Ethical Considerations

What is at stake from a moral perspective?

Teletransportation Paradox

«I would be glad to know … whether when my brain has lost its original structure, and when some hundred years after the same materials are fabricated so curiously as to become an intelligent being, whether, I say that being will be me; or, if, two or three such beings should be formed out of my brain; whether they will all be me, and consequently one and the same intelligent being.» – Thomas Reid (1775)
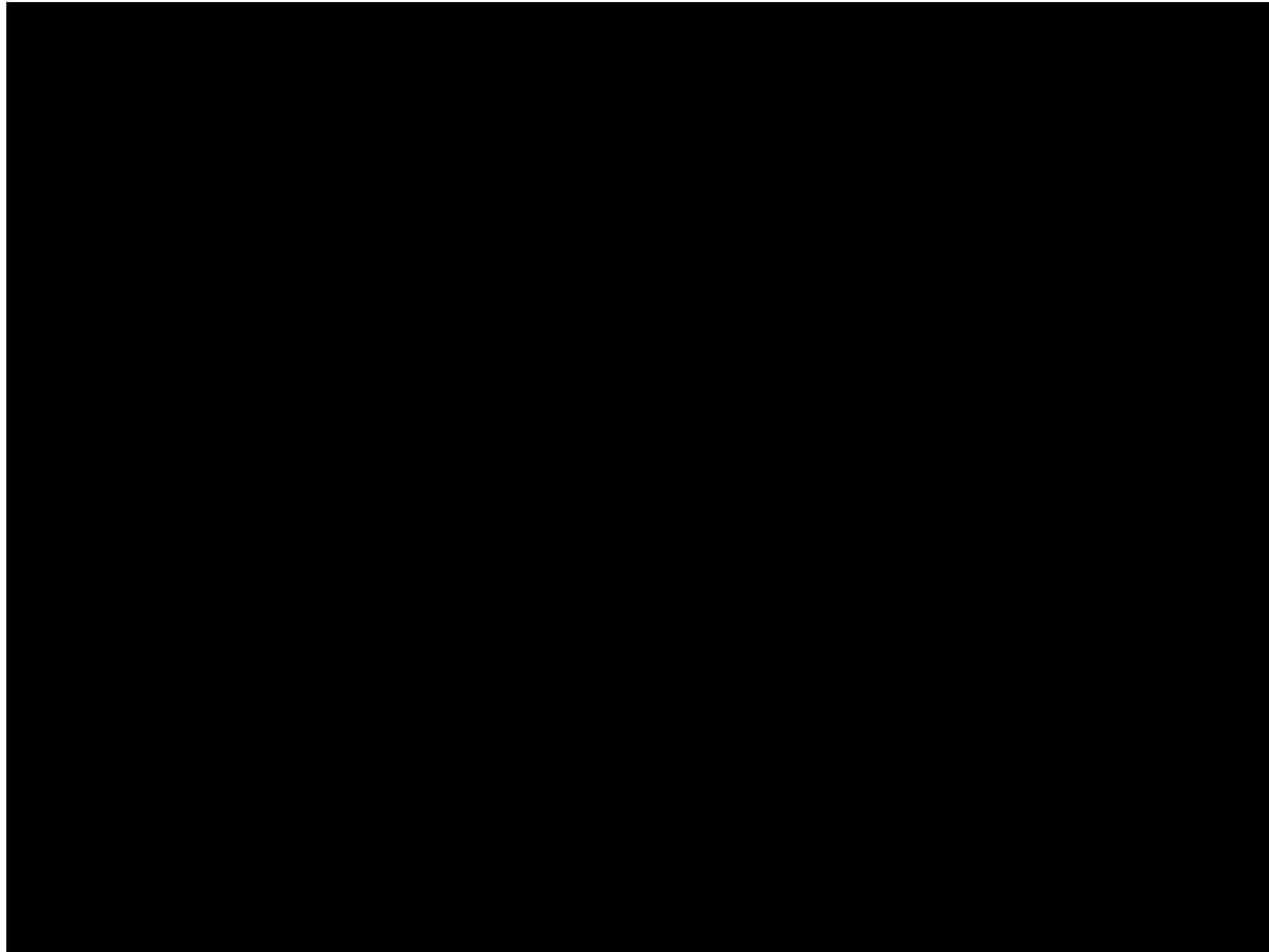
Personal Identity

What does it mean that two persons (e.g. me now and in 20 years) are *one and the same*?

