Verification of Autonomous Systems

Michael Fisher

University of Liverpool, United Kingdom

ICRA, Stockholm, May 2016

Background

Professor of Computer Science

- formal methods, autonomy, proof, programming languages

  `http://www.csc.liv.ac.uk/~michael`

Director of Centre for Autonomous Systems Technology

- cross-displinary centre at University of Liverpool
- involving CS, Engineering, Electronics, Law, Pyschology,...

  `http://www.liv.ac.uk/cast`

Coordinator of UK Network on the Verification and Validation of Autonomous Systems

- funded by EPSRC
- brings together formal verification, testing, user validation, etc

  `http://www.vavas.org`

Interested in: Autonomous Systems

Autonomy:

> *the ability of a system to make its own decisions and to act on its own, and to do both without direct human intervention.*

Even within this, there are variations concerning decision-making:

Automatic: involves a number of fixed, and prescribed, activities; there may be options, but these are generally fixed in advance.

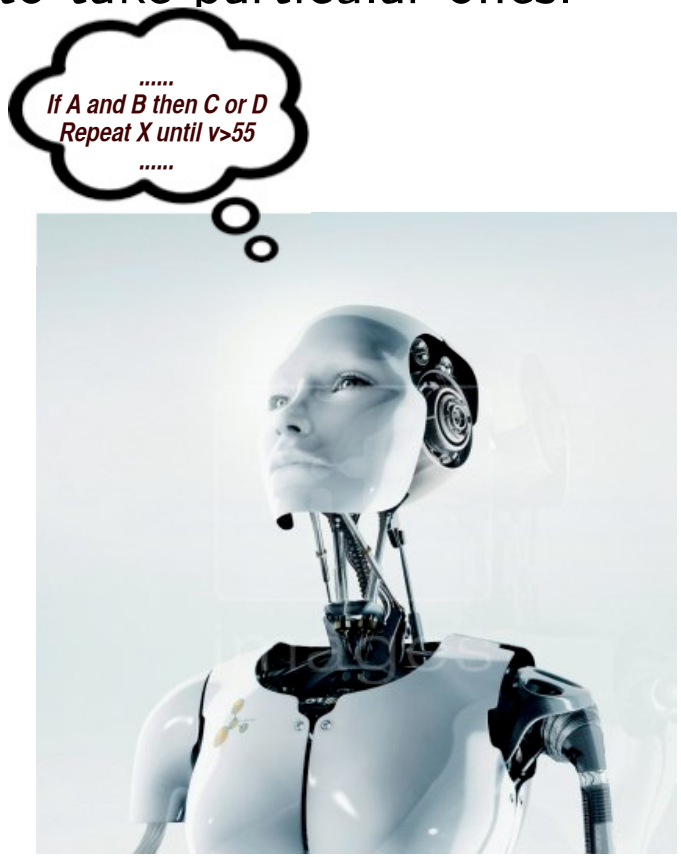
Adaptive: improves its performance/activity based on feedback from environment — typically developed using tight continuous control and optimisation, e.g. feedback control system.

Autonomous: decisions made based on system's (belief about its) current situation at the time of the decision — environment still taken into account, but internal motivations/beliefs are important.

No Psychiatrists for Robots?

With an *autonomous system* we can (at least in principle) examine its internal programming and find out exactly

1. *what* it is "thinking",
2. what *choices* it has, and
3. why it *decides* to take particular ones.



Verifiable Autonomy

Our approach is that

> *we should be certain what the autonomous system* **intends** *to do and how it* **chooses** *to go about this*

Consequently: we should know who (or what) is **responsible**

The *agent* concept captures the core of autonomy, in that it is *able to make its own decisions without human intervention*.

But: this still isn't enough, as we need to know *why*!

A "*rational agent*":

> *must have explicit* **reasons** *for making the choices it does, and should be able to explain these if needed*

Example: from Pilot to Rational Agent

Autopilot can essentially fly an aircraft

- keeping on a particular path,
- keeping flight level/steady under environmental conditions,
- planning routes around obstacles, etc.

