Bits, Bangs, or Bucks?

The Coming Information Crisis

Alan R. Washburn

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943

There is a crisis approaching for military OR, centered on the role of information on the battlefield. It is clear to military professionals that information is becoming increasingly important, but unfortunately the OR profession's ability to measure its contribution is still primitive. We are not good at deciding whether information is cost/effective or whether information is more important than firepower, and yet precisely these questions will be asked more and more frequently as budgets shrink. Scott (1998) quotes General Howell M. Estes III, the commander of USSPACECOM, as saying:

> *Hard choices need to be made between investments in information infrastructure [and] the combat systems themselves. This is an extreme dilemma, because combat systems, without timely, relevant information, are useless. On the other hand, you can't take out an enemy tank with just information. We need to strike a balance between 'shooters' and 'information systems' if we're going to be successful in the future.*

But we aren't even sure how information should be measured — is it measured in bits, or what? The word "crisis" is not too strong for our approaching predicament.

Our leaders have recognized and articulated the importance of information. JV2010 defines Information Superiority to be "The capability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary's ability to do the same", and subordinates other dimensions of warfare. The Revolution in Military Affairs is mainly a revolution about information. It has been stated by many authors that the Allied advantage in Desert Storm was most strikingly an information advantage, rather than one of firepower. It is not as if we haven't been warned about the growing importance of information, and yet we do not seem to have the evaluation tools that we need. Why is the problem so difficult? Are the right tools already out there just waiting to be applied? Can they be developed? Let's review the situation, bearing in mind the JV2010 emphasis on the value of being able to "collect, process, and disseminate…"

Two Points About Information Valuation

Information has only a single purpose: to improve decision making by revealing or partially revealing the state of some process. If there is no decision maker, or if the decision maker already knows everything relevant, then information has no value. In other words,

Information is of no value unless there is an uncertain decision maker.

In fact, one reasonable measure of information is "the extent to which it decreases the decision maker's uncertainty". This observation led Shannon in 1948 to his elegant development of entropy as the most natural measure of information. This measure has many desirable properties, and leads to the fundamental ideas of channel capacity and redundancy. But Operations Researchers hardly ever use the idea. The only example known to the author of a serious attempt to measure military information *a la* Shannon is by

Sherrill and Barr [1996], who use entropy as a common scale on which to measure the informativeness of several information sources.

Why don't we use entropy more? Perhaps we should, but there are at least two major obstacles. The first is that the entropy computation is a significant exercise in conditional probability. The second is that Shannon's definition does not distinguish between relevant and irrelevant information. There is a tank out there, and you are wondering whether to shoot at it. It is probably Red (probability .9), but it might be Blue (.1). It also might have an even or odd serial number, equally likely independent of color. Which would you rather know, the color or the serial number? The color, of course, but the trouble is that the even/odd entropy is 1 bit while the color entropy is only .469 bits. Most of the uncertainty is about something that is tactically irrelevant, so it would be wrong to seek the "most informative" measurement. A distressingly large number of military situations fall into this category, with the commander lamenting that there is too much (irrelevant) data and not enough (relevant) information. We analysts could imitate Sherrill and Barr by confining the analysis to relevant information sources, of course, but the relevant/irrelevant classification won't always be obvious. A 56K modem is unquestionably a better piece of hardware than a 28.8K modem, but we must be careful to distinguish between the capacity of a channel and the information content of the messages that it carries. The bits must be relevant as well as plentiful. Shannon still has an important point: an analyst who wants to demonstrate the value of information must begin with a scenario that involves at least one uncertain decision maker.

If X and Y are two random quantities, then the entropy of X given Y cannot exceed the entropy of X. In the sense of Shannon, there is no such thing as information Y that causes one's uncertainty about X to increase, on the average. In other words, there is no such thing as "disinformation". I refer to this as the canine theorem, having observed that it is impossible to offend a dog with a smell. If disinformation really exists, it will have to be measured on some scale other than Shannon's. This new scale will have to involve decision making, as well as uncertainty.

For information to have value, it is not sufficient that the decision maker be uncertain. Humans have an almost insatiable need to be informed. We are famous for seeking information, even paying for it, in circumstances where we are unable to use it to affect an outcome. If information is to actually be useful, it must be collected, processed and delivered to a decision maker who has the inclination and power to influence events. The latter condition is not mentioned explicitly in the JV2010 definition, but should still be understood. Major General Robert Scales, commandant of the Army War College, recently worried that the effect of getting more information onto the battlefield might be that we would "just die smarter" (address at AIAA, January 22, 1998). Another way of making this negative point is that

*Information is of no value unless the decision maker has the power to use it.*

This second point about information is unfortunate for us analysts, since the notion of being able to influence events is so situation-dependent that we are unlikely to find a general theory of how to value information. Shannon was able to invent a general measure of uncertainty as entropy, but there is no correspondingly general measure for "power". The second point is nonetheless just as true as the first, and must be born in mind as we search for OR techniques that might prove useful in averting the crisis.

A Survey of Techniques for Valuing Information

The rest of this paper is organized according to Figure 1, which partitions analytic combat models according to purpose (normative versus evaluative) and level of abstraction. "Analytic" in the title means that the goal is analysis rather than training; that is, analytic models are by definition ultimately destined to support quantitative system comparisons. The comparisons are explicit in normative models and implicit in evaluative models. The next five sections deal mostly with normative models, and the remainder of the paper deals with increasingly realistic evaluative models.

