2014

Imperfect Situation Analysis: Representing the Role of Error and Uncertainty in Modeling, Simulation and Analysis

Victor Eaton Middleton
Wright State University

IMPERFECT SITUATION ANALYSIS:

REPRESENTING THE ROLE OF ERROR AND UNCERTAINTY IN MODELING, SIMULATION & ANALYSIS

A dissertation submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

By

VICTOR E. MIDDLETON

B.S., Michigan Technological University, 1973
M.S., Michigan Technological University, 1975
M.S., Michigan State University, 1978

2014
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

May 2, 2014

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY SUPERVISION BY Victor E. Middleton ENTITLED Imperfect Situation Analysis: Representing Error and Uncertainty in Modeling, Simulation & Analysis BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Doctor of Philosophy.

____

Frank W. Ciarallo, Ph.D.,
Dissertation Director

____

Ramana Grandhi, Ph.D., Director of the Ph.D. in Engineering Program

____

Robert E. W. Fyffe, Ph.D., Vice President for Research and Dean of the Graduate School

____

Committee on Final Examination

Frank W. Ciarallo, Ph.D.

____

Raymond R. Hill, Ph.D.

____

Yan Liu, Ph.D.

____

Mateen M. Rizki, Ph.D.

____

Mary E. Fendley, Ph.D.

____

David Hudak, Ph.D.

ABSTRACT

Middleton, Victor Eaton. Ph.D., Engineering Ph.D. Program, Wright State University, 2014. Imperfect Situation Analysis: Representing Error and Uncertainty in Modeling, Simulation & Analysis.

Much of traditional modeling, simulation and analysis (MS&A) is supported by engineering models - deterministic, Newtonian physics-based representations of closed systems. Such approaches are not well-suited to represent the intricacies of human behavior. This research advocates and seeks to articulate the concept of a more human-centric approach to MS& A, one that better represents decision-making and other cognitive aspects of human behavior as well as it does physical activity.

It starts with a view of individuals and groups as complex adaptive systems, which are best represented using agent-based modeling. Representation of human behavior through intelligent agents incorporates models of decision-making, knowledge engineering and knowledge representation, as well as the whole gamut of the psychological and physiological interactions of humans with each other and their environment. This representation is exemplified by consideration of situation awareness/situation understanding (SA/SU) as a core element.

This leads to the development of a proof-of-concept simulation of a specific, easily understood, and quantifiable example of human behavior: intelligent agents being spatially "lost" while trying to navigate in a simulation world. This model is named MOdeling Being Intelligent and Lost (MOBIL), noting the ability to be in both of these states is central to the simulation. MOBIL uses a blend of object oriented software principles with agent based modeling to establish the utility of applying the human-

centric approach to analysis.

Applying that simulation in a number of virtual experiments illustrates how it supports investigation into an individual's SA/SU and associated decision-making processes.

Acknowledgments

This dissertation is the result of long years of study and learning, both in academia and in my professional life. Throughout that time I have benefited from the advice, counsel, and friendship of mentors, colleagues, and family. I owe to all of them a debt I cannot possibly ever pay.

I begin with my two foremost academic mentors. My first advisor was Gene Ortner at Michigan Tech, who started me on the path of scholarship and research. My Ph.D. advisor, Frank Ciarallo was truly the perfect fit for me at this time in my life. I would never have come close to completing this dissertation without his encouragement, patience, and perseverance. Most importantly, meeting with him to discuss this research was always a joy, his enthusiasm and willingness to put up with my tendency to prolixity never flagged. His only fault is a sad lack of judgment with respect to NHL allegiance; a fault much overshadowed by his many virtues.

I owe special debts to a number of people for whom or with whom I have worked over the years. Ellen van Son of the TNO and George Mastroianni of the Air Force Academy have contributed significantly to the ideas contained in this work. John D'Errico, Chris Christenson, and Arthur (Dub) Garrett, three distinguished soldiers and equally distinguished military analysts, have taught me much and have been a source of inspiration both professionally and personally. My first boss at the University of Dayton Research Institute, Nick Engler, taught me the ins and outs of contract research, but never lost sight of the fact that doing research should be fun, an attitude I hope to never lose. Matt Herz of the Natick Soldier Center embodied the same spirit. He was both a friend and an advocate for many years; I still miss him. Craig Porter and Bob McIntyre kept me fed for many years through their entrepreneurial skills, and Craig's generosity helped pay for much of my Ph.D work.

I apologize for not singling out the many others with whom I have worked, especially all the folks at UDRI, Jaycor, STI, ORSA Corp, the Natick Modeling and Analysis Team, TRAC-Monterey and the Naval Post Graduate School, I am none the less deeply appreciative of your friendship and fellowship through the years.

I must thank the members of my committee: Ray Hill, whose careful and critical review of this document greatly improved it; Yan Liu and Matt Rizki whose courses helped bridge the 30 year gap in my formal education; Dave Hudak who has been a kindred soul from the day I first met him; and Mary Fendley, who boldly (foolishly?) stepped in at the last moment when I needed another faculty member physically present for my defense. All of the staff at the graduate school, especially Dr. Grandhi and Alysoun Taylor-Hall, greatly facilitated my return to academia after a long hiatus. Many thanks to all at Wright State for your professionalism and help.

Finally, of course, I must acknowledge the continuing and essential support of my family. My mother and father instilled a life-long love of learning, one I believe my wife and I have successfully passed on to our children. Certainly Jennifer, James, and Emily at least learned how to exhort me to “Study hard and make them proud!” I am certainly intensely proud of all three of them. My wife Maris has been the foundation for any success I have had in life; her love and support has been and continues to be the true reward for everything I do.

Table of Contents

Table of Figures

Table of Tables

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

Then took the other, as just as fair,
And having perhaps the better claim,
Because it was grassy and wanted wear;
Though as for that the passing there
Had worn them really about the same,

And both that morning equally lay
In leaves no step had trodden black.
Oh, I kept the first for another day!
Yet knowing how way leads on to way,
I doubted if I should ever come back.

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.
Robert Frost

1 INTRODUCTION

This dissertation represents ideas and aspirations gleaned from more than a quarter of a century in the practice of modeling, simulation and analysis (MS&A) in a number of different fields, but primarily in support of military research and development (R&D). It is in pursuit of a long-held desire to more fully address a critical element in the representation of human behavior in modeling and analysis, and especially with regard to representation of dismounted infantry.

Over the past 25 years, my work as an operations research analyst has increasingly led me to take an approach to modeling and analysis of human behavior that emphasizes human information processing and decision-making.

Fundamental to this approach is the belief that it is no longer sufficient to concentrate principally on the physics-based aspects of military operations, to no longer represent the individual soldier as just a “slow, unarmored tank” (Middleton 2010b).

While the role of physics remains important, it needs to be augmented by a more “human-centric” view that distinguishes between the man and machine, one that explicitly represents the role of human performance moderators on operational capability, e.g.,(Pew and Mavor 1998; Ritter 2000; Mastroianni and Middleton 2001; Hudlicka 2002; Hudlicka 2004; Silverman 2004; van Lent, McAlinden et al. 2004).

In today’s “information age”, there is a need to explicitly represent the role of human perception and decision-making in the physical behaviors that comprise both war and peace, e.g.,(Pew and Mavor 1998; Hill 1999; Deitz 2006; Warwick 2006; Middleton 2010a). This view is further emphasized by the current need to consider non-kinetic as well as kinetic aspects of military operations or other human activities, e.g., (Perumalla and Bhaduri 2006; Hurley, Bucher et al. 2009; Moffat 2011).

Later chapters of this dissertation explain the human centric approach more fully, and describe how it addresses shortcomings in many current models and simulations.

The dissertation consists of five major elements:

1. The articulation of the human-centric approach to modeling, simulation and analysis;
2. The adaption and integration of a multi-disciplinary mix of theoretical concepts into an over-arching conceptual model of intelligent agent decision making;
3. A methodological framework for selective integration of different aspects of these theories;
4. The implementation of that framework into a computer simulation that applies the framework to decision-making with respect to geo-spatial movement; and

5. A series of simulation experiments that demonstrate the viability of the framework as an applied research tool.

The research explores the issues involved in the development and application of that approach. As a proof-of-concept, the research led to the development and application of a representation of incomplete and/or erroneous Situation Awareness/Situation Understanding (SA/SU) in computer simulation. The goal is not to add materially to the theory of SA/SU or other similar concepts, but rather to apply engineering methods to improve representation of these concepts in models and simulations.

The perspective of this dissertation is very much that of Industrial and Systems Engineering (ISE), the application of engineering principles to improve processes and products through an integrated systems approach; one that melds mathematical, physical and social sciences with the engineering design to optimize complex processes. Widely available and used tools: Microsoft Excel; JMP; AnyLogic; and OpenStreetMap, have been combined to produce and apply a new analytical methodology.

This dissertation is directed to the development and application of MS&A tools in support of operations research with respect to human conduct, and it requires the integration of a number of diverse fields of study. Implementing a more human-centric approach to MS&A is an ambitious undertaking; this dissertation focuses on a small part of that undertaking, as shown in Figure 1-1.

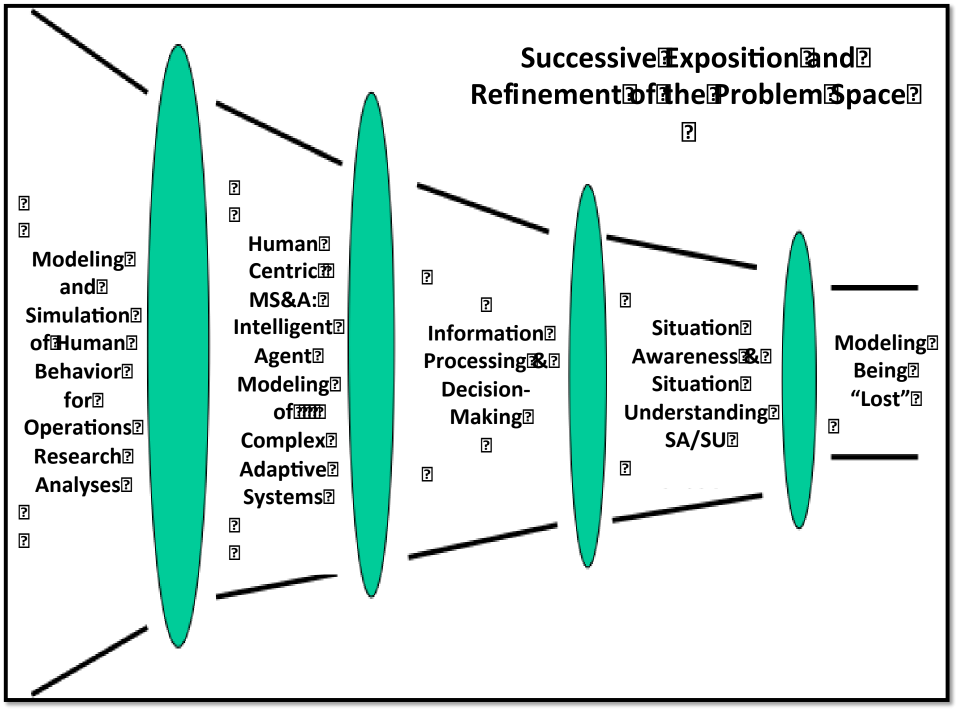


Figure 1-1 Refinement of the Problem Space

The starting point is viewing individuals and groups as complex adaptive systems, which are best represented using agent-based modeling. Representation of human behavior through intelligent agents incorporates models of decision-making, knowledge engineering and knowledge representation, as well as the whole gamut of the psychological and physiological interactions of humans with each other and their environment. This representation is exemplified by consideration of situation awareness/situation understanding (SA/SU) [1] as a core element. This leads to development of a proof-of-concept simulation. Applying that simulation in a number of virtual experiments illustrates how it supports analysis of individual's SA/SU and associated decision-making processes.

---

[1] Throughout most of this dissertation I choose to blur the distinctions between SA and SU into a single over-arching concept following the pragmatic definition of (Adam 1993) "knowing what is going on so I can figure out what to do." Chapter 4 below elaborates on both SA and SU.

1.1 Objectives

This dissertation begins by describing the human-centric approach to MS&A. This approach is essential to meeting the analytic needs of today's military community, but it has application far outside of that community. The human-centric paradigm has been evolving for some years; dating at least to the seminal work by the Military Operations Research Society mini-symposia in the early 90's, e.g.,(Murtaugh 1994), and the comprehensive overview by Pew and Mavor (Pew and Mavor 1998). One of the core objectives of this dissertation is to articulate my view of that paradigm, as developed over a career spanning virtually the entire period of its evolution.

This research develops and demonstrates a methodological framework for the representation of imperfect SA/SU and its effects on mission/task performance. This framework begins with a conceptual model following the dictates of the human-centric approach, one that supports multi-disciplinary investigation of different hypotheses with respect to the nature of, and inter-relationships between, the physical*,* information*,* and cognitive domains[2] of military operations in particular and of human endeavors in general. Heretofore, approaching such hypotheses through conventional models has focused on the performance of tasks, performance that can usually be measured in terms of time and resources. Modeling SA/SU will also allow testing hypotheses related to the *potential* to act, or react, based on changing operational conditions and the contents of the individual's knowledge state.

I instantiate this methodological framework in a simulation model with "plug and play" modules that can support exploration of different schemes for knowledge representation, inference, and decision-making. As appropriate for such different schemes, the model's design allows the integration of functional modules and data structures to:

[2] These domains are described in (Alberts and Garstka 2001; Alberts 2002)

- reflect the uncertainty and error in what individuals know;
- represent how they act on that knowledge, and
- capture metrics that correlate levels of SA/SU with operational outcomes.

The model's intent is to provide a tool that can aid researchers in better understanding of this problem space human behavioral representation in general or in the representation of SA/SU and decision-making in particular. There is an extensive body of research on decision-making under uncertainty[3]; this research provides a methodology for applying elements of that research to improve the representation of decision-making in MS&A. This methodology is an engineering solution to a set of practical problems in the R&D community and the operational world:

- developing information system requirements;
- designing and implementing technological solutions to meet such requirements, and
- evaluating the potential costs and benefits of proposed solutions.

Finally, this research employs the methodology and simulation model to investigate decision-making under uncertainty and error, using that model in a set of constructive simulation[4] experiments to examine how an intelligent agent's behavior is affected by information of varying levels and quality.

1.2 Background

Military analyses have traditionally assumed that the desired effect of military operations is the attrition of materiel and personnel resources (conventional combat power) of the enemy. The attrition-based view of combat modeling is supported by engineering models - deterministic, Newtonian physics-based representations of closed systems. These models typically express combat outcomes through direct measures of

[3] See for example: (Bellman and Zadeh 1970; Safavian and Landgrebe 1995; Lipshitz and Strauss 1997; Klein 1999; Alex 2000; Holmquist and Goldberg 2007; Howard 2007; Middlebrooks and Stankiewicz 2007; Cohen 2008)

[4] Chapter 2.1 discusses live, virtual, and constructive simulation.

attrition warfare, such as the “killer/victim” scoreboard, as expressed by such measures as Loss Exchange Ratios (LERs) and Force Exchange Ratios (FERs). , e.g.,(Shlapak and Davis 1991; Olwell 1997; O’Hanlon 2003; Bowley, Castles et al. 2004; Artelli and Deckro 2008).

Much current thinking, however, emphasizes that the attrition of conventional combat power is simply the means by which some other intended effect is achieved. For example, senior decision makers have suggested that achieving “behavioral” or “cognitive” effects in opposing leaders or combatants is often a goal of military operations in addition to, or at least in conjunction with, force attrition objectives, e.g.,(Alberts and Garstka 2001; Deptula 2001; Roske Jr 2002; Smith 2002).

This is neither a new or revolutionary way of thinking about war, military theorists from Sun Tzu to von Clausewitz have expressed similar ideas. In the focus on the attrition of conventional combat power, however, other effects (cognitive, behavioral) have too often been largely ignored. (Pew and Mavor 1998; Zimm 1999; Horne 2001; Goerger 2002; Lauren 2006; Middleton 2008; NATO/RTO/HFM 2009)

As discussed in Roske[5] and others, attrition modeling approaches provide necessary, but not sufficient tools for analysis of military operations. In addition, the military operations research community needs models and methodologies that reflect the fact that conflict represents an open system, one in which energy, materiel, and information flow across conflict boundaries.

Operational analysis (and analysts) must adapt to encounter and engage the data and methods of the human sciences, and of other disciplines such as history, political science, and economics, in order to make effective use both of the insights these disciplines have to offer. They need to augment application of the physical sciences with

[5] Vincent Roske serving under the Chairman of the Joint Chiefs of Staff as the Deputy Director, J8 (Wargaming, Simulation & Analysis) (Roske Jr 2002)

other, more human-centric disciplines such as psychology and the social sciences. They need to complement and supplement the focus on physical incapacitation of the enemy with consideration of the so-called that are critical to an individual's behavior: will; morale; leadership; training, and values related to ethnic background and national makeup. These concepts relate to virtually all forms of human endeavor, extending far beyond military operations research, e.g., (Wooldridge and Ciancarini 2001; Bonabeau 2002; Macy and Willer 2002; Sanchez and Lucas 2002).

These "soft factors" emphasize the psychological or cognitive components of warfare, recognizing that the will to fight can be as important a consideration as the capability to fight. Analytical paradigms that incorporate these factors must be supported by the ability to measure success against mission objectives such as the disruption of enemy operations without destruction of critical infrastructure, the containment of insurgency, and more generally, the positive projection of force to achieve political, rather than strictly military objectives, e.g.,(Heeringa and Cohen 2000; Lampe, Schwarz et al. 2006; Middleton 2008)

One might think the "human" should have already been the focus of representing the behavior of the warfighter, but to date there has been far more emphasis on the physical actions of that individual and far too little attention to his/her psycho-physiological state and cognitive behaviors. Furthermore, it is one thing to conclude that human factors such as morale, leadership, national identity, and combat experience, play an important role in conflict, and quite another to contend that such factors can and should be taken out of the "too hard to do" box, and represented in military models and simulations. Such simulations are now relied on for a spectrum of applications from operational planning and force employment, to the development and fielding of materiel for those forces; getting it "right" in modeling and simulation is now more important than ever before.

Including these factors in MS&A requires merging elements of engineering, operations research, the behavioral and social sciences, and the humanities, all of which are relevant to modeling and simulation of modern conflict. These disciplines, however, operate under very different epistemologies. In order to derive the most benefit from the new capabilities in analytic tools, model developers and users are confronted with the challenge of understanding, and sometimes reconciling, these various approaches to acquiring and evaluating knowledge, e.g., (Westbury, Wilensky et al. 1998; Alberts, Garstka et al. 2001; Bryant, Johnson et al. 2008).

Meeting this challenge is integral to the human-centric approach, and a good starting point is to begin with experimental psychology and human factors engineering. These disciplines occupy a sort of middle ground between the physical sciences and the more theoretical social sciences, being versed in the empirical scientific methodologies of the laboratory while seeking insight into those aspects of human thought and behavior that have as yet resisted the comfortable certainty found in the "laws" of physics. Topics that inhabit this middle ground appear to be amenable to both quantitative and qualitative methodologies and thus provide a reasonable opening bid for initial forays into human-centric analysis. One such topic is SA/SU. Bringing SA/SU into the methodological framework of models and simulations introduces an essential psychological dimension of human military performance into MS&A.

Furthermore, improving SA/SU is a current emphasis in the development of information systems technologies for both military and civilian applications. MS&A can play a significant role in understanding SA/SU issues: in exploring problems and inadequacies with respect to current capabilities; in assessing the efficacy of new or proposed information technologies and in determining how best to employ these technologies. At present, however, military MS&A suffers from inadequate representation of SA/SU and decision-making, and thus lacks the tools to assess current SA/SU systems or potential improvements from new technologies. In particular, robust representation of SA/SU must address its antithesis, the so-called "fog of war", which is manifested as imperfect SA/SU - information that is uncertain, incomplete, and/or just wrong, e.g.,(Murtaugh 1994; Pew and Mavor 1998) (Ritter 2001; NATO/RTO/HFM 2009)

(Cain and Belyavin 2008).

As evidenced by such programs as the Land Warrior and its successor Nett Warrior, much of the technology being applied to improve SA/SU for the solider is designed to address **incomplete** SA/SU to support decision-making and risk assessment with respect to missing data. This drive to provide more data to the soldier is accompanied by the need to assist in filtering relevant information from large data streams (Hahn and Jezior 1999; Matthews, Shattuck et al. 2001; Smith 2011; Bailey 2012). The focus in this dissertation is to provide the ability to investigate and assess the consequences of **indeterminate**, **incorrect** and **inconsistent** SA/SU, which requires exploring how to recognize and correct SA/SU based on information that is uncertain, imprecise, or just plain wrong.

Incomplete SA/SU is generally easier to address than the other varieties. It appeals to those who would like to provide hardware solutions to SA/SU problems by increasing the amount of information available, the rate at which it can be delivered, and the extent to which it can be more widely disseminated to warfighters. Although it is hard to argue against giving decision-makers more information, there are significant problems with respect to the potential for data overload. Furthermore, simply addressing problems of incomplete SA/SU and missing data discounts equally pertinent issues with respect to the capabilities and fallibilities of the human operator. Humans can (and frequently do) function well with information that is incomplete or imprecise. On the other hand, making decisions based on incorrect or flawed information is almost always a recipe for less than desirable outcomes. Plans based on known data gaps and uncertainties are generally more robust, if only to account for unknown factors. Plans based on wrong information may rely too heavily on fallacious assumptions to optimize outcomes, with potentially catastrophic results. In addition, an incorrect understanding of an operational situation may bias subsequent information processing, and lead to flawed decision-making based on persistent problems with SA/SU. In any dynamic operational situation, it is critical to constantly revisit assumptions and ensure that they are both internally consistent and continue to be supported by new data; failure to abandon or at least adapt incorrect assumptions is an almost certain recipe for disaster.

In ABM, agents (simulated entities) make decisions according to their own individual (and probably imperfect) SA/SU. Each entity has a "perceived truth" knowledge base – an idiosyncratic view of the operational situation, as seen by that individual and obscured by the agent's local "fog of war". This research proposes that monitoring the divergence between this idiosyncratic view and simulation "ground truth" can provide a measure, in quantitative terms, of the degree to which each agent's SA/SU may be imperfect. Furthermore, subjecting each agent to the consequences of acting on its imperfect worldview supports evaluation of the operational costs of uncertain, incomplete and/or incorrect information. It also supports explicit modeling of leader decision-making processes based on such data, of imperfect command and control, and/or imperfect subordinate receipt of and subsequent execution of orders. This kind of modeling is critical for estimation of the benefits of proposed new or modified systems, and/or associated adjustments to tactics, techniques and procedures.

1.3 Approach

This research demonstrates the application of the human-centric paradigm through the use of agent-based modeling (ABM) to explore phenomena associated with SA/SU in a series of simulation experiments. This approach requires translating those phenomena into abstract modeling constructs: systems and entities, with associated behaviors, state descriptors, and static and dynamic features or characteristics. Events, terrain and environmental factors, and any other important elements of the phenomena must be represented as conceptual simulation objects, attributes, and processes, which must then be instantiated in computer code and executed according to an experimental design.

This research explores imperfect SA/SU by modeling a specific, easily understood, and quantifiable example of human behavior: intelligent agents being spatially "lost" while trying to navigate in a simulation world. In this simulation world, such agents each have a unique “mental map” – its idiosyncratic view of its geo-spatial environment. The simulation model is named MOdeling Being Intelligent and Lost (MOBIL), noting that the ability to be in both of these states simultaneously is central to the simulation.

An entity’s decisions are based on its idiosyncratic view, but behavior outcomes are based on ground truth. For example, an entity may “think” the distance between two waypoints on its route of travel may be either shorter or longer than it is in reality; if the entity moves between those waypoints its distance traveled will be determined by the actual value. Thus, as the simulation progresses, the entity’s time required for travel, and its use of resources such as fuel, will be at some divergence from its planned values for these quantities, with possible significant effects on task performance. The rate and degree to which an entity’s expectations diverge from ground truth are measures of that entity’s SA/SU, and quantify the effects of imperfect SA/SU on task performance by measuring the entity’s ability to navigate its environment given various levels of SA/SU.

Current military simulations provide a robust and technologically mature representation of terrain and geo-spatial relationships. In addition, there are now powerful simulation packages that support development and implementation of agent-based models. Integrating such representations into an agent-based model software package provides sound foundation for developing an agent-based model of movement. AnyLogic is such a package and was used to develop the simulation.

In military operations, just as in real life, being “lost” may also be a metaphor for uncertainty as to how one fits into a larger context or world view, not knowing exactly what to do, or worse, where one wants to go and what one wants to accomplish. Using agent-based modeling to represent aspects of the “being lost” in a strictly spatial sense provides a template for dealing being “lost” in the “metaphorical” sense - dealing with imperfect SA/SU in more generic contexts. Thus by studying a concrete example, one can gain insight into the nature of imperfect SA/SU, how individual decision-makers might recognize problems in their SA/SU, how they might seek to correct those problems, and/or strategies they might employ to mitigate the negative effects of imperfect SA/SU.

1.4 <u>Scope</u>

This dissertation addresses the human-centric paradigm through an engineering level modeling approach that blends just enough theory with an understanding of the practical constraints required to address significant real-world questions. Taking such an approach requires recognizing design and implementation trade-offs that must be made in developing and applying simulation tools. Inherent in any model or simulation development is the critical tension between a desire for generality and universality on the one hand, and the benefits of focus and efficiency of specificity on the other hand. This tension is closely associated with tradeoffs between two well-accepted modeling principles:

- correspondence – the closer model features actually correspond to the details of the phenomena being represented, the easier it is to both validate the model based on first principles and to subsequently modify the model to account for new features, and
- parsimony – only include those factors/details that are really needed, models are a simplification of reality precisely because it is easier and generally more useful to consider only those aspects of phenomena that are germane to the specific question being addressed by the model application at hand.

As described above, the approach seeks insights into more complicated problems of human decision-making by exploring a restricted decision space, and focusing on a single but highly representative type of decision: the fundamental question at each stage in a simulated agent's movement: "where to go next?" The question can be at a global or local level depending on the degree of precision needed in selection of the agent's "next position", whether that position represents a "nearest neighbor" point on a regular grid, the degree of advancement along a specific route segment or path, or simply movement in a given direction. The decision elements that must be represented include:

- a goal – an endpoint or location towards which the agent's movement is directed;

- the geo-spatial environment and the agent's interaction with that environment
  - environmental feature characteristics – landmarks, identifying qualitative and quantitative attributes;
  - the agent's perception of the environment – its view of those factors in its world that might influence its movement;
  - ground truth constraints – those factors in the world that restrict or otherwise affect the agent's physical movement;
- inference procedures – the algorithms and/or heuristic methods by which the agent selects that next position, given its goal, its understanding of the environment and its capabilities to move within that environment
- uncertainty and errors – the sources and extent of the agent's potentially imperfect understanding of its environment.

1.5 Research Significance

The Department of Defense has mandated that modeling and simulation play a significant role in research and development of virtually all military technologies, including those for SA/SU, e.g., (Bernstein 1998; Johnson 1998; Davis 1999; Page 2001; Ford and Dillard 2008). Providing better SA/SU for the warfighter is a big driver for the development and application of military information technologies, and the research proposed herein addresses significant gaps in current M&S capabilities to support RD&A of SA/SU capabilities.

1.5.1 Articulation of the Human-Centric Approach

This research focuses on SA/SU as a critical element of the human-centric approach to MS&A. SA/SU is central to the idea that the cognitive state of an individual may help determine success or failure on the battlefield. As such, representation of SA/SU can provide a blueprint for other aspects of human-centric modeling. In showing how aspects of SA/SU can be sufficiently defined in operational terms, and brought into the methodological framework of models and simulations, this approach captures an essential psychological dimension of human military performance. This aspect of the human-centric approach is one of the most fruitful areas for near-term enhancements to models and simulations.

Clearly, the "soft factors" mentioned in Chapter1.2, are critical factors determining not just the outcome of the battle, but of over-all conflict. It is still, however, an important question as to whether it is necessary, possible, or even desirable to include such factors in combat models, and if so, at what levels of detail, aggregation, and analysis. This research addresses this question by seeking to articulate the human-centric paradigm and by demonstrating its potential. This research provides an in-depth explication of the concept of human-centric modeling.

1.5.2 Application of Operations Research Engineering

It employs the novel application of a number of current modeling and simulation methodologies to provide an engineering approach to a complex problem in integrated systems design and assessment.

For example, Value-Focused Thinking (VFT) is a multi-criteria decision-making methodology, primarily applied to help human decision-makers. Typically VFT is applied through a process such as that outlined in Figure 1-1. The process has two major elements, model development and model application. In the first of these, individuals in the problem domain, i.e., people with a stake in solving the problem, are consulted to determine their values, the criteria that will characterize "good" solutions.

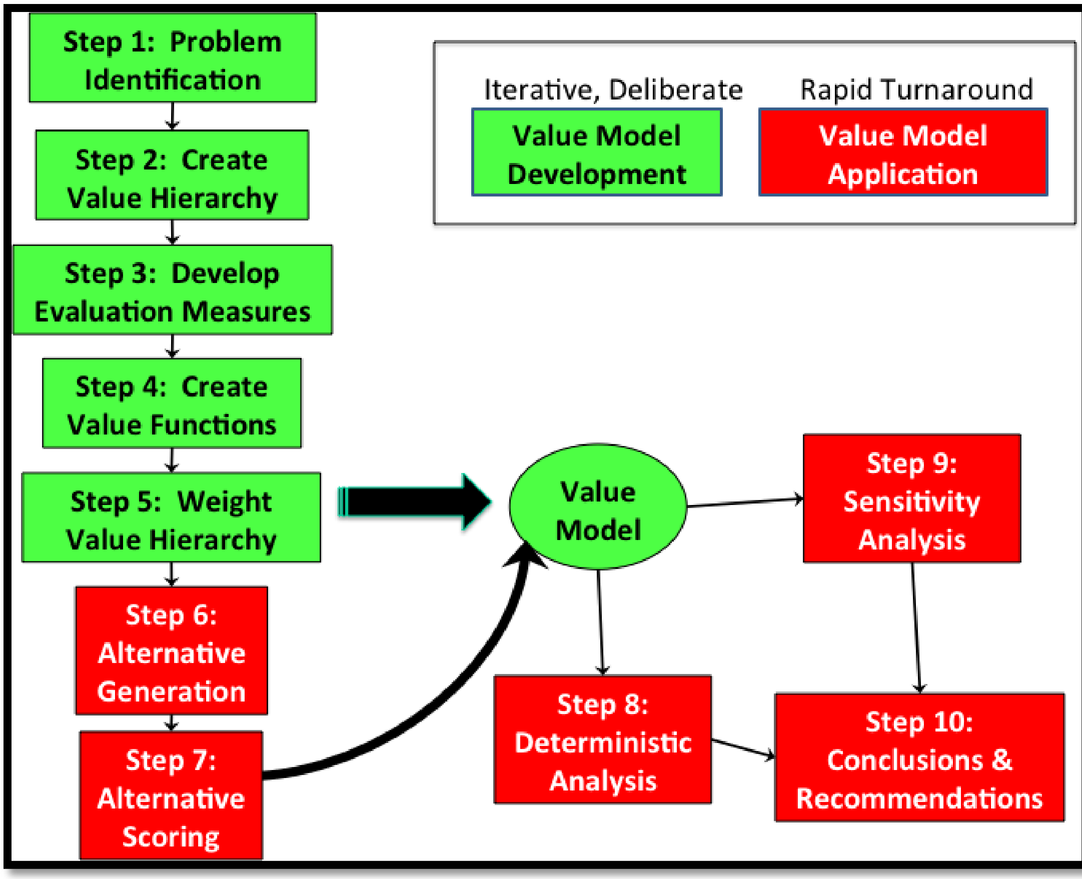


Figure 1-2 10-Step VFT Hierarchy Process (Weir 2009)

A value hierarchy is developed to represent the relative importance of each of these criteria, and measures are defined to quantify the contributions of potential solution alternatives. This process is discussed in more detail in Chapter 8.7 below, where the value hierarchy shown in Figure 1-3 is developed.

As used herein the VFT process is noteworthy on two levels. First, the value hierarchy of Figure 1-3 is embedded in MOBIL code, giving the entity simulated in MOBIL the ability correlate elements of its view of ground truth with its idiosyncratic mental map.

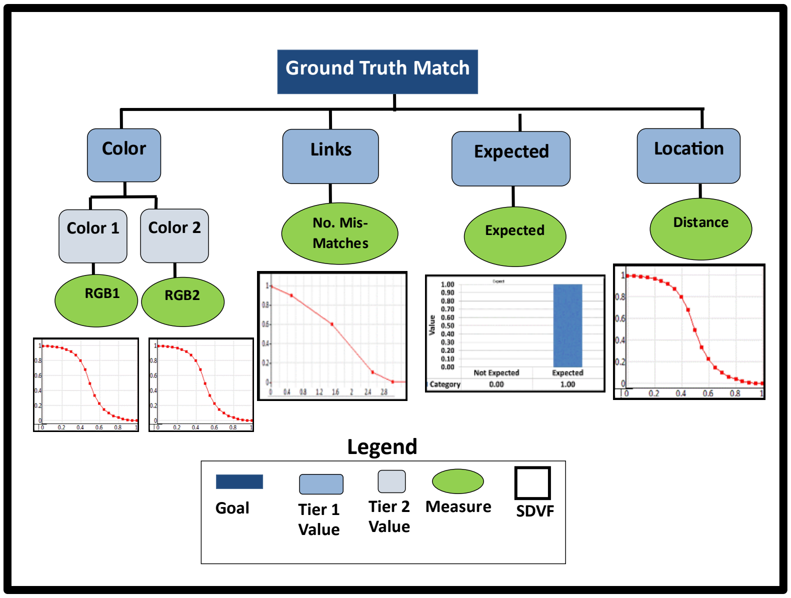


Figure 1-3 MOBIL VFT Hierarchy

As discussed in Chapter 8.7, representing decision processes with VFT is compatible with theories of SA/SU, including: Endsley's three stages of Situation Awareness (Endsley 1995a; Endsley 1995b) and John Boyd's Observe, Orient, Decide and Act (OODA) loop[6] (Boyd 1986; Boyd 1987).

The VFT process, however, is more than just a mechanism for comparing alternatives. The model development part of VFT is a deliberative process that iterates between customer/user interviews and definition of model elements. As described by Keeney, the deliberative process is simple in concept:

> *"You begin with the fundamental objectives that indicate what you really care about in the problem. Then you follow simple logical reasoning processes to identify the mechanisms by which the fundamental objectives can be achieved. Finally, for each mechanism, you create alternatives or classes or alternatives by asking what control you have over that mechanism." (Keeney 1996),*

Not coincidentally, this deliberative process, and the model application process that

[6] Strictly speaking, Boyd's OODA loop is not a model of SA/SU, but a dynamic decision-making process in which the objective is defeat to an adversary by operating at a higher tempo, cycling through one's own OODA loop processes more rapidly than the adversary's update rate. The OODA loop concept, however, does share considerable commonality with more theoretical views of SA/SU.

follows it, together defines an approach very similar to that used in simulation-based analysis. Figure 1-4, which is taken from (Middleton 2008),

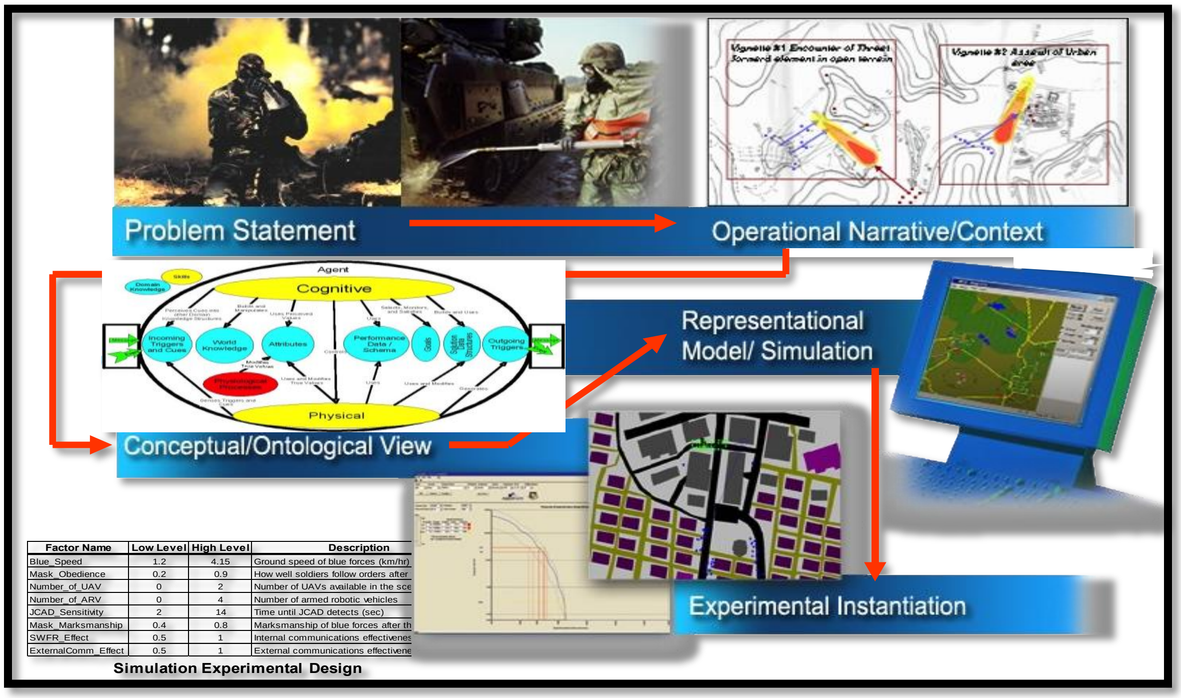


Figure 1-4 Simulation-Based Analysis

depicts simulation based analysis as a five step process:

- Definition of the Problem Statement – Just as in the VFT process, an explicit problem statement captures the analysis goal, placing objectives in the context of key constraints, limitations, and assumptions, with a scope bounded within the resources and time and time available, and most important of all, clear establishment of desired content and form of the answer.
- Operational Narrative/Context - The problem statement provides a skeletal structure for the analysis; the substance or "flesh" is found in operational narratives – use cases expressed through scenarios or vignettes to frame specific elements of the problem in a relevant context. This step basically corresponds to the definition of values in the VFT process; it captures the factors that are important to the individuals seeking to solve the problem. Construction of appropriate narratives requires subject matter experts (SMEs) to provide an experiential understanding of operational situation, and to ensure that the narrative is expressed in "operator-speak" instead of analytical

"geek–speak."

- Conceptual/Ontological View - In the VFT process, one constructs the value hierarchy, which is in essence a model of the value relationships and a formal description of the elements that must be quantified. In simulation-based analysis the operational narrative(s) must be instantiated in models and simulations, which requires translating the narratives into abstract modeling constructs: systems and entities, with associated behaviors, state descriptors, and static and dynamic features or characteristics.
- Representational Model/ Simulation – Both VFT and simulation-based analysis require the adroit choice of measures that translate characteristics of the alternatives being studied into aggregate assessments of their potential to solve the problem under consideration[7].
- Experimental Instantiation – In both VFT and simulation-based analysis, once the model structure has been defined, the alternatives to be studied must be identified and characterized in terms of input data. Typically the model is executed iteratively for sensitivity analysis and insight into the nature and potential of those alternatives. Frequently additional alternatives are developed and studied based on preliminary results.

While both VFT and simulation-based analysis have been defined as multi-step processes, they rarely, if ever, are carried out as a strict sequence of these steps. In general there is considerable interaction and refinement of the model and the alternatives as one works through the processes involved. Both VFT and simulation-based analysis provide a framework for evolving from a usually imprecise and ambiguous problem concept to explicit analytical assessment of potential problem solutions.

The essence of both VFT and simulation-based analysis capture the engineering approach to operations research that is the core of this dissertation.

---

[7] A through discussion of Measures of Performance and Effectiveness (MOPs & MOEs) is provide in Chapter 2.5

1.5.3 Analysis of Simulation Experiments

This dissertation blends object oriented software design and development principles with agent based modeling to establish the utility of applying the human-centric approach to analysis, as evidenced by the results of a series of simulation experiments. For example, as seen in Figure 1-5, a JMP regression tree analysis of MOBIL output shows how one might develop a set of rules for changing goal-seeking strategies based on incremental progress measures.

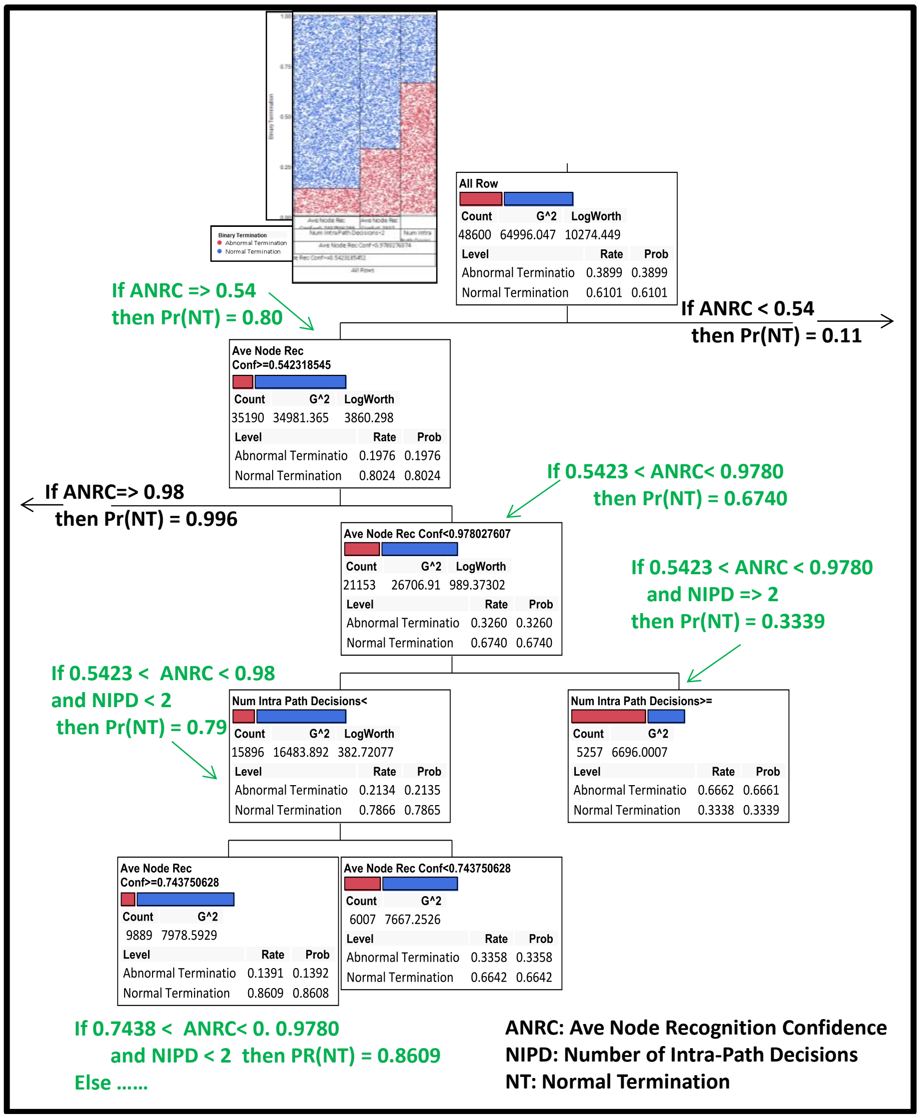


Figure 1-5 Progress Rules Suggested by Analysis of MOBIL Results

1.6 Dissertation Structure

This dissertation has two primary components, the first being the articulation of the human-centric paradigm, with a focus on the roles agent-based modeling, SA/SU, and decision-making have in that paradigm. Chapters 2 through 6 address this component. Chapter 2 begins by describing the world that led me to this research, the world of military operations research. Chapter 3 discusses agent-based modeling and complex adaptive systems, Chapter 4 goes into the details of SA/SU, Chapter 5 examines cognitive architectures and the decision-making process, and Chapter 6 discusses the issues of modeling mobility, route planning, and related issues.

The second main component of the dissertation begins with a brief discussion of the research methodology and the use of simulation experiments in Chapter 7, with a description of the model, MOBIL, itself, in Chapter 8. The next four chapters are devoted to four separate experiments that examine the role of distorting a simulated entity's available information, the entity's ability to reconcile that distorted information with ground truth, an in-depth look at interactions between those two factors, and finally a look at an on-going modification of the model and its application. The dissertation ends with a brief summary and conclusions.

2 THE HUMAN-CENTRIC APPROACH FROM THE MILITARY PERSPECTIVE

This chapter discusses the concept of the human centric approach from the perspective of military MS&A. It provides a brief history of military MS&A, looks at current trends in military operational philosophy, relates those trends to the concept of the individual dismounted combatant as a weapon system, and addresses the need for an engineering model of that weapon system.

2.1 A Brief History of Military MS&A

This history traces military MS&A from sand tables and board games to the birth of operations research (OR) and to the development and application of today's sophisticated computer codes.

2.1.1 War Games From Sun Tzu to World War I

Military M&S traces its origins to war games, which are essentially simulations of military operations. (St Clair 2003; Little 2006) relate that Sun Tzu (circa 500 BCE) is said to have written about the game known to the Chinese as Wei Hai (or Weiqi), a game that may date back as far as 2000 or 3000 BCE. It is still popular today, having migrated sometime around the 7^{th} century A.D. to Japan, where it is known as Igo (giving rise to the English name Go) and from there (albeit much more recently) to the US and other Western countries. "We Hai" literally means "encirclement" and victory is gained by outflanking one's opponent and controlling more territory. It is interesting to note that such strategy is in keeping with Sun Tzu's philosophy of war, as expressed in his treatise on the Art of War (see for example, (Sawyer 2005)). Sun Tzu advises against frontal attack whenever possible, and even counsels that any form of direct combat should be used only as a last resort to achieve military and political objectives[8].

Go player pieces are white and black stones that are alternatively placed at the interchapter of a rectangular grid, and pieces have no intrinsic value of their own but derive their value from their position vis-à-vis other pieces on the board[9]. While the game's rules seem simple, it has as yet resisted attempts to develop the Go-equivalent of the IBM's chess-playing Deep Blue, a computer program that can defeat human experts. (Stern, Graepel et al. 2004; van der Werf 2004) attribute this difficulty in large part to the lack of good evalution functions for different Go positions and the large number of branching opportunities at each turn in the game. These problems are certainly inter-related, and both are endemic in real world military operations. Aspects of them will crop up in later discussions on the need for appropriate measures of performance and effectivess, and in the desirability of modeling military operations as complex adaptive systems. From a game theory point of view, both Go and chess are strictly determined games; at each turn there are a finite number of discrete and known possible move choices. Thus, there exists an optimal strategy for playing these games that guarantees at least a draw, and possibly outright victory to the individual with the first move. Of course, the problem is that the number of possible moves is too great for practical computation of that strategy, at least given current technology. So, while in theory at least, players have perfect perception of the current situation, what they lack (at least until the end becomes inevitable) is perfect understanding of the outcomes that may be projected from that situation. This difference between perception and understanding is at the core of the current argument over the need to distinguish between SA and SU, and will be addressed further in Chapter 4.

[8] Such counsel should not be taken to mean that Sun Tzu eschewed the use of force; on the contrary he had a reputation for ruthlessness.

[9] see for example www.gobase.org for the rules and other information about the game

(Young 1957; Lee 1990) state that chess is believed to be the oldest form of war game, (which may be true from a Western perspective). Its origin is usually ascribed to India in around the 7th century, where it appeared as a Hindu battle game called Chaturanga (meaning four divisions, or Chaturaji, meaning "four kings or rajahs"). Young describes the game as "*played on a board, with a highly conventionalised map using various pieces to represent the arm of the service then in existence: elephants, horses, chariots, and foot soldiers.*" In today's terms, it would be called a human-in-the-loop (HITL) Monte Carlo simulation, one in which multiple players take turns with outcomes described by the throw of a form of dice.

One of the earliest European board games aimed exploring strategies of warfare was Koengspiel (German for King's game) developed in 1664. As described by Young and others (Lee 1990; Perla 1990; St Clair 2003; Little 2006), it was followed by the development of various other forms of "War Chess" throughout the 18th and 19th centuries. These games provided increasing levels of resolution and detail with respect to the forces represented by player "pieces" and the battlefield environment represented by "board" characteristics. These games included von Reisswitz's 1811 Kriegsspiel (literally War Game, usually referred to in English as Kriegspiel), which replaced gridded game boards with a sand table. While von Reisswitz is not the first to use a sand table[10], this change allowed the representation of terrain in three dimensions, and permitted the free movement of multiple types of forces, represented by colored and labeled blocks, i.e. not constrained to board squares. His son continued development of Kreigspiel, replacing the sand table with large-scale, detailed topographic maps, and revising the rules of the game[11] to calculate combat outcomes, as opposed to his father's approach of determining them through discussion. The

[10] (Smith 2010)for example, reports that commanders in the Roman Empire used sand tables and represented military forces and resources through various physical icons to visualize the battlefield and plan operations

[11] The rules of Kriegspiel were published under the title "Instructions for the Representation of Tactical Maneuvers under the Guise of a Wargame", although sources differ as to the publication date.

younger von Reisswitz also added umpires to determine casualties and settle disputes, and introduced the red and blue color-coding for opposing sides that is still used today. His changes were the forerunners of the various modern concepts of a “battle calculus” to define rules for modeling force interactions, estimating casualties, and measuring force effectiveness, and again speak to the need for appropriate “evaluation” functions in military operational analysis.

US Army Major James Livermore’s *The American Kriegspiel: A Game For Practicing The Art Of War Upon A Topographical Map* (Livermore 1879) is credited with bringing war gaming to the U.S. He introduced logistics and made the first attempt to include such factors as fatigue to the game. He was followed by US Army Lieutenant Charles A. L. Totten, whose *Strategos : A Series of American Games of War Based Upon Military Principles and Designed for the Assistance Both of Beginners and Advanced Students in Prosecuting the Whole study of Tactics, Grand Tactics, Strategy, Military History, and the Various Operations of War* (Totten 1880) proposed adding layers to the game depicting the different facets or levels of war: tactical, operational, and strategic, which according to Little (op cit.) is the first time a ‘hierarchy’ in modeling is seen. The U.S. Naval War College codified U.S. Kriegspiel in 1884 and in 1887 fully integrated war- gaming into the curriculum for all attending officers.

In a 1914 article written in cooperation with the US Army, Scientific American reported that “*gaming was by then used in the instruction of every army in the world*”, although these games were used “*not to see who won, but to get results and experience, and to profit by mistakes made*.” (Perla 1990)

2.1.2 Lanchester and the Origins of Quantitative Analysis

During World War I Frederick Lanchester took a more mathematical approach to analysis of military operations and began the modern trend towards trying to predict combat outcomes through aggregation of a small number of measurable factors. As discussed in detail by (Taylor 1983), Lanchester quantified the combat attrition process through a series of differential equations in which casualty exchange ratios are dependent on the relative size of forces and firepower effectiveness parameters (such as the "a"s and "b"s shown Figure 2-1).

Firer v. Firer **F|F** **“modern combat”**	**FirerTarget v. FirerTarget** **FT|FT** **“ancient combat or area fire”**	**Firer v. FirerTarget** **F|FT** **“mixed combat or defensive combat”**
$\frac{dx}{dt} = -ay$	$\frac{dx}{dt} = -axy$	$\frac{dx}{dt} = -ay$
$\frac{dy}{dt} = -bx$	$\frac{dy}{dt} = -bxy$	$\frac{dy}{dt} = -bxy$

Figure 2-1 Representative Lanchester Equations

These parameters are called attrition rate coefficients and are defined by individual weapon system kill rates. Used in concert with equations that appropriately reflect assumptions in the type(s) of combat under study, they allow examination of the advantages of force concentration and the relative contribution of different weapons/units, including the value of indirect fires.

For example, the “modern” view of warfare shown in Figure 2-1 assumes attrition by concentration of forces, where the rate of attrition of each side is due only to the number and weapons’ firepower of the enemy, i.e. the rate of change of side x is a function of y and vice versa.   This view can be differentiated from “ancient” combat, which is characterized as basically a series of individual duels in which the attrition rate of each side is a function not only of the killing power of the enemy, but also of the number of individuals that enemy faces. In this view, ancient combat is similar to modern area fire in that the attrition rate of one side is due not only to the number and efficiency of the enemy, but also to the density of forces facing that enemy. As a result, in “modern” combat, the concept of defense shifts from parrying the thrust of contact weapons (or shielding oneself against area weapons) to the idea of killing your enemy with your own weapons fire - before he can kill you.

The Lanchester equations (and variants thereon, developed by numerous researchers up to the current day) have been modified to incorporate, among other factors, time-variant attrition rate coefficients, troop movement, and stochastic factors. (Taylor, Yildirim et al. 2000); (Schaffer 1968); (Deitchman 1962);They are an example of modeling of the performance military systems through the systems' "effects" (implicit representation of systems' functions by substitution of expected outcomes) as opposed to modeling the "processes" (explicit representation of those functions) involved in the interactions between different weapons systems and/or force structures. Although the current emphasis in modeling and simulation is to provide explicit representation (and visualization) of such combat interactions, the Lanchester equations are still valuable tools for the military analyst. (Bowley, Castles et al. 2003); (MacKay 2008) (Artelli and Deckro 2008) They have, however, been justifiably criticized for not handling the non-linear nature of combat well, and for ignoring many of key factors in battle outcomes, most especially the human element. (Ilachinski 2004); (Ipekci 2002)

2.1.3 The Birth of Operations Research

World War II saw the accepted beginning of Operations Research (OR, also Operational Research as the British first called it) in the need to optimize resources of men and materiel for maximum military effect. Manchester (UK) physicist Patrick Blackett, who is generally heralded as the father of operations research, headed up an OR group attached to the British Anti-Aircraft Command whose principal charge was the improvement in deployment and use of Britain's new radar network. According to (Rider 1994), Blackett distinguished between two general methods of optimization, the a priori method that applied differential equations to selected important variables, and variational methods that used experimental and analytical methods to test the results of variation in such factors as the properties of the weapons or the tactics used.

Both of these approaches are still very much in use today, and Blackett's variational method can be seen as the precursor of the current trend towards virtual experiments in a simulation environment.

2.1.4 The Rise of Technology: Military Operations as Physics

In the years since the end of World War II, computers have taken an ever-increasing role in military operations, as they have in every day life. DARPA and the Internet, to DIS and HLA following the trend to network centric warfare. Little's History and Basics of M&S sums up the progress of the computer era:

> *While the first 50 years of the 20th Century overshadowed a preponderance of design constructs and analysis, the Cold War was still what we now consider manual … analogue … primitive. By 1950, it is still large boards, playing pieces and push pins. While still a novelty for many in the military community, the research industry continually improved their methods of quantitative analysis, bringing the two ever closer into a discipline.*
>
> *Three significant occurrences changed the status quo: the transistor, integrated circuitry and commercially obtainable computers. Once this (sic) became available to the analysis community, M&S increased exponentially. Imagine if you will the notebook computer or personal computer that allows you to read this. Your operating system was only dreamed of in the 1950s. While the large wargaming rooms had mock terrain and push pieces, an equal amount of room was needed for our early models. The logic behind this impetus of modernization was simply one of global political-military competition. In 1958, the Cold War took on a new twist when Russia launched Sputnik. In response, President Eisenhower created a host of organizations in response. One of them is the Defense Advanced Research Projects Agency, which played a role in our field as well as the precursor of the internet.*
>
> *The Naval War College by this time has a computer-assisted wargame taking up three floors of an academic building. Built to accommodate a single wargame, it is torn down completely and reassembled from the very beginning for a subsequent effort. This capability was the natural progression of the wargames developed there by William McCarty-Little in the 1880's. In spite of its obvious drawbacks, this gets the attention of U.S. Navy fleet commanders, who send their staffs to train on it.*
>
> *During the 1960's, IBM roles its mainframe computer called the '7070.' This is the first computer capable of supporting simulations; again a far cry from what you are using now. By 1964, the Naval War College replaces analogue technologies with digital and conducts its first remote wargaming. This is what we in the U.S. call the 'McNamara Era'2.* ***During Vietnam, the overemphasis of quantification distorts decision-making at the highest levels and M&S suffered***

for some years afterwards[12]*. (Little 2006)*

In this era, attempts to quantify military operations focused on ways to “keep score”, from the body counts of Viet Nam to the “killer/victim” scorecards of various force-on-force computer models. The analytical tools of the first four post-WWII decades supported military thinking dominated by the oppressive reality of the Cold War. The overriding focus of operational analysis was on a relatively stable, large-scale, heavily armored, largely symmetrical confrontation across the former inner-German-border.

> *The Cold War was on. The military knew where the next war would be fought - in the Fulda Gap, along the Elbe, between East Germany and West Germany. So, all of the West's military planning was focused on preventing this specific next war, or on winning it. It would be hard to overestimate the amount of time and resources that were devoted to analyzing this single scenario. (Walker 2000)*

This kind of military analysis was well served by a generation of models and simulations that considered materiel systems, and their destruction or incapacitation, as the primary determinants of combat outcomes. The human combatant, when considered at all, was viewed as an intrinsic element of weapons systems/platforms, or as a constituent element of an aggregate force, important mainly as a contributor to force numbers.

In consonance with this view of warfare, this same time period saw a continually increasing emphasis on the role of technology in solving military problems. That role was (and still is) an integral part in the growth of what President Eisenhower referred to as the U.S. military-industrial complex (Eisenhower 1961). The growth of the military-industrial complex was accompanied the application of industrial forms of systems analysis to military affairs. Organizational management cost accounting methods and quantifiable measures of merit are now an essential part of virtually all Department of Defense programs.

[12] Emphasis mine

The integration of these practices from civilian industry to the military has, however, often been uneasy. This unease is perhaps best typified by the experience of Robert McNamara and his "Whiz Kids"[13] in the early sixties. As explained by Air Force Colonel R. Philip Deavel:

> *The dominance of systems analysis in the early 1960s flowed not from the intellectual brilliance of McNamara and the Whiz Kids, though in their hubris they believed so. Their ideas only appeared to shine brightly when compared with the utter inability of the military services to quantify their own objectives, or credibly dissect the methodology of the Whiz Kids. As one of McNamara's analysts succinctly explained their ideological dominance, "Other people had objectives, we had arithmetic."*
>
> *Rather than deal effectively with McNamara on his own terms, the uniformed military tended to dismiss all systems analysts and their civilian advocates, as the proverbial "pencil-necked geeks" who knew nothing of the equally proverbial "real world."*
>
> *...*
>
> *The struggle between McNamara and the officer corps, … is often cast as a contest between military and civilian values. While superficially true, this analysis misses the mark. A long historical view indicates the partisans of both groups represent two separate but equally honorable military philosophies.*
>
> *McNamara and his proteges are the modern disciples of Jomini[14]. Like this great Napoleonic strategist, they view warfare as a cold and precise science. To McNamara, and to Jomini, success goes to the leader with the greatest organizational skill in building and wielding a massed military force. It is* ***warfare as the science of physics[15]****; the ability to concentrate energy and unleash it on an opponent.*
>
> *…*
>
> *At the other end of the philosophical spectrum, the American officer corps are, in the aggregate, disciples of Clausewitz. As such, they view warfare as ultimately a human attribute, an art that can never be completely quantified in a mathematical equation. The firm political support of the nation, flowing*

[13] The "Whiz Kids" was a name given to a group of experts brought into the government by McNamara when he was John Kennedy's Secretary of Defense. The name was taken from a similar group a generation earlier at Ford Motor Company to which McNamara belonged and who are credited with bringing the company to renewed profitability through modern management methods. McNamara was serving as President of Ford when Kennedy appointed him Secretary of Defense and his new Whiz kids were a group of young primarily Ivy League analysts with relatively little military experience.

[14] Baron Antoine Henri de Jomini (1779-1869) one of Napoleon's best generals, author in 1838 of Precis de l'Art de Guerre later printed in English as The Art of War (not to be confused with the similarly titled works of Sun Tzu)

[15] Emphasis mine

through the iron will of the commander energizes the force and cuts through the fog and friction of war. It is a philosophy that gives little credibility to those who would predict success or failure based upon the laws of physics or calculations of economic efficiency. (Deavel 1998)

The concept of warfare "*as the science of physics*" was heavily reflected in the analytical models of this era, which saw enhanced representation of the physics of the battlefield through the addition of more and more detail to describe the accuracy and lethality of weapon systems, and the mobility of large weapon platforms. Models of these systems and their operating environment became the basis for weapon simulators - training devices where individuals could practice their interface with the system. Ever increasing computer technology was applied to providing the level of resolution needed to adequately represent the battlefield physical environment in simulating the interaction of the large, fast moving platforms.

2.1.5 War Gaming in CyberSpace and the Virtual Human

As the technology supporting weapon system simulators advanced, it was natural to try and link them together, and to promote a heterogeneous battlefield for joint service interaction. In 1985 DARPA started development of SIMNET, a standard for distributed interactive simulations(Brock, Montana et al. 1992; Cosby 1995), and a decade later the Department of Defense began requiring that new simulations and simulators be compatible with High-Level Architecture (HLA) a newer, evolving standard to ensure compatibility and interoperability (Page 1998; Straßburger, Schulze et al. 1998; Davis and Anderson 2004). Human operators being an integral part of weapon system simulators, distributed simulations have also embraced the human-in-the-loop (HITL) as a key participant in the world of distributed simulation. By the turn of the century the concept of war gaming had substituted cyberspace for the game board and sand table (Belanich, Sibley et al. 2004; Pellegrino and Scott 2004). Such gaming is not, of course, not limited to the Department of Defense, multi-player first-person-shooter games are immensely popular on the Internet. The commercial gaming industry is a multi-billion dollar business and drives much of the development of advanced artificial intelligence capabilities, in an effort to make simulated game entities competitive with their human adversaries.

As a final note, it is somewhat ironic that the Department of Defense's embrace of computer technology for virtual war gaming has re-emphasized the role of the human decision-maker and thus supports the view of "*warfare as ultimately a human attribute, an art that can never be completely quantified in a mathematical equation*"(Deavel 1998).

2.2 Beyond Attrition Modeling

The attrition-based view of combat modeling that dominated the last century is not, of course, the only approach to understanding of the nature of armed conflict. Network-Centric Warfare (Money 2001), the "maneuverist approach" (Zimm 1999; Zimm 2001) and "Effects Based Operations - EBO" (Deptula 2001; Batschelet 2002; Roske Jr 2002; Smith 2002; Wagenhals and Levis 2002; James and Daniels 2005; Phister Jr, Fayette et al. 2005) are other conceptual viewpoints that have been advanced to help frame discussion about the most effective ways to plan and conduct military operations, as well as to assess the success of such operations.

2.2.1 Network Centric Warfare

The Network Centric Warfare concept holds:

> *Warfare takes place simultaneously in and among the physical, the information, and the cognitive domains.*
>
> *Physical Domain: The physical domain is the traditional domain of warfare. It is domain where strike, protect, and maneuver take place across the environments of ground, sea, air, and space. It is the domain where physical platforms and the communications networks that connect them reside. Comparatively, the elements of this domain are the easiest to measure, and consequently, combat power has traditionally been measured primarily in this domain. Two important metrics for measuring combat power in this domain, lethality and survivability, have been and continue to be benchmarks for measuring the effectiveness of combat operations.*
>
> *Information Domain: The information domain is the domain where information lives. It is the domain where information is created, manipulated, and shared. It is the domain that facilitates the communication of information among*

warfighters. It is the domain where the command and control of modern military forces is communicated, where commander's intent is conveyed. Consequently, it is increasingly the information domain that must be protected and defended to enable a force to generate combat power in the face of offensive actions taken by an adversary. And, in the all-important battle for information superiority, the information domain is ground zero.

Cognitive Domain: The cognitive domain is the domain of the mind of the warfighter and the supporting populous. This is the domain where many battles and wars are won and lost. This is the domain of intangibles: leadership, morale, unit cohesion, level of training and experience, situational awareness, and public opinion. This is the domain where commander's intent, doctrine, tactics, techniques, and procedures reside. Much has been written about this domain, and key attributes of this domain have remained relatively constant since Sun Tzu wrote The Art of War. The attributes of this domain are extremely difficult to measure, and each sub-domain (each individual mind) is unique. (Money 2001)

By recognizing that information and cognitive domains complement and supplement the focus on physical incapacitation of the enemy, NCW incorporates the psychological and/or cognitive components of warfare, and acknowledges the key that information technologies and decision support tools play in today's military operations.

2.2.2 Maneuver Warfare vs. Attrition Warfare

Military operational concepts today accept that disruption of an opponent's operational tempo, and especially the ability to interfere with his Observation, Orientation, Decision, Action (OODA) loop[16] (Boyd 1986; Boyd 1987), may be a key operational objective. Accordingly, analysis of such operations requires the ability to measure success in achieving such mission objectives as the interruption or forced cessation of enemy operations, without destruction of critical infrastructure, the containment of insurgency, and more generally, the positive projection of force to achieve political, rather than strictly military objectives[17].

Alan Zimm's essay "A Causal Model of Warfare" in (Horne 2001) states:

---

[16] The OODA loop is described in more detail in Chapter 4 below.

[17] Chapter 2.5 addresses the need for new metrics, and how the methodological framework proposed here needs to support generation of analytical data to apply them.

> *Ever since Sun Tzu set brush to paper in the sixth century B.C. thoughtful people have struggled to discover the key to victory in armed conflict. At various times victory was supposed to be reached through the death of the enemy's leaders, capture of their capitol, occupation of territory, or defeat of the enemy's army. At various times each of these paths was accomplished but failed to achieve victory.*
>
> *Since the middle 1800's – roughly corresponding to the dissemination of Clausewitz's On War – most military organizations have assumed that victory comes through the defeat of the enemy's army, a defeat accomplished by the destruction of soldiers and equipment. However, the Marine Corps, particularly since the 1989 issue of FMF1, Warfighting, and the subsequent development of Operational Maneuver… from the Sea[18], has advanced a different approach. Marine Corps doctrine targets destruction of "the cohesion of the enemy system," and identifies the mechanism of victory as "panic and paralysis, (Zimm 1999)an] enemy who has lost the ability to resist."[19] This has been presented as a contrast between "Maneuver Warfare" and "Attrition Warfare." " (Horne 2001)*

(Zimm 1999) (Lind 1985; Zimm 2001) further frame the "maneuverist vs. attrition approach as one in which the concepts of "will" , "morale," the "moral dimension of combat," and related ideas complement and supplement the focus on physical incapacitation of the enemy. These concepts recognize that the will to fight can be as important a consideration as the capability to fight. Addressing them analytically requires quantifying their effects, measuring success according to mission objectives that combine the disruption of enemy operations with avoidance of the destruction of critical infrastructure, that strive for the containment of insurgency without requiring the annihilation of the insurgents, and more generally, that posit the positive projection of force to achieve political, rather than strictly military objectives.

2.2.3 Effects Based Operations

The EBO concept similarly supports the view of the importance of the "soft factors" as critical factors in the outcome of military operations.

---

[18] (Corps. 1996)

[19] cites (MCDP1 1997)

There are a variety of modeling techniques that are used to relate actions to effects. With respect to effects on physical systems, engineering or physics based models have been developed that can predict the impact of various actions on systems and assess their vulnerabilities. When it comes to the belief and reasoning domain, engineering models are less appropriate. The purpose of affecting the physical systems is to convince the leadership of an adversary to change its behavior, that is, to make decisions that it would not otherwise make. Thus, the effects on the physical systems influence the beliefs and the decision making of the adversary. Because of the subjective nature of belief and reasoning, probabilistic modeling techniques such as Bayesian Nets and their influence net cousin have been applied to these types of problems. Models created using these techniques can relate actions to effects through probabilistic cause and effect relationships. Such probabilistic modeling techniques can be used to analyze how the actions affect the beliefs and thus the decisions of the adversary.

Thus the EBO concept results in a shift in focus. Instead of focusing on the servicing of a well defined a priori target list, we focus on the effects that we wish to achieve. The target list still exists and includes both hard and soft targets: from weapons systems, to C2 nodes, to leadership nodes, to infrastructure nodes, to the contents of communications. But the target list is only an intermediate construct, a means to an end, which can change rapidly as the effects we wish on the adversary are being achieved or not. Indeed, the list of possible actions we can take is now much larger as it includes all instruments of national (or coalition) power: political, military, or humanitarian; physical or ideological. The availability of all instruments gives us much flexibility in trying to achieve the desired effects and to avoid undesirable ones. But it also makes the Course of Action (COA) problem and the subsequent planning problem much harder. There are now many alternatives, many choices. The choice of a set of actions, their sequencing, and their time phasing become a problem in their own right. (Wagenhals and Levis 2002)

2.3 The Evolution of The Warrior System

The 20th century was characterized by an explosion of technology, much of it driven for or applied to war. From the Wright Brothers to the stealth bomber, from WWI tanks to the M1 Abrams, and from the early days of radio to current C4ISR[20] systems, military applications of 20th century technology produced increasing more capable and more complex mobile weapon systems. The costs of large weapon platforms are now measured in the billions of dollars. By comparison the individual soldier received

[20] Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance

relatively short shrift. Research, development, acquisition and application of materiel for the Infantryman were at far lower levels of expenditures than for "war machines". Such expenditures as were made for the Infantryman focused on individual pieces of equipment. These so-called equipment "eaches" tended to follow a stove-piped process; with largely separate development efforts, each primarily aimed at providing the individual soldier/marine with either better weapons or better body armor. Materiel solutions for the individual dismounted combatant were not viewed as integrated parts of larger weapon systems in the way that equipment provided to crews of planes, tanks, ships and submarines was.

2.3.1 The Soldier as a System

In the late 80's and early 90's, however, the Department of Defense sought to correct this oversight with the development of a new approach, the concept of the "soldier as a system".

> *Just as Desert Storm ended, the 1991 Army Science Board (ASB) conducted a summer study regarding how much Soldiers should carry and, most important, who should decide what and how much. This study, entitled "Soldier as a System" (SaaS), came to the following five conclusions: (1) The requirement to properly equip the Solder for combat is as complex as those of other programs such as the Abrams tank, Bradley fighting vehicle, Patriot missile system, and Black Hawk helicopter programs; (2) Existing Soldier equipment mismatches due to lack of integration are reducing combat efficiency and endangering Soldiers; (3) The planned "Block Change" concept of equipping the force (no new equipment is fielded until enough is procured for the entire Army) is an outdated concept; (4) Promising new technological capabilities should be exploited to ensure battlefield overmatch for the American Soldier; (5) The Army should develop and employ experimentation (wargaming and simulations) with emphasis on future Solder system threats[21].(Lockhart 2006)*

This concept was driven both by the Army's desire to obtain levels of Congressional funding for soldier equipment more on par with that expended for other weapons "systems", and the growing realization that the equipment for the individual had

---

[21] Lockhart cites (Haley 1991) **Soldier as a System;** 1991 Army Science Board Summer Study, Final Report; U.S. Department of the Army; December 1991, see also (Middleton 1991) and (Middleton 2001)

reached the point where the "eaches" needed to function in concert. The Soldier as a System sought to integrate individual components into a synergistic weapon system, making sure that sensors, comms and weapons were compatible in terms of such interoperability considerations as power requirements and ergonomics. (Haley 1991; O'Keefe, MIddleton et al. 1992; Wojcik 1996; Middleton, Sutton et al. 2000; Lockhart 2006)

2.3.2 Modeling the Soldier as a System and the Tactical Small Unit

One of the first steps in defining an integrated soldier system was identifying its individual components, the so-called soldier "illities": lethality, survivability, mobility, command and control, and sustainability. These are shown are often described as members of a hierarchy or branches of a tree, as shown in Figure 2-2.

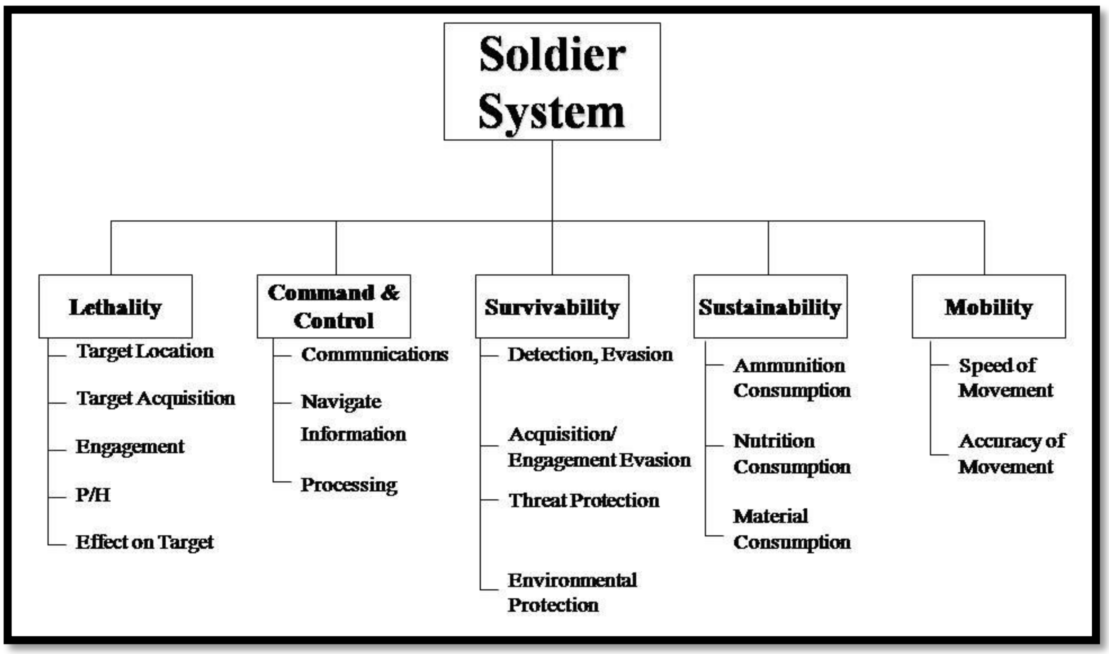


Figure 2-2 Soldier System Hierarchy circa 1991

Modeling the Soldier as a System therefore meant modeling these components and the interplay between them. The US Army's Natick Research Development and Engineering Center (NRDEC) developed Integrated Unit Simulation System (IUSS) for this purpose in the early 90s. (Middleton 1991) The IUSS initially focused on trying to measure trade-offs between operability and survivability, which required modeling soldiers' psycho-physiological states and their ability to perform discrete mission tasks. The IUSS was a task-network model, i.e., one in which entities were constrained to follow a specific sequence of tasks as part of each mission.

As shown in Figure 2-3, tasks were defined as scripted sets of behaviors for a given entity. These entities in the IUSS were usually tactical small units (TSUs): platoons, squads or fire teams, reflecting the fact that individual soldiers virtually never function in isolation, but as coordinated elements of larger forces. Accordingly, mission capability needs to be examined in the context of individuals operating within the context of their unit; their functionality augments and is augmented by other individuals, each providing additional capabilities (and technology).

