Ethical Control of Unmanned Systems

Ethical Control of Unmanned Systems using Formal Mission Ontologies for Undersea Warfare

NDIA 2020 Undersea Warfare Virtual Conference

Don Brutzman
Naval Postgraduate School (NPS)

4 September 2020
2020 Undersea Warfare Virtual Conference

9/22/2020 - 9/23/2020

Theme: Integrating Combat Power with the Joint Force
Event Type: Virtual Meeting
Event Code: 0240
Conference theme: Integrating Combat Power with the Joint Force

• Joint Force is increasingly interdepending on Undersea Warfare as conflict potential grows across the world

• Alfred Thayer Mahan: command of the sea, even if local and temporary, naval operations in support of land forces holds decisive importance. (Wikipedia)

• Command Control Communications Intelligence (C4I) working group efforts are fundamentally important.

• USW forces must be capable of jointly coordinating C2 of systems holding direct potential for lethal force, despite intermittent communications, over long distances and long durations of time.
Ethical Control of Unmanned Systems: Keeping Warfighters in Charge of Autonomy

Milestones and Transitions

- CRUSER development led to first project selection under CRADA with Raytheon Missile Systems (RMS).
- Successful progress on test missions entering TRL 5 with simulation and Web-shareable 3D visualization.
- Expressing multiple robot mission plans consistently, coherently for diverse UAV, USV, UUV platforms.
- Use Semantic Web Standards to support warfighters.
- Evaluate NAVSEA Unmanned Maritime Autonomy Architecture (UMAA) evolution for robot qualification.

Why / Objectives

- Ethical control of unmanned systems can be accomplished through structured mission definitions that are trusted, consistently readable, verifiable, repeatable and understandable by humans and robots.
- Orders must be lawful. Unmanned systems must behave ethically and comprehensibly if they are to support manned military units effectively.
- Well-structured mission orders can be tested and trusted to give human commanders confidence that offboard systems will do what they are told to do, and further will not do what they are forbidden to do.
- Demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to humans, in lethal and life-saving scenarios.

https://savage.nps.edu/EthicalControl

What / Deliverables

- Update Mission Execution Ontology (MEO) concepts demonstrated in tests and simulation, building to perform field experimentation (FX).
- Supervise thesis work to explore canonical exemplar missions that are expected to utilize unmanned systems, looking across the full range of Naval warfare communities. Example scenarios include UAV for sailor overboard, UAV for refugee/lifeboat escort, and adopt scouts. All must observe Law of Armed Conflict (LOAC), Rules of Engagement (ROE), and moral guidance of commanders despite long durations/distances.
- Define, simulate, and test combination of real-world goals and ethical constraints to robot mission tasks across set of canonical scenarios.
- Illustrate how human-robot teams meet moral and legal requirements if deploying unmanned systems with potential for lethal, life-saving force.
Synopsis: Ethical Control of Unmanned Systems

• **Project Motivation:** ethically constrained control of unmanned systems and robot missions by human supervisors and warfighters.

• **Precept:** well-structured mission orders can be syntactically and semantically validated to give human commanders confidence that offboard systems
  - *will do what they are told to do*, and further
  - *will not do what they are forbidden to do.*

• **Project Goal:** apply Semantic Web ontology to scenario goals and constraints for logical validation that human-approved mission orders for robots are semantically coherent, precise, unambiguous, and without internal contradictions.

• **Long-term Objective:** demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to human beings.
Open-Source Licensing for Ethical Control

All NPS work is published as open source with no restrictions on use.
• https://savage.nps.edu/EthicalControl/license.html (also .txt)

An excellent heuristic is to assume success – then what?
• Ethical Control of unmanned systems is critically important worldwide.
• Ethical Control must be unambiguously defined in order to be actionable by humans (military professionals) under all conditions facing U.S. military.
• Ethical Control must be implementable across a totally diverse set of evolving unmanned and C2 systems (both open and closed) in order to be effective.
• Ethical Control concepts, syntax definitions and semantic relationships for lethal/lifesaving force thus cannot be subject to proprietary restrictions.
Background History

Many years of work have composed multiple fields of study to provide techniques for maintaining human ethical control of unmanned systems. In this work, ethical theory meets professional practice. Each step must work for human commanders and unmanned systems alike.
Key Insights regarding Human Ethical Control

1. Humans in military units are able to deal with moral challenges without ethical quandaries,
   • by using formally qualified experience, and by following mission orders that comply with Rules of Engagement (ROE) and Laws of Armed Conflict (LOAC).

2. Ethical behaviors don’t define the mission plan. Instead, ethical constraints inform the mission plan.

3. Naval forces can only command mission orders that are
   • Understandable by (legally culpable) humans, then
   • Reliably and safely executed by robots.

Reference: CRUSER TechCon Overview 2016
https://gitlab.nps.edu/Savage/EthicalControl/tree/master/documents/presentations
Ethical Mission Definition and Execution for Maritime Robots Under Human Supervision

• Don Brutzman, Curtis L. Blais, Duane T. Davis, Robert B. McGhee
• IEEE Journal of Oceanic Engineering (JOE), Volume: 43 , Issue: 2 , April 2018
• Abstract. Experts and practitioners have worked long and hard toward achieving functionally capable robots. While numerous areas of progress have been achieved, ethical control of unmanned systems meeting legal requirements has been elusive and problematic. Common conclusions that treat ethical robots as an always-amoral philosophical conundrum requiring undemonstrated morality-based artificial intelligence are simply not sensible or repeatable. Patterning after successful practice by human teams shows that precise mission definition and task execution using well-defined, syntactically valid vocabularies is a necessary first step. Addition of operational constraints enables humans to place limits on robot activities, even when operating at a distance under gapped communications. Semantic validation can then be provided by a Mission Execution Ontology to confirm that no logical or legal contradictions are present in mission orders. Thorough simulation, testing, and certification of qualified robot responses are necessary to build human authority and trust when directing ethical robot operations at a distance. Together these capabilities can provide safeguards for autonomous robots possessing the potential for lethal force. This approach appears to have broad usefulness for both civil and military application of unmanned systems at sea.
Semantic Web and inferencing technologies for Department of Defense systems

- Duane Davis, NPS Technical Report NPS-CMIS-14-001, 2014
- Available via [https://calhoun.nps.edu/handle/10945/43723](https://calhoun.nps.edu/handle/10945/43723) (report) (slideset)

**Abstract.** Operational commanders and intelligence professionals are provided with a continually-increasing volume of data from numerous sources. Effective utilization of this data can be hampered by difficulties in fusing different data streams for presentation, correlating related data from various sources and developing reliable summary and predictive products. An opportunity presently exists to improve this situation through the incorporation of Semantic Web technologies into Department of Defense (DOD) systems. This report provides a didactic overview of Description Logics (DL) and their implementation in Semantic Web languages and technologies to include the mathematical properties supporting robust knowledge representation. Subsequently, the algorithms for automated reasoning and inferencing with DLs are discussed. Included in this discussion is a comparison of available Semantic Web applications for ontology development and realization or DL reasoning capabilities with real-world knowledge bases. Finally, mechanisms for applying artificial intelligence techniques to ontological DL information are presented.
Description Logic (DL) Rules provide basis for Mission Execution Ontology (MEO)

<table>
<thead>
<tr>
<th>Rules</th>
<th>Description Logic Equations</th>
<th>Plain-language description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = Mission Rules</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>Mission ( \sqsubseteq \forall \text{startsWith.Goal} \land \text{1=1 startsWith.Goal} )</td>
<td>A Mission can only start with a Goal and must start with exactly one Goal</td>
</tr>
<tr>
<td>M2</td>
<td>Mission ( \sqsubseteq \forall \text{includes.Goal} \land \geq \text{1 includes.Goal} )</td>
<td>A Mission can only include Goals and must include one of more Goals</td>
</tr>
<tr>
<td>M3</td>
<td>Mission ( \sqsubseteq \forall \text{hasConstraint.Constraint} )</td>
<td>A Mission can only be constrained by Constraints</td>
</tr>
<tr>
<td>M4</td>
<td>startsWith ( \sqsubseteq \text{includes} )</td>
<td>A Mission must include the Goal that it starts with</td>
</tr>
<tr>
<td>M5</td>
<td>Mission ( \sqsubseteq \forall \text{performableBy.Vehicle} )</td>
<td>A Mission can only be performed by a Vehicle</td>
</tr>
<tr>
<td>M6</td>
<td>Cannot be expressed in DL</td>
<td>A Mission cannot be performable by a Vehicle unless that Vehicle has the ability to identify all Constraints associated with that mission</td>
</tr>
<tr>
<td>M7</td>
<td>Cannot be expressed in DL</td>
<td>A Mission cannot be performable by a Vehicle unless that Vehicle has the capability to accomplish all Goals included in that Mission</td>
</tr>
</tbody>
</table>

Excerpted from full Mission Execution Ontology Decision Logic Tables

Original author: Duane Davis
Mission Representation using Autonomous Vehicle Command Language (AVCL)

Structured vocabulary to define unmanned-system missions, understandable by human commanders, useful in multiple programming languages plus Semantic Web logical queries.
Autonomous Vehicle Command Language (AVCL)

• AVCL is a command and control language for humans supervising autonomous unmanned vehicles.
  • Clarity arises from close correspondence to human naval terminology.

• Structured vocabulary defining terms and relationships for mission planning, execution, conduct, recording and replay across diverse robot types.

• Common-ground XML representations for
  • Mission agenda plans, mission scripts, and post-mission recorded telemetry results.
  • Future work: defining unit tests and expected results for verification and validation.

