From Silicon Gate to Microprocessors to AI to Consciousness

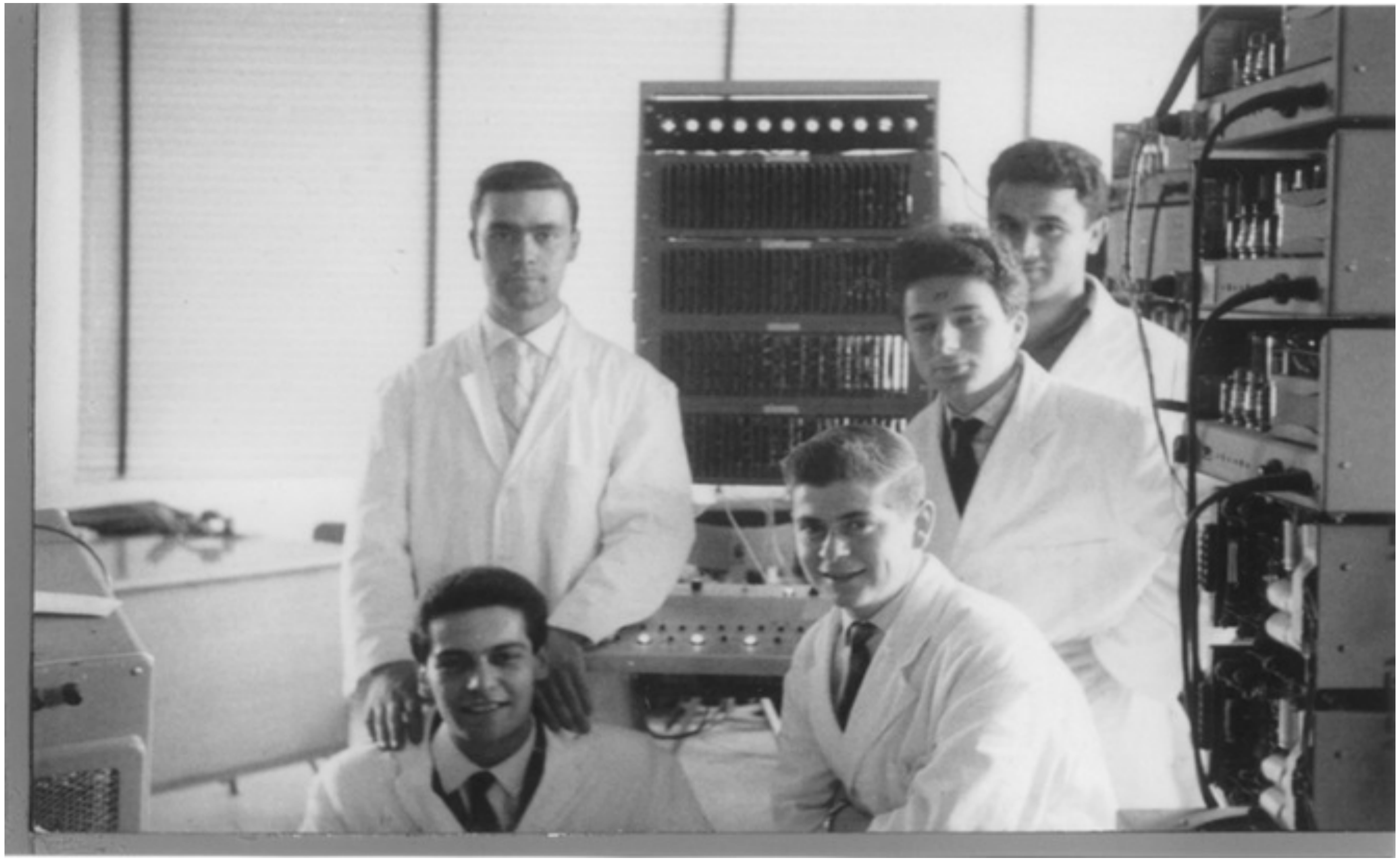
Federico Faggin

Federico and Elvia Faggin Foundation

Background

- **I was born, raised and educated in Vicenza. I graduated "perito radiotecnico" in 1960 and I worked for 15 months at the Olivetti Electronics Lab in Borgolombardo (Mi) in 1960-1961, where I co-designed and built a fully functional experimental programmable computer with germanium transistors**
- I then studied physics at the university of Padua, graduating summa cum laude in 1965
- I joined SGS-Fairchild in Agrate Brianza (Mi) in 1967, where I developed their first manufacturing process for MOS technology and designed their first two commercial MOS ICs
- I moved to Silicon Valley in February 1968 working at the Fairchild R&D Laboratory in Palo Alto, CA
- In 1968, over 95% of all IC sold used bipolar technology. MOS technology was almost universally considered second-best

Olivetti Electronics R&D Lab., 1960-1961



Background

- I was born, raised and educated in Vicenza. I graduated "perito radiotecnico" in 1960 and I worked for 15 months at the Olivetti Electronics Lab in Borgolombardo (Mi) in 1960-1961, where I co-designed and built a fully functional experimental programmable computer with germanium transistors

- **I then studied physics at the university of Padua, graduating summa cum laude in 1965**

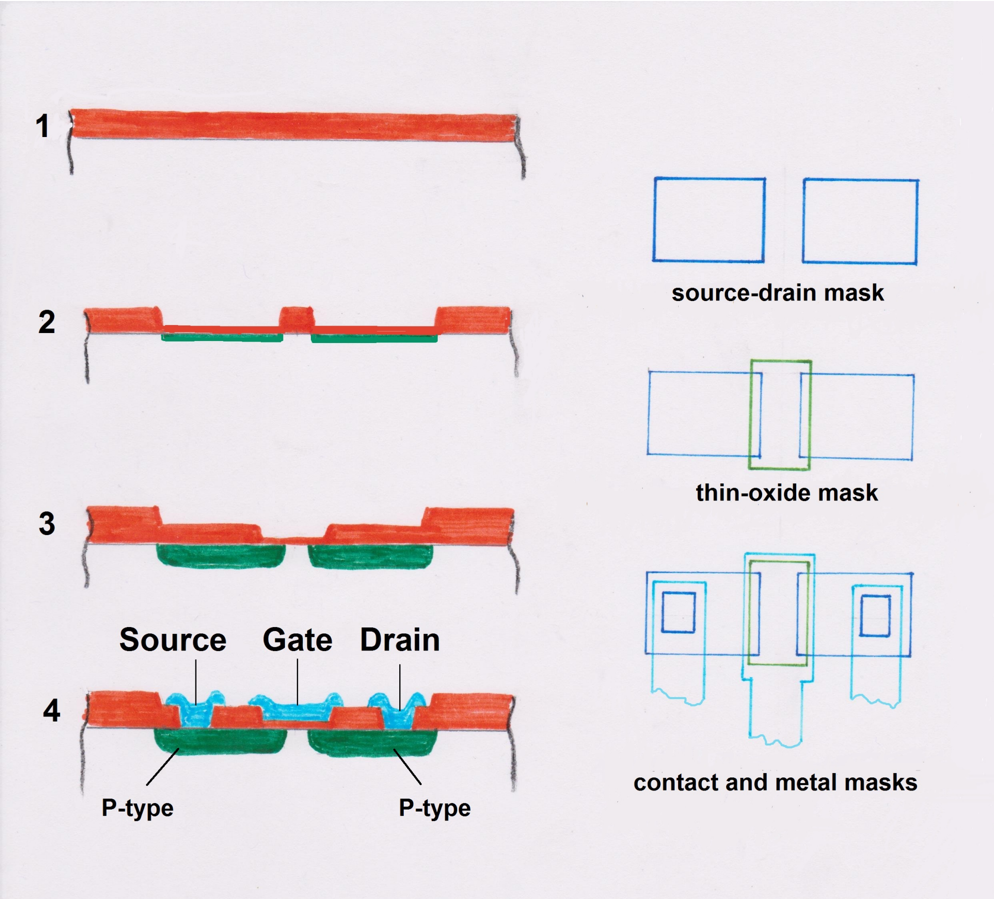
- **I joined SGS-Fairchild in Agrate Brianza (Mi) in 1967, where I developed their first manufacturing process for MOS technology and designed their first two commercial MOS ICs**