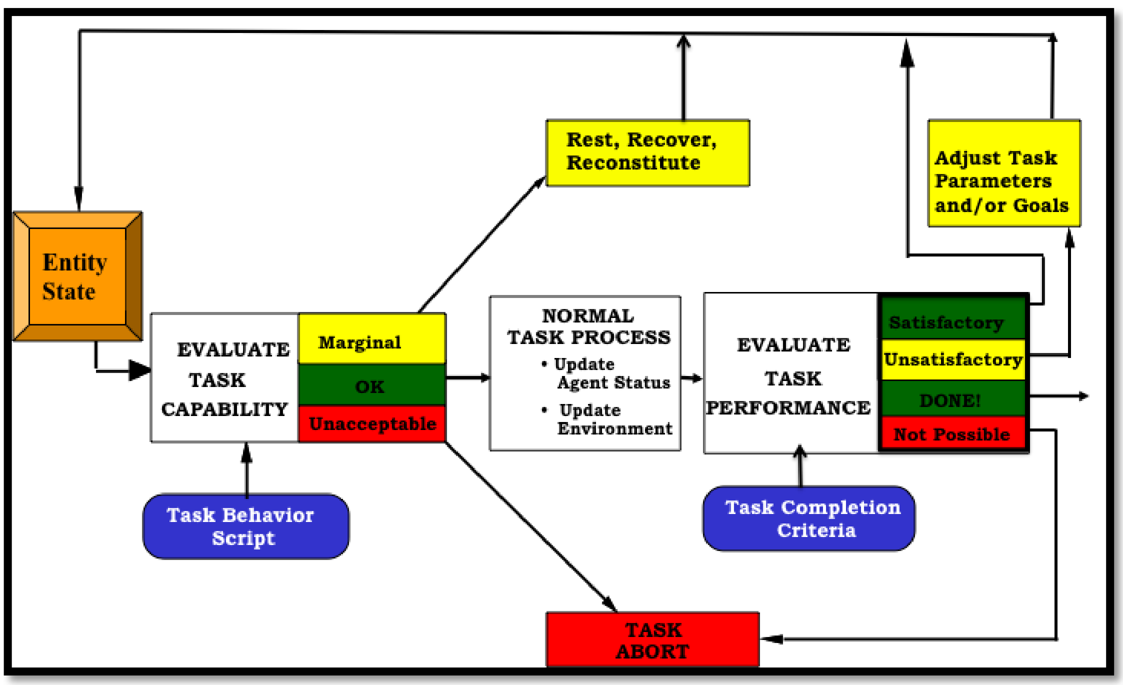


Figure 2-3 Task Performance Evaluation Cycle

Figure 2-4 shows the systems that compose the TSU. A typical TSU is structured around a base tactical element of 7 to 18 fighters organized to perform a fundamental tactical mission or task in a definable area of responsibility (ADR). The TSU can and does encompass several different forms and refers to both friendly and adversary forces; well defined infantry and special forces units, ad hoc reaction teams involved in rear and vital area security, irregular forces, police units and loosely organized street gangs. It is the smallest element that can be assigned to perform basic infantry tasks: patrol; defend or attack a specific tactical objective: provide local security, or perform area or point surveillance.

The IUSS incorporated the Soldier System Hierarchy as part of its breakdown of tasks into different behavioral components. Task progress was largely determined by stochastic draws from performance distributions that were a function of entity capability and the nature of those task behavior components.

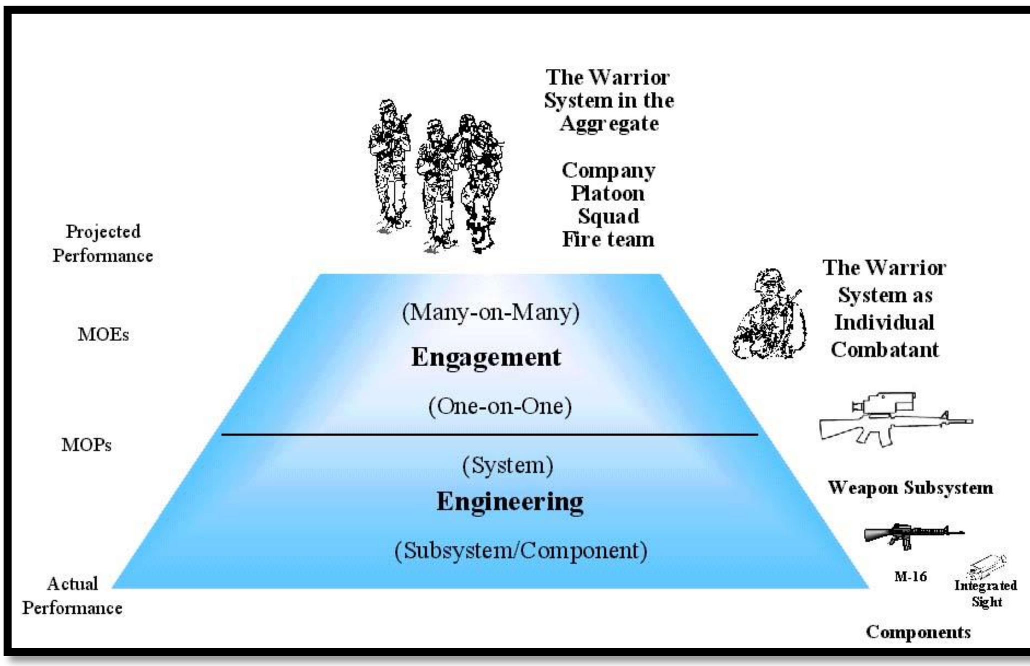


Figure 2-4 The Tactical Small Unit System

In theory, as shown in Figure 2-3, dynamic event response possibilities included aborting the task, adjusting either the completion criteria or the way the task is performed (by adjusting task parameters) or both, or changing, enhancing or renewing entity status to improve performance. In practice changing task completion criteria, in effect changing the task goals, required implementation of a new task sequence altogether, which generally also required new task performance distributions. The ability to react to the behaviors of other entities was primarily limited to the conditions under which those entities could be engaged and the methods of engagement that could be employed.

While the IUSS task structure was limited as a software construct, it did illustrate a core capability that would eventually be more fully realized by its agent-based successor IWARS. Entities in agent-based models must continually assess progress towards currently active goal(s), and respond appropriately whenever that progress appears unsatisfactory.

2.3.3 Human Performance Moderators

Representation of the individual combatant's psycho-physiological state was a key aspect of the IUSS. Factors used to represent behavioral outcomes include task performance time(s), a probability of "successful" task performance, and task accuracy or other measures of "goodness" - how well the task was performed. Human performance moderators were used to reflect computable effects of dynamic entity states on measures of task accomplishment, thus allowing individuals to become adversely affected by fatigue, heat stress, injury or encumbrance of body armor or other protective gear. These moderators can be represented in task network models by adding additional factors to performance distributions and/or modifying the results of draws from those distributions. It is, of course, also possible to use moderators to represent performance enhancers such as rest, rehydration and other recuperative activities. The assumptions implicit in this approach are those of most task network models, that military operations can be modeled as a series of interconnected tasks, that these tasks can in some sense reflect an atomic decomposition of mission performance, and most importantly, that there is adequate definition of the key factors that impact task performance and some reasonable characterization of the distribution of task measure outcomes in terms of these factors.

The use of task networks with moderators is very attractive approach in a server-based distributed simulation environment, because it reduces the consideration of the effects of human behavior on simulation outcomes to a relatively small set of numerical adjustments to simulation performance parameters.

At some levels of aggregation this approach is acceptable, but it does not adequately represent many facets of human behavior such as the "soft factors" discussed above. In particular, it does not represent the highly dynamic reaction of human decisions and behavioral responses to a complex, rapidly changing, battlespace. Task networks and aggregate performance measures are not well suited to situations with multiple competing mission objectives, which can require constant goal conflict resolution and complex behavioral trade-offs. In general aggregate task performance approaches are increasingly contra-indicated in situations where tasks must be viewed less as atomic mission chunks and more as complex feedback loops of human interaction with weapon systems, equipment platforms, other individuals/units, and/or a myriad of other battlefield elements.

2.3.4 The Warrior System

As the Department of Defense began development of an increasing series of new capabilities for the individual dismounted combatant, the Soldier as a System morphed into the more encompassing "Warrior System" [22], which seeks to take advantage of the "digitization of the battlefield" by equipping the individual infantryman with the same kind of information technologies already taken for granted by pilots and tankers. These technologies include heads-up-displays that show the position of friendly forces, militarily relevant aspects of terrain, as well as the best available intelligence on the location of suspected threats and adversaries.

All told, these capabilities have significantly increased the ability of the individual to perceive the environment, to share this perception with others, and to act in concert

[22] See for example (Middleton 2001) FROM RIFLEMAN TO WARRIOR SYSTEM: The Evolution of the Dismounted Close Combatant; SMi Conference on Dismounted Close Combat; London, England

within the TSU to potentially control the battle space far more than ever before.

The warrior system concept has been widely accepted internationally[23], and has been through a number of iterations in the US, starting with the Soldier Integrated Protective Ensemble (SIPE), a three year program begun in 1990 and culminating in an Advanced Technology Demonstration (ATD) (Middleton, Sutton et al. 2000). SIPE was followed by several competing and complementary efforts: The Enhanced Integrated Soldier System (TEISS), the 21st Century Land Warrior (21CLW), the Generation II Soldier (GEN II), all of which contributed in some way to the currently fielded Land Warrior (LW), and the current program for the Future Force Warrior (FFW), now known as the Ground Soldier System (GSS). These systems differed far less in concept than in cost and bureaucratic programmatic issues. All of them sought (or still seek) to end the stove-piping of different systems and sub-systems for individual, providing far more integrated functionality, and, as mentioned above, recognizing the need to treat the individual soldier as a complex weapon system in his/her own right.

2.3.5 Warrior System MS&A Requirements

The warrior system represents the classic "paradigm shift" in the operational view of the individual dismounted combatant, and this paradigm shift calls for a similar shift in modeling and simulation with respect to that individual. The analytical paradigm shift has two large underlying requirements. First, models of individual and small unit behaviors can no longer be satisfied with the "small tank" approach. Models for dismounted soldiers have to represent the world in far greater detail, with higher fidelity than do models of large weapons platforms. The dismounted individual has a far more intimate view of the battlefield than the tanker or pilot: one that cares about terrain at meter resolutions and below; one that cares about threats that can hide

[23] For example, (Housson 2008) discusses programs by the British: FIST (Future Integrated Soldier Technology), Germans: IdZ (Infanterist der Zukunft), Spanish: COMFUT (COMbatiente FUTuro), French: FELIN (Fantassin à Équipements et Liaisons Intégrés), and Italians: Soldato Futuro. See also (Leeuw 1997; HassgÂrd 2002), (Curtis 2000; Hobbs, Goyne et al. 2000), (Underhill 2009) for Dutch, Swedish, Australian, and Canadian examples and perspective.

behind trees and bushes; one that cares about adversaries that can blend into crowds or the local landscape, and so forth. To support an adequate picture of today's digital battlefield for the dismounted soldier or marine, information technologies need to provide an accurate picture of the battlefield a level of resolution commensurate with the environmental factors of interest to that individual. As a consequent, the models and simulations used to evaluate these technologies have to support algorithms and data structures at associated levels of resolution and fidelity. Furthermore, they have to generate the data needed to populate those data structures and drive those algorithms, making them far more data "hungry" than the aggregate models of the Cold War era.

Second, models and simulations must represent how individuals and their units use this mass of data. To adequately explore the operational value of information technologies, analytic tools must now represent the individual combatant not as characterized solely by more or less scripted actions (move, shoot, communicate), but as a decision-making entity, continually evaluating the battlefield dynamic to decide when, where, and how to move, shoot, and communicate. The raw materials for decision-making are data and information; adequate analysis of warrior systems thus requires representing the acquisition, filtration, and fusion of data at levels of detail and significance relevant to the individual dismounted combatant.

2.4 Towards an Engineering Model of the Warrior System?

The decision to treat the individual dismounted combatant as a weapon system also implies a concomitant desire to employ the same kinds of engineering and analytical tools and procedures for the warrior system as for larger, more mechanical weapon systems. Good engineering models exist for large weapons systems, and these models support trade-offs in the design and development of the functional capabilities for these systems.

2.4.1 Engineering Models

Engineering models that describe and predict the operation of systems and subsystems are essential to the development and fielding of today's complex military machines. As mentioned above, such models are generally deterministic and based on Newtonian physics, i.e., systems whose exchanges of mass and/or energy with their environment are constrained to a relatively few, well-known, factors. These models may incorporate stochastic treatment of systems performance, as based on statistical data from measurement of well-defined systems' functions. Their model parameters span the analytically relevant/interesting areas of the problem space, and there is essentially a one-to-one mapping between model features and systems' functions. These features support model verification and validation based on theoretical concepts, and are supported by empirical data on operators/systems' performance.

A design approach that looks to develop materiel for the dismounted combatant by applying the same systems analysis techniques used for tanks, planes and ships faces a fundamental problem: "There is no engineering model of the warrior system".

Certainly, there have been attempts to apply the reductionist modeling approach to break down system capabilities into separate and distinct modules, as evidenced by the Soldier System hierarchy incorporated into the IUSS. The problem with such an approach, however, is that it ignores the complex interactions and inter-dependencies of the hierarchy elements. Exploration of these interactions, and indeed, of the capability areas themselves, has been the subject of considerable research over the last two decades. Much of this research has been focused on cognitive functions, information processing, and psychological aspects of personnel interactions (Haley 1991; Zachary 1998; Banks and Stytz 1999; Heinze 1999; Hill 1999; Curtis 2000; Heeringa and Cohen 2000; Forsythe 2002; Ghazal, Morley et al. 2003; Gratch and Marsella 2003; Pfitzner, Hobbs et al. 2003; Petty, McKenzie et al. 2004; Silverman 2004; Dyer, Wampler et al. 2005; Gluck and Pew 2005; Boylan and Goerger 2006; Betz 2007; Dyer 2009; Underhill 2009) and in fact has led to expanding the C^2 illity, the command and control capability, into C^4ISR: Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance. In addition, many now consider Situation Awareness to be a sixth “illity”, albeit one closely aligned with C^4ISR (Middleton 1999).

2.4.2 Replacing the Soldier System Hierarchy with the Warrior Systems Architecture

The Technology Base Steering Committee’s[24] Warrior Systems Modeling and Simulation Working Group came up with a less structured Warrior Systems Architecture as a replacement for the Soldier Systems Hierarchy. As shown in Figure 2-5, the capability relationship is depicted as a sort of “yin and yang synergistic octopus” to better represent the nature of the interplay between the “illities”.

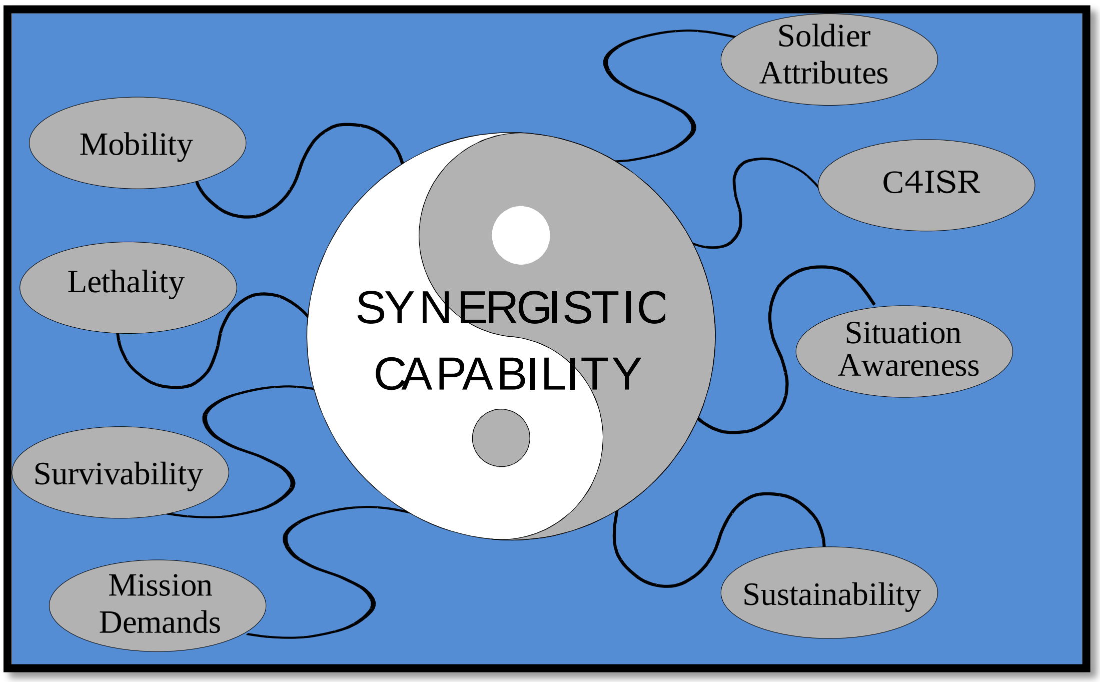


Figure 2-5 The Warrior Systems Architecture as a Synergistic “illity” Octopus (Middleton 1999)

The Working Group was attempting to capture the natural dualities of complementary

---

[24] The Warrior Systems TBESC Technology Based Executive Steering Committee was chartered in December 1997 as an advocate for individual warrior Science and Technology. It’s members include J-8 (Force Structure, Resources, and Assessment Directorate, the Joint Staff US DoD), DDR&E (Director of Defense Research and Engineering US DoD), ARO (US Army Research Office) ARL (US Army Research Laboratories), DARPA (the US Defense Advanced Research Projects Agency), STRICOM (US Army Simulation Training and Instrumentation Command), TSM Soldier (TRADOC System Manger –Soldier, now TRADOC Capability Manager Soldier TCM - Soldier, and The US Army RDEC’s (Research, Development, and Engineering Commands)

opposites within a greater whole. For example, there is always tension between mobility and protection – do I want to try to avoid the enemy bullet, or defeat it when it hits me? Ideally I'd like to do both, but the more effective my armor, the more it weighs me down and slows me down, making me easier to hit. C4ISR gear can warn me and allow me to outmaneuver my adversary; does this mean I can afford to reduce my armor? Is the best protection under some circumstances no protection? The version shown here has been updated[25] with the additional factors of basic soldier attributes (psychological factors such as training, leadership, morale, cultural identity; and physical state factors such as fatigue and heat stress), and mission demands (e.g.; combat intensity, criticality of tasks, and rules of engagement).

The Warrior Systems architecture cross-walks warrior functions to system capabilities and, ultimately, to quantifiable measures of performance and effectiveness (MOPs&MOEs). It can be used to apply the principles of multi-criteria decision analysis and explore the integration of these measures into meaningful tools to assess Warrior System alternatives, culminating in a draft weighted, multi-criteria utility function (or decision aid) to quantify warrior system benefits. It provides a frame of reference and common terminology for warrior systems RD&A. It and the doctrinal publications from which it is derived: the US Army Field Manuals, the Marine Corps Warfighting Publications, and the Marine Corps Doctrinal Publications[26] provide the orthodox definition of the "move, shoot, and communicate" capabilities and behaviors inherent in basic military tasks, and which must be incorporated in any M&S paradigm for Warrior Systems analysis.

As the illity octopus demonstrates, however, the complexities of the warrior system cannot be entirely captured by the strict reductionist approach of orthodox systems analysis.   Accounting for the dynamic and highly non-linear interactions of the cognitive

---

[25] Both because we thought it was important to include these additional elements and because an octopus has eight arms.

[26] In particular, (USAIS 1992; TRADOC 1993; USAIS 1993; TRADOC 1997) (MCCDC 1999) (MCDP6 1996; MCDP1 1997; MCDP5 1997)

and physiological elements that constitute the warrior system requires a more open systems approach, as is discussed in detail below. These complexities are exacerbated further by the nature of current conflicts, especially with regard to irregular warfare and asymmetric combat, in which the interactions between friendly forces, adversaries, and neutrals form a seemingly chaotic dynamic landscape.

How, then, does one develop an engineering model of such a system? An answer lies in upgrading the concept of "engineering" models to include models that allow exploration of virtual systems whose behaviors emerge from general rules of operation, and are not limited to functional capabilities that can be reduced to physics-based algorithms. This concept means incorporating principles from psychology and the social sciences into orthodox physics-based model. It does not mean eliminating the use of physics or the other "hard" sciences, it simply means extending the reductionist approach to support a wider variety of system decompositions. It means, for example, decomposing warrior systems operations into sets of entity or object interactions as is in done in agent based models, and modeling these operations as elements of complex adaptive systems.

2.4.3 The Human-Centric Paradigm: A Marriage between the Physical Sciences, Psychology and the Social Sciences?

The military OR community is increasingly looking to the social sciences to provide answers in the world of insurgencies and asymmetric warfare. (Lucas, Sanchez et al. 2007; Seitz 2008; Alt, Jackson et al. 2009; Pfautz and Toman 2010) The United States Air Force commissioned a study on individual, organizational, and societal (IOS) modeling research programs (Zacharias, MacMillan et al. 2008), which emphasized the multi-disciplinary nature of such models. Sciences from economics to environmental studies (Davis 2007), (Hare and Deadman 2004), have come to view agent-based modeling as an important tool; one which has been widely embraced by the social sciences as way to come to grips with the complexities of human and social behavior. (Axtell 2000), (Macy and Willer 2002), (Bonabeau 2002), (Gilbert and Terna 2000), (Smith and Conrey 2007).

Major acquisition programs within the defense community have increased the use of experimental psychology and human factors engineering and this increase has facilitated and accelerated the incorporation of these factors in some models and simulations (Ramirez 1997; Pew and Mavor 1998; Mastroianni and Middleton 2001; Martin 2005; Warwick 2006).

Aspects of psychology such as human factors that have been accepted in the engineering realm are, however, considerably different from factors such as “will to fight”. The “soft factors” are very different from the rigorous methodologies of most combat models and simulations. Many of these factors are not discussed in the language of science and engineering, as are experimental psychology and human factors engineering, but instead in the vocabulary of history, political science, and even philosophy. In addition, they lack the first principles models of cause and effect that are the foundation of “validated”, physics-based models and simulations. Using agent-based models to address these factors does not mean doing away with the concept of scientific rigor, but there is a need to extend that concept to incorporate a “soft”, incremental focus, where increasing levels of correlation correspond to increased acceptance of predictive validity, and where inconsistency can be accepted as evidence of uncertainty rather than outright error.[27]

Incorporating such factors in combat models and simulations will require contributions from many fields. As former Secretary of the Air Force James Roche states:

“*The original ops researchers understood that to be effective they needed teams of mathematicians, historians, military theorists, psychologists and economists, among others… Somewhere along the line, this was lost as a fundamental concept of military operations analysis…There is a real and significant role for ops research and systems analysis in the campaign in which we are engaged. This campaign also requires knowledge of history,*

[27] This “soft” concept of rigor is evinced in the view of model validation presented in my March 2008 Phalanx article: *“Emergent Validation”: Defensible and Authoritative Modeling, Simulation and Analysis for Complex Adaptive Systems*

economics, religion, finance, psychology, technology, game theory and decision analysis, among many other disciplines." (Roche, 2002)

In particular the human-centric paradigm encompasses the need to assess the human as an information-processing and decision-making entity. This entity is encased in a psycho-physiological substrate, which is shaped by the stressors and other conditions of the battlefield environment, and in turn shapes that environment, or at least aspects thereof.

The human-centric paradigm is motivated by warrior system concepts where effective operations are dependent on efficient command and control functions, including communication, individual and shared SA/SU, and coordinated action. The need to assess current and evolving information technologies, and their application to the acquisition, flow, and use of data and information at the small unit level, drives the need to model the mental state of individuals, and to represent individuals' (and especially commanders') decision processes. Aspects of "behavior" that must be represented include the processing of sensory inputs and the interpretation and integration of sense data into meaningful information about the environment and the operational situation. Such representation encompasses much more than physical behaviors: internal knowledge states and cognitive processes (or their effects); decisions such as target selection and route adjustment, and even non-deliberative processes such as modification of speed and/or posture in response to environmental cues and one's own psycho-physiological state.

2.4.4 From Closed Systems to Open Systems

I suggest that to take development of an engineering model of the warrior system out of the "too hard to do" box, closed systems modeling approaches need to augmented with an open systems modeling methodology that incorporates the "soft" sciences. Such a methodology reflects the fact that human behavior in general and military conflict in particular represent open systems, ones in which energy, materiel, and information flow across the boundaries of the physical, information, and cognitive domains.

The shortcomings of the more orthodox closed system methods have been long recognized. In 1994, the Military Operations Research Society (MORS) had a mini symposium, MORIMOC III, to address “Human Behavior and Performance as Essential Ingredients in Realistic Modeling of Combat”. In that workshop, they charged a working group to:

> *Develop a conceptual approach to incorporating higher level behavioral factors in modeling/analysis to support decision issues. These might include leadership, unit cohesion, communication, battle experience and morale.*
>
> *Consider:*
>
> *Combat models in which opposing forces are composed of multiple units and group phenomena are dominant.*
>
> *Combat models driven by military intention and capabilities rather than by force ratios or attrition.*
>
> *Development of the needed group performance data bases. (Murtaugh 1994)*

The Military Operations Research Society, among others continued to pursue this thread. As alluded to above, in writing for the Society’s Phalanx magazine in 2002 Vincent Roske spoke of the need for new tools to address the class of “open systems” not accessible using traditional operations research tools:

> *“It is as much convenience as perhaps unfortunate fate for us analysts that the industrial revolution bequeathed us an arrogant sense that the world can be thought of as assemblies of closed systems and that we can discern, define, predict behavior, and control those systems. From steam engines to assembly lines to electric power grids we have come to believe that we are in control of our world; that if we push in "here" we can predict what will bulge out over "there." To the extent that we are in control of a system's configuration and behavior, closed system analysis methodologies work well. But what happens when the problems we are analyzing are processes and environments we do not control?*
>
> *….*
>
> *The presence of the human being introduces energy across the systems boundary and produces emergent and adaptive behaviors from the system. This is characteristic of open systems, particularly of complex adaptive systems.” (Roske Jr 2002)*

2.5 <u>MOPs and MOEs</u>

Operations research depends on the use of metrics to assess fitness and /or degrees to which goals may be attained. At the individual and small unit level of simulation these metrics all into two primary categories: Measures of Performance (MOPs) that are used to assess the functioning of individual system elements, and Measures of Effectiveness (MOEs) that are used to determine the way those elements affect behavioral and/or operational outcomes for the total system.

The US Army in DA PAM 5-11 defines:

- Measure of Effectiveness (MOE): A quantitative expression which compares the effectiveness of alternatives in meeting an operational objective or need.
- Measure of Performance (MOP): A defined metric of a component which contributes to basic system effectiveness as described by an MOE. MOPs relate to specific performance characteristics from which data can actually be collected.

From the point of view of this research, MOPs are basically simulation inputs, MOEs simulation outputs - models turn MOPs into MOEs.

2.5.1 The Role of MOPs and MOEs

Picking the "right" MOPs and MOEs is critical to sound analysis, but it is not always straightforward or easy.

An old story tells of interviewing the football coach whose team was dominated by their opponent throughout most of the game, but who manages to win regardless:

"Coach, how do you feel? Your team had less than half the yardage gained of your opponents, only four first downs to their 23, was dominated in time of possession, and drew twice as many penalties."

The coach replied, "The way we play the game today, you win by scoring more points. If you want to change the rules, and declare the team with the most first downs or the most yardage or fewest penalties the winner, then we'll play for those stats. Until then, I'm going to keep trying to win on the scoreboard."

This story has a couple of implications for the practice of MS&A. First, the measures that models use to quantify performance and operational effectiveness both reflect and shape how operations are viewed. How the military (and the model) "keeps score" has significant influence on both strategy and tactics. It provides the basis for acquisition decisions and R&D investment strategy in addition to shaping policy, plans and tactics.

Next, as in the above anecdote, the measures of performance analysts think are important turn out not to be, at least in some instances[28]. So too in military affairs, the things that conventional wisdom says should determine the outcome aren't always the ones that do. At the end of the movie Midway, which portrays the US Navy's victory in a pivotal battle of World War II, one American remarks "It doesn't make any sense … Yamamoto had everything going for him, power, experience, confidence... Were we better than the Japanese, or just luckier?"

In the world of military analysis, just as in sports, prediction is a chancy business. Why is it that the factors that "ought" to predict outcomes, all too frequently don't? In the first place, of course, in any one battle, such as Midway, the answer might be luck. Luck may also contribute to the start of "genetic drift", to the complex vagaries of any individual battle where any one of a myriad of factors can be the deciding element.

For want of a nail the shoe was lost.

For want of a shoe the horse was lost.

For want of a horse the rider was lost.

For want of a rider the battle was lost.

For want of a battle the kingdom was lost.

[28] Perhaps the most notable recent example of the failure of expert judgment based on conventional measures is the experience of the Oakland Athletics as related in the film Moneyball (Armstrong 2012).

And all for the want of a horseshoe nail.[29]

Analysts (and sports writers) can take some solace in the thought that while individual outcomes are hard to predict, averages of aggregate outcomes and long-term trends are far more amenable to assessment and forecasting. As Damon Runyon[30] once remarked, "the race goes not always to the swift, nor the contest to the strong, but that's the way to bet".

The function of MOPs and MOEs is to quantify important problem inputs and outputs. They can support analysis to both find the isolated elements that can lead to critical failure, and to discover the "way to bet" - to find the optimal appropriation of resources for the development and application of technological solutions where such solutions are appropriate.

2.5.2 Current State of the Art

Unfortunately, the things that can be measured easily with orthodox MOPs and MOEs don't capture critical intangibles: morale, leadership, unit cohesiveness and the like. Furthermore, the derivation of MOEs from pertinent MOPs is not generally a straightforward, linear process; the whole is not just a simple sum of the parts. As with the nail, a single critical failure can overwhelm a multitude of successful contributions. On the other hand, the ability of individuals to compensate for faults and problems can counterbalance a multitude of deficiencies. Collective synergy is a key aspect of viewing military operations as complex adaptive systems, and adequate measurement of component contributions and subsequent estimation of overall systems performance are keys to an agent-based modeling approach.

The preferred source of data for MOPs is physical measurements in a laboratory or field setting that describe the engineering level assessment of individual system

[29] Ilanchanski (2004) also uses this nursery rhyme to illustrate the non-linear nature of combat outcomes

[30] American novelist and newspaper writer b.1880, d.1946, best remembered as the creator of "Runyonesque" characters

components or subsystems. Of course, even under laboratory conditions, there are factors that contribute both random noise and systemic sources of variation; summarizing results as statistical distributions allows us to capture this variation. Incorporating these distributions as simulation input parameters then further allows us to account that variation it in the prediction of system performance; the use of such distributions is the basis for Monte Carlo simulation.

Simulation provides assessments of systems performance through representation of system-on-system engagements and larger force-on-force encounters, evaluating mission and battle outcomes, or even theater level concerns, as appropriate. To proceed from MOPs to operational-level measures of effectiveness it is almost always necessary to resort to simulation, whether done in the field with troops engaged in simulated combat, or in a computer, where mathematical models predict the outcomes of mission tasks. While system or component evaluators may prefer the results of field trials, economic realities constrain them to fight the majority of their test battles in the computer.

Just as it makes no sense to speak of experimental data without reference to the tools and metrics used to collect them, MOEs cannot reasonably be separated from their derivative models and simulations. These M&S tools determine the context, resolution, reliability, and applicability of their associated MOEs. As an obvious corollary, M&S tools that do not incorporate the explicit application of information technologies in the representation of SA/SU and decision-making must at the very least be suspect measurement of operational outcomes that depend on these factors.

Of particular significance to the warrior system community and other advocates of the individual dismounted combatant, is the use of MOEs to assess the result of small unit operations. These typically involve force-on-force encounters at the squad and platoon level, and provide the greatest insight into the combat worth of soldier system components or concepts.

In such instances, the term MOP refers to the output of data reduction as applied to the results of experiments or field tests. For example, a field test may produce start and stop times for an individual corresponding to distance covered while equipped with some set of soldier system component alternatives. This translates directly into an MOP, speed. If measurements can be obtained relating the individual's physiological state to speed and other measures of terrain trafficability, mathematical models can be constructed to predict the individual's level of fatigue and percent effectiveness at various mission tasks.

In the case where experimental and field trial data are not available, the analyst can usually obtain subject matter expert (SME) estimates of appropriate MOP values for performance of specific task elements under given conditions. Such values, in combination with whatever empirical data as may be available are often applied in one of the analyst's key tools, parametric variation simulation experiments. Fortunately parametric variation is well suited to the emergent analysis paradigm, and there is in fact a growing body of research on appropriate experimental designs that support efficient exploration of the parameter space through data farming and data mining with computer simulation. (Ipekci 2002; Kleijnen, Sanchez et al. 2005; Liang 2005; Alt 2006; Lucas, Sanchez et al. 2007; Foo 2008; Pearman, Middleton et al. 2009)

There are a number of historical measures that address casualties and other attrition combat outcomes: measures of accuracy and lethality to quantify the contribution of weapon systems to those outcomes; measures of sensitivity and associated probabilities of detection, recognition, and identification to represent the contribution of sensors and target acquisition technologies; measures of durability, fit, protective camouflage, and survivability to represent the contribution of clothing and body armor; and so forth.

Current human factors engineering measures address aspects of human interaction with information systems and subsystems, e.g.; rate of signal processing and other perceptual interface questions; what is lacking is metrics to address the *usability* of data and information. Needed are measures of item/system performance to quantify the resource utilization trade-offs between acquiring data, and filtering, integrating, and understanding them to provide actionable information.

MOPs exist for:

- sensor performance;
- communication network efficiency;
- Human factors engineering of display technologies;
- SAGAT[31]-type quantification of the deltas between perception and ground truth.

Measures as yet don't exist for:

- measures of the "goodness" of data - "quanta" of information[32];
- Time-dependent measures of the utility and fragility of data.

Extant MOEs for conventional combat outcomes include:

- Combat outcomes, e.g; killer victim scoreboard;
- Fire power, e.g; rounds on target;
- Mobility, e.g.; speed attainable, risk of encountering hazards;
- Mission/goal achievement, e.g.; time, %complete.

These metrics don't provide:

- Measures of lost opportunity costs;
- MOEs to quantify the effectiveness of close contact instead of close combat;
- measures of the contributions of situation awareness to greater ability to react to dynamic conditions;

---

[31] Situational Awareness Global Assessment Technique, see Endsley et. al. (1999)

[32] The DoD CCRP Code of Best Practice for Experimentation (Alberts, 2003) suggests 5 measures for information and knowledge: completeness; correctness; currency; consistency; and precision; these may be valid MOPs but they require adequate comparison with ground truth data and further definition of operational context to be meaningful

- Time-dependent measures of the utility and fragility of information.

As may be inferred from this latter list, established metrics of attrition combat effectiveness do not adequately address many aspects of today's extended military operations and especially the effects of information technologies that are the underlying motivation for this dissertation.

As a final note about the current state of the art, while there is certainly need for new MOPs and MOEs, this need does not preclude the usefulness of the older, more orthodox ones, even if they may not support prediction of the outcome of a given battle. In many cases, these older, limited measures remain the best (and/or only) way to assess and improve capabilities and readiness. Then too, it may be the case that the right measures are chosen, but the wrong connections are made between them. The number of first downs the team got may not reflect the outcome of the game; the fact that they couldn't get a first down on their last possession and run out the clock did, because it allowed their opponent the opportunity to make the winning score.

2.5.3 Potential for New MOEs[33]

In traditional attrition-based metrics, the direction and magnitude of desired outcomes is usually unambiguous. It is always better to reduce friendly casualties and increase adversary casualties and destroy more of the adversary's materiel. In current asymmetric conflicts however, it is entirely possible that the direction of change in an MOE may be meaningful only when contextualized by other information or other metrics. For example, a numerical measure of the number of contraband items confiscated in a certain area may not be especially useful when considered alone. A decreasing trend in the number of confiscated items may be indicative of an improving

[33] Much of the material in Chapter 2.5.3 and Chapter4 is the result of collaboration with colleagues over the past several years, most notably Dr. George Mastroianni of the US Air Force Academy, and Ellen N. van Son–de Waard of the TNO Prins Maurits Laboratorium, Rijswijk, The Netherlands

security climate (a positive outcome) or of more effective concealment and evasion (a negative outcome). Correlated measures, perhaps including assessment of confiscation rates in other areas, or of other metrics, may be needed to fully assess the meaning of such measures. Developing metrics that will be useful under such conditions is an immense challenge.

Established attrition metrics are not well suited to missions where combatant status is ambiguous and it may be desirable to avoid casualties to potential adversaries as well as one's own troops. Close "contact" operations such as stability and peacekeeping, for example, share many characteristics with civilian police activities where the use of deadly force is a last resort, and often indicates a failure to keep a situation under adequate control. One potentially fruitful approach might include expanding measures of physical destruction such as lethality and incapacitation to encompass a more comprehensive concept, that of a hierarchy of control. Clearly, the physical destruction of objects, personnel, or places is the most emphatic kind of control, and this is generally the objective of straightforward combat operations. There are, however, other contexts for the application of military power. These include the application of combat (or the threat of the application of that power) to neutralize a real or potential enemy threat, or simply to influence the behavior or potential of adversaries or other participants in the operational environment, may serve as well or better to accomplish the mission. Several qualitative designations of decreasing levels of "control" are shown in Figure 2-6.

These levels can be used to capture such factors as a commander's intent, and to define "close contact" MOEs by relating them to such factors as: size, proximity, maneuverability, and mobility of forces; ability for force projection; situation awareness; force readiness and reaction time. They still, however, boil down to the common sense concept of the ability to have the right force, in the right place, at the right time. To this end, the ability to employ information to optimize movement is a critical contributor to over-all effectiveness, and thus simulation of the appropriate application of dynamic data to movement behaviors is a good starting place to begin exploration of the potential for such metrics.

Attrit	Neutralize	Influence	Monitor
destroy	suppress	restrict	detect
kill	interdict	inhibit/delay	observe
damage	detain	restrain	check
incapacitate	surround	curb	regulate
reduce	contain	lead	inspect

Figure 2-6 Types of Control.

Furthermore, these levels of control now provide greater resolution in addressing the application of that force; success may be defined as much through adroit management of force and force projection as through simply overwhelming an adversary. Here again appropriate adjustment of one's understanding of the dynamic situation is a key.

I believe it is also important to incorporate the concept of opportunity cost into these metrics. The phenomenon of "If I knew then what I know now" is a key to assessment of SA/SU; extinguishing matches requires a good deal less effort than putting out forest fires.

The human-centric paradigm supports assessment of intermediate levels of operational effectiveness in terms these levels of control, using them to characterize the commander's intent with respect to specific mission tasks, evaluating task performance in terms of objective task measures and the extent to which that intent is met. This kind of assessment suggests a hierarchy of measures for:

- “Standard” task performance - behavior according the “script”, e.g.; execution of practiced tactics, techniques and procedures (TTPs) in accordance with commanders’ operational plans.
- Reactive behavior - the ability to flexibly react to operational events.
- Pro-active behavior - the ability to anticipate and effect changes in the operational environment.
- Mission achievement - the extent to which mission objectives and higher commanders' intents are met.

Finally Figure 2-6 almost certainly needs to be expanded to consider the beneficial use of forces, adding a column or columns to address aid and comfort, support and sustainment, construction and development.

3 HUMAN-CENTRIC MODELING TOOLS: AGENT BASED MODELS AND COMPLEX ADAPTIVE SYSTEMS

This chapter provides a brief overview of complex adaptive systems (CAS) and agent-based modeling (ABM) as the tools to implement the human-centric approach, and concludes with the concept of using data farming and data mining with simulation experiments to perform "emergent" analysis, as a way to apply the that approach to the solution of relevant problems. Chapter 3.1 discusses the need for new measures to quantify the key elements of human centric approach and in Chapter 3.2, all of these elements are merged into a synopsis of the human-centric approach to analysis.

The MORIMOC III working group discussed in the last chapter suggested an Object-Oriented Design in which objects (military forces) are possessed of physical and cognitive states, which they change, based on interaction with each other. The working group was primarily concerned with objects that would represent higher echelon forces (battalions, companies and divisions), but the same approach can be applied to represent sub-units, individual weapon platforms or individual dismounted combatants. Furthermore, these recommendations look a lot like suggesting adopting agent-based modeling as an approach that supports exploration of human factors in general and decision-making in particular.

The theory of complex adaptive systems (CAS) provides the philosophical foundation for ABM. It supports augmenting deterministic engineering models, in which "buttons are pushed" and predictable results follow, with models in which autonomous entities interact with their environment and each other in continually evolving adaptations and that are frequently characterized by unexpected ‘emergent” behavior.

3.1 Agent-Based Models

Agent-based models have their foundation in the theory of CAS, and focus on relatively simple interaction rules to achieve complexity. These rules direct how simulation agents deal with each other and their environment. They need only be internally consistent; they do not require a single unified theory of human behavior. In fact, this approach can embrace multiple rule sets, some of which may be based on competing, and even contradictory, social science theories and data[34]. This multi-disciplinary approach also supports representation of individuals with widely diverging belief systems and standards of behavior, a virtual necessity in accommodating the clash of cultures that characterizes the human dimension in much of today’s military operations.

This dissertation is concerned with the specific agent species *Agent Sapiens*, “intelligent” agents, entities that perceive the condition of their environment and select their behaviors to effect changes to that environment. Intelligent agents are[35]:

- goal-oriented - able to build courses of action by taking the initiative to change elements of the world state to desired objectives
- perceptive - able to receive data from their environment, including knowledge of their own state and that of other entities of interest to them,
- active - able to perform actions affecting their environment, and
- autonomous - able to initiate behavior sequences based on internal logic to determine what is appropriate given the perceived environment.

Agents representing combat forces must also generally be:

---

[34] This feature of the approach is one of the central tenets of this dissertation. Its application is the basis of the wide variety of simulation experiments possible with the methodology developed herein, as will be evidenced by the experiment options discussed in Chapter 4 below.

[35] See for example (Wooldridge and Jennings 1995a; Wooldridge and Jennings 1995b; Turner 1998; Ioerger 2000; Wooldridge and Ciancarini 2001; Russell 2003; Hare and Deadman 2004; Jones, Stensrud et al. 2009)

- mobile - able to move around in their simulated environment,
- capable of inferring the intentions of others, the desires and plans of other agents, and
- social - able to share goals, cooperate with or coerce other agents.

A key distinction between agents that are "intelligent" and those that are merely reactive is the concept of maintaining a perceived view of the environment, having "knowledge" of the world based on current and historical data from the agent's sensory input capabilities. Intelligent agents are not omniscient, they do not share the simulation "god's eye" view of the world, rather they gather and interpret data according to their own capabilities[36]. One can characterize the degree of an agent's intelligence based on the extent of its historical sensory database, as well as its capability to use inference to supplement incomplete input data, and/or to resolve uncertain or inconsistent data.

It is important, however, to not over state the "intelligent" element of these software constructs. In fact, part of the rationale for using intelligent agents is their capability to represent such "unintelligent" behaviors as bad decisions, failure to correctly interpret conditions and/or events (poor situation awareness), poor navigation (getting lost), and other cognitively related performance problems.

3.1.1 The Role of Autonomy

The "intelligence" of an agent also speaks to its degree of autonomy. Andrew Ilachinski, one of the pioneers of agent-based modeling and its use in representing CAS, defines autonomous software agents as:

> *The fundamental building block of most models of complex adaptive systems is the so-called adaptive autonomous agent. Adaptive autonomous agents try to satisfy a set of goals (which may be either fixed or time-dependent) in an unpredictable and changing environment. These agents are "adaptive" in the*

[36] This aspect of intelligent agents is another central tenet of the methodology framework developed herein.

sense that they can use their experience to continually improve their ability to deal with shifting goals and motivations. They are "autonomous" in that they operate completely autonomously, and do not need to obey instructions issued by a God-like oracle (Ilachinski 1997). p.13

The most attractive feature of autonomous intelligent agents is not how "smart" they are, but rather that their behavior is controlled internally rather than externally. Instead of being driven by some simulation "god" or master event scheduler, intelligent agents act according to their own perception of the simulation environment. This shift of control from global to local means that simulation setup and execution need not be quite as prescient as before. By providing an autonomous entity a set of stock behaviors and internal selection/decision logic, more scenario contingencies, and a far larger number of contingency combinations, can be addressed, without the need to explicitly account for each one.

In ABM, autonomy is gauged by the degree to which behaviors are not pre-scripted by simulation designers[37], by the number of options available to the agent in response to the perceived environment and by the flexibility the agent has in choosing those options. Current ABM software constructs have to a large extent freed the simulation scenario developer from the need write the scenario as a scripted set of events and turned it more into stage wherein the scenario actors can to some extent ad lib their parts.

3.1.2 Implementations of Agent-based Modeling

Ilanchinski developed several agent-based simulation tools. They include the Irreducible Semi-Autonomous Adaptive Combat (ISAAC) model and its successor EINSTein, Enhanced ISAAC Neural Simulation Tool. (Ilachinski 1997) EINSTein and ISAAC were among the first agent-based models of combat, and were the forerunners of the

---

[37] Bertrand Russell once remarked (or if he didn't he should have) that you could tell how good a philosopher is by how deeply he buried his contradictions. The corollary for ABM modeling of complex systems is establishing how deeply one can bury the scripting inherent in any digital computer simulation. One strives to bury that scripting as deeply as one possibly can, and thus to reduce fundamental ontological relationships to as parsimonious and as simple a set as possible.

Marine Corps' Warfighting Laboratory's Project Albert, whose suite of models grew to include Socrates, Pythagoras, and the New Zealand Defense Technology Agency's Map-Aware Non-Uniform Automata (MANA).

These models have been referred to as distillation models, "a type of computer simulation which attempt to model the critical factors of interest in combat without explicitly modeling all of the physical details." (Horne 2008) [38]

Because these models are small and abstract, "they can "easily be run many times to test a variety of parameter values and get an idea of the landscape of possibilities. The term distillation is added, because the intent is to distill the question at hand down into as simple a representation as possible." (Horne 2008)

In addition:

> *By virtue of their being much easier to run and understand (think: SimCity adapted to a combat situation), they are proving to be effective tools that help capture and scientifically reproduce the ideas of Subject Matter Experts, such as those thinking about tomorrow's concepts, doctrine, and requirements. This suite of entity-based models allow for rapid and highly tailorable changes in entity characteristics and behaviors, quite amenable to, and intentionally designed for rapid, repeatable concept exploration. Project Albert develops a suite, vice a single model, to allow for the testing of robustness of observations across modeling platforms, and because each model has inherent strengths and unique capabilities with regard to each aspect of modeling how entities think, decide, shoot, move, and communicate. (from http://www.projectalbert.org/)*

Agent-based modeling is not a major divergence from popular simulation approaches such as discrete event simulations and systems dynamics. Rather it builds on top of these tools, and any comprehensive combat simulation will incorporate elements of multiple approaches. One must still, of course, deal with questions of fidelity and resolution – what are the key features of those elements represented and

[38] Horne and Schwierz go on to note that these models continue to be developed and recent updates are described in Lauren, M., McIntosh, G., Galligan, D., Anderson, M., 2007. Recent Developments in the MANA (Map Aware Non-uniform Automata) Agent-based Model, Scythe Issue 1: 38-39. Monterey, CA. ; Lampe, T., Schwarz, G., and Wagner, G. 2007. PAX: Designed for Peace Support Operations, Scythe Issue 2: 43-48. Monterey, CA; and Henscheid, Z., Middleton, D. and Bitinas, E. 2007. Pythagoras: An Agent-Based Simulation Environment, Scythe Issue 1: 40-44. Monterey, CA.

what level of detail is appropriate. There are a myriad of different functional elements that must be considered in a combat simulation. They range from the combatants themselves, to their weapons and other equipment, to significant features of their environment such as terrain, weather, physical threats, and cultural milieus. There are a multitude of simulation methodologies for representation of any and all of these elements, choosing the one(s) most suitable for a given application can be a significant challenge. When that application is representation of decision-making, the selection problem can be even more daunting. It is often hard to predict a priori which elements of a given scenario will be critical a given decision, or even what and when decisions will have to be made.

For example, the current agent-based models IWARS[39], Pythagoras[40] and the OneSAF Objective System (OOS) [41] are clock driven simulations that maintain lists of scheduled events. All of them use Monte Carlo draws to determine the outcome of stochastic processes, while also using closed form equations and/or algorithms based on these equations to calculate deterministic results. Military operations are characterized more in terms of a series of goals that must be achieved than as the sequence of tasks required to achieve those goals; driven by the commander's intent interpreted and realized by subordinates[42]. Mission and task performance can incorporate unexpected events and multifaceted simulation object interfaces. These features highlight the difference between task network modeling and the agent-based approach. The successor to the IUSS, IWARS, changes the analytical focus from modeling operations as a sequence of by-the-book tasks to modeling them as the interaction between various

[39] For descriptions of IWARS, the Infantry WARrior Simulation, and illustrations of its application in agent-based analyses, see for example, (Warwick 2006; Bachman, Hester et al. 2008; Gillis 2008; NRDEC 2009)

[40] For descriptions of Pythagoras and illustrations of its application in agent-based analyses, see for example, (Bitinas, Henscheid et al. 2003; Liang 2005; Alt 2006; Corp 2008)

[41] For descriptions of OOS and illustrations of its application in agent-based analyses, see for example, (Parsons 2005; Parsons 2005; Surdu, Parsons et al. 2005; Logsdon 2008)

[42] see for example: (TRADOC 1995; TRADOC 1997; Shattuck and Woods 2000; TRADOC 2003)

entities and the environment.(Middleton 2002; Middleton 2010b) This change did not mean abandoning the representation of Army standard tactics, techniques, and procedures (TTPs) but rather providing a greater degree of autonomy to simulation entities as to when and where to employ these TTPs.

3.1.3 The Role of Error

As stated above, consideration of the individual as an *imperfect* decision-making entity is at the core of the human-centric approach.

> *One of the most important aspects of human performance, which has often been overlooked in models of behavior and problem solving, is errors (although see, for example, Cacciabue et, al.; Freed & Remington, Freedet. al)[43]. There is a consensus building about the definition of errors -- for most people an error is something done that was not intended by the actor, that was not desired, and that placed the task/system beyond acceptable limits (e.g., Senders & Moray)[44].*
>
> *Part of the reason for omitting errors from models of behavior is the fallacy that they are produced by some special error-generating mechanism that can be bolted on to models once they are producing correct behavior on the task at hand. Often, however, the actions that precede errors would have been judged to be correct if the circumstances had been slightly different. In other words, as Mach[45] observed, knowledge and error both stem from the same source.*
>
> *Evidence shows that novices and experienced personnel will often make the same errors when exposed to the same circumstances. The difference lies in the ability to notice and recover from these errors. Experienced personnel are more successful at mitigating errors before the full consequences arise. In other words, it is the management of errors that is important and needs to be trained (Frese & Altmann)[46], rather than vainly trying to teach people how to prevent the inevitable. (Ritter 2001)*

Mica Endsley, whose three stage view of situation awareness is discussed in detail below, developed a taxonomy of SA related errors:

SA Error Taxonomy

Level 1: Failure to correctly perceive information

[43] cites (Cacciabue, Decortis et al. 1992; Freed, Shafto et al. 1998; Freed and Remington 2000)

[44] cites (Senders and Moray 1991)

[45] cites (Mach 1905)

[46] cites (Frese and Altman 1989)

Data not available
Data hard to discriminate or detect
Failure to monitor or observe data
Misperception of data
Memory Loss

Level 2: Failure to correctly integrate or comprehend information
Lack of or poor mental model
Use of incorrect mental model
Over-reliance on default values
Other

Level 3: Failure to project future actions or state of the system
Lack of or poor mental model
Over-projection of current trends
Other

General
Failure to maintain multiple goals
Habitual schema (Endsley 1999)

All of the above are desirable features to capture in *Agent Sapiens.*

Miler and Shattuck, whose Dynamic Model of Situated Cognition (DMSC) is also discussed below, found it necessary to consider how ground truth, which in their model is completely accurate, constantly updated, and identical for all entities involved, is subject to a series of distorting lenses that can propagate error and inaccuracy at any and all stages of individual and unit information processing. (Miller and Shattuck 2006)

Henricksen and Indulska explore the role of error in background context with respect to human perception and information processing:

> *Context-aware applications typically assume that the context information upon which they rely is complete and accurate. However, this assumption is usually unjustified, as sensed context information is often inaccurate or unavailable as a result of noise or sensor failures, while user-supplied information is subject to problems such as human error and staleness.*
>
> *Consequently, usability problems arising from reliance on imperfect context information are sometimes observed in context-aware applications. For example, Benford et al. recently presented an interesting discussion of the implications of using imperfect location data in an mixed-reality game[47]. They*

[47] cites (Benford, Anastasi et al. 2003)

noted that errors in location information often led to confusion when game players were observed to jump around unpredictably in the virtual environment, but were also exploited by sophisticated users for tactical advantage.

Clearly, context-aware applications must be developed with an understanding of the problems inherent in gathering reliable context information, and also of the attendant design issues. In this paper, we explore these challenges. We characterise various types and sources of imperfect context information, present a set of novel context modelling constructs that accommodate these, outline a software infrastructure that supports the management and use of imperfect context information, and describe our experiences with using the context modelling approach and infrastructure. (Henricksen and Indulska 2004)

3.1.4 Building Agent-Based Models Reductionism vs. Synthesism

The correspondence vs. parsimony dynamic described in Chapter 1.4 is perhaps even more acute with respect to ABM. Standard methods of decomposing the entity functions and behaviors into independent modules require ignoring or understating dynamic interactions between the modules, which runs counter to the ABM approach. In general, it is difficult, if not impossible, to separate the function of equipment such as sensors and weapons from the functioning of the human equipment operator. Furthermore, human behaviors themselves typically have both cognitive and physiological components and again, it is not generally easy to delineate between mind and body. Physiological stressors such as fatigue can certainly affect cognitive capability, and conversely mental states and attitude can affect physical performance. In the world of conventional modeling and simulation life is made simpler by assuming things like independence and sequential causality[48]; unfortunately human behavior seems to resist these assumptions. The problems of trying to separate effects into neat little pieces or functional/behavioral models are further exacerbated by the problem of supporting data. Experimentally, one can measure performance outcomes for equipment/operator pairings, generalizing these results to other equipment configurations and/or operators with different skill sets and varying psycho-

[48] for more on these concepts see for example: (Pearl 1987; Schauble, Glaser et al. 1991; Chandrasekaran 1994; Pearl and Robins 1995; Rönngren and Liljenstam 1999)

physiological states is problematic at best. Similarly, for most human behaviors physiological state variables can be measured more or less directly; cognitive state(s) and decision processes must generally be inferred from such observables, somewhat confounding the development of independent effects mechanisms.

One approach to representation of human behaviors is a compromise achieved by combining top-down, reductionist approaches, with bottom-up synthesis. The reductionist approach decomposes well-understood phenomena into component pieces until an appropriate level of fidelity is reached. For example, the typical "two-person duel" engagement process can be broken down to target acquisition, weapon operation, and damage assessment. These functions can be further decomposed into detection, recognition, and identification, weapon aim and operation, projectile fly-out, projectile impact, and so on.

This reductionist approach works best when conditions (expressed as state variables for the objects involved) are relatively static. When conditions become more dynamic, the assumption of independence of the functional modules (necessary for the reductionist approach) is violated. For example, if a target is detected, but acquisition is lost prior to weapon fire, current models simply restart the acquisition process; they ignore the fact that re-acquisition of a previously known target is very different from the initial search and detection of that same target. Under such an approach the "two-person duel" can take on a "whack-a-mole" aspect, with each individual ducking out of sight in turn before his adversary can get off a shot. Less humorous, but equally unrealistic, is the case where an adversary can fire first because of the differential between the time required to re-initialize the acquisition/engagement cycle and the shorter time to engage given the ability to focus in on that adversary's last known position.

The ability of each side to react and adapt to what the other is doing, and has done, is perhaps the greatest advantage of synthesist approach[49]. Entity interactions are based on both action and reaction, and the history of previous interactions has the potential to influence current and future behaviors. Of course, there is no prohibition against including such histories in the reductionist functional breakdown, except that such a breakdown requires predicting all the possible sequences of such interactions and soon runs afoul of the combinatorial explosion of action/interaction possibilities. The agent-based synthesist approach avoids this problem (or at least delays it) by not assuming that the every sequence of possible interaction states is known a priori, as is the case with a strictly reductionist approach.

3.1.5 Agent-based Modeling and Analysis

While agent-based models provide a way to address the human dimension of combat and its uncertainties through emergent analysis:

> *"ABM is a mindset more than a technology. The ABM mindset consists of describing a system from the perspective of its constituent units. A number of researchers think that the alternative to ABM is traditional differential equation modeling; this is wrong, as a set of differential equations, each describing the dynamics of one of the system's constituent units, is an agent-based model. A synonym of ABM would be microscopic modeling, and an alternative would be macroscopic modeling."….*
>
> *One may want to use ABM when there is potential for emergent phenomena, i.e., when:*
>
> 1. *Individual behavior is nonlinear and can be characterized by thresholds, if-then rules, or nonlinear coupling. Describing discontinuity in individual behavior is difficult with differential equations.*
> 2. *Individual behavior exhibits memory, path-dependence, and hysteresis, non-markovian behavior, or temporal correlations, including learning and adaptation.*

[49] Ilichanski argues that agent-based modeling actually combines aspects of both the reductionist and synthesist approaches in something he calls the collectivist approach, where there continual feedback between the behavior of (low-level) combatants and the (high-level) command structure. (Ilachinski 1996) (Ilichanski 1997)

3. *Agent interactions are heterogeneous and can generate network effects. Aggregate flow equations usually assume global homogeneous mixing, but the topology of the interaction network can lead to significant deviations from predicted aggregate behavior.*

4. *Averages will not work. Aggregate differential equations tend to smooth out fluctuations, not ABM, which is important because under certain conditions, fluctuations can be amplified: the system is linearly stable but unstable to larger perturbations.*

Interestingly, because ABM generates emergent phenomena from the bottom up, it raises the issue of what constitutes an explanation of such a phenomenon. The broader agenda of the ABM community is to advocate a new way of approaching social phenomena, not from a traditional modeling perspective but from the perspective of redefining the scientific process entirely. According to Epstein and Axtell, ***ABM may change the way we think about explanation in the social sciences. What constitutes an explanation of an observed social phenomenon? Perhaps one day people will interpret the question, 'Can you explain it?' as asking 'Can you grow it'***[50]*." " (Bonabeau 2002)p.2*

As an example, Macy and Willer discuss the work of Craig Reynolds in modeling movement of a population of artificial "boids".

"Reynolds' computational method is called agent-based modeling. Had Reynolds chosen instead to write a top-down program for the global behavior of the flock, he might still be working on it. By choosing instead to model the flock from the bottom up, based on agent-level interaction, he was able to produce highly realistic flight formations using very simple rules.[51] *Note that Reynolds did not model the flock, nor did he model isolated birds. He modeled their interaction, at the relational level.*

Agent-based models (hereafter ABMs) of human social interaction are based on this same theory-building strategy. Sociologists have traditionally understood social life as a hierarchical system of institutions and norms that shape individual behavior from the top down. Interest in ABMs reflects growing interest in the possibility that human groups, like flocks of birds, may be highly complex, non-linear, path-dependent, and self-organizing. We may be able to understand these dynamics much better by trying to model them, not at the global level but instead as emergent properties of local interaction among

[50] cites Epstein J. M. & Axtell, R. L., (Op cit.).

[51] Referring to Reynolds CW. 1987. Flocks, herds, and schools: a distributed behavioral model. Comput. Graphics 21:25–34 . Reynolds used three simple rules that described properties of separation, alignment, and cohesion.

adaptive agents who influence one another in response to the influence they receive." (Macy and Willer 2002)] p144

3.1.6 Complex Adaptive Systems

There is no single accepted definition of a complex adaptive system, although there is reasonable agreement as to the general characteristics of such a system. (Holland 1992), (Holland 2006) (Johnson 2001), (Waldrop 1992) First, of course, a complex adaptive system is a complex system. As shown in Figure 3-1, such systems are most often an open system possessed of a high degree of structure, with many components having dynamic interactions.

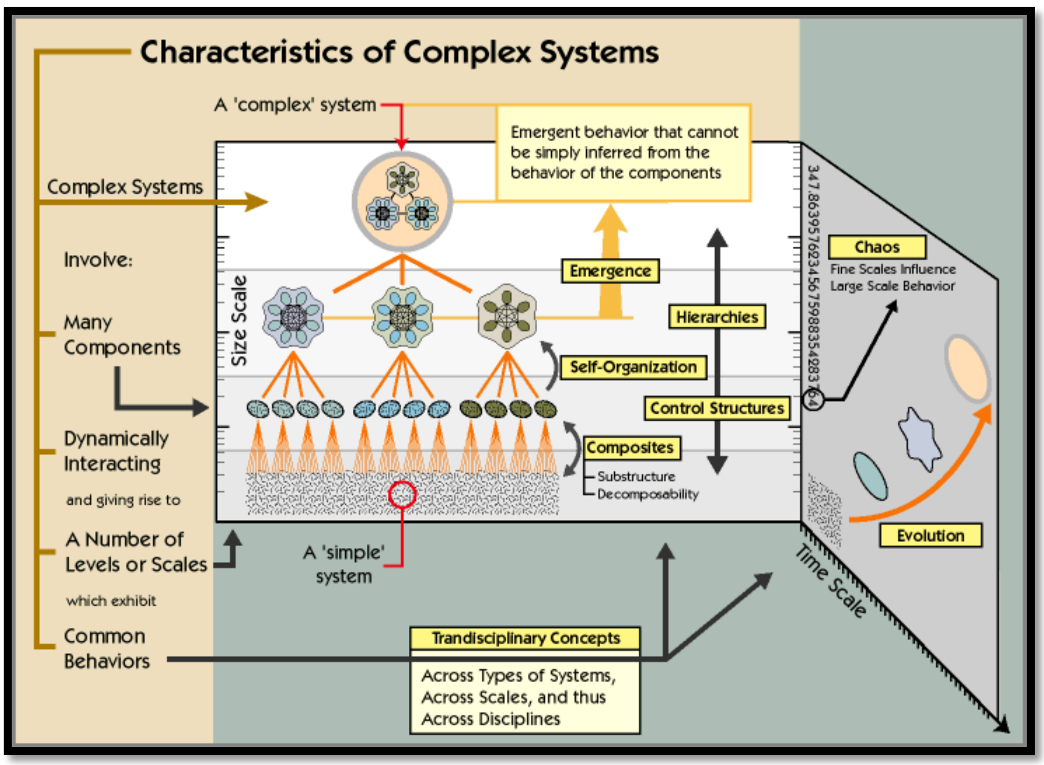


Figure 3-1 Characteristics of Complex Systems[52]

System behavior generally depends not only current state variables, but also on the

[52] Illustration by The Complexity & Artificial Life Research Concept for Self-Organizing Systems in conjunction with New England Complex Systems Institute of Cambridge, MA. www.idiagram.com/examples/complexity.html

history of the system. System relationships and behaviors are frequently highly non-linear and chaotic[53], meaning they can be highly sensitive to small perturbations in input conditions. The complex and dynamic nature of system component interactions frequently leads to emergent behavior, and the system is said to be adaptive when this emergent behavior results in the system(s) coevolving with the environment and other systems. Adaptation is a concept taken from the biological view of evolution and implies the presence of a "fitness" function or functions that support "selection" of those characteristics or behaviors of the system that enable it to best "fit" in its environment.

3.1.7 Military Operations as Complex Adaptive Systems

Certainly the view of military conflict as a complex system is well supported by these attributes.

> *"According to practically any definition of the term "complexity," war qualifies as a complex phenomenon. In what could qualify as an excellent description of complexity theory, Clausewitz wrote: The military machine—the army and everything related to it—is basically very simple and therefore seems easy to manage. But we should bear in mind that none of its components is of one piece: each piece is composed of individuals, every one of whom retains his potential of friction...A battalion is made up of individuals, the least important of whom may chance to delay things or somehow make them go wrong." (Schmitt 1997)*

The MOVES[54] Institute at the Naval Postgraduate School (NPS) is a significant proponent of applying ABM to represent military operations as CAS. Joerg Wellbrink, Mike Zyda, and John Hiles of MOVES refer to Ilanchinski's work, (and quote from it) as follows:

> *... Ilachinski challenged the almost century-old theory of conventional wisdom that combat is driven by force-on-force attrition rate] by arguing that land combat can (and should) be modeled as a complex adaptive system. He transferred complexity theory into the military domain and showed that land*

[53] Chaotic in the mathematical sense is further described, for example in (Gleick 1987)or (Kauffman 1993).

[54] MOVES is an acronym for Modeling, Virtual Environments, and Simulation

combat properties resemble the properties of CAS (in (Ilachinski 1997)). His work has generated a lot of interest in combat modeling, especially because tactical behaviors such as flank maneuvers, containment, encirclement, and "guerilla-like" assaults emerged out of his implementation.

"In ISAAC, the "final outcome" of a battle—as defined, say, by measuring the surviving force strengths—takes second stage to exploring how two forces might "co-evolve" during combat. A few examples of the profoundly non-equilibrium dynamics that characterizes much of real combat include: the sudden "flash of insight" of a clever commander that changes the course of a battle; the swift flanking maneuver that surprises the enemy; and the serendipitous confluence of several far-separated (and unorchestrated) events that lead to victory. These are the kinds of behavior that Lanchesterian-based models are in principle incapable of even addressing. ISAAC represents a first step toward being able to explore such questions [55].

They claim:

The paradigm for combat modeling has fundamentally changed and improved insights into the processes. These types of simulation systems will enhance the capabilities exploring policy and concept development as well as force structure development. (Wellbrink, Zyda et al. 2004)

An open systems paradigm that sees modern conflict as CAS supports exploration of the potential for emergent phenomena in such systems. This support includes representation of:

- Emergence - complex patterns of behavior and structural relationships emerge from apparently simple rules of interaction;
- Perception-based entity behaviors – entities that generally react to the environment and other entities only as they perceive them, not based on ground truth. Entity behaviors are thus a function of the entity's SA and understanding, which incorporate both immediate perception and the entity's history. With explicit representation of an entity's observe, orient, decide and act (OODA) loop, analysts can explore the contributions disrupting an opponent's OODA loop can make towards conflict outcomes:
- The effects of history and feedback loops on entity behaviors - effects of an entity's actions are fed back to it and can affect its future behavior. Fitness

[55] (Ilachinski 1997) p.226)

functions incorporating both negative (damping) and positive (amplifying) feedback determine the development and persistence of emergent behaviors.

- Network centric behaviors and echelons of command as local entity interactions - orders and information are explicitly transmitted and understood (or misunderstood), allowing for the introduction of network errors and communication problems into simulation experiments.
- Non-linear entity interactions – reflecting the often chaotic nature of especially asymmetric conflict, conflict outcomes may not be attributable to simple cause and effect relationships. Small changes in behaviors or seemingly minor interactions may cause large effects or no effect at all.
- Autonomously operating units in highly fluid operations - commanders specify their intent and rules of engagement, but do not attempt to rigidly control subordinate units leaving them to react and/or adapt. Adversaries who may not have a rigidly defined or enforced command and control structure will still have goals and patterns of engagement and/or other behaviors, and thus also function as autonomous entities.

Certainly the use of an open systems approach to combat modeling is not without its own problems. The closed system view essentially sees outcomes as deterministic; giving up a closed system approach does mean acknowledging at best a limited ability to predict those outcomes. There is still a role for closed systems analysis, but it is most useful when input conditions and systems operations options are tightly constrained.

This disclaimer, however, suggests that simply making the case that small unit combat operations can be modeled as CAS is not sufficient. This dissertation is looking at only one small aspect of human behavior in military operations, and one that can be studied without introducing the myriad of complications due to interaction with adversaries. Is it still reasonable to claim that 1) the behavior at question falls under the rubric of CAS, and 2) even if it does, is there sufficient value in doing so to justify having to deal with the attendant complications and inherent uncertainties?

To support the “yes” answer to this question, consider:

- while one would hope that navigation, or directed movement, in a known, static environment reflects deterministic, closed system behavior; the same cannot be said for directed movement guided by uncertain and/or erroneous information, and such movement certainly embodies most or all of the above CAS characteristics;

- the value of applying the CAS/ABM paradigm is can be seen in the fact that representation of decision-making under imperfect SA/SU is one of the quintessential aspects of modeling human behavior that have been in the "too hard to do box" until CAS and agent-based modeling came along. By supporting representation of the consequences of acting on imperfect SA/SU, the CAS/ABM paradigm allows analysts to correlate different levels of that phenomenon with accepted measures of mission outcomes.

Being "lost" in the sense of acting under erroneous or incomplete information, can certainly give rise to emergent behavior; perhaps the most obvious example of which would be the tendency of someone who is lost to wander in a circle. In the world of on-line games, both humans and computer-generated actors (even those with relatively sophisticated AI behaviors), fall into detectable patterns of decision-guided behaviors. Furthermore, if an individual's environment is changing in ways that affect movement, that individual's movement behaviors generally adapt to these changes. The need for adaptation is particularly pertinent when changes reflect differences between the individual's pre-conceived notions about the environment, and what that individual finds to be ground truth. Discovery of such differences is a key characteristic of an open system; new information is flowing into the system, or at least to that part of the system that reflects the individual's perceived view of the environment. Finally, the individual's reaction to that discovery often takes on chaotic and non-linear aspects, including abrupt discontinuities in the choice of type and/or direction of movement – I come to the top of a hill, see a landmark, and realize I'm headed the wrong way entirely.

3.1.8 Complex Adaptive Systems and Emergent Analysis

As mentioned in Chapter 2.2.3 and Chapter3.1.4 above, the notion of influence is one that occurs frequently in complexity science and agent based modeling. Studying the way such influences propagate through a simulation scenario leads to new analytical paradigm, which could be referred to as "emergent analysis". This process may produce outcomes, conflicts and conflict resolutions, as might not have been a priori identified in the analyst's view of the system.

Much of traditional analysis consists of testing a specific hypothesis against scenarios differing only in the value of an "independent variable", and iterated to provide statistical comparisons. Under the emergent analysis approach, as depicted in Figure 3-2, this "comparison test" approach to analysis may be supplemented or replaced with research or study for the sake of discovery, the "what if" or "what happens when" kind of investigation. Such experimentation can be just as valid (or perhaps more importantly, just as useful) as experiments narrowly structured to accept or reject a null hypothesis.

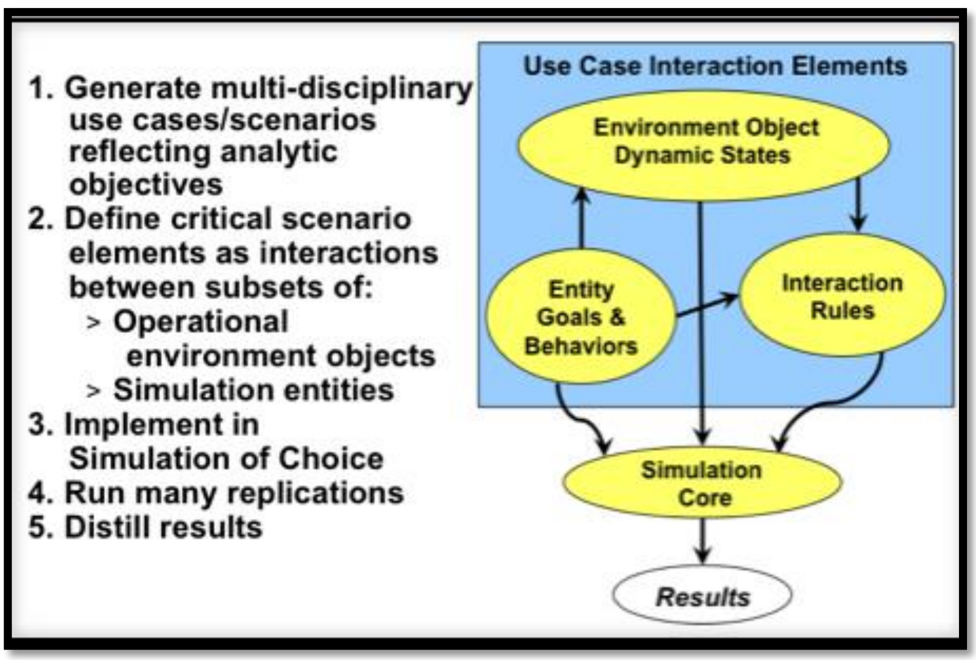


Figure 3-2 Emergent Analysis

This paradigm addresses the disconnect between deterministic models (i.e., individual object/entity behaviors) and the real-world variability due to (the essentially infinite) potential factor interactions. This paradigm is also attractive in that it supports a view of more autonomously operating units in the highly fluid situations emblematic of asymmetric warfare. Under this view, commanders would specify their intent and rules of engagement, but not attempt to rigidly control subordinate units form one central point of command.

With respect to the dissertation model MOBIL, this view results in providing a moving entity a set heuristics to solve its "function of mind": to decide where to move next.

Emergent analysis offers the potential for operational analyses to become more like operations themselves: when necessary, predictability is traded for unexpected, yet valuable insight and adaptation. When real people take part in realistic scenarios, as in live exercises or actual combat, unexpected events and outcomes occur frequently, often requiring on-the-fly adjustments. The nuances and perturbations in plans and operations introduced by human intelligence, and responses to them, are the very essence of military art. Representing this aspect of conflict more satisfactorily is an important challenge facing operations research. As a Navy modeling and simulation review puts it:

> *"Running a simulation for one set of fixed conditions is generally not satisfactory since there are often large uncertainties throughout the system…An important research area, then, is developing ways to use modern computer power to explore the space of simulation outcomes and to search for interesting regimes..." (Naval Research Board, 1997)*

The emergent analysis paradigm uses multiple simulation experiments to explore the behavior of multifaceted phenomena; emergent complex behaviors or properties are not a property of any single phenomenological factor, nor can they easily be predicted or deduced from isolated examination of individual factors or any single simulation replications. Analysts may need to be alert to the insights and anomalies that arise in such simulation runs, in order to capitalize on them, and perhaps move the analysis in an unanticipated, but potentially fruitful direction. For those unaccustomed to applying rigorous scientific methods to the study of fundamentally unruly phenomena such as human behavior, the struggle to identify and balance the subjective and objective elements in the process can be a daunting challenge.

The paradigm must deal with questions of problem scope that are related to the number of interactions between components of a system, which is a function of the number of components considered. The number of interactions can thus increase combinatorially with the number of components and can become quite large. Emergent analysis addresses the problem by employing current techniques in data farming and data mining to isolate and focus in on “interesting” aspects of the phenomena being studied.(Fayyad, Piatetsky-Shapiro et al. 1996; Brady and Starr 2002; Hewawasam, Premaratne et al. 2007; Horne 2008; Middleton 2009; Furnas 2012)

In the June 2008 issue of Wired magazine[56], editor-in-chief Chris Anderson describes the new "Age of the Petabyte," in which the data used for search and analysis exceed terabyte quantities and extend into petabytes. He claims that with the advent of the “age of petabyte”, the scientific method is outdated, unsuited to working with petabyte magnitudes of data, and that:

> *"The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world.* ***Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.." (emphasis mine****)*

As evidence of his claim he cites the success of Google, which:

> *“conquered the advertising world with nothing more than applied mathematics. It didn't pretend to know anything about the culture and conventions of advertising — it just assumed that better data, with better analytical tools, would win the day. And Google was right. ….*
>
> *Google's founding philosophy is that we don't know why this page is better than that one: If the statistics of incoming links say it is, that's good enough. No semantic or causal analysis is required. … it can match ads to content without any knowledge or assumptions about the ads or the content. Speaking at the O'Reilly Emerging Technology Conference this past March, Peter Norvig, Google's research director, offered an update to George Box's maxim[57]: "All*

[56] (Anderson 2008)

[57] George Box’s maxim is: “All models are wrong, some models are useful

models are wrong, and increasingly you can succeed without them." [58]

For military modelers and analysts does this view of the age of petabyte mean having to give up cherished (and not so cherished) models and simulations? The vast increase in available data certainly has (and should have) significant impact on scientific analysis, but just as certainly it isn't the end of the scientific method or the utility of models. Instead it supports a shift in focus, from scientific experimentation as hypothesis and test, towards scientific experimentation as exploration and discovery. Note that this is precisely the aim of data farming and data mining, which use simulation experiments to generate mass quantities of data, trolling these data for patterns, trends, insight and meaning. As in much of science, the process begins with induction and correlation, from which one hopes to infer cause and effect, but must sometimes be satisfied with increased understanding and some hope of predictive validity.

