THE INTERNATIONAL

C2 JOURNAL

VOLUME 4, NUMBER 1, 2010

SPECIAL ISSUE

Agility and Interoperability for 21st Century Command and Control

GUEST EDITOR

Sandeep Mulgund
MITRE

Tactical Agent Model Requirements for M&S-based IT→C2 Assessments

Jim Davidson
Alex Pogel

CCRP
Focus & Convergence for Complex Endeavors

THE INTERNATIONAL C2 JOURNAL

About the Journal

The International C2 Journal was created in 2006 at the urging of an international group of command and control professionals including individuals from academia, industry, government, and the military. The Command and Control Research Program (CCRP, of the U.S. Office of the Assistant Secretary of Defense for Networks and Information Integration, or OASD-NII) responded to this need by bringing together interested professionals to shape the purpose and guide the execution of such a journal. Today, the Journal is overseen by an Editorial Board comprising representatives from many nations.

For more information

Visit us online at: www.dodccrp.org
Contact our staff at: publications@dodccrp.org

CCRP
Focus
& Convergence
for Complex Endeavors

Tactical Agent Model Requirements for M&S-based IT→C2 Assessments

Jim Davidson and Alex Pogel (New Mexico State University, USA)

Abstract

The continuing spotlight on the C2 domain, specifically on information and the progressive organizational paradigms (*edge organizations*), and the corresponding acquisition focus on IT, necessitates that the defense operations research community create analytic capabilities regarding IT's support to C2. Modeling and Simulation (M&S) offers a cost-effective method for exploring issues within the C2 domain. The most common use of M&S assets is to assess whether IT variations correspond to variation in force effectiveness (FE); we call these *IT→FE experiments*, and while they require that M&S assets have explicit models of inputs and outputs to C2 and state change actuation, they do not require explicit models of core (internal) C2 processes. On the other hand, experiments designed to directly assess IT's support of C2—what we term *IT→C2 experiments*—do require explicit models of core internal C2 processes, from which measurements can be produced that provide a basis for comparison of IT variations. One such measurement is goal evaluation accuracy. In this paper, we define three familiar categories of goal statements—purpose, task, and primitive action—using sufficiently specific domain referents that the terms are clearly differentiated from one another. The definitions are intended to aid development of unambiguous and testable requirements for software models of commanders, an important step in the design of M&S assets intended for IT→C2 experiments based on the commander's goal evaluation accuracy. To help clearly express these modeling requirements we refer to and refine Albus' 4D/RCS model of an intelligent agent by indicating that the "commanded tasks" in his model may take any one of the three goal forms. After this goal categorization we then present a brief survey of well-known M&S applications, examining each for their instantiation of the goal categories, and then contrast these developments

with another M&S application that explicitly instantiates two of the three categories. Further we describe how this application was recently used in an IT→C2 experiment. Finally we return to Albus' 4D/RCS model to develop additional commander agent requirements derived from the concept of *organizational agility*, especially the *adaptation* and *innovation* features that figure so prominently in agent cognitive processes. The ontological developments in this paper are understood as an important first step toward establishing modeling requirements for core processes of virtual commanders, in a move toward the ability to conduct direct M&S-based IT→C2 experiments across the entire C2 approach space, based on the fundamental measurement of *goal evaluation accuracy*.

Introduction

The continuing spotlight on the C2 domain, specifically on information and the progressive organizational paradigms (edge organizations), and the corresponding acquisition focus on IT, necessitates that the DoD operations research community create analytic capabilities regarding IT's support of decision makers. Typical responses to this analytic need are experiments conducted with agent-based M&S assets that, at root, integrate four model types: models of information inputs to virtual commanders (sensing and networks), models that produce decisions of the virtual commanders, models of actuation of state-changing mechanisms under the direct or indirect control of virtual commanders, and models of the environment itself. Such M&S-based approaches provide cost-effective and highly controllable experimentation methods, and the constructed experiments balance analytic requirements with current modeling capabilities.

To date, the most common use of M&S assets is to assess whether IT variations correspond to variation in force effectiveness (FE); we call these activities *IT→FE experiments*. Pursuit of robust M&S-based FE assessments must continue since major acquisition decisions require FE assessments and M&S applications provide cost-effective analytic solutions. However, there is another type of experiment that has not received as much attention, where the focus is more directly on decision making outcomes. Irrespective of decision making variables such as genius, nationality, experience, and training, all military decision makers require relevant, timely, and accurate information regarding foundational domain constructs to succeed, e.g., information regarding mission, enemy, terrain, time, troops available, and civil considerations (METT-TC factors). Therefore, in addition to pursuing further developments regarding FE outcomes, analysts must also devote attention to intermediate assessments that are designed to *directly* assess IT's support of C2; we call these activities *IT→C2 experiments*. Successes in this direction should enable eventual experiments of the form IT→C2→FE.

The development of capabilities for direct IT→C2 assessments is especially crucial in the network-centric era, an era in which many key acquisition decisions and future force performance expectations derive from IT developments and procurements. And while M&S assets used in FE assessments incorporate explicit models of inputs and outputs to C2 and state change actuation, they do not require explicit models of core (internal) C2 processes (Ilchanski 1997; Pearce et al. 2003; Galligan et al. 2004; Yang et al. 2006; Moffat and Fellows 2008). Indeed, the C2 decision models are usually expressed as uniform rule sets, decision tables, or game-theoretic algorithms that are designed by analysts outside the simulation environment; none of these algorithmic constructions are developed with an expectation that the values witnessed within the decision algorithms (during simulation execution) will be worthy of being measured as an experimental outcome. Yet without explicit models of core C2 processes, these IT→FE experiments are forced to (indirectly) infer C2 results, based on the assumption that improvements in force outcomes must be (at least in part) due to better decision making; in essence, these applications are not designed to directly assess IT→C2 outcomes. Experiments focused on IT→C2 require M&S assets with explicit models of core internal C2 processes, from which measurements can be produced that provide a basis for comparison of IT variations, and they do not necessarily require inclusion of FE measurements.

The use of force outcomes as the indirect basis for C2 evaluation is an understandable response to the underdeveloped state of the C2 Ontology (C2O), as the ontology is not yet sufficiently developed that we can determine software requirements by simply choosing from a structured mosaic of well-established mechanisms.[1] This is a fundamental obstacle to the creation of correct, unambiguous, complete, verifiable and traceable (IEEE 1998) requirements for virtual commanders that are to be incorporated within agent based M&S applications, especially those constructed for IT→C2 experiments, with or without eventual reference to FE outcomes. Fully developing the C2O is an open problem for military domain researchers, and many previous C2O efforts have produced differing levels of descriptive granularity, successfully highlighting various isolated aspects of the C2 domain.[2] The development of a complete and accredited C2O is obviously difficult, based on the massive scope of the domain, the lack of transparency of tactical decision makers' internal thought processes including their informal interaction methods (such as personal communication between superiors and subordinates), and most importantly a *lack of criteria* for defining completion of various aspects of the ontology itself. Of course, this paper will not solve the problem of completing the C2O, and instead focuses on a more modest aim, the refinement of an absolutely fundamental aspect of

1. Indeed human problem solving within complex and epistemically uncertain military environments is still not well understood and the domain itself is under a constant state of flux given competition between opposing sides where self interests are pronounced and where the stakes are very high; these conditions suggest that the C2O may never properly be considered complete.
2. See for example the growing literature in the Command and Control Research Program (CCRP) and the US Army's own C2 doctrinal literature, most notably Field Manuals: FM 3-0 (Operations), FM 6-0 (Mission Command: Command and Control of Army Forces), FM 5-0 (Army Planning and Orders Production) and FM 5-0.1 (The Operations Process). See also, for example, the US Army Battle Command System.

the C2O, namely goals. In the realm of C2 assessments, a fundamental measure of IT effectiveness is the extent to which IT supports a commander's accurate evaluation of goal attainment status at any time, abbreviated as *goal evaluation accuracy.* Clearly the commander's process of evaluating goal involves establishing inferences and judgments between the METT-TC factors, all fed by information provided by IT assets.

In this paper then, we define three familiar categories of goal statements—*purpose*, *task*, and *primitive action*—using sufficiently specific domain referents such that all three terms are clearly differentiated from one another. The definitions are intended to guide the development of unambiguous and testable requirements for software models of commanders, an important step in the design of M&S assets to be used in IT→C2 experiments based on virtual commanders' goal evaluation accuracy.

In Section 2, we begin by noting the central nature of goals in the achievement of control within any form of organizational hierarchy, a centrality that is emphasized strongly in the progressive organizational control concepts of focus and convergence (Alberts and Hayes 2006). We then build upon the C2O by defining three goal categories. We discuss the relationship between the three goal categories and their impact on decision allocation. We also note that the definitions provide a refinement of Albus' 4D/RCS reference model architecture (Albus 2002) to include specific goal categories in the characterization of a virtual commander's relations with superiors, peers, and subordinates in an organization.

In Section 3, we briefly review a number of well-known M&S developments, including ISAAC, EINSTEIN, MANA, WISDOM, and WISE, and note that all offer potential for FE experiments, but none explicitly model core C2 internal processes from which audits can indicate accuracy of commanders' goal evaluation. This state of affairs is not a mark against any of these M&S efforts, as software development is always oriented toward very specific purposes, e.g., theorizing about complex adaptive systems and emergent behavior. We then conclude section 3 by contrasting these FE-oriented M&S developments with a summary of the System of Systems Survivability Simulation (S4), a simulation that has been created for direct support of IT→C2 susceptibility and vulnerability experiments based in part on virtual commanders' goal evaluation accuracy.

In Section 4, we describe a recent experiment using the S4 M&S asset, highlighting the fact that the ability to implement metrics for key measures of merit, namely assessments of IT support of decision making, was completely dependent on the explicit modeling of *purpose* as a goal category in S4's virtual platoon leaders.

In Section 5, we note a consequence of two cognitively rooted aspects of agility (*adaptation* and *innovation*, of Alberts' and Hayes' six features) that directly depend on the categories of goals present within the processing of a virtual commander. Further we express these consequences as additional modeling requirements for *agile* virtual commanders, by returning to and extending Albus' 4D/RCS conceptualization of an intelligent agent. From these additional *agility based requirements*, we derive other important IT→C2 experimental consequences relating to allocation of decision rights, patterns of interaction and distribution of information. We end section 5 by assessing the extent to which S4 satisfies these additional organizational agility requirements, highlighting successes and failures. Finally we conclude by describing future work on a continuing expansion of the C2O for implementation in the S4, in the area of virtual agents' interactions within an organization.

In summary, the ontological developments in this paper are understood as an important first step toward establishing modeling requirements for core processes of virtual commanders, including software agents advanced enough to express organizational agility. When virtual commanders are constructed to satisfy these requirements, we believe the value gained will be the ability to conduct M&S-based

IT→C2 experiments across the entire C2 approach space, ranging from traditional militaries to edge organizations, all based on the fundamental measurement: goal evaluation accuracy.

