11TH INTERNATIONAL COMMAND AND CONTROL RESEARCH TECHNOLOGY SYMPOSIUM
(11TH ICCRTS)
Symposium Theme:
COALITION COMMAND AND CONTROL IN THE NETWORKED ERA

Applying Models Of Decision Making Under Uncertainty To Simulations Of Military Command And Control Systems

In the topical area of:
C2 CONCEPTS AND ORGANIZATIONS

Sam E. Middlebrooks, Ph.D.
Human Research and Engineering Directorate
U.S. Army Research Laboratory
Fort Hood, Texas
Sam.Middlebrooks@otc.army.mil

Brian J. Stankiewicz, Ph.D.
Department of Psychology & Center for Perceptual Systems
The University of Texas at Austin
bstankie@mail.utexas.edu

Point of Contact:
Dr. Sam E. Middlebrooks
Human Research and Engineering Directorate
U.S. Army Research Laboratory
91012 Station Avenue
Fort Hood, Texas76544-5073
254-288-9572
Sam.Middlebrooks@otc.army.mil

Keywords: Partially observable Markov decision process (POMDP), Bayesian statistics, command and control (C2), command and control system (C2S), computer simulation, discrete event simulation, decision making under uncertainty (DMUC).

Applying Models Of Decision Making Under Uncertainty To Simulations Of Military Command And Control Systems

Sam E. Middlebrooks, Ph.D.
Human Research and Engineering Directorate
U.S. Army Research Laboratory
Fort Hood, Texas
Sam.Middlebrooks@otc.army.mil

Brian J. Stankiewicz, Ph.D.
Department of Psychology & Center for Perceptual Systems
The University of Texas at Austin
bstankie@mail.utexas.edu

Abstract.

Most military decision making requires a sequence of actions. These actions may include aspects of intelligence gathering, troop movement, artillery fire, etc. Typically these actions are tied to a specific goal that might include securing a region or disrupting the enemy forces. Furthermore, when making the necessary decisions to reach the specific goal, there may be a great deal of uncertainty about the situation. Where exactly are the enemy troops? What is their objective? The use of Bayesian statistics makes it possible to compute optimal performance for military-like situations. This research develops a model that provides the theoretical best performance that can be achieved in the task. An example task that can illustrate this condition is a seek-and-destroy mission in which there is an enemy at an unknown location within a certain region. Using reconnaissance that is imprecise and artillery that is not always accurate the mission is to destroy the enemy. Each action (intelligence gathering and artillery) comes at a specific cost. Furthermore, succeeding at destroying the enemy generates a reward and declaring "Mission Accomplished" when the enemy is still alive generates a significant cost.

The current paradigm for the description and understanding of the nature of command and control (C2) system (C2S) operations and performance within the U.S. Army is undergoing a radical change. Tactical battlefield C2 is extremely complicated to orchestrate and conduct in an effective manner. With the introduction of a myriad of new information systems, sophisticated new weapons with unprecedented capabilities for lethality, new requirements for battlefield integration, and the total reorganization of force structures into a new modular concept, the need for effective understanding of how this force structure can work effectively as a system entity increases dramatically. Indeed, the C2S has become complicated to the point as to escape the ability for intuitive understanding of how individual components or subsystems can improve or degrade the operation of the overall system. The goal of this research is to understand the cognitive limitations associated with sequential decision making with uncertainty in these types of situations through predictive computer simulation. By combining empirical research investigating optimal decision making under uncertainty with evolving simulations of military command and control, the potential now exists to correlate optimal decision making performance with actual outcomes on the battlefield. The benefit of these simulations will be to enable the evaluations of new command and control subsystems that take into account optimal decision making performance as an evaluation metric.

1. Introduction.

The traditional paradigm for the description and understanding of the nature of command and control (C2) system (C2S) operations and performance within the U.S. Army is currently undergoing a radical change. The U.S. Army Field Manual FM 6-0 (Army, 2003) defines the C2S as "the arrangement of personnel, information management, procedures, and equipment and facilities essential for the commander to conduct operations." Tactical battlefield C2 is an extremely complicated action to orchestrate and conduct in an effective manner in its own right. However, with the introduction of of new information systems such as the Army Battle Command System (ABCS) (Army, 2002), sophisticated new weapons now exist with unprecedented capabilities for lethality and requirements for battlefield integration. As they are contributing to a total reorganization of force structures into the new modularity concept, the need for effective understanding of how this system can work effectively as a system entity increases exponentially. The fact is that the complexity of the modern C2S has surpassed the ability for an intuitive understanding of how individual components or subsystems can improve or degrade the operation of the overall system. This situation poses the question of how to further develop and improve the performance of the C2S without making changes that might actually degrade its effectiveness. From this it becomes apparent that some systematic approach is needed to predict and evaluate the effects of changes, additions, and improvements on this system will have on its overall ability to conduct battlespace management.

1.1. The Command and Control System's Demands for Decision Making.

Previous research on this topic (Middlebrooks, 2003; Middlebrooks et al., 1999a; Middlebrooks and Stankiewicz, 2006; Wojciechowski, Plott, and Kilduff, 2005) has developed a paradigm for the systematic evaluation of the C2S from the system level viewpoint. A basic premise of this approach is that all observable characteristics of live tactical operations centers (TOCs) in the field can be used in the development of quantitative predictive models of various system components for use in a simulation of the complete C2S. Some of these characteristics include such things as the quantity and quality of communications messages through the TOC from various digital systems, quality and timeliness of intelligence information about the enemy, numbers and expertise levels of team members present in the TOC at any given time, individual and group interactions, physical layout of the TOC, and environmental conditions, to name a few. However, a striking limitation of this approach is in its inability to simulate and predict cognitive performance of individuals and teams in areas such as situation awareness, knowledge elicitation, error generation, individual vs. team performance, and decision making. An ability to predict the optimal decision required for success can be extremely useful as a performance measure for use in describing the overall effectiveness of the system.

1.2. The Model of Optimal Decision Making.

This research integrates basic research in decision making being conducted at the Psychology Department of the University of Texas with applied research in Unit of Action (UA) TOC operations being conducted at the Fort Hood Field Element of the U.S. Army Research Laboratory. It is the goal of this research to develop predictive simulations of the C2S UA performance that can be used in the evaluations of changes to the system or the addition or modification of system subcomponents. For example, what might the effect be on the overall ability of the UA TOC to conduct battlespace management from the addition of a new intelligence system that allows information about the enemy to have a maximum age of one hour before it becomes obsolete versus a maximum age of 4 hours? One intuitive conclusion that could be deduced from this new capability is that it would significantly increase the commander's understanding of the enemy situation because the information is always more

current than before. However, this intense stream of information might cause the commander to become more focused on the instantaneous situation on the battlefield and lose situation awareness of longer term developments with a resulting degradation in the ability to make effective decisions on how to react to the threat. The effective ability of predictive simulations of these types of environments is based on how well they account for the myriad of variables stemming from both physical activities and the human's cognitive ability to react to those variables. This current research is a first step in allowing simulations of system performance to account for limitations in human cognitive performance abilities.

2. Method.

The methodology used in the operation of this model involves two components. The first is what is defined as the belief vector. This represents the current knowledge that the operator has about the system. Utilizing this knowledge the operator decides what to do next or what action to take in the pursuit of the mission goal. This model is generalizable to many different decision oriented situations where the operator is confronted with a goal directed task. In the pursuit of this task the operator may decide to seek information about the condition of the current situation, they may decide to take some action to achieve the goal, and at some point they will make a decision that they have achieved the goal thus terminating the decision mission sequence. An example of this scenario is provided in the hospital emergency room where a doctor is faced with the mission goal of successfully treating a sick patient. The doctor may administer a medication (take an action) or may perform a medical test (seek information) to attempt to identify the patient's condition. A complete series of medications and tests may be performed before the doctor achieves the belief that the patient has been cured. If the doctor terminates the sequence before the patient is cured the patient may die. If the doctor prolongs the sequence beyond the point that the patient is cured a substantial unjustified cost is the result. This medical sequence is an illustration of gathering information and taking actions until a belief is achieved that the goal has been reached. Another example is provided in the military context where a ground force commander is given an order to seek out and find and destroy an enemy that is located at an unknown location. The commander may seek information, an example being to fly an unmanned aerial vehicle (UAV) reconnaissance mission, or they may take a direct action to destroy the enemy, an example being to fire artillery at a suspected location occupied by the enemy. An entire sequence of UAV and artillery missions may be performed in some goal directed order until the commander achieves a belief that the enemy has been destroyed. At this point the commander decides that the mission is complete because the enemy is believed destroyed and decides to terminate the action. Modeling this process is accomplished through the use of Markov decision process analysis to determine the belief vector and the use of conditional probability logic to determine the next action to take based on an evaluation of the current belief vector.

2.1. Determination of the Belief Vector.

It is important to recognize that most decisions that are made are not "one off" decisions in which the decision is made and then the rewards reaped or the punishment endured. Instead, most decisions that are made have future ramifications and affect the options and decisions that are available later. One challenge faced by any decision maker is the uncertainty that the decision maker has about the true state of the system. In most circumstances the true state of the system is unknown or hidden. That is, it cannot be directly observed. For example, in military decisions often there is uncertainty about an enemy's position, strength and morale. Given that the true state is hidden, there are things that can be done to reduce the decision maker's uncertainty about these states. For example, the decision maker may attempt to determine the enemy's position by sending reconnaissance to a location where the enemy is believed to be

located. When the reconnaissance returns with either an "enemy sighted" or "enemy not sighted" report the decision maker must update their belief about the location of the enemy.

If the observations and actions were all deterministic, updating a belief would be relatively simple. However, in almost all conditions, the observations and actions are probabilistic. That is, the probability of getting an observation given the true state of the environment is not necessarily 0.0 or 1.0. Or, in the example above, there is a certain non-zero probability that the reconnaissance mission was sent to the right location and will miss the enemy and send a report of "enemy not sighted". Furthermore, there may be a non-zero probability that the reconnaissance mission falsely sent a report of "enemy sighted" (or false-alarmed) when the enemy was not actually at the location.