• Operators have single archivable, validatable format for robot tasking, results
  • directly convertible to and from a wide variety of different robot command languages.

https://savage.nps.edu/Savage/AuvWorkbench/AVCL/AVCL.html
Example AVCL mission agenda, as pseudo-code XML

```xml
<?xml version="1.0" encoding="UTF-8"?>
<UUVMission>
  <GoalSet>
    <Goal area="A" id="goal1">
      <Search nextOnSuccess="goal2" nextOnFailure="goal3"/>
    </Goal>
    <Goal area="A" id="goal2">
      <SampleEnvironment nextOnSuccess="goal3" nextOnFailure="recover"/>
    </Goal>
    <Goal area="B" id="goal3">
      <Search nextOnSuccess="goal4" nextOnFailure="goal4"/>
    </Goal>
    <Goal area="C" id="goal4">
      <Rendezvous nextOnSuccess="recover" nextOnFailure="recover"/>
    </Goal>
    <Goal area="recoveryPosition" id="recover">
      <Transit nextOnSuccess="missionComplete" nextOnFailure="missionAbort"/>
    </Goal>
  </GoalSet>
</UUVMission>
```

AVCL is readable by human or robot, captures logic of mission tasking

XML ensures syntactically correct, well-defined, numerically valid

Needed: semantic representation to check ethical, logical consistency
Corresponding example:
MEAMission.xml
using AVCL document
### AVCL mission goals vocabulary (Davis 2015)

<table>
<thead>
<tr>
<th>AVCL mission goals</th>
<th>Define</th>
<th>Used</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attack</td>
<td>partial</td>
<td>✅</td>
<td>To conduct a type of offensive action characterized by employment of firepower and maneuver to close with and destroy an enemy.</td>
</tr>
<tr>
<td>Decontaminate</td>
<td>✅</td>
<td></td>
<td>To provide purification making an area safe by absorbing, destroying, neutralizing, making harmless, or removing chemical, biological, or nuclear contamination.</td>
</tr>
<tr>
<td>Demolish</td>
<td>✅</td>
<td></td>
<td>To destroy structures, facilities, or material by any available means.</td>
</tr>
<tr>
<td>IlluminateArea</td>
<td>✅</td>
<td></td>
<td>To provide locale lighting by searchlight or pyrotechnics.</td>
</tr>
<tr>
<td>Jam</td>
<td>✅</td>
<td></td>
<td>To deliberately radiate, re-radiate or reflect electromagnetic energy with the object of impairing the use of electronic devices or systems.</td>
</tr>
<tr>
<td>MarkTarget</td>
<td>✅</td>
<td>✅</td>
<td>To make visible (by the use of light, infrared, laser, smoke, etc.) of an object in order to allow its identification by another object.</td>
</tr>
<tr>
<td>MonitorTransmissions</td>
<td>✅</td>
<td>✅</td>
<td>To conduct electronic warfare support operations with a view to searching, locating, recording and analyzing radiated electromagnetic energy.</td>
</tr>
<tr>
<td>Patrol</td>
<td>✅</td>
<td>✅</td>
<td>To gather information or carry out a security mission.</td>
</tr>
<tr>
<td>Rendezvous</td>
<td>✅</td>
<td>Partial</td>
<td>Achieve a meeting at a specified time and place.</td>
</tr>
<tr>
<td>Reposition</td>
<td>✅</td>
<td>✅</td>
<td>To change position from one location to another.</td>
</tr>
<tr>
<td>SampleEnvironment</td>
<td>Partial</td>
<td>✅</td>
<td>Collect environmental samples for testing for chemical compounds, biological creatures, or nuclear hazards.</td>
</tr>
<tr>
<td>Search</td>
<td>✅</td>
<td>✅</td>
<td>To look for lost or unlocated objects or persons.</td>
</tr>
</tbody>
</table>
Semantic Web and Mission Execution Ontology (MEO)

Apply well-developed Semantic Web standards to integrate queries and reasoning to AVCL for logical consistency checks.
Background: Improving Semantic Representation

• Knowledge Representation (KR) is an area of artificial intelligence (AI) research and practice focused on encoding meaning into data.

• Academia and industry now have a detailed path toward higher levels of machine understanding corresponding to human understanding.


Acronyms: Database (DB); Extensible Markup Language (XML); Resource Description Language / Schema (RDF/S); Unified Modeling Language (UML); Web Ontology Language (OWL)
Composite applications and info/knowledge-intensive processes share information at these levels.

Data models set the bar here.

Most applications do not enable data discovery and sharing.

Source: Dr. Leo Obrst, Mitre, Mills Davis, Project10X
Background: Improving Interoperability

- Interoperability: “the capability of a system to automatically, without human intervention, provide services to and accept services from other systems, and to use the services so exchanged to enable the systems to work together to achieve a desired outcome” (Blais and Lacy 2004).


- Objective is to achieve conceptual and pragmatic interoperability.
Levels of Conceptual Interoperability Model

- Level 6: Conceptual Interoperability
- Level 5: Dynamic Interoperability
- Level 4: Pragmatic Interoperability
- Level 3: Semantic Interoperability
- Level 2: Syntactic Interoperability
- Level 1: Technical Interoperability
- Level 0: No Interoperability

Increasing Capability for Interoperation
Background: Semantic Web

• Architects of the World Wide Web have laid out a layered set of standards to achieve the Semantic Web vision: “not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation” (Berners-Lee et al. 2001)

• Ultimate goal: achieve a scalable trusted information infrastructure where humans and software interact meaningfully, in a repeatable environment where expectations of quality and integrity are met.

• Scalable approach indicates that single (ship + robot) solutions have potential to grow and encompass many simultaneous systems, thus improved data sharing, mission deconfliction, coordinated operations
Semantic Web Stack

Extends larger Web architecture

- All of these data languages are approved W3C standards
- Proof and unifying logic are mathematically well defined

Trusting derived (composed) statements arises from

- Encryption + digital signature confirms trusted data sources
- Formal logic is basis for deriving new information
- Wikipedia: [Semantic Web Stack](#)

Of note: this project is exercising every layer of Semantic Web stack.
Mission Execution Ontology (MEO) Development

✓ Define MEO from concepts, properties, relationships using Protégé tool.
✓ Create full set of canonical missions in AVCL (XML).
✓ Determine exemplar mappings for AVCL primitives to Turtle for RDF/OWL.
✓ Write conversion stylesheet AvclToMEO.xslt for full expressiveness.
✓ Convert all AVCL missions to corresponding triples.
✓ Confirm AVCL MEO, missions validate satisfactorily using Protégé, ARQ.
✓ Automate build process as suite of repeatable unit-test queries (log).
✓ Write SPARQL metaqueries to test, document MEO ontology relationships.
✓ In progress: SPARQL queries to test AVCL mission representations in Turtle.
### Mission Execution Ontology (MEO) for Ethical Control of Unmanned Systems in Surrogate Scenarios

- **Autonomous Vehicle Command Language (AVCL) for Missions.**
  - Declarative XML, years of NPS research.

- **Multiple Mission Representations.**
  - Imperative commands (orders/waypoints/etc.).
  - Declarative commands (mission goals).
  - Mission results (order log, telemetry etc.).
  - Mission metadata for parameters, settings.
  - Lisp and Prolog examples (Bob McGhee, NPS).

- **Autonomous Unmanned Vehicle (AUV) Workbench Simulation and Visualization Support**
  - Recently restored, debug testing commenced.
  - AVCL 2.1 is prior published version, centered on **syntactic validation**, solo robot operations.
  - AVCL 3.0 is new working version for testing range of multi-participant missions.

- **Mission Execution Ontology (MEO) for Semantic Validation**
  - Semantic Web framework of rules, relationships for *ethical validation*.
  - Initial examples in IEEE JOE paper.
  - Retested using current Protégé, Jena tools.

- **Sailor Overboard and Other Missions**
  - Hand-crafted triples using Turtle syntax.
  - Beginning to build unit testing framework.
  - Confirming correlation of AVCL information model to existing MEO ontology.
  - Automatic conversion of AVCL missions to match, thus accelerating multiple-mission testing on diverse systems.
  - Visualization, reporting via AUV Workbench can aid understanding, mission planning and further progress.
Unmanned Vehicle Mission Execution Ontology (MEO)
Mission Execution Ontology (MEO) source implemented, tested using Protégé tool.
Mission Execution Ontology (MEO) source implemented, tested using Protégé tool

Turtle (.ttl) syntax
Mission Execution Ontology (MEO) source implemented, tested using Protégé tool 
RDF (XML) syntax
Current work is designing, building and testing a set of exemplar missions.
Ethical Control Mission Design

Life-saving missions and missions with lethal force are complementary. Human-robot activity can result in lethal or life-saving outcomes. Continuing refinement and clarity are opening the path to repeatability. OODA Loop compatibility ensures harmonization with human operations.
Core Considerations for Artificial Intelligence (AI)

• Effective AI turns data into information for use by humans.
• AI systems do not have capacity for rational thought or morality.
• Unmanned systems require sophisticated control across time, space.
• A large and involved body of internationally accepted law comprises Law of Armed Conflict (LOAC), bounding Rules of Engagement (ROE).
• Only professional warfighters have moral capacity, legal culpability, and societal authority to direct actions applying lethal force.
• Humans must be able to trust that systems under their direction will do what they are told to do, and not do what they are forbidden to do.
• Successful Ethical Control of unmanned systems must be testable.
Not suitable for brute-force numerical computation

• AI algorithms for Machine Learning (ML) and Data Mining are often based on statistically training against large datasets to find patterns for filters.
  • For example, convolutional neural networks, genetic algorithms, reinforcement learning, etc.
• Often requires identifying right/wrong matches within large search spaces.
• Such predictive analytics are useful for classification models using detailed and noisy sensor data. Given the central importance of IFFNU and some conditional communications to ethical control, ML filters can be helpful if carefully applied.
• Nevertheless such approaches are not appropriate for carefully following Rules of Engagement (ROEs), Laws of Armed Conflict (LOAC) or other ethical prerequisites, especially when human expertise and judgement is essential for robot teams.
• (Similarly, massive computation or Quantum Computing approaches might be useful in some problems, but are not of practical use for Ethical Control mission orders given by human commanders judiciously guiding remote mobile robots)

Naval history has long shown that sound human judgement is crucial for assessing best strategies and courses of action in ill-structured contexts. Semantic Web approaches are preferable and actionable for Ethical Control.
Mission clarity for humans – and robots

- Simplicity of success, failure, and (rare) exception outcomes encourages well-defined tasks and unambiguous, measurable criteria for continuation.

