CS 152 Computer Architecture and Engineering

Lecture 21: The Future and Closing Remarks

Dr. George Michelogiannakis
EECS, University of California at Berkeley
CRD, Lawrence Berkeley National Laboratory

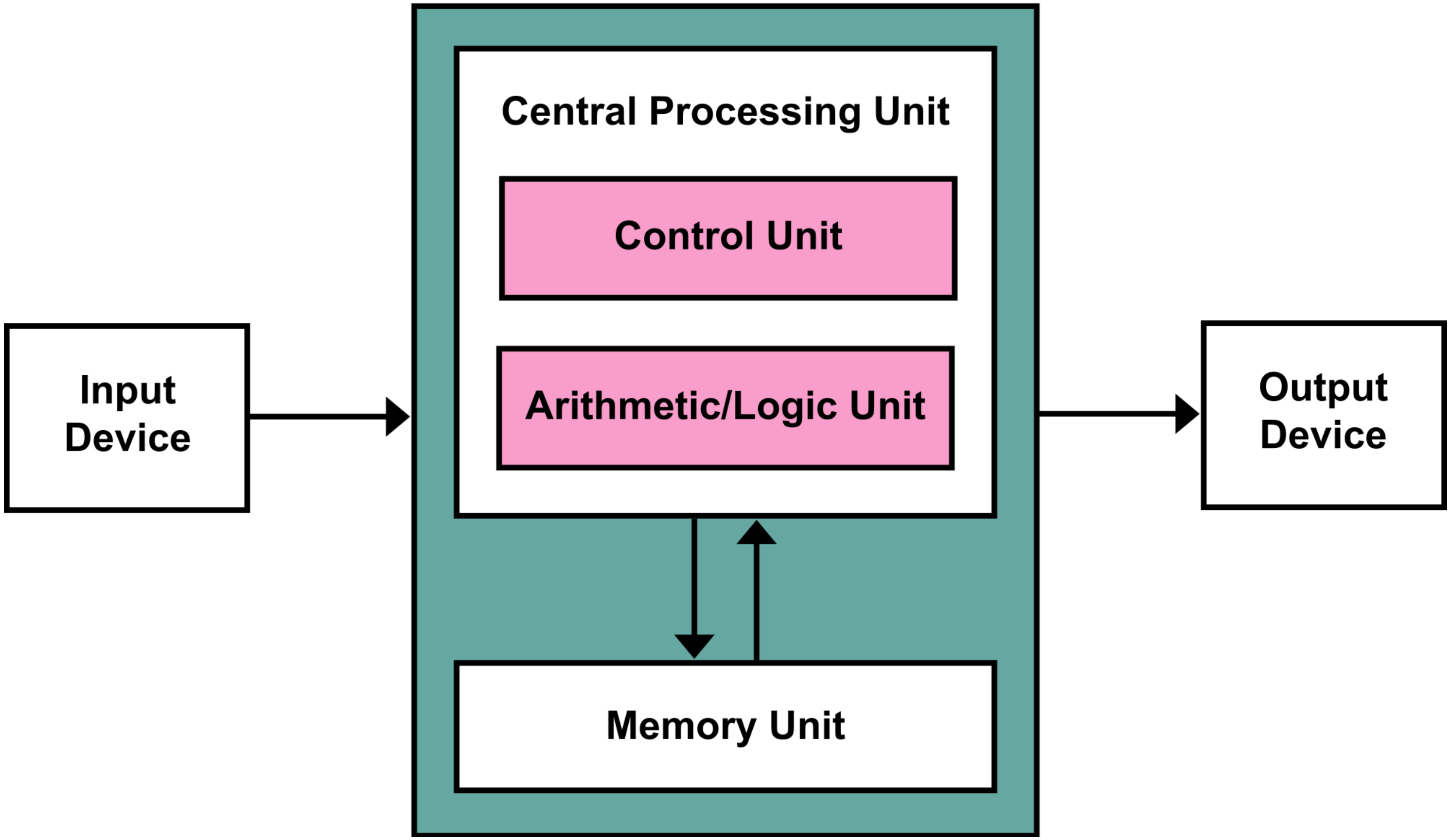
http://inst.eecs.berkeley.edu/~cs152

Administrivia

- Final quiz on Wednesday
  - Show up on time!
- Thursday Colin will have office hours instead of discussion
- Lab 5
- 10:00am today is class evaluation