Emergent analysis thus begins with looking for patterns, the way in which interactions and influence propagate through a simulation scenario, and from these patterns seeks to find the causes and effects of emergent behaviors. Such behaviors frequently result from seemingly arbitrary aggregations of influence of some factors over that of others, which appear a priori to be equally important. The nature of complex systems frequently manifests itself it such chaotic behavior, where minor perturbations of initial conditions have significant effects on ultimate outcomes. Such phenomena is seen in instances where behaviors or attributes initially chosen at random tend to reinforce themselves in what might be termed the principle of "them what has gets". For example, in genetic algorithms one can observe the process of genetic drift as

> *"eventual convergence around one optimum. … imagine that we have two equally fit niches, and a population of 100 individuals equally divided among them. Eventually because of the random effects of selection, it is likely that we will obtain a parent population consisting of 49 of one sort and 51 of the other. … we are increasingly likely to select individuals from the second niche. This*

---

[58] (Anderson 2008)

> *effect increases as the two subpopulations become unbalanced, until eventually we end up with only one niche represented in the population." (Eiben and Smith 2003)p155*

Genetic drift is a process where an initially random application of influence shapes the nature of an outcome, based on simple rules (e.g., select parents randomly from the population), and does so in a way that is not easily reducible to a set of hierarchal procedures.

In a decision-making context this kind of behavior can often be seen in the tendency to "double-down" on previous choices. As a decision-maker has more time and resources invested in a particular course of action (COA), there is a natural tendency to avoid abandoning or altering that course. Evidence that may come to light suggesting that the COA is ineffective (or even disastrously wrong) can often be ignored or mis-interpreted to conform to initial assumptions.

One of the most critical decisions for stochastic simulations is the degree of independence one wishes to assume for sequential events. If, for example, one assumes probability of detection of a given target to be independent of previous detection attempts, the actual probability of detection can become a function of the frequency of Monte Carlo sampling, and thus an artifact of model update rates. If one samples often enough, the assumption of independence and the binomial probability distribution make it likely that even an improbable event will occur eventually. In many (most?) cases, however, detection attempts are not independent; the outcomes of previous detection attempts do have an effect on successive events. As a result, one may need to adjust the probability distributions for Monte Carlo behaviors based on that simulation history. Such is the case in the target engagement process discussed above, if one fails to adjust probability of acquisition based on previous history, the probability of detection is likely to be understated.

When there is a fitness function or some other mechanism further influencing the outcome of stochastic behavior choices, the overall system being modeled tends to adapt or evolve in directions favored by these mechanisms. One such mechanism is the use of goal-driven behaviors, i.e., agents "choose" behaviors to achieve specific goals. Those combinations of behaviors that succeed in helping to meeting these goals, and those agents who select these behaviors, "emerge" as the dominant outcomes in simulation experiments.

3.2 The Human-Centric Paradigm: Review and Recapitulation

This chapter completes the task of articulating the reasons for and the nature of the human-centric paradigm. It looks at the foundational elements of that paradigm, amplifying and summarizing the pertinent points from the discussion above. Chapter 3.2.2 summarizes the core aspects and requirements and Chapter 3.2.3 concludes with a brief over-view of current analytical practice with respect to the paradigm.

3.2.1 *Foundational Elements for the Human-Centric Paradigm*

The previous chapters have laid out a foundation for the human-centric paradigm:

- providing the historical evolution of analytical tools leading up to the human-centric concept;
- defining the motivation for this concept in terms of changes in military philosophy and operations;
- relating the development of human-centric M&S with the evolution of the Warrior System;
- outlining the need for human-centric M&S with respect to R&D assessment requirements for the information technology applications that have accompanied those changes;
- describing the ABM/CAS paradigm that is the enabling methodological foundation for human-centric M&S, and
- discussing the metrics required for performance of human-centric analysis.

Formulation of the human-centric paradigm requires integrating these points into a comprehensive view of human-centric analysis,

One historical approach to modeling the human has been to adapt methodologies originally developed for large weapons platforms (Middleton 2001; Middleton 2010b).

This oft-stated view of the human as a slow unarmored tank has some validity for modeling the physics-based aspects of movement, target detection, employment of ballistic projectiles, message transmission and the like, but it falls woefully short of describing the complexity of human behavioral response to the dynamics of the battlefield. In particular, this view gives short shrift to representation of the dismounted Infantry combatant.

The increasingly asymmetric nature of combat brings questions of individual behavior to the forefront, as does the current push for a smaller, more agile force structure. A the same time, current priorities in the employment of armed forces give greater emphasis to crisis and peace support operations, as well as other, less exclusively conventional combat oriented, missions. This changing emphasis can be referred to as a shift from “close combat” toward “close contact” operations (van Son 2003). From an RD&A perspective this shift brings with it the need to augment traditional models of combat, which focus on the each side’s capabilities for attrition of the others forces and resources, with more nuanced analysis of military operations that supports goals other than the destruction of an adversary. This shift will bring with it the need to augment direct measures of attrition warfare to reflect this much broader range of mission concerns. Development of metrics to quantify these effects, and definition of the human factors that contribute to them, are major challenges.

The multiplicity of human factors that contribute to combat outcomes is a further complication. Some of these human factors are quite compatible with the kind of quantitative analysis characteristic of existing combat models, for example: aspects of human performance related to sensory processing; information processing; and motor performance. These topics, in fact, are the ones generally subsumed under the rubric of “human factors”, reflecting the more narrow common usage of that term, and many of them can already be found in combat models and simulations. To an extent, when humans act as sensors, this behavior can be modeled in the same terms used to represent electro-optical sensors. For example, the behavior of humans acting as sensors can to some extent be modeled in the same terms used to represent electro-optical sensors (CERDEC 2005).

M&S tools to support analysis now need to represent the ways in which individuals make decisions, the highly context-sensitive information requirements associated with those decisions, the capabilities for, and costs of, meeting those requirements, and the consequences for failing to do so. Efficient command and control functions, including communication, shared situation awareness, and support for coordinated action, are critical to a more agile force, and create an ever-increasing demand for information. The need to assess current and evolving information technologies, and their application to the acquisition, flow and use of data and information at the small unit level, drives the need to model the mental state of individuals, and to represent individuals' (and especially commanders') decision processes. Internal cognitive states and processes must be modeled to support realistic soldier behavior occurring at multiple levels of analysis. These include complex planning and high-level strategic or tactical decisions, intermediate-level decisions such as target selection and route adjustment, and even non-deliberative processes such as modification of speed and/or posture in response to environmental cues and one's own psycho-physiological state.

Assessment of operations on the digital battlefield, and especially the impact of technology on those operations, requires representation of the way the individuals and organizations acquire and use information. This includes the need to represent the processing of sensory inputs, and the interpretation and integration of sense data into meaningful information about the situation. It also includes the need to represent the effects of the individual's mental state on behavior, and conversely, the need to represent the effects of the individual's physical state on his or her mental state, and thus on his or her behavior and task performance.

Agent-based modeling techniques appear to be a good fit for representing human use of information, including inferences and decisions based on imperfect and even incorrect information. Such techniques support a view of military operations as complex adaptive systems and a move to an “emergent” analysis paradigm. In this paradigm one specifies the behavioral capabilities of entities and/or groups, defines a (hopefully small) number of possible interactions between entities/groups, and then sees how the agents react to achieve individual and group goals. This process may produce outcomes, conflicts and conflict resolutions, as might not have been a priori identified in the analyst’s view of the system.

This paradigm is also attractive in that it supports a view of more autonomously operating units in the highly fluid situations emblematic of asymmetric warfare. By defining the behaviors of individual, goal-driven entities, ABM provides a manageable and understandable way to study military systems of interest: warrior systems; adversaries, and associated aggregate force structures.

3.2.2 Summary

The human-centric approach to MS&A can be summed up as expanding representation of the individual, up to now viewed primarily as an actor - characterized by actions (move, shoot, communicate), to include consideration of that individual as a decision-making entity - one who continually evaluates the battlefield dynamic to decide when, where, and how to move, shoot, and communicate.

The human-centric paradigm represents the individual as an entity that ***imperfectly***:

- Perceives its Environment
- Possesses and Employs Knowledge
- Makes Decisions
- Orders Action
- Takes Action
- Monitors Mission/Task Progress
- Regulates/Adjusts Performance

This view logically leads to a series of requirements that include:

- The need to distinguish between an individual’s perceived world view and ground truth, which in turn requires representing acquisition and processing of data by that individual, modeling the human as a perceiving entity, incorporating error and uncertainty in both the processes of perception and the resultant world view;

- The need to support for each individual a knowledge base of relatively static information that defines a scenario, the entities within that scenario, and what those entities know about themselves and the scenario. Again, this knowledge base should be able to accommodate both error and uncertainty;
- The need for inference schemes supporting data filtration and fusion to allow the individual to correlate perceived data with its knowledge base, and interpret all these data with respect to a continually varying experiential frame of reference;
- The need to represent schema used to represent inference[59] and/or decision-making[60];
- The need for data structures that support implementation of decisions into actions There are three basic forms: **Course of Action** (COA) options; **Behavior parameters** - targets and target priority lists, types and rates of fire, shoot/no shoot decision thresholds for engagement; routes and waypoints or direction vectors for movement, speed and movement formations; and **Communications** – Situation reports (SitReps) to other units, especially command units, directives to subordinates, unit coordination, request for fire or other support.
- The need for physical models of behavior, as are already extant in most simulations;
- The need to represent the psycho-physiological state of the individual insofar as this state affects both decision-making and physical actions;
- The need to represent behavior moderators that alter the psycho-physiological state of the individual in concert with physical models of behavior and of the environment;

---

[59] Following the general lead of (Davis, Shrobe et al. 1993)I am using inference in the generic sense as a way to get new information from old, rather than as limited to sound logical inference.

[60] The topic of decision-making is worthy of a Chapter in its own right, and I provide one, Chapter 5.2 below.

- The need for metrics to express and assess actions taken in terms of their effectiveness towards accomplishing goals; and
- The need for feedback loops modifying performance parameters as appropriate to the individual's perceived progress towards goals.

Agent-based modeling is a useful construct to represent military operations, and human endeavors in general, as complex adaptive systems. Any simulation that satisfies the above requirements would by definition construe ABM and CAS.

Finally, the human-centric paradigm also includes the use of emergent analysis and the conduct of simulation experiments that are intended more to explore and understand the nature of operations than to predict operational outcomes.

3.2.3 Current Status

Stating that the human-centric approach "has not as yet been fully defined and articulated", does not mean that it is not a significant aspect of current military operational analyses. While there is not a formal organizationally sanctioned program towards human-centric M&S, there has been progress on most, if not all, of the foundational elements listed above. Certainly simulations such as IWARS and Pythagoras are designed and implemented following the basic ideas of human-centric analysis, and a number of the other Department of Defense efforts mentioned in passing above support those ideas.

As mentioned above, current operational philosophies such as EBO and NCW embody much of the human-centric emphasis on the equating the importance of the information and cognitive domains of warfare with those of the physical domain. This view continues, now also framed as "kinetic" vs. "non-kinetic" operations[61].

Simulations such as OOS have explicitly designed their actors with both behavioral and physical agent components, thus distinguishing between cognitive functions and

---

[61] See for example (Lowe and Ng 2004; Darley 2006)

strictly physical actions. In OOS, “*behavioral agents provide command and control capabilities, such as planning, plan execution, and situation assessment*”, while “*physical agents are the “middlemen” between behaviors, the physical world, and physical models*.” OOS also distinguishes between behavioral and physical models: “*behavioral models answer behavior agents’ questions and represent the reasoning of agents*” whereas “*physical models provide physical capabilities, such as mobility, weapons, vulnerability, sensing, and communications. They represent the effectors and perceptors of simulated platforms and the physics of the simulated world*.”(Logsdon 2008)

This approach is similar to the one used in this research. MOBIL simulates and entity through two agents: one that is constrained to act in the ground-truth environment, and another that directs the actions of the first based on a perceived (and usually distorted) view of that ground-truth environment.

As with any model, this approach is a simplification of reality. As discussed above in Chapter 2.3.3, the separation between mind and body is not clear-cut, and the distinction blurs as fatigue and other stressors degrade both the physical and mental capabilities of an individual.

The human-centric approach is currently manifested in a number of efforts to introduce and incorporate behavior moderators such as fatigue and stress into MS&A. Tanks and airplanes don't get hot and tired, they don't make decisions influenced by their psycho-physiological status, but their operators do. Several methodologies have been proposed for tracking behavior moderators and their effects on task and mission performance. See for example:(Gillis and Hursh 1999; Dubois 2000; Ritter 2000; Hudlicka 2002; Hudlicka 2002; Jones 2002)

The “soft” factors can affect both cognitive and physical behaviors. Representing them requires moving beyond the physics of the battlefield to an understanding of human physiology and psychology. "Good" modeling of the individual must take into account the concept of humans as processors of information that accept inputs (scenarios, environmental conditions, missions) and produce outputs (decisions, actions, mission performance). Pew and Mavor state the need for: “an integrative model that subsumes all or most of the contributors to human performance capacities and limitations” (Pew and Mavor 1998). Much still needs to be done to achieve such a model.

The chart shown in Figure 3-3 was prepared FY2000 while working as the chief scientist on the development of IWARS for the U.S. Army NRDEC (Middleton 2000). The chart is divided into the major soldier system capability areas, as then defined for the soldier as a system. The key take-away from this chart is that the physical function of separate elements of the soldier system are fairly well represented in models and simulations, while the integrated function of multiple systems, especially with respect interactions featuring significant cognitive behaviors, are not.

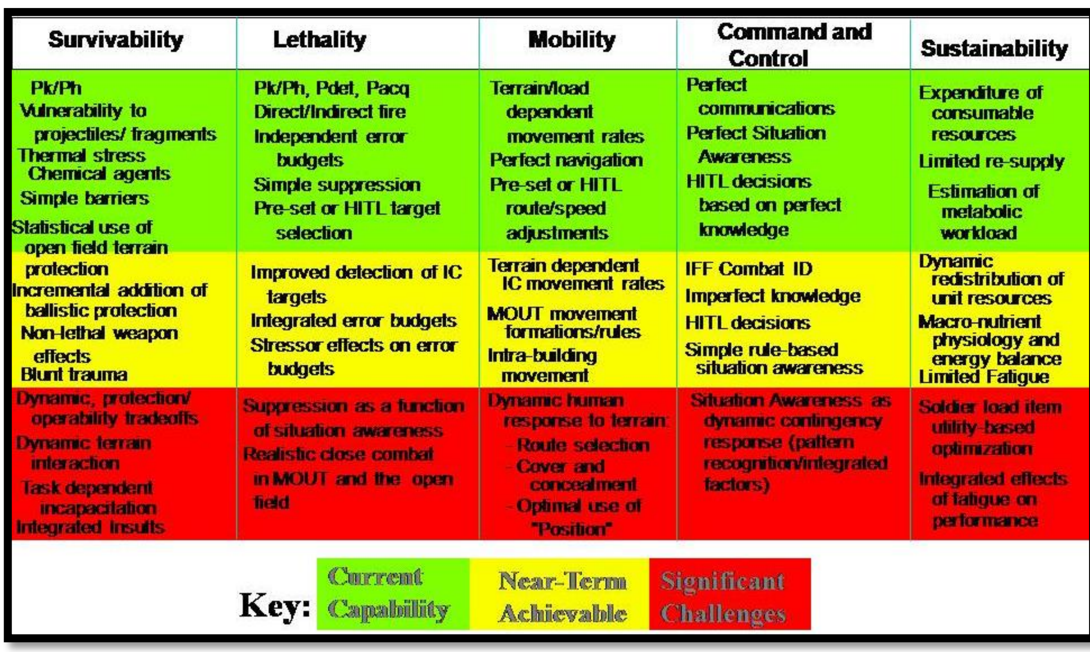


Figure 3-3 Year 2000 View of State of the Art for Individual Combatant Modeling

It is safe to say that the items listed in red are still significant challenges today, as evidenced by the proceedings of any of the recent simulation conferences[62], all of which are replete with papers documenting work being done in current modeling and simulation efforts are striving to meet the challenge of adequately representing the actions, tactics, techniques, and procedures of Infantrymen and Infantry units at the individual, fire team, squad, platoon, and company levels. In fact, in today's environment some of the challenges are even greater. Irregular warfare and cultural geography are two of current hot button topics, both of which require broad interdisciplinary efforts, stretching from sociology and cognitive psychology to military affairs to geographical information systems (GIS). Such an interdisciplinary approach is a natural fit with the modeling of small unit combat as complex adaptive systems.

The "soft factors" are at the core of most of these remaining challenges. Human behavior is different in principle from many of the combat processes modeled in most of our simulation systems. Ballistics and aerodynamics may be complicated, but they are also well-understood and based on physical laws. Human behavior may or may not be rooted in something more than purely physical processes, but human behavior is simply not lawful in the same way that physical events are lawful. Humans are indeed subject to certain regularities in both cognition and behavior that arise from the peculiarities of our construction, but it seems that they can choose to violate at least some of those regularities when it suits them. To better represent human behavior in models and simulations, one must acknowledge that stochastic models, random draws, table look-ups, simple probabilistic errors, and other techniques commonly used in military models and simulations need to be supplemented by new and different methods tailored to the special characteristics of distinctly human behavior.

[62] See for example, the Simulation Interoperability Standards Organization's Simulation Interoperability Workshops (SIW) or Behavior Representation in Modeling and Simulation (BRIMS), at http://www.sisostds.org/index.phpsee, the Institute for Operations Research and Management Sciences INFORMS ONLINE at http://www.informs.org/index.php?c=9&kat=MEETINGS, the Winter Simulation Conferences at http://www.wintersim.org/, The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) at http://www.iitsec.org/; or any of the publications of the Military Operations Research Society (MORS) http://www.mors.org/, to name just a few

3.3 Applying the Human-Centric Approach: Modeling Decision-Making and SA/SU

At this point the literature review and background material turns from the generic view of the human centric approach to MS& A to a more specific aspect of that approach. As described above, the human-centric approach to modeling and simulation embraces both philosophy and methodology. The rest of this dissertation narrows the focus from these over-arching issues of "why" and "how" to deal with the still vast arena of decision-making and situation awareness. Subsequent Chapters will further constrain the discussion to a concrete example of how SA and decision-making interact, complete with a computer model of that interaction.

SA/SU is a practical next step in the admittedly difficult task of pursuing the goals of human-centric modeling. SA/SU is one of the elements of the intermediate ground occupied by human factors that are neither as nebulous as "will to fight" nor as rigorously measurable as sensory capabilities and specific task performance (Mastroianni 1996; Mastroianni and Middleton 2001; van Son 2003; Middleton 2008). This middle ground includes many of the concepts that occupy space in the "red" sector of Figure 3-3.

In order to express any of these concepts in terms concrete enough to be included in constructive simulation, one must first define them in terms of quantitative attributes, find ways to measure or estimate those quantities, and finally, relate those measures to individual and unit behaviors that affect operational outcomes. This chapter discusses this process with respect to SA/SU, exploring SA/SU as a quantifiable aspect of an individual's cognitive state and how that state may be related to success or failure on the battlefield. Representation of SA/SU potentially provides a blueprint for other aspects of human-centric modeling. Showing how aspects of SA/SU can be sufficiently defined in operational terms, and brought into the methodological framework of models and simulations captures an essential psychological dimension of human military performance. Furthermore, this human-centric "middle ground" is one of the most fruitful areas for near-term enhancements to models and simulations.

4 SITUATION AWARENESS: A KEY CONCEPT[63]

Given the need to focus on the representation of SA/SU in simulation, one must first demonstrate that such representation can be done rigorously enough to support the analytic integrity of studies with such simulations. Considering this contention begins by confronting the more basic question of exactly what is meant by “situation awareness”, including what, if anything, is the distinction between SA and SU, “situation understanding”? Definitions of SA abound, and the concept is itself is not without detractors. There is still, for example, debate as to whether SA represents a state of knowledge, the processes required to develop that state, or both.

According to what is probably the most widely accepted view of SA, that of Endsley, SA has three components: perception of elements in the environment, integration of these elements into a comprehensive situation assessment, and projection of that assessment into the future. (Endsley 1995a) She contends that SA is a characteristic of the operator that can be measured independently of the level of ongoing performance and argues that SA is a concept is distinct from performance. She points out that one can have poor SA and still perform well if demands are low; one can also have good SA and perform poorly if the task demands are simply too great. Others, such as Mastroianni, hold that SA is closely related to performance. He points out that good

---

[63] As with Chapter 2.5.3, this Chapter owes much to discussions and other communications with long-time colleague and friend Dr. George Mastrioanni of the Department of Behavioral Sciences and Leadership at the US Air Force Academy as well as with scientists and simulationists at the TNO Prins Maurits Laboratorium, Rijswijk, The Netherlands.

performance is generally linked with the notion of having good SA, and, for example, when an airplane makes a controlled descent into (instead of onto) the ground, it is accepted as *prima facie* evidence of bad SA (not to mention bad karma). Certainly requirements for improved SA are based on the assumption that better SA will lead to better performance. (Mastroianni 1996)

Representing SA and SU requires explicitly modeling the knowledge state of simulated individuals and the dynamic processes by which that knowledge is developed, maintained, updated and applied. This representation must also focus on a critical and sometimes overlooked aspect of SA, the need to represent errors and inconsistencies in the individual's knowledge base. Gauging the incremental benefits of improved SA is difficult without comparison to a baseline that may contain "bad" SA. Representation of inaccurate and incorrect data is essential for many aspects of information warfare. We may, for example, want to estimate the effects on operational outcomes of using information technologies to ensure that the knowledge base of an enemy is contaminated with erroneous and misleading information. It is equally critical to discover the effects of errata in our own knowledge base.

4.1.1 Measuring SA

> *A concept is of heuristic value in proportion to how well it can be measured. (Matthews, Pleban et al. 2000)*

The term "situation awareness" (SA) originated in the field of aviation human factors, but has been now adapted to the much different contexts, such as infantry operations. Of all the warrior system capability areas, SA is the most hypothetical construct. SA is not a "thing" that can be easily measured, like speed, numbers of casualties or even temperature, which while it can't be "seen", still admits to sensible (if arbitrary) ways to measure and express it. Temperature also responds in a reliable and predictable way to specific manipulations. SA is more akin to a construct like intelligence, which is defined mainly by the instrument used to measure it. There is still controversy (after decades of research and study) as to whether there is such a thing as generalized intelligence. Most people think that what we refer to as intelligence is really a complex constellation of traits, abilities, experiences, and knowledge. What one thinks about intelligence depends a lot on what one chooses to include in the definition and how one decides to measure it.

The concept of SA has much in common with such a view of intelligence; both are probably complex constructs embodying a range of traits and abilities. There is, however, much less agreement about how to measure SA, or even how to define it. There is no dearth of opinions; a multitude of surveys and SA collections have explored the diversity of viewpoints and the considerable body of research performed over the last 20 years. For example:

- Vidulich, Vogel, Dominguez and McMillan (1994) provide a number of papers and annotated bibliography of SA;
- Graham and Matthews (1998) present the collected papers from the 1998 Infantry Situation Awareness Workshop held at Fort Benning, Georgia;
- Endsley et.al (2000) documents the methods and findings of the Infantry Situation Awareness (SA) project for the U.S. Army Research Institute for the Behavioral and Social Sciences, conducted to develop a model and measures of SA for the unique aspects of the Infantry operational environment;
- Breton and Rousseau (2001) report the results of a literature survey of SA made in the context of the Soldier Information Requirements Technology Demonstrator (SIREQ-TD) Project for Defence Reseach and Development Canada;
- Banbury and Tremblay (2004) bring together 41 contributors studying and applying SA from a cognitive perspective;
- Breton, Tremblay and Banbury (2007) provide a critical evaluation of the psychometric properties of the available individual and team situation awareness (SA) measurement tools and techniques and to evaluate the applicability of the SA measurement tools found in the literature to C2 environments.

Breton and Rousseau (2001) summarize their survey of SA measurement tools:

From the analysis made on the SA measurement tools, it is interesting to note that most of them are used to compare a given SA content (actual SA) with a reference point (achievable SA). The measurement focuses on the identification of the actual SA content and the missing part of the SA content required to reach the achievable content. The identification of the missing part may have an important influence on the development of new support systems and training programs.

They note that Klein[64] presents four reasons why SA is important:

- SA appears to be linked to performance

- Limitation in SA may result in errors

- SA may be related to expertise

- SA is the basis for decision-making (with respect to Recognition Primed Decision-making)

and go on to state:

In all cases, the focus is on the extra benefits a hyperproficient agent can get by taking advantage of the situation. As stated before, the SA concept and its measurement were initially developed in the context of explaining operator or pilot mistakes that were hard to understand otherwise. A relative consensus emerged about SA being a helpful concept, namely to explain errors in complex and dynamic environments. It is only recently, on the impulsion of the U.S. Army, that research on SA is seen as a major contributor to a strategic advantage on the battlefield. For the infantry, the focus on SA is not so much on error reduction but on obtaining the strategic advantage in the field.

If SA is to be used as a criterion of an individual's capability, it must be measurable SA, or at least one must be able to convincingly argue that changes or additions to the Warrior Systems have the desired effect on SA. However, as Scott Graham of the U.S. Army Research Institute notes, actually demonstrating enhanced SA in systems is not easy, "*in part because of difficulties in operationally defining and measuring SA… in even establishing the level of baseline SA for system comparisons.*"(Graham, Matthews et al. 1998)

If one takes the aforementioned "*focus on the extra benefits a hyperproficient agent*

[64] Klein, G (2000). Analysis of situation awareness from critical incident reports. In M.R. Endsley, D.J. Garland (ed.s), Situation Awareness and Measurement (pp.51-71). Mahwah, NJ: Lawrence Erlbaum Associates

can get", then SA effects should be measurable in simulation experiments that compare such agents to less proficient agents. Fortunately, a simulation designer can develop agents at virtually any level of proficiency, given, of course, a framework and model constructs for representing that proficiency.

4.1.2 Definitions

Such a framework begins with an accepted idea of what SA is, and definition of its constituent elements. A representative sample of SA definitions from the above, and other sources, includes:

- Adam (1993) "knowing what is going on so I can figure out what to do".[65]
- Endsley (1995a) "...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future".
- Adams, Tenney and Pew (1995) "...the up-to-the-minute cognizance required to operate or maintain a system".
- Flach (1995) "...the congruence between the subjective interpretation of an event and objective measures of the actual event".
- Smith and Hancock (1995) "adaptive, externally directed consciousness".
- US Army TRADOC (1998) "ability to have accurate and real-time information on friendly, enemy, neutral, and noncombatant locations; a common, relevant picture of the battlefield scaled to specific level of interest and special need".
- Klein (2000) "perception of reactions to a set of changing events."
- Blackwell and Redden (2000) "the warrior's ability to quickly perceive and then discriminate between facets of the tactical environment, to accurately assess and reassess the where, when and why of that environment, to then know and understand the nature of the tactical situation and to extrapolate near term courses of action based on this understanding".

---

[65] Taken from (Breton and Rousseau 2001), citing (Adam 1993)

The four best known theories relating to SA/SU are Endsley's three elements (perception, understanding and projection) of SA, (Endsley 1995); John Boyd's Observe, Orient, Decide and Act (OODA) loop (Boyd, 1986, 1987); the Joint Director's of the Labs (JDL) Data Fusion Model (Llinas et.al. 2004), and the Dynamic Model of Situated Cognition (Miller and Shattuck, 2004, 2006). All of these support a view as an iterative process of data input, data assessment, decision-making and action. This view serves as a basis for simulation of SA/SU in agent-based modeling.

Many current SA initiatives employ Endsley's three elements of SA: data acquisition, comprehension of what those data might mean, and projection/prediction of future states and consequences. These three elements, shown in Figure 4-1, are the most widely accepted definition of SA, featured prominently role in such network centric warfare applications as collection of intelligence data, analysis of information from sensors, and the use of network operations for analysis and operational decision aid functions. Figure 3.1.2-1 also shows several of the factors that Endsley holds influence SA, both in terms of the variation in individual "*abilities to acquire SA given the same input*, and in *terms of the degree to which a given system provides the needed information and the form in which it provides it"*.

Endsley's definition can be compared to John Boyd's concept of the Observe, Orient, Decide, and Act (OODA) loop[66] as in Figure 4-2; and the Joint Director's of the Labs (JDL) Data Fusion Model as shown in Figure 4-3.

In all of these figures the role of the processes involved in "assessment"," "orientation" or "understanding" is highlighted, to emphasize the fact that these processes are those most in need of development from an ABM point of view.

[66](Boyd 1986; Boyd 1987)unpublished briefings

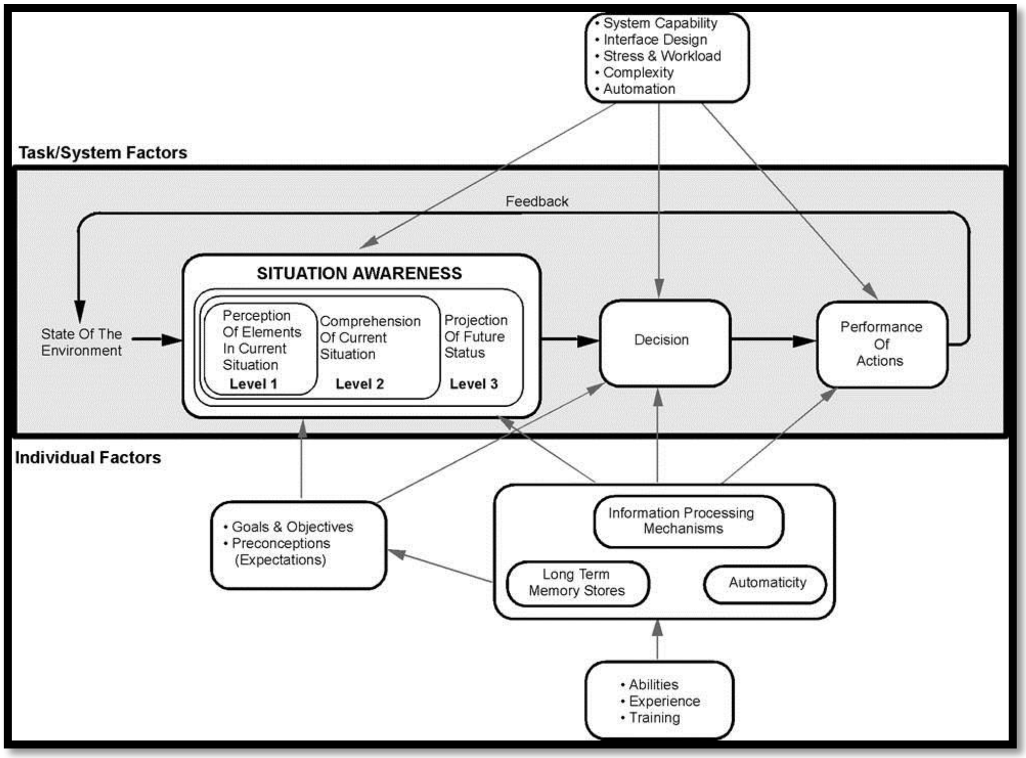


Figure 4-1 Endsley's Three Levels of SA[67]

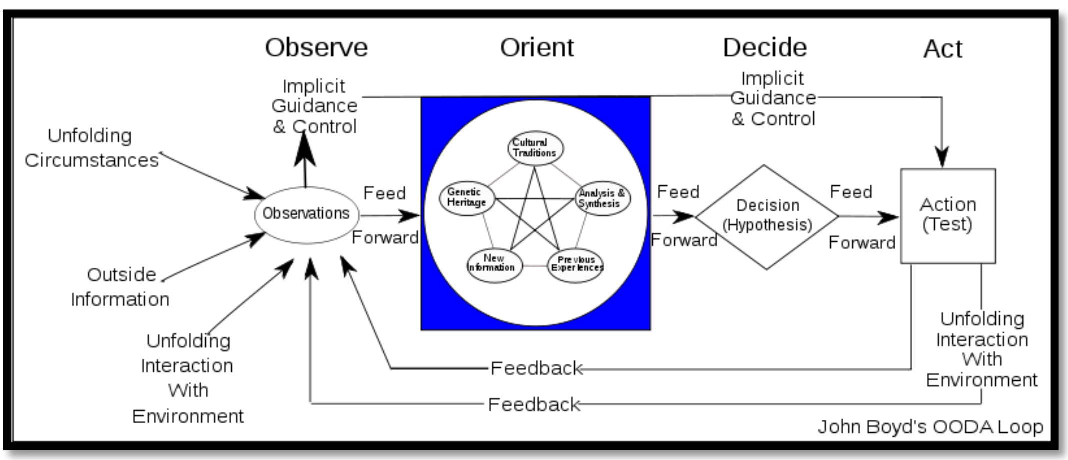


Figure 4-2 John Boyd's OODA Loop[68]

---

[67] Adapted from (Endsley 1995a).

[68] Full diagram of figure originally drawn by John Boyd, adapted by Patrick Edwin Moran and available under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation at http://en.wikipedia.org/wiki/File:OODA.Boyd.svg

The importance of these processes is reflected by the fact that it is now frequent US Army practice to distinguish between situation "awareness" - *immediate knowledge of the conditions of the operation, constrained geographically and in time*, and situation "understanding" - *the product of applying analysis and judgment to relevant information to determine the relationships among the mission variables to facilitate decision making.*[69]. While this distinction is probably valid, as stated earlier it is more convenient to blur the boundaries between SA and SU into Adam's more practical *"knowing what is going on so I can figure out what to do"*(Adam 1993) .

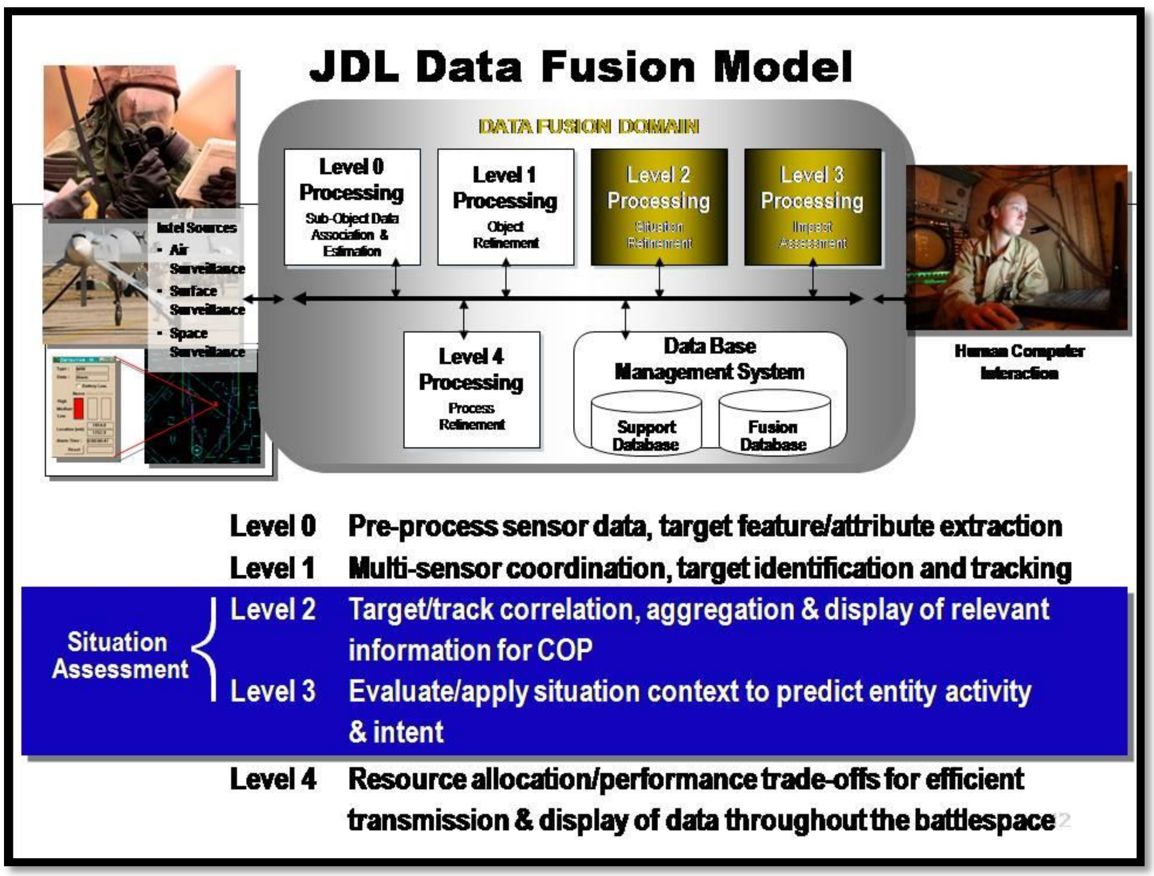


Figure 4-3 The JDL Data Fusion Model[70]

---

[69] US Army Field Manual 3.0 Operations, February 2008

[70] (Steinberg, Bowman et al. 1999), (Llinas, Bowman et al. 2004)

The Dynamic Model of Situated Cognition (DMSC), shown in Figure 4-4, was developed by Miller and Shattuck as an attempt to represent relationships between technology and humans in a system. This model represents the perception of ground truth as a function of sensor systems, the capture of those data by command and control systems, and the (possibly imperfect or erroneous) processing of these data into Endsley's three levels of SA.

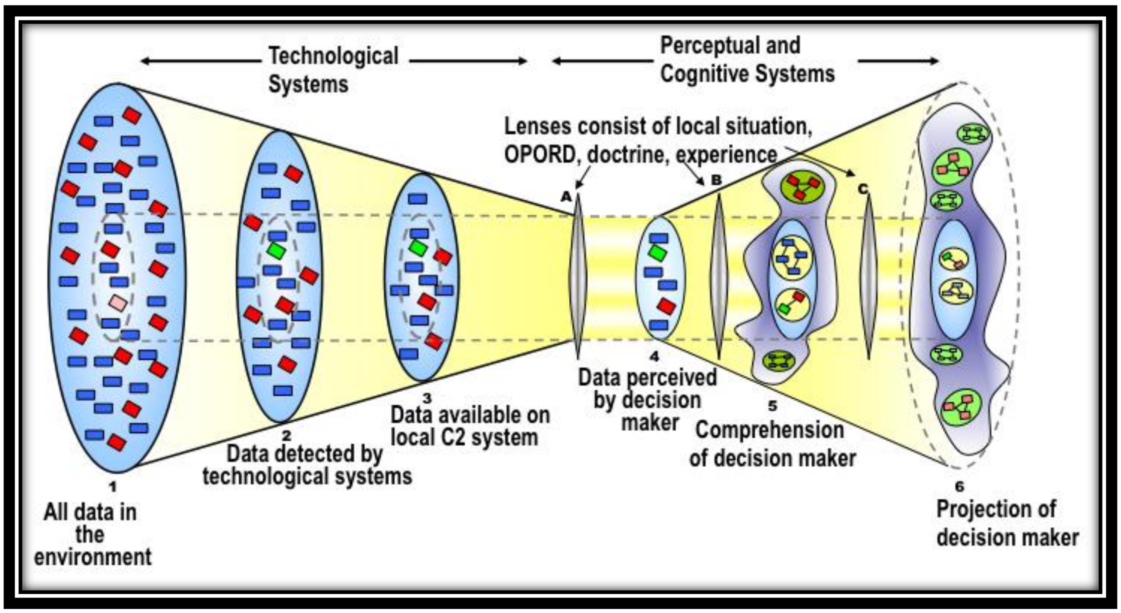


Figure 4-4 Miller and Shattuck's Dynamic Model of Situated Cognition [71]

As described by (Ntuen 2006) DMSC is a model

> *in which data flow from the environment, through sensors and other machine agents to the human agents in the system. This approach overcomes the biases which are inherent in analytical methods focusing almost exclusively either on machine agents or on human agents. The DMSC posits that there are various stages of technological and cognitive system performance (*as shown in 3.1.2-4*). On the technological side, all the data in the environment, data detected by technological systems (e.g., sensors), and data available on local command and control systems (C2; e.g., workstations) are included. Each of these stages includes a subset of what was included in the preceding stage. Building upon this*

[71] C2 System: Command and Control System – referred to in this dissertation as C4ISR
OPORD: Operations Order – mission as defined by higher command

technology are the perceptual and cognitive systems offered by the human operator.

Shattuck and Miller themselves describe the other half of the model as;

Ovals 4, 5 and 6 on the right side of the model represent the perception of data elements, the comprehension of the current situation (sometimes called a mental model) and the individual's projection of current events into the future. These three ovals correspond to situational awareness Levels 1, 2, and 3 in the scientific literature (Endsley, 2000). When the model was first introduced as a tool to assist in understanding laboratory simulations of military scenarios, the three lenses (A, B, and C) consisted of only four things: the local situation, the military operational order (OPORD), military doctrine, and the experience of the operator.

There was an acknowledgement that distortions in the lens could result in inaccurate perceptions (Oval 4), comprehensions, (Oval 5), or projections (Oval 6). It was recognized that once inaccurate data were accepted into any stage of the model, this inaccuracy would be propagated throughout the remaining ovals, leading to inaccurate conclusions and potentially faulty decisions. (Miller and Shattuck 2006)

Shattuck's and Miller's model is the most appropriate for the purposes of this research, although one can certainly see the commonality in Figure 4-1, Figure 4-2, Figure 4-3, and Figure 4-4 as reflecting a common view of simulation awareness/situational understanding (SA/SU). I believe Adam's definition of SA is applicable to all four of these models, and this common view of an iterative process of data input, data assessment, decision-making and action. This view serves as the basis for simulation of SA/SU in agent-based modeling. The difficulty for simulation programmers is, of course, in the definition of software constructs that describe how entities go about "knowing what's going on" and "figuring out what to do". Endsley and Boyd, provide a formal theoretical basis to support the model/simulation developer, but they don't really provide much more in the way of practical guidance in how to design and build software modules to implement the theory. The JDL model takes a more engineering oriented approach, but is too narrow in focus, reflecting its origins in a specific application, which was data fusion and filtering in the processing of incoming target tracks.

There is considerable interest in developing a workable framework to represent

C4ISR, information processing and SA/SU. The Human Factors and Medicine Panel of NATO's Research and Technology Organisation (NATO/RTO/HFM 2009) sought to "*demonstrate how HF* (human factors) *models could interface with operational models by enumerating the basic, existing models and how they fit into a generic framework.*" The NATO panel addressed a version of the stage model for information processing as shown in Figure 4-5:

> *SA capitalizes on the working and long term memories since it uses declarative knowledge (reference models stored in long term memory). It also exploits reasoning and sense making to understand information in the context of the reference model, and perceptual attention is focused to search for confirming information. Finally, reasoning is used to infer consequences of possible events, evaluating the expected outcome. In the decision making cycles, the metric for optimization is established by reasoning and by exploiting episodic memory. Obtaining SA thus is using the full range of cognitive functions and can be considered a higher order process, while cognitive research attempts to attribute steps in this process to distinct mental functions, as depicted in" Figure3.3.6-5. (NATO/RTO/HFM 2009)*

Again, as discussed above, for modeling and simulation this discussion comes down to the fact that a model requires quantifiable inputs, MOPs, quantifiable outputs, MOEs and some sort of functional mapping between the two.

Much soldier modernization equipment proposed or considered has as either a primary or ancillary purpose enhancement of soldier SA. Devices such as head-up displays, integrated communication and navigation systems, and sensor suites are often explicitly justified on the grounds that these systems will improve SA.

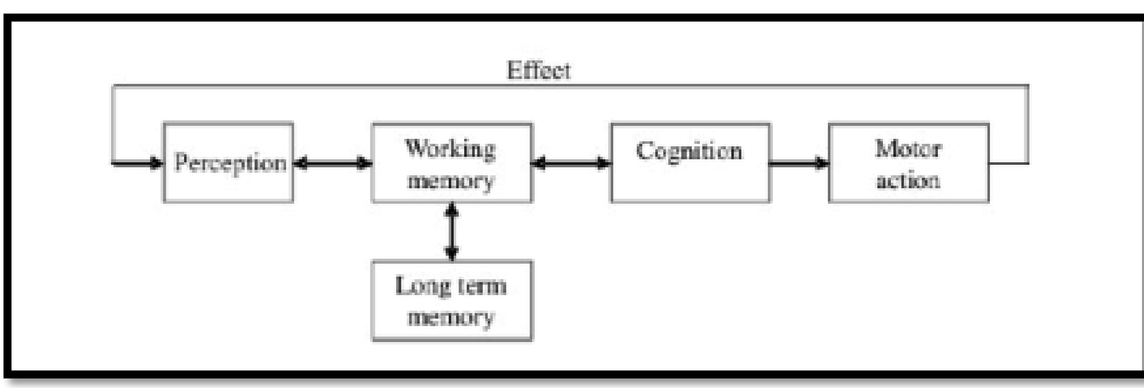


Figure 4-5 Modified Stage Model for Human Information Processing (NATO/RTO/HFM 2009)

Again, this model bears a strong resemblance those in the four previous figures, and, in fact the NATO HFM panel goes on to say:

> *Defining SA for military applications as the static spatial awareness of friendly and enemy troop positions is too simplistic and ignores many important militarily relevant details. … Activities with high levels of complexity and detail such as mission rehearsal, staff and procedural training, SOP and doctrine development will require more comprehensive models that capture and use SA in a more human-like manner. In other cases where the cognitive processes are not central to the problem being addressed, more mechanistic, AI approaches may suffice. Prediction of operational outcomes and mission rehearsal will also depend on the level of representation resolution, with perception and decision making becoming increasingly important as the simulation focuses on individuals or small groups. Credible SA will require adequate models of perception to explain and predict the effects of the cognitive processes involved in each of Endsley's three stages …. This is different from but overlaps current AI approaches to CGF that tend to rely on a complete knowledge of the environment and precise extrapolation of events. For example, it has been observed in practice that during times of stress and uncertainty, there is frequently a failure to consider an adequate number of facts to ensure accurate SA, including failure to verify assumptions in making decisions, failure to weight information on its quality, failure to interpret information (Level 2) and failure to make predictions (Level 3: Fallesen, 1993[72]).*
>
> *Perceptual and attention models should reflect human capabilities, limiting HBR to information a human operator would have access to rather than providing complete knowledge of the environment or allowing precise extrapolation of events. Such models may result in missed or misperceived environmental cues, with a consequent error in judgment based on correct reasoning with faulty data. To make the resulting performance plausible, purely stochastic approaches are unlikely to be an adequate representation of such processes for most applications, underlying the need for models from the human sciences in the relevant domains.*
>
> *Beliefs in the state of the world that evolve from the SA process could depend on some form of reasoning either about what will happen next or how current events evolved. Alternatively, pre-programmed or instance-based memory models may come into play to predict future events as representations of expert*

[72] Fallesen, J. J. (1993). Overview of Army Tactical Planning Performance Research (No. Technical Report 984). Alexandria, VA.: U.S. Army Research Institute for Behavioral and Social Sciences.

knowledge or experiential learning in decision making. Planning models and exploration (mental simulation) to select appropriate responses should reflect human biases as well as capabilities rather than exhaustively searching the problem space for an optimal solution (NATO/RTO/HFM 2009)

A robust framework for SA/SU will have to support the above mentioned techniques, as well as such other decision-making models as neural networks, fuzzy cognitive maps, Naturalistic Decision-Making and Recognition Primed Decision-Making. Such a framework would also have to incorporate or (or at least accommodate) models of perception, search and target acquisition, battle damage assessment, self-awareness (psycho-physiological state, available resources, mission goals and tasks) and the like. All of these models consist of both processes and state variables, some of which (most of which?) are extremely context (i.e., situation) dependent. While there is considerable dissension within the SA community as to whether SA is itself a process/ set of processes or a state/set of states, representation of SA/SU and how it affects operations will certainly require both. The argument as process or state comes into play, however, in the selection of appropriate MOPs and MOEs.

One of the major difficulties in working with SA is to avoid confusion between SA knowledge and the underlying cognitive processes such as perception, memory, attention, categorization, or decision-making. The difficulty is particularly acute when SA has to be measured. In agreement with Adams, Tenney, & Pew (1995), Endsley limited the term "situation awareness to the achieved knowledge (state) about a situation. She proposed the expression "situation assessment" to designate the cognitive processes that produce the knowledge (state).

....

In the context of the development of a SA definition, one is then left with a double problem. On the one hand, if SA is a state, it is essential to give a precise definition of the knowledge that defines the state. There should be a certain mapping between a situation schema and a knowledge one. If one is to improve SA, the elements of the situation critical for SA should be specified, and the SA content definition should follow from these elements. On the other hand, if SA depends on a set of processes that are not an intrinsic part of SA as a state but on which SA depends, it becomes important to specify which processes are essential to SA. (Breton and Rousseau 2001)

4.1.3 The Dynamic Knowledge State

Finally, this chapter addresses representation of knowledge, which must further distinguish between an individual's perceived world view and ground truth as the key to augmenting extant representations of SA/SU and decision making. Such representation requires design, development, and implementation of three inter-related elements:

- knowledge structures to characterize each entity's ego-centric knowledge of the operational environment;
- algorithms and heuristics to populate, maintain, and update those structures; and
- inference schemes creating information from perceptual data and knowledge histories and the employment of this information to make and execute operational decisions.

(Davis, Shrobe et al. 1993) describe knowledge representation as among other things: " *a set of ontological commitments, that is, an answer to the question, In what terms should I think about the world?*" as well as " *a fragmentary theory of intelligent reasoning expressed in terms of three components: (1) the representation's fundamental conception of intelligent reasoning, (2) the set of inferences that the representation sanctions, and (3) the set of inferences that it recommends", and "a medium for pragmatically efficient computation… supplied by the guidance that a representation provides for organizing information to facilitate making the recommended inferences*".

The agent-based modeling community as yet lacks an effective ontology for agent HBR, and such an ontology is a basic requirement for any integrative view of cognitive architectures.

4.1.4 Knowledge State

Maintaining SA/SU in simulations of military operations implies that such awareness will result in better informed (and therefore one hopes more positive) actions to shape the battle space in order to achieve particular goals. An individual's knowledge state is thus crucial as both a repository for information gleaned from the environment, and as the constantly changing raw material from which tactically appropriate actions may be generated. As the knowledge state affects the decisions and actions taken by the soldier, it affects simulation events, thereby potentially modifying the incoming stream of stimuli in subsequent play. Rules, heuristics, fuzzy modeling techniques or other approaches can be used to integrate the contents of the knowledge state, and to drive the behavior of the soldier, supported by decision feed back and affecting sensory and perceptual processes (by altering scan patterns) and cognitive processes (by altering attentional allocation)[73].

For a constructive simulation to represent an individual's knowledge state (or at least the most significant aspects thereof) is ambitious but within reach based on existing knowledge about human information processing and available simulation technology. The knowledge state is envisioned as a dynamic register in which relevant information is maintained and processed by the individual combatant. Sensory events are posted to the knowledge state based on the physical characteristics of the sensory stimuli, the soldier state, and prevailing environmental conditions. Based on the relevance of the stimuli in the scenario, attention and cognitive processing resources may be allocated to the newly acquired sensory data to maintain it in the register, or the stimulus may simply pass through our senses and not be retained, as most are not. Pew and Mavor suggest that to "Include explicitly in human behavior representation a perceptual "front end" serving as the interface between the outside world and the internal processing of the human behavior representation" is an important short-term goal (Pew and Mavor 1998).

A wide range of artificial intelligence techniques are available to represent extracting relevant knowledge from the events that do flow into the knowledge state, thereby enabling the projection of that state into the future, by triggering actions or by adding new elements to the knowledge state. Pew and Mavor list many technologies, including rule-based expert systems, case-based reasoning, and Bayesian belief

[73] See for example: (Middleton 2008; Langley, Laird et al. 2009)

networks, which could be employed to support such information processing (Pew and Mavor 1998).

It is important to note that it is not necessary to develop a general model of human cognition to support the representation of situational awareness at a level compatible with the analysis of most operational questions. Situation awareness requires information processing, but is a far more limited construct than the generalized human reasoning systems commonly discussed in the field of artificial intelligence. According to Pew and Mavor, "The Panel[74] cannot overemphasize how critical it is to develop situation-specific models…The situations and tasks faced by humans in military domains are highly complex and very specific. Any effective model of human cognition and behavior must be tailored to the demands of the particular case." (Pew and Mavor 1998)

Focusing specifically on representing situation awareness in relevant operational contexts will help make these problems tractable.

Developing a workable representation of SA/SU in simulations designed to address dismounted operations is an important step in improving analytic capabilities in this domain. As Pew and Mavor suggest, "Once high priority modeling requirements have been established, we recommend sustained support in focused areas for human behavior model development." (Pew and Mavor 1998, p. 336). The concept of situational awareness clearly captures something that many people think is essential about warfare; incorporating it into simulations can be the first step toward a new generation of human-centric modeling and simulation tools.

Representing SA/SU and an individual's dynamic knowledge state in a simulation system will require a clearly articulated conceptual model of SA/SU, supporting data,

[74] The Panel on Human Behavior and Command Decision Making: Representations for Military Simulations, established by the National Research Council (NRC) in 1996 in response to a request from the Defense Modeling and Simulation Office (DMSO) of the U.S. Department of Defense.

and a strategy for computational representation within the simulation. At least some of the information of which SA/Su is composed is already present in many simulation systems, or may be easily inserted in them. Representation of the complex interrelationships among modulating factors and sensory and perceptual filters, which will help determine what information is available in the knowledge state, will require an ambitious synthesis of existing representation strategies and scientific knowledge. Connecting the knowledge state to the actions of the simulated entity in a way that captures the intelligent and autonomous features of human behavior that are missing from current approaches presents the most significant challenge. Finally, new metrics, especially for information acquisition and information processing, must be identified and implemented.

Building a system that could produce and track the expected knowledge state of an individual combatant in a dismounted combat scenario in a manner consistent with what is known about human sensory capabilities, information processing and the performance of soldiers under operational conditions is no small task. No such system yet exists, nor is it clear that such an all-encompassing system is even necessary to support analysis of military operations. It is clear, however, that representation of information processing and decision-making is of ever increasing importance to operational analysis.

The simulation developed for this dissertation provides an architecture that supports modular inclusion of specific aspects of human information processing and the application of that information to decision-making and behavioral response to the operational environment. This architecture does not represents a unique, or even particularly novel capability with respect to these aspects of human behavior representation, but the modules implemented and the treatment of error and uncertainty they incorporate do provide a novel and important addition to the current state of the art with respect to representation of human behavior in simulation.

The architecture employed can be adapted to augment the representation of SA/SU in many current simulations. The key to such augmentation is the incorporation of model features that further distinguish between an individual's perceived worldview and ground truth. Incorporating these features requires the design, development, and implementation of three inter-related elements:

1. data structures to characterize each entity's perceived knowledge of the operational environment;
2. algorithms and heuristics to populate, maintain, and update those data structures; and
3. inference schemes employing these data to represent operational decisions.

This modular architecture is consistent with Boyd's OODA loop (Boyd 1986) as shown in Figure 4-6.

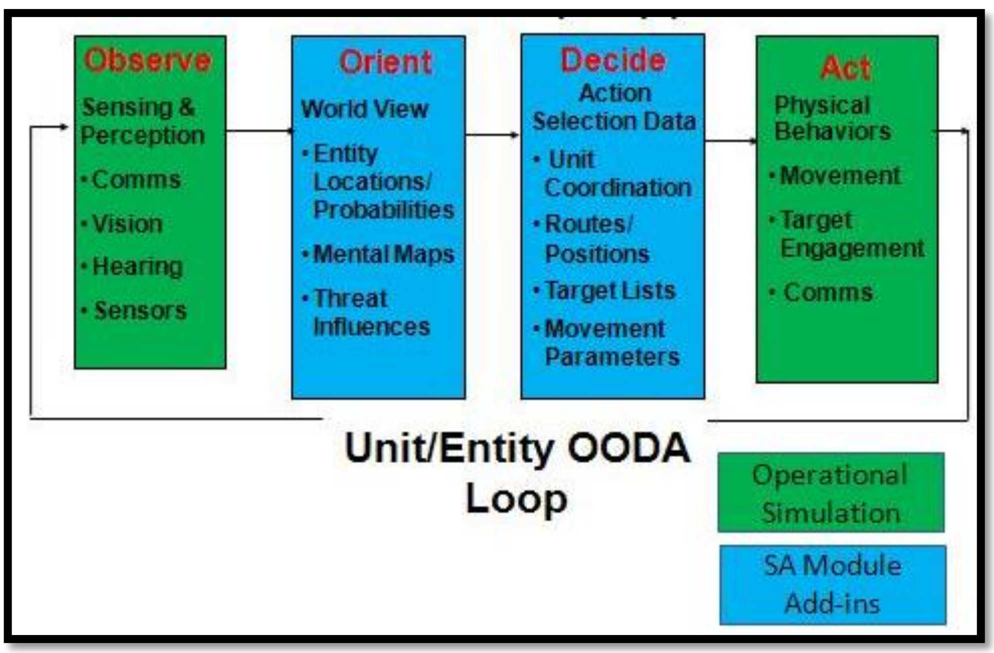


Figure 4-6 Modular OODA Loop Approach (Middleton 2010a)

In this approach, the three elements listed above are encapsulated in modules that constitute the "Orient" and "Decide" components of the OODA Loop, the blue boxes of the figure.

This approach provides a controlled interface between new SA/SU capabilities and extant simulation processes. The sense/perception processes native to host simulation entities allow those entities to "observe" their virtual world as before, providing data on the simulation environment and the objects in it. The new "orient" modules interpret those data though (potentially imperfect) filters to populate and update world view data structures unique to each entity. As an example, an entity may observe another entity that it previously would have identified according to its force association and any threat value. New "orient" filters could "translate" entity sightings into levels of evidence for associating that entity with a given force or threat intent. Similarly such filters could add imprecision and/or error to the sighting entity's perception of the sighted entity's location. Inference routines could evaluate evidence from multiple sources, resulting in attributes of the sighted entity described as degrees of membership in fuzzy sets as opposed to the generally crisp (e.g., friend or foe, within range, at objective) options currently available.

The "oriented data" is now information that is used by the decision logics of the "Decide" module to choose and direct those entity behaviors deemed most likely to achieve entity/group goals. The host simulation "Act" capabilities carry out these behaviors and determine effects on other entities and the environment.

I believe that a simulation system that includes situation awareness (and situation understanding) is applicable to a much wider range of mission environments than a system that relies more narrowly on measures associated with the lethality of weapons systems. Situation awareness could be a useful measure in a conventional dismounted combat scenario as a way to provide so-called "one-sided metrics". These are metrics of performance that are of interest even if no contact is made with the enemy, and are especially relevant to peace support operations. The availability of such metrics could help establish analytic isomorphism with training environments, where one-sided metrics are routinely applied, and enable valuable cross-fertilization between the training and simulation worlds. SA/SU thus offers an opportunity to link the worlds of training, operations, and simulation by creating a common set of measures useable in all three domains.

Collective synergy is particularly important in the world of SA/SU, where information fusion and synthesis are key ingredients for comprehension. SA/SU affects the factors that incorporate any cognitively related use of military systems or system components, from identification of friend or foe to complex course of action decision-making. Our MOPs and MOEs challenge is figure out how to “keep score” with respect to SA; how to quantify SA in terms of dynamic, measurable attributes of individuals and organizations (the things MOPs measure) and to determine its role in behavioral, and ultimately, operational outcomes (the things MOEs measure.)

Representation of the knowledge state can also add sophistication to the tactical play and decision-making of simulated entities. Rules, heuristics, fuzzy modeling techniques or other approaches can be used to integrate the contents of the knowledge state into the behavior of the soldier. As the knowledge state affects the decisions and actions taken by the soldier, it affects simulation events, thereby potentially modifying the incoming stream of stimuli in subsequent play. The decisions taken may feed back and affect sensory and perceptual processes (altering scan patterns) and cognitive processes (by altering attentional allocation).

5 COGNITIVE ARCHITECTURES AND THE FUNCTION OF MIND

My own views on representation of cognition are heavily influenced by (Franklin 1995). He addresses cognitive architectures (among a number of other things) from the perspective of exploring what is meant by the concept of "mind". He looks at four broad approaches to the study of mind: psychology, artificial intelligence, neuroscience, and mechanisms of the mind. This last is his term for "*artificial systems that exhibit some properties of the mind by virtue of internal mechanisms*", and which he considers replacing with the term robotics. He takes a number of positions on the concept of mind:

- *Mind is better viewed as a continuous as opposed to a Boolean notion*. Since Franklin views the principle function of mind as "deciding what to do next", a continuous view of the decision process cannot be limited to the concept of totally irreversible one-time choices, but must also incorporate feedback loops to regulate behavior.
- *Mind is aggregate rather than monolithic*. Cognition includes processes for communication between long-term memory and short-term memory, between different senses/perceptual processes, and between the different constraints of multiple behaviors taking place in parallel.
- *Mind is enabled by a multitude of disparate mechanisms*. The research in this dissertation is in concert with Franklin's lack of support for unified theories of cognition, a one-size-fits-all cognitive architecture is not a good idea. (Sowa 2000) refers to the sum of knowledge in people's heads as "*fluid, heterogeneous, ever changing, … better characterized as* ***knowledge soup****". The soup, may contain many small chunks, corresponding to the typical frames,*

rules and facts in AI systems; it may also contain large chunks that correspond to entire theories. The theories should be internally consistent, but they may be inconsistent with one another".

- *The overriding task of mind is to produce the next action*. Here Franklin cautions not to read too much into the terms "task" and "next". "Next" is not restricted to single, discrete kinds of action, and task simply means "*Producing actions is just what minds do. A consequence is that **minds are properties of autonomous agents***."

Cognitive architectures provide the framework or structure for implementing the "mind" of an autonomous agent. Desr4 Such a framework must then be "integrated" with other aspects of an agent-based model or simulation, an idea discussed at length in Pew & Mavor. Such integration requires a common set of assumptions and consistency of scale. Byrne speaks of the integrative nature of cognitive architectures as meaning "*they include attention, memory, problem solving, decision making, learning, and so on. .... Instead of asking "how can we describe this isolated phenomenon?" people working with cognitive architectures can ask "how does this phenomenon fit in with what we already know about other aspects of cognition?*"(Byrne 2003). Such thinking is also evident in Ritter's discussion of hybrid architectures, cognitive architectures with more than one type of knowledge representation:

> *These architectures can be created in three ways. One way is by adding symbols to a connectionist representation. … Another way is by adding sub-symbolic representations to a symbolic architecture, for example, declarative memory strength in ACT-R and reinforcement learning in Soar 9. Here, the weighting you give to the rules and declarative memory elements is essentially a way to create a symbolic architecture that has changes that are small and gradual. Most hybrid architectures are realized in these two ways, but there are also examples where one of the levels is a genetic algorithm, fuzzy logic, or other representation. Hybrid in this case does not mean symbolic and sub-symbolic, but typically symbolic and some higher level or orthogonal representations or process is used to supplant the strengths of the base architecture. (Ritter 2008)*

While a popular trend in cognitive architectures has been on symbolic methods (e.g., Soar and ACT-R, for this dissertation there is more utility in techniques such as Naturalistic Decision Modeling (NDM), and in the potential for fuzzy logic and neural networks. These latter techniques can more easily and more fully capture the continuous feed back loops that characterize much of human behavior. In particular, they appear more amenable to representation of the OODA loop and the desire to defeat an adversary's OODA loop.

The capabilities of cognitive architectures that are most important for addressing this research, i.e., those of most immediate concern to developing/enhancing autonomous agent simulation of military operations, are defined below.

5.1.1 Decision making and choice

Decision-making is frequently looked at as a discrete event, with alternatives considered, a choice made, and that choice acted on. In the world of discrete event simulation this view is certainly justified at some level, even continuous processes are broken into atomic chunks of activity, and the scheduling of the next event represents a decision of some sort. It is useful, however, to consider decisions as falling into three broad, albeit overlapping, categories:

- prescriptive plans, e.g., course of action selection, scheduling and coordination of entity/unit tasks;
- reaction to unanticipated events, e.g., detour around obstacle/threat, engage an adversary, call for fire; and
- modification of current behavior parameters, e.g., how fast to move, in what direction, choose which targets to engage when, adjust aim points and rates of fire.

Of these, the first is best related to capabilities for problem solving and planning, which, as far as this dissertation is concerned, is a "bridge too far" with respect to the current state of AI for simulation. At present planning is best dealt with at the set-up stage of simulation scenarios, where an analyst outlines mission goals, as well as commander's intent and rules of engagement (ROE), which act primarily as constraints on agent behaviors as those agents try to achieve their goals while reacting to the dynamic circumstances of their environment.

Langley et. al. states that “to support decision making, a cognitive architecture must find some way to represent alternative choices or actions… and some process for selecting among those alternatives”. (Langley, Laird et al. 2009) Again, as can be seen with respect to NDM, this statement is not precisely true. NDM decision-making can be seen as pattern or prototype recognition, with some possibility for adjusting prototype features or parameters. One could refer to these patterns or prototypes as decision alternatives, but NDM expressly eschews the analytic assessment of alternative courses of action that I see as the core of Langley et.al. Furthermore, there is little or no consideration of making “bad” decisions in Langley, or consideration of the imperfections in both SA/SU and/or that may be due to time pressure and/or stress. (Zimm 1999) describes a number of problems observed in decision-makers under stress, including:

- changing from deliberative to reactionary modes;
- relying on a fraction of available information, with a bias towards that which is familiar and corresponds to earlier perceptions over that which is relevant and/or unexpected;
- making more mistakes but being less likely to acknowledge them; and
- increasing micro-management of subordinates.

Representing these tendencies towards “imperfect” decision-making and even irrational decision making (panic) is critical to providing a robust simulation test bed for SA/SU technologies.

5.1.2 Perception and situation assessment;

Stimuli from the environment impinge on the simulated soldier. In reality, individuals are constantly bombarded with stimuli affecting all five traditional senses. In the simulation, the only stimuli that need be considered are those that are generated within the simulation and tracked by entities as behavioral and decision-making cues. These stimuli include visual and acoustic weapons signatures, various kinds of communications, and perceptions about the state of the physical environment.

The extent to which stimuli need to be represented is a function of the degree to which the operations of various types of sensors and signal processing need to be explicitly considered, and such consideration is in turn a function of the types and degrees of information processing that are to be considered by the simulation. Stimuli that are present in the environment may or may not be relevant to individual agents depending on many factors, including the intensity of the stimulus, the contrast between the stimulus and the background, the attentional state of the agent, and many others. Measures of effectiveness or metrics associated with sensory and perceptual processes will include time and accuracy of detection and acquisition. Sensory and perceptual processes may be modulated by such factors as scene clutter, suppression, fatigue state, and simultaneous task demands. Stimuli that pass this stage are those that are "noticed" and subjected to further processing.

Once stimuli have been sensed, they are subjected to further processing. If the stimulus is a message with semantic content, it must be comprehended. The delay and accuracy of processing such messages is a metric associated with this stage of information processing. If attention is required for another task ongoing simultaneously, processing may take longer. Other modulators might include such factors as sleep deprivation, fatigue, or other environmental factors.

Any architecture that supports sensor-related perception "must confront issues of **attention** … and deal with the issues that sensors are often noisy and provide at most an inaccurate and partial picture of the agent's surroundings". (Langley, Laird et al. 2009) Imperfect perception and inferences from a history of such perception will provide each agent with an individual ego-centric view of the environment, which will serve as the basis for that agent's conscious selection of behaviors. Since that ego-centric view differs from both ground truth and the views of other agents, each individual agent will make mistakes, and potentially different mistakes from those of other agents.

Target detection and engagement is an important feature in simulations that explore adversary interactions at he small unit level. These simulations, however, look at perception primarily in terms of the effectiveness of either biological or mechanical sensors[75]. The contributions of the human operator have not been considered to any great extent, target detection and other perception functions are treated as basically sensor-driven processes, with human operator effects represented only implicitly through empirically derived sensor performance distributions or explicitly through human-in-the-loop interaction with a sensor. Now, however, the need arises to also simulate the human as the perceiving entity, to represent explicit operator/sensor interaction for autonomous agents using their cognitive capabilities monitor and control sensor operation, and in particular the role of perceiver's dynamic situation awareness in facilitating the scan and focus processes.

In this view the individual functions as an information-processing and decision-making entity, communicating and coordinating with others, and continually evaluating the battlefield dynamic to decide when, where, and how to sense, move, shoot, and communicate. Human operators direct and focus sensor functions in response to previously known targets, intelligence data, present and past cues, observed or projected enemy behaviors, with all of these data fused into "situation awareness". The individual must correlate signature data with information from other sources, interpret all these data in the context of a continually varying experiential frame of reference, and fuse them with data inferred from secondary observations and/or projected from past sightings rather than overtly perceived from direct continuous surveillance.

The diverse elements of the individual's situation awareness impact sensor operation by suggesting where and how to "look", and are often a key as to what is actually "seen". Thus the history of past observations and projection of anticipated events play key roles in both the perception process itself and in the interpretation of

[75] See for example models such as the Integrate Unit simulation system (IUSS), the Infantry Warrior Simulation (IWARS), Pythagoras, the Map Aware Non-uniform Automata (MANA), COMBAT XXI or references such as (O'Kane 2004; Schumacher 2005; Darken 2007; Darken 2007; Kleiner, Carey et al. 2007; Matsopoulos 2007; Adam, Carter et al. 2008)

what is perceived. The ability to even discern or distinguish perceptual cues is biased by past history and current expectations.

5.1.3 Execution and Action;

As mentioned above, an integrated cognitive architecture must share common assumptions with the behaviors or actions of the autonomous agents it supports. It must also share data about the problem space with these behaviors, if only to direct which actions to undertake and how to perform them. Such data structures include course of action options and selection as well as behavior parameters, which include targets and target priority lists, types and rates of fire, shoot/no shoot decision thresholds for engagement; routes and waypoints or direction vectors for movement, speed and movement formations.

5.1.4 Interaction and Communication.

Effective autonomous agents are often described as needing to be both insightful - capable of inferring the intentions of others, determining the desires and plans of other agents, and social - able to share goals, cooperate with or coerce other agents. These capabilities could well be subsumed under the immediate proceeding set, as simply special type of execution and action. They also require specialized common data structures, such as situation reports (SitReps) to other units, especially command units, directives to subordinates, unit coordination communication with verbal, written, visual, or tactile, and requests such as call for fire or other support.

5.2 Military Decision-making

> *Commanders continuously combine analytic and intuitive approaches to decision-making to exercise battle command. Analytic decision-making approaches a problem systematically. The analytic approach aims to produce the* ***optimal*** *solution to a problem from among the solutions identified. The Army's analytic approach is the military decision-making process (MDMP). In contrast,* ***intuitive decision-making*** *is the act of reaching a conclusion that emphasizes pattern recognition based on knowledge, judgment, experience, education, intelligence, boldness, perception, and character. This approach focuses on assessment of the situation vice comparison of multiple options. It relies on the experienced commander's and staff member's intuitive ability to*

*recognize the key elements and implications of a particular problem or situation, reject the impractical, and select an **adequate** solution.(TRADOC 2008)*

The TRADOC Field Manual (FM) reflects two basic kinds of decision-making in command and control. The first is characterized by the deliberative approach of the MDMP, as shown in Figure 5-1. It provides rational, deliberate, time-consuming sequential kinds of optimization in which multiple clearly different course of action (COA) alternatives are posited, assessed, and the "best" one chosen. In addition to the MDMP, examples of the more deliberative optimizing techniques include Multi-Criteria Decision tools, Multi-Attribute Utility Theory and Value-Focused Decision-Making [76]. Such tools may use influence diagrams, decision trees, and hierarchies of objective to capture all of the pertinent factors involved in a decision

[76] See, for example (Triantaphyllou, Shu et al. 1998) (Clemen and Reilly 2001), (Keeney 1996)

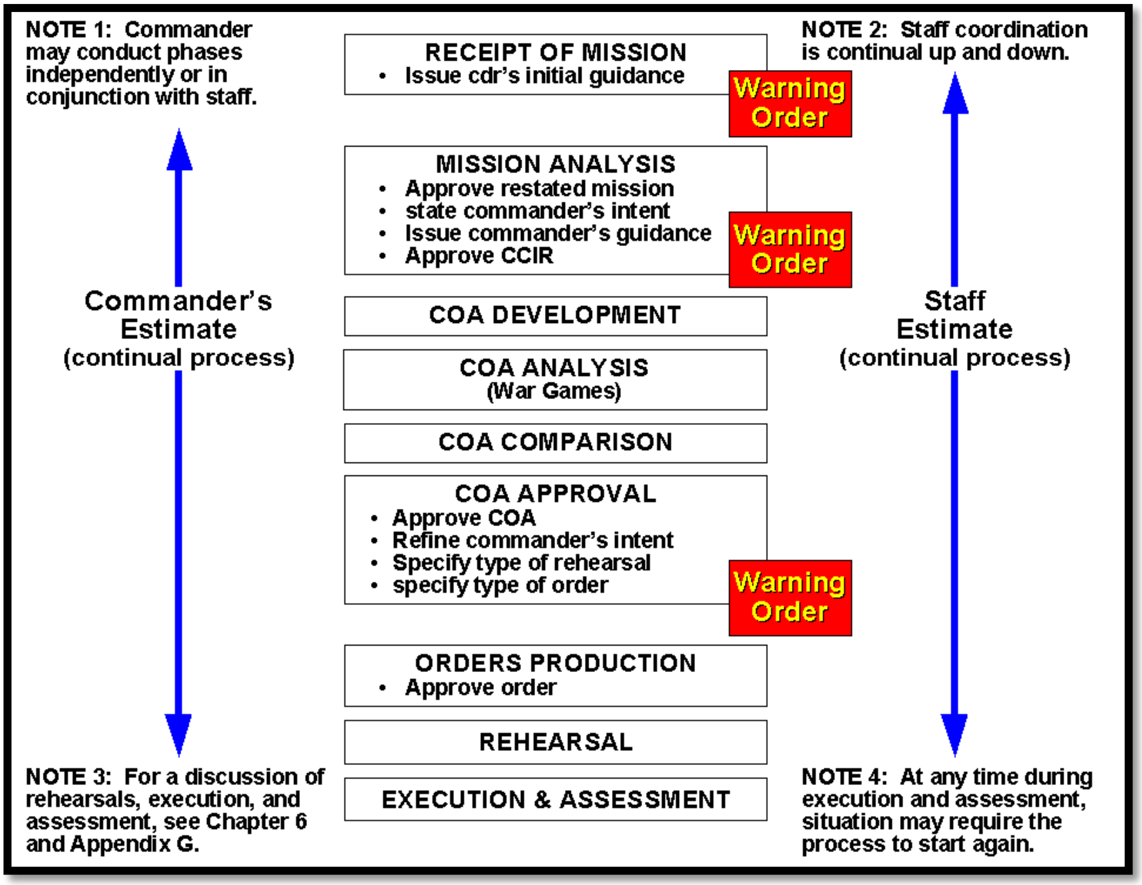


Figure 5-1 The Military Decision-Making Process (from FM 44-100)

The alternative is decision-making characterized by rapid, dynamic adjustment of behavior following known patterns. It is more reactive than deliberate, emphasizing the first usable alternative rather than trying to pick the best one. It must work with uncertain, often rapidly changing data, and supports continual matching of expectations against developing outcomes followed by adjustments based on the quality of that match. It is satisficing[77] rather than optimizing.

---

[77] A term used in economic theory to describe how people make rational choices between options open to them and within prevailing constraints. In 1957 Herbert A. Simon argued that decision-makers can rarely obtain and evaluate all the information which could be relevant to the making of a decision. Instead, they work with limited and simplified knowledge, to reach acceptable, compromise choices ('satisficing'), rather than pursue 'maximizing' or 'optimizing' strategies in which one particular objective is fully achieved.(Simon 1959) Satisficing is sometimes also referred to as a strategy of disjointed incrementalism. The adoption of satisficing models instead of maximizing models of behaviour has been found useful in the theory of the firm and

This alternative is perhaps best typified by Naturalistic Decision Making such as Klein's Recognition Primed Decision Theory[78]. In general, the first kind of decision-making addresses less time critical activities such as planning, consequence management through pre-positioning of assets and other logistics considerations, while the second is more characteristic of crisis management, tactical response, and rapid reaction. This is reflected, for example, by the greater importance of MOPs and MOEs related to the quality, completeness, and confidence of data in the one hand, and the emphasis on timeliness and salience of data on the other.

---

corporate behaviour. For example, to maximize its profits a firm needs complete information about its costs and revenues, which in practice is available only after the event. Satisficing models replace the search for the optimum outcome, which may be unattainable, with rules of thumb and compromises which work well enough. from (Marshall 1998).

[78] See for example Naturalistic Decision-Making (NDM) (Lipshitz, Klein et al. 2001 & Salas 2001) and Recognition Primed Decision-Making (RPD) (Klein 2008).

