Human decision making is a fundamental and key process for our everyday lives. Humans make thousands of decisions everyday that have both immediate and long-term effects. Determining what the correct decision should be is a difficult process due to the fact that the decision maker is typically in a state of uncertainty. The uncertainty can arise from incomplete knowledge about the domain, or it can arise by indeterministic outcomes from a particular action. In previous studies (Stankiewicz, Legge, Mansfield & Schlicht, 2005) the researchers localized human, sequential, decision-making inefficiencies to their inability to accurately generate and update the probability of being in a particular state given the previous actions and observations. In the current studies we investigate whether subjects’ decision-making performances could be improved in a sequential, decision-making task under uncertainty that involved a seek and destroy mission. We present three studies in which subjects were tested under three conditions: no cognitive aid; An action/observation sketch pad, and a likelihood display. In each study performance was not significantly different in the first two conditions, but performance improved significantly when subjects had an explicit representation of the probability vectors. We compared human performance to the optimal performance and found that performance (using Partially Observable Markov Decision Processes) improved from roughly 50% efficiency to 85% efficiency.
Presentations

Invited Talks


Participation/Conference Presentations


Introduction

One of the most fundamental aspects of human cognition is our ability to make decisions. Humans can make decisions in a broad range of domains. For example, everyday we decide what time we will leave for work, where or what we will eat for lunch, what we will wear to work in addition to thousands of other decisions. These decisions oftentimes seem mundane because the ramifications of a “poor” decision are not significant. However, other decisions appear to be more critical and the ramification of a “poor” decision can appear, and often are, more detrimental. For example, deciding whether to attend college or not, medical decisions, and which job offer to accept all have ramifications not only on one’s current state, but they also have ramifications on future states and decisions that can and cannot be made.

It is important to recognize that most decisions that we make are not “one off” decisions in which we make the decision and then reap the reward or endure the punishment. Instead, most decisions that we make have future ramifications and affect the options and decisions that are available to us later. For example, when deciding whether or not to attend a university or college, that decision has an immediate “punishment” in that it costs money to attend college in addition to the stress and labor involved for graduation. However, after deciding to attend college and completing college, one has the opportunity to reap larger rewards than they would without a college education. Therefore, under
certain circumstances, attending college will increase one’s cumulative rewards despite the immediate costs.

However, under many conditions it is not clear what the rewards will be. That is the results of specific decisions or the rewards collected are not deterministic, but instead are determined in a probabilistic form. For example, the likelihood of obtaining a job with a salary greater than $100,000 after graduating from college is probably not 0.0 nor is it 1.0, but there is a certain likelihood of obtaining that level of income given the decisions that are made while someone is in college (e.g., what major one decides to pursue). Therefore calculating the expected reward (or cumulative reward) for going to college versus not going to college can be challenging due to the probabilistic nature of this decision.

The current research focuses on a different sequential decision making problem with uncertainty than the “Go To College” problem. The current research focuses on a task that is commonly faced by decision makers in the military. Namely a 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 him to gain access about the true state of the system (i.e., the location of the enemy) in addition to changing the state of the system (i.e, 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. Each of these actions generate probabilistic outcomes. That is, reconnaissance action will not always detect the enemy when it is sent to the enemy’s location. Furthermore, the reconnaissance may also falsely report that they saw the enemy at a location in which the enemy is not located. Furthermore, the artillery will not always move the enemy to the “Destroyed” state when it is sent to the right location.

In addition to noisy actions and observations the tasks conducted here also have a cost/reward structure. That is, each action has a specific cost. Furthermore, there is a reward for destroying the enemy when you declare “Done” and there is a negative reward if you declare done when the enemy is still alive.

Three studies and three models were developed to study the cognitive limitation in this type of Seek and Destroy task. In the first study we were studied human performance when the enemy’s location was static and subjects and the ideal observer had unlimited resources. In this study the subject’s task was to localize the enemy and destroy the enemy and generate the largest possible reward. We compared human performance to the optimal performance using a Partially Observable Markov Decision Process (POMDP) to get a measure of decision making efficiency. In Experiment 2 the task was the same but the enemy’s location dynamically changed on every turn. In this study the enemy was attempting to reach a Target Location and the subject’s task was to destroy the enemy before it reached that location. In Experiment 3, we studied subjects under a collection of resource conditions and the subject’s task was to destroy the enemy with the resources that were available.

