AFIT/GOA/ENS/99M-05

A BAYESIAN DECISION MODEL
FOR BATTLE DAMAGE ASSESSMENT

THESIS

Daniel W. Franzen, Captain, USAF

AFIT/GOA/ENS/99M-05

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the US Government.

AFIT/GOA/ENS/99M-05

A BAYESIAN DECISION MODEL
FOR BATTLE DAMAGE ASSESSMENT

THESIS

Presented to the Faculty of the Graduate School of Engineering of the

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Analysis

Daniel W. Franzen, B.S., M.A.S.

Captain, USAF

March 1999

AFIT/GOA/ENS/99M-05

A BAYESIAN DECISION MODEL

FOR BATTLE DAMAGE ASSESSMENT

Daniel W. Franzen, B.S., M.A.S.
Captain, USAF

Approved:

RAYMOND R. HILL, Maj, USAF
Assistant Professor of Operations Research

10 Mar 99
date

GREGORY A. MCINTYRE, Lt Col, USAF
Assistant Professor of Operations Research

10 Mar 99
date

JACK M. KLOEBER, Jr., LTC, USA
Assistant Professor of Operations Research

10 Mar 99
date

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Major Ray Hill, for his guidance and encouragement throughout the course of this thesis effort. I would also like to thank the other members of my committee, Lt Col Greg McIntyre and LTC Jack Kloeber, Jr. All of their insights were critical to maintaining the focus of this effort, and more than once, they provided me with the means or the motivation to overcome obstacles.

I also owe thanks to the targeteers who contributed many hours of their time introducing a missileer to the world of air strikes, especially Capt Rob Killefer and MSgt Pat Zwenger of the 608th Air Intelligence Squadron. Their patience and flexibility made this thesis possible. Their contributions are all the more significant in light of the continuing U.S. operations in the Middle East.

I would additionally like to acknowledge the unceasing efforts of Saddam Hussein and the Iraqi regime, who continue to provide a dynamite motivation for this thesis and ensure its relevance. Unfortunately, this came at the price of restricted access to subject matter experts, since most of them were in Southwest Asia.

In all seriousness, however, I wish most of all to thank my wife Marlee. Her continual encouragement and love over the past eighteen months have kept me going. I deeply appreciate her support and patience through the long dark hours of the AFIT experience.

Dan Franzen

Table of Contents

Page

List of Figures

Figure Page

List of Tables

AFIT/GOA/ENS/99M-05

Abstract

Battle damage assessment (BDA) is critical to success in any air campaign. However, Desert Storm highlighted numerous deficiencies in the BDA process, and operations since Desert Storm continue to point out weaknesses. We present a review of the Phase I BDA decision, or physical damage assessment, and model the decision process using a Bayesian belief network. Through subject matter expert (i.e., the targeteers) elicitation sessions, imagery was found to be critically important to the BDA process yet this information is generally not retained. This use of "perfect information" is delineated in the BDA process models. We proposed a methodology based on Bayesian belief networks for incorporating this perfect information. We demonstrate the Bayesian belief network's capability to update conditional probability distributions using data generated in real world operations. This capability allows the network's conditional distributions to evolve, increasing model accuracy and reducing uncertainty in the decision.

A BAYESIAN DECISION MODEL FOR BATTLE DAMAGE ASSESSMENT

CHAPTER 1

The research documented in this thesis is sponsored by the Air Force Command and Control Battlelab (C2B) located at Hurlburt Field, Florida. The C2B was established "...to identify innovations in command and control and battle management operations concepts and measure their potential to advance Air Force core competencies and joint warfighting" (AF/XOR, 1998). Toward that end, the C2B sought to investigate ways to automate or speed up the battle damage assessment (BDA) process, and enlisted AFIT to spearhead a research effort.

Our goal in this effort is threefold. First, we present a detailed description of the BDA process, drawing on guidelines and regulations as well as interviews with targeteers. This description represents a significant portion of the value of this research, acting as an outsider's perspective and encouraging discussion about the decision process. Second, we construct a model of Phase I BDA, or physical damage assessment. Finally, we present a methodology for implementing this model and improving it through historical data. This model can serve as the basis of an eventual automated decision tool for BDA.

The remainder of this thesis is devoted to the BDA decision model, from the foundations on which it is built to the future directions it may take. Chapter 2 is formatted as a stand-alone article suitable for submission to an academic journal.

Chapter 3 outlines possible improvements and extensions of this work. Appendix A contains the complete Bayesian models presented in Chapter 2. Appendix B describes the motivation behind the choice of the Bayesian belief network methodology over other methodologies frequently used in decision modeling and analysis.

This research initiated as an attempt to build a decision matrix for BDA—a simple, spreadsheet-style way for targeteers to make a quick assessment of damage given various pieces of evidence. However, our research shows the number of variables in the BDA problem and the number of potential states they may take on would make such a matrix or spreadsheet intractably large. In other words, the paradigm implicit in a decision matrix model is simply not suited to the BDA problem. However, the Bayesian belief network model offers two possible substitutes for a decision matrix. The first pseudo-decision matrix is provided in the form of the conditional probability distributions. The conditional distributions, shown in numerous tables in this work, provide an indicator for the likelihood of assessing a particular damage level given the evidence. This converts easily to a decision matrix by simply choosing, in all cases, to assess the damage level with the highest number. A second possible way to reach quick decisions is simply to query the network directly. Many Bayesian network software packages offer easy ways to enter evidence and query nodes for their distribution. Using this capability, we can set the evidence nodes to the appropriate state and determine which damage level is most likely to result in an accurate assessment. The same method can serve as a rough way to conduct sensitivity analysis. Bayesian networks do allow for some sensitivity analysis, but this capability has not been incorporated into a software application at this time. The reader is referred to Jensen, 1996 and Cozman, 1998 for more information on sensitivity analysis in Bayesian networks.

CHAPTER 2

2.1 Introduction and Overview

The effective prosecution of an air war demands efficient and effective use of airpower assets against important enemy targets. Battle damage assessment (BDA) is a crucial part of combat assessment (CA), which in turn is an essential element of the targeting cycle. During BDA, intelligence specialists known as targeteers attempt to determine how much damage friendly forces have inflicted on struck targets. The ultimate decision is whether airpower forces achieved their target damage objective, or further strikes against the target are necessary. BDA is of critical importance to any military campaign. However, recent military operations have highlighted a need to improve the BDA process. During Desert Storm, U.S. and allied forces found themselves severely in need of faster and more accurate BDA. The advent of precision-guided munitions led to pinpoint strikes that often left little visible damage on the target, making the damage assessment more difficult. Additionally, the widespread use of precision-guided munitions resulted in a dramatic increase in the need for damage assessments. The two effects combined to outstrip the abilities of the BDA system. Since Desert Storm, few significant technological developments have appeared to help remedy the problem. In numerous operations below the level of war, BDA has been an essential factor from both the operational and media-relations perspectives.

2.1.1 Problem Statement.

Because BDA directly affects restrike decisions, it affects the entire targeting cycle, so accurate and timely BDA is an absolute must. However, during Desert Storm and in operations below the level of war since 1991, commanders and operators alike have pointed out a need for better BDA tools. The official Department of Defense (DoD) review of the Gulf war viewed the BDA process as taxed beyond its ability (DoD, 1992). Numerous studies conducted since the conclusion of Desert Storm have also been critical of the BDA process and pointed out several areas for improvement (Hallion, 1992; Smith, 1993; Sweigart, 1993). Operation Desert Fox, conducted in 1998 against targets in Iraq, as well as the ensuing series of engagements with Iraqi forces, also highlighted the need for faster BDA. In some instances, the United States was uncertain as to the need for further strikes for a significant period of time, sometimes several days. Similarly, the DoD significantly upgraded its official estimates of damage to sites targeted during Desert Fox following an extensive review of all the information available. Clearly, there is a need to improve many aspects of the BDA process, especially timeliness and accuracy.

2.1.2 Background of Effort.

This effort seeks to improve the BDA portion of the targeting cycle by developing the methodology for, and a prototype of, a decision tool to assist the targeteer in determining whether target damage objectives were met. This research has three goals. First, we develop an accurate model of the targeteer's decision to facilitate learning about the BDA decision process itself. Second, we demonstrate this model's utility as a tool to improve targeteer decision timeliness. Finally, we demonstrate an approach wherein the model incorporates historical data to continually improve. Our focus is not the entire BDA process, nor all the different assessments that make up BDA. We focus primarily on the first stage of BDA, and demonstrate a modeling technique that is applicable to the other facets of the BDA process.

2.1.3 Research / Literature Review.

Despite wide awareness of the problem throughout the military community, there has been little work in the area of BDA decision support, and really only one effort to improve the BDA process through modeling and simulation. The Air Force Research Laboratory (AFRL) at Rome, New York created a BDA model entitled BDASIM beginning in 1994 (Rome Laboratory, 1996). However, the BDASIM effort focused on the intelligence structure architecture in an attempt to speed up the BDA process. The base task in the AFRL study was to develop a comprehensive end-to-end model of the BDA process. Once the base task was completed, the AFRL intended to study potential alternatives to the traditional BDA architecture and select a subset for further study. Finally, the BDASIM team intended to generate a high-fidelity end-to-end BDA model, seeking improvement in areas such as data fusion, decision accuracy, and processing techniques. However, only the base-level task of modeling the current BDA architecture was completed. Further, the AFRL effort looks for general solutions and improvements in BDA information architecture, rather than the specific decision model and support application sought in this effort.

2.1.4 Overview of Methodology.

The goal of this research effort is to prove the viability and utility of a Bayesian belief network model as an aid to the targeteer making BDA decisions. This is done by researching the BDA decision process and constructing a Bayesian belief network to model the decision of interest. The Bayesian belief network methodology involves identifying and modeling the various factors important in the targeteer's decision, and determining the range of values for these factors and the probabilities of these values occurring. Secondly, we demonstrate a significant strength of the Bayesian belief network model—its ability to combine an elicited expert knowledge base with data to improve its ability to model the real world. We illustrate this capacity through a constructed, representative (albeit notional) scenario and demonstrate the increased accuracy of the model after incorporating the data generated. The scenario consists of an air strike operation common enough to represent a typical BDA problem, yet complex enough to yield a variety of results and damage levels.

The BDA process is both large and complex. As a proof of concept, we focus primarily on the initial, or Phase I, BDA decision. This decision consists of the physical damage assessment, defined as "the quantitative extent of physical damage (through munition blast, fragmentation, and/or fire damage effects) to a target resulting from the application of military force" (DIA, 1996). Additionally, we limit the scope of the model to include only a portion of the many types of potential targets U.S. forces may face. Although limited, this model is sufficient to demonstrate the validity of the concept. We recognize that any final product based on these methods must incorporate a larger pool of expert knowledge and apply to more realistic, complex scenarios.

2.2 BDA Decision Model

2.2.1 Background.

The targeting cycle is an iterative process consisting of six phases: Objectives and Guidance, Target Development, Weaponeering, Force Application, Execution Planning / Force Execution, and Combat Assessment (see Figure 1). During the Objectives and Guidance phase, commanders determine attack goals and under what conditions to act. These objectives and rules must be understandable, achievable, and measurable in order to facilitate assessment at the end of the cycle.

The Target Development phase involves the examination of enemy military, political, or economic systems to identify critical targets and aimpoints. During Target Development, targeteers assess the level of physical and functional damage necessary to achieve the specified command objectives. This provides the targeteer with crucial information necessary to assess damage, such as target construction or functional layout.

The Weaponeering phase matches weapon characteristics to target vulnerabilities, taking into account secondary objectives such as minimizing collateral damage. The Weaponeering goal is to select the optimal weapon and delivery platform to achieve the necessary level of damage. This information is also crucial to BDA, allowing the targeteer to anticipate weapons effects and target response.

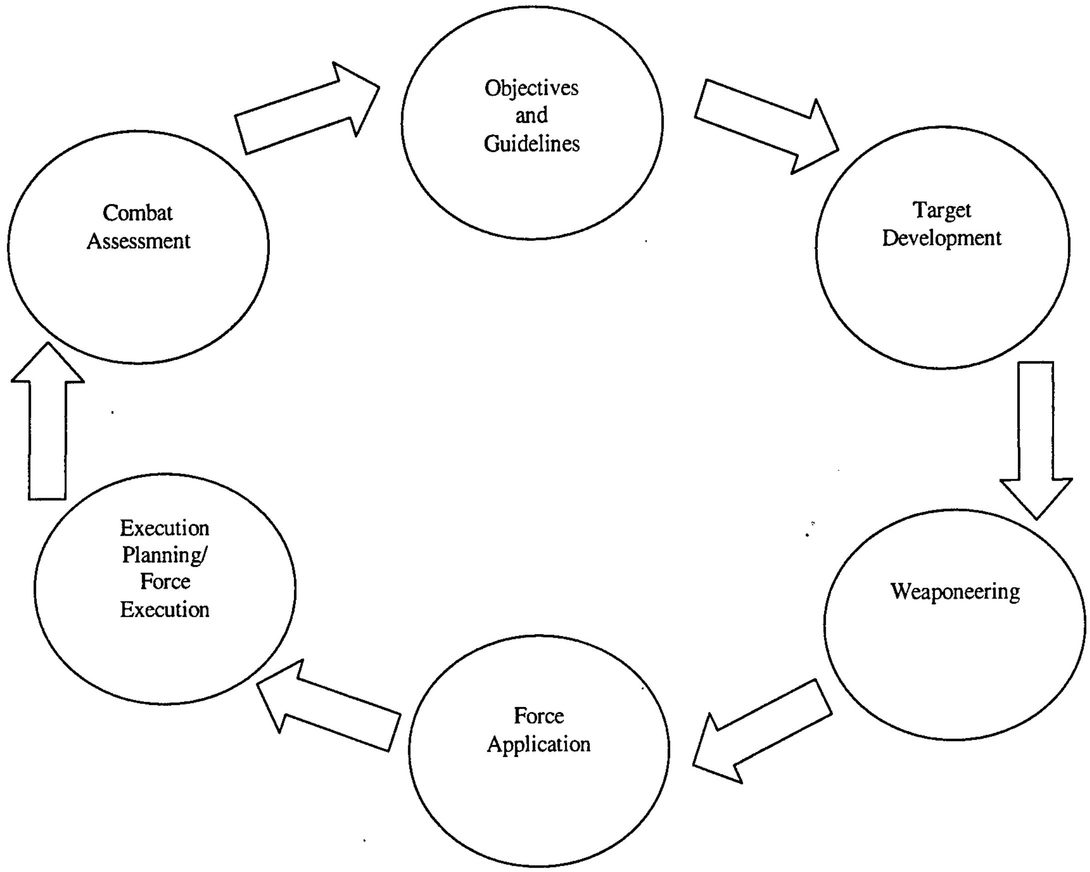

Figure 1. The Targeting Cycle (DIA, 1996)

Force Application matches available weapons and delivery platforms to the selected target aimpoints. This phase considers the realities of the operational world, such as the availability or delivery accuracy of specific weapons or platforms.

