Measuring the Value of High Level Fusion

Louis R. Moore, III
The RAND Corporation
1700 Main Street
PO Box 2138
Santa Monica, CA 90407-2138
310-393-0411 x7578
Louis_Moore@rand.org

Daniel Gonzales
The RAND Corporation
1200 Hayes Street
Arlington, VA 22202-5050
703-413-1100
Daniel_Gonzales@rand.org

ABSTRACT

In most current ground force combat simulations, the operational movements and command intent of forces follow prescribed, inflexible objectives and plans. Because of this limitation, the value of advanced intelligence, surveillance, and reconnaissance (ISR) and high-level fusion is reflected only in better targeting and not in improved operational-level command and control (C2). RAND has developed an agent interaction-based constructive simulation called the Ground C4-ISR Assessment Model (GCAM) to help examine the contributions of C4-ISR to ground forces. In GCAM simulated ground force commanders make decisions on the basis of shared awareness derived from information produced by Level 1 (Identify/Correlate), Level 2 (Aggregate/Resolve), and Level 3 (Interpret/Determine/Predict) fusion processes. In this way simulated ground commanders can adapt their plans in response to perceived changes in enemy capability, activity, or intent, or to perceived changes to the battlefield environment. This paper details the representation of high-level fusion processes used in GCAM and developed with the support of the U.S. Army Model Improvement Program. Those processes are modeled using the Assistant Secretary of Defense for Networks and Information Integration (ASD-NII) Decision Support Center (DSC) Multi-INT fusion study "Knowledge Matrix" methodology. The information or knowledge added by high-level fusion and analysis of raw sensor data from multiple sources is represented in this methodology by increased information quality levels for activity, capability, and intent. This research will allow military analysts to demonstrate the utility and the relative importance of improved C2 and high-level fusion capabilities for Army and Joint forces.

1. Introduction

In most current ground force combat simulations, the operational movements and command intent of forces follow prescribed, inflexible objectives and plans. Because of this limitation, the value of advanced intelligence, surveillance, and reconnaissance (ISR) and high-level fusion is reflected only in better targeting and not in improved operational-level command and control (C2). To help examine the contributions of C2, Communications, Computers, and ISR (C4ISR) to ground forces the RAND Corporation has developed an agent interaction-based constructive simulation called the Ground C4ISR Assessment Model (GCAM).

GCAM is a time-stepped, multi-sided, stochastic simulation of combat in a theater context with representations of C2 at multiple echelons on all sides. GCAM is intended to support the analysis of C4ISR issues. Figure 1.1 depicts many of the features of GCAM entities.