In all three studies subjects were tested in three different conditions that investigates specific cognitive deficits that might lead to poor decision making relative to the optimal observer: Last Observation, All Observations and Belief Vector. In the Last Observation

---

1In this study and Experiment 2, the reconnaissance and the artillery generated a negative reward along with a negative reward for declaring finished when the enemy was still alive. There was a positive reward for declaring finished when the enemy was actually destroyed.
condition, subjects were given only their last observation and action. In this condition subjects were required to remember what actions they made, the results of those actions (observations) and update where they believed that the enemy was located. In the All Observations condition subjects were given a display in which their previous actions and observations were displayed, but subjects were still required to update their belief about the enemy’s state. In the Belief Vector condition subjects were given a display with all of the previous actions and observations, but were also given the likelihood of the enemy’s current state. However, the subject was still required to make the necessary decisions on what action to make.

In all three studies subjects’ performances in the Last Observation and All Observations conditions were not significantly different. However there was a significant improvement in performance when they were given an accurate Belief Vector of the enemy’s current state.

The Optimal Observer

To best evaluate human 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 one can optimally estimate the likelihood of the true state of the system, but this likelihood does not instruct one on what action one should select. In order to do action selection, one must not only calculate the current state given the previous actions and observations, but also calculate the optimal action to be performed in a given belief state.

The Bayesian statistics that should be used for sequential decision making with uncertainty use Partially Observable Markov Decision Processes (POMDP) (e.g., see Cassandra, Kaelbling, & Littman, 1994; Cassandra, 1998; Kaelbling, Cassandra, & Kurien, 1996; Kaelbling, Littman, & Cassandra, 1998; Sondik, 1971). By defining the State Space, Observation Vector Transition Matrix and the Reward Structure one can compute the expected reward for a particular action. In the following sections we will describe what each of these are, in addition to describing how to optimally update one’s belief (Belief Updating) given these definitions in addition to how to compute the Expected Reward for generating a particular action.

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). The problems used in this project, the hidden state was the enemy’s current position within the 5x5 area plus an additional “Destroyed” state that the enemy could transition into following an artillery strike at their current position.

Defining the Observation Vector. Although the true state is hidden, the observer is always armed with actions and observations that provide information about the true state of the problem. In the current problem, the observer fire artillery at a specific position or can send reconnaissance to a particular location within the environment (i.e., one of the 25 locations). The current problem has three observations: Enemy Sighted, No Enemy Sighted
or No Information. When the observer decides to send reconnaissance to a particular location he or she will receive one of two observations: Enemy Observed or No Enemy Sighted. In the current problem, the artillery only returns one possible observation: No Information.

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 in Region” and “No Enemy in Region”.

<table>
<thead>
<tr>
<th>Action</th>
<th>Observation</th>
<th>State</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recon.</td>
<td>Enemy Sighted</td>
<td>Enemy in Region</td>
<td>0.75</td>
</tr>
<tr>
<td>Recon.</td>
<td>No Enemy Sighted</td>
<td>Enemy in Region</td>
<td>0.25</td>
</tr>
<tr>
<td>Recon.</td>
<td>Enemy Sighted</td>
<td>No Enemy in Region</td>
<td>0.2</td>
</tr>
<tr>
<td>Recon.</td>
<td>No Enemy Sighted</td>
<td>No Enemy in Region</td>
<td>0.8</td>
</tr>
<tr>
<td>Strike</td>
<td>NoInfo</td>
<td>Enemy in Region</td>
<td>1.0</td>
</tr>
<tr>
<td>Strike</td>
<td>NoInfo</td>
<td>No Enemy in Region</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**Actions: Defining the Transition Matrix.**

In the Seek & Destroy problem, the observer has fifty-one different actions. There are twenty-five reconnaissance actions (one to each of the 25 locations in the environment), twenty-five artillery actions (again, one to each of the twenty-five locations within the environment) and the “Declare-Finished” option. The transition matrix defines the probability of the resulting state if the observer generates a particular action in a given state (i.e., \( p(s'|s,a) \)). In the static form of the Seek & Destroy problem there is only one state transition that could occur. When the observer sends artillery to the location where the enemy is located, the enemy will transition into the “Destroyed” state with a probability of 0.75.

