Specifying Autonomy Levels for Unmanned Systems: Interim Report

Hui-Min Huang, James Albus, Elena Messina
National Institute of Standards and Technology
Intelligent Systems Division
Gaithersburg, Maryland 20899-8230
hui-min.huang@nist.gov,
james.albus@nist.gov,
elena.messina@nist.gov

Robert Wade
AMRDEC, TUV/SED
Huntsville, Alabama 35801
robert.wade@us.army.mil

Woody English
Titan Systems Corporation
Huntsville, Alabama 35801
woody.english@titan.com

ABSTRACT

The viability of Unmanned Systems as tools is increasingly recognized in many domains. As technology advances, the autonomy on board these systems also advances. In order to evaluate the systems in terms of their levels of autonomy, it is critical to have a set of standard definitions that support a set of metrics. As autonomy cannot be evaluated quantitatively without sound and thorough technical basis, the development of autonomy levels for unmanned systems must take into account many factors such as task complexity, human interaction, and environmental difficulty. An *ad hoc* working group assembled by government practitioners has been formed to address these issues.

The ultimate objectives for the working group are:

- To determine the requirements for metrics for autonomy levels of unmanned systems.
- To devise methods for establishing metrics of autonomy for unmanned systems.
- To develop a set of widely recognized standard definitions for the levels of autonomy for unmanned systems.

This paper describes the interim results that the group has accomplished through the first four workshops that the group held. We report on the initial findings of the workshops toward developing a generic framework for the Autonomy Levels for Unmanned Systems (ALFUS).

1. INTRODUCTION

Unmanned vehicles have been fielded in several domains in the recent past, ranging from battlefields to Mars. Most major efforts have been funded by various U. S. Government agencies. As the number of programs for developing unmanned systems (UMS) accelerates within government, there is a growing need for characterizing these systems. Individual government agencies have begun these efforts. The Department of Defense Joint Program Office (JPO), the U.S. Army Maneuver Support Center, and National Institute of Standards and Technology (NIST) have, in separate but related efforts, described levels of robotic behaviors for the Army Future Combat Systems (FCS) program [1, 2, 3]. The Air Force Research Laboratory (AFRL) has established Autonomous Control Levels (ACL) [4]. The Army Science Board has described a set of levels of autonomy [5]. It is imperative that these and other agencies leverage each other's efforts and aim at a government wide consistent approach.

Technologic advances are providing a fertile landscape for rapid evolution of mobile robotic technology. Military and civilian agencies continue to expand the role that unmanned systems serve. As Government agencies continue to specify unmanned system capabilities for future systems, numerous applicable standards will evolve to address developmental and life-cycle issues including cost, time to market, interoperability and terminology. There are increasing demands for common terminology for describing an unmanned system and standard metrics for evaluating the UMS. Recognizing the needs, practitioners from Departments of Commerce, Defense, Energy, and Transportation (in alphabetical order) and

their supporting contractors assembled at NIST in July 2003 and formed[1] the ALFUS *ad hoc* working group to address these issues. More specifically, ALFUS will provide the unmanned systems community with the terminology for prescribing and evaluating the level of autonomy that an unmanned system can achieve. A web site has been established to facilitate interaction and information sharing [6].

When presented with the question "What is autonomy?" responses vary widely. At the philosophical level, choice becomes synonymous with autonomy. From a scientific perspective, intelligence will likely pervade the conversation. The ALFUS Working Group provides the following definition [7]:

> **Autonomy**
> (A) The condition or quality of being self-governing [8].
> (B) A UMS's own ability of sensing, perceiving, analyzing, communicating, planning, decision-making, and acting, to achieve its goals as assigned by its human operator(s) through designed human-robot interaction (HRI). Autonomy is characterized into levels by factors including mission complexity, environmental difficulty, and level of HRI to accomplish the mission.