6 MODELING MOVEMENT UNDER IMPERFERCT SA/SU

This dissertation focuses on mobility as the Warrior Systems capability area in which to explore the operational effects of imperfect SA/SU. The mobility capability area combines the physical behaviors of movement, with a rich spectrum of fairly well understood decision processes, from route planning and route following to various way finding activities. It features the dynamic use of data as individuals update position data and their understanding about their environment as they move about in that environment. It allows ample exploration of the ways in which data may be incomplete, uncertain and incorrect data. It supports definitive measures of the cumulative effects of imperfectly perceiving data, imperfectly understanding what those data mean, and imperfectly estimating course of action outcomes based on them. Finally it presents a significant challenge in trying to simulate an easily understood but difficult to model mobility-related phenomenon, that of being spatially "lost" while trying to navigate in a simulation world.

Being "lost" is in fact a very complex phenomenon. It involves determination of where an entity actually is, where it thinks it is, where it is going and where it thinks it is going. The entity can be in one of several states: knowing where it is and where its going; being uncertain about one or both of these factors; being actually wrong about one or both of the factors; being right about them, but thinking its wrong and vice versa; recognizing or not recognizing whether it is wrong (i.e., knowing that it is lost), and having some set of potential corrective actions to take in the eventuality of being lost or believing it's lost. All of these correspond to various levels of SA/SU.

6.1 Mobility Basics

Mobility reflects the ability to move and control terrain, as measured by speed and/or distance. Current M&S capabilities support definition of individual movement speed in response to such as factors terrain grade and trafficability, with limited consideration of physiological state (as measured by core temperature, max VO2, heart rate, etc.) (Mastroianni and Middleton 1996). Algorithms exist to support dynamic obstacle avoidance and some use of terrain features, and the use of path finding/planning tools such as Dijkstra's algorithm and the A* algorithm are fairly common (Reese and Stout 1999; Reece 2000; Reece 2003) .

These M&S capabilities make it possible to define intelligent agents that can regulate, either through self-pacing or in response to command guidance, their speed, their movement formations, the frequency with which they take rest breaks, and other aspects of dismounted movement. Simulations employing such agents can be used to optimize the trade-offs between energy expenditure, threat avoidance, and mission demands. (Mastroianni and Middleton 2001), (Middleton 1996; Middleton 2002) At present, a key concept for the Warrior System mobility area capability is the network centric force and its employment of global positioning systems (GPS) and Geographic Information Systems (GIS) (Murdock 2002), (Mitchell 1999). Again, in order to assess the benefits of these technologies, baselines of imperfect information are required, for example, permitting entities to "get lost" as a response to poor situation awareness, bad decision-making or combination thereof.

6.2 Intelligent Agents and Terrain

Intelligent agents must react to, and make use of, terrain and cultural features in decision-making and other simulated behaviors. This imposes a need for information not found in traditional terrain databases and relating to attributes of environmental features that impact on decision-making and/or physical interactions with those features. Current efforts in IWARS development, for example, are addressing this information gap by enhancing data bases to include "semantic" terrain, enhancing the attribute lists of current features and/or creating new terrain artifact objects to represent such features as: areas of cover and concealment, areas of tactical importance, danger areas, choke points, etc. (Stanzione and Johnson 2006). The information provided by such attributes would be typically known to a human observer, but not easily inferred by an autonomous agent in a constructive simulation. The benefits of augmenting terrain with these data include more robust adversary behaviors in training simulators and more robust modeling of intelligent agents on all sides in constructive simulations. In particular, these enhancements support the position object concept, which is central to how autonomous agents within IWARS interact with terrain (NRDEC 2005).

6.2.1 Terrain Representation for way finding, route planning, and navigation

Representation of military mobility and maneuver, or any other kind of directed movement, must define a route search space, which needs to be characterized with appropriate "cost" functions that relate both static and dynamic aspects of the movement space to an underlying representation geo-spatial environment.

As in virtually any kind of heuristic optimization method, route-planning methods may be improved by cycling between global and local approaches to the search space. For example, network-based graph-theoretic approaches may be useful for general routing of a force or force component, while local use of a free-space approach[79] may be necessary to adjust that movement in response to dynamically observed conditions.

6.2.2 Models of Spatial Data

There are two fundamental data models for the representation of geo-spatial information, reflecting two different views of such information.[80] The first treats the

---

[79] See for example (Reece 2000; Pisula, Hoff III et al. 2004)

[80] See for example (Duckham, Mason et al. 2000; Worboys and Duckham 2004; Richter and Klippel 2005; Duckham, Lingham et al. 2006; Klippel, Worboys et al. 2008)

information as continuous data defined on an underlying coordinate system, with point-by-point characterization of phenomena of interest. The second treats the geo-spatial domain as populated by discrete identifiable entities or objects, each of which is associated with geo-spatial references. The first view sees the data as an attribute of location, while the second sees location as a function of the data. There are computational advantages and disadvantages associated with each viewpoint. The first viewpoint leads to raster (also called field-based, tessellation-based, or image-based) database models, while the second corresponds to vector (also called object-based or feature-based) database models.

Raster-based models tend to be less computationally efficient, requiring large amounts of storage to account for point-by-point information. Vector-based models provide for much more compact storage of information, but do not lend themselves as easily to high-resolution representation of phenomena. Vector-based models more naturally support geometric, topological and set-oriented mathematical operations related to discrete properties such as set membership or discrete object relationships. They easily support hierarchies of object oriented model attributes such as inheritance. Raster-based models, on the other hand, are the natural province of the description of the variation of attributes over a region, and can more easily be adapted for some aspects of fuzzy set treatment of natural phenomena.[81]

Both raster data and object feature are incorporated in the ArcGIS geodatabase, which is the foundation for the Battlespace Terrain and Reasoning Awareness-Battle Command (BTRA-BC) framework developed for the U.S. Army Corps of Engineers (USACE) Engineer Research and Development Center (ERDC). BTRA-BC is a program aimed at increasing the effectiveness of the Battle Command and Military Decision Making Process (MDMP). According to (Visone 2005):

[81] For more detailed description of raster and vector approaches, see (Gold, 1992), (Gahegan and Leeb 200), (Worboys and Duckham 2004)(Gold 1992; Gahegan and Leeb 2000; Worboys and Duckham 2004)

"BTRA focus is on the development of six (6) information generation components and four (4) decision tools addressing terrain and weather effects. Each of these components utilizes terrain feature data, digital elevation models, current and forecasted weather and information regarding tactics, techniques and system performance. BTRA analytic components generate information products addressing:

- *Suite of Line-of Sight capabilities that incorporate weather attenuation*
- *Cover, Concealment and Obstacles*
- *Advanced mobility analysis*
- *Spatial Operational Compartment and Positions of Advantage (Key Terrain) for specific force types/tasks*
- *High fidelity weather/terrain effects of mobility and signature physics*
- *Digital ground and air Modified Combined Obstacle Overlays (MCOOs) supporting interactive route analysis."*

(Zeiler 2010) describes the ArcGIS geodatabase as having "*a data model that is implemented as a series of data tables holding feature classes, raster datasets, and attributes*." He goes on to describe the fundamental geometrical "shapes" of the geodatabase: points, lines and polygons. Of most interest to this dissertation methodology are routes, "*a special type of line feature that have measure values … that can represent any unit of measurement, you choose, including miles, kilometers… for distance, or hours … for time intervals.*" The geodatabase also supports rasters, "*a sampling of one or many attributes of attributes of continuous phenomena on a rectangular array of equally sized cells…. Continuous surfaces modeled with rasters include elevation, rainfall, temperature, contaminant level… and population density.*"

6.2.3 Route Finding[82]

The ArcGIS geo-database is typical of the description of terrain for route finding purposes in a combination of metrical/Euclidean and topological terms. Euclidean

---

[82] much of the material in this Chapter has been adapted from (Reece 2000)

schemes focus on straight-line distances between features of interest, while topological schemes describe spatial relationships (e.g., adjacency, connectivity, and containment) between such features. In both cases terrain is often overlaid with covering polygons, which can be regular tessellating polygonal tiles (triangles, squares or hexagons), or irregular polygon covering schemes such as Voronoi diagrams.

In strictly Euclidean schemes node-to-node “distance” metrics are based on regular grid coordinates, while more generic topological approaches can reflect a myriad of relational factors, such as trafficability, the availability of cover/concealment, and/or influence ambits based on the proximity of geo-political configurations, static and/or dynamic adversary threats, and the like.

6.2.4 Graph theoretic shortest path approaches

One of the most popular approaches to route finding uses arc-node graphs and shortest path algorithms, e.g., A* or Dijkstra’s algorithm. Nodes specify waypoints along a path, with arcs describing the connections between these nodes. In the simplest case, the nodes are only defined for points of interest, with the arcs representing possible connections. Arc costs from one node to another can reflect any and all of the “distance” metrics described above, and can be used in “shortest” path algorithms to determine the optimal path

Simple graphs consist of a finite number of vertices (also called nodes) connected by edges (also called arcs). These graphs may be directed, i.e., the edges have direction and can be “one-way”, or undirected. Undirected graphs are referred to as symmetric – their matrix representation has the property that $e_{i,j} = e_{j,i}$ whereas directed graphs are asymmetric. Mathematically, these graphs represent only the topological nature of underlying network; they are only concerned with which nodes are connected to which other nodes. In order to make them more useful as a search space representation, the edges or arcs are usually associated with one or more values, such as symbolic labels or a numeric attributes (capacity, distance, etc.). As shown in Figure 6-1, A* algorithms and the like use these values to define the network cost functions for finding optimal paths with respect to cost-related criteria.

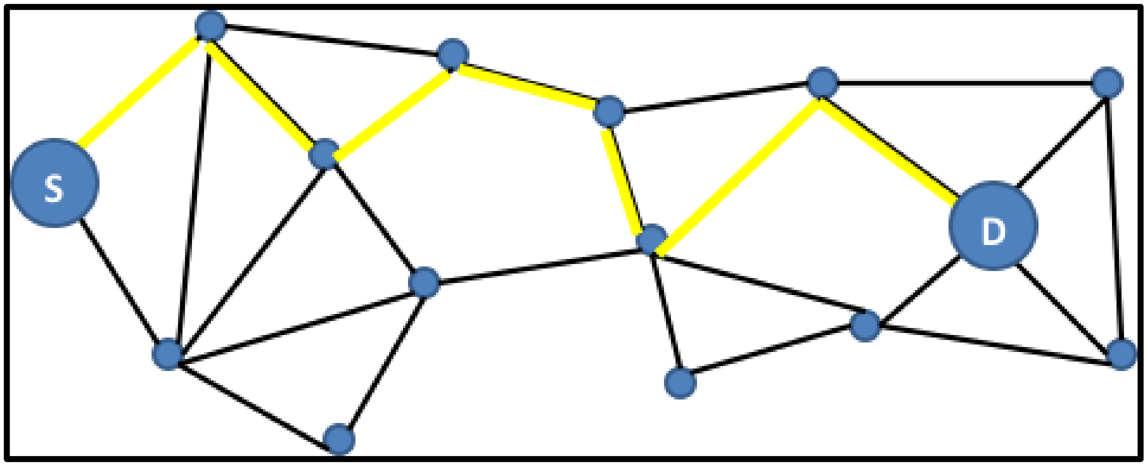


Figure 6-1 "Shortest" Path Determination with an Arc/Node Network

6.2.5 Terrain Tiling and Voronoi Diagrams.

Tiling approaches, also referred to as cell decomposition approaches, model the underlying terrain, including free space and obstacles, as a cell grids made up of convex polygons. The simplest of these grids use one the regular tessellating polygons (i.e., triangles, squares or hexagons), but there are a number of irregular tiling schemes as well. A refinement of such schemes is the use of Quadtrees, where the space is divided into squares, with some of these being recursively divided into smaller squares as may be necessary to provide adequate resolution for the phenomena under study, i.e., when the characteristics of interest are not sufficiently close to being homogeneous within a square.

The grid polygons can be used as the nodes of an arc node network, with arc connections based on polygon adjacency relationships. In the case of Euclidean tessellation approaches, nodes typically coincide with the polygons or tiles covering the space, with arcs for each shared boundary line. Other examples include skeleton approaches, which:

> *"reduce free space to a network of one-dimensional lines. Common representations are visibility graphs and Voronoi diagrams. A visibility graph is a collection of lines that connects the visible vertices of obstacles with each other. A Voronoi diagram is the set of points equidistant from two or more*

objects. Path planning with these skeletal free space representations involves finding a path from the start point to the nearest skeleton line, doing likewise with the goal point, and then using a graph search technique to find the lowest cost path from start to goal along the skeleton. The visibility graph solution yields a true shortest path, taut-rope solution (at least for the portion of the path on the graph), while the Voronoi diagram solution yields a path that stays as far as possible away from the obstacles.

A variation of this planning technique is used to find "short" paths in ModSAF[83] *and CCTT, another U.S. Army simulation system*[84]*. In these systems, no representation of free space is built. Instead, candidate free space routes are generated and evaluated. The first candidate route goes directly from the start to the goal. If it intersects an obstacle, it is rejected but two other candidate routes are created by computing "skirt" points around the obstacle in each direction. If the skirting paths intersect obstacles, additional routes will be generated by skirting those obstacles, and so forth. The route selected will be the candidate that doesn't intersect obstacles and is the shortest (or best by some other criteria)."(Reece 2000)*

The use of Voronoi diagrams to navigate around obstacles is illustrated in Figure 6-2. It divides the drawing into regions around each obstacle that are shaped so that the borders of the regions are equidistant from the two nearest obstacles. The paths indicated show boundaries of influence from each obstacle. If one wanted to address different levels of influence to each obstacle, one would simply increase obstacle size to include a buffer zone.

In addition:

The Ordinary Voronoi Diagram or OVD has been suggested as an alternative to overcome some of the limitations of conventional geographic data models (Gold, 1992a; Okabe et al., 1992). The main strength of this approach is that it is, to some extent, an integration of both vector and raster models. It explicitly encodes topology (spatial adjacency: in the form of shared vertices)—like the vector approach, but also provides a space-filling model—like the raster approach. In other words, all space is fully occupied, and fragmented into tiles (usually equivalent to zones of influence) around each discrete map object. Therefore, every location in the space can be assigned to at least one of the members in any underlying point dataset. Another strength of the Voronoi

83 (Smith, 1994)

84 Campbell, et.al 1995).

approach is that it permits many operations to be performed in a local fashion, rather than global, using the explicit spatial adjacency relationships (Gold, 1994). As a result, the tessellation can be maintained dynamically using local updates following from any changes. Since topology construction is computationally expensive, this is an appealing property and is especially useful when the objective is to allow users to make reasonable decisions quickly, rather than providing a globally optimum solution (Gold, 1993). Finally, Voronoi constructions form an integral part of many useful discrete interpolation methods (e.g. Watson, 1992). (Gahegan and Leeb 2000)

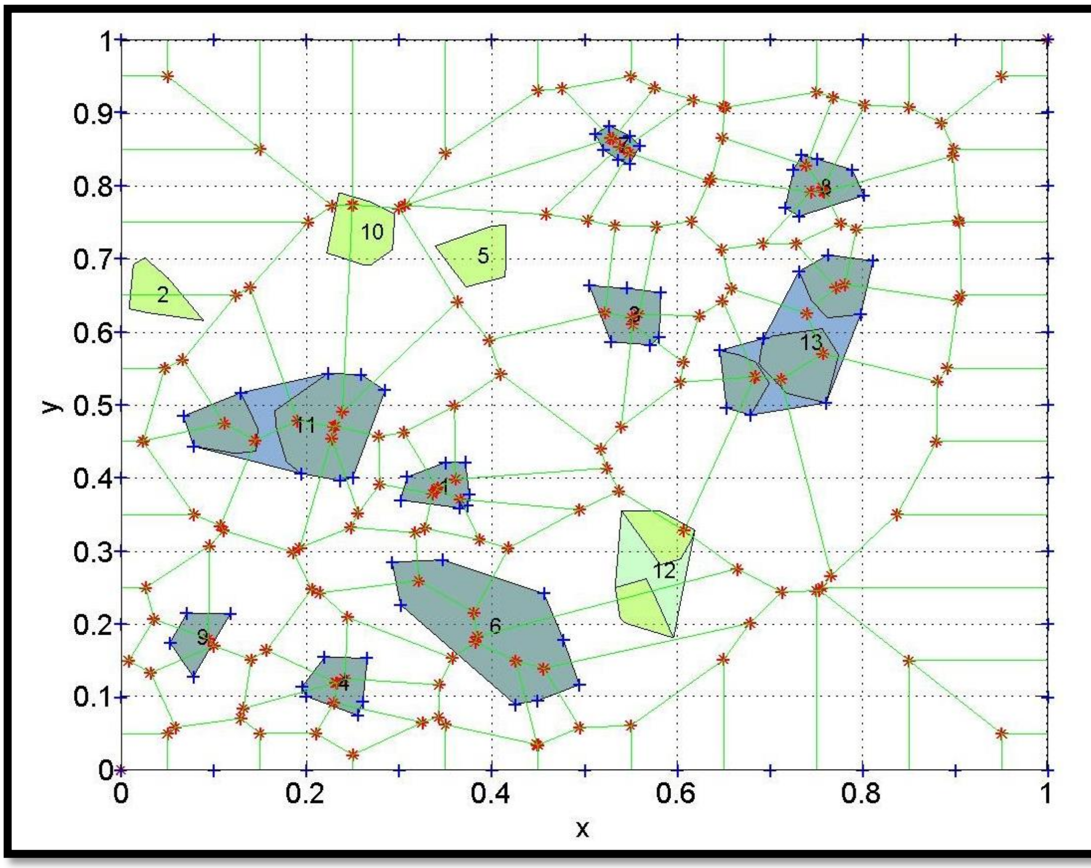


Figure 6-2 Voronoi Diagram Avoiding Obstacles from (Kim and Bhattacharya 2007)

6.2.6 Potential Field "Attractor/Repulsor" Schemes

Another approach, one that does not require explicit space-filling characterization of the terrain, is to attach potential fields to objects/features of interest. Features that must be avoided, such as obstacles, are considered to have repulsive potential fields around them, while desirable route features, such as cover and concealment are considered to have attractive potential fields around them. In both cases the potential field strength is inversely proportional to the distance from the feature; while there is also a uniform attractive force to the goal. In such cases, entity movement takes on a kind of “rubber band” aspect, where a rubber band connecting the entity to the goal position is stretched according the sum of the attractive and repulsive vectors encountered at any point in time. Such an approach is subject to being trapped in a local minimum or sink, but there are various methods for correcting such a problem.

A popular form of this approach frequently used in the computer gaming world is the use of influence maps (IMs). IMs are generally overlaid on a regular tiling scheme and describe the relative influence of each tile with respect to one or more operational factors. For example, in the case of two competing forces, a positive influence for a given with respect to one of the forces indicates some advantage to that force while a negative influence indicates an advantage to the other force. Influences are normalized to some standard, and color-coding map tiles by influence provides an easy visualization of the extent of each force’s control of the tactical situation. The bulk of the problem is figuring out how to calculate the influence of any given tile, and how these values may change dynamically throughout the game, or actual operational scenario.

7 RESEARCH METHODOLOGY

7.1 Introduction

This dissertation documents research designed to establish and illustrate the efficacy of an agent-based, “human-centric” approach to modeling simulation and analysis (MS&A). The simulation developed to support this research is simply a means to this end. The research itself focuses on the investigation of decision-making under imperfect SA/SU in experiments performed with that simulation. A series of simulation experiments explore a specific, easily understood, and quantifiable example of the impact of imperfect SA/SU on human behavior: intelligent agents being spatially “lost” while trying to navigate in a simulation world.

These simulation experiments vary aspects of an agent’s perceived worldview to study how a mistaken understanding of ground truth affects achievement of the agent’s goals. They provide insight into multiple aspects of decision-making as affected by problem complexity, information quality, risk tolerance, and decision strategies. The experimental framework is characterized by multiple, often interacting, elements. At the highest level of organization these elements can be structured into three main groups:

- Type of simulation experiment performed;
- Characteristics of the arc/node networks explored by the agent; and
- Nature of the agent decisions explored.

7.2 Types of experiments

Two types of simulation experiment are used to observe and analyze the decisions made by the agents and the effects of those decisions on goal achievement. Interactive simulations provide detailed insight into agent behavior for specific model input values;

parametric experiments provide a more thorough basis for statistical assessment of that behavior over a wide range of possible input values. Both employ Monte Carlo simulations, providing random sampling by selection of numerical draws from known or hypothesized statistical distributions.

When operating in interactive simulation mode, an entity's movement in ground truth can be seen. That entity's generally imperfect view of ground truth, i.e., its mental map, can also be seen in parallel, as the simulation executes, as can any changes or updates to that mental map. The simulation can be paused at any time and the internal values of model variables examined. This form of simulation execution provides the most complete and detailed view of the entity's behavior. It is useful for verifying and debugging model code, but more importantly it supports intuitive exploration of the causes and effects of specific entity decisions.

When the simulation is executed in parametric mode, specific combinations of parameter values define the sets of experimental conditions that comprise different experimental cases. Each of these cases is executed for a number of Monte Carlo trials and a wide variety of output data are recorded. Commercial-off-the –shelf, COTS, software such as JMP and Excel are used to analyze these data by:

- Collating descriptive statistics, which summarize the experimental outcomes;
- Appling inferential statistics, which identify significant factors within the results; and
- Supporting exploratory data analysis, which seeks to provide insight into the behavior of the simulated entity.

All three of these types of statistical activity are important to scientific investigation based on agent-based modeling as is discussed in detail above in Chapter 3.1. All three of them will be prominently featured in the experimental results provided below. The kinds of investigation performed herein, however, reflect a shift in focus from scientific experimentation as hypothesis and test, towards scientific experimentation as exploration and discovery[85]. Such experimentation is precisely the aim of data farming and data mining, which use simulation experiments to generate mass quantities of data, trolling these data for patterns, trends, insight and meaning.(Fayyad, Piatetsky-Shapiro et al. 1996; Horne 2001; Shneiderman 2002; Witten and Frank 2005; Han and Kamber 2006; Vaughan 2006; Horne 2008; Middleton 2009) As in much of science, exploration of the data looks to induction and correlation for help in inferring cause and effect, but often results just in increased understanding and direction for further research.

7.3 The Simulation Software Environment

The experimental simulation is implemented in the simulation package AnyLogic®, which supports the principal methods of simulation shown in Figure 7-1: Discrete Event (also known as process-centric), Agent-Based, and System Dynamics modeling.

Agent-based modeling has been described in detail above. In this approach to simulation, each agent's behaviors are defined by functions internal to that agent, which permits the use of different models for knowledge representation, inference and decision-making.

Discrete event simulation is similar to the agent-based approach in that it seeks to approximate the continuous behavior of systems with a series of distinct, separate events that drive the evolution of the system over time. System dynamics models, on the other hand, rely primarily on differential equations to describe a system in of terms feedback loops and time delays. AnyLogic® supports a graphical development approach for each of these methods, implemented through the use of state charts in the case of agent-based models, stock flow diagrams in the case of systems dynamics models, and process flowcharts in the case of discrete event simulation.

AnyLogic® supports clock driven simulations that maintain lists of scheduled events,

[85] See also (Hooker 1994)

and can use Monte Carlo draws to determine the outcome of stochastic processes, while also using closed form equations to calculate deterministic results for specific types of agent behavior. AnyLogic® is a Java platform and model development with AnyLogic® consists of building the appropriate graphical structures and customizing their function with Java code.

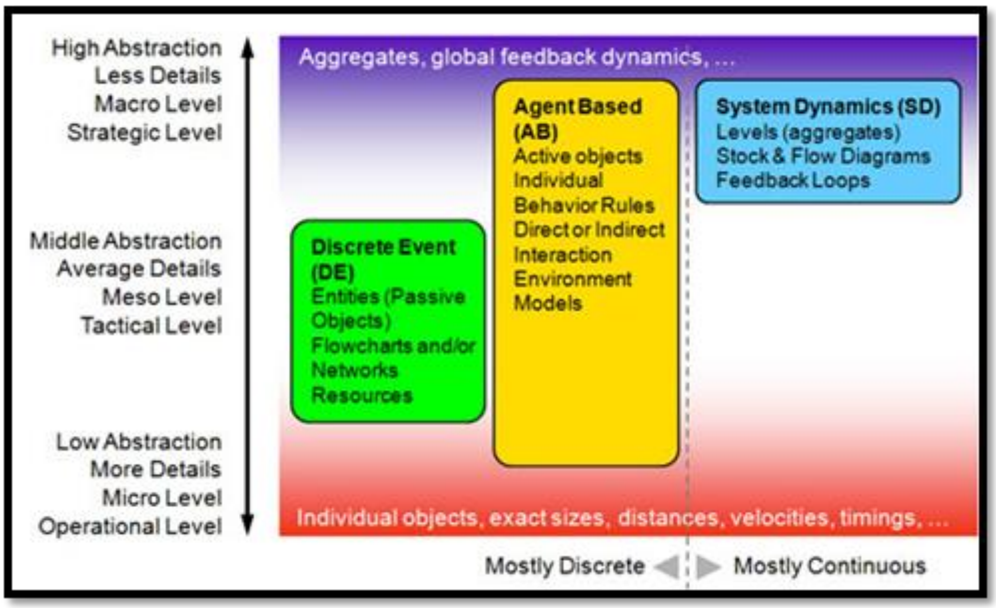


Figure 7-1 Principal Methods of Simulation[86]

Model development in AnyLogic® consists of selecting basic program element templates from its graphical user interface (GUI) menus and writing Java language code to complete the templates with desired model functionality. There is a menu for general model constructs such as parameters, functions, and a variety of data structures with which to define model variables. There is a menu for the specific state chart constructs needed in agent-based modeling and another for systems dynamics features. The AnyLogic® GUI also provides a presentation menu through which the user can

[86] Illustration by Sergey Suslov, 20 Dec. 2009, released to the public domain through Wikimedia Commons

define his/her own GUI for the model he/she is developing with AnyLogic®.

1. **FUTURE WORK: EXPANDED USE OF ANYLOGIC® CAPABILITIES[87]:**

The agent-based features of AnyLogic® are sufficient for the current dissertation research. Eventually, however, integrating the other simulation paradigms with the agent-based structures should provide an even more powerful mechanism for exploring the problem space. For example, it would be desirable to use systems dynamics methods to represent soldier physiological processes, such as fatigue and heat stress, that affect basic decision processes. Similarly, both systems dynamics and discrete event modeling are useful to represent changing aspects of the battlefield environment and the threats faced by the soldier.

[87] This dissertation in many ways breaks new ground with respect to the simulation of human behavior and there a large number of potential avenues for future work. Here and throughout the rest of this document, those opportunities will be designated with the phrase **Future Work** and a **Title**, highlighted in red, with the explanatory text italicized.

8 THE SIMULATION MODEL

MOBILE represents an entity trying to navigate in a simulated world. The entity has a unique "mental map" – its idiosyncratic view of its geo-spatial environment and makes decisions based on this idiosyncratic view, with behavior outcomes based on ground truth. The simulated entity attempts to find its way on an arc/node network, moving from a given start node in the network to a designated end node. It can (1) generate a route plan, a sequence of nodes and arcs from the start point to the end goal; (2) it can employ local search techniques to seek the goal; or (3) it can employ some combination of global/local strategies. Should an agent attempting to follow a route become "lost", i.e., diverge from its chosen route, it can switch to search tactics to either attempt to rejoin its route or to otherwise achieve the end goal.

Being "lost" carries with it the idea of some kind of failure in the entity's mental map, either in its structure, in its content, in its registration/correlation with ground truth, or in some combination of all of these faults. The types and degrees of such failure are examples of poor SA/SU, and the extent to which they affect the entity's ability to achieve its end goal provides a measure of the value of SA/SU.

Actual humans can find their way from one point to another with very rudimentary and/or inaccurate maps. They can frequently function satisfactorily with ambiguous and unclear directions. An effective simulation of getting or being "lost" should take into account this human resilience with respect to the effects of imperfect SA/SU, and seek to delineate between when SA/SU imperfections do and don't matter.

Model requirements for a useful simulation of being "**lost**" include:

- The capability for an entity's view of where it is and where it is going to be different from ground truth.
- An error taxonomy that reflects both types of being lost and degrees of

"lostness" within those types;

- The mechanisms by which an individual achieves different states of being lost; and
- The mechanisms by which an individual recognizes and attempts to correct being lost.

Under the agent-based structures of the model, a single entity is represented by two inter-connected agents. The first of these is the **ground truth (GT) agent**, who moves physically in the world of ground truth reality, and the second is **the voice-in-head (VIH) agent**, who represents the entity's decision-making capabilities and who maintains the entity's perceived view of the world.

The entity's world is an arc/node network, but the key to the model is the fact that the GT agent operates on the arc/node network that represents the "real" world of the entity, while the VIH agent maintains an idiosyncratic view of that network, a "**mental map**" that represents its own particular, generally distorted, view of ground truth geography.

Route planning and route following decisions are made with respect to the VIH mental map, while actual movement takes place on the ground truth network. The GT agent maneuvers around the GT network and reports to the VIH agent the characteristics of the GT network as the GT agent experiences them. The VIH agent monitors the GT agent's progress, and compares the state of the VIH mental map to the GT network characteristics reported by the VIH agent. The VIH agent can use these data to update the mental map, but the agent and its map are always subject to possible misperception and/or misinterpretation of the GT data. The underlying philosophy is very much in keeping with Endsley's phase of SA as discussed above in Chapter 4**.**

2. FUTURE WORK: SIMULATION OF MISPERCEPTION

The model at present can represent the entity's inability to reconcile ground truth information with its mental map, but there is no explicit representation of the entity's perceptual processes or capabilities and hence no explicit representation of how those processes might be influenced by biases induced *by previous states of the mental map. In the real world an individual's capability to perceive the environment is very much a function of both the psycho-physiological state of that individual and the amount of "noise" present in the environment. There is a good body of research in human factors upon which to base future work incorporating perceptual capabilities and agent knowledge states into the model. For example ,* The attention/situation awareness model in *Figure 8-1 The attention/situation awareness model* from*(McCarley, Wickens et al. 2002) shows "a sequence of events (upper left) are attended (center) to a degree that is degraded by workload. Attended events provide evidence for the belief module (box at lower right) a belief that decays over time. The SA belief then contributes to a choice, at the bottom."*

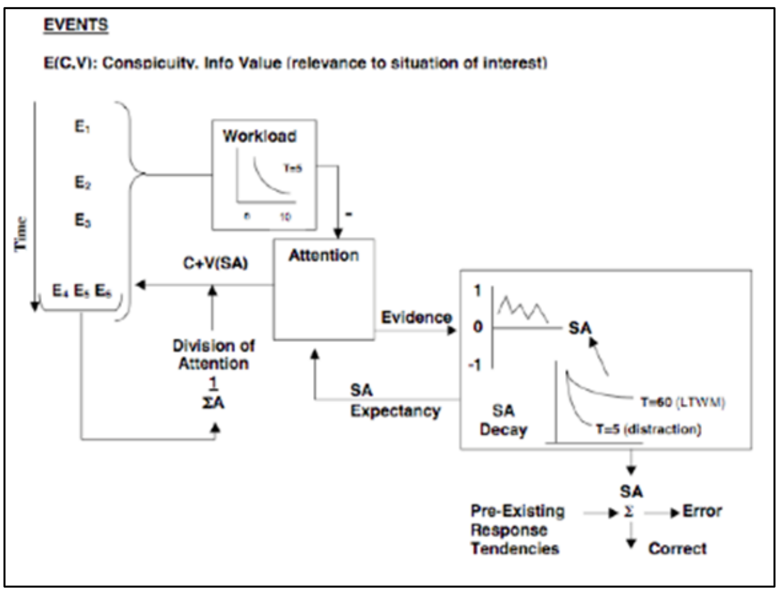


Figure 8-1 The attention/situation awareness model

8.1 Decision-making

The VIH agent makes decisions at multiple levels. The first level can be characterized as deciding whether to employ a global or a local strategy. Following a global strategy consists of two parts:

- Route Planning, finding a sequence of nodes and connecting arcs that will take the agent from its start point to its end goal, and
- Route Following – recognizing the elements of the planned sequence and adhering to them to achieve the goal.

Local strategies are characterized by Way-Finding, parsing a global route into a sequence of one or more choices that will ultimately lead the agent to its goal. Typically the agent will attempt global routing as a first option and adopt way-finding behaviors when faced with a failure in either of the planning or following elements of the global route. There are several different way-finding strategies available to the agent, dependent on how the agent became lost and what information is available and credible. Among them are:

- Returning to a previous known node in the global route;
- Seeking an unvisited node in the global route;
- Seeking a landmark;
- Moving in what is believed to be the general direction of the goal; and
- Random movement.

Earlier chapters identified the need to define software constructs that describe how agents go about "knowing what's going on" and "figuring out what to do". Implementation of cognitive architectures in terms of specific data structures and inference processes provides the ability to study the implications different theories of cognition and perception may have with respect to an individual's worldview, the inferences that individual makes with respect to perceived ground truth, and the decisions that result.

The simulation model's data structures and inference capabilities are designed to accommodate a wide variety of cognitive architectures, as long as those architectures can support three basic decision elements of movement strategies:

- Where am I?

  The simulation process most often used to answer this question occurs when the agent has arrived at a GT node. The agent compares candidate nodes from its

VIH map to find the one most closely matching the GT Node.

- Where do I want to go next?

  Ideally the agent will have a “next node” goal in a planned route sequence. Failing that, the agent may be able to aim for an intermediate goal - a landmark, unvisited node in the global route, or other predefined waypoint. At present, such intermediate goals are limited to nodes on the agent’s VIH map; MOBIL represents only point features and arcs between them. Future extensions of the simulation could provide linear features, or area features, e.g., roads, rivers or other bodies of water, as well as other geographic or geo-political regions and their associated boundaries.

  An important component of “where do I want to go next” decisions is one or more measures of progress towards the ultimate goal. Currently the simulation focuses primarily on the degree of confidence the agent has that it is following the global route. Various other measures are collected as experimental outputs and can be explored for use by the agent in selecting or changing its movement strategy.

- How do I get there?

  When this decision is made in the context of an agent at a node in ground truth, it becomes the question of selecting which ground truth arc connection most closely matches the VIH map arc leading to the node that is the result of its “where do I want to go next” decision. In this instance, the agent’s candidates will be the set of ground truth arcs available at its current ground truth location and it will select the one that most closely matches its VIH choice.

8.2 <u>Model Environment</u>

As described above, the environment in which the simulated entity operates consists of two different views of an arc/node network. First is the ground truth view, which is the true state of all the objects in the environment. Ground truth represents the environment that affects and is affected by an entity’s physical behaviors. Second is the entity's worldview, the VIH, which defines the state of the environment and objects in it as the entity perceives them.   This worldview is the basis for the entity's cognitive behaviors, wherein it decides what to do next.

The first step in setting up a simulation experiment is the selection of these two views, and the first step to that is selecting the ground truth arc/node network that will be the entity's operational environment.

The simulated arc/node networks can be characterized by three different sets of features:

- spatial – such as positional data expressed on an xy-coordinate system;
- topological – the spatially invariant features of the network describing node connectivity; and
- semantic – non-spatial features that aid in distinguishing nodes and arcs from each other, represented in the simulation as line types and widths and/or colors.

All of these features can be distorted, either individually or in combination to provide the entity with its idiosyncratic VIH world-view that differs from ground truth. Furthermore, these distortions can be randomly allocated to individual nodes and/or arcs, or applied in some sort of systematic fashion, and/or assigned some degree of uncertainty or "fuzziness".   The entity's VIH arc/node network can also differ from the GT arc/node network simply in the amount of information provided, from a sparse sketch of the network map to a very detailed, albeit incorrect, network map.

For each experiment, the entity's VIH map, its mental image of ground truth, is created by distorting the ground truth map through a series of Monte Carlo processes. Selected nodes can have their location changed by random perturbations in their X and Y coordinates. Actual nodes from the GT network may be missing in the VIH network.

Initial experiments assume a completely connected arc/node network for both Ground truth and mental maps. This connectivity means the entity can use Dijkstra's algorithm to generate a route plan according to its mental map.   Should an agent attempting to follow a route become "lost", i.e., diverge from its chosen route, it can switch to search tactics to either attempt to rejoin its route or to achieve the end goal. Later experiments will allow missing nodes to disconnect the mental map network, meaning a planned route cannot always be found, requiring different goal seeking strategies from the simulation start.

Each experimental run is deemed a success if the entity finds its way to the goal and a failure otherwise. Critical aspects of the state of the entity when failure occurs are recorded to allow a more detailed study of failures.

8.3 Arc/Node Network Map Structure

The knowledge structure used to represent the simulation arc/node networks is based on the format of OpenStreetMaps[88]. There are three major data types in this structure:

- Nodes – map locations, defined by (x,y) coordinates;
- Ways – a series of sequentially linked nodes – arc links between nodes are defined as way segments; and
- Relations – groups of nodes and/or ways that are related some way, sharing some set of common attributes.

Each instance of these types has a unique reference id. Each of them may or may not have additional information provided through a series of tags. OpenStreetMap has a currently accepted set of standard tags and associated values, but one can define virtually any set of tags and tag values that might be useful. For example, the node characteristics used to compare VIH and GT Nodes as described in Chapter 8.7 below are defined as tags for both GT and VIH maps.

3. FUTURE WORK: DISTORTION OF OPENSTREETMAP FEATURES:

Using the OpenStreetMap tag structure supports a much wider degree of variation between ground truth and the voice in head map, by altering the

[88] http://wiki.openstreetmap.org/w/index.php?title=Main_Page&oldid=941300

tag values of selected nodes, ways or relations on the VIH map, either systematically or at random. In addition, tags can be set to represent uncertainty in any node or arc attribute by providing parameter values for a stochastic draw. OpenStreetMap is also a source for real-world maps that can be used in the simulation to represent ground truth and altered as desired to represent the VIH map.

4. FUTURE WORK: NETWORK COMPOSITION AND COMPLEXITY:

Network complexity is one of many potential areas for future study, and there is an extensive literature on network complexity (see for example (Kaimann 1974; Hall and Preiser 1984; Bonchev and Buck 2005) on which to base such work. In addition, network representations of real world environments are easily adaptable for study, one such network being the roads, paths, and tunnel system at Wright State University.

There is a fairly extensive body of literature of the topic of network complexity in a topological sense, see for example, (Mackaness and Beard 1993; Molenaar 1998; Hayes 2000; Blondel, Gajardo et al. 2004; Bonchev and Buck 2005; Zhu and Wilson 2005; Wu 2006). Topological factors of potential importance in assessing movement task complexity include:

- *The number of possible paths from the start goal to the end goal, which may be quantified by looking at the minimum, maximum, average and/or percentile values across all of the possible start/end goal pairings in the network. These values can be theoretically calculated by taking successive powers of the network incidence matrix. For larger networks it may be more reasonable to approximate them with Monte Carlo methods.*
- *The number of paths with the potential for a "dead end' or a "sink" based on a wrong turn or other route following errors.*
- *The number and location of "hub" nodes.*
- *The degree to which the complexity of a GT graph may differ from that of the VIH map as a function of the topological changes made to the VIH map.*

For simplicity, initial experiments explore variants of three basic networks: a square grid, a repeating polygonal shape with various node interconnections, and a grid with a number different connectivity features. These three maps shown in Figure 8-2, Figure 8-3 and Figure 8-4

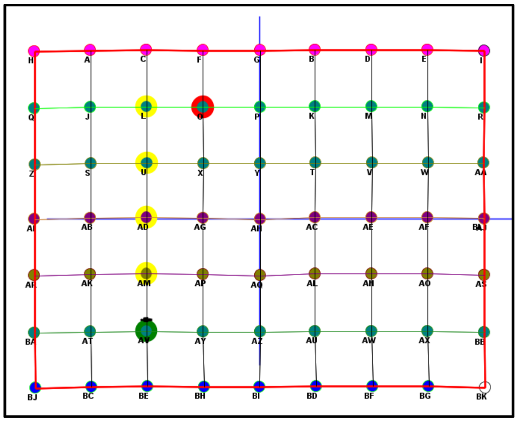


Figure 8-2 Square Grid Map

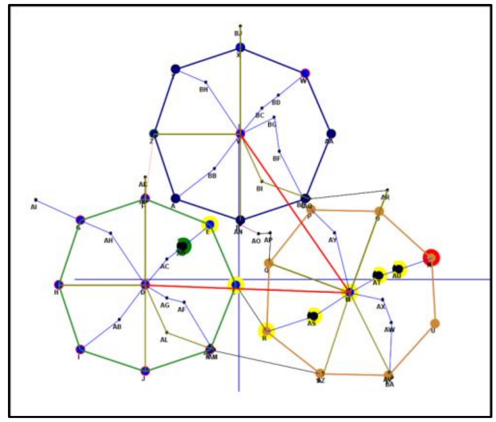


Figure 8-3 Polygon Map

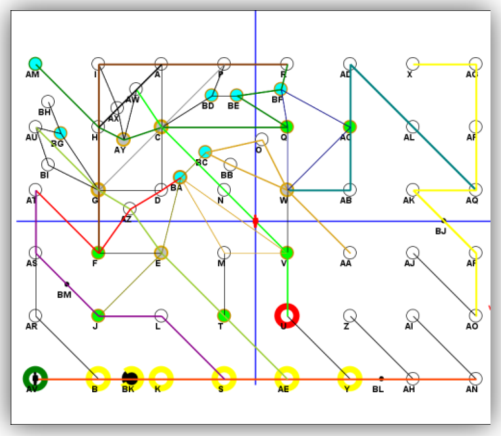


Figure 8-4 Multiple Feature Map

8.4 Map Distortion Process

The simulation supports two methods of defining the VIH mental map, the first being internal to the simulation - the application of random distortions to the arcs and nodes of the GT map, and the second being external to the simulation - reading in user-defined files of distorted networks.

In either case, the distortions themselves can be the result of:

1) **spatial deformations** – warping the network by applying location errors to

node x and y coordinates, with the associated bending and stretching of the arcs that connect those nodes, and/or

2) **topological modifications** to the basic network structure by

- changing the connections between nodes,
- deleting from the VIH network nodes and/or arcs that are present in the GT network, and/or
- adding VIH nodes and/or arcs that aren't actually present in the GT network.

Mathematically, the mental map represents a functional transformation from the GT network to the VIH network. Viewed in this way the first type of distortion, spatial deformation, results in a one-to-one transformation from the GT network to the VIH network that is topologically invariant. In this case, the GT and VIH networks share the same incidence matrix with respect to node connectivity, and there is a direct correspondence between each of the arcs and nodes of the GT network and those of the VIH network. In other words, in this initial model of distortion, only the distance between nodes and the angles between them, i.e. the bearing of the arcs connecting them, are affected. Network connectivity relationships, as defined by the network topology, are unchanged by this spatial distortion of the VIH network. As a result, the arc/node elements that comprise routes between any two nodes are the same for both the GT and the VIH agents. The distance relationships of those routes may change, however, thus affecting the definition of the "shortest" path between those nodes. In such a case, although the entity may have a path that will allow it achieve its goal, it may be said to be "lost" in the sense that this route does not correspond precisely to ground truth and as a consequent may delay the time otherwise required to achieve that goal.

The entity's GT and VIH maps each have their own relative coordinate system. If the entity's SA/SU is perfect, the location of each node in both the GT and VIH maps is the same with respect to its appropriate origin. When the VIH Map is distorted spatially, selected node X and/or Y coordinates are changed in the VIH Map. These changes modify the spatial relationship between nodes, i.e., changing the distance and angles of connecting arcs. Input parameters control the probability of node error, Pr(NE), and the extent of the X and Y coordinate errors, X and Y Er Lim. A Monte Carlo random draw determines first if a node is to be distorted, and then separate draws determine the extent of the errors in X and Y, which are calculated according to the formulae:

newX = oldX + XandYEr Lim + (uniform_pos()*(-2.0*(XandYErLim);
newY = oldY + XandYEr Lim + (uniform_pos()*(-2.0*(XandYErLim).

where **uniform_pos()** generates a random number uniformly between zero and one.

The topological nature of the network can also be distorted by omitting ground truth nodes from the mental map. As stated above, initial experiments assume a completely connected arc/node network for both ground truth and mental maps, so for these initial experiments any node paths that would be disconnected by removed nodes are reconnected by joining the removed node's predecessor and successor nodes in those paths. The input parameter Pr(MN) defines the probability of a missing node, i.e, determining whether a given GT node is omitted from the VIH map.

Figure 8-5, Figure 8-6 and Figure 8-7 show examples of distorted VIH maps for each of the three maps described above.

The node labels of the VIH map are the same alphabetic characters as those of the corresponding GT network nodes, with a small "v" appended to distinguish them from the labels of the "true" nodes.

Removed nodes are shown colored light pink and corresponding arcs to and from those nodes are shown as dashed light pink lines. Arcs used to connect predecessor/successor nodes to the removed nodes are shown as dashed green lines. The nodes and arcs removed from the VIH map still exist in ground truth and hence can be encountered by the agent as it traverses the network. The new VIH connector arcs do not correspond to any actual arcs in ground truth, but the agent may confuse these arcs with those ground truth arcs that have been removed from the VIH map. In such a case, the agent would anticipate that those ground truth arcs connect in the predecessor/successor nodes specified by the VIH map, and make movement choices accordingly.

5. Future Work: Structured or Systematic Distortion in VIH Maps

In many cases, such as when using real-world maps, it is probably appropriate to make more structured distortions than the random mechanisms employed in current experiments. Systematic distortions, however, should be accompanied by more complex logic for the agent to recognize the nature of the distortion and potentially update/correct its VIH map accordingly.

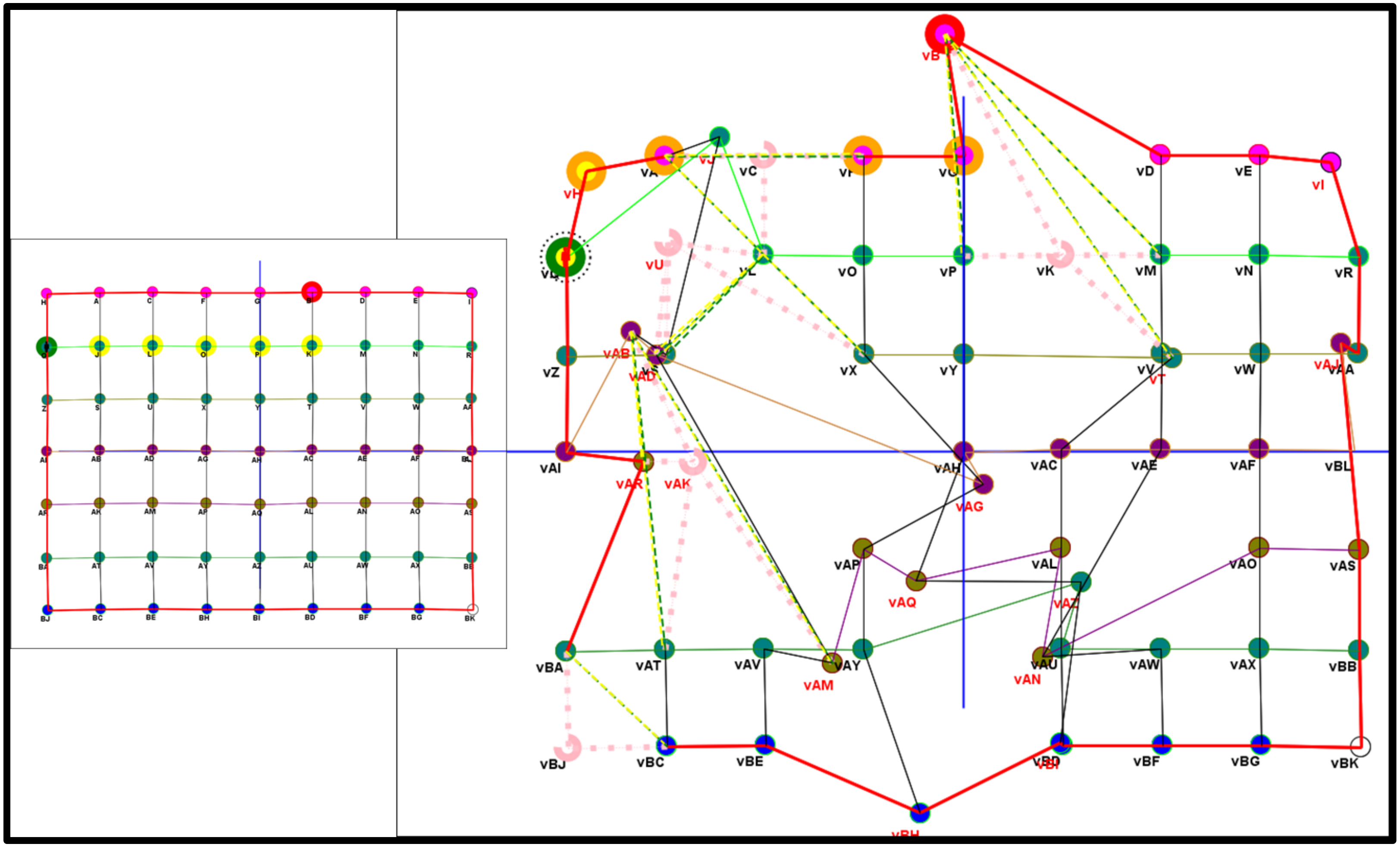


Figure 8-5 Distorted Square Map

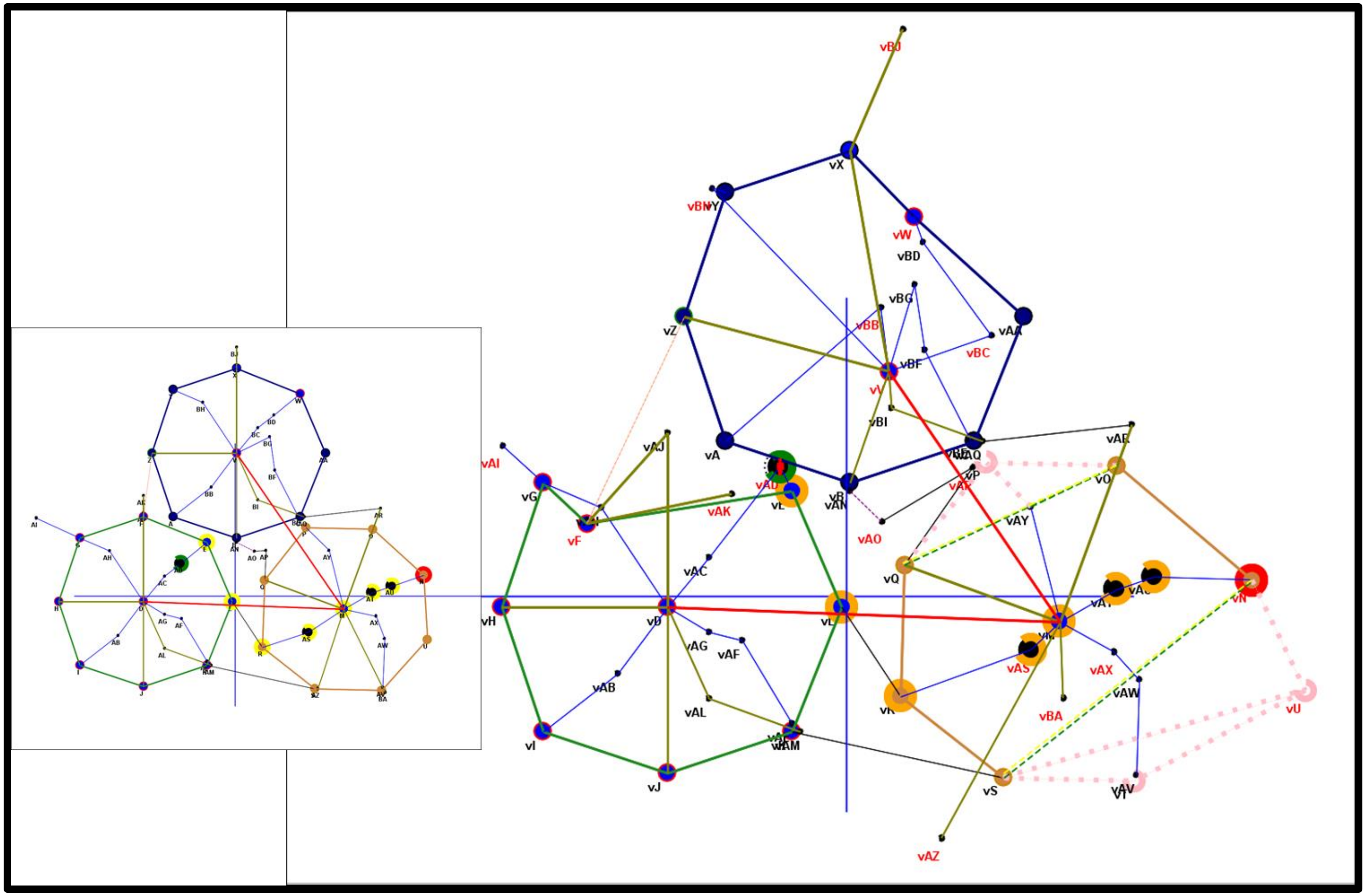


Figure 8-6 Distorted Polygon Map

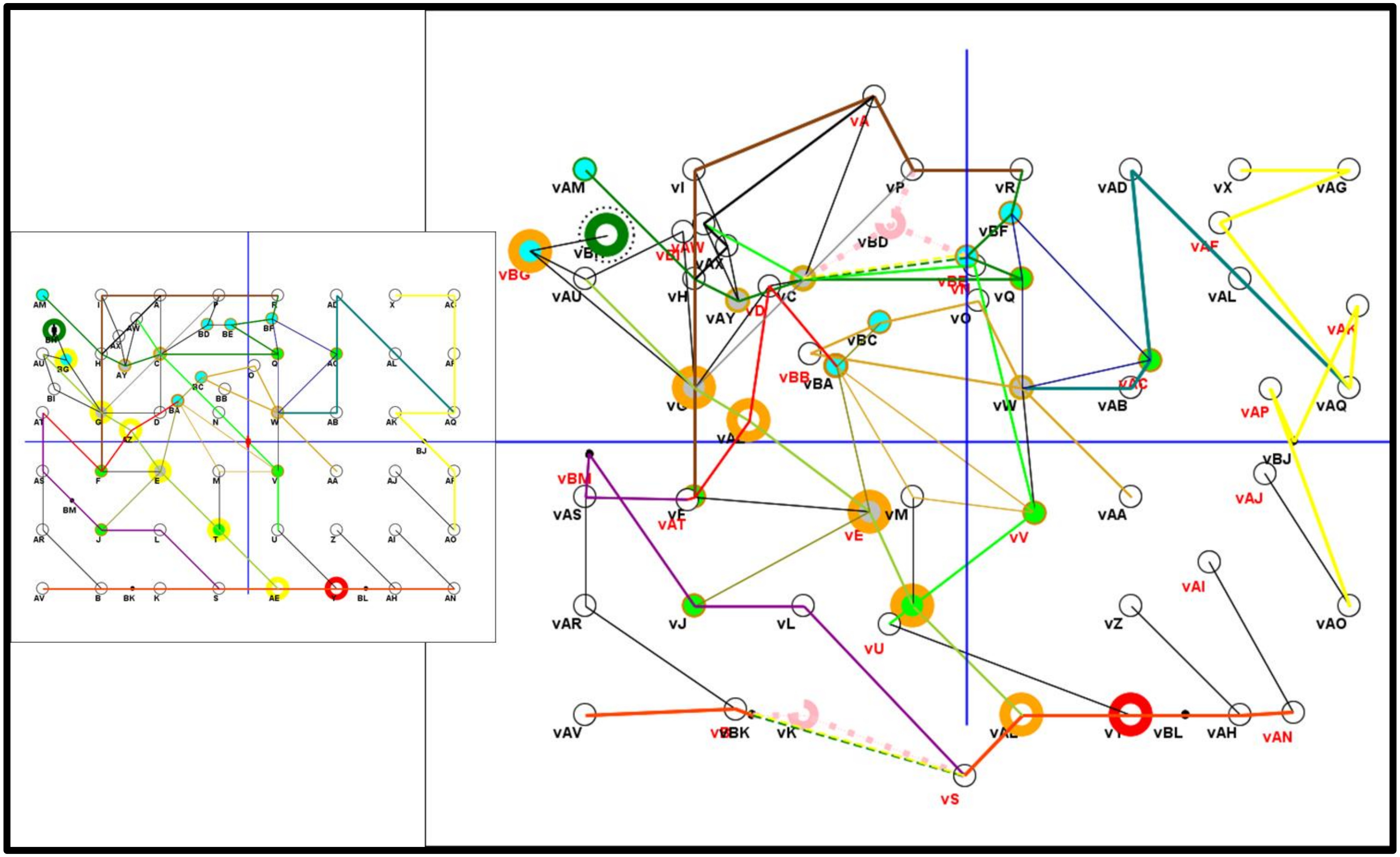


Figure 8-7 Distorted Multi-Feature Map

8.5 Model Structure

The model as currently implemented has three principal elements. The first of these is **Main**, which is the default active object class shared by all AnyLogic® programs - it handles initial model setup details and some of the model book-keeping functions. The GT arc/node network is initialized in Main. The other two elements are (1) the active object classes for the aspects related to the physical movement of the entity, the **GT Agent** and (2) the decision-making aspect of the entity, the **VIH Agent.**

6. Future Work: Entity Interaction

Under the current model, the GT agent is limited to following directions and observing the ground truth environment; it has no other entities with which to interact, nor can it affect the environment. Adding other entities is easily done in AnyLogic®, so, for example the model could be altered to add goals for seeking or avoiding other entities to more robustly represent military operations. Entity interactions are also possible through the definition of further agent state logics.

8.5.1 GT Agent

At any point in the simulation, the GT agent is either:

1. At a node;
2. Moving along an arc at a constant speed; or
3. Stopped at a point along an arc.

Figure 8-8 shows the model state chart diagram for the GT agent.

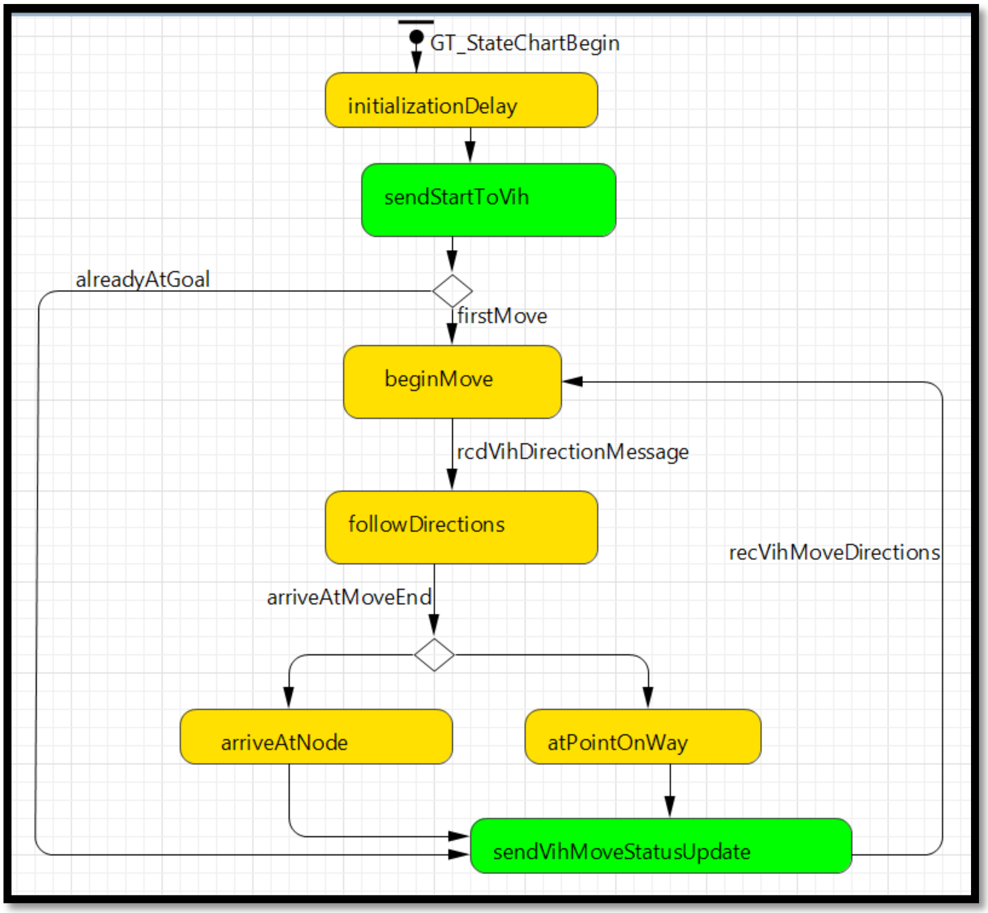


Figure 8-8 GT Agent State Chart

In AnyLogic® state charts are developed using the GUI state chart menu, which provides templates for the state chart elements. Shown here are simple state blocks indicated by the rounded rectangles. An AnyLogic® state block is essentially a container for the model-specific Java code that defines the behavior of the agent while it is in the given state. Color-coding distinguishes between states associated with basic agent behaviors (in yellow) and the states that represent the need to communicate with the VIH agent (in green). State transitions (lines with arrows) and decision points (diamonds) are additional basic AnyLogic® state diagram elements. They define when the agent should enter or leave a state and to which state it will next transition. These also are augmented with Java code to implement specific features of the model. As the figure shows, the transitions that fire each increment of movement cycle are the ones triggered by messages received from the VIH agent.

The states shown here are relatively simple constructs. AnyLogic® also permits the definition of compound states and history states, in which the agent “remembers” where it was in the compound state, and can resume its prior behavior, as for instance it might want to do after leaving the state for an interrupt.

As shown here, the GT agent begins the simulation by sending a message to the VIH agent containing the GT node that is the entity’s start position. The GT agent then follows the directions provided by the VIH agent. At each movement step the VIH agent selects the GT arc along which the GT agent will move, and the GT agent follows that arc until it either arrives at a GT node or exceeds the distance that the VIH agent is willing to travel looking for that node. The GT agent then relays where it is and what it observes to the VIH agent, initiating the next step in the movement cycle.

8.5.2 VIH Agent

The VIH agent state chart is considerably more complex than that of the GT agent, and is presented here in two parts. The initial behavior of the VIH agent is described in Figure 8-9, which shows the first components of the AnyLogic® VIH agent state chart.

The VIH agent receives the GT start/goal node pair from the GT agent and establishes its VIH map. Current experiments create the VIH map by reading in the GT map from an input file and then distorting that map as specified by input parameters. This procedure supports parametric experiments with respect to the various distortion options described in Chapter 8.4. Alternatively, the model can import an already distorted version of the GT map file to explore the use of different movement selection logics in a common environment.

The VIH agent receives the GT start/goal node pair from the GT agent and establishes its VIH map. Current experiments create the VIH map by reading in the GT map from an input file and then distorting that map as specified by input parameters. This procedure supports parametric experiments with respect to the various distortion

options described in Chapter 8.4. Alternatively, the model can import an already distorted version of the GT map file to explore the use of different movement selection logics in a common environment.

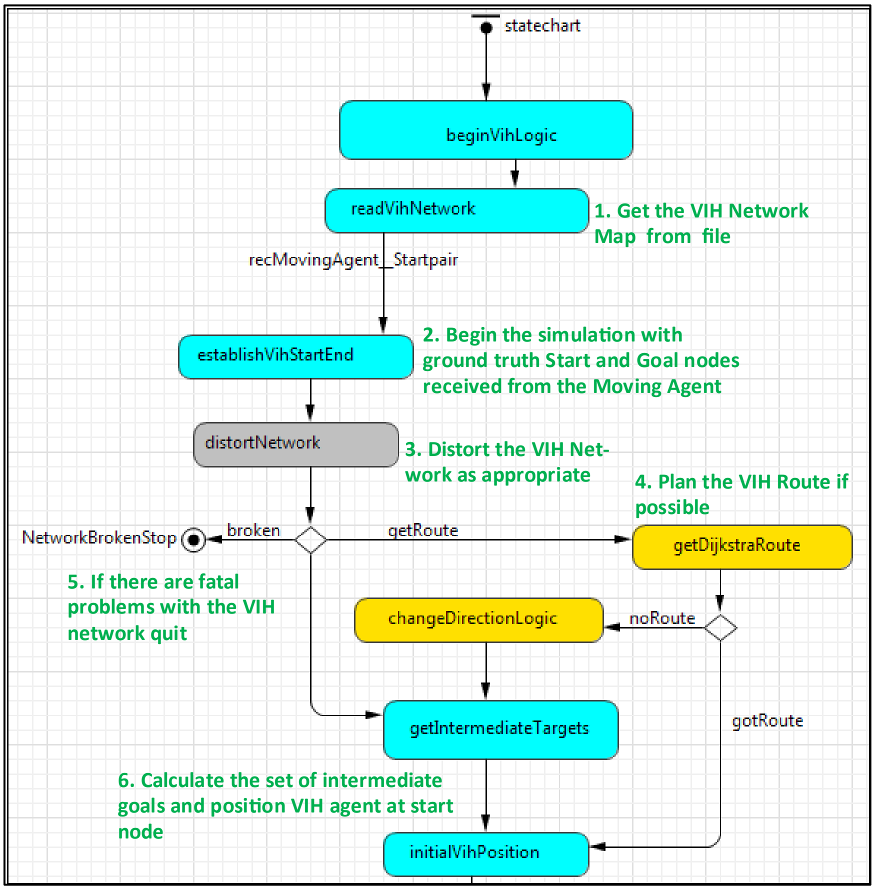


Figure 8-9 Initialization of the VIH Agent

The VIH agent then seeks to plan its route from the start node to the goal node. The default logic for route planning is the use of Dijkstra's algorithm to find the shortest path through the network, as calculated using the arc-node connections and arc lengths (inter-node distances) of the VIH mental map. The simulation tracks both the "true" shortest path from start to end goal, as would be calculated from the GT Network, and the VIH path, the shortest path based on the VIH mental map. The divergence between these two is one metric for the level of effective distortion in the mental map. It can then be related in the experiments to the effect of distortion on entity behavior through selection of a less than optimal path.

7. Future Work: Route Planning Options

At present the model uses only Euclidean distance as the Dijkstra algorithm cost function but there is no problem incorporating other factors into this function. In particular, the use of OpenStreetMap tags can be used to define such arc link characteristics as: speed limits; trafficability constraints; and/ or uncertainties/risk factors - stochastic hazards whose presence is based on a Monte Carlo draw.

The simulation allows for abnormal termination in case of fatal errors in the arc/node network, as might be caused by corrupted input files. In the case of none-fatal problems with the VIH map, e.g., a disconnected network where the VIH agent cannot determine a path to the goal, the agent uses a set of intermediate goals to direct its movement. These intermediate goals can also be useful if the VIH finds itself off of its planned route and needs to institute a local search or way finding strategy in lieu of path following.

Figure 8-10 shows the final components of the AnyLogic® VIH agent state chart, which describe the agent's decision processes in response to the information received from the GT agent's movement.

Entity movement requires of a sequence of decisions: each time the GT agent reaches a node, the VIH agent must interpret its position on the VIH map and make decisions about where to go next. i.e., which arc link to take. If the GT agent travels along an arc without reaching a node where one is expected, it must decide whether to continue or quit. These decisions are a function of the entity's current decision-making mode, the amount of distortion in the entity's mental map, and the threshold parameters set by the user to persist in the face of ground truth deviation from the conditions anticipated according to that mental map. Later experiments will introduce a number of additional mental map features and entity decision options based on those features.

As shown in decision block 7, if the VIH agent finds itself to be at a GT node, it can identify that node with a corresponding VIH map node, either correctly or incorrectly, or fail to recognize the GT node as matching any VIH node.

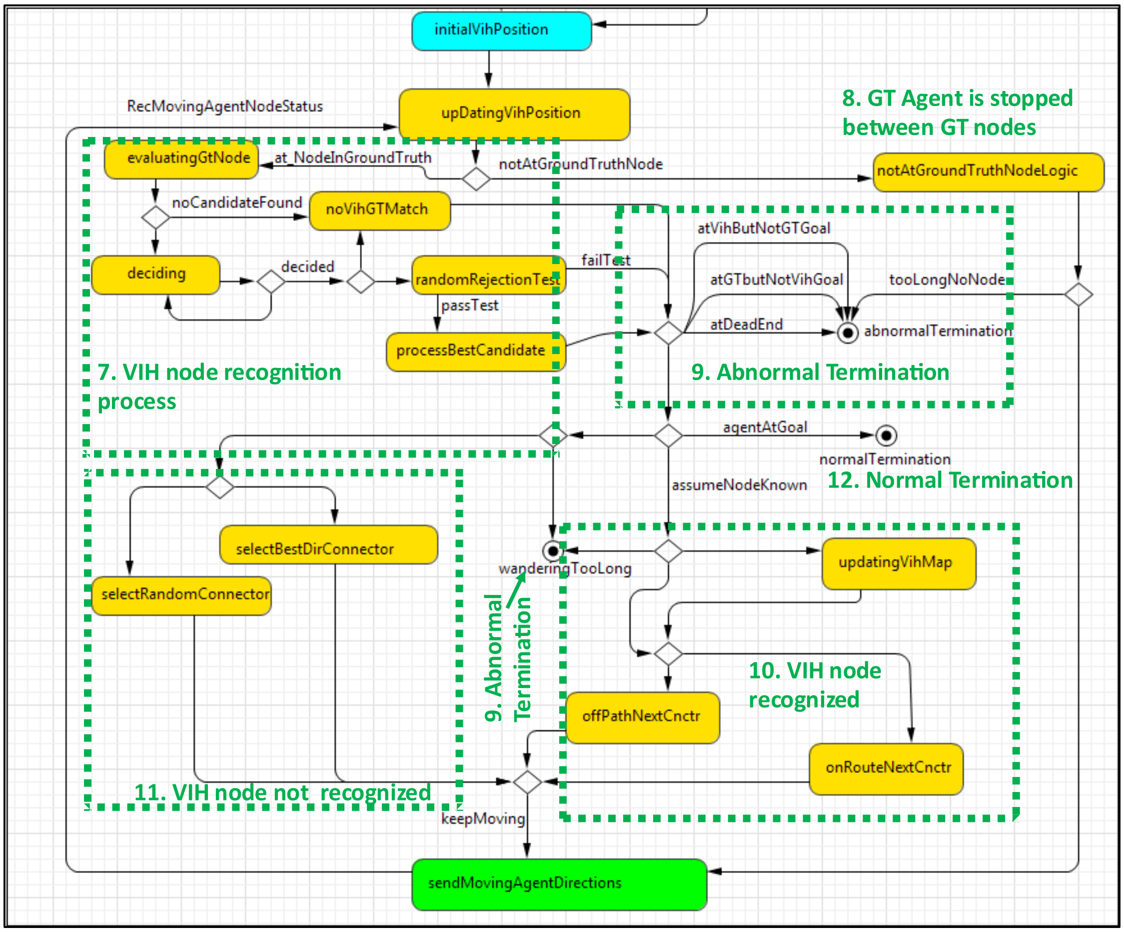


Figure 8-10 VIH Decision Processes

Should the GT agent not be located at a GT node, it must be on a GT map arc. If that arc is recognized as on the VIH map, it has a movement goal. The goal is the VIH map node in which the given arc terminates. The agent must decide whether to continue on the arc searching for that node, return to the start node of the arc and seek another link path, or simply quit as being unwilling to search further.

Block 9 details the conditions under which the entity will terminate the experiment without achieving its goal:

- dead end – the entity finds itself at a ground truth node from which it has tried all possible exit paths;
- lost no node – the entity travels too far on an arc without finding a ground truth node, reaching a maximum distance set by the user as a threshold parameter;
- wandering too long – the entity has visited too many ground truth nodes, exceeding a user-set threshold parameter, without being able to identify itself as on its planned path (i.e., the entity's perceived movement history has not agreed with the sequence of VIH nodes it believes leads to its goal);
- failure to recognize a true goal – the entity finds its way to the goal node in ground truth, but fails to recognize it; or
- accepts a false goal – the entity erroneously identifies a ground truth node as its true goal and stops.

Block 10 shows the decisions the VIH agent must make if it "recognizes" its current GT node position as a VIH node. If that node coincides with a node on the VIH Dijkstra path, the agent will attempt to follow the VIH arc link to the next node on the path. If that node does not coincide with a node on the VIH Dijkstra path, it will select the VIH arc link whose terminal node it believes best advances it towards its current intermediate goal. In both of these cases, the VIH agent must determine which of the available GT arcs most closely corresponds to the chosen VIH arc link.

The **Dijkstra route** is the sequence of nodes and arcs leading from the start point to the final goal with the shortest total distance. Each node along the route is linked to the next in the sequence by the arc that connects the two. Each time the GT agent arrives at a GT node, the VIH agent seeks to identify the VIH node that most closely corresponds to that GT node. If the entity is successfully following the VIH Dijkstra route, that node will be the *expected node* in the route sequence. If that node is also the end goal, the simulation ends. Otherwise the VIH agent will then attempt to identify the GT arc that most closely corresponds to the VIH arc leading to the next node in the VIH route sequence and direct the GT agent to follow that GT arc until it either arrives at another GT arc, or stops at some point along the arc, having exceeded the distance the entity is willing to go in searching for such a node.

If the GT agent's stopping point is a GT node that does not match the appropriate Route node, or if the GT agent was unable to arrive at a GT node, the entity has lost its way with respect to the planned route, and must either quit, terminating the simulation without successfully meeting its goal, or switch to another movement paradigm.

The current model provides two movement paradigms in addition to Dijkstra route following. The three basic VIH movement paradigms provide different ways of selecting from among the GT node arc choices available at the end of the GT movement. Should the GT agent's movement stop short of a GT node, the VIH agent must decide to either: (1) retreat to the originating node of its current GT arc or to (2) continue forward in search of the terminal node of that arc; in either case applying one of the three basic paradigms.

Both of the alternatives to Dijkstra route-following assume the VIH agent can define one or more intermediate goal nodes as local search targets. These **intermediate targets** then take on the role of the Dijkstra sequence "next node" in determining where the VIH agent wants to direct the GT agent to go. At present the Intermediate targets may consist of the as yet unreached Dijkstra route nodes and the end goal, but the model is designed to also allow the user to identify specific VIH nodes as potential intermediate targets. This capability will, for example, support subdividing the network into regions as might represent geographical features such as locales separated by natural barriers, political boundaries, or military areas of influence.

The first alternate paradigm is the **best direction selection** in which the VIH agent finds the intermediate target node perceived to be the closest to the VIH agent's current position, calculates the compass bearing from that position to the selected intermediate target, and then chooses the GT arc that best corresponds to the VIH compass bearing.

In the current version of the model the definition of "closest" intermediate node is the Euclidean distance calculated using the VIH mental map coordinates of the entity's current position and those of the intermediate target node. Future versions of the model could replace or augment this calculation by taking into account other considerations, such as potential movement constraints imposed by geographical features such as terrain barriers and trafficability, and/or location of potential threats or other elements the entity might want to avoid.

Similarly, in the current model the calculation of GT arc correspondence with the desired VIH bearing is based on the difference between the desired VIH bearing and the compass direction of the GT arc. The available arc having the smallest absolute difference from the desired bearing is selected for the next movement. Future versions of the model could again replace or augment this calculation by taking into account other potential arc characteristics, such as identifying a set of contiguous arcs as representing a road or highway, or identifying the arcs as belonging to or connecting geographical regions.

The final movement paradigm in the current model is the **random draw**; the VIH agent picks at random from the GT arc choices available. While this paradigm can be used at any time in the simulation process, its most effective use is as a last resort local search for an intermediate target and/or a landmark of some kind. If the randomly moving agent finds itself at an intermediate target, it can attempt to calculate a new Dijkstra route, and revert to a more deliberate movement decision process. Block 11 shows the options the VIH agent has if it does not recognize the current GT node as on its VIH map. It can either select randomly from the arc links available at that GT node, or select the one that most closely corresponds to the perceived direction of the VIH agent's next intermediate goal.

The logic by which the VIH agent determines its next move is summarized in Figure 8-11, which shows the states described above and the options available to the VIH agent in each of those states.

