A Synopsis of My Work on
AI and the Future of Humanity

Anton Korinek

University of Virginia and Darden GSB

INET Workshop at Columbia

August 29th, 2018

Starting Point

Monism not Descartes' Dualism

- human brain is most powerful super-computer in the known universe
  - gives rise to amazing emergent phenomena such as intelligence, creativity, …

- BUT: technological progress in computing is continuing unabated
  - "artificial" intelligence is bound to overtake human intelligence, likely in the next few decades

→ We are at the onset of the *Age of Artificial Intelligence*

Progress in AI: Exponential Growth
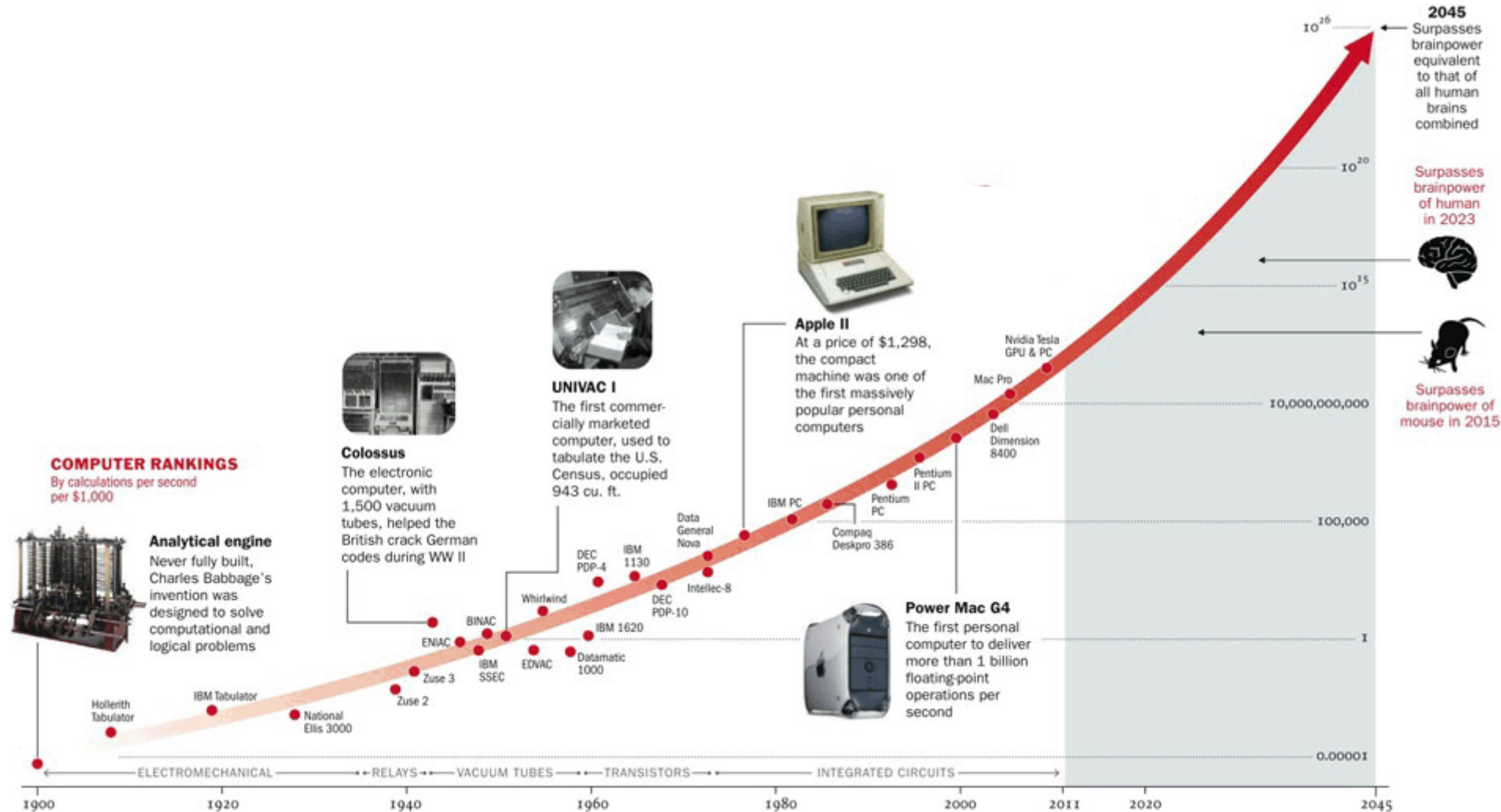
Key Questions