2. THE ALFUS FRAMEWORK

In the short time the ALFUS Working Group has existed, it has formulated, through consensus, a framework in which the levels of autonomy can be described. The framework addresses the technical aspects of unmanned systems but could extensively affect the financial and life-cycle aspects. Thus, the Autonomy Levels for Unmanned Systems (ALFUS) framework includes the following elements:

- Terms and Definitions: A set of standard terms and definitions that support the autonomy level metrics.
- Detailed Model for Autonomy Levels: A comprehensive and detailed specification for determining the autonomy. The audience is technical users of UMSs.
- Summary Model for Autonomy Levels: A concise, scalar presentation of Autonomy Levels. The audience is executives and end users (In the DoD domain, these would include combat leadership, program managers, unit leaders, and soldiers).
- Guidelines, Processes, and Use Cases: A process to translate the detailed, technical ALFUS model into the summary model as well as guidelines to apply the generic framework to specific ALFUS models. A number of use cases may be generated to demonstrate the application processes.

Working Group consensus suggests that, at the end-user level, the autonomy level definitions should be mission specific to be most useful. At the other extreme of the spectrum, it would be very beneficial to have a generic framework that applies to the wide differences in unmanned system domains that include ground vehicles, air vehicles, undersea vehicles, surface vessels, and littoral water robots. Figure 1 demonstrates these Framework concepts. A continuing challenge to the ALFUS Working Group is to obtain a truly generic model. In the interim, the working group will develop ALFUS using a spiral software development approach. The first iteration will address the Army FCS needs.

[1] The concept was incepted by Hui-Min Huang of NIST and Bruce Clough of AFRL and their colleagues in a series of email exchanges discussing the issue.

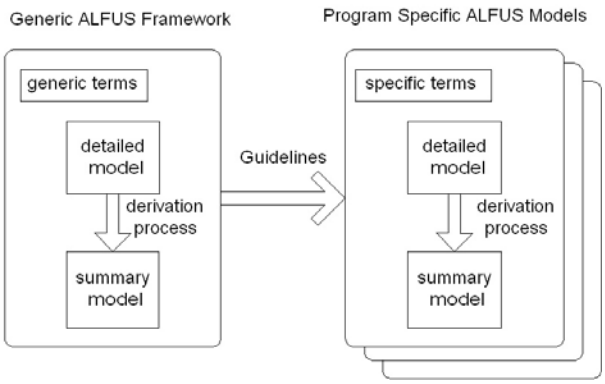

Figure 1: ALFUS framework, the construct

2.1 Terms and Definitions

Terms and definitions provide a core vocabulary necessary in the discourse of unmanned systems. They are key to establishing a common description of Autonomy Levels for Unmanned Systems. These terms will provide a basis for metrics for system performance evaluation on which the Autonomy Level for a UMS will be determined. The approaches in this terminology effort are:

- Leverage existent work and adopt existent definitions. This expedites the Group's effort in proceeding with its core objective, the autonomy level model. Modifications to the existent definitions may be necessary to fit the objectives of this working group.
- Consider the cultural factor, for example, how people are using the terms, to ensure a seamless transition of the outcome to the users.
- Establish separate sections for generic terms and domain-specific terms.

The main generic terms that we have defined or adopted to support the ALFUS framework include: autonomy, environment, fusion, human robot interface (HRI), mission planning, mode of operation, perception, robotic follower, situation awareness, task decomposition, and unmanned system (UMS). The main domain specific terms include: cooperative engagement, sensor to shooter, and unattended ground sensors. Version 1.0 of the terminology has been published [7]. The following are among the key definitions providing the basis for the ALFUS framework:

Detailed Model for Autonomy Levels. A comprehensive set of metrics that represent multiple aspects of concerns, including mission complexity, environmental difficulty, and level of HRI that, in combination, indicate a UMS's level of autonomy.

Summary Model for Autonomy Levels. A set of linear scales, 0 through 10 or 1 through 10, used to indicate the level of autonomy of a UMS. This model is derived from the UMS's Detailed Model for Autonomy Levels.

2.2 Detailed Model

The participants converged on the notion that the autonomy levels involve multiple aspects, including:

- Autonomy relates to multiple technical areas and subsystems.
- Task complexity and adaptability to environment are key aspects.
- Nature of collaboration with humans is important, including levels of involvement and different types of interaction.
- Qualitative Factors: e.g. whether and how do the following affect a UMS's autonomy levels: mission success rate, response time, and precision/resolution/allowed latencies [9].