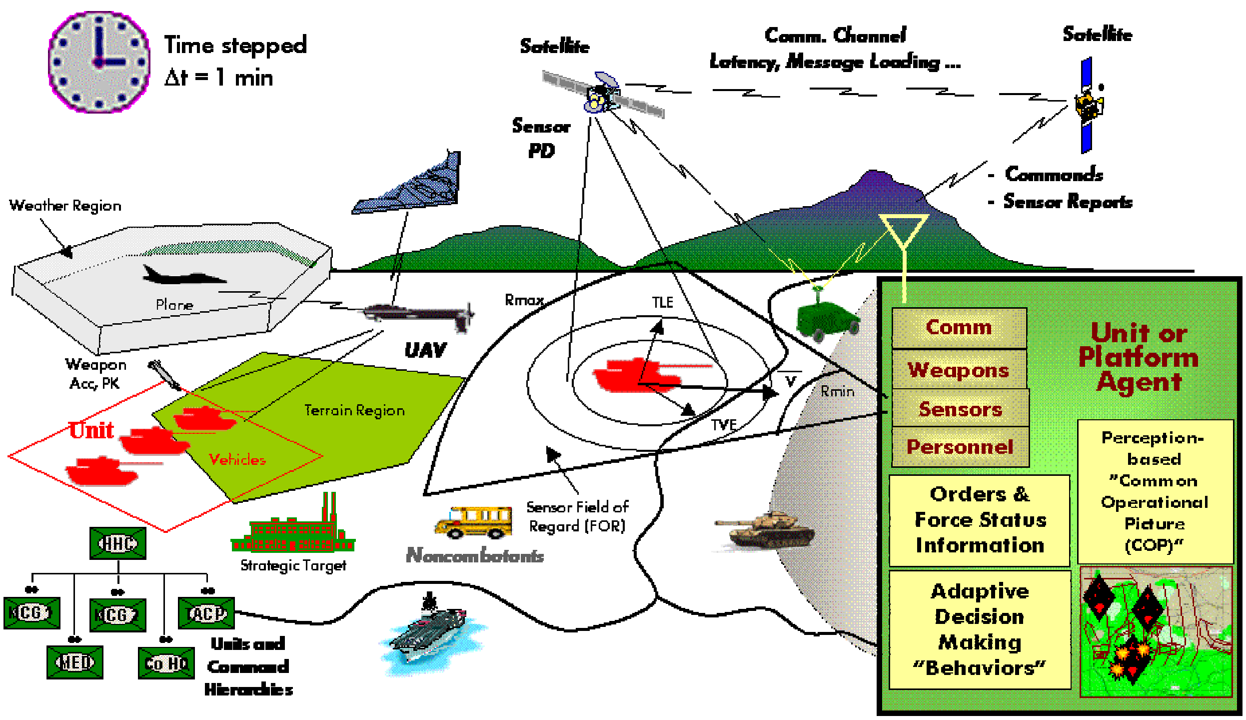

Figure 1.1 Depictions of GCAM Entities

Areas or volumes having common characteristics represent the terrain and environment probabilistically. Agents in GCAM may be platforms, such as satellites, aircraft, ships, vehicles, school buses, etc. or units existing in a command hierarchy. Agents are governed by their own individual behaviors and external orders from superiors. Agents may mount weapons and sensors and communicate with other agents over communications channels. In GCAM, agents, playing the role of commanders, make decisions on the basis of a perceived individual and shared battle space awareness called the common operational picture (COP) [1]. These decision-making agents start with an initial course of action (COA) and maneuver plans along with their own perceived Intelligence Preparation of the Battlefield (IPB). The agents use their own organic sensors, communications with other entities, and fusion processes to derive their perceptions. In this the way those agents can adapt their plans in response to perceived changes in the enemy or to the battlefield environment.

If we characterize combat simulations with respect to their representation of decision-making and modeling approach, Figure 1.2 presents where some current combat simulations would lie with respect to each other and GCAM. The arrows in Figure 1.2 try to capture where the simultion in question is heading in its development. Distilled decision making features simple rules governed by exogenous influences. Rational decision-making represents deliberate planning based on internal perception and complex sequences of behaviors. The equation-based simulation approach concentrates on the integration of observables (equations) through time. Equation-based simulations frequently involve the solution of large difference equations such as those commonly called Lanchester equations [2]. On the other hand, in agent-interaction based simulations overall patterns emerge from outcomes that evolve through the interactions and adaptations of many agents. These patterns are significantly more complex than the behaviors of the individual agents would imply. In [3] agent interaction-based simulations using distilled decision-making were called "Dot Wars.