Section 2. Goal Categories

In any organization, control trickles down the hierarchy through the construction and delivery of goal statements, from superior to subordinate, level by level. This is the case across all possibilities expressed in the *C2 approach space* (Alberts 2009), ranging from *traditional* organizations to *edge* organizations. In many traditional organizations, goals provide over-specified descriptions of future actions to be taken by subordinates that too often represent improper management of uncertainty. This is not to say that rigid control is never necessary as some contexts may very well require tight control (e.g., a microprocessor assembly line or operation of a nuclear power plant). In the realm of full spectrum operations however, control and decision authority will invariably call for wide distribution of a general intent, combined with forces' use of disciplined initiative in seeking the overall endstate. Indeed progressive organizational designs are essential for the 21st century and beyond. But even under the most progressive of forms, edge organizations, that are imbued with focus and convergence, goals remain a central referent: "In brief, …focus provides the context and defines the purposes of the endeavor; convergence is the goal-seeking process that guides actions and effects" (Alberts 2007). In other words, *focus* is the existence of the goal, "a synthesis of how the situation is perceived and understood, including perceptions about the nature of the endeavor (strategies and plans) that are appropriate for the situation," while *convergence* is the process of seeking the goal, in which "the emphasis is placed squarely on improving the value-view of current and future states rather than achieving some specific result." Another implication of convergence is the expectation that the participants in an endeavor will alter their *in situ* environments, consistent with the roles they are to fulfill in achieving the organizational focus. This will necessarily include altering their internally produced or organizationally assigned sub-goals.

However, goals do not take the same form across the C2 approach space; indeed, there is a strong connection between specificity of goal and degree of decision allocation. Further, information distribution and patterns of interaction between members of an organization also depends on specificity of goals. For an organization to express agility, goals must be formed carefully to indicate *why* a mission is undertaken, but *not* expressing *what* is to be done, nor *how* it is to be done; therefore a high degree of abstraction in the expression of goals, nonetheless providing full expression of the bare minimum beneficial effects to be gained, is essential for maximum decision allocation exercised by subordinate leaders. By contrast, detailed command or classical C2 (a command strategy still exercised by many contemporary military organizations) is conducted through delivery of highly specific goals, that spell out what is to be done and how it is to be done, and it is now understood that commanders using such specific goal statements will often assert too much control over subordinates in the wrong context. In recognition that tight control is inappropriate in many 21st century contexts several armies (e.g., US, UK, Australia) have now adopted *mission command doctrine*, a distributed command and control strategy that is rooted in the German concept of *Auftragstaktik*. The essence of mission command is "disciplined initiative" in which subordinates are expected to carry out their mission orders with a high degree of initiative and latitude in execution, consistent with the commander's intent (or more recently, common intent). However, since command styles are apt to vary widely from highly controlled to very loose control (i.e., more or less decision allocation, even within the same organization) M&S applications designed to support IT→C2 assessments must account for the differences since each command style is likely to have its own corresponding IT configuration, and the various IT configurations are likely to show more or less susceptibilities to threats such as electronic warfare and computer network attacks.

In order to decompose and assess the full range of C2 strategies that IT is meant to support, including decision allocation, we must characterize the different types of goals that can be passed from superior to subordinate, that is, we must categorize goals.

During our past literature reviews we have found that many goal forms and goal statements lack the level of specificity needed for formalization in an agent-based software model. For example in US Army doctrine we find concepts that share the term *task*: tactical mission task, mission-essential-task, collective task, and individual task. The emphasis on tasks, especially in the context of training for full spectrum operations, has led to their encapsulation in the US Army Universal Task List and even in the Joint Universal Task List at DoD level. Despite the effort, task (as a goal category) definitions remain ambiguous and are not easily translated into code for consumption by virtual commanders. For example the definition of the US Army tactical mission task (FM 3-90 2001) *disrupt*, reads:

> [It is] a tactical mission task in which a commander integrates direct and indirect fires, terrain, and obstacles to upset an enemy's formation or tempo, interrupt his timetable, or cause his forces to commit prematurely or attack in a piecemeal fashion.

Whereas the definition for *interdict* reads:

> [It is] a tactical mission task where the commander prevents, ***disrupts***, or delays the enemy's use of an area or route.

Clearly it is problematic when developing software requirements for virtual commanders if one definition of task includes the term of another task in its own definition.

Yet, even in software representations of commanders with explicit, unambiguous computations corresponding to goal evaluation, it is still true that the term *goal* could be interpreted broadly too, e.g., to mean "this agent's algorithm attempts to make the value of this function as great as possible." In some circles (rational agent theory) this would be correct, since utility functions are considered the basis for all decision making, and there are no presumed theoretical restrictions on what is expressed by a utility function. However, when software is developed for an agent model to be used for assessment of IT→C2, based on a virtual commanders' goal evaluation accuracy, the formulation and the interpretation of goals in the eyes of superior and subordinate becomes an important feature of the model itself: if the goal of an agent (a component of an algorithm) is not sensibly interpretable by an analyst employing the M&S asset, the experiment conducted will necessarily lose analytic power. Where, so often, focus in software efforts (regardless of domain) is placed on computational efficiency, the act of *modeling* requires deliberate attention to expressive power. Therefore, the aim of the remainder of this section is to categorize goals with three detailed definitions; the definitions should allow us to understand not only when a software agent possesses and acts upon a goal, but also to interpret what kind of goal the agent is acting upon. As mentioned above, a by-product of this categorization is that we can then refine the concept of the decision allocation spectrum to show how goal forms correlate to allocation of decision making authority.

At the basic level, doctrinal goal descriptions take the form *task* and/or *purpose*. For example US Army doctrine (FM 3-90 2001; FM 1-02 2004) provides these definitions:

a. Purpose (or *why* in the mission statement): the reason for a unit's task, activity or endeavor (dictionary synonyms include object, aim, goal, or end). For example: ***protect*** *the move of the humanitarian aid convoy to village D from the insurgents on hill B.*

b. Task (e.g., tactical mission task; *what* in a mission statement): The specific activity performed by a unit while executing a form of tactical operation or form of maneuver. It is the minimum essential effects to accomplish the purpose, usually expressed with reference to an enemy force or terrain (dictionary synonyms include job or assignment). For example: ***contain*** *insurgents inside the Khyber Pass*.

To develop the C2O with respect to goal statements including over-controlling goals, we are going to add another category called *primitive action*, which we define as the direct application of an asset to achieve a desired state change. Although there is no explicit reference to *primitive actions* in doctrine, we do find implicit references in various equipment user's manuals and how-to guides, in training and evaluation literature on tactics, techniques and procedures common to the domain, and in technical manuals pertaining to sequences of actions necessary to perform work such as diagnosing faults on complex components and systems, including hardware and software.[3] Some examples of primitive action goals include: a Team leader's command to a subordinate to dash across the street and enter a building; a Stryker commander's instruction for a loader to re-load the main gun ammunition ready box; a psychological operations section chief's command for a subordinate to print a thousand information operation leaflets; a maintenance supervisor's instructions to a mechanic to run diagnostic test 7230 on the turret distribution box in a Bradley Fighting Vehicle.

In its broadest form, the term *goal* merely indicates an effect, some state of some referent that is desired by the subject agent, and in general, such a desired state could be considered sufficient as a foundation for evaluation of current and future states by agents directly controlling assets (as in the primitive action case); however, such a broad characterization is insufficient for development of requirements for virtual commanders whose primary state-changing control mechanisms are indirect, expressed through modification of evaluation functions of other agents, that is, through the delivery of goals

3. See any of the US Army's Equipment Maintenance Technical Manuals. TM 8-4110-002-14&P: Unit, Direct Support, and General Support Maintenance Manual; Refrigerator, Solid State, Biological Model DLA-50T (1998) is an excellent example of an extensive set of primitive actions.

to subordinate agents. Intelligent agents, sufficient for experimenting on IT→C2 across the C2 approach space, must use appropriate goal categories to allocate decision authority and must be aware of risks associated with the assignment of goals to subordinates, never assigning goals that will result in no apparent benefit for the risk involved. Thus, we define three goal categories, namely *purpose*, *task*, and *primitive action*:

> A *purpose* is a general description of desired, directly beneficial futures for an OWNFOR referent, representing an aggregation of a multitude of beneficial concrete future states the OWNFOR referent may achieve;
>
> A *task* is a general description of desired futures for a non-OWNFOR referent or for an OWNFOR referent from which benefit can only be indirectly inferred, representing an aggregation of a multitude of concrete future states the referent may achieve;
>
> A *primitive action* is a specific description of a future state that is directly achievable by direct application of assets.

The first distinction to draw from the definitions is that purposes and tasks are expressed in general terms that reflect a variety of specific outcomes, so they inherently involve some degree of abstraction in the description of the desired futures, while actions only describe a specific future state that can be realized by an application of assets. Purposes and tasks are distinguished from one another by the degree of directness with which some benefit is expressed by the goal: a general goal is a purpose if it expresses some direct benefit to an OWNFOR referent; otherwise it is referred to as a task.

There are two crucial notions implicit in the definition of purpose. In a purpose statement, received in a high risk, epistemically uncertain environment, the perception of benefit by the receiving agent generates willingness to accept risk in the pursuit of purpose despite the dangers and uncertainties. For example, a force may be willing to put their own lives at risk if they know by doing so that they will be protecting a friend, since trust, honor and loyalty between comrades is highly prized. Furthermore, the generality required of the description constituting a purpose leads to broad applicability and persistent relevance. For example, a purpose to *protect* a sister unit is likely to remain persistently relevant owing to the receiving unit's desire for honor and loyalty among peers, but also abstract enough to allow for wide variety in execution; the allowance for independence of thought and action during execution is especially important in high risk, uncertain contexts). Ultimately, a commander's choice of an appropriate level of generality will determine the effectiveness of a purpose: too concrete a purpose will lose relevance when natural situational variation arises, requiring that further purposes be delivered, with a cost incurred from additional communication (usually via IT) and comprehension of the new goal descriptions; too general a purpose will not provide subordinates a criteria against which they can judge their alternatives. Selecting a level of abstraction that assures purpose is persistent is a judgment call for the creator; this judgment *must be* made by agents who evaluate candidate purposes and choose one to deliver to subordinates. Importantly the requirement that purpose refer to a beneficiary (the referent) makes explicit the benefit toward which all activity is oriented, including expenditures, asset application, and risk undertaken.

