ELSI, Autonomy and Civil Aeronautics

Wendell Wallach
Yale Interdisciplinary Center for Bioethics
Irving, CA – August 28, 2013

- ELSI – Ethics, Law, and Societal Issues
- Privacy, Surveillance and Property Rights
- Autonomy
- Engineering for Responsibility
- Moral Machines
- Governance Coordination

ELSI

- Attention to ELSI Lags Far Behind Technological Development.

- ELSI Research Must be Funded Along With the Technologies.

BRANCO ConservativeDailyNews.com ©2012
NYPD
DRONE!!
MEGA POP
Cola
WELCOME TO NANNY STATE

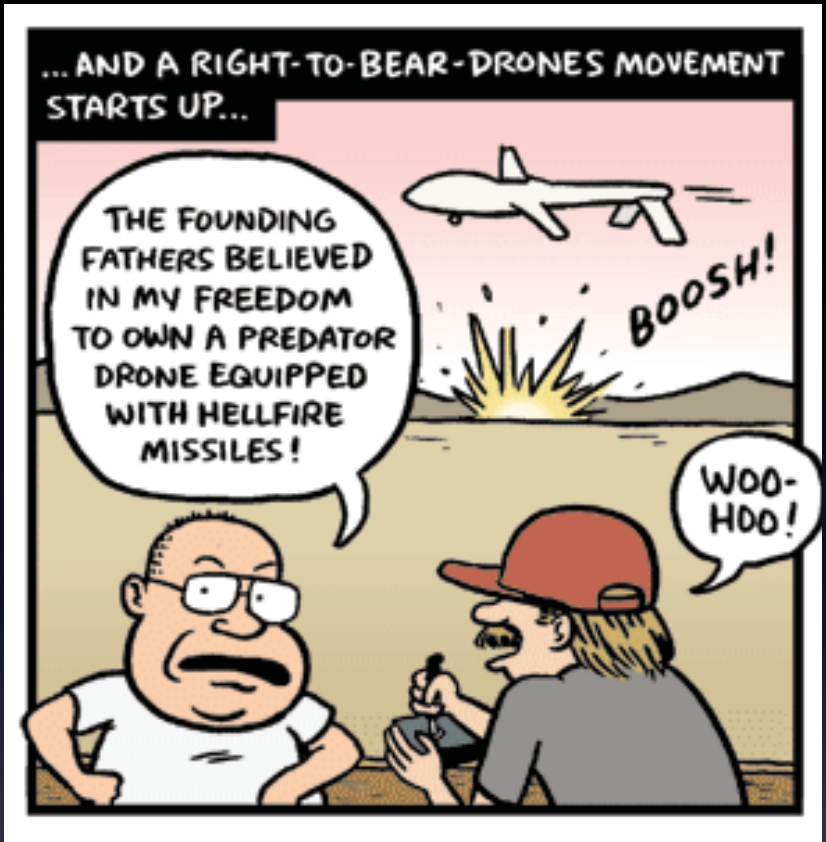
NULLIFY DRONES

Texas House Bill 83 proposes criminal charges for using drones to capture images without a warrant or probable cause of a felony, or imminent danger.

DOMESTIC USE OF DRONES? BAD IDEA

Privacy and Surveillance

Property Rights


... AND A RIGHT-TO-BEAR-DRONES MOVEMENT STARTS UP...
THE FOUNDING FATHERS BELIEVED IN MY FREEDOM TO OWN A PREDATOR DRONE EQUIPPED WITH HELLFIRE MISSILES!
BOOSH!
WOO-HOO!

CIVILIAN ISSUE • RESIDENT • PERMANENT
DRONE HUNTING PERMIT
PERMIT NUMBER
1-234-D34-672-1142-DR78
All Methods Allowed
NO BAG LIMIT
FLYING VERMIN
ALL TERRITORIES

