From Optimal to Robust COAs: Challenges in Providing Integrated Decision Support for Simulation-Based COA Planning

A White Paper

B. Chandrasekaran
Laboratory for AI Research
Department of Computer & Information Science
The Ohio State University
Columbus, OH 43210
Email: Chandra@cse.ohio-state.edu

February, 2005

Abstract. With the increasing availability computer-based simulations, military Courses of Action (COA) planning can go beyond the doctrinally required three distinct COAs to accomplish the mission. In principle, planners can generate, simulate, compare and select from a large number of COAs, thus vastly increasing the chances for finding good COAs. However, any planning based on simulations has to face an *intrinsic* problem of simulation models, namely, that there is an inevitable gap between the models and reality. Comparing and selecting COAs on the basis of simulation results can be problematic in the face of simulation model gaps and errors. The gap between model and reality is not an issue for computer simulations alone; as Secretary Rumsfeld's now-famous remark about "unknown unknowns" suggests, simulations that human minds perform are also subject to the same problem. However, humans have evolved a set of heuristics by which they sometimes critique their models and simulations to test their robustness. The awareness of this gap is not common in research on decision support for planning involving computer simulation. In addition, since much of the simulation takes place rapidly in the computer, the planner is not even aware of possible gaps, and there is a tendency to place too much reliance on the results of the simulation, and hence on the plans based on them. A related issue is in how to handle the uncertainties in the models used in simulation. A standard approach has been to assume probability distributions and make decisions based on expected values for outcomes of interest. In fact what the planner needs is not quite the expected values, but rather what effect the uncertainties have on what dimensions of the outcome to what extent, and correspondingly, how to make the desired outcomes be less sensitive to the uncertainties. Both of these issues – handling the gaps as well the uncertainties – require a shift in point of view from *optimality* to *robustness*. In order to realize the full potential of the vastly increased search spaces made possible by computer simulation, it is essential that the decision support system empower the planner to explore the plans for robustness of the selected plans. The research challenge is to identify, and incorporate as part of decision support systems, a variety of techniques by which the selected plan can be tested for sensitivity to

various model assumptions and uncertainties.   We describe a research program to realize the vision of an integrated decision support system that supports robust planning through the plan-execute-replan cycle.

1. Introduction

Over the last five years, under ARL and AFRL support, we have been researching, and developing technologies for, decision support for COA planning. During this research, we have become aware of the degree to which advances in computer modeling and simulation, along with ever decreasing costs of computing, have made possible revolutionary changes in planning military operations, with the potential to empower the planner to explore much larger decision spaces, i.e., consider more COAs, and, for each COA, obtain more accurate estimates of performance, than before. The COAs that result from this process are more likely to achieve the strategic and tactical goals of the Blue forces. Increasing interest in Effects-Based Operations (EBOs) has placed even more of a premium on advances in simulation and modeling in support of COA planning. We have developed an architecture for multi-criterial decision support, that makes use of simulation as an essential component of COA evaluation, and helps the decision-maker to select optimal COAs. This architecture is part of a framework that views decision support as integrated for the planning life-cycle as a whole – generating, simulating, and selecting COAs, exploring the decision space, monitoring during execution, and replanning. This architecture has been demonstrated as an aid to COA planning in a number of Army and EBO problems.

The gap from simulation models and reality. Because of our experience in decision support for COA planning, we have also become aware of new challenges that arise from the inevitable gap between simulation models and reality. Because of this gap, the decision maker cannot take at face value the simulation results about the performance of a COA. Potentially missing aspects of reality cast a shadow over the simulation results. While the gap from the model and reality is not the same as the issue is of uncertainties that are modeled in terms of probabilities, there is also the question of how to make use of the probability distributions in COA planning. Most approaches cast the issue as one of obtaining the estimated values, and some may even bring in variances, but the decision-maker is also interested in phenomena that are not high in probability but offer opportunities or warn of dangers. Perhaps, he can take additional measures to exploit opportunities, or to avoid dangers, but he first needs to know of them. The decision support system should help the decision maker explore the decision space for such opportunities and dangers, and similarly help him analyze a COA with respect to its sensitivity to various simulation assumptions.

From Optimality to Robustness. While most decision support frameworks are based helping the planner select the optimal plan, i.e., one that scores the most desirable values on the various criteria of interest, what appears to be an optimal plan based on simulation models may in fact be quite far optimal, since simulations require models, and models are *in principle* incomplete and possibly inaccurate. The incompleteness of models is especially a problem in planning EBO’s, since simulating EBOs requires modeling social and psychological factors of leaders and complex organizations, and such models are especially likely to have significant gaps. In addition to missing aspects of reality, there is the inevitable uncertainty in the environment. Even when we model them in terms of probabilistic distributions, comparing COAs based on their expected values is unlikely to be the best approach, since the planner might be interested in taking or avoiding risks based on less likely but possible opportunities or dangers. All of this calls for a change in point of view. In the face of considerable uncertainties and incomplete models, what planners want or should want is not so much an optimal decision, but a *robust* one, one with a good chance of achieving his goals in spite of the many uncertainties. The decision support framework should thus enable the planner to explore how robust the COA's of interest are. Exploring for robustness involves supporting activities such as: identity which parameters or variables a performance criterion is sensitive to and over which range of values; identify ways in which knowledge of such variables can be the basis for launching selective critiquing of aspects of the model for completeness; identify how to design COAs such that early stages of the plan give useful information that may be used to fix gaps in the model and improve the plans for the later stages; etc.