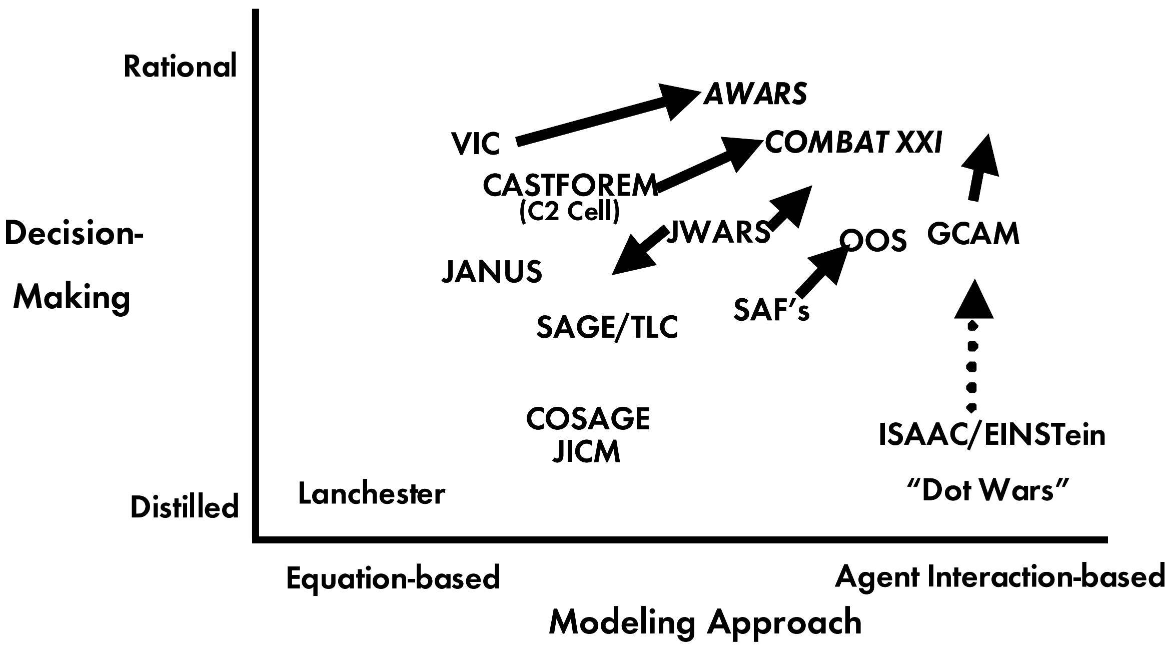

Figure 1.2 Relationship of GCAM to Other Simulations

Modern network enabled warfare phenomena are not well understood but may be characterized by the dynamic interactions of autonomous information-oriented decision-making entities. Therefore, rational agent interaction-based modeling seems natural. GCAM is built upon the System Effectiveness Analysis Simulation [4] (SEAS), a toolkit for building such simulations. The quality of GCAM agent decisions is a function of the quality and the accuracy of the perceived situational awareness available to them. Thus, analysts may assess the impact of the quality of information produced by the ISR and fusion architectures on the command decision-making process and ultimately on combat itself examining the outcomes of battles fought with alternative architectures. An adequate representation of the fusion process is key to this evaluation. This paper details the representation of high-level fusion processes used GCAM.

2. Fusion

Fusion may be defined as "a series of processes performed to transform observational data into more detailed and refined information, knowledge, and understanding [5]." The fusion process observes significant events and battlefield entities performing various actions. We distinguish in our model between four general types of entities: infrastructure and facilities (buildings, roads, bridge etc.), pieces of equipment (tanks, trucks, etc.), aggregates (units, collections, organizations, etc.), and elements with structured relationships such as an order of battle. Based upon the Joint Directors of Laboratories (JDL) Data Fusion Model [6] first proposed in 1985 under the guidance of the Department of Defense (DoD), we partition fusion into six, not necessarily sequential, levels.

Level 0 fusion organizes discrete pieces of sensed data into forms that can be used by the process. Level 1 fusion processes sensed data to identify that discrete entities or events have been observed, correlates and combines like information, and resolves information conflicts. The output from Level 1 fusion is a set of discrete observed battlefield entities with information about the type, location, movement, identity, status, and capability, along with an evaluation of the quality of this information.

Based upon the Level 1 products, IPB order of battle templates, and knowledge about the environment, Level 2 fusion aggregates discrete entities into larger objects that are interacting. It interprets entity events and actions and hypothesizes what events may occur next. The outputs from Level 2 are aggregated as well as inferred entities, observed and derived events and actions, and a collection of hypotheses on what events will happen in the future. In addition, quality assessments of these products will be available.

Level 3 fusion projects current situation into the future to predict intent and courses of action. Level 4 assesses and controls the fusion process to drive improvements in the process. It allocates resources to satisfy collection needs. Level 5 fusion provides the fusion process with customer feedback and control through visualization of the fusion products and determination of Priority Intelligence Requirements (PIR).

Many simulations represent the Level 1 fusion of location and movement information for single entities with an eye to targeting them with weapons. The representation described is this paper concentrates on higher-level (Level 1, 2, and 3) processes. Separate representations of Level 4 and 5 fusion would use the results of the higher-level process representation.