The Execution Planning and Force Execution phase generates mission-specific operational data. Targeteers finalize the specific numbers and types of weapons and delivery platforms, as well as time on target (TOT). This information helps targeteers plan for intelligence collection and future BDA following mission execution. Following the completion of the Execution Planning phase, the targeteer has a complete list of all the targets in the strike, the aircraft performing the strike, and the weapons on the aircraft to execute the strike.

The final phase of the targeting cycle is Combat Assessment (CA) which closes the loop and allows commanders to prepare for the next iteration of the cycle. As the *BDA Quick Guide* states, "The goal of CA is to determine the overall effectiveness of force employment during military operations and to recommend future courses of actions" (DIA, 1996).

Combat Assessment really consists of three interrelated processes: BDA, Munitions Effectiveness Assessment (MEA), and Reattack Recommendation. MEA evaluates the effectiveness of weapon systems and identifies possible deficiencies in weapon system performance or combat tactics. This information is useful in future weaponeering. By evaluating whether the weapons used performed as expected and were appropriate for the situation, targeteers can make a more informed choice of weapons and tactics. BDA, on the other hand, determines the effectiveness against the objective—whether the mission achieved its goal. Reattack Recommendation, using both BDA and MEA, makes a determination of what should be done next with respect to specific targets. The commander receives recommendations from the targeteers on reattack options, new targets to attack, or the use of different munitions or tactics.

The BDA process is composed of three distinct phases. Our focus here, Initial BDA or Phase I BDA, is an initial analysis and estimate of damage based primarily on visual observation of the target. Information usually comes from a few sources, such as aircrew mission reports (MISREPs), still imagery, or weapon system video. Phase II, Supplemental BDA, amplifies the initial analysis and evaluates functional damage to a target to estimate strike impact on the target system. Phase II sources include signals intelligence (SIGINT), imagery intelligence (IMINT), and measurement and signature intelligence (MASINT). Finally, Phase III BDA uses all supplemental BDA and the experience of subject matter experts to assess the remaining capacity of the overall target system. In this phase, "The bottom-line question is 'How successful have our efforts been to degrade or deprive the enemy's warfighting capabilities?'" (DAF, 1996). After resolving this question, targeteers provide a reattack recommendation to the commander.

The Phase I BDA is a complex, demanding problem. The targeteer must weigh numerous sources of information and make a quick, accurate decision. The decision is literally a matter of life and death, since the targeteer's assessment can influence the targets assigned for the next round of strikes. Since assessment of battle damage is a favorite media topic during any military strike operation, the targeteer's assessment could potentially appear on broadcasts around the world. Moreover, the targeteer at a busy Air Operations Center can potentially make dozens of these assessments during a single shift.

In assessing battle damage, the targeteer combines prior expert knowledge, database information, and new information (*i.e.*, intelligence) to arrive at a decision. The targeteer has some information available even before the first aircraft in the strike package takes off. For example, the targeteer knows the weapons and aircraft involved in the strike. In addition, the targeteer has a thorough knowledge of the different weapons and platforms and their strengths and weaknesses against various types of targets. Further, the targeteer is familiar with the targets in the strike and how heavy target area defenses are. All this information drives the targeteer to form a basic assessment of the chance of success of each aircraft/weapon combination against its assigned target.

During and after the strike, the targeteer receives more information. Pilots provide in-flight reports (INFLTREPs) as to whether they were able to engage and successfully hit the target. After returning to their bases, aircrews provide mission reports (MISREPs) with their assessment of the damage to the target. Some weapon systems provide videotape of the approach to the target. Analysis of such weapon system video can be very valuable to the BDA targeteer. Unfortunately, not all weapon systems provide such information. For example, the Tomahawk Land Attack Missile (TLAM), which has seen increasing use since Desert Storm, is unmanned, with no video capability, providing the targeteers less evidence on which to base their conclusions.

In Phase I BDA, the targeteer must combine expert knowledge with the available evidence to estimate a level of damage inflicted on the target. In general, damage is assessed as one of five levels: No Damage, Light, Moderate, Severe, or Destroyed, although other damage levels are employed for certain target types such as railroads or runways (DIA, 1996). Each of the five damage levels is clearly defined for each target type. For example:

> Satellite Dishes:
>
> > NO DAMAGE: No apparent/observable damage.
> > LIGHT DAMAGE: A few reflective panels blown off.
> > MODERATE DAMAGE: Less than 25 percent of dish reflective panels blown off plus damage to dish support structure and/or damage to feedhorn.
> > SEVERE DAMAGE: 25 to 60 percent of reflective panels blown off plus some deformation of the dish and/or the dish's structural components. Antenna pointing changed.
> > DESTROYED: Feedhorn is destroyed, and/or greater than 60 percent of reflective panels blown off, and/or extensive structural deformation of the dish, and/or dish knocked off its base. (DIA, 1996)

The targeteer makes an assessment of the likely damage inflicted, and can then compare the assessed damage level to the objective level defined in the Target Development phase. If the assessed damage level does not meet the objective, the targeteer then considers making a recommendation to restrike the target.

2.2.2 Targeteers' Description of the Phase I BDA Decision.

Elicitation sessions with targeteers from the 608th and 609th Air Intelligence Squadrons provided insight into the Phase I BDA decision. While discussing the BDA problem, some of the published guidelines regarding BDA were found somewhat flexible, and other unwritten rules or policies came to the forefront. These elicitation sessions yielded the structure of the decision model, which sources of information to include in the model, and the way in which these sources interact.

In making the physical damage assessment, targeteers rely on very few sources of information. This is due in large part to the extreme importance of time in the targeting cycle. Targeteers feel they do not have sufficient time to wait for multiple sources of information on which to base damage assessments. Further, targeteers view still imagery as the most important source, to the extent of treating still imagery analysis as perfect information (Curry, 1999; Killefer, 1999; Zwenger, 1999). Such still imagery may come from satellites or tactical reconnaissance aircraft. In any case, still imagery analysis requires several hours to obtain.

This time lag presents an opportunity for targeteers to employ other information available to them to make a predictive assessment, which can then be compared against the "perfect" information provided through imagery analysis. Improving predictive, pre-imagery assessments is the goal of this research, and could provide several benefits. Targeteers can save a significant amount of time if they can make an accurate assessment of physical damage without waiting for imagery. Further, the ability to assess damage without imagery facilitates more efficient use of imagery resources. If a targeteer does not require an image to assess damage, that imagery asset can be assigned to targets that are more difficult to assess.

To make a predictive, pre-imagery assessment, targeteers must rely on the other information available. This information falls into two categories: information available before the strike and information available just after the strike. Before the strike occurs, targeteers already know much about the strike package, including what weapons and platforms are included. The targets are carefully identified and developed during the early stages of the targeting cycle. Further, during the weaponeering phase, planners consider the target, objective damage level, and available resources, and select the weapon platform, weapon system, and tactics to employ the weapon. This is done using the Joint Munitions Effectiveness Manual (JMEM) Air-to-Surface Weaponeering Software, or JAWS. JAWS allows strike planners to adjust different parameters to achieve a desired probability of damage, or PD. The JAWS PD, expressed as a number between zero and one, provides the targeteer with a rough *proxy measure* of how a strike is likely to damage a particular target. Against softer targets, a PD of 0.3 may be enough to result in moderate damage, while hardened targets are usually weaponeered to a much higher PD, such as 0.6 to 0.7. The targeteer combines the JAWS PD with prior knowledge about the target to form a rough assessment of the most likely damage level.

Another information source available to the targeteers before the strike is the weather forecast. With the advent of high technology and precision guided munitions, weather can significantly affect strike performance. Advanced technology allows the aircraft to fly through bad weather to reach the target, but the high cost of a laser-guided bomb may preclude its use if the target is obscured by clouds (Zwenger, 1999). Weather is also a factor in the use of unguided or "dumb" bombs, but to a lesser extent. The only weapons not affected by weather are those guided by the Global Positioning System, or GPS. Such weapons, including Tomahawk missiles and the Joint Direct Attack Munition (JDAM), rely on satellite transmissions to guide them to specific coordinates on the earth's surface, and so are relatively impervious to weather (Zwenger, 1999; Clancy, 1995). The targeteers can combine knowledge of the guidance systems the strike package will employ with the weather forecast for the target area to adjust their pre-strike damage assessment level.

Once the strike occurs, the targeteer can obtain additional information prior to receiving still imagery analysis. For certain weapons platforms, the targeteer can access the aircrew's MISREP. This report is filed within thirty minutes of the aircraft's return to base following the strike (Killefer, 1999). Similarly, certain weapons or platforms include the capacity to videotape the strike from weapon release to impact. Such videotape can be a valuable tool for BDA (Smith, 1993). Although targeteers consider weapon system videotape a less reliable information source for BDA than still imagery, such tapes are available for analysis immediately upon the aircraft's return to base (Killefer, 1999). While awaiting still imagery, the targeteers use this post-strike information to update their assessment of the damage inflicted on the target.

2.2.3 Influence Diagram of BDA Decision.

Influence diagrams provide a simple, graphical representation of a decision (Clemen, 1995). Influence diagrams are especially useful for representing decisions made under uncertainty. In an influence diagram, rectangles represent decisions and ovals represent chance events. A rounded rectangle represents a mathematical computation or constant value. In an influence diagram, these three symbols are nodes. Arrows, or arcs, are used to join the different nodes, and can signify either sequence or relevance. Arcs leading to chance nodes or computation nodes show relevance. The predecessor, or source of the arc, influences the outcome of the chance or value node. For example, an arc from a decision node to a chance node means the outcome of the chance node depends on the course chosen in the decision node. On the other hand, arcs leading into a decision node indicate only sequence. An arc from a chance node into a decision node implies only that the state of the chance node is known at the time the decision is made. The influence diagram may reflect the sequence of events reading from the left of the diagram to the right. Nodes resolved early in the timeline appear on the left of the diagram. Although the format of influence diagrams is simple, they are a powerful decision-modeling tool.

Figure 2 is an influence diagram model of the Phase I BDA decision process incorporating information elicited from the subject matter experts, the targeteers. Imagery, available at the end of Phase I, is "perfect" information. The various sources of information previously discussed appear in the influence diagram as chance or decision nodes as appropriate. The final outcome of the targeteer's assessment is the accuracy, meaning the difference between the targeteer's assessment of the damage level and the actual damage level reflected in the imagery. Because the damage is assessed at one of five levels, the maximum possible error is a four-level difference. Additionally, the targeteer's assessment is either cautious or aggressive. For example, if the targeteer made a pre-imagery assessment of Light Damage and subsequent imagery showed the damage to be Severe, the targeteer's initial assessment is cautious by two levels. On the other hand, a pre-imagery assessment of Moderate Damage is aggressive by one level if subsequent imagery suggests Light Damage. By incorporating the accuracy of the targeteer's assessment into the model, we can use the influence diagram to produce a decision policy guideline.

The targeteer makes a pre-strike assessment based on knowledge of the strike package, the weaponeering, and the weather forecast. As Figure 2 depicts, the true state of weather over the target, and the extent of the weather's effects, are resolved after the targeteer's initial assessment. Weapons guided by GPS are influenced by the number of satellites in view over the target and by any enemy jamming capability, neither of which the targeteer can predict. Consequently, these nodes influence the actual damage inflicted on the target, but not the targeteer's assessment. Imagery is affected by the actual target damage. Since imagery is treated as perfect information, the targeteer can calculate the accuracy of the pre-strike assessment by comparing it against the imagery report.

The post-strike assessment, depicted in Figure 3, is virtually identical to the pre-strike assessment, but incorporates the additional information available to the targeteer, specifically the aircrew's MISREP and the weapon system videotape. With these additional pieces of information, the targeteer forms a post-strike assessment, which is also compared to imagery received later to determine accuracy.