Multi-Criterial Nature of Planning. It is now commonplace that decisions are *multi-criterial*, in particular that at least some of the criteria might be contending criteria, i.e., improving on one criterion might involve losing on another. Widely used tools such as Dynarank provide some form of support for However, decision support techniques are often based on psychologically unrealistic requirements such as weights on criteria, and in any case, they are not especially designed to exploit the power of computing to generate large numbers of decision alternatives. Finally, the multicriteriality of the decisions should run throughout the multiple stages of the planning cycle, not just the main selection phase.

In this white paper, we outline the conceptual foundations of a decision framework that can help satisfy the above desiderata. The central concern is to help the decision maker protect himself from the gap between models and reality, by empowering him to explore how robust the COAS are if the world turned out to be different from what the models said it was. We also briefly describe a number of specific research tasks that would advance our understanding of how to build and use such decision support systems.

2. Integrated Plan Support Architectures

COA planning does not exist in a vacuum. It is embedded in a decision making cycle that starts with the mission and information about the current situation, and any plans selected for execution need to be monitored, modified as necessary during execution in response to the developments, and finally, the DM has to abstract from the execution lessons learned that would be helpful for future COA planning. During our investigations on COA decision support, under ARL and AFRL support, we have developed a broader vision of COA decision support, as illustrated in Figure 1. It shows the various subtasks associated COA planning: generating COA alternatives, given the mission and the situation, evaluate the plans by simulation or other methods, explore the decision space for insights that could be useful during COA monitoring and replanning, compare COAs and select one to execute, and monitor execution and replan as needed, based on insights obtained during the decision space exploration stage. Decision support may be provided for any of the subtasks, but it is important that the basic framework must be flexible enough to make use of different solution strategies and associated software for each of the subtasks. For example, there are alternative simulation systems available for COA simulation and the decision framework should be able to work with them. The ovals in Figure 2 indicate some of the technologies we have developed in our Laboratory with ARL and AFRL support. We will briefly describe some of them later in the document.

3. Comparing COAs Multi-Criterially: The S-F-V Architecture

Let us consider a simple version of the problem: we have a set of *n* alternative COAs, represented simply as D_1, D_2, .. D_n, and that each COA has been evaluated along a set *m* of criteria, $C_{1,}$ C_2,..C_m. (Mnemonically, D's represent decision alternatives and C's stand for criteria labels.)

Assumptions: Criteria values are known and are comparable. When the value of a criterion is not known for a COA, and the criterion is relevant, the COA cannot be effectively compared. For now, let us assume that for each COA that is being considered, the values of all it criteria variables of interest are known. We also assume that the criteria values satisfy the comparability criterion: given two values, either we prefer one to the other, or we are equally happy with the two values. That is, "the values cannot be compared" is excluded.

