

On the untenable simulation argument, or No, you are NOT living in a computer simulation

Thorvald Henriksson

Uppsala Universitet, Institutionen för informationsteknologi

Uppsala Science fiction club

In the beginning of the 2000s, famous Swedish transhumanist Nick Bostrom [Bostrom 2003] proposed and claimed to answer the following conjecture in the affirmative way.

From the next three propositions,

a. the humankind will be quite possibly extinct well before reaching the transhumanistic stage of evolution (also known as the Technological Singularity).

b. any posthuman (transhuman) civilization, for some reasons, keep the ancestral simulation of its own primordial (pre-Singular) history inappropriate for itself.

c. reality surrounding us *is* the virtual restoration —

*at least **one** is true.*

This conjecture became famous as the Bostrom simulation argument. Bostrom also considered it crucial for possible evaluation of the Fermi paradox and investigation of the possible nature of the Technological Singularity.

Of course, case 1) and 2) false and 3) true quickly became the most interested in by the cyberpunk fans, newspaper reporters and transhum movement activists, especially 'cause the first Bostrom's works in this field turned out (probably *not incidentally*) time-coincident with the Wachowskis' *Matrix* prominent block-buster trilogy. The next decade (I write this article in December 2012, just days before the "end of world") brought to life a lot of articles and (as a rule, SF-)works dedicated to Bostrom's concept in critical or supportive and even in-depth investigating way. Here I recall novels by Iain McDonald (*Brasyl*) and Ken MacLeod (*The Restoration Game*), Greg Egan's novellette (*Crystal Nights*) as well as recent award-winning *noir*-detective *Osama* by Lavie Tidhar. When looking on the earlier works consistent with Bostrom's conceptual framework, one could praise Haruki Murakami's cyberpunk novel *Sekai-no ovari to haadaboירוdo-wandaarando* (*Hardboiled Wonderland and the End of the World*).

Bostrom simulation argument can be rendered in roughly mathematical way as follows. Let me introduce the following notation:

f_{ph} — fraction of all human-level technological civilizations that survive to reach a posthuman stage and ability to reconstruct their own history in the limitless virtual sensorium environment,

f_{int} — the fraction of posthuman civilizations that are interested in running ancestral simulations (or that contain at least some individuals who are interested in that and have sufficient resources to run a significant number of such simulations),

N — average number of ancestral simulations run by a posthuman civilization,

N_{int} — average number of ancestral simulations run by such especially interested civilizations,

H — average number of individuals that have lived in such a civilization before it reaches a posthuman stage,

then

$$N = f_{int} \times N_{int}$$

So, according to Bostrom, actual fraction of all (human-like) inhabitants of the virtual simulation is as follows:

$$f_{sim} = \frac{f_{ph}NH}{H + f_{ph}NH} = \frac{f_{ph}f_{int}N_{int}}{1 + f_{ph}f_{int}N_{int}}$$

Bostrom followed Tipler in premise that computing power of posthuman civilizations is so immense that N_{int} may turn out almost infinitely big. With such premise he claimed to have formulated the core of the simulation argument: from the following three propositions 1) $f_{ph} \approx 0$, 2) $f_{int} \approx 0$, 3) $f_{sim} \approx 1$ at least one is true.

Naturally, we can interpret Bostrom's simulation argument (and that we can see from this last formulation) in such way that Reality surrounding us is actually Virtual fake run by some ultrapowerful posthuman civilization for uneven reasons, probably scientific or religious.

Let me say here that such claim is nowhere as new when compared to some Hinduistic myths (where all world is considered fragile creation of the one moment of god's dream) or even South-American Amazon basin Indians' myth described by Iain McDonald in *Brasyl*: the least claims that the Jaguar made the world, but not very well; and it ended on the third day and we — the world, everything we think is real — are just the dreams of the third night. But on the third day's morning the Jaguar will awake and start to, UNIX-like speaking, unmount disks mounted to root.