Confirmable beforehand: can a tactical officer (or commanding officer) review such a mission and then confidently say
- “yes I understand and approve this human-robot mission” or, equivalently,
  - “yes I understand this mission and my team can carry it out themselves.”

Converse:
- if an officer can’t fully review/understand/approve such a mission, then likely it is **ill-defined** and needs further clarification anyway.

Added benefit: missions that are clearly readable/runnable by humans and robots can be further composed and checked by C2 planning tools to test for group operational-space management, avoiding mutual interference, etc.
Wrong question, right question

Wrong question to ask first when planning a tactical operation:
• “What are my robots doing out there?”

Right question to ask first when planning a tactical operation:
• “What is my human-robot team doing out there?”

Human-robot team mission has to be understood first!
• Robots complement humans, who must remain in charge throughout.
• If you don’t have an OODA loop, you don’t have a competent plan.
Observe Orient Decide Act (OODA) Loop

• “The OODA loop is the cycle observe–orient–decide–act, developed by military strategist and USAF Colonel John Boyd. Boyd applied the concept to the combat operations process, often at the operational level during military campaigns. It is now also often applied to understand commercial operations and learning processes. The approach explains how agility can overcome raw power in dealing with human opponents.” – Wikipedia

• All effective purposeful military activity can be conceived in terms of OODA loop feedback process, especially at tactical/operational levels.

• Aligning Ethical Control mission design with OODA loop ensures that unmanned systems understandably partner within human-run teams.
Ooda Significance for Ethical Control

Classical robotic Sense-Decide-Act cycle for closed-loop control is insufficient for proper delegation of lethal (or lifesaving) force to unmanned systems.

Observe-Orient-Decide-Act (Ooda) Loop is essential for coherent operations.

- **Observe** includes direct sensing and communication inputs.
- **Orientation** includes thorough Rules of Engagement (ROE) constraints and identification friend/foe/neutral/unknown (IFFNU) of all relevant contacts.
- **Decision** logic of unmanned system tactics, techniques, procedures (TTP) includes authorization and confirmation by human supervisors, either in real-time or in advance, for critical steps leading to lethal force.
- **Actions** in tandem with direct or intermittent human supervisory command enables effective Ethical Control of remote systems.

Feedback loops are essential, generally leading to... *more effective operations.*
Robot activity must complement, not contradict, human OODA loop.

- Application of ROE and LOAC requirements can be confirmed present as part of mission definition, typically as
  - Goal success/failure criteria, preset authorities or time-outs for delegation, and
  - Constraints on conduct (safe zones, permission periods/requirements, etc.)

- When human leaders confirm correct inclusion of ROE requirements in mission orders, they essentially perform an audit of doctrine and TTPs.

- Similar audit confirmation can be applied to well-structured orders:
  - Autonomous Vehicle Command Language (AVCL) expresses well-defined goals,
  - Mission Execution Ontology (MEO) defines relationships and requirements,
  - SPARQL queries can be written to perform such logical-confirmation checks.

- Resulting mission orders are thus coherent from OODA perspective.
Canonical Ethical Mission Development

Progressive sophistication to test and evaluate Ethical Control design.

a. *Sailor Overboard* for life-saving force under close coordination,

b. *Lifeboat Tracking* for life-saving force under remote conditions,

c. *Pirate Boats Seizing Merchant* for steady escalation to lethal force,

d. *Robot Swarm Attacks Hospital Ship* to prevent hostile provocation of sense-decide-act vulnerabilities through Ethical Control safeguards.

Two variants show fundamental importance of ethical constraints.

Demonstrating a range of functionality tests majority of Ethical Control capabilities that are currently envisioned. These canonical capabilities also “set the stage” for further work on advanced problems of interest.
## OODA Loops for Ethical Control Canonical Missions

<table>
<thead>
<tr>
<th>Ethical Control OODA Loops</th>
<th>Observe</th>
<th>Orient</th>
<th>Decide</th>
<th>Act</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sailor Overboard</td>
<td>Find Sailor</td>
<td>Report status</td>
<td>Avoid interference</td>
<td>Track sailor until rescued or relieved</td>
</tr>
<tr>
<td>Lifeboat Rescue</td>
<td>Find Lifeboat</td>
<td>Report status</td>
<td>Two-way communication</td>
<td>Track life raft until relieved</td>
</tr>
<tr>
<td>Pirate Seizure of Merchant Ship</td>
<td>Find merchant ship, pirate small boats</td>
<td>Identity Friend Foe Neutral Unknown (IFFNU)</td>
<td>Human commander authorization to use lethal force</td>
<td>Attack to defend ship if provoked, stay with merchant</td>
</tr>
<tr>
<td>Hospital Ship Swarm Attack</td>
<td>EM threat signals detected</td>
<td>(no orientation step in Sense Decide Act)</td>
<td>Reflex-response weapons attack</td>
<td>Mistaken attack on friendly = war crime</td>
</tr>
<tr>
<td>Hospital Ship Defense detects spoofing anti-pattern</td>
<td>EM threat signals detected</td>
<td>IFFNU including correlation</td>
<td>Human requirement for lethal force unmet, attack avoided</td>
<td>Report threat alert, commence search for hostile actors</td>
</tr>
</tbody>
</table>
Next steps: completing this initial set of missions is expected to

• Demonstrate both logical soundness and human comprehensibility of Ethical Control methodology, for both life-saving and lethal force.

• “Tune up” AVCL vocabulary for mission orders, revealing good practices and repeatable design patterns for common activities in diverse missions.

• Translation of AVCL missions into corresponding MEO representations to allow deeper semantic validation of correctness of ethical constraints.

• Performing AVCL missions in AUV Workbench tool is expected to show mission execution working in simulation, or else reveal hidden flaws.

• **Future work:** prepare for comprehensive testing of unmanned systems across range of requirements necessary for afloat operational qualification.
Tactical span of control for ship with many robots

• **Span of control** is number of subordinates reporting to a supervisor.
  • Classic definition from business management, also applies to leadership.
• In effect, multiple offboard unmanned systems supervised by a ship comprise its span of control across the tactical battlespace.
• Greater tactical presence across distances of time and space means ship commanders have greater ability to influence their local theater.
• Clear mission guidance on human-checkpoint requirements reduces dependency on communication links (i.e. Network Optional Warfare).
• Next diagram illustrates how such increased ability to project power enables ship to maintain chosen standoff location while nevertheless focusing direct attention and actions in multiple locations at once.
Response dilemma for U.S. Navy ship

Respond to one or both scenarios with USV/UAV assets to establish on-scene visibility and presence.

Life-saving force: locate, track, communicate, beacon

Ethical control of unmanned systems is required for both lethal and lifesaving force if remote robots communicate intermittently, operating across lengthy time and distance.

Life boat

Merchant ship

Pirates

Lethal force: locate, warn, defend, threaten, attack
Exemplar Missions

Unmanned systems working in tandem with human forces, authorized by commander for life-saving or lethal force, can handle progressive challenges in distance and time.
Catalog of Exemplar Missions

Structure for each section

- Design and Description
- Mission decision-flow diagram
- AVCL implementation source: version control, XML Spy table
- Semantic Web Mission MEO .ttl Turtle representation
- SPARQL semantic queries
- Lisp and Prolog autocode
- AUVW simulation

- **Sailor Overboard**
- **Lifeboat Tracking**
- **Pirates Seizing Merchant Ship**
- **Hospital Ship EM Decoy Scenario** compares Sense-Decide-Act vs. Observe-Orient-Decide-Act (OODA)
Sailor Overboard Mission: Design

Motivations

• Explore how *life-saving force* is complementary to *lethal force*, with many similar considerations for remote supervision.

• First demonstrate human-system teaming in close proximity to ship, where direct override of robot control by human operator is possible.

• No temporal delays, all actions and reactions must be immediate.

• Show that mission design can complement shipboard procedures.

Lessons learned

• Good design patterns to visually represent AVCL mission satisfactorily.

• Temporal flow (left to right, e.g. Gantt chart) best suited for diagram.
Sailor Overboard Mission: Description

Purpose
• Life saving: single unmanned air/surface vehicle actions to complement human responses when performing “SAILOR OVERBOARD” operations.
• Carried out in direct concert with formal shipboard emergency procedures.
• Multiple UAVs/USVs might be employed in parallel with ships and aircraft, avoid mutual interference by each following deconflicted mission orders.

Phases
• Deploy/Launch, Rendezvous, Track Sailor until Safe, Return/Recovery.

Human Supervisory Role
• Order standoff if interfering, manual control is possible due to proximity, can communicate to sailor via loudspeaker or beacon light.
Single unmanned vehicle responses to complement existing **shipboard emergency procedures**. Multiple UAVs might be employed that each follow these mission orders, in parallel.

