

Terry Allard PhD
tta1020@gmail.com

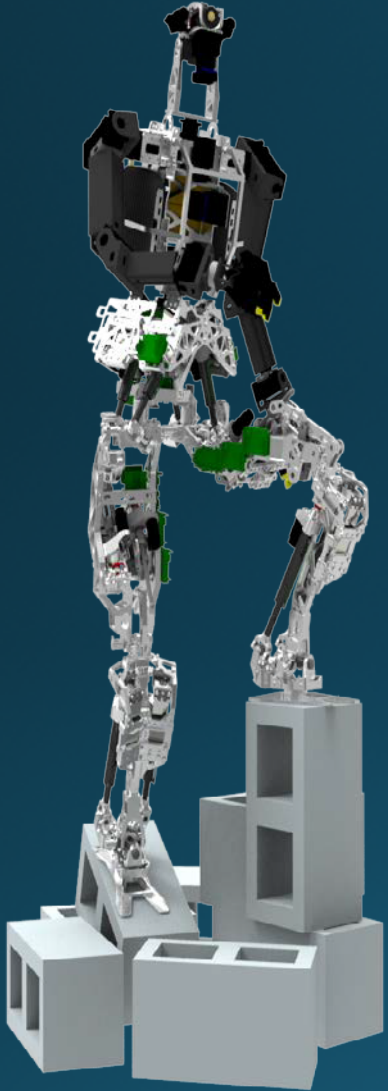
Chaos and Complex Systems Seminar
University of Wisconsin-Madison
21 November 2017

Evolution and Agency

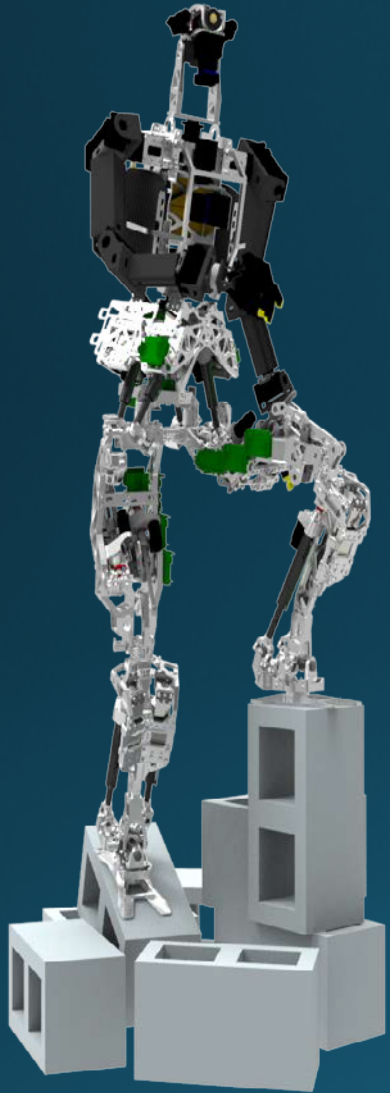
Artificial Intelligence

Artificial Intelligence

- Many Recent Successes
- Possible Long-term Risks



Overview



- Why is AI important?
- What is Intelligence?
- What is Artificial Intelligence?
- What are we afraid of?
- When is it going to happen?
- What could it look like?
- What does it mean to be human?

Why is AI Important?

US lead in AI technology is being targeted

**“Artificial intelligence is the future, not only for Russia, but for all humankind.”
“Whoever becomes the leader in this sphere will become the ruler of the world.”**

--Vladimir Putin

reported by state-funded RT, 1 Sept2017

“[China] will become the world’s leader in AI by 2030”

--People’s Republic of China State Council policy

reported by NY Times, 20July2017

but . . . Fake News, Social Media Spoofing, Bots (software robots)

What are the long-term risks?