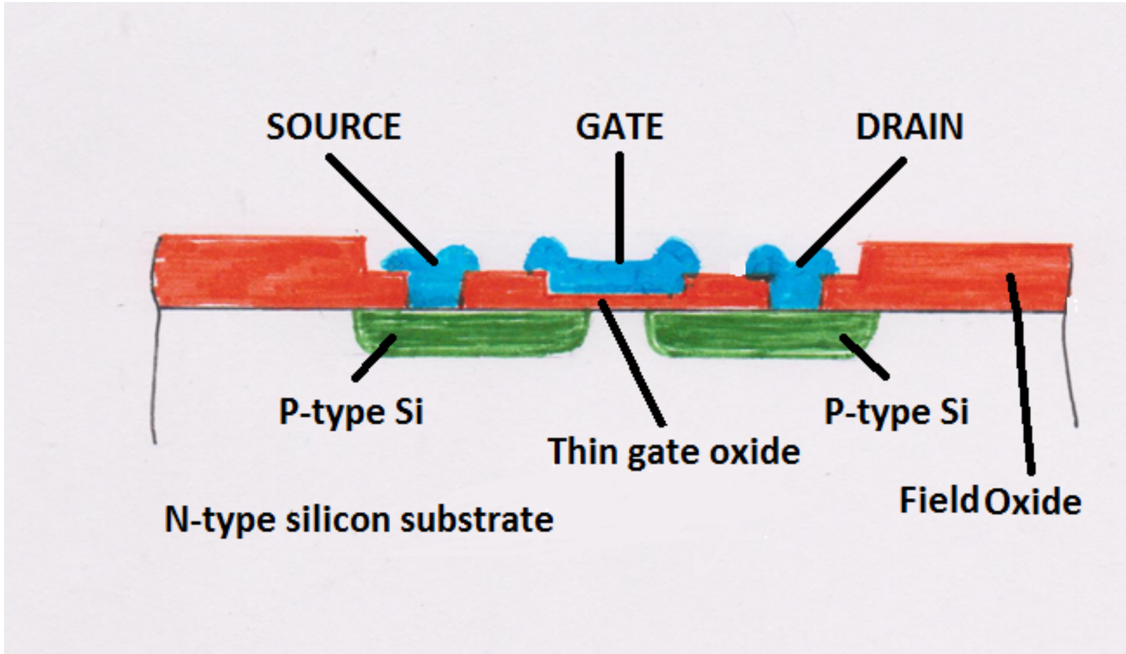
- **I moved to Silicon Valley in February 1968 working at the Fairchild R&D Laboratory in Palo Alto, CA**

- **In 1968, over 95% of all IC sold used bipolar technology. MOS technology was almost universally considered second-best**

High-Threshold Voltage P-channel MOS Transistors

State-of-the-art in 1968

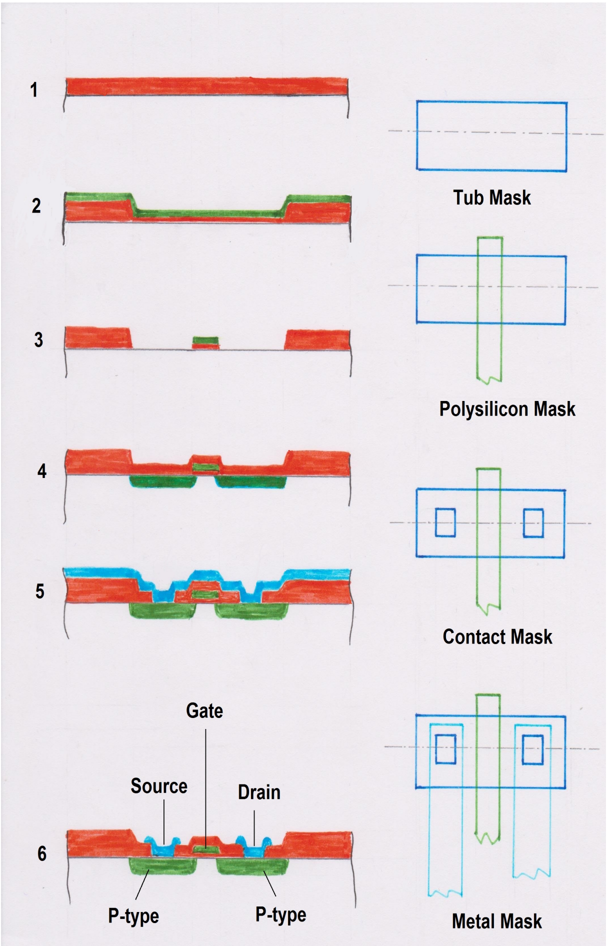




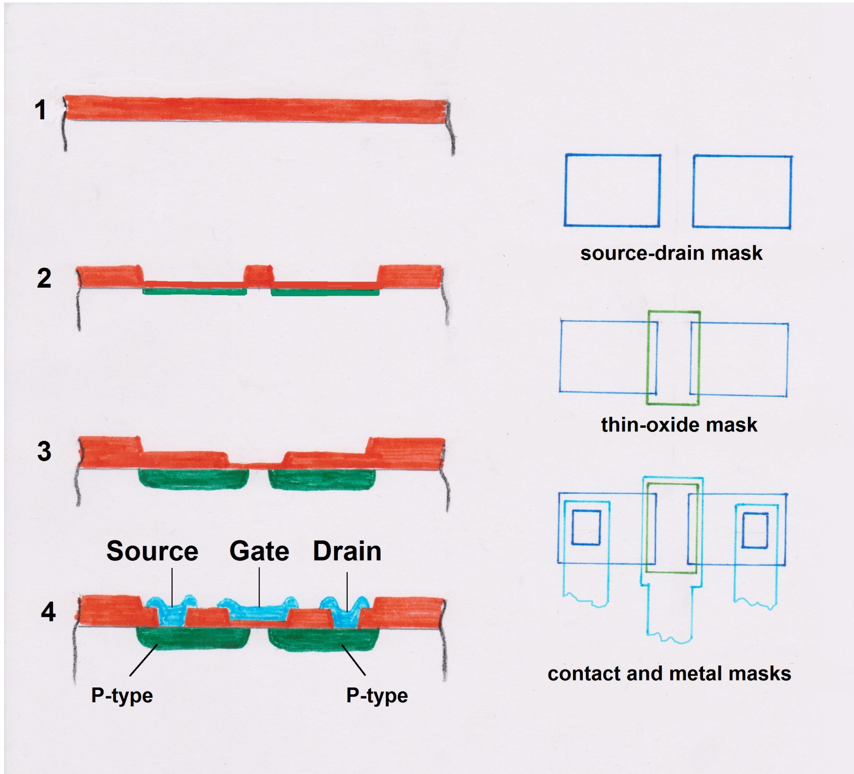
- **In 1968 the industry was producing MOS ICs using P-channel MOS transistors with high-threshold voltage – very slow and still somewhat unreliable**
- **The need for a better MOS technology with higher speed and reliability required lower threshold voltages and self-aligned gates**
- **The concept of self-alignment had been advanced in 1966 by Robert Bower of Hughes Aircraft, and in 1967, John Sarace and collaborators at Bell Labs were the first to create self-aligned MOS transistors with amorphous silicon, but their transistors had all their sources connected together**
- **When I joined the Fairchild R&D Lab as project leader of the SGT, Tom Klein had already made some experiments to determine the work-function difference between heavily P-doped amorphous silicon and single-crystal silicon**
- **He determined that a P-channel MOS transistor with silicon gate would have a threshold voltage 1.1 volt less than an equivalent transistor with aluminum gate**
- **However, nobody at Fairchild or at Bell Labs had yet figured out the process architecture necessary to fabricate self-aligned silicon gate transistors suitable for ICs. In fact, nobody at Fairchild even knew how to precision-etch thin layers of amorphous silicon. My job was to invent/develop the missing ingredients and integrate the existing ideas into a process that worked**

