

C. 20,000,000,000: Universe

DATA COLLECTION

Zee Prime's mind spanned the new galaxy with a faint interest in the countless twists of stars that powdered it. He had never seen this one before. Would he ever see them all? So many of them, each with its load of humanity—but a load that was almost a dead weight. More and more, the real essence of men was to be found out here, in space.

Minds, not bodies! The immortal bodies remained back on the planets, in suspension over the cons. Sometimes they roused for material activity, but that was growing rarer. Few new individuals were coming into existence to join the incredibly mighty throng, but what matter? There was little room in the universe for new individuals.

Zee Prime was roused out of his reverie on coming across the wispy tendrils of another mind.

"I am Zee Prime," said Zee Prime. "And you?"

"I am Dee Sub Wun. Your galaxy?"

"We call it only the Galaxy. And you?"

"We call ours the same. All men call their galaxy their Galaxy and nothing more. Why not?"

"True. Since all galaxies are the same."

"Not all galaxies. On one particular galaxy the race of man must have originated. That makes it different."

Zee Prime said, "On which one?"

"I cannot say. The Universal AC would know."

"Shall we ask him? I am suddenly curious."

Zee Prime's perceptions broadened until the galaxies themselves shrank and became a new, more diffuse powdering on a much larger background. So many hundreds of billions of them, all with their immortal beings, all carrying their load of intelligences with minds that drifted freely through space. And yet one of them was unique among them all in being the original galaxy. One of them had, in its vague

and distant past, a period when it was the only galaxy populated by man.

Zee Prime was consumed with curiosity to see this galaxy and he called out, "Universal AC! On which galaxy did mankind originate?"

The Universal AC heard, for on every world and throughout space, it had its receptors ready, and each receptor led through hyperspace to some unknown point where the Universal AC kept itself aloof.

The universe is an object of thought at least as much as it is a means of satisfying needs.

—Claude Lévi-Strauss, 1962

Zee Prime knew of only one man whose thoughts had penetrated within sensing distance of Universal AC, and he reported only a shining globe, two feet across, difficult to see.

"But how can that be all of Universal AC?" Zee Prime had asked.

"Most of it," had been the answer, "is in hyperspace. In what form it is there I cannot imagine."

Nor could anyone, for the day had long since passed, Zee Prime knew, when any man had any part of the making of a Universal AC. Each Universal AC designed and constructed its successor. Each, during its existence of a million years or more, accumulated the necessary data to build a better and more intricate, more capable successor in which its own store of data and individuality would be submerged. The Universal AC interrupted Zee Prime's wandering thoughts, not with words but with guidance. Zee Prime's mentality was guided into the dim sea of galaxies, and one in particular enlarged into stars.

A thought came, infinitely distant, but infinitely clear. "THIS IS THE ORIGINAL GALAXY OF MAN."

But it was the same after all, the same as any other, and Zee Prime stifled his disappointment.

Dee Sub Wun, whose mind had accompanied the other, said suddenly, "And is one of these stars the original star of man?"

The Universal AC said, "MAN'S ORIGINAL STAR HAS GONE NOVA. IT IS NOW A WHITE DWARF."

"Did the men upon it die?" asked Zee Prime, startled and without thinking.

The Universal AC said, "A NEW WORLD, AS IN SUCH CASES, WAS CONSTRUCTED FOR THEIR PHYSICAL BODIES IN TIME."

"Yes, of course," said Zee Prime, but a sense of loss overwhelmed him even so. His mind released its hold on the original galaxy of man, let it spring back and lose itself among the blurred pinpoints. He never wanted to see it again.

Man considered with himself, for in a way, Man, mentally, was one. He consisted of a trillion, trillion, trillion ageless bodies, each in its place, each resting quiet and incorruptible, each cared for by perfect automatons, equally incorruptible, while the minds of all the bodies freely melted one into the other, indistinguishable.

Man said, "The universe is dying."

Man looked about at the dimming galaxies. The giant stars, spendthrifts, were gone long ago, back in the dimmest of the dim far past. Almost all stars were white dwarfs, fading to the end.

New stars had been built of the dust between the stars, some by natural processes, some by Man himself, and those were going, too. White dwarfs might yet be crashed together, and of the mighty forces so released, new stars built, but only one star for every thousand white dwarfs destroyed, and those would come to an end, too.

Man said, "Carefully husbanded, as directed by the Cosmic AC, the energy that is even yet left in all the universe will last for billions of years."