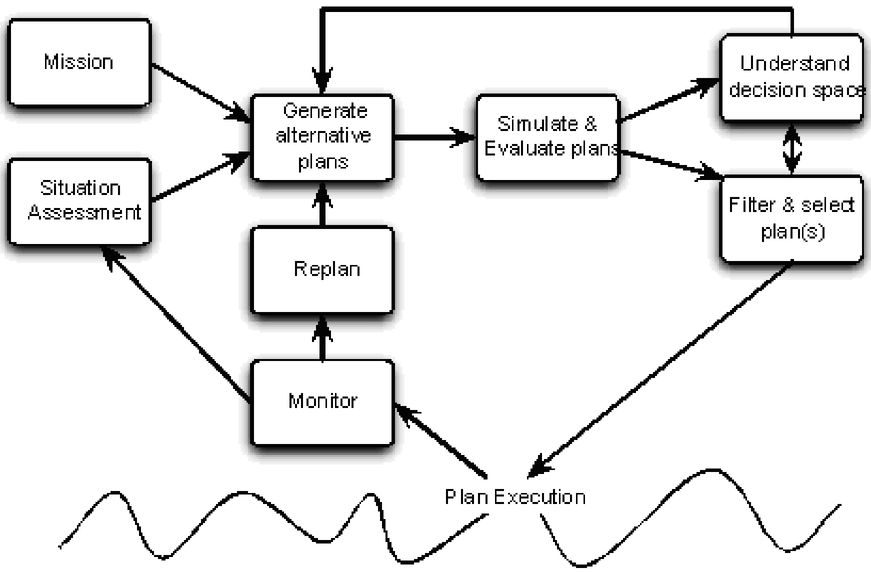


Figure 1. Subtasks of COA planning and monitoring

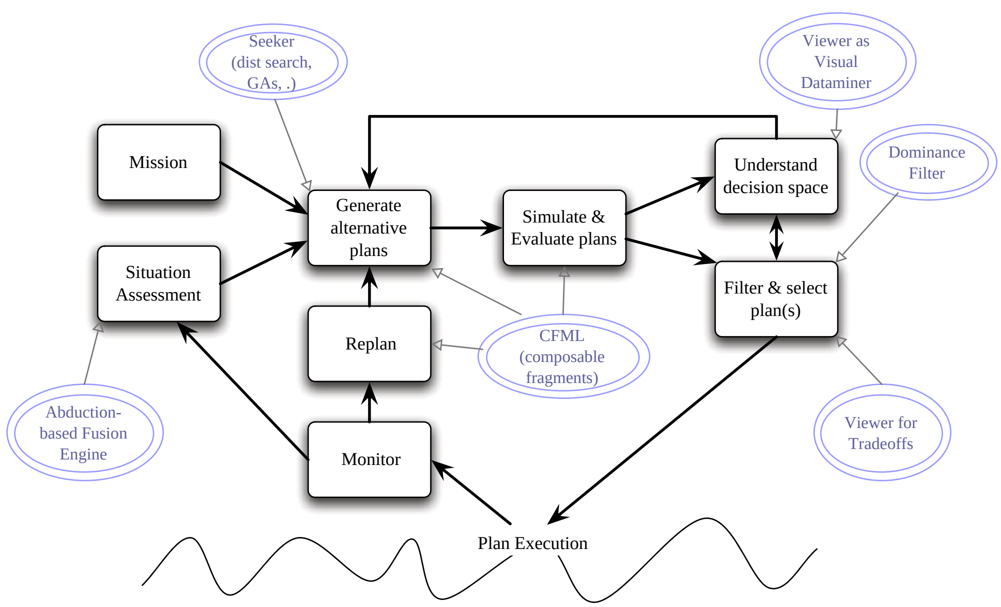


Figure 2. A vision of an integrated planning architecture, where the ovals represent some of the technologies currently available from our laboratory.

With the above two constraints, and given two alternatives D_i and D_j, one alternative may be better than another on some criteria and worse than the other in some other criteria. An enormous literature exists (see [1] for a summary, and [2] for a book-length review) on multi-criterial decision-making. The central idea in multi-criterial comparison is *Pareto Optimality*.

Pareto Dominance. An alternative A *Pareto-dominates,* or for simplicity, *dominates* alternative B if A is at least as good as B in all the criteria of interest and is better than B in at least one criterion.

Pareto-Optimal subset. Given a set of alternatives D and a set of criteria C, an element d_i of D belongs to the Pareto-optimal subset (with respect to C), iff there is no element $d_j \varepsilon$ D, $d_j \neq d_i$, such that d_j dominates d_i. The elements of the Pareto set do not dominate each other, i.e., for any pair, each of them is better than the other in at least one criterion.

When Pareto dominance is applied to a set of alternatives, typically more than one alternative will survive the dominance filtering. The percentage of survivors asymptotically goes to zero as the number of alternatives increases, showing that the filter scales up well. In one experiment [3] involving close to two million decision alternatives and four criteria, only 0.06% survived Pareto filtering. The percentage of survivors slowly grows with *m*, the number of criteria; if A dominates B on *m* criteria, and an *m+l*-th criterion is introduced, the chances of A dominating B go down.

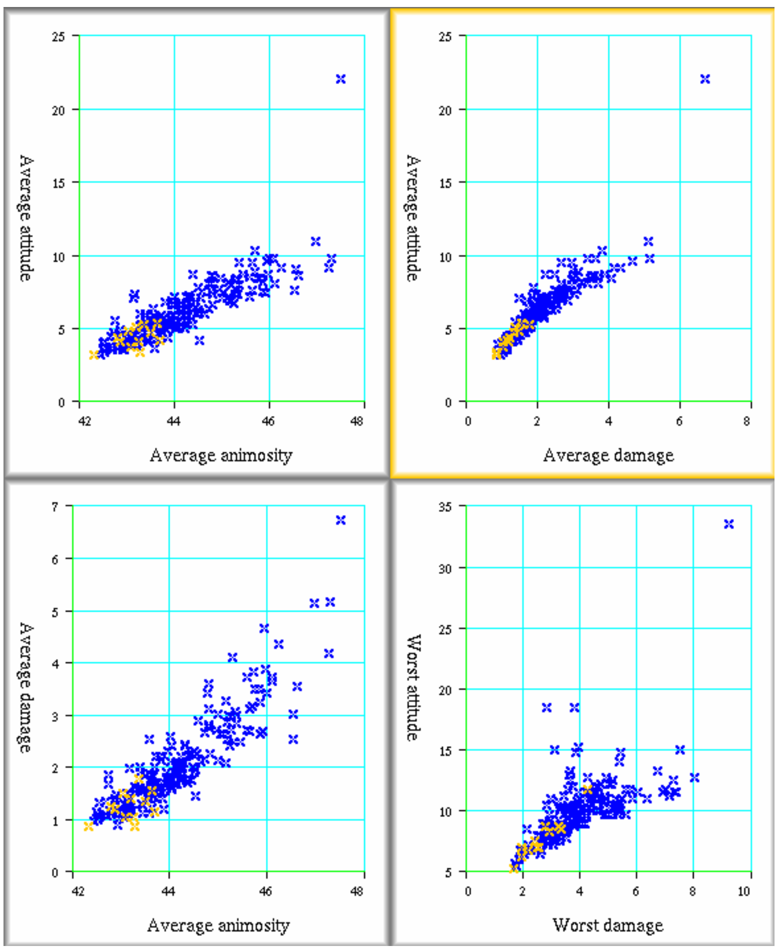
We have focused on Pareto-optimality since it is a kind of gold standard for multi-criterial choice. Any approach that doesn't guarantee that the selected alternative is in the Pareto set is in a clear sense sub-optimal, though it may have other virtues such as simplicity of decision making.

