21st Century Simulation:
Exploiting High Performance Computing and Data Analysis

Dan M. Davis
Information Sciences Institute, USC
Marina del Rey, California
ddavis@isi.edu

Garth D. Baer
Oracle Corporation
Culver City, California
garth.baer@oracle.com

Thomas D. Gottschalk
California Institute of Technology
Pasadena, California
tdg@cacr.caltech.edu

ABSTRACT

This paper identifies, defines, and analyzes the limitations imposed on Modeling and Simulation by outmoded paradigms in computer utilization and data analysis. The authors then discuss two emerging capabilities to overcome these limitations: High Performance Parallel Computing and Advanced Data Analysis. First, parallel computing, in supercomputers and Linux clusters, has proven effective by providing users an advantage in computing power. This has been characterized as a ten-year lead over the use of single-processor computers. Second, advanced data analysis techniques are both necessitated and enabled by this leap in computing power. JFCOM's JESPP project is one of the few simulation initiatives to effectively embrace these concepts. The challenges facing the defense analyst today have grown to include the need to consider operations among non-combatant populations, to focus on impacts to civilian infrastructure, to differentiate combatants from non-combatants, and to understand non-linear, asymmetric warfare. These requirements stretch both current computational techniques and data analysis methodologies. In this paper, documented examples and potential solutions will be advanced. The authors discuss the paths to successful implementation based on their experience. Reviewed technologies include parallel computing, cluster computing, grid computing, data logging, OpsResearch, database advances, data mining, evolutionary computing, genetic algorithms, and Monte Carlo sensitivity analyses. The modeling and simulation community has significant potential to provide more opportunities for training and analysis. Simulations must include increasingly sophisticated environments, better emulations of foes, and more realistic civilian populations. Overcoming the implementation challenges will produce dramatically better insights, for trainees and analysts. High Performance Parallel Computing and Advanced Data Analysis promise increased understanding of future vulnerabilities to help avoid unneeded mission failures and unacceptable personnel losses. The authors set forth road maps for rapid prototyping and adoption of advanced capabilities. They discuss the beneficial impact of embracing these technologies, as well as risk mitigation required to ensure success.

ABOUT THE AUTHORS

Dan M. Davis is the Director, JESPP Project, Information Sciences Institute, University of Southern California, and has been active for more than a decade in large-scale distributed simulations for the DoD. While he was the Assistant Director, Center for Advanced Computing Research, California Institute of Technology, he managed Synthetic Forces Express. He has also served as a Director at the Maui High Performance Computing Center. He served in the USMC on active duty and is a Commander, Cryptologic Specialty, U.S.N.R.-Ret. He has been the Chairman of the Coalition of Academic Supercomputing Centers and received a B.A. and a J.D., University of Colorado, Boulder.

Garth D. Baer is a technical analyst who is studying the impact of policy on technology as well as the changes technology makes on policy formulation and implementation. A Principle Support Engineer at Oracle Software Corporation, he develops new web-based applications and troubleshoots production database issues. Earlier, he was a Mission Control Engineer for Milstar Communications Satellites at Lockheed-Martin. He recently was invited to join a multi-university group developing a vision statement for DoD policy on M&S. He received Bachelors in Physics, Univ. of Colorado, Boulder and a Masters in Technology Management, Colorado Technical Univ.

Thomas D. Gottschalk is a Lecturer in Physics at the California Institute of Technology and a Member of the Professional Staff, Center for Advanced Computing Research there at Caltech. For the last decade, much of his research has been on the use of parallel computers to simulate various physical phenomena. His instructional duties include his upper division course on Statistics for Caltech Physics Graduate students. He received a B.S. in Physics from Michigan State University and a Ph.D. in Theoretical Physics from the University of Wisconsin.

21st Century Simulation:
Exploiting High Performance Computing and Data Analysis

Dan M. Davis
Information Sciences Institute, USC
Marina del Rey, California
ddavis@isi.edu

Garth D. Baer
Oracle Corporation
Culver City, California
garth.baer@oracle.com

Thomas D. Gottschalk
California Institute of Technology
Pasadena, California
tdg@cacr.caltech.edu

BACKGROUND AND INTRODUCTION

This paper will discuss the immediate need to augment simulations, offer effective enhancements, and show how these enhancements can best be implemented.

The Need to Improve Simulations

For thousands of years, leaders have used various representational methods to prepare for the defense of their societies. These have ranged from the venerable game of chess to complex electronic emulations of combat. When threatened, there is an understandable pressure to use what has proven reliable in the past and there is a countervailing desire to make use of effective new techniques. Two of the promising technologies available to defense leaders today are high performance parallel computing and advanced data analysis.

Contemporary analysts are faced with increasing pressure to provide more opportunities to both analyze the present dangers and train for the future operations. The vacant battlefield of yesterday is being replaced by the crowded urban warfare environment of today, populated with non-combatants for whom there is an increased sense of responsibility. Weapons of increased destructive power and refined targeting capabilities make it both possible and necessary to honor this sensitivity. Planners and trainers must have access to simulations of unfettered scale that are built on increasingly sophisticated environments, with better emulations of foes and more realistic civilian populations. The coordination and synergy of these simulation and analytical capabilities are necessary to deliver insights for the analyst and trainee.

There are well-recognized limitations that restrict the full exploitation of what the DoD calls Forces Modeling and Simulation (FMS), (HPCMP, 2004). This paper focuses on two:

- The inherent constraints of current computer-use paradigms
- The restrictions found in traditional techniques of data analysis

In order to meet the two-fold test of reliability and efficacy, new capabilities designed to overcome these limitations must provide sufficient improved utility to warrant the risk and effort expended in adopting them.

The nature of the adoption process is critical. The correct approach will lead to early productivity, low risk and continued utility. A well-thought-out plan, following the proven paths of analogous analytical disciplines, will reduce cost and accelerate benefits. Disciplines of interest include academic research fields investigating physical and biological phenomena. After several decades of using high performance parallel computing and advanced data analysis techniques in these areas, the pitfalls to avoid and the productive paths to follow have been clearly established.