Bostrom's simulation argument is, of course, fruitful for art, especially for building plots, as we can see from aforementioned works. But, well, in the long shot it will render it senseless, even debacle-alike for a priori demiurgical truthfulness of author's cosmi. Intuitively, Bostrom's simulation argument seems scientifically uninteresting and had been criticized from such premises. Really, Bostrom's conjectures lead us to logically inevitable need to treat all laws of Nature, described by Science, in extremely critical, even solipsistic way. But such laws are *efficient* — many-times proven by practice. So, all our experience, all life-actions, all that what is called life in usual sense turns out at least questionable. Maybe mathematics remains, as a tool that is specially designed for describing not only physical, but even unphysical world, whereas Bostrom's propositions fill the word *unphysical* by true horror.

Analysis of publications related to subject (see, e.g., [Beshard 2004]) leads to conclusion that Bostrom's simulation argument is usually criticized on the ground of its rather wild implications in practical action context instead of the mathematical ground. Generally speaking, such method is not quite efficient. David Deutsch noted that evidence for principal realizability of the computing machine (calculator, computer, supercomputer etc.) cannot be found in mathematics or logic. The reason why we find it possible to construct, say, electronic calculators, and indeed why we can perform mental arithmetic, cannot be found in mathematics or logic. The reason is that the laws of physics *happen to* permit the existence of physical models for the operations of arithmetic such as addition, subtraction and multiplication. If they did not, wrote Deutsch, these familiar operations would be non-computable functions. We might still know of them and invoke them in mathematical proofs (non-constructive!), but we could not perform them [Deutsch 1985].

For example, below I'll mention Chaitin's constant; and it's non-computable [Chaitin 1977]. But, from the other hand, every digit of Chaitin's constant can, of course, be computed, just because of computer program which "prints" such digit always exists.

Bostrom argues us to use not by mathematics but by metamathematics, as well as states, without hard support, that computing machines used in everyday work and scientific computations actually compute unreal results, unreal in sense that they might IN REAL produce some another results. Why? For whom? These questions were (and are) never answered by Bostrom. He proposes futile process of comparing one unphysical structure with other, slightly less unphysical and, on some "true" level, possessing overLord-ing validity. Such level, according to him, is inhabited by the "System Architects" of our Universe and their "team leads".

From the mathematical point of view, such problem is equivalent to words comparison problem, while the latter one — to Gödel theorem.

When speaking in rather imaginative language, one could recall Zeno's paradoxes and Ming-jia Chinese philosophers' conjectures equivalent to them. They were interested by problems of infinite division and describing a set by its properties.

Say, Gongsuen Long (Kung-sun Lung), Chinese philosopher, rejected the equivalence of notions with intuitively identical content and different descriptions ("a white horse is not a horse", he often said). Interestingly, he paid customs for his white horse as for the one when being stopped in travel on the frontier post [Hansen 1976, Reiman 1980–1981]!

In principle, aforementioned is sufficient for fustigation of Bostrom's claims. But from thence I'll try to mark the method for critical investigation of Bostrom's simulation argument on the ground of quantum informatics. I didn't find completely analogous, when speaking of argumentative framework used, articles or discussions in the world's infosphere. If even I reinvent the wheel, in what I don't have faith, I hope, *in good faith*, to enrich notional kit connected to problem stated by two new thought experiments. Below I'll discuss them in some detail.

Imagine morning on Earth and look at a flat in Stockholm high-rise house, where after dark, stormy and challenging night the girl, named Ingrid, and her boyfriend are sleeping. The boyfriend's name is irrelevant for our thought experiment, but what about the girl? Her name reminds us of Ingrid Bergman, one of best of the best actresses in the cinema history, as well as of the simulational-computational matrix (in-grid): here we acquiesce it all is a Bostrom-designed fake.

