Simulation Argument

Pratik Saptal[#1], Tushar Kamble[*2], Laukik Limbukar[#3], Ameya Wad [#4]

Computer Department, Mumbai University
Terna Engineering College, Mumbai, India

[1]pbsaptal@gmail.com
[2]tshrtnmk8@gmail.com
[3]laukik2424@gmail.com
[4]wadameya@gmail.com

Abstract— Nick Bostrom, a philosopher presented the exasperating idea that we are now living in a computer simulation, in 2003. Although Nick Bostrom's argument is structured to include a "hypothesis," it is not sorted that his proposition can be accounted as a properly scientific hypothesis. Bostrom's argument is engaged seriously by accounting for philosophical and scientific positions that have implications for Bostrom's principal thesis. Discussions from Heidegger, Einstein, Heisenberg, Feynman, and Dreyfus are consider that relate to modelling of structures of thinking and computation. In consequence of this accounting, given that there seems to be no reasonably admissible evidence to count for the task of falsification, one concludes that the computer simulation argument's hypothesis is only speculative and not scientific.

Keywords-- Bostrom; simulation hypothesis; Heidegger; Dreyfus; Feynman; falsifiability

I. INTRODUCTION

Science fiction as well as some forecasts by serious technologists and futurologists predicts that enormous amounts of computing power will be available in the future. Let us consider for a moment that these predictions are correct. One thing that later generations might do with their super-powerful computers is run detailed simulations of their forebears or of people like their forebears. Because their computers would be so powerful, they could run a great many such simulations. Suppose that these simulated people are conscious (as they would be if the simulations were sufficiently fine-grained and if a certain quite widely accepted position in the philosophy of mind is correct). Then it could be the case that the vast majority of minds like ours do not belong to the original race but rather to people simulated by the advanced descendants of an original race. It is then possible to argue that, if this were the case, we would be rational to think that we are likely among the simulated minds rather than among the original biological ones. Therefore, if we don't think that we are currently living in a computer simulation, we are not entitled to believe that we will have descendants who will run lots of such simulations of their forebears. That is the basic idea. The rest of this paper will spell it out more carefully.

In 2003, Nick Bostrom argued that at least one of several propositions is likely to be true:

(1) The human species is very likely to go extinct before reaching a "posthuman" stage, i.e., *(fp≈0)*;
(2) Any posthuman civilization is extremely unlikely to run a significant number of simulations of their evolutionary history (or variations thereof), i.e., *(fI≈0)*;
(3) We are almost certainly living in a computer simulation, i.e., *(fsim≈1)*.

(Bostrom 2003) Proposition (3) is characterized as the simulation *hypothesis*, thus only a part of Bostrom's simulation argument. The argument is thus basically a statement of possibilities:

Either (fp≈0) or (fI≈0) or (fsim≈1)—such that, Bostrom claims, we should distribute our credence more or less evenly among them.1 But this leaves us with the problem of justification, i.e., why we should distribute our belief in one or another of these propositions more or less evenly.

Bostrom uses the words "living in" to stipulate that, whatever "we" *are* is to be understood in terms of the following points concerning a posthuman simulation of present-day humans (Bostrom 2003):

(1) “a computer running a suitable program would be conscious.”

(2) “it would suffice for the generation of subjective experiences that the computational any irregularities.

(3) “Simulating the entire universe down to the quantum level is obviously infeasible, unless radically new physics is discovered. But in order to get a realistic simulation of human experience, much less is needed – only whatever is required to ensure that the simulated humans, interacting in normal human ways with their simulated environment, don’t notice any irregularities.”

4) “a posthuman simulator would have enough, much less is needed – only whatever is required to ensure that the simulated humans, interacting in normal human ways with their simulated environment, don’t notice in the appropriate domain on an as-needed basis. Should any error occur, the director could easily edit the state of any brains that have become aware of an anomaly before it spoils the simulation. Alternatively, the director could skip back a few seconds and rerun the simulation in a way that avoids the problem.”

II. DISCUSSION

Despite these observations, Bostrom is not adequately clear how we are to understand the term ‘simulation’. For example, we conceive of humans as having “organic” intelligence (assuming here some mind-brain interaction) and distinguish this from “artificial” intelligence. The former is associated with a biological (carbon-based) entity, while Bostom anticipates the latter is of a different material substrate. Bostrom allows for posthumans implementing a process of mechanized intelligence that operates on some kind of material substrate that need not be organic. There would then be a causal relation of orders of being: A simulation, S, (3rd order), i.e., what is being processed, is operationally dependent on a model, M, (2nd order), i.e., the programming, that re-presents a reality, R, (1 st order), i.e., what is fundamentally real and, as such, is the presupposition of any model, thus: (R → M → S, i.e., if and only if there is that which is fundamentally real can there then be a model which is the representation of that reality by way of a programming experienced as a simulation). Hence, when he says humans may be “living in” a simulation, Bostrom means that literally:
They have their being only as 3rd order artificial intelligence processes and they are not “really” biologically independent organic intelligent entities such as we presently understand the members of the set, Homo sapiens, to be. All that we are, all that we think and do, whether seemingly mental or corporeal activity, all are the manifestations of a simulation, or said otherwise, what posthumans would call “ancestor-simulations.”