Limitations Imposed on Modeling and Simulation by Current Computing Paradigms

The FMS community has become accustomed to waiting for the additional power represented by the doubling of circuit devices on a computer chip every 18 months. Being able to move from the floor of the gym at the Naval War College (see Figure 1*)* onto the vastly larger canvas of a digital computer terrain database was a momentous leap.

Figure 1. *1930's - U.S. Naval War College personnel conducting simulated campaigns on a gym-sized floor.*

The more distant horizons, such as global-scale, high-resolution terrain environments, seem out of reach. The FMS community has an opportunity to overcome this unnecessarily limited vision.

Terrain databases are now available in multiple resolutions and for nearly every area of the globe. Using workstation and PC technology hosted on LAN configurations, truly incredible advances have been made in our ability to provide a realistic and geographically appropriate environment for conducting large operations (Ceranowicz, 2002). Even these capabilities, however, are often limited in two important dimensions: resolution and total area. As the areas of interest broaden for both the analysts and policy makers, the need to have access to representations of any terrain becomes more imperative.

An example of an important feature of current computer practice to which the community has become accustomed is the constraint imposed by the limits of individual processing speed. The desire to represent sophisticated behaviors requires ever-increasing processor power, and this is magnified by the desire to run multiple instances of non-deterministic simulations to evaluate the range of outcomes (Horne, 1999). The need to represent tens of thousands of entities that are "aware" of each other also impacts performance. In one class of this "awareness," entities are within a range where they can see each other. A much more extreme case is now of concern: the high altitude intelligence platform with sensors that can "see" virtually every entity in an entire theater of war. Current programming, as exemplified in the SemiAutomated Forces (SAF) programs, handles this location and awareness issue by running an inter-visibility calculation every few milliseconds. Obviously, with a huge number of entities, this represents a huge compute burden. Current practice shows this type of situation can be simulated on a typical single processor of present-day (2004) capacities at only a few hundred vehicles. On a network of similar PCs on a LAN, experience seems to indicate that the total vehicle count is limited to a few tens of thousands - not even enough for military vehicles in a large battle.

Moreover, modern battlefields are rarely located on remote plains, and the battles in urban areas are not fought with the destructive abandon of World War II, as in Stalingrad. Instead, the modern analyst is looking for ways to achieve national goals while operating in populated urban areas with negligible loss of non-combatant life, minimal destruction of civil infrastructure, and reduced losses to friendly forces. For that reason, the simulations-enabled analyst is faced with the challenge of trying to understand how modern intelligence platforms can view a city full of vehicles and other entities. Clearly, something on the order of a million civilian entities approaches realism; a few thousand does not.

Limitations Imposed by Traditional Data Analysis

Similar constraints are observed when using only the traditional methods of data analysis. Historically, the validation of the insights gained from simulation are not infrequently lost by virtue of the imposition of accepted views.

Analytical approaches have not changed much over the intervening decades. With all of our increased sophistication in electronically produced simulations, one very common method of strategic deliberation remains the observation, logging and analysis of simulation outcomes by subject matter experts (SMEs). The authors maintain that adopting and implementing analytical techniques used in the behavioral sciences and operations research should enable these experts to be even more valuable.

In trying to understand the output of simulations similar to Project Albert, one is faced with a virtual flood of information (Brandstein, 1998). This flood presents problems in collection, collation, and consideration. SAF and Albert programs are driven by the application of a series of pre-established probability tables for many of their activities, *e.g.* accuracy of fire, damage occasioned by weapons strike, mechanical failures. Against these tables, a random number is applied and the resultant action is implemented. This results in a non-deterministic simulation. Analysis can be much enhanced, if the simulation is run multiple times, with the resultant outcomes appropriately analyzed.

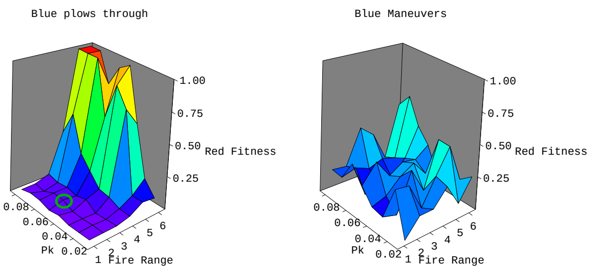

Figure 2 . *Three-dimensional representations of the effects of two parameters* (Brandstein, 1998).

As the analysts consider the outcomes, they are faced with assessing the impact of several parameters simultaneously. These multi-dimensional solution spaces become difficult to visualize, as *n* exceeds 3. (see Figure 2) Tabular representations of "killer-victim scoreboards” are not uncommonly seen as insufficient to produce all of the insights that are necessary.

In addition to the data that is collected from running analytical simulations, there is also a huge amount of data that is or could be collected from simulations run for training purposes. This issue of the appropriate and improved use of the data collected for purposes other than its analytical content is treated in more detail later.

Future Needs of Analysts

As has been shown by advancing a few examples, the analysts of today are faced with problems for which current technology implementations are not adequate. While analysts have done heroic duty in developing work-arounds for these limitations, the degree to which they are missing important insights remains unquantified, but disturbing. By analogy, the physical sciences found themselves in a very similar position early in the days of simulations on computers and the subsequent analysis of physical phenomena. Their experience suggests that the current practice of reducing resolution and geographical scope of FMS scenarios is leading in the wrong direction. It may be robbing the analysts of insights that could be extracted from more sophisticated and detailed models running on larger terrain databases.

The FMS community is also increasingly experiencing pressure to simulate some of the more complex forms of human behavior. One area of great interest in military science is the actual role of the individual soldier and commander (Ben-Ari, 1998). More computing power will be required to be able to deliver analytical and training platforms that can model emotion in a useful way (Garlan, 1993).

While FMS analysts now have access to much more computing power than they did a few decades ago, there is evidence that significant additional capacity could be implemented, with a high degree of confidence in its expanded utility, reliability, and stability. It is neither necessary nor desirable to move into the future using untested technology nor is it wise to duplicate already available, useful programs. Others have broken the ground that the FMS community can now till.