The SGT Process Architecture

Silicon gate

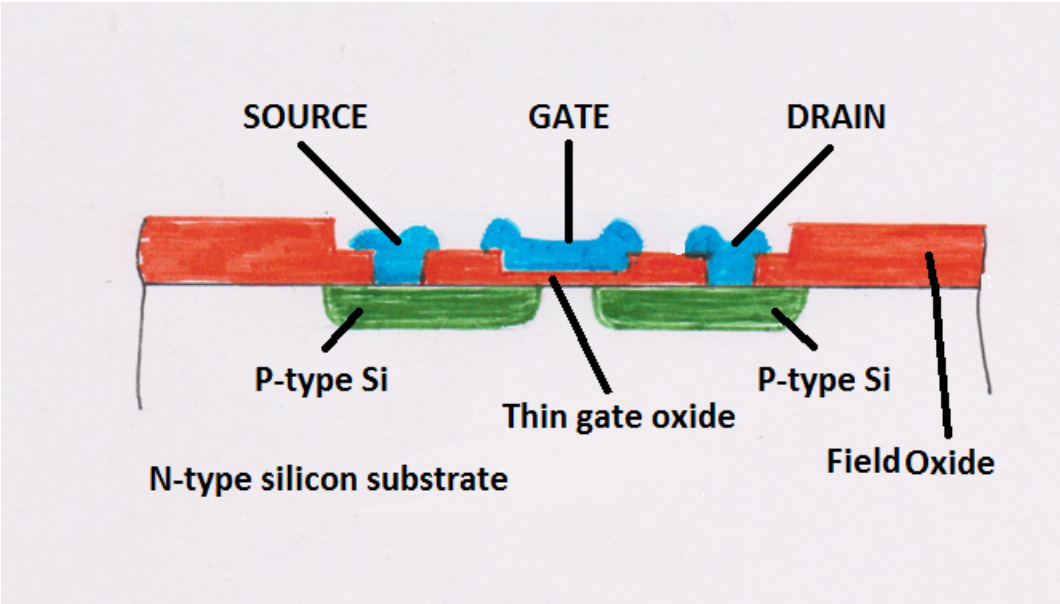


Aluminum Gate

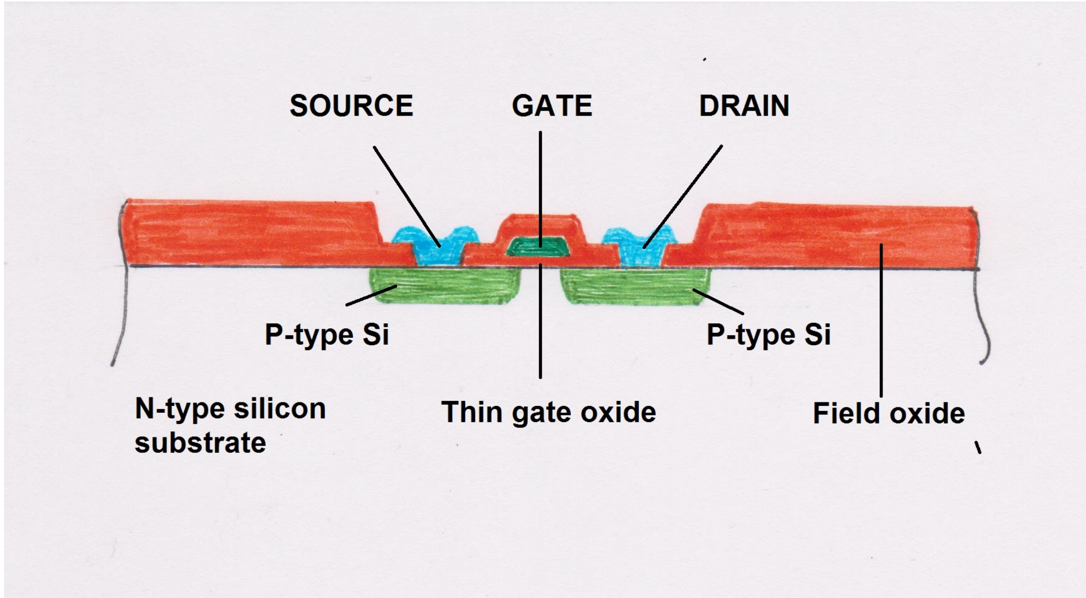


P-channel MOS Transistors

Aluminum gate



Polysilicon gate

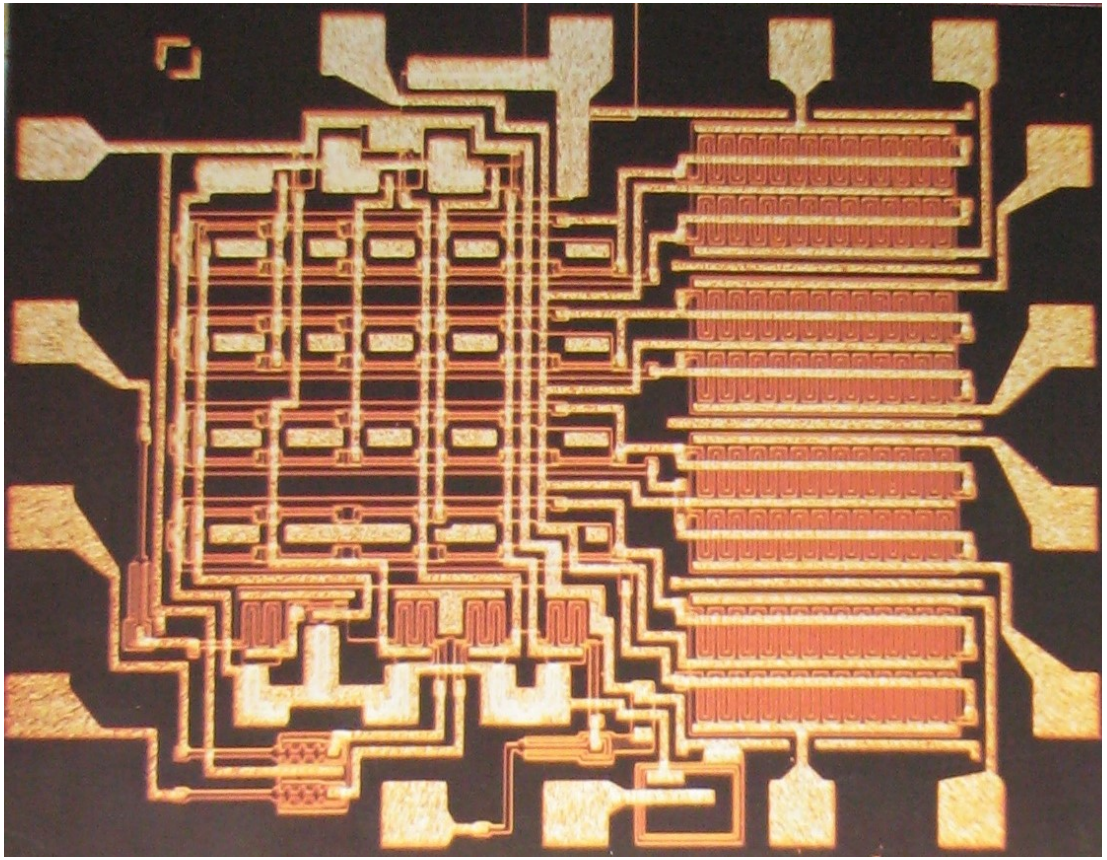


Filling the Missing Ingredients

- **The next step was to design an IC to prove the superiority of the SGT compared with the incumbent technology**
- **The chip chosen was the Fairchild 3705 an 8-channel analog multiplexer with decoding logic that was difficult to manufacture**
- **I designed the SGT version of the 3705 called 3708 and got working devices in July 1968**
- **I found that amorphous silicon would occasionally break at the oxide steps and using poly-silicon obtained by vapor deposition of silane solved the problem**
- **Leakage current was greatly improved by using phosphorous gettering and by the end of 1968 the 3708 was sold commercially**
- **The MOS operating division resisted adopting the SGT claiming that the layout density was inferior to metal gate and bootstrap loads could not be made**
- **I showed how to properly layout silicon-gate circuits with and without an additional invention of mine called buried contact and I solved the bootstrap load obstacle one year later**

Fairchild 3708 – 8-Channel Analog Multiplexer

World's first commercial IC with silicon gate: sold in Dec. 1968: 5x faster, 3
lower on-resistance, 100x less leakage than 3705 (aluminum-gate versior