Given that the observation and actions are probabilistic, updating a belief given an observation and an action can become cognitively difficult. Furthermore, evaluating the added benefit of a specific piece of equipment that changes these probabilities can also become difficult. This research focuses on a task that is commonly faced by decision makers in the military — namely, the seek-and-destroy task. In this task the decision maker is trying to localize and destroy an enemy within a specific region. At the decision maker's disposal are actions that allow information to be gained about the true state of the system (i.e., the location of the enemy) in addition to changing the state of the system (for example, moving the enemy from being at a specific location to the state of Destroyed). The former actions are reconnaissance actions and the latter are artillery actions. The outcomes of these actions are probabilistic. That is, reconnaissance actions will not always detect the enemy when a sensor is sent to the enemy's location. Furthermore, the reconnaissance may also falsely report that the enemy is seen at a location in which the enemy is not located. In addition, the artillery will not always kill the enemy when striking it which is characterized as moving the enemy from being alive at a certain location to the "Destroyed" state.

2.1.1. The Optimal Observer.

To best evaluate performance in a task that leads to uncertainty and probabilistic actions, it is useful to define the optimal performance within the task. The optimal performance can be calculated using Bayesian statistics. However, due to the nature of the current type of task, simple Bayesian statistics are insufficient. That is, with simple Bayesian statistics the likelihood of the true state of the system can be optimally estimated. However, this likelihood does not indicate what action should be selected. In order to do action selection, not only must the current state be calculated given the previous actions and observations, but also the optimal action to be performed in a given belief state must be calculated. This is done to verify whether a belief state has a particular probability distribution across all of the possible states in the environment.

A variation on classical Bayesian statistics that may well add some additional predictive power for sequential decision making under uncertainty is the *Partially Observable Markov Decision Processes* (POMDP) (Cassandra, 1998; Cassandra, Kaelbling, and Kurien, 1996; Cassandra, Kaelbling, and Littman, 1994; Kaelbling, Littman, and Cassandra, 1998; Sondik, 1971). By defining the *State Space, Observation Vector, Transition Matrix* and the *Reward Structure*, the expected reward for a particular action can be computed. In the following sections a description of these actions is provided. In addition, a description of how to optimally update an individual's belief (*Belief Updating*), given these definitions, is provided.

An Ideal Observer Model provides optimal performance given the information available in the task. Typically ideal observers <u>are not proposed as models of human cognition</u>. Instead, the ideal observer provides a benchmark by which to compare human performance. More specifically, these models illustrate what optimal performance should look like. When human

performance matches that of the ideal-observer model, it can be concluded that the human is making use of all of the information in the task. When the human under-performs the ideal observer, specific discrepancies between the human data and the ideal data may illuminate the constraints imposed by the human information-processing system.

Ideal observer analysis is not new to this research and has been previously used to understand perceptual functions from the quantum limits of light detection (Hecht, Shlaer, and Pirenne, 1942) to many forms of visual pattern detection and discrimination (Geisler, 1989), to reading (Legge and Hooven, 2002; Legge and Klitz, 1997) object recognition (Liu and Knill, 1995; Tjan, Braje, Legge, and Kersten, 1995; Tjan and Legge, 1998) eye movements (Najemnik and Geisler, 2005) and also in reaching tasks (Trommershäuser, Gepshtein, Maloney, Landy, and Banks, 2004).

2.1.2. Defining the State Space.

In all problems that are solved using a POMDP architecture, there are a set of *states* that the problem can be in. In a POMDP problem, the true state ($State_{True}$) is not directly observable (i.e., it is hidden). For the work in this project, the hidden state is defined as the enemy's current position within a 2 dimensional grid. This grid of location state spaces is supplanted by an additional "Destroyed" state that the enemy could transition into following an action to destroy it such as an artillery strike at its current position. Thus, the size of the grid can be characterized as:

$$(X + Y) + z, \text{ where both } X \,\&\, Y > 1 \text{ and } z = 1.$$

In this case, X is the size of the location grid in the X dimension, Y is the size of the location grid in the Y dimension, and Z is the dead state which is always equal to 1. Using this nomenclature a 5x5 location grid state space this yields a 26 state space. A 4x4 location state grid gives a 17 state space, a 3x3 location state gives a 10 state space, a 2x2 location state gives a 5 state space, and so on. These differing state space sizes are illustrated in Figures 1 and 2.

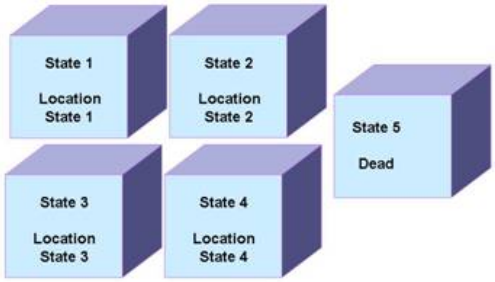


Figure 1 – 2x2 Location State Space

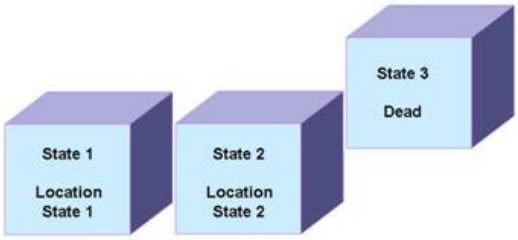


Figure 2 – 2x1 Location State Space

2.1.3. Defining the Belief Vector.

Although the true state is hidden, the operator typically employs actions and observations that provide information about the true state of the problem. In a 26 state space as shown in Figure 1, the operator can fire artillery at a specific position or conduct reconnaissance to a particular location within the environment (i.e., one of the 25 location states). In this model a reconnaissance action provides two possible observations: "*Enemy Sighted"* or, "*Enemy Not Sighted"*. An artillery strike only provides an observation of "*No Information"* meaning that while the artillery strike was conducted this action provides no information on the resulting condition of the enemy resulting from it as only a reconnaissance mission can observe the condition of the enemy. This replicates the fact that the artillery firing unit does not see the effects of its fires because it is an indirect firing unit and is not able to see where the artillery rounds fall. It must rely on forward observer assets to report back what is termed 'battle damage assessment (BDA)' in military jargon. An illustration of a reconnaissance asset might be a forward observer (FO) on the ground or an unmanned aerial vehicle (UAV) that provides the BDA.

This model assumes that the observer has a belief probability between 0% and 100% that the enemy exists in one of the states within the total state space at any given period in time. It is noted that residing in one of the location states is mutually exclusive of residing in the destroyed state. I.e., if the enemy is 'alive' in one of the location states it cannot be 'dead' in the destroyed state and vice versa. The destroyed state is considered to be an "absorbing state" in that once the enemy transitions from being alive in a location state to being dead in the destroyed state it cannot return to a live location state. The set of the belief probabilities for all the states in the state space is termed the **belief vector** (BV). For the simple 3 state space example in Figure 3 the BV could be represented as:

$$[B_{\text{Location 1}}, B_{\text{Location 2}}, B_{\text{Destroyed}}]$$

For this case assume that the enemy is alive with an equal probability of residing in one of the location states. Thus, the BV becomes:

$$[0.5, 0.5, 0.0].$$

2.1.4. Defining the Transition Matrix.

In this Seek and Destroy problem, the observer, in this case the military commander, has a number of possible different actions that can be taken. For actions in a 5x5 location grid state space there are 25 possible reconnaissance actions (one to each of the 25 locations in the environment), 25 possible artillery actions (again, one to each of the 25 locations within the environment) and the action to declare 'mission complete' when it is believed that the enemy has been destroyed for a total of 51 possible actions. The transition matrix defines the probability of the resulting state given that the observer executes a particular action in a specified state (i.e., $p(s'|s,a)$) where s' is the resulting state, s is the existing state, and a is the action taken. In the static form of the Seek and Destroy problem there is only one state transition that can occur for any of the possible actions.

When the commander fires artillery to where the enemy is located, the enemy may be killed with a certain probability which will cause it to transition into the "Destroyed" state. Sample probability estimations for use in this discussion are shown in Tables 1-3. These values are estimates only and are not to be construed as factual. The probabilities to be used in actual analyses are dependent on the scenario conditions at the time of the actual action sequence and are left as model input parameters to be employed during the course of simulation studies that use this model.

Table 1 – The set of actions and their observations for the current Seek & Destroy task.

(The observations for the reconnaissance action are dependent upon whether the enemy is actually within the viewing region of the reconnaissance. Thus, the two possible states are "Enemy Present" and "Enemy Not Present".)

Action	Observation	Condition	Probability
Recon	Enemy Sighted	Enemy Present	0.75
Recon	Enemy Not Sighted	Enemy Present	0.25
Recon	Enemy Sighted	Enemy Not Present	0.2
Recon	Enemy Not Sighted	Enemy Not Present	0.8

Table 2 – Probabilities for Observation from Artillery Strike

Action	Observation	Condition	Probability
Strike	NoInfo	Enemy Present	1.0
Strike	NoInfo	Enemy Not Present	1.0

Table 3 – Probabilities for Killing Enemy from Artillery Strike

Action	Result	Condition	Probability of Dead
Strike	Probability of Enemy being killed.	Enemy Present	0.75
Strike	Probability of Enemy not being killed.	Enemy Present	0.25
Strike	Probability of Enemy being killed.	Enemy Not Present	0.0

2.1.5. Updating the Belief Vector.

Given an initial probability distribution over the state space, the Observation Matrix and the Transition Matrix, hypotheses can be generated about the current state of the problem following an action and the returned observation. The general form of Bayes' rule (Trueman, 1977; Walpole, Myers, and Myers, 1998; Wine, 1964), as shown in Equation 1, is used as a basis to develop this relationship.

$$P(B_r \mid A) = \frac{P(B_r \cap A)}{\Sigma_{i=1,k} P(B_i \cap A)} = \frac{P(B_r)\, P(A \mid B_r)}{\Sigma_{i=1,k} P(B_i)\, P(A \mid B_i)} \qquad \text{for } r = 1,2, \ldots k$$

Equation 1 – General Form of Bayes' Rule

where,

P	-	probability
A	-	State A
B	-	State B
|	-	'such that', or, 'given'
$\in$	-	Probability theory - all state spaces
$\cap$	-	Boolean AND
$\cup$	-	Boolean OR

2.1.5.1. Bayesian Updating Rule.

Using Bayes' Rule POMDP expressions are derived to simulate sequential decision making with uncertainty. An updated BV is computed by performing a particular action to account for the current condition of the State Space, a Transition Matrix for moving from one state to the next, the application of a BV generated by the results of past actions, and an Observation Vector of

past information elements obtained from previous observations of the state space. This Bayesian Updating Rule is expressed as:

$$p(s'|b,o,a) = \frac{p(o|s',b,a)\, p(s'|b,a)}{p(o|b,a)}$$

Equation 2 – Bayesian Updating Rule

where

s'	=	true state (of the condition being present within the total of all states, S), represented as: $s' \in S$
b	=	prior belief vector
o	=	observation
a	=	action that was generated

nomenclature. The term p(s'|b,o,a), is read as:

The probability of s' being true, 'such that' or 'given' the Boolean conditions of 'b' AND 'o' AND 'a'.