Given a set of Pareto-optimal alternatives, choosing one over another requires making additional trade-off judgments. We will shortly mention a technology that enables an analyst to bring to bear these additional trade-off preferences.

Reference [1] reviews the various proposals for how to choose an alternative in the face of multiple criteria of interest. A common starting point is to express in some way the relative importance of the various criteria. One approach is to order the criteria in terms of importance, choose the best on the most important criterion, and if there are ties, choose the best on the second-most important criterion, and so on. Note that this approach does not usually result in a final selection that is Pareto-optimal. Another common approach is *weighting* the criteria. In this approach, the DM is asked to indicate the relative importance of the various criteria by assigning weights to them. A linear sum of the criteria values (assuming that the criteria values are numerical) may then be constructed, and the COA or COAs with the best value for this weighted sum can then be chosen. It is worth noting that all the alternatives that would be found best for some linear combination of weighted criteria are in the Pareto-optimal set.

Optimizing a quantity that is a linearly weighted sum of criteria is mathematically convenient, but it is hard for the DM to come up with a set of weights that they are comfortable with. The reason is that the relative importance of the various criteria is not the same in different parts of the decision space. A commander might be willing to trade off .05 probability of mission accomplishment for a 10% decrease in friendly casualties when the mission accomplishment probability is 0.9 and estimated friendly casualties are high, but he might have an entirely different attitude to the relative ratios when the mission success probability is 0.6 and the estimated casualties are low. A fixed set of weights does not capture this non-linearity across the space.

Systems such as Dynarank combine some elements of ordering the criteria and focusing



Alternatives COAs in the Viewer. Each scatter plot contains the same COA alternatives (x's), but displays them with their values on different pairs of criteria. The DM can examine trade-offs in the various scatter plots and make selections on any of the plots. The selected ones appear in the different color in all the scatter plots, so the DM can examine how the selected ones fare with respect to other criteria of interest.

on the most important, and of weighting. In the following we briefly describe a technology – we call it the Seeker-Filter-Viewer or S-F-V technology – that was developed in our Laboratory over the last several years with support from DARPA, ARL and AFRL [3-6].

The Filter in the technology is the Pareto Filter as we just described, and the Viewer environment enables the DM to view the alternatives as scatter plots along different criteria, or as trade-off scatter plots for pairs of criteria. In Fig. 3, a set of alternative COAs are shown in the Viewer, as four scatter plots each for one pair of criteria. Each scatter plot contains all the alternatives. He can select one or a subset in any one of the scatter plots – the selected ones change color in the scatter plot on which the selection was made as well as in the other plots, so the DM can see how the selected ones fare with respect to other trade-offs. He can select a subset in one, and all the scatter plots are instantly updated. At this point, he can move to consider trade-off on another plot.
In contrast to the weighting schemes, the Viewer environment does not force the DM to articulate a priori a set of weights. He is always considering concrete alternatives with concrete values for the criteria, so he can decide what kind of trade-off makes sense to him *at that point* in the decision space. Typically, DMs look for sweet spots – situations where for a small penalty in one of the criteria, he can see he would get a large improvement on another criterion of interest. DMs can narrow their choices by making a series of trade-off –based selections in each of the plots.

This brief introduction to the S-F-V architecture is sufficient for our current purposes.

We remark in passing a crucial assumption: that the criteria values are known exactly. If they are only known to some degree of error, then the comparisons of criteria values for determining dominance, and hence decision about dominance, may be in error, because of the stochastic nature of many of the phenomena -- the weather, what the enemy would do, equipment malfunctions, etc., or because the simulation model was incomplete or inaccurate. We will revisit this issue in Sec. 4.3, but point out that, if an estimate of the error can be made in advance, a range of modifications to the Pareto Dominance filter can be made. Alternative filters for this situation and their mathematical properties are discussed in [1].

4. Issues in Simulation-Based Planning

We consider three issues:

1. Evaluating incompletely specified COAs. Simulation forces specifying details, whereas often in human planning, we can compare incompletely specified COAS of certain types.

2. Issues involving uncertainty in the models.

3. Issues involving missing elements in the models.

4. 1. Evaluating Incompletely Specified COAs

We consider two types.

4. 1. 1. COAs Specified as Schemas

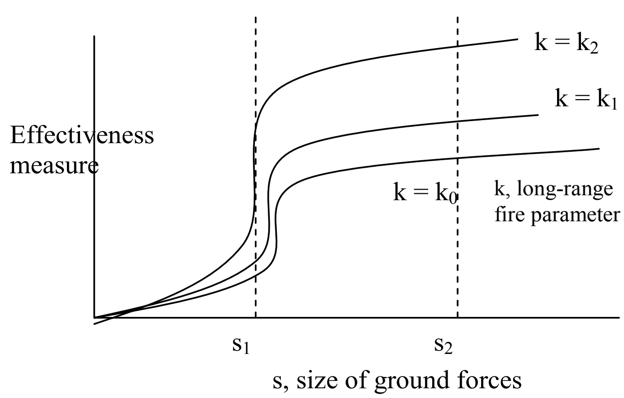
Comparing COAs requires that we have evaluated them on the criteria of interest, and doing this by simulation requires, as a rule, that the COAs are completely specified. However, often we do compare incompletely specified COA's, such as the following options in a given situation:

- COA1: "Introduce ground forces as a trip wire, but use long range fires to stop the enemy if he proceeds to engage the ground forces. (Hope that the presence of ground forces deters the enemy, but accept the probability of a catastrophic event that the trip wire is tripped, the long range fires are not effective in stopping enemy and the ground forces suffer great loss.)"