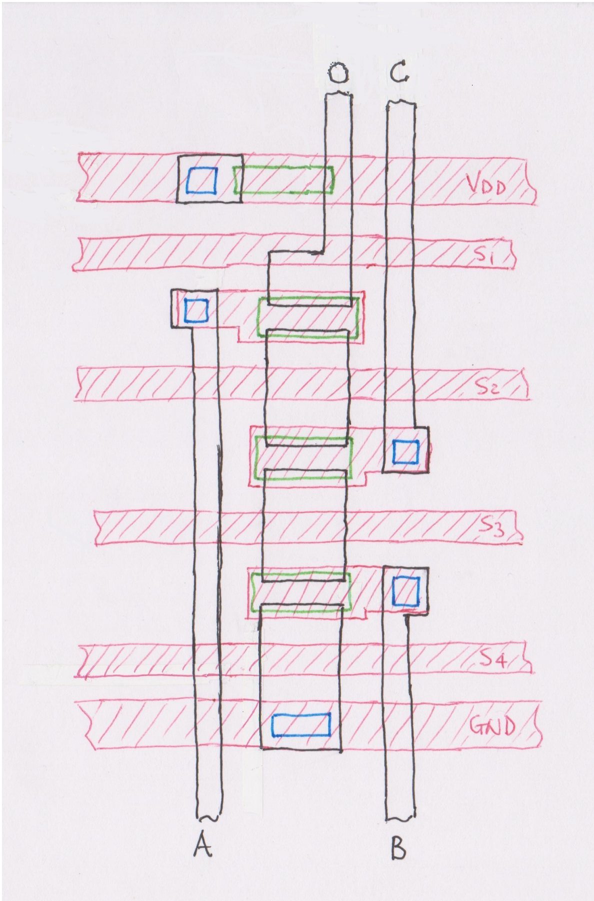


Filling the Missing Ingredients

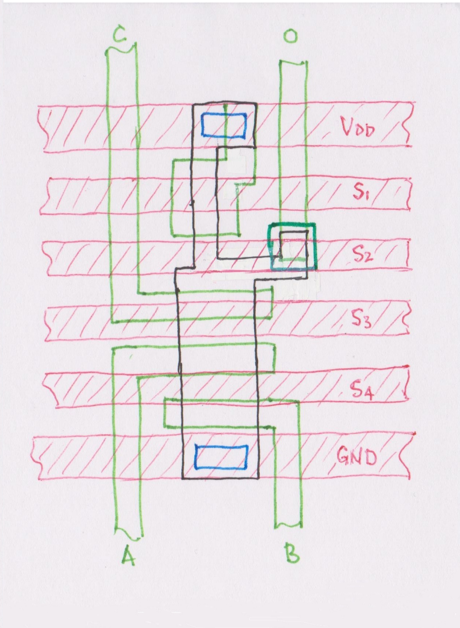
- The next step was to design an IC to prove the superiority of the SGT compared with the incumbent technology
- The chip chosen was the Fairchild 3705 an 8-channel analog multiplexer with decoding logic that was difficult to manufacture
- I designed the SGT version of the 3705, called 3708, and got working devices in July 1968
- I found that amorphous silicon would occasionally break at the oxide steps and using poly-silicon obtained by vapor deposition of silane solved the problem
- Leakage current was greatly improved by using phosphorous gettering and by the end of 1968 the 3708 was sold commercially
- **The MOS operating division resisted adopting the SGT claiming that the layout density was inferior to metal gate and that bootstrap loads could not be made**
- **I showed how to properly layout silicon-gate circuits with and without an additional invention of mine called buried contact and I solved the bootstrap load obstacle one year later**

Comparison Layout of a 3-input NAND gate

Metal gate

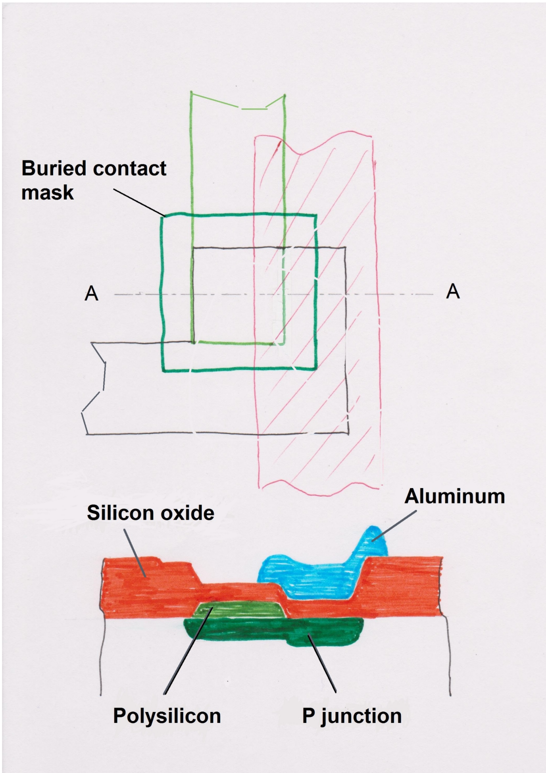


Silicon gate (50% smaller)

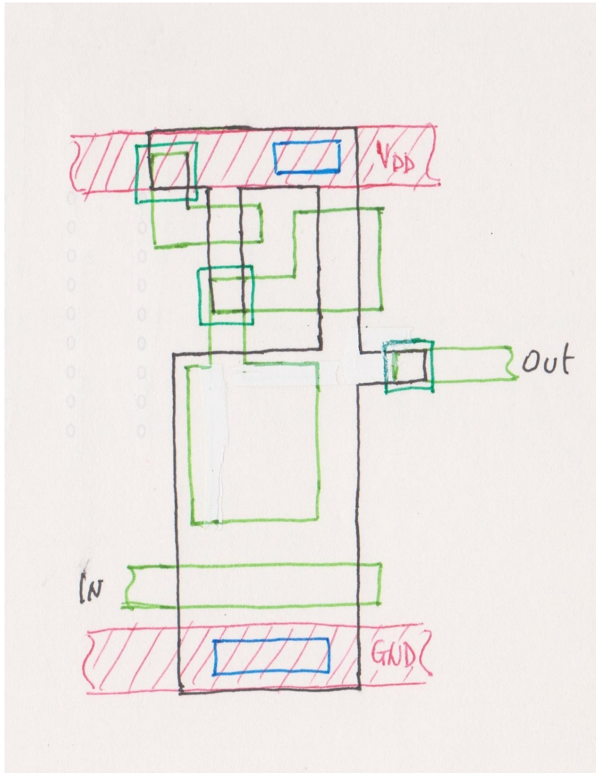


Buried Contact and Bootstrap Load with SGT

BURIED CONTACT: Direct contact between polysilicon and junctions. It allows metal interconnections to run over the contacts



BOOTSTRAP LOAD: It allows achieving outpu voltage equal to Vdd dynamically. Metal line (not shown) can run over the circuit



What Made the SGT Second to None

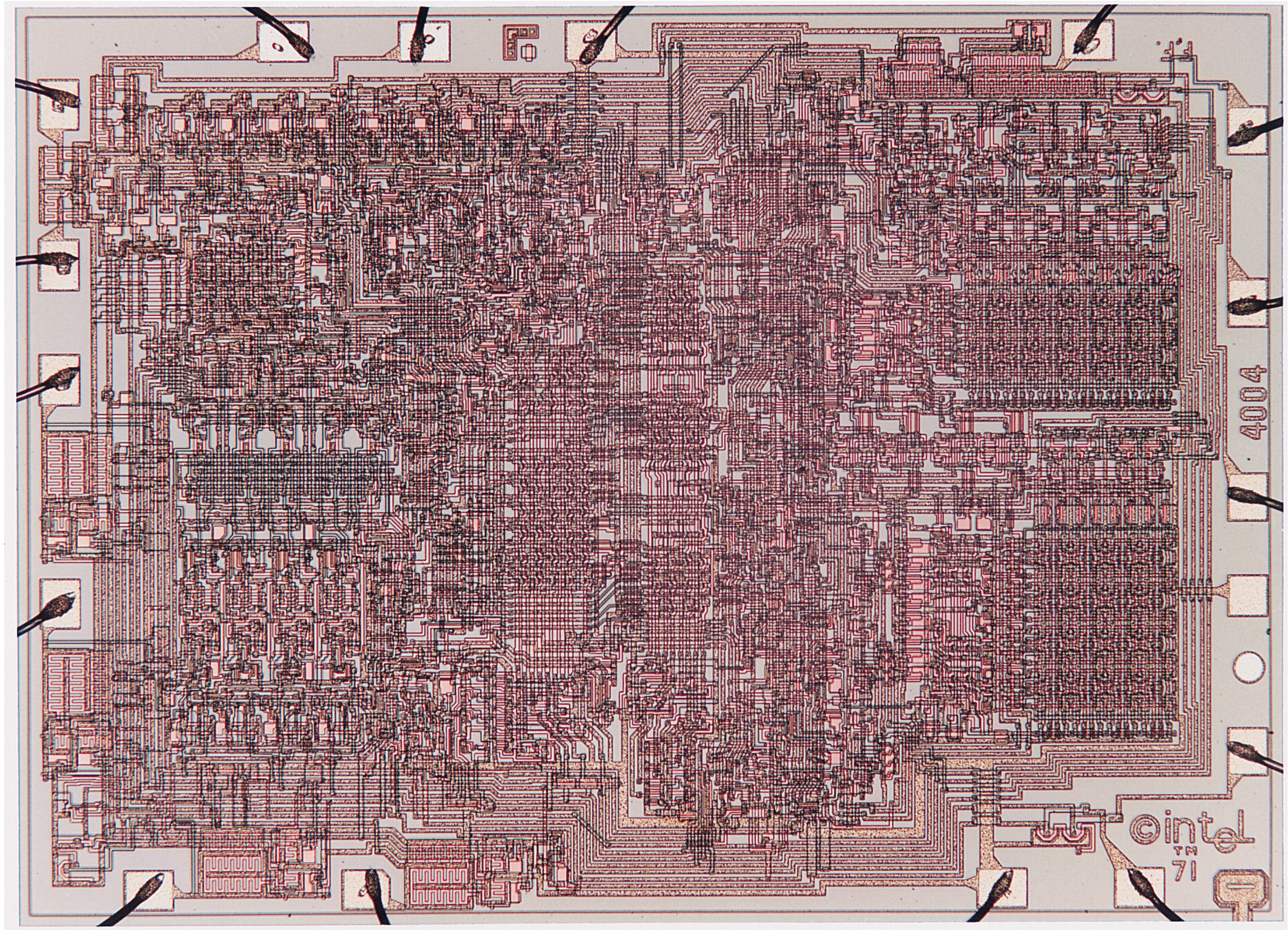
- New process architecture
- Etching solution for silicon layer
- Replaced amorphous silicon with poly-silicon
- Phosphorous gettering, buried contact, and bootstrap loads
- With all these inventions and improvements the SGT became by far the best technology in the world and eventually ended up replacing the incumbent *bipolar* technology
- The SGT made possible DRAMs, microprocessors, non-volatile memories, and CCD imagers. It was universally adopted by the semiconductor industry and it allowed the scaling of MOS from 8000 nm down to 45 nm
- Today flash memories still rely on SGT because silicon dioxide is the best insulator known