HIGH PERFORMANCE COMPUTING

Originally, computer scientists considered that the only way to speed up the process was to accelerate the CPU. At a certain point, however, it seemed obvious that the technology would not be able to keep improving processors to do calculations faster by increasing clock speed. They would similarly not be able to continue to make each clock cycle more effective by adding functions to the processor (Moore, 1965).

That led to theorizing about parallel computation and harnessing more than one computer to work on the same task. A generalized theory of parallel processing effectiveness was advanced by Amdahl, in which he carefully described the speed-up that one would expect (Amdahl, 1967). Starting with work at Caltech on the Intel Delta, with 512 64-bit processors, increasing numbers of simulations have been effectively parallelized for the big machines, with high-speed inter-node communications fabrics. As this size was orders of magnitude greater than the early limits theorized, this class was often referred to a Massive Parallel Processors (MPPs).

Table 1. *World’s fastest supercomputers***.**

Rank	Site Country/Year	Computer / Processors Manufacturer	R_{max} R_{peak}
1	**Earth Simulator** Japan/2002	**Earth-Simulator** / 5120 NEC	35860 40960
2	**Los Alamos National Lab** U.S./2002	**ASCI Q - AlphaServer, 1.25 GHz** / 8192 HP	13880 20480
3	**Virginia Tech** U.S./2003	**1100 Dual 2.0 GHz G5/ Infiniband/GigE**/ 2200 Self-made	10280 17600
4	**NCSA** U.S./2003	**P4 Xeon 3.06 GHz, Myrinet** / 2500 Dell	9819 15300
5	**Pacific NW National Lab** U.S./2003	**Integrity Itanium2 1.5 GHz, Quadrics** / 1936 HP	8633 11616
6	**Los Alamos National Lab** U.S./2003	**Opteron 2 GHz, Myrinet** / 2816 Linux Networx	8051 11264
7	**Lr. Livermore National Lab** U.S./2002	**MCR Linux Cluster Xeon 2.4GHz, Quadrics** / 2304 Linux Networx/Quadrics	7634 11060
8	**Lr. Livermore National Lab** U.S./2000	**ASCI White, SP Power3 375 MHz** / 8192 IBM	7304 12288
9	**NERSC/LBNL** U.S./2002	**SP Power3 375 MHz 16 way**/ 6656 IBM	7304 9984
10	**Lr. Livermore National Lab** U.S./2003	**xSeries Cluster, Xeon 2.4GHz, Quadrics**/ 1920 IBM/Quadrics	6586 9216

The Top 500 Supercomputers list presents rankings in order of performance using LINPAC, a common benchmark. The top ten of the list for 2003 is reproduced above in Table 1. The number of processors follows the name (van der Steen, 2003). Note that the least amongst these has 1,920 processors and that the biggest, but not most powerful, has 8,192.

This list covers only the supercomputers that are engaged in work that can be publicly acknowledged.

Linux Clusters: The Beowulf Concept

This last cost-based definition brings us to the next important concept: commodity clusters, or Beowulfs. Dr. Thomas Sterling propounded and popularized large-scale parallel computing through the use of cheap commodity components: CPUs, power supplies, RAM, internode communications, operating systems and software (Sterling, 1999). By taking best advantage of the cost benefits of mass production, he collected and organized mass numbers of commodity processors. These are usually Intel architecture PCs, with communications between them using gangs of low-cost Ethernet switches (or the more expensive but more powerful cluster communications switches).

Figure 3 *The IBM Linux cluster at the Maui High Performance Computing Center.*

Beowulfs typically use the largely free operating systems like Linux and the GNU series of compilers (see Figure 3). For internode communications programming, there are a number of languages following the Message Passing Interface (MPI) standard. These may be obtained without paying the expensive license fees that are typical with some of the proprietary supercomputers.

The Beowulf technology is not the most effective one for computing that requires exceptionally high-speed serial computation, exceptional floating-point power, or exceptionally low latencies for their internode communications. Clusters are fortunately very useful for most programs. Unlike the Cray series that were very high-speed vector machines and required liquid cooling, the Beowulf series are now universally air cooled, requiring only sufficient machine room cooling to remove the heat from the amassed processors. The avoidance of the efficient, but expensive, CPU/liquid-coolant interface is an incredible cost savings. A typical price for a significantly sized cluster is on the order of a few thousand dollars per node.

Grid Computing

If processors can be amalgamated to produce more power locally, then there could be even more power made available if remote computers could be similarly connected to provide additional processors. The previously mentioned concept of scalability clearly comes into play and one new concept must be considered. Most of the clusters and supercomputers discussed so far have been homogeneous, *i.e.* all of the processors are the same. If grid computing entails using clusters and processors from different sites, then the likelihood of homogeneity falls rapidly. Fortunately, Beowulfs have been remarkably tolerant of heterogeneity and data will be later adduced to show the capabilities of grids made up of the Beowulf Linux clusters and the proprietary supercomputers.

Grid computing usually conveys the concept of using a Wide Area Network (WAN), frequently the Internet itself, to connect remotely located SPPs, both supercomputers and Linux clusters. The landmark work on this innovation was done by Ian Foster and Carl Kesselman (Foster, 1997). In order for all of these diverse and dispersed assets to be useful, there must be methods of coordinating, initiating, and controlling them. The tool developed by Foster and Kesselman is called Globus and is generally recognized as a very effective way to approach this type of distributed high performance computing.

Another, more localized version of this concept, is that of using all of the idle PCs on an organization’s LAN. This involves running processes on the various PCs making those processors available to the central user when they are not in use by the PCs “owner.” When the owner interfaces with his computer in any way, it immediately suspends the remote process and redelivers control to the owner. One popular program providing this service is Condor (Litzkow, 1998.) This technique is a natural choice for a type of computing that does not need cycles on demand. One nationally distributed use of this concept is doing signals processing as part of the search for extraterrestrial intelligence.