- COA2: "No ground forces, all work done by long range and air fires to halt the enemy."

These COA's are *structurally distinct*, but they are really COA *schemas*, since implementing either of them requires making a number of specific commitments, e.g., size and location of ground forces, specifics of long range fires, etc., for COA1 and specifics of the long range and air fires for COA2. To simulate the above COAs, specific instantiations of the variables would need to be made. In fact, by varying the values assigned to the variables, numerous specific COAs can be generated from each of the schemas.

Nevertheless, for many purposes, intelligent discussion can proceed without instantiating the schemas, i.e., the schemas specify real options that can be evaluated. The theoretical question of interest is the following: What enables us to compare certain types of schemas whereas for other COAs specified in schema form, we won't be able to proceed without instantiation?



Sharp Phase Transition

We propose the *Sharp Phase Transition Hypothesis* to understand the issue. Suppose there is a critical range of values of variables over which the COA instances differ in performance, but beyond that range the performance doesn't vary much. For example, in COA1, there are two values, s_1 and s_2, for the size of the ground force such that below s_1 it really doesn't serve as a trip wire psychologically, and above s_2, it is not a trip wire either, but a defending force, so the COA conceptually belongs to a different class (see figure). Further, let us say that the outcome measures of interest do not vary significantly for almost the entire range s_1 to s_2, except for sharp phase changes around s_1 and s_2. If these conditions are satisfied, COA1 can indeed be treated as essentially one COA, with the understanding that the intended instantiation is a value for troop size in the range s_1 to s_2 (and similarly for other variables). The COA can then be simulated or otherwise evaluated based on instantiations that correspond to the appropriate ranges.

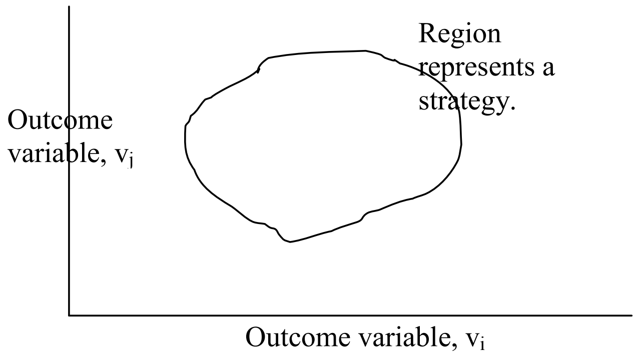
One might mention in this context the technique called *qualitative simulation* that has been developed in the AI community, which also dispenses with the requirement for complete specification of a situation. Under certain conditions, qualitative predictions can be made. However, this technique is limited in its application even to moderately complicated situations, since ambiguities start proliferating when simulation is carried out over several temporal instants.

4. 1. 2. COAs that have branches to be chosen at run time

COAs are often specified with branches that cannot be instantiated until run time (or, in simulation, until the COA has run in the simulation for some time) and specific events have or have not occurred. That is, actions in the COA depend on earlier events, including enemy actions. In this case, the COA cannot be treated as the equivalent of multiple COAs, each for one set of instantiations. This is because each of these COAs cannot be independently chosen and executed – each depends on the occurrence of specific earlier events in the world (or in simulation). Conceptually, the COA with event-dependent branches is one COA – it encapsulates one approach to the task.

However, there is still the issue of how to compare two such COAs? Should probabilities be assigned to the branches, and an overall, probabilistically weighted, value obtained for the criteria? If such COA's performance depends on enemy COAs, the situation is even more complicated. For each simulation of the Blue COA for example, should a number of different Red COAs be executed and the values for the criteria weighted probabilistically, or using a mini-max rule as in Chess-playing programs?

Again, I propose a hypothesis that applies at least to classes of situations. The hypothesis is that often a COA is planned such that the commitments made to the COA parameters before the occurrence of events that trigger the different branches determine an abstract region in the space of possibilities. The intention is that any particular set of events that actually happen in execution, including the actions corresponding to the various branches and their consequences, will be within this region. This region corresponds to the strategy of the COA in some sense, and the various branches are simply ways of making sure that the actual evolution of the activity stays in this region. This is something like making a plan to follow a certain route to get somewhere, and during the execution of the route making additional commitments, such as zigzagging to avoid a puddle, or to take a short detour because of an accident. These local commitments still keep us on the high-level specification of the route. Similarly for a high-level specification of a COA that might at run time involve branches and different commitments in different branches, but they all belong conceptually to the same COA strategy. The strategy anticipates a range of options by the enemy, e.g., and depending on what he actually does, specifies friendly actions that have a high chance of meeting the overall objectives. Similarly, the strategy might anticipate a range of weather events, and might provide measures depending on which weather events actually happen.



High-level parameters with the likelihood that outcome variables stay in the region. At run time, if events cause them to stray from the region, additional parameters are instantiated to bring their values back to the region.