From Silicon Gate to Microprocessors

- **In July 1968, Bob Noyce and Gordon Moore left Fairchild to start Intel, a company dedicated to developing semiconductor RAM to replace the magnetic core memories used in computers**
- **Intel "adopted" the SGT and I joined Intel in 1970 to lead the "Busicom Project," a family of 4 chips of which one became the world's first microprocessor**
- **No chip of that complexity and speed had been done before**
- **Everything that I'd ever learned and invented before was necessary to design the 4004. In March 1971, the first commercial microprocessors were sold to Busicom for their desktop printing calculator**
- **I successfully lobbied Intel's management to sell the microprocessor to the entire market for all applications**
- **The SGT with buried contact and bootstraps also empowered the first dynamic RAM, the Intel 1103**
- **The Intel 1103 and the 4004 showcased the power of the SGT to the rest of the industry that had resisted adopting it. Within a few years, every semiconductor manufacturer in the world switched to the SGT**

The Intel 4004 – World's First Microprocessor

April 1970 – March 1971

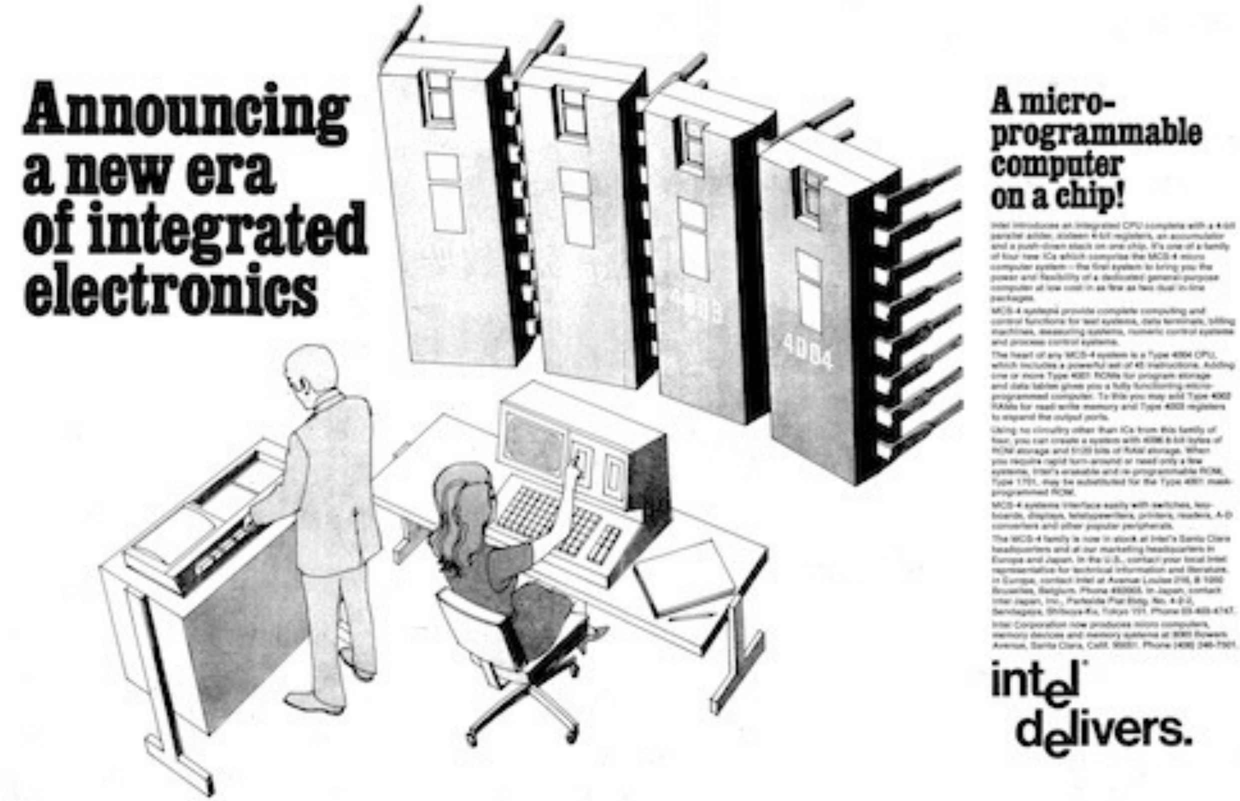


From Silicon Gate to Microprocessors

- In July 1968, Bob Noyce and Gordon Moore left Fairchild to start Intel, a company dedicated to developing semiconductor RAM to replace the magnetic core memories used in computers
- Intel "adopted" the SGT and I joined Intel in 1970 to lead the "Busicom Project," a family of 4 chips of which one became the world's first microprocessor. No chip of that complexity and speed was ever done before
- Everything that I'd ever learned and invented before was necessary to design the 4004. In March 1971, the first commercial microprocessors were sold to Busicom for their desktop printing calculator
- **I successfully lobbied Intel's management to sell the microprocessor to the entire market for all applications**
- The SGT with buried contact and bootstraps also empowered the first dynamic RAM, the Intel 1103
- The Intel 1103 and the 4004 showcased the power of the SGT to the rest of the industry that had resisted adopting it. Within a few years, every semiconductor manufacturer in the world switched to the SGT