The fusion process in current and future Army intelligence units the will be conducted by five foundational domains - four single-source and one integrating.

1. The Imagery and Geospatial Information Domain handles Imagery Intelligence and Geospatial Information.
2. The Signatures Domain handles Measurement and Signatures Intelligence.
3. The Signals Domain handles Signals Intelligence.
4. The Human Domain handles Human Intelligence and Counterintelligence.
5. The Integrating Domain performs all source analysis by integrating, planning, guiding, and tasking all other domains in conducting all intelligence tasks.

Observations are assumed to be the result of Level 0 Fusion as part of the sensing process. As indicated in Figure 2.1 observations may be then processed by one of the single source domains and then the integrating domain or directly by the integrating domain. Fusion levels 1 through 5 occur in each domain.

3. Knowledge Matrix

One purpose of fusion is to reduce the amount of data on the COP. But the product of a fusion process should also have higher quality than the input to the process. Thus, to represent fusion adequately in a simulation we must represent not only the product but also its quality. Our goal is to represent the fused observation and also the improvement in the quality of that observation added by the fusion process.

The "Knowledge Matrix" [7] is a tabular framework for capturing the quality of a piece of observational data. Each column of the matrix represents a different type of knowledge (Location, Track, Identity, Activity, Capability, Intent, etc.) about the data. Within each column, each row represents decreasing levels of quality for the type of knowledge represented by the column. Figure 3.1 presents an example of a knowledge matrix with descriptions of the levels of quality.

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	5 m.	Vectors & Patterns	Object Hierarchy	Precise Actions	All Elements	All Objectives
4	10 m.	Vectors	Object	Specific Actions	Many Details	Major Objectives
3	20 m.	Velocity	Classify	Identifiable Actions	Some Details	Primary Objectives
2	100 m.	Toward or Away	Categorize	Single Action	General Information	General Objectives
1	1 km.	Moving	Discriminate	Unidentifiable Actions	Minimal Information	Single Objective
0	10 km	Detect	Detect	Detect	Detect	Detect

Figure 3.1 Knowledge Matrix Cell Descriptions

The entry in each cell of the knowledge matrix for a piece of data is the likelihood that the data achieves the level of quality or better for that cell. It depends on the sensor or process that generated the piece of data, the environment, and the target itself. This information may be derived from sensor characteristics, expert opinion, or played parametrically. A knowledge matrix may be used to portray the quality of any piece of information about battlefield entities. We distinguish between four types of entities: infrastructure (buildings, roads, bridge etc.), equipment (tanks, trucks, etc.), aggregates (units, collections, etc.), and elements of an order of battle.

Figure 3.2 is an example of a knowledge matrix. If we were to adopt a 90% threshold for quality this data would indicate, see the shaded cells, a vehicle moving away, say, with ~100 meter location error, general capability and unknown activity and intent. In this example and extra row of ones has been appended to emphasize that each column can be treated as a cumulative distribution function for a discrete random variable. In the remainder of this paper if this last row is not explicitly stated it should be assumed to exist. In addition, for knowledge matrix G, we will refer to the likelihood in the cell for Quality Level i of column j as $G_{i,\,j}$.

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	0.0	0.0	0.0	0.0	0.0	0.0
4	0.3	0.0	0.0	0.0	0.0	0.0
3	0.8	0.0	0.0	0.0	0.0	0.0
2	0.9	0.7	0.9	0.0	0.0	0.0
1	0.95	0.9	0.95	0.0	0.9	0.0
0	0.99	0.95	0.98	0.0	0.92	0.0
-1	1.0	1.0	1.0	1.0	1.0	1.0

Figure 3.2 Example Knowledge Matrix

Suppose a scout has been assigned to watch a specific Named Area of Interest (NAI). Vehicles observed in that NAI are highly likely to be following a specific Course of Action (COA). Figure 3.3 presents a possible knowledge matrix for a detection of such a vehicle.