Figure 2 demonstrates a perspective of the current thrust of the ALFUS detailed model. This model comprises three axes, namely, difficulty of the environment, complexity of the mission, and operator interaction (inversely proportional – less interaction is more autonomous). The autonomy level of a particular UMS can be represented with a triangular surface with certain values on the three axes. This model suggests vectors, as opposed to a single scale, to characterize unmanned system autonomy levels.

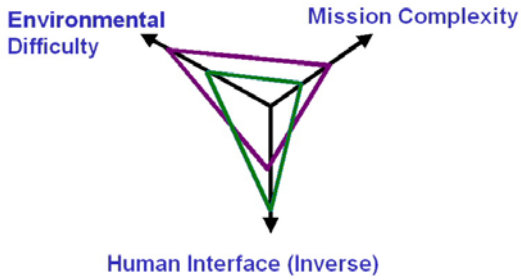

Figure 2 Autonomy level for unmanned systems, detailed model

2.3 Summary Model and Its Derivation Process

Ultimately, the intent for ALFUS is to convey high-level characteristics of Unmanned Systems to Engineering Managers, Procurement Officers, Government Officials, Corporate Leadership, and other non-technical parties. In the DoD domain, these would include combat leadership, program managers, unit leaders, and soldiers. It is important that these users are and will remain the centerpiece of our concern. The metrics that we develop should use the languages that these types of users speak and be consistent with a culture that these types of users live in. This working group should interact with users within individual programs so that the particular cultural issues are addressed adequately in considering representations.

Basing the Summary Model on the Detailed Model provides the target audience a certainty in the foundation and formulation of the level of autonomy assigned. An issue that was discussed within the working group was whether the autonomy levels should be characterized using numbers. It was suggested that modes, instead, might be better characterization for UMS autonomous behaviors. Higher autonomy may not be characterized with stepwise capability increase of equal amounts, as numbers would indicate. One example given is that the JRP Master Plan describes four modes of operations, which are not expressed in numbers. A counter argument is that, users, meaning the soldiers and combat leaders in the DoD domain, relate to numbers better. In this regard, a simple 0 (or 1) through 10 scale was the consensus.

This ultimately could be a domain dependent issue. It has been suggested that, in the automobile industry, people may be used to the notion of “classes” of the vehicles. Perhaps the autonomy levels in this domain could leverage this notion, as opposed to numbered levels. It is the intent of the working group to investigate all possible approaches. In terms of converting the detailed model to the summary model, a simple way is to add up all the metric measurements in the detailed model. Sophisticated algorithms involving statistics or logarithm could also be used.

3. CURRENT FOCUS

Having established the general approach for the definition of the Autonomy Levels for Unmanned Systems, the working group is now faced with the challenge of establishing the comprehensive set of metrics to support the Detailed Model for Autonomy Levels. At this time the Detailed Model is composed of three primary groupings: *Mission Complexity, Environmental Difficulty, and the Level of HRI*. It is the intent of the working group to populate each group with an exhaustive set of measurable factors. These factors, in addition to a formalized approach for their analysis (*statistical analysis, "weighted average", distribution*), will aid in the formulation of a consistent and repeatable measure(s) of "autonomy" for communicating functionality within the unmanned systems framework.

3.1 Mission Complexity

It was a consensus view that the autonomy levels for particular UMSs are specified and evaluated according to the missions and tasks that the systems are capable of performing. An unmanned system that is capable of performing a security surveillance task is regarded as having a higher level of autonomy than a system that is only able to perform a point-to-point driving task.

We have developed four major categories of metrics for measuring the complexity of missions. They are:

1) Tactical Behavior: The composition and structure of the involved tasks provide an essential measure for the complexity of a mission. The particular metrics include:
    - Number of different types of major subtasks
    - Numbers of supervisor-subordinate levels within the UMS
    - Number of decision points

    Note that, for UMSs that employ solution paradigms other than hierarchical task decomposition, these metrics can be weighed low or out, depending on the metrics' relevance to the UMS architecture.