Look: Ingrid and his boyfriend awake, think for some time of who knows what, and then Ingrid, being responsible and accustomed to wake up early, stands up and goes for a coffee. Leaning to the wall, she charges Jura Impressa X7 with freshly grinded coffee, pours milk to cappuccinator, places coffee cup under the spout and waits. For a few moments Ingrid falls out of the simulated reality and enters simulation within simulation called the somnolence. During these moments his boyfriend is leaving bed and making his way to kitchen too.

Ingrid rouses herself wide awake when hearing his steps, sees coffee's ready, stretches hand for the cup, but 'cause of uneven coordination of movements after the doze she burns her fingers, jerks and throws the cup down. The cup falls on the floor and breaks. Ingrid's boyfriend materializes at kitchen's door the moment Ingrid reaches for the cup and fails to stop her (well, how could he know it was necessary to stop her girlfriend?)

The situation aforescribed is verisimilar, easy to imagine in ordinary life, so it passes basic Bostrom ancestral simulator restoration veracity test.

Let's leave Ingrid or (what might be more courteous) her boyfriend to absterge the coffee puddle and clean out shards of broken cup whilst I translate these events to software code language. Speaking in terms of intra-Bostrom ancestral simulator process interacting, we may say that participants of the thought experiments become processes having in possess pointers for segments of shared memory space, so while one process (Ingrid) creates the segment, the second (her boyfriend) gains access to it.

In the first moment of imitated (according to Bostrom's opinion) time (although such Bostrom's concept is not so constructive too, because of, as Feynman noted, time is always imitated, not modelled, in computing simulations) *code string*, which defines expected time of coffee preparation as well as its temperature for Ingrid, contained prognostic evaluation of these parameters carried upon by the simulated girl's mind. But in the second moment, when that code string was reached by process representing Ingrid's boyfriend's simulated mind, limiting number has been written there already: Ingrid's mind detected the true temperature of coffee and jerked her hand, throwing down the poor cup. Ingrid's boyfriend's mind *failed* to undo this reaction. For want of a nail, cup is broken, shards are scattered, coffee is leaked out.

Thus we may say that process, which is entangled with Ingrid's boyfriend, detects this limit *after* running all code fragment.

Let's imagine the same program run repeats: Ingrid's boyfriend "prepares" the second thought experiment and asks her blue not and gain new coffee instead of coffee from the broken cup (see Zen ko'an ersatz, yeah?)

The situation aforescribed is verisimilar, easy to imagine in ordinary life, so it passes basic Bostrom ancestral simulator restoration veracity test.

But, evidently, this run is unlike previous: neither Ingrid prepares new portion of coffee, nor she tastes coffee prepared by her boyfriend till the time pain relieves, shards are cleaned off and coffee puddle absterged. Also possible she won't drink any coffee further this morning in Stockholm. The reader may say she might turn out slut suffering from analgesia, but *I* refuse to run such version of thought experiment.

Beyond questions, boyfriend's proposition is clever and correct according to Bostrom ancestral simulator validity criteria, but nonetheless, Ingrid's very probable refusal seems expectable and understandable, when it comes to human emotions.

For what reason the program gives us different results? The participants of thought experiments keep for themselves shared access to memory segment; almighty ruling software didn't erase their privileges, didn't delete this segment at all, but only displaced it, transposed to imitated past. Ingrid and her boyfriend undoubtedly remember how the girl stretched hand for the cup too early, how the cup turned out so hot that shit hits all the morning after-intercourse fun. The limit stated by coffee'n'code segment is now detected by processes called *Ingrid* and *Ingrid's boyfriend* at the early stage of run, well before entering the conditional loop.