The general description that comprises a task does not present literal expression of direct benefit to OWNFOR, and as such this goal form is not specifically rooted in the psychology of people at war (Grossman 1993; Ardant du Picq 2009); that is, people fight best when they can plainly see how their fighting matters to their cohorts' survival. Supported by historical evidence and psychological studies, it is widely acknowledged that people fight because they cannot let other people down.[4] Significantly the content of a typical *task* goal

4. See for example Leonard Wong et al. "Why They Fight: Combat Motivation in the Iraq War." This study is an impressive survey and historical review of combat motivation completed for the Strategic Studies Institute at Carlisle Barracks, PA, July 2003.

does not make explicit how one person's effort helps another. Of course, some benefit should arise from a task, but this benefit must be abducted by the stakeholder. In practice, provided a task is present, commitment to a task is justified through inferences on the part of the organization or its members that achievement of the task will make achievement of the purpose more likely. On the other hand, if a purpose is not provided, this inference becomes a matter of guesswork, of abduction, based primarily on what purpose might be implicit in the determination of this task.

Primitive actions are instances of state changes derived from the available state changing capabilities of assets that are under the direct control of the agent. If an agent can drive a vehicle, the vehicle has the ability to change the position of the agent, so an action based on this capability could "Move from point A to point B;" if an agent can fire round type R from weapon W at any point within 3km of its current location, a primitive action based on this capability could be "place round R at point (x, y, z) (that is 1.2km away)."

In both generality and importance to OWNFOR, purposes are greater than tasks, and tasks are greater than primitive actions. We summarize these relationships as the *goal ordering principle*, expressed as a multi-part inequality:

$$\text{Purpose} > \text{Task} > \text{Primitive Action}.$$

The relation > should be read as "are served by," so the principle indicates that lower level goals always serve higher level goals. If an agent or team has a purpose, then the tasks they adopt (or were assigned) are believed to serve the purpose; correspondingly, with any task, the primitive actions are believed to serve the task. It is apparent from this sequence that the category of goal delivered to a subordinate agent or team has a great impact on the degrees of freedom available to the agent or team in fulfilling the goals. After a display of examples of the interactions, we further discuss these degrees of freedom.

Example 1. Consider a platoon with the following goals: purpose is "protect platoon Y from enemy tank fire, while it advances to hill 62"; task is "contain enemy tanks south of a line from benchmark 591, east to benchmark 788"; primitive action is "engage 2nd enemy tank from left."

Example 2. Consider a multinational battalion force: purpose is "The US Embassy is to remain safe from car bombings during the diplomatic meetings"; task is "search all vehicles entering the US Embassy compound"; primitive action is "open the trunk of the car."

Example 3. Consider a multinational humanitarian relief organization: purpose is "ensure citizens of city Z receive an uninterrupted flow of food and medical aid during the monsoon season"; task is "block rising river adjacent to food warehouse"; primitive action is "fill and stack sand-bags."

In Example 1, the purpose will clearly be evaluated by examining expected or actual enemy tank fire present in current and future states of platoon Y, during its advance to hill 62, while the task is considered sufficient to achieve the purpose because the enemy tank firing range is known and keeping them south of the designated line will disallow them from firing upon platoon Y; finally, to keep enemy tanks sufficiently south of platoon Y, a member of the platoon will engage the 2nd enemy tank from the left.

At the risk of overemphasizing goal definitions and our previous goal categorization, we want to explain the necessity to do so when developing agent-based models. As can be inferred from the definitions, an agent operating on a purpose goal is clearly going to require a completely different evaluation (and hence require different information) than an agent operating on a task goal, and similarly for tasks versus primitive actions. For example, suppose an agent X is given a purpose to protect platoon A from enemy B during platoon A's move to hill 123. Clearly agent X will need to concern itself with platoon A's execution, and the movement, disposition and threat posed by enemy B (at any time) against platoon A, while platoon A moves to hill 123. If platoon A changes direction, agent X will have to assess how platoon A's change in execution impacts its current and future ability to protect platoon A from enemy B, and adjust accordingly. One can see from this example that communication, coordination and cooperation between platoon A and agent X is essential for agent X to evaluate its goal attainment status. Consequently, IT plays a crucial role in helping agent X perform its evaluation by providing the medium for the vital, dynamic information exchange that must take place between platoon A and agent X and IT must help agent X interpret environmental changes that impact its ability to protect platoon A.

Now assume that agent X is instead given the task to *seize* hill 123. In this case agent X will need to evaluate the threat sitting on hill 123 and judge how its unit will move to (with minimal risk) and take hill 123 from enemy forces so ensconced. In this example agent X's judgment fundamentally revolves around how its unit is to fight and survive while taking a hill the enemy presumably wants to keep. IT will need to support agent X's evaluation, but the information requirements are very different from the previous example. More importantly agent X will likely avoid risk since it has no compunction to take on more risk than is absolutely necessary since the capture of the hill offers no apparent direct benefit to anyone. From these two simple examples we can see how goal categories offer very different agent evaluation and IT requirements, and the examples indicate how agent goals can directly impact patterns of interaction, decision allocation, and information distribution within an organization.

With the previous discussion as backdrop we now refine Albus' 4D/RCS intelligent agent model (Albus 2004) to express differences in core cognitive processes of an agent when it is acting on a particular goal category (see Figure 1). Note that *commanded task* (i.e., goal), appearing at upper right of the diagram, is provided as input to an agent (presumably the output of a higher level agent), and *commanded actions* (i.e., subgoals), appearing at lower right of the diagram, are provided as output of the agent (to become input to lower level agents). The goal categories referred to in the model are task and primitive action, respectively; however, this restriction is not appropriate if we expect to model the entire C2 approach space. Therefore, we extend Albus' 4D/RCS model by allowing that the incoming and outgoing goals can be from any of the goal categories.

For the remainder of the paper we will continue to use Albus' reference model to express implications of goal categories for intelligent (simulation) agents, especially *agile* intelligent agents. Significantly, the US National Institute of Standards and Technology's (NIST) Intelligent Systems Division (ISD) has been developing this model of intelligent agents for over 30 years. The most recent version of the original reference control architecture, now named 4D/RCS, adds time to each of the other modules: Sensory Processing, World Modeling, Value Judgment, and Behavior Generation. The model has served as the conceptual foundation for the implementation, testing, and revision of several robotics applications within US and German defense organizations (hence the *Operator Interface* appearing at the right hand side of Figure 1). Also the model integrates concepts from theories in cognitive psychology, semiotics, neuroscience, and artificial intelligence (Albus 2006). There are other models of intelligent agents, e.g., SOAR, with its focus on learning, and ACT-R, with its focus on primitive elements of thought, but we prefer the emphasis on command and control in Albus' reference model, specifically the communication of goals from superior to subordinate (and so on), along with its ability to represent core concepts that are identifiable within military commanders. Indeed, the key notions of many other models of intelligent agents already appear, at least implicitly, within Albus' model. For example, belief-desire-intention (BDI) concepts (Cohen and Levesque 1990) are embedded inside Albus' modules (World Model, Value Judgment, and Behavior Generation, respectively), but Albus does not require that modal logic be used to encode constraints on BDI.

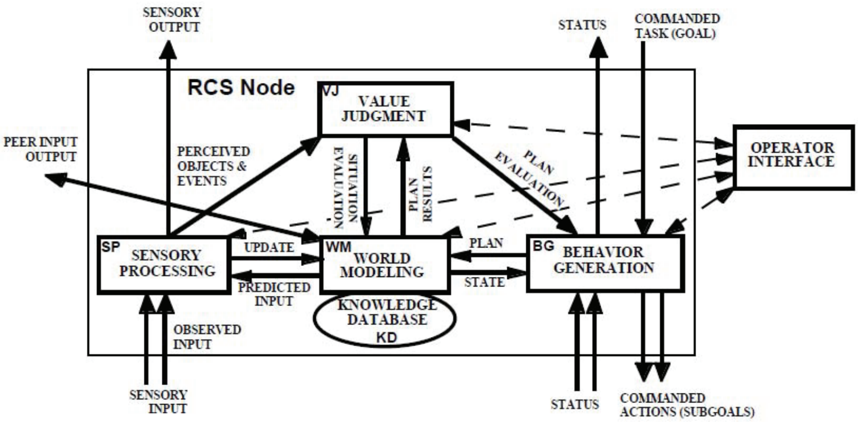

Figure 1. Internal view of Albus' structure of a typical 4D/RCS node.

Also consider Klein's Recognition Primed Decision (RPD; Klein 1998) model illustrated in Figure 2. Klein's concept "Experience the situation in a changing environment" is refined by combining Albus' Sensory Processing and World Model modules and "recognition has four by-products" is consistent with a combination of Albus' World Model, Value Judgment, and Behavior Generator modules. Similarly Klein's concept "evaluate, mentally simulate and will it work" is refined by combining Albus' World Model and Value Judgment modules. Note however, that the 4D/RCS makes sensory perception explicit and internal to the agent, and the model includes interactions with other agents; these concepts are missing from the RPD model. More importantly RPD is a model of expert decision making, by decision makers familiar with the contexts in which they decide. Yet we know from historical evidence (especially recent history) that military decision makers will often be called upon to make decisions in contexts for which they have no previous experience (hence commanders are often novice decision makers).[5] One example illustrates the point very well:

> "I am a combat infantryman. You want me to fire and maneuver; I can fire and maneuver—anywhere, in any terrain, anywhere you want to do it. Here, I have had to learn how sewage works. In my AO [Area of Operations], I can brief you where all my pumps are, all my manholes, and where my sewage is broke."[6]

5. See, for example, the Institute for Defense Analyses report by Tillson, John C. et al., 2005. <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA442427&Location=U2&doc=GetTRDoc.pdf> This in-depth report for the *Under Secretary of Defense for Personnel and Readiness* highlights the DoD need to train adaptive leaders for 21st century asymmetric threats.

6. Wong, Leonard. 2004. *Developing Adaptive Leaders: The Crucible Experience of Operation Iraqi Freedom*. Strategic Studies Institute. US Army War College. Carlisle, PA. p. 6. <http://www.strategicstudiesinstitute.army.mil/pdffiles/PUB411.pdf>

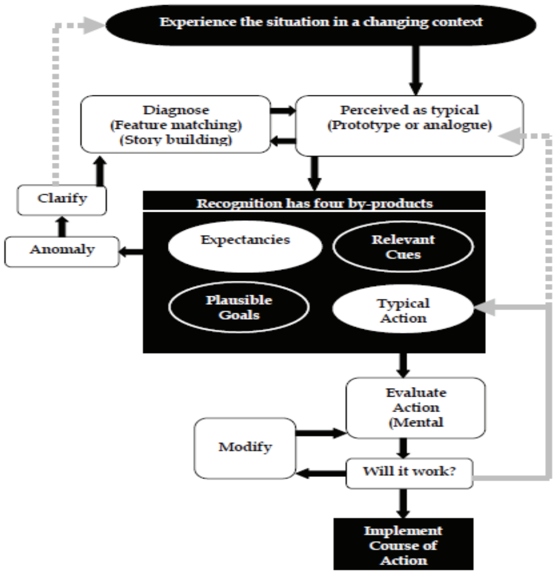

Figure 2. Klein's RPD as adapted by Elliott; Australian Defence and Technology Organisation 2005.