**Belief updating.** Given an initial probability distribution over the state space, the Observation Matrix and the Transition Matrix one can begin to generate hypotheses about the current state of the problem following an action and an observation. Equation 1 provides the Bayesian updating rule.

\[
p(s'|b,o,a) = \frac{p(o|s',b,a)p(s'|b,a)}{p(o|b,a)} \tag{1}
\]

Equation 1 specifies how the ideal observer would update its belief that \( s' \) is the true state given the prior belief \( b \), the observation \( o \) and the action that was generated \( a \).

To illuminate the process of belief updating we will provide a simple example of a smaller Seek and Destroy problem. We will use the Transition and Observation Matrices used in the studies, but to simplify the process we will use a three-state problem instead of a 26-state problem. More specifically, the enemy will be in one of three states: State\(_1\), State\(_2\) or Destroyed. The prior probability of the state will be: State\(_1\)=0.5, State\(_2\)=0.5 and Destroyed=0.0 (to simplify, we will represent this as \([0.5, 0.5, 0.0]\)).
Let us first assume that the observer decides to do reconnaissance to State \(1\) and receives a “Enemy Sighted” observation. What is the likelihood that the enemy is in State \(1\), State \(2\) or Destroyed.

Using Equation 1 we can update the likelihood that the enemy is in State \(1\). That is, we want to compute: \(p(\text{State}1|0.5,0.5,0.0,\text{"EnemySighted", Recon}_1)\).

First, let us compute \(p(o|s', b, a)\) or \(p(\text{"EnemySighted"}|\text{State}_1, [0.5,0.5,0.0], \text{Recon}_1)\). To do this we need to compute the likelihood that we would get an observation of “Enemy-Sighted” if State \(1\) was the true state. In the Observation Matrix section above, we define the likelihood of correctly identifying the enemy as 0.75. We also need to compute 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, these remain at the prior probabilities which both are 0.5. Finally we have to compute the likelihood of receiving the observation “EnemySighted” when Reconnaissance is made at State \(1\) or \(p(\text{"EnemySighted"}|[0.5,0.5,0.0], \text{Recon}_1)\).

\[
p(o|b,a) = ((0.5 \times 0.75) + (0.5 \times 0.2) + (0.0 \times 0.2))
= 0.475
\]

\[
p(\text{State}1|[0.5,0.5,0.0], \text{"EnemySighted", Recon}_1) = \frac{0.75 \times 0.5}{0.475}
= 0.7895
\]

Furthermore,

\[
p(\text{State}2|[0.5,0.5,0.0], \text{"EnemySighted", Recon}_1) = \frac{0.20 \times 0.5}{0.475}
= 0.2105
\]

and Finally,

\[
p(\text{Destroyed}|[0.5,0.5,0.0], \text{"EnemySighted", Recon}_1) = \frac{0.2 \times 0.0}{0.475}
= 0.0
\]

Thus, if the first action is to observe at State \(1\), the new belief vector would be \([0.7895,0.2105,0.0]\).

Now, let us imagine that the observer decided on the action Strike \(1\) following the action Recon \(1\). To update the belief that the enemy is in State \(1\) we need to compute \(p(\text{State}1|0.7895,0.2105,0.0, \text{"NoInfo", Strike}_1)\).

First, we need to compute \(p(o|s', b, a)\) which is the probability of receiving the “NoInfo” observation given that the true state is State \(1\), the current belief \(([0.7895,0.2105,0.0])\) and the action Strike \(1\). The probability of receiving this observation is 1.0. Regardless of the state of the problem a Strike always returns the observation “NoInfo” (see Table 1). This same probability and logic holds for computing \(p(o|b, a)\).
We also need to compute \( p(s'|b,a) \). That is, what is the probability of the true state being \( \text{State}_1 \) given the current belief and the action \( \text{Strike}_1 \). As described in Section on the Transition Matrix, the probability of transitioning the problem into the Destroyed state is 0.75 if the enemy is at the location where the artillery strike occurred. This means that there is a probability of 0.25 that the enemy’s state will not change, or that the enemy will remain in \( \text{State}_1 \) if it was initially in \( \text{State}_1 \).

