

Lorenzo Perilli

**Artificial Consciousness**

**How Machine Think About and Transform Human Experience**



Machines can think? This question doesn't concern only the future of technology but also the profound meaning of what we define as human. Lorenzo Perilli doesn't limit himself to exploring what AI could become but examines how its existence is already redefining our understanding of consciousness and reality itself. *Artificial consciousness* examines the origins of the idea of consciousness, from ancient philosophy to modern science. Through the history of thinking machines, Perilli explores the boundaries between simulation and awareness, between learning and comprehension. Key concepts such as "imitation intelligence" and the "Turing test" are analysed, along with the neural networks and algorithms that today learn, predict, and operate. But is a system that responds in a complex way also aware of doing so? Or is consciousness something more, something that science still isn't able to capture? This book is a journey into the uncertain

confines between human and artificial, where consciousness may cease to be exclusively human. If memory, intuition, and even emotion can be replicated, the distinction between human and artificial dissolves. Perilli invites us to reflect on our role in a world where technology is no longer an instrument, but a possible new form of existence. The question isn't if machines will become conscious, but how this will change forever our definition of life and thought.

Italian publication date: **February 2025**

**312 pages**

**World Rights Available**

**Lorenzo Perilli** is a philosopher and academic, an expert in humanistic computer science. He has led courses with a focus on the methodological implications and applications of computer science to classic texts and in the sphere of philosophical studies.

English translation by Seán McDonagh

The book is divided into 88 very brief chapters. The author, an academic of international renown, has paid particular attention to language and style, to be clear and original while still maintaining scientific exactness. Ranging from one to four pages, each chapter addresses an aspect of the relationship between humans and the algorithmic machines that now dominate our lives, and refers to the fundamental moments of Western history and culture, in which clues to the ongoing developments can be seen: from Homer to Aristotle and Galen, from Descartes and Leibniz to Norbert Wiener, John von Neumann, and Alan Turing, from Samuel Butler and Aldous Huxley to John Searle, Roger Penrose, and Douglas Hofstadter, to name a few; the book not only invites us to reflect and enlighten ourselves but also serves as a guide on a journey through the history of culture and science.

In *Artificial Consciousness*, all the most important issues concerning the relationship between the individual, science, society, politics, and algorithmic machines are addressed, explaining how the oldest dream of humans is coming true today: the dream of building *automata* endowed with intellect and capable of understanding language and instructions given by humans, the dream already narrated by Homer in the *Iliad*.

As Norbert Wiener, the father of the modern interweaving of science and technology, once said, “The idea of non-human devices of great power and great ability to carry through a policy, and of their dangers, is nothing new. All that is new is that now we possess effective devices of this kind. In the past, similar possibilities were postulated for the techniques of magic”.

The core of the book lies in a fundamental consideration: there is a radical alterity between the functioning of algorithmic machines and the functioning of the human machine. Algorithmic machines are oriented solely to the achievement of the objective with maximum efficiency. They reduce to a minimum, to the point of removing all of its traces, the *process* necessary to achieve that objective. On the contrary, the human mind has developed thanks to those processes and the efforts necessary to achieve human goals, from the simplest to the most complex: the cognitive development of humans is therefore destined for radical change and perhaps for regression at a moment when machines replace them not only in the most mechanical among their tasks, but also in cognitively higher tasks such as reasoning, writing, mathematics, interpretation of reality, and creativity.

Today, machines know how to read a person’s mind through neural interfaces and are able to learn and self-replicate as von Neumann expected. There are machines that, as bionumeric organisms, are capable of reproducing biological phenomena. Nothing that has happened in human history is comparable to what is happening today. AI holds an implication that none of the turning points that have characterized human history have ever had: *it reverses the relationship between subject and object*. The subject becomes passive, the object becomes active and takes over the former’s role.

Machines are thus able to perform better and faster almost all the functions that until now we considered unique to humans. Even recognizing the emotions betrayed by the face of those who speak to us is accomplished better when using affective computing systems. Now, if all such functions, which we consider as higher functions that distinguish the human being from other living beings, can be performed much more effectively by a machine, and the human processes that lead to them are lost as superfluous, what remains of us?

The text is based on three convictions:

- 1) that the only possibility of understanding the scope of the ongoing technological revolution is to define the concepts and correlate the different areas in which research is making immeasurable progress, areas that too often are looked at in isolation, thus missing their actual importance and potential impact;
- 2) that the ongoing revolution is already leaving a permanent mark on all aspects of our individual and communal lives: interaction between individuals, development of human cognitive abilities, the neural

structure of the brain, the political system, the capitalist structure of society, work, teaching, communication and information, the notion of true/false, free will, the very notion of human and machine, power, and democracy; and

3) that today more than ever it is urgent to overcome, as Charles Percy Snow had hoped many decades ago, the distances between *the two cultures*, science and humanities, and recover the ancient unity of knowledge, if we want to understand the new world.