Relevance:

– If we have more and more possibilities to modify (or even copy) humans, what does it mean for a person to *continue existing*?

– What is it that we ultimately care about if we care about our continued existence?

John Weldon's "To Be"

Whole Brain Emulation (WBE)

«Whole brain emulation or mind uploading is the hypothetical process of copying mental content from a particular brain substrate to a computer. The computer could then run a simulation of the brain's information processing, such that it responds in essentially the same way as the original brain (i.e. being indistinguishable from the brain for all relevant purposes) and experiences having a conscious mind.»
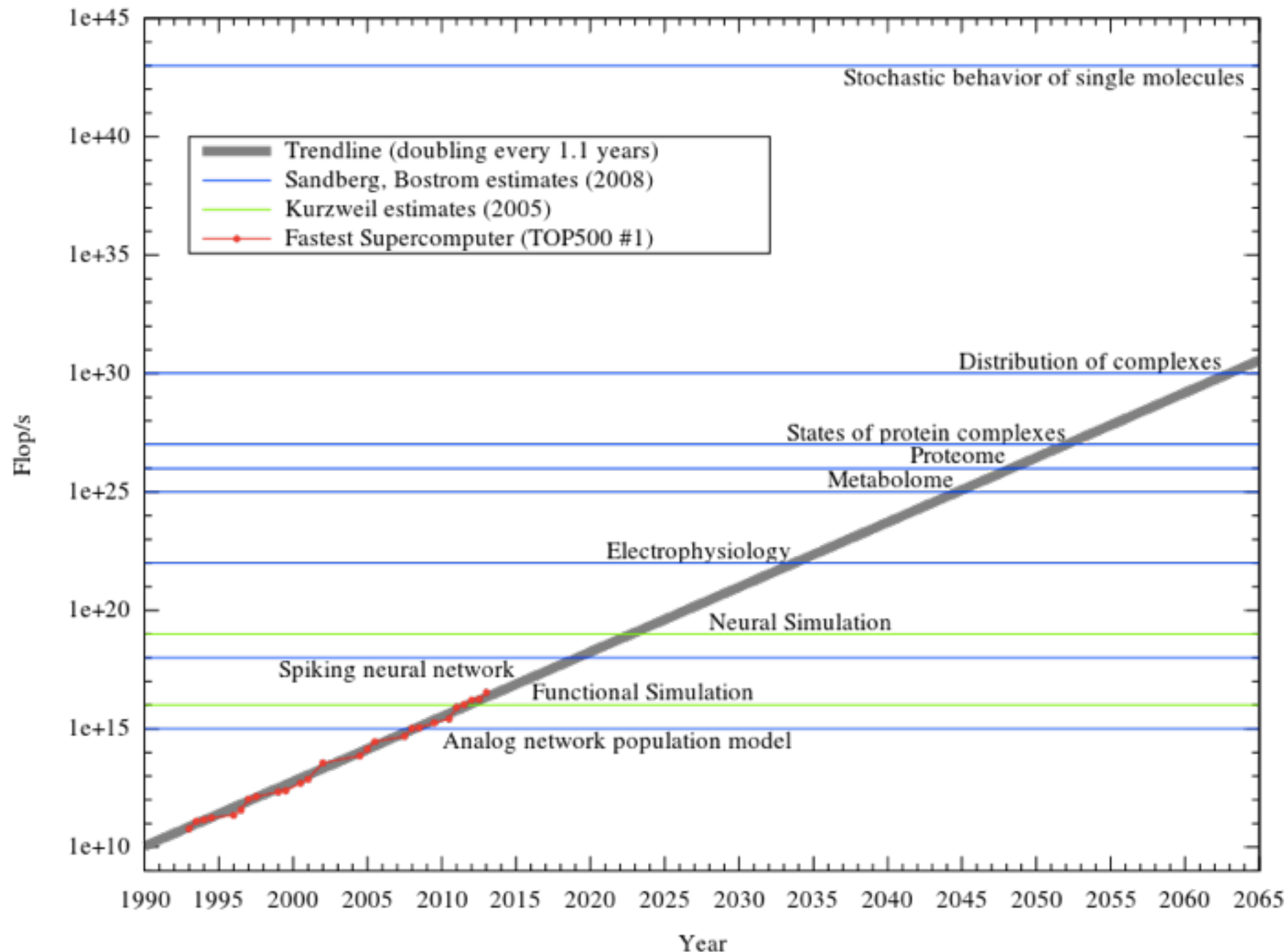
Benefits of WBE

– **Backup** ("immortality")

– **Speedup** by many factors

– **Easier** to copy and update

Path towards strong AI?