    A mission [7] may be typically decomposed into levels of subtasks until reaching the actuator level task commands [10]. This concept is described extensively in the NIST 4D/RCS architecture [10]. The combination of the number of subtasks after the first-level decomposition and the number of levels of decomposition should provide system technical staff a good measure of the complexity of the mission without requiring exhaustive details of a complete task decomposition structure. In experimental systems, there may well be multiple ways to perform task decomposition. However, in a more established and operational domain, there tends to be commonly used task vocabulary that could provide a basis for the task decomposition.

    Number of decision points indicates how the subtasks are used in the mission execution. In simple cases, the subtasks may be executed only once to accomplish the mission goals, whereas complex situations may require the execution of multiple instances of some of the subtasks, possibly concurrently.

2) Coordination and Collaboration: A mission with a higher level of complexity typically requires a higher level of coordination and collaboration among the components or subsystems. From a system perspective, a UMS that is able to perform a high level of coordination and collaboration should be regarded as having a high level of autonomy.

    The metrics in this category include:
    - Number of participating entities
    - Enabling interfaces: types of data, frequency, number of channels, idling time due to data dependency.

3) Performance: A UMS's ability to achieve mission goals with high efficiency and accuracy through its planning and execution components indicates the UMS's autonomous capability. The specific metrics to measure the performance include:
    - Mission planning and analysis capability at pre-mission, during mission, and post mission stages
    - Allowed latencies and errors

    Mission planning capability means the UMS's capability to generate plans to achieve the desirable states for the three stages, namely, Ready, Goal, and Standby, respectively. Allowed latencies and errors means the UMS's capability to execute the plans to achieve and maintain at the Goal state. The measure has to be statistically sound. The allowed errors contain spatial and temporal aspects. Note that effectiveness and efficiency of the generated plans should be a part of the measure.

4) Sensory Processing/World Modeling: The perception requirements for particular missions and the dependency on external information indicate levels of complexity of the missions. The metrics include:
    - Situation awareness required
    - Information independence

Task planning and performance require corresponding perception capability, covering sensing, information modeling, knowledge update, through event detection. Information independence is proportional to the UMS's level of autonomy.

The effort to develop the mission complexity model continues, with a particular focus on the measures. The following table shows our current results and the direction that we are heading. Note that plenty of flexibility is provided, on weights and types of measures, to suit Programs' priorities:

Metric	**Weight (Low/ Mid/ High)**	**Default Measure**	**Extended Measure (**user definable, the descriptions in this column are intended to serve as a reference**)**
1. Tactical Behavior			
1.1. Number of different types of major subtasks	M	Low/ medium/ high	min: 0 subtask, at lowest level low: 1 mid: 2-4 high:5-9--typical human effective performance (G. Miller). adv:>9
1.2. Number of supervisor-subordinate levels	M	Low/ medium/ high	min:1--self only low:2 mid:3 high:4--single entity adv:>4--organization
1.3. Number of decision points	M	Low/ medium/ high	min: 1--one step to goal low: # of subtasks; sequence through all subtasks once, nominally mid: 10X; subtasks may be required for up to 10 times high: 100X; subtasks may be required for up to 100 times adv:>100X; indefinite
2. Coordination and Collaboration			
2.1. Single or multi entity mission: a. number of participating entities	M	Low/ medium/ high	min: 1--self low: 2--2 UMSs mid: up to two hierarchical levels or 5 UMSs high: up to three hierarchical levels or 11 UMSs adv: >3 levels 11 UMSs
2.2. Enabling interfaces: a. types of data—discrete events, state space, images b. frequency c. number of channels d. synchronization--inter dependency, idle time	M	Low/ medium/ high	* Ability to conduct many types and high frequency communication, less idling.