Equation 2 specifies how the ideal observer would update the belief that *s*' is the true state given the prior belief (*b*), the observation (*o*) and the action that was generated (*a*).

2.1.5.2. Update the Belief Vector for 1st Action- Perform $Recon_1$ at $State_1$.

To illustrate the process of belief updating the simple 2x1 location state space of Figure 3 will be utilized. Here, the enemy will be associated with one of three states: $State_1$, $State_2$, or $State_{Dead}$. For this case assume that the enemy is alive with an equal probability of residing in one of the location states. Thus, the BV becomes:

[0.5, 0.5, 0.0]

meaning that there is a 50% probability of the enemy being believed to be in $State_1$, a 50% probability of the enemy being believed to be in $State_2$, and a 0% probability of the enemy being believed to be in $State_{Dead}$. Assume that the enemy actually is located in $State_1$ and that the observer decides to do a reconnaissance to $State_1$ and receives a "Enemy Sighted" observation. The first task is to determine what the observer's belief is resulting from this action for the enemy being located in $State_1$, $State_2$ and $State_{Dead}$.

Using Equation 2 the belief likelihood that the enemy is in $State_1$ is computed. That is, the desire is to compute the belief probability that the enemy is in State1 given the current BV, the current observation, and the current action, or: *p(State1| [0.5, 0.5, 0.0], "EnemySighted",Recon1).*

Computing the separate components of Equation 2:

First, compute $p(o|s',b,a)$ or p("EnemySighted"|$State_1$, [0.5, 0.5, 0.0],$Recon_1$). To do this the likelihood of obtaining an observation of "Enemy-Sighted" if $State_1$ was the true state is needed. From Table 1 above, the likelihood of correctly identifying the enemy as 0.75 is selected.

Next, compute *p(s'|b,a)* or the likelihood of the true state being $State_1$ given the previous belief and the action of $Recon_1$. Because there is no transition possible to $State_{Dead}$ from a recon mission these remain at the prior probabilities of 0.5.

Finally, compute *p(o|b,a)* the likelihood of receiving the observation "EnemySighted" when reconnaissance is made at $State_1$ or p('EnemySighted'|[0.5,0.5, 0],$Recon_1$).

These calculations are:

$\underline{p(o|s',b,a)}$, for State 1 = p('Enemy Sighted' | True State Belief, [0.5, 0.5, 0.0], $Recon_1$)

= Probability of Enemy sighted given Belief that enemy was at State1 and Recon1 was performed at State1 =
= $\underline{0.75}$,
from Table 1

$\underline{p(s'|b,a)}$, for State 1 = p(State1 is true state | ([0.5, 0.5, 0.0], Recon1))

= Probability of State1 being the true state given Belief that State1 is true state and Recon1 showed enemy present in State1 = s' =
= $\underline{0.5}$,
from assumption of equal probability that the enemy has an initial probability of being at one of the 2 location states.

p(o|b,a) , for State1 = p('Enemy Sighted' | ([0.5, 0.5, 0.0], Recon1))

= (Probability of Enemy in State1 x Probability of Enemy Sighted When Present) + (Probability of Enemy in State2 x Probability of Enemy Sighted When Not Present) + (Probability of Enemy Being Dead x Probability of Enemy Sighted When Not Present)
= 0.5x0.75 + 0.5x0.2 + 0.0x0.2 = .375 + 0.1 = $0.\underline{475}$

Thus, $p(State_1 \mid [0.5, 0.5, 0.0], \text{"EnemySighted"}, Recon_1) =$

$$p(s'|b,o,a) = \frac{\underline{p(o|s',b,a)\,p(s'|b,a)}}{p(o|b,a)} = \frac{\underline{0.75 \times 0.5}}{0.475} = \underline{0.7895}$$

Likewise, $p(State_2 \mid [0.5, 0.5, 0.0], \text{"EnemySighted"}, Recon_1) =$

$$= \frac{\underline{0.2 \times 0.5}}{0.475} = \underline{0.2105}$$

And, $p(State_{Dead} \mid [0.5, 0.5, 0.0], \text{"EnemySighted"}, Recon_1) =$

$$\frac{\underline{0.2 \times 0.0}}{0.475} = \underline{0.0}$$

Thus, if the first action is to observe, i.e., perform a UAV reconnaissance mission, at $State_1$, the new belief vector would be:

[0.7895, 0.2105, 0.0].

The interpretation of this BV is that the enemy has a 0.7895 probability of being believed to be at $State_1$ (present in $Cell_1$), a 0.2105 probability of being believed to be at $State_2$ (present in $Cell_2$), and a 0.0000 probability of being believed to be at $State_3$ (Dead). Since these probabilities must account for the total belief state of the operator they must therefore sum to 1.0. Performing this check sum:

CS = 0.7895 + 0.2105 + 0.0000 = 1.0000, therefore Checksum verification passed.

2.1.5.3. Update the Belief Vector for 2nd Action- Perform $Strike_1$ at $State_1$.

Now assume that the second Action is to conduct an artillery strike at $State_1$ which is represented as $Strike_1$. In order to update the BV with the belief that the enemy is in $State_1$ as a result of this new action, determine the probability that the enemy is at $State_1$ given the BV from the first action ($Recon_1$), [0.7895, 0.2105, 0.0], and the new Action, $Strike_1$, recognizing that the only observation from an artillery strike is that the strike was fired which provides the observation 'NoInfo'. Thus, the new probability

$$p(State_1|[0.7985, 0.2105, 0.0], \text{“NoInfo”}, Strike_1)$$

is computed. Calculating the components for the updated BV component for $State_1$ from Equation 2:

p (o|s',b,a) = p ('NoInfo' | True State Belief, [0.7895, 0.2105, 0.0], Strike1) =
= Probability of 'NoInfo' given current Belief Vector and Strike1 was performed at State1 =
<u>= 1.0</u>,
because an artillery strike will always return a report of 'NoInfo' simply meaning that the artillery strike was fired with no other information provided.

<u>p(s'|b,a)</u> = p($State_1$ | [0.7895, 0.2105, 0.0], $Strike_1$)
= Probability of 'Enemy Not Dead' being the true state given the current Belief Vector and the Action of $Strike_1$ being fired.
From Table 3, the probability of the enemy being transitioned to dead if artillery is fired at the location containing the enemy, or in this case, $Strike_1$ being to $State_1$ is = 0.75.
Conversely, if the artillery strike, $Strike_1$, into $State_1$ does not kill the enemy with the enemy remaining in a state of 'Enemy Not Dead', from Table 3 the probability becomes (because of the 3 states in the state space which are $State_1$ (present or not present in cell 1), $State_2$ (present or not present in cell 2), and $State_3$ (belief that the enemy is Dead or Not Dead), the deduction is that the enemy is believed to be alive (i.e., Not Dead) in State3) = 0.25.
Thus, the probability that the enemy's state will not change, or that they will remain alive in $State_1$ is equal to 0.25 times the probability that the previous $Recon_1$ sighted the enemy in $State_1$, or 0.7895 from p($State_1$| [0.5,0.5,0.0], 'Enemy Sighted', Recon1), above.
Therefore, p(s'|b,a) , for $State_1$, following $Strike_1$ is:
p(s'|b,a) = 0.7895 x 0.25
<u>= 0.1974</u>

p(o|b,a) = p(('NoInfo' | ([0.7895, 0.2105, 0.0], Strike1))
= Probability of 'NoInfo' given current Belief Vector and $Strike_1$ being performed at $State_1$ =
<u>= 1.0</u>, because an artillery strike will always return a report of 'NoInfo' simply meaning that the artillery strike was fired with no other information provided.

Employing equation 1 to determine the BV:

$$p(s'|b,o,a) = \frac{p(o|s',b,a)\, p(s'|b,a)}{p(o|b,a)} = \frac{1.0000 \times 0.1974}{1.0} = \underline{0.1974}$$

Thus, $p(State_1 \mid [0.7895, 0.2105, 0.0], \text{“NoInfo”}, Strike_1) = \underline{0.1974}$
Likewise, $p(State_2 \mid [0.7895, 0.2105, 0.0], \text{“NoInfo”}, Strike_1) = \underline{0.2105}$
Likewise, $p(State_{Dead} \mid [0.7895, 0.2105, 0.0], \text{“NoInfo”}, Strike_1) = \underline{0.5921}$

Thus, after the 2nd iteration, where the action was to fire artillery strike #1 into $Cell_1$ ($State_1$), called $Strike_1$,

the BV now becomes = [0.1974, 0.2105, 0.5921]

which is interpreted to mean a 0.1974 probability of the enemy being believed to be alive in $State_1$ ($Cell_1$), a 0.2105 probability of the enemy being believed to be alive in $State_2$ ($Cell_2$), and a 0.5921 probability of the enemy being believed to be Dead or in $State_{Dead}$. Performing the check sum verification:

CS = 0.1974 + 0.2105 + 0.5921 = 1.0000, therefore Checksum verification passed.

2.2. Determination of the Action Sequence.

The determination of which action to take at each iteration of the model is made in order to create the statistical optimal end state effect. The optimal end state is defined as the end state reached by the fewest action sequences with the most optimal reward value possible. Reward structure will be discussed later in this paper. Thus, the selection of the action sequence is achieved through a deterministic evaluation of the previous Bayesian state space and BV utilizing conditional probability logic. To begin this discussion the following definitions are made:

Δ ≡ (Delta) Declare Threshold; if the Belief Probability of Enemy Destroyed ≥ Δ, then Declare. Δ only refers to the DEAD state.

σ ≡ (Sigma) Shoot Threshold; if the Belief Probability of Enemy Destroyed ≥ σ, then Shoot, otherwise perform Recon. σ only refers to the LOCATION states.

Contrast ≡ The probability of the enemy being in one state relative to all other (location) states (Dead state is therefore excluded from the Contrast determination). This is a calculated value referred to as a Conditional Probability (CP).

Thus, if the belief probability of the enemy being Destroyed ≥ Δ, then the action decision will be to Declare mission complete and end the mission. However, if $\Delta \leq 0.9$ then go through Recon vs. Shoot logic according to the σ threshold. Note that the Δ threshold only applies to $State_3$, the Dead State, and the σ threshold only applies to the location states, $State_1$ and $State_2$. Therefore, the Δ and σ thresholds are independent and not directly related. For the purposes of discussion the following assignments are made. These are for reference only as were previous assignments in Tables 1-3 and are left as input parameters to the model to set its level of tolerance and to reflect the scenario conditions in effect at the time of the model invocation.