The 4D/RCS conceptualization allows for both kinds of decision makers, expert and novice. Lastly, Albus' model integrates the *individual* level with the *organizational* level, specifically multiple levels of interacting agents, an organizational characterization that neither BDI nor RPD account for. Setting aside completeness and generalizability of the Albus model, our refinement of the goal categories indicates that the goal inputs, the goal outputs, and the peer-to-peer interaction in 4D/RCS can involve any of the previously described goal categories. Therefore, we can infer from the goal ordering principle that decision allocation is realized through freedom in goal-setting; this is not to be confused with resource allocation that is determined independently of the granting of broad or little goal-setting freedoms (NATO 2006). More specifically, in terms of our goal categories, the degree of decision allocation witnessed in an organization is largely dependent on the category of goals employed in the superior-subordinate relationships: a respectively high/medium/low degree of decision allocation is granted when goals are expressed in the form of purpose/task/primitive action statements, respectively.

Section 3. Goal Types Instantiated in Agent-Based Models

Across the range of military M&S models that we have surveyed, we have not seen explicit representation of goal categories and specifically *purpose* evaluation in the dynamic cognitive processes of virtual commanders. Many M&S developments and intelligent agent models do account for goals but these refer to *tasks* and/or *primitive actions* (Gonzalez et al. 2005; Kraker et al. 2009; Kewley 2004; Cheng 2007; Bock 2000), not *purpose*. For example three M&S assets devoted to the discovery of emergent behavior in warfare are the Irreducible Semi-Autonomous Adaptive Combat (ISAAC), and its extension the Enhanced ISAAC Neural Simulation Toolkit (EINSTEIN) (Ilachinski 1997; 2004), the Map Aware Non-uniform Automata (MANA) (Lauren and Stephen 2002), and Warfare Intelligent System for Dynamic Optimization of Missions, Version II (WISDOM-II) (Yang et al. 2006). These models implement agents whose decision making processes are based on externally pre-determined rule sets that agents follow during execution of tasks and primitive actions. The aggregated behavior of homogenous, rule-following agents, operating on tasks and primitive actions in relatively simplified worlds results in unexpected, unpredictable behavior. However interesting the emergent results are, there is a notable absence of *explicit agent goal evaluation* in these M&S applications, neither with respect to subordinates' achievement of self-determined goals (e.g., evaluate avoidance of threat line-of-sight while attempting to move from point A to point B), nor the agent's achievement of goals ordered from above (e.g., evaluate, at any time, achievement of the task: seize hill 123, or evaluate and protect platoon A from threat B); consequently, in none of these simulations can an agent's goal evaluation accuracy be computed. Any goal category that may be implicitly present, in the mind of the rule set developer (e.g., "these agents are trying to capture the flag"), is not explicitly evaluated by the agents themselves.

Another M&S example with more mature representation of concepts in the C2O is the Wargame Infrastructure and Simulation Environment (WISE) (Pearce, Robinson, and Wright 2003) developed in the United Kingdom (UK). WISE (Fellows et al. 2004) "is an agent based analytical tool used for the analysis of combat but with a heavy emphasis on the command and control aspects of manoeuvre warfare," that was designed to support a variety of UK defense analysis needs such as cost benefit analysis, net-centric analysis (i.e., digitization of the battlefield), doctrine and tactics, as well as futuristic C2 concepts. A unique aspect of WISE is the facility for human-in-the-loop interaction (hence Wargame in the acronym) or constructive, closed-loop simulation mode. In terms of agent decision making representations, WISE incorporates abstractions of "deliberate" and "rapid" planning (Moffat and Mason 2000; Moffat 2007; Moffat and Fellows 2008). We note from scenarios explored within the WISE by *UK Defence Science and Technology Laboratory* analysts (Moffat 2007; Moffat and Fellows 2008), that the goal categories distributed between higher (deliberate planners) to lower (rapid planners) agents generally follow the form *tasks* and/or *primitive actions*. Despite the game-theoretic computations performed before a simulation is executed, in order to seed the planners with rich decision logic, in the constructive mode of WISE's applications, the agents do not explicitly evaluate their goals, and consequently, goal evaluation accuracy cannot be computed from simulation outputs.

In summary, many of these agent-based models were designed for IT→FE experiments in order to theorize about complexity when warfare is viewed through the lens of complex adaptive systems and emergent phenomena. The aforementioned applications serve their purposes well, but they were not specifically designed to assess *how well IT configurations support a virtual commander's evaluation of its goal attainment status.* Consequently these applications do not require agents to be aware of or evaluate all three goal categories. In our work we are concerned with directly assessing IT→C2, specifically we want to know how perturbations to IT (e.g., via electronic warfare or computer network attacks) impact C2 performance. For the remainder of this section then, we describe an M&S development motivated by a desire to produce IT→C2 assessments, namely the System of Systems Survivability Simulation (S4). The S4 presents a virtual commander modeling paradigm that instantiates two of the goal categories but incorporates agent evaluations of all three.

The S4 model was developed for the Army Research Laboratory by New Mexico State University's Physical Science Laboratory. As Berstein et al. (2006) described, "each agent's revision process includes information and knowledge from local sensors, remote sensors via the network, its engagements and encounters, and its general situational awareness and understanding." This agent-based simulation acknowledges the stages of situated cognition described in the Dynamic Model of Situated Cognition (DMSC) (Miller and Shattuck 2004) and analysis of simulation outputs use the DMSC framework to develop and categorize key measures of merit and explanatory metrics. Two key modules in S4 that represent the DMSC stages are the communications module and the decision-making process (DMP) module. The communications module can introduce delay and loss of information, and agents' decision outputs, and more importantly the evaluations within their decision processes, are, in turn, sensitive to the quality, content and timeliness of the information provided to them by IT.

Platoon leaders represent the most advanced decision modules in S4, as their decision processes include explicit representations of purpose and task goals delivered to them by their virtual company commanders. As such we will focus our remaining discussion on the platoon DMP. Through Figures 3, 4, and 5, and their accompanying narratives, we present an overview of the S4 hierarchy and a platoon leader's decision making process.

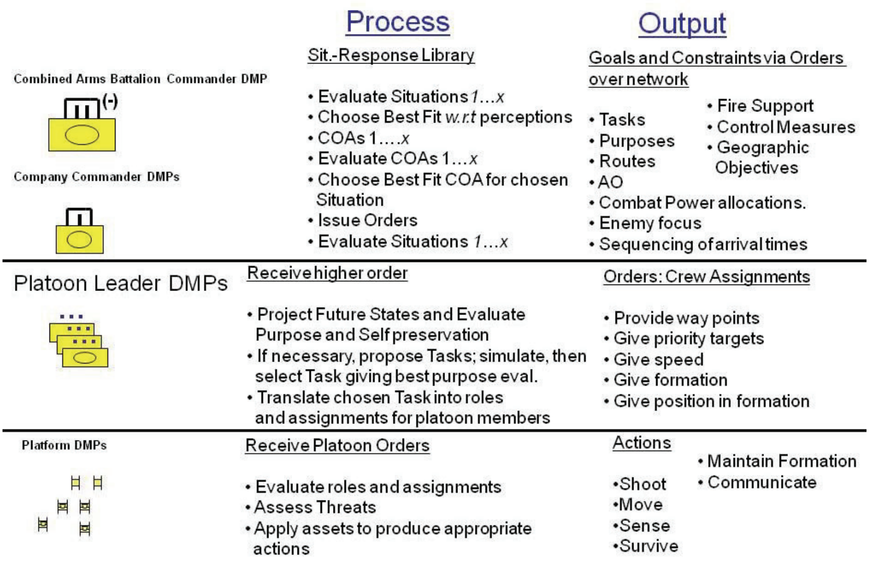

Figure 3. S4 Agent Hierarchy

In Figure 3, we see the platoon leader DMP resting below company commanders but above platform DMPs. In brief, each platoon leader performs a decision (i.e., purpose) evaluation at least once a minute (this time interval value is configurable), and more often when it receives information relating to a *critical threat* (a metric we will return to later in the paper). As indicated in Figure 3, the platoon leader's role is to accept an order from the company commander and produce orders for his vehicle commanders. An order from company level includes mission information such as an initial task, a fixed purpose (unless the company commander changes orders later), a direction of attack, an area of operations, allocation of combat power, a geographic objective for orientation, sequencing with sister platoons in terms of arrival times at company determined objectives, the platoon's enemy focus, fire support target reference points, and other control measures such as lines of departure and phase lines requiring automatic status reports. The platoon leader uses its perception of the situation and the contents of its order to project future states and evaluate the states relative to its assigned purpose and unit survival. If a purpose and/or survival evaluation is sufficiently problematic, the platoon leader proposes tasks to solve its goal given its identified critical threats. The platoon leader then projects future states out to a configurable horizon based on its unit's execution of the proposed tasks (respectively) and ultimately chooses the task giving the best purpose/survival evaluation provided the winning task differs from the current task. Note that this feature of the platoon DMP differs from Klein's RPD model wherein the expert decision maker chooses the first course of action that it believes will solve the problem (aka Herbert Simon's "satisficing"). The platoon DMP more closely approximates Cohen's Metacognition Model (Cohen et al. 1996) by allowing for mental simulation (i.e., brainstorming/wargaming) by the agent in epistemically uncertain situations, and this sequence of alternative evaluations represents an implementation of the interactions between the Behavior Generator, World Model, and Value Judgment modules in Albus' reference model architecture. Once the platoon leader completes its decision process, it will eventually communicate roles and assignments to its platoon members consistent with the winning task (Davidson et al. 2006).