Announcing the 4004 to the World

November 1971

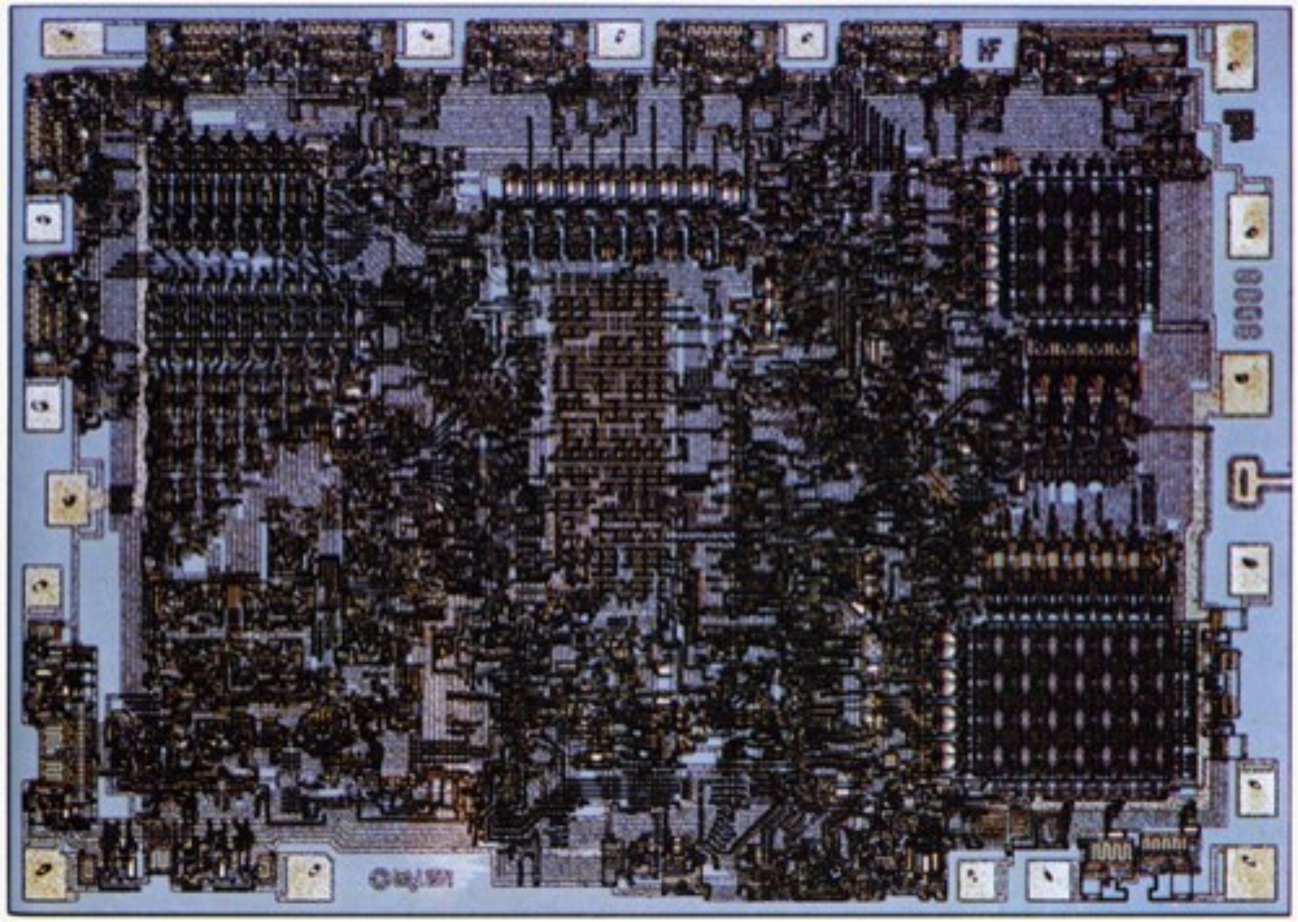


From Silicon Gate to Microprocessors

- In July 1968, Bob Noyce and Gordon Moore left Fairchild to start Intel, a company dedicated to developing semiconductor RAM to replace the magnetic core memories used in computers
- Intel "adopted" the SGT and I joined Intel in 1970 to lead the "Busicom Project," a family of 4 chips of which one became the world's first microprocessor. No chip of that complexity and speed was ever done before
- Everything that I'd ever learned and invented before was necessary to design the 4004. In March 1971, the first commercial microprocessors were sold to Busicom for their desktop printing calculator
- I successfully lobbied Intel's management to sell the microprocessor to the entire market for all applications
- **The SGT with buried contact and bootstraps also empowered the first dynamic RAM, the Intel 1103**
- **The Intel 1103 and the 4004 showcased the power of the SGT to the rest of the industry that had resisted adopting it. A few years later, every semiconductor manufacturer in the world switched to SGT**

