The Temporal Singularity: time-accelerated simulated civilizations and their implications

Giacomo Spigler

The Biorobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy

ISTITUTO
DI BIOROBOTICA

Scuola Superiore
Sant'Anna

The Temporal Singularity (1/4)

- Old idea in AGI #1: **speeding up the execution** of an agent to improve its problem solving capabilities (e.g., [Vinge, 1993]).
  - More subjective time available -> shorter solving time in the external world.
- Not only valid for AGI. *Example*: current Reinforcement Learning agents.

Architecture	CPUs	GPUs[1]	FPS[2]	
Single-Machine			Task 1	Task 2
A3C 32 workers	64	0	6.5K	9K
Batched A2C (sync step)	48	0	9K	5K
Batched A2C (sync step)	48	1	13K	5.5K
Batched A2C (sync traj.)	48	0	16K	17.5K
Batched A2C (dyn. batch)	48	1	16K	13K
IMPALA 48 actors	48	0	17K	20.5K
IMPALA (dyn. batch) 48 actors[3]	48	1	21K	24K
Distributed				
A3C	200	0	46K	50K
IMPALA	150	1	80K	
IMPALA (optimised)	375	1	200K	
IMPALA (optimised) batch 128	500	1	250K	

[1] Nvidia P100 [2] In frames/sec (4 times the agent steps due to action repeat). [3] Limited by amount of rendering possible on a single machine.

The Temporal Singularity (2/4)

- Old idea in AGI #2: **collective intelligence**. Many agents coming up with different solutions or working together (e.g., [Solomonoff, 1985]).

- In the extreme, we can think of simulating entire "civilizations", possibly pursuing complex sets of goals like general progress in science and technology.

- NOTE: we refer to agents and civilizations as "simulated" only to mean that they experience a simulated environment in contrast to the "real" external world, but they would technically be as real as any agent outside the simulation.

The Temporal Singularity (3/4)

- If a group of simulated agents unfolding at faster than external-time can start a *Technological Singularity*, the resulting progress in computing technology **may allow to speed up** the simulation (which in turn would accelerate progress, which would further speed up the simulation, etc...)

- Normal exponential growth (e.g., 2x for each unit of time):

  1, 2, 4, 8, 16, …

- Time-accelerated exponential growth (e.g., 2x for each unit of subjective time, with hardware updates every unit of external time):

  1, 2, 8, 2048, …

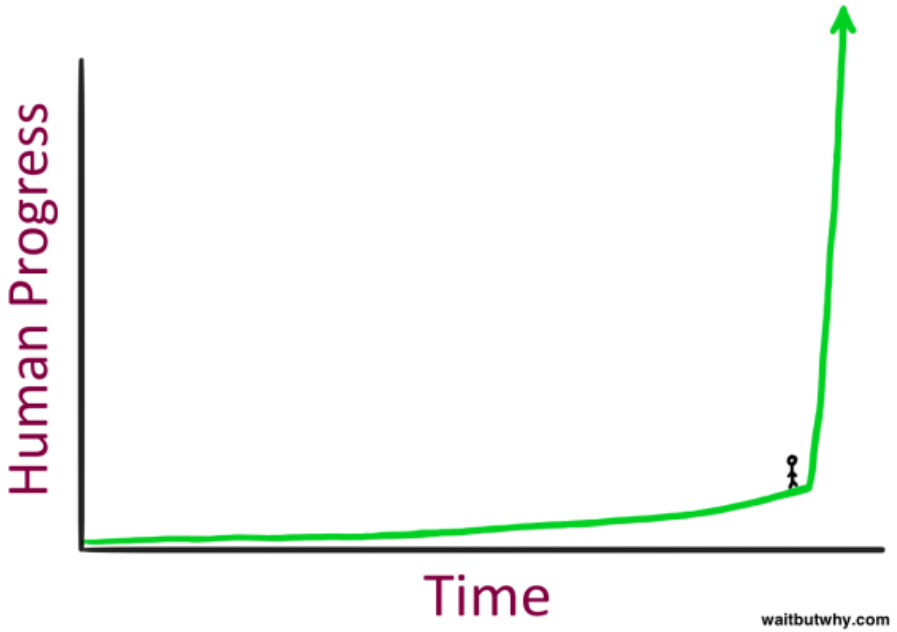
The Temporal Singularity (4/4)

- Assuming the growth in temporal speed-ups holds for some time, then the simulations may progress to producing entire civilizations spanning thousands or millions of years or even more in an arbitrarily short external time interval.
- The extreme time-dilation observed from the outside and its rate of growth would look like a **Temporal Singularity**.
- **Temporal Singularity**: the moment in time when a minimal simulation capable of beginning the runaway exponential growth of its own temporal speed-ups is started.