If the above hypothesis applies to a COA, it doesn't mean that the actual events and the branches chosen do not differentially contribute to the various criteria evaluations. For example, certain events might indeed result in additional casualties, and additional materiel expenditure. However, the criteria corresponding to the most strategic aspects of the plan would be relatively less affected by the specific events and the branches. So, at least with respect to these criteria, different branches would have similar values.

For those criteria for which actual branches taken might make a significant difference, e.g., one branch may involve a significant increase in civilian damage than another, even though both might meet the mission objective, following tradition, certainly the best and worst case values of a criterion could be taken into account. However, it is important to keep the following caution in mind. Given two criteria, C_1 and C_2, there may be correlations between them in such a way that those branching choices in which the COA does well on C_1, may perform poorly on C_2. It is thus not enough to generate best, worst, etc. values for the two criteria separately. It will be useful to get an understanding of how the criteria relate to each other.

4.2. Uncertainty and Gaps in the Simulation Models

The difference between uncertainty and gaps is that in the former, we know what we don't know and model it as an uncertainty characterized by a probability distribution, while in the latter, we either aren't aware of important aspects of reality, or have mistakenly decided that they are not relevant.

In the case of uncertainty, the standard approach is to use the probability distributions on the various uncertainty elements to run Monte Carlo simulations, and come up with expected values for the various outcomes of interest. However, as Bankes [8] points out, difficult decisions about the future cannot be made on the basis of expected value for at least two reasons. The first is that if the outcome measures are correlated, then individual expected values will give a false picture of the future, but this can be taken care of by a more sophisticated stance towards computing the joint expected values. More seriously, however, expected values-based assessments fail to indicate both dangers and opportunities that may lie nearby, and possibilities for driving the future states to avoid dangers and exploit opportunities. What the decision maker really needs is an understanding of how sensitive the outcomes are to the assumptions about the uncertainty. For example, it would be good to know if the likelihood of taking the objective changes much or little for a given COA whether or not it rains, even though the actual uncertainty about the rain might be quite large. Even more importantly, the decision maker might be able to determine that while the probability of achieving the objective varies considerably depending on whether and how much it rains, if subgoal S is achieved, then rain doesn't affect the final outcome very much, and that increasing a certain COA parameter to say 8, increases the chances of achieving S independent of rain. What the planner has done in this case is to explore the robustness of a COA in the context of uncertainties in the environment. Another example is a potential large uncertainty in enemy position. The decision maker might explore a how to make a COA robust with respect to this uncertainty. It would be good for him to discover that it is possible to design a COA such that the initial stages of the COA are largely oriented towards getting more information about the location, and such that based on the information, later stages of the COA can be assigned appropriate parameters. Again, the goal is less to simply obtain estimated values for the outcomes than to explore the space around a COA to see whether it is possible to make it more robust.

In the case gaps, the issues are similar, but not identical. Ultimately, there is no real guarantee that we would be able to identify all the relevant gaps in the models and fix them. But the situation is not hopeless. If we start with the assumption that the model we have is generally good, but potentially missing important components, we can adopt the strategy of identifying regions of certainty and completeness which might then be used to suggest areas of the model that we are less sure of in terms of completeness. In some cases, the results of simulation in certain regions of the space might suggest missing or incorrect assumptions. For example, in planning an EBO, a prediction of the consequences of certain rare actions on the part of the Blue forces might be counter-intuitive or contradicted by empirical evidence, though the same model might make what appear to be reasonable predictions for run-of-the-mill Blue actions. This oddity might raise questions about specific aspects of the model that should call for further examination.

An important consideration is that we are not abstractly interested in critiquing or improving the simulation model, but with respect to evaluating one or a small set of COAs. This is likely to focus the need for additional investment in model-building to specific questions, reducing the amount of work involved.

5. More on Robustness of COA's

Whether in the COA domain or in car design, any evaluation of choice alternatives can requires use of *models*. Performance equations, simulations, etc., all make use of models of the relevant part of the world where we intend to implement the choices. This reliance on models is a property of all prediction, whether by computer programs or by thinking.

There are two fundamental sources of uncertainty in prediction. The first is the fact the world is stochastic in nature, so at best we can only get a probability distribution of the values for the various criteria of interest. The second problem is more serious: models are in principle incomplete, and even in the dimensions that are included, they may be inaccurate to various degrees. These problems imply that when a COA is deemed to be superior to another COA based on criteria values predicted by model-based simulations, this decision could be wrong, and sometimes grievously so; the rejected COA might in fact be better. Is there anything that a decision support system can do to protect a DM against this?

These are issues of *robustness* of COAs. There are two different but related robustness concepts:

i. Alternatives A and B have approximately similar expected outcomes, but the one whose ratio of upside/downside is higher is the more robust one[1].
ii. Alternatives A and B have approximately similar expected outcomes, but the one whose outcomes are less sensitive to simulation assumptions and uncertainties is the more robust one.

The robustness we have in mind to cope with model errors is type (ii) above. While everyone would agree, when pressed, that indeed simulations and reality differ, I don't think people in the field are sensitized to the extent to which simulation-based decision comparisons can be problematic, and hence the decision finally chosen may be deeply flawed as a result of this mismatch between reality and simulation. While there is no absolute guarantee against this problem – as mentioned earlier, human thinking itself is based on world models and so limits of simulation apply to human thinking in general – nevertheless, decision support systems ought to empower the DM to evaluate how sensitive a decision alternative is with respect to their specific assumptions in the simulation. We regard development of such systems as the next major challenge for decision support.