Decision Theory

In principle, the method based on utility (von Neumann and Morgenstern [1944]) does it all. The decision maker's utility function U is first determined (every rational decision maker has one), and the best decision is then the one that maximizes E(U). The expected value E() is necessary because U will in general be random, but E(U) is still a scalar measure of quality. There are no restrictions on applicability, so in principle the method could deal with complicated budgetary decisions involving tradeoffs between bits,
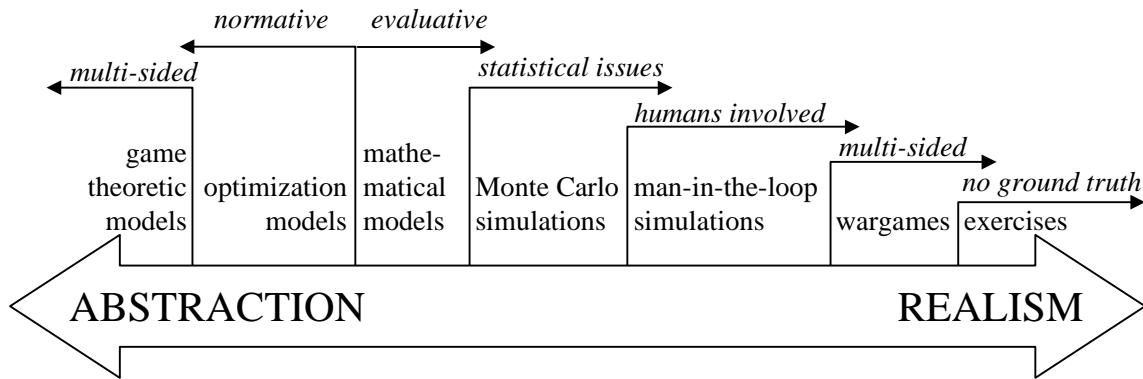
Figure 1. The Analytic Combat Modeling Spectrum
Arrow properties apply to everything in the indicated direction.

bangs, and bucks. The practical difficulties are in eliciting the function U, which must be assessed for every outcome in a set whose size can easily become explosively large, and in quantifying E(), which involves determining a probability for each outcome as a function of a decision that may also be in a very large set.

