Verifying Autonomy

Leon Kester
Some Applications of Autonomy

› Mobility
  › Cooperative autonomous vehicles.

› Military Vehicles
  › Underwater mine hunting
  › Swarms for surveillance
  › (Semi-) Autonomous ship cooperating in a task group
  › Meaningful human control for LAWS
Developments in Autonomy: Automation

- Algorithms for Situation assessment
- Algorithms for Situation management

- Developments: dealing with more complex situations and better cooperation between autonomous systems

- Advantages: fast, transparent

- Disadvantage: not dealing well with failures and unexpected situations
Developments in Autonomy: Learning

- Online learning (learning on the job)
- Offline learning (training)

- Developments: faster learning (faster computers, efficient algorithms)
- Advantages: very flexible tool
- Disadvantages:
  - still very slow for complex systems in complex situations
  - difficult to understand how it works and why it works
  - inherently unpredictable and more easy to deceive (online)
  - not dealing well with failures and unexpected situations (offline)
Consequences of this view for verification and role human

▶ Difficult to derive sub-system performance measures

▶ System has to be verified in many, many different situations

▶ Human as fail-safety agent
  ▶ When the autonomous system is failing the human intervenes.

▶ Human as a moral agent
  ▶ The human makes moral decisions and intervenes when autonomous system does not act in a morally acceptable way.
Developments in Autonomy: self-determination

- It involves self-assessment
  - What are my resources: sensors, computation, communication, actuators?
  - How well am I doing?

- And self-management
  - Now that I know and understand myself and my environment how should I manage myself in order to reach my goals?

- Advantages
  - Self-assessment and management can be used for sub-system verification
  - Faster than learning and can deal better with failures/unexpected situations
  - Can better explain to the human what it is doing and why it is doing it