Proceedings of the 2003 Performance Metrics for Intelligent Systems (PerMIS) Workshop, Gaithersburg, MD, August 16 - 18, 2003.

Autonomy Level Specification for Intelligent Autonomous Vehicles: Interim Progress Report

Hui-Min Huang, Elena Messina, James Albus

National Institute of Standards and Technology
Intelligent Systems Division
Gaithersburg, Maryland 20899-8230
{ hui-min.huang, elena.messina, james.albus }@nist.gov

Abstract
As unmanned systems become widely employed in various applications, it is critical to have a set of standard definitions and metrics for specifying and evaluating the systems in terms of their levels of autonomy. Developing autonomy levels for unmanned systems is a complex issue that has to take into account many factors such as task complexity, human interaction, environmental difficulty, mission and system dependence, and quality factors. We report on a workshop that addresses this issue.

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 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.

These efforts have been driven by the advance of mobile robotic technology and the expanded roles that unmanned systems are playing in military and civilian situations. As Government agencies specify unmanned system capabilities, it is critical to have a set of standard definitions. These definitions can also provide a basis for metrics for system performance evaluation. An ad hoc government working group has been formed to address these issues.

The ultimate objectives for the working group are:

- To determine the needs 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 briefly summarizes the first workshop. Follow-on workshops are planned, starting with one in mid-September 2003.

2. The First Workshop Summary

The first Workshop was held on July 18, 2003 at NIST. Participants included representatives from the Department of Defense Office of the Secretary of Defense (DoD OSD), Defense Advanced Research Projects Agency (DARPA), U.S. Air Force, multiple organizations within the U.S. Army, Department of Commerce (DOC), Department of Energy, and Department of Transportation (DOE).

Given the complexity of the problem and the diversity of the application domains, the following near-term objectives were set up for the first meeting:

- To exchange viewpoints and begin to grapple with the technical issues.
- To establish a forum for discussion and exchange of ideas.

The participants presented the autonomy requirements from their particular programs. The aforementioned, existent autonomy level definitions were presented and discussed.

The meeting resolved to aim at a first version of an autonomy level model by September 2003 to answer a critical Army requirement. To accomplish this objective, a small group representing different domains, including DOD, DOE, DOT, and DOC has been assembled and charged to:

- define a set of terms,
- define methods for establishing metrics for autonomy, and
- draft an autonomy levels for unmanned systems (ALFUS) model.

3. Motivations for Greater Autonomy

A brief summary of the motivators for higher autonomy in unmanned system is:

- Autonomy has been motivated by lack of bandwidth. It is not realistic to expect to be able to teleoperate a vehicle remotely due to the large amount of data that needs to be sent back from the vehicle to the human operator. Bandwidth is a limited resource in most circumstances. There are also instances where communications must be curtailed for stealth reasons and where communications are not available (e.g., deep within tunnels).

- Safety for personnel is another major motivator. If an unmanned system requires close attention and monitoring by human personnel (possibly nearby) these humans are potentially in peril due to either lack of focus on their own safety as they tend to the vehicle or because of their physical proximity to a dangerous area.

- Potential system mission effectiveness is limited by cognitive workload. Currently fielded unmanned systems require close control by highly trained human operators. In some cases, multiple operators are required for a single vehicle. The humans are therefore limited in their ability to perform their own tasks within the mission.

4. Requirements Expressed by Participants

Given that autonomy is desirable in unmanned systems, several of the requirements as expressed by the various participants that drive the working group include:

- Definitions and Framework: A common set of definitions and framework are needed for measuring technology goals and for advancement in terms of system autonomy.

- Metrics and Benchmarks: Conformance metrics or benchmarking are needed for Government to evaluate autonomy that contractors propose.

- Autonomy Level Definitions: The Army Future Combat Systems (FCS) requires autonomy level definitions to be able to specify its unmanned system requirements. While existent references, such as the Joint Robotic Program UGV Master Plan, contain useful information, upgrades are needed.

- Standard Definitions: Definition of success for this group would be the establishment of standard definitions of UGV autonomy that describe UGV mobility, control, and behavior to support combat developments, technology development, and joint communications.

- Quantitative Measures: Participants also welcomed the emphasis on metrics. As technology for unmanned systems matures, the community needs to become more rigorous in its evaluations and definitions.   Quantitative measures are necessary.