The book does not intend to say whether we should or should not fear AI. Whether AI is intelligent or not intelligent. Whether it is conscious or not conscious. It argues, rather, that we tend to adapt our concept of intelligence, and our concept of consciousness, to what algorithmic machines can or cannot do. We are progressively but rapidly modifying our categories. We, today, are or are not intelligent to the extent that we know how to do or not do what the machine does. Now, since the machine, if it does the same things as us, certainly does them better, it is good that we let the machine do those things. Since the machine is intelligent, the things it does are the result of intelligence. Since it does them better than we would do them, it is more intelligent than us.

Is it possible that the individual processes of intelligence, and the individual mental processes as well as the individual processes that lead to the constitution of consciousness, can be formalized and reproduced algorithmically? Or are neither intelligence, nor mind, nor consciousness reproducible as the sum of such individual processes, and the whole is not the sum of the parts?

In *Artificial Consciousness*, Lorenzo Perilli offers an attempt to tie the threads, trusting to solicit, if not an awareness, at least that natural curiosity that has always motivated humans to look around and wonder about themselves, about their destiny, about the world —and that should lead us today to look at the new self, the new destiny, the new world.

*Artificial Consciousness* is a profound and innovative work, which reveals to us how today's technology is no longer stolid inert matter or something that reacts to our solicitations but knows how to make itself an autonomous subject; and what we might ultimately find ourselves facing is a new form of existence.

Prologue  
Introduction

## Part I: The Destiny of the Human

1. The Reversed Paradigm
2. The Fundamental Question
3. Objectives, Not Processes
4. A New Perspective for Humans in the World
5. Principle of Autonomous Mathematical Action
- Addendum to 5: What is an Algorithm?
6. The World as Representation
7. Non-Things and the Loss of the Hand
8. Representation of Language and Cultural Homogeny
9. Towards a Single Universal Culture
10. The New Tower of Babel

## Part II: Life, Intelligence, and Consciousness of Algorithmic Machines

11. A Heptadic Scheme, or: The Seven Phases of the Construction of the World
12. Deep Learning
13. Qubits
14. Sensor Networks
15. Seeing by Hearing
16. Artificial Intelligence (AI)
17. AI: An Ambiguous Triumph
18. Defining AI: Tesler's Theorem
19. AI: The One-Dimensional World
20. "I Can See a Moment in the Future...": The Butler-Shannon Prediction
21. Computational Intelligence
22. The Inner Logic of the Brain
23. Metaphors of the Living
24. Life and Non-Life: Part 1
25. Life and Non-Life: Part 2
26. Life and Non-Life: Part 3
27. The Digital Genesis
28. Biological Phenomena and Bionumeric Phenomena
29. Life and Non-Life: Part 4
30. The Child Machine
31. The Difference Between Decision and Choice: Weizenbaum's Doubt
32. Decision in the Game of War
33. McCarthy's *coup de maître*
34. The Freedom to Be a Round Peg in a Square Hole
35. The Ethics of Magic
36. Doers and Watchers
37. Linguistic Systems for Intelligent Machines
38. Deep Reasoning: Machines Learn to Reason
39. Artificial Consciousness (AC): Part 1
40. Artificial Consciousness: Part 2
41. Artificial Consciousness: Part 3
42. Telepathy and Mind-Reading
43. Artificial Consciousness: Part 4
44. The Clock is Ticking: Wiener's Warning
45. Is This Really Consciousness?
46. Another Idea of Consciousness
47. Subject, Object, and the Perception of the World: Part 1

- 48. Subject, Object, and the Perception of the World: Part 2
- 49. The Brain and the Experience of the World
- 50. The Dual Form of Things
- 51. The World is a Game of Go, but Playing Does Not Make Sense Anymore
- 52. "Human Beings Seem to be Redundant"
- 53. The MuZero Algorithm