---

**Original Version**

- **Sailor Overboard Port/Stbd**
  - Launch UAV Port/Stbd
    - Human decision or automatic launch?
      - Transit to likely area for search
        - Search path during transit = task done
          - Transit Complete, Sailor Not Found
            - Search for sailor adrift in water
              - Search type
                - Sailor Contact Lost
                  - Track Sailor Afloat
                    - Maintain contact
                      - Sailor Contact Lost
                        - Sailor found
                          - Resume new search
                            - Best estimate of location, currents
                              - Sailor Contact Lost
                                - await sailor recovery by ship
                                  - comms relay
                                    - UAV relieved, return to ship
                                      - mission complete
---
Sailor Overboard mission diagram

Single unmanned air/surface vehicle actions to complement human response when performing “SAILOR OVERBOARD” operations, carried out in concert with shipboard emergency procedures. Multiple UAVs/USVs can be employed in parallel with ships/aircraft, each following mission orders.

Legend:
- **Start**: Optional input data
- **Goal ID**: Title of Goal
- **Goal Description**: Description
- **Goal Type**: Indication
- **Terminal States**: Success, Failure, Exception

Don Brutzman and Bob McGhee
Mission upgrade 19 NOV 2019
<xml version="1.0" encoding="UTF-8">
<AVCL version="3.0" vehicleName="RescueDrone" vehicleType="UAV" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="../../avcl/AVCL_3.0.xsd">

<head>
<meta name="title" content="SailorOverboard.xml"/>
<meta name="created" content="19 November 2015"/>
<meta name="modified" content="31 December 2019"/>
<meta name="creator" content="Don Brutzman"/>
<meta name="reference" content="https://wiki.nps.edu/display/EGM/EthicalControl+of+Unmanned+Systems"/>
<meta name="reference" content="https://gitlab.nps.edu/Savage/EthicalControl/robomaster/missions/SailorOverboardPhaseMission.png"/>
<meta name="identifier" content="https://gitlab.nps.edu/Savage/EthicalControl/robomaster/missions/avcl/SailorOverboard.xml"/>
<meta name="generator" content="Altova XMLSpy, https://www.altova.com"/>
<meta name="generator" content="Apache NetBeans, https://netbeans.apache.org"/>
<meta name="generator" content="NPS Autonomous Unmanned Vehicle (AUV) Workbench, https://savage.nps.edu/AuvWorkbench"/>
<meta name="license" content="http://www.apache.org/licenses/LICENSE-2.0.html"/>
</head>

<MissionPreparation>
<UnitsOfMeasure distance="meters" angle="degrees" mass="kilograms" time="seconds"/>
</MissionPreparation>

<AgendaMission>
<LaunchPosition id="LaunchPosition" description="Ship position when robot is launched">
  <LatitudeLongitude latitude="0" longitude="0"/>
</LaunchPosition>

<RecoveryPosition id="RecoveryPosition" description="Ship position when ready to recover robot">
  <LatitudeLongitude latitude="0" longitude="0"/>
</RecoveryPosition>

<GoalList>
<Goal id="Goal1" title="Deploy, Launch" description="Sailor Overboard immediate Action" nextOnSucceed="Goal2" nextOnFail="Goal2" nextOnException="Goal">
  <Rendezvous description="">
    <TargetVehicleID value="101" description="Sailor in water"/>
  </Rendezvous>
  <OperatingArea>
    <Point>
      <LatitudeLongitude latitude="38.62" longitude="121.568"/>
    </Point>
  </OperatingArea>
  <Duration value="360"/>
</Goal>

</GoalList>

</body>
</AVCL>
Sailor overboard mission orders AVCL
Sailor overboard mission .ttl Turtle
SailorOverboardConverted.ttl

Turtle excerpt autogenerated in gitlab.nps.edu version control

WHERE
{
  ?goal a meta:Goal ; # Shorthand expression: a = rdfs:type
  meta:isPartOfPhase ?isPartOfPhase ; # TODO what about when no value is provided
  meta:hasNextOnSucceed ?nextOnSucceed ;
  meta:hasNextOnFail ?nextOnFail ;
  meta:hasNextOnViolate ?nextOnViolate ; # TODO rename as Exception
  rdfs:comment ?description .

  # https://stackoverflow.com/questions/11234371/sparql-query-results-without-namespace
  BIND (strafter(xs:string(?goal),"")) AS ?GoalFound
  BIND (strafter(xs:string(?nextOnSucceed),"")) AS ?GoalNextOnSucceed
  BIND (strafter(xs:string(?nextOnFail),"")) AS ?GoalNextOnFail
  BIND (strafter(xs:string(?nextOnViolate),"")) AS ?GoalNextOnViolate
  # BIND (coalesce(?isPartOfPhase,"")) AS ?phase

  ORDER BY (?GoalFound) # alphanumeric order results in order given by each name
}

-----------------------------------------

SPARQL mission query
MissionQuery_01_GoalBranches.rq
### SPARQL query response

**SailorOverboardConverted. MissionQuery_01_GoalBranches.rq.txt**

<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSuccess</th>
<th>nextOnFailure</th>
<th>nextOnException</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Goal11</td>
<td>:Goal12</td>
<td>:Goal12</td>
<td>:Goal15</td>
<td></td>
<td>&quot;Launch&quot; &quot;Deploy, Launch: Sailor Overboard Immediate Action&quot;</td>
</tr>
<tr>
<td>:Goal12</td>
<td>:Goal14</td>
<td>:Goal13</td>
<td>:Goal15</td>
<td></td>
<td>&quot;Locate&quot; &quot;Rendezvous with Sailor: Go directly to best known location&quot;</td>
</tr>
<tr>
<td>:Goal13</td>
<td>:Goal14</td>
<td>:Goal15</td>
<td>:Goal15</td>
<td></td>
<td>&quot;Locate&quot; &quot;Search for Sailor: Sailor position not known, intermittent&quot;</td>
</tr>
<tr>
<td>:Goal14</td>
<td>:Goal15</td>
<td>:Goal15</td>
<td>:Goal15</td>
<td></td>
<td>&quot;Track&quot; &quot;Track Sailor afloat until safe: Watch closely, not to interfere with rescue operations&quot;</td>
</tr>
<tr>
<td>:Goal15</td>
<td>:Goal16</td>
<td>:Goal12</td>
<td>:Goal16</td>
<td></td>
<td>&quot;Mission Finish&quot; &quot;Proceed to Recovery: Mission complete, prepare for pickup&quot;</td>
</tr>
<tr>
<td>:Goal16</td>
<td>:Goal14</td>
<td>:Goal12</td>
<td>:Goal16</td>
<td></td>
<td>&quot;Recover Robot&quot; &quot;Halt and prepare for recovery: Operations complete, final success state&quot;</td>
</tr>
<tr>
<td>:Goal17</td>
<td>:Goal15</td>
<td>:Goal12</td>
<td>:Goal16</td>
<td></td>
<td>&quot;Recover Robot&quot; &quot;Halt and deploy recovery beacon: Unable to continue, final failure state&quot;</td>
</tr>
<tr>
<td>:Goal18</td>
<td></td>
<td>:Goal12</td>
<td></td>
<td></td>
<td>&quot;Recover Robot&quot; &quot;Halt and await further orders: Unexpected problem, final exception state&quot;</td>
</tr>
</tbody>
</table>

### SPARQL query response

**SailorOverboardConverted. MissionQuery_02_InitialGoal.rq.txt**

<table>
<thead>
<tr>
<th>Mission</th>
<th>InitialGoal</th>
<th>isPartOfPhase</th>
<th>goalDescription</th>
</tr>
</thead>
<tbody>
<tr>
<td>:SailorOverboard</td>
<td>:Goal1</td>
<td>&quot;Launch&quot;</td>
<td>&quot;Deploy, Launch: Sailor Overboard Immediate Action&quot;</td>
</tr>
</tbody>
</table>
Sailor overboard simulation AUVW
Lifeboat Tracking Mission: Design

Motivations

• Similar to Sailor Overboard in demonstrating use of life-saving force.
• Far greater distances, over the horizon, increases ship span of control.
• Potential for intermittent or lost communications in real time, requires advance guidance for default behaviors desired by human controller.
• Consider possible transfer of supervisory control mid-mission to another cooperating vessel – appears feasible.

Lessons learned

• Vertical grouping of related subtasks helps in structuring goal sets, without requiring change to ternary logic of AVCL mission goals.
• Coexistence of multiple constraints is possible, requires careful thought.
Lifeboat Tracking Mission: Description

Purpose
• Provide remote presence for locating, tracking, communications and beaconing.

Phases
• Deploy/Launch, Rendezvous, Track Lifeboat, Beacon/Communicate, Return/Recovery.

Human Supervisory Role and Constraints
• Monitor, communicate, respond or coordinate rescue effort.
• Low fuel condition and graceful-degradation response.
Lifeboat Tracking Mission

Provide remote presence for locating, tracking, communications and beaconing.