3. Performance			
3.1. Mission planning and analysis: a. pre-mission: achieve/maintain Ready state; proficiency b. during mission: achieve and maintain Goal state; proficiency c. post mission: achieve and maintain Standby state & required post processing; proficiency	M	Low/ medium/ high	* % of tasks (per metric 1.1) planned * higher or lower than human performance * real-time replanning required * allow for change of mission
3.2. Allowed latencies and errors: a. accomplishment rate: achieve/maintain Goal state; proficiency b. spatial errors c. temporal errors	M	Low/ medium/ high	* % of tasks (per metric 1.1) accomplished * success rate could be categorized in 20% increments from 0%. * spatial err could be categorized as: < footprint, up to 5 ums size, up to 10 size * temporal err could be categorized using % of entire mission time
4. Sensory Processing/World Modeling			
4.1. Situation awareness required: a. fusion/perception levels required b. types of objects recognized c. map resolutions d. sizes of rule/case bases/engines	M	Low/ medium/ high	Per 5 fusion/perception levels as in Terminology [7]: min:1 low:2 mid:3 high:4 adv:5?
4.2. Information independency	M	Low/ medium/ high	* Dependencies on external information, such as GPS, indicating low levels of autonomy.

As the number of control levels in a system increases, the multiple tactical behaviors that the lower level subsystems perform may be integrated into single behaviors with a higher level of abstraction [10]. For example, when the task is for a team of UMSs to conduct security surveillance at a certain area, at the individual vehicle level, we could say that the vehicle A has ALFUS-5[2] for mobility, ALFUS-3 for the Reconnaissance, Surveillance, and Target Acquisition (RSTA) function, and ALFUS-4 for communication. Vehicle B may have different ALFUS capabilities. However, at the higher, Section level, the autonomy level should not be characterized as Vehicle #A in section 1 has ALFUS-5 for mobility, etc. Instead, the ALFUS should be specified such as Section Alpha has autonomy ALFUS-3 for the bounding over-watch behavior. Section Bravo has ALFUS-5 for convoying. At an even higher level, joint behaviors including aerial vehicles may be identified.

[2] We use the hypothetical indices, without elaboration, only as an illustration and do not imply establishing any ALFUS metrics at this point.

3.2 Environmental Difficulty

The measure of Environmental Difficulty is complex and highly impacted by the other measures. This measure is decomposed into categories including Static Environment, Dynamic Environment, Electronic/Electromagnetic Environment, Mobility, Mapping and Navigation, Urban Environment, Rural Environment, and the Operational Environment. Each category is further described with a draft set of specific measurable factors. The granularity of the factors within each category must still be determined. For example, although a UMS can maneuver through smoke, the level of visibility must remain greater than 5 meters and the obscurant limits the safe speed of the vehicle. Many such inter-dependencies and issues with the detail of the measure arise. The following table represents the initial breakout of the primary categories for the Environmental Difficulty set of measures.

Static Environment	Measure the ability of the unmanned system to operate within a known, non-moving, geo-referenced area. The categories Urban and Rural Environment also include static environment variables, but address man made static entities only. Static Environment variables include terrain type, soil characteristics, water depth, terrain elevation and elevation change characteristics.
Dynamic Environment	Assess the unmanned system's ability to detect and negotiate changes in the environment while minimizing the impact on mission goals. Dynamic Environment variable metrics include frequency of obstacles, density of obstacles, detection and use of access points, and human interaction.
Electronic/Electromagnetic Environment	The unmanned system's ability to communicate and function with respect to the impact of electromagnetic fields, and/or any other energy source both hostile and friendly. Specific measures to be accounted for within this category include the UMS's ability to withstand communication drop-outs, jamming, magnetic fields, EMP, and multi-path.
Mobility	The impact of the environment on the UMS includes common metrics such as range, turn radius/rate, max roll/pitch, shock, vibration, acceleration and deceleration. Ideally, these measures will be collected for a full 6 DOF environment.
Mapping and Navigation	The resolution of the required a priori data for the environment such as maps, elevations, etc., impacts the vehicle's Level of Autonomy. Also included in this category are navigational aids used by the UMS such as GPS and air traffic control interfaces.
Urban Environment	The Urban Environmental factors account for the measurable impact of traffic, road conditions, road variation, traffic rules, control points, and any other man-made mobility constraints and choices. Additional factors of the Urban Environment may be listed in other, more specific, categories.
Rural Environment	Rural Environment variables include vegetation, fences, walls and other barriers, and biological factors such as wildlife and domestic animals. As with the Urban Environment, other factors may be listed in other categories.
Weather	This measure accounts for variables such as sea state, wind speed, pressure, humidity, visibility (due to atmospheric conditions), lighting conditions, and any other natural phenomenon that might impact mobility.
Operational Environment	The Operational Environment differs from Mission Complexity in that it does not account for the tactical or strategic attributes of the mission, but captures the factors that force change, temporary and/or intermittent, on the mission. Factors listed within this category include enemy fire, decoys, and change in the operational tempo.

