An Experimental Testbed for Battle Planning

Barry Bodt, Joan Forester, Charles Hansen, Eric Heilman, Richard Kaste, Janet O’May
Simulation Concepts Branch
U. S. Army Research Laboratory
Attn: AMSRL-IS-CD
Aberdeen Proving Ground, MD  21005-5067
(410) 278-7781
rck@arl.mil

Abstract

This paper describes a modular evaluation testbed for Army Course Of Action (COA) Analysis (COAA), with emphasis on research into assessing the feasibility of COAs developed using a simplistic wargaming mechanism. Central to the project is the use of a realistic combat simulation to produce detailed COA evaluation. Components of the testbed include the COA generator Fox-GA (developed under the auspices of ARL’s Federated Laboratory and based on genetic algorithm technology) and the evaluation platform Modular Semi-Automated Forces (ModSAF), a widely used modular combat simulation. Emphasis is placed on COA elements, transformation of Fox-GA COAs to ModSAF scenarios, experimental challenges, and statistical approaches to assessing the execution results.

1. Introduction

The Simulation Concepts Branch of the U. S. Army Research Laboratory (ARL) is responsible for creation of techniques and systems to assist battlespace commanders in the military decision making process. The primary focus of the ongoing Course of Action (COA) Analysis (COAA) project is merging operational reality with simulation to benefit the battle staff at division and below.

The project applies military planning and combat simulation software to the evaluation of automated COA generation tools. Envisioned is a testbed that will enable the assessment of COAs in a simulated operational environment. Initially, one COA generation tool, Fox-GA by the ARL Federated Laboratory, and one combat simulation, Modular Semi-Automated Forces (ModSAF) by Lockheed-Martin, comprise the prototype testbed. The project is exploring statistical analysis and experimental design techniques that might enable simulated exercises to be utilized as a part of COAA. The project is a work in progress; preliminary results are detailed in this paper.

2. Challenges

Currently, U. S. Army doctrine specifies planning and analysis methods that incorporate metrics based on feasibility, suitability, acceptability, and uniqueness. Contemporary planning, which is time-intensive, identifies a COA that can be used to complete given mission objectives,
1. REPORT DATE  
2000

2. REPORT TYPE

3. DATES COVERED
00-00-2000 to 00-00-2000

4. TITLE AND SUBTITLE
An Experimental Testbed for Battle Planning

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Army Research Laboratory, ATTN: AMSRL-IS-C, Aberdeen Proving Ground, MD, 21005-5067

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:
   a. REPORT unclassified
   b. ABSTRACT unclassified
   c. THIS PAGE unclassified

17. LIMITATION OF ABSTRACT

18. NUMBER OF PAGES 17

19a. NAME OF RESPONSIBLE PERSON

Standard Form 298 (Rev. 8-98)  
Prescribed by ANSI Std Z39-18
minimize friendly casualties, and position military forces to maintain the battlefield initiative during follow-on engagements.

Automated COA generation software and COAs generated by such tools are rarely field tested. It is understood that several elements from existing military metrics will be instrumental in development of a COA evaluation attribute set. Establishment of computer generated COA validity is a major project challenge. Fundamental elements of battle plan analysis such as the principles of war, current doctrine, the tenets of military leadership, and battlefield tactics, techniques and procedures are being considered. Tools that rapidly evaluate COAs with respect to doctrinal criteria, extract positive and negative aspects, identify interrelationships among planning factors, and develop measures of risk will improve planning techniques.

Reliable translation of attribute metrics into quantitative values describing COAs will enable a battlefield commander to quickly choose a COA that best fits his intent for the mission. Hence, application of experimental techniques that reasonably assign values to COAs is another project challenge. Rapid COA development requires abstracted computer algorithms designed to provide the commander an expanded selection of alternatives when planning a battle. Given a set of attribute metrics, various statistical methods can be used to evaluate COAs.