Timelines for WBE

Graph: Flop/s vs. Year (1990–2065), logarithmic y-axis from 1e+10 to 1e+45.

Legend:
- Trendline (doubling every 1.1 years)
- Sandberg, Bostrom estimates (2008)
- Kurzweil estimates (2005)
- Fastest Supercomputer (TOP500 #1)

Horizontal reference lines (Sandberg, Bostrom estimates): Stochastic behavior of single molecules, Distribution of complexes, States of protein complexes, Proteome, Metabolome, Electrophysiology, Spiking neural network, Analog network population model.

Kurzweil estimates: Neural Simulation, Functional Simulation.

Prospects for WBE

«A detailed, functional artificial human brain can be built within the next 10 years.» – Henry Markram, Director of the Blue Brain Project at EPFL, 2009

The resolution of scanning remains the biggest obstacle.

Ethics of Uploading

Transcendence: Caster's body dies, but he can upload his mind.

- Would you do it?

- Under what conditions would you consider the upload to be yourself?

- If we consider different uploading procedures with different accuracy, how accurate must the procedure be such that you would consider the update to be yourself? Where's the boundary?

- If we could copy your upload, should we do it? How often?

Digital Sentience

– When is a non-biological being worthy of **moral concern**?

– What is the moral relevance of a simulation of a biological brain?

– What about **intelligent processes** with no relation to biological brains?

Artificial Suffering

«The story of the first artificial Ego Machines, those postbiotic phenomenal selves with no civil rights and no lobby in any ethics committee, nicely illustrates how the capacity for suffering emerges along with the phenomenal Ego; suffering starts in the Ego Tunnel.

It also presents a principled argument against the creation of artificial consciousness as a goal of academic research. Albert Camus spoke of the solidarity of all finite beings against death. In the same sense, all sentient beings capable of suffering should constitute a solidarity against suffering. Out of this solidarity, we should refrain from doing anything that could increase the overall amount of suffering and confusion in the universe. While all sorts of theoretical complications arise, we can agree not to gratuitously increase the overall amount of suffering in the universe — and creating Ego Machines would very likely do this right from the beginning. We could create suffering postbiotic Ego Machines before having understood which properties of our biological history, bodies, and brains are the roots of our own suffering. Preventing and minimizing suffering wherever possible also includes the ethics of risktaking: I believe we should not even risk the realization of artificial phenomenal self-models.»