5. Elements of ALFUS Definitions

It became clear during the initial workshop and in subsequent interactions that the definition of a system's autonomy level is multi-faceted and has various ultimate purposes. Several overall definitions of autonomy were proposed, including:

> Autonomy is having free will.
>
> Autonomy is the ability to observe, orient, decide, and act in an environment without outside (human) assistance.
>
> A system's autonomy is its own ability to achieve its goals.

These definitions require further iterations.

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

- Autonomy relates to multiple technical areas or subsystems: mobility, control, and tactical behaviors
- Task complexity and adaptability to environment are key aspects—washing machines should not be considered to have high levels of autonomy.
- Nature of collaboration with humans is important: levels of involvement and different modes of interaction.
- Relevance of various quality factors: how do the following affect UMS's autonomy level: mission success rate, response time, precision/resolution/tolerances, and environmental difficulty.

Figure 1, although not a consensus view, demonstrates a perspective. Degree of difficulty of the environment, Degree of difficulty of the mission, and Inverse of operator workload constitute the three axes. The autonomy level of a particular UMS can be represented with a triangular surface with certain values on the three axes. This is one way of representing ALFUS, but it also implies that vectors, as opposed to a single scale, may be better suited to characterize unmanned system autonomy levels.

The following sections elaborate on these elements.

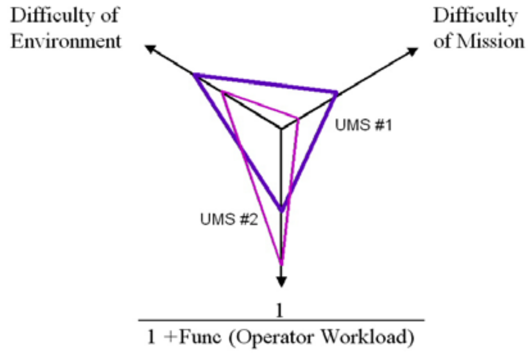

Figure 1: Complexity of autonomy level definitions

5.1 Human Factors

The human and robot interaction (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 [6, 7] were mentioned as potentially helpful to the objectives of this workshop.

The Army states that the control that soldiers have over systems is critical and that the Army does not foresee a system that does not require HRI.

It was suggested that human and robot involvement should not be complementary fractions that add up to one. Rather, preliminary thinking was that, at higher levels of control, HRI is at higher levels of abstraction and such interactions, when designed properly, might greatly enhance system capabilities. Further exploration along this line is needed.

Additional HRI factors that could contribute to the autonomy level definitions include:

- whether a UMS can initiate and assume authority,
- what is required of humans in order to support a UMS conducting a particular mission,
- how should operator skill levels and types be factored in their assignment to particular UMS's, and
- whether { number of vehicles / number of operators } or { person hours / flight hour } might be sound metrics.

It was left for further investigation as how human factors affect ALFUS definitions.

5.2 Missions Dependence and Task 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.

As a system expands its configuration and includes higher levels in the control hierarchy, the multiple tactical behaviors that lower level subsystems perform may be integrated into single behaviors with a higher level of abstraction [8]. For example, when the task is for a team of UMSs to conduct security surveillance on the NIST grounds, at the individual vehicle level, we could say that the vehicle A has ALFUS-5[1] 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 overwatch behavior. Section Bravo has ALFUS-5 for convoying. At an even higher level, joint behaviors including aerial vehicles may be identified.

5.3 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.

5.4 Environmental Difficulty

It was pointed out that ALFUS should consider the environmental factors. For example, for the task of road following for a UGV, the road could be:

- A road with or without clear markings.
- Road with or without same or opposite direction traffic.
- Light vs. heavy traffic.
- Straight vs. curve roads.
- Urban roads vs. freeways.
- Well paved vs. snow or ice covered roads.

It is possible that multiple ALFUS designations are needed to address all of these road conditions, but this requires further investigation.

5.5 Cost and Technology Readiness

It was pointed out that, cost, affordability, as well as the maturity of the technology enabling particular ALFUS levels, ought to be taken into account when considering autonomy levels. This requires further investigation.