The event described looks like casual and common, but seems catastrophic exception from the point of test and evaluation of historical simulator's developers' work. Apparently they may — MUST, as we'll see later — evaluate a solution, delete the affected segment before termination of current run and free locked computational space (although almost unlimited: Bostrom clearly states that each of colossal stack of Restoration Virtualities contains myriad of participants and actors; every action is different, each in myriad ways). To organize correct, from the restorator's human or posthuman point of view, interaction of processes called *Ingrid* and *Ingrid's boyfriend*, one could insert semaphores or mutexes governing access to shared space to read and write, semwait and sempost. As early as in Edsger Dijkstra's works [Dijkstra et al., 1972] semaphore approach was proposed (probeer te verlagen and verhogen semaphore operations, in Dijkstra's notation). Thus, changes in enclosed segment may be caught, corrected and undone.

Please imagine virtual girl and her virtual partner in the virtual kitchen of the virtual flat of the virtual house of the virtual Stockholm on the surface of the virtual planet, simulating orbital rotation round simulated star in the simulated Universe. Can the System Architects, who design, compile and run all that fuzz, be really so stupid to think a priori that pointer-centered actions in the simulator's memory space will be always atomic in computational nature? If rejecting the Architects and taking into their place some Wolfram-Friedkin's cellular automaton, nothing changes at all. Indeed, the Architects can also be simulated, according to *reductio ad absurdum* over Bostrom's ideas.

But even after insertion and activation of such semaphores Ingrid coffee machine paradox stays possible, *conceivable*, imaginable. What does that mean? Even if semaphores are caught, all work is for nothing, without purpose, and the Architects' objection will be overruled by free will and chaos at end. When UNIX-language speaking, software code (probably) used in the thread signal masks processing nonetheless returns not that one mask what the Architects need for optimizing a termination of the signal handler conjugated with Ingrid's mind, but some different mask.

Whether it exists, reason for such stupidity?

Or not?

How could it exist at all?

How could the Architects *know*, where it is need to position semaphores?

Bostrom, unlike me, doesn't interest in such questions.

Let me codify the crucial question. How could at all almighty program, which is necessarily designed to imitate historically evidenced autonomy of thinking in various individual inhabitants of ancestral simulation, *know* that prognostic evaluation of coffee cup's temperature by Ingrid's mind must differ from the one according to external Virtual Reality scope?

This event is non-algorithmic at all. Even from the pure mathematics, its probability cannot be computed: the previously-mentioned problem is equivalent to *Entscheidungsproblem*, and, as a matter of fact, probability for Ingrid to burn her fingers when touching the coffee cup corresponds to one (and we cannot know which) member of the infinite Chaitin's constants' regression thread.

As I noted above, every digit of Chaitin's constant can, of course, be computed, just because of computer program which "prints" such digit always exists, but one couldn't compute even one constant of that regression as a whole. If transcending to high-enclosed system "Ingrid is thinking about Ingrid preparing the morning coffee, touching the coffee cup under the coffee machine spout and jerking hand 'cause of burn and breaking the cup...", we might use this argument recursively.

Additionally, I'd like to underline that *Ingrid-prepares-the-coffee* plot, non-algorithmic in general, and so unfeasible in the simulation, must be synthesized by ruling software inside the pre-simulated Ingrid's mind and brain, giving birth to yet another abstraction layer. Indeed, what exactly must be simplified for processing the model of Ingrid's model within the Ingrid's model? Taking into account new metaprogramming language characterizable by unique set of Chaitin's constants, one would always expect a thing that could make matters even worse. Mutually enclosed Virtual Realities, as Zeno or Gongsuen Long would predict, will gradually slow till the complete halt of infinite division, something probably alike DLL hell.

Thus, proposition about simulated nature of Ingrid's fingers' burn becomes excession, so I reject it stating that Ingrid jerks her hand in the Real.

By the way, Earth is statistically undistinguished from the rest of the Universe, so it seems Virtual nature of surrounding world's true nature is very unlikely to reveal just here, in neither starting nor final sentence of the universal Book of Truth. From this point of view, Bostrom simulation argument is nowhere as better than Alain de Lille's famous phrase — "God is an intelligible sphere whose centre is everywhere and whose circumference is nowhere".