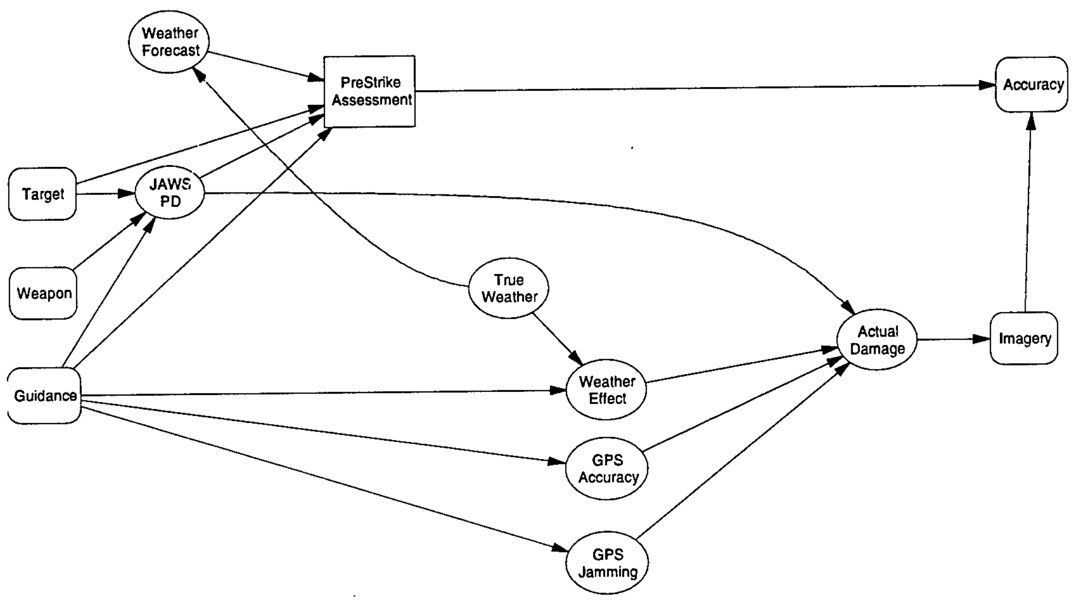

Figure 2. Phase I BDA Influence Diagram (Pre-Strike)

2.2.4 Bayesian Belief Network Model of BDA Decision.

Bayesian belief networks are an increasingly popular tool for decision modeling. Recent advances in theory and computer capability have reduced many obstacles to the use of Bayesian belief networks, also known as belief networks or Bayesian nets (Jensen, 1996). A Bayesian belief network is an efficient way to encode the joint probability distribution for a set of variables using an easily understood graphical format (Heckerman, 1995). Although Bayesian belief networks resemble influence diagrams in appearance, they do not convey the same information.

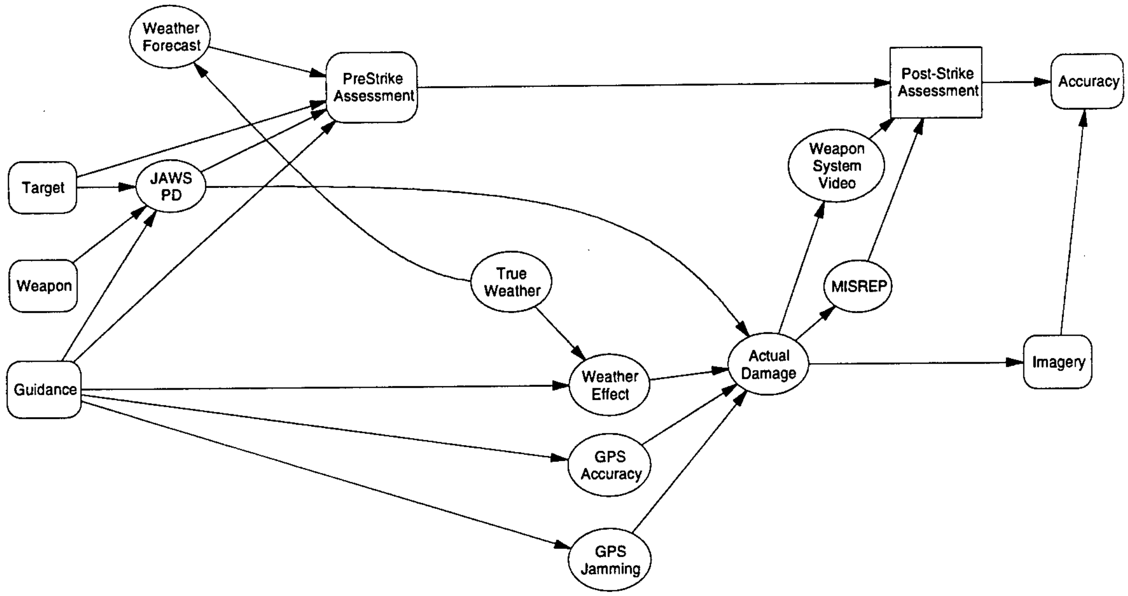

Figure 3. Phase I BDA Influence Diagram (Post-Strike)

Bayesian networks encode the joint probability distribution for a set of variables in a graphical format. Each variable in the decision of interest is represented as a node, as in an influence diagram, and arcs depicting dependencies connect nodes. Each node consists of a finite number of mutually exclusive and collectively exhaustive states and fully contains the conditional probability distributions for each of those states (Jensen, 1996). Thus, nodes with dependence arcs leading into them will reflect the conditional distribution of that variable. The overall network, including all nodes and conditional distributions, therefore encodes the joint probability distribution for all the variables in the problem of interest.

The keys to understanding Bayesian networks are the concepts of conditional probability distributions and probability calculus. Most people intuitively understand the way one event can affect the probable outcome of another event. Mathematically, this is referred to as the conditional probability. If the outcome of event B is influenced by Event A, then we speak of the probability of B given A, denoted $p(B|A)$. Probability calculus employs two key axioms to help calculate conditional probabilities. The first is the definition of conditional probability: $p(A|B) = p(A \text{ and } B) \times p(B)$. When A and B are mutually exclusive, $p(A \text{ and } B) = p(A) + p(B)$. This leads directly to Bayes' Rule: $p(B|A) = p(A|B)\, p(A) / p(B)$, which allows us to reverse the conditional probability. These conditional probability calculations are simple in theory, but can quickly grow to be intractable in even moderately complex problems (Jensen, 1996). The Bayesian belief network provides a compact method to encode the different variables in a problem of interest, as well as dependencies between variables and the inherent conditional probabilities.

The characteristic that makes Bayesian networks such a powerful tool is the capability to perform inference given evidence. In other words, Bayesian networks provide the capacity to update the probability of any node of interest when specific knowledge is available as to the state of a particular variable. This enables the user to enter evidence into the network and query the network for the probability distribution of any of the nodes in the network. Inferences are performed through repeated applications of Bayes' Rule. Thus, by taking some initial amount of evidence in conjunction with the previously assessed, or prior, conditional distributions, the user can obtain a posterior probability distribution for the variable of interest. This posterior distribution can be used to make probability statements as credibility intervals, as a point estimate, or to predict future data (Ramoni and Sebastiani, 1998b).

Figures 4 and 5 are the pre-strike and post-strike Bayesian belief networks, respectively, corresponding to the influence diagrams of Figures 2 and 3. We consider the pre-strike model and the post-strike model separately to focus on improving the accuracy of each assessment once perfect information becomes available.

Both Bayesian network models closely resemble their influence diagram counterparts, but the differences merit discussion. Some nodes of the influence diagram are collapsed into nodes in the Bayesian belief network model without affecting the accuracy of the model. The Weapon and Target decision nodes are collapsed into the JAWS PD node because the JMEM Air-to-Surface model takes the weapon, target type, and weapon guidance into account. We also remove the influence from the Guidance node to the JAWS PD node.

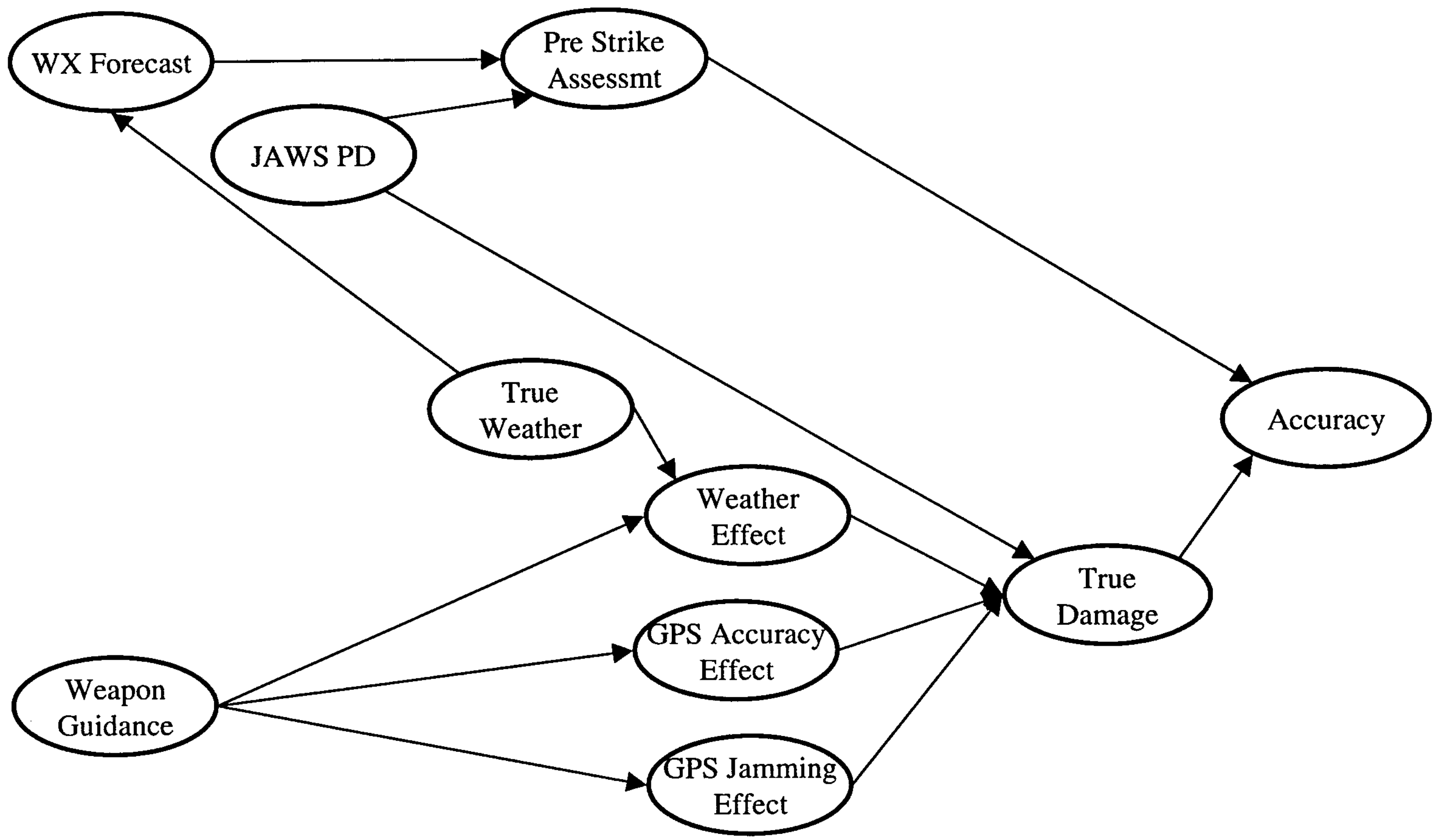

Figure 4. Phase I BDA Bayesian Belief Network (Pre-Strike)

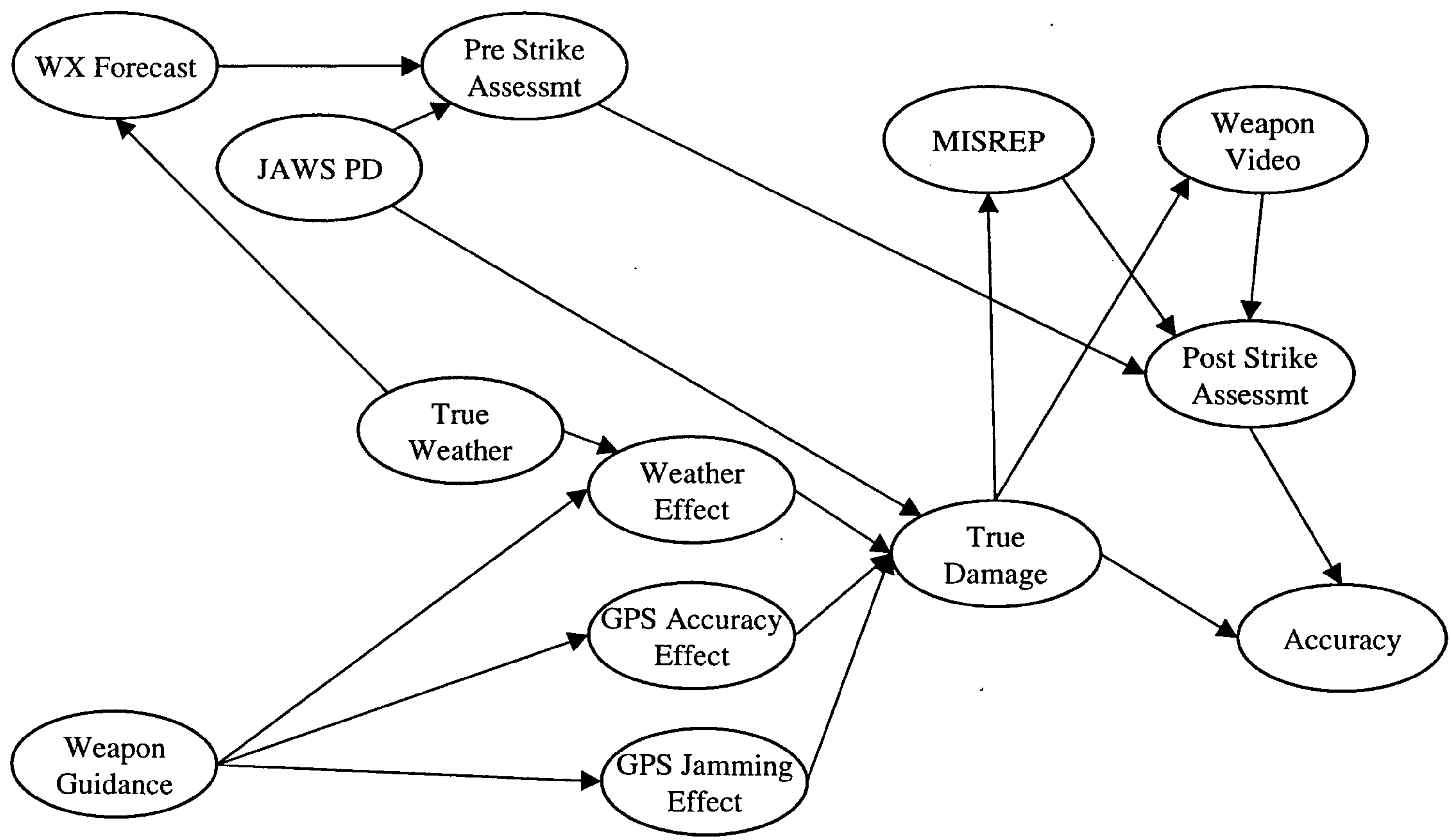

Figure 5. Phase I BDA Bayesian Belief Network (Post-Strike)