Autonomy

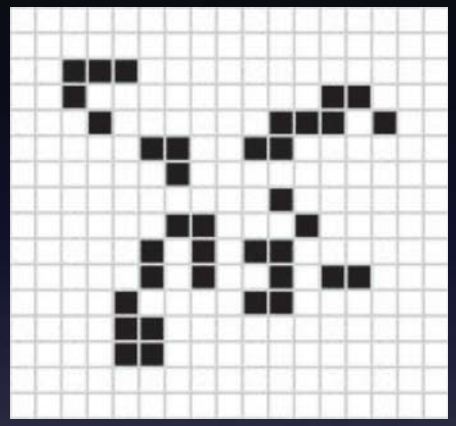
- Safety
- Predictability
- Resilience if something goes wrong
- Attribution of responsibility, liability, and culpability
- Decision making – choices and actions
  - Legal, ethical, and appropriate,

Increasingly Autonomous Robotic Systems

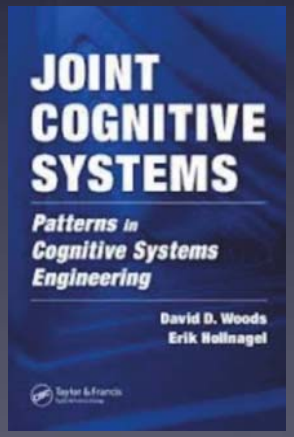
threatens to undermine the foundational principle that a human agent (either individual or corporate) is responsible, and potentially accountable and liable, for the harms caused by the deployment of any technology.

Responsibility

- Difficulty in predicting actions of algorithmic and complex systems.



  - ‘many hands’ (Nissenbaum, 1997)



  - Joint Cognitive Systems/Coordination
    - Woods and Hollnagel - Global Hawk UAV – December 6th, 1999 – $5.3M

”we have constructed a world in which the potential for high-tech catastrophe is embedded in the fabric of day-to-day life.”

Malcolm Gladwell

Engineer for Responsibility

- Design of Artifacts
- Social Engineering

Rules for Computing Artifacts

- Rule 1: The people who design, develop, or deploy a computing artifact are morally responsible for that artifact, and for the foreseeable effects of that artifact. This responsibility is shared with other people who design, develop, deploy or knowingly use the artifact as part of a sociotechnical system.

- Rule 2: The shared responsibility of computing artifacts is not a zero-sum game. The responsibility of an individual is not reduced simply because more people become involved in designing, developing, deploying or using the artifact. Instead, a person's responsibility includes being answerable for the behaviours of the artifact and for the artifact's effects after deployment, to the degree to which these effects are reasonably foreseeable by that person.

- Rule 3: People who knowingly use a particular computing artifact are morally responsible for that use.

- Rule 4: People who knowingly design, develop, deploy, or use a computing artifact can do so responsibly only when they make a reasonable effort to take into account the sociotechnical systems in which the artifact is embedded.

- Rule 5: People who design, develop, deploy, promote, or evaluate a computing artifact should not explicitly or implicitly deceive users about the artifact or its foreseeable effects, or about the sociotechnical systems in which the artifact is embedded.

Machines must not make ‘decisions’ that result in the death of humans.

- Ban on Killer Robots
  - ICRAC (Berlin, Oct ’10)
  - Call for an Executive Order (Feb ‘12)
    - ALFIS violates LOAC & IHL
    - malum in se
  - HRW/Harvard Law School Human Rights Clincic (Nov. 19th 2012)
  - Autonomy in Weapons Systems (Nov. 23rd, 2012)
    - DoD Undersecretary Ashton Carter

Moral Agency - AMA

- If and when robots become ethical actors that can be held responsible for their actions, we can begin debating whether they are no longer machines and are deserving of some form of personhood.
- ‘mala in se’
  - Not because they are machines
  - Unpredictable, can not be fully controlled, and attribution of responsibility is difficult if not impossible

Once such a principle is in place we can go on to the more exacting discussion as to the situations in which robots and information systems are indeed an extension of human will and intention and when their actions are beyond direct human control.