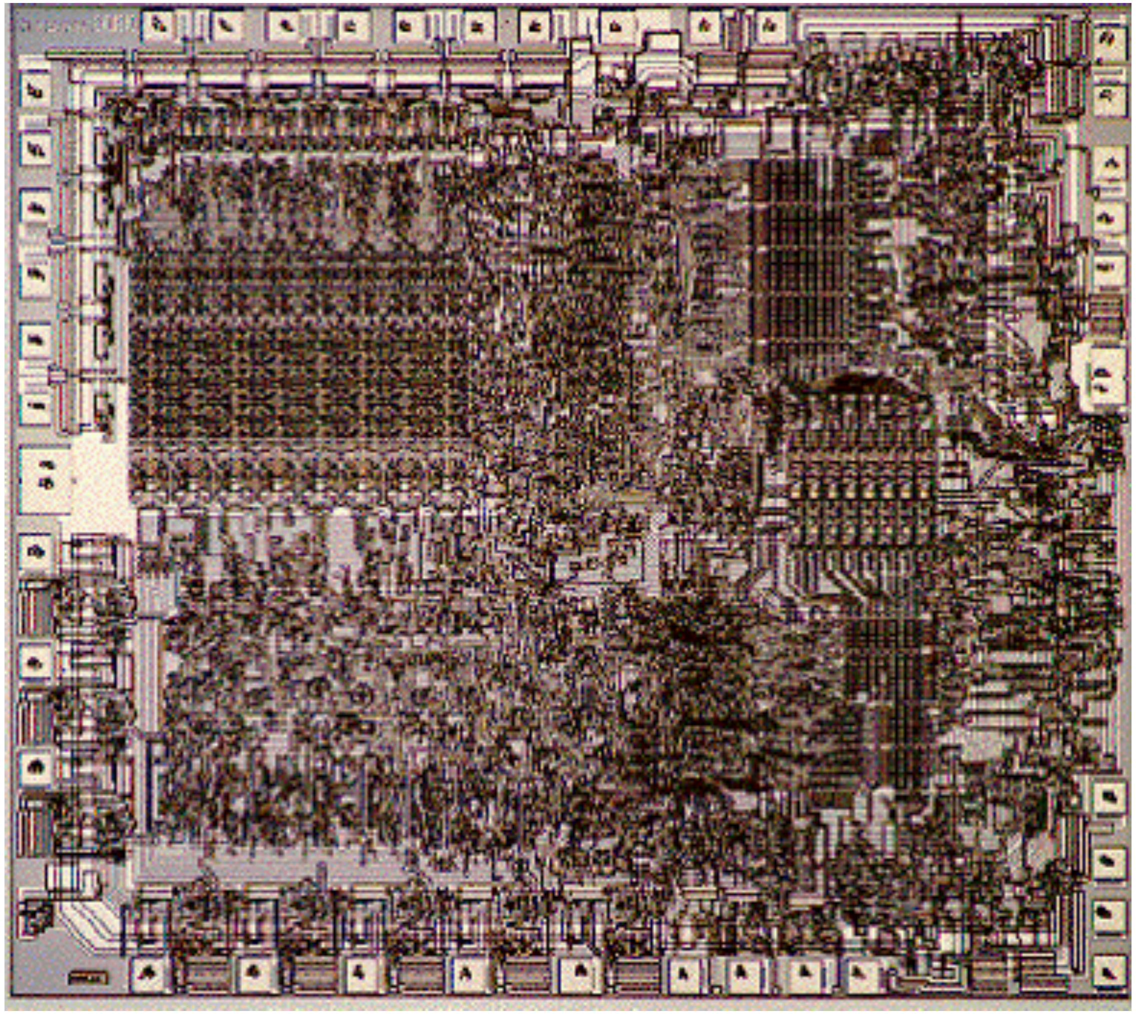
The World's First 8-bit Microprocessor

The Intel 8008 – April 1972



The First 2nd Generation Microprocessor

The Intel 8080 – March 1974

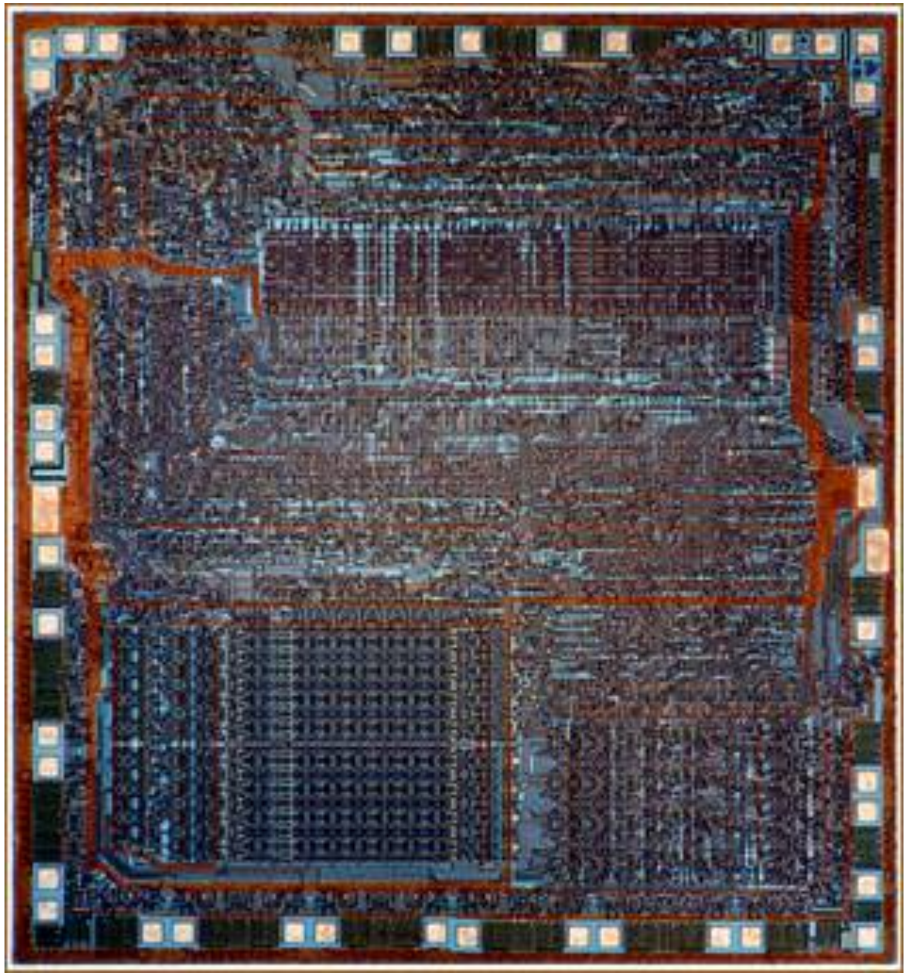


From Chip Designer to Entrepreneur

- **In November 1974, I started my first company, Zilog, with Ralph Ungermann**
- **Zilog was the first company entirely dedicated to the emergent microprocessor and microcontroller market**
- **Zilog introduced the Z80 family of chips starting in May 1976 with the Z80-CPU, followed by the Z80-PIO, Z80-CTC, Z80-SIO, and Z80-DMA**
- **Zilog introduced the Z8 microcontroller in 1978, and the 16/32 bit Z8000 microprocessor in 1979**
- **The Z80-CPU and the Z8 are still in volume production today**
- **Zilog was acquired by Exxon Enterprises in 1981**

The First 3rd Generation Microprocessor

The Zilog Z80-CPU – March 1974



From Chip Designer to Entrepreneur

- In November 1974, I started my first company, Zilog, with Ralph Ungermann
- Zilog was the first company entirely dedicated to the emergent microprocessor and microcontroller market
- Zilog introduced the Z80 family of chips starting in May 1976 with the Z80-CPU, followed by the Z80-PIO, Z80-CTC, Z80-SIO, and Z80-DMA
- **Zilog introduced the Z8 microcontroller in 1978, and the 16/32 bit Z8000 microprocessor in 1979**
- **The Z80-CPU and the Z8 are still in volume production today**
- **Zilog was acquired by Exxon Enterprises in 1981**

Founder-CEO of Cygnet Technologies (1982)

Voice-Data Cosystem – "The other half of the PC" – March 1984

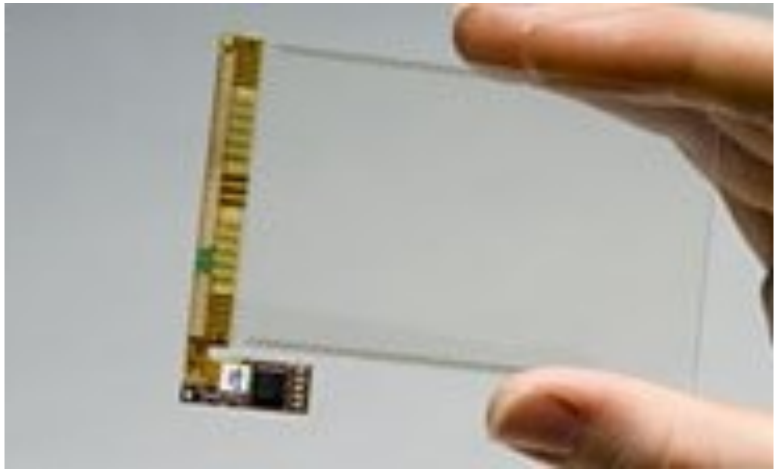
From Microprocessors to Artificial Intelligence

- **Co-founded Synaptics in 1986 to do research in artificial neural networks (NN) realized with analog floating-gate technology**
- **Carver Mead from Caltech joined Synaptics in 1987 (neuromorphic sensory chips)**
- **Explored applications of NN and developed the technology for creating cognitive computers by *emulating* brain structures**
- **Synaptics developed the I-1000 first OCR chip for Verifone (optical imager plus two NNs for MICR character recognition. Achieved 99.999% accuracy)**
- **Led the team that came up with the Touchpad and Touchscreen idea in 1992. Decided to change course and pursue Touchpad in 1993 and Touchscreen in 2001**

Synaptics, Inc.

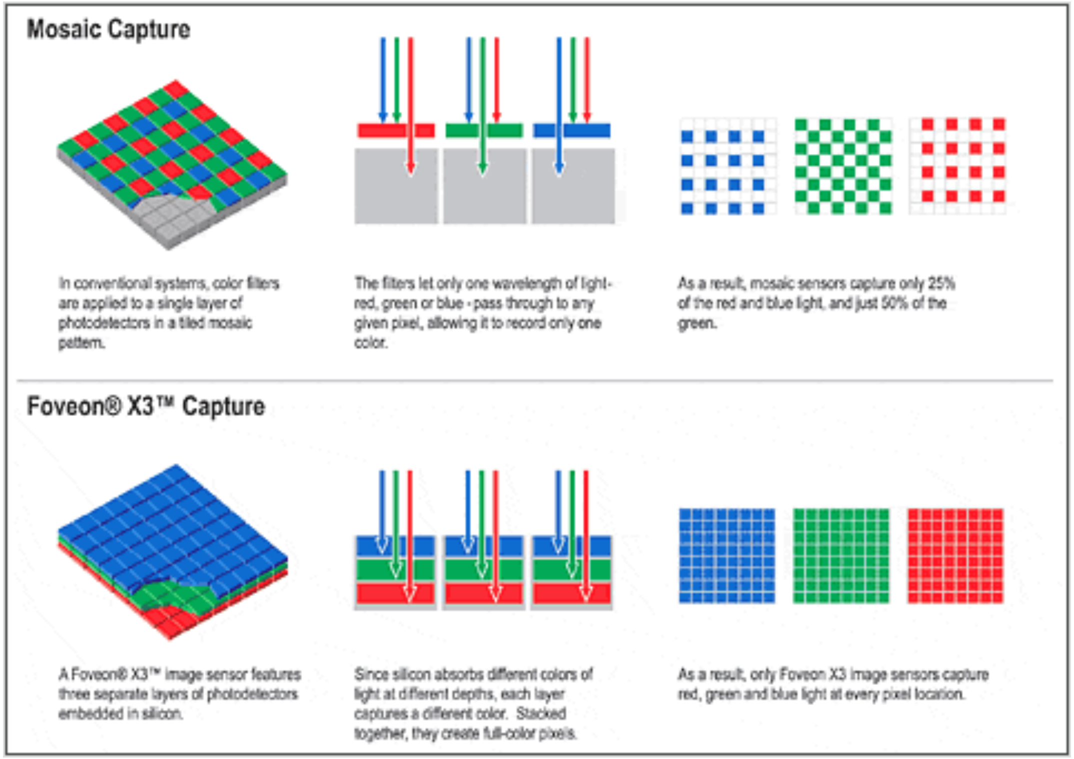
Touchpad (1994)

Touchscreen (2001)



- **I was CEO of Synaptics from 1987 to 1999 and then Chairman of the Board until 2009**
- **In 1997 Carver Mead started Foveon, Inc. to develop a professional digital camera with 3 CMOS imagers and a prism. National Semiconductors and Synaptics were the first investors**
- **In 2000 Dick Merrill of Foveon came up with the idea of stacking the image sensors to reduce the chip size**
- **In 2002 Sigma introduced the Sigma SD9 DSLR camera with the Foveon X3 technology**
- **In 2003 I became CEO of Foveon and decided to focus the company on camera-phone applications**
- **In 2008 Foveon was acquired by Sigma**