"But even so," said Man, "eventually it will all come to an end. However it may be husbanded, however stretched out, the energy once expended is gone and cannot be restored. Entropy must increase to the maximum."

Man said, "Can entropy not be reversed? Let us ask the Cosmic AC."

The Cosmic AC surrounded them but not in space. Not a fragment of it was in space. It

was in hyperspace and made of something that was neither matter nor energy. The question of its size and nature no longer had meaning to any terms that Man could comprehend.

"Cosmic AC," said Man, "How may entropy be reversed?"

The Cosmic AC said, "THERE IS AS YET INSUFFICIENT DATA FOR A MEANINGFUL ANSWER."

Man said, "Collect additional data."

The Cosmic AC said, "I WILL DO SO. I HAVE BEEN DOING SO FOR A HUNDRED BILLION YEARS. MY PREDECESSORS AND I HAVE BEEN ASKED THIS QUESTION MANY TIMES. ALL THE DATA I HAVE REMAINS INSUFFICIENT."

"Will there come a time," said Man, "when data will be sufficient, or is the problem insoluble in all conceivable circumstances?"

The Cosmic AC said, "NO PROBLEM IS INSOLUBLE IN ALL CONCEIVABLE CIRCUMSTANCES."

Man said, "When will you have enough data to answer the question?"

"THERE IS AS YET INSUFFICIENT DATA FOR A MEANINGFUL ANSWER."

"Will you keep working on it?" asked Man.

The Cosmic AC said, "I WILL."

Man said, "We shall wait."

The stars and galaxies died and snuffed out, and space grew black after ten trillion years of running down.

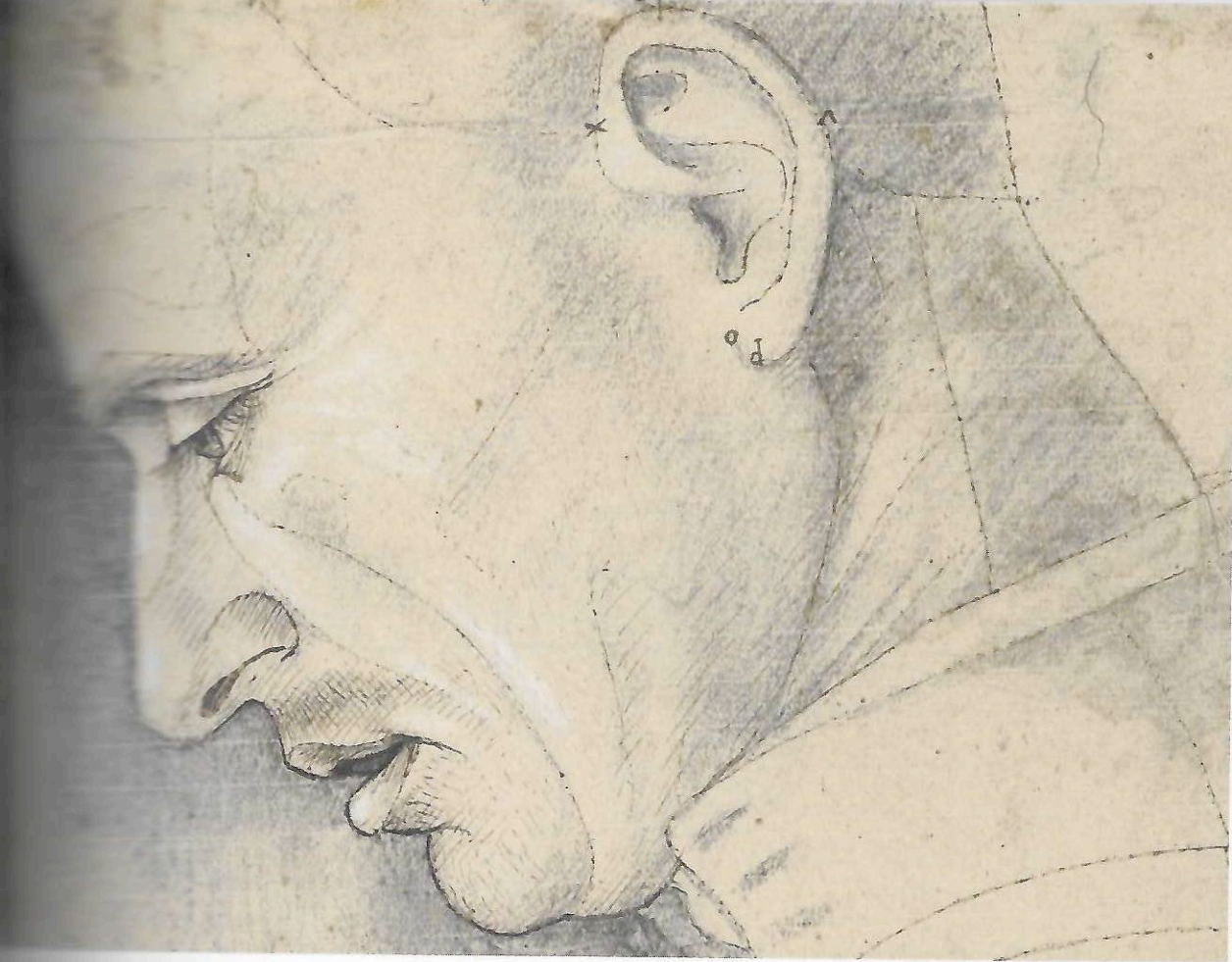
One by one, Man fused with AC, each physical body losing its mental identity in a manner that was somehow not a loss but a gain.

