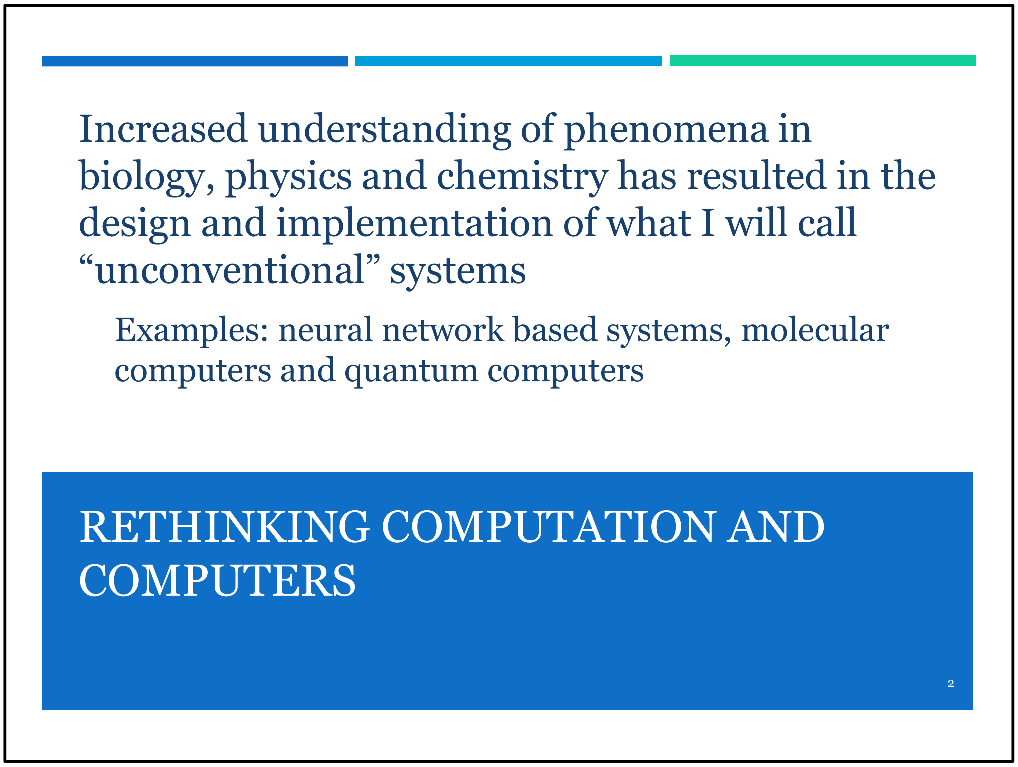
UNCONVENTIONAL COMPUTER SYSTEMS

UB'S COMPUTER SCIENCE AND ENGINEERING'S 50TH ANNIVERSARY

BRUCE SHRIVER

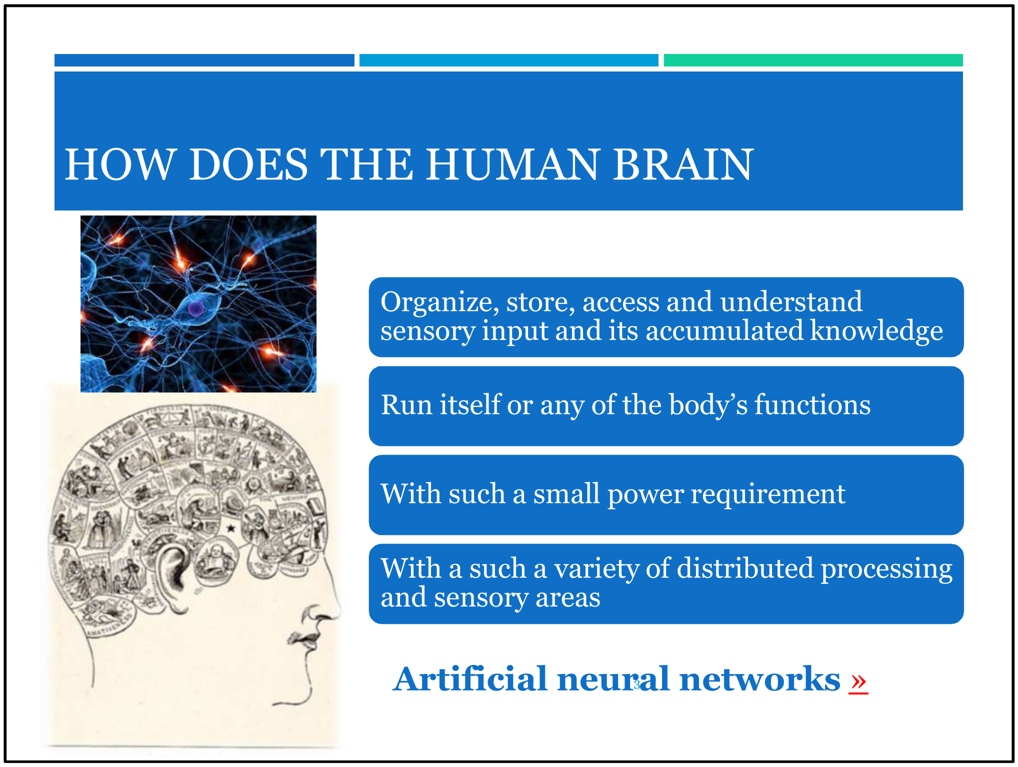
SEPTEMBER 2017



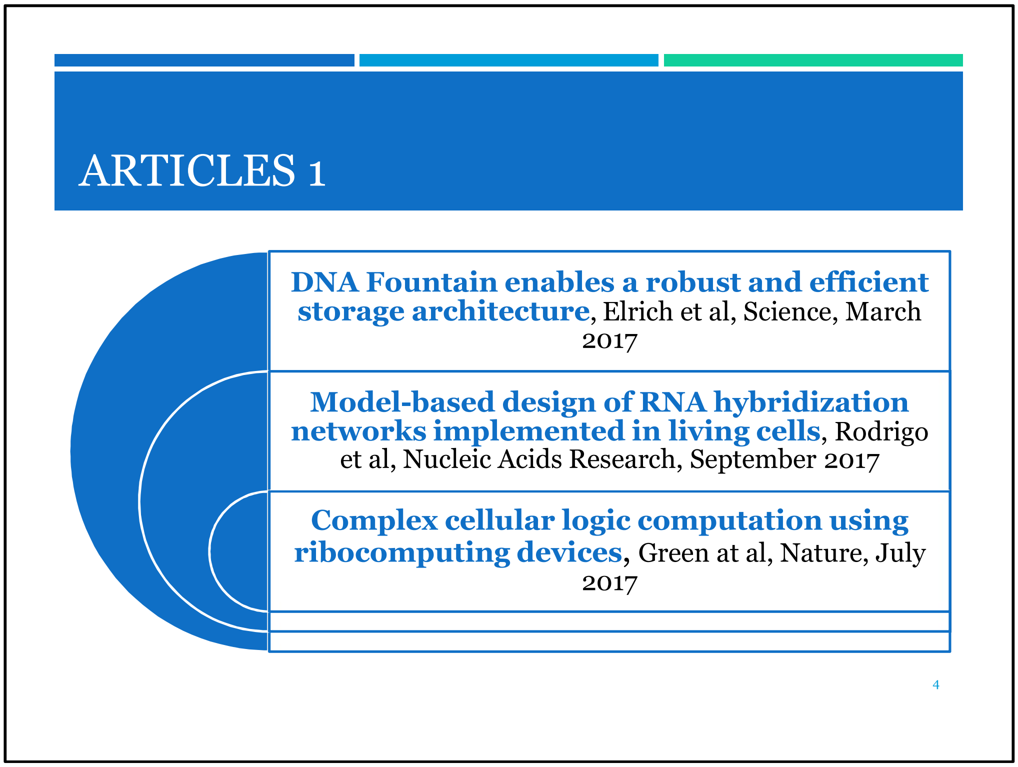
Such systems are "unconventional" in that they are not based on the abstract machine model of Turing or on the electronic computer model which has become know as the Von Neuman Architecture.

When he was involved in designing a silicon retina and cochlea, Carver Mead said, "Engineers would be foolish to ignore the lessons of a billion years of evolution." Mead. as you know, is a brilliant engineer and scientist. He worked with Hopfield and Feynman, helping to create three new fields: Neural Networks, Neuromorphic Engineering, and the Physics of Computation. He