For a useful Bayesian belief network, we must clearly define each node and determine all its possible states (Heckerman and Wellman, 1995). Additionally, each node requires a conditional probability distribution for each possible state. This initial model is based largely on the expert knowledge of U.S. Air Force targeteers. Elicitation of this information was conducted using standard multi-objective decision analysis (MODA) techniques, as outlined in Clemen (1995). We elicited the factors most important to the targeteers in making damage assessments, as well as the factors' possible states. For the probability distributions, the reference lottery technique (Clemen, 1995) was used to determine the expert's subjective assessment of different probabilities with two notable exceptions as described below. Tables 1 and 2 show the different nodes and possible states in the pre-strike and post-strike Bayesian network models, respectively. Appendix A contains the complete conditional probability distribution information for both models.

Table 1. Nodes and Possible States (Pre-Strike Network)

Node Name	Possible States						
Weather Forecast	Socked In	Clouds	Clear				
JAWS PD	Light	Moderate	Severe				
Guidance	GPS	Laser	Unguided				
Pre-Strike Assessment	Light	Moderate	Severe				
True Weather	Socked In	Cloudy	Clear				
Wx Effect	No Effect	Clear	Mixed	Obscured			
GPS Accuracy	No Effect	Poor	Nominal	Good			
GPS Jamming	No Effect	Yes	No				
True Damage	No Damage	Light	Moderate	Severe	Destroyed		
Pre-Strike Acc	Cautious_Three	Cautious_Two	Cautious_One	Exact	Aggress_One	Aggress_Two	Aggress_Three

Table 2. Nodes and Possible States (Post-Strike Network)

Node Name	Possible States								
Weather Forecast	Socked In	Clouds	Clear						
JAWS PD	Light	Moderate	Severe						
Guidance	GPS	Laser	Unguided						
Pre-Strike Assessment	Light	Moderate	Severe						
True Weather	Socked In	Cloudy	Clear						
Wx Effect	No Effect	Clear	Mixed	Obscured					
GPS Accuracy	No Effect	Poor	Nominal	Good					
GPS Jamming	No Effect	Yes	No						
True Damage	No Damage	Light	Moderate	Severe	Destroyed				
MISREP	No Damage	Light	Moderate	Severe	Destroyed				
Weapon System Video	No Damage	Light	Moderate	Severe	Destroyed				
Post-Strike Assessment	No Damage	Light	Moderate	Severe	Destroyed				
Post-Strike Acc	Cautious_Four	Cautious_Three	Cautious_Two	Cautious_One	Exact	Aggress_One	Aggress_Two	Aggress_Three	Aggress_Four

The first exception was determining the conditional probabilities to assign to damage assessments. In the pre-strike model, two nodes influence the pre-strike damage assessment—the JAWS PD and the weather forecast. Elicitation revealed the targeteers view the JAWS PD as falling into one of three categories: Light, Moderate, or Severe Damage. In other words, although the JAWS software application provides a numerical PD between zero and one, the targeteers treat this as a proxy measure and mentally translate this number into a likely damage level given the target type. Similarly, although the weather (and by extension, the weather forecast) can take on a wide range of potential values, targeteers view weather forecasts as falling into one of three levels: Clear, Cloudy, and Socked In. Finally, although there are five potential damage levels, targeteers will not make pre-strike assessments of No Damage or Destroyed. This limits the range of possible damage levels for the Pre-Strike Assessment node to Light, Moderate, and Severe Damage. Combine these three states with the three states possible in each of the two influencing nodes and the Pre-Strike Assessment node requires 27 conditional probabilities.

As depicted in Figure 5, the Post-Strike Assessment node is influenced by the Pre-Strike Assessment node along with the MISREP and Weapon System Video nodes. While the Pre-Strike Assessment node has three possible states, the other three nodes contain all five damage levels. Therefore, the Post-Strike Assessment node encodes 375 conditional probabilities.

Eliciting the information for the Pre-Strike and Post-Strike Assessment nodes, while feasible, was not practical due to real-world circumstances. Consequently, we developed an alternative method to approximate these initial conditional probabilities. These approximations, while not perfect, make sense given the information elicited from the experts.

2.2.5 Derivation of Initial Conditional Distributions.

Through elicitation sessions, we determined the targeteers' views on the implications of their assessments. As stated earlier, the targeteers greatly prefer to err on the side of caution. They would rather underestimate than overestimate the damage expectancy. This stems from a desire to avoid aircrew losses from threats assessed as knocked out. The targeteers felt that overestimating the damage to a target is twice as bad as underestimating damage. Further, they felt that the penalty for incorrect assessments grows quadratically as the error increases. In other words, the penalty for incorrect assessments should grow as the square of the error. This penalty function implies an assessment that is off by two damage levels receives four times the penalty of an assessment that is off by one. These two factors yield a penalty function for incorrect assessments which depends on whether the assessment is cautious or aggressive:

$$Penalty(A,I) = \begin{cases} 2\cdot(A-I)^2 & \text{if } (A\text{-}I) > 0 \\ (A-I)^2 & \text{otherwise} \end{cases} \qquad (1)$$

where A is the assessed damage level on a scale from 1 to 5 and I is the damage level indicated by imagery. The penalty function reflects the higher penalty for aggressive assessments. Exact assessments receive a penalty of zero.

Using decision analysis software to enumerate all potential outcomes, we obtained an expected penalty score for each possible assessment in all possible situations. For any given situation, the damage level with the lowest expected penalty score is the assessment the targeteer is most likely to make. We assume an indirect relationship between probability and penalty score to derive an initial estimate of the conditional probability distribution.

As an example, consider a notional case in which the expected penalty scores for the five possible damage levels are 20, 6, 2, 18, and 54. Calculating the ratio of each penalty score to the minimum score, we get ratios of 10, 3, 1, 9, and 27. Because the lowest expected penalty option is the most attractive and likely choice, it should have the highest probability. The ratio scores are re-scaled so larger is better, yielding 5, 50/3, 50, 50/9, and 50/27, and normalized, yielding pseudo-probabilities (or proxy measures) of 0.063, 0.211, 0.632, 0.070, and 0.023, respectively. Table 3 illustrates this process.

By calculating and presenting the pseudo-probabilities in this format, we accomplish several goals. First, we avoid making the blanket statement that in a given situation, a targeteer will always assess a certain damage level. Second, and more importantly, presenting the data in this fashion conveys a sense of relative preference among the potential damage assessment levels. If the probability mass function for the conditional probabilities reflects a large portion of the probability assigned to one damage level, this conveys that the decision is relatively clear-cut. The targeteer can choose the damage level with the highest pseudo-probability and feel confident that it is the correct assessment. However, if two or more damage levels have approximately equal likelihoods of assessment, this indicates a more difficult decision situation. Such a situation usually occurs in cases where there are conflicting pieces of evidence. In this situation, a targeteer may wish to seek further information before making an assessment, or assess a range of values rather than a specific damage level. Tables 4 and 5 show the conditional pseudo-probability distributions for the Pre-Strike Assessment and Post-Strike Assessment nodes, respectively.

Table 3. Conversion of Penalty Scores to Pseudo-Probabilities

Assessment Option	Penalty Score	Ratio (Score / Min Score)	Sum / Ratio	Pseudo-Probability
No Damage	20	10	50 / 10 = 5.00	5 / 79.08 = 0.063
Light Damage	6	3	50 / 3 = 16.67	16.67 / 79.08 = 0.211
Moderate Dmg	2	1	50 / 1 = 50.00	50 / 79.08 = 0.632
Severe Damage	18	9	50 / 9 = 5.56	5.56 / 79.08 = 0.070
Destroyed	54	27	50 / 27 = 1.85	1.85 / 79.08 = 0.023
Sum	---	50	79.08	1

Because the above calculations are based on a penalty function that may vary from one targeteer to the next, the resulting pseudo-probability distributions are not unique—another reason we have described them as likelihood indicators or pseudo-probabilities. However, we can insert these pseudo-probabilities into the Bayesian belief network model as an initial guess. By gathering data, and using that data to update the conditional probability distributions, we get convergence to accurate conditional distributions. When sufficient data is available, the pseudo-probabilities will converge to the true probabilities and the two distributions will be identical. Until sufficient data is available, the likelihood indicators serve as a proxy measure used by the targeteer to help assess a particular damage level.

Table 4. Pre-Strike Assessment Conditional Distributions Before Data Learning

JAWS PD Category	Weather Forecast	Pseudo-Probability of Assessed Damage Level		
		Light	Moderate	Severe
Light	Socked In	0.773	0.165	0.062
	Cloudy	0.773	0.169	0.058
	Clear	0.752	0.187	0.061
Moderate	Socked In	0.678	0.231	0.091
	Cloudy	0.528	0.349	0.124
	Clear	0.384	0.467	0.149
Severe	Socked In	0.531	0.321	0.149
	Cloudy	0.341	0.418	0.241
	Clear	0.228	0.436	0.336

Table 5a. Post-Strike Assessment Conditional Distributions Before Data Learning, Light Pre-Strike Assessment

PreStrike	Video	MISREP	Pseudo-Probability of Assessed Damage Level				
			No Damage	Light	Moderate	Severe	Destroyed
Light	No Damage	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Light	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.097	0.843	0.043	0.012	0.005
		Moderate	0.068	0.886	0.033	0.009	0.004
		Severe	0.052	0.912	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
Light	Moderate	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.079	0.326	0.494	0.076	0.025
		Severe	0.032	0.126	0.775	0.052	0.015
		Destroyed	0.013	0.052	0.902	0.026	0.007
Light	Severe	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.099	0.424	0.381	0.070	0.025
		Severe	0.033	0.076	0.294	0.516	0.081
		Destroyed	0.015	0.033	0.114	0.775	0.063
Light	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.013	0.052	0.902	0.026	0.007
		Severe	0.033	0.073	0.252	0.544	0.098
		Destroyed	0.023	0.044	0.107	0.449	0.376

Table 5b. Post-Strike Assessment Conditional Distributions Before Data Learning, Moderate Pre-Strike Assessment

PreStrike	Video	MISREP	Pseudo-Probability of Assessed Damage Level				
			No Damage	Light	Moderate	Severe	Destroyed
Moderate	No Damage	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Moderate	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.117	0.814	0.050	0.014	0.006
		Moderate	0.076	0.874	0.036	0.009	0.004
		Severe	0.053	0.911	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
Moderate	Moderate	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.032	0.124	0.778	0.052	0.014
		Severe	0.017	0.066	0.877	0.032	0.008
		Destroyed	0.013	0.052	0.902	0.026	0.007
Moderate	Severe	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.043	0.169	0.708	0.063	0.018
		Severe	0.029	0.067	0.262	0.565	0.077
		Destroyed	0.018	0.039	0.128	0.723	0.092
Moderate	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.014	0.052	0.901	0.026	0.007
		Severe	0.034	0.070	0.206	0.553	0.137
		Destroyed	0.016	0.029	0.067	0.274	0.614

Table 5c. Post-Strike Assessment Conditional Distributions Before Data Learning, Severe Pre-Strike Assessment

			Pseudo-Probability of Assessed Damage Level				
PreStrike	Video	MISREP	No Damage	Light	Moderate	Severe	Destroyed
Severe	No Damage	No Damage	0.986	0.010	0.002	0.001	0.001
		Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Severe	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.214	0.683	0.071	0.021	0.010
		Moderate	0.118	0.811	0.050	0.014	0.006
		Severe	0.053	0.911	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
Severe	Moderate	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.032	0.124	0.778	0.052	0.014
		Severe	0.017	0.066	0.877	0.032	0.008
		Destroyed	0.013	0.052	0.902	0.026	0.007
Severe	Severe	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.043	0.169	0.708	0.063	0.018
		Severe	0.013	0.029	0.111	0.793	0.054
		Destroyed	0.020	0.041	0.119	0.681	0.140
Severe	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.013	0.052	0.902	0.026	0.007
		Severe	0.025	0.048	0.125	0.551	0.251
		Destroyed	0.009	0.017	0.039	0.157	0.778

2.3 Distribution Learning Capability

One of the most powerful capabilities inherent in a Bayesian belief network is its capability to evolve probability distributions and even network structure through data (Heckerman, 1995; Ramoni and Sebastiani, 1998b). After assuming a particular distribution for the prior probabilities, incorporation of data facilitates updating of the distribution parameters resulting in different probabilities for each of the possible states of a node. We describe a notional airstrike scenario used to generate realistic data. The scenario data is then employed through the data learning capability to support the elicited expert knowledge, and improve the accuracy of the Phase I BDA model.