As is to be naturally expected, Bostrom ancestral simulator is a universal quantum computer (UQC), or \mathcal{Q} , in usual notation of quantum informatics. Let me recall its definition according to Deutsch-Bennett-Turing model [Deutsch 1985]. The state of \mathcal{Q} is a unit vector in the Hamiltonian space \mathcal{H} spanned by the computational basis states, and the dynamics of UQC are summarized by an unitary operator U on \mathcal{H} . Operator U obeys to Schrödinger evolution equations

$$|\Psi(nT)\rangle = U^n |\Psi(0)\rangle \quad (n \in \mathbb{Z}^+)$$

$$U^\dagger U = U U^\dagger = \hat{1}.$$

Here I use standard notation.

The computation begins from an empty state (in our Universe, it may be the moment of Big Bang as well as any other moment of time, because of, according to Bostrom, memory connected to the previous moments can be easily implanted *after* the start by the almighty ruling program), $t = 0$.

Evolution of the UQC memory's registers is as follows:

$$|\Psi(0)\rangle = \sum_m \lambda_m |0; 0; \mathbf{m}\rangle$$

$$\sum_m |\lambda_m|^2 = 1.$$

Here $\lambda_m = 0$, whenever an infinite number of the memory's elements, \mathbf{m} , adopts non-zero values.

UQC is called universal and can be treated as functioning computer in the sense that only a finite-dimensional unitary evolution transformation need be effected at every step, whereas a quantum computer has an infinite-dimensional state space by definition. It allows UQC to simulate any other quantum computer $\mathcal{Q}[\mathcal{U}^+, \mathcal{U}^-]$, with arbitrary precision needed.

It may seem (not mention problems of decoherence and possible Penrose-type objective collapse) that such property possessed by UQC makes it natural candidate for almighty ruling machine of the ancestral simulator; and in *Brasyl* twice-mentioned Iain McDonald utilizes just that, calling UQC *Parousia*. But here I must state that infinite hierarchy of Bostrom's mutually enclosed Virtual Realities does require infinite-dimensional evolution transformation!

Let me recall now that some physical experiments seem to indicate existence of granulous, foam-like spacetime structure at Planck scale, i.e., existence of some limit of length and distance (ℓ_{Pl}), at that and (theoretically speaking) below that Lorentz invariance breaks. By the way, human mind's possibility to imagine pre-Planck scale, at least describe it in mathematical form of inequalities or dimensional correlations, if not to visualize, tends to indicate by itself that we are living in the realms of the Real.

Bostrom, to say, attributes such data to Plato-like play of shadows run by the Architects.

From the Ingrid coffee machine paradox, we have already seen that ancestral simulation process is recursive. Deutsch showed that no recursive function can be computed by UQC on average more rapidly with the help of quantum software than with classical one, and established that UQC quantum parallelism cannot be used to improve the mean running time of parallelizable algorithms. It means that, for the Virtual technologies and time-compression ratios involved between the simulation and inside world (which, in turn, can by itself turn out the simulated world), infinitely craftful, clockwork-ish time scale adjustment would be impossible.

Additionally, one could recall that some actions of human mind very probably possess a non-algorithmic nature (as in [Penrose 1999] showed). Even if one cannot calculate the probability to be simulated in some particular place at some particular time, one may ask questions about the most probable observational results under the condition that one is simulated there.

I am born in Uppsala, and if everybody around speaks Swedish, it is not surprising. But if I find that everybody around me speaks Russian, I will be really surprised, and I will try to come with some explanation. Here I cite and paraphrase [Linde 2007] to show again that Virtual ancestral simulation machinery must, so say, leak and fail all around us to be consistent with our *ordinary* life, let alone extraordinary events.

Such incompetence I reckon imbecile.