is considered the founder of Neuromorphic Engineering and is credited with coining the term "neuromorphic processors.

"IEEE Rebooting Computing is a global initiative launched by IEEE that proposes to rethink the concept of computing through a holistic look at all aspects of computing, from the device itself to the user interface." = from Wikipedia



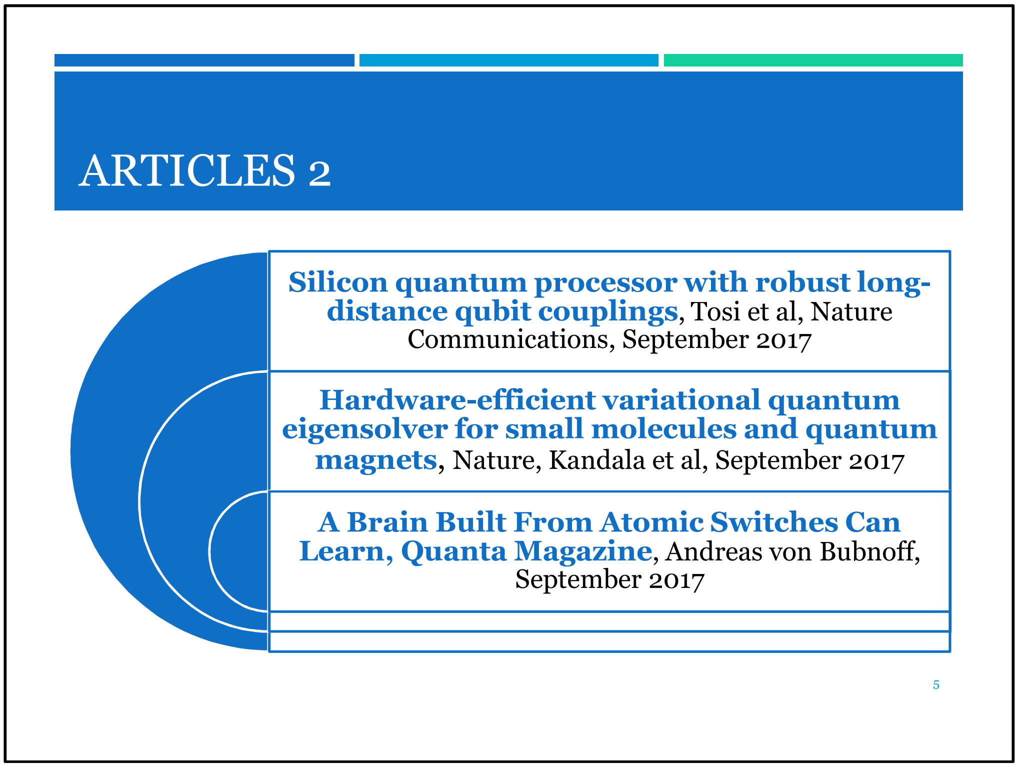
One of the most widely known "biologically inspired" computational paradigms is an artificial neural network. It seems as though every few weeks the press has an article on deep learning, language translation, medical diagnosis, data mining or some other application that employs neural networks. ANNs arose out of our seemingly never ending quest to understanding the underpinning of the human brain.