This issue is drawing attention in decision theory in general [8, 9]. For example, in [9] the authors remark, "Reliance by decision makers on formal analytic methodologies can increase susceptibility to surprise as such methods commonly use available information to develop single-point forecasts or probability distributions of future events. In doing so, traditional analyses divert attention from information potentially important to understanding and planning for effects of surprise." They use the term "deep

[1] This definition was given to me by Paul K. Davis (personal communication).

uncertainty" to refer to the situation where the system model is uncertain. They propose an approach that they call *robust adaptive planning* (RAP), which involves creation of large ensembles of plausible future scenarios. They call a strategy robust if it "performs reasonably well, compared to the alternatives, over a wide range of plausible scenarios." The idea is that, using a user-friendly interface, the DM looks for alternatives that are robust in this sense, rather than optimal in the traditional sense.

There is no universal solution, or even attitude, to the problem of protecting oneself against surprises. According to historian Eric Bergerud, "Bismarck in particular never thought that events could be predicted with precision. When a policy was pursued a range of outcomes could be expected. The trick was to develop policy where the minimum outcome (today we might call it a worst case scenario) was acceptable. If a triumph ensued, great. If it was something in between, don't die of surprise." This assumes that the goal is to minimize the worst case scenario. Not all missions can be approached this way. There are times when one might wish maximize the best, taking chances of bad outcomes. The decision support system has to support exploring the space of alternatives under a variety of attitudes, conservative or risk-taking, as the commander deems appropriate.

Ultimately there is no guaranteed protection against missing crucial knowledge. However, DM's are rarely in the situation where they know nothing and often they have some idea of what they are unsure of. For these kinds of situations, technologies that enable the DM to develop robust decisions, or at least explore the robustness of the decisions they are considering, will be very useful.

6. Decision Support Systems for Building Robust COAs: A Research Program

6. 1. Heuristics for Decision Support.

Given that the mindset we are proposing is one of exploring the COA space for robust COAs, what techniques/technologies can we develop along with a discipline of using them in the appropriate way? In what follows, we propose a research program intended to explore a set of ideas.

We plan to explore the design of capabilities for a DSS in order to support the following strategies:

1. Vary model assumptions & simulate over statistical contingencies of each model.

    Model assumptions about enemy in particular

    Search for assumptions, and *transition regions*, that outcomes of interest are most sensitive to.

Examples:

"In my model, I am assuming that the mullahs would rather give up their support of the Sadr faction than risk extending US presence close to their border. How does the likelihood of a deal change if I give up this assumption?"

"In the COA for disrupting enemy communication network, what are the variables highly correlated with the outcome variable, and which model assumptions do they depend on?" In this case, once we identify the assumptions, we can devote additional empirical resources to verify or test the assumptions for completeness and accuracy.

2. Explore COA modifications so that certain commitments may be made or changed as execution starts and information about reality becomes available. That is, design COAs such that initial results can be used to gather information, which in turn is used to specify parameters for later parts of the COA. Example:
"We are assuming enemy strength is weak here, but if it turns out wrong, we can reinforce from here."

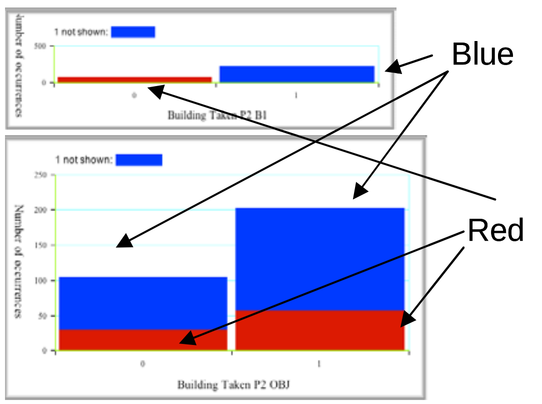
3. Based on insights about events on edge of phase transitions, think about modifying model so as to make or not make the events happen. Example: From an analysis of a simulation, we find that when variable V takes values above say V_0, an outcome variable of interest radically changes in value in an unsatisfactory direction. This then can be the basis for explicitly examining the domain to see which of the phenomena in the domain might have the effect of moving V beyond the critical value and identifying ways in which V can be kept within the desirable range.

Notice that many of these strategies are in fact heuristics used in human planning on an everyday basis. We usually have a sense of trust or lack of trust in our models of the world in mental simulation, and adopt strategies of early least commitment, information gathering, consistency and reasonable values checking, and so on.

6. 2. A Tool for Performing Robustness Studies

In [6], we discuss the use of the Viewer in the S-F-V architecture to explore sensitivities of one variable on another. Specifically in that application, a MOUT COA was simulated about 300 times over a wide range of statistical contingencies. The COA was intended to take a building labeled Objective, along with a set of minor objectives of taking other buildings labeled B1 through B4. A natural question one might ask is whether taking buildings B1 through B4 has a positive or negative effect on the likelihood of taking Objective. On one hand, resources allocated to the minor objectives might take away from the resources needed for taking the major Objective. On the other hand, taking these buildings might draw the resources of the defenders, leaving Objective easier to take. In this case, everything depends on the specifics of the case.