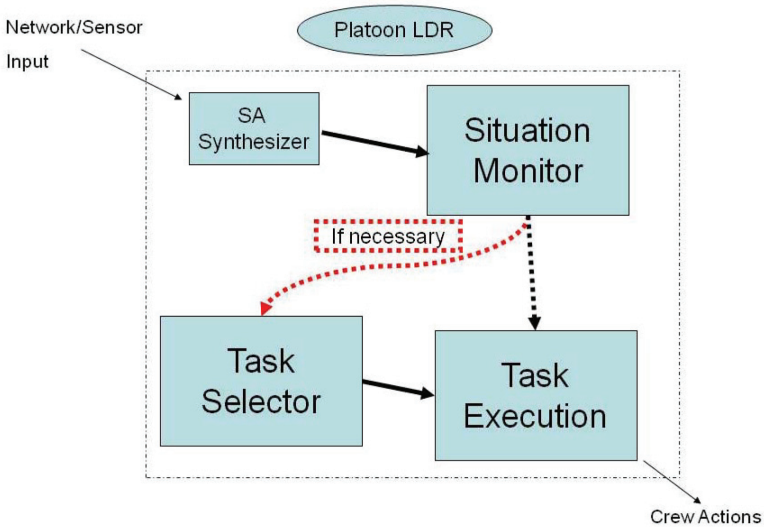

Figure 4. S4 Platoon Leader DMP

Referring to Figure 4, we now address the primary components within the platoon DMP's core internal processes (excluding Task Execution since this component concerns execution of tasks by subordinates, but not the DMP's purpose evaluation or task selection processes). The first is Situation Awareness Synthesizer (SAS); in short SAS tailors the platoon DMP's incoming information (via sensors and network) corresponding to the area of operations assigned to it by the company DMP; the SAS is an implementation of Albus' Sensory Processing module. This information is evaluated in the Situation Monitor module (SM) with respect to the platoon leader's purpose and unit survival. SM is the component that amalgamates all incoming information (including orders from the company DMP via the network) from SAS and evaluates that information to produce a situation; SM is an implementation of Albus' World Model and Value Judgment modules. SM uses a state building machine to forecast futures (out to a 15 minute horizon–this is configurable) regarding the execution of the platoon's current task, implementing

World Model, and evaluates the futures against the backdrop of its currently perceived threats and understood friendly activity, implementing Value Judgment.

With respect to IT→C2 experiments, the most important evaluation SM performs on enemy information concerns degree of threats to the platoon's purpose, and SM determines that one or more of these threats may become the platoon's critical threat. For example, SM may conclude that one of three threat systems is projected to travel to within 500 meters of the protected platoon, and have line of sight at that position, while the other two are projected to remain several kilometers away. Based upon SM's purpose evaluation it may send a request to task selector (TS), e.g., if a new threat is deemed to be a serious hazard to protection of beneficiary platoon under the current task then SM will request a task, or if a new purpose goal arrives from the company DMP, then SM will request a new task, etc. In this case, the *brainstormer* subcomponent of TS proposes relevant tasks for the problem futures (see the decomposition in Figure 5); the TS and task execution components in the platoon DMP are implementations of Albus' Behavior Generator module. The *wargamer* subcomponent within TS then "mentally simulates" the execution of the brainstormer's proposed tasks by reusing the state building machinery of SM and the purpose evaluation mechanism. Ultimately, wargamer scores each proposed task according to its purpose and unit survival satisfaction ranking. Finally the task scores produced by wargamer are delivered to the *task chooser* subcomponent of TS that compares the task scores and simply chooses the best task (assuming that the proposed winning task can overcome the platoon DMP's inertia for its current task).

Figure 5 highlights the interactions between the components in the design. As can be expected, the interactions between the various DMPs (as there are many within the Blue and Red organizational hierarchies), in response to their enemy counterparts and other aspects of the physical environment, produces substantial tactical variety during the execution of many hundreds of runs. The variety stems not only from the DMP constructions themselves and their interactions, but also from the fact that the information provided to commanders, through the software models of DMSC stages, is never constant or perfect. The M&S applications surveyed at the beginning of this section also produce substantial variety, normally expressed as the emergence produced (in FE terms) from the aggregate behaviors of many autonomous, relatively homogeneous, localized, rule following agents. By contrast, variety and unpredictability within S4 stems from the aforementioned opposing agents' goal categories and evaluation of goals (particularly at the platoon level), and their execution. Further the accuracy of their evaluations is a reflection of the quality of their information provided by the IT. In short an agent's perceptions and its goal evaluation accuracy depends quite directly on the IT. This should come as no surprise since S4 was designed to support assessments of IT→C2, especially when IT is perturbed by enemy action such as electronic warfare or computer network attacks.

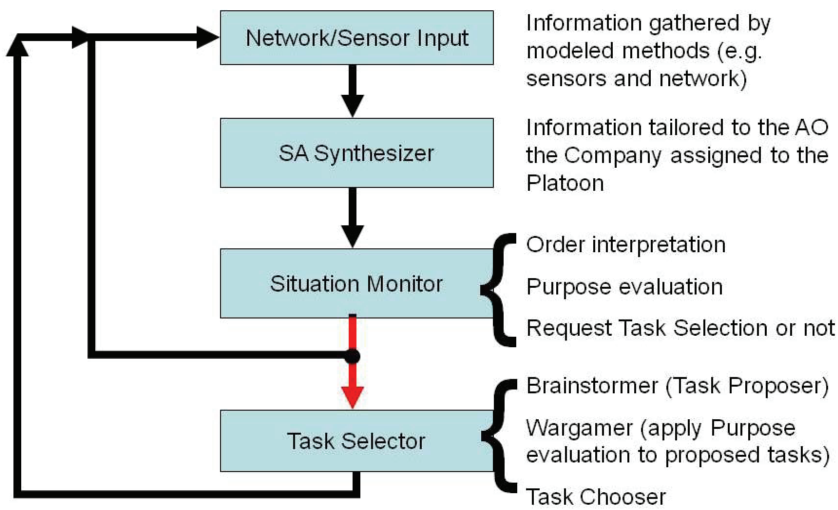

Figure 5. Platoon Leader Process Breakdown

With respect to the network represented in S4, communications are modeled over multiple networks for both Blue and Red forces. Depending on the study issues, the level of fidelity required for analysis can be set within the simulation. These levels range from a *No Attenuation* setting that enables perfect communications for all messages regardless of range and power to a *Fully Attenuated* setting that includes the use of complex algorithms (specifically, Terrain Integrated Rough Earth Model, [TIREM]) to determine dB loss over irregular terrain. The ability to vary simulation IT configurations (thereby exploring various IT perturbations) naturally supports our primary assessment method. For each simulation agent A, denoted SimA (e.g., a platoon DMP), there is a ground truth agent, denoted GTA, that makes decisions, including evaluations of threats to goal attainment, using the same algorithms as the SimA, but using ground truth data as input to the algorithms. The GTA decisions have no impact on the simulation, but are recorded for later comparison with those of the subject SimA, supported by various IT configurations. This methodology is called the Objective Information System Assessment (OISA) method and further details on the comparisons enabled by the method have been reported elsewhere (Davidson et al. 2008).

In the next section we transition from the conceptual description given here to a recent application of S4 in an IT→C2 experiment, in particular using the Situation Monitor (specifically, the purpose evaluation module) of the platoon leader model. These results have been reported elsewhere (Hudak et al. 2008), but serve here to highlight an analyst's ability to use the S4 M&S asset to directly perform IT→C2 assessments, because of the explicit representation of goal categories, virtual commanders' goal evaluation, and IT configuration control in the S4.

Section 4. An IT→C2 Experiment Based on Purpose Evaluation Accuracy

These experimental results were produced at New Mexico State University, using the S4 M&S application, as part of a doctoral dissertation in Industrial Engineering, and were reported fully (in Hudak et al. 2008). The experiment focused on assessing the relationship between information provided by the technological systems and a virtual commander's (specifically, a PLT leader's) ability to accurately understand the battlefield and make decisions. The research attempted to answer two questions: (1) how do different degrees of sensor density, sensor quality, and communication network quality affect the quality of information provided to the decision-maker; and (2) how does this change in information affect decisions made by the simulation agents, in comparison to their corresponding ground truth agents? The experimenters used the OISA method to assess decision making: information provided to the subject Simulation Agent (SimA) by IT was judged to be high quality if the SimA judgments corresponded with the Ground Truth Assessor's (GTA) judgments in the following sense (Hudak et al. 2008): "As a generalization of the traditional notion of value of information, we assert that value is dependent upon the relevancy and accuracy of the information provided to a decision-maker…(based on) the principle that relevant and accurate information provided to a decision maker always increases utility."

The experiment employed a full factorial design, over three key factors expressing IT variants: Sensor Quality, Sensor Density, and Communications. *Sensor quality* had three levels (quality = 80/90/100% of baseline capability), *sensor density* had three levels (Low/Medium/High force density of intelligence, surveillance and reconnaissance [ISR] units), and *communications* had three levels (80/90/100% transmission efficiency in the communications domain). The outcome variables were decision metrics, indicating levels of perception, comprehension and projection, consistent with the Endsley's SA Levels 1, 2, and 3 (Endsley 2000). Metric values were computed as goal evaluation accuracy scores for explicit judgments of the PLT leader, by comparing the PLT leader's (SimA's) judgments at numerous decision points with the assessor agent's (GTA's) judgments at these same decision points.

The experiment scenario depicts 1st Battalion in a high intensity battle, with Team Infantry, consisting of two infantry carrier platoons and one tank platoon, in the spotlight. The company's purpose is to protect the advance of Team Tank through a mountain defile (Tidzi Pass), thus enabling Team Tank to continue its advance to another key mobility corridor further to the east (Farsi Pass). The capture of Farsi Pass should allow for the unimpeded advance of the battalion main effort to ultimately secure an international airport east of Farsi Pass. The factor settings of sensor density low, medium, and high equated to 1, 2, or 3 reconnaissance platoons, respectively, that provide intelligence reports to Team Infantry. The threat force consisted of ten insurgents within Tidzi pass and a company-size militia armor force. The enemy's purpose goal is to protect forces at the international airport that are finishing their defensive preparations by delaying the advance of 1st Battalion through Tidzi Pass for five hours. The decision processes of the second platoon leader, Team Infantry, were of particular interest during the experiment. Figures 6 and 7 below illustrate the Battalion and Team Infantry scenarios modeled in the S4 for this experiment.

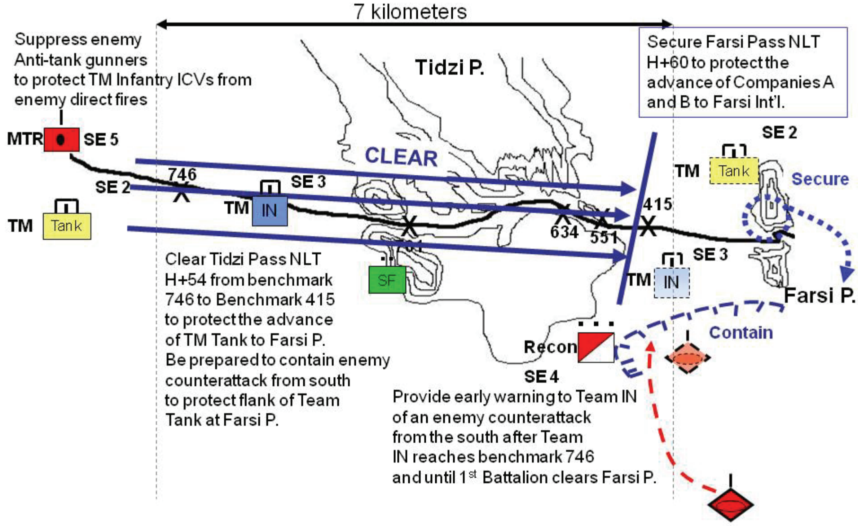

Figure 6. 1st Battalion Concept of Operation. Note the purpose and initial task goal of Team Infantry.