Assume the following assignments:

Δ = 0.9.
σ = 0.75.

2.2.1. Conditional Probability Logic.

For the case under investigation here the action decision logic can be viewed as a Conditional Probability (CP). The probability of the enemy being in one of the 2 location states ($State_1$ or $State_2$), given that the enemy is not destroyed, i.e., not present in $State_3$, can be represented as:

$P(S_1 \mid \text{Enemy Not Destroyed}) = P(S_1 \mid !S_3)$, where the ! symbol represents Boolean 'NOT'

Employing the form of Bayes' Theorem above (Equation 3), this becomes:

$$P(S_1 \mid !S_3) = \frac{P(!S_3 \mid S_1)\, P(S_1)}{P(!S_3)}$$

Equation 3 – Conditional Probability Initial Form

Since the probability of the enemy not being dead if they are in $State_1$ is equal to 1.0, meaning that he is alive, is represented as:

$$P(!S_3 \mid S_1) = 1.0,$$

And the probability of the enemy not being dead, $P(!S_3)$ is equal to the sum of the probabilities of being in one of the location states, or in this case $[P(S_1) + P(S_2)]$, Equation 3 now becomes:

$$P(S_1 \mid !S_3) = \frac{1.0 \times P(S_1)}{P(S_1) + P(S_2)} = \frac{P(S_1)}{P(S_1) + P(S_2)}$$

Equation 4 – Conditional Probability Expression

2.2.2. Calculation of Example Action Sequence Using Conditional Probability.

The action decisions made by the model during each iteration of the logic are made by evaluating the CP using the σ and Δ thresholds. After each action a new BV is calculated to be used in the next action decision. For the sample σ and Δ threshold values of 0.75 and 0.90 the action sequence for the first 5 actions is computed for verification of the computer simulation runs that will be made using this model. These parameter choices provide for an action sequence for the first 5 actions of:

1) $Recon_1$ to S_1.
2) $Strike_1$ to S_1.
3) $Recon_2$ to S_2.
4) $Strike_2$ to S_1.
5) $Recon_3$ to S_2.

See the Appendix for tables that illustrate the action calculations that determine this sequence.

2.3. Implementation in C3TRACE.

To implement this model in a computer simulation, the programming environment of Command, Control, and Communications: Techniques for the Reliable Assessment of Concept Execution (C3TRACE) (Kilduff, Swoboda, and Barnette, 2005; Plott, 2002; Plott, Quesada, Kilduff, Swoboda, and Allender, 2004) is employed. C3TRACE, developed through funding by the Human Research and Engineering Directorate of the U.S. Army Research Laboratory, is an adaptation of the commercial discrete event programming language Micro Saint Sharp™ (Schunk and Plott, 2004). While the basic Micro Saint Sharp™ programming language allows task based computer simulations of real world systems and processes to be represented,

C3TRACE has embedded data structures that augment Micro Saint Sharp™ to allow for detailed representation of U.S. Army command and control systems.

The optimal decision making model described in this paper allows existing computer simulations of command and control systems configured around task performance analysis (Middlebrooks, 2001; Middlebrooks, 2003; Middlebrooks, 2004; Middlebrooks et al., 1999a; Middlebrooks et al., 1999b; Middlebrooks and Williges, 2002; Wojciechowski et al., 2005), to now be structured to incorporate optimal decision making as a performance metric using the belief updating model. The steps in this process resemble the well known Observe-Orient-Decide-Act (OODA) model (Belknap, 1996; Boyd, 1982; Salas, Morgan, Glickman, Woodard, and Blaiwes, 1986). The decision actions in this model consist of gathering information, updating the belief about the environment or state space, taking an action to accomplish an objective in the state space, and then making a decision of whether to continue the mission or terminate it with an assessment of either mission success or failure. An example in a military command and control scenario employs a unmanned aerial vehicle (UAV) to gather the intelligence, artillery to take an action to destroy an enemy somewhere within the state space, and belief updating to evaluate the situation after each action and then either repeat the sequence or declare 'mission complete'.

2.3.1. Design of the C3TRACE Simulation.

C3TRACE programs are implemented using discrete event language constructs common to any Micro Saint Sharp™ simulation program. The top level of a command and control sub-workgroup within a sample organization is shown in the example depicted in Figure 4. Here messages received by the radio operator are distributed according to their subject content. Situation Reports (SITREP) are passed to the S3 Operations officer, Logistics Reports are passed to the S4 Logistics officer for action and so on. If, for example, a mission directive such as seek out and destroy an enemy, is received it is passed to the commander for action. There are different reactions that might be experienced to such a directive. The commander might communicate back to the originating authority to clarify information, an initial estimate of the situation before taking action might be performed, an updating of the situational awareness before taking action might be performed, or, the mission might be undertaken as directed. In this case, as depicted in the green box in Figure 4, what is referred to as the Decision Making under UnCertainty (DMUC) process would be initiated to execute the mission.

Figure 5 illustrates the optimal decision process that is modeled. As stated, this process is very similar to the OODA model. This diagram represents an iterative process where the decision maker makes an initial estimate of the situation and then begins an iterative process of either gathering additional information (flying a UAV mission) or taking an action to destroy the enemy (firing artillery). When the commander believes that the enemy has been destroyed, a mission complete decision is made and the results of the decision are realized. If the enemy was destroyed and the decision maker made that correct assessment, then a positive reward resulting from a good decision is applied to the performance of the overall system. If the enemy was not destroyed and the decision maker believed that it was destroyed, then a negative battlefield outcome is applied to the simulation. Likewise, if the enemy was destroyed but the decision maker believed it was not, then the results of poor decision making are applied. This process of iterative action can be generalized to similar scenarios where information is gathered (Observe), belief updating occurs (Orient), decisions are made for mission success (Decide), and actions are taken to accomplish the mission (Act). The examples of employing a UAV and firing artillery are used here to simply provide a tangible example of how this type of activity might occur.

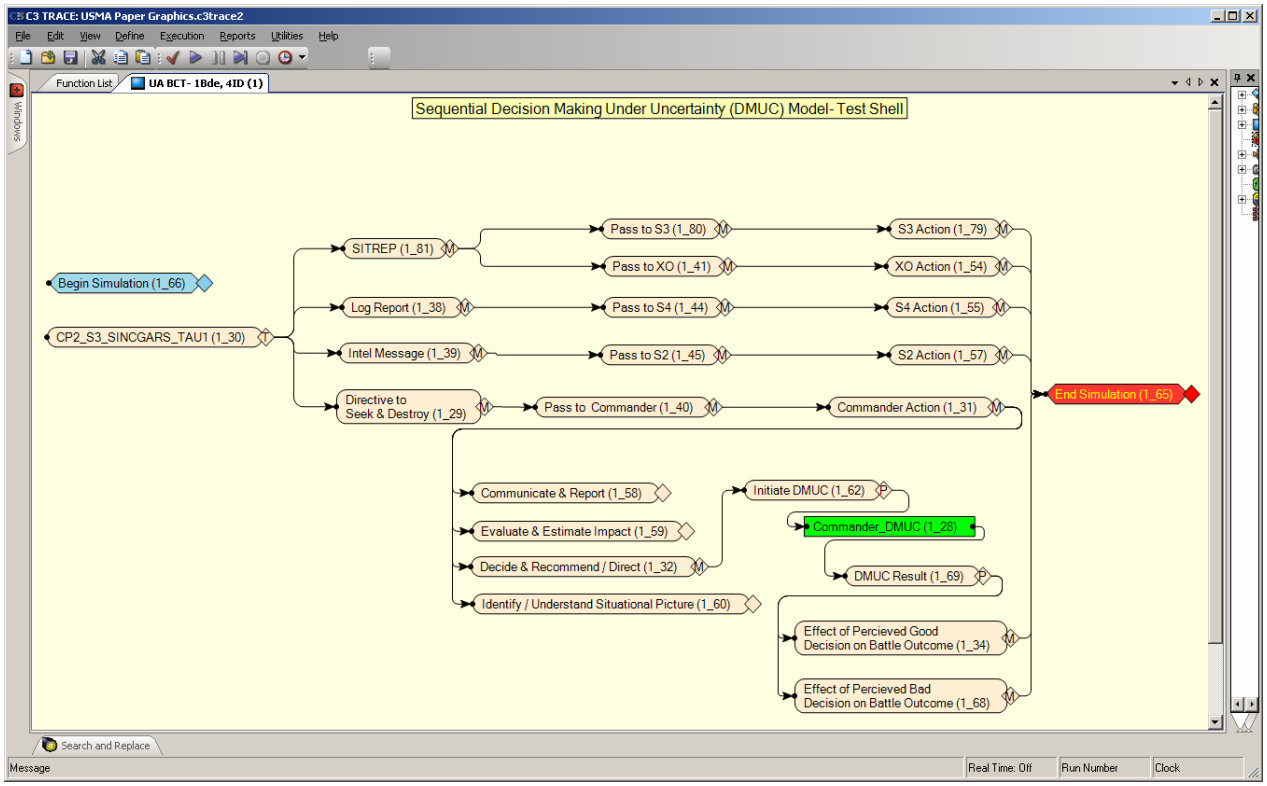


Figure 3 – C3TRACE Command and Control Simulation Vignette

Referring to Figure 5, the top level logic for this model can be examined. After initiating the decision sequence and performing an initial estimate of the situation, the commander updates the BV, defined as the belief about the current situation regarding the enemy, and then begins an iterative process of looking for information or taking an action to accomplish the mission. When this process has reached some level of belief that the mission is accomplished, the commander terminates the action and completes the decision process by declaring that the mission is either a 'success' or a 'failure'.

If the initial desire is to obtain additional information, a UAV is sent to a specified location to attempt to locate the enemy. The UAV is the information gathering or battle damage assessment (BDA) tool available to the commander to update the BV about the enemy. If the target is already dead from previous artillery action, then there is no correct location for the enemy because it does not exist as it is dead. If the enemy is alive and the UAV is sent to the correct location then it has a probability, according to Table 1, of either detecting or not detecting the enemy according to the accuracy of the UAV. From this it will either correctly or incorrectly report that the enemy was found. Likewise, if it is sent to a location where the enemy is not located, or if the enemy is already dead, it may correctly or incorrectly report the enemy sighted again according to Table 1. The values in Table 1 are only sample estimates for use in the development of this model and do not represent any actual system currently in existence. During actual use of this model these parameters are set to represent the actual detection characteristics of the information gathering entity being evaluated. After the UAV mission is flown the commander evaluates the report from the UAV through the process of updating the BV and, using this new information, decides what process to invoke next.