We demonstrated in that paper how the Viewer can be used to derive insights about such dependencies. The Figure below shows how it is used for answering the question for building B1. This is really a set of hypotheses, one for each building. In Figure 5, the planner has used the Viewer to call up two histogram plots, one displaying the number of runs in which B1 was and was not taken, and the other displaying the number of runs in which Obj was and was not taken. The heights correspond to the number of runs in each category. The planner has selected the occurrences corresponding to B1 being taken (the right set in the plot, top of Figure 5). (This selected set appears in blue in the Viewer, and the remainder appears in red; in the figures, color labels are used. In the plot of Obj taken (bottom plot), the same runs appear in the same color in the two histograms. Looking at the histogram for Obj not taken (histogram on the left at the bottom of the figure), among these cases, approximately a third are red and the remainder blue, i.e., in a third of the cases where Obj was not taken B1 was also not taken. Similarly, looking at the histogram on the right, in approximately a third of the cases where Obj was taken, B1 was not taken. So, a heuristic conclusion is that taking or not taking B1seems to have about the same impact on the chances of taking Obj. This kind of analysis may be repeated for other buildings.



Plots showing B2 taken vs not taken, and Obj taken and not taken. See text for explanation of colors.

We think the Viewer could be similarly useful in investigating the sensitivity of outcomes to model assumptions. If the DM is especially unsure of an aspect of the model, by generating a range of scenarios with varying assumptions about that aspect of the model, and exploring the sensitivity of the outcomes to these assumptions, the DM can develop a sense of how robust his COA might be with respect to this aspect of uncertainty.

6. 3. Network Disruption Planning: Concrete Application to Drive Research

We propose to drive the research by exploring the issues in the context of a concrete, realistic EBO application. The one we have in mind, and we have some experience with, is *COA planning for network disruption*: a set of nodes and links form a model of a network. A variety of networks, such as transportation and communication, may be modeled in this way. The nodes and links have differing values. There are various costs and benefits associated with disrupting the various nodes and links. The task is to come up with a list of nodes and links to be disrupted. We have earlier experimented with this domain for generating COAs by using evolutionary computation. This domain has a good deal of potential richness to it: increasingly complex models of enemy capabilities for defending and rebuilding nodes and links, psychological and political factors, may be built within the overall application.

Model gaps may exist about enemy capabilities and potential responses, and thus it would provide opportunities for modeling simulation errors of various kinds. Further, results of earlier disruption plans may be used for model-updating – one way to protect against model gaps – and replanning in various ways, providing an opportunity to investigate various issues in integrated plan support.

7. Discussion

This white paper makes a simple, but we think important, claim: that while computational simulation has opened up the prospects for revolutionary advances in COA planning, it also has the potential to result on mistaken decisions because of certain problems intrinsic to simulation and the decision maker leaving the details of simulation to the computer and thus not critically examining the assumption. We have proposed a change in perspective in which the goal is to produce robust COAs that are less sensitive to the problems with simulation that we have identified. We propose an initial research program whose goal is to demonstrate a set of techniques for exploring the robustness of COAs.

Bibliography

1. Iyer, N.S., *A Family of Dominance Filters for Multiple Criteria Decision Making: Choosing the Right Filter for the Decision Situation*, in *Computer & Information Science*. 2001, The Ohio State University: Columbus, OH. p. 169 + iv.
2. Miettinen, K. M., *Nonlinear Multiobjective Optimization*. International Series in Operations Research and Management, ed. F.S. Hillier. 1999: Kluwer Academic Publishers.
3. Josephson, J.R., et al. *An Architecture for Exploring Large Design Spaces*. in *Proceedings of National Conference of the American Association for Artificial Intelligence*. 1998. Madison, Wisconsin.
4. Chandrasekaran, B., et al. *"Mining Simulation Data for Insights about a Decision Space: Application to an Urban Combat COA*. in *SPIE Enabling Technologies for Simulation Science Conference, The Defense and Security Symposium*. 2004. Orlando, Florida: Bellingham, WA: SPIE.
5. Hillis, D., et al., *Collaborative Visualization Tools For Courses Of Action (Coa) In Anti-Terrorist Operations: Combining Coevolution And Pareto Optimization*, in *Army Research Laboratories Collaborative Technology Alliance Symposium*. 2003, U.S. Army Research Laboratories: University of Maryland Conference Center, Adelphi, MD. p. CD-Rom.
6. Kaste, R., et al., *From Simulation To Insights: Experiments In The Use Of A Multi-Criterial Viewer To Develop Understanding Of The COA Space*, in *Army Research Laboratories Collaborative Technology Alliance Symposium*. 2003, U.S. Army Research Laboratories: University of Maryland Conference Center, Adelphi, MD. p. CD-Rom.
7. Davis, P.K., *Effects-Based Operations (EBO): A Grand Challenge for the Analytical Community*. 2001, Santa Monica, CA: RAND Corporation.
8. Bankes, S.C., *Tools and techniques for developing policies for complex and uncertain systems.* PNAS, 2002. **99** (suppl. 3): p. 7263-7266.
9. Lempert, R., S. Popper, and S. Bankes, *Confronting Surprise.* Social Science Computer Review, 2002. **20**(4): p. 420-440.