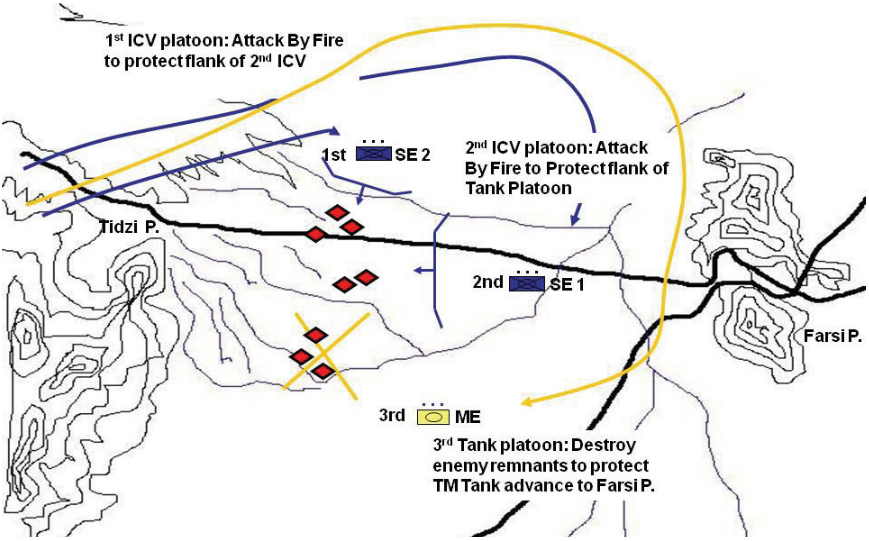

Figure 7. Team Infantry Course of Action 3 (of 5). Note the purpose and initial task goal of 2nd platoon.

Each platoon leader was modeled by an agent with platoon DMP described in the previous section. The experiment employed a measure called *situation awareness positional accuracy* (SA Pos Accuracy) implemented as a metric based on a computation of the completeness, or percent of all enemy forces known, combined with a distance-based accuracy score with the qualifier that for a perfect accuracy score, per enemy tank, the tank's position must be known within 100 meters of ground truth. This kind of measure of information input to a decision process is also used by many other M&S applications; the only noteworthy aspect here is that SA Pos Accuracy in S4 is computed using the OISA method, so that the standard against which accuracy is ultimately computed can be easily varied, when alternative assessor agents are of particular interest to the analyst. An example of an alternative assessor agent is the perfect communications agent (PCA), programmed to receive every message intended for the subject agent, though possibly not received by the subject agent; along with the GTA, the PCA is implemented in S4.

Another measure employed in the study, of more direct interest for this paper, is *critical threat unit accuracy* (CTU accuracy) that represents the accuracy of the comprehension of the platoon leader in determining the enemy entities that are a threat to the platoon leader's purpose (a protection goal, referring to a peer platoon). This measure is implemented as a metric by computing the percentage of entities within the GTA platoon leader's CTU array that are also in the SimA platoon leader's CTU array: the GTA provides the objective expression of which enemy entities should be understood as threats to purpose, while the SimA is judged according to which of those (GTA) entities he has identified as threats to purpose. Finally for projection, a platoon assigned objective accuracy (not to be confused with purpose) measure was created, that compared the distance between the final objective (i.e., way point) given by the SimA platoon leader to his subordinates to the final objective the GTA platoon leader would have given to his subordinates; the way point is the grid coordinate ordered by the platoon leader to his subordinate vehicle commanders, at which the platoon is to execute the task selected. This measure was implemented as a metric by computing the distance in meters between the way point for the SimA task execution and the way point for the GTA task execution.

Based on 53 replications per point in the full factorial design, the sensor quality factor demonstrated impact on the amount of area the reconnaissance force could survey and the number of intelligence reports submitted. However, this impact was small in comparison to the sensor density and communications factors, so we report only outcomes in the sensor density and communications levels. As seen in Figure 8, compared to the low sensor density level, the high sensor density level covered only about 38% more of the AO, but produced nearly 70% more intelligence reports. The low communications level led to about 36% more delay than the high level and slightly decreased the number of intelligence reports.

Communications Level	Sensor Density Level	Average (%) of AO Observable	Average Number of Intelligence Reports	Average Amount of Delay (seconds) to Transmit Reports
Low	Low	48	15.6	82
	Medium	56	19.7	81
	High	66	26.2	98
Medium	Low	48	16.1	71
	Medium	56	20.4	76
	High	66	27.5	80
High	Low	48	18.0	60
	Medium	56	21.5	72
	High	66	31.1	72

Figure 8. Experimental results for technological metrics, across various parameter values.

Figure 9 indicates outcomes for the CTU accuracy metric for the 1st platoon leader (whose goal was to protect the 2nd platoon). The experiment indicated that the competition for network resources, which arose naturally from the voice protocol model employed within S4, created delays in communication with larger ISR forces.

Though more information was gathered by the forces, it was slower to be distributed to the platoon leader in question, and the platoon leader's goal evaluation accuracy (modeled as CTU accuracy) suffered as a consequence. *In this context*, delay was more important than spatial coverage, indicating that the spatial coverage provided by the sensor density low level was already sufficient to report threats to platoon leaders. A more thorough investigation of the interaction between IT variations and decision maker's goal evaluation accuracy, to provide robustness of the final assessments, would include a variety of operational contexts used with the same factor and level combinations.

What is crucial here, and our primary motivation for reviewing the results of this IT→C2 experiment (Hudak et al. 2008), is that the explicit nature of the representation of goal categories and their evaluations in the cognitive process of this virtual platoon leader was necessary for the experiment to focus on the accuracy with which these evaluations were made—without the explicit models, these particular measurements would not be possible.

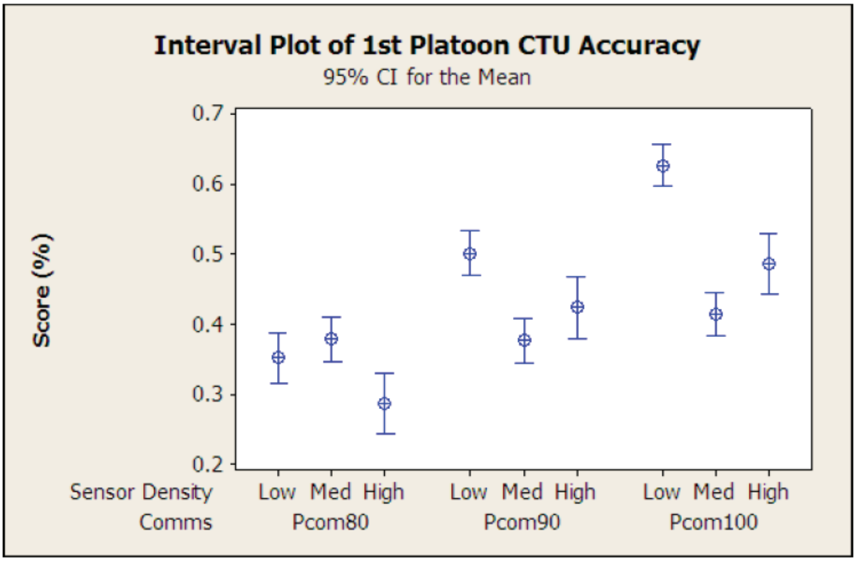

Figure 9. Confidence Intervals for Comprehension metric values in (Hudak et al. 2008)

Section 5. Additional Requirements for Agile Agents

In this section, we briefly review the notion of agility (Alberts and Hayes 2006) and note that a consequence of the previous goal category definitions is that we can refine two aspects of agility (*adaptation* and *innovation*, within Alberts and Hayes' six features). These two features of agility (i.e., the extent to which an agent, or an organization, manifests adaptivity or innovativeness) directly depend on the categories of goals present within the cognitive processing of a virtual commander. With this observation in mind, we further extend Albus' 4D/RCS model of an intelligent agent by identifying requirements for a specific model of an *agile agent.* From these additional requirements for agile agents, we also derive important IT→C2 experimental consequences relating to allocation of decision rights, patterns of interaction, distribution of information, and to IT representations themselves. Later we assess the extent to which S4 satisfies the additional agility requirements, highlighting successes and failures.

As the reader has undoubtedly noted, we have frequently referred to Alberts and Hayes, and Albus throughout the paper; this section will be no different. For that reason it is worthwhile to distinguish between the concept of agility and Albus' reference model of an intelligent agent. Alberts and Hayes have decomposed the concept of organizational agility into its relevant parts; converting ambiguous terms from human language into precise definitions with corresponding syntax is not an easy task but it is essential for the development of agent models and more generally the C2O. Significantly Alberts and Hayes describe for us *what* organizational agility is but do not go so far as to tell us *how* organizational agility happens. On the other hand Albus is not concerned with defining or describing the concept of organizational agility *per se*. Rather Albus' model is an attempt to describe an architecture for the cognitive processes organic to an intelligent agent, virtual or human, agile or not, experienced or not. As such Albus' model is a conceptual design that when fully developed and implemented would ostensibly lead to intelligent behavior, but the model does not go so far as to describe the required contents of each of the modules nor does it indicate the algorithms that would cause the modules to work in harmony to produce intelligence. In both cases, analysts are left to do more work—and rightfully so, as a reflection of the appropriate level of generality selected by these respective conceptual model builders—that is, work that defines how organizational agility and intelligence are actually produced in a specific context. To design, build and test agent-based M&S applications for IT→C2 assessments we must distill these conceptual and descriptive characterizations into requirements and specifications for eventual algorithmic form. The aim of this section then is to synthesize concepts originating at these respective sources, in an initial attempt to identify what the requirements would be for an agile, tactical decision making agent, within a specific M&S application.

Recent research indicates that an organization is *agile* if it has the ability to sustain progress toward goals in the face of various impediments (Alberts and Hayes 2003). The six features of agility suggest *what kinds* of impediments (robustness and resilience) can exist and also *how* an organization's progress is sustained, in terms of both cognitive processes (adaptation and innovation) and execution (responsiveness and flexibility). In the list below, we reorder the six terms, starting with the identification of impediments, proceeding to solution of the problem posed by the impediments, and finally resulting in effective execution of the solution. Further, we rephrase the six features to disambiguate the terms, to isolate the features that relate directly to goal categories.

