a consequence of the aging of the "Boomer" generation. The current epidemic of obesity, especially among children, has ominous implications for associated chronic illnesses and their cost of care. Left unchecked, growth in healthcare spending would crowd out all other government functions and dominate the costs of operating a business. These outcomes are not compatible with a functioning economy and consequently changes will be made in the healthcare segment of the economy. The key issue is the nature of the changes and the processes that introduce them; will the changes be planned and orderly, or will they be disruptive and chaotic?

Without adequate decision support tools, we have limited ability to fathom root causes and to make informed policy. Approximately 45 millions Americans lack health insurance – either through inability to pay or by choice – whereas in most industrialized nations coverage is universal. The ability to understand the demographics of healthcare coverage and the consequences—both positive and negative—of any policy-driven changes to the way healthcare insurance is utilized in the US is a goal of effective decision support tools.

Although any number of ideas (sometimes conflicting) have been proposed as potential solutions to the impending healthcare crisis, tools for predicting the impact of proposed policies on healthcare costs and the adequacy of patient care have been severely lacking Particularly frustrating to policy-makers and the general public is the observation that so many changes to the system have resulted in seemingly the opposite of their intended effect, often increasing costs while reducing the overall quality of outcomes. Even the value and causes of these increased costs are often poorly understood.

The healthcare system is characterized by complex interactions between large numbers of patients, hospitals, insurance companies, and governing bodies. High-performance computing coupled with agent-based modeling will provide a tractable simulation of the US healthcare system as opposed to more traditional methodologies that are overwhelmed by the complexities of this problem. Successful development of such a tool would be extremely useful in performing studies that capture the interplay, cost dynamics, and merits associated with various policy decisions, methods of healthcare delivery, and myriad other factors.

2. Objectives and Considerations for a National Model

Scope

Our modeling initiative is motivated by the desire to facilitate informed public debate on alternative policies for how we, as a nation, provide healthcare to our population. This motivation has two principal implications for the model. The first is the domain of the model. Existing models in the health policy domain are typically health finance models – they measure who has health insurance, who is paying for health care, where the money flows are. What they don't measure is health. The underlying goal of a healthcare policy is to insure that the population has the opportunity to have a healthy life. In order to compare policies along this dimension, our model has to forecast a much broader range of measures than just health economics, but rather health outcomes as well. These will be discussed in greater detail in a later section.

Openness

The second major implication for our model is the issue of openness. Existing models, such as those from the Lewin Group, the Urban Institute, and others are essentially black boxes – the owners reveal the inputs and the outputs, but only the most rudimentary information about the internals of the models is publicly disclosed. As a consequence, it is impossible to reconcile the outputs of two models when they fail to agree. Robust public discussion on policy alternatives cannot survive if competing models provide radically differing results and it is impossible to determine the validity of the model. A successful policy analysis model must be open, at a variety of levels, to achieve the role we seek. The first, and most literal layer of openness is the source code. The source code must be publicly visible so that the actual implementation is visible to anyone of sufficient skill to read it. This enables an outsider of sufficient skill to independently verify any claim we make about the code.

Given the complexity that is inherent in a system of this scale, it is essential that the code be documented to make independent review practical. This documentation must exist both at the code level, primarily in terms of comments, but should also exist in the form of design documents and technical reports on the underlying models, their mathematics, and implementation, verification, and validation issues.

Comprehensive analysis of interesting policies will almost certainly require a considerable volume of input data covering demographic, economic, geographic, and other factors. Not only must data sources be identified and the data acquired, it will often have to be reformatted in a manner compatible with the requirements of the model. In order to prevent data acquisition from becoming a major impediment to widespread use of the model, documented input data sets, particularly for non-controversial data such as census data, must be provided with the executable model itself. Wherever possible, input formats should be consistent with standard sources of data. When not possible, tools for selecting and transforming standard data sources into the required format should be developed and supplied with the model distribution.

The objective of supporting robust, open debate dictates that we place as little barrier as possible to widespread availability of the executable model. This leads to the conclusion that the model should be available with no licensing fee. This does not imply, however, that the model is distributed into the public domain. There are numerous “open-source” licensing models, ranging from the GPL with its “viral” effect, requiring improvements to be licensed under identical terms, to those that mainly preclude others from asserting rights to what has already been developed under the license. The choice of exact license terms will have to be an early decision. Perhaps more important is to control the *name* of the model and how it is used. For example. $\TeX$ is a registered trademark of the AMS. As such, they can control what is and is not called $\TeX$. Our model should be managed in a similar manner. By creating legal ownership of the name, we can, at least in theory, control what may or may not be called by that name. Someone making a change to the model or a derivative model may be barred from describing their results as being a product of the model. If we have achieved our goal of being a credible and neutral prediction model, then failure to contribute a new module or data set to the model community will lead to a devaluation of the derivative model and a general disinterest in results produced using the derivative. Thus peer pressure, rather than legal mandate, becomes the enforcement mechanism.