In all cases the GT arc link selected by the VIH agent and, where known, the expected distance to the next node are transmitted to the GT agent, who will travel along that arc until it reaches another GT node or exceeds its expected searching distance.

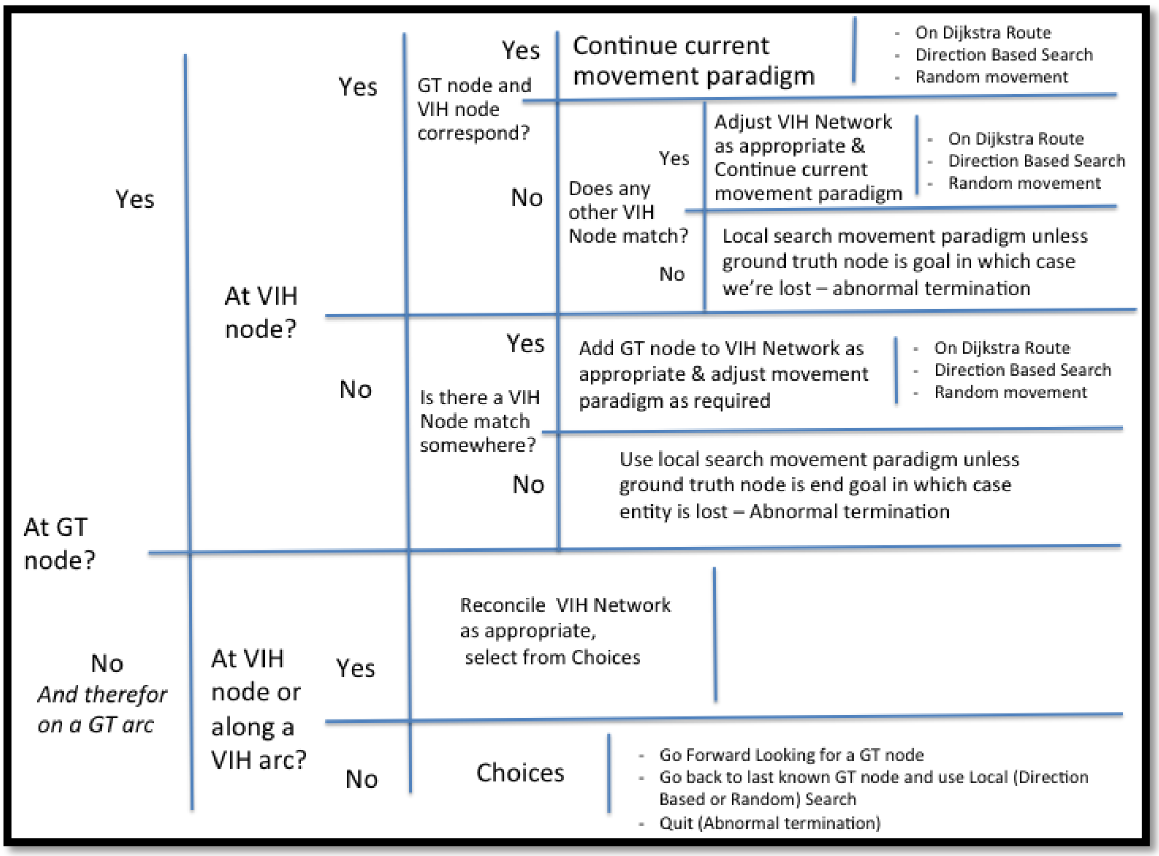


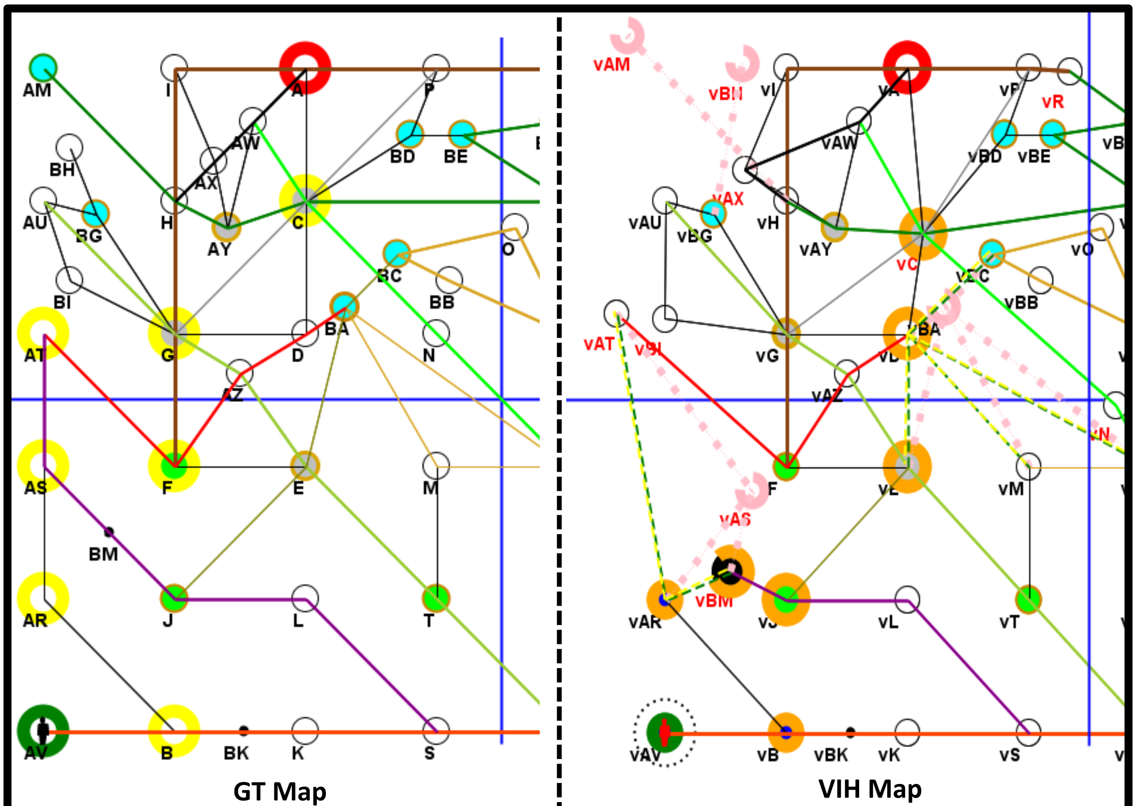
Figure 8-11 VIH Movement Selection Logic

8.6 Route Planning and Route Following

THE FIRST STEP IN THE AGENT'S GOAL SEEKING PROCESS IS ROUTE PLANNING, DEFINING A PATH FROM THE START NODE TO THE GOAL NODE. DIJKSTRA'S ALGORITHM IS USED TO FIND A SEQUENCE OF CONNECTED NODES CONSTITUTING THE "SHORTEST"[89] PATH FROM THE START NODE TO THE GOAL NODE. THE YELLOW-HIGHLIGHTED NODES IN Figure 8-12 SHOW THE GROUND TRUTH MAP PATH

[89] Here the shortest path is defined to minimize path Euclidean distance, but the algorithm is easily modified to consider factors such as movement speed and risk of delays or other constraints on movement.

FROM NODE AV TO NODE A, AND THE ORANGE-HIGHLIGHTED NODES ON THE VIH MAP FOR THE ASSOCIATED VIH NODES vAV AND vA. THE DECISION-MAKING ELEMENT OF THE ENTITY, THE VIH AGENT, HAS NO KNOWLEDGE OF THE GROUND TRUTH DIJKSTRA PATH. THAT PATH IS ONLY AVAILABLE TO THE ANALYST, AND IS PROVIDED TO SUPPORT ASSESSMENT OF HOW WELL THE AGENT DOES USING THE VIH MAP AS ITS MOVEMENT GUIDE.



GT_Dijkstra	9 nodes	AV	B	AR	AS	AT	F	G	C	A
VIH_Dijkstra	9 nodes	vAV	vB	vAR	vBM	vJ	vE	vD	vC	vA

Figure 8-12 GT and VIH Dijkstra Routes

In the case shown here, the VIH map is missing GT node AS, which is on the true shortest path to the goal. As seen in Figure 8-13, when the agent arrives at GT Node AR, it has two possible choices for continuing, the arcs linking AR to AS or to B. The agent tries to follow its VIH path which includes the arc it believes links vAR to vBM, but this link has no ground truth analogue.

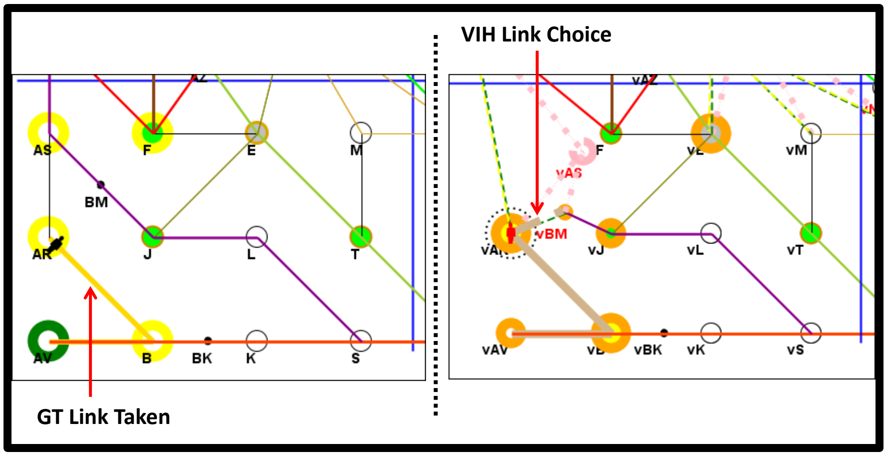


Figure 8-13 Arc Link Options at Node AR

In the current situation the agent arrives at AR from B, but the link back to B most closely corresponds to its VIH target, linking vAR to vBM, so it returns to B and ends up eventually in a dead-end situation. A contributing factor to the agent's failure may well be the nature of the arc node network, which in this case provides very limited link options for the agent at certain nodes, such as AR. One could speculate that the agent might have had more success if this problem occurred on the regular square grid map of Figure 8-2, which provides more options for movement at each node. As will be seen in later experimental results, there is some evidence for this premise. When all other factors are equal, the agent generally has better goal achievement statistics for the square grid map than for either the polygon or the multi-factor maps.

8. Future Work Backtracking Logic

This particular failure of the agent to reach the goal could be corrected adding by logic to avoid back-tracking to previously visited nodes, but such logic should then also consider those occasions when backtracking is in fact desirable. For example, should the agent stray from its optimal path by taking a wrong turn (i.e.; choosing the incorrect node link), its best strategy might be to return to a previously known location.

8.7 Node Recognition

As the agent moves through the ground truth arc/node network, it attempts to follow its VIH path plan and reconcile its VIH Map view of its position with its ground truth observations. A primary component of the reconciliation process occurs when the agent arrives at a GT node, and tries to match that GT node with a corresponding VIH node. There are several steps to this process.

First, the agent ascertains, which, if any, VIH Map nodes are candidate matches for the GT node by finding all of the VIH nodes within a given distance of the currently perceived VIH agent location. This distance is controlled by an input parameter that should generally be set to an amount at least as great as the potential distortion in VIH node X and Y coordinates (if it is known to the agent!). Setting the parameter too small might lead to failure to consider the actual Gt-VIH match, setting it too large could lead to confusing another candidate with the actual match.

Figure 8-14 shows the candidate nodes for an agent at GT node B in the above example. For scale purposes, the map considers each integer change along the X and Y axis as a single unit, for example nodes AV and B are 100 units apart. The distance parameter in this instance is set at 125 units, yielding candidate VIH nodes: vAv, vB, vBk, vK, and vJ as potentially matching GT node B.

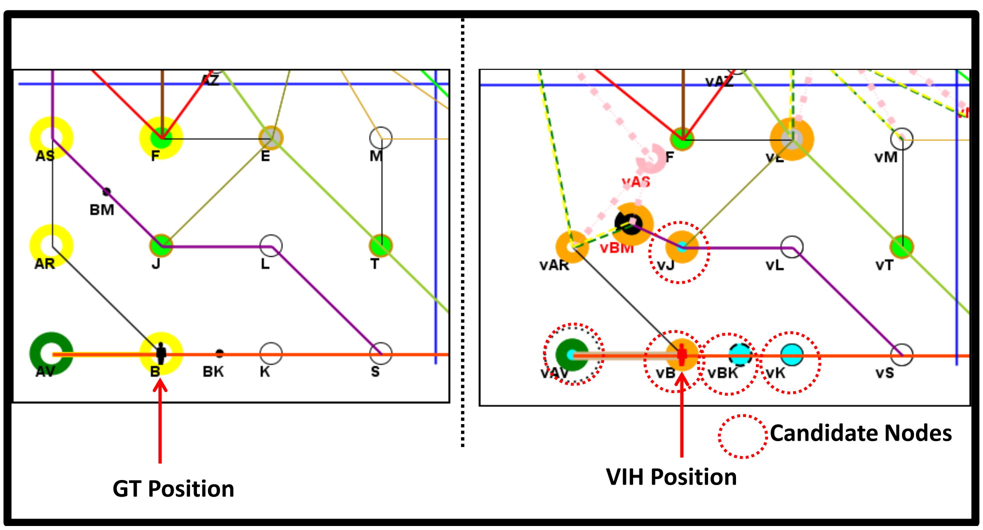


Figure 8-14 Candidate Nodes

The candidates are compared to the true GT node based on several characteristics:

- Offset distance: the distance in the VIH coordinate system of the candidate node from the VIH location corresponding to the true GT Node position.
- Expectation: a binary value set to one if the VIH node corresponds to the node the agent expects to find at the end of the link taken and zero otherwise.
- Arcs/Links Comparison: difference between the number of links per quadrant of the GT Node and the candidate node
- Color 1 and Color 2 Comparison: each node has two color characteristics designed to represent generic factors. The difference between the GT Node and the candidate node for each color is calculated by comparing RGB values (integers between 0 and 255) for each.

Each of these characteristics is quantified according to the measures in the value based hierarchy shown in Figure 8-15. The resulting values are aggregated into a single overall GT- VIH correspondence value. The best value is then tested against a confidence limit threshold. If it passes, the agent “recognizes” it as its VIH node location, which it assumes to be the same as its GT node position. If it fails, the agent does not associate a VIH node with its current position.

The value hierarchy approach used at AFIT defines single decomposition value functions (SDVFs) for each of the measures in the hierarchy. These functions convert the range of possible input values for each measure to outputs on a unit-less zero to one scale. SDVFs can take virtually any functional form desired, for example:

- constant returns to scale – linear functions;
- decreasing returns to scale – concave functions;
- increasing returns to scale – convex functions;
- combinations of the above – “S” shaped functions;
- categorical functions – discrete value step function;
- conditional functions – classical “if then else” statements; or
- fuzzy logic-based conditional functions – “if then else” statements using

fuzzy set theory.

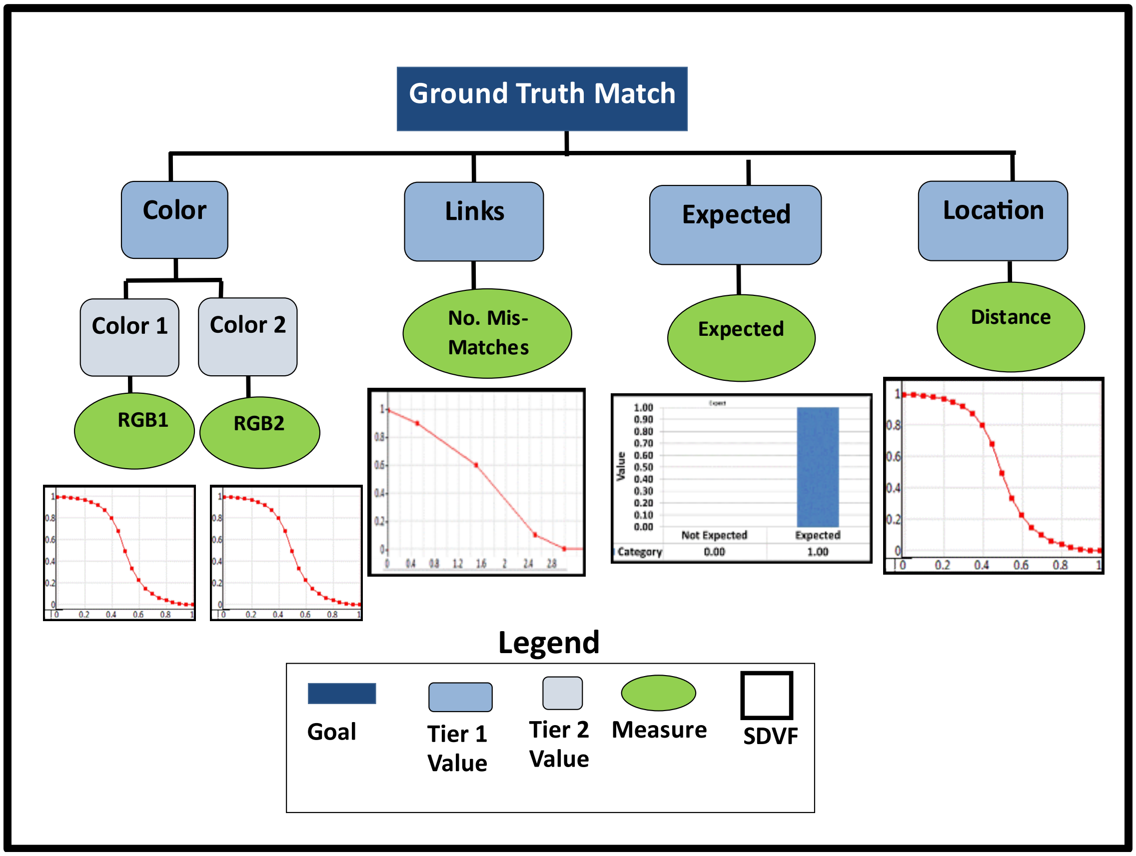


Figure 8-15 Node Recognition Value Hierarchy

At present the model uses relatively simple functional forms, each of which are discussed below.

8.7.1 Distance Match

Construction of the arc/node map in the VIH coordinate system begins with replicating the GT map with a linear offset from the GT coordinate system. Thus there is a translation of the GT origin to a VIH origin with all of the distance relationships between the objects in the two system remaining the same, as measured from their respective origins. Distortion of the elements in the VIH system changes their relationship to the VIH origin and to each other. The extent of this distortion is measured by calculating the distance from distorted VIH element positions to their original offset from the GT coordinate system. Comparing the relative position of a candidate VIH node to a GT node is computed as:

```
distanceToGTNodeLocation = calcPointToPointDistance((VIHNode.X()-
VIHNetworkXoffset),(VIHNode.Y()-
VIHNetworkYoffset),gtNode.X(),gtNode.Y());
```

where the `calcPointToPointDistance` function calculates the Euclidean distance between the two points. When this distance is zero, the position of the two nodes is a perfect match, as would be the case if the VIH node is a non-distorted copy of the GT node. The model divides the distance by an upper limit parameter to linearly convert the distance variation to a fraction between zero and one, where one is the maximum possible distortion value, and distances that are greater than or equal to the upper limit parameter are assumed equal to one. This distance fraction value is used as the abscissa for the S- shape curve shown in Figure 8-16. The curve could, of course, have been constructed using the actual distance values as the abscissa, but separating the two calculations allows reuse of the same standard curve table for other measures.

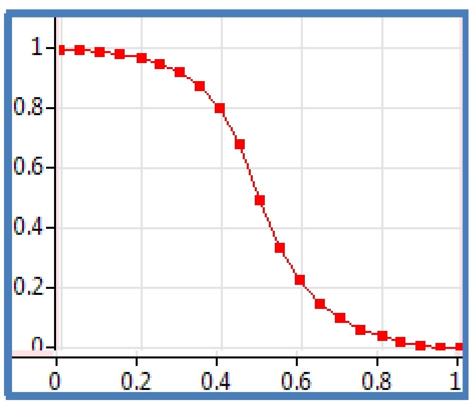


Figure 8-16 S-Shaped Value Curve

8.7.2 Expectation Match

There is an extensive literature on the role of expectation in situation awareness and how individual expectations bias a variety of human decision processes. The candidate selection value hierarchy includes a measure to account for expectation bias. This measure allows favoring a candidate node in selection of a GT node match if it is the node the agent expects to find at a given point on its route. As shown in Figure 8-17, it is a binary measure, taking on only the values zero and one. It provides a simple example of the categorical measures supported by the value hierarchy methodology, which supports multiple categories, requiring only that the sum of their values is equal to one.

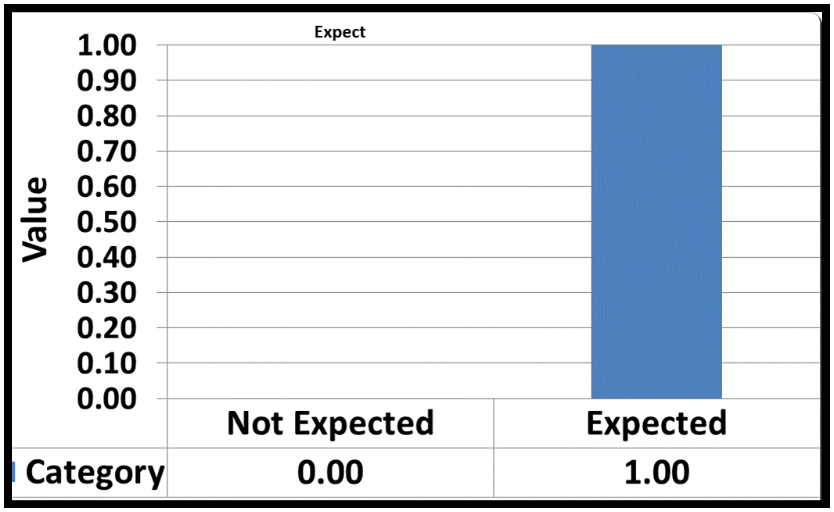


Figure 8-17 Binary Expectation Measure

8.7.3 Links Match

A key characteristic of any node is the number of arcs that lead from it or into it, as well as the directions those arcs take in linking the node to other nodes. Comparison of two nodes with respect to arc links is based on the number of links in each quadrant of a coordinate system with its origin at the node. As shown in Figure 8-18, Anylogic coordinate systems are oriented with the positive Y-axis below the X-axis as opposed to the Western tradition of placing it above the X-axis. As also shown in the figure, bearing angles are calculated as the counter-clockwise offset from the positive Y-axis.

I use a fuzzy quadrant membership correspondence in order to not over-exaggerate direction differences of links that may be only slightly distorted, but which might cross quadrant boundaries, i.e., links that are close to the boundary of two quadrants are considered to be partially in each of those two quadrants. Each arc link is associated with the angle, α, it makes with the initial quadrant boundary, as defined by the counter-clockwise initial limit of the quadrant and calculated as the remainder in integer division of the bearing by 90^o. Examples are shown in Figure 8-18. The result, unlike the reference angles used for standard trigonometric functions, can thus be oriented to either the X or Y-axis as appropriate.

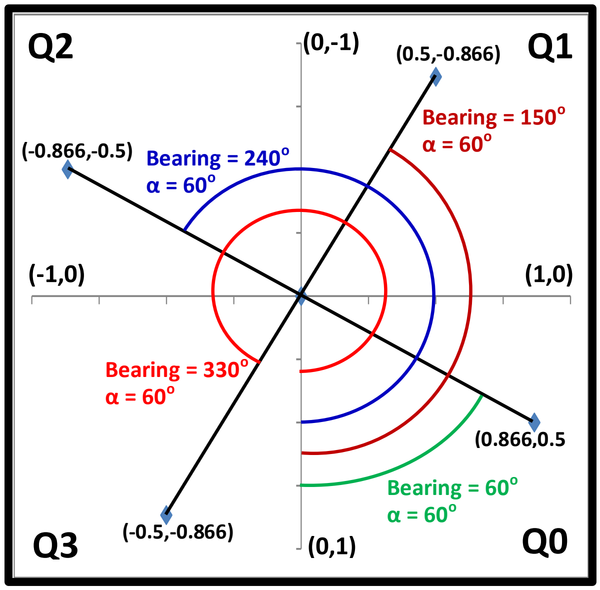


Figure 8-18 Calcuation of Bearing and Quadrant Membership Angles on the AnyLogic Coordinate System

Given the angle α, quadrant membership is calculated as shown in Figure 8-19. In this implementation of the membership function, angles between 20^o and 70^o are assumed to be fully in the quadrant, i.e., quadrant membership equal to one. Angles from 0^o to 20^o are assumed to take on linearly increasing membership in the given quadrant from 0.5 to 1.0 and linearly decreasing membership in the previous quadrant from 0.5 to 0.0. Similarly angles between 70^o and 90^o are assumed to take on linearly decreasing membership from 1.0 to 0.5 in the given quadrant and increasing membership in the next quadrant from 0.0 to 0.5. Experimentation with minor modifications of the membership function has shown that the model results and conclusions are not sensitive to its exact form.

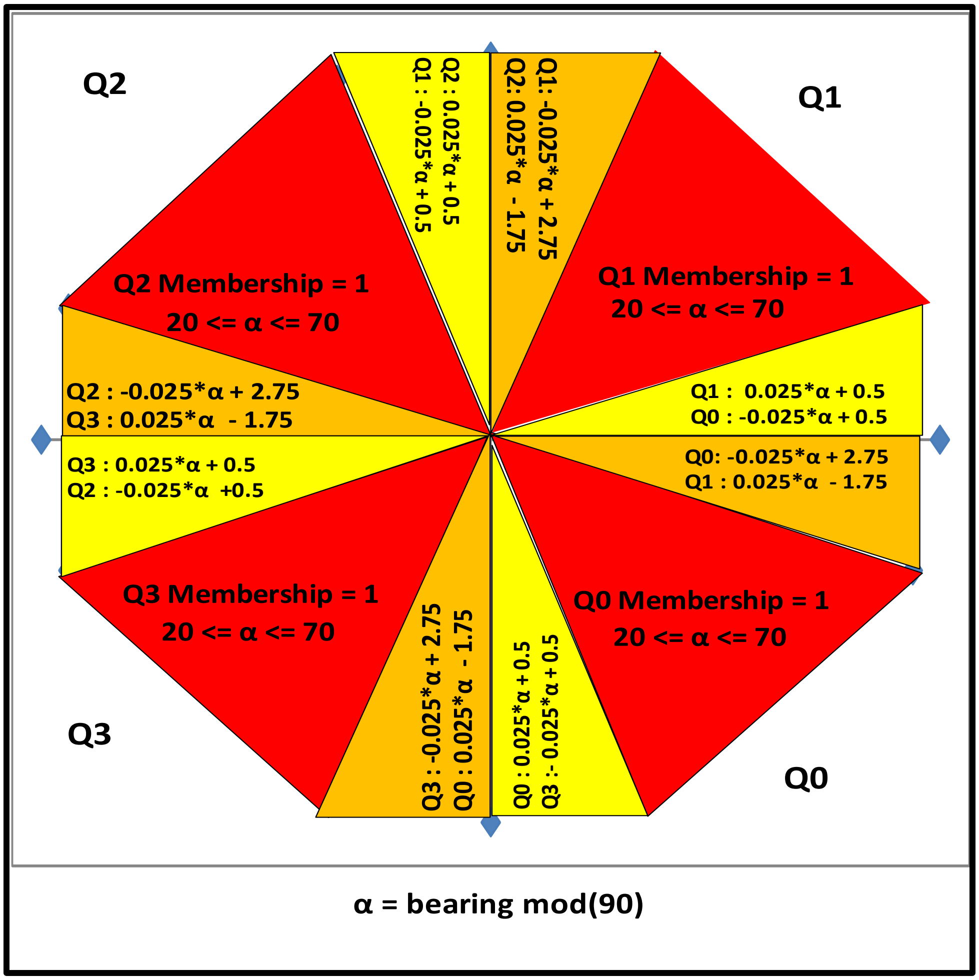


Figure 8-19 Quadrant Membership Calculation

The resultant fractional membership values for all node links are added by quadrant and stored as the number of links per quadrant. Comparison of nodes by link matches calculates the difference in the fractional membership values by quadrant for the two nodes, with the total difference over all four quadrants representing the links difference between the two nodes. This difference value is used as the x-axis value for the value hierarchy measure shown in Figure 8-20 .

Here the measure is implemented as a piece-wise linear function with the upper limit set to 1.0 after the links difference exceeds 3.0. This function could, of course, have been implemented using the two-step procedure from the distance difference match measure. The use of the piece-wise linear function is primarily to illustrate that form of measure, and again, in this instance the specific form has little or no effect on results and conclusions.

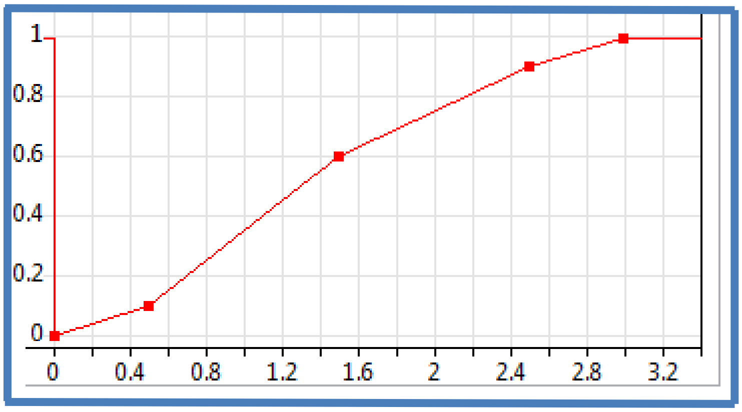


Figure 8-20 Links Difference Match Measure

A more comprehensive implementation of the links match measure would look at a number of other potentially distinguishing features. To name just a few:

- is the link part of a specific road, highway or other wayfare;
- does the link have specific terrain attributes affecting trafficability;
- is it possessed of, or does it border, visually distinguishing features?

9. FUTURE WORK: LINKS FEATURES

Exploration of such features is a potentially rich source of research, with respect to all aspects of agent movement: route planning, route following, node recognition and way finding.

8.7.4 Color Match

There are a potentially infinite number of features with continuous or semi-continuous membership measures that could characterize nodes: degree of urban/rural qualities; population; extent of risk from adversary action or natural causes; availability of desired resources; trafficability constraints, and so on. Any or all of such features could be distorted in the VIH map, and might be considered while attempting to match VIH candidate nodes to GT nodes. Rather than selecting specific instances of these, two generic factors were implemented as color characteristics.

These characteristics are designated as Color 1 and Color 2 attributes of each node and can take default values as the line and fill color of the node's AnyLogic presentation object. Each color attribute actually specifies three separate values, the description of the color according to its RGB components. In RGB format, the red, blue, and green components of a color are given as an integer from 0 to 255, where 0 indicates no contribution from this primary color, and 255 indicates that it is present in its maximum intensity. As with the links difference, for each node being compared, the differences in each component for a color attribute are calculated and summed. As with the distance match component, this sum is normalized by dividing by the maximum possible difference, which for any two colors is 765, and the result used as the x-axis value for the measure illustrated in Figure 8-16.

8.7.5 Candidate Node Selection

For each tier of the value hierarchy branch, values are calculated as a weighted sum of the values from the immediately preceding branch elements. Weights are established as model input parameters and are subject to the constraint that for all the branches connecting at that tier the weights must sum to one. Thus at each tier, each branch of the value hierarchy contributes a normalized and weighted value to the tier above. For the value hierarchy of Figure 8-15, the individual normalized measures for the two color attributes are weighted and summed to a single value between zero and one, and that value weighted and summed with all the normalized measure values of the other attributes.

Table 8-1 Candidate Node Comparison

GT Node	VIH Candidate	Distance Match	Links Match	Expectation Match	Color 1 Match	Color 2 Match		GT - VIH Correspondence		Fit Delta
	vB	1.00	1.00	1.00	1.00	1.00		1.00		0.00
	vK	0.04	0.65	0.00	1.00	1.00		0.54		0.46
B	vBK	0.75	0.65	0.00	0.00	1.00		0.48		0.52
	vAV	0.04	0.25	0.00	1.00	1.00		0.46		0.54
	vJ	0.04	0.25	0.00	0.00	0.00		0.06		0.94

For the example of Figure 8-14,

Table 8-1 Candidate Node Comparison shows the value hierarchy GT-VIH correspondence of the candidate nodes.

In this instance, VIH node vB, with no distortion from the ground truth map in any way, has a perfect match value of 1.0. Hence vB will be selected as the VIH node matching the current GT node. Since this node is on the VIH path to the goal, the agent will next choose the link it believes leads to the next node on the route. As noted above, for this particular example, the agent fails to find the correct link when it reaches the next node in its path and instead backtracks. The complete node history for this example is: AV, B, AR, B, AV; the entity ends up back at its start node and dead ends. At that point the GT agent has no untaken links left to try. In this particular case, the VIH agent's node recognition was perfect, it recognized all of the nodes it encountered, but the entity still failed to reach its goal.

8.7.6 Recognition Results

For post-simulation analysis the model classifies recognition results in several categories:

- True Positive – the recognized node corresponds to the GT node;
- True Negative – none of the candidate nodes correspond to the GT node;
- False Positive – the recognized node does not actually correspond to the GT node;
- False Negative – the agent does not recognize a VIH node candidate when one actually corresponds to the GT node;
- Recognition Fail – sum of the results for false negative, true negative, and rejection of a true positive based on a Monte Carlo draw against an input parameter to simulate random errors in recognition.

The number and relative frequency of node recognition decisions resulting in each category are provided as part of the simulation output.

9 EXPLORATION OF VIH MAP DISTORTION EXPERIMENTS

This set of simulation experiments explores the effects of mental map distortion on agent movement. The simulation DOE was designed to emphasize variation in spatial distortion, both with respect to the number of nodes whose position in modified on the VIH Map and with respect to extent of x and y coordinate shifts.

Some topological distortion is introduced by removing randomly selected nodes from the mental map, but nodes to which they connect are reconnected to each other to maintain overall network connectivity.

It is possible to represent distortion in node color characteristics as well, but this capability was not used for this set of experiments.

9.1 Experimental Setup

Simulation trials were executed for each map, varying all of the parameters listed in Table 9-1, with each combination of parameter values run for fifty trials.

The simulation experiments used the three basic ground truth maps shown in Figure 8-2, Figure 8-3 and Figure 8-4. The first three parameters listed in the table were used to construct the different VIH maps for each trial, specifying the amount and extent of distortion of the given GT map as described above in Chapter 8.4. The next two values given identify the node match weights and the resulting match weight coefficients in the node recognition value hierarchy as described above in Chapter 8.7. Finally, the last row provides the three values used for GT-VIH Threshold confidence limit as described in that chapter.

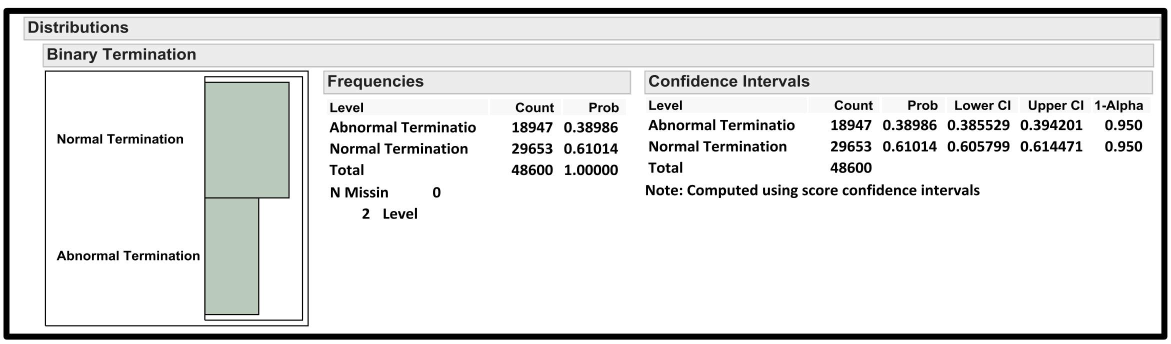
Each of the 36 different combinations of distortion parameters was run against each of the three threshold confidence limit values, with 150 trials for each case, resulting in 36x3x150 trials for each of the three GT map, for a total of 48,600 simulation runs.

Table 9-1 Experiment 1 Parameter Values

Parameter	Parameter Values
Pr(NE): Probability node location (X and Y values) is distorted, i.e., probability node X and Y values will vary from Ground Truth on the mental map	**0.1, 0.3, 0.5, 0.7**
X and Y Er Lim: maximum amount of distortion in X and Y	**25, 75, 125**
Pr(MN): Probability of missing nodes i.e., probability that a selected node will be removed from the mental map	**0.0, 0.15., 0.3**
All Match Weights[90]	**1**
All Match Coefficients	**0.2**
GT-VIH Correspondence Threshold	**0.65, 0.75, 0.85**

9.2 Results: Assessment of Normal vs. Abnormal Termination

The first question of interest for each trial is whether the entity succeeds in finding its goal. As shown in Figure 9-1, the entity's overall success rate is 61%.



Frequencies

Level	Count	Prob
Abnormal Terminatio	18947	0.38986
Normal Termination	29653	0.61014
Total	48600	1.00000
N Missin 0		
2 Level		

Confidence Intervals

Level	Count	Prob	Lower CI	Upper CI	1-Alpha
Abnormal Terminatio	18947	0.38986	0.385529	0.394201	0.950
Normal Termination	29653	0.61014	0.605799	0.614471	0.950
Total	48600				

Note: Computed using score confidence intervals

Figure 9-1 Frequency of Normal and Abnormal Termination Over All Trials

[90] The match weight values are used to determine match coefficients, the match coefficients for each criteria equals the match weight of that criterion divided by the total of all match weights. As shown here, with all the match weights are set to 1, all the match coefficients equal 0.2

9.2.1 Statistical Significance[91]

It is important to begin a discussion of the statistical treatment of these simulation experiments by reiterating the over-all goal of the human-centric approach, which is not to predict human behavior but to study it. Model fit techniques are employed primarily to provide insight into the nature of the results, and are used to determine when and how input parameters affect those results. In cases where models are fit to nominal or categorical results, the concern is less with the degree of predictive classification provided by those models than with the explanatory insights supported.

For categorical data, JMP calculates a Logistic Fit Model instead of the standard least squares fit for numerical data. Logistic regression fits nominal *Y* responses to a linear model of *X* terms. The standard JMP model for nominal or categorical data uses logistic regression based on maximum-likelihood estimation, selecting the set of X coefficients, $\boldsymbol{\beta}_i$, that maximize the probability of the model output matching the Y_j values actually observed in the experimental results.

For binary response levels, the function is:

$$\Pr(Y = r_1) = \left(1 + e^{-X\beta}\right)^{-1}$$ where r_1 is the first response

OR EQUIVALENTLY:

$$log\left(\frac{\Pr(Y=r_1)}{\Pr(Y=r_2)}\right) = X\beta$$ where r_1 and r_2 are the responses.

In the case where r>2, i.e, more than two nominal responses, JMP fits r-1 sets of linear model parameters of the following form:

$$log\left(\frac{\Pr(Y=j)}{\Pr(Y=r)}\right) = X\beta_j$$

[91] The details on logistic fit in this chapter are taken from JMP documentation found at http://www.jmp.com/support/help/Introduction_to_Logistic_Models.shtml#97437

As stated above, the fitting principal of maximum likelihood means that the βs are chosen to maximize the joint probability attributed by the model to the responses that did occur. This fitting principal is equivalent to minimizing the negative log-likelihood:

$$-logLikelihood = \sum_{i=1}^{n} -\log(\Pr(i\text{th } row\ has\ the\ y_j th\ response))$$

JMP provides several tests to compare the fit of the specified model with subset or superset models, as illustrated in Figure 9-2.

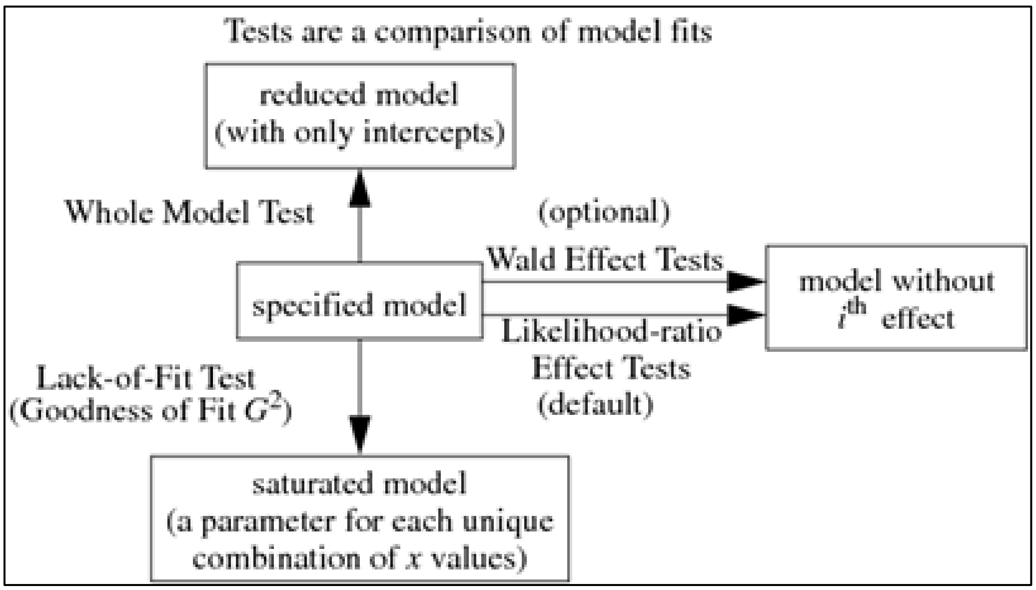


Figure 9-2 Relationship of Statistical Tables[92]

If a test shows significance, then the higher order model is justified:

- Whole model tests: if the specified model is significantly better than a reduced model without any effects except the intercepts.
- Lack of Fit tests: if a saturated model is significantly better than the specified model.

---

[92] Figure from http://www.jmp.com/support/help/The_Logistic_Fit_Report.shtml

- Effect tests: if the specified model is significantly better than a model without a given effect.

For the complete model fit of all 16,200 trials considering all two-way interactions between the five variables (Map Type, Pr(NE), XandY Error Limit, Pr(MN), and Confidence Threshold Limit) that were varied for this experiment. The logistic fit model converged in gradient in five iterations, with the results shown in Table 9-2.

Table 9-2 Whole Model Test

Whole Model Test
Nominal Logistic Fit for Binary Termination
Converged in Gradient, 5 iterations

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	9038.962	20	18077.92	<.0001
Full	23459.061			
Reduced	32498.024			
RSquare (U)	0.2781			
AICc	46960.1			
BIC	47144.7			
Observations	48600			

Measure	Training	Definition
Entropy RSquare	0.2781	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.4212	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.4827	∑ -Log(ρ[j])/n
RMSE	0.3982	$\sqrt{\sum(y[j]-\rho[j])^2/n}$
Mean Abs Dev	0.3165	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.2358	∑ (ρ[j]≠ρMax)/n
N	48600	n

The Whole Model Table tests the null hypothesis that all of the $\boldsymbol{\beta}_i$s are zero; it compares the whole-model fit to the model that omits all the regressor effects except the intercept parameters. The test is analogous to the Analysis of Variance table for continuous responses. The negative log-likelihood corresponds to the sums of squares, and the Chi-square test corresponds to the *F* test.

The JMP Whole Model Table uses the terminology shown in Table 9-2:

- **Difference** - the difference between the Reduced and Full models. It measures the significance of the regressors as a whole to the fit.
- **Full** - the negative log-likelihood for the complete model.
- **Reduced** - the negative log-likelihood that results from a model with only intercept parameters. Here the –LogLikelihood for the reduced model that includes only the intercepts is 32498.024.
- **DF -** associated degrees of freedom (DF) for the Difference between the Full and Reduced model. Here, there two-way factorial design of the fit considers the possible one-way and two-way combinations of variables, which is 15, but since the map type is a categorical values with 3 possible values, it contributes two degrees of freedom for each of the 5 combinations in which it is considered, yielding a total of 20 degrees of freedom.
- **Chi-Square -** Likelihood-ratio Chi-square test for the hypothesis that all regression parameters are zero. It is computed by taking twice the difference in negative log-likelihoods between the fitted model and the reduced model that has only intercepts.
- **Prob>ChiSq -** probability of obtaining a greater Chi-square value by chance alone if the specified model fits no better than the model that includes only intercepts. Here the value is far less than 0.001 indicating a very high degree of significance to the contributions of the full model.
- **RSquare (U) -** the R^2 value, which is the ratio of the **Difference** to the **Reduced** negative log-likelihood values. It is sometimes referred to as *U*, the uncertainty coefficient. **RSquare** ranges from zero for no improvement to 1 for a perfect fit. As in this case, a **Nominal** model rarely has a high **Rsquare**, and it has a **Rsquare** of 1 only when all the probabilities of the events that occur are 1.
- **Measure -** gives several measures of fit to assess model accuracy.

Of these measures provide, probably the most useful is the **Misclassification Rate,** the rate for which the response category with the highest fitted probability is not the observed category. In this instance, the misclassification rate of 0.2358 of the fitted model can be compared to the misclassification rate of 0.3898, which would be obtained in one simply guessed that the result on any trial would be the most common response over all the trials. In this case that response would be Abnormal Termination, which occurs approximately 61% of the time. Table 9-3 shows the confusion matrix for the Logistic Regression Fit Model calculated here.

Table 9-3 Confusion Matrix

Confusion Matrix		
Actual/Predicted	**Abnormal Termination**	**Normal Termination**
Abnormal Termination	**11469**	**7478**
Normal Termination	**3981**	**25672**

As similar view of the effectiveness of the model is provided by the Receiver Operating Curve (ROC), shown for the current model in Figure 9-3. In this instance, the area under the curve (AUC) is 0.83838. This value shows that the model provides a fair amount of predictive capability, but still allows for further refinement through the consideration of other factors.

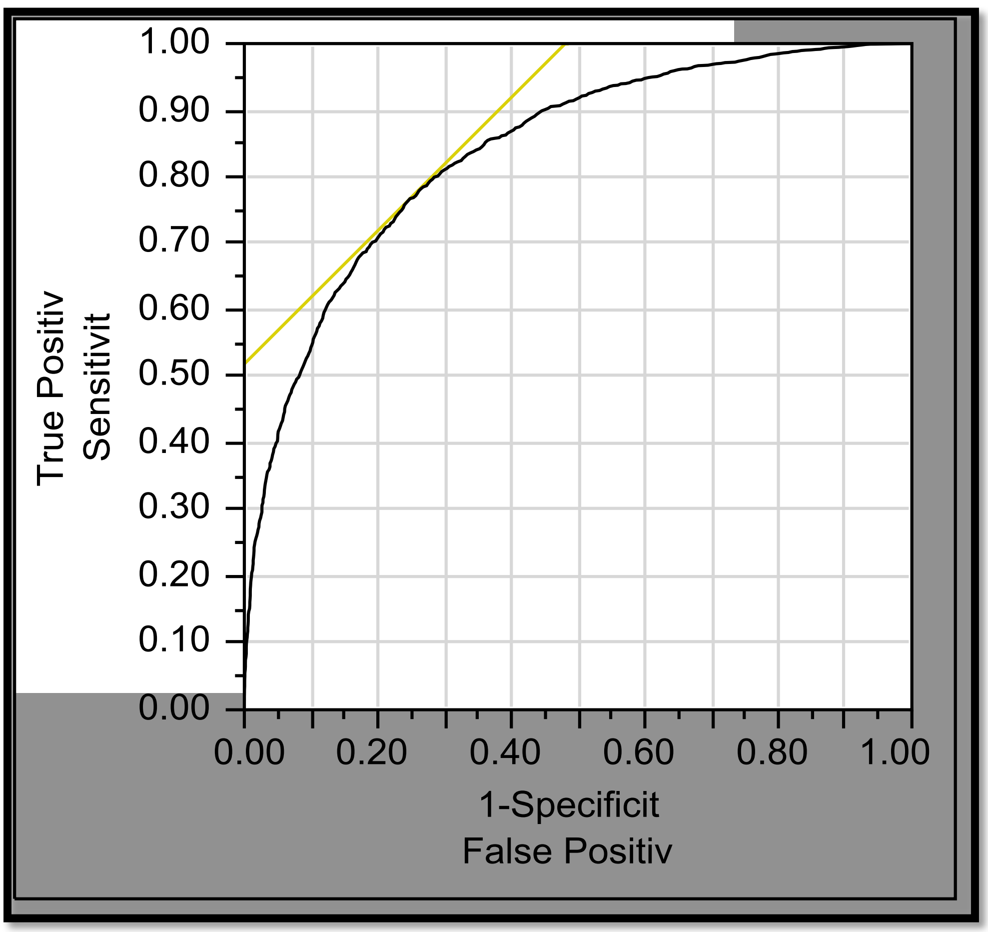


Figure 9-3 Logistic Fit Model ROC Curve

The model's capability can be further quantified through provides a Lack of Fit test, sometimes called a Goodness of Fit test. JMP calculates lack of fit using a pure-error negative log-likelihood, constructing categories for every combination of the regressor values in the data (the Saturated line in the Lack Of Fit table), and it testing whether this log-likelihood is significantly better than the Fitted model. JMP results are shown in Table 9-4.

Table 9-4 Lack Of Fit

Lack Of Fit				
Source	**DF**	**-LogLikelihood**	**ChiSquare**	**Prob>ChiSq**
Lack Of Fit	**303**	**1391.866**	**2783.733**	**<.0001**
Saturated	**323**	**22067.195**		
Fitted	**20**	**23459.061**		

Here there are 323 **Saturated** degrees of freedom, which is one less than the number of unique populations sampled in the experiment, i.e., 16200/50 -1. The Fitted degrees of freedom value is the number of parameters not including the intercept, which again in this case is 20. The Lack of Fit DF is the difference between the Saturated and Fitted models, in this case 323 -20 = 303.

The Lack of Fit table lists the negative log-likelihood for error due to Lack of Fit, error in a Saturated model (pure error), and the total error in the Fitted model, and again uses Chi-square statistics test for lack of fit. For the fitted logistic model, the lack of fit Chi-square is highly significant (Prob>ChiSq < 0.001) and supports the conclusion that more complex terms need to be added to the model.

Finally, it is also desirable to test the value each variable provides to the model fit. The JMP Effect Likelihood tests, for the $\boldsymbol{\beta}_i$ corresponding to each model term, the null hypothesis that $\boldsymbol{\beta}_i$, is equal to zero. If the *p* value for a given coefficient $\boldsymbol{\beta}_i$ is small, the null hypothesis should be rejected, implying that the variable X_i does influence the probability that Y_i will be predicted correctly. The Likelihood-ratio Chi-square tests are calculated as twice the difference of the log-likelihoods between the full model and the model constrained by the hypothesis to be tested (the model without the effect).

Table 9-5 lists the model terms and shows them all to be significant at the 95 % confidence level and all but one to be significant at the 99.9% confidence level.

Table 9-5 Effect Likelihood Tests

Effect Likelihood Ratio Tests				
Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Map	2	2	993.756	<.0001
Pr(NE)	1	1	3885.475	<.0001
XEr Lim	1	1	10354.970	<.0001
Pr(MN)	1	1	216.407	<.0001
GT-VIH Corspnd Thrshd	1	1	1295.929	<.0001
Map*Pr(NE)	2	2	7.706	0.0212
Map*XEr Lim	2	2	84.468	<.0001
Map*Pr(MN)	2	2	20.762	<.0001
Map* GT-VIH Corspnd Thrshd	2	2	50.613	<.0001
Pr(NE)*XEr Lim	1	1	945.211	<.0001
Pr(NE)*Pr(MN)	1	1	32.135	<.0001
Pr(NE)* GT-VIH Corspnd Thrshd	1	1	102.186	<.0001
XEr Lim*Pr(MN)	1	1	173.145	<.0001
XEr Lim* GT-VIH Corspnd Thrshd	1	1	337.804	<.0001
Pr(MN)* GT-VIH Corspnd Thrshd	1	1	20.324	<.0001

9.2.2 Analysis of Parameter Effects

Having established the significance of the distortion model parameters, it is next appropriate to examine the nature of their effects and their possible interactions.

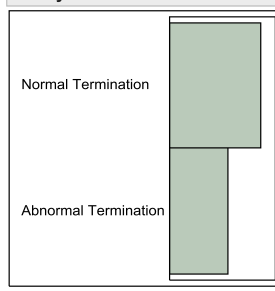
Figure 9-4 shows a breakdown of the normal termination probability observed in Figure 9-1 , now calculated as the observed frequency of normal terminations for each of the combinations of distortion error parameters. As one would expect, as the amount of total distortion increases, the fraction of normal terminations decreases. Given that each entry in the Figure 9-4 represents more than 1000 model trials, and given the size of frequency differences shown, one can conclude the differences are significant in both a statistical and a practical sense.

Further insight into the parameter effects is obtained by looking at the individual logistic fits of binary termination conditions by each of the distortion factors applied. Figure 9-5 shows the results of these individual fits, which indicate that for the cases observed, variation in the spatial distortion factors appears to translate to greater variation in termination condition than does variation in distorting the topology by removing nodes but leaving the network connected. Topological variation is revisited in Chapter 12, where the possibility of a disconnected VIH network is explored.

All Maps Equal Match Coefficients Each Case 1350 Replications		Probability of Node X and Y Error				
Probability of Missing Nodes	**Amount of X and Y Error**	**0.15**	**0.3**	**0.5**	**0.7**	**Row Average**
0	25	0.984	0.963	0.951	0.950	0.962
	75	0.830	0.624	0.459	0.361	0.568
	125	0.747	0.443	0.259	0.140	0.397
0.15	25	0.920	0.904	0.884	0.841	0.887
	75	0.776	0.600	0.466	0.378	0.555
	125	0.701	0.421	0.247	0.131	0.375
0.3	25	0.853	0.849	0.823	0.800	0.831
	75	0.770	0.594	0.464	0.340	0.542
	125	0.662	0.429	0.251	0.151	0.373
Column Average		0.805	0.647	0.534	0.455	0.610

Distributions

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	18947	0.38986
Normal Termination	29653	0.61014
Total	48600	1.00000

Figure 9-4 Probability of Normal Termination

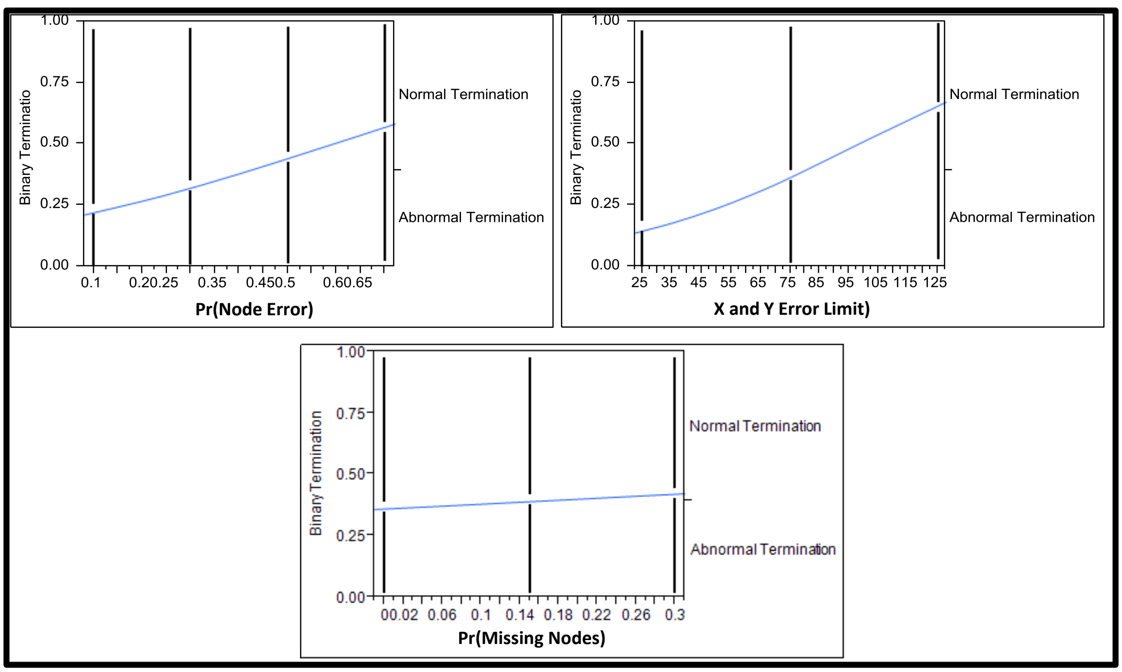


Figure 9-5 Separate Logistic Fit of Y by X for Binary Termination by Pr(NE), X and Y Error Limit, and Pr(MN)

In the cases shown here, the VIH agent had the ability to use all five of the GT-VIH match factors in determining whether any candidate VIH node passed the GT- VIH correspondence threshold. The experiments did look at one key recognition process parameter, altering the threshold limit for accepting a VIH Node as match to a GT Node.

As shown in Figure 9-6, there is little variation between the threshold at 0 .60 and that at 0.75, but considerable difference when the threshold moves to 0.90. While it is premature to make any specific conclusions at this point, it is worth noting that with all of the match coefficients set to 0.20, complete failure in one of them, as for example the binary-valued expectation coefficient, would result in a match failure at the 0.90 level but would not at the other two. Later experiments explore the effects of the match coefficients and threshold values in more detail.

All Maps GT-Vih Corspnd Thrshld 0.6 Each Case 450 Replications		Probability of Node X and Y Error				
Probability of Missing Nodes	Amount of X and Y Error	0.15	0.3	0.5	0.7	Row Average
0	25	0.969	0.969	0.929	0.940	0.952
	75	0.909	0.778	0.589	0.487	0.691
	125	0.824	0.560	0.344	0.207	0.484
0.15	25	0.902	0.860	0.869	0.813	0.861
	75	0.842	0.718	0.620	0.524	0.676
	125	0.740	0.491	0.318	0.202	0.438
0.3	25	0.833	0.836	0.787	0.787	0.811
	75	0.816	0.709	0.629	0.467	0.655
	125	0.711	0.496	0.324	0.216	0.437
Column Average		0.839	0.713	0.601	0.516	0.667

Normal Termination GT-VIH Correspondence Threshold = 0.6

All Maps GT-Vih Corspnd Thrshld 0.75 Each Case 450 Replications		Probability of Node X and Y Error				
Probability of Missing Nodes	Amount of X and Y Error	0.15	0.3	0.5	0.7	Row Average
0	25	0.984	0.933	0.947	0.947	0.953
	75	0.920	0.740	0.624	0.520	0.701
	125	0.820	0.509	0.333	0.193	0.464
0.15	25	0.918	0.900	0.882	0.831	0.883
	75	0.822	0.727	0.596	0.529	0.668
	125	0.733	0.504	0.344	0.167	0.437
0.3	25	0.851	0.858	0.822	0.796	0.832
	75	0.824	0.729	0.593	0.482	0.657
	125	0.689	0.509	0.333	0.196	0.432
Column Average		0.840	0.712	0.608	0.518	0.670

Normal Termination GT-VIH Correspondence Threshold = 0.75

All Maps GT-Vih Corspnd Thrshld 0.9 Each Case 450 Replications		Probability of Node X and Y Error				
Probability of Missing Nodes	Amount of X and Y Error	0.15	0.3	0.5	0.7	Row Average
0	25	1.000	0.987	0.978	0.962	0.982
	75	0.660	0.353	0.164	0.076	0.313
	125	0.598	0.260	0.098	0.020	0.244
0.15	25	0.940	0.951	0.902	0.878	0.918
	75	0.662	0.356	0.182	0.080	0.320
	125	0.631	0.267	0.078	0.024	0.250
0.3	25	0.873	0.853	0.860	0.818	0.851
	75	0.671	0.344	0.171	0.071	0.314
	125	0.587	0.282	0.096	0.042	0.252
Column Average		0.736	0.517	0.392	0.330	0.494

Normal Termination GT-VIH Correspondence Threshold = 0.9

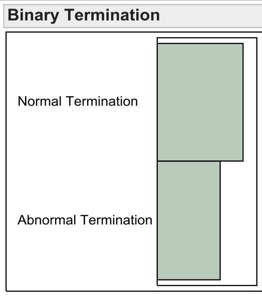
Figure 9-6 Normal Termination by GT-VIH Correspondence Thresholds

9.2.3 Normal Termination by Map

The three GT maps used in the experiment were constructed so as to offer different degrees of challenge to the entity moving on them, Figure 9-7 shows that this goal was achieved with respect to the square grid map, on which normal termination occurred significantly more often than on the other two. In this instance the polygonal map had fewer normal terminations than the multi-feature map with 95% significance, but only barely so, as the confidence intervals for the two values fail to overlap by 0.57747 (the lower confidence interval limit for the multi-factor map) – 0.56340 (the upper confidence interval value for the polygonal map) = 0.01407.

Distributions Map=MF Map

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	6845	0.42253
Normal Termination	9355	0.57747
Total	16200	1.00000

N Missin 0

2 Level

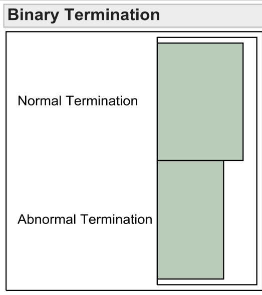
Confidence Intervals

Level	Count	Prob	Lower CI	Upper CI	1-Alpha
Abnormal Terminatio	6845	0.42253	0.414944	0.430155	0.950
Normal Termination	9355	0.57747	0.569845	0.585056	0.950
Total	16200				

Note: Computed using score confidence intervals

Distributions Map=Poly Map

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	7073	0.43660
Normal Termination	9127	0.56340
Total	16200	1.00000

N Missin 0

2 Level

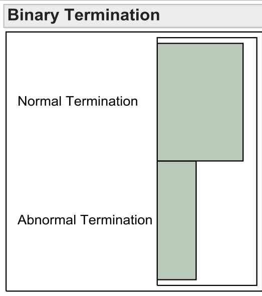
Confidence Intervals

Level	Count	Prob	Lower CI	Upper CI	1-Alpha
Abnormal Terminatio	7073	0.43660	0.428984	0.444256	0.950
Normal Termination	9127	0.56340	0.555744	0.571016	0.950
Total	16200				

Note: Computed using score confidence intervals

Distributions Map=Sqr Map

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	5029	0.31043
Normal Termination	11171	0.68957
Total	16200	1.00000

N Missin 0

2 Level

Confidence Intervals

Level	Count	Prob	Lower CI	Upper CI	1-Alpha
Abnormal Terminatio	5029	0.31043	0.303353	0.317601	0.950
Normal Termination	11171	0.68957	0.682399	0.696647	0.950
Total	16200				

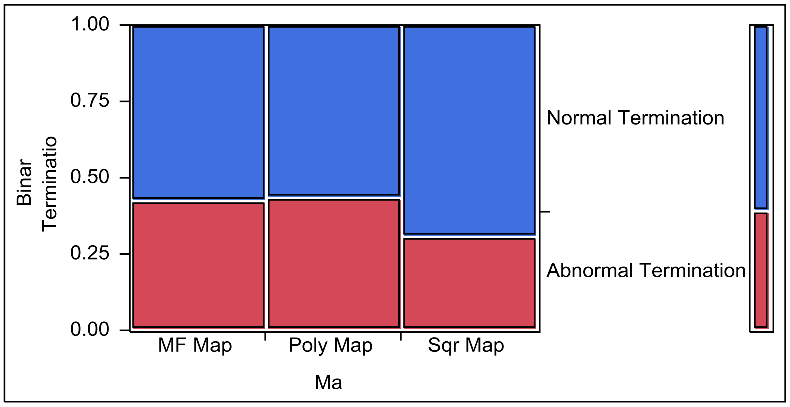
Note: Computed using score confidence intervals

Figure 9-7 Termination Condition by Map

The significance of the differences between the maps is also shown by the JMP contingency analysis in Figure 9-8.

Contingency Analysis of Binary Termination By Map

Mosaic Plot



Contingency Table

Binary Termination

Map	Count Total % Col % Row %	Abnormal Termination	Normal Termination	
	MF Map	6845	9355	16200
		14.08	19.25	33.33
		36.13	31.55	
		42.25	57.75	
	Poly Map	7073	9127	16200
		14.55	18.78	33.33
		37.33	30.78	
		43.66	56.34	
	Sqr Map	5029	11171	16200
		10.35	22.99	33.33
		26.54	37.67	
		31.04	68.96	
		18947	29653	48600
		38.99	61.01	

Tests

N	DF	-LogLike	RSquare (U)
48600	2	330.82821	0.0102

Test	ChiSquare	Prob>ChiSq
Likelihood Rati	661.656	<.0001 *
Pearson	651.169	<.0001 *

Figure 9-8 Contingency Analysis of Binary Termination by Map

9.2.4 Regression Tree Analysis

JMP also supports the use of partitioning and regression tree analysis to fit a model to experimental data. In partition analysis classes of the dependent variable, in this case binary termination mode, are divided according to their probability of occurrence given different values of the independent variables from the design of experiments as per Table 9-1.

This process begins with the construction of contingency tables based on the observed frequency of dependent variable responses compared to the expected frequency under the null hypothesis that response rates are independent of the various parameter values. JMP® software implements this process by iteratively exploring all of the possible two factor contingency tables for combinations of the parameter levels and response possibilities. For each possible contingency table JMP® calculates the G^2 statistic:

$$G^2 = 2\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}\log\left(\frac{n_{ij}}{\mu_{ij}}\right)$$

where μ_{ij} is approximated by

$$\frac{row\ sum{*}col\ sum}{total} = \{\left(\sum_i n_{ij}\right)\left(\sum_j n_{ij}\right)\}/\{\sum_i \sum_j n_{ij}\}.$$

The formula finds the expected number in each cell by multiplying the proportion of the col contributions to the overall sum times the row size. When partitioning data, the G^2 is exploring evidence for the null hypothesis that the row and col variables are statistically independent, the larger the statistic, the greater the evidence they are dependent. The ideal situation would be to partition the data so that the column data, which represent the response variables or different experimental outcomes, are completely determined by the values of the row data, which are the factors of the experimental design. The degrees of freedom in a test of independence are equal to (number of rows)−1 × (number of columns)−1

The largest resultant G^2 value in each iteration defines the partition split for that iteration. Figure 9-9 shows how this process works.

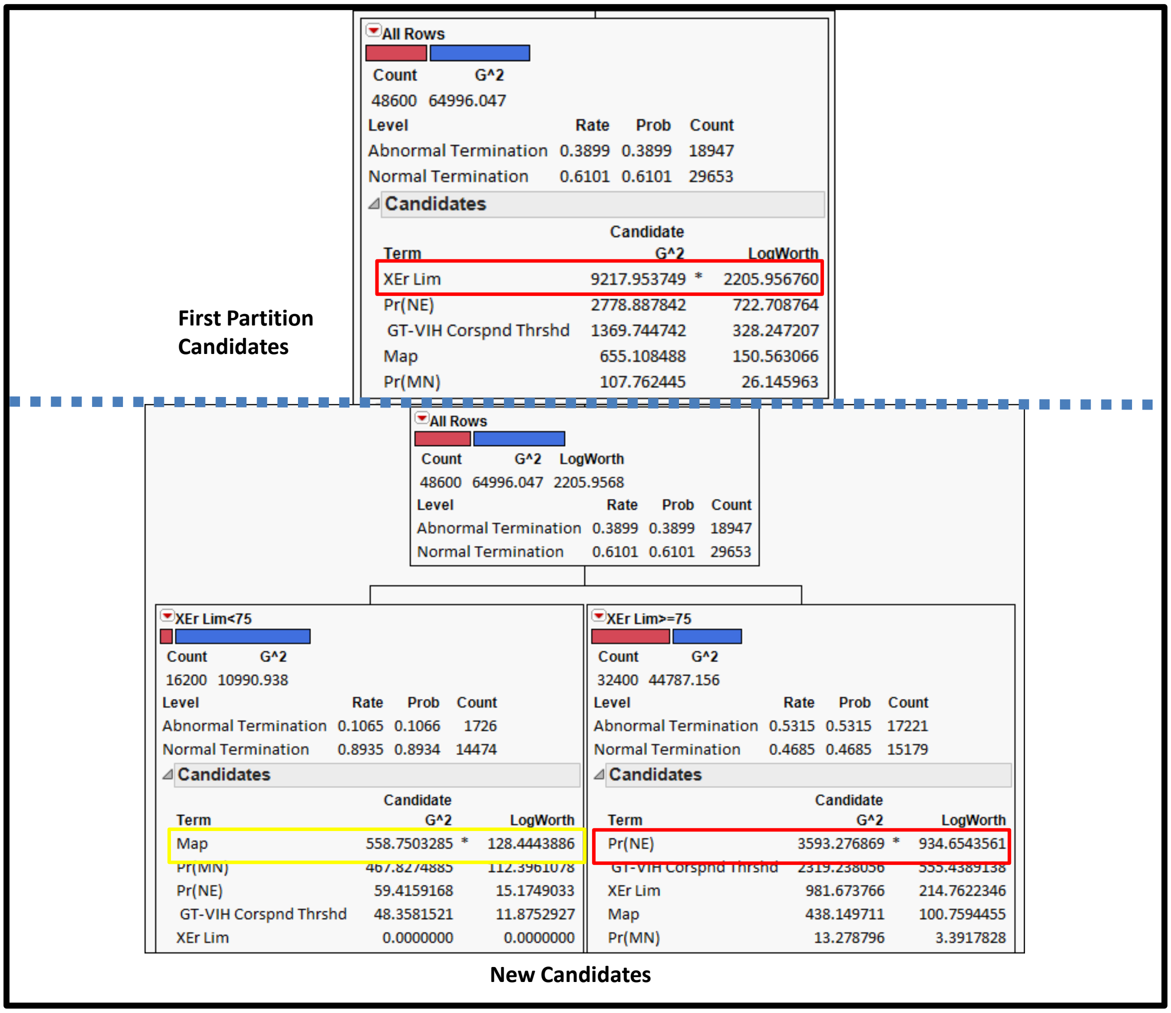


Figure 9-9 First Steps in the Regression Tree Partition Process

At step 1 G^2 values are calculated for all possible splits among the values of the eight variables under consideration. In this case the X and Y Error Limit produces the largest G^2 value. As recalled from Table 9-1, the possible X and Y Error Limit values are 25, 75, and 125, and JMP determines the best split partitions the data into two classes, class 1 where the Error Limit is <75, i.e., equal to 25, and class 2 where the error limit is =>75, i.e., = to either 75 or 125. The process is repeated for each of the two new classes. The best splits are found to be:

- Class 1 on Map - although not shown here, the split groups the Polygonal and

Multi-Factor Maps together and isolates the Square Map by itself; and

- Class 2 on Pr(NE) – which will end up differentiating between values less than 0.5, i.e., 0.1 and 0.3 and values greater than or equal to 0.5, i.e., 0.5 and 0.7.

Here the split of class 2 provides the best result for the next partition.

Figure 9-10 shows the regression tree after ten splits. As one can see, there is no requirement for symmetry in the partition tree; at any split any class which has not yet been partitioned is considered for the next partition.

Figure 9-11 provides a different graphical view of the partition tree after the same ten splits, while Table 9-6 provides a breakdown of the leaf structure of the tree. This table provides an in-depth view of the nature of the splits and the relative contributions of each parameter to the tree.

Table 9-6 Leaf Report and Leaf Contributions to the Regression Tree

Leaf Report

Response Prob

Leaf Label	Abnormal Termination	Normal Termination
XEr Lim<75&Map(Sqr Map)	0.0325	0.9675
XEr Lim<75&Map(MF Map, Poly Map)&Pr(MN)<0.15	0.0548	0.9452
XEr Lim<75&Map(MF Map, Poly Map)&Pr(MN)>=0.15	0.1881	0.8119
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)<0.3	0.1958	0.8042
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)>=0.3&XEr Lim<125	0.2667	0.7333
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)>=0.3&XEr Lim>=125	0.4885	0.5115
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.3	0.3652	0.6348
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.3	0.6895	0.3105
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd<0.9&XEr Lim<125	0.4450	0.5550
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125	0.7351	0.2649
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd>=0.9	0.9081	0.0919

Response Counts

Leaf Label	Abnormal Termination	Normal Termination
XEr Lim<75&Map(Sqr Map)	175	5225
XEr Lim<75&Map(MF Map, Poly Map)&Pr(MN)<0.15	197	3403
XEr Lim<75&Map(MF Map, Poly Map)&Pr(MN)>=0.15	1354	5846
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)<0.3	1057	4343
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)>=0.3&XEr Lim<125	720	1980
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd<0.9&Pr(NE)>=0.3&XEr Lim>=125	1319	1381
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.3	986	1714
XEr Lim>=75&Pr(NE)<0.5& GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.3	1862	838
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd<0.9&XEr Lim<125	2403	2997
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125	3970	1430
XEr Lim>=75&Pr(NE)>=0.5& GT-VIH Corspnd Thrshd>=0.9	4904	496

Column Contributions

Term	Number of Splits	G^2
Map	1	558.750
Pr(NE)	3	4616.469
XEr Lim	3	10460.014
Pr(MN)	1	399.031
GT-VIH Corspnd Thrshd	2	2849.825

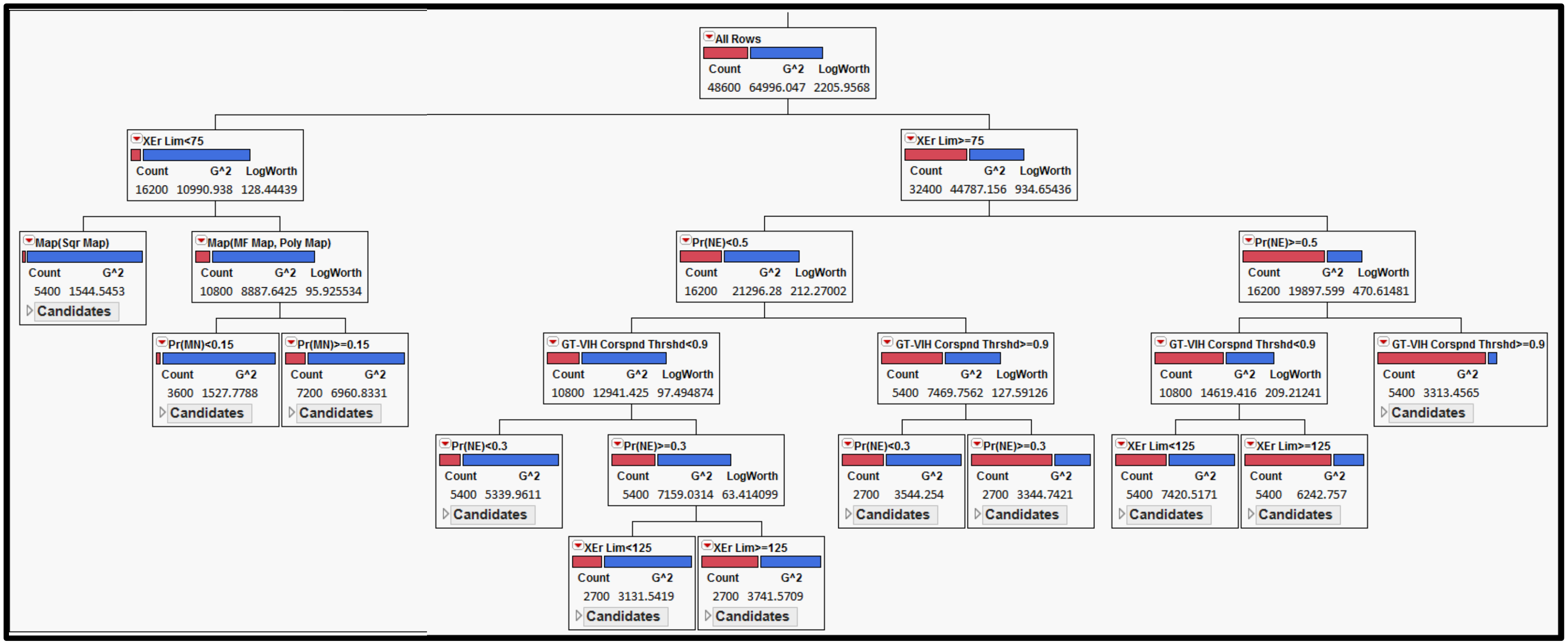


Figure 9-10 Regression Tree after Ten Splits

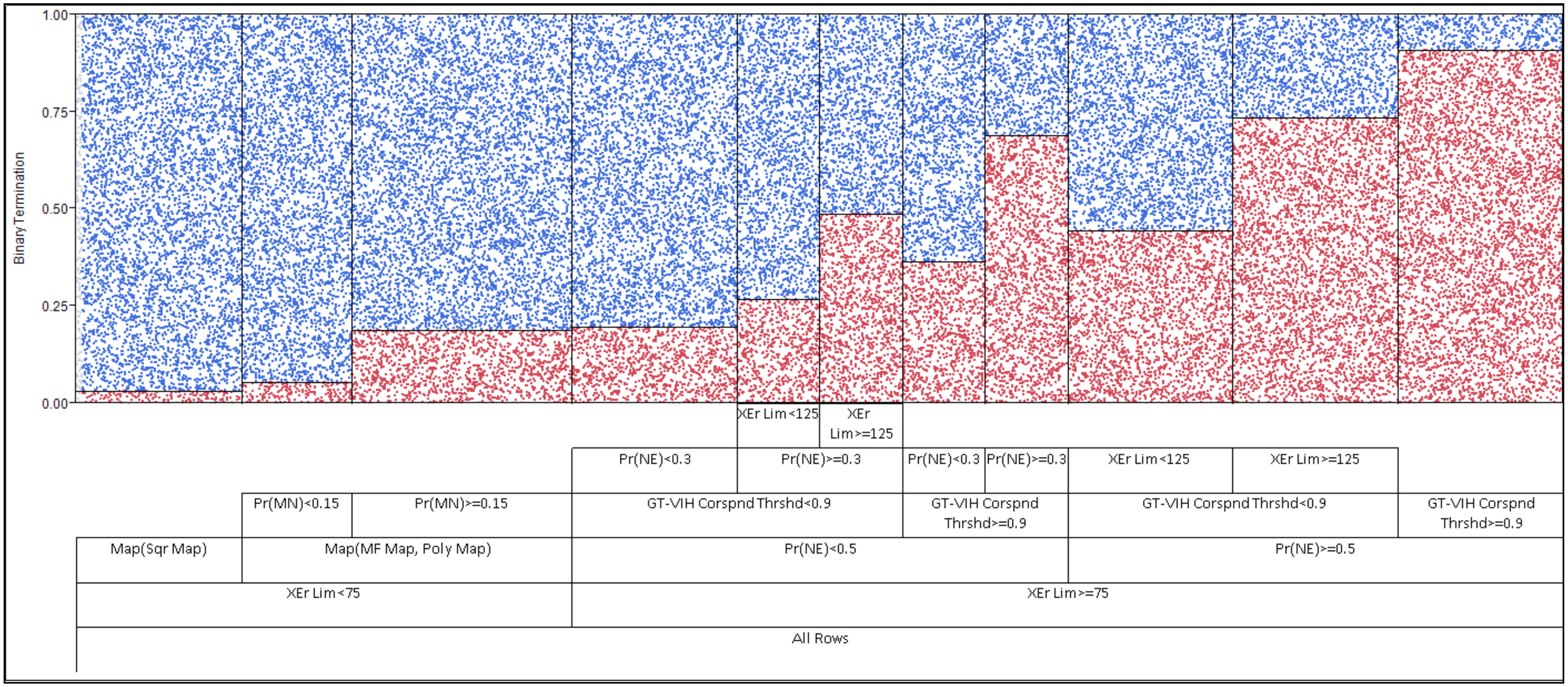


Figure 9-11 Regression Tree Partition Graph (Ten Splits)

Both views of the tree further show the interaction between the GT-VIH correspondence threshold and the spatial distortion parameters. As one would expect from the previous analysis, correspondence threshold splits first pair 0.6 and 0.75 in one group and 0.9 in another.