At this time, the working group is not directly addressing specific inter-dependencies between the various measures. The granularity of the measure, however, will be proposed and used for development of the model for the Autonomy Levels. As the model matures, many changes in both the categories and factors of the data set and the granularity of measures

will occur. The population of the first draft of the Environmental Difficulty factors is currently in progress. Examples from the Environment Difficulty group are presented to further exemplify the challenge.

Soil Characteristics	General description of the type of soil the unmanned vehicle can traverse without damage and/or loss of traction. Sand Clay Grass Rock Gravel Pebble
Terrain Elevation Change	Average change in elevation of terrain accounting for both frequency and amplitude Vertical Change in meters from trough to peak (0 to 10,000) Horizontal distance in meters between peaks (0 to 10,000)
Frequency of Obstacles	Determine the impact of system performance based on the occurrence of obstacles with respect to time (during mobility). This measure represents the ability of the system to process sensory input into the obstacle classification and/or negotiation processes. Greater than 5 minutes 5 minutes to 120 second intervals 0 to 1 Hz … > 10 Hz
Density of Obstacles	Obstacle density is the measure of occurrences of anomalies within a 1000m x 1000m area. These anomalies qualify as obstacles only if they are not present in a priori data sets. This is a measure of the number of obstacles a system can detect, track, and avoid (if necessary) in the determined area. 1000 500 to 1000 250 to 500 … less than 10

From the example in the previous table, it can be seen that numerous factors appear to be left out. The type of soil, for instance, does not fully quantify the impact of that soil on the vehicle system. Clay might be easily traversable until a hard rain makes it all but impossible for a light skid-steer UMS to perform simple maneuvers. The classification of elevation data, presented as an average for an area, would not at first glance appear to have an impact on the Autonomy Level of Unmanned System. However, combined with minimum and maximum factors, the measure provides a basis for determining the fit of a robot to a particular terrain. Further, these factors do not provide any quantities to process within a mathematical model.

3.3 Level of HRI

At the initial meeting of the ALFUS Working Group a great deal of discussion evolved around the relationship between the Level of Human Robot Interaction (HRI) and the Autonomy Level for the Unmanned System. For simple unmanned systems this association remains fairly linear. The introduction of planning and coordination algorithms into the robots forces variances in the measure of HRI. Rather than collecting factors based on intervals between instructions or bandwidth for command and control, the measures must now account for the time before and after a mission the operator-computer interaction is required for mission completion: duration (seconds) and frequency (Hz) of human

controller interaction pre-mission, during mission, and post-mission. For highly experimental robots, the time of the scientists and engineers must be added to that of the operators. So, for five minutes of unsupervised operation, the machine might require five days of direct human interaction. As this will not be the case for most unmanned systems utilizing the ALFUS framework, the model must allow for the omission of such attributes.

More interestingly, however, is the relationship between communications and autonomy. Communications with the human operator, as measured within the Level of HRI group, and all other communications would appear to have much the same impact on the level of autonomy as does the level of human interaction. The ALFUS Working Group has identified through other works that this logic is faulty. As indicated earlier, the typical kitchen appliance is highly autonomous when it comes to communications and operator interaction. It rates extremely low when measured against mission elements and environmental difficulty. Another important consideration is the communications between multiple unmanned systems. Although highly autonomous, coordinated flight and cooperative behaviors require a high degree of communications albeit not only with human operators.