Foveon, Inc. X3 technology (2002)
Stacked image sensors without color filters



- I was CEO of Synaptics from 1987 to 1999 and then Chairman of the Board until 2009
- In 1997 Carver Mead started Foveon, Inc. to develop a professional digital camera with 3 CMOS imagers and a prism. National Semiconductors and Synaptics were the first investors
- In 2000 Dick Merrill of Foveon came up with the idea of stacking the image sensors to reduce the chip size
- In 2002 Sigma introduced the Sigma SD9 DSLR camera with the Foveon X3 technology
- **In 2003 I became CEO of Foveon and decided to focus the company on camera-phone applications**
- **In 2008 Foveon was acquired by Sigma**

From AI to Consciousness

- **While studying neuroscience in the 1980's I asked myself how could consciousness emerge from the operation of the brain. I thought that if consciousness emerges from a complex system, I should be able to create a conscious computer**
- **I started investigating how such a goal could possibly be achieved. And the more I thought about it, the more impossible the task became**
- **I decided to investigate the nature of consciousness as a personal project. After 20 years of intense study, I concluded that consciousness cannot emerge from inert matter**
- **After the sale of Foveon, I decided to dedicate myself full-time to the study of consciousness**
- **Then I found that no research money was available to study consciousness if the hypothesis is that consciousness may be an irreducible property of nature**
- **Therefore, I started the Federico and Elvia Faggin Foundation in 2011 to fund such research in various US institutions**

Part II

What Are the Fundamental Differences Between Human and Artificial Intelligence?

What Is Artificial Intelligence?

- **Artificial intelligence (AI) is about machines that learn, solve problems, and behave with intelligence *similar to* humans**
- **At the beginning, AI was about top-down, procedure-driven *programming* and expert systems. Neural networks (NN) with their bottom-up, data-driven *learning* started early on with the *Perceptron*, but were considered the wrong approach until recently**
- **NN are structures that mimic how the brain processes information. During the last 5-10 years, machine learning based on NN won over conventional AI**
- **On a class of specific problems, AI may outperform the best human being**
- **On some other problems, like autonomous driving and factory robots, AI may perform like an average human being, but its low cost, *consistent performance,* and 24/7 availability, threaten human employment**
- **Current machine learning involves computers that *simulate* mathematical models of NNs. Exploration of specialized architectures for NNs is now intense**

- **The crucial questions are: Is *understanding* essential to intelligent behavior? Can a machine truly understand? Does understanding require consciousness? And if so, can a machine be conscious?**

- **Today AI poses a serious challenge to our understanding of the nature of life, the nature of consciousness, and the nature of reality. These are no longer philosophical questions. They must be seriously addressed by science**

- **Are we biological computers? And if not, what are we?**

What Is Life?

- **A living organism is an *open* system made of one or more *cells* going through a *life cycle* during which it *autonomously grows*, *adapts*, *reproduces, eats, metabolizes, eliminates, and "recharges its own batteries"***
- **A living cell is the minimum *quantum of life*. A cell is a dynamic system the like of which doesn't exist in machine-land. In a cell, elementary particles, atoms and molecules constantly flow in and out of the organism**
- **A cell can only "self-assemble" within another cell in its entirety *while the organism is running*. It takes a cell to make another cell, prompting the *unanswered* question: "How did the first cell self-assemble?"**
- **A cell is a goal-oriented organism that exchanges *energy, information and structural elements* with the environment. It manipulates elementary particles, atoms, and molecules *individually.* A cell is not a *statistical* system either classical or quantum**
- **Cells have been studied primarily as biochemical, closed, dead, and classical machines. A few scientists believe they are *open, conscious systems doing quantum information processing of a kind we have yet to understand***

What Is Consciousness?

- **In our experience we partake of two realities: the inner world of sensations, feelings, emotions, and thoughts (four *distinct* families of qualia) and the outer world of "objects moving in space and time"**
- **The outer world is brought into the inner world to be *perceived* and *comprehended*. Therefore, a third-person experience is actually a first-person experience of "outer information" processed by the human sensory-brain system**
- **Consciousness is the capacity to have an inner *sentient* experience based on qualia. Without consciousness it would be impossible for each of us to *know* anything about our own existence and the existence of the outer world**
- **I know I am conscious but I cannot prove it. No outer measurement can do it. Therefore, when I say that someone or something is or isn't conscious, I am making an unprovable *attribution***
- **Consciousness is generally attributed to living organisms with a brain, but not to living cells, trees, or inanimate objects. But where is the *boundary* between them?**

What Is Human Intelligence?

- **Human intelligence is the capacity to *comprehend* our *inner and outer* realities and translate that comprehension into behavior *appropriate* to each situation**
- **Intelligence is *both* an *individual* and a *social* affair: we comprehend ourselves and the world by translating inner meaning into outer symbols, and vice versa, as we *interact* with each other and with the world**
- **Most scientists believe that human beings are machines and intelligence is defined in such a way that the difference between humans and machines is insignificant. Intelligence is often attributed to "reasoning"**
- **I believe that the crucial aspect of human intelligence is the part responsible for any *new* comprehension, invention, inspiration, intuition, creation. That part is not "mechanical." It requires consciousness**

Where Is The Difference?

- **Living organisms are completely different than machines but are studied with a physicalist perspective as if they were *dead* machines**
- **When describing AI, we use the same terms used for humans: learning, perception, memory, understanding, recognition, information, symbols, etc.**
- **These terms are intimately bound up with our conscious inner reality in ways not yet understood. When applied to machines, we end up attributing inner reality and comprehension to machine where there is none**
- **When we study something, we normally place boundaries around unbounded aspects of reality. We close what is open thus turning a *holistic* system into a *reductionist* one**
- **We tend to call "random" or "junk" what we do not understand and in our desire to control, we turn probabilistic systems into deterministic ones**
- **In short, we forget the "wave" and turn everything into "particles"**

- **In my view, the crucial aspect of human intelligence (HI) is *comprehension*. Comprehension is a property of *consciousness* that varies with people**
- **Comprehension is a flash of insight, a flash of *meaning*, a flash of knowing, appearing in our consciousness *before* we can translate "it" into symbols. Comprehension is not a symbol. It is about getting the meaning**
- **The best of us get *new* comprehensions and translate them into *new* symbols for the rest of us to understand, thus advancing society**
- **A robot can "learn" and perform pattern *recognition*, but machine pattern recognition is just a blind symbol-to-symbol conversion. It is a *transduction* without comprehension**
- **Humans also perform transductions like machines, but in addition they perform recognition based on conscious comprehension. The truly creative aspects of HI are *not* accessible to a robot**

Information and Symbols

- **One of the concepts that are often improperly applied to living organisms and to human intelligence is *Shannon information***
- **Shannon defined the "quantity of information" carried by a symbol s_i as $-\log_2 p_i$ where p_i is the probability that s_i might appear next in a string of symbols**
- **Shannon information is *abstract information*, information without meaning, appropriate only to computers and communication systems that are not conscious**
- **Shannon also defined *information entropy* as the average quantity of information carried by a string of symbols**
- **Since the negative of information entropy has the same form as *thermodynamic entropy*, physicists have successfully applied "information" to physical systems**
- **When scientists speak of information in living systems, they mean Shannon information. This usage *ipso facto* reduces life and physical reality to a mechanism**

- **In the reality described by abstract information, consciousness *should not exist* in either physical reality, computers, or living organisms**
- **But each of us is conscious, even though no external measurement can prove it. The proof of consciousness is only *within* each of us. *Knowing* can only come from within**
- **Physics describes a meaningless world because the meaning has been defined away – not because the world is meaningless**
- **The existence of consciousness – what gives meaning to existence – cannot be explained with abstract information: meaningless symbols can only be converted into other meaningless symbols. That's all. Physics without consciousness is incomplete**
- **We need to consider the real possibility that ontology and epistemology are inseparable in the physical world as they are within us. Therefore, to explain the nature of consciousness requires a new conceptual framework about the nature of reality**

Conclusion

Physics describes a *virtual reality* world where consciousness cannot exist

Most scientists are not yet taking seriously the *existence* of consciousness and the *crucial fact* that it is consciousness that gives *meaning* to existence. Explaining the nature of consciousness is not somebody else's problem. Only foundational science can solve this riddle!

Blinded by our physicalist assumptions, we do not to see the elephant in the room. We believe that the solution is to improve by 10^6 or 10^{12} how many petaflops per watt we can achieve with our technologies

I believe that consciousness goes way beyond computation and artificial NNs, and that to understand it will generate a profound revision of physics and of our current physicalist worldview

I also believe that the scientific study of consciousness will bring us closer to discovering our true human dimension, helping us solve the *real* problems that humanity is facing