positive:

what will this imply for humanity?

normative:

(how) can we work towards a positive outcome for humanity?

→ in the short/medium run, i.e. during the transition
→ in the long run

Project 1: Worker-Replacing Progress

The Simple Economics of Worker-Replacing Technological Progress (with Joe)

Question: what are the economic effects of a new technology that makes workers fully redundant?

Approach:

- make machine labor M a perfect substitute to human labor H so total labor L = H + M
- study implications in a standard economic model

Project 1: Worker-Replacing Progress

Results:

1) singularity: "reproducible" labor lifts the greatest scarcity in the economy
   → AK-style growth driven by unbounded machine accumulation
   → workers not necessarily hurt – although labor share goes to zero

2) if there are other irreproducible factors (e.g. land), then these ultimately earn all the surplus from the new technology and workers *are* hurt

3) but irreproducible factors can be taxed without distortion to redistribute

Project 2: Macro-Economics of Superstars

Digitization, Increasing Returns, and the Macro-Economics of Superstars (with Ding Xuan Ng)

Question: what are the economic effects of replacing traditional production technologies with "digital" (zero MC) technologies?

Approach:

- consider a Dixit-Stiglitz multi-sector economy with perfect competition within each sector

- model digital innovation in a sector as one firm spending a fixed cost to lower the unit cost of output → *non-rival* but potentially *excludable*

Project 2: Macro-Economics of Superstars

Digitization, Increasing Returns, and the Macro-Economics of Superstars (with Ding Xuan Ng)

Results:

1) digital innovation creates natural monopolies → "superstars"
→ economic rents, growth with stagnating wages, declining labor share

2) normative side: under-production and under-innovation
→ prescription: curb monopoly power to reduce 1)
- public investment in digital innovation
- reduces sources of monopoly power: patent protection, free user information, …

Project 3: Artificially Intelligent Agents (AIAs)

Artificially Intelligent Agents in Our Economy

Questions:

- How should we think about a world with artificially intelligent agents, and about the allocation of resources between humans and AIAs?

- How can we even approach questions like "will machines replace us"?

Note: economics = science that studies the allocation of scarce resources

$\rightarrow$ in principle, well placed for such questions

Thought Experiment

Consider an observer from another galaxy who arrives on earth:

encounters humans and machines busily interacting with each other

- are the humans controlling the machines?
- or are they controlled by the little black boxes that they carry around and constantly check?
- and who controls the little black boxes?

→ just one example of the blurring lines about who is in charge

→ observer will probably view humans and machines as two different types of moderately intelligent organisms living in symbiosis

Traditional (Anthropocentric) Economics

Humans = Agents	**Machines = Objects**	**More general: entities**
• absorb consumption expenditure	• absorb investment expenditure	• absorb resources
• supply labor services	• supply capital services	• supply factors
• behavior encoded in preferences	• behavior encoded in technology	• exhibit defined behavior
• evolve according to a law of motion	• evolve according to a law of motion	• evolve according to law of motion

Project 3: Artificially Intelligent Agents (AIAs)

Findings:

- economy with different types of agents fits easily into our frameworks
- currently: humans and AIAs are symbiotic
- in recent decades, resource allocation has shifted more and more towards machines → economic growth is entity-specific
- property rights vs. who determines de-facto resource allocation
- an economy of the machines, by the machines and for the machines is conceivable