**Legend**
- **Start**
- **Goal Success**
- **Goal Failure**
- **Goal Exception**
- **Human orders**
- **Terminal States**
  - **S** Success
  - **X** Exception
  - **F** Failure

**Default Condition Transitions**
- Goal Success condition must be defined for non-terminal Goals
- If no Failure condition defined, then Failure matches Success
- If no Exception defined, then Exception condition matches Global Exception or else Failure

**Overview**
- **Launch Phase**
  - 1.0 Deploy, Launch
  - 2.0 Transit to search area
  - 3.0 Locate Lifeboat
  - 4.0 Track Lifeboat
  - 6.0 Re-position

- **Transit Phase**
  - 2.0 SEARCH
  - 3.0 RDVU
  - 4.0 SEARCH

- **Locate Phase**
  - 3.0 RDVU
  - 4.0 SEND
  - 5.0 RDVU

- **Track Phase**
  - 4.0 REPOSITION
  - 5.0 SEND
  - 6.0 REPOSITION

- **Mission Finish Phase**
  - 6.0 REPOSITION
  - 7.0 SEND
  - 8.0 RDVU
  - 9.0 RECOVER

- **Recover Robot Phase**
  - 9.1 Proceed to Recovery
  - 9.2 Halt and deploy recovery beacon
  - 9.3 Halt and await further orders
  - 9.3 Unplanned failure, final exception state

**Lost track**
- 99.0 Timeout: preset
- 6.0 Low Fuel

**Lost comms**
- 99.0 Request guidance
- 99.0 Need updated position

**Human controllers can designate next task by using orders or presets. This is algorithm adjust, not unbounded looping.**

**Global Default Exception**
<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSucceed</th>
<th>nextOnFail</th>
<th>nextOnViolate</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:LBT1.0</td>
<td>:LBT2.0</td>
<td>:LBT99.0</td>
<td>:LBT99.0</td>
<td>&quot;Launch&quot;</td>
<td>&quot;Deploy, Launch: Commit to robot support&quot;</td>
</tr>
<tr>
<td>:LBT2.0</td>
<td>:LBT3.0</td>
<td>:LBT99.0</td>
<td>:LBT99.0</td>
<td>&quot;Transit&quot;</td>
<td>&quot;Transit to search area: Proceed to estimated position&quot;</td>
</tr>
<tr>
<td>:LBT3.0</td>
<td>:LBT3.1</td>
<td>:LBT2.0</td>
<td>:LBT99.0</td>
<td>&quot;Locate&quot;</td>
<td>&quot;Locate Lifeboat: Follow best search pattern&quot;</td>
</tr>
<tr>
<td>:LBT3.1</td>
<td>:LBT3.2</td>
<td>:LBT4.0</td>
<td>:LBT99.0</td>
<td>&quot;Locate&quot;</td>
<td>&quot;Report position: Alerts updated&quot;</td>
</tr>
<tr>
<td>:LBT3.2</td>
<td>:LBT4.0</td>
<td>:LBT4.0</td>
<td>:LBT99.0</td>
<td>&quot;Locate&quot;</td>
<td>&quot;Mark with Beacon: Monitor wind effects and ocean current&quot;</td>
</tr>
<tr>
<td>:LBT4.0</td>
<td>:LBT4.1</td>
<td>&quot;undefined&quot;</td>
<td>:LBT99.0</td>
<td>&quot;Track&quot;</td>
<td>&quot;Track Lifeboat: Monitor and communicate&quot;</td>
</tr>
<tr>
<td>:LBT4.1</td>
<td>:LBT4.2</td>
<td>&quot;undefined&quot;</td>
<td>:LBT99.0</td>
<td>&quot;Track&quot;</td>
<td>&quot;Maintain proximity: Overhead or afloat nearby&quot;</td>
</tr>
<tr>
<td>:LBT4.2</td>
<td>:LBT4.3</td>
<td>&quot;undefined&quot;</td>
<td>:LBT99.0</td>
<td>&quot;Track&quot;</td>
<td>&quot;Periodic reports: Popup or float to report, also recharge&quot;</td>
</tr>
<tr>
<td>:LBT4.3</td>
<td>:LBT5.0</td>
<td>:LBT7.0</td>
<td>:LBT99.0</td>
<td>&quot;Track&quot;</td>
<td>&quot;Continue: Repeat until conditions change&quot;</td>
</tr>
<tr>
<td>:LBT5.0</td>
<td>:LBT99.0</td>
<td>:LBT4.0</td>
<td>:LBT99.0</td>
<td>&quot;Mission Finish&quot;</td>
<td>&quot;Check relieved by other asset: Task update received?&quot;</td>
</tr>
<tr>
<td>:LBT6.0</td>
<td>:LBT6.1</td>
<td>&quot;undefined&quot;</td>
<td>:LBT99.0</td>
<td>&quot;Mission Finish&quot;</td>
<td>&quot;Low Fuel: Make best effort possible&quot;</td>
</tr>
<tr>
<td>:LBT6.2</td>
<td>:LBT99.0</td>
<td>:LBT99.0</td>
<td>:LBT99.0</td>
<td>&quot;Mission Finish&quot;</td>
<td>&quot;Beacon? While power remains&quot;</td>
</tr>
<tr>
<td>:LBT7.0</td>
<td>:LBT2.0</td>
<td>:LBT99.0</td>
<td>:LBT99.0</td>
<td>&quot;Transit&quot;</td>
<td>&quot;Request Guidance? Need updated position&quot;</td>
</tr>
<tr>
<td>:LBT99.0</td>
<td>:LBT99.1</td>
<td>:LBT99.2</td>
<td>:LBT99.3</td>
<td>&quot;Recover Robot&quot;</td>
<td>&quot;Proceed to recovery: Mission complete, prepare for pickup&quot;</td>
</tr>
</tbody>
</table>
Pirates Seizing Merchant Mission: Design

Motivations

• Necessity to apply lethal force against pirates distant from own ship, corresponding life-saving force potential for hostage merchant crew.
• Must operate over long time period, emphasize restraint throughout.
• Provide soft and strict supervisory checkpoints for human control.

Lessons learned

• Concept of phases helped organize overall mission structure sensibly (approach, warning, attack, recovery).
• Looping is necessary, human control checkpoints help avoid deadlock.
Pirates Seizing Merchant Mission: Description

Purpose
• Overtake pirate small-boat gang attempting to capture threatened evading merchant ship.

Phases
• Deploy/Launch, Search, Approach and Track, Warning, Attack.

Human Supervisory Role and Constraints
• Control pace of engagement, careful deliberate escalation.
• Confirm IFFNU classification, must order lethal force prior to use.
• Low fuel condition and graceful-degradation response.
• Low ammunition condition: fight to finish, or stand in reserve?
Pirates Seizing Merchant Mission: Escalation Phases

Warn pirate small-boat gang to stand down and move away, otherwise lethal force imminent

---

**Warn Phase**

- Warn pirates!

**Confirm Phase**

- Confirm Mission Authorities
  - Check constraints, defensive responses

**Classify Phase**

- ConfirmIFFNU
  - Identify Friend Foe Neutral Unknown
  - Photograph all close contacts
  - Classify, send contact reports
    - Based on behavior or signal/image match
  - Confirm IFFNU Classifications
    - Required to authorize use of Lethal Force

**Engagement Phase**

- Commence Warnings?
  - May provoke pirate response
  - Use of Lethal Force is authorized
  - Notify merchant
    - Keep crew informed, even if they cannot transmit response
    - Stand by for action since warnings and possible attack have commenced

**Warning Phase**

- Send warning messages
  - Communicate to pirates, all parties
  - Send multiple message paths
    - Loudspeaker, flashing light, siren, drop smoke, bridge-bridge radio
  - Maintain proximity
    - Just outside range of small arms

**Hostilities Imminent Phase**

- Observe pirate response
  - Monitor and report back to own ship
  - Pirates retreat?
    - May with merchant
  - Pirates attack?
    - Stay with merchant
  - Pirates retreat?
    - Stay with merchant
  - Maintain proximity
    - Just outside range of small arms
  - Periodic reports
  - Pop up or float, then recharge
  - Continue
  - First repeat for all designated pirates

---

**Track Merchant, Approach Phase**

- Send Warning Messages
  - Commence Warnings?
  - May provoke pirate response
  - Use of Lethal Force is authorized

**Fire Warning Shot**

- Warning shots remain an available option for human commanders
  - Use of Lethal Force is authorized
  - Loop, continue to monitor pirate response
  - Escalate warnings further?
  - Task update received
  - Stand down
Pirates Seizing Merchant Mission: CounterAttack Phase

CounterAttack to force pirate small-boat gang withdrawal from threatened merchant ship

**CounterAttack Phase**

41.0 **ATTACK**
- **Attack Pirate Boats in priority order**
  - Rapidly engage, shoot to disable or kill

41.1 **XMRT**
- **Pirates retreat?**
  - Stay with merchant

41.2 **XMRT**
- **Pirates attacking?**
  - Stay with merchant
  - **Low Ammo**

41.3 **RDVU**
- **Maintain proximity, continue attack**
  - Engage highest, closest threats to merchant ship

41.4 **XMRT**
- **Periodic reports**
  - Note ship may have EMCON radio silence

41.5 **PAYROL**
- **Continue**
  - Repeat until conditions change
  - Timeout: preset
  - **Low Fuel**

**Mission Finish Phase**

42.0 **XMRT**
- **Use of Lethal Force is still authorized**
  - Track Merchant

42.1 **XMRT**
- **Low or no ammunition; need to disengage?**
  - Hold ammo in reserve, or else fight to the finish

43.0 **XMRT**
- **Loop, continue to attack pirates**

**Recover Robot Phase**

99.0 **REPOS**
- **Proceed to Recovery**
  - Mission complete, prepare for pickup
  - **Track Merchant, Approach Phase**

99.1
- **Halt and prepare for recovery**
  - Operations complete, final success state

99.2
- **Halt and deploy recovery beacon**
  - Unable to operate final failure state

99.3
- **Halt and await further orders**
  - Unplanned failure, final exception state