Even more serious objection against Bostrom's argument can be preferred from the UQC (i.e., universal simulator)'s \mathcal{Q} -logical depth notion. Deutsch and Bennett [Bennett 1983, Deutsch 1985] define \mathcal{Q} -logical depth of a quantum state as the running time of the shortest quantum program that would generate the state (of ancestral simulation, I assume) from a blank input (or, perhaps, without losing generality one could use the harmonic mean of the running times of the ensemble of such interacting programs). This measure is immanent for each "world" of that ensemble and caused, in Bostrom's terms, by the deep (or beyonder, if you prefer such way of motion) Architects' actions.

When Bostrom's ancestral simulation is running, *matrioshka* of enclosed Virtual Realities needs, for proper functioning, its \mathcal{Q} -logical depth to increase monotonously. Standard C language

function longjmp is in some aspects analogous to such feature.

But, according to Bennett, the arrow of time, at least psychological, even better epistemological and evolutionary, does necessarily require the Q -logical depth of the Universe is at a global minimum, at least non-decreasing, from the start! Thus, Bostrom ancestral simulator must simultaneously possess ever-increasing AND globally minimal Q -logical depth. In the aforescribed Ingrid coffee machine paradox, two subsystems of the virtual environment (one for coffee machine and cup, as well as another, for Ingrid's mind) represent regions with different time-trends of Q -logical depth's changing. Hence, we come to contradiction.

Again even more, and even in our Reality, Q -logical depth as a property of a quantum state is not necessarily well-defined for quantum subsystem (and this is expected, say, when all the information about the system, which is accessible to observer, reside and transfers entirely in and by trans-systemic correlations).

The simplest example of this is quantum Zeno effect (here Q -logical depth needed is Q -logical depth of the state from which decay products are born), more prominent is quantum cryptography (like the time-delayed choice quantum eraser experiment, where retrocausality problems, totally ignored by Bostrom, are also involved).

It is often said that Bostrom-type UQC yet can be achieved in the most latter stage of evolution of the Universe, as expanding gives place to contraction; such hypothetical epoch is known as the Big Crunch. Whether such scenario really possible or not, we can't verify by now. Recent Higgs boson discovery (which, of course, needs be verified too), as well as cosmostatistical data concerned the Universe-dominating matter and dark matter's energy density, seem to exclude such possibility. The Universe will expand forever, with some acceleration. Yet let's try to investigate whether Bostrom-type UQC computations meet necessary criteria in the crunching Universe or not. Lloyd in [Lloyd 2001] showed that the total number of operations that theoretically can be performed by UQC in the Universe some 700Ky after the Big Bang within a co-moving volume is

$$\Phi = (2E/\pi\hbar) \int_{\tau_0}^{\tau_1} \sqrt{\tau_1/\tau} d\tau = (4E/\pi\hbar)(\tau_1 - \sqrt{\tau_1\tau_0})$$

Here E is the total energy available for computation in the radiation-dominated Universe at time τ_1 . Note that Lloyd equation turns out symmetrical versus time-inversion and, hence, may be addressed to the Big Crunch's epoch when a closed universe contracts to a singularity. From this, we see that only finite number of computational operations can be performed in a Bostrom ancestral simulator. To overrule this objection we need the amount of energy of the Universe goes to infinity as $1/\tau$ or even faster, as the Universe returns to primordial singularity. It seems that quantum gravity neglects such possibility: naïve surmises upon accumulation of all energy in the Universe for purposes of creation of the Universal "Omega point" computer quickly fail.

Yet another interesting possibility remains. Inflationary expanding of the Universe may catalyze its rip to causally-unconnected sectors that would have a much higher free energy and bit capacity than connected system; such realms would be separated by event horizons as long as the inflation itself persists. In principle, process described, known as the Big Rip, can catalyze the sundering of the Universe to ultimate ensemble of Bostrom simulators, Leibniz monads-like to some extent. Such plot was utilized — again — by Iain McDonald in *The Tear*.