Parallel Data Handling

The two foci of this paper are parallel computing and advanced data handling techniques in FMS. While much of the issue of data handling is appropriate for the section on data analysis, some portion of it is more closely tied to parallel processing. Parallel distributed processing both enhances and encumbers data collection, storage and retrieval. One of the major daily uses of high performance computing is the rapid processing of huge masses of transactional data by retailers and financial institutions, an indication of its value in this arena.

Like parallel processing, there is an extensive experience base in parallel data handling. At the Information Sciences Institute, a distributed data system has been developed to support the SAF simulations. Client applications communicate using RTI routines and data that is identified is stored on local disks at each node. A central aggregator acts to query the tasks, when desirable, and collects all information at the end of the simulation. The data content is then analyzed and archived. As it is new technology, the techniques for maximizing the utility of the parallel capabilities are not universally practiced, but the expertise is easily accessed.

Figure 4. *Tertiary storage tape silos at USC.*

ADVANCED DATA ANALYSIS

As simulations have moved from the gym floor to the computer, a similar change has taken place in the means of assessing the results of the exercises. When the SAFs were first used to train tank crews, the most important factor was face validity. As long as the tanker trainee perceived the representations as realistic, the simulation was considered to be a success. Now that the community has moved from a few vehicles to more than one million vehicles, the need for a more elaborate approach has become clear. Policymakers and leaders of the simulation community now seek new ways to exploit the data being collected. (Dubik, 2003).

Additionally, this country no longer has the luxury enjoyed in past wars of taking months to mobilize technology for defense efforts, and learning from early combat experience to hone later tactics. Today's battlefield is much more technologically loaded, complex and fast-paced (Cebrowski, 2000). It follows that there is a need for more complex, faithful and illuminating simulations of future battlespaces. The insights needed must be more timely and of greater specificity, in order to defend against new foes who are less identifiable, less predictable and more capable of attacking asymmetrically.

One of the first issues of concern is defining just what the simulation community and government leaders should and can extract from the simulations. Rather than considering this issue *de novo*, much can be learned from the Operations Research approach (Kleijnen, 2001). Many of their techniques have already been implemented on SPPs and their rigorous analysis of critical parameters is very useful.

Advances in Database Technology

As data sets have grown exponentially larger and more complex, so also has the technology grown to query that data and return useful and timely result sets. While the expenditures of the DoD are not insignificant in this field, much of the productive innovation is being delivered out of the commercial database market and much of the intellectual leadership resides on the campuses of the U.S.'s top research universities. Search engines such as the currently pre-eminent "Google" respond rapidly and accurately in non-rigorous, but demanding, civilian situations. The military analyst, while retaining the same needs for excellent interface, speed, accuracy, relevancy and scope, also requires a greater assurance that data ascertained represents sufficient, and accurate results of relevant materials. The high performance computing centers provide a common ground where these diverse database professionals meet. Synthesizing the advances from all of these disparate fields arguably provides the synergy necessary to meet the rapidly expanding needs of the FMS community.

Data Mining

Data mining techniques are defined here as the extraction of useful patterns and modes from data sets that are often large. More particularly we, and others, use the term to specially imply the extraction of insights from databases for which that data structure was not originally designed.

Some authors have described data mining as lying at the intersection of statistics, machine learning, data management, pattern recognition, artificial intelligence and other related disciplines. The authors see it as the application of myriad techniques to accomplish its goals, but not subsuming all of these techniques into itself. Its focus on "…unsuspected relationships ..." and summarizing data in "… novel ways that are both understandable and useful …" (Hand, 2002) is the capability that is seen as most promising for FMS data analysis.

Data mining can be more generally said to require some significant effort in each of the following tasks:

1. initial data analysis to gain understanding of organization and visualization possibilities
2. an attempt at describing a loose-fitting, but acceptable model of the data under analysis
3. the creation of a model capable of predicting the results and the relationship of those results to certain input parameters
4. the final analysis of the data sets with the final product being not only the discovered relationships, but also the real-world insights that such relationships support

Simulations of the order discussed in this paper may generate as many as 1,600,000,000 data points. With such vast amounts of data, not all useful analysis can be done real-time, nor is it optimally productive to do so. The common result is that reams of recorded data sets are discarded as too cumbersome to be of analytical use. Data mining tools offer promise in that they allow the analyst to find useful information and patterns amidst the mass of data points even after the simulation is completed

With the power of scalable parallel processor supercomputers, once simulation results have been characterized and values ascribed to various outcomes, the recursive analysis of the data will undoubtedly find useful new views of that which critical to the outcome. For example, using data from numerous iterations of a flight simulation designed solely for training, one might find a pattern of inexperienced pilots tending to overshoot their targets. Without making the effort to analyze such data or to create effective tools for sifting through the vast amount of noise to find useful information, the opportunity to discover such useful patterns is lost. Data mining tools help to isolate not just the story from the activity, but the wisdom to be gained there from.

The data mining process does require efforts beyond traditional simulation analysis. Normally data mining requires all or some of the following:

1. Achieving a thorough understanding of the representations' inherent characteristics and organization (*e.g.*, parameters of the entities, descriptions of their activities).
2. Selecting methods of defining and comparing the data in such a way that it will yield quantifiable results that can be compared (*e.g.*, losses, mission success, time).
3. Discovering, defining and applying an algorithm to compare results with input parameters (multi-variate studies of data sets).
4. Analyzing and implementing those data management techniques that will enable and facilitate steps one through three.

The tasks above need not make demands on the structure of the data nor the means for attaining it. By its very nature, data mining presents low cost opportunities for gains in insight and understanding from simulations of almost any sort with little to no impact on or cost to the simulation itself.

Data mining has historically proven to be an effective tool in numerous fields. Trigon Blue Cross Blue Shield uses data mining techniques to identify early indicators of serious disease, thus allowing them to effectively treat patient before they become seriously ill. Data mining methods helped retailer Williams-Sonoma save millions in advertising costs without losses in sales by creating a targeted system for catalog distribution. Using data mining technologies, banks have developed better credit scoring models that more accurately predict applicants that may default on loans. In science, data mining techniques have been used to identify new binary stars by using radio telescope data collected for mapping, but which serendipitously contain the characteristic oscillation frequencies of such stars yet undiscovered (Moore, 1998). Similarly, the authors believe that unsuspected insights that will save lives, money and missions lie deeply imbedded in the data being generated today by FMS.