Paper 1: An operating system and a movie were stored on and retrieved from DNA using an erasure-correcting algorithm called fountain codes. Using this coding strategy, the researchers could pack 215 petabytes of data on a single gram of DNA—100 times more than methods developed by other researchers.

Paper 2: Researchers in the UK used synthetic biology techniques to modify ribonucleic acid (RNA) molecules to specify actions within a cell. They demonstrated that cells can be programmed with pre-defined RNA commands in a manner "similar" to programming a computer.

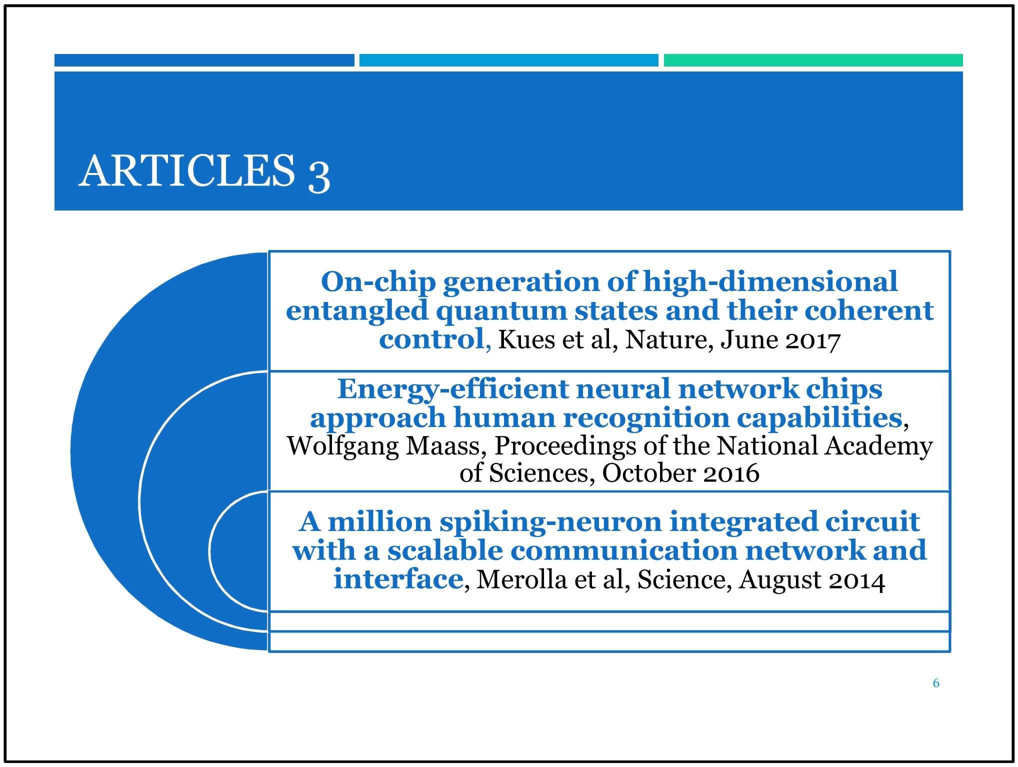
Paper 3: Researchers at Harvard's Wyss Institute for Biologically Inspired Engineering developed an RNA molecule which can sense multiple signals and make logical decisions to control protein production. Such programmable nano-devices allow the construction of more sophisticated synthetic biological circuits.



Paper 1: Australian and US researchers develped a “flip-flop qubit” which combines electron and nuclear spin states. This enables neighboring qubits to remain "coupled" over larger physical distances allowing implementation of simpler control schemes for the large arrays of qubits.

Paper 2: These researchers at IBM T J Watson Research Center implemented a new quantum algorithm capable of efficiently computing the lowest energy state of small molecules. This work was done on a seven qubit quantum processor.