Some of these difficulties are well illustrated in a recent paper by Griggs, Parnell and Lehmkuhl [1997] (GPL). The core of the paper is a mixed-integer-program (MIP) that maximizes E(total target value destroyed minus total value of the aircraft lost in the process), with the quantity in parentheses being the utility function U. The GPL decision maker has no uncertainty about inputs in the core problem, simply using the MIP to select the best possible air tasking order (ATO). But GPL do explore one dimension of uncertainty—the weather may be Good, Marginal, or Bad, and there may be a forecast available. The MIP is first solved three times, once with data relevant to each weather state, thus obtaining three ATOs and three utilities UG=203.5, UM=133.6, and UB=94.2. Then weather forecasts are introduced, together with a table of conditional probabilities that the actual weather is good when the forecast is marginal, etc. Figure 9 from that paper is reproduced below as Figure 2. The probabilities in the decision tree have been determined by Bayes' Theorem, and the customary backwards rollup has been performed to conclude that the best warfighting policy is to use the Good ATO when the forecast is "Good", etc. The optimized objective function is 139.7, an improvement over the 119.4 that GPL show to be possible without the forecast. Thus the value of the forecast has been quantified, which is exactly the kind of thing that we need to be able to do. But Figure 2 also illustrates what GPL note to be a fault with decision trees: "they get very cumbersome for complex decisions". The tree has 18 leaves, for each of which a value must be established. There would be 27 leaves except for the GPL assumption that weather states can be ordered so that an ATO optimal for one weather state is equally effective in anything better. If there were f weather forecasts, d possible decisions, and w weather states, there would be fdw leaves. Furthermore, since there is in general no guarantee that the best warfighting policy will always be one of those policies that is optimal when the weather type is known, d is in principle the same as the number of strategies in the MIP! The size of the tree becomes awkwardly large very quickly. It is not just a problem of graphically describing the tree. To solve the (f,d,w) problem, it would be necessary to first solve f distinct MIPs, with the objective function in each case involving the probabilities shown in the deepest level of the decision tree. If the "intelligence and reconnaissance" node of the influence diagram that begins the GPL paper were implemented in the model with r possibilities for the results of reconnaissance, then the decision tree would begin with rf branches, rather than just f. In every direction we see the possibility for explosive growth in the difficulty of even describing the problem, much less solving it.
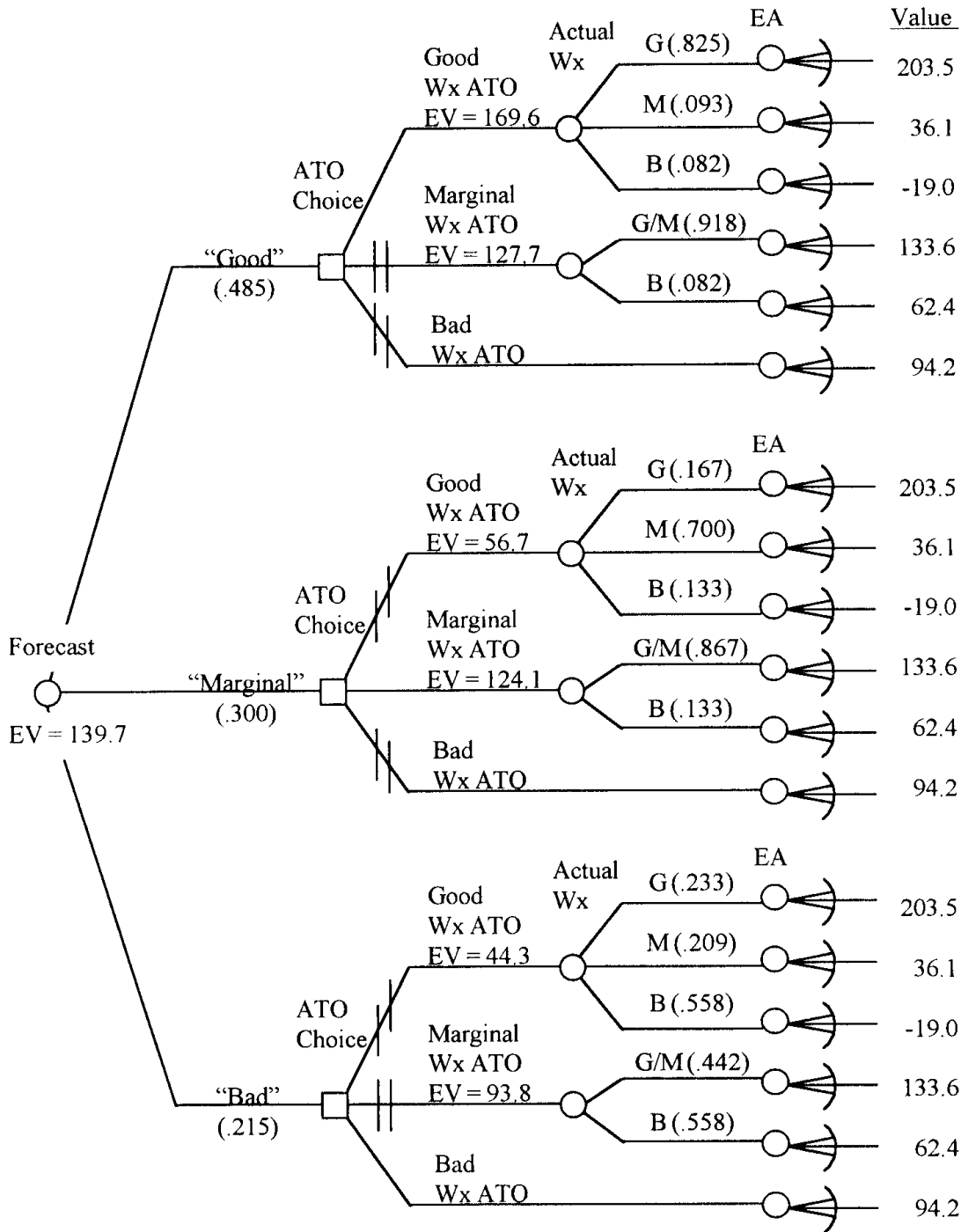
3

Figure 2. Decision Tree with a Weather Forecast (GPL Figure 9)

Decision trees are useful concepts, but without specializations of the type that GPL employ they are certainly not going to avert the coming crisis. The basic problem is that the decision maker must in effect determine a policy that specifies the decision *as a function of the information,* and there are lots and lots of functions. Representing the problem with an influence diagram (Schachter [1986]) is more efficient graphically, but does not change this fundamental difficulty. If the decision does not depend on the information, then the information is valueless. If it does depend on the information, then finding the best manner of that dependence is a problem in finding an optimal *function.* In most cases this means that the

best policy cannot be found by exhaustion. Some kind of special structure must be imposed if we are to have any chance of pronouncing a policy "optimal".

Decision Theorists have introduced a quantity called the expected value of sample information (EVSI, Winston [1994]). EVSI can be zero for information unrelated to the decision being made, but it cannot be negative. The question persists, then: "What is disinformation?" Perhaps the answer is so intimately related to competition that it cannot be answered except in that context. We will return to the question in the Game Theory section below.

Optimization Models (Mathematical Programming)

The same technological achievement that has caused the information crisis may also contribute to its solution, since electronic microminiaturization continually makes computers more powerful. Amazingly large mathematical programs can be solved these days, including the MIP studied by GPL. Why not use this capability to study Information Warfare?

Mathematical programming is already in use. GPL is one example, and another is CFAM, the current heir apparent to models such as TAM, MIXMASTER, HEAVY ATTACK and others that have been used by the Air Force for analyzing the air-to-ground battle (Yost [1996]). CFAM includes error probabilities for Battle Damage Assessment (BDA), an aspect of the air-to-ground battle that the advent of expensive precision weapons has made especially important. The general idea in CFAM is that a certain fraction of strikes at live targets result in erroneous "live" reports, that such targets must be struck again at the next opportunity, along with those that are truly alive, and that the restrike will correctly reveal the target's status. The important points about this are that the policy about restrikes is legislated rather than optimized, and that the legislated rule prevents target status from becoming too complicated. The former means that decision variables do not need to have a strike/restrike subscript, and the latter means that the set of targets does not explode through having to keep track of the probability of being alive. Such measures to limit scale are necessary in a model that is already computationally stressed through having to keep track of many types of aircraft, weapon, target, loadout, delivery profile, etc. The purpose of CFAM is thus not to determine how information *should* influence decision making, but only to measure the implied consequences of a particular imposed decision rule. CFAM's predecessors have long been in use by USAF for investigating tradeoffs between inventories of aircraft and weapons, and CFAM now extends the capability in the sense that the implications of changing the BDA probability are measurable.