Restated, an organization is said to be *agile* if it has the ability to sustain progress toward goals:

a. *Despite external impediments, based on environment or enemy* (robustness),

b. *Despite internal damage short of catastrophic loss of function, based on change of own state* (resilience),

c. *By changing goals, including the implementation of changes to organizational structures that will achieve those goals* (adaptation),

d. *By creating new capabilities and/or effects from out of existing organizational structures and resources* (innovation),

e. *By executing in a manner that employs timely information* (responsiveness); and

f. *By seamlessly transitioning between a breadth of existing capabilities* (flexibility).

As previously mentioned, Albus' model of the internal components of an intelligent agent (see Figure 1) can be extended in order to realize further requirements for *agile agents*. The first extension is the determination of requirements for the Sensory Processing (SP), World Model (WM), Value Judgment (VJ), and Behavior Generation (BG) modules, based on our rephrasing of the six features of agility (A - F above):

a. To express *robustness*, the agent's SP module must deliver current external state changes in its environment to the WM module;

b. To express *resilience*, the agent's SP module must deliver current internal state changes it (or its unit) has personally suffered to the WM module;

c. To express *adaptation*, an agent's BG module must be able to reset current goals for subordinates, in the form of purposes, tasks, or actions appropriate to the agent's role in the organization, by delivering plans to WM and VJ and choosing between them based on the returned evaluations;

d. To express *innovation*, an agent's BG module must be able to create new assets from existing assets and generate goals, in the form of previously unobserved purposes, tasks, or actions;

e. To express *responsiveness*, there are multiple module interactions that must occur: before delivering Commanded Actions (goals), an agent's BG module must command IT assets to resolve uncertainty values from VJ associated with possible plans; and

f. To express *flexibility*, an agent's WM module must produce future states, based upon predictions formed from a combination of input from the SP module and the KD module, and the BG module must continuously produce potential plans to resolve issues present in those future states.

Notice that the refinements of *adaptation* and *innovation* both refer to goals that can take the form of any of the three categories discussed in Section 2. However, we believe that organizational agility is enhanced when the goal category exchanged between superiors and subordinates is purpose, since it is most general and provides the greatest degree of decision allocation (because of the goal ordering principle of Section 2: Purpose > Task > Primitive Action). We also claim that when purposes are the preferred form of goal exchange, then peer-to-peer cooperation is likely to be highest since achievement of one's own goals depends upon others' achievement of their goals, necessitating cooperation and coordination between self and others. Thus another dimension of the C2 approach space, peer-to-peer collaboration, depends upon the use of purpose as the primary goal category. Furthermore, a corollary to the previous relationship is that when superiors use purposes, subordinates will manifest high levels of collaboration and information sharing. The various uses of goal categories in controlling subordinates within an organization also implies that IT will need to be flexible enough to support a variety of C2 configurations depending on the control methods imposed on a system.

For example in the extreme detailed command case where goals primarily take the form of tasks and primitive actions, IT will need to directly support the top nodes in the hierarchy. These nodes will need most of the bandwidth to manage all of the down-up information exchange that will be required to micromanage forces and their execution. On the other hand, if purpose exchange is the preferred method of control then IT will need to be configured to provide maximum support to peers, with little remaining bandwidth required for the top level nodes since they will primarily need to "listen in" on the developing situations, decisions, and interactions of subordinates; fine tuning execution when needed. Further, each IT variant presumably will have its own unique susceptibilities and vulnerabilities that will impact each virtual commander's goal evaluation accuracy.

Our inescapable conclusion is that IT→C2 experiments (especially those focused on agility) require goal evaluations (specifically purpose evaluations), IT variations, and measures of commanders' goal evaluation accuracies within organizational contexts. In reconsidering the earlier survey of M&S developments, we note that many of these applications [MANA, ISAAC, WISE, etc.] may be considered implementations of Albus' conceptual model for intelligent agents. However, with the agility requirements added these same applications often fail to satisfy the extended model of an agile agent, specifically these applications do not explicitly model the goal categories necessary for agility (especially purpose), lack agent projection of future states, lack agents' evaluation of futures, lack agent control of information system assets, and they lack other C2O concepts indicated by the full description of organizational agility. Within S4 itself, the development of the platoon DMPs only satisfies two key requirements of the extension of Albus' model to agile agents, namely the evaluation of purpose goals and the projection of futures.

Despite the advances in the development of the S4 platoon leader, the current range of S4 DMP implementations requires a different set of metrics (for IT→C2 experiments) unique to each type of virtual decision maker. This is because the DMPs are not consistent throughout the hierarchy, and each DMP has a different computational representation. To create a uniform assessment method for goal evaluation accuracy, all the DMPs at the different levels of the S4 hierarchy should differ only in the scale and scope of the problems they solve, in the assets they employ, and in the roles they fulfill, but not in the decision process. In this way an analyst using the S4 M&S application can apply the same suite of metrics to all DMPs, irrespective of their placement in the hierarchy, to produce an experimental design that treats the organization uniformly.

Yet despite these shortcomings, trace elements of organizational agility can already be found throughout the S4 model and also in some of the models previously mentioned. For example *robustness* is directly supported within S4. The platoon DMP's evaluation of alternative tasks can be triggered by external state changes due to enemy or terrain. Further each agent possesses a suite of sensors and network devices that provide for the agent's perceptions of its environment, specifically those state changes of immediate and direct concern to the agents themselves and of indirect concern to other agents on their teams. Recent efforts were also made to include high fidelity ballistics damage information within the S4 model. A consequence is that the simulation now supports *resilience* wherein each agent becomes aware of, responds to and reports its own internal capability degradations as a function of damage incurred. *Adaptation* is explicitly supported at the level of platoon leaders (while only indirectly and loosely supported at the battalion, company and platform levels) as described in the previous section. *Innovation* is not supported at all and is perhaps our biggest challenge, assuming that this element of agility is required to perform IT→C2 experiments. *Responsiveness* is loosely supported through the platoon leader's use of Class I UAVs to survey ground five minutes ahead of its unit's current position, but this situational use of IT assets does not occur elsewhere in the DMP hierarchy. The biggest challenge here is to model an agent's information uncertainty and its re-application of IT to fill in missing, unclear, incomplete or suspect information before it decides and issues orders. Currently S4 agents always possess certain information, however inaccurate it might be. Finally *flexibility* is also not supported. Here too the challenges are substantial. It will be a long road to model an agent's predictions of futures that will require the agent to conceive of new ways of working to solve a vast number of problems that could be foreseen in those predicted futures (e.g., fighting insurgents one day and building school houses the next). Again, despite shortfalls, we continue to expand the C2O and implement in S4 those concepts relevant to IT→C2 experiments, especially *how well do various IT configurations support virtual commanders' goal evaluations.*

Section 6. Summary and Future Work

In this paper, we defined three familiar categories of goal statements—purpose, task and primitive action—using sufficiently specific domain referents such that the three terms are clearly differentiated from one another. The definitions are intended to serve as an important start toward unambiguous and testable requirements for software models of commanders, a necessity in the design of M&S assets to be used in IT→C2 experiments based on the commanders' goal evaluation accuracy. We then surveyed familiar M&S developments and noted their lack of instantiation of goal categories (specifically purpose), goal evaluations and projections of futures, and presented a summary description of the S4. We also examined how the S4 was used to support a recent IT→C2 experiment. Finally we developed requirements for agile agents, derived from the concept of *organizational agility,* which we expressed via Albus' 4D /RCS model of an intelligent agent and then reconsidered advances in the S4 in light of the additional requirements. By constructing virtual commanders that satisfy the requirements in this paper, we believe it will be possible to conduct direct M&S-based IT→C2 experiments across the entire C2 approach space, ranging from traditional militaries to edge organizations, all based on the fundamental measurement: *goal evaluation accuracy*.

In the future we expect to expand our capability to perform more sophisticated IT→C2 experiments, yet modeling challenges remain in the domains of uncertainty, information systems' interdependencies, individual agent decision making behavior, and agent roles and interactions within organizations. Our immediate future will focus on the latter challenge. We want to account for and model organizational interactions (such as coordination, cooperation and collaboration) between virtual commanders (and their subordinates) in all types of organizations, including edge organizations. Continuing our use of Albus' conceptual framework we will endeavor to express requirements for inter-agent interactions between intelligent agents. The extensions will further expand the C2O and support modeling of the three key features of an edge organization: broad dissemination of information, unconstrained patterns of interaction, and maximum allocation of decision rights to subordinates.

References

Alberts, David S., and Richard E. Hayes. 2003. *Power to the Edge: Command and Control in the Information Age.* Washington, DC: Command and Control Research Program (CCRP).

Alberts, David S., and Richard E. Hayes. 2006. *Understanding Command and Control*. Washington, DC: CCRP.

Alberts, David S.. 2007. Agility, Focus, and Convergence: The future of Command and Control. *The International C2 Journal* 1(1): 1-30.

Alberts, David S. 2008. Agility, Focus, and Convergence: Adapting C2 to the 21st Century. Inaugural meeting of the *Focus, Agility and Convergence Team* (FACT). Vienna, VA. [Presentation] Available from CCRP website. <http://www.dodccrp.org/html4/events_fact.html>

Alberts, David S. 2009. Command and Control Maturity and Agility: A Way Forward. Presented to *AFCEA-GMU Critical Issues in C4I*. Washington, DC. <http://c4i.gmu.edu/events/reviews/2009/slides/Alberts-slides.pdf>

Albus, James S. 2002. 4D/RCS, a Reference Model Architecture for Intelligent Unmanned Ground Vehicles. Proceedings of the *SPIE 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando, FL.

Ardant du Picq, Charles. 2009. *Battle Studies: Ancient and Modern Battle.* Ed. Colonel John N. Greeley and Major Robert C. Cotton, trans. 1947. Dodo Press. Isbn 13: 9781409975076.

Beeker, Emmet R., III. and Ernest H. Page. 2006. A Case Study of the Development and Use of a Mana-Based Federation for Studying US Border Operations. Proceedings of the *2006 Winter Simulation Conference.* Monterey, CA.

Bonabeau, Eric. 2002. Agent-Based Modeling: Methods And Techniques For Simulating Human Systems. *PNAS* 99(3). <http://www.redfish.com/research/bonabeau_abm.pdf>

Bernstein, Raymond Jr., Richard Flores, and Michael Starks. 2006. Objectives and Capabilities of the System-of-Systems Survivability Simulation (S4). *Army Research Laboratory Technical Report*, ARL-TN-260: White Sands Missile Range, NM. <http://www.arl.army.mil/arlreports/2006/ARL-TN-260.pdf>

Bock, Conrad. 2000. Goal-Driven Modeling. *Journal of Object Oriented Programming* 13:5. Repr., New York, NY: 101 Communications.