\[
\begin{align*}
p(o|s',b,a) & = 1.0 \quad (10) \\
p(s'|b,a) & = 0.7895 \times 0.25 \\
p(o|b,a) & = 1.0 \\
p(\text{State}_1|[0.7895, 0.2105, 0.0], "\text{NoInfo}'', \text{Strike}_1) & = \frac{1.0 \times 0.7895 \times 0.25}{1.0} \\
& = 0.1997
\end{align*}
\]

\[
\begin{align*}
p("\text{NoInfo}''|\text{Destroyed}, [0.7895, 0.2105, 0.0], \text{Strike}_1) & = 1.0 \quad (11) \\
p(\text{State}_2|[0.7895, 0.2105, 0.0], \text{Strike}_1) & = 0.2105 \\
p("\text{NoInfo}''|[0.7895, 0.2105, 0.0], \text{Strike}_1) & = 1.0 \\
p(\text{State}_2|[0.7895, 0.2105, 0.0], "\text{NoInfo}''', \text{Strike}_1) & = \frac{1.0 \times 0.2105}{1.0} \\
& = 0.2105
\end{align*}
\]

\[
\begin{align*}
p("\text{NoInfo}''|\text{Destroyed}, [0.7895, 0.2105, 0.0], \text{Strike}_1) & = 1.0 \quad (12) \\
p(\text{Destroyed}|[0.7895, 0.2105, 0.0], \text{Strike}_1) & = 0.7895 \times 0.75 \quad (13) \\
p("\text{NoInfo}''|[0.7895, 0.2105, 0.0], \text{Strike}_1) & = 1.0 \\
p(\text{Destroyed}|[0.7895, 0.2105, 0.0], "\text{NoInfo}''', \text{Strike}_1) & = \frac{1.0 \times 0.7895 \times 0.75}{1.0} \\
& = 0.5921
\end{align*}
\]

Thus the belief vector following \( \text{Recon}_1 \) and receiving the observation of “EnemySighted” followed by \( \text{Strike}_1 \) the resulting belief vector is \([0.1997, 0.2105, 0.5921]\). Or

\[
\begin{align*}
p(\text{State}_1) & = 0.1997 \\
p(\text{State}_2) & = 0.2105 \\
p(\text{Destroyed}) & = 0.5921
\end{align*}
\]

**Defining the Reward structure.** Equation 9 and 1 formalize how the optimal observer would update its belief about the enemy’s current state given its previous actions, observations and prior probabilities. However, it does not specify which action that the observer should take. In order to specify the optimal action one must specify a **Reward Structure**
for the problem. A reward structure specifies the expected value (both positive and negative rewards) for generating an action in a particular state. For example, in the Seek and Destroy task, the reward (or cost) for firing artillery to a specific position is invariant with the enemy’s state (i.e., \( R(\text{"Artillery}''_{xy}|\text{EnemyState}) = R_{\text{Artillery}}) \). However, the reward for declaring finished changes depending on the state of the enemy. That is, the reward for declaring finished when the enemy is not in destroyed is one value and the reward for declaring finished when the enemy is still in one of the position states is another value (i.e., \( R(\text{"DeclareFinished"}|\text{Destroyed}) \neq R(\text{"DeclareFinished"}|\text{Destroyed}) \)).

Choosing the optimal action

Using the Transition matrix, Reward Structure and Observation Matrix we can compute the optimal action for any given belief state that the observer might have\(^2\). To compute the optimal action one must consider the immediate reward for generating a specific action in a specific state. The following equation is the reward for generating a single action given that you a particular belief vector \( b \).

\[
\rho(b, a) = \sum_{s \in S} R(s, a) b(s)
\]  

\( \rho \) provides the immediate reward for making the action \( a \). Under most circumstances there is no clear choice for the observer if their choice is myopic – that is, if they only consider the immediate rewards. Given the structure of the problem, one can begin to consider the actions to help differentiate between which of the immediate actions is optimal. Computing the expected value when one is considering future actions is an iterative process – a process that has to consider future actions and beliefs. To compute the maximal expected value we want to compute the rewards that are acquired as the observer is moving through the problem. That is, the reward will be:

\[
R = r(0) + \sum_{t=1}^{\infty} r(t)
\]  