Scalability

In the long-run, the model we are developing will be required to produce high-fidelity forecasts of complex policy proposals on a national scale and over a long (25+ year) horizon. The computational demands of this scope have the potential to be substantial. We are cognizant that only a small fraction of our target user community will have access to world-class supercomputers, should the model require this level of resource to produce results in acceptable run times. While we will endeavor to make the model as efficient and compact as possible, we will not shy away from using the most powerful computers in our reach, *if the model requires it.* That said, we do recognize the equal importance of making the model usable by the much larger community that does not have access to supercomputers costing tens of millions of dollars. To this end, the model will support scalability across a variety of dimensions to facilitate its use on machines from supercomputers on down to desktop PCs. The simplest dimension is time – by running longer on less powerful machines, we need to be able to produce the same results produced in shorter time spans on the supercomputers. Another dimension is geographic extent. Rather than modeling the entire U.S., the model must be able to support state and regional analysis, with an appropriate proxy for the balance of the country. By sacrificing fidelity (hopefully at predictable levels), one can reduce the number of agents (the model resolution), the time resolution, or spatial resolution, leading to lower time and/or memory requirements on less capable platforms.

Another notion of scalability lies in the level of expertise required to operate the model. The overall project design should anticipate the need to support users at a wide range of levels of experience and understanding. It is largely inherent that the required level of sophistication of the modeler will be proportional to the sophistication of the analysis to be performed. Analyses with novel requirements will require experienced users, perhaps requiring a capability to extend the model, develop new data sources, or provide statistical analysis techniques to measure output validity. However, we can, with appropriate effort, support narrow or targeted analyses by even the least sophisticated users. For example, if a specific proposal is under consideration in Congress, we can use the model to develop a response surface model to measure the model's behavior as a small number of variables are modified. A web-based interface could then be provided to the pre-computed response surface, enabling anyone with web access to experiment with the model, developing a personal understanding of the dynamics as the allowable parameters are changed.

Educational Initiatives

If we are to achieve our goal of improving the level of discourse on healthcare policy issues, it is essential that we develop a community of technicians who can operate the model and explain it's outputs, that we train health policy analysts about the power and limitations of modeling, and finally that we educate policy-makers on how to integrate modeling and modeling results into their decision-making processes. We expect the bulk of the educational initiatives will be carried out by academic partners who are better situated to perform this aspect of the program. We are engaged in on-going discussions with the University of Texas, particularly the LBJ School, and the Leonard Davis Institute for Health Economics at the University of Pennsylvania in this area.

We anticipate two levels of courses for detailed user training, one for general users and a second for developers who desire to extend the model. Both would begin with an overview of the model and discussion of the principles underlying the model, such as the agent structure, choice behavior, geospatial impacts, economic behavior, and other key concepts. The users course will then focus on how to translate policies into the required input format, how to find the appropriate input data to run the model, the literal mechanics of running the model on a range of platforms, how to scale your query to available computational resources and input data, and finally how to interpret and present results. After the overview, the developers course will focus on details of the code structure – identification of the major modules, the object structure, and program flow. We will cover how these key factors are impacted by problem scale and platform capability. The course will sketch out how various sorts of model extensions are implemented and work through at least one case in detail.

Training for policy analysts will exist at two levels. At the more fundamental of the two, we expect our university partners to develop curricula incorporating computational modeling into advanced degree programs in policy analysis. The materials would include course outlines, text books, and case studies. The other training thread would be enhancement/enrichment courses for practicing policy analysts covering similar topics, but in a short-course format.

The final component of the educational program is directed towards the policy-makers themselves. The target audience would be senators and representatives, their key staffers, agency executives, and their counterparts at the state level, such as governors, state secretaries of health, etc. The goal of the executive-level training is to teach policy-makers how to use models in decision-making – what is modeling, what are the limits, how much to trust the output, what the statistical terms mean, and other similar topics. These courses will have to be short, in some cases perhaps no more than an hour lecture during an orientation program for new governors or senators.

3. Modeling approach

Healthcare policy modeling has been dominated by micro-simulation models. This approach is relatively easy to implement, has only modest data requirements, and is reasonably well-suited to single-stage analyses focused on health finance issues. The model used by the Lewin Group, a well-known consulting group in the healthcare policy field, is described as a micro-simulation model. (Relating to our earlier discussion on openness, almost nothing is published about the Lewin Group’s model, making it impossible to accurately describe it’s internal structure or operation.)