Enabling an automated environment capable of producing a number of COAs, against which a commander can compare mission objectives and battlefield intelligence, is another challenge. The prototype testbed must incorporate methods to better associate computer-generated COAs with reality while avoiding expensive field exercises. However, the limits of experimental methodology and digital combat simulations must be firmly established to ensure reasonable planning, especially in a hectic tactical operation center for a mission where lives are at stake.

The application of abstract processes to the detail found in reality has always been a challenge for the combat simulation community. Yet simulations enabling better understanding of the combat environment do exist, with improvements occurring daily. Harnessing the potential of simulations will increase combat power and help ensure retention of battlefield initiative through a reduction in planning time requirements.

3. Methodology

A competent battle staff that includes many subject matter experts can accomplish the military COA creation process within a short period of time. Normally a maximum of three COAs will be considered for any battlefield mission due to limitations of staff assets, including time. Automated COA generation has long been desired for easing the staff’s planning load and enabling the examination of many mission alternatives. However, automated generation efforts have been hampered due to the fluid nature of situational awareness requirements and the flexibility required for exploiting opportunities. COA augmentation decisions must sometimes be made within seconds of receiving new information.

ARL and its Federated Laboratory partners have produced software tools that assist experts in the creation of COAs. The initial COAA project focuses on one of these tools, Fox-GA, as part of a testbed (see Figure 1) designed to evaluate the applicability of automated COA generation to
actual battlefield requirements. (Other tools include the Combat Information Processor, which facilitates visualization of Fox-GA output, and OWL, a postprocessor that will be discussed later.) The COAA testbed contains four process parts with associated computer software, hardware, and data. These parts, described below, are: 1) Automated COA Generation, 2) Scenario Translation, 3) Experimentation, and 4) Statistical Evaluation.

Figure 1. Course of Action Analysis Testbed

3.1 Automated COA Generation

ARL, in conjunction with the Universities of Illinois and Minnesota, has created a software tool to produce COAs. The program, called Fox-GA (for Fox-Genetic Algorithm; hereafter referred to simply as Fox), is “an intelligent planning support tool designed to rapidly generate a variety of coarse-grained, high quality, friendly courses of action for military planners.”1 (See Appendix A for a discussion of Genetic Algorithms.)

The Fox user interface accepts up to six possible threat COAs, a friendly military unit organization structure, and parameters describing a friendly battlefield mission. Fox begins by evaluating several friendly COA base cases against each enemy COA using an internal

---

wargamer. The better-rated COA cases are then submitted to a genetic algorithm that may modify them to produce new cases. The amalgamated case set is then stripped of similar COAs and again subjected to the simulation for evaluation. After a number of iterations, Fox reports a set number (up to ten) of friendly COAs that are rated the highest.

Although developed as a proof of concept prototype and currently limited to a particular scenario and set of military units, Fox is a candidate for the COAA project. Moreover, the Fox developers, as members of the ARL Federated Laboratory consortium, are available for questions concerning developmental decisions.

In its final form, the rapid execution and flexible user interface Fox provides will enable the battle staff to consider a greater set of COAs for presentation at mission briefings. Fox also will enable the commander to examine the details used to generate a COA via a visual rendering of a combat simulation featuring units, routes, objectives and expected attrition portrayed on the mission terrain. The flexibility of using an entire COA or individual ideas from many COAs will allow the commander to consider an increased number of mission-planning directives.

3.2 Scenario Translation

Migration of Fox COAs to a different combat simulation requires a working knowledge of battlefield mechanics. The staples of a scenario, such as placement on the terrain, force structure, control measures (for example, movement routes and objectives), and battlefield doctrine, must be well understood in relation to a generated COA. Understanding the assumptions of the Fox programmers is an important consideration that requires both direct questioning and access to computer codes.

The version of Fox used in the initial effort contains a single scenario layout. The terrain is that of the U. S. Army National Training Center (NTC), Fort Irwin, CA. NTC terrain features a relatively open and hilly desert environment suitable for combat between brigade-sized military units. In the scenario, a friendly brigade deliberately attacks a threat force battalion.