- It is difficult to imagine what it would look like, as even a single century of progress at present rate is challenging to forecast. What about millions of years, during which potentially any questions we may ever ask could have been answered.
- See [Hutter, 2012] for a nice discussion of what a Singularity may look like from inside vs outside the simulation.

Singularity vs Temporal Singularity

- The Temporal Singularity is a type of Technological Singularity.
- Difference: the most common definition of Technological Singularity requires a runaway increase in the cognitive capabilities of the artificial agents, while the Temporal Singularity does not (strictly) require it. The Temporal Singularity only requires a *runaway growth of the temporal speed-ups* of the simulation.



A Few Implications

- Shared with the Technological Singularity:
  - Progress in science and technology.
  - Economic advantages for the first entity to achieve it and/or for humanity.
- Further advantages:
  - Simulated civilizations need to be stable for a very long (subjective) time -> study of societal designs and hints for our own long-term survival.
  - Implications on the future of intelligent civilizations (see next slides)
- Dangers:
  - General issues with AI boxing, made worse by the difficulty to understand what's going on inside without losing the benefits of accelerated time.
  - Existential risks due to sudden technological progress: e.g., what would have happened if we abruptly produced not just the technology, but a full stockpile of thermonuclear weapons during the middle ages?

The Fermi Paradox (1/3)

- What is the Fermi Paradox?
  - "*Apparent contradiction between the lack of evidence and high probability estimates for the existence of extraterrestrial civilizations.*"
- **(1)** Because of its advantages, if a Temporal Singularity is achievable, it may be *inevitable* that all intelligent civilizations that develop computing and technology would ultimately produce it.
  - Time-accelerated simulations could thus be part of some or all possible intelligent civilizations, with advantages like achieving "subjective immortality", for individual agents or for their civilization as a whole.
  - Possible implication of an abundance of post-biological civilizations in the universe.
- **(2)** External time would be an important resource, and the speed of galactic colonizazion and communication, which is already considered slow, may become unbearable. Future civilizations may then prefer to *avoid large-scale galactic colonization*.

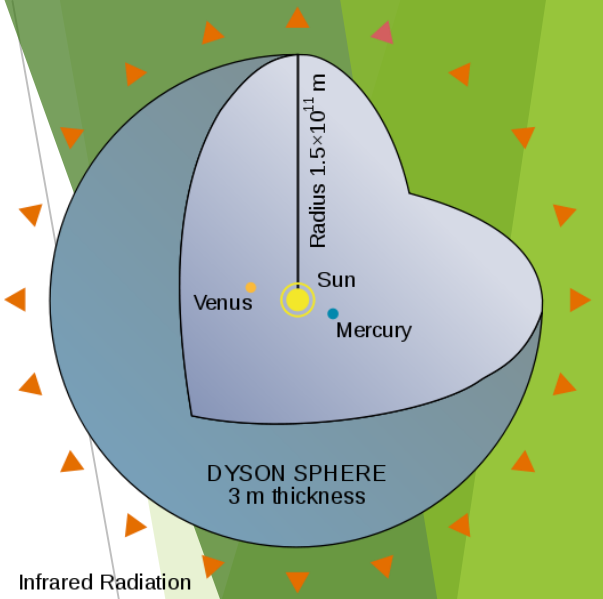
The Fermi Paradox (2/3)

- **(3)** Transcension Hypothesis (Smart, 2012): advanced civilizations may strive for information gain and processing by migrating to increasingly dense and miniaturized spaces, and eventually approach black holes.
  - Opposite dynamics as the Temporal Singularity: civilizations trying to *forward time travel* to a point where all civilizations can meet via black hole merging. Subjective time is slowed down instead of accelerated.
- **(4) Great Filter**: an obstacle that prevents most civilization from surviving or progressing.
  - The Temporal Singularity as a problem: sudden development of technology and AI boxing.
  - The Temporal Singularity as a solution: advanced technology and extreme temporal speed-ups may help overcome traditional Great Filters, like asteroid impacts or the Berserker scenario (i.e., an advanced civilization attacking any newly emerging civilization).