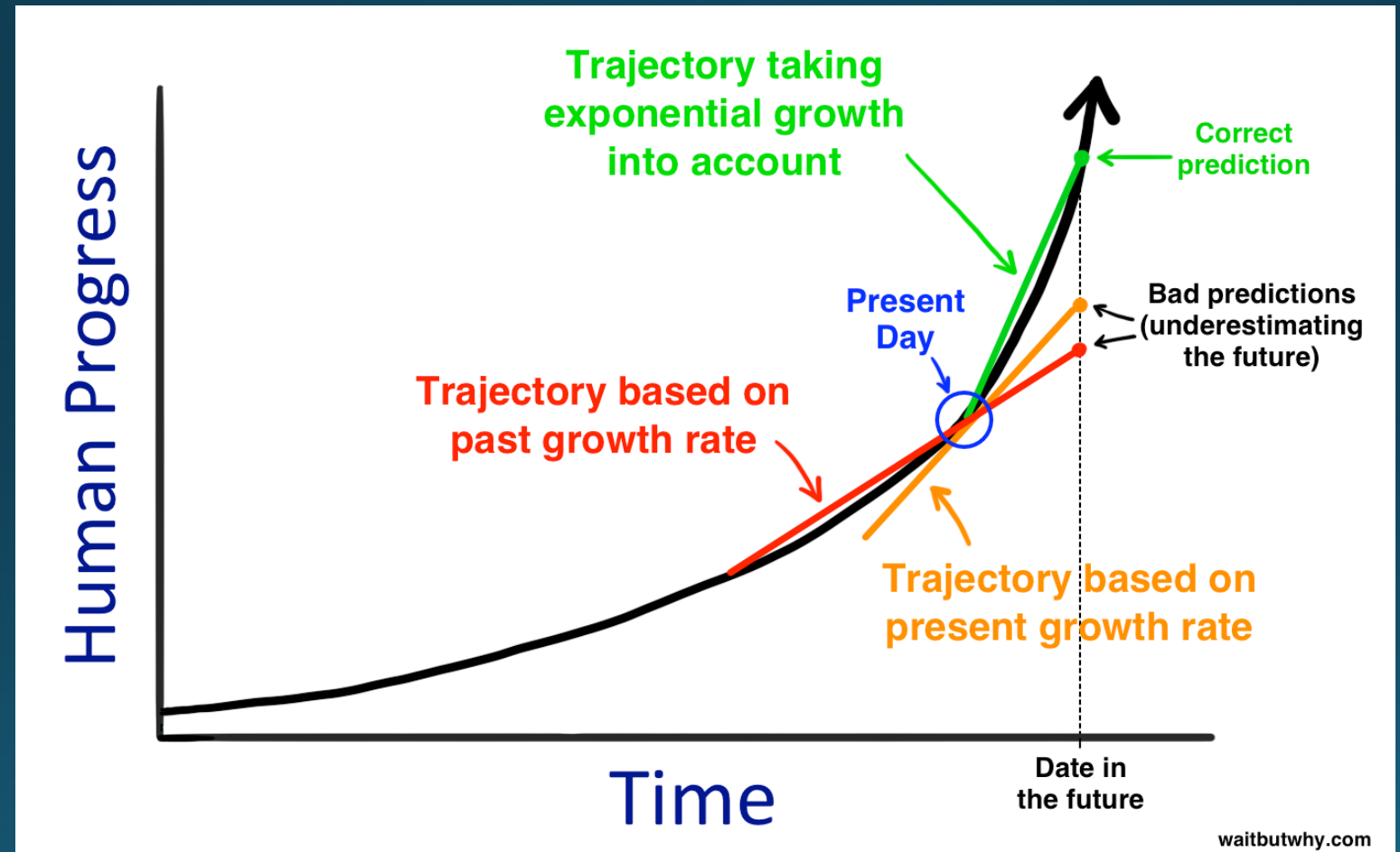
- Smart people concerned about the emergence of Super-Intelligence
 - ✓ *Stephen Hawking, Bill Gates, Elon Musk*
AI could be an existential threat to the human race
 - ✓ *Nick Bostrom (University of Oxford, Future of Humanity Institute)*
If it can be invented, it will be invented; It's just a matter of time
- Machine capability for self-improvement not limited by biology
 - ✓ Evolution of machine intelligence could outpace human potential

**INTELLIGENT MACHINES AS THE DOMINANT
LIFE FORM ON THE PLANET ???**

Could it happen soon?

People are Linear Thinkers

Your Intuition is a poor guide to the future.



Technology is Accelerating

but . . .

Still many obstacles to Deployment
e.g., Infrastructure change is slow

See counterarguments from Rodney Brooks:
“The Seven Deadly Sins of AI Predictions”\MIT Technology Review
2017, Vol 120 | No.6, pp. 79-86

What is Intelligence?

- ***Ability to acquire and apply knowledge and skills***
 - Huge range from “Simple” (e.g., Insects) to Complex (Human)
- ***Human Intelligence emerges from***
 - ✓ Pattern recognition to Storytelling
 - ✓ Problem-solving & Reasoning
 - ✓ Insight and Generalization
 - ✓ Fluid Intelligence / Working Memory
 - ✓ Procedural & Episodic Memory
 - ✓ Natural Language
 - ✓ Motor and Perceptual Learning
 - ✓ Habit formation even for complex skills
 - ✓ Social Communication
 - ✓ Implicit Biases
 - ✓ Artistic Expression
 - ✓ Language Skills
 - ✓ Spatial Cognition
 - ✓ Emotional Awareness
 - ✓ Empathy . . .

What is Intelligence?

- **Ability to acquire and apply knowledge and skills**
 - Huge range from "Simple" (e.g., Insects) to Complex (Human)

- **Human Intelligence emerges from**

- ✓ Pattern recognition to Storytelling
- ✓ Problem-solving & Reasoning
- ✓ Insight and Creativity
- ✓ Fluid Intelligence
- ✓ Procedural Memory
- ✓ Natural Language
- ✓ Motor and Perceptual Learning
- ✓ Habit formation even for complex skills
- ✓ Communication
- ✓ Cognitive Biases
- ✓ Artistic Expression
- ✓ Language Skills
- ✓ Spatial Cognition
- ✓ Emotional Awareness
- ✓ Empathy . . .