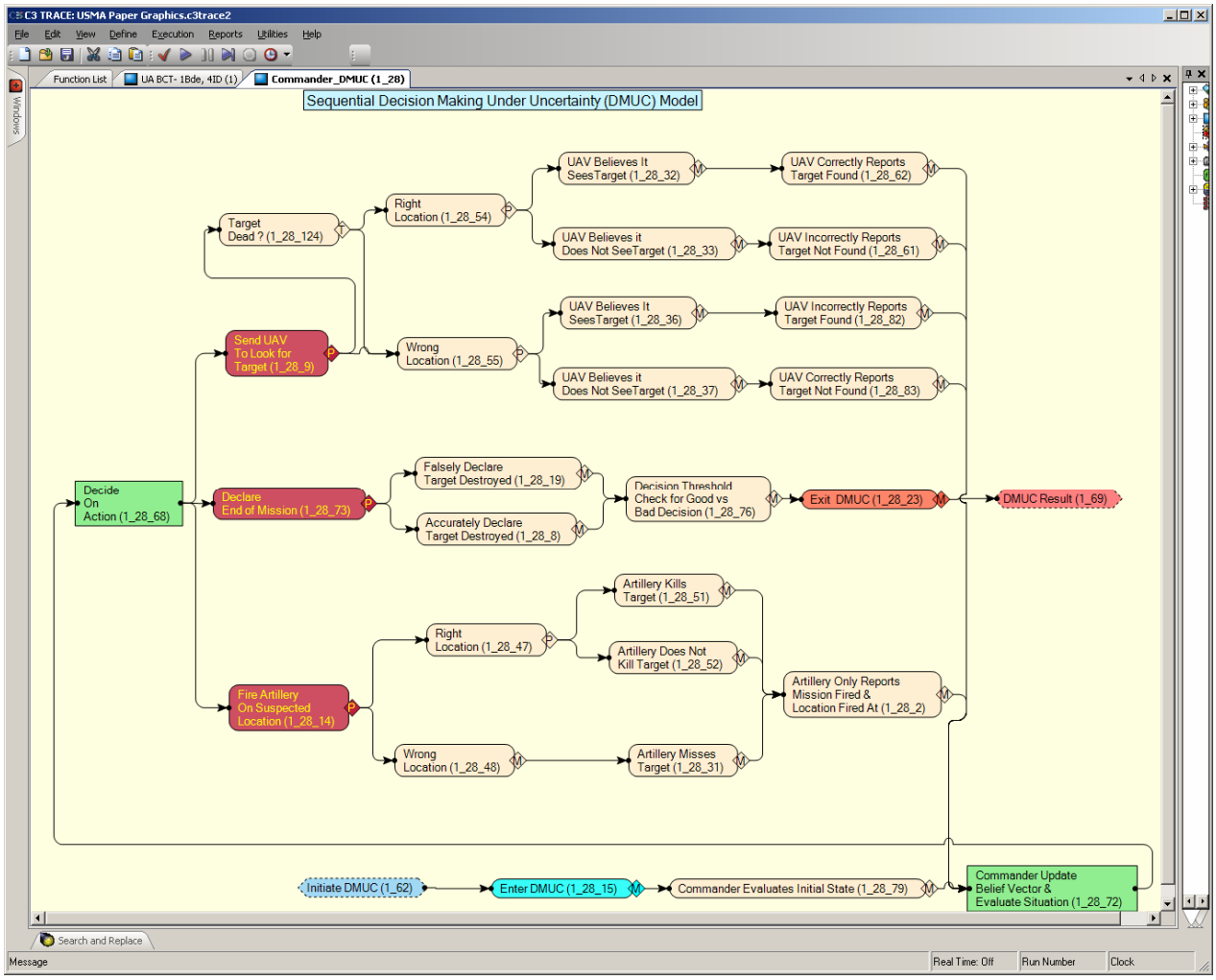


Figure 4 – Optimal Decision Making Under Uncertainty Model

If the commander decides to fire artillery (which is representative of taking a positive action to do something to accomplish the mission) then the probability exists that either the right or wrong location will be fired upon. If the artillery fires on the wrong location then the only outcome will be to miss the target. If the correct location is fired upon then the artillery will either kill or not kill the enemy according to the circular area of probability for the type of artillery fired. Independent of where the artillery is fired, the only report that is sent back to the commander is that the artillery fired upon the location specified or 'No Information' concerning BDA. This represents the fact that artillery is an indirect fire weapon and the firing unit never actually sees the target. It is up to the forward observer, or in this case the UAV, to report back as to the actual target situation, i.e., to provide the BDA. It is then up to the commander to evaluate the firing data and information from previous UAV reconnaissance missions to decide if to continue the mission or declare the enemy is dead and end the mission.

When the commander believes that the enemy has been destroyed, then mission complete is declared. Then the commander is faced with either the rewards of a successful, i.e., good decision sequence where the enemy was killed meaning that the mission accomplished, or the effects of a bad decision where the enemy was not killed meaning that the mission was not accomplished.

2.3.2. Belief Updating Logic.

Figure 6 illustrates the inputs feeding the sequence of evaluating the current situation, and updating of the belief vector and the resulting choice for the next action.

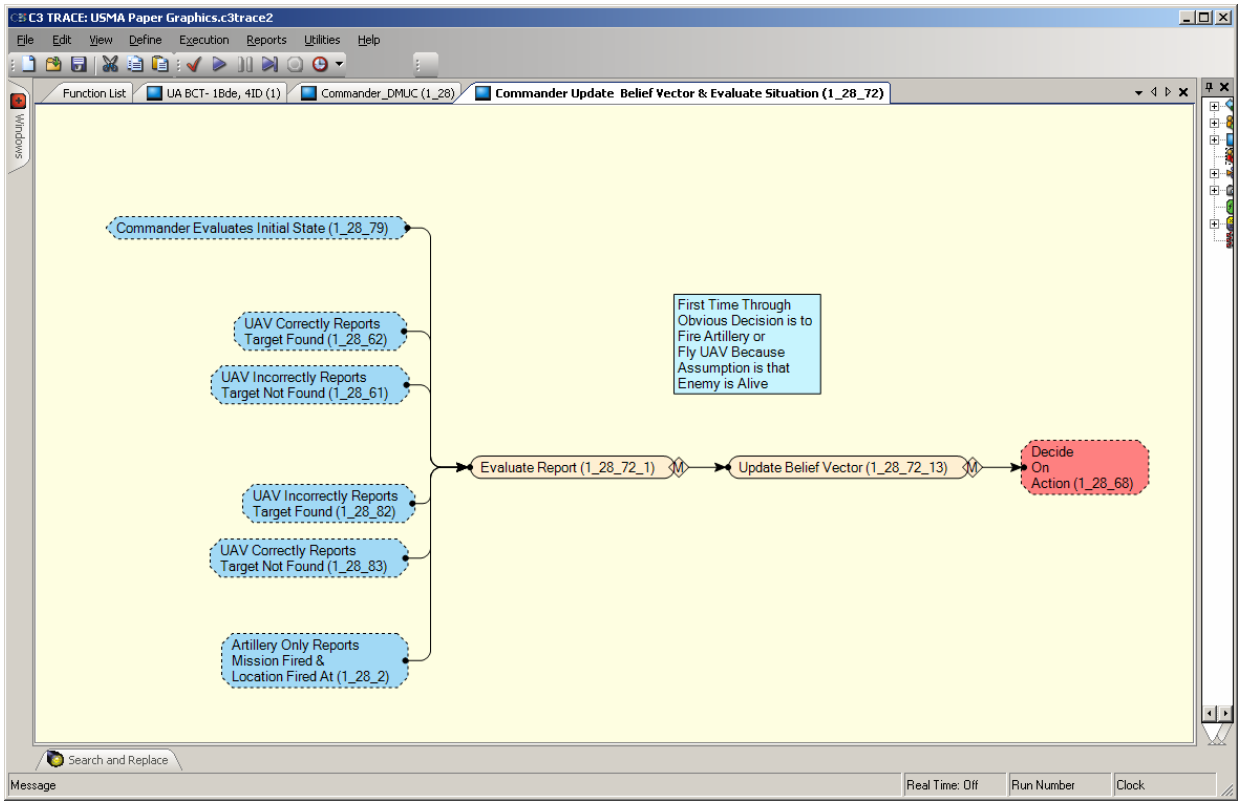


Figure 5 – Task Diagram for Belief Updating

The code script in the beginning effect of C3TRACE task 'Update Belief Vector' of Figure 6 closely follows the BV updating logic described above. An annotated description of this logic as implemented in the C3TRACE computer simulation takes the form:

Definitions:

BV - Belief Vector

Variables in Equation 2:

- s' - True state within the total state space.
- b - Prior belief for that state.
- o - Current observation.
- a - Action that was generated.
- In-State - Probability of a state being transferred in to. This is equal to 0 for location states because for a static enemy condition a location state can never be transferred in to.
- Out-State - Probability of a state being transferred out from. This is equal to 0 for the dead state because once the enemy is dead it can not be transferred back to alive.

Components of Equation 2:

- PoGs'ba - Probability of o Given s' & b & a.
- Ps'Gba - Probability of s' Given b & a.
- PoGba - Probability of o Given b & a.
- Ps'Gboa - Probability of s' Given b & o & a = (PoGs'ba * Ps'Gba) / PoGba => Eq. 2.

If the action generated was to recon, then update the BV from the recon mission:

Do for each state:

Calculate PoGs'ba from Table 1 lookup.

Calculate Ps'Gba from previous belief probability of the enemy in the state occupied.