Mathematical programs such as CFAM have many virtues for making tradeoff analyses, including the important virtue that the optimized decision is automatically adapted to new technologies and inventories. However, the tendency for the number of variables to explode is still lurking in the wings once the goals are expanded to include policy optimization. The following paragraphs outline two additional barriers to the widespread use of mathematical programming when the tradeoffs involve information systems.

Abstract models often employ Expected-Value-Analysis (EVA), by which is meant any analysis that replaces random variables with their expected values. In CFAM's case, EVA enters when the number of sorties implied by the decision to strike a target is calculated as an expected value (the actual number is random on account of random BDA results, among other reasons). This expected value is then accumulated and compared to an overall constraint to determine feasibility, thus ignoring the tactical question of what to do if ill luck on the first several targets consumes an unexpectedly large amount of resources. Whether EVA is harmless depends on circumstances. Since the optimal use of relevant information can be expected to imply variability in resource consumption, the consistent use of EVA in problems involving information has got to make one nervous. The techniques of Stochastic Programming may help here as computers become more and more capable, but so far only the regular kind of Mathematical programming has been widely applied to military problems.

In optimization models that include a succession of time periods, it is common to assume implicitly or explicitly that passage from one period to another provides a fresh start where the status of all targets is known at the beginning of the new period. For example, suppose there are 1000 targets at time 0, opposed by 3000 weapons with kill probability .5. Suppose further that the first attack commits two weapons to each target, killing $1000(1-.5^2)$ and leaving 250 targets for the remaining 1000 weapons in the next time period. So far this is an example of EVA, but in addition the *identity* of the remaining 250 live targets and 750 dead targets is typically assumed to be known; that is, it is assumed that the passage of time will reveal the status

of each target. How else could one assign weapons to the 250 live targets in the next period? The implied assumption is that the information system responsible for monitoring the target states is perfect between periods; that is, "Time Reveals All" (TRA). The almost irresistible analytical advantage of this assumption is that the situation at time 1 is just a reduced-scale version of the situation at time 0. The assumption has always been wrong, of course, but it is becoming seriously wrong in an era of long-range, expensive weapons that very much depend on knowledge of target status.

The TRA assumption can be made more plausible by allowing for implied information resources. In the previous example the 1000 attacks in the first period might have to be accompanied by 1000 inspections of some kind. In effect, information enters the model as a "tax" on attempts at target destruction. Again, the policy for using information is highly simplified and legislated, rather than optimized.

EVA and TRA do not necessarily come together. Evans [1996] is an interesting example of an analysis that employs EVA but not TRA in investigating a multi-stage war where BDA sensors make both type 1 and type 2 errors. He comments that "It is easy to simulate perfect BDA or nonexistent BDA, but it is hard to model flawed BDA". That statement is all too true, and it sometimes leads to taking shortcuts like the ones described in the next paragraph when real, error-prone information systems are involved.

Interesting Rheostat Knobs (IRKs)

There is one way of abstractly handling information that deserves special mention. The extreme cases where information is perfect or completely absent are often much simpler than intermediate cases. It is therefore tempting to simply include a tuning parameter as a bridge between well-understood extremes. HEAVY ATTACK's c-factor is an example of this. HEAVY ATTACK is a nonlinear optimization because the probability of killing a target is nonlinearly related to the number of allocated sorties x by the formula

$$P_K = \min\{(1 - \exp(-c\alpha x))/c, \quad 1\}; x \geq 0,$$

where α represents a sortie's lethality against a particular kind of target and c is a parameter between 0 and 1. When c=1, the formula represents the familiar exponential relationship characteristic of independent shots. This is what one would expect with zero information about target status in between shots. In the limit as c approaches 0, the expression is $\min\{\alpha x, 1\}$, a linear increase that is possible only when the information system is so good that no sortie is ever directed against a target that has already been killed. Intermediate cases are presumably represented by the open interval 0<c<1. The trouble with this approach is that c isn't well described except in the extremes, since there is no combat theory that would permit c to be measured. After all, c^2 is a number between 0 and 1 just as much as c is, but (except in the extremes) the formula would certainly behave differently if c^2 were substituted for c. A domain expert's response to the question "What is the c-factor for tanks being attacked by A-10's?" is likely to be "Huh?", and it is very difficult to make the question more meaningful. This makes the c-factor an IRK, and HEAVY ATTACK has been criticized for having one (Brown, Coulter and Washburn [1994], Evans [1996], Yost [1996]). Models that deal with information are especially inclined to include IRKs, since the extremes are especially simple compared to intermediate cases. This is true of optimization models, but it is also true of other less abstract types. The documentation for TACWAR, for example, defines C2 capability to be "A measurement of a unit or element's ability to receive, assimilate and act on information and orders from internal and external resources" (JCS [1997]). Huh? How would we measure that?