The combat scenario parameters are highly abstract to enable rapid COA evaluation. Aggregate company-sized units are represented as numerical base strengths. Other factors affecting combat are either added to or multiplied by the unit base strength. Friendly companies are aligned along a maximum of three approach routes. Threat forces are positioned on “Lines of Defensible Terrain” (LDT), identified as strings of roughly adjacent narrow areas or choke points. Force attrition occurs at the intersection of “horizontal” approach routes and “vertical” LDTs. (See Figure 2)

ModSAF provides methodologies for creating and controlling military entities on a simulated battlefield. These entities can move, fire, sense, communicate, and react without operator

---

2 For example, the base strength of a friendly M2 mechanized infantry company is 3.0. If the company receives support, the factor is translated into a percentage of the force and added to the base: a 30% support factor brings the company to a strength of 3.9 via (3.0 x 0.3) + 3.0.

intervention. Each entity is given extensive capabilities. For example, a vehicle can drive over terrain avoiding obstacles, fire its weapon at opposing entities, use sensors to detect electronic warfare emissions, and send radio messages to other friendly entities while executing a mission. The goal of ModSAF is to replicate the outward behavior of simulated units and their component vehicle and weapon systems to a level of realism sufficient for training and combat development. ModSAF enables an improved representation of the relatively abstract Fox battles without the expense of actually fielding military units at the NTC. Reliance on the validated ModSAF simulation is intended to improve the value of battle results data.

Figure 2. Fox Combat Viewer

The Fox user interface was employed to ascertain the composition of units for Blue and Opposition forces, as shown in Figures 3 and 4 respectively. The Fox abstract force strength values were not directly applicable to ModSAF, which incorporates physical models and discrete attrition equations at the individual vehicle level. Fox provides for support values that can enhance a unit’s strength, but does not provide specific details to define the values in terms of combat support assets. The support values were not translated into the ModSAF scenario.

---

ModSAF movement routes were created using the LDTs and objective points detailed in Fox and drawing straight lines between the points. ModSAF units, consisting of company-level structures similar to those in the Fox scenario, were deployed with each battalion center of mass roughly equivalent to the position of the Fox aggregate battalion taskforce. Since the Fox scenario features a friendly attack mission, ModSAF friendly forces moved on the routes in a bounding overwatch attack posture. The threat company level units were moved into hasty defensive positions at locations gained from the Fox interface.

The terrain area used by Fox is currently not entirely available for the ModSAF scenario. Although a large terrain database was obtained from the U. S. Army Topographic Engineering Center, some of the Fox units are playing off of the terrain. This problem has yet to be solved;
possible options include moving all entities onto terrain or determining Fox’s actual reliance on terrain features.

Translation of the Fox scenario into a form usable by the ModSAF combat simulation was initially a manual process. The Fox combat viewer (see Figure 2) was used to establish the terrain coordinates of mission control measures and unit positions. (See Figure 5) All locations are expressed in degrees of latitude/longitude. The composition of each Fox friendly battalion taskforce (containing a number of company unit structures) and threat company was recorded. Mission control measures were plotted via the user interface on the NTC terrain representation within ModSAF. Taskforce units were reduced to company structures and placed with a center of mass analogous to the Fox unit positions. Movement routes included in ModSAF unit task orders were approximated given the Fox objectives and LDT positions. The ModSAF scenario layout was then saved for future experimentation. (See Figure 6)

![Figure 5. Fox to ModSAF Scenario Translation Values](image-url)
3.3 Experimentation

3.3.1 Platform Establishment

Initial efforts focused on implementation of an experimental COAA testbed. Version 1.0 of the Fox software was installed and used to create an initial scenario. The detailed combat simulation, ModSAF version 5.0, was installed on several systems, including four super computers (SGI Origin 2000s).