Bostrom’s extended argument presupposes historical relation between a species of posthumans and contemporary humans, such that
(1) Posthumans are objectively real beings (1 st order),
(2) Contemporary humans are simulated beings (3rd order),
(3)There is a universe (i.e., a physical reality) that is objectively real (1 st order), although the perceived universe of the simulated beings may be nothing more than a simulation. Bostrom conjectures:

…later generations…with their super-powerful computers [might] run detailed simulations of their forebears or of people like their forebears…They could run a great many such simulations. Suppose that these simulated people are conscious…[It] could be the case that the vast majority of minds like ours do not belong to the original race but rather to people simulated by the advanced descendants of an original race. It is then possible to argue that, if this were the case, we would be rational to think that we are likely among the simulated minds rather than among the original biological ones. (Bostrom 2003)

He concludes, “Therefore, if we don’t think that we are currently living in a computer simulation, we are not entitled to believe that we will have descendants who will run lots of such simulations of their forebears.” (Bostrom 2003) But there are questions begging here: Why should anyone think we are currently living in a computer simulation? Why would anyone believe, or want to believe, that we will have descendants who will run many simulations of forebears such as ourselves? And how is it that the former proposition entitles one to believe the latter?
It is not Bostrom’s conclusion that interests us here. Since we do not normally think—and do not find it normatively rational to think—that we are currently “living in” a computer simulation, rather than each human today being an objective, material (biological, organic) reality, then it is by no means problematic to

us to be concerned with the likelihood of ancestor-simulations or our logical "entitlement" to any such belief. As a matter of what at the least appears to us to be objectively probable real fact, we may or may not have descendants (posthumans) who will run simulations of humans. There is no epistemological obligation to believe this proposition, although we may entertain it, at minimum, as a prediction having probability value (where truth value=1 and probability-value is >0 but <1).

III. THE LOGIC OF BOSTROM'S PROPOSAL

At this point it may be useful to clarify what seems to be the logic of Bostrom's "proposition (3)," i.e., that we are almost certainly living in a computer simulation. Two claims are central:

1. Posthumans (who live at a real space-time TF) run simulations of humans (who live at real space-time TP), but the latter are not simulations at real space-time TP. We give a symbolic notation for this proposition, thus: *PTF (rs) HTP ≠ Sim.*
2. Posthumans (who live at a real space-time TF) run simulations of humans (who "live" at virtual space-time TP, i.e., in programmed computational design features only,) who are thus mere simulations of humans and thus not really biological humans. We give a symbolic notation for this proposition, thus: *PTF (rs) Sim VTP ≠ H.*

Bostrom claims, "…the main computational cost in creating simulations that are indistinguishable from physical reality for human minds in the simulation resides in simulating organic brains down to the neuronal or sub-neuronal level…[For] a realistic simulation of human history, we can use ~1033-1036 operations as a rough estimate." He argues further, "we would be rational to think that we are likely among the simulated minds rather than among the original biological ones."

Bostrom's foregoing statements point to two distinct propositions:

3. (x)[(Hx → ~Simx)], meaning 'For all x, if x is a human, then x is not a simulation.' That is to say, all humans such as we know them are real minds and not simulated minds.
4. (x)[LMx → (LMx = Mx)], meaning 'For all x, if x is a mind like ours (i.e., x has a only a simulated consciousness and is not a real consciousness), then x with a mind like ours is really/identically our mind.' That is to say, LMx, the mind like ours, is a virtual reality (better, a virtuality) having virtual experiences. Mx, i.e., a (non-biological) human who believes s/he is a real mind, has only virtual experiences that deceive him/her into having and holding this false belief. In short, Mx is a virtuality and not a reality.

Given (3) and (4) above, we are presented with a disjunctive syllogism, the second premise of which remains undetermined and in q

5. *P1: (3) v (4)*

   P2: ~(3)? /~(4)? [undecided (yet to be determined) negation of one of the disjuncts]

   C: (4) [if P2 is decidedly ~(3)]; (3) [if P2 is decidedly ~(4)]

Alternatively, one could argue, there is an x such that 'some humans (seemingly living today) are simulations' and 'some humans (really living today) are not simulations.' This can be represented in symbolic notation,

thus: (∃x)[(H*x → Simx) ● (Hx → ~Simx)]

The question thus arises: How do we falsify the hypothesis that, 'Some humans are simulations'?

IV. CONCLUSION

A technologically mature "posthuman" civilization would have enormous computing power. Based on this empirical fact, the simulation argument shows that *at least one* of the following propositions is true: (1) The fraction of human-level civilizations that reach a posthuman stage is very close to zero; (2) The fraction of posthuman civilizations that are interested in running ancestor-simulations is very close to zero; (3) The fraction of all people with our kind of

experiences that are living in a simulation is very close to one.

If (1) is true, then we will almost certainly go extinct before reaching post humanity. If (2) is true, then there must be a strong convergence among the courses of advanced civilizations so that virtually none contains any relatively wealthy individuals who desire to run ancestor-simulations and are free to do so. If (3) is true, then we almost certainly live in a simulation. In the dark forest of our current ignorance, it seems sensible to apportion one's credence roughly evenly between (1), (2), and (3).

Unless we are now living in a simulation, our descendants will almost certainly never run an ancestor-simulation.

REFERENCE

1) Beane SR., Davoudi Z, Savage, MJ. 2012. "Constraints on the University as Numerical Simulation." arXiv: 1210.1847v2 [hep-ph] 9 Nov 2012; NT@UW-12-14 INT-PUB-12-046; http://arxiv.org/pdf/1210.1847v2.pdf.
2) Bostrom N. 2003. "Are You Living in a Computer Simulation?" Philosophical Quarterly. 53(211): 243-255.
3) Bostrom N. 2005a. "The Simulation Argument: Reply to Weatherson." The Philosophical Quarterly. 55(218): 90-97.
4) Bostrom, N. 2005b. "Why Make a Matrix? And Why You Might be in One." http://www.nickbostrom.com; In: Irwin, W. ed., More Matrix and Philosophy: Revolutions and Reloaded Decoded. Open Court.