Systems dynamics is being used with increasing frequency in modeling healthcare, often with policy implications. Systems dynamics is attractive for several reasons. One is that it is particularly well-suited to representing models that incorporate feedback mechanisms. The second is that it captures dynamic behavior over time. The third is that models can be developed through graphical interfaces, eliminating the need to learn a programming language. The primary drawback of systems dynamics models is that they represent mean behavior of the system; only with great difficulty can they be used to express the range of behaviors or effects on a population of entities. Thus, in a health policy example the systems dynamics models might tell us what fraction of the population is insured over time, but is not well-suited if we need to know how the coverage varies by race, income, geographical location, or other demographic factors. Another limiting factor for systems dynamics models is that the work best and are most easily implemented when their feedback mechanisms can be expressed as well-behaved, closed form mathematical functions. With sufficient effort either of these limitations can be overcome, but the required effort generally offsets the benefits of using the systems dynamics approach. Systems dynamics models have been successfully developed for facility management and disease management types of problems. They do not have much history in application of the sorts of policies that would be considered as a national healthcare policy.

We are planning to use an agent-based paradigm for our implementation. The agents are small computational entities (or function) that are programmed with a set of behaviors and are then set free to work with each other according to a set of interaction rules. There are several benefits to the agent-based approach. First is the very natural mapping of the model to the underlying reality that we are attempting to model. This feature is particularly helpful in explaining the model to policy-makers and impacted communities or individuals. In addition, this natural mapping can facilitate addition of new features to the model because of the natural analogy between the real-world behavior and it's representation in the model. A second strength is emergent behavior, a term that is much abused and we use here with a certain reticence. Emergent behavior, as we use it here, is the property of a system, in this case a simulation, to exhibit unprogrammed, apparently complex behavior as a consequence of the interactions of agents programmed with basic or primitive behaviors. One of the canonical examples of emergent behavior is flocking of birds. In some of the earliest work on computational agents and emergent behavior, Reynolds [2] demonstrates that three simple steering behaviors by individuals leads to very lifelike and seemingly purposeful flocking behaviors by simulated "birds". The third strength is the natural scalability of agent-based models. The fidelity and scope of the model is tightly coupled to the number of simulated agents, facilitating our goal of making the model useful at a variety of scales. Finally, agent-based modeling quite naturally produces distributional information on outcomes, in contrast to the systems dynamics approach, while also revealing dynamic effects and feedback behaviors, in contrast to micro-simulation. Probably the greatest weakness of the agent-based approach is the hurdle it creates to efficient parallelization for execution on massively-parallel (MP) supercomputers. Extremely careful problem formulation is required to enable a good mapping onto the most common architectures in MP machines.

4. Model Domain

General Concepts

The very nature of healthcare policy has implications for modeling practice that are independent of our choice of modeling approach. The first is the requirement that our simulation tools provide both distributional information and dynamic effects. In analyzing alternative policies it is clearly not sufficient to look only at average behavior or steady state behavior. Average behavior ignores inequities among populations. Average behavior may not even apply to anyone in the population. In bi-modal distributions, the average may fall in a place where there is zero probability. Steady-state behavior has the possibly severe shortcoming that it may only come after passing through unacceptable intermediate states and is there completely irrelevant. Even when all intermediate states are acceptable, they may not be desirable or preferable to the sequence of states of an alternative proposal with inferior steady-state behavior, or the time to reach the steady-state behavior leaves you in inferior intermediate states for so long that the overall policy value is less than that of an alternative with a less preferred steady-state, but one that is reached quickly.

Policy robustness measures the degree to which the policy produces the desired outcomes over a range of scenarios that represent possible futures. Will the policy work when the economy is strong? weak? average? Will we get the same results if there is a major flu outbreak in the next 10 years? The model must be capable of running a range of scenarios so that the scope of impacts can be witnessed. Policy analyst training must be sure to cover this concept and how it is tested using models.

Sensitivity analysis is a related, but subtly different concept from robustness. Sensitivity analysis seeks to determine which factors impact the policy outcome. Understanding sensitivity is essential to understanding where to invest your resources in developing data sources, resolving modeling ambiguities, or increasing resolution in some dimension. Sensitivity is a function of the model, the policy, and the outcome parameters we are interested in. As a consequence, some terms will have a high importance in almost every problem, while others may be important for one question, but not for another. Developing a broad understanding of sensitivity is essential so that we don't have to perform a complete scope of tests every time we have a new policy to test.