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	0.0	0.0	0.0	0.0	0.0	0.0
4	0.3	0.0	0.0	0.0	0.0	0.2
3	0.8	0.0	0.0	0.0	0.0	0.6
2	0.9	0.7	0.9	0.0	0.0	0.9
1	0.95	0.9	0.95	0.0	0.9	0.9
0	0.99	0.95	0.98	0.0	0.92	0.9

Figure 3.3 Knowledge Matrix for a NAI

Most simulations, both stochastic and deterministic, have and use the information available in a knowledge matrix. In a stochastic simulation, if G is the knowledge matrix for an entity that has been detected, for each Knowledge Type j, one would generate a uniform random variable, u_j, and find the sampled Quality Level, $q_j = G_j^{-1}(u_j)$, where $G_j^{-1}(x)$ is the highest Quality Level i such that $x \le G_{i,j}$. That sampled quality level becomes the perceived knowledge about the entity being observed and would be reported for use by the other processes in the simulation. In addition, the vector of sampled uniform random variables, $u = (u_j)$, is associated with the observation and knowledge matrix.

In a deterministic simulation, if n_k entities of type k are detected, then $n_k G_{i,j}$ type k entities are reported at Quality Level i of Knowledge Type j. For example, if a sensor in a deterministic simulation detects 10 tanks and Figure 3.3 is the knowledge matrix for those tanks then 0, 3, 8, 9, 9.5, and 10 tanks would be reported with location errors less than 5 m., 10 m., 20 m., 100 m., 1 km., and 10 km., respectively. The vector of detections, $n = (n_k)$, would be associated with the knowledge matrix. Unfortunately, many simulations fail to associate the knowledge matrix information with the observation data after the data is produced. As we will see in the next subsection maintaining the knowledge matrix is central to our representation of fusion.

Our representation of the fusion process involves four steps. The first is to "age" the knowledge matrices. Then candidates for fusion must be determined. The third step combines the observations and their associated knowledge matrices. The last step infers quality improvements that are supported by the information in the fused knowledge matrix.

4. Knowledge Matrix Aging

The quality of a piece of data becomes less accurate as time advances. This is due to the propensity of battlefield entities to unpredictably change their state or behavior. Stationary entities begin to move. Moving entities stop. Infrastructure is destroyed and rebuilt. We thus need to "age" the likelihoods in a knowledge matrix over time. If knowledge matrix $G(t_1)$ was obtained at time t_1 then the aged knowledge matrix for time t_2, $G(t_2)$, is formed by multiplying each $G(t_1)_{i,j}$ by $\exp\{R_{i,j}(t_1 - t_2)\}$ for some value of $R_{i,j}$ that depends on the quality level, type of knowledge, type of target, and the environment. Possible values for $R_{i,j}$ are given in [6]. For example, $R_{i,1}$ is 0.069 per minute for moving vehicles for all quality levels indicating that the likelihood for the quality levels of the location of a moving vehicle halve every 10 minutes.

In addition, the standard deviations, σ, of the circular location errors for moving entities grow proportionately with time even though those entities do not change their direction or behaviors. That is, Δt after a moving entity has been detected $\sigma^2 = TLE^2 + TVE^2\,\Delta t^2$ where TLE is the target location circular error standard deviation and TVE is the target velocity circular error standard deviation for the detection. The likelihoods for the location of moving battlefield entities should also be recomputed as well as aged. The quality associated to the location moving entities may diminish quickly.

5. Determining Fusion Candidates

The second step in the representation of fusion [7] is to determine if two observations are candidates for fusion. Two observations of same type battlefield entities are candidates for fusion if their aged Quality Level 3 likelihoods for Location knowledge must meet some threshold, β, and their estimated locations are close. That is, the square of the distance between their estimated locations is less than $\{\sigma_1 2 + \sigma_2 2\}\chi^2(\alpha)$ where σi is the standard deviation of the circular location error for observation i and $\chi^2(\alpha)$ is the value that a chi-squared variable with two degrees of freedom exceeds with probability α. The parameters α and β vary by type of entity. For infrastructure observations β may equal 25% while for equipment observations β may equal 75% or 90%. Usually α equals 10% or 5%.