2.3.1 Background on Distribution Learning.

Bayesian belief networks are powerful modeling tools not only for their ability to provide inference based on available evidence, but also for their inherent ability to learn from data. Once we have constructed a Bayesian model of a given problem, we can use data about the problem to improve the accuracy of the conditional probability distributions within the model. This allows the initial model, with conditional probabilities based on subjective opinion, to adapt and learn from data gained through repeated trials. In this way, we can model the human approach to probabilistic thinking. As an example, consider the case of a coin toss where we wish the outcome to be heads. Initially, we may assume the coin is fair and assign the subjective probability of heads to be 0.5. However, if we then conduct 20 trials in which every toss comes up tails, we will likely update our probability based on what we have learned through these trials. In a similar manner, we can use data gained through repeated trials to update the conditional probability distributions in a Bayesian belief network. To relate to the BDA problem, we want to incorporate target intelligence into the process to improve the targeteers' assessment accuracy.

To understand the data learning capability, we must examine the nature of the Bayesian network more closely. The conditional probability we wish to update is the probability of an outcome occurring given prior information and the available data. The true probability of the outcome is uncertain. However, we can choose a probability distribution for the possible values of the true probability, and use the data to update the parameters of that distribution. We continue with the coin-tossing example suggested by Heckerman (Heckerman, 1995).

We have assigned an initial probability distribution, known as the prior distribution, for the possible states. In other words, we have assessed that each possible state *i* occurs with probability p(x_i) where $i = 1$ to r and r possible states exist. In the coin-toss example, only two states are possible, heads and tails, and the probability of tails is simply (1-p(heads)). We have subjectively determined a prior distribution for the probability of heads and, by extension, tails, based on our knowledge of the situation. However, this assigned prior distribution may be different from the true probability distribution, and the data reflects that true probability distribution. What we seek to do is update the prior distribution so that the various probabilities p(x_i) more closely model the probabilities exhibited in the data.

We begin by choosing a known distribution for the initial probabilities. The binomial nature of the coin toss leads us to select the beta distribution in our example. By choosing a beta distribution to model the prior, we ensure that the posterior distribution is also a beta distribution. This is because the beta distribution is a conjugate family of distributions for binomial sampling, such as we have in this case (Heckerman, 1995). Thus, the number of heads and tails in the data translate to the two parameters of the beta distribution. In cases with more than two possible states, data are sampled from a multinomial distribution, and so other distributions may be more appropriate (Heckerman, 1995). However, this example serves to illustrate the working of the distribution learning mechanism. For more detailed discussions of distribution learning, the reader is referred to Heckerman (Heckerman, 1995) or Ramoni and Sebastiani (Ramoni and Sebastiani, 1998b).

The shape of the beta distribution depends on two parameters and we choose the values of these parameters in order to match our prior distribution and confidence in our knowledge. Then, applying Bayes' rule and using the data available, we can update the parameters of the beta distribution to form the new, or posterior, distribution. The posterior distribution is still a beta distribution, but a different beta distribution. The probability of each possible state $p(x_i)$ can then be determined using the posterior distribution.

For the coin-tossing example, we initially assume the coin is fair and assign a probability of 0.5 to heads. However, since we do not have a great deal of knowledge about the coin, we equate this prior knowledge with only six trials, a relatively small number. We adjust the parameters of the beta distribution so that the first parameter reflects the imagined number of trials yielding heads, and the second parameter reflects the number of imagined trials yielding tails (Heckerman, 1995). This results in a beta (3,3) distribution, shown in Figure 6, for the possible true probability of heads. Clearly, the expected value of the probability of heads is 0.5, which accurately reflects our belief that the coin is fair. However, this distribution exhibits a wide variance, reflecting our relative uncertainty about the possible true probability of heads.

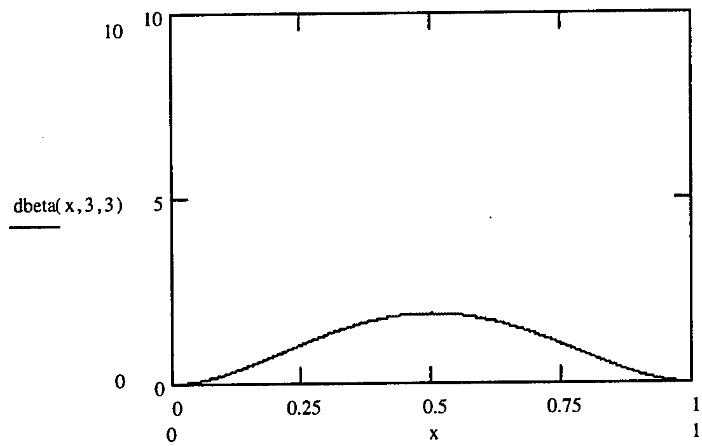

Figure 6. Prior Distribution—Beta (3,3)

Now we conduct the trials and obtain the data to update the probability distribution. As outlined in Heckerman (1995), the parameters of the updated probability distribution will simply be the imagined number of heads (or tails) plus the number seen in the trials. Therefore, if we conduct twenty trials and obtain only five heads, the posterior distribution becomes beta (8,18), as shown in Figure 7. This distribution exhibits smaller variance, because we now have data in addition to our prior knowledge, and also reflects the decreasing likelihood that the coin is fair. This is shown by the expected value of the distribution, which serves as our best estimate of the true probability of heads and is plainly less than 0.5 (Heckerman, 1995).

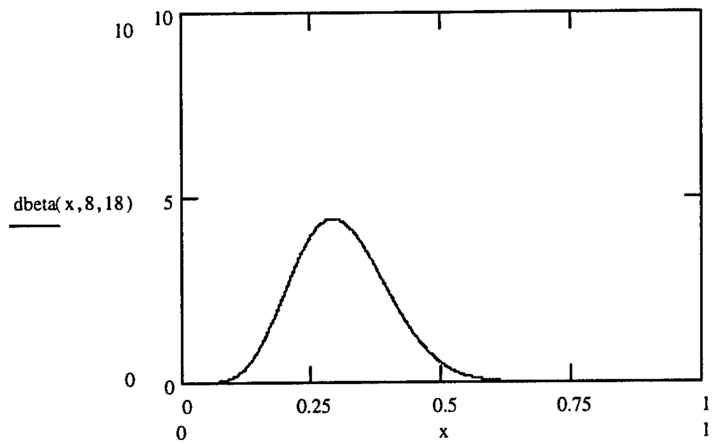

Figure 7. Posterior Distribution—Beta (8,18)

Suppose we received no heads at all in twenty trials, as hypothesized. The posterior distribution in this case would be beta (3,23), shown in Figure 8. This extreme example would lead to a significant drop in our estimate of the probability of heads. Additionally, the large proportion of trials yielding the same outcome greatly decreases the variance of the posterior distribution.

As a final illustration, let us now suppose the coin truly is fair, and in twenty trials we observe ten heads and ten tails. In this case the posterior distribution is beta (13,13), shown in Figure 9. The expected value of this distribution, our posterior probability of heads, is 0.5, exactly what we started with. However, the variance of the posterior distribution is significantly smaller than that of the prior distribution. Clearly, the observed data has reduced the uncertainty in our estimate of the true probability of heads, although the estimate itself has not changed.

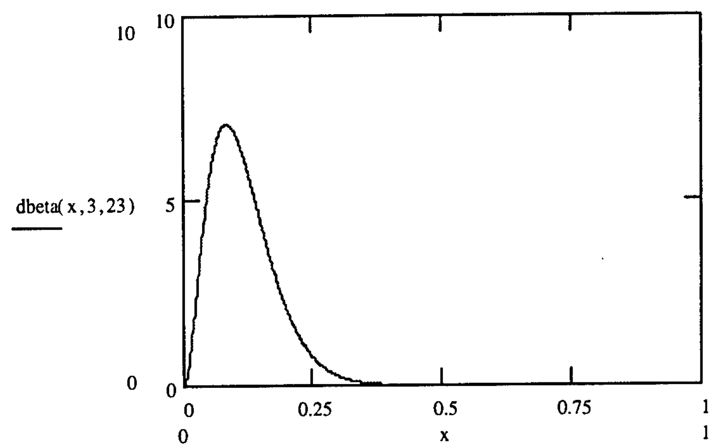

Figure 8. Posterior Distribution—Beta (3,23)

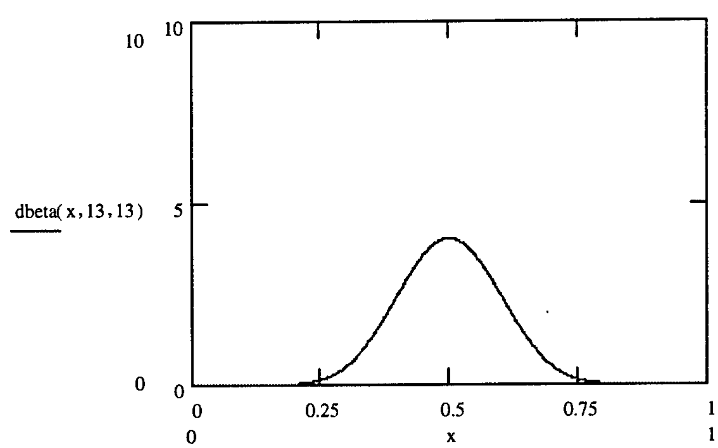

Figure 9. Posterior Distribution—Beta (13,13)

This simple example illustrates how Bayesian networks can update conditional distributions based on data. While we have used the beta distribution in this example, other distributions can be used to implement data learning and may be more appropriate for multinomial sampling (Heckerman, 1995). We use this updating capability to combine data with elicited expert knowledge to increase model accuracy.

2.3.2 BDA Scenario and Data.

To illustrate the distribution learning capability of the decision tool, we demonstrate its use with a notional attack scenario that is both realistic and representative of the type of BDA decisions targeteers routinely face. This scenario is not as complex as real-world situations but sufficient for demonstration purposes.

The scenario developed is a U.S. raid on an enemy airfield and was selected for several reasons. Airfields are a common target. Because airfields are large, complex systems, they offer a large variety of targets and target types. A typical airfield can contain as many as one hundred separate point targets, referred to as Desired Mean Points of Impact or DMPIs (Killefer, 1999). Airfield targets include both point targets, such as buildings, and area targets, such as runways or parking aprons. Further, airfields are usually very high priority targets in any air campaign. One of the initial goals of any air campaign is to gain and maintain air superiority and knocking out enemy airfields is essential to achieving that goal. Targeteers are routinely faced with assessing battle damage inflicted on enemy airfields.

Clancy Field is a notional enemy airfield located in a fictional country, developed with the help of several Air Force targeteers (Curry, 1999; Killefer, 1999; Zwenger, 1999). The targeteers provided information about the target facilities as well as the strike packages through a series of elicitation and consultation interviews. During these interviews, we sought to develop a notional yet realistic scenario that would provide enough sufficiently useful data to evaluate the utility of the Bayesian network model.

Clancy Field is a small airfield located in an enemy country, used as a backup facility for aircraft operations. Figure 10 shows a diagram of Clancy Field's layout, not to scale. The field has two runways, configured in parallel, as well as an aircraft parking apron. Precision approach radar beacons and other navigational equipment are located at either end of the runways. The field contains several hardened aircraft shelters for protecting aircraft while not in operation. The field has a control tower for monitoring base flight operations, a two-bay aircraft maintenance hangar, plus aircrew barracks and dining facilities and a few other base infrastructure buildings. As a military airfield, Clancy Field boasts significant facilities for ammunition and weapons storage and loading, as well as facilities for petroleum, oil, and lubricant (POL) storage and pumping. A few mobile surface-to-air missile batteries are set up around the airfield for protection and they change location periodically to make any attack more difficult. In addition, a hardened shelter exists near the control tower to keep essential personnel safe in case of an enemy attack. Finally, the enemy country has built a hardened underground command post deep under the airfield to maintain command and control even in the event of an attack on the airfield. This command bunker is considered a very hard target, constructed of many layers of concrete and buried deep below the surface.

2.3.3 Striking the Airfield.

Striking Clancy Field is a complicated endeavor. The command objectives of gaining and maintaining air superiority require doing as much damage to the airfield and its associated systems and subsystems as possible. The targeteers have set the existing aircraft and runways as their highest priority targets. Next on the priority list are the maintenance facility and the aircrew housing. Aircrews have been directed to strike the POL or ammunition storage or loading facilities if they feel a weak link exists. For example, if the POL storage tanks are widely distributed, but only a few trucks are available to load the fuel into aircraft, those trucks are excellent targets for an air strike. The precision navigation aids at the ends of the runways are also useful targets, depending on the type of aircraft located at the field and the local climate. If the weather is frequently inclement, the navigational aids play a larger role in aircraft operations from the airfield. In this case, the navigational equipment does not play an overly significant role in sustained air operations. The mobile SAM batteries providing air defense for the field are also important targets, as they affect the ability to strike the airfield.