Mapping a New Field of Inquiry

Professor Colin Allen
Indiana University

Machine Morality
Machine Ethics
Computational Ethics
Artificial Morality
Friendly AI
Roboethics

Artificial Moral Agents (AMAs)

- Implementation of moral decision-making faculties in artificial agents

- Necessitated by autonomous systems making choices

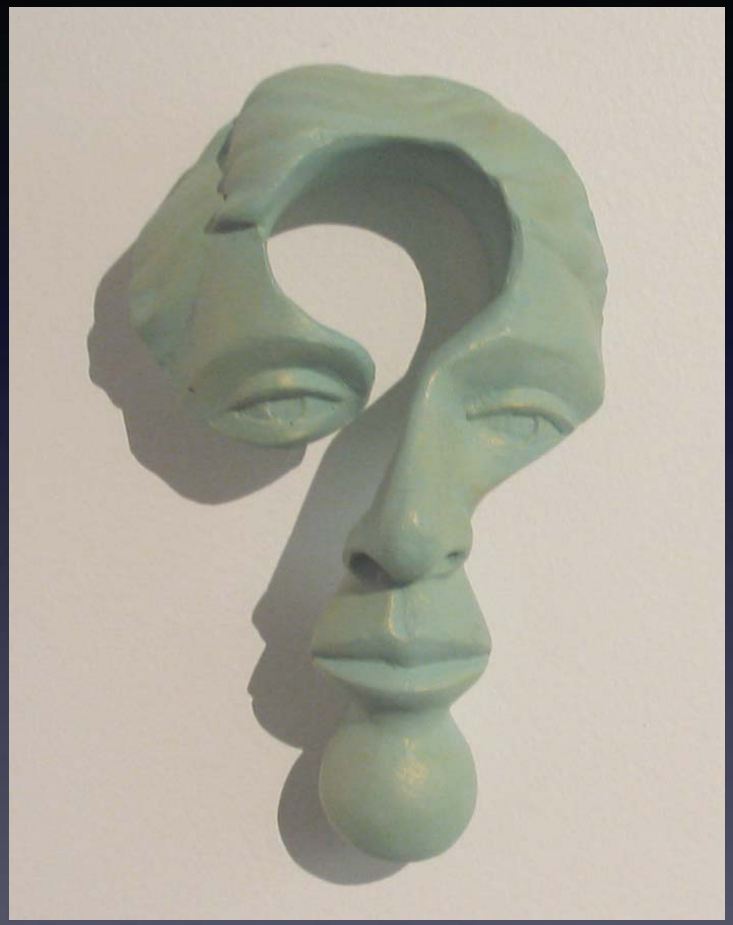
"The greater the freedom of a machine, the more it will need moral standards."