The HRI issues, in terms of how they affect the unmanned systems and how they affect the ALFUS definitions, were discussed at length. Relevant HRI studies [11, 12] were mentioned as potentially helpful to the objectives of this workshop. Further effort in establishing a set of metrics for the Level of Autonomy with respect to human interaction was performed by the ALFUS HRI Subgroup. This subgroup established the following high-level categories for capturing metrics to account for the impact of HRI on autonomy:

> Human Intervention: This metric captures the frequency, duration, and number of unplanned robot initiated interactions. Intervention is "an unanticipated action or input by the user to help complete a task" (as defined by ALFUS WG). Normal interaction is measured by Operator Workload.
>
> Operator Workload: Operator workload: Measures the workload associated with normal (i.e. planned) operation of the UMS. This measure is captured for Pre/Post Mission workload and Mission workload. The NASA TLX [13] is a common measurement tool for operator workload that includes six categories (Physical Effort, Mental Effort, Temporal, Performance, Frustration, and Overall Mental and Physical Effort).
>
> Operator Skill Level: Both the UMS Operator and Support Personnel are included in this category. These measure capture the training and education level. The higher the autonomy level of the robot, the less skill is required of the operator.
>
> Operator to UMS Ratio: This measure captures the ratio of operators to unmanned systems. The larger number of robots one operator controls, the higher the level of autonomy is assumed for the robots.

Two other metrics identified by the HRI subgroup are Logistics and Ease of Use. Currently these metrics are not included in the HRI ALFUS matrix. The logistics measures deal with system specific aspects of the unmanned system such as effort to launch and retrieve, control station attributes and communications. Ease of Use measures that are not subjective will be captured in the other categories.

3.4 Additional Concerns

We have determined that our first model should focus on the aforementioned three axes. Additional axes, however, may be required for the future versions. The concerns may include:

3.4.1 System Dependence

It was suggested that good characterization of UMS capabilities is critically important for the system autonomy specification. A question was brought up for further investigation: whether small and large robots should be separately evaluated in terms of their autonomy levels.

Implementation also may affect the ultimate autonomy ratings or evaluation frameworks. It was pointed out that different system control approaches, e.g., reactive sensor based behavior and deliberative knowledge based behavior might lead to different autonomy frameworks.

3.4.2 Cost and Technology Readiness

It was suggested whether cost, affordability, as well as the maturity of the technology enabling particular ALFUS levels, should be taken into account when considering autonomy levels. This requires further investigation.

3.5 Perceived Application Models for Autonomy Level Framework

Various types of users may employ the autonomy level framework at various levels of detail. As mentioned, a corporate executive or battlefield commander may only need to know a concise index showing the UMSs' autonomy level. Engineering staff may need the full detail of the presented Detailed model to test and evaluate a UMS. Project management personnel might need a representation that is between the two extremes. Discussions have begun on how to summarize and present the metrics in such a format.

We envision that each of the detailed metrics axes, as described in the earlier sections, can be summarized into a one to ten scale, indicating the autonomy level from the particular perspective. Figure 3 demonstrates the effect.

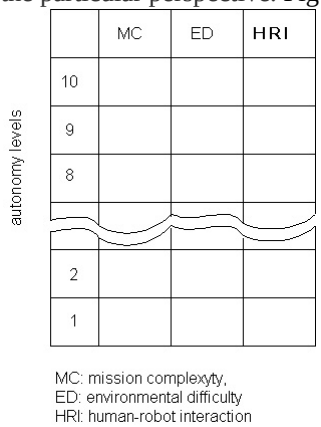

Figure 3: Autonomy levels matrix

We further envision specifying each of a UMS's major functions with this matrix. Figure 4 demonstrates a particular vehicle being specified or evaluated with the matrix.

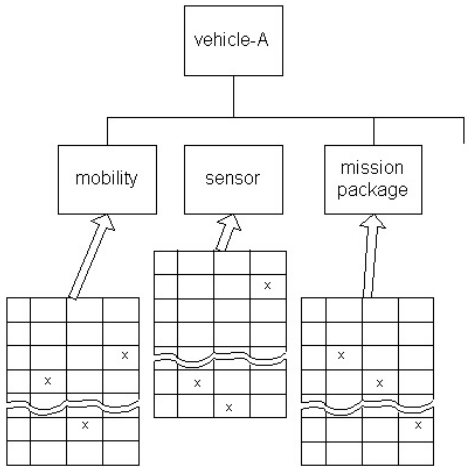

Figure 4: Applying the autonomy level matrix to a UMS

4. SUMMARY AND PLANS

It is recognized that the issue of autonomy levels for unmanned systems is extremely complex. The Autonomy Levels for Unmanned Systems working group has established a structure for the autonomy level framework and has accomplished an initial set of metrics and measures for the framework.