The knowledge matrices for each observation must be scored using the threshold value, β, from above. For Location knowledge, if $G_l^{-1}(\beta)$ is level 5,4,3, or 2, the score for Location is 12. If $G_l^{-1}(\beta)$ is level 1, 0, or –1 the score for Location is 8, 4, or 2.4, respectively. For Identification and Tracking (if moving) knowledge, if $G_l^{-1}(\beta)$ is level 5 or 4, the score is 12. If $G_l^{-1}(\beta)$ is level 3, 2, 1, 0, or –1 the score is 9.6, 7.2, 4.8, 2.4, or 0, respectively. The score for Identification and Tracking is approximately 2.4 times the expected quality level for the column. We use this modified scoring method since it considers more of the information in the column.

For each of the knowledge matrices sum the scored values and divide by either 24 or 36 (if moving). This yields a correlation probability, $P1$ or P_2, for knowledge matrix 1 or 2. The likelihood that both knowledge matrices come from observations of the same entity is $P_1 \blacksquare P_2$. In a stochastic simulation the two observations will be fused if a sampled uniform random variable is less than $P_1 \blacksquare P_2$. In a deterministic simulation the two observations will be fused if $P1 \blacksquare P2$ is greater than 50%. If the fusion test fails we will pick the observation with the highest correlation probability. The last step allows us to discard observations with poor quality. In [7] only observations for aggregate or order of battle entity types are scored. We score the observations for all entity types.

6. Combining Knowledge Matrices

The third step in the representation of fusion [7] is to combine the observations to be fused along with their knowledge matrices. Suppose two observations with associated knowledge matrices G and H have been selected for fusion. We will assume that both observations have been obtained in a probabilistically independent fashion, and the knowledge matrix of the fused observation will not contain likelihood values less than those of G or H.
An implicit formulation for the fused knowledge matrix, $F=(F_{i,j})$ may be derived as

$$F_{i,j} = 1 - (1 - G_{i,j})(1 - H_{i,j}).$$

Alan Steinberg suggested this equation as an extension to The Dempster-Schafer Theory of Evidence [8]. For a stochastic simulation, if $g=(g_i)$ and $h=(h_i)$ are the uniform sampling vectors associated with G and H respectively, the fused vector $f=(f_j)$ is

$$f_j=1-(1-g_j)(1-h_j)[1-\ln\{(1-g_j)(1-h_j)\}].$$

The sampled quality levels for each type of knowledge, j, may be obtained from $F^{-1}(f_j)$. This is an implicit approximation since using the actual fusion process would produce an observation with a knowledge matrix containing likelihoods no less than those of F. Preliminary experiments seem to indicate the difference between explicit and implicit likelihoods is not great. If one wishes to substitute for the implicit computation of f_j the calculation from an explicit fusion process, say, using a Kalman filter [9] for Location knowledge, one may so. The appropriate entries in the knowledge matrix should then be recomputed. Either way the resulting fused observation will have a knowledge matrix better than the either of the fused observations.

Suppose Figure 3.3 is the knowledge matrix, G, for one observation and Figure 6.1 is the knowledge matrix, H, for an observation on a commander's common operational picture (COP)[1][10].

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	0.0	0.0	0.1	0.3	0.0	0.0
4	0.2	0.0	0.6	0.4	0.0	0.0
3	0.7	0.0	0.8	0.5	0.7	0.0
2	0.9	0.7	0.9	0.6	0.8	0.3
1	0.95	0.8	0.95	0.7	0.9	0.4
0	0.99	0.9	0.98	0.8	0.92	0.5

Figure 6.1 COP Observation Knowledge Matrix

Further suppose that both are candidates for fusion and that the fusion test has been passed. The fused knowledge matrix, F, is presented in Figure 6.2.