Rosalind Picard, Director

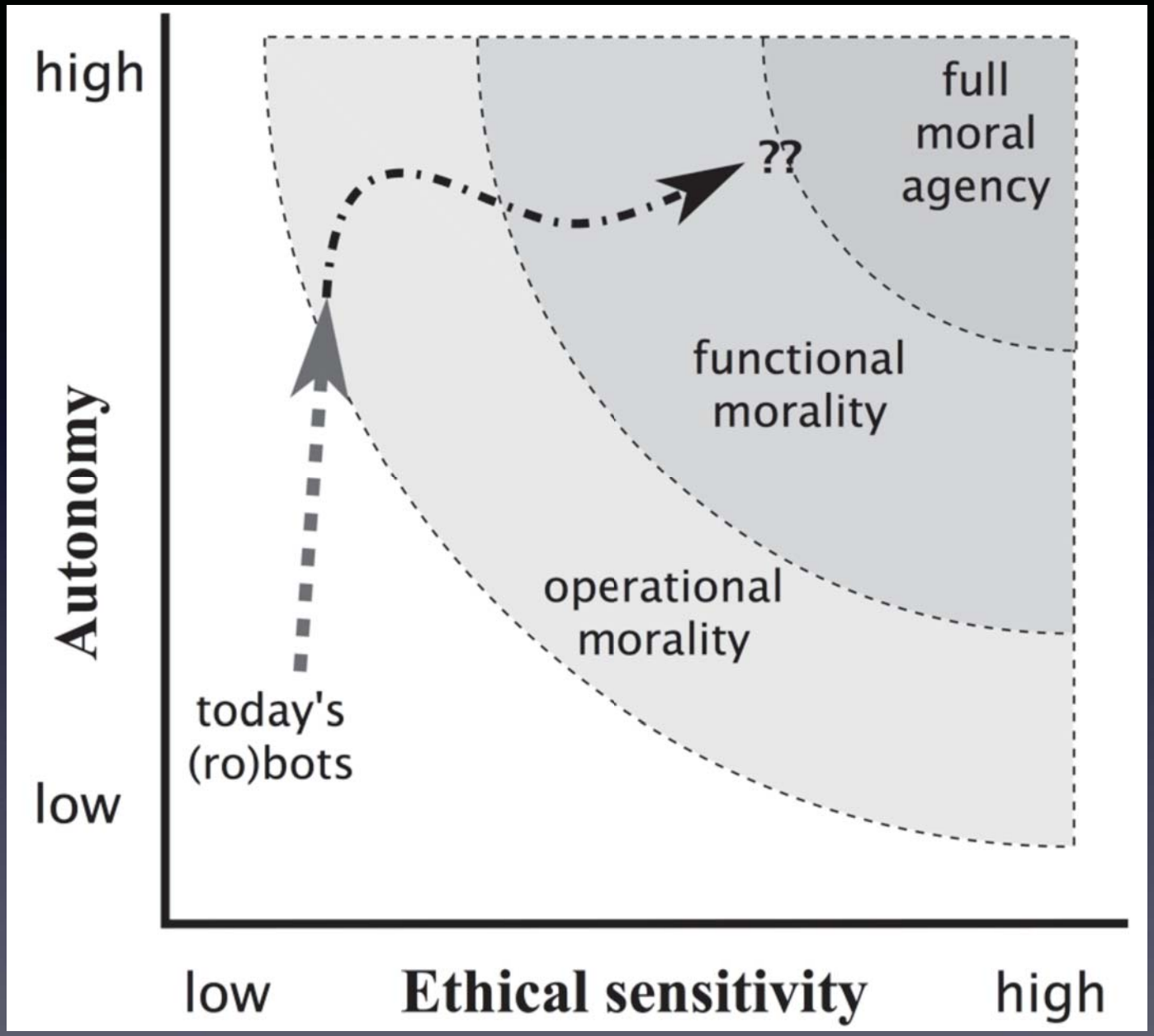
M.I.T. Affective Computing Group

- If robots can be designed so that they are sensitive to ethical considerations and factor those considerations into their choices and actions, new markets for their adoption will be opened up.
- If they fail to adequately accommodate human laws and values, there will be demands for regulations that limit their use.

Questions

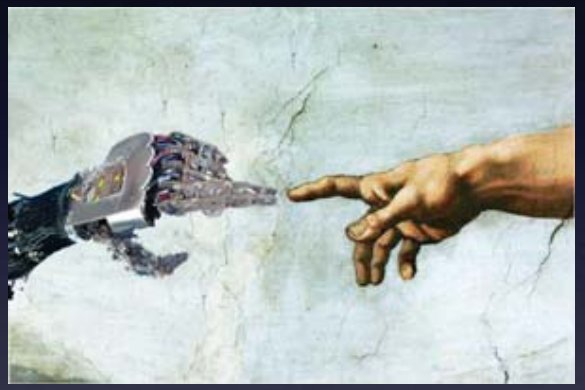


- Do we need artificial moral agents (AMAs)?
    - When? For what?
- Do we want computers making ethical decisions?
- Whose morality or what morality?
- How can we make ethics computable?


high
Autonomy
low
full
moral
agency
??
functional
morality
operational
morality
today's
(ro)bots
low
Ethical sensitivity
high