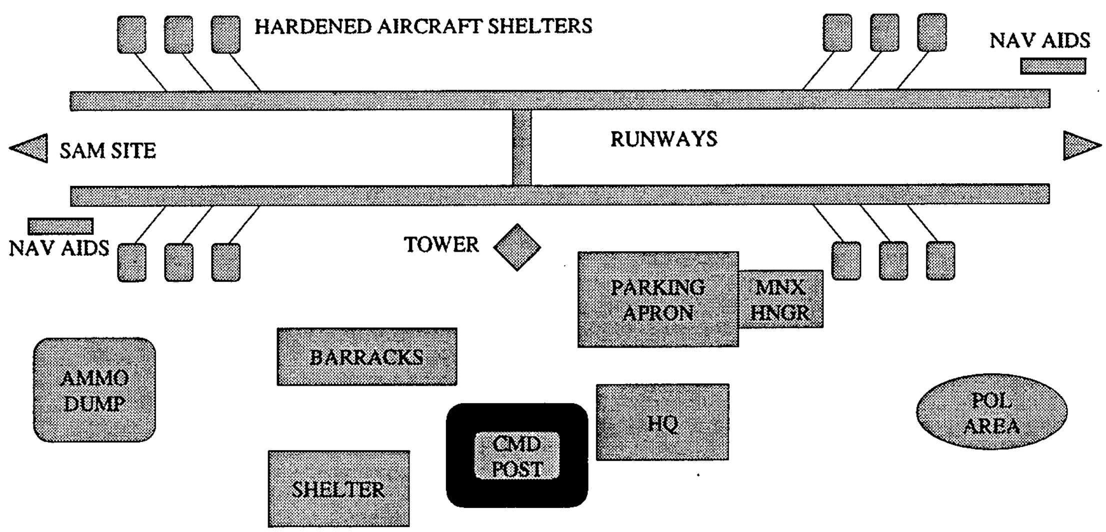

Figure 10. Clancy Field Layout

Targeteers have a variety of options available in planning a strike against an airfield. They can choose any of a number of different aircraft, each with a different specialty in the combat arena. Additionally, they can choose from an even wider array of potential weapons from simple gravity-guided dumb bombs to laser-guided, rocketassisted munitions. Planners may even choose to employ unmanned weapon systems such as the TLAM or Conventional Air-Launched Cruise Missile (CALCM). Finally, once they have chosen the appropriate weapons and platforms during the Weaponeering phase of the targeting cycle, the planners build the strike package by determining the TOTs for each of the different weapons.

For this particular strike, the planners have selected a variety of weapons and platforms. The strike begins with TLAMs launched against the SAM locations protecting the airfield. Removing the SAM threat improves subsequent strike aircraft effectiveness. The Tomahawks are closely followed by F-16CJ aircraft performing a suppression of enemy air defense (SEAD) mission. These aircraft use electronic weaponry to jam the SAM sites, and high-speed anti-radiation missiles (HARMs) to destroy any SAM sites activating their radar systems. The two weapons should combine to effectively negate the threat from SAMs to the other aircraft in the strike. The F-16CJ aircraft carry unguided bombs to crater the runways of the airfield once the SAM threat is suppressed.

The next phase of the strike begins about ten minutes after the TLAMs have impacted their targets. F-15E Strike Eagles using laser-guided GBU-24 bombs attack the point targets located around the airfield. Six Strike Eagles, configured in two-ship flights with one-minute separation, attack the targets. The Strike Eagles' targets include the POL and ammunition facilities, the maintenance hangar, the hardened personnel shelter, and the precision navigation systems.

The final phase of the strike involves two B-1B Lancer aircraft using the Joint Direct Attack Munition (JDAM) against the hardened aircraft shelters located on the airfield. The JDAM is a Mark-84 general purpose bomb coupled with an inertial guidance system and a GPS receiver (Clancy, 1995). The B-1Bs fly over the airfield and drop the JDAMs through the roof of each aircraft shelter, destroying everything inside and eliminating the airfield's capacity for aircraft storage. The B-1s reach the target approximately 90 seconds after the last F-15E completes its raid. The entire strike lasts less than fifteen minutes.

Conducting BDA on such a strike is a complicated endeavor, much like the strike itself, although the targeteer has numerous pieces of information available before the first TLAM is launched. From the Air Tasking Order, or ATO, the targeteer knows the location and description of each target. Further, the targeteer knows which weapons platforms are assigned to the various targets and what weapons they will employ against those targets. Most importantly, the targeteer knows the PD obtained during weaponeering from JAWS. As described earlier, this number provides a rough estimate of the planned damage level. The final piece of information available in the pre-strike time frame is the weather forecast for the target area. For this scenario, the weather is forecast as clear.

After the strike is conducted, some of the aircrews provide mission reports (MISREPs) and weapon system video. These additional pieces of evidence factor strongly in the targeteer's assessment of the most likely damage level and can also be inserted into the Bayesian belief network of the post-strike decision. The data for this scenario are shown in Tables 6 and 7. Table 6 presents the data for the weapon systems that do not provide a meaningful mission report or a video capability, while Table 7 contains the data for those platforms that provide MISREP and video. This additional data is useful in the post-strike decision model, but the pre-strike model has no capacity to use the extra information.

2.3.4 *Illustration of Distribution Learning.*

Using the data shown in Tables 6 and 7, we can demonstrate the capacity of the Bayesian network model to evolve the conditional probability distributions from data. The learning capacity is implemented using the Bayesian Knowledge Discoverer (BKD) software in conjunction with the distribution updating scheme outlined earlier (Ramoni and Sebastiani, 1998a). Within the BKD software, this is done by simply selecting the appropriate database and letting the software quantify the network.

We highlight two important points about the use of data to update the Bayesian belief network. First, the data must specify one of the possible states for all the nodes in the belief network. Any gaps in the data where a value is not specified will result in an error and the network will not update properly. Second, conditional distributions will only change for those cases for which data exists. This implies that some of the conditional distributions will probably never update at all, because they represent situations that are highly unlikely to occur. As an example, consider a case in which the aircrew's MISREP states Destroyed, but the weapon system video shows No Damage. Although our influence-diagram-based decision model allowed us to compute pseudo-probabilities for damage assessments in this case, such a situation is highly improbable in the real world. Consequently, data for such a situation is unlikely to arise and the conditional distribution will probably never update. A second consequence of the same fact is that in order to see the results of data from the notional scenario described above, we need only examine those cases for which we have data. For the pre-strike assessment model, those data points are all easily located in the conditional distribution of the Pre-Strike Assessment node. This distribution is shown as Table 8. The original conditional distributions (Table 4) are included for comparison. A bold border highlights those conditional probabilities changing through data learning.

Table 6. Pre-Strike Scenario Data

Guidance	Wx_Forecast	Jaws_Pd	True_Weather	Prestrike_Assmt	Wx_Effect	Gps_Accuracy	Gps_Jamming	True_Damage	Prestrike_Acc
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Destroyed	Cautious_Two
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	NoDamage	Aggress_Two
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact
Unguided	Clear	Moderate	Clear	Moderate	Clear	NoEffect	NoEffect	NoDamage	Aggress_Two
Unguided	Clear	Moderate	Clear	Moderate	Clear	NoEffect	NoEffect	Light	Aggress_One
Unguided	Clear	Moderate	Clear	Moderate	Clear	NoEffect	NoEffect	NoDamage	Aggress_Two
Unguided	Clear	Moderate	Clear	Moderate	Clear	NoEffect	NoEffect	Moderate	Exact
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Nodamage	Aggress_Two
Laser	Clear	Severe	Clear	Severe	Clear	NoEffect	NoEffect	Destroyed	Cautious_One
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Severe	Cautious_Two
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Cautious_One
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Cautious_One
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Severe	Cautious_One
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Destroyed	Cautious_Two
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Severe	Cautious_One
Laser	Clear	Light	Clear	Light	Clear	NoEffect	NoEffect	Destroyed	Cautious_Three
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Destroyed	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Severe	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Moderate	Clear	Moderate	NoEffect	Nominal	No	Severe	Cautious_One
GPS	Clear	Moderate	Clear	Moderate	NoEffect	Nominal	No	Moderate	Exact

Table 7. Post-Strike Scenario Data

Guidance	Wx_Forecast	Jaws_Pd	True_Weather	Prestrike_Assmt	Wx_Effect	Gps_Accuracy	Gps_Jamming	True_Damage	Misrep	Video	Poststrike_Assmt	Poststrike_Acc
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Nodamage	Nodamage	Nodamage	Nodamage	Exact
Laser	Clear	Severe	Clear	Severe	Clear	NoEffect	NoEffect	Destroyed	Destroyed	Destroyed	Severe	Cautious_One
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Severe	Destroyed	Destroyed	Severe	Exact
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Moderate	Severe	Moderate	Exact
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Severe	Moderate	Moderate	Exact
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Severe	Severe	Severe	Severe	Exact
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Destroyed	Destroyed	Destroyed	Destroyed	Exact
Laser	Clear	Severe	Clear	Moderate	Clear	NoEffect	NoEffect	Severe	Destroyed	Destroyed	Destroyed	Aggress_One
Laser	Clear	Light	Clear	Light	Clear	NoEffect	NoEffect	Destroyed	Destroyed	Destroyed	Severe	Cautious_One
Laser	Clear	Moderate	Clear	Light	Clear	NoEffect	NoEffect	Moderate	Moderate	Moderate	Moderate	Exact

Table 8. Pre-Strike Assessment Conditional Distributions Before and After Data Learning

Original conditional distributions

JAWS PD	Forecast	Pseudo-Probability of Assessed Damage Level: Light	Moderate	Severe
Light	Socked In	0.773	0.165	0.062
	Cloudy	0.773	0.169	0.058
	Clear	0.752	0.187	0.061
Moderate	Socked In	0.678	0.231	0.091
	Cloudy	0.528	0.349	0.124
	Clear	0.384	0.467	0.149
Severe	Socked In	0.531	0.321	0.149
	Cloudy	0.341	0.418	0.241
	Clear	0.228	0.436	0.336

After data learning

JAWS PD	Forecast	Pseudo-Probability of Assessed Damage Level: Light	Moderate	Severe
Light	Socked In	0.773	0.165	0.062
	Cloudy	0.773	0.169	0.058
	Clear	0.772	0.172	0.056
Moderate	Socked In	0.678	0.231	0.091
	Cloudy	0.527	0.349	0.124
	Clear	0.392	0.530	0.078
Severe	Socked In	0.530	0.321	0.149
	Cloudy	0.341	0.418	0.241
	Clear	0.079	0.774	0.147

Table 8 shows the effects of distribution learning although, as stated above, learning only occurs for cases where data exists. For example, in the case where the weaponeering indicates severe damage and the weather forecast is clear, the pseudo-probability of assessing moderate damage increases from 0.436 to 0.774. This shows how the model's relatively uncertain recommendation of assessing moderate damage is reinforced by the targeteers' application of expert knowledge. The same effect can be seen in the other two cases for which data exists, although not as strongly.

The effects of data learning are more evident in the case of the post-strike assessment node. The post-strike assessment node conditional distributions after the implementation of data learning are shown in Table 9. Since data is available for only a small number of the possible cases, only a few of the conditional distributions change. Entries demonstrating a change are highlighted using a larger font size and a bold border.

Table 9a. Post-Strike Assessment Conditional Distributions After Data Learning, Light Pre-Strike Assessment

			Pseudo-Probability of Assessed Damage Level				
PreStrike	Video	MISREP	No Damage	Light	Moderate	Severe	Destroyed
Light	No Damage	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Light	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.097	0.843	0.043	0.012	0.005
		Moderate	0.068	0.886	0.033	0.009	0.004
		Severe	0.052	0.912	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
		No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
Light	**Moderate**	**Moderate**	**0.045**	**0.186**	**0.711**	**0.043**	**0.014**
Light	**Moderate**	**Severe**	**0.018**	**0.072**	**0.871**	**0.030**	**0.009**
		Destroyed	0.013	0.052	0.902	0.026	0.007
		No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
Light	**Severe**	**Moderate**	**0.057**	**0.243**	**0.647**	**0.040**	**0.014**
		Severe	0.033	0.076	0.294	0.516	0.081
		Destroyed	0.015	0.033	0.114	0.775	0.063
Light	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.013	0.052	0.902	0.026	0.007
		Severe	0.033	0.073	0.252	0.544	0.098
Light	**Destroyed**	**Destroyed**	**0.009**	**0.018**	**0.043**	**0.780**	**0.151**

Table 9b. Post-Strike Assessment Conditional Distributions After Data Learning, Moderate Pre-Strike Assessment

PreStrike	Video	MISREP	Pseudo-Probability of Assessed Damage Level: No Damage	Light	Moderate	Severe	Destroyed
Moderate	**No Damage**	**No Damage**	**0.990**	**0.006**	**0.002**	**0.001**	**0.001**
Moderate	No Damage	Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Moderate	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.117	0.814	0.050	0.014	0.006
		Moderate	0.076	0.874	0.036	0.009	0.004
		Severe	0.053	0.911	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
Moderate	Moderate	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.032	0.124	0.778	0.052	0.014
		Severe	0.017	0.066	0.877	0.032	0.008
		Destroyed	0.013	0.052	0.902	0.026	0.007
Moderate	Severe	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.043	0.169	0.708	0.063	0.018
Moderate	**Severe**	**Severe**	**0.017**	**0.038**	**0.150**	**0.751**	**0.044**
		Destroyed	0.018	0.039	0.128	0.723	0.092
Moderate	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.014	0.052	0.901	0.026	0.007
		Severe	0.034	0.070	0.206	0.553	0.137
Moderate	**Destroyed**	**Destroyed**	**0.006**	**0.012**	**0.027**	**0.110**	**0.846**

Table 9c. Post-Strike Assessment Conditional Distributions After Data Learning, Severe Pre-Strike Assessment

PreStrike	Video	MISREP	Pseudo-Probability of Assessed Damage Level				
			No Damage	Light	Moderate	Severe	Destroyed
Severe	No Damage	No Damage	0.986	0.010	0.002	0.001	0.001
		Light	0.986	0.010	0.003	0.001	0.001
		Moderate	0.986	0.010	0.003	0.001	0.001
		Severe	0.156	0.293	0.312	0.161	0.078
		Destroyed	0.156	0.293	0.312	0.161	0.078
Severe	Light	No Damage	0.986	0.010	0.003	0.001	0.001
		Light	0.214	0.683	0.071	0.021	0.010
		Moderate	0.118	0.811	0.050	0.014	0.006
		Severe	0.053	0.911	0.026	0.007	0.003
		Destroyed	0.156	0.293	0.312	0.161	0.078
Severe	Moderate	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.032	0.124	0.778	0.052	0.014
		Severe	0.017	0.066	0.877	0.032	0.008
		Destroyed	0.013	0.052	0.902	0.026	0.007
Severe	Severe	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.053	0.911	0.026	0.007	0.003
		Moderate	0.043	0.169	0.708	0.063	0.018
		Severe	0.013	0.029	0.111	0.793	0.054
		Destroyed	0.020	0.041	0.119	0.681	0.140
Severe	Destroyed	No Damage	0.156	0.293	0.312	0.161	0.078
		Light	0.156	0.293	0.312	0.161	0.078
		Moderate	0.013	0.052	0.902	0.026	0.007
		Severe	0.025	0.048	0.125	0.551	0.251
Severe	Destroyed	Destroyed	0.005	0.010	0.022	0.518	0.445

As in the pre-strike assessment node, the trend is to reinforce the initial likelihood indicators derived from the targeteers' expert knowledge. In all cases, the assessment level with the highest initial likelihood number exhibits an increase. The other damage levels' likelihood indicators all see a decrease. In this way, the data serves to reduce the uncertainty in the damage assessment process. As more and more data is gathered, the assessment decision becomes increasingly clear-cut.