Verification and validation (V&V) and uncertainty quantification (UQ) are an essential component of any modeling initiative. A successful program will need to consider these topics throughout all phases of the project.

Outputs

Although we can't anticipate every possible output parameter that will be necessary to analyze the range of policies future users will want to evaluate, we can expect the following factors to be of interest:

- Mortality
- Morbidity
- Access
- Cost
- Quality of care
- Longevity
- Quality of life
- Workforce impacts
- Economics impacts (perhaps to serve as input to independent economic models)

As noted earlier, point estimates of these parameters is not adequate. The model must produce the distribution of these parameters over a variety of control variables, such as age, race, gender, socio-economic status, geographic location, educational level, and marital status. With so many output parameters and controlling variables, the modeling suite should include tools for querying and analyzing the output datasets and a capability to output data in formats usable by commonly available tools, such as spreadsheets or statistical analysis packages. Formats unique to the model or proprietary to a single vendor should be avoided whenever possible.

Inputs

Like the outputs, we cannot predict every input that will be needed to model every policy someone will throw at the model, but we can anticipate that the following baseline inputs will be generally required:

- Population/demographic data (this includes census-like data about individuals, households, employment, education, geographic location)
- Employer data
- Qualitative choice models/parameters

Because of our interest in health outcomes, in general we can expect to require a variety of health effects data that can used to predict the health impacts of policy choices. These might include

- Disease scenarios (e.g., robustness to pandemic flu, AIDS)
- Health response data/models (e.g., heart disease vs. cholesterol medication compliance, impact of co-pay on compliance)

For the foreseeable future we expect our model the general economy to be very basic, so we either need to provide key input assumptions about the course of the economy over the course of the simulation, or link to an external model that produces the relevant economic responses.

And finally, we need to provide the policy itself as an input. Some sort of *policy markup language* will be developed. An editor for writing in this markup language and for validating the resulting policies is a high-priority task.

5. Implementation approach

The model will be implemented in C++ using an object-oriented approach. Python will be used as a scripting language where appropriate.

Code will be managed from a publicly accessible SVN (or similar) repository system. Commit authority will be closely controlled.

Our implementation paradigm will lean towards the agile approach more than the traditional waterfall implementation. The agent-based perspective naturally lends itself to a highly modular design and facilitates the incremental implementation dictated by the agile paradigm. Going beyond being modular, we will develop a plug-in style interface that facilitates the substitution of behaviors in the agent models, gathering of additional statistics, and addition of novel functionality.

Development of V&V/UQ strategies throughout the implementation will be an essential task.

6. The Software Ecosystem

Although the simulation kernel probably is the hardest and technically most critical component of the project, our goal of impacting how policy is developed and evaluated requires a suite of tools to surround the kernel. In an open source environment such as we propose, adoption of some key strategies will encourage the growth of these supporting codes from others out side the project. Among these strategies are

- Use of standard, and especially non-proprietary, data formats for input and output.
- Provision of a well-documented plug-in architecture to facilitate development of model extensions.
- Well-documented source available from an easily accessible repository running a widely used source-code management system.
- Examples and tutorials on developing code extensions and plug-ins.
- Clearly delineating, separating, and minimizing platform dependent code.

We expect the supplemental codes to include input editors for defining runs, run managers for controlling execution, particularly when running large numbers of scenarios or replications on multiple hosts, and output viewers and analysis tools for managing the model output, which will have the potential to be extremely large.

The model we propose will have substantial data input requirements. Must of this data will be independent of the policy we wish to evaluate. To facilitate model use, it is essential that we provide accessible repositories of formatted, license-free data for use by anyone capable of running the model. Similarly, analyses performed using the model have the potential to produce large output datasets that would be expensive to replicate and are of potential use to a community of analysts. We should provide a means to store output datasets, a method to locate cases of interest, and a means to retrieve the data. All output datasets, and especially those in the repository, have to be annotated with metadata that gives sufficient precision to the specification of the input data and run-time parameters to enable independent replication of the results.

References

1. WHO World Health Report, 2006, Statistical Annex http://www.who.int/entity/whr/2006/annex/06_annex1_en.pdf
2. Reynolds, Craig, http://www.red3d.com/cwr/boids/

Distribution:

1	MS-0123	LDRD, 1011
1	MS-0370	George Backus, 1431
1	MS-1316	David Strip, 1412
2	MS-0899	Technical Library, 4536 (1 electronic and 1 hardcopy)
2	MS-9018	Central Technical Files, 8045-1 (1 electronic and 1 hardcopy)