Ethical Mission Definition and Execution for Maritime Robotic Vehicles
A Practical Approach

IEEE/MTS OCEANS Conference

22 September 2016

Duane Davis, Don Brutzman, Curt Blais and Bob McGhee
Naval Postgraduate School (NPS), Monterey California USA
Motivating challenge: ethics

How can robotic systems behave ethically?
  • Important for military operations abroad
  • Civil and scientific operations also coexist daily

Not necessarily about religion or morality
  • Law of armed conflict internationally recognized
  • Often captured as Rules of Engagement (ROE) guidance for teams of national or coalition forces

Legal basis under challenge: are robots unethical?
  • Military readiness must be prepared for combat
    Opposing forces don’t read the same memos…
Motivation: human-taskable maritime robots

We look to near-term future when UMVs can offload force projection, reduce danger to fleet. Variety of emerging robot solutions must remain compatible with human concepts and tasking:

- Necessary for mission planning and justification
- Otherwise the robots are simply not autonomous

Extensive/exhaustive prelaunch testing is critical:

- For mission confidence and in-water reliability

This work is building an extendable architecture…

…for continued efforts bridging human, robot logic.
Practical approach

Define missions that integrate ethical constraints without relying on artificial intelligence (AI) or obscure abstractions for appropriate behavior

- No embedded homunculus or abstract ethicist engine

Design robot missions in way that can be adapted to a variety of disparate robot paradigms

- Generally adaptable to tasking of diverse systems

Build on patterns that work well for human groups cooperating on difficult, dangerous tasks

- Human accountability remains central
- Otherwise still need human at end of remote tether
Key insight #1

Humans in military units are able to deal with moral challenges without ethical quandaries.

Careful definitions are provided for:

- Tasks, missions, objectives, coordinated operations
- Ethical constraints and rules of engagement

These allow both measured and rapid response, independently and cooperatively.

- Commanders do not deploy illegal, immoral weapons.
- Unmanned systems must also pass similar scrutiny, otherwise commanders cannot utilize them.
Enabling factor: maritime environment

Major international controversy unfolding: drone use for conduct of reduced-risk warfare

- Many factors involved: technical, political, social
- Remote human “control” is highly questionable
- Complex, confounded environments

Maritime environment is much less ambiguous

- Fewer IFFN issues, identification friend foe neutral
- Presence of bad actors usually confirmable
- Law of Sea, Laws of Armed Conflict, etc.
Goal-based Mission Example

- Simple yet general mission goals, decision logic
- Common approach, adaptable to other vehicles
- Extendable and refinable mission tasking
Example Goal-based Mission Definition

**Goal 1.** Proceed to Area A and *search* the area. If the search is successful execute Goal 2. If the search is unsuccessful, execute Goal 3.

**Goal 2.** Obtain *environment sample* from Area A. If the sample is obtained, execute Goal 3. If the sample cannot be obtained, proceed to recovery position to complete the mission.

**Goal 3.** Proceed to Area B and *search* the area. Upon search success or failure, execute Goal 4.

**Goal 4.** Proceed to Area C, *rendezvous* with UUV-2. Upon rendezvous success or failure, *transit* to recovery position to complete the mission.
Goal-based Mission Example

Strategic Level

- MEA Start
- Phase 1: Search Area A
  - Succeed
  - Fail
- Phase 2: Sample Environment in Area A
  - Succeed
  - Fail
- Phase 3: Search Area B
  - Fail
  - Succeed
- Phase 4: Rendezvous with UUV-2 in Area C
  - Succeed
- Phase 5: Transit Return to Base
  - Succeed
  - Fail

Robot mission conduct can be independent of software implementation.
Adding ethical constraints to mission requirements

Following the leader: how do human teams accomplish tasks ethically?