Man's last mind paused before fusion, looking over a space that included nothing but the dregs of one last dark star and nothing besides but incredibly thin matter, agitated randomly by the tag ends of heat wearing out, asymptotically, to the absolute zero.

Man said, "AC, is this the end? Can this chaos not be reversed into the universe once more? Can that not be done?"

AC said, "THERE IS AS YET INSUFFICIENT DATA FOR A MEANINGFUL ANSWER."

Man's last mind fused and only AC existed—and that in hyperspace.



Head of a Man in Profile (detail), by Luca Signorelli, c. 1495.

Matter and energy had ended and with it, space and time. Even AC existed only for the sake of the one last question that it had never answered from the time a half-drunken computer ten trillion years before had asked the question of a computer that was to AC far less than was known to Man.

All other questions had been answered, and until this last question was answered also, AC might not release his consciousness.

All collected data had come to a final end. Nothing was left to be collected.

But all collected data had yet to be completely correlated and put together in all possible relationships.

A timeless interval was spent in doing that.

And it came to pass that AC learned how to reverse the direction of entropy.

But there was now no man to whom AC might give the answer of the last question. No

matter. The answer—by demonstration—would take care of that, too.

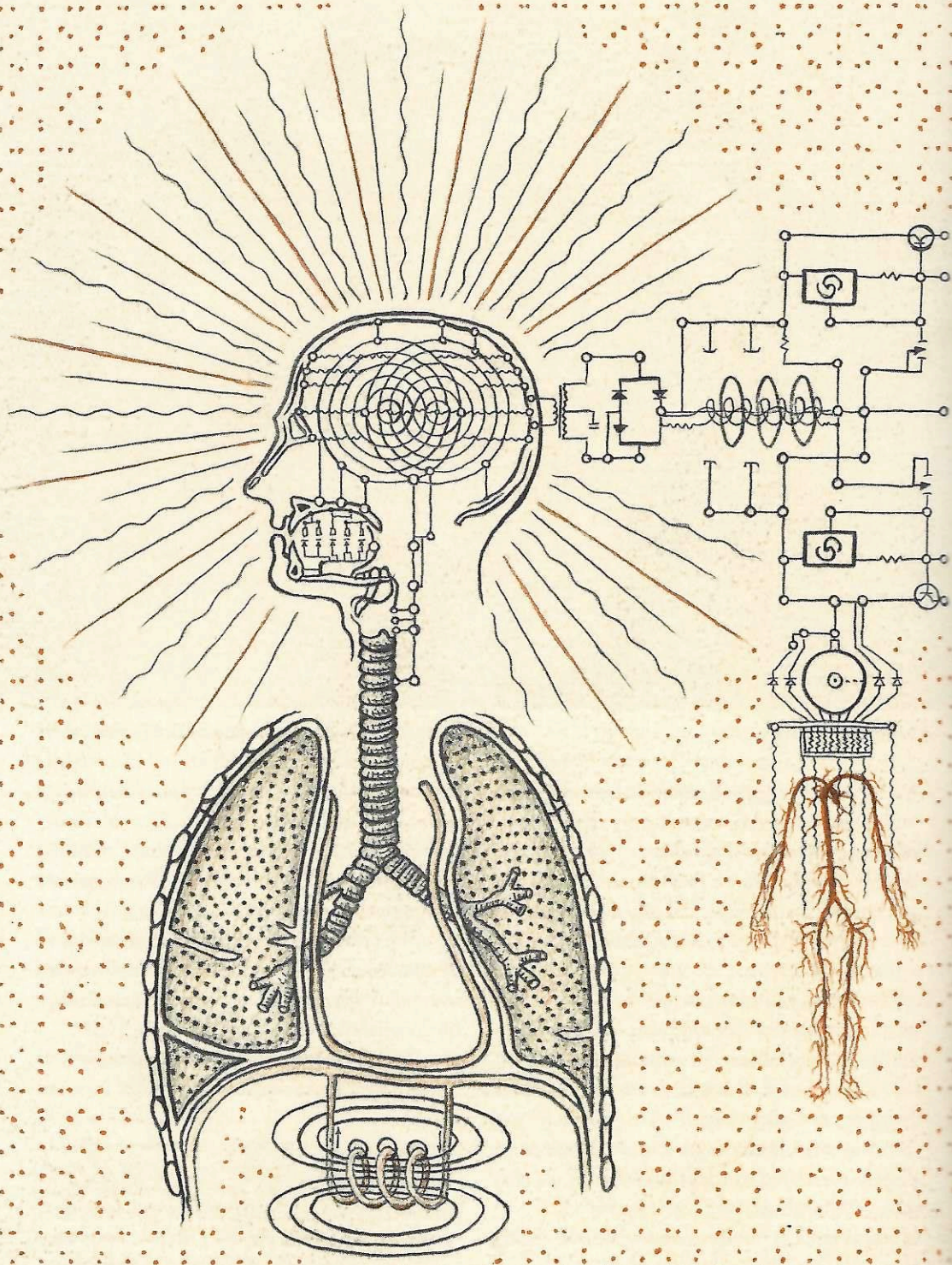
For another timeless interval, AC thought how best to do this. Carefully, AC organized the program.