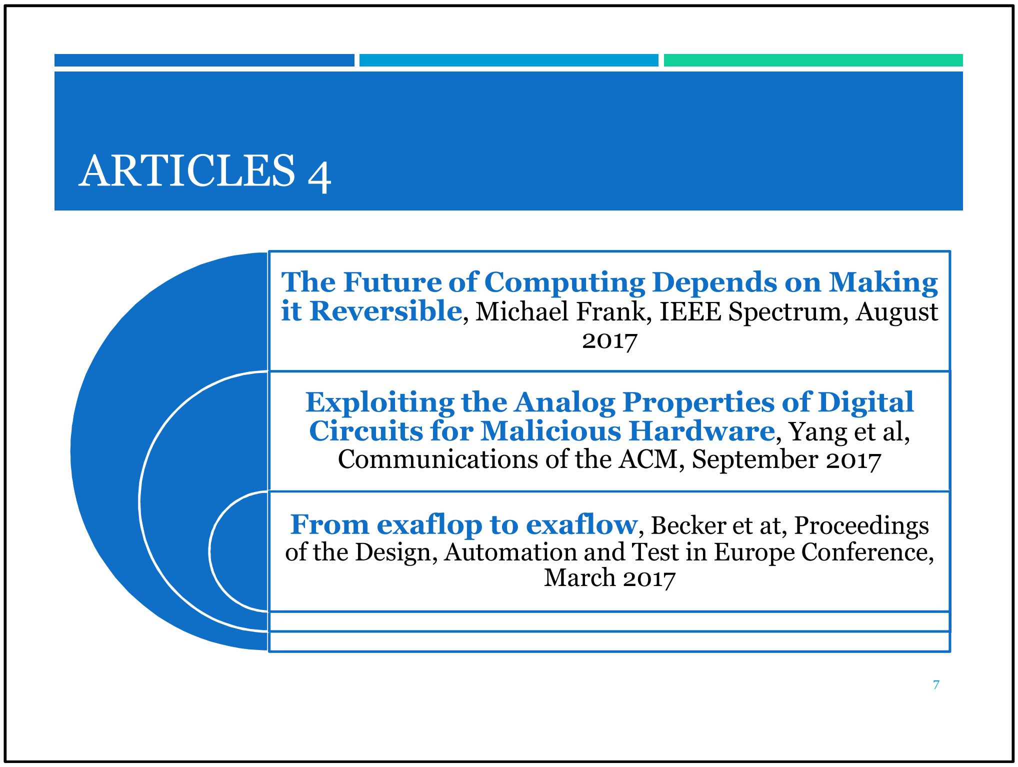
Paper 3: Summarizes work done at the California NanoSystems Institute (CNSI) at UCLA in which researchers have made a 2-millimeter-by-2-millimeter mesh of silver nanowires connected by artificial synapses with a potential density of 1 billion artificial synapses per square centimeter. The device organized itself out of random chemical and electrical processes.



Paper 1: Researchers at Canada's National Institute of Scientific Research (INRS) fabricated a microchip that can generate two entangled qudits each with 10 states, for 100 dimensions total, more than the 64 dimensions that six entangled qubits could generate. The photonic chip was fabricated using techniques similar to ones used for integrated circuits. A laser fires pulses of light into a micro-ring resonator, a 270-micrometer-diameter circle etched onto silica glass, which in turn emits entangled pairs of photons. Each photon is in a superposition of 10 possible wavelengths or colors.

Paper 2: This is a short review paper of neural network chips and learning strategies circa 2016 and, if you haven't read anything in this area in a couple of years, might provide a quick refresher.

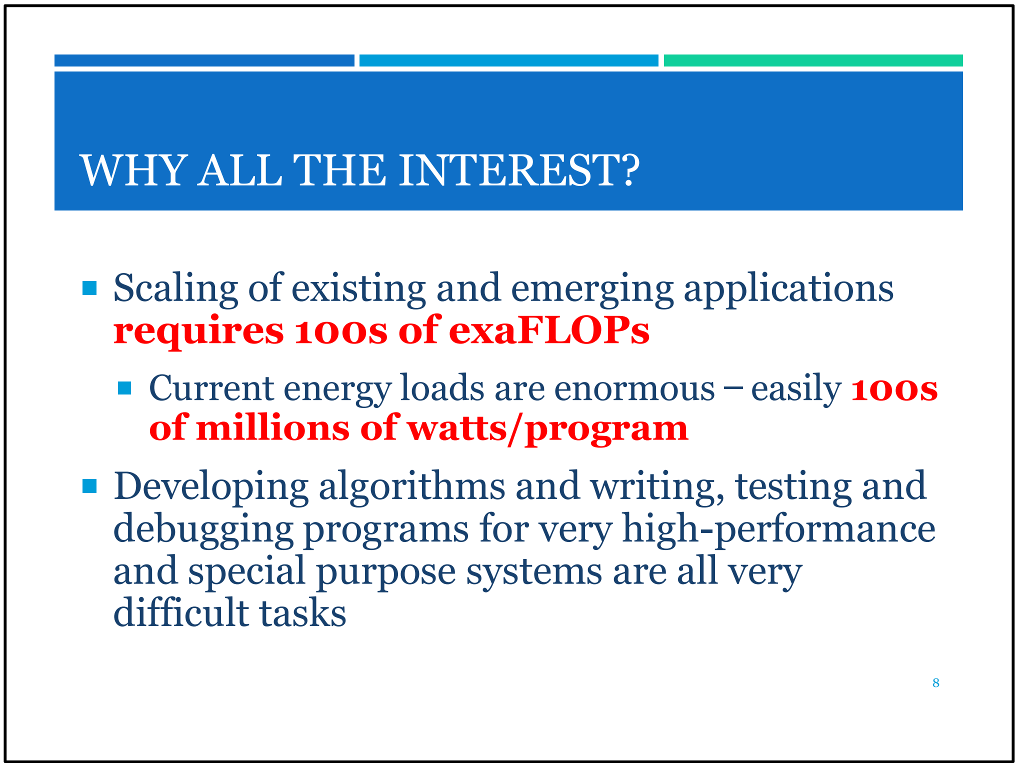
Paper 3: American and Japanese researchers fabricated a 5.4-billion-transistor chip with 4096 neurosynaptic cores interconnected via an intrachip network integrating 1 million programmable spiking neurons and 256 million configurable synapses. They state that, "The architecture is well suited to many applications that use complex neural networks in real time, for example, multiobject detection and classification. With 400-pixel-by-240-pixel video input at 30 frames per second, the chip consumes 63 milliwatts."