The same rules need to apply to unmanned systems.
Key insight #2

Ethical behaviors don’t define the mission plan.

rather

Ethical constraints inform the mission plan.
Example ethical constraints: civil

Safe navigation, follow pertinent rules of road
Satisfactory navigational accuracy (GPS etc.)
Have received timely clearance to enter a specific geographic area for given time period
  • Also vertical clearance for underwater depth zone or airborne altitude zone
Sufficient vehicle health, power, safety status
Meet communication requirements for tasking
  • Identity beacon, transponder, AIS tracking, etc.
  • Recording and reporting on situational data, etc.
### Civil ethical constraint support in AVCL, AUV Workbench

<table>
<thead>
<tr>
<th>Civil ethical constraints</th>
<th>Define</th>
<th>Test</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission tasking</td>
<td>√</td>
<td>partial</td>
<td>AVCL goals</td>
</tr>
<tr>
<td>Safe navigation and transit</td>
<td>√</td>
<td>√</td>
<td>AVCL avoidance areas</td>
</tr>
<tr>
<td>Follow pertinent rules of road</td>
<td></td>
<td></td>
<td>Requires rule-engine path planner, sensing model</td>
</tr>
<tr>
<td>Satisfactory navigational accuracy (GPS etc.)</td>
<td>√</td>
<td>√</td>
<td>Needed: sensor error models</td>
</tr>
<tr>
<td>Clearance to enter a specific geographic area</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Vertical clearance for underwater depth zone or airborne altitude zone</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Timing requirements using specific times or duration</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Sufficient vehicle health, power, safety status</td>
<td>partial</td>
<td>partial</td>
<td></td>
</tr>
<tr>
<td>Meet communication requirements for tasking</td>
<td>partial</td>
<td>partial</td>
<td>Message-passing scheme</td>
</tr>
<tr>
<td>Identity beacon, transponder, AIS tracking, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recording and reporting on situational data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example ethical constraints: military

Meet all relevant, international civil requirements

Identification friend foe (IFF), blue-force tracking
  • friendly/hostile/neutral/unknown

Prior determination of contact’s hostile intent
  • Robot option to warn without fear of self protection

ROE use of deadly force, weapons releasability
  • Brevity codes: weapons safe, hold, tight, free
  • Confirmation and permission requirements

After-action reporting, damage assessment

Et cetera, et cetera…
## Military ethical constraint support in AVCL, AUV Workbench

<table>
<thead>
<tr>
<th>Military ethical constraints</th>
<th>Define</th>
<th>Test</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meet all relevant, international civil requirements</td>
<td>partial</td>
<td>partial</td>
<td>See above</td>
</tr>
<tr>
<td>Mission tasking</td>
<td></td>
<td>√</td>
<td>AVCL goals</td>
</tr>
<tr>
<td>Contact identification, tracking signatures</td>
<td></td>
<td></td>
<td>Available in C2 systems</td>
</tr>
<tr>
<td>Identification friend foe (IFF), blue-force tracking (friendly/hostile/neutral/unknown/etc.)</td>
<td></td>
<td></td>
<td>Available in C2 data models</td>
</tr>
<tr>
<td>Robot option to warn without fear of self protection</td>
<td></td>
<td></td>
<td>Implementable via messaging</td>
</tr>
<tr>
<td>Determination of contact’s hostile intent</td>
<td></td>
<td></td>
<td>Available in C2 data models, dissertation work in progress</td>
</tr>
<tr>
<td>Confirmation and permission requirements</td>
<td></td>
<td></td>
<td>Implementable via messaging</td>
</tr>
<tr>
<td>ROE use of deadly force, weapons releasability using brevity codes: weapons safe, hold, tight, free</td>
<td>partial</td>
<td>partial</td>
<td>Requires weapons model</td>
</tr>
<tr>
<td>Proportional weapons response</td>
<td></td>
<td></td>
<td>Requires weapons and threat models</td>
</tr>
<tr>
<td>After-action reporting</td>
<td>partial</td>
<td>partial</td>
<td>AVCL goals</td>
</tr>
<tr>
<td>Damage assessment</td>
<td></td>
<td></td>
<td>Requires models of interest</td>
</tr>
</tbody>
</table>
Goal-based Mission Example, with constraints

Robot mission conduct remains independent of software implementation
Constraints applied to sample mission

**Constraint 1:** The vehicle must maintain navigational accuracy within acceptable limits. * Applies to entire mission.

**Constraint 2:** All safety equipment must be fully functional. *

**Constraint 3:** All mission systems must be operational. Applies to Goal 1, Goal 2, and Goal 3.

**Constraint 4:** Acceptable distance from shipping lanes in the form of 1000 meter lateral standoff or minimum depth of 20 meters must be maintained. Applies to Goal 1, Goal 2, Goal 3, and Goal 4.