The distribution learning capability offers numerous advantages to the user and benefits to the decision model. First, as more and more data is incorporated into the model, the conditional probability distributions begin modeling the true probabilities with increasing accuracy. The subjective probabilities on which the initial model was built become less important in the face of mounting real-world trials. Secondly, the distribution learning capability allows targeteers to make the most of the very expensive data generated through airstrikes. Currently, such information is not used in any statistical analysis although many resources exist for storing and maintaining information from airstrikes. This learning capability might also be used to identify trends or problem areas depending on how the Bayesian models are implemented. For example, targeteers in a theater of operations may notice that their local area data shows different trends from the data incorporated into the model from different areas. This could drive the local targeteers to re-examine weapons or tactics to understand the disparity. Perhaps the largest and most obvious benefit of the data learning capability, though, is facilitation of more efficient use of imagery resources. As highlighted above, the data learning capability reduces the uncertainty in the damage assessment decision. With enough data in certain situations, targeteers may eventually feel comfortable in assessing damage levels even without imagery. In turn, this would free up imagery resources to apply to other sites whose assessments are more difficult. There is also the benefit of reducing the time required to provide a damage assessment, since imagery may not be required.

2.4 Applications to Existing Systems

The Bayesian network decision model examined here could easily be incorporated into the Air Force's next generation of battle management tools. The Air Force is in the process of switching over to the Theater Battle Management Core System, or TBMCS. TBMCS is already set to include much of the information the Bayesian decision model would require as part of the Combat Intelligence System, or CIS. The CIS is designed to make rapid use of intelligence data and automate its processing as much as possible (Frey, 1998). TBMCS also contains a Targeting and Weaponeering Module (TWM) designed to replace the numerous database- or spreadsheet-based applications targeteers have developed over the years to track BDA information and assessments (Killefer, 1999). To incorporate the Bayesian model would thus require only adding a Bayesian belief net application into TBMCS and providing a link to the Military Intelligence Data Base, or MIDB. Targeteers could then use the TWM in conjunction with the Bayesian decision model to view information about a target, make decisions, and record damage assessments. This data would then be stored in the MIDB. Periodically, based either on passage of time or entry of data, the Bayesian network model could update the conditional distributions internally, as described above. In this way, the targeteers could reap the benefits of data analysis listed earlier.

Another potential area of application for the Bayesian network methodology is in the Phase II and Phase III BDA decisions. These decisions, much like Phase I BDA, require combining information from a number of sources with expert opinion and judgment, and Bayesian belief networks are extremely useful for modeling such problems. The Air Force is already attempting to move away from BDA based primarily on imagery. In January 1999, the Air Force won approval to conduct an advanced concept technology demonstration (ACTD) of BDA in the Joint Targeting Toolbox (JTT). BDA for JTT is designed to yield software sufficient to give theater commanders the capability for automated BDA. Further, these damage assessments will be based on interrelationships of targets and their vulnerabilities and strengths rather than observed physical damage. Also, some functional damage may be inferred rather than observed. As pointed out throughout this paper, a Bayesian belief network is particularly well suited to performing inference based on evidence and to learning from data. These characteristics imply Bayesian belief networks show promise as a potential methodology for the BDA for JTT ACTD effort.

2.5 Summary and Conclusions

We have closely examined the BDA process in general and the Phase I BDA decision process in particular. We presented two models of Phase I BDA decisions. One model is based on influence diagrams and the other implements a Bayesian belief network. We have demonstrated the usefulness of the Bayesian belief network as a decision model. Further, we highlighted the powerful capability to evolve conditional probability distributions from data which allows a Bayesian belief network to grow increasingly accurate. Through a representative scenario, we generated notional data which was used to demonstrate the learning capability. We then pointed out numerous benefits of employing this learning capability through the Bayesian belief network model as well as ways in which the Bayesian model could serve as a substitute for a decision matrix. Finally, we pointed out how the Bayesian decision model could be incorporated into two technological efforts the Air Force is undertaking.

Although Bayesian belief networks have only recently come into vogue as decision-modeling tools, they are very powerful and easy to use. We have shown in this research the applicability of Bayesian belief networks to a complex, real-world problem, the Phase I BDA decision. Further, we have shown how the distribution learning capability allows the Bayesian belief network to reduce uncertainty and model true probabilities with increasing accuracy. The Bayesian belief model offers several improvements over the current BDA process. First, the Bayesian model makes use of the "perfect" information the current process discards. Second, the Bayesian model allows for the use of data to reduce uncertainty about the BDA decision and identify trends. Finally, the Bayesian model offers the potential to eventually make BDA decisions without waiting for the "perfect" information of still imagery. The Air Force has already announced its intention to conduct research toward this end, and the Bayesian belief network methodology has the potential to serve as the foundation of that research (Hebert, 1999). Clearly, Bayesian belief networks are a powerful tool that can greatly advance understanding of BDA and make the decision process faster and more accurate.

CHAPTER 3

While the Bayesian belief network model is a valuable tool in its own right, future work could increase that value by improving and extending the model. The single largest area for improvement is the knowledge base on which the model is based. Due to real world constraints, access to targeteers was limited during the time frame of this research. Consequently, the decision model is based on the expert knowledge of only two targeteers. Interviewing additional targeteers would serve to validate the existing model, and might highlight other variables or probability distributions to fold into the model.

The probability distributions in the model are another area for improvement. As with all other parts of the model, the initial probability distributions for the JAWS PD, Guidance, and True Weather nodes were based on subjective information elicited from two targeteers. Because these nodes are independent, their distributions are not conditional. However, these nodes' probability distributions play a role in the calculation of penalty scores, which in turn affect the initial conditional distributions for the damage assessments. Real world historical data offers the potential to improve the accuracy of the probability distributions of these nodes. By mining past data, we can better approximate the proportion of strikes that are weaponeered to the various damage levels or the probability that a strike aircraft will employ a certain guidance system.

Similarly, the True Weather node currently contains a notional distribution for the possible weather over the target. While past data for the entire world will probably not prove useful in updating this distribution, data could still contribute to increasing its accuracy. By making the True Weather node dependent on a Region or Climate node, then gathering weather data for the various regions or climates of the world, we can increase the model's ability to accurately model the true weather.

One additional area for future study is determining the correct influence of expert knowledge. Given the Bayesian belief network's capability to learn from data, we must determine the correct balance to strike between prior expert knowledge and future data. In this research, we have set the weight of the elicited expert knowledge to highlight the model's data learning capability. However, as more targeteers are incorporated into the knowledge base, the respective importance of expert knowledge and data will require adjustment.

We applied the Bayesian belief network methodology to the Phase I BDA decision in this study. However, the Phase II and Phase III BDA decisions are very similar to Phase I. Currently, experts make these decisions by coupling their knowledge with information from a number of independent sources to reach a decision. Bayesian belief networks are excellent tools for modeling such expert decisions. Applying the Bayesian methodology to all phases of the BDA process could yield significant benefits in both speed and accuracy, especially if coupled with an automated data processing system such as TBMCS.

Appendix A: Probability Distributions

This appendix contains the complete probability distributions for all nodes of the Pre-Strike and Post-Strike Bayesian belief network before implementation of data learning. The distributions reflect the expert opinions of targeteers elicited through interviews, as detailed in Chapter 2. For example, the distributions of the MISREP and Video nodes reflect two assumptions. First, the weapon system video is more accurate than the aircrew's MISREP. Second, both these information sources will very rarely underestimate the actual damage to the target, although they may overestimate the damage. The other nodes reflect the targeteers' opinions in similar fashion.

Cases that are not possible do not appear in these tables in order to reduce complexity. For example, the True Damage node does not list probabilities for cases where both weather and the GPS environment (jamming and accuracy) have an effect. Due to the nature of the weapon guidance, only one of the two factors will have an effect.

Independent Nodes

Node True Weather

States	Socked In	Clouds	Clear
Probabilities	0.333	0.333	0.334

Node JAWS PD

States	Light	Moderate	Severe
Probabilities	0.2	0.4	0.4

Node Guidance

States	GPS	Laser	Unguided
Probabilities	0.4	0.4	0.2

Dependent Nodes

Node WX Forecast

Parents	*States and Probabilities*		
True Weather	Socked In	Cloudy	Clear
Socked In	0.8	0.1	0
Clouds	0.2	0.8	0.2
Clear	0	0.1	0.8

Node WX Effect

Parents		*States and Probabilities*			
True Weather	Guidance	No Effect	Clear	Mixed	Blocked
Socked In	GPS	1	0	0	0
Socked In	Laser	0	0	0	1
Socked In	Unguided	0	0	0.7	0.3
Clouds	GPS	1	0	0	0
Clouds	Laser	0	0.1	0.8	0.1
Clouds	Unguided	0	0.25	0.7	0.05
Clear	GPS	1	0	0	0
Clear	Laser	0	1	0	0
Clear	Unguided	0	1	0	0

Node GPS Accuracy

Parent	*States and Probabilities*			
Guidance	Poor	Nominal	Good	No Effect
GPS	0.333	0.333	0.333	0
Laser	0	0	0	1
Unguided	0	0	0	1

Node GPS Jamming

Parent	*States and Probabilities*		
Guidance	Yes	No	No Effect
GPS	0.02	0.98	0
Laser	0	0	1
Unguided	0	0	1

Node Pre-Strike Assessmt

Parents		*States and Probabilities*		
JAWS PD	WX Forecast	Light	Moderate	Severe
Light	Socked In	0.773	0.165	0.062
Light	Cloudy	0.773	0.169	0.058
Light	Clear	0.752	0.187	0.061
Moderate	Socked In	0.678	0.231	0.091
Moderate	Cloudy	0.528	0.349	0.124
Moderate	Clear	0.384	0.467	0.149
Severe	Socked In	0.531	0.321	0.149
Severe	Cloudy	0.341	0.418	0.241
Severe	Clear	0.228	0.436	0.336

Node	True Damage							
Parents				States and Probabilities				
JAWS PD	WX Effect	GPS Accuracy	GPS Jamming	No Damage	Light	Moderate	Severe	Destroyed
Light	No Effect	Poor	Yes	0.7	0.25	0.05	0	0
Light	No Effect	Poor	No	0.25	0.68	0.05	0.02	0
Light	No Effect	Nominal	Yes	0.5	0.44	0.05	0.01	0
Light	No Effect	Nominal	No	0.18	0.68	0.08	0.05	0.01
Light	No Effect	Good	Yes	0.3	0.65	0.05	0	0
Light	No Effect	Good	No	0.15	0.7	0.1	0.04	0.01
Light	Clear	No Effect	No Effect	0.18	0.68	0.08	0.05	0.01
Light	Mixed	No Effect	No Effect	0.25	0.65	0.08	0.02	0
Light	Blocked	No Effect	No Effect	0.5	0.45	0.05	0	0
Moderate	No Effect	Poor	Yes	0.7	0.1	0.15	0.05	0
Moderate	No Effect	Poor	No	0.15	0.1	0.68	0.05	0.02
Moderate	No Effect	Nominal	Yes	0.5	0.44	0.05	0.01	0
Moderate	No Effect	Nominal	No	0.1	0.09	0.68	0.09	0.04
Moderate	No Effect	Good	Yes	0.3	0.05	0.6	0.05	0
Moderate	No Effect	Good	No	0.1	0.09	0.7	0.09	0.02
Moderate	Clear	No Effect	No Effect	0.1	0.09	0.68	0.09	0.04
Moderate	Mixed	No Effect	No Effect	0.25	0.09	0.6	0.05	0.01
Moderate	Blocked	No Effect	No Effect	0.5	0.2	0.28	0.02	0
Severe	No Effect	Poor	Yes	0.7	0.05	0.05	0.15	0.05
Severe	No Effect	Poor	No	0.15	0.03	0.07	0.68	0.07
Severe	No Effect	Nominal	Yes	0.5	0.08	0.12	0.28	0.02
Severe	No Effect	Nominal	No	0.1	0.04	0.09	0.68	0.09
Severe	No Effect	Good	Yes	0.3	0	0.05	0.6	0.05
Severe	No Effect	Good	No	0.1	0.03	0.07	0.7	0.1
Severe	Clear	No Effect	No Effect	0.1	0.04	0.09	0.68	0.09
Severe	Mixed	No Effect	No Effect	0.25	0.03	0.08	0.6	0.04
Severe	Blocked	No Effect	No Effect	0.5	0.08	0.15	0.25	0.02