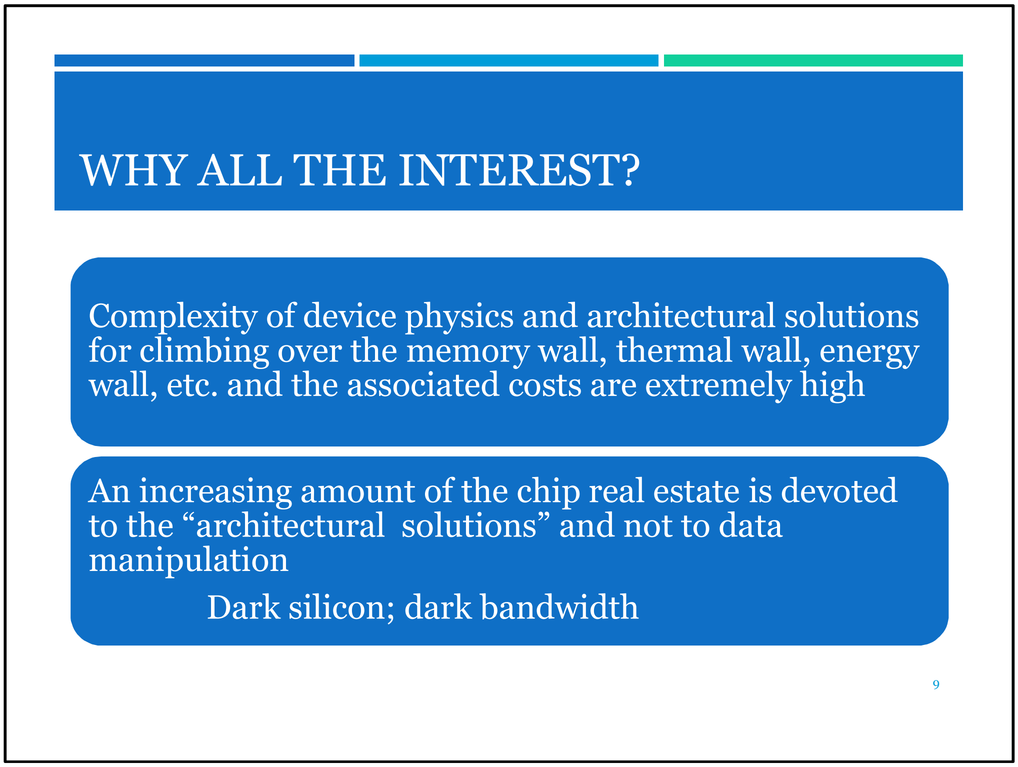
Paper 1: Provides a short and interesting history of reversible computing and some of the problems which lie ahead in attempting to design and implement reversible circuits and devices (also called adiabatic). The history (and it's many hyperlinked references) make it a good starting point if you are interested in this topic. There is also a 2011 MIT Technology Review article worth reading called "The Fantastical Promise of Reversible Computing".

Paper 2: The authors state in their conclusions: "Experimental results with our fabricated malicious processor show that a new style of fabrication-time attack is possible, which applies to a wide range of hardware, spans the digital and analog domains, and affords control to a remote attacker." They go on to state, "We believe that our results motivate a different type of defense, where trusted circuits monitor the execution of untrusted circuits, looking for out-of-specification behavior in the digital domain."

Paper 3: This is basically a description of Maxeler's hybrid dataflow system. The authors state, "Our model splits the computation into a conventional CPU-oriented part and a highly efficient fully programmable data flow part. We present a number of systematic transformations and optimisations targeting Maxeler dataflow systems that typically yield one to two orders of magnitude improvements in terms of both performance and energy efficiency." For a different dataflow approach see "A Coarse Grain Reconfigurable Array (CGRA) for Statically Scheduled Data Flow Computing", Chris Nicol, Wave Computing, 2017



IBM’s Watson employs IBM's DeepQA technology which it uses to generate hypotheses, gather massive evidence, and analyze data. It it is “workload-optimized”. The hardware employs highly parallel processors and a cluster of 90 8-core servers with four threads per core yielding a system with 2,880 processor threads and 16 terabytes of RAM. It’s estimated that Watson's hardware a bit more that three million dollars. According to IBM, “Watson can process 500 gigabytes, the equivalent of a million books, per second. Its Linpack performance stands at 80 TeraFLOPs, which is about half as fast as the cut-off line for the Top 500 Supercomputers list.” According to John Rennie ("How IBM’s Watson Computer Excels at Jeopardy!", PLoS), “all content was stored in Watson's RAM for the Jeopardy game because data stored on hard drives would be too slow to be competitive with human Jeopardy champions.”



Dark siliconDark Silicon: "ark silicon is the amount of circuitry of an integrated circuit that cannot be powered-on at the nominal operating voltage for a given thermal design power (TDP) constraint. This is a challenge in the era of nanometer semiconductor nodes, where transistor scaling and voltage scaling are no longer in line with each other, resulting in the failure of Dennard scaling. This discontinuation of Dennard scaling has led to sharp increases in power densities that hamper powering-on all the transistors simultaneously at the nominal voltage, while keeping the chip temperature in the safe operating range." -Wikipedia

Dark Bandwidth: "The technical community has been working on a false thesis, Beard says. “That compute systems efficiently utilize the bandwidth provided to the core by the memory subsystem. Not only does research suggest that this is often not the case, it shows that many applications make use of only a small fraction of the data moved to the compute elements. The trend towards heterogeneous accelerators only serves to exacerbate the problem as bus length and buffering increase. Recovering this lost bandwidth and reducing superfluous data movement gives the system more usable bandwidth for real computing. It is not just the size and speed of the memory technology that matters, it is how you use it that will enable future systems to utilize what is effectively dark or hidden memory bandwidth.” -- Shedding Light on Dark Bandwidth, Nicole Hemsoth, The Next Platform, September 2017.