As increasingly complex simulations produce larger and larger datasets, data mining techniques will help the analyst sift through that mountain of data in order to find and quantify useful relationships and patterns. Increased computing power and faster computational capabilities only increase the opportunity to find useful patterns. While the authors do not represent that data mining will solve all problems nor discover all relationships of interest, they do accept the notion that it has the potential of discovering many new relationships, some of which may enable significant new capabilities or prevent monumental losses.

Evolutionary Computing

Another area of significant opportunity lies in the application of the techniques described by the Fogels in their work on Evolutionary Computation. (Fogel, 2000) Many of the new battlefield challenges represented by the relationships of the data described above are far removed from the current understanding of defense strategies. They will not be observed, presumed or described by even the most rigorous analysis of the data. Novel and asymmetric threats are continually and rapidly evolving. These new threats are being driven by groups whose one remaining effective weapon may be their tactical innovation and the resultant element of surprise. In this they are aided and abetted by their remoteness from the defense analysts in topics such as their value system, goals, training, and *zeitgeist*.

The family Fogel presents a way to examine a virtually unlimited horizon of possibilities by using techniques that perturb the physically accurate simulations of the world without regard to the constraints of the expectation or creativity. They replace the rule-based foundation of the Monte Carlo simulations with the concept of an entity that is able to freely roam the range of possibilities, with an appropriate feedback loop to help in optimizing the path to the goal. Basing their work on the areas of artificial intelligence, expert systems and neural net training, the evolutionary computer scientists further look to the biological paradigms popularized by Charles Darwin in his work on the evolution of animals. This group eschews slavish imposition of genetic rules and prefers to let electronic intelligence finds its own path in parallel with biologic evolution.

Applying the concepts laid out by these evolutionary computational scientists has the promise of establishing unimagined methods and threats. Should the evolutionary computing process result in the identification of an unnoticed vulnerability or the determination of a new threat, the defenses could be altered, steps could be taken to ameliorate the losses, or contravening punitive actions aimed at the attackers could be imposed.

Genetic Algorithms

In a variant of the work by the Fogels, David Goldberg reports significant success in applying more stringently biologic rules to his analysis (Goldberg, 2002). He sees the genetic evolutionary driver as having been tested over the millennia and therefore not likely to be deficient. His application of genetic rules is similarly successful in the test phases of his work. He feels the insights he gains are more likely to be in accord with the behaviors observed in actual life. Dr. Goldberg has used his techniques to model both organizational entities such as small populations and physical phenomena such as gas pipelines. His approach does suggest a very supportable relationship between his data and the observed data in the population under study and the pipeline under observation.

The selection between evolutionary computing and genetic algorithms can be left to the user as an exercise. In each case, the identification of a novel concept would have to sustain the challenge of reason and the governmental vetting process prior to funding a new defense or the acceptance of a new approach. The *caveat* to be remembered is not to disregard novel approaches and valid insights.

Monte Carlo Analyses

Many of the simulations in use by the services today rely heavily upon Monte Carlo techniques. These simulations have a pre-established rule set and distribution or likelihood for each major activity as was described above. As noted earlier, these simulations are not deterministic and often the same basic initial definition is executed several times (hundreds of runs are not uncommon) to examine the distribution of the final outcomes, (Horne, 1999). This work is often analyzed by plotting out a series of two dimensional solution spaces on a three dimensional graph, as in Figure 2, and visually identifying the optima and their relation to one another for each pair and then estimating the interrelation of the multivariate group.

Based on the work of a physicist at Caltech, the OTCI company has developed a tool that can quantify the degree to which the input parameters affect the final outcome. This can be done in *n* dimensions, which would be an improvement on the visual analytical procedure outlined above. Further, this procedure yields very interesting results in fewer runs, sometimes orders of magnitude fewer (Johnson, 1999). The technology is currently implemented for financial analyses, but could be "ported" over to battlefield simulation analyses with a high expectation of efficacy and a reasonable hope for better analytical products.

.

IMPLEMENTATION EXAMPLES

There are both examples of successes and of well-documented plans for how the techniques reviewed above can be implemented in the future, including three easily envisioned ways to approach the scalable parallelization of a simulation. First, design the code as a well-parallelized program from the beginning. Second, after reviewing existing code, completely rewrite an existing code base in a scalable parallel manner. Third, take the code as it is and implement a new "wrapper" around the code that makes it scalable. Two of the noted implementations have been seen in the intelligent agent, non-deterministic variety of simulations: SF Express/JESPP and Project Albert.

In SF Express, the ModSAF code was enhanced with communications routers written in Message Passing Interface (MPI). The routers enabled scalability both within the SPP mesh and across the nation. This was a successful example of distributed, heterogeneous supercomputing. Scalability was measured in comparing the times experienced in communications activity as the size of the sample increased.

In JESPP, the team was asked to make the JFCOM's JSAF simulation much more scalable (see Figure 5) and portable to Linux clusters of the 256 node class. JFCOM needed to "field" more than a million vehicles in an urban setting.

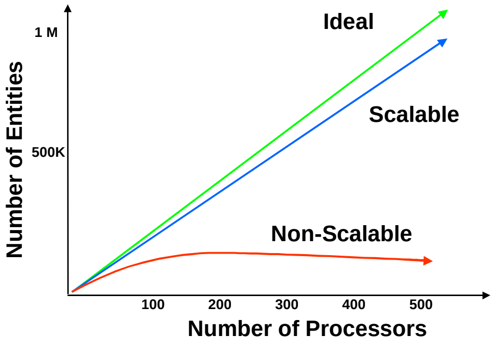

Figure 5. *Notional Scalability using Mesh Routers*

This was necessary to test many concepts, including the need to assess various simulated sensor platforms and associated systems in their ability to discriminate combatants from non-combatants.