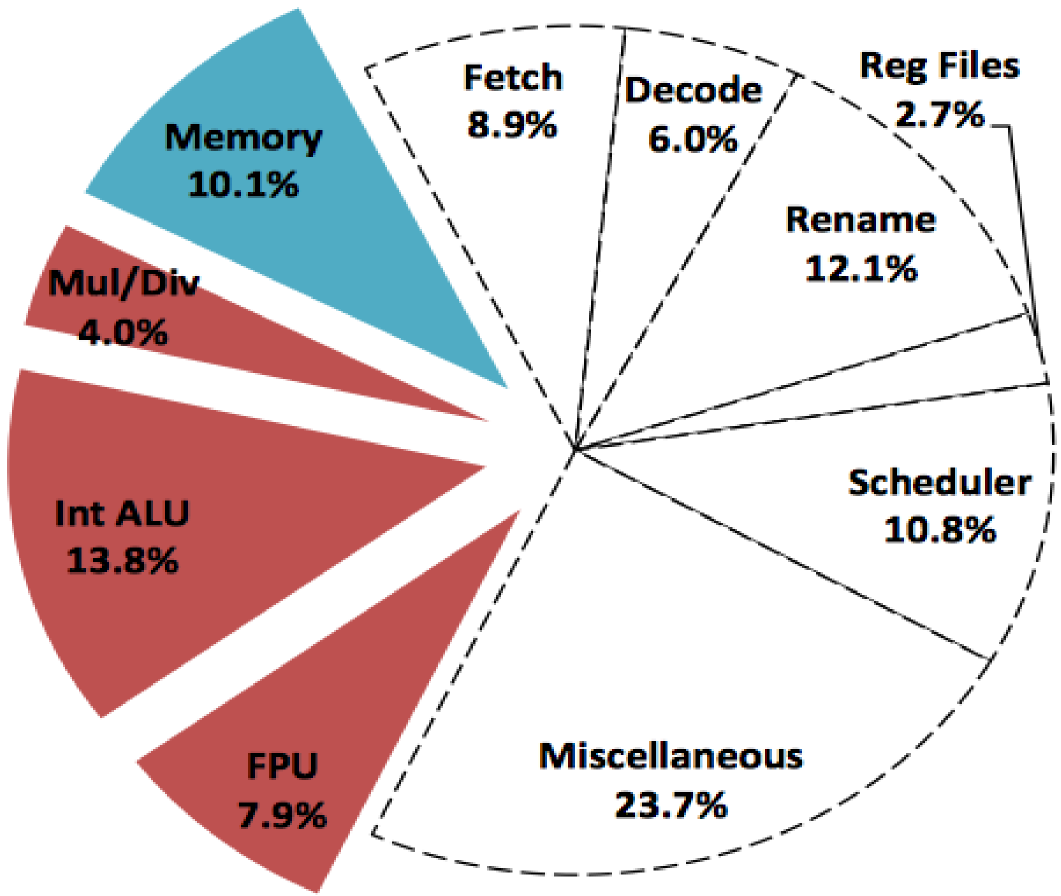
So Far We Talked About

- Von Neumann architectures
- Consist of:
  - ALU, registers, program counter, instruction unit, memory to store data, I/O mechanisms, external mass storage



The Problem With Programmability

- [Cong, 2014]. This is power breakdown for a superscalar out-of-order processor, typically configured

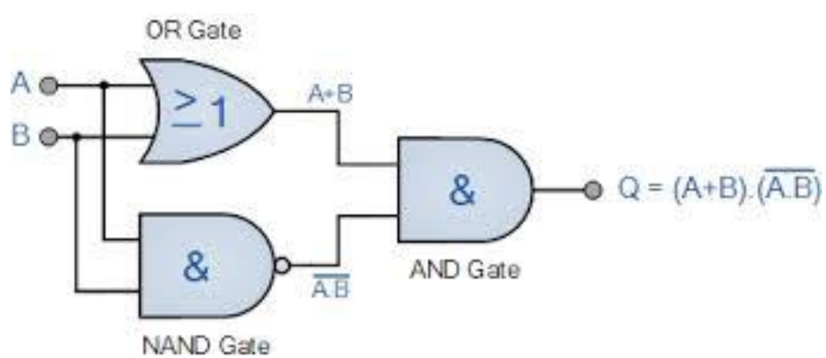


Accelerators

- Specialized hardware that is fixed function
  - One application
  - Or one function common to many applications (e.g., FFT)