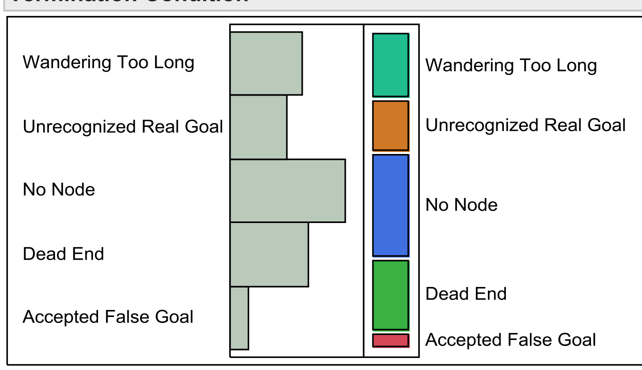
It is easy to see the effects of the partitioning process in Figure 9-11, where each successive partition reflects the degree to which the results differ from class to class. The width of the columns in the figure also highlight the relative proportion of the experimental results covered in each partition class.

9.3 Refinement of Abnormal Termination Results

The previous chapters looked at a binary breakdown of simulation experiment outcomes as either normal termination, i.e., the entity successfully arrived at its goal node, or abnormal termination, i.e., the entity failed to reach its goal. As discussed above in Chapter 8.5.2, there are a total of five different ways the entity can fail to reach its goal: dead end; lost no; lost no goal; failure to recognize a true goal; or accepts a false goal. Figure 9-12 shows the relative occurrences of each of these categories for the 18947 trials (39.0 of the total) that resulted in abnormal termination.

Distributions

Termination Condition



Frequencies

Level	Count	Prob
Accepted False Goal	1055	0.05568
Dead End	4357	0.22996
No Node	6388	0.33715
Unrecognized Real Go	3155	0.16652
Wandering Too Long	3992	0.21069
Total	18947	1.00000

N Missin 29653

5 Level

Figure 9-12 Frequencies of Abnormal Termination

A prime rationale for the investigation of these results is to seek out the conditions that lead to failure in the hope that such research will help define strategies to eliminate or minimize any or all of the types of failure. To further that end, the simulation model provides output data to support examination of the details of entity behavior with respect to any of these failure categories. For example, the most prevalent category, Lost No Node, indicates that the entity found itself to traveling on an arc and failing to find a node where it expected one. This condition occurs when:

- the spatial distortion moves a node so the VIH map considers that node to be much closer to the arc link's terminal node than it is in ground truth;
- the entity has failed to recognize a VIH node corresponding to its actual ground truth position, and takes an arc link, while not knowing the distance to that arc link's terminal node;
- the entity knows where it is on both the GT Map and the VIH, but spatial and/or topological distortion in the VIH map cause it to makes an incorrect arc link choice, i.e., failing to recognize the ground truth arc link that actually corresponds to its desired VIH path choice.

The entity's physical movement element, the GT agent moves in response to the directions provide it by the entity's cognitive component, the VIH agent. If the GT agent is at a GT node, these directions consist of specifying the bearing of the GT arc link chosen by the VIH agent, and a distance to travel along that arc. Two input parameters control the distance value given.

The first of these values is a multiplier applied to the VIH distance to the arc link's expected terminal node, if this node is known. This multiplier establishes a stopping threshold as a fraction of the expected VIH distance the VIH agent is willing to exceed, the idea being that the greater the expected distance, the more tolerant the agent will be of error. For the current experiment this value is set to 1.2, indicating the agent has a 20% error tolerance.

The second value is specified by the Keep Searching Distance parameter and is at present a constant value set to 75. Both parameter values were chosen after exploration with some preliminary experiments, and obviously can be further explored through other parametric experiments.

10. Future Work Movement Constraint Thresholds and Risk Tolerance:

These two distance parameters are accompanied by other movement constraint thresholds whose effects can be studied in conjunction with varying the effects of the entity's risk tolerance. Thorough exploration of such factors requires model modifications to introduce uncertainty factors in node and arc link recognition factors and other characteristics. The simulation model is designed to accommodate such features, but new code and additional input data features are required to implement them.

The GT agent travels along the chosen arc link until it either encounters a GT node or exceeds the distance set by the VIH agent. In either instance, it stops and requests further direction from the VIH agent, following the logic shown in Figure 8-10 above. If the GT agent is at a GT node, the VIH agent executes the code in block 7 of that figure, which requires a node recognition decision by the VIH agent. If the GT agent is stopped along an arc the VIH agent enters the state “notAtGroundTruthNodeLogic”, which requires an intra-Path decision. The simulation model tracks the number of times each of these decision states occurs. Node recognition decisions have a number of possible outcomes, as detailed in Chapter 8.7.6, and the frequency of each of these outcomes is provided as output. At present, the intra-path decision can have only two outcomes: if it has exceeded its distance thresholds declare a Lost No Node condition and quit, or else keep going until does exceeds those thresholds or encounters a GT node, whichever comes first.

As would be expected, there is a significant correlation between the number of intra-path decisions and abnormal termination in general (shown in Figure 9-13), and in the Lost No Node Category in particular (shown in Figure 9-14).

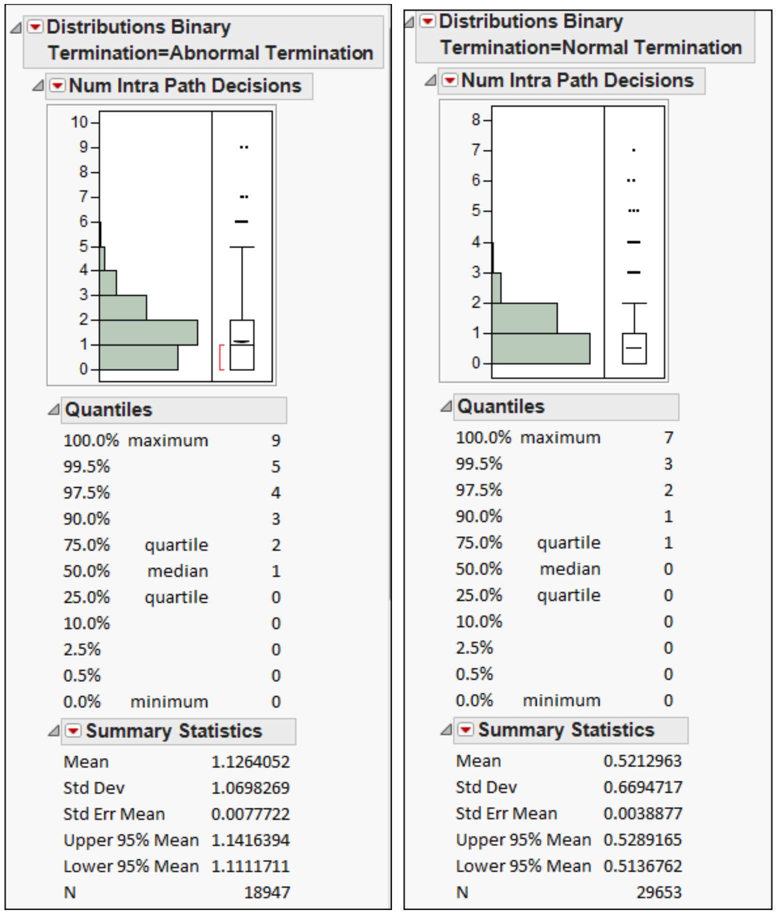


Figure 9-13 Frequency of Intra-Path Decisions by Normal and Abnormal Termination Conditions

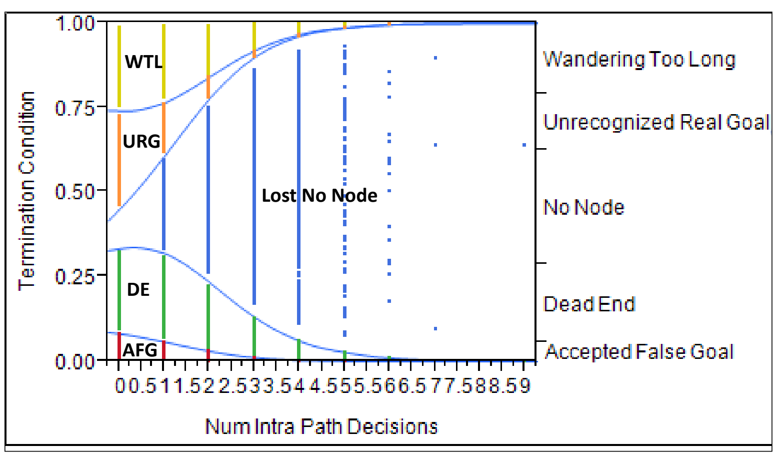


Figure 9-14 Logistic Fit of Abnormal Termination Categories by number of Intra-Path Decisions

The JMP Nominal Logistic Fit Model of Abnormal Termination Categories with respect to the parameters of the simulation DOE, again reveals high significance with respect to the whole model test (Table 9-7) and most of the two-way interactions (Table 9-9). Again, however, the lack of fit test, Table 9-9, shows there are significant effects not accounted for in the logistic fit model.

Table 9-7 Abnormal Termination Categories JMP Nominal Logistic Fit

Whole Model Test Nominal Logistic Fit for Abnormal Termination Converged in Gradient, 7 iterations				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	1963.546	80	3927.093	<.0001
Full	26305.569			
Reduced	28269.115			
RSquare (U)	0.0695			
AICc	52779.9			
Misclassification Rate	0.632			
Observations	18947			

Table 9-8 Abnormal Termination Categories Effect Likelihood Ratio Tests

Effect Likelihood Ratio Tests				
Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Map	8	8	652.321	<.0001
Pr(NE)	4	4	25.700	<.0001
XEr Lim	4	4	410.606	<.0001
Pr(MN)	4	4	136.758	<.0001
GT-VIH Corspnd Thrshd	4	4	1303.300	<.0001
Map*Pr(NE)	8	8	43.762	<.0001
Map*XEr Lim	8	8	42.964	<.0001
Map*Pr(MN)	8	8	8.152	0.4187
Map* GT-VIH Corspnd Thrshd	8	8	133.318	<.0001
Pr(NE)*XEr Lim	4	4	40.370	<.0001
Pr(NE)*Pr(MN)	4	4	6.026	0.1972
Pr(NE)* GT-VIH Corspnd Thrshd	4	4	342.320	<.0001
XEr Lim*Pr(MN)	4	4	33.937	<.0001
XEr Lim* GT-VIH Corspnd Thrshd	4	4	550.605	<.0001
Pr(MN)* GT-VIH Corspnd Thrshd	4	4	38.468	<.0001

Table 9-9 Abnormal Termination Categories Fit Lack of Fit

Lack Of Fit				
Source	DF	-LogLikelihood	ChiSquare	Prob>ChiSq
Lack Of Fit	1168	1008.269	2016.538	<.0001
Saturated	1248	25297.3		
Fitted	80	26305.569		

Figure 9-15 Abnormal Termination Categories JMP Logistic Fit Lack of Fit

Further evidence of this lack of fit is given by the confusion matrix, Table 9-10 as one sees the model vastly over predicts the most frequently occurring category (Lost No Node) and under predicts everything else.

Table 9-10 Abnormal Termination Categories Logistic Fit Confusion Matrix

Confusion Matrix						
Actual/Predicted	Accepted False Goal	Dead End	No Node	Unrecognized Real Goal	Wandering Too Long	Row Sums (Actual Observed)
Accepted False Goal	83	127	681	116	48	1055
Dead End	16	945	2521	203	672	4357
No Node	79	789	4646	214	660	6388
Unrecognized Real Goal	31	424	1687	284	729	3155
Wandering Too Long	10	620	2155	193	1014	3992
Column Totals (Predicted)	219	2905	11690	1010	3123	18947

For more insight, one can again turn to a partition tree graph, as shown in Figure 9-16. The most dominant effect seen in this tree is the split on the GT-VIH Correspondence Threshold. When this threshold is equal to 0.9, there are virtually no false positives in recognizing the simulation goal, while they are observed when this threshold is relaxed. More false positives are also seen of the Square Grid Map, which is not surprising as the map's regularity increases the similarity between nodes.

The relative contributions of each combination of parameter values used to determine these first splits is shown in Figure 9-17, which shows the partition tree leaf report.

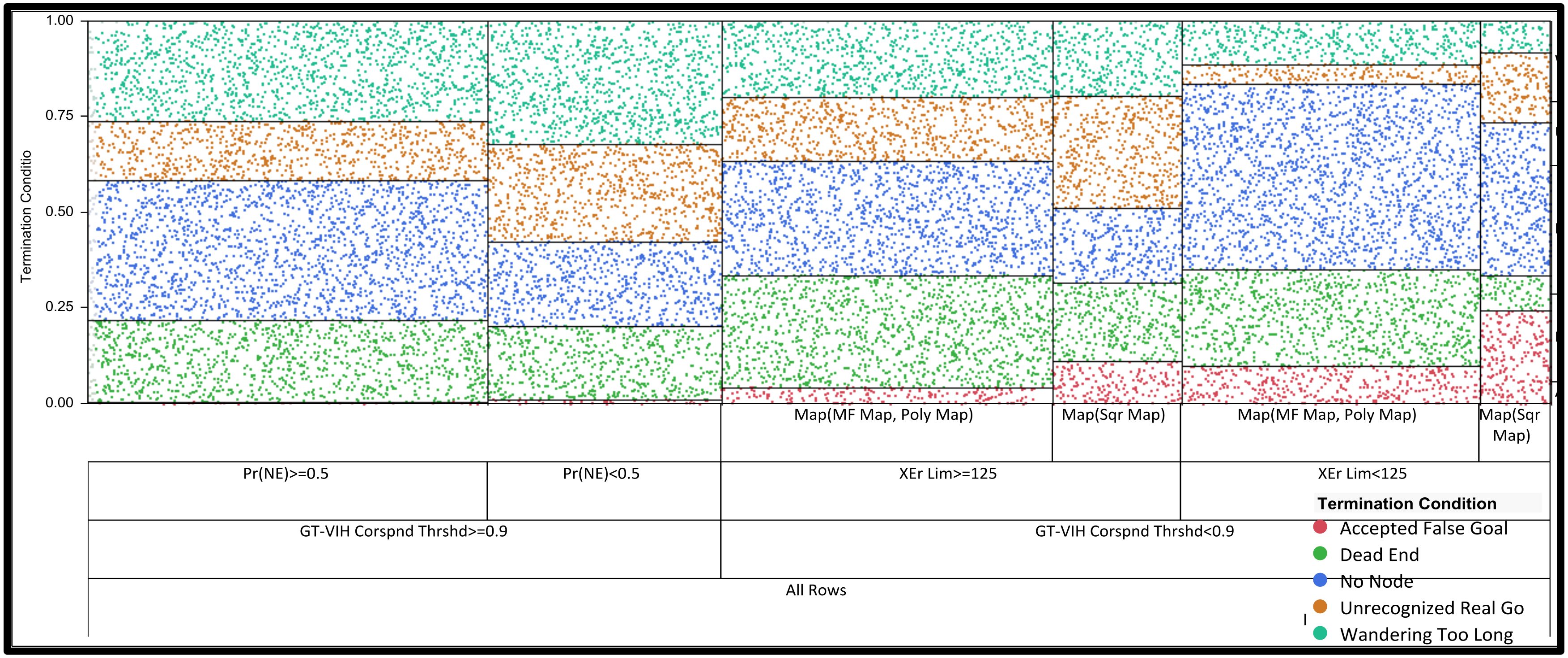


Figure 9-16 Abnormal Termination Categories JMP Logistic Fit Partition Tree

Leaf Report

Response Pro

Leaf Label	Accepted False Goal	Dead End	No Node	Unrecognized Real Goal	Wandering Too Long
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim>=75&Map(MF Map, Poly Map)	0.0018	0.2617	0.3575	0.1351	0.2439
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim>=75&Map(Sqr Map)	0.0050	0.1161	0.4263	0.1812	0.2714
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim<75	0.0480	0.2472	0.0527	0.2616	0.3905
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.5&Map(MF Map)	0.0010	0.2692	0.1750	0.2261	0.3287
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.5&Map(Sqr Map, Poly Map)	0.0186	0.1479	0.2487	0.2693	0.3154
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(MF Map, Poly Map)& GT-VIH Corspnd Thrshd>=0.7	0.0272	0.2912	0.2617	0.1712	0.2487
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(MF Map, Poly Map)& GT-VIH Corspnd Thrshd<0.75	0.0621	0.2952	0.3333	0.1655	0.1439
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(Sqr Map)	0.1125	0.2052	0.1962	0.2911	0.1950
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(MF Map, Poly Map)&XEr Lim>=75	0.0799	0.2680	0.4637	0.0737	0.1148
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(MF Map, Poly Map)&XEr Lim<75	0.1500	0.2121	0.5317	0.0010	0.1052
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(Sqr Map)	0.2447	0.0922	0.3986	0.1831	0.0814

Response Count

Leaf Label	Accepted False Goal	Dead End	No Node	Unrecognized Real Goal	Wandering Too Long
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim>=75&Map(MF Map, Poly Map)	6	866	1183	447	807
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim>=75&Map(Sqr Map)	8	185	680	289	433
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)>=0.5&XEr Lim<75	13	67	14	71	106
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.5&Map(MF Map)	1	294	191	247	359
GT-VIH Corspnd Thrshd>=0.9&Pr(NE)<0.5&Map(Sqr Map, Poly Map)	36	286	481	521	610
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(MF Map, Poly Map)& GT-VIH Corspnd Thrshd>=0.7	59	631	567	371	539
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(MF Map, Poly Map)& GT-VIH Corspnd Thrshd<0.75	132	628	709	352	306
GT-VIH Corspnd Thrshd<0.9&XEr Lim>=125&Map(Sqr Map)	187	341	326	484	324
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(MF Map, Poly Map)&XEr Lim>=75	220	738	1277	203	316
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(MF Map, Poly Map)&XEr Lim<75	167	236	592	1	117
GT-VIH Corspnd Thrshd<0.9&XEr Lim<125&Map(Sqr Map)	226	85	368	169	75

Column Contributions

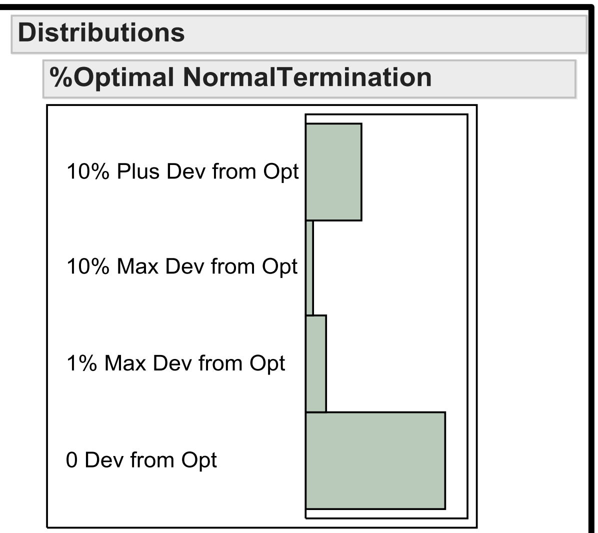
Term	Number of Splits	G^2
Map	4	856.9516
Pr(NE)	1	260.9070
XEr Lim	3	1252.8080
Pr(MN)	0	0.0000
GT-VIH Corspnd Thrshd	2	1325.0206

Figure 9-17 Abnormal Termination Categories Partition Tree Leaf Report & Column Contributions

9.4 Assessment of Normal Termination

While the simulation model does not explicitly determine separate categories for normal termination, it does provide sufficient output detail to assess different degrees of successful outcomes. The most obvious of these distinctions is the comparison of the actual distance traveled compared to the optimum distance, as measured by the shortest possible ground truth path. Figure 9-18 divides normal termination into categories:

- Zero deviation from optimum ;
- 0 < deviation from optimum <= 0.01;
- 0.01 < deviation from optimum <= 0.1; and
- 0.10 < deviation from optimum.



Frequencies

Level	Count	Prob
0 Dev from Opt	18433	0.62162
1% Max Dev from Opt	2788	0.09402
10% Max Dev from Op	1092	0.03683
10% Plus Dev from Op	7340	0.24753
Total	29653	1.00000

N Missin 18947
4 Level

Figure 9-18 Frequency of Normal Termination Optimality Categories

The figure shows the relative occurrences of each of these categories for the 29653 trials (61% of the total) that resulted in normal termination. The whole model fit test for a JMP nominal logistic fit again is statistically significant at the 99.9% confidence level, but with a misclassification rate of 0.371, the model essentially predicts that all of the outcomes will fall in the zero deviation from optimum category. The two-way factor effects likelihood tests are similarly all statistically significant with only one exception.

The most obvious rationale for this outcome is simply that frequently the GT path and the VIH path consist of the same nodes, and as long as the entity stays on its path, it will travel the ground truth distance. Some evidence to support this hypothesis is given by examination of the model output for how often the entity strayed from its intended path. For each trial the model output gives the number of times the VIH Agent recorded it was at a node not on its path. Figure 9-19 shows this count plotted against the normal outcome alternatives.

While the vast majority of the time true optimality coincides with zero nodes off path, there is not a strict one-to-one relationship between these two model outcomes. An interesting element in Figure 9-19 is the number of times the entity was never off its path and still failed to achieve its optimal outcome, as well as the number of times it left its path yet did meet the optimum.

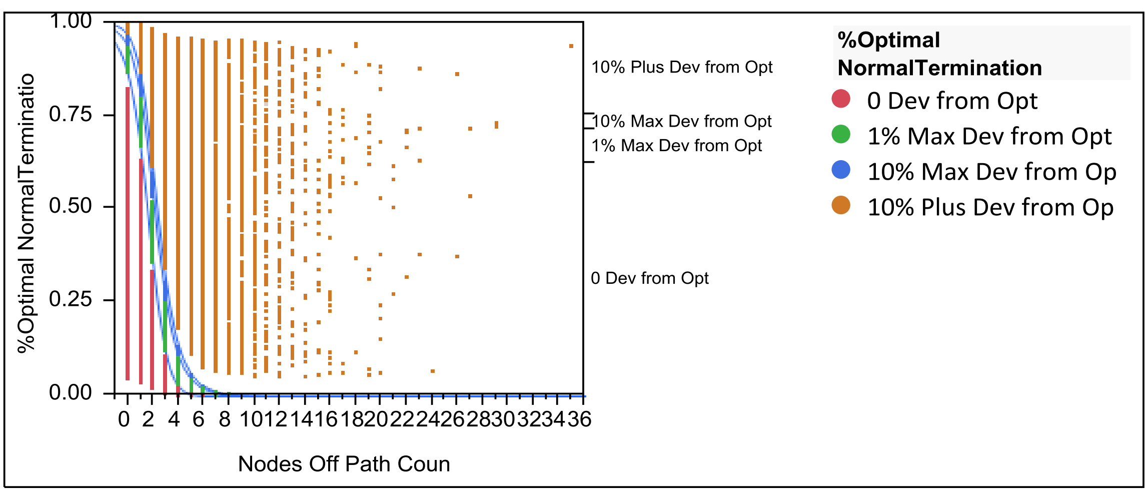


Figure 9-19 Number of Nodes of Path Count for Normal Termination Outcomes.

The first explanation for both of these phenomena lies in the fact that the Off Path Count refers to the entity's VIH path, while the achievement of path optimality is calculated by comparing the true distance travelled to the true optimum, length of the ground truth shortest distance path.

Figure 6-20 shows the results from Figure 9-19 broken down into those cases where the GT-VIH Optimums are the same and those where they differ. As expected, if the two paths distances are the same, meaning they are most probably the same path, then a zero Off Path Count ensures an optimal termination result.

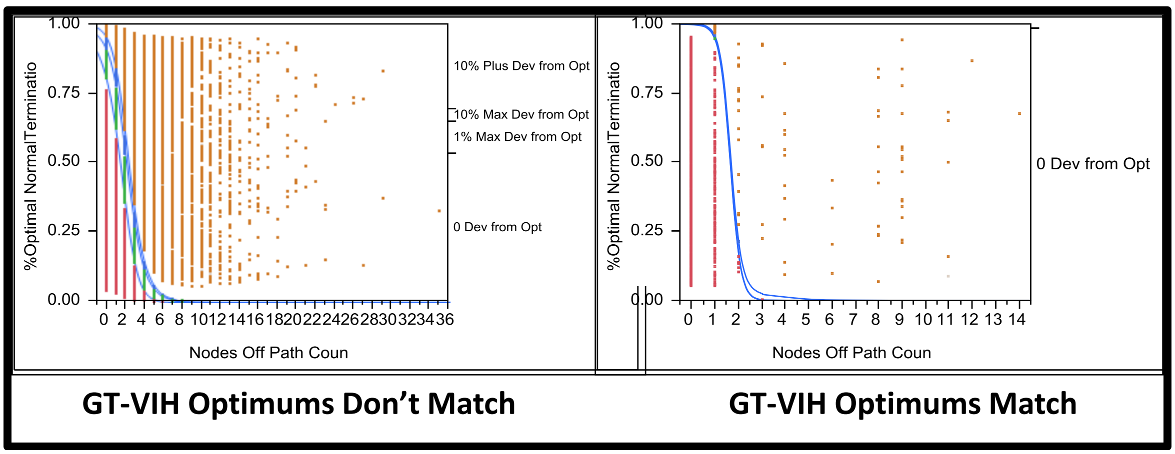


Figure 9-20 Comparison of Node Off Path Counts by GT-VIH Path Match Criterion

It is still possible in both cases to stray from the path and still have an optimal result. During interactive trials of the simulation model two circumstances where this condition can occur have been observed. In the first the VIH and GT paths don't match, but the entity mistakenly (but fortuitously) moves along the actual ground truth shortest path. In the second, the entity happens on a portion of the arc/node network where regularity provides one or more equidistant path segments, each leading to the goal.

The dominance of the zero deviation from optimum category can be seen again in Figure 9-21, which shows a partition tree for the normal outcomes as affected by the various input parameters. Figure 9-22 provides a small tree view for explanation of the partitions that cannot be read in the previous figure. One has to split the regression tree more than 20 times to get columns in which that category is not the most likely.

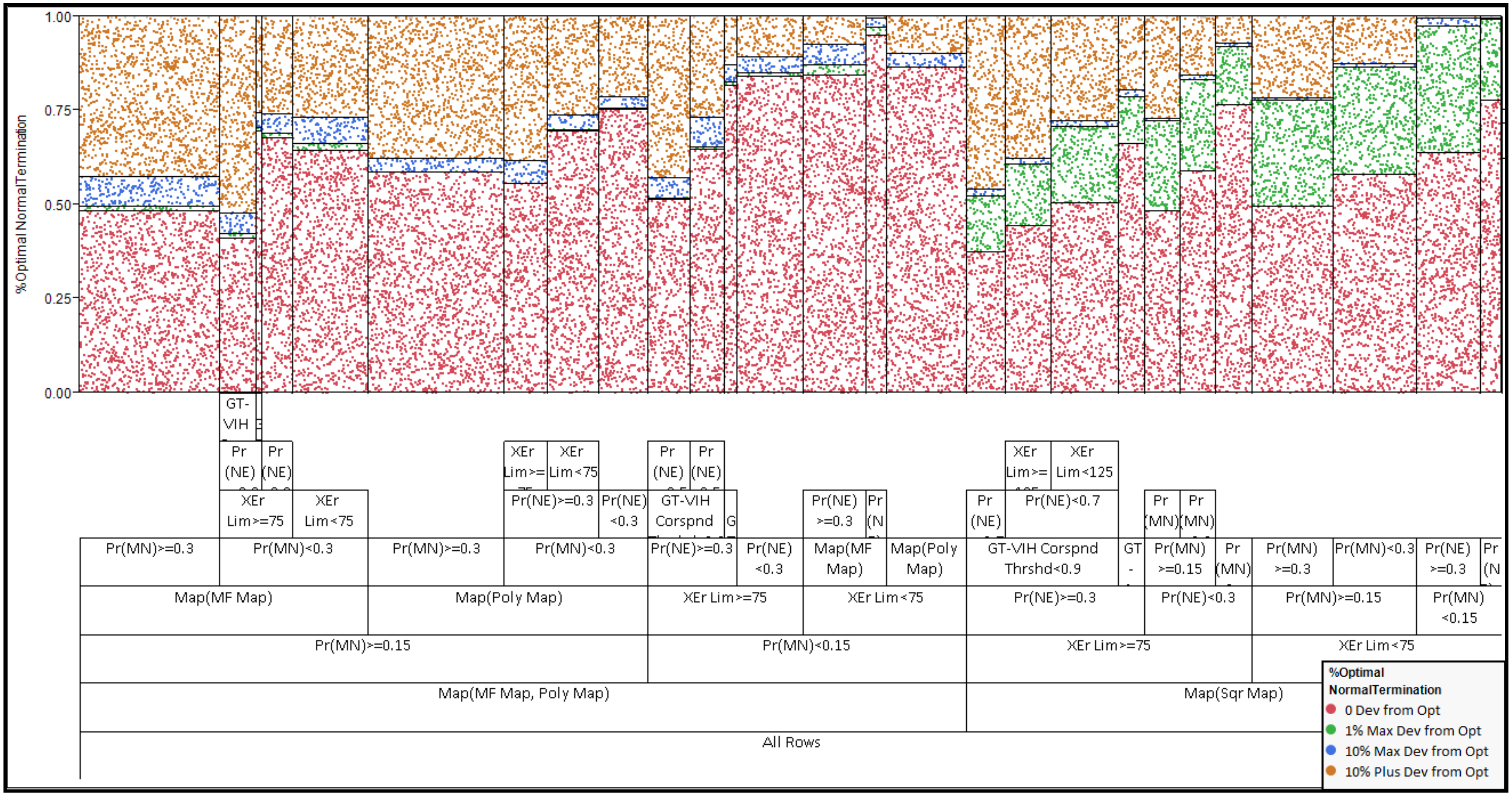


Figure 9-21 Graph of Partition for Termination Optimums after 25 Splits

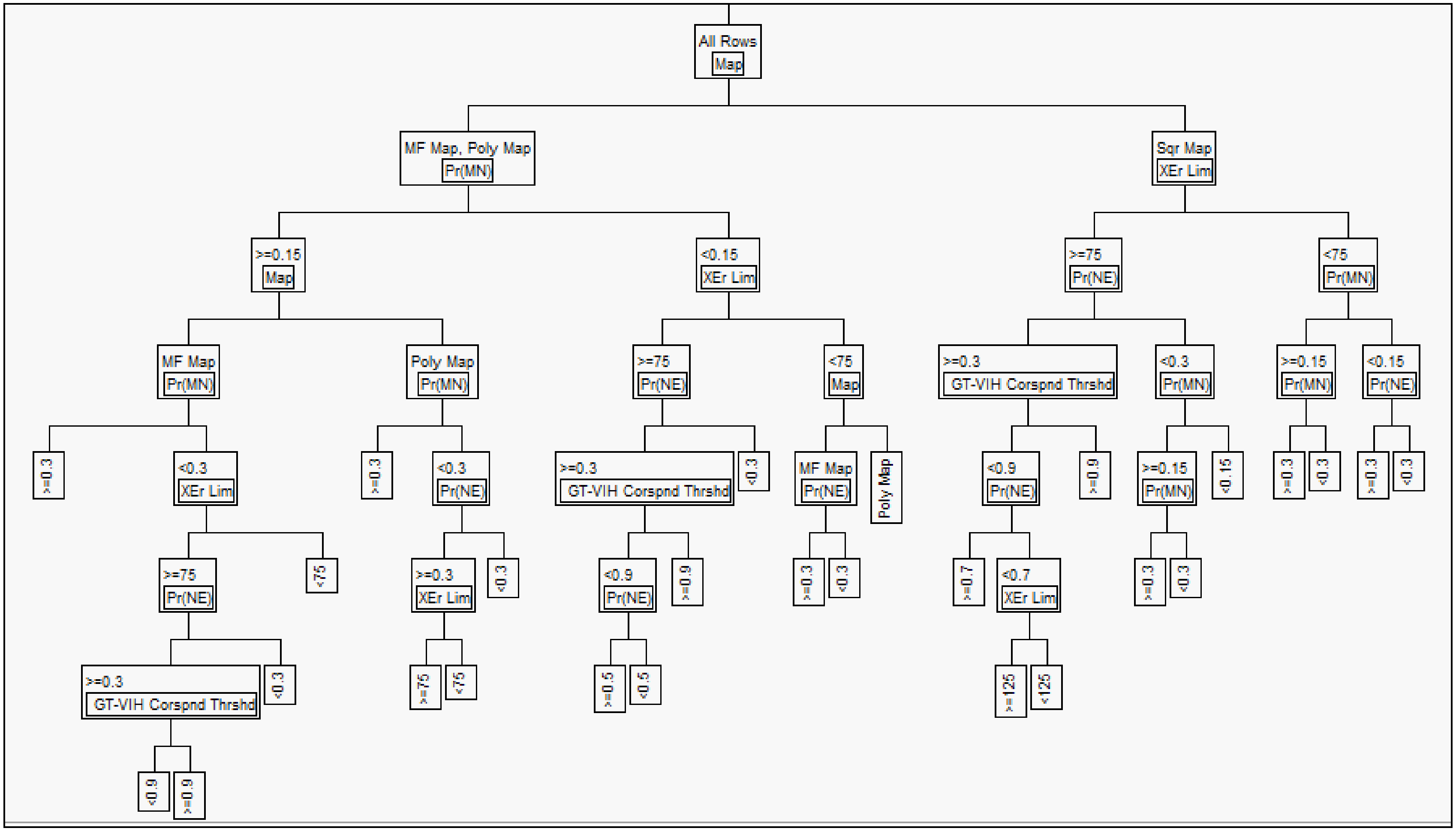
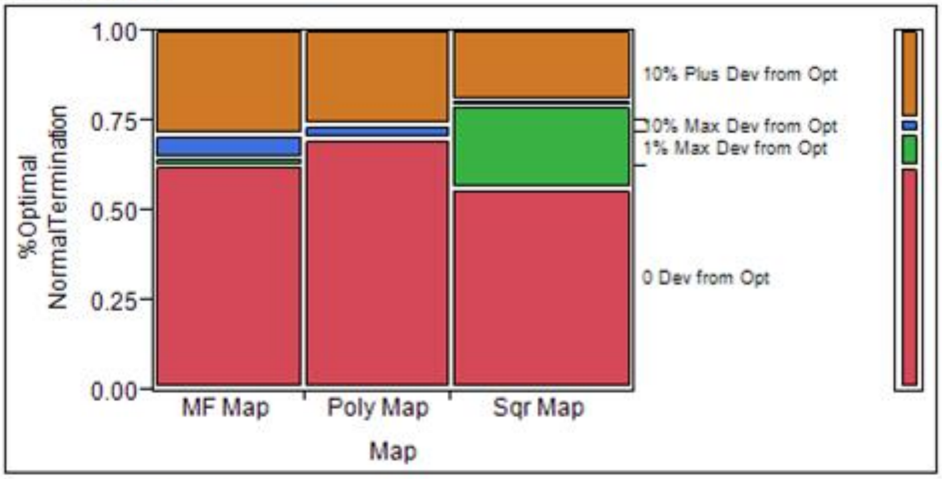


Figure 9-22 Small Tree View of Partition for Termination Optimums

A final insight into the interaction between the entity and its environment can be gleaned by looking at a contingency analysis of the normal termination alternatives explored in this chapter, as is shown in Table 9-11.

Table 9-11 Contingency Analysis of Normal Termination Alternatives

Mosaic Plot



Contingency Table

%Optimal NormalTermination

Map	Count Total % Col % Row %	0 Dev from Opt	1% Max Dev from Opt	10% Max Dev from	10% Plus Dev from	
	MF Map	5865 19.78 31.82 62.69	150 0.51 5.38 1.60	606 2.04 55.49 6.48	2734 9.22 37.25 29.23	9355 31.55
	Poly Map	6339 21.38 34.39 69.45	10 0.03 0.36 0.11	350 1.18 32.05 3.83	2428 8.19 33.08 26.60	9127 30.78
	Sqr Map	6229 21.01 33.79 55.76	2628 8.86 94.26 23.53	136 0.46 12.45 1.22	2178 7.34 29.67 19.50	11171 37.67
		18433 62.16	2788 9.40	1092 3.68	7340 24.75	29653

As before, there is a significant difference between the Square Map and the other two maps, but here one can also notice that the Square Map is much more prone to moderate deviations from the optimal route. In other words, the Square Map appears to be more "forgiving" of entities that stray somewhat from there chosen path. A similar breakdown by Map of the Off Path Node Analysis of Figure 9-19 is shown in Figure 9-23 also shows this moderate tolerance for leaving the VIH planned route.

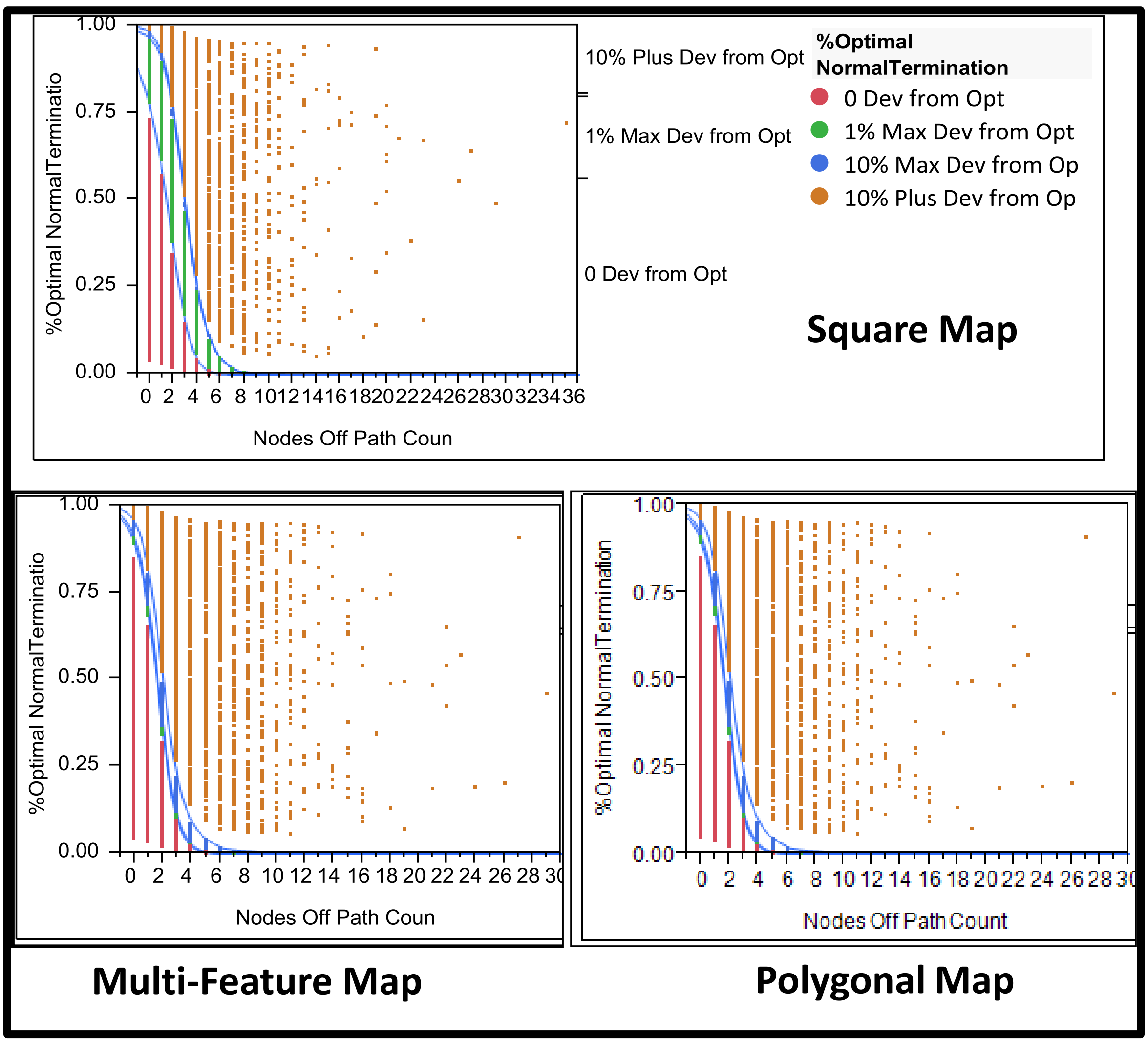


Figure 9-23 Breakdown by Map of Off-Route Node Counts vs Normal Terminations

9.5 Measuring Progress Towards the Goal

The final question for this set of experiments is whether they can contribute to modification/construction of entity strategies based on measures of progress towards the goal. The previous chapters looked at the ways in which an entity could fail and ways in which the "goodness" of its success could be measured. The question now is whether indications of potential failure or sub-optimal performance are available to the entity during the simulation. Most of the statistics examined in the previous chapters are not part of the VIH agent's knowledge; all it "knows" it what it surmises ground truth to be.

It may be possible to develop rule sets, fuzzy cognitive maps, and other inference schemes so that over time an entity can build up a knowledge base supporting modification of its mental map. At present, however, the VIH agent is limited to tracking its position relative to its starting point and updating the map characteristics (primarily position) of the nodes it recognizes. The entity begins its movement without even any idea of the extent to which its map may be distorted, so any of the previously observed correlations between distortion parameters and probability of success are not useful to the VIH agent.

The only measures currently available to that VIH agent are based tracking the decisions it makes and on the degree of confidence it has in those decisions. At present confidence values are only calculated for node recognition decisions, but similar VFT hierarchies could be set up for selection of inter-node links and making intra-link decisions as to whether to keep going on a currently selected link.

The model tracks a running average of entity confidence over a user-selected number of previous decisions. This value is currently available to model user during interactive simulation experiments as shown in Figure 9-24. The graph in the upper left corner of the figure tracks a number of statistics on node confidence including the value of the current best fit, the average of the best fits, and the threshold value for acceptance or rejection of the best fit value for the current node.

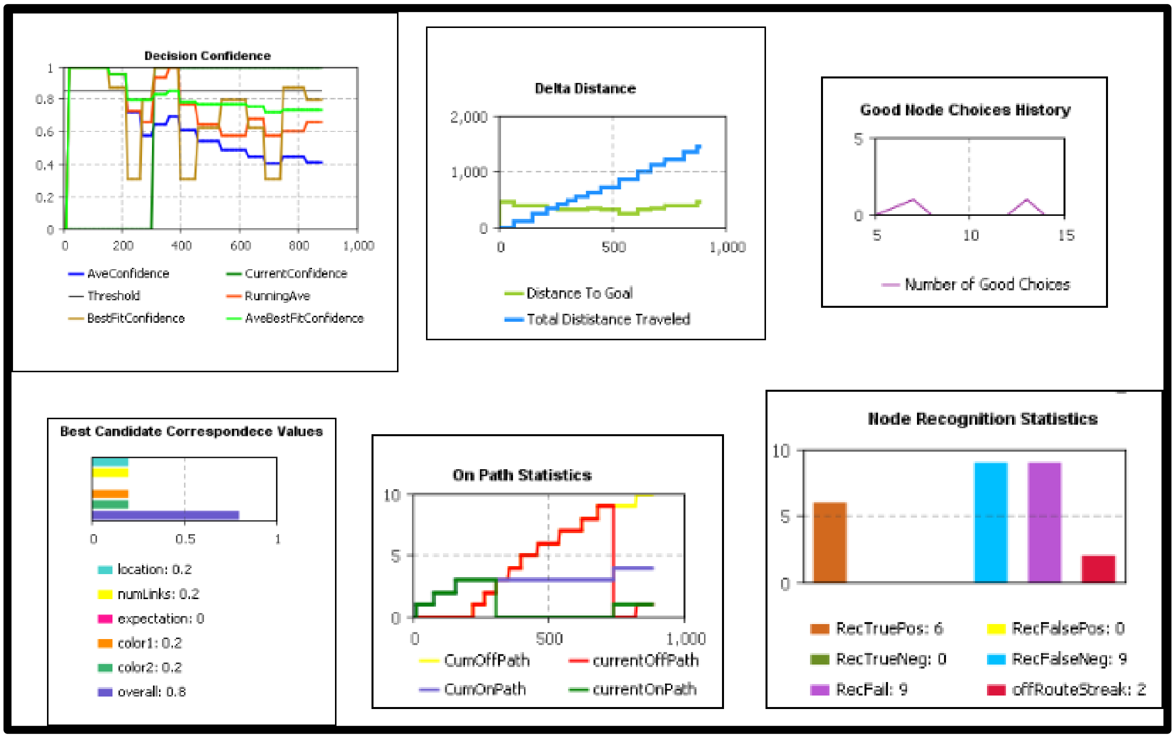


Figure 9-24 Interactive Model Data Outputs

Beneath that graph is one showing the node recognition values associated with the current best candidate. The right side of the figure shows two graphs, one of which provides information to the user on the ground truth nature of the decisions made by the VIH agent, and therefor information that is not accessible to the VIH agent.

The other graph on the right does, however, provide information that might be useful to the VIH agent. Each time the VIH agent makes a node recognition decision, the model checks to see if other node candidates have a GT-VIH correspondence level within a user-selected delta of the best candidate node. In effect, the agent is tracking the number of "good" alternatives it has to choose from, the idea being that the potential for error increases when the agent must select between nearly equal candidates. While not used by the agent at present, these kinds of decision history data could be employed by a "lost" agent to determine when and where it might have made an error.

11. Future Work: “Learning” from Past Mistakes - Application of Multi-Criteria Decision Methodologies to Support Adaptive Entity Behaviors

It is important that any agent-based entity used to represent human behavior displays the ability to “learn” from its environment – to adapt its cognitive behaviors as well as its physical ones to changes in that environment. The VFT approach taken herein can support such adaptation through dynamic SDVFs and VFT hierarchy match weights. Giving the agent the capability to asses the quality of its previous decisions, or any other behaviors, suggests the capability to possibly correct mistakes, e.g., backtracking to the node where it made a “wrong turn” by selecting an inappropriate link, or simply calculating a new “best” path. The potential exists, however, to go beyond simple error correction and use evaluation of historical data modify its behaviors. In this case such modification would consist of updating its mental map, and/or altering elements in its VFT decision process – the SDVFs and/or weight coefficients that define decision outcomes.

Another graph provides on/off path statistics, which as noted earlier on based on the VIH agent’s perception of whether it is on the VIH map, and hence is information available to the agent in making strategy decisions. Unfortunately, as also discussed earlier, this information may or may not correlate to whether the agent is making optimal progress to its goal.

The same is true for average node recognition confidence, while it correlates highly with success or failure, as shown in Figure 9-25, it doesn’t guarantee an outcome.

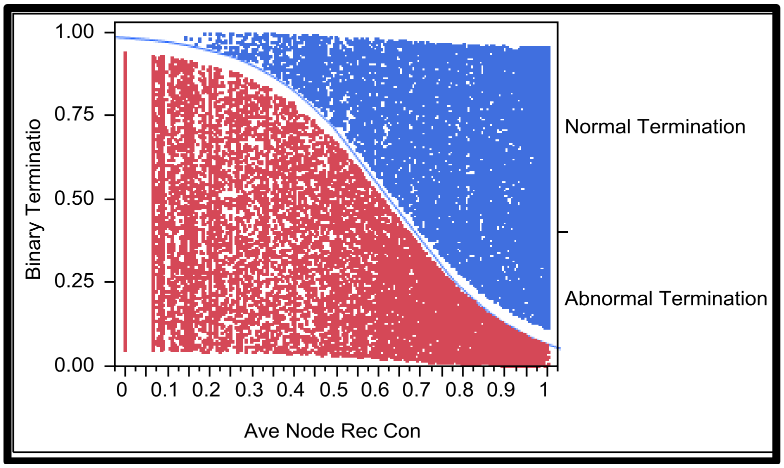


Figure 9-25 Logistic Fit Binary Termination by Average Recognition Confidence

The potential for information available to the VIH agent to assist in evaluating its chances for success is shown by looking at a logistic fit to binary termination outcomes. The values used are:

- Average Node Recognition Confidence;
- Number of Off Path Nodes Observed;
- Number of Intra-Path Decisions Made;
- GT-VIH Correspondence Threshold; and
- Map (under the belief that the agent can define some measure of map complexity or difficulty.

The results suggest strongly that these values can support prediction of success the resultant model has:

- a misclassification rate of 0.1345;
- an AUC for the ROC curve equal to 0.92;
- a lack of fit probability > chi-square equal to 1.0; and
- high significance for all of the factors and two-way factor interactions.

A note of caution is in order; the values used to fit this model were obtained only after the simulation ended, so the extent to which they can be correlated against incremental progress must yet be gauged.

12. Future Work: Assessment of Incremental progress

Further experiments are needed to explore how dynamic the incremental progress measures suggested above are:

- *Do they reflect slow, steady declines in overall performance, or do they define abrupt discontinuities in the probability of success, and how robust are any strategy-changing thresholds?*
- *What are the levels of interaction between the discrete and continuous measures?*
- *Do they help distinguish between high-risk high-reward and more cautious strategies?*

Fit model techniques can suggest when progress may or may not be at acceptable levels. For example, a regression tree partition analysis on the potential interim progress metrics listed above is shown in Figure 9-26, Figure 9-27, and Figure 9-28. Figure 9-26 shows the clear delineation the partition splits make with respect to probabilities of normal termination or entity "success". Figure 9-27 details the quantitative values for each split and Figure 9-28 isolates on the middle portion of the graph to show how these split sequences can be easily translated into if-then-else rules with respect to the probability of success. The VIH agent could use such rules to determine if and when it might re-evaluate and/or change its goal-seeking strategy.

Continuous variable data such as that provide by average node recognition confidence seem to be well adapted to the definition of fuzzy set membership criteria, which could then be applied to fuzzy if-then-else rules or/and fuzzy cognitive maps. Whichever inference scheme one chooses to apply, the result could be matched to an appropriate strategy. For example, the entity could abandon a fruitless path-following behavior and search for a local intermediate goal such as a landmark.

Given a number of course of action to choose from, one could also apply the VFT methodology to select the best candidate. In this case, the measures of intermediate progress would serve as SDVFs for the VFT hierarchy.

While at present the simulation model has only a limited number of such strategies, these experimental results do suggest the methodology employed herein. It can be applied to evaluate current strategies, to seek to adapt them to new circumstances, or assist in the development and employment of new strategies all together.

13. Future Work: Further Investigation into Incremental Measures of Progress

There are certainly other measures that could be calculated during the simulation as measures of incremental progress. They include: comparison of distance travelled against expected path distance, indications of the extent and degree of distortion in expected path attributes; and evidence of systemic distortion in expected map features. Quantifying these elements would allow for their use by the agent in selecting or changing its movement strategy.

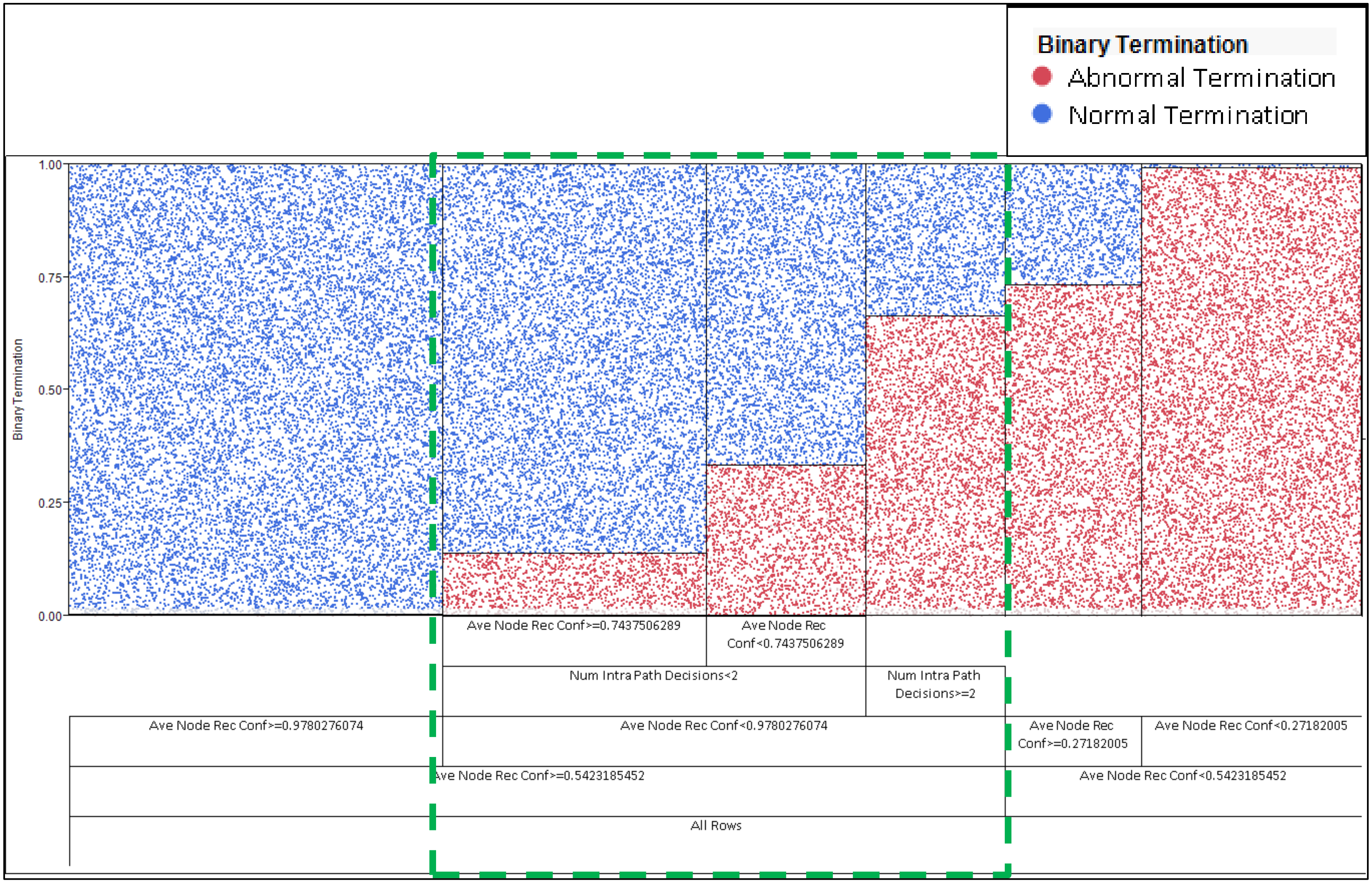


Figure 9-26 Partition Tree Graph on Potential Progress Metrics

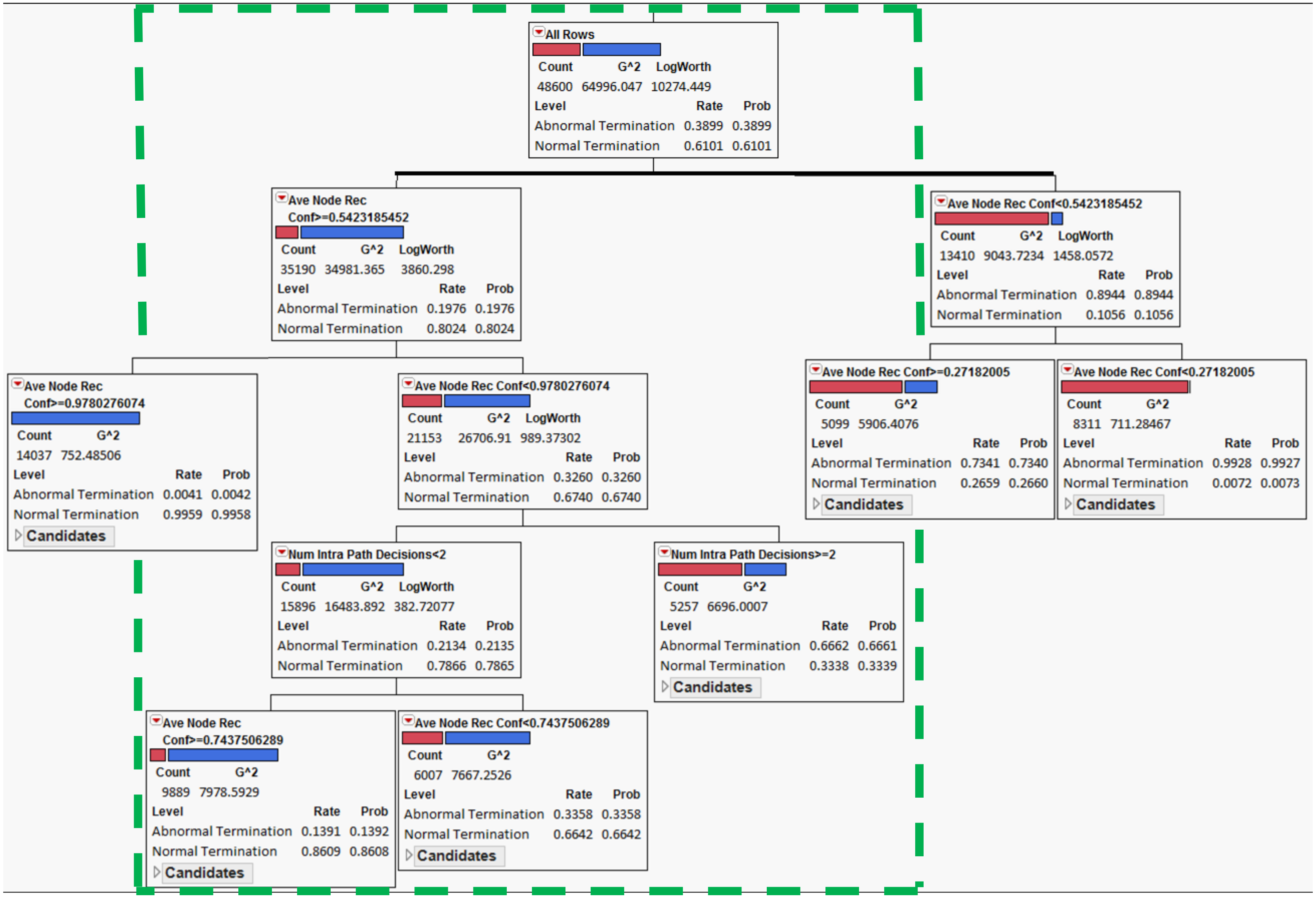


Figure 9-27 Partition Tree of Potential Progress Metrics

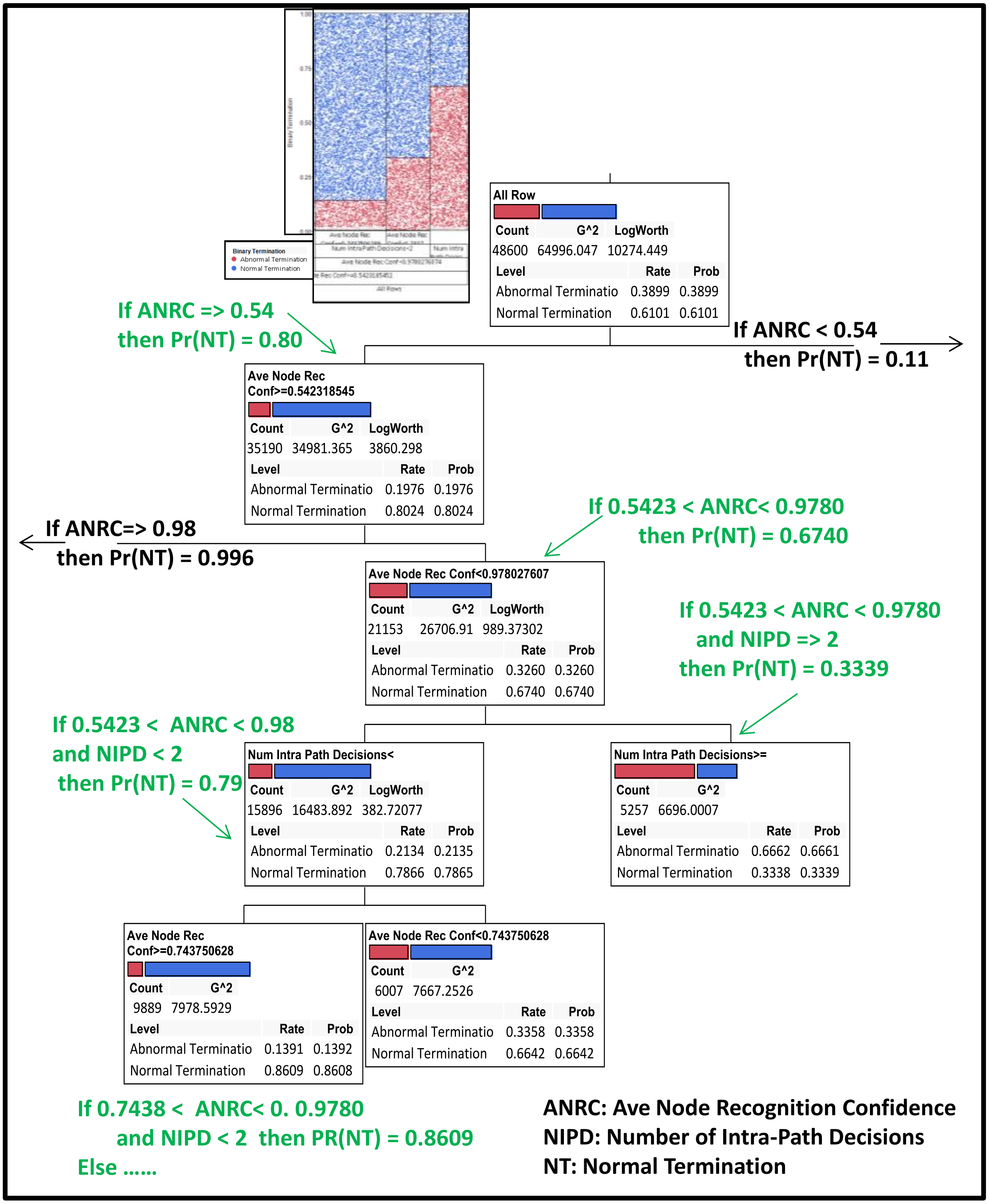


Figure 9-28 If-Then-Else Rules Suggested by Partition Tree Analysis

9.6 Summary of Distortion Experiment Findings

The simulation model can produce data sufficient to discriminate between output classes based on model parameter values as indicated by the preponderance of statistically significant effects found in various fit models. This capability to determine parameter and parameter interaction effects supports the search for explanatory hypotheses as to the causes of entity behavior. The use of interactive simulations in combination with parametric experiments both supports the development of such hypotheses and provides evidence as to their validity.

This set of experiments shows the expected trends entity success/failure rates as functions of the extent and degree of VIH map distortion – more errors in the mental map decrease the probability that the entity will reach its goal.

It is useful to go beyond binary classification of success and failure to explore entity behavior and the model provides adequate mechanisms for breaking down these classes into sub-classes that support explanatory investigation into such behavior. Of particular interest are the trials where the entity does not achieve the success optimum, the minimal distance ground truth path. These trials represent the instances where changes to the agent's mental map and movement strategies could potentially improve performance. They represent the fertile ground for improving SA/SU.

Data produced by the model can support incremental assessment of entity progress towards goals and has the potential to support adaptive goal-seeking strategies based on such assessment.

The model can accommodate any map structure that can be expressed as a planar arc/node network, and shows the ability to investigate the effects of network complexity on entity behavior.

10 EXPLORATION OF NODE RECOGNITION FACTORS

This set of simulation experiments takes an in depth look at the VIH agent's ability to recognize ground truth nodes in terms of their correspondence with nodes on its VIH map.

10.1 Experimental Setup

Simulation trials were executed for each map, varying all of the parameters listed in Table 10-1, with each combination of parameter values run for 40 trials.

Table 10-1 Experiment Parameter Values

Pr(Ne)
0.1
0.3
0.5

X

X and Y Error Limit
75
125

X

DM Coef	EM Coef	LM Coef	CA Coef
0	0	0	0
0	0	0	1
0	0	0.33	0.67
0	0	1	0
0	0.25	0.25	0.5
0	0.33	0	0.67
0	0.5	0.5	0
0	1	0	0
0.2	0.2	0.2	0.4
0.25	0	0.25	0.5
0.25	0.25	0	0.5
0.33	0	0	0.67
0.33	0.33	0.33	0
0.5	0	0.5	0
0.5	0.5	0	0
1	0	0	0

3 X 2 X 16 = 96

The first two parameter boxes listed in the table were used to construct the different VIH maps for each trial, specifying the amount and extent of distortion of the given GT map. The larger matrix identifies the combinations of match weight coefficients used in the node recognition value hierarchy, and provides the variation in node recognition capabilities that is the subject of this experiment. All other parameter values were held constant, with Pr(MN) = 0.1, GT-VIH threshold confidence limit = 0.75, the good fit delta = 0.15, the VIH Candidate distance = 100, and the keep searching upper limit = 50.

Each of the 96 different parameter combinations in Table 10-1 was run for each of the three maps. The simulation experiments used the same three basic ground truth maps as before, in this case producing a total of 11,520 simulation runs.

10.2 <u>Results: Assessment of Binary Termination Conditions</u>

As before, the first question of interest for each trial is whether the entity succeeds in finding its goal. As shown in Figure 10-1, the entity's overall success rate is 52%, considerably lower than the 61% rate in seen in the distortion experiments of Chapter 9.

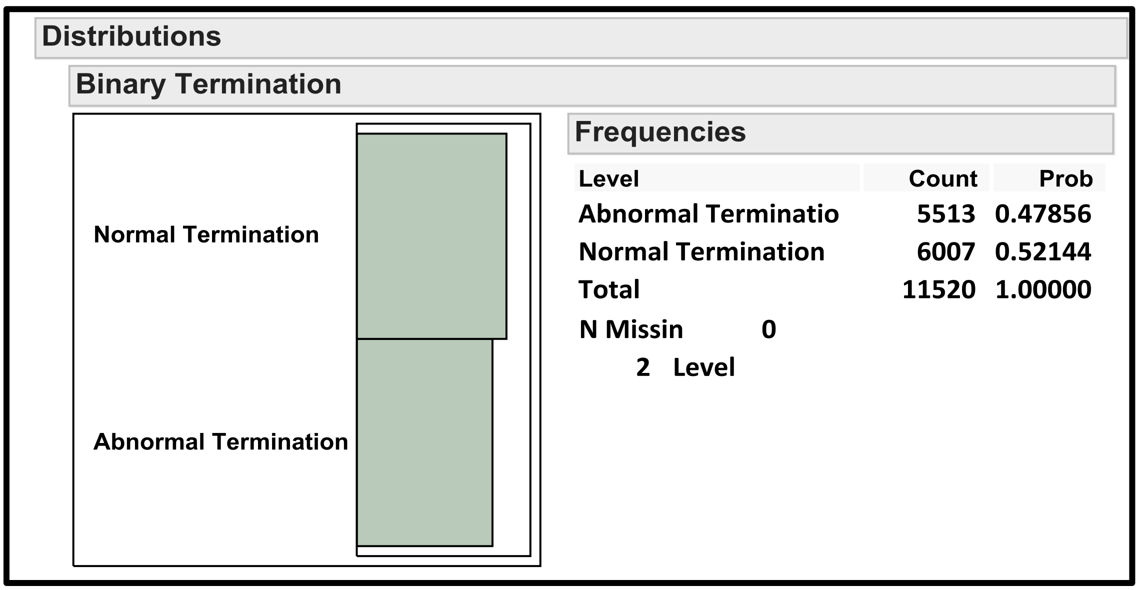
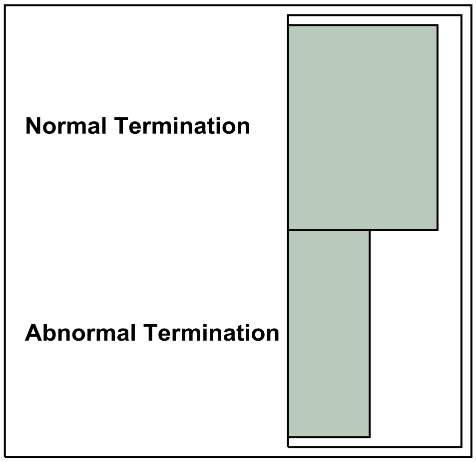


Figure 10-1 Normal vs. Abnormal Termination All Trials

In the instance where all of the match coefficient weights in the VFT node recognition hierarchy were zero (720 of the 11520 trial), it was impossible for the entity to succeed – even if the entity managed to reach its goal, it wouldn't recognize it. Those trials constitute 6.25% of the total so they account for most of, but not the entire, drop in entity performance rates from the experiments of Chapter 9.

Distributions

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	255	0.35417
Normal Termination	465	0.64583
Total	720	1.00000

N Missin 10800

2 Level

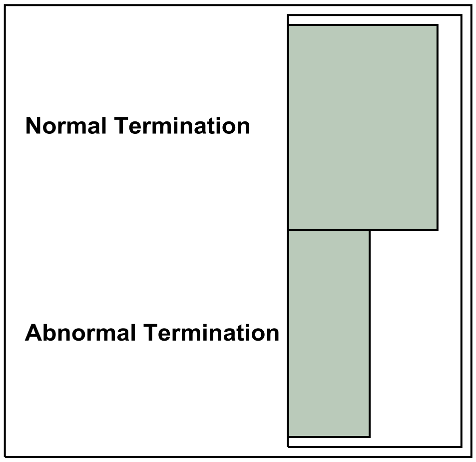
Figure 10-2 Termination Results for All Node Recognition Coefficients Equal

shows the most direct contrast with the "all coefficients zero" or "no recognition case"- the 720 trials where no match coefficient is zero. For this condition to hold, all of the coefficients must be equal, as they were in the experiments of the previous chapter. In this "no zero coefficients" case the success rate is approximately 65%, an improvement over the 52% rate of the over-all data set that includes the "no recognition" case, and an indication of the importance of the VIH agent's ability to recognize nodes.

If the "all node recognition match coefficients zero" trials are ignored, the overall success rate in these experiments is the 55.6% listed in Figure 10-2, which can also be compared to the 61% success rate observed in Chapter 9. The DOE matrix for those experiments included higher probabilities of node distortion than are present in these current simulations, but it also included cases with smaller limits in the size of changes in X and Y coordinate values, and cases in which there was no topological distortion, so comparison of success rates is not straightforward.

Distributions

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	255	0.35417
Normal Termination	465	0.64583
Total	720	1.00000

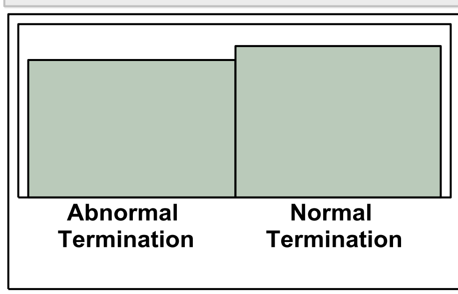
N Missin 10800

2 Level

Figure 10-2 Termination Results for All Node Recognition Coefficients Equal

Distributions

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	5513	0.47856
Normal Termination	6007	0.52144
Total	11520	1.00000

N Missin 0

2 Level

Figure 10-3 Termination Results Excluding the All Coefficients Zero Trials

Given evidence that node recognition capability does affect performance, the experiments of this chapter are designed to investigate whether some of the recognition capabilities are more important than others, and if so, which and to what degree.

Examination of outcomes for each of the four node recognition match coefficients[93],

[93] There are actually five such coefficients in the current VFT node recognition hierarchy, but for convenience the two color attributes have been given the same coefficient and their collective value is aggregated with the other three factors: Distance

as shown the JMP Y by X Contingency Analysis of Binary Termination in Figure 10-4, indicates that the “all node recognition coefficients equal” cases result in the highest probability of success. This case also shows that relying only on the distance match or color attribute measures does not appreciably differ from the all trials composite result and in fact is a little worse.

Interpreting Figure 10-4 is not straightforward, as the different cases shown are not independent of one another. More clarity is possible by looking at JMP Nominal Logistic Fit and JMP Regression Tree Analyses.