Previous implementations on LANs did not simulate more than 30K vehicles. In making the code more scalable and running it on a series of Linux clusters, the JESPP team was able to achieve more than 1,000,000 vehicles (Lucas, 2003 and Helfinstine, 2003). The approach they used is worth future consideration. It entailed the careful study of the simulation code in use, JSAF, and then constructing a system of very scalable software routers to make the code capable of effectively using the hundreds of processors in large Linux clusters. The code base itself was not significantly impacted and the software routers were designed to accommodate as best they could the almost daily changes in JFCOM's growing and dynamic needs, which caused frequent modifications to the underlying JSAF code. To achieve scalability, the effort engaged computational scientists with extensive backgrounds in physical science simulations who had the intellectual and creative skills to rapidly understand and to effectively enhance the code.

Project Albert has taken a different track. From the beginning, the code was constructed with a kind of built-in scalability. The basic idea of Brandstein and Horne was that Albert would not have a fully fleshed out simulation, but would be convey the "essence" of the activity running in a very small module that can be run over and over. The Project Albert crew has worked very closely with the parallel-computing experts at the Maui High Performance Computing Center. The code base is kept quite small by design. It has less than ten per cent of the lines of the JSAF code.

While this *de novo* approach has the benefits of elegance and simplicity, it is, by definition, only open to the developers who are producing entirely new programs. The authors find that the FMS community frequently adapts and expands existing code and the one recent major new, "bottoms-up" FMS program was recently terminated. Nonetheless, developers of new systems should resist being seduced by the ease of single-processor designs, as experience has now shown that there will be pressure to expand along the dimensions of complexity, resolution, and magnitude, hence requiring or benefiting from parallel processing.

In looking at appropriate data analysis techniques to implement, one issue that must be settled is that of the users' goals in this analytical process. As this seems rather intuitive, there is an inclination to skip, or at least slight, this step. Previous scholarship bears careful attention. Going outside the narrow confines of FMS, it is clear that Operations Researchers (OR) have studied comparable issues for a century. Some members of the OR community have been active in helping simulation groups understand their goals and analyze their results (Sanchez, 2002).

Ongoing Work

Many of these programs are ongoing. JFCOM has committed to the JESPP project for the foreseeable future. The award of the Distributed Center cluster to JFCOM will provide a natural home for the technical life of the cluster. Additional clusters can be enlisted to provide more processing power, and the expandable capabilities of the JESPP scalable routers produce the ability for the JSAF code to utilize all of the processors that can be gathered for an exercise. An extension of this work into the analysis of homeland security issues for air traffic control has been advanced and is considered a well-founded application of the concept of taking code as is and enhancing it with augmenting wrappers.

The Albert project continues to be a vital part of the study of maneuver warfare. In addition to studying new ways of utilizing large parallel computers, Albert seeks out new ways to analyze the huge amounts of data presented by the multiple run method. As the project increases the number of important variables, the difficulty in visualizing multi-dimensional solution spaces may find resolution in the work of Monte Carlo simulators in the financial community.

This generates an analytically valuable data surfeit that may be found in the vast quantities of data that could be collected from training simulations. As pilots, sailors and tankers train in their simulators, their activities' data make a fertile field for other analyses.

Adopting either of the two methods discussed above, designing for scalability or augmenting with scalable wrappers, should produce several benefits. First, both should create not only early scalability but imbue the code with an ability to scale further and make early use of new processor technologies. Second, experience has shown that the parallelization process itself frequently improves the serial code and, not infrequently, leads to insights into the subject phenomena. Third, careful application of these techniques should not disturb the development or use of the delivered code. Fourth, cost, schedule and performance can be kept in balance.

There are actions that will reduce potential disruptions and produce the best results. The most important of these may seem obvious, but it is not infrequently overlooked - the reliance on experienced parallel computational scientists. Parallel programming is a unique skill-set. Attempts to automate the process of parallelizing code have not been particularly fruitful, especially in programs where coarse-grained parallelism is appropriate. A research and development group seeking to make their code scalable would be well advised to identify a successful effort implementing comparable code on an SPP and then engage the parallel programmers who were responsible and who have exhibited a transferable aptitude.

TECHNOLOGY IMPACT

It is the firm conviction of the authors that the technology detailed above will prove to be a vital asset for the FMS community and then have an essential impact on the defense of the nation. The necessity of dealing with the commingling of combatants and non-combatants, the current mandates to conduct operations with minimal disruption of civilian infrastructure, and the ability to wage effective warfare against an asynchronous enemy all will be addressed more completely using the advanced techniques discussed.

However, the FMS community has not shown as much acceptance of these technologies as might have been expected. At the 2003 IITSEC meetings, only the JFCOM papers on the use of Linux Clusters evidenced implementations in everyday use (Lucas, 2003; Helfinstine, 2003; and Williams, 2003). The three other papers mentioning these topics (Pratt, 2003; Schiavone, 2003; and Mielke, 2003) took valid, but much more theoretical, perspectives. This year may not show much of an increase. A review of the submitted titles for IITSEC 2004 reveals the lack of a single mention of the terms Beowulf, supercomputer, parallel processing, data mining, evolutionary computing, sensitivity analysis or high performance.

The ability of the analyst to distinguish between non-combatants and enemy forces hiding among them relies on increasingly effective sensors, well-designed analytical systems, and advanced training in realistic environments. Current limitations in resolution, entity count and sophistication of behavior interfere with all of these. Simulation experimenters report that analysts engaged in early exercises had so few civilian entities in their environment that they were inclined to opt for destruction of all vehicles under observation, when there was doubt as to their identity. Reasonable choices were also restricted when the number of civilians was smaller than the number of enemy combatants, a condition driven by the lack of compute capacity on platforms consisting of PCs on LANs.

The lack of sophistication also can render an exercise less meaningful. Human operators are very sensitive to behavior differences. If computer constraints enforce very simplistic behaviors on modeled civilian vehicles, the operators quickly can distinguish them from the more complex behavior capabilities of the combatant vehicles, *e.g*. if the simulation controllers turn off collision avoidance to save on inter-visibility calculations, the operators will quickly perceive that any vehicle that passes right through another is not a combatant. Neither good training nor good analytical input can result from similarly constrained conditions.