Translation and execution of the Fox scenario within ModSAF resulted in several observations and challenges for setting up the initial testbed. While the threat force structure was easily translated, the definition of friendly force structures took several iterations. The initial force consisted of 418 scenario entities, including dismounted infantry and vehicles, of which 358 were friendly forces. The existing local network architecture could not support scenario execution. Two implementations were used to reduce network overhead. First, vehicles were placed on the terrain without the additional infantry, decreasing the number of friendly entities to 210. A lower number of entities resulted in fewer network packets. Second, a network packet bundling technique also diminished the number of network transmissions.
Determining the correct amount of computational power necessary to handle the large scenario was another challenge. The ModSAF scenario was originally built on a Sun Microsystems UltraSPARC60, a system with two 296 MHz UltraSPARC-II processors. The UltraSparc60, however, could not complete the scenario. Additional systems were added to enable simulation execution in a distributed environment. The UltraSPARC60 displayed the plan view and did not support scenario entities. Five SGI systems were used to share simulation entities; three SGI O2s with one 175 MHz (IP32) R10000 processor each, one SGI Maximum Impact with one 195 MHz (IP29) processor, and one SGI Infinite Reality Onyx with four 194 MHz (IP25) processors.

The modified network and hardware configuration was still insufficient to gather data of completed scenarios. The next step was to obtain accounts on ARL’s Major Shared Resource Center’s high performance computers. The ModSAF software was installed on four SGI Origin 2000 systems. Each system has at least 32 processors (250 MHz IP27 R10000) and 32 gigabytes of main memory. Since one system supported successful scenario completion and data collection, network transmissions were disabled. Project data were compiled using one Origin 2000 system.

3.3.2 Combat Simulation Data Collection

Data collected from Fox-generated COAs were processed within ModSAF and will be used to determine COA “goodness.” The higher fidelity found in ModSAF enables a detailed accounting of interactions between scenario units. Progress toward objectives can readily be measured with consideration given to unit behaviors imposed on simulated entities. These examples of accumulated data represent metrics supporting the statistical design.

Data were collected using a numerical measure to represent attrition and mission fulfillment. Metric data were generated from iterations of both the Fox wargamer and ModSAF simulation. Detailed unit damage and locations were recorded to gain an understanding of projected losses and mission completion. The numerical strength loss results from Fox and the actual vehicle damage accumulated during ModSAF exercises were noted.

As can be seen in Figures 7 and 8, Fox and ModSAF attrition were measured as a percentage of force lost. Mission completion through occupation of terrain by a unit in Fox was determined by its icon position within the objective terrain area.

Analogous measures for the higher fidelity ModSAF simulation were recorded by examining the condition and position of each vehicle. Individual vehicle damage was represented as a decimal value, position as an integer. There are five categories of damage to a vehicle: undamaged, firepower kill, mobility kill, firepower and mobility kill and catastrophic kill, and three possible terrain placements: at LDT, at objective or neither.
3.4 **Statistical Evaluation**

3.4.1 **Objectives**

The planned statistical evaluation has two parts. The initial focus is comparability, with regard to wargamed results, between ModSAF and Fox. The Fox wargamer has not been validated, but it has been continuously improved by expert opinion from experienced officers. The cited studies incorporate subjective quality assessments of recommended COAs as part of usability experimentation during software development. Fox is now sufficiently mature to begin careful consideration of simulated battle results from implementation of suggested COAs. Since ModSAF is itself a simulation and not universally accepted as truth, agreement between Fox and ModSAF does not constitute validation. However, owing to the widely accepted use of ModSAF, a comparison of Fox to a ModSAF baseline is instructive. This part of the evaluation was accomplished by playing one specific friendly COAs against one specific enemy COAs in both Fox and ModSAF. Similarity or difference can be determined on the basis of the available Fox measures of friendly attrition, threat attrition, and accomplishment of objective.