Calculate PoGba from:

(Probability of Enemy in State Action To * Table 1 Lookup for that State) +

(Probability of Enemy in State Action Not To * Table 1 Lookup for that State) +
(Probability of Enemy in State Dead * Table 1 Lookup for condition applicable to State Dead)
Calculate Ps'Gboa from: (PoGs'ba * Ps'Gba) / PoGba
Equate the BV component for that state = Ps'Gboa.
End Do

Else if the action generated was to Shoot, then update BV from the shoot mission:
Do for each state:
Calculate PoGs'ba from Table 2 lookup which will always = 1.0.
Calculate Ps'Gba from:
[Probability (In-State) x (previous probability for State Struck] +
[(1.0- Probability (Out-State) x (previous probability for State Occupied)]
Calculate PoGba from Table 2 lookup which will always = 1.0.
Calculate Ps'Gboa from: (PoGs'ba * Ps'Gba) / PoGba
Equate the BV component for that state = Ps'Gboa.
End Do

End

2.3.3. Action Decision Making Logic.

The code script in the beginning effect of C3TRACE task 'Evaluate Report' of Figure 6 closely follows the CP logic described previously. An annotated description of this logic as implemented in the C3TRACE computer simulation takes the form:

Definitions:

CP - Conditional Probability

Sigma - σ, Shoot Threshold; if the Belief Probability of Enemy Destroyed $\geq \sigma$, then Shoot, otherwise recon. σ only refers to the LOCATION states.

Delta - Δ, Declare Threshold; if the Belief Probability of Enemy Destroyed $\geq \Delta$, then Declare. Δ only refers to the DEAD state.

Variables in Equation 6:

PS_1 - Previous BV component for S_1.

PS_2 - Previous BV component for S_2.

Ps_1GNs_3 - Probability of S_1 given NOT S_3 = probability of the enemy not being dead

- $PS_1 / (PS_1 + PS_2)$ => Eq. 6.

Do
Calculate $CP_1 = PS_1 / (PS_1 + PS_2)$
Calculate $CP_2 = PS_2 / (PS_1 + PS_2)$
End Do
If $((CP_1 \& CP_2) < \sigma)$ then
If $(CP_1 > CP_2)$ then Recon at S_1
If $(CP_1 < CP_2)$ then Recon at S_2
If $(CP_1 = CP_2)$ then Recon at Random pick of S_1 or S_2
End If
If $(CP_1 > \sigma)$ then SHOOT at S_1
If $(CP_2 > \sigma)$ then SHOOT at S_2
If $((CP_1 > \sigma) \& (CP_2 > \sigma))$ then SHOOT at Random pick of S_1 or S_2
Calculate new BV
If ((BV component for S_3) $> \Delta$) then
Declare mission complete
Else continue processing and go to next iteration

3. Discussion / Results.

The simplest version of the POMDP model state space design as shown in Figure 3 is used in this paper to evaluate and demonstrate the logic through the computer simulation in C3TRACE. While the state space that consists of two location states, $State_1$ and $State_2$, and one status state,

$State_{Dead}$, can seem trivial and unrelated to any actual human performance condition, even this simple arrangement can relate to actual performance. The seek and destroy mission looking to destroy an enemy residing at some unknown location can be characterized as looking or shooting at the enemy at the right or wrong location before declaring that the enemy has been destroyed. Thus, a simple form of the state space such as this can form the basis for developing logic that can be expanded after verification to much larger location state spaces.

A means for evaluating the performance of the simulation is to implement a reward structure (RS) consisting of an explicit cost for taking different actions. There would be a certain cost for conducting a reconnaissance, another greater cost for conducting an artillery strike. There would also be a reward if the mission complete declaration is made when the enemy has actually been destroyed and a corresponding large cost assessed when mission complete is called when the enemy has not been destroyed.

3.1. Control Parameters.

The Δ and σ control parameters for the BV and CP calculations allow the model to respond to settings for aggressiveness by the operator in making decisions to recon or to shoot and to reflect the operator's confidence on when a successful mission has occurred. The σ control parameter is used to set the recon versus shoot threshold criteria for the performance of the simulation. Values of σ during analytical runs of the simulations can be varied to represent the complexity and decision threshold conditions of the scenario being simulated. The Δ control parameter is used to set the decision threshold criteria for the performance of the simulation. Values of Δ during analytical runs of the simulations can be varied to represent the complexity and decision threshold conditions of the scenario being simulated.

3.2. Action Sequence Assessment.

In order to examine the CP logic associated with actions in this state space, an example of actions and the resulting belief vectors will be examined. The assumptions are that the enemy is located in $State_1$ and that it is static, i.e., not moving. Initially, there is an equal probability in the belief of the commander that the enemy could be in either of the location states and a belief that the enemy is alive. The initial belief vector is thus [0.5, 0.5, 0.0], meaning a 50% chance of being in location $State_1$, a 50% chance of being in location $State_2$, and a 0.0% chance of being in $State_{Dead}$, i.e., the enemy is alive.

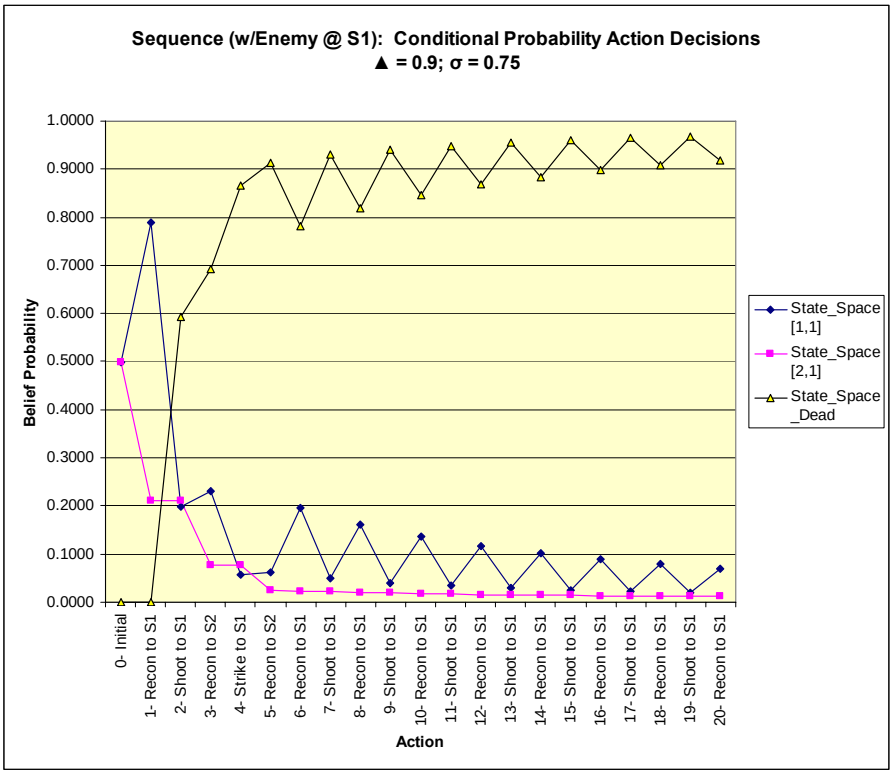
Assume that the control parameter values are initially set to:

$\sigma = 0.75$, i.e., if the Belief Probability of Enemy Destroyed (in regard to a location state) $\geq \sigma$, then Shoot, otherwise Recon.

$\Delta = 0.90$, i.e., if the Belief Probability of Enemy Destroyed (in regard to the dead state) $\geq \Delta$, then Declare mission complete.

These assumed values are for example only and are not to be construed to represent any actual system.

Applying these parameters to the BV and CP logic generates the sequence of actions as shown in Figure 7 and Table 9. Even though the BV component for $State_{Dead}$ exceeds Δ at iteration 5, the model run was continued for 20 iterations to illustrate the action sequence asymptotic relationships.



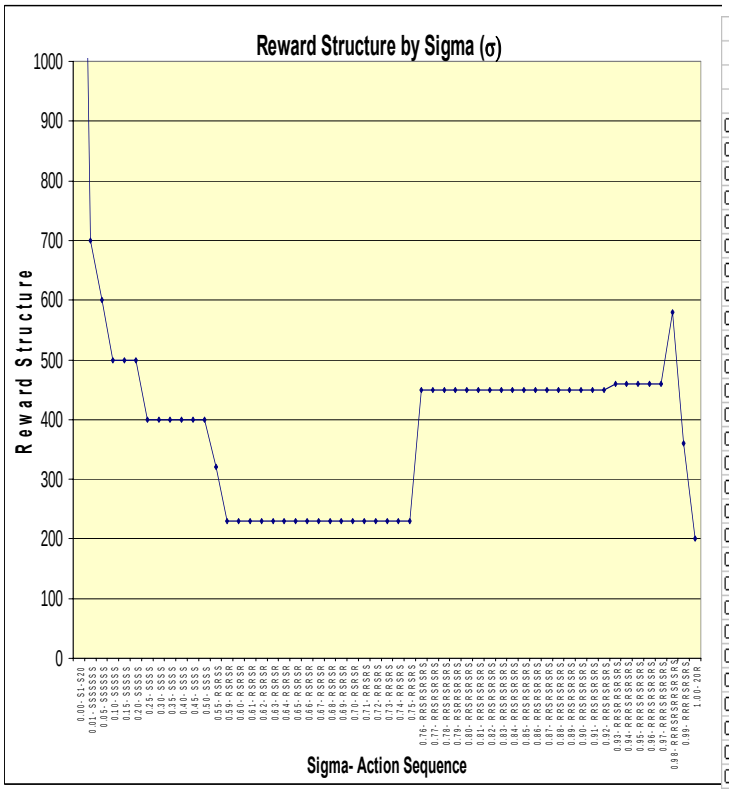
Action	State_Space [1,1]	State_Space [2,1]	State_Space _Dead
0- Initial	0.5000	0.5000	0.0000
1- Recon to S1	0.7895	0.2105	0.0000
2- Shoot to S1	0.1974	0.2105	0.5921
3- Recon to S2	0.2308	0.0769	0.6923
4- Shoot to S1	0.0577	0.0769	0.8654
5- Recon to S2	0.0609	0.0254	0.9137
6- Recon to S1	0.1957	0.0217	0.7826
7- Shoot to S1	0.0489	0.0217	0.9293
8- Recon to S1	0.1617	0.0192	0.8192
9- Shoot to S1	0.0404	0.0192	0.9404
10- Recon to S1	0.1364	0.0172	0.8463
11- Shoot to S1	0.0341	0.0172	0.9487
12- Recon to S1	0.1169	0.0158	0.8673
13- Shoot to S1	0.0292	0.0158	0.9550
14- Recon to S1	0.1015	0.0146	0.8840
15- Shoot to S1	0.0254	0.0146	0.9600
16- Recon to S1	0.0889	0.0136	0.8974
17- Shoot to S1	0.0222	0.0136	0.9641
18- Recon to S1	0.0786	0.0129	0.9086
19- Shoot to S1	0.0196	0.0129	0.9675
20- Recon to S1	0.0699	0.0122	0.9179

Figure 6 – 1st 20 Action Decisions for $\Delta = 0.90$ & $\sigma = 0.75$

Activating the Δ control parameter causes the simulation to declare mission complete after iteration 5 shown in Table 9 with a $State_{Dead}$ BV component = 0.9137 which is just over the Δ threshold of 0.90. This results in a five action sequence of Recon-Shoot-Recon-Shoot-Recon to Declare. If a reward structure (RS) is implemented with a cost of 10 combat power points to Recon and 100 combat power points to shoot then the cost of this action sequence would be (3x10) + (2x100) = 230.