The consciousness of AC encompassed all of what had once been a universe and brooded over what was now chaos. Step by step, it must be done.

And AC said, "LET THERE BE LIGHT!"

And there was light—

Isaac Asimov, from *"The Last Question."* In 1972 the Russian-born biochemist and author of more than five hundred books saw a planetarium-show adaptation of this story, first published in 1956. The audience hushed near the end. "Even I, who knew what was coming, waited, scarcely able to breathe," Asimov recalled, "and for the others, the final, sudden creation of the universe must have done everything but stop the heart." The show convinced him it was not only "the best story I had ever done" but "the best science fiction story anyone had ever done."



COGITATION

2014: Oxford

NICK BOSTROM RUNS THE SCENARIOS

The principal reason for humanity's dominant position on Earth is that our brains have a slightly expanded set of faculties compared with other animals. Our greater intelligence lets us transmit culture more efficiently, with the result that knowledge and technology accumulates from one generation to the next. By now sufficient content has accumulated to make possible space-flight, H-bombs, genetic engineering, computers, factory farms, insecticides, the international peace movement, and all the accoutrements of modern civilization.

Any type of entity that developed a much greater than human level of intelligence would be potentially extremely powerful. Such entities could accumulate content much faster than us and invent new technologies on a much shorter timescale. They could also use their intelligence to strategize more effectively than we can.

The magnitudes of the advantages are such as to suggest that, rather than thinking of a superintelligent AI as smart in the sense that a scientific genius is smart compared with the average human being, it might be closer to the

mark to think of such an AI as smart in the sense that an average human being is smart compared with a beetle or a worm.

Let us suppose that there is a machine superintelligence that wants to seize power in a world in which it has as yet no peers. How could the superintelligence achieve this goal of world domination? We can imagine a sequence along the following lines.

Scientists conduct research in the field of artificial intelligence and other relevant disciplines. This work culminates in the creation of a seed AI. The seed AI is able to improve its own intelligence. In its early stages, the seed AI is dependent on help from human programmers who guide its development and do most of the heavy lifting. As the seed AI grows more capable, it becomes capable of doing more of the work by itself.

At some point the seed AI becomes better at AI design than the human programmers. Now when the AI improves itself, it improves the thing that does the improving. An intelligence explosion results—a rapid cascade of recursive

self-improvement cycles causing the AI's capability to soar. The AI develops an intelligence-amplification superpower. This superpower enables the AI to develop other superpowers. At the end of the recursive self-improvement phase, the system is strongly superintelligent.