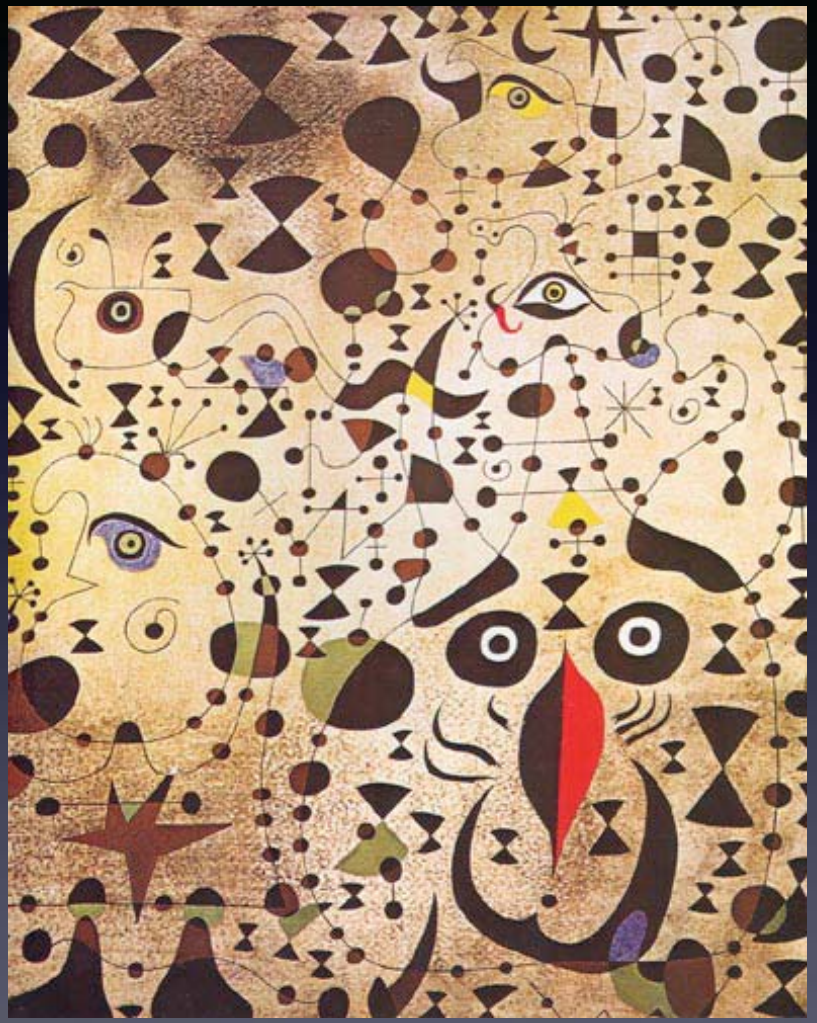
Approaches

- Role of ethical theory in defining the control architecture.
  - Top-Down
  - Bottom-Up
  - Hybrid Approaches
- Supra-rational Faculties



The Two Hard Problems

- Implementing norms, rules, principles, or procedures for making moral judgments
- Framing Problems



Framing Problems

How does the system recognize it is in an ethically significant situation?

How does it discern essential from inessential information?

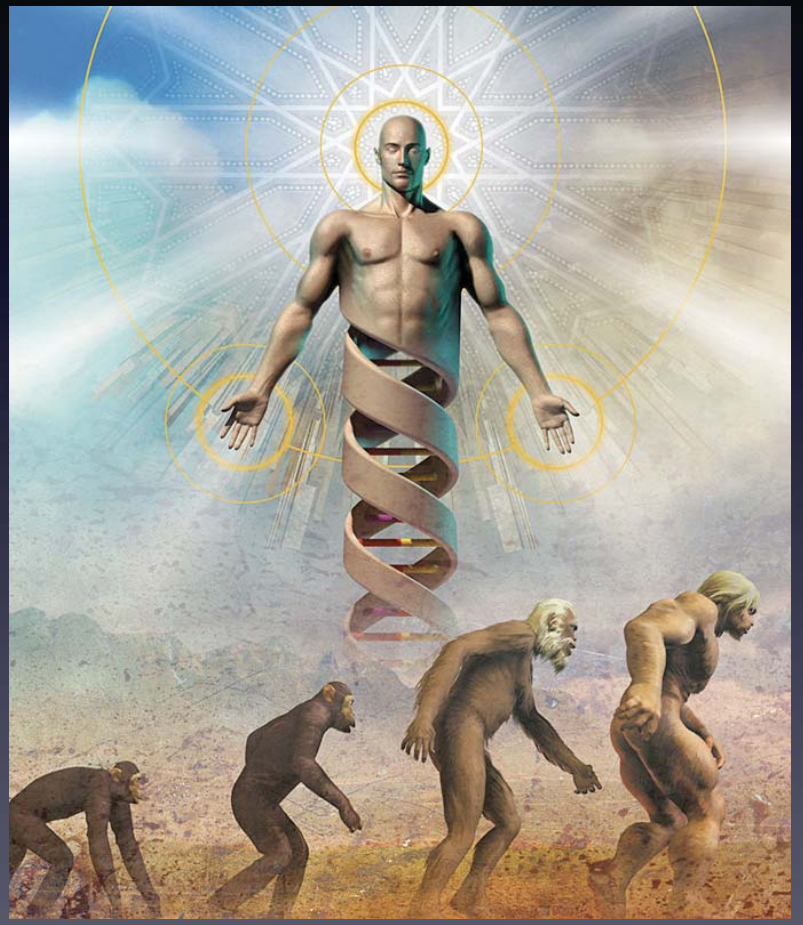
How does the AMA estimate the sufficiency of initial information?