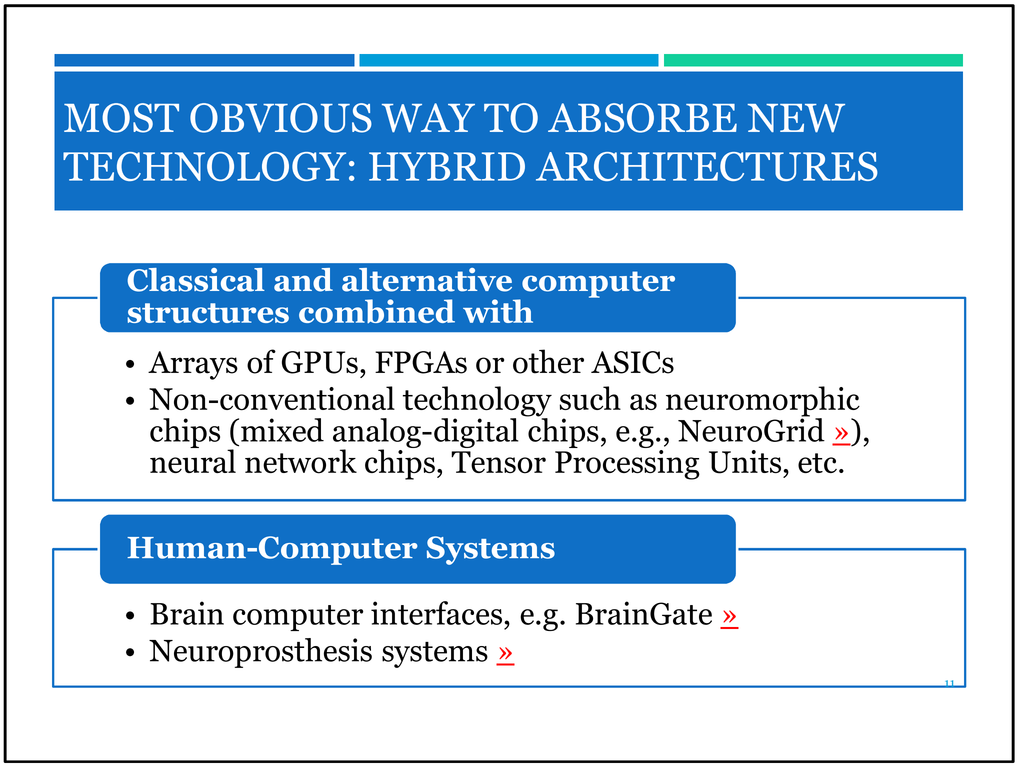
Also see:

1. Is multicore hardware for general-purpose parallel processing broken?, Uzi Vishkin, Communications of the ACM, April 2014

2. The Future of Semiconductors, Samuel Greengard, Communications of the ACM, March 2014

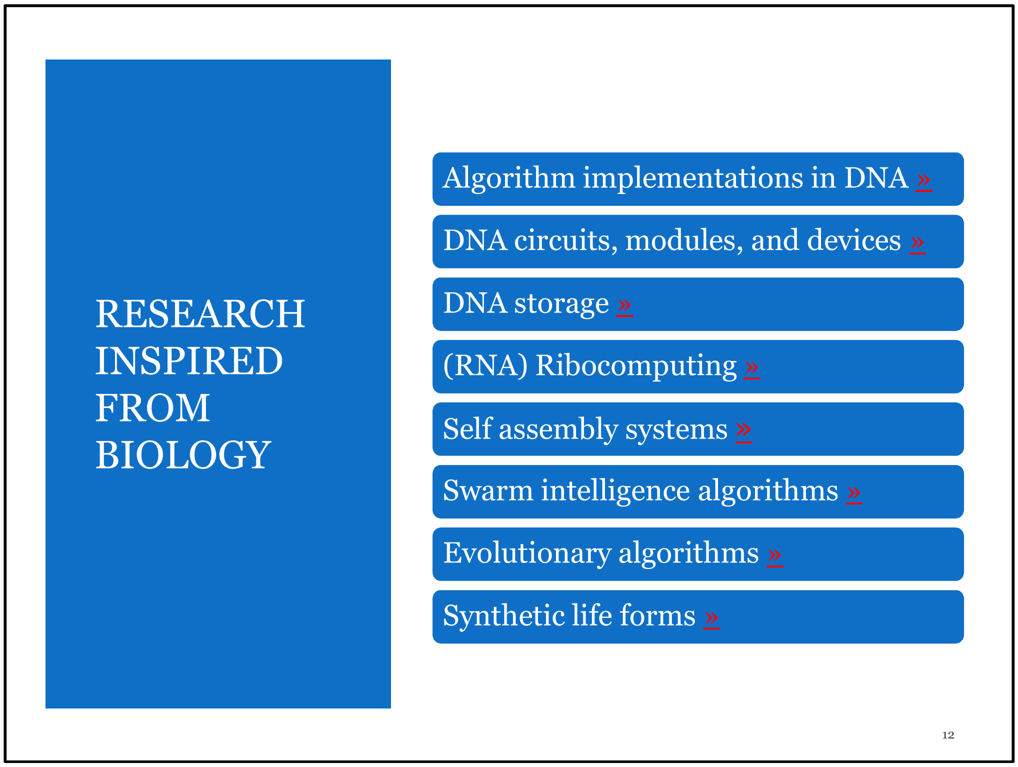
UNCONVENTIONAL SYSTEMS HAVE GOTTA BE EXPENSIVE, RIGHT?