... Common Sense and Strategic Thinking are hard problems

What is Artificial Intelligence (AI)?

- *Machine Computing systems that perform tasks normally requiring Human or Animal Intelligence*
 - ✓ Embodied (e.g., robots, drones) or disembodied (software)
 - ✓ Software / Hardware / Wetware
 - ✓ Deep Learning and pattern analysis; Fast and powerful
 - ✓ Usually networked to other resources including sensors, algorithms, data, other AI's, and people

Current Trajectory of AI

- ***Narrow AI***

- Happening now, special-purpose, ubiquitous, successful

“As soon as it works, no one calls it AI any more.”

--John McCarthy, AI pioneer (1927-2011)

- ***General-Purpose aka Human-Level AI***

- Aspirational but within reach at some levels

- ***Super AI***

- Indefinite Future (2030 to 2230 ???)

see Tim Urban's *Wait But Why* article at

<https://www.inverse.com/article/7852-ai-revolution-road-superintelligence>

Narrow AI examples

Tremendous Progress

- Google Search / Siri
- Google Natural Language / Speech Recognition
- Google Translate
- Google Maps / Navigation
- Google Image Recognition
- Google . . . (notice a pattern here?)

**but ... Amazon,
Facebook, Apple,
Microsoft, Baidu**

Narrow AI examples

Allard Nov 21, 2017
AI Evolution & Agency

Tremendous Progress

- Google Search / Siri
- Google Natural Language / Speech Recognition
- Google Translate
- Google Maps / Navigation
- Google Image Recognition
- Google . . . (notice a pattern here?)
- Chess / Go players
- IBM Watson big data analytics
- Self-Driving Cars
- Aviation Autopilots
- Mars Exploration Rovers
- Nest Home Thermostats
- ***Russian bots . . .***

. . . Bots don't have to be smart to be effective. Viral transmission!

General or Human-Level AI

General Purpose Information Processing Machines and Systems

- ✓ Generalize from specifics to similar or novel situations
- ✓ Model-based interpretation of incomplete, ambiguous, implicit or erroneous information
- ✓ Learning, Adaptation, Communication, Collaboration, Interaction

e.g., Cognitive Science successes

- ✓ Education: Intelligent Tutoring Systems
- ✓ Engineering: Human and Machine simulation
- ✓ General theories of Human Information Processing
- ✓ Neuroscience, Psychology, Computer Science, . . .

Super Intelligent AI

Doesn't exist -- yet

- *Beyond Human Capability*
in Speed, Power, Complexity, Learning
- *Networked Systems of Systems*
- *Self-coding, Evolving ?*
- *Fully Autonomous Agency ?*

What is Agency ?

Capacity to act with personal consequence for the actor

✓ Unconscious & Involuntary AND / OR Purposeful & Goal-directed

Capabilities

1. Sense and Respond
2. Learn and Adapt
3. Improve and Evolve

Drivers

1. Survivability and Self-Sustainment
2. Curiosity and Completion
3. Self-Improvement and Efficiency

Levels

1. Dependent on human direction / control
2. Autonomous within defined parameters
3. Independent of human direction / guidance

Fear of Super AI Agency

Loss of Control?

- ***Run-away Self-improvement / Evolution***
 - ✓ Beyond Human capability
 - ✓ Unbounded by biology
- ***Fully Autonomous, Independent of human control or influence***
 - ✓ No “Off” Switch
 - ✓ Unintended consequences / motivations
- ***Lack of Human Values – Lack of Valuing Humans***

When could Super AI happen ?

- Current research on the nature of human intelligence ongoing
 - Machine systems increasingly capable for more complex tasks
 - Narrow-AI and networked systems are pervasive
-
- *Big data analytics, Self-organizing Systems and Networks* powering industrial “Deep Learning” and “Artificial Neural Networks” today
 - *Automated Machine Learning* is an expanding field (Google AutoML)
 - *BUT . . . Many obstacles to technology deployment*

What would it look like?

- ***AI's will not be single, independent entities***
 - ✓ Pervasive, distributed, multi-layered and networked
 - ✓ Incremental progress - not AI sentience breakthrough
- ***Human-Level AI is not a required step to Super AI***
 - ✓ Artificial Intelligence is not a simple continuum
 - ✓ Federated human-machine systems
- ***Multiple levels of agency ?***
 - ✓ No single "off switch" allowing humans to pull the plug
 - ✓ People will accept, expect and embrace interdependence

What does it mean to be Human?

People will become more fully networked with machine systems

- ✓ Prosthetics (Perception, Cognition, Social Interaction)
- ✓ Personal Monitoring; Neural Human-Machine Interaction
- ✓ Artificial Avatars and Assistants, Caregivers, Companions, . . .

**“We will transcend all of the limitations of our biology.
That is what it means to be human – to extend who we are.”**

--Ray Kurzweil

National Geographic cover story, 5 April 2017

Many Successes and Long-Term Possibilities *A new society of humans and machines ?*