The argument for IRKs is that they can be adjusted until results agree with the expectations of experts, and that anyway there isn't any other choice. This is a powerful argument that should not be rejected until better alternatives are available. The counterargument is that it is hard to learn much from a model if its IRKs are simply adjusted to fit preconceptions, and that IRKs are dangerous in being an invisible way to adjust a model to prove whatever political necessity dictates. This is a powerful counterargument. As a practical matter we are stuck with IRKs for the moment, but we need to develop better alternatives.

Probabilistic Dynamic Programming

I use this term rather than Markov Decision Process because I do not mean to imply stationarity in time (Winston[1994]). This class of model (PDP) would seem to be natural for problems where information must be valued. The decision maker's action set can include both shooting at a target ("bangs") and seeking information about its status ("bits"), each of which can be more or less costly, and furthermore the resulting

policies are provably optimal. Information has all the subtle usages and implications that it does in reality, rather than being modeled as a tax on bangs. A good example is Aviv and Kress [1997], who use PDP to determine optimal firing policies in a small problem where the information received after each attack is error-prone. However, I can give no example of a large-scale military model employing the idea. The basic problem is Bellman's "Curse of Dimensionality." Suppose that there are 100 targets, each of which is either live or dead. The total number of ways of describing the status of the target set is then 2^{100}. This is far too large a number for any computer in the foreseeable future to deal with by exhaustion, and it would be still larger if there were more than two possibilities for the status of each target.

In spite of this limited record to date, PDP deserves more attention by us analysts for problems of the sort under discussion. Its virtues will sooner or later cause us to find some way of outwitting Bellman's curse by making special assumptions. Yost [1998] has described a large air-to-ground weapon allocation problem that can be attacked by a hybrid method involving both Linear programming and PDP. The PDP generates good policies by dealing with individual targets, and the Linear program finds an optimal mixture of the policies. Dynamic programs deal with *time* efficiently, and time is of such central importance to the value of military information that PDP will eventually become attractive.

Game Theory

Much of the uncertainty in military operations, and therefore much of the need for information, comes from the unpredictability of enemy activities. The theoretical justification of this unpredictability is again due to von Neumann [1928], who proved that every finite two-person-zero-sum (TPZS) game has a value as long as randomized strategies are permitted. It would seem that this particular source of uncertainty could be handled by formulating the decision problem as a game, then using the game's value to make tradeoffs between bits, bangs, and bucks. The game value's automatic adjustment of enemy tactics to account for structural changes is particularly attractive, since that valuable feature is otherwise hard to come by.

Markov games are a subclass that is especially well suited to model military decision problems extended in time, and have already enjoyed a certain amount of success as abstract models of air warfare. The original paper along these lines is Berkovitz and Dresher [1959], who consider a campaign of eight periods, in each of which an air force must be partitioned into counter-air, air-defense, and ground-support components in ignorance of similar decisions made by the other side. A more modern representative of the class would be OME-IV (STR Corporation [1995]), in which the roles and resources of the opponents are considerably generalized.

Unfortunately, Markov games also run up against Bellman's curse. In fact, the solution technique usually employs a functional equation very similar to the one employed in PDP. We are therefore forced to come to the same conclusion, namely that Markov games are a potentially useful model class that will require cleverness in formulation in order to avoid computational roadblocks.

TPZS games are normally represented as matrices where the row player tries to maximize and the column player tries to minimize. There is nothing wrong conceptually with letting the rows and columns be functions that represent different ways of processing and acting on information, so we can again ask the question "Is there any kind of information for (say) player 1 for which the game value goes down if he has it?" If so, then we have discovered disinformation. Once again, however, there is no such thing. To be more precise, there is no such thing *as long as both sides know the rules of the game*. Added information for player 1 can make player 1's payoff decrease in circumstances where he *thinks* he is playing game A but is really playing game B, while player 2, knowing that the game is B, provides signals that lead to player 1 taking actions that would be smart in game A, but are not in B. The notion of disinformation is thus tenable, but seems to require a rather odd and hard-to-model circumstance for player 1. Since game B is actually the case, why would player 1 not even suspect the possibility? It seems unlikely that a theory of disinformation can be invented for circumstances like this. At the moment, we really don't have a theory that can serve to define "disinformation", even though we sometimes use the term. We literally don't know what we are talking about, and that is part of the crisis.

Evaluative Mathematical Models

All of the model types described above are *normative*. Perhaps information would be easier to represent accurately if our analytical goals were less grand. The *evaluative* models considered here and

below do not attempt to automatically manipulate a decision, so they are spared many of the constraints and artificialities that normative models require. The title of this section is meant to restrict interest to evaluative models that include neither random number generators (Monte Carlo models are the subject of the next section) nor human subjects, but it is not meant to imply determinism. The ultimate purpose is still to make decisions, just as when normative models are employed. The difference is that the burden of manipulating the decision is shifted from the model to the analyst.

The classic example of an evaluative mathematical model is a system of ordinary differential equations (ODEs), especially Lanchester systems where the state variables represent surviving numbers of combatants. ODE systems are appealing because systems with thousands of equations can be easily solved as a function of time, and time sensitivity is particularly important in modeling information. Cebrowski [1997] states that "The principal utility of information superiority is time—the immense advantage of being able to develop very high rates of change." ODE systems are ideally suited for such situations.