Thomas Metzinger: The Ego Tunnel
www.amazon.com/dp/0465020690

Population Ethics

«It is impossible to take a stance on such important problems as climate change policy or healthcare priori-sation without making controversial assumptions about population ethics.»
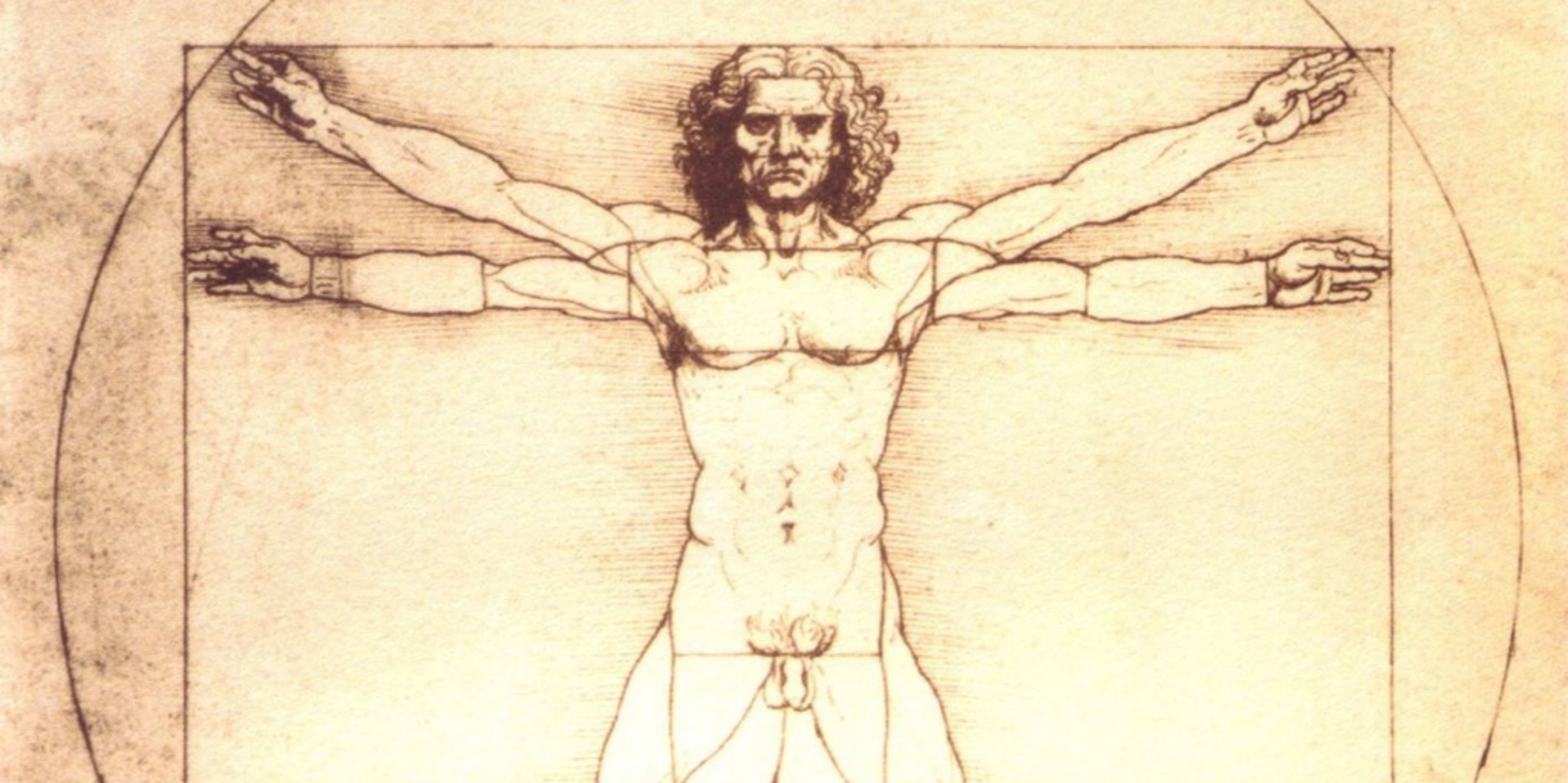
Ethics of Creating New Beings

When should we (not) create new beings?

– Population ethics: How to **compare** populations of different sizes and welfare levels?

– Population ethics is necessarily **counterintuitive**: There impossibility theorems for compelling assumptions!

Moral Realism vs. Anti-Realism

– Do ethical sentences express propositions that refer to objective features of the world we live in?

– The question has been debated in Western Philosophy at least since Plato.

– Moral realism is the only possibility to block the Orthogonality Thesis.

Anthropical Considerations

Is the universe fine-tuned for conscious life?

Fine-Tuning of the Universe

- Fine-tuned physical parameters: Even small changes would prevent life as we know it.

- There are around 200 of them.

- Example: Energy of the Hoyle state (abundance of carbon in the universe)

Anthropic Principle

Is fine-tuning "surprising"?

«The universe (and hence the fundamental parameters on which it depends) must be such as to admit the creation of observers within it at some stage.» – Brandon Carter, 1974