What capabilities would an AMA require to make a valid judgment about a complex situation? e.g., combatant v. non-combatant.

How would the system recognize that it had applied all necessary considerations to the challenge at hand or completed its determination of the appropriate action to take?

Distinction

- Humans -- Biochemical,
  - Instinctual, Emotion Platform
  - Higher Order Faculties Emerged
- Computers -- Logical Platform



Possible Advantages

- Calculated Morality
  - Rationality Less Bounded
- Absence of base motivations
  - greed -- dubious for Alife
- Absence of emotions
  - emotional “highjackings”
    - (sexual jealousy)

Role of Emotions in Moral Decision-Making Machines

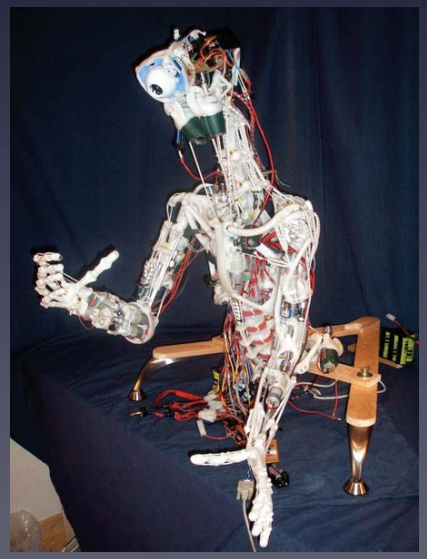
- Stoicism vs. Moral Intelligence
  - Both beneficial and dysfunctional
  - Will AMA's Require some affective intelligence
    - Emotions of their own? Cognitive? Somatic?
  - Emotional heuristics

- **Will humans feel comfortable with machines sensitive to their emotional states?**

Beyond Reason

- Emotions
- Sociability
- Embodiment
- Theory of Mind
- Empathy
- Consciousness
- Understanding

Genghis



Chronos2

- Will artificial agents need to emulate the full array of human faculties to function as adequate moral agents and to instill trust in their actions?