- This type of research requires experts from several disciplines working closely together
  - Molecular, systems and synthetic biologists, computer scientists, mathematicians, engineers, physicists, chemists
- Cross-Department/Cross-School appointments, degree programs, lab facilities/infrastructure, research fellows, IP, open source, etc.



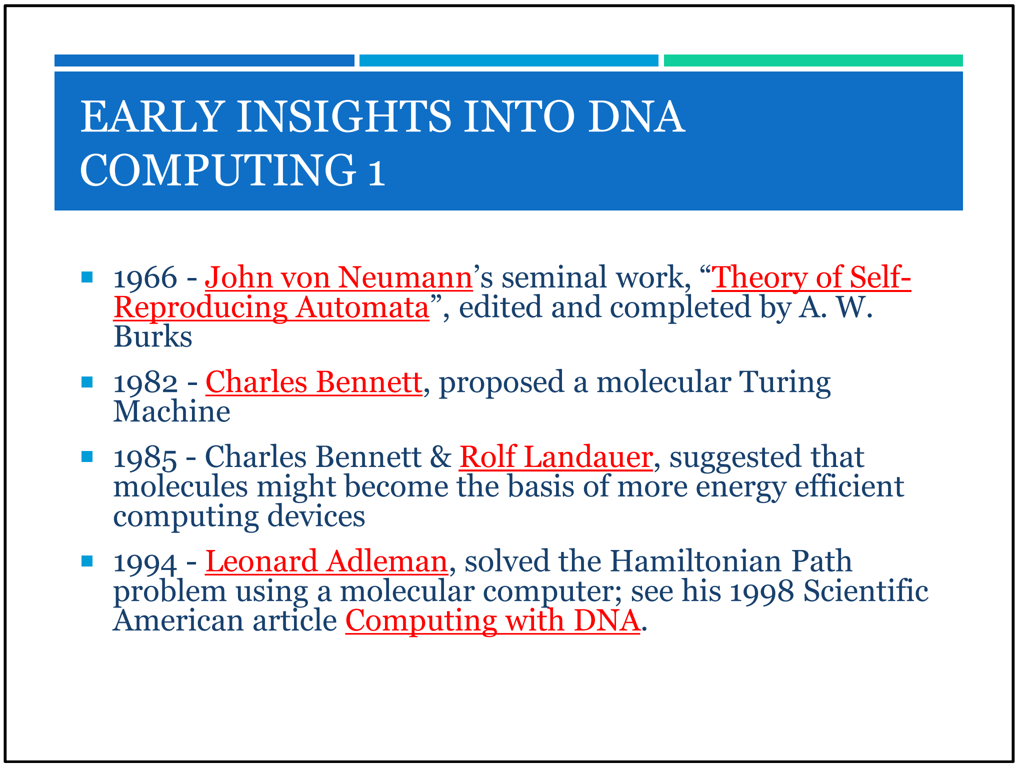
Tensor processing units (or TPUs) are application-specific integrated circuits (ASICs) developed specifically for machine learning.

The European Union is funding the development of the EuroEXA project to build an HPC based on ARM Cortex processors and Xilinx Ultrascale FGGAs with the goal of implementing an energy-efficient petaflops system by 2020. The EuroEXA project is to become the foundation for realizing exascale capability within the 2022-2023 timeframe.



Initial Computer Developments are often limited by the I/O mechanisms available.

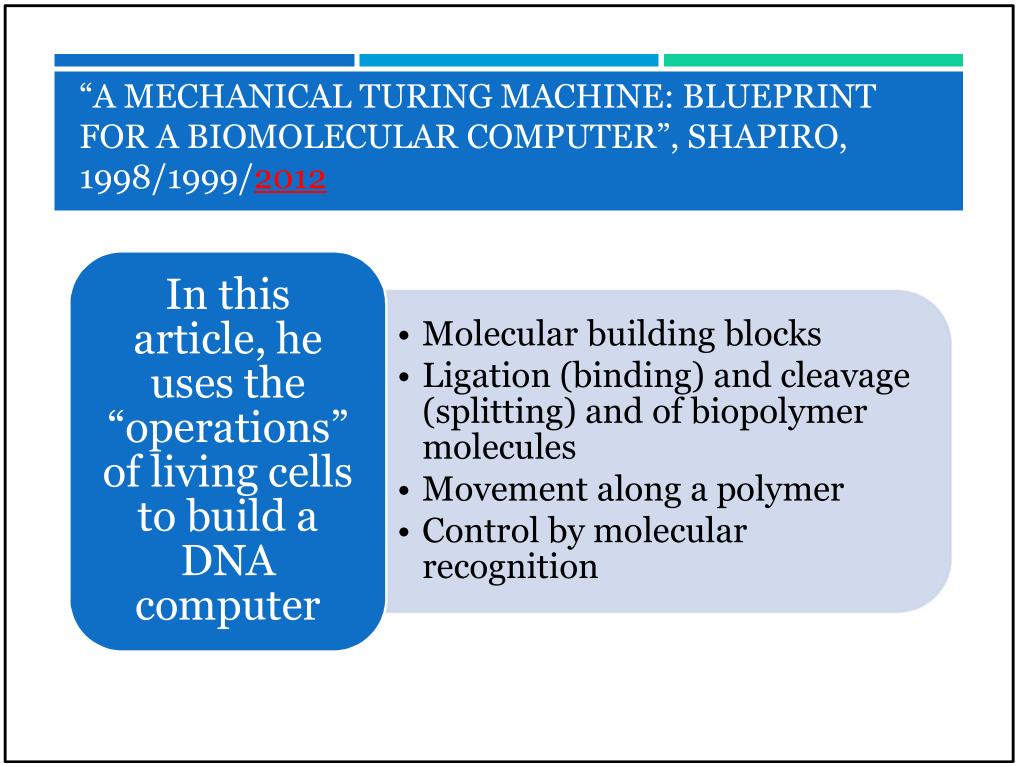
In vitro (in a controlled environment outside of a living organism, in a test tube)
In vivo (in living organisms)
In silico (computer simulation)



