SAND2020-7355C

Artificial Intelligence and Autonomy in Space: *Balancing Risks of Unintended Escalation*

July 22, 2020

Nancy K. Hayden

Jason Reinhardt, Mallory Stewart, J.D. Doak, Thushara Gunda, Kelsey Abel

Sandia National Laboratories*

0:00-0:30

The Changing Space Domain: increasingly contested, congested, and competitive with proliferation of new actors and technologies. New space assets face both natural and man-made threats. AI and Autonomous systems increasingly deployed as critical enablers in space systems.

- AI functions: planning and scheduling, navigation, and classification.

- AI-informed deterrence scenarios rely on communication, capability, and credibility viz-a-viz these functions for clear signaling to avoid escalation.

- Issues that have arisen in terrestrial applications of AI may also be present in space.
    - Explainability
    - Vulnerability
    - Performance

What trade-offs between explainability, performance, and vulnerability in AI in space systems will best support deterrence, and which ones may increase the risk of escalation?

Approach:

- Discrete event modeling
- Embedded ABM (informed by game theory) within System Dynamic model
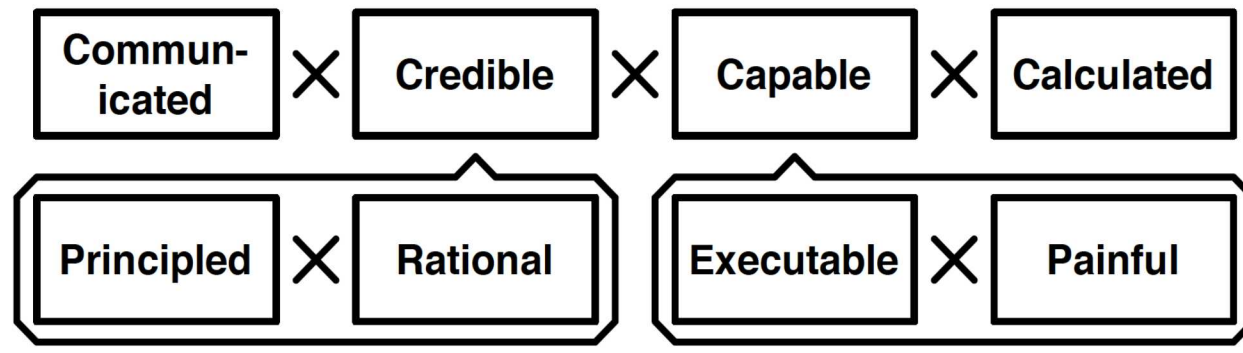
Dynamic Hypothesis:

- Deterrence stability landscapes in contested and congested space will depend on appropriate trade-offs in explainability, performance, and vulnerability.
- Robust solution space requires (at least) two-player perspective

Use Cases:

- Space control (classification of abnormal behaviors)
- Space domain awareness (planning and scheduling; navigation for and traffic management)

A Framework for Deterrence Effectiveness

Communicated × Credible × Capable × Calculated

Principled × Rational   Executable × Painful

Bonin, Ben and Jason Reinhardt, "Deterrence & Nuclear Strategy: An Imperfect Overview" (2019). Sandia National Laboratories, Albuquerque, NM. SAND2019-9183 PE.

2:00-3:30

3:30-5:00