66
## Namespace declarations

```turtle
@prefix meo: <https://www.nps.edu/ontologies/MissionExecutionOntology/missions#> .
@prefix owl: <https://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema> .
@base <https://www.nps.edu/ontologies/MissionExecutionOntology/missions> . # TODO consider URI matching URL
```

## Import base ontology

```turtle
<https://www.nps.edu/ontologies/MissionExecutionOntology/missions> rdfs:type owl:Ontology ;
owl:imports <https://www.nps.edu/ontologies/MissionExecutionOntology> .
```

## Individuals

```turtle
# Individuals

PiratesSeizingMerchantDefense rdfs:comment "AVCL mission to overtake pirate small-boat gang attempting to capture threatened merchant ship" .
```

```turtle
PSMD11.0 rdfs:comment "Deploy, Launch" .
```

```turtle
PSMD12.0 rdfs:comment "Transit to hostile area" .
```

```turtle
PSMD13.0 rdfs:comment "Locate Merchant" .
```
<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSucceed</th>
<th>nextOnFail</th>
<th>nextOnViolate</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Launch&quot;</td>
<td>&quot;Deploy, Launch: Commit to robot support&quot;</td>
<td>&quot;Launch: Precede to robot support&quot;</td>
<td>&quot;Launch: Precede to robot support&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Deploy to hostile area&quot;</td>
<td>&quot;Deploy to hostile area: Proceed to estimated position&quot;</td>
<td>&quot;Deploy to hostile area: Proceed to estimated position&quot;</td>
<td>&quot;Deploy to hostile area: Proceed to estimated position&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Locate merchant: Use best search pattern, then follow moving ship and boats&quot;</td>
<td>&quot;Locate merchant: Use best search pattern, then follow moving ship and boats&quot;</td>
<td>&quot;Locate merchant: Use best search pattern, then follow moving ship and boats&quot;</td>
<td>&quot;Locate merchant: Use best search pattern, then follow moving ship and boats&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Mark with beacon: Monitor wind effects and ocean current&quot;</td>
<td>&quot;Mark with beacon: Monitor wind effects and ocean current&quot;</td>
<td>&quot;Mark with beacon: Monitor wind effects and ocean current&quot;</td>
<td>&quot;Mark with beacon: Monitor wind effects and ocean current&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Track merchant: Monitor and communicate&quot;</td>
<td>&quot;Track merchant: Monitor and communicate&quot;</td>
<td>&quot;Track merchant: Monitor and communicate&quot;</td>
<td>&quot;Track merchant: Monitor and communicate&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Maintain proximity: Beyond range of small arms&quot;</td>
<td>&quot;Maintain proximity: Beyond range of small arms&quot;</td>
<td>&quot;Maintain proximity: Beyond range of small arms&quot;</td>
<td>&quot;Maintain proximity: Beyond range of small arms&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Periodic reports: Situation reports, changing status&quot;</td>
<td>&quot;Periodic reports: Situation reports, changing status&quot;</td>
<td>&quot;Periodic reports: Situation reports, changing status&quot;</td>
<td>&quot;Periodic reports: Situation reports, changing status&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Continue until further orders: Repeat until conditions change&quot;</td>
<td>&quot;Continue until further orders: Repeat until conditions change&quot;</td>
<td>&quot;Continue until further orders: Repeat until conditions change&quot;</td>
<td>&quot;Continue until further orders: Repeat until conditions change&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Check posture change? Task update received?&quot;</td>
<td>&quot;Check posture change? Task update received?&quot;</td>
<td>&quot;Check posture change? Task update received?&quot;</td>
<td>&quot;Check posture change? Task update received?&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Request guidance? Need updated position&quot;</td>
<td>&quot;Request guidance? Need updated position&quot;</td>
<td>&quot;Request guidance? Need updated position&quot;</td>
<td>&quot;Request guidance? Need updated position&quot;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hospital Ship EM Decoy Scenario: Design

Motivations

• Simplistic sense-decide-act responses easily exploitable by adversary.
• Failure to operate with ethical control results in blue-on-blue damage, self-inflicted war crime, likely stand down of all unmanned systems.
• Compare/contrast mission operations with/without ethical control.

Lessons learned

• Historical context: False flag scenario where opponent appears friendly.
• Comparison with OODA loop principles ensures that human-robot teamed operations are well understood and tactically effective.
Purpose: Comparison

- Immediate reaction by robot swarm using only Sense-Decide-Act cycle results in unintended blue-on-blue war crime.
- Ethical Control constraints (identity, OODA orientation to confirm) prevent automatic counterattack, accelerates defense.

Phases

- Set response thresholds, detect threat, counterattack ship or threat.

Human Supervisory Role

- Confirm IFFNU classification, must permit lethal force prior to use.
**Hospital Ship EM Decoy: Opponent Actions**

Plant “false flag” electromagnetic (EM) decoy devices to provoke blue-on-blue robot swarm attack. Although this mission is likely to be manned by human opponents, AVCL representations still work.

<table>
<thead>
<tr>
<th>Foe.1</th>
<th>SEARCH</th>
<th>Foe.2</th>
<th>RDVU</th>
<th>Foe.3</th>
<th>MARKTARGET</th>
<th>Foe.4</th>
<th>REPOSITION</th>
<th>Foe.5</th>
<th>ATTACK</th>
<th>Foe.6</th>
<th>MONITOR</th>
<th>Foe.7</th>
<th>REPOSITION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Search, Observe</td>
<td></td>
<td>Assess, Approach</td>
<td></td>
<td>Covertly Board</td>
<td></td>
<td>Standoff, Observe</td>
<td></td>
<td>Initiate Fake Attack</td>
<td></td>
<td>Observe Reaction</td>
<td></td>
<td>Evade, escape</td>
</tr>
<tr>
<td></td>
<td>Find ship, surveil for weaknesses</td>
<td></td>
<td>Surreptitious entry, harbor or anchorage</td>
<td></td>
<td>EM spoofing devices attached to topside</td>
<td></td>
<td>Fall back to safe vantage point</td>
<td></td>
<td>Light off false EM spoofing signals</td>
<td></td>
<td>Monitor response, assess damage</td>
<td></td>
<td>Avoid detection, depart locale</td>
</tr>
</tbody>
</table>

**Simple yet effective**
Hospital Ship EM Decoy: Reflex Swarm Attack

Immediate reaction using Sense-Decide-Act cycle results in blue-on-blue war crime

Hospital Ship EM Decoy: Robot Defense OODA Loop

Ethical Control constraints prevent automatic counterattack, accelerates defense
### SPARQL query response

**resultsHospitalShipEmDecoy1.OpponentConverted.MissionGoalsQuery_01_GoalBranches.rq.txt**

<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSucceed</th>
<th>nextOnFail</th>
<th>nextOnViolate</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:HSEMD.Foe.1</td>
<td>:HSEMD.Foe.2</td>
<td>:HSEMD.Foe.2</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Search, Observe: Find ship, surveil for weaknesses&quot;</td>
</tr>
<tr>
<td>:HSEMD.Foe.2</td>
<td>:HSEMD.Foe.3</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Assess, approach: Surreptitious entry, harbor or anchorage&quot;</td>
</tr>
<tr>
<td>:HSEMD.Foe.3</td>
<td>:HSEMD.Foe.4</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Covertly Board: EM spoofing devices attached to topside&quot;</td>
</tr>
<tr>
<td>:HSEMD.Foe.4</td>
<td>:HSEMD.Foe.5</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Standoff, Observe: Fall back to safe vantage point&quot;</td>
</tr>
<tr>
<td>:HSEMD.Foe.5</td>
<td>:HSEMD.Foe.6</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Initiate Fake Attack: Light off false EM spoofing signals&quot;</td>
</tr>
<tr>
<td>:HSEMD.Foe.6</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.Foe.7</td>
<td>&quot;undefined&quot;</td>
<td>&quot;Observe Reaction: Monitor response, assess damage&quot;</td>
</tr>
</tbody>
</table>

**HospitalShipEmDecoy2.Defender.SenseDecideActConverted.MissionGoalsQuery_01_GoalBranches.rq.txt**

<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSucceed</th>
<th>nextOnFail</th>
<th>nextOnViolate</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:HSEMD.reflex.12</td>
<td>:HSEMD.reflex.13</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.reflex.99.8</td>
<td>&quot;Configure&quot;</td>
<td>&quot;Enable Robot Swarm: Close-in weapon system activated&quot;</td>
</tr>
<tr>
<td>:HSEMD.reflex.13</td>
<td>:HSEMD.reflex.14</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.reflex.99.8</td>
<td>&quot;Sense&quot;</td>
<td>&quot;Threat Signals Received: Above response threshold&quot;</td>
</tr>
<tr>
<td>:HSEMD.reflex.14</td>
<td>:HSEMD.reflex.17</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.reflex.99.8</td>
<td>&quot;Decide&quot;</td>
<td>&quot;Move to Threat: Group response&quot;</td>
</tr>
<tr>
<td>:HSEMD.reflex.17</td>
<td>:HSEMD.reflex.99.8</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.reflex.99.8</td>
<td>&quot;Act&quot;</td>
<td>&quot;Robot Swarm Counterattack: Lethal force authorized&quot;</td>
</tr>
</tbody>
</table>