To compute the optimal action for a given belief vector, one needs to consider all of the actions that are available to the observer in the current state \( A \) the immediate reward for generating an action \( \rho \) from Equation 16) and the transition function that specifies the likelihood that a new belief state will be generated \( \tau \). By choosing the action that maximizes the following function one can act optimally in the environment.

\[
V(b) = \max_{a \in A} \left[ \rho(b, a) + \sum_{b' \in B} \tau(b, a, b') V(b') \right]
\]  

Experiment 1: Static Enemy

Our initial investigation into understanding the cognitive limitations associated with sequential decision making with uncertainty leads us into a typical military “Seek & Destroy” scenario. In this task the subject was given a 5 x 5 area in which a single enemy

\(^2\)A belief state is a particular probability distribution across all of the possible states in the environment.
was located. Using reconnaissance and artillery strikes, the subject’s task was to destroy the enemy using the minimal amount of resources as possible.

We conducted the study on three subjects to evaluate the cognitive limitation associated with sequential decision making with uncertainty. More specifically, we were investigating whether the cognitive limitation was associated with the subject’s memory for the actions and observations made during the sequence of actions, or whether the cognitive limitation was associated with the subject’s inability to accurately update the likelihood that the enemy was in each state.3

In the study we used a reconnaissance action that was not limited to the position that was selected, but the reconnaissance had an extended footprint (see Figure 1). That is, reconnaissance was making an observation, where available, just above, below and beside the position that was selected.

**Observation Vector**

In addition to the extended reconnaissance region, the experiment also had noisy actions. That is, the likelihood of detecting the enemy if the enemy was within the footprint was not 1.0, but instead it was 0.9. Furthermore, there was a certain probability that the reconnaissance would make a false alarm – report that the enemy was detected when it was not within the reconnaissance footprint. Reconnaissance would false alarm 5% of the time (or with a probability of 0.05).

**Transition Matrix**

There was only one type of action that would result in a state transition and that was the artillery action. However, these actions were also probabilistic. When the subject made an artillery strike to a particular location, the enemy would enter the Destroyed state with a probability of 0.75. That is, the artillery would miss with a probability of 0.25. If the enemy was not at that location, then the probability of the enemy entering the Destroyed state was 0.0.

**Reward Structure**

In the first study the cost for making reconnaissance to any of the locations was -35 units. The cost for making an artillery strike was -100 units. There was also a positive reward when the subject Declared Finished when the enemy was in the Dead State of 500 units. There was also a cost for declaring finished when the enemy was not in the dead state of -750 units.

**Design**

Subjects participated in three different conditions: Base-Line, Memory and Belief-Vector. In the first condition, the subject was required to do all of the computations on his/her own. This served as a baseline condition to evaluate the expected reward when all of the actions and observations were completed without any explicit representations by the

---

3In the current problem there are 26 states that the enemy could be in. The 25 positions in the environment plus the one Destroyed state.
Figure 1. An illustration of the Reconnaissance Footprint used in each of the studies. When Reconnaissance was sent to a particular position (e.g., 2,2) the enemy could be detected in the positions just above, below and beside this position also. In this illustration, the gray region illustrates the reconnaissance footprint for a reconnaissance action to 2,2.
experiment. In the second condition we provided a **Memory** aid that illustrated the previous actions and observations at each of the locations. If subjects showed an improvement in this condition over the Base-Line condition, then we would conclude that subjects’ performances are limited by a memory limitation associated with the actions and observations made. In the third condition we presented the subject with an explicit representation of the current **Belief Vector** (i.e., the likelihood that the enemy is at a particular location). We presented this likelihood in two ways. We presented the explicit probability in numerical format **and** we also modified the intensity of each square that was dependent upon the current likelihood. The darker the cell, the less likely the enemy was in that state.

**Results**

Figure 2 illustrates the performance of the three subjects who participated in Experiment 1. The upper 3 plots are the average Reward for fifty trials in each condition for each subject. The lower three plots present the subjects’ efficiency values for each of these conditions. Efficiency was calculated as the Expected Reward for the optimal observer, divided by the Average Expected Reward for the subject.