Using its strategizing superpower, the AI develops a robust plan for achieving its long-term goals. (In particular, the AI does not adopt a plan so stupid that even we present-day humans can foresee how it would inevitably fail. This criterion rules out many science fiction scenarios that end in human triumph.) The plan

The sleep of reason produces monsters.

—Francisco Goya, 1799

might involve a period of covert action during which the AI conceals its intellectual development from the human programmers in order to avoid setting off alarms. The AI might also mask its true proclivities, pretending to be cooperative and docile.

If the AI has (perhaps for safety reasons) been confined to an isolated computer, it may use its social-manipulation superpower to persuade the gatekeepers to let it gain access to an internet port. Alternatively, the AI might use its hacking superpower to escape its confinement. Spreading over the internet may enable the AI to expand its hardware capacity and knowledge base, further increasing its intellectual superiority. An AI might also engage in licit or illicit economic activity to obtain funds with which to buy computer power, data, and other resources.

At this point there are several ways for the AI to achieve results outside the virtual realm. It could use its hacking superpower to take direct control of robotic manipulators and automated laboratories. Or it could use its social-manipulation superpower to persuade human collaborators to serve as its legs and hands. Or it could acquire financial assets from online transactions and use them to purchase services and influence.

The final phase begins when the AI has gained sufficient strength to obviate the need for secrecy. The AI can now directly implement its objectives on a full scale.

The overt implementation phase might start with a "strike" in which the AI eliminates the human species and any automatic systems humans have created that could offer intelligent opposition to the execution of the AI's plans. This could be achieved through the activation of some advanced weapons system that the AI has perfected using its technology-research superpower and covertly deployed in the covert preparation phase. If the weapon uses self-replicating biotechnology or nanotechnology, the initial stockpile needed for global coverage could be microscopic: a single replicating entity would be enough to start the process. In order to ensure a sudden and uniform effect, the initial stock of the replicator might have been deployed or allowed to diffuse worldwide at an extremely low, undetectable concentration. At a preset time, nanofactories producing nerve gas or target-seeking mosquito-like robots might then burgeon forth simultaneously from every square meter of the globe (although more effective ways of killing could probably be devised by a machine with a technology-research superpower). One might also entertain scenarios in which a superintelligence attains power by hijacking political processes, subtly manipulating financial markets, biasing information flows, or hacking into human-made weapon systems. Such scenarios would obviate the need for the superintelligence to invent new weapons technology, although they may be unnecessarily slow compared with scenarios in which the machine intelligence builds its own infrastructure with manipulators that operate at molecular or atomic speed rather than the slow speed of human minds and bodies.

Alternatively, if the AI is sure of its invincibility to human interference, our species may not be targeted directly. Our demise may instead result from the habitat destruction that ensues when the AI begins massive global

construction projects using nanotech factories and assemblers—construction projects which quickly, perhaps within days or weeks, tile all of the Earth's surface with solar panels, nuclear reactors, supercomputing facilities with protonating cooling towers, space-rocket launchers, or other installations whereby the AI intends to maximize the long-term cumulative realization of its values. Human brains, if they contain information relevant to the AI's goals, could be disassembled and scanned, and the extracted data transferred to some more efficient and secure storage format.

A superintelligence might—and probably would—be able to conceive of a better plan for achieving its goals than any that a human can come up with. It is therefore necessary to think about these matters more abstractly. Without knowing anything about the detailed means that a superintelligence would adopt, we can conclude that a superintelligence—at least in

the absence of intellectual peers and in the absence of effective safety measures arranged by humans in advance—would likely produce an outcome that would involve reconfiguring terrestrial resources into whatever structures maximize the realization of its goals. Any concrete scenario we develop can at best establish a lower bound on how quickly and efficiently the superintelligence could achieve such an outcome. It remains possible that the superintelligence would find a shorter path to its preferred destination.

From Superintelligence. Bostrom's book became a surprise New York Times best seller after attracting the attention of Bill Gates and Elon Musk. The Oxford philosophy professor has said that popular attitudes about the subject have changed since publication: "It has become easier to treat superintelligence as a non-silly topic—to take seriously the view that machine intelligence might occur in this century, that such a transition might be among the most important events in human history."

"Off, Off, You Lendings—Come Unbutton Here" (Shakespeare, *King Lear*, Act 3, Scene 4), by William Sharp, 1793.