But even under such conditions infinite increase of the Q -logical depth remains unrealistic.

The maximal energy density and bit capacity are restricted on Planck scale: so-called holographic principle allows us to store only $\approx c^2/H^2\ell_{Pl}^2$ within the horizon which area is c^2/H^2 , where H is the Hubble parameter. Even if the latter parameter changes in time due to background scalar field evolution, corresponding change of scale factors of the expanding of the Universe will be indemnified, in the informational sense, for the horizon energy density, acting as Planck firewall: communication between causally separated regions will render impossible.

Donald Kingsbury notes [Kingsbury 2001] that Tipler-type ancestral simulator requires infinite information storing capacity within volumes less than ℓ_{Pl}^3 ; in general, his remark is equivalent to conclusions from the Lloyd equation.

Interestingly, total number of bits available for Bostrom-type UQC is *at the same scale* as the number of bits (10^{90}) available in the modern, expanding Universe, if every degree of freedom of every particle, even including gravitational ones, would be registered. It can be calculated from the conventional black-body model and the maximum entropy condition:

$$\mathcal{J} = (4/3 \ln 2) \Phi^{3/4}$$

as long as the Grand Unification threshold, of 2×10^{16} GeV, is not achieved.

So when looking at nether realms of the contracting, massless radiation-dominated Universe, we don't gain any specific computational advantages! The limit factor of the computational power's increase is somewhere near 2 (more strictly, 1.92). Of course, that is absolutely insufficient for Tipler and Bostrom's arrogant plans.

In this quick and short note I can't review some other interesting alternatives to the simulation argument, for which Fermi paradox and transition to the Technological Singularity are also points of interest. By the way, the possibility of such transition itself is negated by Bostrom due to the Architects' disagreement. But I can't omit one especially breakthrough concept formulated by Vernor Vinge, who himself coined the term Technological Singularity. Vinge described the Universe where speed of information propagation and processing depends from observer's position in cosmi. Of course, even theoretical conversion of the Universe as a whole to Bostrom-type UQC tends to impossible according to Zones of Thought [Vinge 1991–2011] model by Vinge.

Acknowledgements

It's my pleasure to thank Anita Linden, Ulrika Folkesson and Kevin Lindström for many fruitful discussions.

In any topic touching upon this article please contact me via e-mail:

thorvald.henriksson@yahoo.com.

References

- Dijkstra et al., 1972 — E. W. Dijkstra, O.-J. Dahl, C. A. R. Hoare, *Structured Programming*, 1972.
Hansen 1976 — C. D. Hansen, *Philos. East and West*, **26**, 1976/2.
Chaitin 1977 — G. J. Chaitin, *IBM J. Res. Dev.*, **21**, 350.
Reiman 1980 — F. Reiman, *ibid.*, **30**, 1980/3.
Reiman 1981 — F. Reiman, *ibid.*, **31**, 1981/4.
Bennett 1983 — C. H. Bennett, *On various measures of complexity, especially "logical depth"*, Aspen IBM Conf. Lect. Report, 1983 (later reprinted).
Deutsch 1985 — D. Deutsch, *Proc. Roy. Soc. Lond. A*, **400**, 97.
Vinge 1991–2011 — V. Vinge, *A Fire Upon the Deep*, 1991. *A Deepness in the Sky*, 1999. *The Children of the Sky*, 2011.
Penrose 1999 — R. Penrose, *The Emperor's New Mind*, 1999.
Kingsbury 2001 — D. Kingsbury, *Psychohistorical Crisis*, 2001.
Lloyd 2001 — S. Lloyd, arXiv:quant-ph/0110141v1, 2001.
Bostrom 2003 — N. Bostrom, *Phil. Quart.*, **53** (2003), 211, 243.
Beshard 2004 — F. Beshard, *Refutations of the Simulation Argument*, 2004.
Linde 2007 — A. Linde, arXiv:hep-th/0611043v3, 2007.