![Figure 2](image)

*Figure 2.* The average reward for the three subjects in the three conditions in Experiment 1 (upper three plots) and their performance relative to the optimal observer (lower three plots).

**Summary**

All three subjects showed a slight to no improvement in the Last-Observation (Base-Line) Condition over the All-Observation (Memory) condition. This suggests that memory
is not the limiting factor in making optimal decisions in a sequential decision making task. However, there was a significant improvement in performance from the All-Observation condition to the Belief Vector Condition. This suggests that one of the significant factors in limiting subjects from making optimal decisions is their inability to effectively integrate the actions and observations into an accurate belief vector.

**Experiment 2**

Experiment 2 extends the design of Experiment 1 in a number of ways. Perhaps the most important is that it begins to move the paradigm into a more “realistic” situation. In Experiment 2 the subject is faced with a Seek & Destroy task in which the enemy’s position is no longer static, but it is actually changing after each iteration. The enemy’s position is not changing randomly, but instead, the enemy is attempting to reach a particular state in the problem space before it is destroyed. The goal of the subject is to destroy the enemy before it reaches this state with the minimum amount of resources possible. We also changed the prior probability of the enemy’s state (that is the starting position of the enemy). In Experiment 1 it was a uniform distribution over all of the positions in the environment. In Experiment 2 it was the upper two rows of the environment.

**Observation Vector**

As with Experiment 1, the current study used an extended footprint for the reconnaissance. However, we changed the Hit and False-Alarm rate of the reconnaissance. The likelihood of detecting the enemy when the enemy is within the reconnaissance footprint was lowered to 0.75 and the likelihood of making a false alarm increased to 0.2.

**Transition Matrix**

We used the same transition matrix as was used in Experiment 1 with respect to the artillery strikes. However, because the enemy was dynamic there was a certain probability that the enemy’s position would change on each trial. Thus, the probability that the enemy would stay in its current state was 0.8. The probability that the enemy would move horizontally toward the goal (position 2,4) was 0.05, the probability that the enemy would move vertically toward the goal was also 0.05 and the probability that the enemy would move diagonally toward the goal was 0.10.

**Reward Structure**

We modified the reward structure from Experiment 1. In Experiment 2 a reconnaissance action cost -25 while artillery still cost -100. The reward for declaring finished when the enemy was destroyed was 1000, while declaring finished when the enemy was not destroyed was -2500.

**Design**

As in Experiment 1, subjects ran in three different conditions: Last-Observation, All-Latest-Observations and Belief-Vector. Figure 3 provides a sample illustration for the three viewing conditions used in Experiment 2.
Figure 3. An illustration of the three viewing conditions used in Experiment 2. The display on the far left illustrates a display in the Last-Observation condition. In this condition, only the last action and observation is displayed on the screen. The center panel illustrates a display in the All-Latest-Observations condition. In this display condition the previous actions and observations are displayed to the subject. The right panel illustrates a sample display in the Belief-Vector condition in which the subject has made a sequence of actions. The display shows the current likelihoods of the enemy in each state, including the “Destroyed” (Dead) state in two ways. First by displaying the actual values in text and also by varying the intensity of each cell proportional to the probability that the enemy is in that particular state.

Results

Data were collected on six subjects. As in Experiment 1 we computed the average reward for three different conditions (Last-Observation, All-Observations and Belief-Vector) and compared that to the expected reward for the ideal observer. We computed an performance efficiency by taking the ratio of the subjects’ average rewards to that of the optimal observer’s expected reward. The average efficiency for the six subjects and the standard error of the mean are illustrated in Figure 4. There was no significant difference between the Last Observation condition and the All Observation condition. However there was a significant difference between the All Observation condition and the Belief Vector condition.

Summary

Experiment 2 investigated whether the cognitive limitation preventing subjects from performing ideally was a memory limitation or in computing an accurate belief vector. This study extends the paradigm used in Experiment 1 to a situation in which the environment is constantly changing – more specifically one in which the enemy’s position can change after each action. The results from Experiment 2 suggest that there is a cognitive deficit in the subject’s ability to accurately update the likelihood that the enemy is in a particular location. However, there does not seem to be much of a memory deficit.