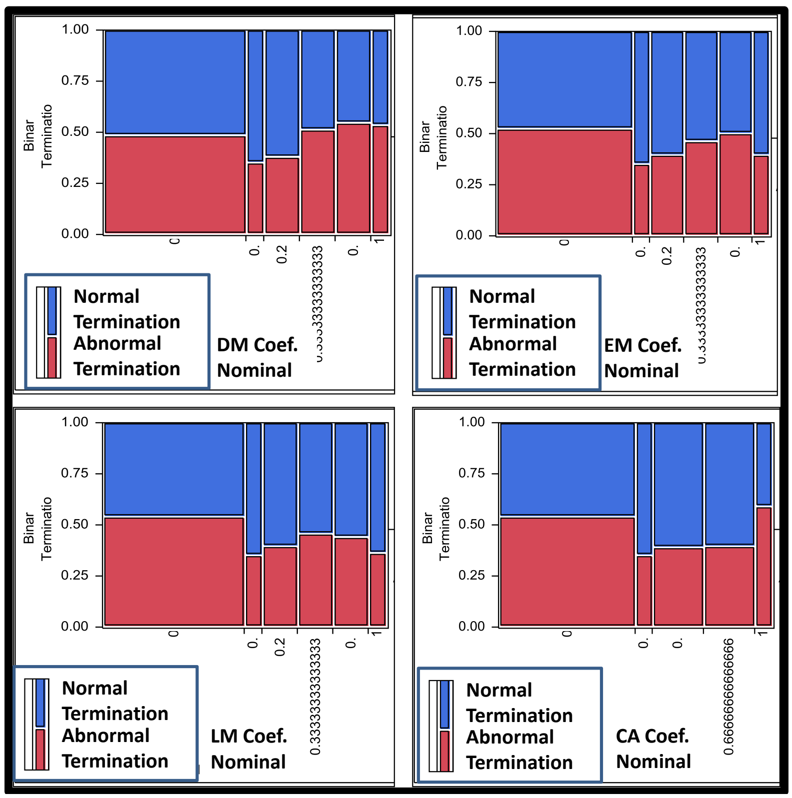


Figure 10-4 Fit of Binary Termination by Node Recognition Match Coefficients

Table 10-2 shows the JMP Whole Model Fit is statistically significant. The model has a misclassification rate of 29%, an improvement of almost 20% over just selecting the

---

Match, Expectation Match, and Link Match, which occupy the same level of the hierarchy. As a result the single color coefficient value reported is always twice that of the other factors in the simulation DOE matrix of Table 10-1.

most common outcome, normal termination. The confusion matrix in Table 10-3 shows the Logistic Regression Model over-predicts the most likely result, success/normal termination, and under-predicts the less likely result, failure/ abnormal termination.

Table 10-2 Whole Model Test: Binary Termination of Node Recognition Match Coefficient Experiments

Whole Model Test Nominal Logistic Fit for Binary Termination Converged in Gradient, 22 iterations				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	1642.7768	35	3285.554	<.0001
Full	6331.6836			
Reduced	7974.4604			
RSquare (U)	0.206			
AICc	12735.6			
Misclassification Rate	0.2869			
Observations	11520			

Table 10-3 Nominal Logistic Model Confusion Matrix: Binary Termination of Node Recognition Match Coefficient Experiments

Confusion Matrix			
Actual/Predicted	Abnormal Termination	Normal Termination	Row Sums (Actual Observed)
Abnormal Termination	3584	1929	5513
Normal Termination	1376	4631	6007
Column Totals (Predicted)	4960	6560	11520

True Positive	77.09%
True Negative	65.01%
False Positive	27.74%
False Negative	29.41%

	Predicted	Actual
Abnormal	43.06%	47.86%
Normal	56.94%	52.14%

The model does show statistical evidence of Lack of Fit, as provided in Table 10-4.

Table 10-4 Lack of Fit Test: Binary Termination of Node Recognition Match Coefficient Experiments

Lack Of Fit				
Source	DF	-LogLikelihood	ChiSquare	Prob>ChiSq
Lack Of Fit	252	159.1733	318.3465	0.0029
Saturated	287	6172.5103		
Fitted	35	6331.6836		

The Effect Likelihood Ratio Test results are provided in Table 10-5. They show that all of the match coefficient parameters have significant main factor effects and most of them have significant two-way interactions as well.

Table 10-5 Chapter Experiments Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Map	2	2	0.000	1
Pr(NE)	1	1	0.000	0.9988
X and Y Error Limit	1	1	0.000	0.9967
DM Coef	1	1	386.026	<.0001
EM Coef	1	1	562.030	<.0001
LM Coef	1	1	687.249	<.0001
CA Coef	1	1	1021.729	<.0001
Map*Pr(NE)	2	2	4.558	0.1024
Map*X and Y Error Limit	2	2	0.986	0.6106
Map*DM Coef	2	2	0.000	1
Map*EM Coef	2	2	0.000	1
Map*LM Coef	2	2	0.000	1
Map*CA Coef	2	2	0.000	1
Pr(NE)*X and Y Error Limit	1	1	13.199	0.0003
Pr(NE)*DM Coef	1	1	0.000	0.9998
Pr(NE)*EM Coef	1	1	0.000	0.9998
Pr(NE)*LM Coef	1	1	0.000	0.9999
Pr(NE)*CA Coef	1	1	0.000	0.9999
X and Y Error Limit*DM Coef	1	1	0.000	1
X and Y Error Limit*EM Coef	1	1	0.000	0.9999
X and Y Error Limit*LM Coef	1	1	0.000	0.9999
X and Y Error Limit*CA Coef	1	1	0.000	0.9999
DM Coef*EM Coef	1	1	54.446	<.0001
DM Coef*LM Coef	1	1	5.579	0.0182
DM Coef*CA Coef	1	1	59.965	<.0001
EM Coef*LM Coef	1	1	6.336	0.0118
EM Coef*CA Coef	1	1	71.959	<.0001
LM Coef*CA Coef	1	1	56.110	<.0001

The spatial distortion parameters represented do not have significant main factor effects, but they do show significant effects when interacting with each other. They are, in fact, the first values on which the regression tree splits, as seen in Figure 10-5 and Figure 10-6. These splits are then followed by a variety of partitions on node recognition match criteria values, which do not appear to follow any consistent pattern with respect to the order in which the criteria are selected.

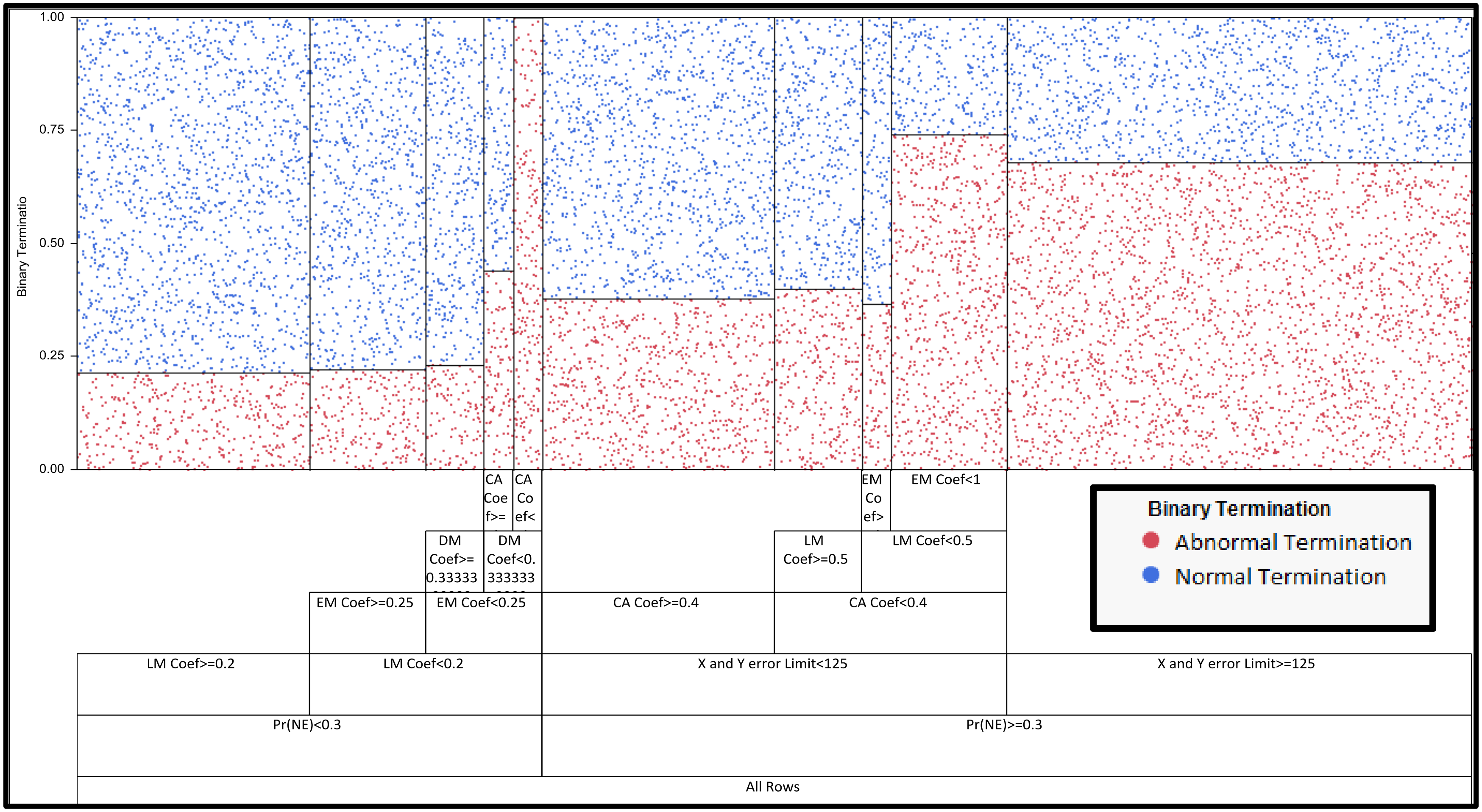


Figure 10-5 Regression Tree Graph: Partition of Binary Termination by Node Recognition Match Coefficients

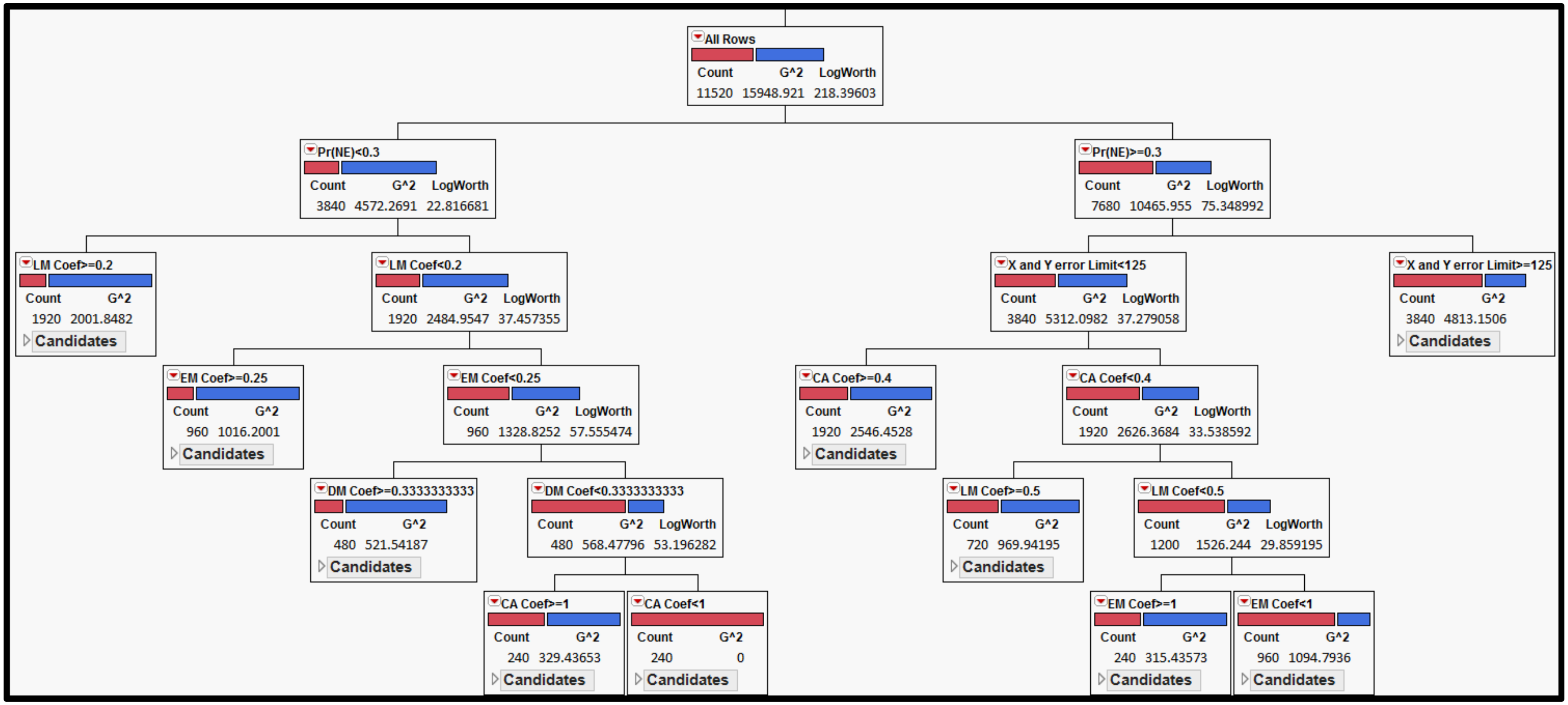


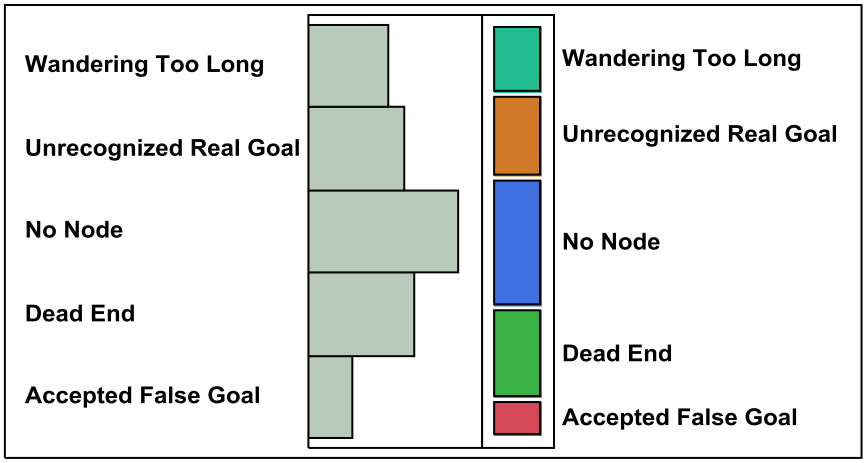
Figure 10-6 Regression Tree: Partition of Binary Termination by Node Recognition Match Coefficients

10.3 Refinement of Abnormal Termination Results

For the set of experiments in this chapter, the breakdown of the abnormal termination into the five “categories of failure” is illustrated in Figure 10-7.

Distributions

Termination Condition



Frequencies

Level	Count	Prob
Accepted False Goal	511	0.09269
Dead End	1232	0.22347
No Node	1735	0.31471
Unrecognized Real Goa	1110	0.20134
Wandering Too Long	925	0.16779
Total	5513	1.00000

N Missin 6007

5 Level

Figure 10-7 Node Recognition Match Coefficients: Abnormal Termination Breakdown

The unrecognized goal condition occurs almost 10% of the time. It represents the success rate if the VIH agent has some, but very limited, node recognition capability, i.e., restricted to only recognizing its goal.

These values do not differ dramatically from those in Chapter 9. The probability of unrecognized real goals and the probability of accepted false goals both increase by about 3.5%, while the probability of wandering too long (without recognizing a node on the planned path) decreases by almost 5% and the probability of no node (lost in between nodes) decreases by around 2%.

Table 10-6 Abnormal Output Comparison

	Node Recognition Experiments			Spatial Distortion Experiments			
Termination Condition	**Count**	**Percent of Abnormal**		**Count**	**Percent of Abnormal**		**Difference**
Accepted False Goal	**511**	**9.27%**		**1055**	**5.57%**		**3.70%**
Dead End	**1232**	**22.35%**		**4357**	**23.00%**		**-0.65%**
No Node	**1735**	**31.47%**		**6388**	**33.72%**		**-2.24%**
Unrecognized Real Goal	**1110**	**20.13%**		**3155**	**16.65%**		**3.48%**
Wandering Too Long	**925**	**16.78%**		**3992**	**21.07%**		**-4.29%**
Total Abnormal	**5513**	**100.00%**		**18947**	**100.00%**		
Abnormal Percent of All Trials	**47.86%**			**38.90%**			**8.96%**
Total Normal	**6007**			**29653**			
Normal Percent of All Trials	**52.14%**			**61.10%**			**-8.96%**

One would expect that the 720 trials (6.25% of the total) in which no recognition is possible should lead to an increase in unrecognized real goals.

In 450 (.93%) of the Chapter 9 trials the probability of missing nodes, Pr(MN), was zero, which should have contributed to higher probabilities of success in those trials, as all of the cases in this current experiment had a constant Pr(MN) =0.01. On the other hand, the current experiments do not consider the highest probabilities of node error, Pr(NE) and missing nodes (0.7 and 0.3 respectively), which combined represent almost 2% of the Chapter **Error! Reference source not found.** results, and one would presume these values would decrease the probability of over-all success. Figure 10-8 and Figure 10-9 show how Pr(NE) and Pr(MN) affect the different categories of abnormal termination in the Chapter 9 experiments.

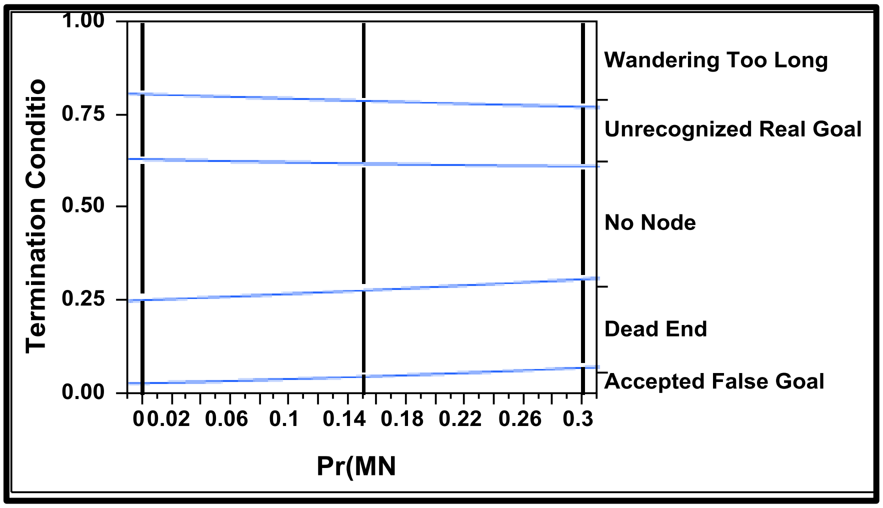


Figure 10-8 Chapter 9 Distortion Experiments: Fit Abnormal Termination Conditions by Pr(MN)

As the Pr(MN) increases so does the frequency of wandering too long, which helps explain the decrease in that category seen in the current Chapter's results. Similarly the greater number of no node results at Pr(MN) = 0 appears to overcome the decrease seen at Pr(MN) =0.3, which would result in the slight increase in no node responses of Chapter101010 results over those of Chapter 9.

On the other hand, varying Pr(NE) shows little effect in the kinds of abnormal termination even though it has strong effect on the aggregate results.

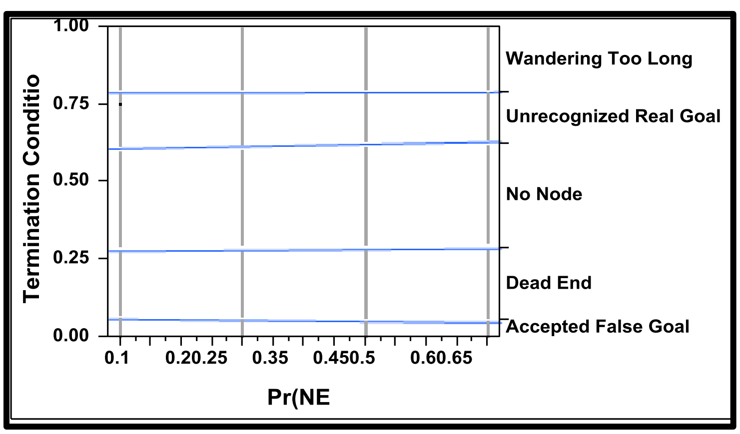


Figure 10-9 Chapter 10 Experiments: Fit Abnormal Termination Conditions by Pr(NE)

Table 10-7 compares the abnormal results of Chapter 10and Chapter 9.

Table 10-7 Chapter 10 and Chapter 9 Abnormal Results Comparison

	Node Recognition Experiments			Spatial Distortion Experiments			
Termination Condition	Count	Percent of Abnormal		Count	Percent of Abnormal		Difference
Accepted False Goal	511	9.27%		1055	5.57%		3.70%
Dead End	1232	22.35%		4357	23.00%		-0.65%
No Node	1735	31.47%		6388	33.72%		-2.24%
Unrecognized Real Goal	1110	20.13%		3155	16.65%		3.48%
Wandering Too Long	925	16.78%		3992	21.07%		-4.29%
Total Abnormal	5513	100.00%		18947	100.00%		
Abnormal Percent of All Trials	47.86%			38.90%			8.96%
Total Normal	6007			29653			
Normal Percent of All Trials	52.14%			61.10%			-8.96%

The table demonstrates the capability of the model to focus on specific characteristics of the simulated ground truth and perceived environments, and on how they affect decision-making behaviors.

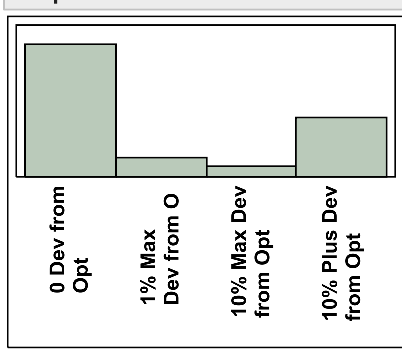
14. Future Work: Sensitivity Analyses of Node Recognition Decisions

AFIT provides an Excel add-in VFT tool that can be used to replicate specific node recognition decisions, evaluating the candidate VIH nodes for their correspondence with a given GT node. For example, the tool can explore shifting relative hierarchy branch weights to determine when relative ranking of alternatives would change. This tool would allow an in-depth sensitivity analysis of the node recognition VFT hierarchy with respect to both match *coefficients and SDVF measure parameters, but it is only practical to run the tool off-line from the simulation model. Using the tool to conduct sensitivity analysis would therefor require modifying the simulation to record relevant data for use in such an analysis. The simulation currently flags node recognition errors, i.e., false positives, false negatives, and situations where no candidate passes recognition threshold values, but at present it only maintains a record of the number of occurrences of these conditions. The capability to export the relevant Gt and VIH candidate node characteristics could easily be added to the model. Automatically identifying the most "interesting" decision errors for analysis out of thousands of trials is not so straightforward, which simply means that such data would be best gathered operating the simulation in its interactive mode.*

10.4 <u>Assessment of Normal Termination</u>

Figure 10-10 shows results for normal termination using the categories developed in Chapter 9. The results for the current node recognition factor experiments are very similar to those of the spatial distortion experiments of Chapter 9. The frequencies for each category differ by around one to two percent between the two sets of experiments.

%Optimal Normal Termination



Frequencies

Level	Count	Prob
0 Dev from Opt	3637	0.60546
1% Max Dev from Opt	510	0.08490
10% Max Dev from Op	238	0.03962
10% Plus Dev from Op	1622	0.27002
Total	6007	1.00000

N Missin 5513
4 Level

Figure 10-10 Node Recognition Experiments: Frequency of Normal Termination Optimality Categories

Figure 10-11 provides another view of how normal termination categories are affected by the node recognition match coefficients.

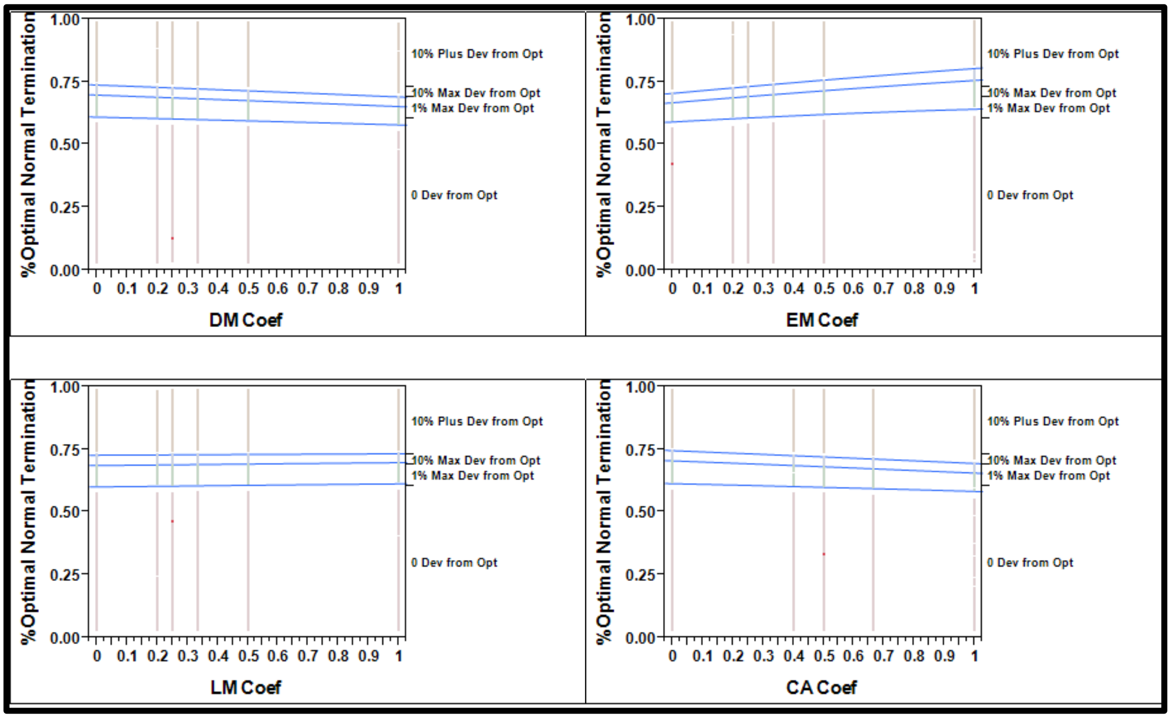


Figure 10-11 Fit of Normal Termination Categories by Node Recognition Coefficients

Figure 10-12 shows the partition tree graph of a logistic regression tree for the current series of experiments. The tree shows much the same structure as the binary regression tree of Figure 10-5 and Figure 10-6.

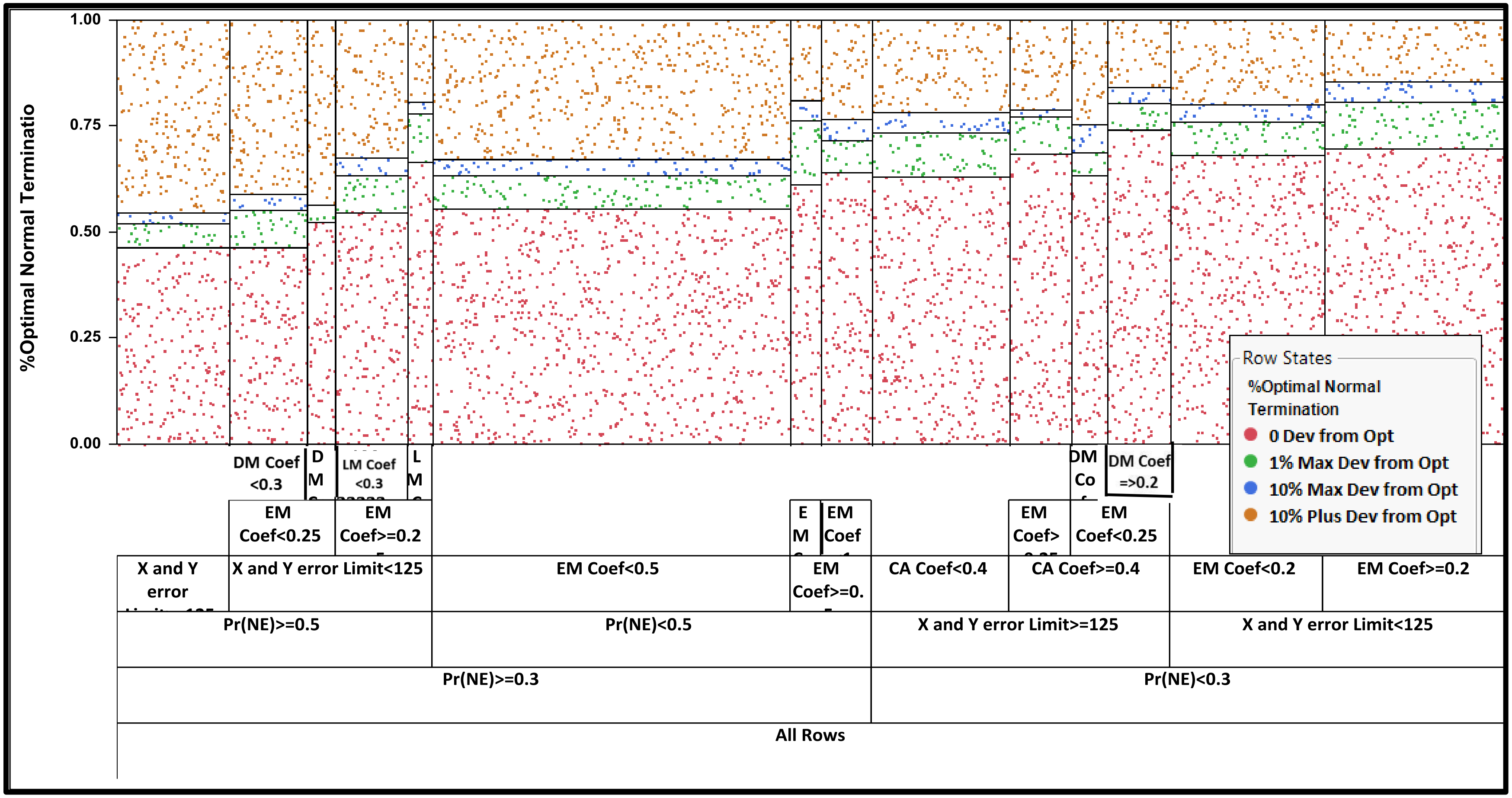


Figure 10-12 Node Recognition Experiments: Partition of Normal Termination Categories by Node Recognition Match Coefficients

In this instance, the expected node match coefficient is noticeably more prominent in the early splits than the other match coefficients, which may be evidence of the importance of staying on the planned best route. Further evidence is provided by Figure 10-13, which shows the effects on optimal performance of both straying from the VIH path, and the VIH agent's confidence in having recognized GT nodes on its VIH map.

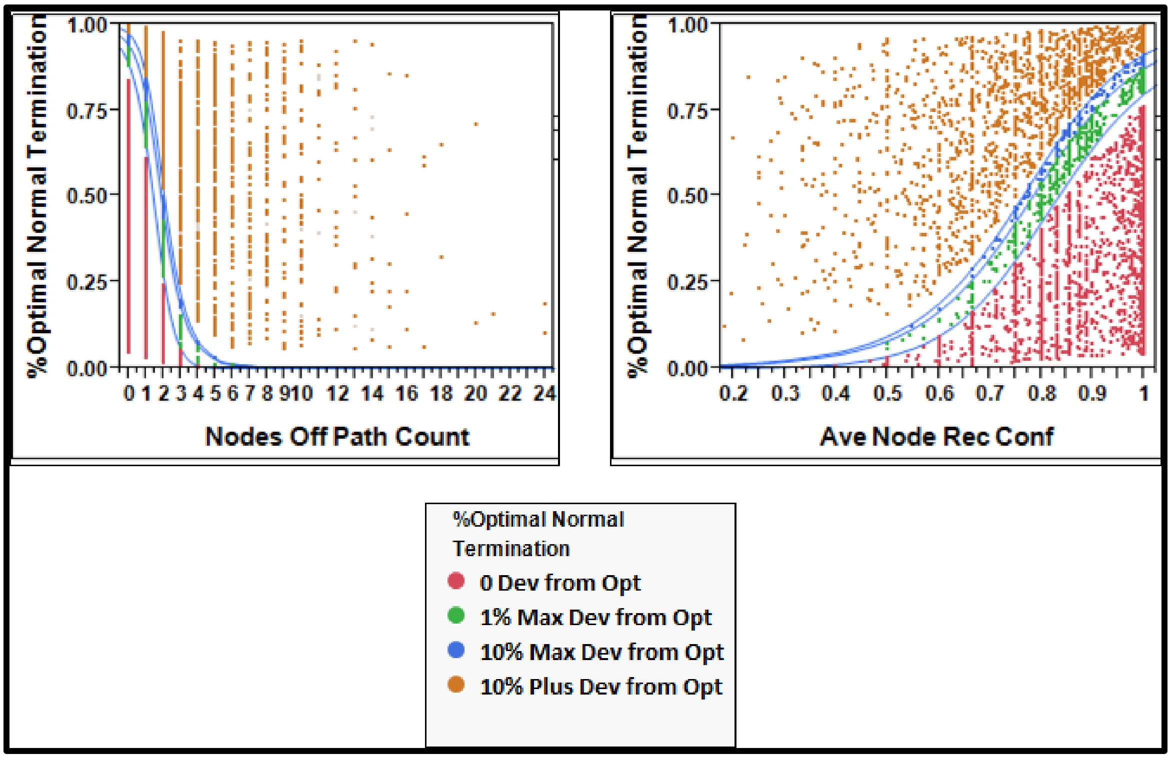


Figure 10-13 Percent Optimal Performance as a Function Nodes Off Path Count and Average Node Recognition Confidence

Figure 10-14 shows that the Square Map is more prone to tolerate moderate deviations from the optimal route, a similar result to that seen in Chapter 9. A breakdown by map of the off path node analysis of Figure 10-13 is similar to that provided in Figure 9-23. It shows essentially the same map interaction effect with nodes off path counts seen in that earlier figure.

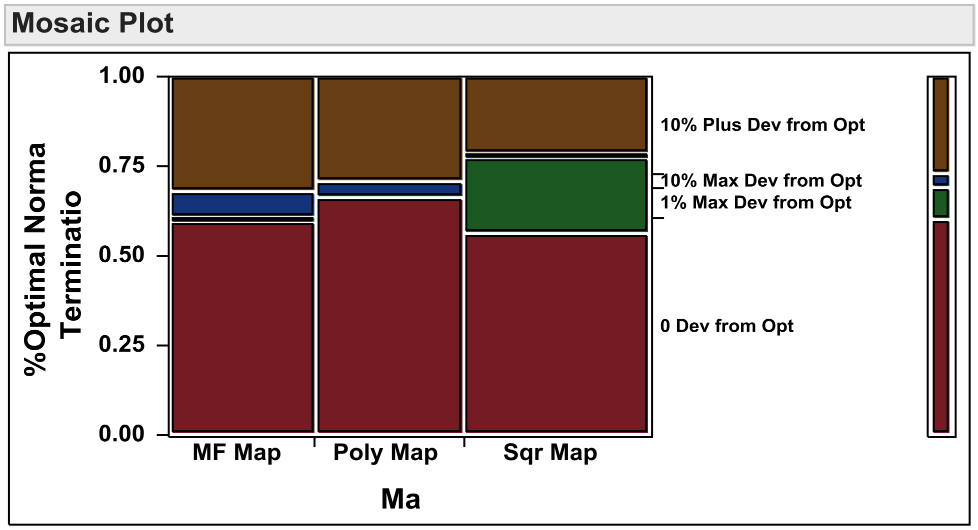


Figure 10-14 Contingency Analysis of Normal Termination Alternatives

10.5 Measuring Progress Towards the Goal

As in Chapter 9, the potential for information available to the VIH agent to assist in evaluating its chances for success is shown by looking at a logistic fit to binary termination outcomes. The values used are[94]:

- Average Node Recognition Confidence;
- Number of Off Path Nodes Observed;
- Number of Intra-Path Decisions Made; and
- Map (under the belief that the agent can define some measure of map complexity or difficulty.

The results again suggest strongly that a regression model fit to these values can support prediction of success in MOBIL. Table 10-8 gives the results for nominal logistic fit models, both for all trials in the experiment and for those trials excluding the outcomes for the case in which all node recognition coefficients were set to zero. The only substantial change from the results of Chapter 9 is that here there is strong

[94] GT-VIH Correspondence Threshold was not included in this analysis, as it was in the analysis performed Chapter 9 because all values were held constant for the current set of experiments.

evidence of model lack of fit, indicating that there is significant probability of model improvement through the consideration of additional factors.

Table 10-8 Nominal Logistic Fit: Binary Termination Outcomes by Information Available to the VIH Agent

	All Trials	All Zeros Case Excluded
Misclassification Rate	0.1839	0.1921
ROC AUC	0.8192	0.9047
Lack of fit Pr > chi-square	<.0001	<.0001

Table 10-9 gives the significance of the individual factors and their two-way interaction. The two instances where there is no evidence of significant interaction between the type of map and the information factors used suggest the need for further investigation with maps that show more difference in some of the node recognition characteristics. For example, it would be interesting to explore maps that had different degrees of color coding for nodes and arcs.

Table 10-9 Effect Likelihood Ratio Test for Table 10-8 Results

Effect Likelihood Ratio Tests						
			All Trials		All Zeros Case Excluded	
Source	Nparm	DF	L-R ChiSquare	Prob > ChiSq	L-R ChiSquare	Prob > ChiSq
Map	2	2	28.762	<.0001	33.235	<.0001
Nodes Off Path Count	1	1	18.673	<.0001	29.221	<.0001
Ave Node Rec Conf	1	1	2939.245	<.0001	2371.290	<.0001
Num Intra Path Decisions	1	1	791.137	<.0001	761.565	<.0001
Map* Nodes Off Path Count	2	2	9.531	0.009	7.936	0.019
Map* Ave Node Rec Conf	2	2	0.536	0.765	0.070	0.966
Map* Num Intra Path Decisions	2	2	7.823	0.020	9.605	0.008
Nodes Off Path Count* Ave Node Rec Conf	1	1	27.719	<.0001	25.888	<.0001
Nodes Off Path Count* Num Intra Path Decisions	1	1	1.193	0.275	0.124	0.725
Ave Node Rec Conf* Num Intra Path Decisions	1	1	151.311	<.0001	205.116	<.0001

10.6 Summary of Node Recognition Experiment Findings

Again the simulation model has demonstrated it can produce data sufficient to discriminate between output classes based on model input parameter values, supporting the search for explanatory hypotheses as to the causes of entity behavior.

The entity success/failure rate trends of this set of experiments show the importance of node recognition factors in the probability that the entity will reach its goal. There are, however, only a few instances where they distinguish between the value of different node recognition characteristics, the most notable of these instances being the observations with respect to expectation matches.

15. Future Work : Arc/link Recognition Factors

As discussed above in Chapter in the current model the calculation of GT arc correspondence with VIH arc/link candidates is based only the bearing is based on the match between desired VIH path elements and the compass direction of available GT arc. VIH arc/link decisions concern both the desired VIH arc/link, i.e., the link most likely to lead to the next intermediate node goal, and the ability to match that arc/link to the choices available at the current GT node. The model could first augment decision-making capability by providing a larger set of potential arc characteristics, such as identifying a set of contiguous arcs as representing a road or highway, or identifying the arcs as belonging to or connecting geographical regions. The model's incorporation of the OpenStreetMap structures discussed in Appendix A would facilitate these kinds of improvements and also allow more substantive consideration of real world environments. A second, and more substantial, set of model improvement would require the model to apply these augmented arc/link capabilities in the use of both types of arc/link decisions, and would require more sophisticated logic with respect to "next move" choices.

11 FACTOR INTERACTION EXPERIMENTS

The previous two chapters each provided details on effects of MOBIL's spatial distortion and node recognition parameters. This chapter briefly looks at a massive set of trial runs designed to provide more specifics on the interaction between these two sets of parameters.

11.1 <u>Experimental Setup</u>

These trials were restricted to the use of the multi-factor map, and varied the parameters shown in Figure 11-1.

Pr (NE)
0.1
0.3
0.5
0.7

Pr (MN)
0.00
0.15
0.30

X and Y Error Limit
0.00
0.15
0.30

GT-VIH Corspnd. Thshd
0.60
0.75
0.90

CA Std. Dev.
0
20
40

X X X X X

DM Coef	Em Coef	LM Coef	CA Coef
0.00	0.00	0.00	0.00
0.00	0.00	0.00	1.00
0.00	0.00	0.33	0.67
0.00	0.00	1.00	0.00
0.00	0.25	0.25	0.50
0.00	0.33	0.00	0.67
0.00	0.50	0.50	0.00
0.00	1.00	0.00	0.00
0.20	0.20	0.20	0.40
1.00	0.00	0.00	0.00

4 X 3 X 3 X 3 X 3 X 10 = 3240

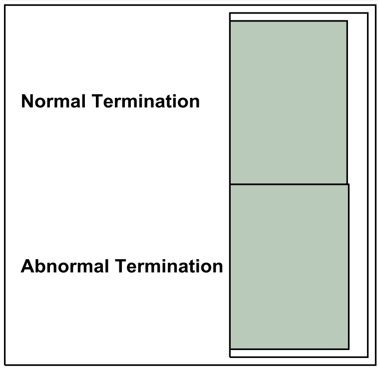
Figure 11-1 Factor Inter-Action DOE

Again the first three boxes on the left give the parameter used to distort the VIH map, as does the box labeled CA Standard Deviation, whose values are used to distort the color attributes of nodes. The larger matrix identifies the combinations of match weight coefficients used in the node recognition value hierarchy, and the GT-VIH Correspondence Threshold box gives the values for acceptance or rejection of the best candidate node. As the figure indicates, there were 3,240 distinct cases, each of which was executed for 50 trials resulting in a total of 162,000 MOBIL trials. Figure 11-2 shows the fraction of normal and abnormal outcomes observed. There results are statistically significant, but don't appear to represent any practical distinction between the two types of outcomes.

11.2 Results: Assessment of Binary Termination Outcomes

Distributions

Binary Termination



Frequencies

Level	Count	Prob
Abnormal Terminatio	81388	0.50240
Normal Termination	80612	0.49760
Total	162000	1.00000

N Missin 0

2 Level

Confidence Intervals

Level	Count	Prob	Lower CI	Upper CI	1-Alpha
Abnormal Terminatio	81388	0.50240	0.49996	0.50483	0.950
Normal Termination	80612	0.49760	0.49517	0.50004	0.950
Total	162000				

Note: Computed using score confidence intervals

Figure 11-2 Factor Inter-Action Experiments: Binary Termination

The breakdown of normal termination results based on the amount of spatial and topological distortion of the VIH map in Table 11-1 confirms the trends seen in Chapters 10 and 9, as distortion increases the probability for success decreases.

Table 11-2 provides a similar breakdown for the cases varying node recognition factors, the VFT hierarchy match coefficients and the recognition acceptance thresholds. The all match coefficients case again provides no capability for the entity to successfully achieve its goal. For these cases at least, the expectation factor and the match in number of arc/links by quadrant provide by far the best that node recognition capabilities. This conclusion is enhanced by the JMP fit of Y by X graphs of Figure 11-3.

Table 11-1 Probability of Normal Termination as a Function of VIH Map Distortion

MF Map All Trials Each Case 4500 Replications		Probability of Node X and Y Error			
Probability of Missing Nodes	Amount of X and Y Error	0.1	0.3	0.5	0.7
0	25	0.846	0.806	0.771	0.742
	75	0.746	0.548	0.419	0.313
	125	0.649	0.348	0.192	0.109
0.15	25	0.756	0.719	0.681	0.655
	75	0.687	0.518	0.404	0.305
	125	0.578	0.335	0.189	0.105
0.3	25	0.687	0.652	0.620	0.614
	75	0.627	0.474	0.369	0.294
	125	0.538	0.324	0.186	0.108

Interestingly, in these cases the recognition threshold does not appear to by an important factor.

Table 11-2 Probability of Normal Termination as a Function of Node Recognition Factors

Node Recog. Match Coef.				GT-VIH Corrspd. Thrshld.			
DM Coef	EM Coef	LM Coef	CA Coef	0.600	0.750	0.900	Row Ave.
0.00	0.00	0.00	0.00	0.000	0.000	0.000	0.000
0.00	0.00	0.00	1.00	0.316	0.319	0.325	0.320
0.00	0.00	0.33	0.67	0.565	0.597	0.616	0.593
0.00	0.00	1.00	0.00	0.580	0.595	0.605	0.593
0.00	0.25	0.25	0.50	0.610	0.618	0.625	0.618
0.00	0.33	0.00	0.67	0.559	0.617	0.616	0.598
0.00	0.50	0.50	0.00	0.593	0.622	0.624	0.613
0.00	1.00	0.00	0.00	0.564	0.558	0.555	0.559
0.20	0.20	0.20	0.40	0.604	0.627	0.471	0.567
1.00	0.00	0.00	0.00	0.524	0.513	0.509	0.515
			Co. Ave.	0.492	0.507	0.495	0.498

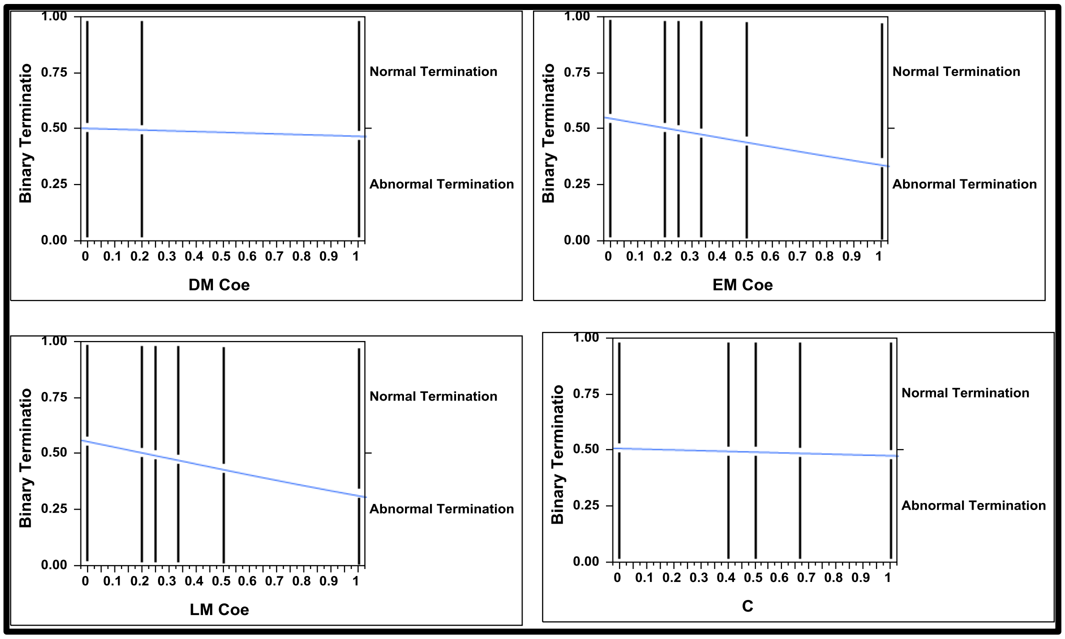


Figure 11-3 Fit of Binary Termination by Node Recognition VFT Coefficients

Table 11-3 shows the statistics for the JMP nominal regression fit of binary outcomes by the experimental factors.

Table 11-3 Whole Model Test: Binary Termination of Factor Interaction Experiments by All Case Factors

Converged in Gradient, 23 iterations				
Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	34160.82	43	68321.65	<.0001
Full	78127.16			
Reduced	112287.98			
RSquare (U)	0.3042			
AICc	156346			
Misclassification Rate	0.2414			
ROC AUC	0.84374			
Observations	162000			

Lack Of Fit				
Source	DF	-LogLikelihood	ChiSquare	Prob>ChiSq
Lack Of Fit	3196	3927.336	7854.672	<.0001
Saturated	3239	74199.824		
Fitted	43	78127.16		

Considering all of the factors produces a fit that improves the classification rate from about 50% for selection of the most probable outcome to better than 75%. Further evidence of the degree of fit model efficacy is given by the 0.84 area under the ROC curve.

The confusion matrix in Table 11-4shows the Logistic Regression Model over-predicts success/normal termination, and under-predicts failure/ abnormal termination.

Table 11-4 Nominal Logistic Model Confusion Matrix: Binary Termination Factor Interaction Experiments by All Case Factors

Confusion Matrix			
Actual/Predicted	Abnormal Termination	Normal Termination	Row Sums (Actual Observed)
Abnormal Termination	59512	21876	81388
Normal Termination	17238	63374	80612
Column Totals (Predicted)	76750	85250	162000

True Positive	77.09%
True Negative	65.01%
False Positive	27.74%
False Negative	29.41%

	Predicted	Actual
Abnormal	47.38%	50.24%
Normal	52.62%	49.76%

Table 11-5 list all of the significant factor interactions for the JMP fit model. The probability of missing nodes and the node recognition confidence threshold do not appear on the list anywhere, while all four of the node recognition match factors appear as main effects.

Shown in Figure 11-4 and Figure 11-5 is a regression tree partition of the Factor Interaction Experiments by all the case factor with more than 30 splits. It provides an eclectic mix of partition split factors, which again do not appear to follow any consistent pattern with respect to the order in which the criteria are selected.

Table 11-5 Nominal Logistic Model Significant Factor Interactions: Binary Termination Factor Interaction Experiments by All Case Factors

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
DM Coef	1	1	386.026	<.0001
EM Coef	1	1	562.030	<.0001
LM Coef	1	1	687.249	<.0001
CA Coef	1	1	1021.729	<.0001
Pr(NE)*X and Y Error Limit	1	1	13.199	0.0003
DM Coef*LM Coef	1	1	5.579	0.0182
DM Coef*CA Coef	1	1	59.965	<.0001
EM Coef*LM Coef	1	1	6.336	0.0118
EM Coef*CA Coef	1	1	71.959	<.0001
LM Coef*CA Coef	1	1	56.110	<.0001

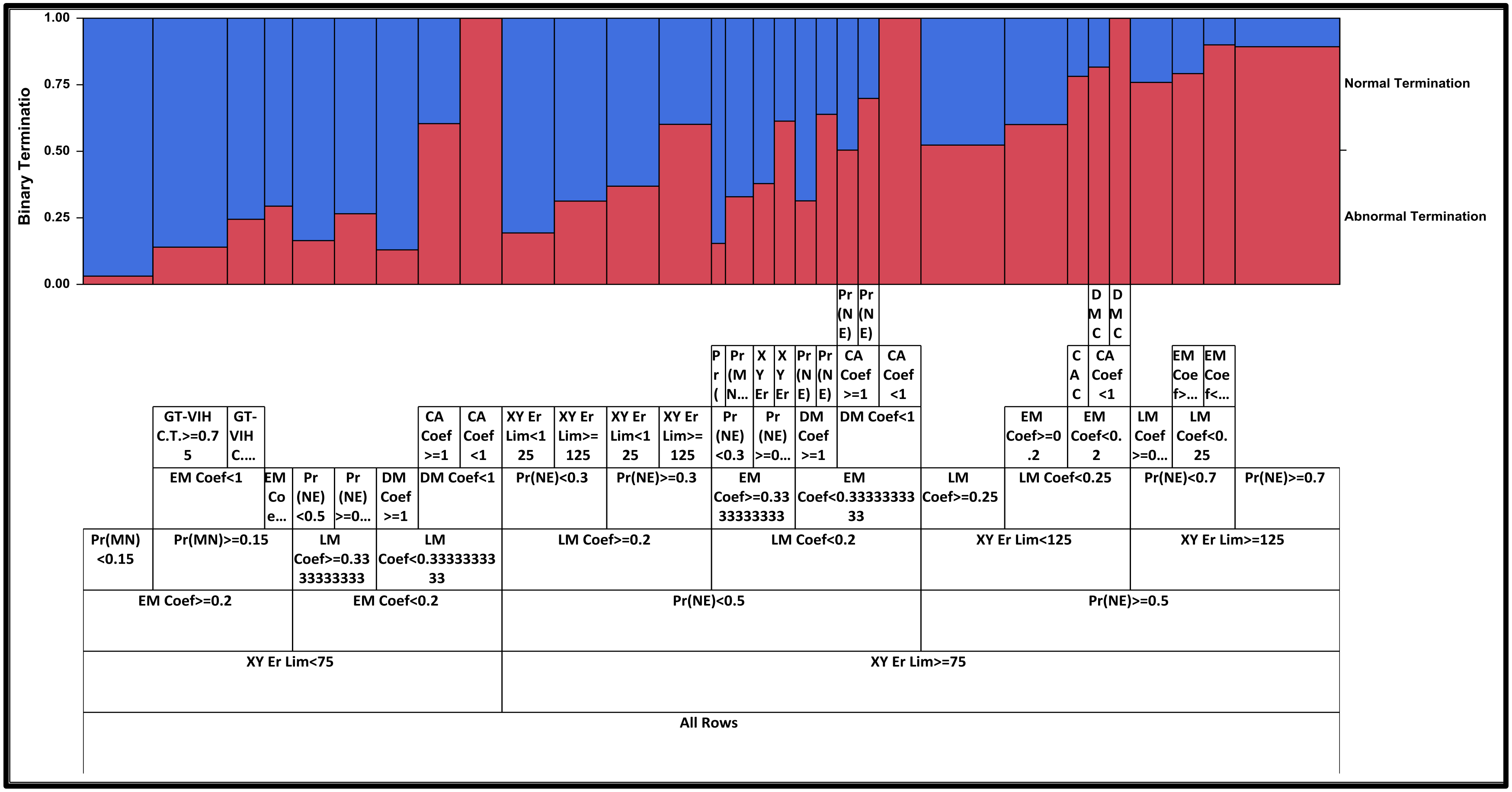


Figure 11-4 Regression Tree Graph: Partition of Binary Termination Factor Interaction Experiments by All Case Factors

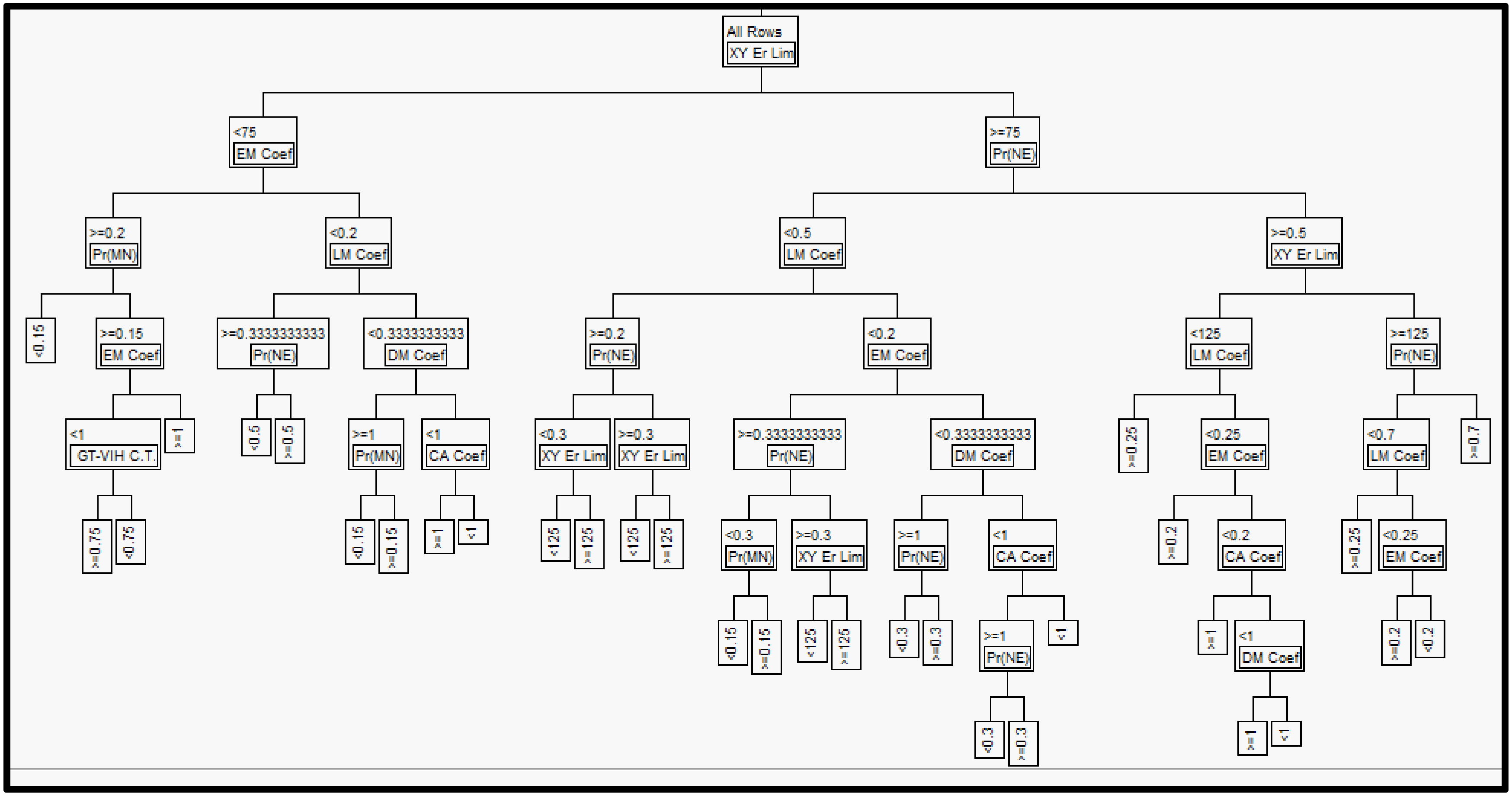


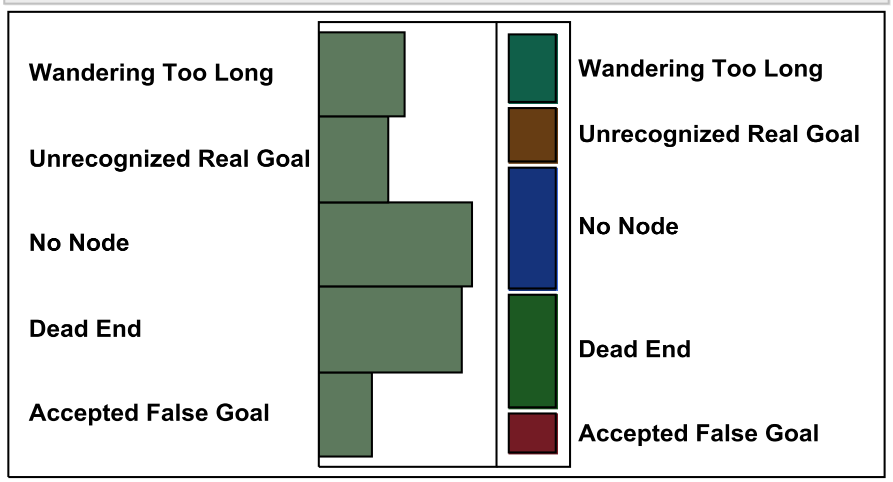
Figure 11-5 Regression Tree Small Leaf Node Graph: Partition of Binary Termination Factor Interaction Experiments by All Case Factors

11.3 Refinement of Abnormal and Normal Termination Results

As in the two previous chapters, both abnormal and normal termination results can be further broken down into categories of failure and categories of success. Results are shown in Figure 11-4 and Figure 11-5.

Distributions

Termination Condition



Frequencies

Level	Count	Prob
Accepted False Goal	8712	0.10704
Dead End	22952	0.28201
No Node	24522	0.30130
Unrecognized Real Goa	11229	0.13797
Wandering Too Long	13973	0.17168
Total	81388	1.00000

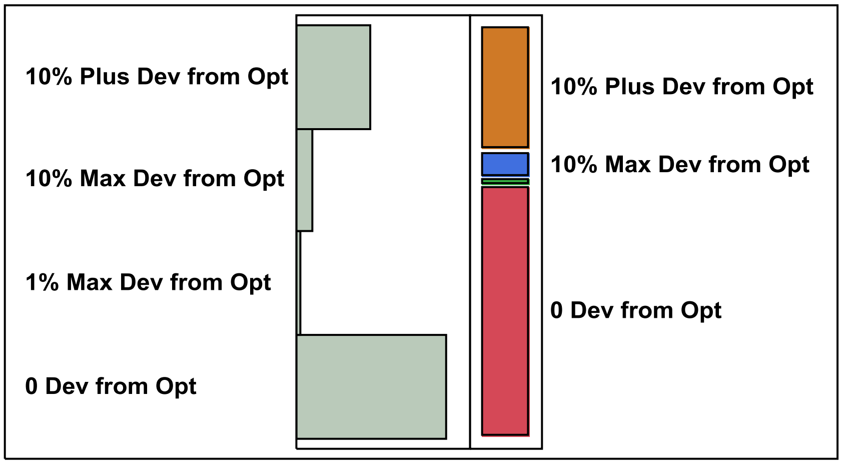
N Missin 80612

5 Level

Figure 11-6 Factor Interaction Experiments: Abnormal Termination Breakdown

Distributions

% Optimal Termination



Frequencies

Level	Count	Prob
0 Dev from Opt	49471	0.61369
1% Max Dev from Opt	1246	0.01546
10% Max Dev from Op	5388	0.06684
10% Plus Dev from Op	24507	0.30401
Total	80612	1.00000

N Missin 81388

4 Level

Figure 11-7 Factor Interaction Experiments: Contingency Analysis of Normal Termination Alternatives

11.4 <u>Summary of Factor Interaction Experiments</u>

These experiments essentially confirmed the characteristics and trends of the two previous chapters, as can be seen in Table 11-6 and Table 11-7.

Table 11-6 Three Chapter Comparison of Abnormal Termination Conditions

	Node Recognition Experiments			Spatial Distortion Experiments			MF Map Multiple Interactions	
Abnormal Termination Condition	Count	Percent of Abnormal		Count	Percent of Abnormal		Count	Percent of Abnormal
Accepted False Goal	511	9.27%		1055	5.57%		8712	10.70%
Dead End	1232	22.35%		4357	23.00%		22952	28.20%
No Node	1735	31.47%		6388	33.72%		24522	30.13%
Unrecognized Real Goal	1110	20.13%		3155	16.65%		11229	13.80%
Wandering Too Long	925	16.78%		3992	21.07%		13973	17.17%
Total Abnormal	5513	100.00%		18947	100.00%		81388	100.00%
Abnormal Percent of All Trials	47.86%			38.90%			50.24%	
Total Normal	6007			29653			80612	
Normal Percent of All Trials	52.14%			61.10%			49.76%	

Table 11-7 Three Chapter Comparison of Normal Termination Conditions

	Node Recognition Experiments			Spatial Distortion Experiments			MF Map Multiple Interactions	
Normal Termination Condition	Count	Percent of Normal		Count	Percent of Normal		Count	Percent of Normal
0 Dev from Opt	3637	60.55%		18433	62.16%		49471	61.37%
1% Max Dev from Opt	510	8.49%		2788	9.40%		1246	1.55%
10% Max Dev from Opt	238	3.96%		1092	3.68%		5388	6.68%
10% Plus Dev from Opt	1622	27.00%		7340	24.75%		24507	30.40%
Total Normal	6007	100.00%		29653	100.00%		80612	
Normal Percent of All Trials	52.14%			61.10%			49.76%	
Total Abnormal	5513			18947			81388	
Abnormal Percent of All Trials	47.86%			38.90%			50.24%	

12 META-ROUTE EXPERIMENTS

The set of experiments discussed in this chapter explore the use of the most recent update to MOBIL, a capability referred to as meta-routes. The term “meta” is used here to denote an abstraction of a concept, hence meta-routes refers to a more abstract route concept than the explicit routes that MOBIL calculates using Dijkstra’s algorithm on the arc/node structures representing GT and VIH maps. Meta-routes define a set of intermediate goals to support entity goal-seeking when a complete route is not available.

12.1 Meta-Routes

Prior experiments with the model have made the assumption that the arc node network of the entity’s VIH map is connected, i.e., that the VIH could plan its route using Dijkstra’s algorithm to find the shortest path to its goal. The meta-route concept relaxes that assumption and allows the entity to plan a route using potentially disconnected landmarks to fill “holes” in its VIH map.

Figure 12-1 illustrates the designation of a set of Multi-Feature map nodes as landmarks.

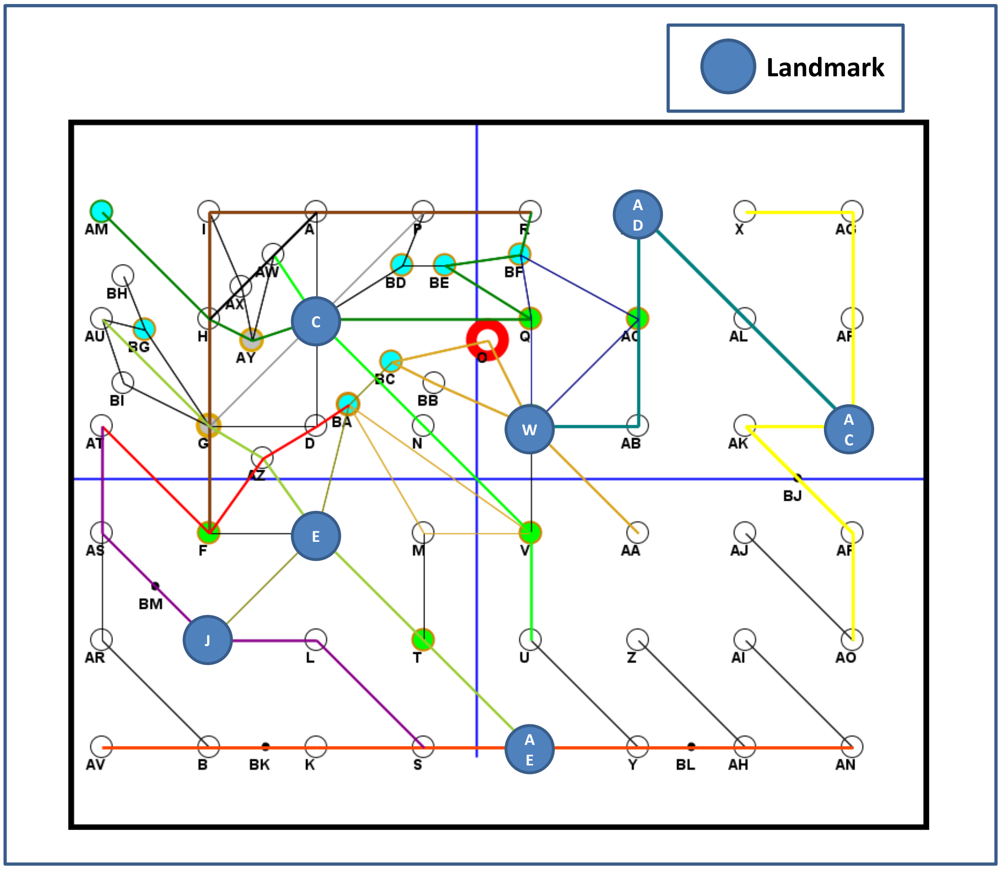


Figure 12-1Example Landmarks for the Multi-Feature Map

If a random draw on Pr(MN) divides the map's arc node network into disconnected subsets, such that the entity's start node and end goal are in different subsets, the VIH agent will plan a meta-route between the disconnected elements of the map using the landmarks. GT nodes that are designated landmarks have two features that support the meta-route concept:

- The landmarks are assumed to be present on the VIH map, i.e., they are not candidates for removal based on the result of the Monte Carlo draw for selection of missing nodes and must be present as nodes on the VIH map.
- The landmarks are assumed to be recognizable by the VIH agent, i.e., the node recognition VFT value for the landmark as a candidate VIH node is set to 1.0 when the entity is at the landmark's GT node location.

If the start node and end goal are disconnected, the entity picks the landmarks closest to the two of them and constructs a meta-route beginning at the start node and ending at the end goal via those landmarks.

MOBIL uses a landmark adjacency matrix to define what can be referred to as meta-connections. Meta-connections are not associated with actual arcs on the GT or VIH maps; they are more correctly understood as a "next neighbor" relationship. The landmark adjacency matrix is used by a meta-route specific Dijkstra algorithm to find a series of landmarks as intermediate goals from the start node to the goal node. After these intermediate goals are established, the standard Dijkstra function is applied to determine if there is an actual VIH node/link path between any two of the landmarks. If such a path exists it is integrated into the meta-route.

A formal mathematical description of the meta-route consists of a sequence of elements of the form $p_i = (s_i, n1_i \ldots nj_i \ldots e_i, \text{null})$ where s_1 is the start node of the element p_1, e_1 is the closest landmark node to s_1 , and $s_{i+1} = e_i$. The nj_i are a sequence of VIH nodes connecting s_i and e_i , if such a sequence exists or a null other wise. The presence of a null anywhere in the sequence indicates that the VIH agent must go into local search mode to find the next landmark in the sequence. The last e_i is the end node. In this way the meta-route allows the agent construct a route between connected and unconnected elements of the VIH map, with local search between the unconnected nodes.

The concept allows considerable flexibility in defining potential goal-seeking strategies for the VIH agent, in that the landmarks are not assumed to be inter-connected. For example, the nodes may be assumed to be connected to one another as shown in Figure 12-2, where they form the endpoints of a polygon. In this case, s_i, e_i elements of the meta-node elements must be neighboring polygon boundary endpoints.

It is important to note once again, that the meta-route links (highlighted as bold red lines) are not necessarily associated with actual GT or VIH arc/links (although such arc/links may or may not exist), they only indicate the meta-connections of the landmark adjacency matrix.

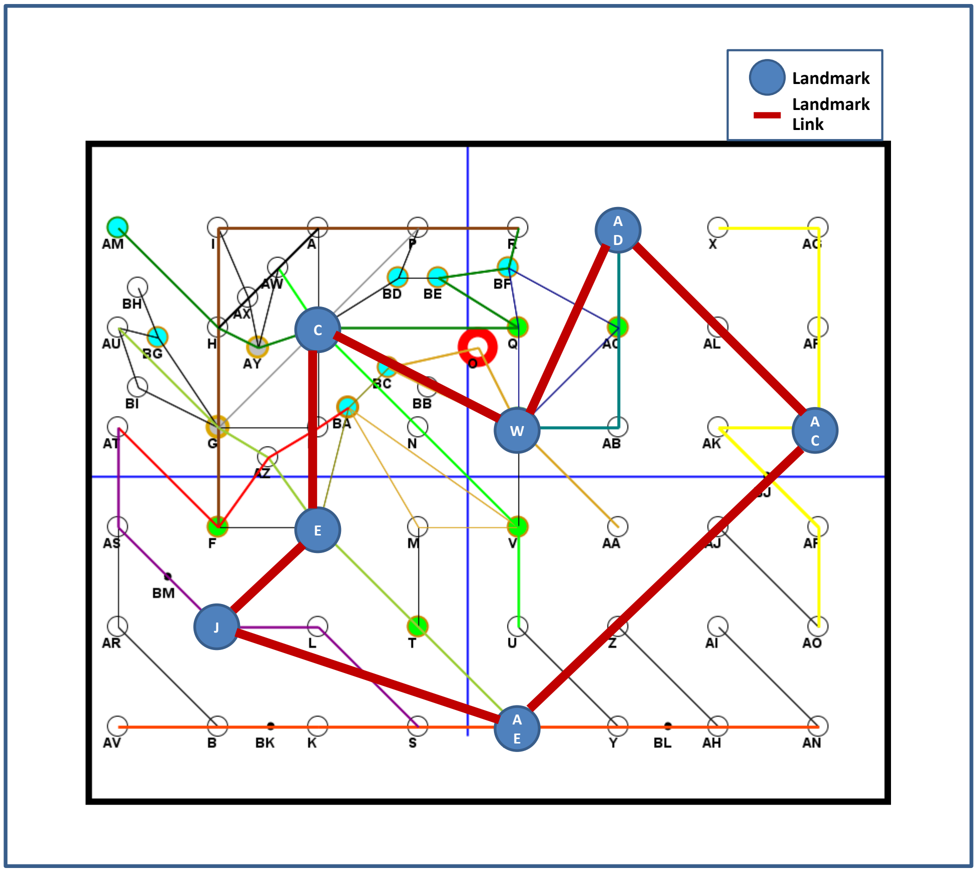


Figure 12-2 Landmarks Connected as Polygonal Endpoints

12.2 Simulation experiments using landmarks

MOBIL was run for a series of experiments using the Multi-Factor Map, the nodes shown in Figure 12-1 and three different landmark connection schemes:

1) The one illustrated in Figure 12-2, which is referred to as 2x2 links;

2) One in which none of the landmarks were assumed to be connected to one another, i.e., every entry in the landmark adjacency matrix is zero. This lack of connectivity means the VIH agent was in local search mode from the start node to the end node with none of landmarks as potential meta-route s_i, e_i element nodes. This map is referred to as start-end (SE) only.

3) One which all of the landmarks are assumed connected to each other, i.e., every entry in the landmark adjacency matrix is one. This degree of connectivity results in the agent selecting one and only one landmark as a meta-route element when

the start node and end goal are disconnected on the VIH map.

For these experiments the Pr(NE) was given three values : 0.25; 0.50; and 0.75.

The X and Y error limit was held constant at 100.

The Pr(MN) was given three values: 0.3; 0.45; and 0.6.

The GT-VI correspondence threshold was set to either 0.75 or 0.9.

All of the node recognition match coefficients were held constant at equal weights for each factor.

12.3 <u>Results: Binary Termination Outcomes</u>

Since missing nodes, start nodes, and goal nodes are selected at random, the disconnection of the VIH Map with respect to paths from the start node to the goal node is also random. Figure 12-3 shows the experimental results across all the cases, and separates out those cases where the VIH map was disconnected, i.e., those cases where the VIH agent had to rely on a Meta-route strategy to attempt to find its goal.

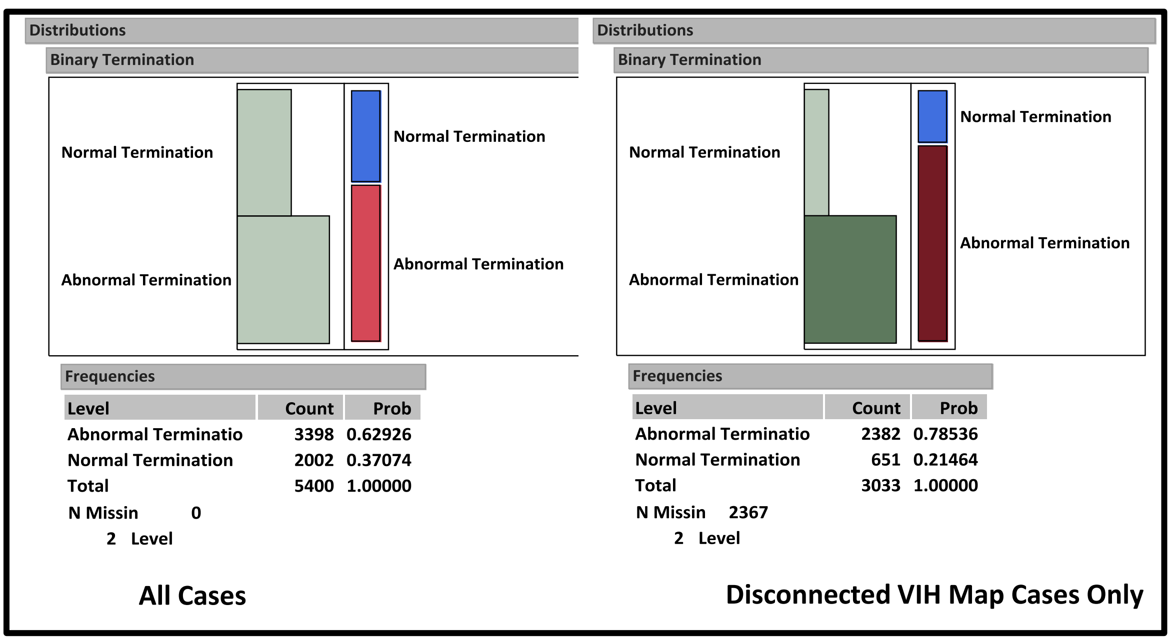


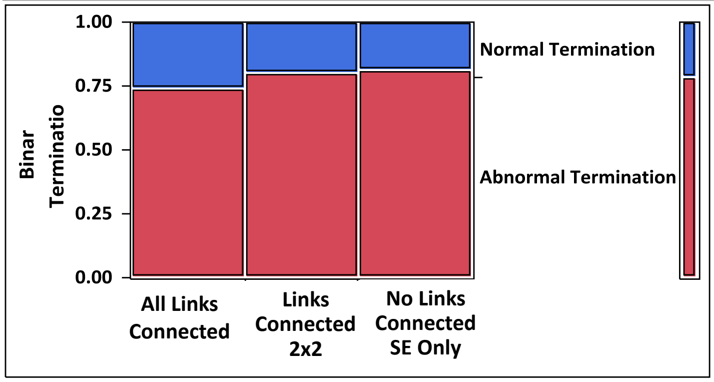
Figure 12-3 Binary Termination Results for the Meta-Routes Experiments

The disconnected cases are approximately 56% of the total. Figure 12-4 shows the effects of the three different landmark link schemes used in this set of experiments, and

Figure 12-5 shows the effects of the input parameter levels on these results.