3.3. Reward Structure.

Evaluation of the RS as σ is varied from 0.00 to 1.00 provides an indication of the 'cost of doing business' based on how aggressive the decision maker is in making action choices. Figure 10 shows a profile of the RS over this range with σ being incremented in increments of 0.10. Table 11 showing an expanded view of the information in the X axis.



Recon Cost	10		
Shoot Cost	100		
σ, Sigma- Sequence	RS	σ, Sigma- Sequence	RS
0.00- S1-S20	2000	0.74- RRSRS	230
0.01- SSSSSSS	700	0.75- RRSRS	230
0.05- SSSSSS	600	0.76- RRSRSRSRS	450
0.10- SSSSS	500	0.77- RRSRSRSRS	450
0.15- SSSSS	500	0.78- RRSRSRSRS	450
0.20- SSSSS	500	0.79- RSRRSRSRS	450
0.25- SSSS	400	0.80- RRSRSRSRS	450
0.30- SSSS	400	0.81- RRSRSRSRS	450
0.35- SSSS	400	0.82- RRSRSRSRS	450
0.40- SSSS	400	0.83- RRSRSRSRS	450
0.45- SSSS	400	0.84- RRSRSRSRS	450
0.50- SSSS	400	0.85- RRSRSRSRS	450
0.55- RSRSS	320	0.86- RRSRSRSRS	450
0.59- RSRSR	230	0.87- RRSRSRSRS	450
0.60- RSRSR	230	0.88- RRSRSRSRS	450
0.61- RSRSR	230	0.89- RRSRSRSRS	450
0.62- RSRSR	230	0.90- RRSRSRSRS	450
0.63- RSRSR	230	0.91- RRSRSRSRS	450
0.64- RSRSR	230	0.92- RRSRSRSRS	450
0.65- RSRSR	230	0.93- RRSRRSRSRS	460
0.66- RSRSR	230	0.94- RRRSRSRSRS	460
0.67- RSRSR	230	0.95- RRRSRSRSRS	460
0.68- RSRSR	230	0.96- RRRSRSRSRS	460
0.69- RSRSR	230	0.97- RRRSRSRSRS	460
0.70- RSRSR	230	0.98- RRRSRSRRSRSRS	580
0.71- RRSRS	230	0.99- RRRRSRSRS	360
0.72- RRSRS	230	1.00- 20R	200
0.73- RRSRS	230		

Figure 7 – Action Cost from Reward Structure Varied by σ

The intent is to minimize the cost of doing business by performing the least costly sequence of actions to achieve the desired belief that the enemy has been destroyed. In an analytical use of this model tailoring the reconnaissance and strike asset capabilities so that they support an action sequence of Recon-Shoot-Recon-Shoot-Recon to achieve the belief threshold specified would allow the system to be tailored for optimal performance along this parameter. Here, the optimal performance occurs over the range of $\sigma \cong 0.59$ to 0.75 with a resulting action cost of 230.

4. Conclusions / Future Work.

The current work has established a model that supports a computer simulation capable of determining and optimizing optimal decisions under conditions of uncertainty according to evaluations of the Belief Vector about the current state, action decisions based on Conditional Probability logic, and optimal performance determination through the evaluation of action cost from the Reward Structure. Thus, the C3TRACE simulation employing this model has the ability to make action decisions based on conditional probability evaluations of the belief state representing the current situation which are oriented toward a goal directed optimal outcome and then recognize when the belief has been achieved that the outcome has been reached.

This paper demonstrates the logic of this model through the evaluation of the most simple of state spaces which consists of a 2x1 location state matrix and a single status state of the dead condition for a total 3 state space system. Future work will expand the location state space matrix to 2x2 and 5x5 as shown in Figures 1 and 2 along with more sophisticated enemy actions for moving versus static operations and goal directed movement activities.

6. References.

Army. (2002). ABCS-LRG 6.2. Army Battle Command System Leader's Reference Guide. U.S. Army Special Text. Fort Leavenworth, KS.

Army. (2003). FM 006-000 Mission Command: Command and Control of Army Forces. U.S. Army Field Manual. Washington, D.C. (Distribution: Approved for Public Release, Distribution is Unlimited).

Belknap, M. (1996). Military Decision Making in the Information Age. Unpublished Monograph, Naval War College, Newport, RI. Joint Military Operations Department. Ad-a307-648-6-xab.

Boyd, J. R. (1982). A Discourse on Winning and Losing (Seminal Work Introducing the OODA Loop). Unpublished briefing and essays, Air University Library, Maxwell Air Force Base, U.S. Air Force, Document no. MU 43947, August 1987. http://www.d-n-i.net/second_level/boyd_military.htm#discourse.

Cassandra, A. R. (1998). Exact and Approximate Algorithms for Partially Observable Markov Decision Processes. Unpublished Dissertation, Brown University.

Cassandra, A. R., Kaelbling, L. P., & Kurien, J. A. (1996). Acting under Uncertainty: Discrete Bayesian Models for Mobile-Robot Navigation. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems.

Cassandra, A. R., Kaelbling, L. P., & Littman, M. L. (1994). Acting Optimally in Partially Observable Stochastic Domains. In Proceedings of the Twelfth National Conference on Artificial Intelligence. (AAAI), 1023--1028. http://citeseer.ist.psu.edu/cassandra94acting.html Seattle, Washington. MIT Press.

Geisler, W. S. (1989). Ideal-Observer Theory in Psychophysics and Physiology. Physica Scripta, 39, 153-160.

Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, Quanta, and Vision. Journal of General Physiology, 25, 819-840.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and Acting in Partially Observable Stochastic Domains. Artificial Intelligence, 101, 99-134.

Kilduff, P. W., Swoboda, J. C., & Barnette, B. D. (2005). Command, Control, and Communications: Techniques for the Reliable Assessment of Concept Execution (C3TRACE) Modeling Environment: The Tool. U.S. Army Research Laboratory, Human Research and Engineering Directorate. Aberdeen Proving Ground, MD. ARL Technical Report ARL-MR-0617.

Legge, G. E., & Hooven, T. A. (2002). Mr. Chips 2002: New Insights from an Ideal-Observer Model of Reading. Vision Research, 42, 2219-2234.

Legge, G. E., & Klitz, T. S. (1997). Mr. Chips: An Ideal-Observer Model of Reading. Psychological Review, 104, 524-553.

Liu, Z., & Knill, D. C. (1995). Object Classification for Human and Ideal Observers. Vision Research, 35(4), 549-568.

Middlebrooks, S. E. (2001). Experimental Interrogation of Network Simulation Models of Human Task and Workload Performance in a U.S. Army Tactical Operations Center. Unpublished Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Middlebrooks, S. E. (2003). The Compass Paradigm for the Systematic Evaluation of U.S. Army Command and Control Systems Using Neural Network and Discrete Event Computer Simulation. Unpublished Dissertation, Department of Industrial and Systems Engineering. Virginia Polytechnic Institute and State University, Blacksburg, VA. U.S. Army Research Laboratory Reprint Report ARL-RP-110 (November 2005).

Middlebrooks, S. E. (2004). Neural Net and Discrete Event Simulations of C2 Performance in Army Tocs. In Proceedings of the Human Factors and Ergonomics Society's 48th Annual Meeting, September 20-24, 2004 (pp. 513-517). New Orleans, LA: HFES.

Middlebrooks, S. E., Knapp, B. G., Barnette, B. D., Bird, C. A., Johnson, J. M., Kilduff, P. W., Schipani, S. P., Swoboda, J. C., Wojciechowski, J. Q., Tillman, B. W., Ensing, A. R., Archer, S. G., Archer, R. D., & Plott, B. M. (1999a). Cohost - Computer Simulation Models to Investigate Human Performance Task and Workload Conditions in a U.S. Army Heavy Maneuver Battalion Tactical Operations Center. In Proceedings of the HFES 43rd Annual Meeting (pp. 242-246). Houston, TX: Human Factors and Ergonomics Society.

Middlebrooks, S. E., Knapp, B. G., Barnette, B. D., Bird, C. A., Johnson, J. M., Kilduff, P. W., Schipani, S. P., Swoboda, J. C., Wojciechowski, J. Q., Tillman, B. W., Ensing, A. R., Archer, S. G., Archer, R. D., & Plott, B. M. (1999b). Cohost (Computer Modeling of Human Operator System Tasks) Computer Simulation Models to Investigate Human Performance Task and Workload Conditions in a U.S. Army Heavy Maneuver Battalion Tactical Operations Center. Human Research and Engineering Directorate, U.S. Army Research Laboratory, Aberdeen Proving Ground, MD 21005. ARL Technical Report ARL-TR-1994; DTIC ADA368587.

Middlebrooks, S. E., & Stankiewicz, B. J. (2006). Toward the Development of a Predictive Computer Model of Decision Making under Uncertainty for Use in Simulations of U.S. Army Command and Control Systems. U.S. Army Research Laboratory Technical Report ARL-TR-3719. DTIC ADA443462.

Middlebrooks, S. E., & Williges, R. C. (2002). Experimental Design Interrogation of Network Simulation Models of U.S. Army Command and Control Centers. In Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society (pp. 458-462). Baltimore, MD: HFES.

Najemnik, J., & Geisler, W. S. (2005). Optimal Eye Movement Strategies in Visual Search. Nature, 434, 387-391.

Plott, B. (2002). Software User's Manual For: C3TRACE (Command, Control and Communication - Techniques for Reliable Assessment of Concept Execution). U.S. Army Research Laboratory, Human Research and Engineering Directorate, Aberdeen Proving Ground, MD.

Plott, B., Quesada, S., Kilduff, P., Swoboda, J., & Allender, L. (2004). Using an Information-Driven Decision Making Human Performance Tool to Assess U.S. Army Command, Control, and Communication Issues. In Proceedings of the Human Factors and Ergonomics Society's 48th Annual Meeting, September 20-24, 2004 (pp. 2396-2400). New Orleans, LA: HFES.