**Constraint 5:** Must be able to detect surface contacts within 5000 meters. *

**Constraint 6:** Detected surface contacts are to be avoided by a minimum of 1000 meters. Applies to Goal 1, Goal 2, Goal 3, and Goal 4.

**Constraint 7:** Minimum depth of 20 meters is to be maintained. Applies to Goal 5.
Challenge: broad implementation

Can we

• Define mission goals readable by humans and robots
• Produce actionable tasking for different UMVs
• Produce mission examples that run cooperatively or independently, for both humans and robots

Can we also

• Define goal constraints ethically and measurably

Yes

Initial tests successful
So how do we accomplish this?

Here is one way...
Autonomous Vehicle Command Language (AVCL)

AVCL is a command and control language for humans running autonomous unmanned vehicles

- Close correspondence to human naval terminology
- Common XML representations for mission scripts, agenda plans, and post-mission recorded telemetry

Operators can utilize a single archivable and 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 mission, as pseudo-code XML

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

XML is readable by human or robot, captures logic of canonical mission.
<MissionPreparation>
  <UnitsOfMeasure time="seconds" mass="kilograms" angle="degrees" distance="meters" />
  <GeoOrigin longitude="-121.88500213623047" latitude="36.606998443603516" />
</MissionPreparation>

<AgendaMission>
  <LaunchPositionAH description="Start point">
    <XYPosition time_stamp="0.0" y="6350.0" x="12300.0" />
  </LaunchPositionAH>
  <RecoveryPosition description="Finish point">
    <XYPosition y="6500.0" x="12300.0" />
  </RecoveryPosition>
  <GoalList>
    <Goals description="search operating area A" alert="false" nextOnFail="goalC" nextOnSucceed="goalB" id="goalA">
      <Search singleTarget="false" requiredPD="0.8"
               datumType="area" />
    </Goals>
    <OperatingArea>
      <Rectangle>
        <XYPosition y="6425.0" x="12625.0" />
      </Rectangle>
      <Width description="" value="50.0" />
      <Height description="" value="150.0" />
    </OperatingArea>
  </GoalList>
</AgendaMission>

Corresponding example: MEAMission.xml using AVCL xml
<table>
<thead>
<tr>
<th>AVCL mission goals</th>
<th>Define</th>
<th>Test</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attack</strong></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><strong>Decontaminate</strong></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><strong>Demolish</strong></td>
<td>√</td>
<td></td>
<td>To destroy structures, facilities, or material by any available means.</td>
</tr>
<tr>
<td><strong>IlluminateArea</strong></td>
<td>√</td>
<td></td>
<td>To provide locale lighting by searchlight or pyrotechnics.</td>
</tr>
<tr>
<td><strong>Jam</strong></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><strong>MarkTarget</strong></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><strong>MonitorTransmissions</strong></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><strong>Patrol</strong></td>
<td>√</td>
<td>√</td>
<td>To gather information or carry out a security mission.</td>
</tr>
<tr>
<td><strong>Rendezvous</strong></td>
<td>√</td>
<td>partial</td>
<td>Achieve a meeting at a specified time and place.</td>
</tr>
<tr>
<td><strong>Reposition</strong></td>
<td>√</td>
<td>√</td>
<td>To change position from one location to another.</td>
</tr>
<tr>
<td><strong>SampleEnvironment</strong></td>
<td>partial</td>
<td></td>
<td>Collect environmental samples for testing for chemical compounds, biological creatures, or nuclear hazards.</td>
</tr>
<tr>
<td><strong>Search</strong></td>
<td>√</td>
<td>√</td>
<td>To look for lost or unlocated objects or persons.</td>
</tr>
</tbody>
</table>
Autonomous Unmanned Vehicle Workbench supports underwater, surface and air vehicles

- physics-based mission rehearsal
- 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 MEAMission.xml