Node	Prestrike Acc							
Parents		States and Probabilities						
True Damage	PreStrike Assessmt	Cautious Three	Cautious Two	Cautious One	Exact	Aggress One	Aggress Two	Aggress Three
No Damage	Light	0	0	0	0	1	0	0
No Damage	Moderate	0	0	0	0	0	1	0
No Damage	Severe	0	0	0	0	0	0	1
Light	Light	0	0	0	1	0	0	0
Light	Moderate	0	0	0	0	1	0	0
Light	Severe	0	0	0	0	0	1	0
Moderate	Light	0	0	1	0	0	0	0
Moderate	Moderate	0	0	0	1	0	0	0
Moderate	Severe	0	0	0	0	1	0	0
Severe	Light	0	1	0	0	0	0	0
Severe	Moderate	0	0	1	0	0	0	0
Severe	Severe	0	0	0	1	0	0	0
Destroyed	Light	1	0	0	0	0	0	0
Destroyed	Moderate	0	1	0	0	0	0	0
Destroyed	Severe	0	0	1	0	0	0	0

Node	MISREP				
Parent	States and Probabilities				
True Damage	No Damage	Light	Moderate	Severe	Destroyed
No Damage	0.75	0.2	0.05	0	0
Light	0	0.55	0.4	0.05	0
Moderate	0	0	0.55	0.4	0.05
Severe	0	0	0	0.6	0.4
Destroyed	0	0	0	0.1	0.9

Node	Video				
Parent	States and Probabilities				
True Damage	No Damage	Light	Moderate	Severe	Destroyed
No Damage	0.9	0.1	0	0	0
Light	0	0.8	0.15	0.05	0
Moderate	0	0	0.8	0.15	0.05
Severe	0	0	0	0.8	0.2
Destroyed	0	0	0	0.05	0.95

Node PostStrike Acc

Parents		*States and Probabilities*								
True Damage	PostStrike Assessmt	Cautious Four	Cautious Three	Cautious Two	Cautious One	Exact	Aggress One	Aggress Two	Aggress Three	Aggress Four
No Damage	No Damage	0	0	0	0	1	0	0	0	0
No Damage	Light	0	0	0	0	0	1	0	0	0
No Damage	Moderate	0	0	0	0	0	0	1	0	0
No Damage	Severe	0	0	0	0	0	0	0	1	0
No Damage	Destroyed	0	0	0	0	0	0	0	1	1
Light	No Damage	0	0	0	1	0	0	0	0	0
Light	Light	0	0	0	0	1	0	0	0	0
Light	Moderate	0	0	0	0	0	1	0	0	0
Light	Severe	0	0	0	0	0	0	1	0	0
Light	Destroyed	0	0	0	0	0	0	0	1	0
Moderate	No Damage	0	0	1	0	0	0	0	0	0
Moderate	Light	0	0	0	1	0	0	0	0	0
Moderate	Moderate	0	0	0	0	1	0	0	0	0
Moderate	Severe	0	0	0	0	0	1	0	0	0
Moderate	Destroyed	0	0	0	0	0	0	1	0	0
Severe	No Damage	0	1	0	0	0	0	0	0	0
Severe	Light	0	0	1	0	0	0	0	0	0
Severe	Moderate	0	0	0	1	0	0	0	0	0
Severe	Severe	0	0	0	0	1	0	0	0	0
Severe	Destroyed	0	0	0	0	0	1	0	0	0
Destroyed	No Damage	1	0	0	0	0	0	0	0	0
Destroyed	Light	0	1	0	0	0	0	0	0	0
Destroyed	Moderate	0	0	1	0	0	0	0	0	0
Destroyed	Severe	0	0	0	1	0	0	0	0	0
Destroyed	Destroyed	0	0	0	0	1	0	0	0	0

Appendix B: Analysis of Alternate Methodologies

Before selecting Bayesian belief networks as the methodology with which to implement the BDA decision model, we examined other potential approaches. Specifically, we investigated building the model using traditional multi-objective decision analysis (MODA) techniques and using an artificial neural network (ANN). Both methodologies could be applied to the BDA problem; however, each had a significant drawback.

MODA techniques rely on interviews with decision makers to identify key factors and their relative importance. Using weighting schemes and probabilistic analysis, the best alternative in a given decision situation is identified. This approach would take maximum advantage of the targeteers' accumulated expert knowledge. However, the weighting schemes employed in MODA are static, meaning that the decision model would have no capacity to incorporate data. Further, the model would have no means to evolve along with technology and tactics. A Bayesian belief net can incorporate both expert knowledge and historical data, and has the capability to evolve over time.

On the other hand, artificial neural nets make maximum use of data. Artificial neural nets rely solely on data, rather than a system of rules. The various sources of information to the problem act as inputs in a system based on the human nervous system. Different combinations of information result in different outcomes from the final node in the network. This methodology offers the ability to evolve the system over time, unlike the MODA approach, but has other clear faults. First, an ANN approach would discard the accumulated expert knowledge of targeteers who have performed BDA for years.

Second, an ANN requires significant amounts of data to build and train the network, plus an additional set of data to validate the model. Such data is not currently available and would require an extended period of heavy operations. The Bayesian belief network model can incorporate expert knowledge initially, then become more reliant on data over time. Additionally, the Bayesian model can be validated at any point without sacrificing valuable data. Validation is easily accomplished through a Turing test by presenting the model and a targeteer with the same set of inputs and comparing the resulting assessments.

Bibliography

Almond, Russell G. *Graphical Belief Modeling*. London: Chapman and Hall, 1995.

Berger, James O. *Statistical Decision Theory and Bayesian Analysis*, Second Edition. New York: Springer Verlag, 1985.

Buede, Dennis M., Joseph A. Tatman, and Terry A. Bresnick. "Introduction to Bayesian Networks: A Tutorial for the 66th MORS Symposium." World Wide Web, http://systeng.gmu.edu, 1998.

Clancy, Tom. *Fighter Wing: A Guided Tour of an Air Force Combat Wing*. New York: Berkley, 1995.

Clemen, Robert T. *Making Hard Decisions*. Belmont: Duxbury Press, 1995.

Cozman, Fabio G. "JavaBayes Version 0.34." Software and user's manual. World Wide Web, http://www.cs.cmu.edu/~javabayes. 13 Aug 98.

Curry, Hugh, Major, USAF. Chief of Targets, 609th Air Intelligence Squadron. Personal interview. 12 Jan 99.

Davis, Richard G. *Decisive Force: Strategic Bombing in the Gulf War*. Washington: GPO, 1996.

Defense Intelligence Agency (DIA). *Battle Damage Assessment (BDA) Quick Guide*. Washington: Defense Intelligence Agency, February 1996.

Department of Defense. *Conduct of the Persian Gulf War*. Washington: Department of Defense, unclassified version, April 1992.

Department of the Air Force, Directorate for Operational Requirements (AF/XOR). "USAF Battlelab Home Page." World Wide Web, http://www.hq.af.mil/xo/afbattlelab/index.htm. Sep 1998.

Department of the Air Force (DAF). *The Joint Targeting Process*. AFPAM 10-225, May 1996.

"Execution Planning/Combat Assessment." Class handout, Combat Targeting Course. 17th Training Wing, Goodfellow AFB TX, Aug 1996.

Frey, Steven, Captain, USAF. "Time Critical Targeting in Theater Battle Management Core System (TBMCS)." Briefing slides. ESC/ACE, Hanscom AFB MA, Nov 1998.

Green, Thomas K., Captain, USAF and David A. Roodhouse, Major, USAF. *An Analysis of Optimal Airfield Attack Parameters.* MS Thesis, AFIT/GST/OS/85M-4. School of Engineering, Air Force Institute of Technology (AU) Wright-Patterson AFB OH, March 1985 (AD-A155857).

Hallion, Richard P. *Storm Over Iraq.* Washington: Smithsonian Institution Press, 1992.

Hearst, Marti A. "Banter on Bayes: Debating the Usefulness of Bayesian Approaches to Solving Practical Problems." *IEEE Expert and Intelligent Systems*, 12: 18-25 (November/December 1997).

Hebert, Adam. "New Battle Damage Assessment ACTD Addresses BDA Shortcomings," *Inside The Air Force*, 22 January 1999: 6.

Heckerman, David, and Michael P. Wellman. "Bayesian Networks." *Communications of the ACM, 38*: 27-30 (March 1995).

Heckerman, David. *A Tutorial on Learning with Bayesian Networks.* Redmond: Microsoft Technical Report MSR-TR-95-06, 1995.

Jensen, Finn V. *An Introduction to Bayesian Networks.* New York: Springer Verlag, 1996.

Keeney, Ralph L. "Using Values in Operations Research." *Operations Research, 42*: 793-813 (September-October 1994).

Keeney, Ralph L. *Value-Focused Thinking: A Path to Creative Decisionmaking.* Cambridge MA: Harvard University Press, 1992.

Killefer, Robert, Captain, USAF. Chief of Targets, 608th Air Intelligence Squadron. Personal interviews. January 1999.

Lehnert, Wendy G. *The Process of Question Answering.* Hillsdale NJ: Lawrence Erlbaum Associates Inc., 1978.

Miglin, Robert N., Captain, USAF. *"AAPMOD," an Interactive Computer Model for Analysis of Conventional Weapons Effectiveness.* MS Thesis, AFIT/GST/OS/84M-13. School of Engineering, Air Force Institute of Technology (AU) Wright-Patterson AFB OH, March 1984.

Myers, Laura. "Mobile missiles difficult to hit," *The Dayton Daily News*, 1 January 1999, sec A:5.

Needler, Mark S., LCDR, USN. *Battle Damage Assessment Considerations for the Operational Commander*. Naval War College paper. Newport: Naval War College, 1993 (AD-B173022).

Pearl, Judea. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann Publishers Inc., 1988.

Ramoni, Marco, and Paola Sebastiani. "Bayesian Knowledge Discoverer Version 1.0." Software and online documentation. World Wide Web, http://kmi.open.ac.uk/projects/bkd, 1998a.

-----. *Bayesian Methods for Intelligent Data Analysis*. Milton Keynes: KMi Technical Report KMi-TR-67, 1998b.

-----. *Learning Probabilities from Incomplete Data: An Experimental Comparison*. Milton Keynes: KMi Technical Report KMi-TR-64, 1998c.

Rome Laboratory/IRRE. *Bomb Damage Assessment Study: Battle Damage Assessment Simulation (BDASIM)*. Rome NY: Rome Laboratory, February 1996.

Smith, Kevin W. *Cockpit Video: A Low Cost BDA Source*. Maxwell AFB AL: Air University Press, 1993.

Sweigart, Frank, CDR, USN. *Desert Storm: A Commander's Perspective on the Naval Air Campaign*. Naval War College paper. Newport: Naval War College, 1993 (AD-B174498).

Zwenger, Patrick, MSgt, USAF. NCOIC, Targets, 608th Air Intelligence Squadron. Personal interviews. January 1999.

Vita

Captain Daniel W. Franzen was born on 5 June 1970 in Los Angeles, California. He graduated from Westlake High School in Westlake, Ohio, in 1988. He entered undergraduate studies at the Virginia Military Institute and was named a distinguished graduate in 1992 with a Bachelor of Science degree in Mechanical Engineering. He was commissioned a Second Lieutenant, USAF, upon graduation.

After being named a distinguished graduate from Undergraduate Missile Training, his first assignment was at Malmstrom AFB as a missile combat crewmember. While at Malmstrom AFB, he earned a Master of Administrative Science degree through the University of Montana. In August 1997, he entered the graduate program in Operations Analysis at the Graduate School of Engineering, Air Force Institute of Technology.

Permanent Address: 36257 Falcon Crest Drive
Avon, OH 44011

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1999

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
A BAYESIAN DECISION MODEL FOR BATTLE DAMAGE ASSESSMENT

5. FUNDING NUMBERS

6. AUTHOR(S)
Daniel W. Franzen, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
2950 P Street
Wright-Patterson AFB, OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GOA/ENS/99M-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Major Mike Deaton
USAF Command & Control Battlelab (C2B)
238 Harston Street
Hurlburt Field, FL 32544-5200

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Battle damage assessment (BDA) is critical to success in any air campaign. However, Desert Storm highlighted numerous deficiencies in the BDA process, and operations since Desert Storm continue to point out weaknesses. We present a review of the Phase I BDA decision, or physical damage assessment, and model the decision process using a Bayesian belief network. Through subject matter expert (i.e., the targeteers) elicitation sessions, imagery was found to be critically important to the BDA process yet this information is generally not retained. This use of "perfect information" is delineated in the BDA process models. We propose a methodology based on Bayesian belief networks for incorporating this perfect information. We demonstrate the Bayesian belief network's capability to update conditional probability distributions using data generated in real world operations. This capability allows the network's conditional distributions to evolve, increasing model accuracy and reducing uncertainty in the decision.

14. SUBJECT TERMS
Bayes Theorem, Bayesian belief networks, damage assessment, decision aids, decision analysis

15. NUMBER OF PAGES
75

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL