U.S. ARMY COMBAT CAPABILITIES DEVELOPMENT COMMAND – ARMY RESEARCH LABORATORY

Robotics Collaborative Technology Alliance (RCTA) Program

Collaborative Alliance Manager: Stuart Young (ARL)
Consortium Manager: Dilip Patel (GDLS)
Enable the teaming of autonomous systems with soldiers

Provide fundamental science underpinnings of autonomous systems for the Army

Soldiers/Unmanned System Teaming:
- Combat multiplier
- Team member
- Heterogeneous groups
- Following commander’s intent

From micro-systems to combat vehicles
AUTONOMY TO ENABLE MULTI-DOMAIN OPERATIONS

To provide rapid and mobile ISR and mission support for the Dismounted Soldier beyond the eye of current national assets

Caves, Strategic Bunkers, Subterranean

Jungles and Under Canopies

Complex and Urban
ASSUMPTIONS DRIVING RESEARCH TOPICS

Military Relevant Environments:
- Highly Unstructured to Austere
- Limited to no a priori knowledge
- GPS denied and Low/no Light
- Complex RF environments
- Dynamic and Hostile

Technical Challenges:
- Robust operation across domains
  - Urban, subterranean, indoor/outdoor, etc
- Constantly changing environments
  - Varying lighting, dynamic objects
- High Operational Tempo - from cover, NLOS
- Soldier transportable systems
  - SWaP and processing constrained
  - Gust tolerant & Complex terrain
- Localization in Austere environments
- Long Duration GPS denied navigation
- Communications
- Human in the loop – heterogeneous teaming

Military Relevant Missions:
- Rapid and mobile Soldier ISR
  - Who, what, when and where
- Building search and clearance
  - People, CBRNE, etc
- Route clearance
  - Culvert/bridge search
- Perimeter surveillance
- Explosive Ordinance Disposal
THE RCTA VISION

Making unmanned systems an integral part of the small unit team

Systems that:
- Understand the environment
- Learn from experience
- Adapt to dynamic situations
- Possess a common world view
- Communicate naturally
- Conduct useful activity
- Can act independently, but within well prescribed bounds

Through research to enable and advance:
- Abstract reasoning
- Learning by example
- Reinforcement learning
- Semantic perception
- Communication through language
- Human behavior modeling
- Agile 3-D mobility at operational tempo
- Human-like manipulation

From tool to teammate
The unique hybrid architecture combining cognitive upper and geometric lower layers – bridging the gap between artificial intelligence & control

- Introduction of abstractions to facilitate human-robot communication – behavior specification based on structured language
- Algorithms to permit semantic labeling of objects, behaviors and their relationships
- A focus upon learning, to include learning without reliance on large sets of training data
- A broad description of the environment that goes beyond placing objects in four-dimensional space to include concepts such as object compliance and signature

**Coupled with enhancements to facilitate teaming**

- Natural modes of communication – voice and gesture
- Exploration of unconventional mobility modes
- Human-scale manipulation
For a simple case in a largely uncluttered environment:

• Parse natural (structured) language instructions into machine executable actions
• Interact naturally (voice/gesture) with teammates to disambiguate if required
• A semantic understanding of object classes, although computational requirements result in slow OPTEMPO
• Fuse data from multiple organic sensors and reasoning about mission-specific constraints
• Learn to manipulate arbitrarily shaped objects while considering environmental constraints
• Perform basic maneuver in three-dimensional and constrained environments

“Navigate to the back of the building that is behind the car.”

Experience Graphs transitioned (to “MoveIt”, open source SW platform for manipulation & mobility)
WHAT’S NEW

Thrust 1: Optempo Maneuvers in Unstructured Environments
- Capability 1: Optempo mobility in dynamic scenes
- Capability 2: Optempo mobility in rough terrain
- Capability 3: Optempo mobility in confined spaces

Thrust 2: Human-Robot Execution of Complex Missions
- Capability 1: SA in Unstructured Environments
- Capability 2: Distributed Mission Execution
- Capability 3: Verified and Trusted Execution

Thrust 3: Mobile Manipulation
- Capability 1: Manipulation in cluttered spaces
- Capability 2: Manipulation of unknown heavy objects

Scientific Challenges
Example: Thrust 2, Capability 1
- Use and implement computational models and algorithms to improve semantic perception by focusing on “cognitive perception” and dynamic scene understanding across length scales
- Use natural language through user-centered design communication interfaces that are user tested to support situational awareness

Common HW & SW Tools
STATE OF THE ART

Thrust 1 - Optempo Maneuvers in Unstructured Environments

- **Robustness**: Techniques are designed for favorable observation conditions, motion patterns, and data quality
- **Speed**: Focuses on learning models (e.g., using deep learning) offline from a large amount of annotated data
- **Environments**: Structured environments with limited dynamic interactions does not address scenarios with complex, unstructured geometry and arbitrary motions.

Thrust 2 - Human-Robot Execution of Complex Missions

- **SA in Unstructured Environment**: Robots can not *autonomously* explore unstructured environments and create *situational awareness*
- **Natural Language**: Lack of *natural language* to share information to achieve *collaborative mission execution with humans*.
- **Semantic Understanding**: Robots have inadequate semantic understanding of environments, constraints and objectives to carry out complex missions
- **Trust**: Robots have yet to become a *highly transparent system* that sets adequate levels of trustworthiness

Thrust 3 – Mobile Manipulation

- **Complex 3D Environments**: Precise position control of known objects within known environments, e.g. DARPA Robotics Challenge, Light payloads <5lbs, Requires fair amount of operator input, not robust to failures, Outdoor operation
  - Vision guided manipulation with fixed based table top robotic arms. E.g. DARPA ARM-S Challenge, Amazon Picking Challenge, Google, Restricted to indoor settings
Thrust 1 - Optempo Maneuvers in Unstructured Environments

• **Robustness**: Understanding dynamic environments, motions, and activities including (predicted) motions, behaviors, including occlusions and complex visibility. Anticipate motion and activities - Understand how to move relative to teammates, Perform perception functions robustly

• **Optempo Collaboration**: Communicate description of situation to human teammate, generating planning directives from human teammates and reporting observations to human teammates, and human-assisted planning through complex terrain configurations.

• **Complex Environments**: Gaits, control and planning in dynamic environments, terrain-dependent locomotion models and control schemes, Dynamic motion primitives for high-energy self-manipulation, planning for pushing objects when environment is obstructed

Thrust 2 - Human-Robot Execution of Complex Missions

• **Situation Awareness in Unstructured Environments**: Advances in models and algorithms for cognitive perception and dynamic scene understanding across length scales. Advances in natural language and gesture understanding.

• **Distributed Mission Execution**: Advances in structure prediction that allow robot to learn new model components and combine its own reasoning and previous experience from human teammates. Advances in inference to allow a unified computational model. New graphical model representations and algorithms for decentralized model inference of natural language. New computational models using recurrent neural networks (RNN) based approaches to synthesize descriptions to allow robots to explain the rationale behind their decision making/planned actions.

• **Trusted Execution of Verified Missions**: Advances in LTL inference for ensuring correctness of models and algorithms. A corpus of suitable trustworthiness/reliability estimates [i.e., (dis)trust signatures]. Advances in design contexts and design elements that are necessary to support properly calibrated trust in soldier-robot teams.

Thrust 3 – Mobile Manipulation

• **Complex 3D Environments**: Cluttered unknown environments with severe uncertainty, minimal and abstract operator input, an ability to handle heavier payloads and high reaction forces, and adaptive approaches that are robust to failures and learn from interaction.
THRUST 1: OPTEMPO MANEUVERS IN UNSTRUCTURED ENVIRONMENTS

Leads: Martial Hebert (CMU) and Ethan Stump (ARL)
Thrust 1 provides the foundation for a robot teammate operating at optempo with a small group of Soldiers.

To achieve this: we need a system that is: robust, optempo, and operates in unstructured environments.
Thrust Leads:
Brett Kennedy (JPL)
Harris Edge (ARL)
Martial Hebert (CMU)
Raghuveer Rao (ARL)

Vision:
Provides perception, intelligence, mobility, and interaction in support of fast motion in unstructured environments: Dynamic, rough, confined.

Differentiators:
- **Robust:** Uncontrolled observation conditions, rough motion of sensors
- **Optempo:** Decision making on the move, no offline processing
- **Unstructured:** Complex terrain geometry, unmapped environments/situations
Robustness:
SoA techniques are designed for favorable observation conditions, motion patterns, and data quality and cannot be used directly in the unstructured, uncontrolled world envisioned in our scenarios.

Speed:
SoA focuses on learning models (e.g., using deep learning) offline from a large amount of annotated samples and cannot be used in online, on the move scenarios with limited computation.

Environments:
SoA: structured environments with limited dynamic interactions does not address scenarios with complex, unstructured geometry and arbitrary motions.
SCIENTIFIC CHALLENGES

- Online operation and decision-making with limited computation
- Online planning and decision-making in dynamic environments
- Detailed understanding of unstructured dynamic environments
- Robust perception functions with respect to rough motions
- Fast mobility and interaction in complex, unstructured terrain
- Human-assisted planning
- Robust perception in confined, unstructured environments
- Online planning in confined, dynamic environments
- Interaction with teammates and other agents
**THRUST 1 – OPTEMPO MOBILITY**

**CAPABILITIES**

- Understand environment statically
- Recognize/predict motion and activities
- Perform perception functions robustly
- Plan based on dynamic representation

- Efficient motion in rough terrain
- Environment adaptive planning
- Human assisted planning
- Next-gen quadruped

- Robust perception in confined, unstructured environments
- Planning in confined, dynamic environments
- Interacting with teammates and other agents

**Integration**

- Integrated system for mobility in dynamic environments
- Configuration for confined spaces; planning + interaction SW
- Move through crowded environment
- Move through 3D terrain using unique mobility
- Interact with human for planning through difficult terrain
- Move through 3D indoor clutter (stairs, obstacles, ...)
- Natural interaction with teammates and other agents

**2019 Capstone Functionality**

- Move through indoor clutter (stairs, obstacles, ...)
- Natural interaction with teammates and other agents
THRUST 2: HUMAN-ROBOT EXECUTION OF COMPLEX MISSIONS

Leads: Nick Roy (MIT) and Ethan Stump (ARL)
Human-Robot Execution of Complex Missions

**Thrust Area Leads:**
- Nicholas Roy (MIT) & Florian Jentsch (UCF)
- Ethan Stump (ARL) & Bill Evans (ARL)

**Thrust Vision:**
Develop human-robot teaming that goes beyond teleoperation and waypoint navigation by incorporating mission intent, rich understanding of the world, and collaborative execution.

**Capability Focus:**
- Semantic perception of the environment
- Ability to reason about the environment in the context of mission objectives, based on a rich knowledge base of mission types
- Natural, multi-modal communications that provide the Soldier with critical information when it is needed, in the format that is most useful
- Collaborative problem solving between Soldier and robot(s)
- Trustworthy robot capabilities that are appropriately used by the Soldier
Human and group of robots collectively form plan and share information as needed during mission.

Human supervisor monitors robots throughout mission.

Mission is specified as a goal configuration.

State of the art: Supervisory control.

Research Goal: Human-robot Execution of Complex Missions.

Complexity of mission:
- Mission is specified as high-level objective, over long length- and time-scales.

Complexity of teaming:
- Human and group of robots collectively form plan and share information as needed during mission.
Robots do not have the capability to autonomously explore unstructured environments and create support situational awareness.

Robots are yet to use natural language to share information to achieve collaborative mission execution with a human teammate:

- Update the teammate through the appropriate communication modality by adapting to environmental demands or mission context.

Robots have yet to use semantic understanding of environments, constraints and objectives to carry out complex missions:

- Surveil a location.

Robots have yet to become a highly transparent system that sets adequate levels of trustworthiness.
THRUST 2: SCIENTIFIC CHALLENGES

• Capability 1 Situation Awareness in Unstructured Environments
  – Requires advances in models and algorithms that allow focused “cognitive perception” and dynamic scene understanding across length scales
  – Requires advances in natural language and gesture understanding, integrated with user-centered design communication interfaces that are user tested to support situational awareness

• Capability 2 Distributed Mission Execution
  – Requires advances in structure prediction that allow robot to learn new model components and combine its own reasoning and previous experience from human teammates
  – Requires advances in inference to allow a unified computational model that supports the team reasoning jointly about imperative commands and world knowledge.
  – Requires new graphical model representations and algorithms for decentralized model inference of natural language instructions
  – Requires new computational models using recurrent neural networks (RNN) based approaches to synthesize descriptions to allow robots to explain the rationale behind their decision making and describe their planned actions

• Capability 3 Trusted Execution of Verified Missions
  – Requires advances in LTL inference for ensuring correctness of models and algorithms
  – Requires a corpus of suitable trustworthiness/reliability estimates [i.e., (dis)trust signatures].
  – Requires advances in design contexts and design elements that are necessary to support properly calibrated trust in soldier-robot teams under varying degrees of task demand and imposed cognitive workload.
THRUST 2: THREE CAPABILITIES

Thrust 2: Human-Robot Execution of Complex Missions
(co-leads: Roy & Jentsch [consortium]; Evans & Stump [ARL])

1. Situation awareness in unstructured environments (J. Oh, CMU)
   - “Robot, go tell me what’s out there”

2. Distributed mission execution (E. Stump, ARL; D. Barber, UCF)
   - “Robot, let’s go together and secure that street corner.”

3. Trusted execution of verified missions (B. Evans, ARL)
   - “Robot, I think you can do that. Can you?”
THRU? 2 – HUMAN-ROBOT EXECUTION OF COMPLEX MISSIONS

CAPABILITIES

Semantic scene understanding in unstructured environments
Dynamic scene understanding in unstructured environments
Synthesizing semantic content into human-understandable descriptions
Design of system for effective teaming
Collaborative mission planning and execution
Distributed situational awareness and communication
Share multi-modal dialogue
Verifiable planning
Maintaining correctness and trust through communication
System design to maintain trust

Output Products: Perception algorithms, cognitive architectures, multi-modal comm. systems, criteria and procedures for training, V&V, and trust assessment

Integration

2019 Capstone Functionality

Semantic perception
Reasoning about the environment
Natural, multi-modal comms
Collaborative problem solving
Trustworthy robot capabilities
THRUST 3: MOBILE MANIPULATION

Leads: Siddhartha S. Srinivasa (UW), Kostas Daniilidis (U Penn), and Chad Kessens (ARL)
INTRODUCTION TO THRUST AREA MANIPULATION

Thrust Area Leads:
- Kostas Daniilidis, UPenn
- Siddhartha S. Srinivasa, UW
- Chad Kessens, ARL

Thrust Vision:
Provide an autonomous unmanned ground team member with the ability to manipulate objects in human habitable spaces with geometric, spatial and inertial uncertainty.

Capability Focus:
- **Place objects** such as sensors in **confined spaces**
- **Conduct ingress**: door with curtain, sliding door, with different sorts of handles.
- **Search and retrieve specific items**, and open storage elements or containers to search them.
- **Manipulate and maneuver heavy objects** (>20lbs).
Current State of Art:

1. Precise position control for known objects within known environments
   - E.g. DARPA Robotics Challenge.
   - Light payloads < 5 lbs.
   - Requires fair amount of operator input.
   - Not robust to failures.
   - Outdoor operation

2. Vision guided manipulation with fixed base table top robotic arms.
   - E.g. DARPA ARM-S Challenge, Amazon Picking Challenge, Google
   - Light payloads < 5 lbs.
   - Restricted to indoor settings

Scientific Challenges:
- Towards cluttered unknown environments with severe uncertainty.
- Towards minimal and abstract operator input
- Towards an ability to handle heavier payloads and high reaction forces
- Towards adaptive approaches that are robust to failures and learn from interaction.
SCIENTIFIC CHALLENGES

Problem Space & Focus

Known Geometrical Properties

Known Inertial Properties

Capability 1

m ≤ 5 lbs

geometric and spatial uncertainty

Shape ?

Capability 2

m ≥ 20 lbs

inertial uncertainty

Shape ?