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	0.0	0.0	0.1	0.3	0.0	0.0
4	0.44	0.0	0.6	0.4	0.0	0.2
3	0.94	0.0	0.8	0.5	0.7	0.6
2	0.99	0.91	0.99	0.6	0.8	0.93
1	0.998	0.98	0.998	0.7	0.99	0.94
0	0.9999	0.995	0.9999	0.8	0.994	0.95

Figure 6.2 Fused Knowledge Matrix (F)

If the uniform sampling vector for G is g=(0.75, 0.84, 0.9, 0.6, 0.52, 0.3) and for H is h =(0.38, 0.24, 0.6, 0.52, 0.48, 0.3) then the fused uniform sampling vector is f=(0.56, 0.62, 0.83, 0.49, 0.40, 0.16). The sampled quality level vector of the resulting fused observation is $F^{-1}(f)$ =(3,2,2,3,3,4). That is, the observation used by the simulation will have a 20 meter circular location error, with known actions, capability, and primary objective. The Identity of the observation can only be classified as tracked or wheeled in the simulation.

7. Knowledge Matrix Inference

At this point we have obtained our goal to represent not only how observations are fused but also how that fusion improves the quality of that observation. But more is possible in the fourth step of the process. The likelihoods in the fused knowledge matrix, and thus the information reported to the simulation, may be improved using inference rules patterned after expert system judgment. Reference [7] presents a set of rules derived from the U.S. Army All Source Analysis System (ASAS). For example, one of these rules states that if, at the 90% threshold, the quality level for Location is 3, i.e., $F_{Loc}(3) \geq 0.9$, and the quality level for Track is 2, then the quality levels for Identity, Activity, and Capability at the 90% threshold may be increased to 3, 2, and 2, respectively. The resulting inferred knowledge matrix is presented in Figure 7.1. The improved cells have been shaded.

	Type of Knowledge					
Quality Level	Location	Track	Identity	Activity	Capability	Intent
5	0.0	0.0	0.1	0.3	0.0	0.0
4	0.44	0.0	0.6	0.4	0.0	0.2
3	0.94	0.0	0.9	0.5	0.7	0.6
2	0.99	0.91	0.99	0.9	0.9	0.93
1	0.998	0.98	0.998	0.9	0.99	0.94
0	0.9999	0.995	0.9999	0.9	0.994	0.95

Figure 7.1 Inferred Knowledge Matrix

Now the Identity of the vehicle revealed to the simulation using $F^{-1}(f)$ may be classified. This would be an important improvement if this vehicle were actually a civilian automobile or an armored personnel carrier. Other such inference rules are implemented in GCAM.

8. Summary

This paper has presented the underlying methodology for the representation of fusion that has been implemented in RAND's GCAM simulation of ground combat. It is based upon the concept of a knowledge matrix that captures the quality of an observation. It is applicable to infrastructure, equipment, aggregate, and order of battle observation types. In addition, the implicit fusion algorithm we have presented allows the quality improvements due to fusion to be represented and used by decision-making entities in the simulation.

9. References

[1] TP525-3-0.1, “Objective Force Battle Command (C4ISR) Concepts,” TRADOC, October 31, 2002.

[2] James G. Taylor, “Lanchester Models of Warfare, Volumes I and II,” Operations Research Society of America, 1983.

[3] Vincent P. Roske, Jr, “Opening Up Military Analysis: Exploring Beyond The Boundaries,” Phalanx, June, 2002.

[4] SEAS homepage, January 07, 2004, http://www.teamseas.com (as of March 12, 2004)

[5] “Objective Force Fusion,” Directorate of Combat Developments, U.S. Army Intelligence Center, Fort Huachuca, AZ, March 14, 2003.

[6] Alan N. Steinberg, Bowman, C.L. and F.E. White, Jr., “Revisions to the JDL Data Fusion Model,” Proceedings of the 3rd NATO/IRIS Conference, Quebec City, Canada, 1998.

[7] Hans Keithley, “Multi-INT Fusion Performance Study,” DCS-00-02, Joint C4ISR Decision Support Center, 2000.

[8] G. Shafer, “A Mathematical Theory of Evidence,” Princeton Univ. Press, Princeton, NJ, 1976.

[9] P. S. Maybeck, “The Kalman Filter – An Introduction for Potential Users,” TM-72-3, Air Force Flight Dynamics laboratory, Wright-Patterson AFB, Ohio, June 1972.

[10] DCGS-A ORD, Version 1.0, October, 2003.