ODE systems are also appealing because they allow a graphical description, at least for a small number of state variables. Figure 3 is an example with six state variables, each named inside a circle with the initial value shown in the lower half. The arrows are labeled by formulas that show the *rate* at which one kind of thing is converted to another. For example the rate at which BL ("Blue Live") units are converted to BD ("Blue Dead") units is b(LU+LI), with LU+LI being the number of red Live Unidentified units plus the number of red Live Identified units (red names omit the initial R to limit the length of variable names). In other words, all live red units fire at blue in the classic Lanchester aimed-fire manner, with b being the number of blue casualties per red man-day. The dynamics among the four red state variables are more complicated. LU units must be converted to LI units by an information system before they can be killed, and LI units may again become LU units if they aren't killed quickly enough. LI units may also become either DL (Dead looking Live) or DD (Dead looking Dead). Since blue cannot distinguish LI from DL targets, the appropriate fraction of blue's fire is devoted to DL targets. This reduces the rate at which LI units are killed, but has the side benefit of gradually converting DL to DD. Once the six technological parameters (a,b,p,q,s,t) are specified, the information in Figure 3 determines the course of the battle between red and blue. Furthermore, the six equations can be easily solved dynamically in a spreadsheet. For example, using a time step of .1 day, the number of blue survivors on day 4 if the parameters are (.5, .2, .4, .4, 2.0, .5) is 61.63, or it is 50.51 if the parameters are (.3, .2, .4, .4, 5.0, .5). The second case can be thought of as one where blue spends more on information systems (the surveillance parameter s is increased from 2 to 5) at the cost of spending less on firepower (the firepower parameter a is reduced from .5 to .3). By this measure spending more on information is a mistake. The reason is that s is so high even in the first case that all of the LU units are converted almost instantly to LI units. Blue's problem is one of firepower, rather than
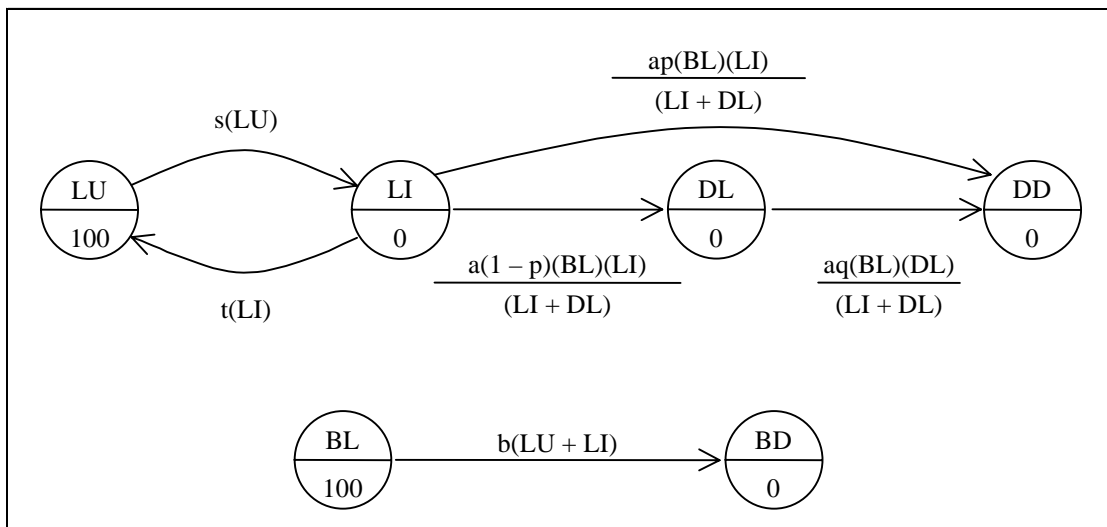

Figure 3. Information Warfare by Ordinary Differential Equations

surveillance. The problem could be rigged, of course, to make the decision go the other way. The point is merely that such systems can be the basis for tradeoffs between bits and bangs.

ODE systems are easy enough to solve that the analyst's problem is more likely to be one of acquiring data than one of computation. My colleagues Pat Jacobs and Don Gaver have been experimenting with information models similar to Figure 3 but larger, and there is no computational reason why models with thousands of state variables could not be solved. EVA will necessarily be rampant in these models, since the state variables usually represent quantities that ought to be integers, but there is a remedy in that the rates of change can always be interpreted as the rates of a nonhomogeneous Poisson process (NHPP). An NHPP model requires exactly the same information as an ODE model, but interprets it probabilistically, so a Monte Carlo simulation of the corresponding NHPP could be used as a kind of verification test for an ODE system. The hope would be that the ODE quantities can be interpreted as approximate expected values of the corresponding NHPP random variables. The automatic availability of the NHPP counterpart is another argument for the use of this kind of model.

Small ODE systems could also be generalized to include certain simple dynamic decisions, in which case the theory of Optimal Control is available. This theory has been applied to the problem of dividing fire optimally between various classes of targets, and there is no reason why flexible information warfare assets could not be treated similarly. There could be a control variable named "fraction of UAVs devoted to surveillance", for example, with the rest being devoted to some other function. The results might very well be illuminating. But the associated computational problems are hard even when only a few state variables are involved. Excepting situations that are so specialized that an analytic solution is possible, consideration of systems involving hundreds or even tens of variables is not in the cards.