Additionally, the not uncommon reliance on SME reviews of simulations, while effective and useful, may be missing valuable insights. These insights might otherwise lead to new strategic concepts or prevent overlooking significant vulnerabilities. Not yet having faced the unknown enemy of the future, not knowing its mind-set, and not having the luxury of learning at a leisurely pace, the simulation community would be well-advised to take advantage of the expanded capabilities presented above in the section on advanced data analysis techniques.

Orderly retrieval of information using the latest database techniques will assist human analysts in pursuing intuitive leads. The innovative techniques representing data mining can be invoked to extract even more esoteric concepts and bring these to the attention of the analysts for confirmation and analysis. This gives real hope for identifying asymmetric tactics that might not be foretold by traditional military analysis. The concepts of evolutionary computation, genetic algorithms, and Monte Carlo sensitivity analyses also show promise in making sure nothing is missed in the search for security.

A Development Path: Successful Rapid Prototyping

Transitions from current simulation methods to full exploitation of present and near-term computational capabilities and practices take effort and significant experimentation. It is perhaps best to illustrate the process with a particular example: a suite of large field-of-view sensors attempting to detect isolated "suspicious" behaviors within a large population of normal (i.e., "civilian") entities. This has been, in fact, a major thrust of ongoing JSAF developments, with the Simulation of the Locations and Attack of Mobile Enemy Missiles (SLAMEM™) surveillance/tracking software system fed by detections from simulated civilian vehicles (euphemistically called "clutter") within the JSAF simulation

.

The JSAF/SLAMEM combination has so far been rather fruitful. For present purposes, it is sufficient to consider three particular items:

1. In order to support large numbers of clutter entities, the clutter models within JSAF had to be quite simplistic, with, *e.g*., very little "self-awareness" among clutter entities.
2. While the SLAMEM-JSAF system exploits a number of clever procedures to distribute much of the computational burden (in particular, some of the simulated signal processing), the tracking and situation-assessment procedures within SLAMEM were originally done on a single processor, thus providing a significant constraint on the size of the underlying simulated scenario.
3. The very large numbers of simulated detections within a typical SLAMEM-JSAF were largely "unexploited", beyond the immediate task of driving track formation and feeding operator displays.

There are a number of straightforward technology "patches" for many of these problems, including parallelized tracking algorithms, a much richer, distributed database system supporting data mining and "discovery" activities. Incremental developments along these paths are inevitable. The problem, of course, comes with the word "incremental." The standard practice of inserting pieces of computational technology, as though one were simply using higher clock rate processors, drives the system along a path dictated by "ease of insertion" rather than ultimate end-user needs. In the authors' opinion, it is definitely progress, but it is unlikely to be progress that will ever catch up to available capabilities.

Consider, again, the conceptual SLAMEM problem. At a high-enough level, the outcome of present experimentation must point to the desired or idealized product: Operators are watching displays of highly processed tracking results, looking for indications of both "suspicious behavior" and reactions to interdiction activities. Human interpretation of these data will always be subjective. The details that could be provided by scaled computational power alone are overwhelming, if not overwhelmingly useless. Operators needed dynamic access to the available data at several scales of both "resolution" and "historicity" in order to assign likelihoods to the important bottom-line issues of asymmetrical combat.

It can be argued that the ongoing FMS development path would not reach this goal (or rather, if it does, it will do so very, very slowly). This is not to say that standard practices should be abandoned! The authors maintain that incremental development and implementation is the only sane way of improving the state of the art while maintaining capabilities.

The authors suggest a parallel development track is needed, emphasizing the "top-down" approach with the goals of identifying: 1) inherent limitations in standard practices and 2) technology needed to resolve the identified problems. The intent of the second point must be clarified/emphasized. Rather than ask the implicit "standard practices" question ("What incremental capabilities can be added through readily available technology?"), a very different question must be asked for optimal implementation: "What technology is needed to achieve required capabilities?" Viewed from the perspective of the idealized ultimate user, the system to which JSAF/SLAMEM activities are pointing must be database driven and must address the following questions: "What information will best aid the decision maker?" and "What automated discovering and mining procedures are needed to make this data perceivable to the decision maker?"

CONCLUSION

September 11, 2001, exemplified the penalty for not adequately preparing for unexpected threats. While the techniques proposed in this paper may not have averted that tragedy, the authors maintain those techniques will increase the opportunity for the analysts to discover such threats and work out the best way to defend against such destruction. Considering the huge losses that the nation incurred from that one attack, the efforts required in implementing the described techniques pale in comparison.

Data interpretation is the critical task in the war on terrorism. Simulation systems, especially simulation systems for training, must be based on this cornerstone and must provide a real-time laboratory for refining and exploiting advances in data mining over the last decade.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the members of the ISI JESPP team who have contributed to this paper through their efforts and their intellectual stimulation. Much of the success reported here came from the Joint Experimentation on Scalable Parallel Processor project, initiated, directed and funded by the Joint Forces Command and to a very large degree conducted on the compute assets of the Maui High performance Computing Center and other members of the High Performance Computing Modernization Program. Without the support and encouragement from all of the above, none of this would have been possible.

REFERENCES

Amdahl, G.M. (1967). Validity of the single-processor approach to achieving large scale computing capabilities. *AFIPS Conference Proceedings*, 30, 483-485.

Ben-Ari, E. (1998). Mastering Soldiers: Conflict, Emotions and the Enemy in an Israeli Military Unit. *New Directions in Anthropology,* 10. Oxford: Berghahn Books.

Brandstein, A. (1998). Data Farming: A Meta-technique for Research in the 21st Century. *Maneuver Warfare Science* Quantico, VA: United States Marine Corps Combat Develop Command, 93-99.

Brunett, S., Davis, D., Gottschalk, T., & Messina, P. (1998) Implementing Distributed Synthetic Forces Simulations in Metacomputing Environments, *The Seventh Heterogeneous Computing Workshop*, Orlando, FL.

Cebrowski, A., & Garstka, J. (1998) Network Centric Warfare: Its Origin and Future, *Naval Institute Proceedings*, 124(1), 28-35.