The Fermi Paradox (3/3)

- **(5)** A prediction:
  - The Temporal Singularity implies an extreme increase in the power used by a civilization in short time spans.
  - Abrupt transition of host civilizations from Type-I to Type-II (Kardashev scale) in decades rather than millennia.
  - Possible construction of Dyson spheres in a very short time.
  - Present day telescopes are already capable of detecting even partial potential Dyson spheres (up to some distance from us) via deviations in the host star's infrared spectrum.
  - The search should look at changes in the star spectra over time.

- **(6)** *Bonus implication*: if the Temporal Singularity is inevitable, and science/technology were somehow bounded, then all intelligent civilizations past this stage would be roughly at the same level of development.



Conclusions

- Positive feedback between runaway growth in scientific and technological progress, as in the Technological Singularity, and increase in the temporal speed-ups of the simulations, may lead to simulated civilizations unfolding much faster than the external world.
- Great advantages, but also great dangers.
- Deep implications on the future of our own civilization and potential other advanced civilizations in the universe.

- Medium-term relevance of the topic: a back-of-the-envelope calculation suggests that the computing power required to start the process may be available in few decades.
  - Also, because of the competitive advantages of running such a simulation, it may be inevitable that if it will ever be technically possible to create it, it will be created.

Questions?

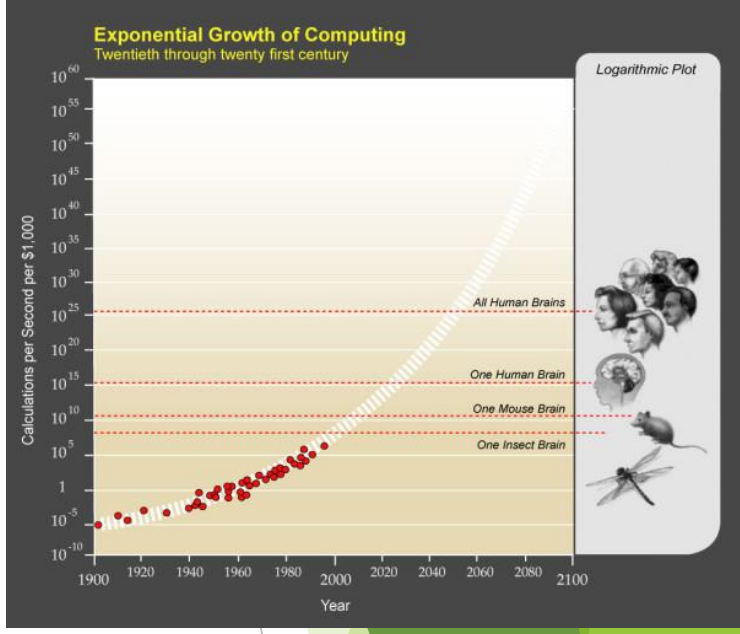
giacomo.spigler@santannapisa.it

Supplementary Slides

Feasibility (1/2)

- AGI may not be required, at least not immediately.
- Perfect or imperfect simulation? Is fooling the agents required?
- Computational requirements:
  - AGI estimates based on the estimated computational power of human brains. ~10^13 to ~10^18-10^25 FLOPS, with a common estimate ~10^16 FLOPS (tens of PFLOPS).
- An imperfect simulation may limit the engineering and technology that can be developed inside. However, even present day development already consists of a large time in simulation (e.g., COMSOL simulator, etc...). Also, software development, CS, math and lots of other research can be performed without any access to the external world. For experiments, the agents may provide all the details of the procedures to be performed in the external world, and then collect the results.

Feasibility (2/2)



- Minimal simulation:
  - ~100x a single agent (# agents x speed-up) ?
- If Moore's Law holds: ~ 10^18 - 10^21 FLOPS should be available by 2020-2040 on supercomputers, and 2055-2075 on PCs. Specialized hardware (e.g., GPUs and TPUs) could make such power available much earlier though, as soon as we get closer to achieving one such minimal simulation and the algorithms and requirements are more clear.
- Allocation: *number of agents* vs *temporal speed-ups* vs *resources per agent*.
- Limitations: temporal speed-ups may not be the most optimal allocation (though it is likely to be, at least at times); perfect simulations may be required; upgrades and maintenance, which could limit the ultimate achievable speed-up, etc...