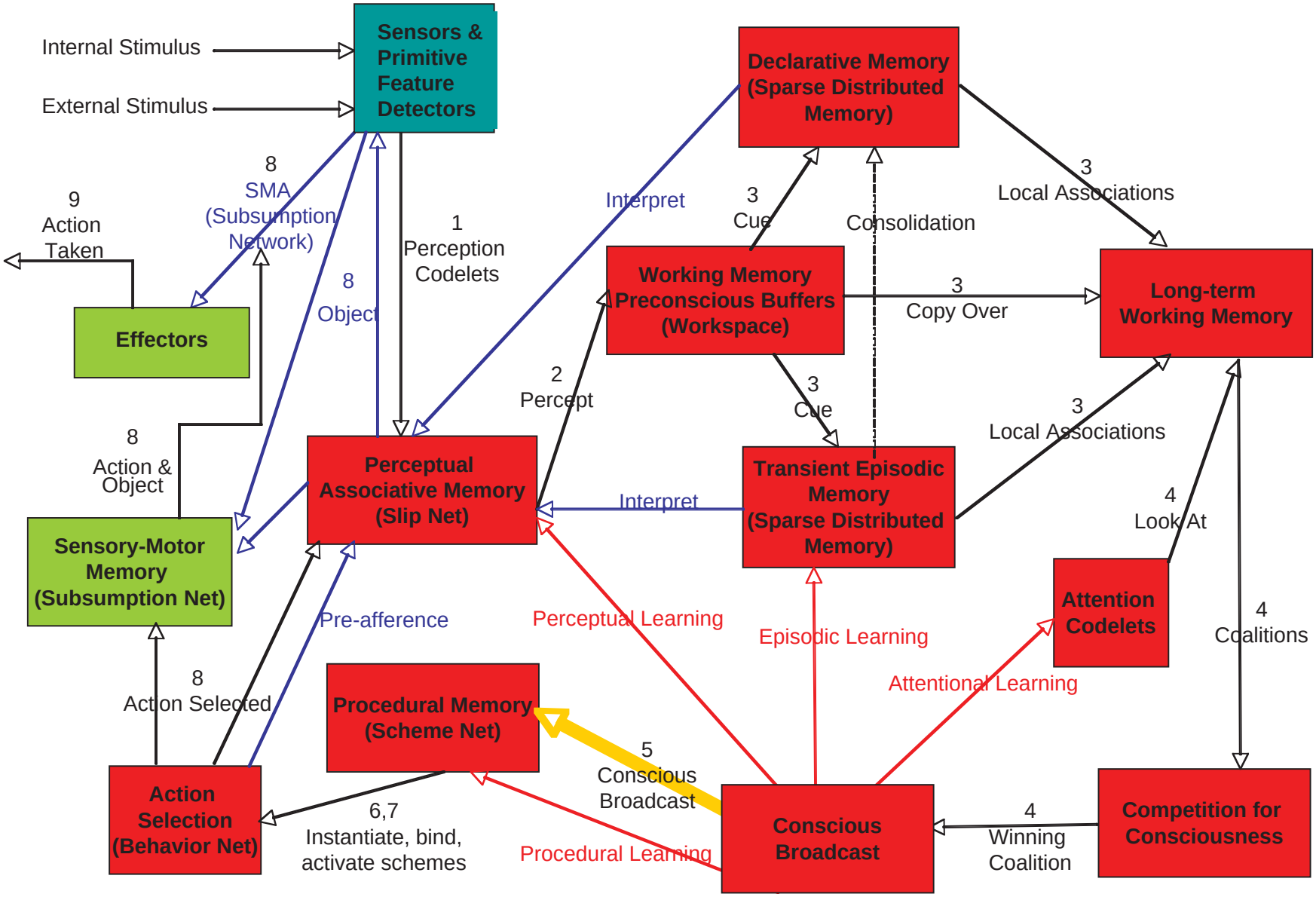
Hiroshi Ishiguru

WHAT HAS BEEN ENGINEERED SO FAR?

- Not Much!
  - Ethical advisors
  - Systems developed for discrete tasks

LIDA Cognitive Cycle

Stan Franklin & Bernie Baars (GWT)



- If there are clear limits in our ability to develop AMAs or manage robots, then it will be incumbent upon us to recognize those limits so that we can turn our attention away from a false reliance on autonomous systems and toward more human intervention in the decision-making process of computers and robots.

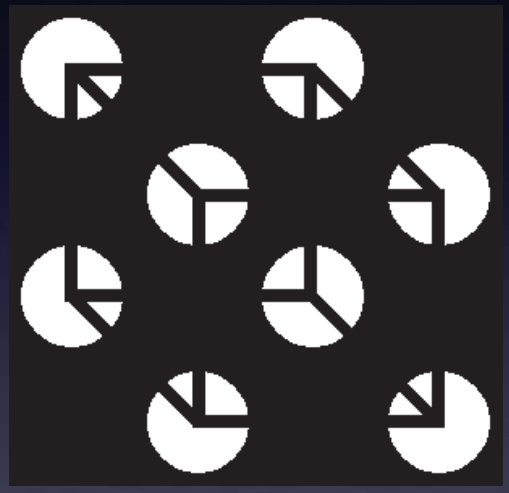
Existing Policy Mechanisms

- Govt. Financing of Research
- Laws & Regulations
  - Oversight
- Professional Codes of Ethics
- Research Ethics
- Laboratory practices and procedures
- Insurance Policy

- Govt. agencies
  - FDA, CDC, EPA, etc.
- University and Hospital
  - IRBs, IACUCs, Biosafety Committees, & ESCROs