Ceranowicz, A., Torpey, M., Hellfinstine, W., Evans, J. & Hines, J. (2002) Reflections on Building the Joint Experimental Federation. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, Orlando, FL.

Dubik, J. (2003) Comments made by Major General James M. Dubik, J9, US Joint Forces Command, at I/ITSEC in Orlando, FL.

Foster, I. & Kesselman C. (1997) Globus: A Metacomputing Infrastructure Toolkit. *Intl J. Supercomputer Applications*, 11(2), 115 –128.

Fogel, D. (1995) *Evolutionary Computation*. New York: IEEE Press.

Gibson, D. (2003) Casualty Estimation in Modern Warfare. *Army Logistician*, Nov-Dec 2003 35(6), 34. Fort Lee, Virginia.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Boston, Massachusetts: Addison-Wesley Publishing Co.

Garlan, D. & Shaw. M. (1993) *An Intro to Software Architecture: Advances in Software and Knowledge Engineering*. World Scientific Publishing, I.

Graebener, R., Rafuse, G., Miller, R. & Yao, K-T. (2003) The Road to Successful Joint Experimentation Starts at the Data Collection Trail. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, Orlando, FL.

Hand, D., Mannila, H., Smyth,. P. (2001). *Principles of Data Mining (Adaptive Computation and Machine Learning).* Boston, MA: MIT Press.

Harnad, S. (1989) Minds, Machines and Searle. *Journal of Theoretical and Experimental Artificial Intelligence*.1, 5-25.

Helfinstine, W. Torpey, M. & Wagenbreth, G. (2003) Experimental Interest Management Architecture for DCEE. *Proceedings of the Interservice / Industry Training, Simulation and Education Conference*, Orlando, FL.

Hill, R. W., Gratch, J., & Rosenbloom, P.S. (2000). Flexible Group Behavior; Virtual Commanders for Synthetic Battlespaces. *Proceedings of the Fourth International Conference on Autonomous Agents*, Barcelona, Spain.

HPCMP (2004) Computational Technology Areas. Forces Modeling and Simulation CTA, (Online)May 13, 2004, http://www.hpcmo.hpc.mil/Htdocs/CHSSI/cta_description.html

Horne, G. (1999) Maneuver Warfare Distillations: Essence not Verisimilitude. *Proceedings of the 1999 Winter Simulation Conference. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans*, Phoenix, AZ

Hull, J. (2002). *Options, Futures and Other Depravities, 5th Ed.* New York: Prentice Hall.

Johnson, E., Bergman, L. & Spencer, B. (1999). Intelligent Monte Carlo Simulation and Discrepancy Sensitivity. In P.D. Spanos (ed.), *Computational Stochastic Mechanics (pp. 31-39).* Balkema, Rotterdam.

Kevrekidis, I., Gear, C., Hyman, J., Kevrekidis, P., Runborg, O. and Theodoropoulos, D. (2003) Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*. 1, 715-762.

Kleijnen, J., Sanchez, S., Lucas, T., & Cioppa, T., in print, A User's Guide to the Brave New World of Designing Simulation Experiments, under revision for *INFORMS Journal on Computing*.

Krishnaprasad, S. (2001) Uses and Abuses of Amdahl's Law. *The Journal of Computing in Small Colleges* 17(), 288–293.

Litzkow, M., Livny, M. & Mutka, M. (1988) Condor - A Hunter of Idle Workstations. *Proceedings of the 8th International Conference of Distributed Computing Systems*, 104-111.

Lucas, R. & Davis, D. (2003) Joint Experimentation in Scalable Parallel Processors. *Proceedings of the Interservice / Industry Training, Simulation and Education Conference*, Orlando, FL.

Marshall, S.L.A. (2000). *Men Against Fire: The Problem of Battle Command*, Norman, Oklahoma: University of Oklahoma Press.

Messina, P. C., Brunett, S., Davis, D. M., Gottschalk, T. D. (1997, April) Distributed Interactive Simulation for Synthetic Forces, In J. Antonio, (Chair), *Mapping and Scheduling Systems, International Parallel Processing Symposium*, Geneva, Switzerland.

Mielke, R. & Phillips, M. (2003) Development and Application of an Academic Battle Lab. *Proceedings of the Interservice / Industry Training, Simulation and Education Conference*, Orlando, FL

Moore, G. (1965) Cramming more components onto integrated circuits. *Electronics* 38 (8), 114-117

Moore, R., Prince, T., Ellisman, M. *(1998)* Data-Intensive Computing and Digital Libraries. *Communications of the ACM*.

Pratt, D. & Henningger, A. (2003) Load Balancing for Distributed Battlefield Simulations: Initial Results. *Proceedings of the Interservice / Industry Training, Simulation and Education Conference*. Orlando, FL

Sanchez, S.M. and Lucas, T.W. (2002) Agent-based Simulations: Simple Models, Complex Analyses, Invited paper. *Proceedings of the 2002 Winter Simulation Conference*. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sanne, J. (1999). *Creating Safety in Air Traffic Control*. Unpublished doctoral dissertation, Institute of Tema Research, Linköping University, S-581 83 Linköping, Sweden.

Schiavone, G. Dolezal, M., Tracy, J., Secretan, J., & Mangold, L. (2003) Beowulf Supercomputing for Mobile Applications. *Proceedings of the Interservice / Industry Training, Simulation and Education Conference*. Orlando, FL

Sterling, T., Salmon, J., Becker, D. & Savarrese, D. (1999) *How to Build a Beowulf*. Boston, MA: MIT Press.

USNews. (1988, August 15) *Was stress the villain in the air tragedy?* U.S. News and World Report, 105(7), 13

Van der Steen, A., & Dongarra, J. (2003) Overview of Recent Supercomputers. *Top 500 Supercomputer Sites*. (Online) February 12, 2003, http://www.top500.org/ORSC/2003/

van Lent, M. & Laird, K. (1998). Learning by Observation in a Complex Domain. *Proceedings of the Knowledge Acquisition Workshop*. Banff, Canada.