Salas, E., Morgan, B. B., Jr., Glickman, A. S., Woodard, E. A., & Blaiwes, A. S. (1986). Measurement of Team Behaviors in a Navy Training Environment. Human Factors Division, Department of the Navy. Technical Report NTSC TR-86-014. DTIC ADA185237.

Schunk, D., & Plott, B. (2004). Use of Discrete Event Simulation to Model Human Performance. In Proceedings of the Human Factors and Ergonomics Society's 48th Annual Meeting, September 20-24, 2004 (pp. 2126-2130). New Orleans, LA: HFES.

Sondik, E. J. (1971). The Optimal Control of Partially Observable Markov Processes. Unpublished Dissertation, Stanford University, Stanford, California.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human Efficiency for Recognizing 3-D Objects in Luminance Noise. Vision Research, 35(21), 3053-3069.

Tjan, B. S., & Legge, G. E. (1998). The Viewpoint Complexity of an Object Recognition Task. Vision Research, 15/16, 2335-2350.

Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2004). Optimal Compensation for Changes in Effective Movement Variability. Nature Neuroscience Under Review, under review.

Trueman, R. E. (1977). An Introduction to Quantitative Methods for Decision Making. (2 ed.). New York, NY: Holt, Rinehart and Winston, Inc.

Walpole, R. E., Myers, R. H., & Myers, S. L. (1998). Probability and Statistics for Engineers and Scientists. Upper Saddle River, New Jersey 07458: Prentice Hall.

Wine, R. L. (1964). Statistics for Scientists and Engineers. Englewood, Cliffs, NJ: Prentice Hall.

Wojciechowski, J. Q., Plott, B., & Kilduff, P. (2005). Human Performance Model Development of a Battalion Tactical Operations Center. U.S. Army Research Laboratory, Human Research and Engineering Directorate. Aberdeen Proving Ground, MD. ARL Technical Report ARL-TR-2635.

7. **Appendix –** Conditional Probability Calculations for First 5 Action Sequences**.**

7.1. Action #1: Make Choice to Perform $Recon_1$ to S_1.

The first action decision to be made by the model is whether to recon or shoot at S_1 or S_2. This is based on the initial conditions of the state space where the BV has been previously defined as [0.5, 0.5, 0.0]. It is also noted that the enemy is occupying $State_1$. Using Equation 6 the numerical evaluation of the CP, based on the limitations of the σ and Δ thresholds, is performed resulting in a first action to perform a reconnaissance mission to $State_1$. The results of the action is to generate a new BV = [0.7895, 0.2105, 0.0000].

Table 8 – Conditional Probability Calculations for 1st Action: $Recon_1$ to S_1

Initial Belief Vector	CP	Exceed Δ Threshold ?	Exceed σ Threshold ?	Choice	Observation
[0.5,0.5,0.0]	$\left[\frac{S_1}{S_1+S_2}, \frac{S_2}{S_1+S_2}\right]$ $\left[\frac{0.5}{0.5+0.5}, \frac{0.5}{0.5+0.5}\right]$ = [0.5 , 0.5] ∴For Action where Dead = 0.0, the denominator must sum to 1.0, and the total contrast must sum to 1.0.	NO Because DEAD state with a BV value = 0.0 is less than Δ which is equal to 0.9.	NO Because neither location cell has a value > 0.75 as each cell has a contrast ratio value = 0.5.	As both S_1 & S_2 meet recon criteria with value = 0.5, Select, RAN (i) Assuming Random Pick=1. ∴ Choice = RAND recon (S_1), or perform a recon to S_1, or perform RECON1.	ENEMY SIGHTED From this action of RECON1 to S_1 when enemy @ S_1. Generating a Belief Vector = [0.7895,0.2105,0.0].

7.2. Action #2: Make Choice to Perform $Strike_1$ to S_1.

Using the BV from Action #1 of [0.7895, 0.2105, 0.0000], the CP is now evaluated to select Action #2 to be to perform an artillery strike to S_1 generating a new BV = [0.1974, 0.2105, 0.5921].

Table 9 – Conditional Probability Calculations for 2nd Action: $Strike_1$ to S_1

Previous Belief Vector	CP	Exceed Δ Threshold ?	Exceed σ Threshold ?	Choice	Observation
[0.7895, 0.2105, 0.0]	$\left[\frac{S_1}{S_1+S_2}, \frac{S_2}{S_1+S_2}\right]$ $\left[\frac{0.7895}{0.7895+0.2105}, \frac{0.2105}{0.7895+0.2105}\right]$ = [0.7895 , 0.2105] ∴For Action where Dead Belief = 0.0, the denominator must sum to 1.0, and the total contrast must sum to 1.0.	NO Because DEAD state with a Belief Vector value = 0.0 is less than Δ which is equal to 0.9.	YES Because cell S_1 has a Contrast Ratio value > 0.75	Because S_1 and only S_1 meets SHOOT criteria, Select, Shoot S_1 because S_2 @ 0.2105 is < σ at 0.75. For future cases where there might be multiple location states exceeding σ, set up the general selection of **RAND Shoot (S_1).** In this case the choice is to SHOOT at S_1.	NO INFO (From this action of STRIKE1 to S_1.) Generating a Belief Vector = [0.1974, 0.2105, 0.5921]

7.3. Action #3: Make Choice to Perform $Recon_2$ to S_2.

Using the BV from Action #2 of [0.1974, 0.2105, 0.5921], the CP is now evaluated to determine Action #3 to be to perform a reconnaissance mission to S_2 generating a new BV = [0.2308, 0.0769, 0.6923].

Table 10 – Conditional Probability Calculations for 3rd Action: $Recon_2$ to S_2

Previous Belief Vector	CP	Exceed Δ Threshold ?	Exceed σ Threshold ?	Choice	Observation
[0.1974, 0.2105, 0.5921]	$\left[\frac{S_1}{S_1+S_2}, \frac{S_2}{S_1+S_2}\right]$ $\left[\frac{0.1974}{0.1974+0.2105}, \frac{0.2105}{0.1974+0.2105}\right]$ = [0.4839 , 0.5161] ∴For Action where Dead Belief ≠ 0.0, the denominator will not sum to 1.0, but the total contrast must still sum to 1.0., or 0.4839 + .5160 = 1.0	NO Because DEAD state with a Belief Vector value = 0.5921 is less than Δ which is equal to 0.9.	NO Because neither cell has a CR value > 0.75 with S_1= 0.4839 and with S_2= 0.5160	Because neither S_1 or S_2 meets σ criteria, select, RAN (i) of set of cells w/ largest contrast value, in this case only S_2 @ 0.5160 is in the set of cells containing the largest contrast value, thus select a recon into S_2, i.e., perform RECON2 into S_2.	ENEMY NOT SIGHTED From this action of RECON2 to S_2 when enemy @ S_1. Generating a Belief Vector = [0.2308, 0.0769, 0.6923]

7.4. Action #4: Make Choice to Perform $Strike_2$ to S_1.

Using the BV from Action #3 of [0.2308, 0.0769, 0.6923], the CP is now evaluated to determine Action #4 to be to perform an artillery strike to S_1 generating a new BV = [0.0577, 0.0769, 0.8654].

Table 11 – Conditional Probability Calculations for 4th Action: $Strike_2$ to S_1

Previous Belief Vector	CP	Exceed Δ Threshold ?	Exceed σ Threshold ?	Choice	Observation
[0.2308, 0.0769, 0.6923]	$\left[\frac{S_1}{S_1+S_2}, \frac{S_2}{S_1+S_2}\right]$ $\left[\frac{0.2308}{0.2308+0.0769}, \frac{0.0769}{0.2368+0.0769}\right]$ = [0.75008 , 0.2499] ∴For Action where Dead Belief ≠ 0.0, the denominator will not sum to 1.0, but the total contrast must still sum to 1.0., or 0.750081 + .249919 = 1.0	NO Because DEAD state with a Belief Vector value = 0.6923 is less than Δ which is equal to 0.9.	YES Because cell S_1 has a Contrast Ratio value = 0.750081 which is > σ at 0.75.	Because S_1 and only S_1 meets SHOOT criteria, Select, Shoot S_1 because S_2 @ 0.0769 is < σ at 0.75. For future cases where there might be multiple location states exceeding σ, set up the general selection of **RAND Shoot (S_1).** Thus select SHOOT at S_1, i.e., perform STRIKE2 at S_1.	NO INFO (From this action of STRIKE2 to S_1.) Generating a Belief Vector = [0.0577, 0.0769, 0.8654]

7.5. Action #5: Make Choice to Perform $Recon_3$ to S_2.

Using the BV from Action #4 of [0.0577, 0.0769, 0.8654]., the CP is now evaluated to determine Action #5 to be to perform a reconnaissance mission to S_2 generating a new BV = [0.0609, 0.0254 ,0.9137]. As the new BV component for $State_3$ at 0.9137 now exceeds the Δ threshold of 0.90, the model makes the decision to terminate with a declaration of 'Mission Complete' with the belief that the enemy has been destroyed.

Table 12 – Conditional Probability Calculations for 5th Action: $Recon_3$ to S_2

Previous Belief Vector	CP	Exceed Δ Threshold ?	Exceed σ Threshold ?	Choice	Observation
[0.0577, 0.0769, 0.8654]	$\left\lceil \frac{S_1}{S_1+S_2}, \frac{S_2}{S_1+S_2} \right\rceil$ $\left\lceil \frac{0.0577}{0.0577+0.0769}, \frac{0.0769}{0.0577+0.0769} \right\rceil$ = [0.4287 , 0.5713] ∴For Action where Dead Belief ≠ 0.0, the denominator will not sum to 1.0, but the total contrast must still sum to 1.0., or 0.4287 + .5713 = 1.0	NO Because DEAD state with a Belief Vector value = 0.8654 is less than Δ which is equal to 0.9.	NO Because neither cell has a CR value > 0.75 with S_1= 0.4287 and with S_2= 0.5713	Because neither S_1 or S_2 meets σ criteria, select, RAN (i) of set of cells w/ largest contrast value, in this case only S_2 @ 0.5713 is in the set of cells containing the largest contrast value, thus select a recon into S_2, i.e., perform RECON3 into S_2.	ENEMY NOT SIGHTED From this action of RECON3 to S_2 when enemy @ S_1. Generating a Belief Vector = [0.0609, 0.0254, 0.9137]

Note:

Δ threshold now exceeded with DEAD state Belief Vector value = 0.9137 which is greater than Δ at 0.9, therefore next action will be to DECLARE.