Unknown Geometrical Properties

Unknown Inertial Properties

CY 16

CY 17-18

CY 19

KNOWN OBJECTS

Problem Space & Focus

UNKNOWN OBJECTS
SCIENTIFIC CHALLENGES

Problem Space & Focus

Known Geometrical Properties

Known Inertial Properties

Current State of Art In Mobile Manipulation
e.g. DARPA Robotics Challenge

Unknown Geometrical Properties

Unknown Inertial Properties

Scientific Goals

Capability 1

m <= 5lbs

geometric and spatial uncertainty

Shape ?

Capability 2

m ?

inertial uncertainty

>= 20 lbs

Known Geometrical Properties

Unknown Geometrical Properties

Scientific Goals

UNCLASSIFIED
INTRODUCTION TO CAPABILITIES

Manipulation (Kostas/Larry)

1. Capability - Manipulation in cluttered human habitable spaces (Max Likhachev)
   1. Where are objects?
   2. How to grasp?
   3. How to manipulate?

2. Capability - Manipulation of unknown heavy objects (Sisir Karumanchi)
   1. How to perceive and maneuver heavy objects?
   2. How much force can be applied to or expected from the environment?
   3. How to enable physical interactions via force feedback?
THRUST 3 – MOBILE MANIPULATION

CAPABILITIES

Geometry learning of rigid and deformable objects
Planning with Uncertainty in the Model & Perception
Learning Symbolic Planning Models for High-DOF Systems
Generalizing grasps across objects and tasks
Techniques for fast human-like manipulation

Integration

2019 Capstone Functionality

Roman Platform Full-body Manipulation and World Interaction Behaviors

Search & retrieve Specific items
Maneuver ingress openings

Roman Platform Force Feedback Behaviors

Maneuver heavy objects (>20lbs)
Lift stretcher with human support

Next-gen mobile manipulation behaviors
Mobility aided manipulation & novel actuators
Learning novel force-based mobile manipulation
Learning Mechanical and Geometric Models of Unknown Objects Online
Human-Robot Co-Manipulation

Manipulation in cluttered human habitable spaces
Manipulation of unknown heavy objects

Learning in cluttered human habitable spaces
Learning Symbolic Planning Models for High-DOF Systems
ARL Sciences of Maneuver

- Focus on soldier-robot teaming
- Enable maneuver in unstructured, dynamic environments
- Fundamental technology applicable to air, ground, and maritime domains

DBS Summer Study

"Sustain research focus on systemic challenges" – Focus on “gap technology” not explored by the commercial sector

Robotics and Autonomous System (RAS) Strategy

Far-Term Priorities:
- “Improve situational Awareness…” – through rapid learning, inclusion of contextual information
- “Improve maneuver with advancements to unmanned combat vehicles” – maneuver in 3-D environments

3rd Offset - Leveraging US Ally’s Investment
# RCTA PLATFORMS

**Common Perception Sensor Suit, Computers, and Power Supply**
Velodyne/Hokoyu Ladar, Adonis and Realsense cameras, 4 quad i7 computers, and standard batteries.

<table>
<thead>
<tr>
<th></th>
<th>Husky</th>
<th>RoMan</th>
<th>RCTA Robosimian</th>
<th>LLAMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCTA Qty</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Speed</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>*****</td>
</tr>
<tr>
<td>Rough Terrain</td>
<td>**</td>
<td>***</td>
<td>*****</td>
<td>*****</td>
</tr>
<tr>
<td>Dexterity</td>
<td>*</td>
<td>****</td>
<td>*****</td>
<td>**</td>
</tr>
</tbody>
</table>

LLAMA (Legged Locomotion & Mobility Adaptation)
ARL Robotics Collaborative Technology Alliance:

• A cooperative agreement between ARL & 8 Partners+
• Conducts fundamental research to create autonomous robots that can effectively team with Soldiers in small units
• It conducts research in four technical domains: Perception - Intelligence - Human-robot interaction - Manipulation & Mobility
• Three multi-disciplinary thrusts to achieve greater levels of autonomy
  • Optempo Maneuvers in Unstructured Environments
  • Human-Robot Execution of Complex Missions
  • Mobile Manipulation
• Alliance conducts periodic technology assessments to assess progress and facilitate transition
• Further information can be found at http://www.arl.army.mil/www/default.cfm?page=392