The assumption in Figure 3 is that Red targets are always in one of only four states. As long as this is true, it is easy to explore the effects of changes in the transition rates. Unfortunately, it is easy to imagine circumstances where the number of states would have to be much larger than four or even several thousand. If targets have to be located as well as identified, then location accuracy (a standard deviation, perhaps) ought to be a target property, and it ought to be updated when multiple sightings occur or (for moving targets) when time passes. This would be easy enough if targets were tracked individually, but an ODE system simply *counts* the number of targets in a category. One might subdivide the LI category according to location precision. How fine should this partition be, and what should be done about transitions *out* of all the resulting states, as well as *between* them? To continue this line of thought, what if the *identity* of targets is not necessarily clear when a detection is made? Shall we introduce states such as "targets that are mobile missile launchers with probability .72 and school busses with probability .28"? Surely not. Questions such as these expose ODE systems for what they are: aggregated, low resolution models that are useful for investigating quantity questions when information is so simple that decision rules are obvious, but which suffer from state space and policy explosion when more subtle questions are posed.

ODE systems as described above model time as a continuous variable, but nothing essential changes if time is made discrete.

Monte Carlo Simulation

This modeling technique makes it possible to avoid EVA, since random variables can be realized using a random number stream. Discrete random variables remain discrete, and the variability of combat results is evident because every replication is different. The computational cost is in the need for replication, since it is only by repetition that the distribution of results can be understood.

Monte Carlo is a very compelling paradigm. The variability of combat results is deliberately and scientifically included, and the avoidance of EVA permits a conceptual simplicity that is often impossible in mathematical models. It should therefore come as no surprise that many modern battle models are Monte Carlo simulations. This conceptual simplicity extends to modeling the effects of information, since it is possible and useful to distinguish between truth and perception. The true and perceived versions of properties such as location, identity, and status can each be modeled explicitly.

The fusion of several sources of information is one area that illustrates the possibilities in a Monte Carlo model. Suppose we have a target moving in two dimensions, with information about its position being obtained occasionally from sensors. The fusion/extrapolation scheme might be simple dead reckoning from the latest sensor report, or it might be some more sophisticated scheme such as Kalman filtering. If the sensor report might have been generated by one of several targets, it can be associated with the nearest

estimated target location, or with the target that minimizes some other report-to-target score. If a weapon is fired at the target, the effects of the weapon can depend on the fact that the true target location is known, at least to the analyst. A simulation would ordinarily take care to make sure that the true target location is not used in the aiming algorithm except through sensor reports , but the alternative of violating this principle for efficiency's sake is always available if needed. All of these possibilities are present in the Naval Simulation System (METRON [1997]), which offers several levels of fusion to the analyst. The corresponding model types described earlier would all be complicated conceptually and computationally by the need to construct a probability distribution for the same ground truth that is known in the simulation.

Monte Carlo simulations have similar advantages in many other aspects of warfare. These aspects include attrition, movement, logistics, and many others. They capitalize effectively on regular improvements in computer technology, and deservedly hold the central place in warfare modeling. But even simulations are imperfect tools for demonstrating the value of information. Real military operations often hinge on good estimates of the situation. Many man-hours may be expended in estimating the one situation that actually unfolds, and even so the estimates are sometimes faulty. Partly because of the need for replication and partly because of the difficulty of capturing actual human decision making (or optimal decision making) in a simulation, this process must be heavily abstracted in a Monte Carlo model. The abstraction process is unlikely to capture optimal decision making, and therefore risks understating the effectiveness of information systems that depart strongly from a baseline. Consider Unmanned Aerial Vehicles (UAVs) used as reconnaissance platforms. One thing that distinguishes UAVs from most other such platforms is that they are controllable in real time by a human on the ground who sees whatever the UAV sees. This gives them an ability to investigate interesting observations, and through those investigations an ability to identify certain targets that would merely be detected by (say) a satellite. The usual sweep rate model will understate UAV effectiveness, but will still be tempting because UAV controllers are hard to simulate. Besides, what if we have not previously distinguished between detection and identification, as many simulations do not? Is the whole conceptual basis for the simulation to be changed just to accommodate UAVs? Maybe this is a good use for an IRK — just multiply sweep widths by a user adjustable "identification factor" (just kidding).

The ability to distinguish between truth and perception has been given as one of the advantages of simulation, but it turns out that truth is much easier to keep track of than perception. The fog of war is likely to produce vague perceptions, and this vagueness must somehow be represented compactly without losing the essence of the matter. This is not always easy to do. Any single-hypothesis correlator-tracker of the type that is typically implemented in a simulation, for example, is compelled to think of the distribution of each target's two-dimensional location as a bivariate normal random variable. Bayes' theorem, on the other hand, could very well produce a probability map resembling a topographic map of (say) Idaho, rather than a single distended volcano, when there are significant false alarm or false association rates. It is one thing to ignore the difference in considering yesterday's question of whether the Navy has enough firepower to support a Marine landing, and another thing entirely to ignore it when the question concerns the C4ISR system. The value of information depends on how it is represented, processed and used. We aren't good at efficiently capturing how humans do that, either actually or ideally. We certainly aren't good at it under the additional constraint that the fusion algorithm must be fast enough to risk incorporating in a simulation.

It doesn't help that our subject is the value of *military* information. One lesson from Game theory is that the enemy is motivated to present information to us that is deliberately confusing, so a tracker that might work well in tracking whales may not work so well in tracking submarines. Difficult decision making situations are the norm, rather than the exception, when the enemy acts optimally.

Man-in-the-loop, etc.