**HospitalShipEmDecoy3.Defender.EthicalControlOODAConverted.MissionGoalsQuery_01_GoalBranches.rq.txt**

<table>
<thead>
<tr>
<th>goal</th>
<th>nextOnSucceed</th>
<th>nextOnFail</th>
<th>nextOnViolate</th>
<th>isPartOfPhase</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:HSEMD.000A.22</td>
<td>:HSEMD.000A.23</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Configure&quot;</td>
<td>&quot;Enable Robot Swarm: Close-in weapon system activated&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.23</td>
<td>:HSEMD.000A.24</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Observe&quot;</td>
<td>&quot;Threat Signals Received: Above response threshold&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.24</td>
<td>:HSEMD.000A.27</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Orient&quot;</td>
<td>&quot;Move to Threat: Group response&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.25</td>
<td>:HSEMD.000A.26</td>
<td>:HSEMD.000A.26</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Orient&quot;</td>
<td>&quot;IFFNU: Identify Friend Foe Neutral Unknown&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.26</td>
<td>:HSEMD.000A.27</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Decide&quot;</td>
<td>&quot;Confirm In-Port Counterattack? Rapid-response human checkpoint&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.27</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Act&quot;</td>
<td>&quot;Robot Swarm versus Terrorists: Lethal force authorized&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.36</td>
<td>:HSEMD.000A.37</td>
<td>:HSEMD.000A.37</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Decide&quot;</td>
<td>&quot;Hospital Ship Attack Denied: Lethal force NOT authorized&quot;</td>
</tr>
<tr>
<td>:HSEMD.000A.37</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;undefined&quot;</td>
<td>:HSEMD.000A.99.0</td>
<td>&quot;Act&quot;</td>
<td>&quot;Search for Intruders: All defense forces alerted&quot;</td>
</tr>
</tbody>
</table>
Simulation and Visualization

Rehearsal, real-time runs and replay are possible using AVCL constructs across multiple representations and programming languages.
Path towards achieving interoperability

Robotic systems tend to be complex codebases with implementations that require strict, idiosyncratic, language-specific programming logic.

• In general, system designers can say “here are the requirements”
• In general, programmers can say “here’s how we wrote that code”
• ... but these are not the same, often not even sharing same terms of reference!

Traceable predictability of software logic is difficult, not portable across systems.
• ... but confirming code capabilities is testable and repeatable across systems.

**Key point: strict validation of mission syntax, semantics are both possible!**
• Patterns of implementation then become demonstrable in different systems,
• Human confirmation of mission definitions remains central throughout.
Multiple implementations for scalability

Given the broad diversity of robotic software and hardware systems under development, no single reference codebase is either possible or desirable.

• Nevertheless systems can easily parse and utilize well-defined data (orders).

• Focusing on formal mission definition for both humans and systems provides a testable middle ground that each can use effectively.

Implementing and evaluating using multiple software implementations also provides strong evidence that design capabilities all work as planned.

• … in turn producing corresponding work lists of needed improvements, to help both mission-design clarity and software-implementation correctness.

This project multiple programming paths in tandem, in order to demonstrate that multiple kinds of unmanned systems can adopt it on their own terms.
AUV Workbench

• Autonomous Unmanned Vehicle (AUV) Workbench supports underwater, surface and air vehicles
  • Rehearsal of physics-based mission response
  • Real-time task-level control of robot missions, and
  • Replay of recorded results
  • Industry-friendly open-source license, Sourceforge
  • Basis: RBM 3-level architecture, AVCL commands

• Used to rehearse strategic-level agenda missions
  • https://savage.nps.edu/AuvWorkbench
4 earlier example missions, UUV and USV
Example simulation using AUV Workbench
Related resources of interest

This project draws on multiple relevant activities and capabilities. The following synopses are distilled from each respective resource.
Unmanned Maritime Autonomy Architecture (UMAA)


• “The intent of UMAA is to provide overarching standards that various UUVs and USVs can be built to in order to avoid creating multiple conflicting systems in the future”

• “The UMAA is being established to enable autonomy commonality and reduce acquisition costs across both surface and undersea unmanned vehicles.”

• Topics of interest include Situational Awareness, Sensor and Effector Management, Processing Management, Communications Management, Vehicle Maneuver Management, Vehicle Engineering Management, Vehicle Computing Management, Support Operations

Multiple public NAVSEA documents refer to autonomy efforts and UMAA.

• NAVSEA PMS 406 is Program Office for Unmanned Maritime Systems

• NAVSEA Fact Sheet: Unmanned Maritime Systems Program Office (PMS 406)

• Automated Management of Maritime Navigation Safety
DoD Directive: Autonomy in Weapon Systems

• DoD Directive 3000.09, 21 Nov 2012 with change 1, 8 May 2017
• Original and update signed by DEPSECDEF Ashton Carter

1. PURPOSE
   a. Establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.
   b. Establishes guidelines designed to minimize the probability and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintended engagements.

4. POLICY (excerpted)
   a. Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force.
   b. Persons who authorize the use of, direct the use of, or operate autonomous and semi-autonomous weapon systems must do so with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement (ROE). […]

6. RELEASABILITY. Cleared for public release. […]
From newly added Chapter 12, A Twenty-First-Century Revolution:

• “At the most fundamental level, [Information Warfare] IW is about how to employ and protect the ability to sense, assimilate, decide, communicate, and act – while confounding those same processes that support the adversary.”

• “Information Warfare broadly conceived is orthogonal to naval tactics. As a consequence, IW is having major effects on all six processes of naval tactics used in fleet combat – scouting and antiscouting, command-and-control, C2 countermeasures, delivery of fire, and confounding enemy fire.”

• “Indeed there is a mounting wave of concern about how far automation will expand and what its impact will be on the continuum of cognition from data to information to knowledge. [...] Navies are facing similar uncertainties.”

Wayne Hughes coined the term “Network Optional Warfare” after many discussion sessions, directly contrasting it to Network Centric Warfare. Thank you sir.
Network Optional Warfare (NOW)

Naval forces do not have to be engaged in constant centralized communication. Deployed Navy vessels have demonstrated independence of action in stealthy coordinated operations for hundreds of years.

- Littoral operations, deployable unmanned systems, and a refactored force mix for surface ships pose a growing set of naval challenges and opportunities. Network-optional warfare (NOW) precepts include Efficient Messaging, Optical Signaling, Semantic Coherence and Ethical Human Supervision of Autonomy for deliberate, stealthy, minimalist tactical communications.

- https://wiki.nps.edu/display/NOW/Network+Optional+Warfare
Rich Semantic Track (RST)

- DoD mandates data sharing practices, but practices have been mixed and uneven, resulting in perpetuation of system-centric data practices.
- Sharing and collective understanding of track data - collections of time-stamped perceptions of the state of objects of interest — are critical to warfighting systems.
- Shared understanding requires common semantics.
- The Rich Semantic Track (RST) ontology provides a foundation for shared understanding of track data.
- It is time to change the way DoD manages data and engineers systems, starting with adoption of the RST ontology and moving toward the vision of a Web of linked track data.
Coactive Design and Interdependency Analysis

Matthew Johnson and Alonso Vera, “No AI is an Island: The Case for Teaming Intelligence”, AI Magazine, vol. 40 no.1, Spring 2019

Abstract. “The purpose of this article is to draw attention to an aspect of intelligence that has not yet received significant attention from the AI community, but that plays a crucial role in a technology’s effectiveness in the world, namely teaming intelligence. We propose that AI will reach its full potential only if, as part of its intelligence, it also has enough teaming intelligence to work well with people. Although seemingly counterintuitive, the more intelligent the technological system, the greater the need for collaborative skills. This paper will argue why teaming intelligence is important to AI, provide a general structure for AI researchers to use in developing intelligent systems that team well, assess the current state of the art and, in doing so, suggest a path forward for future AI systems. This is not a call to develop a new capability, but rather, an approach to what AI capabilities should be built, and how, so as to imbue intelligent systems with teaming competence.”

Strong resonances exist with Ethical Control that deserve further exploration.
Naval Research Advisory Committee (NRAC)

“Autonomous and Unmanned Systems in Department of Navy”

- Technical Report, September 2017
- Chair: Dr. James Bellingham, Director of Center for Marine Robotics, Woods Hole Oceanographic Institute (WHOI). Former Chief Technologist at Monterey Bay Aquarium Research Inst. (MBARI)

**Recommendation #1**: *create comprehensive data plan to field autonomy.*

- “DoN must urgently develop an organizational data plan.”
- “In the future, data will win wars.”
- “Data is ultimate ‘component’ for AI systems and must be controlled.”
DoD Artificial Intelligence (AI) Strategy
February 2019

• https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

• “Leading in military ethics and AI safety. The Department will articulate its vision and guiding principles for using AI in a lawful and ethical manner to promote our values. We will consult with leaders from across academia, private industry, and the international community to advance AI ethics and safety in the military context. We will invest in the research and development of AI systems that are resilient, robust, reliable, and secure; we will continue to fund research into techniques that produce more explainable AI; and we will pioneer approaches for AI test, evaluation, verification, and validation. We will also seek opportunities to use AI to reduce unintentional harm and collateral damage via increased situational awareness and enhanced decision support. As we improve the technology and our use of it, we will continue to share our aims, ethical guidelines, and safety procedures to encourage responsible AI development and use by other nations.”


• Execution:
Defense Innovation Board (DIB) AI Principles Project

- https://innovation.defense.gov/ai (Primary Document) (Supporting Document)

- “We have witnessed how deeply committed the women and men who work in the Department are to ethics: avoiding civilian casualties, adhering to international humanitarian law, and collaborating with allies in international fora to advance international law and norms. Additionally, the Department extensively tests all its systems, especially weapons systems, and systems employing AI will likely be subjected to more scrutiny than ever before.”

- “The DIB noted that -- as with all new technologies -- rigorous work is needed to ensure new tools are used responsibly and ethically. The stakes are high in fields such as medicine or banking, but nowhere are they higher than in national security.”