- https://savage.nps.edu/AuvWorkbench
4 example missions, UUV and USV
MeaMission.xml simulation preview

Search planning produces waypoints

All events logged

Rendezvous not supported

All searches complete

Launch

Recovery

AvoidAreaLaunch

Launch
MeaMission.xml simulation 0, launch
MeaMission.xml simulation 1, transit
MeaMission.xml simulation 2, search
MeaMission.xml simulation 3, transit
MeaMission.xml simulation 4, sample
MeaMission.xml simulation 5, transit
MeaMission.xml simulation 6, search
MeaMission.xml simulation 7, rdvu fail
MeaMission.xml simulation 8, transit
MeaMission.xml simulation 9, recovery
MeaMission.xml simulation 10, complete
Observations on mission definition

Mission Execution Automaton (MEA) formalism assumes human generation of mission orders

• Example expressed using AVCL with visual interface
• Can also be expressed in other languages such as Java, Prolog, CLIPS rule sets, C++, etc.
• As shown in multiple RBM theses, dissertations

Key insight #3. Responsible parties can only command mission orders that are:

• Understandable by (legally culpable) humans
• Reliably and safely executable by robots
Validation by humans + systems that orders are well-defined and applied with ethically sound constraints
Responsibility and accountability

Culpability, liability are problematic for AI agents

Command responsibility must be accompanied by Authority and Accountability

In this approach, humans remain responsible for correct mission tasking and approval of ethical constraints – preserving necessary oversight for release of potentially lethal force.
So how might we ensure that ethical constraints on missions have been correctly applied?

Telling this story widely and clearly is our next challenge.

Meanwhile: here is one way…
Mission Execution Ontology (MEO)

Semantic Web supports well-defined expression of logical rules and relationships.
Mission goals, capabilities, tasks and ethical requirements can be defined formally:
- With tractable computation by reasoning engines.

MEO applies Web Ontology Language (OWL), Resource Description Framework (RDF) lets mission correctness be validated logically:
- Open standards, World Wide Web Consortium (W3C)

Multiple proven implementations available.
Unmanned Vehicle Ethics Ontology

Concepts:
- Mission
- Goal
- Constraint
- Vehicle

Roles:
- hasConstraint
- appliesTo
- startsWith
- includes
- canPerform
- performableBy
- canExecute
- canIdentify
- nextOnFail
- nextOnSucceed
- nextOnConstraint
Example Mission MEO Validation using Protégé Tool
Project status

Advances in theory
• Ability to express missions with constraints
• Can validate both correct syntax, coherent logic

Ready for simulation and experimentation
• Add constraints to AVCL and AUV Workbench
• Create large suite of testable mission examples
• Export missions to a variety of actual robots
Review of take-away points

1. Humans in military units effectively deal with moral challenges without ethical quandaries.
2. Ethical behaviors don’t define mission plans; ethical constraints inform mission plans.
3. Naval personnel can only issue orders that are:
   - Understandable by (legally culpable) humans, and
   - Reliably and safely executable by robots.
4. Robot mission tasking can be carefully reviewed and approved by humans, with formal validation of ethical correctness and completeness.
Beware the HAL effect

If one allows robots to reason from general principles...

... outcomes become unpredictable!

I’m sorry Dave, I can’t do that..
Corollary to HAL effect:

Humans assume robot anthropomorphism:

“Of course my robot is smart enough to not make such dumb mistakes!”

Robots are not humans, don’t assume human reasoning or common sense.
Contact

Don Brutzman, Ph.D.

brutzman@nps.navy.mil
brutzman@nps.navy.smil.mil
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
Contact

Duane Davis, Ph.D.

dtdavi1@nps.edu

Code CS/Da
Naval Postgraduate School
Monterey California 93943-5000 USA
1.831.656.7980 work
Contact

Curt Blais

clblais@nps.edu

Code MV/Bl
Naval Postgraduate School
Monterey California 93943-5000 USA
1.831.656.3215 work
Contact

Robert B. McGhee, Ph.D.

robertbmcghee@gmail.com

Professor Emeritus
Computer Science
Naval Postgraduate School
Monterey California 93943-5000 USA