Human pilot makes high-level decisions, such as

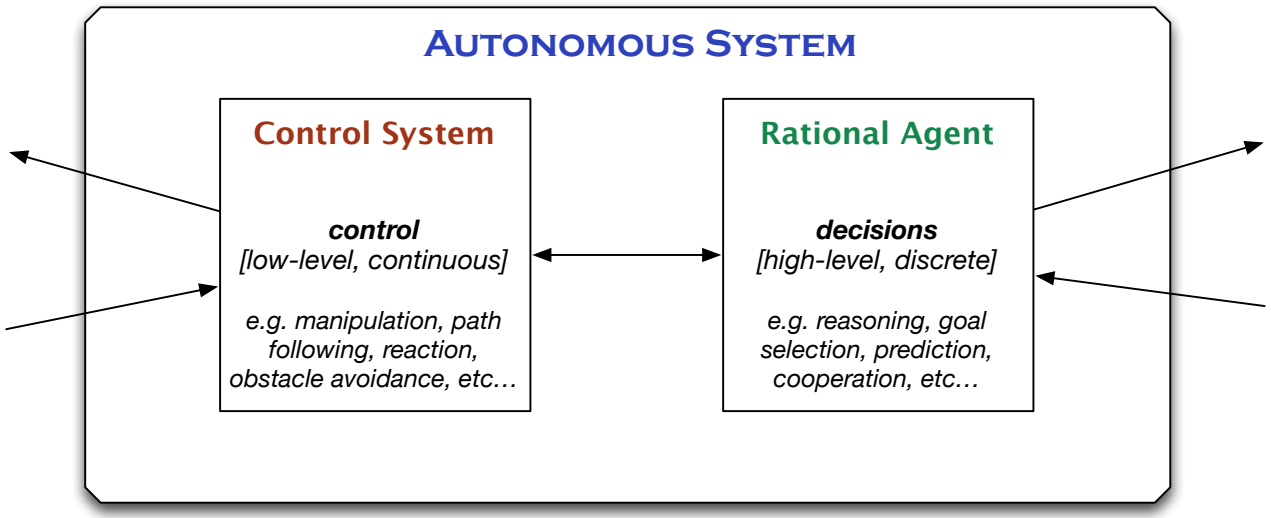
- where to go to,
- when to change route,
- what to do in an emergency, etc.

Rational Agent now makes the decisions the pilot used to make.

Verification of Autonomous Systems

We verify the rational agent within the system's architecture.

Importantly, this allows us to verify the *decisions* the system makes, not its outcomes.



In summary:

> *We cannot prove what the system will achieve, since interactions with the real world are always uncertain, but we can prove what (and why) it will try to achieve.*

Summary

Verification examples:

- UAV certification
- domestic robotic assistants
- autonomous vehicle platooning
- formation-flying satellites
- human-robot teamwork
- robot ethics

Opportunities for research:

- Requirements — what should our system do/decide?
- Logics — do we have logics to specify/verify requirements?
- Architectures — what about more opaque architectures?
- HRI — representing/verifying teamwork, trust, etc?

Sample Relevant Publications

- Dennis, Fisher, Aitken, Veres, Gao, Shaukat, Burroughes. Reconfigurable Autonomy. *Kunstliche Intelligenz 28(3):199-207*, 2014.
- Dennis, Fisher, Slavkovik, Webster. Formal Verification of Ethical Choices in Autonomous Systems. *Robotics and Autonomous Systems 77:1-14*, 2016.
- Dennis, Fisher, Webster. Verifying Autonomous Systems. *Communications of the ACM 56(9):84–93*, 2013
- Dennis, Fisher, Lincoln, Lisitsa, Veres. Practical Verification of Decision-Making in Agent-Based Autonomous Systems. *Journal of Automated Software Engineering 23(3):305-359*, 2016.
- Webster, Cameron, Fisher, Jump. Generating Certification Evidence for Autonomous Unmanned Aircraft Using Model Checking and Simulation. *Journal of Aerospace Information Systems 11(5):258–279*, 2014.
- Webster, Dixon, Fisher, Salem, Saunders, Koay, Dautenhahn, Saez-Pons. Toward Reliable Autonomous Robotic Assistants Through Formal Verification: A Case Study. *IEEE Trans. Human-Machine Systems 46(2):186-196*, 2016.