Contingency Analysis of Binary Termination By Map

Mosaic Plot



Tests

N	DF	-LogLike	RSquare (U)
3033	2	9.4070190	0.0060

Test	ChiSquare	Prob>ChiSq
Likelihood Rati	18.814	<.0001 *
Pearson	19.178	<.0001 *

Figure 12-4 Meta-Route Results as a Function of Landmark Link Cases

VIH Map Disconnected Cases Only			All Cases		All Landmarks Inter-Connected		Landmarks Linked 2 by 2		SE Only No Landmark Links	
Pr(NE)	Pr(MN)	GT-VIH CT	Count	PerCent Normal	Count	PerCent Normal	Count	PerCent Normal	Count	PerCent Normal
0.25	0.30	0.75	25	21.19%	12	29%	10	26%	3	8%
0.25	0.30	0.90	29	23.20%	11	30%	8	17%	10	24%
0.25	0.45	0.75	33	18.75%	14	23%	8	14%	11	19%
0.25	0.45	0.90	49	30.82%	19	38%	11	22%	19	32%
0.25	0.60	0.75	65	28.63%	24	32%	19	27%	22	27%
0.25	0.60	0.90	62	29.25%	27	39%	20	26%	15	22%
0.50	0.30	0.75	14	14.00%	8	27%	3	7%	3	11%
0.50	0.30	0.90	27	21.60%	10	24%	8	18%	9	23%
0.50	0.45	0.75	29	16.57%	14	25%	7	12%	8	13%
0.50	0.45	0.90	38	20.54%	13	22%	15	25%	10	16%
0.50	0.60	0.75	55	24.89%	22	30%	18	24%	15	21%
0.50	0.60	0.90	56	27.05%	23	32%	21	28%	12	20%
0.75	0.30	0.75	16	13.79%	5	13%	4	10%	7	19%
0.75	0.30	0.90	16	13.01%	8	16%	5	16%	3	7%
0.75	0.45	0.75	33	18.13%	13	20%	10	18%	10	17%
0.75	0.45	0.90	23	0.14	10	20%	8	14%	5	9%
0.75	0.60	0.75	41	19.71%	19	26%	14	21%	8	12%
0.75	0.60	0.90	40	19.14%	13	18%	10	16%	17	23%

Figure 12-5 Normal Termination Results as a Function of Input Parameter Levels

Table 12-1 shows the JMP Whole model fit to these parameters

Table 12-1 Whole Model Fit Results

Disconnected Meta-Routes Only	
Summary of Fit	
RSquare	0.862005
RSquare Adj	0.738795
Root Mean Square Error	3.052668
Mean of Response	12.05556
Observations	54

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	25	1629.9074	65.1963	6.9962	<.0001
Error	28	260.9259	9.3188		
C. Total	53	1890.8333			

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
case id	2	2	196	10.5164	0.0004
Pr(NE)	2	2	245.7778	13.1872	<.0001
Pr(MN)	2	2	1036	55.5867	<.0001
GT-VIH CT	1	1	15.5741	1.6713	0.2067
case id*Pr(NE)	4	4	23.8889	0.6409	0.6378
case id*Pr(MN)	4	4	25.3333	0.6796	0.6119
case id*GT-VIH CT	2	2	3.7037	0.1987	0.8209
Pr(NE)*Pr(MN)	4	4	33.5556	0.9002	0.4771
Pr(NE)*GT-VIH CT	2	2	36.5926	1.9634	0.1592
Pr(MN)*GT-VIH CT	2	2	13.4815	0.7233	0.494

Case ID Letter Comparison			
Level			Least Sq Mean
all landmarks Interconnected	A		14.722222
landmarks 2x2		B	11.055556
SE only no links		B	10.388889

The landmark link scheme, Pr(NE), and Pr(MN) are significant factors, with the landmark link scheme differentiating between the case where all the landmarks are inter connected and the other two cases, which are not significantly different. The all landmarks interconnected case had a higher normal termination rate than the other two cases.

12.4 Summary of Meta-Route Experiments

These experiments show a significant, but limited, value of landmarks as intermediate goals. In the current cases, recognition factors play a larger part in the ultimate success of the goal-seeking entity, but much more research needs to be done with respect to landmarks that might divide the global map into recognizable subsets. In such a case it is likely that the landmarks would provide more information as to the nature of map locales, and thus greater guidance for entity movement.

Ultimately, the meta-route concept is designed to serve as the foundation for dividing the VIH map into zones that allow the VIH agent to use the zone boundaries as navigation aids, e.g, follow the river until you come to a bridge, or to navigate from one zone to another. The distance metric associated with meta-route specific Dijkstra algorithm can be as simple as Euclidean distance (as determined in VIH coordinate system) or it can be modified to incorporate point, line, or area characteristics that may be associated with individual landmarks to make them more desirable meta-route choices. For example, a landmark in an area where the VIH map provides a better guide to local movement (e.g., more detail in VIH node/link relationships, and more reliable data) may be a more desirable intermediate goal than a landmark where the VIH map data are sparse. In similar fashion they could reflect areas that might be attractive/ unattractive due to geo-political alignments or the risk of adversary encounters.

13 SUMMARY AND CONCLUSIONS

The Robert Frost poem on page 1 of this dissertation captures the essence of our need for SA/SU. When two roads diverge, we look down each one as far as we can, we make a choice, and we live with the consequences.

While this dissertation focuses on imperfect SA/SU, at its core it is about decisions, the information upon which those decisions are based, and the consequences of having made them. It provides a methodological framework to quantify and analyze elements of the decision process. For such a methodology to be useful in determining how to support decision makers by improving SA/SU, it must address the multiple facets of the question, "What makes a decision hard?" The list is long, and includes:

- Problem complexity
    - Does the decision-maker have to make a number of inter-related choices?
    - Are there large numbers of alternatives for each choice?
    - Is it difficult to distinguish between the alternatives?
    - Can very similar appearing alternatives have vastly different outcomes?
    - Are the characteristics of the alternatives highly dynamic?
    - Are there multiple alternatives that can lead to "good" outcomes?
    - Are there "robust" alternatives?
- Information Support
    - What is the quality of the decision-maker's initial SA/SU and what is the decision-maker's estimate of that quality?

- Does the decision-maker have sufficient information to support making "good" choices?
- How difficult is it to extract data from the environment?
- Can relevant data be distinguished/filtered from irrelevant data?
- How difficult is to understand the data that can be extracted, i.e. how hard is it to turn "data" into "information"?
- To what degree are the available data reliable and how well can the decision-maker assess that degree of reliability?
- Does the decision-maker have timely feedback from previous decisions?

* Risk/Reward Factors
  - Are there high-risk/high-reward alternatives?
  - Are risk/reward factors well known?
  - What kind of uncertainty does the decision-maker face?

This dissertation has balanced dealing with both sufficient breadth and depth of these questions to show that they are consonant with, and can be incorporated within, the methodological framework developed herein.

In so doing, the goals set forth in Chapter 0 have all been met. The human-centric paradigm has been articulated, both within the context of military MS&A and with respect to its potential for operations research outside the bounds of military application.

The software developed for the dissertation, MOBIL has shown its ability to support investigation of SA/SU issues, and the associated experiments demonstrate how it might be used to develop goal-seeking strategies based on the risk-taking levels of the goal-seeker.

Certainly, there is much left to pursue and this research addresses only a very small part of the potential of the human-centric paradigm, as should be clear from the depth and breadth of the Future Work items listed throughout the dissertation and enumerated in Appendix B below.

Throughout this effort, the problem space explored is best defined by two primary dimensions, the first of which addresses the complexity of the movement task facing an entity on an arc-node network, and the second which characterizes the quality of the information upon which that entity bases its movement decisions. The node recognition process and the degrees of VIH map distortion serve as a surrogate here for an individual's SA/SU.

As used herein the quality of information encompasses the dynamic nature of the VIH agent's world view and is consistent with Endsley's three phases of SA

While the entity represented in the present version of MOBIL has only a limited number of movement decision strategies, ultimately I would like partition the problem space according to the effectiveness of different strategies. Figure 13-1 shows a conceptual view of how such partitioning might work. In general the critical tradeoff between strategies is the degree to which their effectiveness is task and context specific, and their robustness. The first of these characterizes the extent to which the strategy provides the optimal course of action for dealing with the exact set of circumstances of performing the specific task in the given environment. The robustness criteria, on the other hand, captures the extent to which the strategy is adaptable and agile, able to find a course of action that may not be the most efficient way to perform the specific task under current conditions, but is less likely to fail if those conditions are uncertain or ambiguous, and/or should any of the task parameters change during execution.

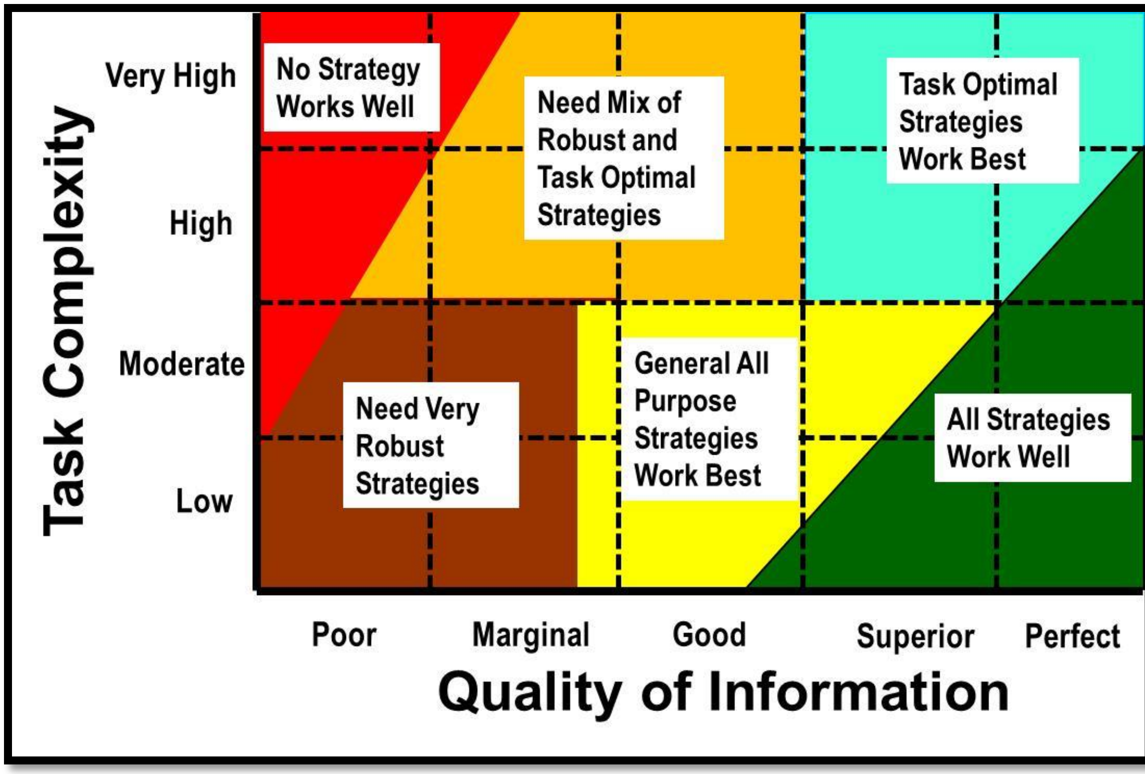


Figure 13-1 Partitioning the Problem Space

Meeting this ultimate goal requires development of a set of metrics s:

- Measures of task complexity;
- Measures of information quality;
- Measures of task performance as a function of the strategies employed.

The complexity of the movement task is a function of a number of different factors, including:

- The nature of the arc/node network;
- The dimensions of the task – the number and nature of the criteria that have to be considered in optimizing task performance; and
- The agility and adaptability required to deal with any dynamic features of either the environment or the task goals.

At present MOBIL supports only one measure of task complexity, the number of nodes in the GT defined optimal path. Figure 13-2 shows the relationship this measure and MOBIL measure of task outcomes for the experiments in Chapter12, but this is another area in need of future work.

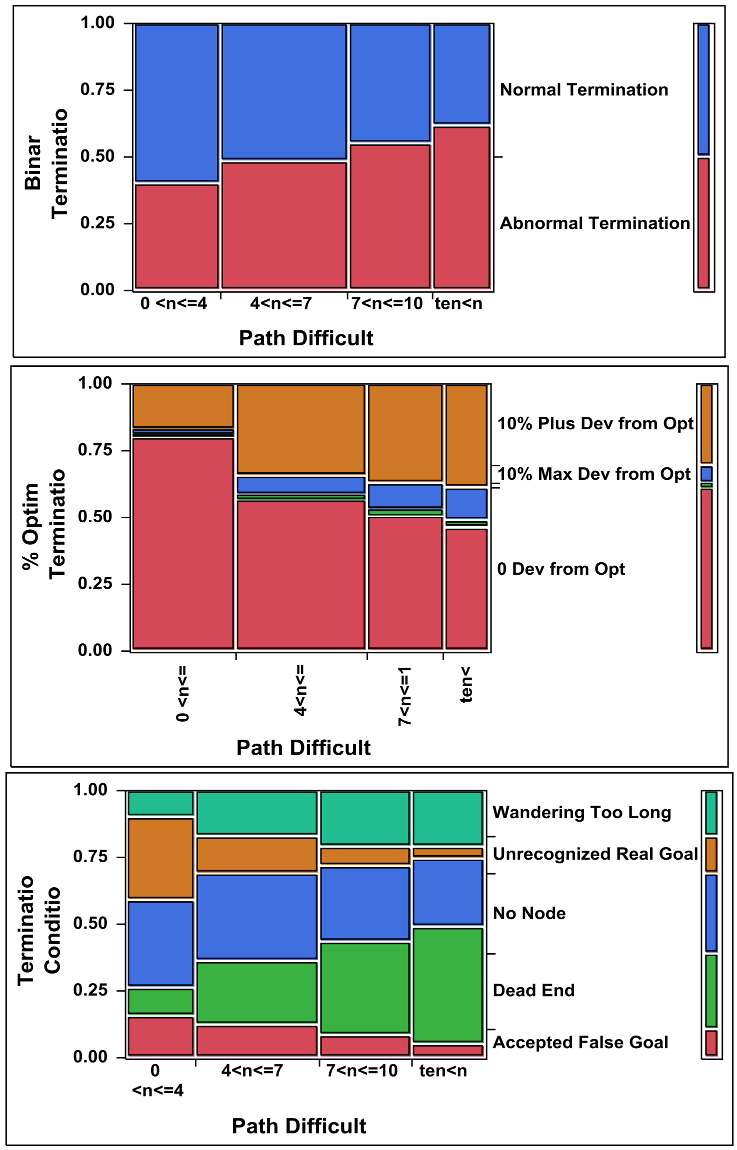


Figure 13-2 Path Difficulty as a Measure of Task Complexity

MOBIL does a better job of addressing information quality and task performance, as evidenced by the analysis of the simulation experiments in chapters 9,10,,11, and 12.

Of course, as suggested in the Future Work paragraphs, there is much work to be done in these areas as well.

In addition, the simulation methodology implemented in MOBIL will be really useful only if one can demonstrate a correspondence between the actions of simulated entities and real world behaviors, and more importantly, if the simulation can provide insight into those behaviors that supports improvement in SA/SU for real world operations.

The current version of MOBIL simulates a single entity interacting with a static environment. The true potential of the agent-based approach used in MOBIL can only be realized in a dynamic environment, with multiple entities, where agent interaction provide the possibility of the emergent behavior characteristic of complex adaptive systems.

Although MOBIL is still be limited to a significantly restricted area of the overall "human-centric" approach, it shows the viability of the human-centric paradigm and articulates the value of its application in supporting the use of modeling, simulation and analysis to address broad spectrum of problems.

APPENDIX A: MOBILE OVERVIEW

MOBILE has three principal elements: Main, GT Agent Object, and VIH Agent Object Class.

Main

The first of these is **Main**, which is the default active object class shared by all AnyLogic® programs - it handles initial model setup details and some of the model book-keeping functions. The GT arc/node network map is initialized in Main. MOBILE has a utility function derived from an earlier version of the model that supports construction of the map using the AnyLogic® graphical user interface. In such a case the utility writes the map out to a file in a format consistent with OpenStreetMaps[95].

Figure A1 shows a portion of the Excel file sheet that provides the map nodes information. The numbers in the header row define the x axis extent of the map, these values support MOBILE calculation of the x-axis offset for presentation of the VIH Map alongside the GT Map when running MOBILE in interactive presentation Mode. The subsequent numbers in these two columns represent node X and Y coordinates.

Figure A2 shows a corresponding Excel sheet for map arc/link data, which are defined in terms of OpenStreetMaps Ways. The nodes belonging to each way must be defined in the Excel node sheet.

---

[95] http://wiki.openstreetmap.org/w/index.php?title=Main_Page&oldid=941300

header row											350		-350
<node	0										-150		-250
	<tag	k=	color1r	v=	255								
	<tag	k=	color1g	v=	255								
	<tag	k=	color1b	v=	255								
	<tag	k=	color2r	v=	0								
	<tag	k=	color2g	v=	0								
	<tag	k=	color2b	v=	0								
	<tag	k=	label	v=	A								
	<tag	k=	style	v=	0								
	<tag	k=	width	v=	1								
	<tag	k=	radius	v=	10								
</node>													
<node	1										-250		250
	<tag	k=	color1r	v=	255								
	<tag	k=	color1g	v=	255								
	<tag	k=	color1b	v=	255								

Figure A1 Sample Node Input Data

</way>					
<way	id=	6			
	<nd	ref=	34		
	<nd	ref=	39		
	<tag	k=	colorR	v=	0
	<tag	k=	colorG	v=	0
	<tag	k=	colorB	v=	0
	<tag	k=	style	v=	0
	<tag	k=	width	v=	1
</way>					
<way	id=	7			
	<nd	ref=	35		
	<nd	ref=	40		
	<tag	k=	colorR	v=	0
	<tag	k=	colorG	v=	0
	<tag	k=	colorB	v=	0
	<tag	k=	style	v=	0
	<tag	k=	width	v=	1
</way>					
<way	id=	8			
	<nd	ref=	25		
	<nd	ref=	33		
	<tag	k=	colorR	v=	0
	<tag	k=	colorG	v=	0

Figure A2 Sample Arc/LInke Input Data

GT Agent Object Class

Figure A3 shows the logic for the GT Agent class, which is defined in detail above in Chapter 8.5.1.

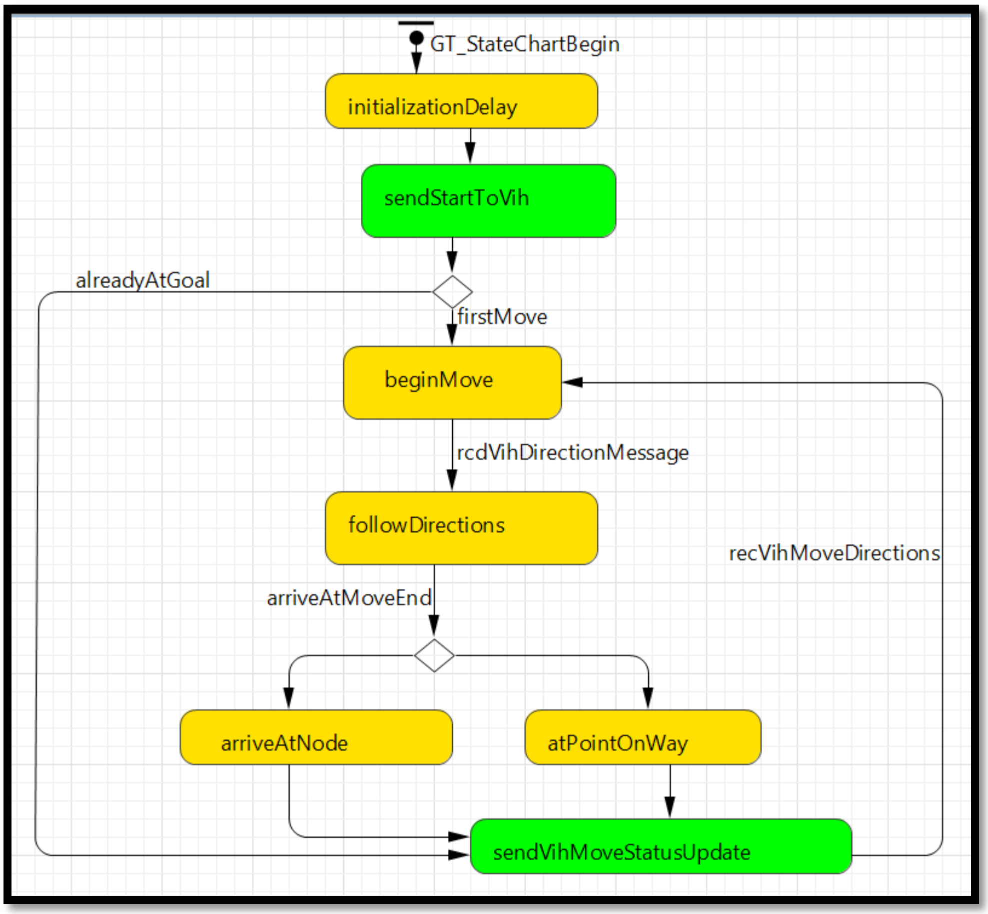


Figure A3 GT Agent Logic

VIH Agent Object Class

Figures A4 and A5 show the logic for the VIH agent, which is described in more detail above in Chapter 8.5.2.

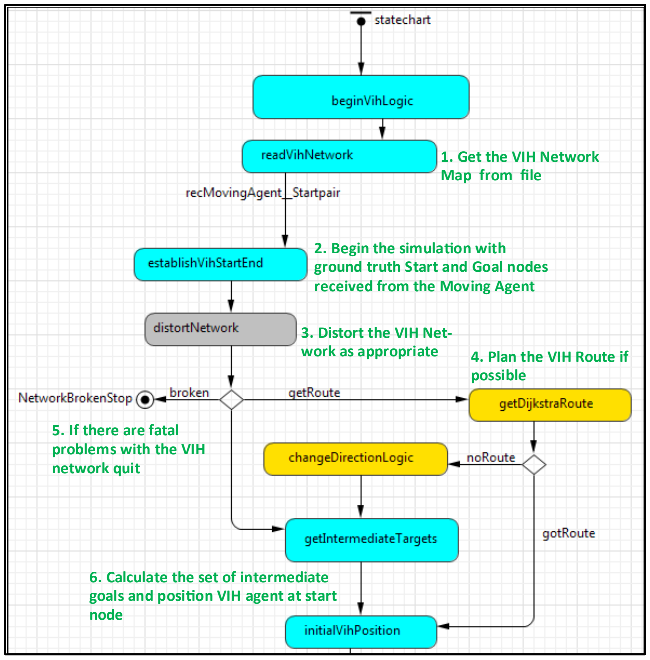


Figure A4 Initialization of the VIH Agent

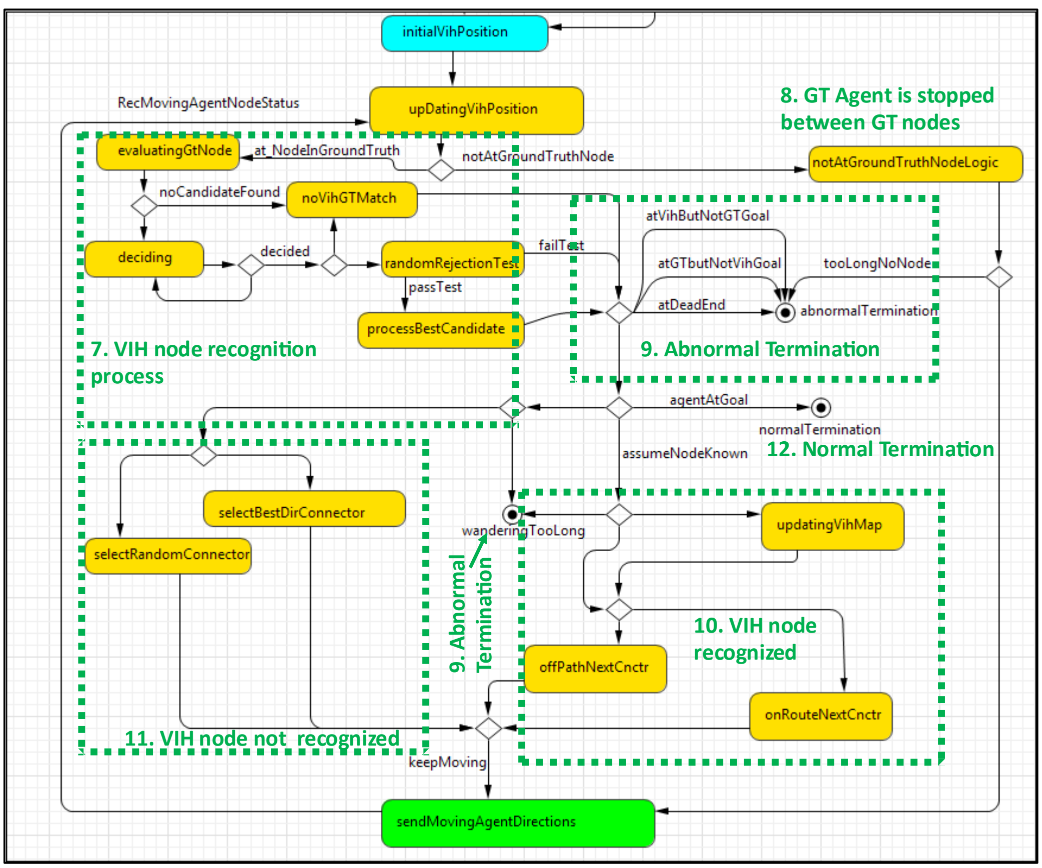


Figure A5 VIH Agent Decision Logic

Model Input/Output

Model Output for interactive simulation runs is described above in Chapter 9.5.

Table A1 Principle Input Parameters

Map Identifier
Probability of Node Error
X Error Limit
Y Error Limit
Direction Logic
Nodes Without a Goal Limit
Node Recognition Error Probability
VIH Candidate Distance
Keep Searching Upper Limit
Probability of Missing Nodes
Color Standard Deviation (for distortion calculations)
Color Distortion Upper Limit
Location Distance Upper Limit
Distance Match Coefficient
Expectation Match Coefficient
Links Match Coefficient
Color Attribute 1 Match Coefficient
Color Attribute2 Match Coefficient
GT-VIH Correspondence Threshold
Update VIH Map Threshold
Auto Update Node Logic Flag
Delta Good Fit Threshold

Table A2 Principal Model Outputs (for Multiple Trial Parameter Experiments)

Observed Termination Condition
MA Start Node
MA End Goal
run statistics
Number of Nodes in GT Dijkstra Path
Number of Nodes in VIH Dijkstra Path
Meta Route Used Flag
Number of GT Nodes Visited
Number of VIH Nodes Visited
GT Dijkstra Path Length
VIH Dijskstra Path Length
Delta VIH Dijkstra - GT Dijkstra
Total Distance travelled
Delta Distance From True Optimum (GT Dijkstra Path Length)
Fractional Delta Distance From Optimum
Delta Distance From VIH Optimum (VIH Dijkstra Path Length)
Fractional Delta Distance From VIH Optimum
Nodes On VIH Path Count
Nodes Off VIH Path Count
Maximum Off VIH Path Streak
Number of False Path Links (due to reconnections for missing nodes)
Node Recognition Decision Counter
Average Node Recognition Confidence
Number of Node Recognition Failures
Number of True Positive Node Recognitions
Number Recognition Decisions for which there is No Node Choice
Number of False Positive Node Recognitions
Number of False Negative Node Recognitions
Number of True Negative Node Recognitions
Percent of Node Recognition Failures
Percent of True Positive Node Recognitions
Percent of False Positive Node Recognitions
Number of Correctly Rejected Updates
Number of Intra-Path Decisions
Landmarks On Flag

The model documentation generated by AnyLogic for MOBIL can be found at:

http://cecs.wright.edu/~fciarall/Middleton/Middleton.Dissertation.Appendices.pdf

APPENDIX B: AREAS FOR FUTURE RESEARCH

Throughout this document a number of areas for future research have been identified. This appendix provides a list of those areas and the pages on which they may be found.

REFERENCES

Adam, E. C. (1993). Fighter cockpits of the future. The 12th IEEE/AIAA Digital Avionics Systems Conference**:** (pp. 318-323).

Adam, G. E., R. Carter, et al. (2008). "Hydration effects on cognitive performance during military tasks in temperate and cold environments." Physiology & Behavior, 93(4-5), 748-756. **93**.

Alberts, D. S. (2002). Code of best practice: experimentation. Information Age Transformation Series Washington D.C., Office of the Assistant Secretary of Defense Command andControl Research Program (CCRP).

Alberts, D. S. and J. J. Garstka (2001). Network Centric Warfare Department of Defense Report to Congress. Washington, DC. July.

Alberts, D. S., J. J. Garstka, et al. (2001). Understanding Information Age Warfare. Washington D.C., The Command and Control Research Program (CCRP) , Office of the Secretary of Defense (OSD)

Alex, C. A. (2000). Process and Procedure: The Tactical Decision-making Process and Decision Point Tactics. Master of Military Art and Sceince, Naval Postgraduate School.

Alt, J. (2006). Exploring Tactics, Techniques and Procedures for a Future Force Warrior Small Combat Unit. Master's, Naval Postgraduate School.

Alt, J. K., L. A. J. Jackson, et al. (2009). "The Cultural Geography Model: Evaluating the Impact of Tactical Operational Outcomes on a Civilian Population in an Irregular Warfare Environment." The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology **6**(4): 185.

Anderson, C. (2008). “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”. WIRED MAGAZINE. **16.07**.

Armstrong, J. S. (2012). "Predicting Job Performance: The Moneyball Factor." Foresight: The International Journal of Applied Forecasting(25).

Artelli, M. and R. Deckro (2008). "Modeling the Lanchester Laws with System Dynamics." The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology **5**(1): 1.

Axtell, R. (2000). Why agents? On the varied motivations for agent computing in the social sciences. Workshop on agent simulation: applications, models and tools, Citeseer.

Bachman, J. T., L. R. Hester, et al. (2008). Analyzing the Marine Corps Infantry Squad Using the Infantry Warrior Simulation (IWARS) 2008 Spring Simulation Interoperability Workshop. Westin Hotel - Providence, Rhode Island.

Bailey, K. (2012). Capability Set 13: Mission Command Technologies Providing Web-based Simplicity. Fort Gordon,GA,30905-5301, U.S. Army Signal Center of Excellence,Army Communicator.

Banks, S. B. and M. R. Stytz (1999). An Approach to Enhance Human Behavior Modeling for Computer-Generated Actors. Advancing Simulation Technology and Training The Simulation Technology and Training (SimTecT99) Conference, Carlton Crest Hotel, Melbourne.

Batschelet, A. (2002). Effects-Based Operations: A New Operational Model?, ARMY WAR COLL CARLISLE BARRACKS PA.

Belanich, J., D. E. Sibley, et al. (2004). Instructional characteristics and motivational features of a PC-based game. Alexandria VA, United States Army Research Institute fro the Behavioral and Social Sciences

Bellman, R. and L. Zadeh (1970). "Decision-Making in a Fuzzy Environment." Management Science **17**(4).

Benford, S., R. Anastasi, et al. (2003). "Coping with uncertainty in a location-based game." Pervasive Computing, IEEE **2**(3): 34-41.

Bernstein, R. (1998). A Road Map for Simulation Based Acquisition: Report of The Joint Simulation Based Acquisition Task Force Director, Test, Systems Engineering, and Evaluation (DTSE&E).

Betz, D. (2007). "Redesigning Land Forces for Wars Amongst the People." Contemporary Security Policy **28**(2): 221-243.

Bitinas, E., Z. Henscheid, et al. (2003). "Pythagoras: A new agent-based simulation system." Technology Review: 45.

Blondel, V. D., A. Gajardo, et al. (2004). "A measure of similarity between graph vertices: Applications to synonym extraction and web searching." SIAM Review: 647-666.

Bonabeau, E. (2002). "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Sciences of the United States of America **99**(Suppl 3): 7280.

Bonchev, D. and G. A. Buck (2005). Quantitative measures of network complexity. Complexity in Chemistry, Biology, and Ecology**:** 191-235.

Bowley, D., T. Castles, et al. (2003). "Constructing a SUITE of Analytical Tools: A Case Study of Military Experimentation." ASOR BULLETIN **22**(4): 2-10.

Bowley, D., T. Castles, et al. (2004). Attrition and suppression: defining the nature of close combat. Edinburgh South Australia Australian Defence Science and Technology Organisation (DSTO) Information Sciences Laboratory.

Boyd, J. (1986). "Patterns of conflict." unpublished briefing, Defence and the National Interest, available at: www. dni. net/second_level/boyd_military. htm (accessed 14 November 2004).

Boyd, J. (1987). A discourse on winning and losing. Maxwell AFB, Ala.: Air University Library, Document No. M-U43947, unpublished collection of briefing slides.

Boylan, G. L. and S. R. Goerger (2006). PEO Soldier simulation roadmap: Continued efforts in implementation. West Point NY, United States Military Academy Operations Research Center of Excellence

Brady, E. and S. Starr (2002). Re-Assessing Dismounted Operations in Complex Terrain using the NATO CoBP for C2 Assessment. RTO SAS Symposium on "Analysis of the Military Effectiveness of Future C2 Concepts and Systems". NC3A, The Hague, The Netherlands, RTO-MP-117.

Breton, R. and R. Rousseau (2001). Situation Awareness: A review of the concept and its measurement, DRDC Valcartier TR. **220**.

Brock, D. L., D. J. Montana, et al. (1992). Coordination and control of multiple autonomous vehicles. IEEE International Conference on Robotics and Automation, Nice, IEEE.

Bryant, M., P. Johnson, et al. (2008). Layered Sensing: Its Definition, Attributes, and Guiding Principles for AFRL Strategic Technology Development, Air Force Research Laboratory Wright-Patterson Air Force Base OH.

Byrne, M. D. (2003). Cognitive architecture. The human-computer interaction handbook. A. J. Julie and S. Andrew, L. Erlbaum Associates Inc.**:** 97-117.

Cacciabue, P. C., F. Decortis, et al. (1992). "COSIMO: a cognitive simulation model of human decision making and behavior in accident management of complex plants." Systems, Man and Cybernetics, IEEE Transactions on **22**(5): 1058-1074.

Cain, B. and A. Belyavin (2008). Inclusion of Human Behaviour Representation in military modelling and simulation. MSG-060 Symposium on "How is Modelling and Simulation Meeting the Defence Challenges out to 2015?" Vancouver, British Columbia 17.

CERDEC (2005). Acquisition Level Definitions and Observables for Human Targets, Urban Operations, and the Global War on Terrorism. Fort Belvoir, VA 22060, US ARMY RDECOM Communications-Electronics Research, Development, and Engineering Center, Night Vision and Electronic Sensors Directorate AMSRD-CER-NV-MSD

Chandrasekaran, B. (1994). "Functional representation and causal processes." Advances in computers **38**(73-143).

Clemen, R. T. and T. Reilly (2001). Making Hard Decisions with Decision Tools®, Duxbury Thomson Learning.

Cohen, I. (2008). "Improving Time-Critical Decision Making in Life-Threatening Situations: Observations and Insights." Decision Analysis **5**(2): 100-110.

Corp, N. G. S. M. S. (2008). Pythagoras 2.0 Alpha Version

Corps., U. S. M. (1996). Operational Maneuver From the Sea. United States Marine Corps Emerging Operational Concepts Maritime Prepositioning Force 2010 and Beyond. Quantico VA, US Marine Corps Combat Development Command.

Cosby, L. N. (1995). SIMNET: An Insider's Perspective. Alexandria VA, Institute for Defense Analyses.

Curtis, N. (2000). Planning for the next generation of soldier modernisation. Land Warfare Conference. Melbourne. Melbourne Australia**:** 314ñ327.

Darken, C. J. (2007). "Computer Graphics-Based Models Of Target Detection: Algorithms, Comparison To Human Performance, And Failure Modes." The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology **4**(Jul 2007): 262-276.

Darken, C. J. J., Brian E. (2007). Computer Graphics-Based Target Detection for Synthetic Soldiers. 16th Conference on Behavior Representation in Modeling and Simulation (BRIMS), Citeseer.

Darley, W. M. (2006). Clausewitz's Theory of War and Information Operations, Army Combined Arms Center Fort Leavenworth KS

Davis, C. (2007). Integration of life cycle analysis within agent based modeling using a case study on bio-electricity. MSc Masters, Institute of Environmental Science at Leiden University and the Energy and Industry Group at the Delft University of Technology.

Davis, J. (1999). Simulation Based Acquisition; A State of the Art Report. Huntsville AL, Defense Modeling Simulation and Tactical Technology Information Analysis Center

Davis, P. K. and R. H. Anderson (2004). "Improving the composability of DoD models and simulations." The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology **1**(1): 5-17.

Davis, R., H. Shrobe, et al. (1993). "What is a knowledge representation?" AI magazine **14**(1): 17.

Deavel, C. R. P. (1998). "The Political Economy of Privatization for the American Military." Air Force Journal of Logistics (AFJL) **XXII**(Number 2, Summer 199): 5-11.

Deitchman, S. (1962). "A Lanchester Model of Guerrilla Warfare ".

Deitz, P. (2006). Decision Support Tools Mission Means Framework and the Advanced Decision Architectures (ADA) Collaboative Technology Alliance (CTA). Warrior Systems Tech Base Executive Committee Conference Natick Soldier Center, MA, Army Materiel Systems Analysis Activity (AMSAA) Aberdeen Proving Ground MD

Deptula, D. (2001). "Effects-Based Operations: Change in the Nature of Warfare." Aerospace Education Foundation, Arlington, VA.

Dubois, R. D. J. M., Robert J. Ph.D. (2000). Conceptual Modeling of Human Behavior (CMHB). Ninth Conference on Computer Generated Forces and Behavior Representation. Orlando Florida.

Duckham, M., J. Lingham, et al. (2006). "Qualitative reasoning about consistency in geographic information." Information Sciences **176**(6): 601-627.

Duckham, M., K. Mason, et al. (2000). A Formal Ontological Approach to Imperfectionin Geographic Information GIS Research UK (GISRUK) 2000.

Dyer, J. L. (2009). Annotated Bibliography of the Army Research Institute's Training Research Supporting the Land Warrior and Ground Soldier Systems: 1998-2009, U.S. Army

Research Institute for the Behavioral and Social Sciences fort Benning GA.

Dyer, J. L., R. L. Wampler, et al. (2005). After action reviews with the ground Soldier system, United States Army Research Institute for the Behavioral And Social Sciences Fort Benning GA.

Eiben, A. E. and J. E. Smith (2003). Introduction to evolutionary computing, Springer Verlag.

Eisenhower, D. D. (1961). "The Eisenhower Farewell." The Bulletin of the Atomic Scientists **17**(3): 90-91.

Endsley, M. (1995a). "Toward a theory of situation awareness in dynamic systems." Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(1): 32-64.

Endsley, M. (1995b). "Measurement of situation awareness in dynamic systems." Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(1): 65-84.

Endsley, M. R. (1999). Situation awareness and human error: Designing to support human performance. The High Consequence Systems Surety Conference Albuquerque, NM

Fayyad, U., G. Piatetsky-Shapiro, et al. (1996). "From data mining to knowledge discovery in databases." AI magazine **17**(3): 37.

Foo, K. (2008). Exploring First Responder Tactics in a Terrorist Chemical Attack. Master of Science in Operations Research Naval Postgraduate School

Ford, D. N. and J. T. Dillard (2008). Modeling the integration of open systems and evolutionary acquisition in DoD programs. 5th Annual Acquisition Research Symposium: Creating Synergy for Informed Change, Monterey, CA, Naval Postgraduate School,Graduate School of Business and Public Policy.

Forsythe, C. X., PG (2002). Human emulation: progress towards realistic synthetic human agents. 11 th Conference on Computer-Generated Forces and Behavior Representation, Orlando FL, SISO.

Franklin, S. (1995). Artificial Minds. Cambridge MA, The MIT Press.

Freed, M. and R. Remington (2000). Making human-machine system simulation a practical engineering tool: An APEX overview. Proceedings of the 3rd International Conference on Cognitive Modelling, Citeseer.

Freed, M. A., M. G. Shafto, et al. (1998). Employing simulation to evaluate designs: The APEX approach. The 2nd Workshop on Human Error, Safety, and System Development.

Frese, M. and A. Altman (1989). The treatment of errors in learning and transfer. IN L. Bainbridge & SA Quintanilla. Developing skills with new technology. L. Bainbridge and S. A. R. Quintanilla. Chichester, UK:, Wiley**:** 65- 86.

Furnas, A. (2012). Everything You Wanted to Know About Data Mining but Were Afraid to Ask. The Atlantic.

Gahegan, M. and I. Leeb (2000). "Data Structures and Algorithms to Support Interactive Spatial Analysis Using dynamic Voronoi Diagrams." Computers, Environment and Urban Systems **24**(6): 509-537

Ghazal, T., R. Morley, et al. (2003). Soldier tactical mission system combat effectiveness.

Systems and Information Engineering Design Symposium, 2003 IEEE, IEEE.

Gilbert, N. and P. Terna (2000). "How to build and use agent-based models in social science." Mind & Society **1**(1): 57-72.

Gillis, D. (2008). Analysis of the Operational Effect of the Joint Chemical Agent Detector Using the Infantry Warrior Simulation (IWARS). Military Operation Research Society (MORS) Sysposium. New London, Connecticut

Gillis, P. and S. Hursh (1999). Using behavior moderators to influence cgf command entity effectiveness and performance. Eigth Conference on Computer Generated Forces and Behavioral Representation, Orlando FL, SISO.

Gleick, J. (1987). "Chaos: Making a new science." New York: Viking Penguin **8**.

Gluck, K. and R. Pew (2005). Modeling human behavior with integrated cognitive architectures: Comparison, evaluation, and validation, Lawrence Erlbaum Assoc Inc.

Goerger, S. R. (2002). Validation and Evaluation of Cognitive Models for Combat Simulations. Doctor of Philosophy in MOVES Dissertation, Naval Postgraduate School.

Gold, C. (1992). "The meaning of "Neighbour"." Theories and methods of spatio-temporal reasoning in geographic space: 220-235.

Graham, S., M. Matthews, et al. (1998). Infantry Situation Awareness. Papers from the 1998 Infantry Situation Awareness Workshop. Alexandria VA, Army Research Institute for the Behavioral and Social Sciences

Gratch, J. and S. Marsella (2003). "Fight the Way You Train: The Role and Limits of Emotions in Training for Combat." Brown Journal of World Affairs **10**(1): 63.

Hahn, R. and B. Jezior (1999). "Urban Warfare and the Urban Warfighter of 2025." Parameters **29**: 74-86.

Haley, D. R. L., Chairman; Dr. Joyce Shields, Vice Chair; Dr. Crystal C. Campbell; Dr. Gerald D. Godclen; Mr. Marvin R. Hotter;Dr. Walter B. LaBerge; General James G. Lindsay (USA-Ret); Mr. Charles L. Malone; Dr. Bruce Montgomery; Dr. Edward J. Powers, Jr.; Dr. Robert E. Weigle; Dr. Stanley C. White (1991). Army Science Board Summer 1991 Study: SOLDIER AS A SYSTEM.

Hall, N. R. and S. Preiser (1984). "Combined network complexity measures." IBM journal of research and development **28**(1): 15-27.

Han, J. and M. Kamber (2006). Data mining: concepts and techniques, Morgan kaufmann.

Hare, M. and P. Deadman (2004). "Further towards a taxonomy of agent-based simulation models in environmental management." Mathematics and Computers in Simulation **64**(1): 25-40.

HassgÂrd, U. (2002). The Lowest Echelon in Network Centric Warfare Master's (Military Equivalent), Swedish National Defense College

Hayes, B. (2000). "Graph theory in practice: Part I." American Scientist **88**(1): 9ñ13.

Heeringa, B. and P. Cohen (2000). An underlying model for defeat mechanisms. Winter Simulation Conference, American Association for Artificial Intelligence (www.aaai.org)

Heinze, C. L., I; Goss, S; Pearce, A (1999). Collaborating cognitive and sub-cognitive processes for the simulation of human decision making. 4th International SIMTECT Conference, Citeseer.

Henricksen, K. and J. Indulska (2004). Modelling and Using Imperfect Context Information. The Second Annual Conference on Pervasive Computing and Communications, Workshop on Context Modelling and Reasoning (PERCOM 2004), Orlando FL, The IEEE Computer Society.

Hewawasam, K. K. R., K. Premaratne, et al. (2007). "Rule Mining and Classification in a Situation Assessment Application: A Belief-Theoretic Approach for Handling Data Imperfections." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on **37**(6): 1446-1459.

Hill, R. W. J. (1999). Modeling Perceptual Attention in Virtual Humans. Eighth Annual Conference on Computer Generated Forces and Behavioral Representation. Orlando FL, Simulation Interoperability Standards Organization (SISO).

Hobbs, W., D. Goyne, et al. (2000). LAND 125 Soldier Combat System (Project WUNDURRA) Australiaís Soldier Modernisation Program: Past, Present and Future. SMI Conference on Next Generation Soldier Technology [Proc], London.

Holland, J. (1992). "Complex adaptive systems." Daedalus **121**(1): 17-30.

Holland, J. (2006). "Studying Complex Adaptive Systems." Jrl Syst Sci & Complexity (2006) 19: 1–8

Holmquist, J. and S. Goldberg (2007). Dynamic Situations: The Soldier's Situation Awareness, NATO RTO-TR-HFM-121-Part-II.

Hooker, J. N. (1994). "Needed: An empirical science of algorithms." Operations Research **42**(2): 201-212.

Horne, G. E. S., Klaus-Peter (2008). Data Farming Around the World Overview. 2008 Winter Simulation Conference

Horne, G. L., Mary, Ed. (2001). Maneuver Warfare Science 2001 Defense Automated Printing Service 1001 Barnett Ave Quantico, VA 22134-5007 USA

Housson, J.-P. (2008). "Soldier Modernisation: European Overwatch." from http://www.soldiermod.com/summer-08/features-europe.html.

Howard, R. A. (2007). "Foundations of Decision Analysis Revisited." Advances in Decision Analysis from Foundations to Applications: 32-56.

Hudlicka, E. (2002). Architecture and Representation Requirements for Modeling Effects of Behavior Moderators. Eleventh Conference on Computer Generated Forces and Behavior Representation (CGF-BR 2002), Orlando FL, SISO.

Hudlicka, E. (2002). "This Time with Feeling: Integrated Model of Trait and State Effects on Cognition and Behavior." Applied Artificial Intelligence **16**(7/8): 611-641.

Hudlicka, E. (2004). Beyond cognition: Modeling emotion in cognitive architectures. Proceedings of the Mahwah NJ, Lawrence Earlbaum.

Hurley, W. J., C. R. Bucher, et al. (2009). Non-Kinetic Capabilities for Irregular Warfare: Four

Case Studies. Alexandria VA, Institiute for Defense Analyses.

Ilachinski, A. (1996). Land Warfare and Complexity, Part I: Mathematical Background and Technical Sourcebook (U), Center for Naval Analyses.

Ilachinski, A. (1997). Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-Life Approach to Land Warfare (U), Center for Naval Analyses.

Ilachinski, A. (2004). Artificial war: Multiagent-based simulation of combat, World Scientific Pub Co Inc.

Ioerger, T. R. V., Richard A.; Yen, John (2000). Modeling Cooperative, Reactive Behaviors on the Battlefield with Intelligent Agents. Ninth Conference on Computer Generated Forces and Behavior Representation. Orlando Florida.

Ipekci, A. (2002). How Agent Based Models Can Be Utilized to Explore and Exploit Non-linearity and Intangibles Inherent in Guerrilla Warfare. Master of Science in Operations Research, Naval Postgraduate School.

James, R. and T. Daniels (2005). Effects Based Operations (EBO) EndState. Rome NY, United States Air Force Research Laboratory Information Directorate Rome Research Site 102.

Johnson, M. V. M., Mark F. ; Szanto, Terence R. (1998). Simulation Based Acquisition: A New Approach. Fort Belvoir VA, Defense Systems Management College

Johnson, S. (2001). Emergence: The connected lives of ants, brains, cities, and software, Scribner.

Jones, R. H., AE; Chown, E (2002). Interfacing emotional behavior moderators with intelligent synthetic forces. Eleventh Conference of Computer Generated Forces and Behavior Representation, Orlando FL, SISO.

Jones, R. M., B. S. Stensrud, et al. (2009). Knowledge-Rich Intelligent Agents for SOF Training. Special Operations Forces Industry Conference (SOFIC) 2009

Kaimann, R. A. (1974). "Coefficient of network complexity." Management Science **21**(2): 172-177.

Kauffman, S. (1993). The origins of order: Self organization and selection in evolution, Oxford University Press, USA.

Keeney, R. (1996). Value-focused thinking: A path to creative decisionmaking, Harvard Univ Pr.

Kim, S. H. and R. Bhattacharya (2007). Multi-Layer Approach for Motion Planning in Obstacle Rich Environments. American Institute for Aeronautics and Astronautics (AIAA) Guidance, Navigation and Control Conference and Exhibit Hilton Head, South Carolina.

Kleijnen, J. P. C., S. M. Sanchez, et al. (2005). "A userís guide to the brave new world of designing simulation experiments." INFORMS Journal on Computing **17**(3): 263-289.

Klein, G. (1999). Sources of power: How people make decisions, The MIT Press.

Klein, G. (2008). "Naturalistic Decision Making." Human Factors: The Journal of the Human Factors and Ergonomics Society **50**(3): 456-460.

Kleiner, M., S. Carey, et al. (2007). Geospatial Battle Management Language: Bridging GIS, C2 and Simulations. ESRI users conference.

Klippel, A., M. Worboys, et al. (2008). "Identifying factors of geographic event conceptualisation." International Journal of Geographical Information Science **22**(2): 183 - 204.

Lampe, T., G. Schwarz, et al. (2006). "PAX: Designed for Peace Support Operations." Scythe Issue **1**(2): 43-48.

Langley, P., J. E. Laird, et al. (2009). "Cognitive architectures: Research issues and challenges." Cognitive Systems Research **10**(2): 141-160.

Lauren, M. (2006). Characterising the difference between complex adaptive and conventional combat models. Arxiv preprint nlin/0610035. Auckland, New Zealand Defence Operational Support Technology Establishment.

Lee, D. B. L. C. U. (1990). "War Gaming." Airpower Journal - Summer 1990.

Leeuw, M. (1997). Stand van zaken Focusprogramma Soldier Modernisation (Progress Report Focus Programme Soldier Modernisation). Rijswijk (Netherlands), Prins Maurits Laboratorium TNO

Liang, L. (2005). The Use of Agent Based Simulation for Cooperative Sensing of the Battlefield. Master of Science in operations Research, Naval Postgraduate School

Lind, W. (1985). Maneuver warfare handbook. Boulder CO 80301, Westview Press.

Lipshitz, R., G. Klein, et al. (2001). "Focus article: Taking stock of naturalistic decision making." Journal of Behavioral Decision Making **14**(5): 331-352.

Lipshitz, R. and O. Strauss (1997). "Coping with uncertainty: A naturalistic decision-making analysis." Organizational Behavior and Human Decision Processes **69**(2): 149-163.

Little, D. (2006). History and Basics of M&S. Washington D.C., United States Army Intelligence Washington D.C. Battle Command Simulation and Experimentation.

Livermore, W. R. (1879). The American Kriegsspiel. A game for practicing the art of war upon a topographical map, Houghton, Mifflin and Co.

Llinas, J., C. Bowman, et al. (2004). Revisiting the JDL Data Fusion Model II. Seventh International Conference on Information Fusion (FUSION 2004).

Lockhart, C. (2006). Soldier as a System. Master of Strategic Studies Degree Strategy Research Project, United States Army War College Strategic Studies Inst. Carlisle Barracks PA.

Logsdon, J. (2008). One Semi-Automated Forces (OneSAF): Capabilities, Architecture, and Processes, United States Army Simulation Training and Instrumentation Command Orlando FL Program Manager Constructive Simulation.

Lowe, D. and S. Ng (2004). Effects-based operations: language, meaning and the effects-based approach. 004 Command and Control Research and Technology Symposium: The Power of Information Age Concepts and Technologies.

Lucas, T. W., S. M. Sanchez, et al. (2007). Defense and homeland security applications of multi-agent simulations, IEEE Press.

Mach, E. (1905). "Knowledge and error: Sketches on the psychology of enquiry [TJ McCormack & P. Foulkes, Trans.]." Boston, MA: Reidel.

Mackaness, W. A. and K. M. Beard (1993). "Use of graph theory to support map generalization." Cartography and Geographic Information Science **20**(4): 210-221.

MacKay, N. (2008). "Lanchester models for mixed forces with semi-dynamical target allocation " Journal of the Operational Research Society (2008), 1 -- 7

Macy, M. and R. Willer (2002). "FROM FACTORS TO ACTORS: Computational Sociology and Agent-Based Modeling." Annu. Rev. Sociol **28**: 143-166.

Marshall, G. (1998). "satisficing". A Dictionary of Sociology.

Martin, G. S., Jeffery; Kwinn, Michael J. Jr. (2005). Acquistion-Based Simulation. Proceedings of the 2005 Winter Simulation Conference

Mastroianni, G. R. (1996). Situation Awareness: A Primer For Materiel Developers.

Mastroianni, G. R. and V. E. Middleton (1996). Computer Simulation of Dismounted Movement. 20th Army Science Conference. Norfolk, VA.

Mastroianni, G. R. and V. E. Middleton (2001). Where Psychology, Physiology, Biomechanics, and Geography Intersect: Modeling Human Performance. Language, Logic, and Logistics: Modeling and Cross-Disciplinary Discourse. New Mexico State University, Las Cruces, New Mexico.

Matsopoulos, A. (2007). Little by Little Does the Trick: Design and Construction of a Discrete Event Agent-Based Simulation Framework. MASTER OF SCIENCE IN OPERATIONS RESEARCH, NAVAL POSTGRADUATE SCHOOL MONTEREY CA.

Matthews, M., R. Pleban, et al. (2000). Measures of infantry situation awareness for a virtual MOUT environment. Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference — October, 2000 Savannah, GA.

Matthews, M., L. Shattuck, et al. (2001). Situation awareness for military ground forces: Current issues and perspectives. Human Factors and Ergonomics Society Annual Meeting Proceedings, Cognitive Engineering and Decision Making Human Factors and Ergonomics Society.

McCarley, J., C. Wickens, et al. (2002). A Computational Model of Attention/Situation Awareness Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society, Santa Monica, Human Factors and Ergonomics Society.

MCCDC (1999). Marine Corps Warfighting Publication (MCWP) 0-1, Marine Corps Operations. Quantico VA, Marine Corps Combat Development Command

MCDP1 (1997). Marine Corps Doctrinal Publication MCDP 1, Warfighting. U. M. Corps. Washington, DC: United States Marine Corps.

MCDP5 (1997). Marine Corps Doctrinal Publication MCDP 5,Planning. U. M. Corps. Washington, DC United States Marine Corps.

MCDP6 (1996). Marine Corps Doctrinal Publication MCDP 6 Command and Control. U. M. Corps. Washington, DC, United States Marine Corps.

Middlebrooks, S. and B. Stankiewicz (2007). The Application of Models of Decision Making During Uncertainty to Simulations of Military Command and Control Systems, United States Army Research Lab Aberdeen Proving Ground MD Human Research and

Engineering Directorate.

Middleton, V. (2010a). Imperfect Situation Awareness: Representing Error and Uncertainty in Modeling, Simulation & Analysis of Small Unit Military Operations. 19th Annual Conference of Behavioral Representation In Modeling and Simulation (BRIMS), Charleston SC.

Middleton, V., K. Sutton, et al. (2000). Soldier Integrated Protective Ensemble (SIPE) Advanced Technology Demonstration (ATD).

Middleton, V. E. (1991). Integrated Unit and Soldier System Survivability and Effectiveness Evaluation Roadmap. Natick Technical Report. Natick MA, Advanced Systems Directorate United States Army Natick Research, Development and Engineering Center.

Middleton, V. E. (2000). FY 2000 NSC Modelng and Analysis Roadmap. Natick Soldier Center Natick MA, Simulation Technologies, Inc.

Middleton, V. E. (2008). "“Emergent Validation”: Defensible and Authoritative Modeling, Simulation, and Analysis for Complex Adaptive Systems." Phalanx, the Bulletin of Military Operations Research **41No.1**(March 2008).

Middleton, V. E. (2009). "“EPOCHS-”: A Meta-Methodology for the Human Dimension;." Phalanx, the Bulletin of Military Operations Research **Vol.42 No.1**(March 2009).

Middleton, V. E. (2010b). Simulating Small Unit Military Operations with Agent-Based Models of Complex Adaptive Systems 2010 Winter Simulation Conference Baltimore MD.

Middleton, V. E., Hennon, C. G., Hayes, T.R. (2002). “The Next Generation Integrated Unit Simulation System: Intelligent Agent Simulation for Warrior Systems”. Eleventh Conference on Computer Generated Forces and Behavior Representation;. Orlando FL.

Middleton, V. E., McIntyre, R.T.III; (2001). “From Rifleman to Warrior System: The Evolution of the Dismounted Close Combatant”. SMi Conference on Dismounted Close Combat. London.

Middleton, V. E. M., G. R, (1996). Using Fuzzy Logic to Model Human Behavior. 64th Military Operations Research Society Symposium;. US Army Combined Arms Center; Ft. Leavenworth KS.

Middleton, V. E. M. G. (2008). Implications of Human Centric Modeling for Operational Analysis. 2008 Conference on Behavior Representation in Modeling and Simulation (BRIMS). Westin Hotel - Providence, Rhode Island, SISO.

Middleton, V. E. O. K., J.A. IV; McIntyre, R. T. III; Christenson, W.M.; (1999). Development of the Warrior System Conceptual Architecture (WSCA). 8th Computer Generated Forces and Behavioral Representation Conference, , . Orlando, Florida. **Proceedings of the 8th Computer Generated Forces and Behavioral Representation Conference, Orlando, Florida, May 1999**.

Miller, N. and L. Shattuck (2006). A Dynamic Process Model for the Design and Assessment of Network Centric Systems. 2006 Command and Control Research and Technology Symposium (CCRTS) The State of the Art and the State of the Practice.

Mitchell, A. (1999). The ESRI guide to GIS analysis. Volume 1: Geographic patterns &

relationships, ESRI Press.

Moffat, J. (2011). Adapting Modeling & SImulation for Network Enabled Operations. Wahington D.C., Office of the Assistant Sceretary of Defense for Networks and Information Integration

Molenaar, M. (1998). An introduction to the theory of spatial object modelling for GIS. London, Taylor and Francis, Ltd.

Money, A. (2001). Report on Network Centric Warfare Sense of the Report, Office of the Secretary of Defense, Assistant Secretary of Defense (C3I). **934:** 106-398.

Murdock, P. (2002). "Principles of War on the Network-Centric Battlefield: Mass and Economy of Force." Parameters: US Army War College **32**(1): 86.

Murtaugh, S., Workshop Chair and Editor (1994). Workshop Proceedongs: Human behavior and performance as essential ingredients in realistic modeling of combat workshop (MORIMOC III). M. O. R. Society, Center for Naval Analyses**:** 26-28.

NATO/RTO/HFM (2009). Human Behavior Representation in Constructive Simulation, North Atlantic Treaty Organisation (NATO) Research and Technical Organisation (RTO) Human Factors and Medicine (HFM) Panel

NRDEC (2005). Infantry Warrior Simulation (IWARS) User Guide Version 1.0 Natick MA, US Army Natick Research, Development, and Engineering Center.

NRDEC (2009). IWARS User's Guide Version 2.0, US Army Natick Research Development and Engineering Center (NRDEC) and US Army Material Sciences Analysis Acitivity (AMSAA).

Ntuen, C. (2006). Cognitive Constructs and the Sensemaking Process. 2006 2006 Command and Control Research and Technology Symposium (CCRTS):The State of the Art and the State of the Practice.

O'Keefe, I., J. A., V. E. MIddleton, et al. (1992). Modeling and Simulation of the Soldier as a System. Natick MA, US Army Natick Soldier Research Development and Engineering Center.

O'Hanlon, M. (2003). "Estimating Casualties in a War to Overthrow Saddam." Orbis **47**(1): 21-40.

O'Kane, B. L. P., GL; Wilson, DL; Bohan, DJ (2004). Cycle criteria for detection of camouflaged targets. RTO SCI Symposium on "Sensors and Sensor Denial by Camouflage, Concealment and Deception", . Brussels, Belgium.

Olwell, D. (1997). "Modeling loss exchange ratios as inverse Gaussian variates: implications." Military Operations Research **3**(1): 51-67.

Page, E. H. (1998). The rise of Web-based simulation: implications for the high level architecture. Winter Simulation Conference 1998., IEEE.

Page, E. H. L., Wendell H. (2001). Architectural Principles for the US Armyís Simulation and Modeling for Acquisition, Requirements and Training (SMART) Initiative. 2001 Winter Simulation Conference.

Parsons, D. S., John LTC.; Jordan, Ben (2005). OneSAF: A Next Generation Simulation Modeling the Contemporary Operating Environment 2005 Euro Simulation Interoperability

Workshop.

Parsons, D. S., JR (2005). The US Armyís Next Generation Simulation Modelling the Response to the World's Future Threat. The Effectiveness of Modelling and Simulation – From Anecdotal to Substantive Evidence Neuilly-sur-Seine, France: RTO.

Pearl, J. (1987). "Evidential reasoning using stochastic simulation of causal models." Artificial Intelligence **32**(2): 245-257.

Pearl, J. and J. Robins (1995). Probabilistic evaluation of sequential plans from causal models with hidden variables. The Eleventh Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc.

Pearman, G. M., V. E. Middleton, et al. (2009). "Studying Chemical Effects on Infantry Operations Using Agent-based Simulation and Emergent Analysis Techniques." Military Operations Resaerch (MOR) Journal.

Pellegrino, J. and A. Scott (2004). The transition from simulation to game-based learning. The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC), NTSA.

Perla, P. P. (1990). The art of wargaming: a guide for professionals and hobbyists, Naval Inst Pr.

Perumalla, K. S. and B. Bhaduri (2006). On accounting for the interplay of kinetic and non-kinetic aspects in population mobility models. European Modeling and Simulation Symposium, Citeseer.

Petty, M., F. McKenzie, et al. (2004). Developing a Crowd Federate for Military Simulation. 2004 Spring Simulation Interoperability Workshop 04S-SIW.

Pew, R. and A. Mavor (1998). Modeling human and organizational behavior: Application to military simulations, National Academies Press.

Pfautz, S. and P. Toman (2010). A Hybrid Model of Ethnic Conflict, Repression, Insurgency and Social Strife. Behavioral Representation in Modeling and Simulation (BRIMS).

Pfitzner, D., V. Hobbs, et al. (2003). A unified taxonomic framework for information visualization. Asia-Pacific symposium on Information visualisation, Australian Computer Society, Inc. Darlinghurst, Australia, Australia.

Phister Jr, P. W., D. Fayette, et al. (2005). CyberCraft: Concept Linking NCW Principles with the Cyber Domain in an Urban Operational Environment. Rome,NY, AF Research Laboratory, AFRL/IFB.

Pisula, C., K. Hoff III, et al. (2004). "Randomized Path Planning for a Rigid Body Based on Hardware Accelerated Voronoi Sampling." Algorithmic and computational robotics: new directions: the fourth Workshop on the Algorithmic Foundations of Robotics.

Ramirez, T. L. (1997). "Modeling military task performance for Army and Air Force personnel wearing chemical protective clothing." Military Psychology **9**(4): 375-393.

Reece, D. A. (2003). "Movement Behavior for Soldier Agents on a Virtual Battlefield." Presence Teleoperators **12**(4): 387-410.

Reece, D. A. K., Matt; Dumanoir, Paul (2000). Tactical Movement Planning for Individual Combatants. Ninth Conference on Computer Generated Forces and Behavior Representation. Orlando Florida.

Reese, B. and B. Stout (1999). "Finding a Pathfinder." Proceedings of the AAAI 99 Spring Symposium on Artificial Intelligence and Computer Games.

Richter, K. and A. Klippel (2005). A Model for Context-Specific Route Directions. Spatial Cognition IV. Reasoning, Action, and Interaction, Springer Berlin / Heidelberg.

Rider, R. R. (1994). Operational Research. Baltimore, The Johns Hopkins University Press.

Ritter, F. (2008). Some Frontiers of Cognitive Modeling: A Modest Research Agenda Exploring Emotions and Usability. University Park, PA 16802 Applied Cognitive Science Lab

College of Information Sciences & Technology

The Pennsylvania State University

Ritter, F. A., MN (2000). Steps Towards Including Behavior Moderators in Human Performance Models in Synthetic Environments. Penn State, School of Information Sciences and Technology, The Pennsylvania State University

Ritter, F. S., NR; Elliman, D; Young, RM; Gobet, F; Baxter, GD (2001). Techniques for Modeling Human Performance in Synthetic Environments: A Supplementary Review, Wright-Patterson Air Force Base, OH: Human Systems Information Analysis Center

Rönngren, R. and M. Liljenstam (1999). On event ordering in parallel discrete event simulation. The Thirteenth Workshop on Parallel and Distributed Simulation (PADS), IEEE Computer Society.

Roske Jr, V. (2002). "Opening Up Military Analysis: Exploring Beyond The Boundaries." Phalanx, June.

Russell, S. J. N., Peter (2003). Artificial Intelligence A Modern Approach 2cd ed. Upper Saddle River, New Jersey, Pearson Education, Inc.

Safavian, S. R. and D. Landgrebe (1995). "Inference and Decision-making with Partial Knowledge." ECE Technical Reports: 140.

Sanchez, S. and T. Lucas (2002). Exploring the World of Agent-Based Simulations: Simple Models, Complex Analyses. 2002 Winter Simulation Conference

Schaffer, M. (1968). "Lanchester models of guerrilla engagements." Operations Research **16**(3): 457-488.

Schauble, L., R. Glaser, et al. (1991). "Causal models and experimentation strategies in scientific reasoning." The Journal of the Learning Sciences **1**(2): 201-238.

Schmitt, J. F. (1997). Command and (Out of) Control: The Military Implications of Complexity Theory Complexity, Global Politics, and National Security. D. Alberts, Czerwinski, TJ. Washington D.C., National Defense University 99-111.

Schumacher, C. (2005). Ground moving target engagement by cooperative UAVs. American Control Conference, 2005, IEEE.

Seitz, T. (2008). Representing Urban Cultural Geography in Stabilization Operations: Analysis of a Social Network Representation in Pythagoras. Master of Science in Operations Research, Naval Postgraduate School.

Senders, J. W. and N. P. Moray (1991). Human error: Cause, prediction, and reduction. from the

Second Conference on the Nature and Source of Human Error, 1983, Bellagio, Italy, Lawrence Erlbaum Associates, Inc.

Shattuck, L. and D. Woods (2000). Communication of intent in military command and control systems. The human in command: Exploring the modern military experience. C. M. a. R. Pigeau, New York : Kluwer Academic/Plenum Publishers, 279-292, 2000.

: 279–291.

Shlapak, D. A. and P. K. Davis (1991). Possible Postwar Force Requirements for Persian Gulf: How Little is Enough?, Rand Corporation.

Shneiderman, B. (2002). "Inventing discovery tools: combining information visualization with data mining." Information Visualization **1**(1): 5-12.

Silverman, B. (2004). Toward Realism in Human Performance Simulation. The Science and Simulation of Human Performance. J. W. Ness, Ritzer, Darren R. & Tepe, Victoria, New York: Elsevier**:** 469-498.

Simon, H. (1959). "Theories of Decision-Making in Economics and Behavioral Science." The American Economic Review **49**(3): 253-283.

Smith, C. (2011). Nett warrior C3Conflict experiment: measuring the effect of battlefield awareness in small units. Monterey, California., Naval Postgraduate School.

Smith, E. A. (2002). Effects Based Operations: Applying Network Centric Warfare in Peace, Crisis, and War, Washington, DC: DoD CCRP.

Smith, E. R. and F. R. Conrey (2007). "Agent-based modeling: A new approach for theory building in social psychology." Personality and social psychology review **11**(1): 87.

Smith, R. (2010). "The long history of gaming in military training." Simulation & Gaming 40th Anniversary Issue **41**(1): 6.

Sowa, J. (2000). Knowledge representation: logical, philosophical, and computational foundations, MIT Press.

St Clair, S. J. (2003). Barriers to Using Models and Simulations (M&S) in Training Forums, united States Army War College Carlisle Barracks PA.

Stanzione, T. and K. Johnson (2006). Tools for the Creation of Semantic Information for Modeling and Simulation, DOD CCRP.

Steinberg, A., C. Bowman, et al. (1999). Revisions to the JDL data fusion model, ERIM International, Inc. 1101 Wilson Blvd. Arlington, VA22209.

Straßburger, S., T. Schulze, et al. (1998). Internet-based Simulation using off-the-shelf Simulation Tools and HLA. 30th Winter Simulation Conference, IEEE Computer Society Press.

Surdu, J. R. B., D. Parsons, et al. (2005). The Three Block War in OneSAF. SimTecT Conference 2005, Sydney, Australia, NTSA.

Taylor, J. (1983). Lanchester models of warfare Vol.s 1 & 2, Operations Research Society of America, Arlington, VA.

Taylor, J. G., U. Z. Yildirim, et al. (2000). Hierarchy-of-models approach for aggregated-force attrition. Proceedings of the 32nd conference on Winter simulation. Orlando, Florida, Society for Computer Simulation International.

Totten, C. A. L. (1880). Strategos: a series of American games of war, based upon military principles and designed for the assistance both of beginners and advanced students in prosecuting the whole study of tactics, grand tactics, strategy, military history, and the various operations of war, D. Appleton and company.

TRADOC (1993). Field Manual No. 100-5 Operations, Headquarters, U.S. Army Training and Doctrine Command,.

TRADOC (1995). TRADOC Pamphlet 525-69 Military Operations: Concept for Information Operations. U. S. A. T. a. D. C. F. M. Headquarters, VA

TRADOC (1997). Field Manual No. 101-5 STAFF ORGANIZATION AND OPERATIONS, Headquarters, U.S. Army Training and Doctrine Command,.

TRADOC (2003). Field Manual No. 6-0 Mission Command:Command and Control of Army Forces, Headquarters, U.S. Army Training and Doctrine Command,.

TRADOC (2008). Field Manual No. 3.0 Operations, Headquarters, U.S. Army Training and Doctrine Command.

Triantaphyllou, E., B. Shu, et al. (1998). "Multi-criteria decision making: an operations research approach." Encyclopedia of electrical and electronics engineering **15**: 175-186.

Turner, R. M. (1998). "Context-mediated behavior for intelligent agents." International Journal of Human-Computers Studies **48**(3): 307-330.

Underhill, R. (2009). "The Multifunctional Materials Needs of the Future Dismounted Soldier." The Canadian Army Journal **12**: 61-74.

USAIS (1992). Field Manual No. 7-8 Infantry Rifle Platoon and Squad, US Army Infanty School.

USAIS (1993). Field Manual No. 90-10-1 An Infantryman's Guide to Combat in Built-up Aeas, US Army Infanty School.

van Lent, M., R. McAlinden, et al. (2004). Enhancing the behaviorial fidelity of synthetic entities with human behavior models. 13th Conference on Behavior Representation in Modeling and Simulation (BRIMS).

van Son, E. M., George; Lotens, Wouter; Middleton, Victor (2003). From Close Combat to Close Contact. US and Germany Technical Exchange. Naval Postgraduate School, Monterey CA.

Vaughan, D. (2006). Exploration of Force Transitions in Stability Operations Using Multi-Agent Simulation. Master's, Naval Postgraduate School.

Visone, D. (2005). Battlespace Terrain Reasoning and Awareness. 2005 ESRI International User Conference, Redlands, California.

Wagenhals, L. and A. Levis (2002). Modeling Support of Effects-Based Operations in War Games, George Mason University Fairfax VA.

Waldrop, M. (1992). The Emerging Science at the Edge of Order and Chaos, Simon and Schuster Paperbacks, New York, New York.

Walker, W. E. (2000). Uncertainty: The challenge for policy analysis in the 21st century, DTIC Document.

Warwick, W. (2006). Agent-based Modeling of Dismounted Infantry Through Inclusion of Perceptions, Inferences and Associations Phase 1. Boulder CO, Micro Analysis and Design.

Weir, J. (2009). Lecture Notes OPER 643 Air Force Institute of Technology Wright Patterson AFB Dayton OH.

Wellbrink, J., M. Zyda, et al. (2004). "Modeling Vigilance Performance as a Complex Adaptive System." The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology **1**(1): 29.

Westbury, C., U. Wilensky, et al. (1998). Knowledge representation in cognitive science: Implications for education. First International Conference on the Learning Sciences and the Challenges of the Information Era. Lima: Peru Lima: Peru.

Witten, I. H. and E. Frank (2005). Data Mining: Practical machine learning tools and techniques 2nd Edition, Morgan Kaufmann.

Wojcik, J. H., C; Wilson, B (1996). Soldier System Architecture (SSA) Multiple Criteria Decision Aid Detailed Development Plan. Natick, MA, US Army Soldier Systems Command, Natick Research, Developemnt and Engineering Center.

Wooldridge, M. and P. Ciancarini (2001). "Agent-oriented software engineering: The state of the art." Lecture Notes in Computer Science: 1-28.

Wooldridge, M. and N. Jennings (1995a). "Intelligent agents: Theory and practice." Knowledge engineering review **10**(2): 115-152.

Wooldridge, M. and N. Jennings (1995b). "Agent theories, architectures, and languages: a survey." Lecture Notes in Computer Science **890**: 1-21.

Worboys, M. and M. Duckham (2004). GIS: A computing perspective, CRC Press.

Wu, Q. (2006). Incremental Routing Algorithms For Dynamic Transportation Networks. Master of Science, University of Calgary

Young, J. P. (1957). History and Bibliography of War Gaming, Johns Hopkins University Chevy Chase MD Operations Research Office.

Zacharias, G., J. MacMillan, et al. (2008). Behavioral Modeling and Simulation: From Individuals to Societies, Natl Academy Pr.

Zachary, W. R., JM; Hicinbothom, JH (1998). Cognitive task analysis and modeling of decision making in complex environments. Decision-Making Under Stress: Implications for individual and team training**:** 315-344.

Zeiler, M. (2010). Modeling our world: the ESRI guide to geodatabase design 2nd ed., Esri Press.

Zhu, P. and R. C. Wilson (2005). A Study of Graph Spectra for Comparing Graphs. British Machine Vision Conference, Oxford Brookes University Oxford, U.K. .

Zimm, A. D. (1999). Modeling Maneuver Warfare:Incorporating Human Factors and

Decisionmaking in Combat Modeling. Eighth Annual Conference on Computer Generated Forces and Behavioral Representation. Orlando FL, Simulation Interoperability Standards Organization (SISO).

Zimm, A. D. (2001). "A causal model of warfare." Military Review **81**(1): 47-53.