Brouse, Edwin C. 2006. The US Army and Constabulary Operations. School of Advanced Military Studies Monograph, US Army Command and General Staff College. Fort Leavenworth, KS. <http://cgsc.cdmhost.com/cgi-bin/showfile.exe?CISOROOT=/p4013coll3&CISOPTR=714>

Cioppa, Thomas M., T. W. Lucas, and S. M. Sanchez. 2004. Military Applications of Agent-Based Simulations. Proceedings of the *2004 Winter Simulation Conference*.

Cheng, Betty H.C. 2007. *Goal Modeling*. [Presentation]. Michigan State University, Department of Computer Science and Engineering. East Lansing, MI. [Online] Available from <http://www.cse.msu.edu/~chengb/RE-491/Lectures/06-goal-modeling.pdf>

Cohen, Philip R. and Hector J. Levesque. 1990. Intention is Choice with Commitment. In *Artificial Intelligence* 42: 213-261.

Daft, Richard L. 2006. *Organization Theory and Design*, 9th ed. Cincinnati, OH: South-Western Educational Publishing.

Cohen, M.S., J. T. Freeman, and S. Wolf. 1996. Meta-Recognition In Time-Stressed Decision Making: Recognizing, Critiquing, and Correcting. *Journal of the Human Factors and Ergonomics Society* 38(2): 206-219.

Davidson, J., A. Pogel, and J.A. Smith. 2008. The Role of Battle Command in Information System Assessments. Proceedings of the *13th Annual International Conference on Industrial Engineering Theory, Applications, and Practice*.

Davidson, J., A. Pogel, J. Harikumar, and C. Lewis. 2006. *Decsion Making Processes: Task 6*. S4 Project deliverable of the Physical Science Laboratory of New Mexico State University, produced under contract to the Army Research Laboratory's Survivability, Lethality, Analysis Directorate. Las Cruces, NM.

Doubler, Michael D. 1988. *Busting the Bocage: American Combined Operations in France, 6 June31–July 1944*. US Army Command and General Staff College, Combat Studies Institute. Fort Leavenworth, KS.

Elliott, Taryn. 2005. Expert Decision-Making in Naturalistic Environments: A Summary of Research. *Defence Science and Technology Organisation (DSTO) Publications*, Land Operations Division, Systems Sciences Laboratory: Edinburgh, South Australia.

Endsley, Mica. 2000. Theoretical Underpinnings of Situation Awareness: A Critical Review. In *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fellows, Susan, P. Pearce, and C. Pernin. 2004. Wargame Infrastructure and Simulation Environment (WISE) Formation Wargame (FW): Command and Control Validation Activities. In *UK Operational Research Society Simulation Workshop (SW2004)*. Birmingham, UK: Hornton Grange.

Ghazal, T., R. Morley, E. Terry, and R. Klingaman. 2003. Soldier Tactical Mission System Combat Effectiveness. Proceedings of the *2003 IEEE Systems and Information Engineering Design Symposium*. Charlottesville, VA.

Girodano, John C., Paul F. Reyhnolds Jr., and David C. Brogan. 2004. Exploring the Constraints of Human Behavior Representation. Proceedings of the *2004 Winter Simulation Conference*.

Gonzalez, Avelino J., Jose Castro, and William Gerber. 2005. Automating the Acquisition of Tactical Knowledge of Military Missions. *Journal of Defense Modeling and Simulation* 2(3): 145-160. <http://www.scs.org/pubs/jdms/vol2num3/Gonzalez-Castro-pp145-160.pdf>

Grossman, David A. 1993. Defeating the Enemy's Will: The Psychological Foundations of Maneuver Warfare. In *Maneuver Warfare: An Anthology*. Novato, CA: Presidio Press, 142-187.

Hudak, D., J. Mullen, and A. Pogel. 2008. Determining the Impact of Information on Decision-Making in Contexts Lacking Well-Defined Utility Functions. Proceedings of the *13th Annual International Conference on Industrial Engineering Theory, Applications and Practice*.

Ilchanski, Andrew. 1997. *Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-Life Approach to Land Warfare.* Center for Naval Analyses, Alexandria, VA.

Institute of Electrical and Electronics Engineers, IEEE. 1998. IEEE Recommended Practice for Software Requirements Specifications, IEEE Std. 830-1998. *Software Engineering Standards Committee of the IEEE Computer Society*: New York, NY.

Kelton, David W. 2000. Experimental Design for Simulation. Proceedings of the *2000 Winter Simulation Conference*.

Kewley, Robert H. Jr. 2004. Agent-Based Model of Auftragstaktik: Self Organization in Command and Control of Future Combat Forces. Proceedings of the *2004 Winter Simulation Conference.*

Klein, G. 1993. A Recognition-Primed Decision (RPD) Model of Rapid Decision Making. In *Decision Making in Action: Models and Methods.* Norwood, NJ: Ablex Publishing Corporation, 138-147.

Kraker, Klaas Jan de., Philip Kerbusch, and Erik Borgers. 2009. Re-usable Behavior Specifications for Tactical Doctrine. Proceedings of the *18th Conference on Behavior Representation in Modeling and Simulation* [paper: 09-BRIMS-006, pp. 15-22], Sundance, Utah.

Lauren, Michael K. and Roger T. Stephen. 2002. Map-Aware Non-Uniform Automata (MANA)–A New Zealand Approach to Scenario Modelling. *Journal of Battlefield Technology* 5(1).

Law, Averill M. and David W. Kelton. 2000. *Simulation Modeling and Analysis*, 3rd Ed., New York, NY: McGraw Hill.

Miller, N. L. and L. G. Shattuck. 2004. A Process Model of Situated Cognition in Military Command and Control. Proceedings of the *2004 Command and Control Research and Technology Symposium (2004 CCRTS)*, San Diego, CA.

Miner, John, B. 1982. *Theories of Organizational Structure and Process*. Hillsdale, IL: Dryden Press.

Mintzberg, Henry. 1979. *The Structuring of Organizations: A Synthesis of the Research.* Englewood Cliffs, NJ: Prentice-Hall.

Mintzberg, Henry. 1981. Organization Design: Fashion or Fit. *Harvard Business Review*, Jan-Feb.

Moffat, James and Susan Fellows. 2008. Looking to the Future: Representing Higher Level Planning in a Wargaming or Synthetic Environment. *The Defence Science and Technology Laboratory.* Farnborough, UK.

Moffat, James and Colin R. Mason. 2000. Representing the C2 Process in Simulations: Modelling the Human Decision-Maker. Proceedings of the *2000 Winter Simulation Conference*. Orlando, FL.

Moffat, James. 2007. Modelling Human Decision-Making in Simulation Models of Conflict. *The International C2 Journal* 1(1): 31-60.

Military Operations Research Society (MORS) Experimentation Community of Practice. 2008. *Experimentation Lexicon,* Version 1.0. <http://morsnet.pbworks.com/f/Experimentation+Lexicon+V1.1.doc>

Nissen, Mark, E. and Raymond R. Buettner. 2004. Computational Experimentation with the Virtual Design Team: Bridging the Chasm between Laboratory and Field Research in C2. Proceedings of the *2004 CCRTS*, San Diego, CA. <http://www.dodccrp.org/events/2004_CCRTS/CD/papers/162.pdf>

North Atlantic Treaty Organization, NATO. 2002. *NATO Code of Best Practice for C2 Assessment.* Rev. ed. NATO Research and Technology Organisation (RTO), Research Group SAS-026. Washington, DC: CCRP. <http://www.dodccrp.org/files/NATO_COBP.pdf>

North Atlantic Treaty Organization, NATO. 2006. *Exploring New Command and Control Concepts and Capabilities.* NATO Research and Technology Organization (RTO), SAS-050 Final Report. <http://ftp.rta.nato.int/public//PubFullText/RTO/TR/RTO-TR-SAS-050///$$TR-SAS-050-ALL.pdf>

Pearce, Paul, Alan Robinson, and Susan Wright. 2003. The Wargame Infrastructure and Simulation Environment (Wise). In *Knowledge-Based Intelligent Information and Engineering Systems, 7th International Conference*.

Perrow, Charles. 1986. *Complex Organizations: A Critical Essay*. 3rd ed. New York, NY: Random House.

Pew, R. W. and A. S. Mavor. 1998. *Modeling Human and Organizational Behavior: Application to Military Simulations*. Washington, DC: The National Academies Press.

Pigeau, Ross, and Carol McCann. 2002. Re-conceptualizing Command and Control. *Canadian Military Journal Spring*: 53-63. <http://www.journal.dnd.ca/vo3/no1/doc/53-64-eng.pdf>

Pressfield, Steven. 1999. *Gates of Fire.* New York, NY: Bantam Books.

Roggio, Bill. 2009. The Long War Journal: US Launches Second Strike Outside of Pakistan's Tribal Areas. *The Long War Journal.* <http://www.longwarjournal.org/archives/2009/03/us_launches_second_s-print.php>

Slim, Viscount William. 2000. *Defeat Into Victory: Battling Japan in Burma and India, 1942-1945*. New York, NY: Cooper-Square Press.

Smith, Roger. 1999. Fundamental Principles of Modeling and Simulation. Proceedings of the *1999 Computer Game Developer's Conference*, San Jose, CA.

Straver, M.C., Vincent, E. and Fournier, P. 2006. Experiences with the MANA Simulation Tool. Defence Research and Development Canada (DRDC), Valcartier. [Technical Memorandum]. <http://pubs.drdc.gc.ca/PDFS/unc51/p525762.pdf>

Thompson, James D. 1967. *Organizations in Action*. New York, NY: Mcgraw-Hill.

United States Army Field Manual, FM 3-90. 2001. *Tactics*. Washington, DC: Headquarters, Department of the Army.

United States Army Field Manual, FM 1-02. 2004. *Operational Terms and Graphics*. Washington, DC: Headquarters, Department of the Army.

van Creveld, M. L. 1985. *Command in War.* Cambridge, MA: Harvard University Press.

Waite, W. F. 2004. Fundamental Concepts of Modeling. Proceedings of the *36th Southeastern Symposium on System Theory, IEEE*: 366-373.

Widder, W. 2002. Auftragstaktik and Innere Führung: Trademarks of German Leadership. *Military Review* 82(5): 3-9

Wiener, Earl L. and David C Nagel. 1988. *Human Factors in Aviation*. San Diego, CA: Academic Press.

Yang, A., N. Curtis, H. A. Abbass, R. Sarher, and M. Barlow. 2006. WISDOM-II: A Network Centric Model for Warfare. In *Complex Science for a Complex World: Exploring Human Ecosystems with Agents*. The Australian National University Press, ch. 8: 149-173. <http://epress.anu.edu.au/cs/pdf/ch08.pdf>

Yang, A., H. A. Abbass, and R. Sarker. 2004. Landscape dynamics in multi-agent simulation combat systems. Proceedings of the *17th Joint Australian Conference on Artificial Intelligence*, LNCS. Cairns, Australia: Springer-Verlag.