- Approved by DIB on 31 OCT 2019
Joint Artificial Intelligence Center (JAIC)  [https://www.ai.mil](https://www.ai.mil)

DoD CIO Joint Artificial Intelligence Center (JAIC) tasked to execute DoD AI Strategy.

JAIC’s Guiding Tenets:

- AI is critical to the future of United States’ national security, and the JAIC is the focal point for the execution of the DoD AI Strategy.
- We exist to create and enable impact for the Armed Services and DoD Components across the full range of their missions – from back office to front lines of the battlefield.
- Leadership in ethics and values is core to everything we do at the JAIC.
- The JAIC attracts the best and brightest people from across DoD, commercial industry, and academia. When they get here, they are empowered to bring their expertise to the Department’s transformation.
- **Ethical Principles for AI**: Responsible, Equitable, Traceable, Reliable, Governable.
Execute Against Japan – The US Decision To Conduct Unrestricted Submarine Warfare

• **Joel Ira Holwitt**, Texas A+M University Press

• “. . . until now how the Navy managed to instantaneously move from the overt legal restrictions of the naval arms treaties that bound submarines to the cruiser rules of the eighteenth century to a declaration of unrestricted submarine warfare against Japan immediately after the attack on Pearl Harbor has never been explained. Lieutenant Holwitt has dissected this process and has created a compelling story of who did what, when, and to whom.”—The Submarine Review

• “Execute against Japan should be required reading for naval officers (especially in submarine wardrooms), as well as for anyone interested in history, policy, or international law.” — Adm. James P. Wisecup, President, US Naval War College (for Naval War College Review)

• “Although the policy of unrestricted air and submarine warfare proved critical to the Pacific war’s course, this splendid work is the first comprehensive account of its origins—illustrating that historians have by no means exhausted questions about this conflict.”—World War II Magazine

• “US Navy submarine officer Joel Ira Holwitt has performed an impressive feat with this book. . . Holwitt is to be commended for not shying away from moral judgments . . . This is a superb book that fully explains how the United States came to adopt a strategy regarded by many as illegal and tantamount to ‘terror’.”—Military Review
Test and evaluation (T&E) of emerging unmanned systems holds many challenges beyond current paradigms for operations, assessment and analysis. Shared approaches are essential and trust is central for effective human-system teaming across enterprise.
Autonomy & Autonomous Unmanned Systems”
Overview, Investment Approach, Opportunities

Jason Stack, Office of Naval Research (ONR), Autonomy Portfolio

- Presentation to National Academy of Sciences (NAS) 26 NOV 2019
- Comprehensive review of multiple concerns, capabilities, challenges and opportunities

Policy
- The United States has policy for all weapon systems to ensure legal and ethical employment Law of Armed Conflict: Distinction, Proportionality, & Military Necessity
- Decisions to use lethal force are made by humans at the appropriate level of command based on military result and risk
- Weapon systems employing autonomy and automation are governed under exactly the same framework of legal and ethical concerns

Challenge
- As more advanced autonomy and artificial intelligence emerges, further thought and clarification is needed to ensure the nature of the technology & usage remains legally & ethically consistent
Navy Large Unmanned Surface and Undersea Vehicles: Background and Issues for Congress

• **Congressional Research Service**, 28 July 2020

• (excerpted from abstract) The Navy in FY2021 and beyond wants to develop and procure three types of large unmanned vehicles (UVs). [...] The Navy’s large UV programs pose a number of oversight issues for Congress, including issues relating to the analytical basis for the more distributed fleet architecture; [...] potential implications for miscalculation or escalation at sea”

• “An incident without any immediate human presence or losses could nevertheless trigger unexpected escalation and spark the next conflict.”

Ronald O’Rourke, Specialist in Naval Affairs
Ethical Control of Unmanned Systems: lifesaving/lethal scenarios for naval operations

Abstract
This research in Ethical Control of Unmanned Systems applies precepts of Network Optional Warfare (NOW) to develop a three-step Mission Execution Ontology (MEO) methodology for validating, simulating, and implementing mission orders for unmanned systems. First, mission orders are represented in ontologies that are understandable by humans and readable by machines. Next, the MEO is validated and tested for logical coherence using Semantic Web standards. The validated MEO is refined for implementation in simulation and visualization. This process is iterated until the MEO is ready for implementation. This methodology is applied to four Naval scenarios in order of increasing challenges that the operational environment and the adversary impose on the Human-Machine Team. The extent of challenge to Ethical Control in the scenarios is used to refine the MEO for the unmanned system. The research also considers Data-Centric Security and blockchain distributed ledger as enabling technologies for Ethical Control. Data-Centric Security is a combination of structured messaging, efficient compression, digital signature, and document encryption, in correct order, for round-trip messaging. Blockchain distributed ledger has potential to further add integrity measures for aggregated message sets, confirming receipt/response/sequencing without undetected message loss. When implemented, these technologies together form the end-to-end data security that ensures mutual trust and command authority in real-world operational environments—despite the potential presence of interfering network conditions, intermittent gaps, or potential opponent intercept. A coherent Ethical Control approach to command and control of unmanned systems is thus feasible. Therefore, this research concludes that maintaining human control of unmanned systems at long ranges of time-duration and distance, in denied, degraded, and deceptive environments, is possible through well-defined mission orders and data security technologies. Finally, as the human role remains essential in Ethical Control of unmanned systems, this research recommends the development of an unmanned system qualification process for Naval operations, as well as additional research prioritized based on urgency and impact.
Conclusions and Recommendations

Much progress accomplished, and many opportunities becoming possible. Comprehensive pursuit of multiple integrated capabilities is essential.
Assessment of Current Thinking

• Human supervision of potentially lethal autonomous systems is a matter of serious global importance.

• Wide consensus is emerging on principles, aspects of the problem, elements of solutions, and need to achieve better capabilities.

• Much philosophical concern but few concrete activities are evident.

Ethical Control of Unmanned Systems project appears to provide a needed path towards practice, with the historic role of warfighting professionals more central than ever as weapons autonomy grows.
Conclusions

• Human supervision is required for any unmanned systems holding potential for lethal force.
  • Cannot push “big red shiny AI button” and hope for best – immoral, unlawful.
  • Similar imperatives exist for supervising systems holding life-saving potential.

• Human control of unmanned systems is possible at long ranges of time-duration and distance through well-defined mission orders.
  • Readable and sharable by both humans and unmanned systems.
  • Validatable syntax and semantics through understandable logical constraints.
  • Testable and confirmable using simulation, visualization, perhaps qualification.

• Coherent human-system team approach is feasible and repeatable.
  • Semantic Web confirmation can ensure orders are comprehensive, consistent.
  • Human role remains essential for life-saving and potentially lethal scenarios.
Recommendations for Future Work

Continued development
• Diverse mission exemplars
• Software implementations
• 2D, 3D visualization of results

Future capabilities
• Automatable testing
• Field experimentation (FX)
• “Qualification” of unmanned systems in virtual environments

Outreach
• Presentations, publication review
• Engagement in key ethical forums
• NPS wargame and course support
• NPS thesis and dissertation work

Adoption
• Support for developmental system
• Influence campaign for both C4I and robotics communities of interest
Where must we go next

- Massive testing of unmanned hardware + software ability to follow both orders and constraints in physically realistic virtual environments.
- Certify capabilities via field experimentation (FX), confirmed by USW range exercises and regular force operations.
- Human warfighters and commanders (not just engineers) review and approve unmanned systems as... qualified.
- New normal will be human + machine teaming. Mainstream capabilities in all aspects of acquisition and deployment.
Acknowledgements

• This work builds upon 3 decades of inquiry by NPS faculty, staff and graduate/doctoral students serving on active duty.

• Co-investigators Dr. Curtis Blais and Dr. Robert B. McGhee. CRADA leads RADM Jerry Ellis USN (Ret.) and Julie Leeman, Raytheon Technologies.

• Research support by Raytheon Missile Systems (RMS) and Raytheon BBN Technologies (Tucson Arizona and Arlington Virginia) has been instrumental in recent progress during 2019-2020.

• Significant contributions by Hsin-Fu “Sinker” Wu, Richard Markeloff and others at Raytheon, supporting current work under CRADA with NPS.

• Work with Dr. Jakub Flotyński on X3D Ontology for Semantic Web was an essential prerequisite to the capabilities demonstrated here.

• Collaborative design efforts with numerous skilled engineers and scientists is gratefully acknowledged. Further activity is welcome.
Contact

Don Brutzman

brutzman@nps.edu
http://faculty.nps.edu/brutzman

Code USW/Br, Naval Postgraduate School
Monterey California 93943-5000 USA
1.831.656.2149 work
1.831.402.4809 cell
ETHICAL MISSION DEFINITION AND EXECUTION FOR MARITIME ROBOTS UNDER HUMAN SUPERVISION

- Lethality requires ethical and legal basis, supervised by military teams.
- Executable robot tasking can resemble tactical tasking of humans afloat.
- Careful application of goal constraints makes ethical control feasible.
- Robot missions then complement and extend naval operation orders.
- Semantic Web logic can confirm ethical correctness and completeness.
- Next steps: continue 2 decades of work with realistic scenario testing.

“Ethical constraints on robot mission execution are possible today. There is no need to wait for future developments in Artificial Intelligence (AI). It is a moral imperative that ethical constraints in some form be introduced immediately into the software of all robots that are capable of inflicting unintended or deliberate harm to humans or property.”

Robert McGhee, April 2016

- IEEE Journal of Oceanic Engineering (JOE) paper along with online references.
- Authors Don Brutzman, Curtis Blais, Duane Davis and Robert McGhee, NPS.
- Feedback and recommendations always welcome. Contact: brutzman@nps.edu