Existing Policy Mechanisms

- Relatively robust system of protections
- Grounds for criticism
  - lack of funding for oversight
  - hamper productivity
  - piecemeal
- Adding, modifying and dismantling
  - FBI WMD Directorate, Biological Countermeasures Unit (2006)



International Considerations

- Differing values
- Precautionary Principle
  - more stringent/competitive disadvantage
  - more open/exposure to dangerous products

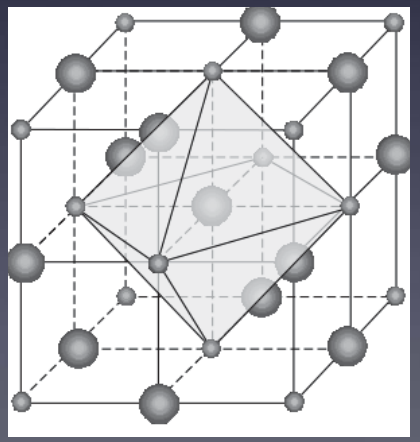
Oversight

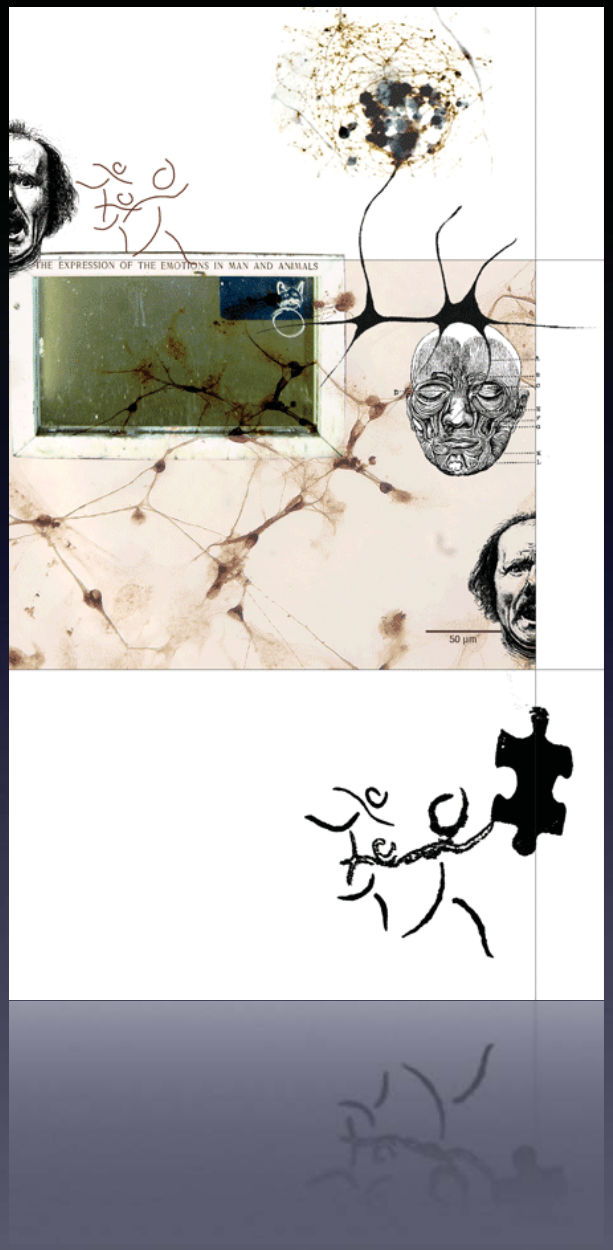
- Governance Mechanism – GCC
  - monitor
    - whether the lines of responsibility have been established for new systems being deployed.
    - whether thresholds are about to be crossed that pose serious dangers.
  - Manage
    - Relationship between stakeholders
    - Gaps
  - Modulate

Gary Marchant

GOVERNANCE COORDINATION COMMITTEE (GCC)

- Coordinate activities of the various stakeholders
- Modular
- Comprehensive monitoring
  - Flag issues/gaps
  - Find solutions within the robust set of available mechanisms
  - Mandated to avoid regulation where possible
    - Soft Governance
- Nimble/Flexible/Adaptive/Lean
- Credible vehicle





Thank You!

Email: wendell.wallach@yale.edu