---


In a related effort the Rockwell Science Center has developed a postprocessor, OWL, to study the sensitivity of battle results to user-defined battlefield conditions, as assessed by the Fox wargamer.\textsuperscript{9} OWL takes the input values and overlays probability distributions around them. Since the battlestaff estimates of important parameters, for example enemy strength, will not be exact, COA development will likely be based on values differing from the true battle conditions. By considering randomly chosen strengths about the estimate and running the wargamer for each, a distribution of battle results can be established. By examining this distribution, a better understanding of input information accuracy can be obtained. OWL does not assess a change in COA quality as a function of user-defined inputs. Rather, OWL introduces an additional stochastic element in the wargamer for a fixed COA, to produce a distribution of battle outcomes. "A comparison of Fox/OWL wargaming results to those of a higher-fidelity battle simulator may be required to assess the adequacy of the wargaming algorithm.\textsuperscript{9,10}

The second part of the planned evaluation employs sensitivity analysis to gain a better understanding of the role of Fox user-input values on COA generation. Questions to be examined include: how do changes in input selections affect the quality of friendly COAs, is the additional planning time provided by Fox beneficial, and how does the risk value affect the recommended COA. This effort assumes reasonable agreement in the initial evaluation. Better use of the Fox tool will come from a more complete understanding of how varying these initial inputs changes the character of the recommended COAs and their observed performance. COAs from Fox developed under a range of initial inputs, will be analyzed in terms of ModSAF results.

3.4.2 Analysis and Discussion

Preliminary simulation results available at the time of writing support only a limited analysis of the first objective, comparing the wargame results for one fixed Friendly COA (FCOA) played against one fixed Enemy COA (ECOA). Ten ModSAF and ten Fox runs were completed. Remaining strengths for friendly and enemy forces and terrain objectives are examined.

In Figure 9, the overall performance of ModSAF and Fox is given for ten runs of each. At first glance, the two wargamers appear to be giving similar results. However, this representation masks some important information, especially as to variability of the strengths. In Figure 10 it is clear there is greater variability among the ten runs of Fox.

In ModSAF the remaining strength percentages for Blue and Opposition forces are clustered about their means of 88% and 30%, respectively, with standard deviations of 8% and 7%, respectively. In contrast the Fox results show standard deviations of 36% and 15% for Blue and Opposition, respectively. The greater variability for Fox results can, in part, be explained by the battle outcome. In three of the four lowest remaining strength values for Blue, the result of the battle was a loss for Blue. The three greatest remaining strengths for Opposition corresponded, not surprisingly, to the Blue losses. The second smallest remaining strength for the Blue force corresponded to the smallest remaining strength for Opposition. Blue won that battle, but only after both forces experienced significant losses. The variations of the remaining strength values


\textsuperscript{10}Ibid.
are comparable between Fox and ModSAF for the Opposition force, if Fox runs in which Blue lost are ignored.

Figure 9. Remaining Force Strengths (Mean %)

Figure 10. Remaining Strength Percentages in ModSAF (MS) and Fox (F)
For Blue (B) and Opposition (O) Forces

More troubling than the issue with variation is the fact that Fox delivered remaining strength percentages with a median value slightly in excess of 100% for Blue. This curious result is due to Fox taking into account extra information. For example, Fox assigns some bonus value for forces being committed at the end of the battle. This extra information was not included when summarizing the ModSAF results. Those values are the percentage of fully functioning units at
the conclusion of battle; however, the justification for these remaining strength values in Fox is unclear.

Mission objectives were satisfied differently in Fox and ModSAF. As previously mentioned, in three Fox runs, Blue lost. In the remaining Fox runs, Blue took only two of the three terrain objectives. In ModSAF, Blue won every battle, taking all three terrain objectives in seven runs and two of the three objectives in the remaining three runs. The same terrain objective proved most difficult for both ModSAF and Fox.

4. Conclusion

The initial effort to link the Fox COA generation software and the ModSAF high-fidelity combat simulation through scenario translation and experimental trials is complete. These components and processes are the foundation of an ARL COAA testbed capable of producing combat performance data for use in statistical analysis. Data resulting from testbed operation have been collected and are the basis for preliminary COA comparisons and analysis.