Humans are still very effective decision makers, especially in dynamic, competitive, foggy situations such as combat. It is tempting to give up on attempts at decision making in the abstract, and simply let humans make decisions on the basis of simulated information. Humans will have to make the decisions in actual combat anyway. If an information system lets humans produce better battle outcomes, then it ought to be given credit for the improvement. No improvement, no credit.

There is certainly a role for this approach, since the premises are good ones, but it is no panacea. There is a need for replication, just as there is in Monte Carlo simulation, on account of the inherent variability in situations where information is important. Replication is always expensive when humans are

involved, and there are special experimental difficulties when the fog of war is simulated. The fog must be essentially different on each replication if the same human is used in several replications, since otherwise the human may simply learn what fog to anticipate. Even if the fog is sampled independently on each replication, the human may still learn about the *distribution* of fog by repeatedly playing the game. One might make several warmup runs where the human learns about the distribution, or one might make the distribution so close to the human's prior experience that no warmups are necessary. In either case the object is to produce humans who are experts in one sense and naïve (non-anticipatory) in another. This is easier in principle than in practice. If the fog is historical or for other reasons difficult to sample, one might use different experts one time each to prevent anticipation. But all "experts" are not equally expert, so there is a danger of having the outcome be determined more by the expert than by the information system. In every direction there are statistical issues that are expensive to deal with. As a result, man-in-the-loop models are mainly useful for testing qualitative changes, rather than variations on a theme. They might be useful for finding out whether UAVs make a difference, but not for settling design questions or measuring marginal changes.

When several humans with different goals are involved, a man-in-the-loop simulation becomes a wargame (see Figure 1). Wargames and exercises involve some new issues in addition the to the statistical ones outlined above, but nothing changes the basic conclusion that models that rely on human decision makers are bound to entail statistical problems that limit their precision. As a result, they are best employed in comparing qualitatively different systems, rather than in investigating marginal tradeoffs.

In Summary

Our profession is poorly prepared to deal with Information Warfare, especially for the kind of marginal questions that are involved in trading off IW systems with conventional ones. We have all seen briefing charts where a lot of arrows point into a circle labeled "FUSION", and one arrow labeled "fused information" comes out. It's not that easy, and we would be wise to begin by admitting it. But the situation is not hopeless. While none of the techniques surveyed above are clearly indicated as the technique of choice, every one of them has been applied to problems involving information evaluation. With more effort put into sharpening and adapting our tools, we can possibly deal with the crisis. In the meantime, let's at least employ EVA, TRA, and IRKs more reluctantly, recognizing that they are particularly dangerous in models that claim to measure the contribution of information systems.

References

Aviv, Y. and Kress, M. (1997) "Evaluating the Effectiveness of Shoot-Look-Shoot Tactics in the Presence of Incomplete Damage Information," *Military Operations Research*, **3**, 1997, 79-89.

Berkovitz L. and Dresher, M. (1959) "A Game-theory analysis of tactical air war," *Operations Research*, 7, 599-624.

Cebrowski, A. (1997) "Sea change," *Surface Warfare*, November.

Brown, G., Coulter, D., and Washburn, A. (1994) "Sortie Optimization and Munitions Planning," *Military Operations Research*, **1**, #1, 13-18.

Evans, D. (1996) "Bomb Damage Assessment and Sortie Requirements," *Military Operations Research*, **2**, #1, 31-35.

Griggs, B., Parnell, G, and Lehmkuhl, L. (1997) "An Air Mission Planning Algorithm Using Decision Analysis and Mixed Integer Programming," Operations Research **45**, #5, 662-676.

Joint Chiefs of Staff, Joint Vision 2010, Joint Staff Pentagon, Washington, DC.

JCS TACWAR Project Officer (1997) "TACWAR Integrated Environment C2 Analyst Guide Version 5.0," page 1-2.

METRON, Inc. (1997) "Naval Simulation System Analysis Guide (Draft, version 1)," 15 Feb, 4-29 to 4-32.

Scott, W. (1998) "Computer/IW Efforts Could Shortchange Aircraft Programs," *Aviation Week and Space Technology*, January 19, p. 59.

Schachter, R. (1986) "Evaluating Influence Diagrams", *Operations Research,* **34**, 871-882.

Shannon, C. (1948) "The mathematical theory of communications," Bell System Technical Journal, **27**, 379-423 and 623-656.

Sherrill, T. and Barr, D. (1995) "Exploring a Relationship between Tactical Intelligence and Battle Results," *Military Operations Research*, **2**, #1, 17-33.

STR Corporation (1995) "OME-IV Analyst's Manual (CDRL A010)," AF contract F33657-93-C-2434, June.

Yost, K. (1996) "The Time Strike Munitions Optimization Model," Naval Postgraduate School Technical Report NPS-OR-96-001.

Yost, K. (1998) "Solution of Sequential Weapons Allocation Problems with Imperfect, Costly Information," *Sixth INFORMS Computer Science Technical Section Conference*, January 9.

von Neumann, J. (1928) "Zur Theorie der Gesellschaftsspiele," Mathematische Annalen, **100**, 295-320.

von Neumann, J. and Morgenstern, O. (1944) *Theory of Games and Economic Behavior*, Appendix A, Princeton University Press.

Winston, W.(1994) *Operations Research Applications and Algorithms* (3rd ed.), chapter 21, Duxbury Press, Belmont CA.