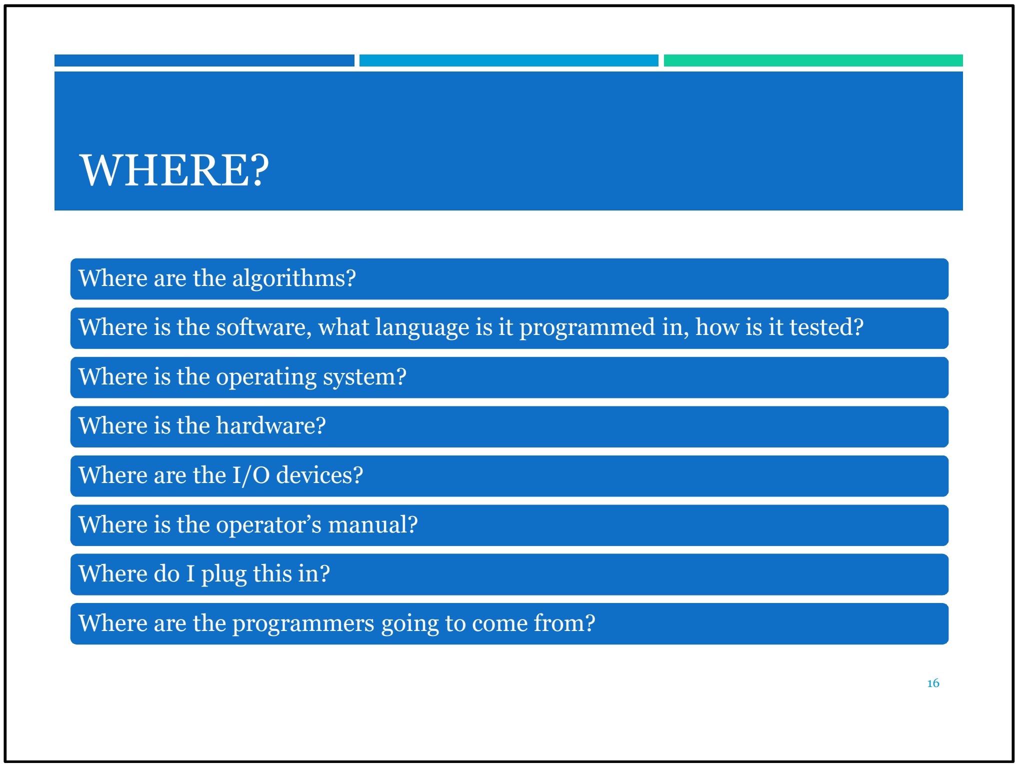
Current synthetic molecular circuits can only deal with (i.e., sense) a small numer of biological signals. Thus they deal with a very incomplete set of conditions within the extremely complex workings of the cell. The circuits contain a variety of parts in the form of different types of molecules, proteins and enzymes which must "work" together. It is very much a "wet lab environment". The to sense and process signals. Identifying m development of of new biological circuits is both time-consuming and difficult.

EARLY INSIGHTS INTO DNA COMPUTING 2

- 1996 - The first DNA Based Computers meeting held at Princeton University (DIMACS DNA Based Computers)
- 1999 - Ehud Shapiro, “A Mechanical Turing Machine: Blueprint for a Biological Computer”. This work is presented in the 2006 Scientific American article “Bringing DNA Computers to Life”
- 2000 - Dirk Faulhammer, Anthony Cukras, Richard Lipton and Laura Landweber solved the Four Knights problem in chess using RNA molecules (article published in the PNAS)



Shapiro: "The Turing machine was perceived for decades as a theoretical widget devoid of practical relevance, especially given the overwhelming success of its younger alternative, the von Neumann stored-program computer architecture. However, as our understanding of molecular cell biology and biochemistry unfolded, it became ever clearer that the concepts underlying the Turing machine are deeply rooted in nature. The Turing machine infinite tape, in which each cell may store one symbol taken from a finite alphabet, cannot be more similar, mathematically, to DNA, a potentially unbounded polymer in which each monomer is one of four letters. Molecular machines such as DNA polymerase, RNA polymerase and the ribosome are most naturally understood as simple finite-state transducers, a special case of the Turing machine."



How do you test, use, and evaluate any unconventional systems?

What is the manufacturing infrastructure required to move unconventional computer systems out of research labs into the commercial marketplace?

Software: molecular and quantum. See, for example, "First quantum computers need smart software", Zeng et al, Nature, 13 September 2017.

CONJECTURES

- Breakthroughs will occur in molecular computing I/O
- Cloud-based access to quantum computing will spur the development of algorithms
- *Augmented reality* will become a tool in multidisciplinary research
- Quantum and Bio-inspired models will complement one another in systems
- Special purpose and application specific architectures continue to be developed unabated »

17

QUESTIONS