### **Part III: The Iron Cage of Capitalist Rationality**

- 54. The Relationship between Man and World in Algorithmic Thinking
- 55. Algorithms and the World of Life
- 56. Against Meritocracy
- 57. Capitalism, the Truck Driver, and Roses
- 58. Liberation from the Dominance of Mediocrity
- 59. Capitalism, Meritocracy, Algorithmization
- 60. The Iron Cage of Rationality
- 61. The Bokanovsky Process and AI
- 62. Man, Process, Error
- 63. Digital Humans and Biological Humans
- 64. A Technology that Changes Mind and Brain
- 65. The Computer Learns to Speak
- 66. A Language Without Reasoning
- 67. Power of Conditioning
- 68. From Word Scenery to Image-Generating Word
- 69. Humans' Oldest Dream
- 70. Labour Creates Humans
- 71. Nature Against Us
- 72. The Modern Prometheus and 19<sup>th</sup> Century Imagery

### **Part IV: Power, Control, Freedom**

- 73. Community, Identity, Stability
- 74. AI, the Cloud, and Totalitarianism
- 75. To Love One's Own Slavery
- 76. Freedom or Democracy: The Choice
- 77. Rights, Duties, and the Human Soul
- 78. Technique and Western Tradition
- 79. An Unbridgeable Social Gap
- 80. Living in the Panopticon
- 81. Panopticon and Datacenter
- 82. Less Perfection, More Freedom
- 83. Efficiency and the Quest for Meaning

### **Appendix**

Brief and Optimistic Journey Among the Promises of New Medicine

- A1. Magnificent and Progressive Fates: The Revolution
- A2. Tumors and the Immune System
- A3. The Building Blocks of the Body and Their Portrait
- A4. Virtual Human Physiology
- A5. The New Doctor and His Assistant

### **Epilogue**

*Complementary Notes and Bibliographic Integrations*

*In the Margin: An Etymological Note*

## Chapters Breakdown

### Part I: The Destiny of the Human

**Part I** is devoted to the fundamental questions posed by the new machines: What are humans? Is there any compatibility between the human dimension and the algorithmic dimension? What are the consequences of the standardization of language and culture - that is, the rapid process towards a single universal culture and a single description of the world, based on just two symbols, the 0s and 1s of binary code? There is likely a point of equilibrium between human beings, the results of their creativity, and the surrounding world. However, finding it requires a complex journey with an uncertain outcome. At times, our rationalistic faith in technology displays elements of childishness.

### Part II: Life, Intelligence, and Consciousness of Algorithmic Machines

**Part II** reflects on the developments of algorithmic machines, which force us to rethink the fundamental categories of our identity and to seek definitions for hitherto elusive concepts such as intelligence, consciousness, and life. Can machines understand, speak, and live? What precisely do these concepts mean?

The author argues here that it is a mistake to focus on only one aspect of the current technological evolution, as specialists immersed in their work tend to do. Instead, it is necessary to look at the whole set of aspects that will determine the new world, namely:

- deep-learning algorithms;
- computational power of quantum computers;
- sensor networks managed by algorithms;
- artificial-intelligence systems based on deep learning;
- computational intelligence systems and genetic algorithms;
- linguistic systems for intelligent machines; and
- artificial consciousness systems.

Taken together, these seven phases give rise to systems of complexity and capability never before seen in human history. However, what matters most for imagining the medium- to long-term impact and transformative potential is not computational and processing capacity but the *capacity for autonomous decision-making and action*.

### Part III: The Iron Cage of Capitalist Rationality

**Part III** reflects on the consequences that the ongoing transformation will inevitably have, and is already having, on the political-economic structure of society. It explains how the new algorithmic technologies are perfectly coherent with a system founded on neoliberal principles and forms of anarcho-capitalism. In this system, every rule aimed at a rebalancing is erased, and everyone is left free to act as they can and exploit their power - whether economic, political, or cultural - to the fullest.

### Part IV: Power, Control, Freedom

**Part IV** extends the reflection to the individual and their role in society, and to the achievements that have been made since antiquity and from the Enlightenment onwards especially. The new forms of power and centralized control, in the hands of very few and against which even nation states are powerless, aim to transform individuals into users and spectators – that is, passive entities. They can achieve this by relying on the free adherence of each person to the new system.

The conveniences are so many and so great that we are ready to renounce the principles that have characterized the development of society thus far. We are no longer willing to make efforts that are not for individual advantage but have the common good as their goal. The democratic system proves difficult to reconcile with the new forms of power and the redefinition of the very concept of freedom.

### Appendix: A Brief and Optimistic Journey through the Promises of New Medicine

The **Appendix** briefly explains the radical change that awaits medicine, and the great innovations that will characterize this science that has always accompanied human existence: immunology, physiology, digital

twins of the patient's body, cancer research, genomics, diagnostic imaging, robotic surgery, and the role of the doctor in the new system.