A significant amount of work remains to be accomplished, including:

- further refinement of the metrics along the three axes and their harmonization.
- devise a method of transforming the three-axis detailed model to the concise, summary model for autonomy levels,
- develop methods to apply the ALFUS framework,
- develop testing evaluation procedures for the ALFUS framework,
- develop use cases and examples to help disseminating the Framework, and
- seek to apply the framework to the Army Future Combat Systems and other unmanned system programs.

Acknowledgement

The authors acknowledge all the other Workshop participants for their contributions to the Framework as well as additional practitioners for forwarding significant comments on the Terminology. The following is a list, barring omission, in alphabetical order: **Air Force**: Bruce Clough, Robert Smith, Jeff Wit. **Army**: Curt Adams, Keith Arthur, Robert Barnhill, Bruce Brendle, Marsha Cagle-West, Jeff Cerny, Mike Dzugan, Woody English, Ray Higgins, Julie Hirtz, Kelley Hodge, David Knichel, Brian Novak, Kerry Pavek, Al Richardson, John Rovegno, Jeff Rowe, Kent Schvaneveldt, Charles Shoemaker, Stephen Swan, Robert Wade. **DARPA**: LTC Gerrie Gage, Doug Gage, Dennis Overstreet. **Navy**: Caesar Mamplata. **DOE HQ**: Tom Weber. **FHWA (DOT)**: Robert Ferlis, Peter Huang. **INEEL (DOE)**: David Bruemmer. **NASA**: Jeremy Hart, Ryan Proud. **NIST**: Brian Antonishek, Tony Barbera, Maris Juberts, Jean Scholtz, Harry Scott, Albert Wavering. **OSD**: Keith Anderson, Jeffrey Kotora.

References

1 http://www.jointrobotics.com/activities_new/FY2003%20Joint%20Robotics%20Master%20Plan.pdf

2 Knichel, David, Position Presentation for the Maneuver Support Center, Directorate of Combat Development, U.S. Army, the First Federal Agencies Ad Hoc Working Group Meeting for the Definition of the Autonomy Levels for Unmanned Systems, Gaithersburg, MD, July, 18, 2003.

3 James Albus, Position Presentation for National Institute of Standards and Technology, Intelligent Systems Division, the First Federal Agencies Ad Hoc Working Group Meeting for the Definition of the Autonomy Levels for Unmanned Systems, Gaithersburg, MD, July, 18, 2003.

4 Bruce T. Clough , "Metrics, Schmetrics! How The Heck Do You Determine A UAV's Autonomy Anyway?" Proceedings of the Performance Metrics for Intelligent Systems Workshop, Gaithersburg, Maryland, 2002.

5 Army Science Board, Ad Hoc Study on Human Robot Interface Issues, Arlington, Virginia, 2002.

6 http://www.isd.mel.nist.gov/projects/autonomy_levels/

7 Huang, H. Ed., *Terminology for Specifying the Autonomy Levels for Unmanned Systems, Version 1.0*, NIST Special Publication 1011, National Institute of Standards and Technology, Gaithersburg, MD, January 2004.

8 Houghton Mifflin Company, The American Heritage Dictionary.

9 Hui-Min Huang, Elena Messina, James Albus, "Autonomy Level Specification for Intelligent Autonomous Vehicles: Interim Progress Report," 2003 PerMIS Workshop, Gaithersburg, MD.

10 James Albus, et al., 4D/RCS: A Reference Model Architecture For Unmanned Vehicle Systems, Version 2.0, NISTIR 6910, Gaithersburg, MD, 2002. http://www.isd.mel.nist.gov/projects/rcs/

11 Sheridan, T. B. Telerobotics, Automation, and Human Supervisory Control. The MIT Press. 1992

12 Drury, J.L., Scholtz, J., and Yanco, H.A., "Awareness in Human-Robot Interactions" to Appear in the Proceedings of the IEEE Systems, Man, and Cybernetics Conference, Washington DC, October 2003.

13 Hart, S. G., & Staveland., L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human Mental Workload. Amsterdam. The Netherlands: Elsevier.