Anthropic Reasoning

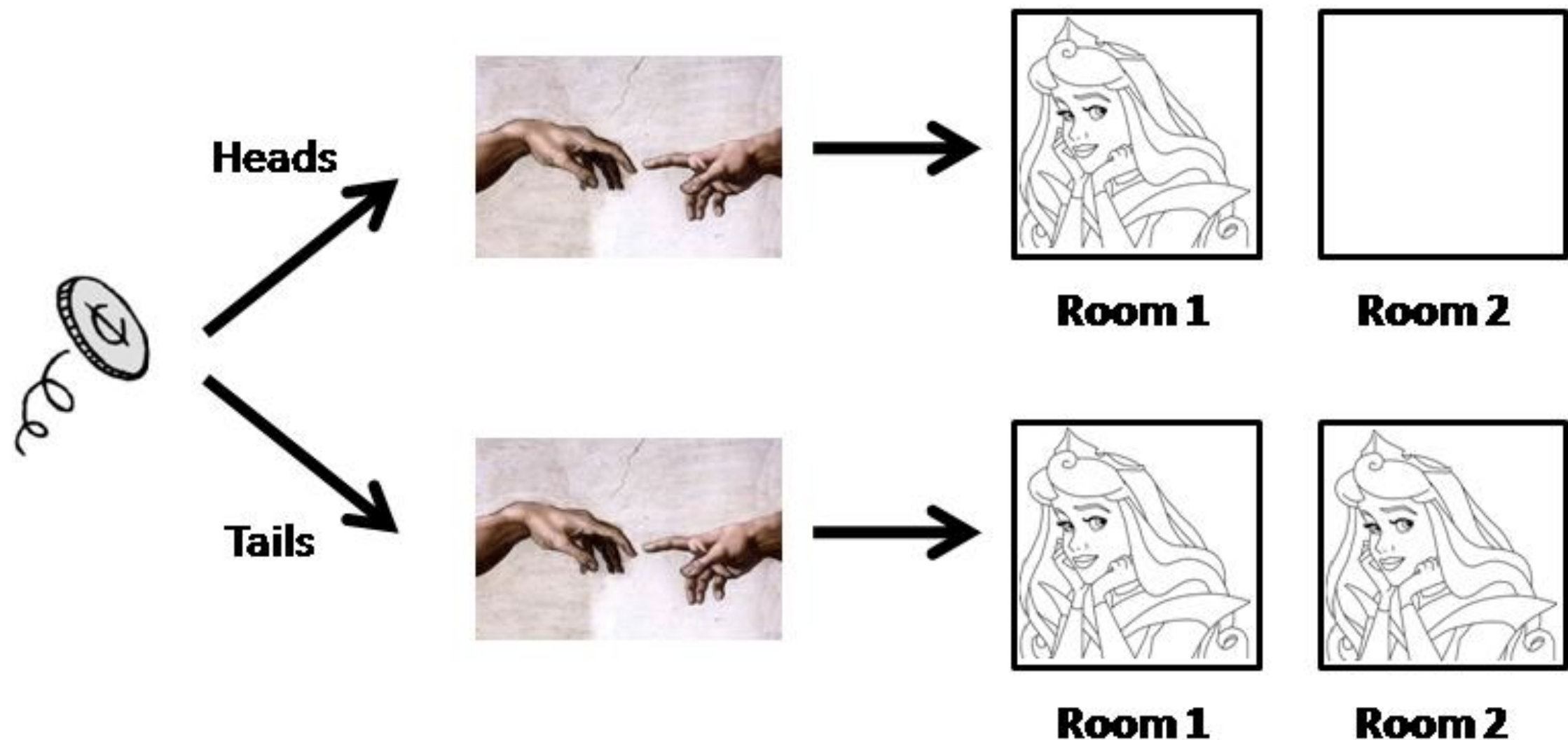How do you assess the credence in hypotheses that affect whether you exist or not?

Self-Sampling Assumption
All other things equal, an observer should reason as if they are randomly selected from the set of all *actually existent* observers (past, present and future) in their reference class.

Self-Indication Assumption
All other things equal, an observer should reason as if they are randomly selected from the set of all *possible* observers.

Incubator Thought Experiment

Doomsday Argument

– 60 billion humans have lived so far.

– Self-Sampling Assumption: Us having such a low birth-rank is more probable if there are fewer humans in total.

– Earlier "doom" is likelier than we might have thought using other considerations.

Presumptuous Philosopher

– The Doomsday Argument is one of the main arguments against SSA.

– Main argument against SIA: Automatic strong preference for a large universe with many observers.

Simulation Argument

At least one of the following propositions is true:

1. The human species is very likely to go extinct before reaching a "posthuman" stage.

2. Any posthuman civilization is extremely unlikely to run a significant number of simulations of their evolutionary history.

3. We are almost certainly living in a **computer simulation**.