5.6 Performance and Quality Factors

The autonomy level metrics might also include the following performance and quality factors:

- precision/error bounds of solutions
- solution efficiency and optimality
- mission success rate
- system response time
- system perception capability in handling poor visibility

6. Representation Issues for ALFUS

The multi-dimensional, task-specific nature of autonomy measures, coupled with the need to communicate the measures to multiple types of audience poses many challenges in terms of how

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

to represent autonomy levels. The Joint Architecture for Unmanned Systems (JAUS) model is an example effort that includes a Reference Architecture for technical readers, a Domain Model for operational readers, and a Strategic Plan for overall guidance [9]. Another example would be 4D/RCS. 4D/RCS is a reference model architecture and engineering methodology that has been used very successfully in the ARL Demo III program. 4D/RCS is currently being used for development of autonomous driving skills and tactical behaviors on the DARPA Mobile Autonomous Robots (MARS) program, the TACOM VTI program, and the FCS Autonomous Navigation System. The highly modular structure of 4D/RCS is suggestive of how performance of various components, such as planners, perception systems, world modeling algorithms, and cognitive reasoning engines might be evaluated individually as well as part of an integrated system [8].

It was discussed that the autonomy level specification may contain at least two facets. One would be comprehensive and detailed specifications aiming at technical users such as staff with various Army various Research, Development, and Engineering Centers (RDEC), Government Contracting Technical Specialists and evaluators, and contract concerns. This specification would encompass all the considerations as described in this section.

Another facet would be a set of concise, numbered indices aimed at use by combat leadership, executives, and operational users like program managers (PMs), unit leaders, and soldiers. There needs to be a sound process to translate the technical ALFUS definitions into languages that these types of users speak and into a culture that these types of users live in. An advantage would be simplicity. An important cultural issue is that combat leaders and soldiers are and will remain the centerpiece of our military force.

This working group should interact with users with individual programs so that the particular cultural issues are addressed adequately in considering representations.

Opposing views were expressed about 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, relate to numbers better. In this regard, a simple 0 or 1 through 10 scale was the consensus. This is a cultural issue to be addressed.

7. Approach

There was consensus that, at the end-user level, the autonomy level definitions should be mission specific to be most useful.

Given the wide differences in unmanned system domains that include ground vehicles, air vehicles, undersea vehicles, surface vessels, and littoral water robots, where autonomy may be defined differently, a great challenge would be whether generic definitions could be obtained.

The notions of generic model and reference model were, however, proposed as worth further investigations. A generic model was presented that attempts to characterize autonomy levels in terms of vehicle groupings, environmental complexity, task complexity, and quality factors. The objective is to explore whether the model can be extended for each individual mission specific ALFUS.

The working group decided to develop ALFUS using a spiral approach. The first iteration should respond to the Army FCS needs for its next ORD cycle, by October 2003. It was also determined that, to expedite the progress, a smaller team of six from various Federal departments and representing application domains should be, and was assembled and charged to:

- define a set of terms,
- define methods for establishing metrics for autonomy, and
- draft an ALFUS model.

The results should provide a focal discussion point for the next workshop, planned to be held at mid September, to respond to the FCS requirements.

It was recommended that the JRP Master Plan, containing relevant term definitions, should be referenced. The NASA originated and Army adopted Technology Readiness Level (TRL), both the scale and the entire method behind it, was mentioned as a possible reference [10].

8. Summary

It is recognized that the issue of autonomy levels for unmanned systems is extremely complex. The workshop has established a continuing forum to allow the community to address the issue.

The first workshop has begun discovering vastly different requirements from participating organizations and programs. The workshop revealed that ALFUS has to take into account many factors such as task complexity, human interaction, environmental difficulty, mission and system dependence, and quality factors. This workshop has set up a solid foundation to allow pursuit of its long-term objectives of devising methods for establishing metrics of autonomy for unmanned systems and developing a set of widely recognized standard definitions for the levels of autonomy for unmanned systems.

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

7 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.

8 4D/RCS Reference Model Architecture @ http://www.isd.mel.nist.gov/projects/rcs/

9 http://www.jauswg.org/

10 John C. Mankins, Technology Readiness Levels: A White Paper. Office of Space Access and Technology, NASA, http://www.hq.nasa.gov/office/codeq/trl/trl.pdf, April 6, 1995