ModSAF and Fox data show consistency as Blue was winning decidedly most of the time in each. Fox’s internal wargamer may be good enough to provide battle results for competing COAs which are approximately accurate in a relative performance sense, but not accurate in a ground truth sense. More runs with different COAs will be completed to evaluate Fox recommendations. Results will be correlated with ModSAF. With differing test conditions involving both ECOAs and FCOAs, a binary logistic regression will be performed in addition to the descriptive summaries discussed previously. More formal analyses will then be possible.

The extension of formal statistical design and optimization methodologies to combat simulation provides a rich research opportunity. Future work will address the realism of Fox battle progressions. FCOAs will be paired against several ECOAs for multiple replications. Using these runs, a reference distribution will be developed. The Fox battle can then be compared in each case, either on a univariate scale for each measure separately or on a multivariate scale for all simultaneously. Another question for future analysis is: if the suggested COAs play through simulation, what range of COA quality would be observed? Here interest is in the variability of COA performance. Since ECOA implementation is not known in reality, the reference distribution for performance should be constructed from multiple runs. The way to approach this is to form tolerance intervals for outcomes of the FCOAs via “the bootstrap,” a popular nonparametric technique.

Lessons learned will lead to operational improvements of the Fox COA generator and wargamer, techniques for automated COA translation and analysis, and extended COA measures of effectiveness. The impact of this work is potentially quite significant. The prototype testbed will evolve into a suite of implemented algorithms for tactical monitoring and decision support. Disbursing “real-world” data and simulation results to battlefield Command and Control (C2) elements, the suite will enable staff planners to quickly modify COAs. Faster planning achieves battlefield initiative.
The goal is to reveal strengths and weaknesses inherent in COA generation software. Scientific evaluation of COA attributes will benefit the soldier by providing increased operational possibilities during planning. The complexity of COA provides fertile ground for development and application of decision aid technology. This work, being done in collaboration with the Command Post XXI Advanced Technology Demonstration, has potential for improving C2 in the near term.
Appendix A

Genetic Algorithms

A common technique used to determine a solution to a problem that does not lend itself to traditional methods is the use of Genetic Algorithms (GA). Embodied in the GA is the ability to generate possible solutions and identify those which tend to be better than the rest. In the case of Fox generated COA's, a solution consists of a fixed set of variables and their assigned values, deemed significant in determining a battle plan’s success.

A fixed number of possible solutions are derived, frequently randomly. The GA is executed for a number of iterations until the fitness of the solution is stabilized.

One iteration of the GA involves:

1. The selection of a set of solutions which frequently will be more fit than the others in the set.

2. The exchange of like solution sections to generate a new solution. The intent is to create a solution better than the solutions from which it was derived.

3. The changing of a number of bits to encourage exploration of possible solutions in other areas of the solution set. (Fox represents a solution as a string of bits.)

Convergence to a single solution is almost guaranteed unless a strategy is used to generate a set of solutions, which will be significantly different. A technique known as “to remove solutions from the solution set that are too similar, allowing solutions which may not be as fit to survive the next iteration. As a result the commander is offered a number of battle plans (solutions) to consider.
Appendix B

Acronym Table

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARL</td>
<td>Army Research Laboratory</td>
</tr>
<tr>
<td>C2</td>
<td>Command and Control</td>
</tr>
<tr>
<td>COA</td>
<td>Course Of Action</td>
</tr>
<tr>
<td>COAA</td>
<td>Course Of Action Analysis</td>
</tr>
<tr>
<td>ECOA</td>
<td>Enemy Course Of Action</td>
</tr>
<tr>
<td>FCOA</td>
<td>Friendly Course Of Action</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>HPC</td>
<td>High Performance Computer</td>
</tr>
<tr>
<td>LDT</td>
<td>Line of Defensible Terrain</td>
</tr>
<tr>
<td>MHz</td>
<td>Megahertz</td>
</tr>
<tr>
<td>ModSAF</td>
<td>Modular Semi-Automated Forces</td>
</tr>
<tr>
<td>MSRC</td>
<td>Major Shared Resource Center</td>
</tr>
<tr>
<td>NTC</td>
<td>National Training Center</td>
</tr>
</tbody>
</table>
References