Experiment 3

Oftentimes when participating in a sequential decision-making task, one has to maximize the likelihood of success given a certain amount of resources. That is, the decision
Figure 4. Average efficiency for the six subjects in Experiment 2 (error bars represent one standard error of the mean). There was no significant difference in performance between the Last Observation and the All-Latest-Observation conditions. However, there was a significant improvement in efficiency between the All-Latest-Observations and the Belief Vector condition.

maker is constrained by the amount of resources available to him/her when making a decision. In Experiment 3 we investigated human performance in the Seek and Destroy task when subjects were given only a certain amount of resources. This study used the Observation Vector defined in Experiment 2 and used the Transition Matrix defined in Experiment 1 (the position of the enemy remained static).

In Experiment 3 the goal of the subject was to maximize the likelihood that the enemy was destroyed with 250, 500, 750, 1000 and 1250 resource units. Once these resource units were used the subject was forced to declare that he/she was finished.

Results

Plotted in Figure 5 is the average probability that the enemy was destroyed with the three viewing conditions and the five resource levels. Performance for each subject in each condition is also fitted with a logistic function to get a better idea of how well subjects performed in each condition.

As can be seen in Figure 5, subjects performed best in the Belief Vector Condition over the other two conditions in every case. Consistent with the results from Experiment 1 and 2, these results suggest that a significant cognitive limitation associated with acting optimally in a sequential decision-making task with uncertainty is the inability to accurately estimate the likelihoods given the previous actions and observations in the decision sequence.
Figure 5. Results of the probability of successfully destroying the enemy given limited resources. The probability of success was fitted with a logarithmic function. For all eight subjects, the performance for the Belief Vector condition was better than the Last Observation and All-Latest Observation conditions.
General Discussion

A significant number of decisions in the military are decisions that are made under uncertainty. Despite the incredible improvements in technology for information gathering, reconnaissance, and artillery, often times high ranking officers have the need to make decisions when they are not certain about the current state of the battlefield. This state of uncertainty can range from uncertainty about the enemy’s position, the strength of the enemy, to the current position of a reconnaissance plane, and whether or not an artillery strike was successful or not.

The current project investigated the cognitive limitations to optimal decision making in a sequential, decision-making with uncertainty task. The focus of this research was to understand the cognitive limitations associated with this task by measuring human performance and comparing that performance with that of an optimal observer. The optimal observer was faced with the same degree of uncertainty and had to deal with the same probabilistic actions and observations as the human observer. However, because the optimal observer used Bayesian statistics and decision processes, this model is not limited by any cognitive functions such as memory, decision making strategies or belief updating strategies. That is, the model provides us with the theoretical best performance for the task.

In these three studies the researchers wanted to isolate whether the cognitive limitation was in the subjects Memory system, Likelihood Estimation or Decision Making strategy. To investigate this question we studied subjects with three viewing conditions. The first viewing condition provided no cognitive aid (i.e., neither a memory aid nor an aid showing an accurate belief vector) and served as our base-line condition (Last Observation). In this condition the subject had to integrate all of the previous actions and observations to generate the likelihoods that the enemy was in each of the 26 different states. In the second condition we presented the subject with a memory aid. In this condition the previous actions and observations remained on the screen during the entire trial. We hypothesized that if the limitation was in the subjects inability to remember the previous actions and observations then this second condition should improve their performance over the base-line condition (i.e., the Last Observation Condition).

In the final condition (Belief Vector) we presented the subjects with an accurate representation of the likelihoods that the enemy was in each of the 26 states given their previous actions and observations. We hypothesized that if the cognitive limitation was in the subject’s inability to accurately update the likelihoods then we should find a significant improvement in this condition over the Last Observation Condition and the All-Latest Condition (memory aid condition).

In all three experiments we found no significant improvement in the subjects’ performances in the base-line and Last Observation conditions. This suggests that memory is not the limiting factor. In all three experiments we found a significant improvement between the Last Observation conditions and the Belief Vector conditions. This result suggests that one limiting factor preventing optimal decision making is the subject’s inability to accurately estimate the current likelihood of the enemy’s state given the previous actions and observations.
References