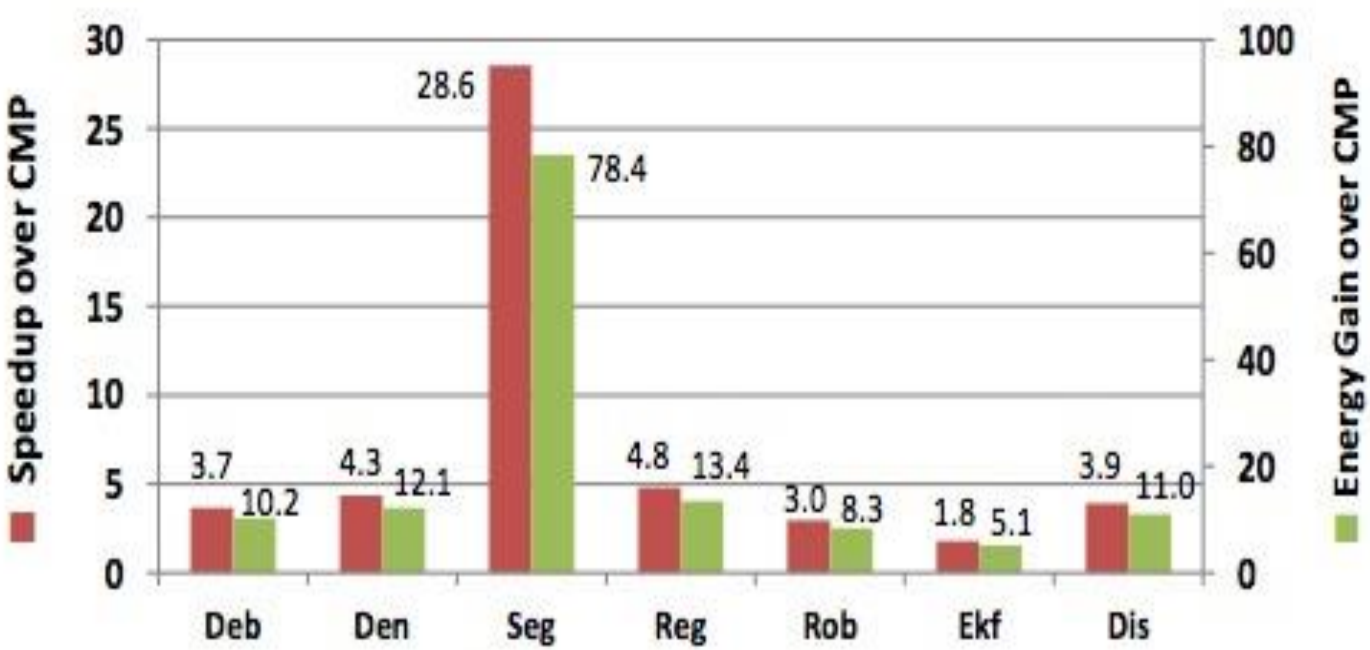
```
typedef unsigned long U32;

U32 cyclic_mac(U32 *p1, U32 *p2)
{
  U32 sum = 0;
  int i;

  for(i = 0; i < BUF_SIZE*4; ++i)
  {
    sum += *p1++ * *p2++;

    if((i % BUF_SIZE) == (BUF_SIZE - 1))
    {
      p1 -= BUF_SIZE;
    }
  }

  return sum;
}
```

```
; Enabling modulo addressing for r0
lbf 0x1, moduen
; Setting modulo factor for r0
lbf 64, modi

; Loop prologue
mpy     (r0).dw+1, (r1).dw+1
mpypa  (r0).dw+1, (r1).dw+1, a0

rep       127
; Loop body
mac      (r0).dw+1, (r1).dw+1, a0

ret{ds1, t}
; Disabling modulo addressing for r0
lbf 0x0, moduen
```

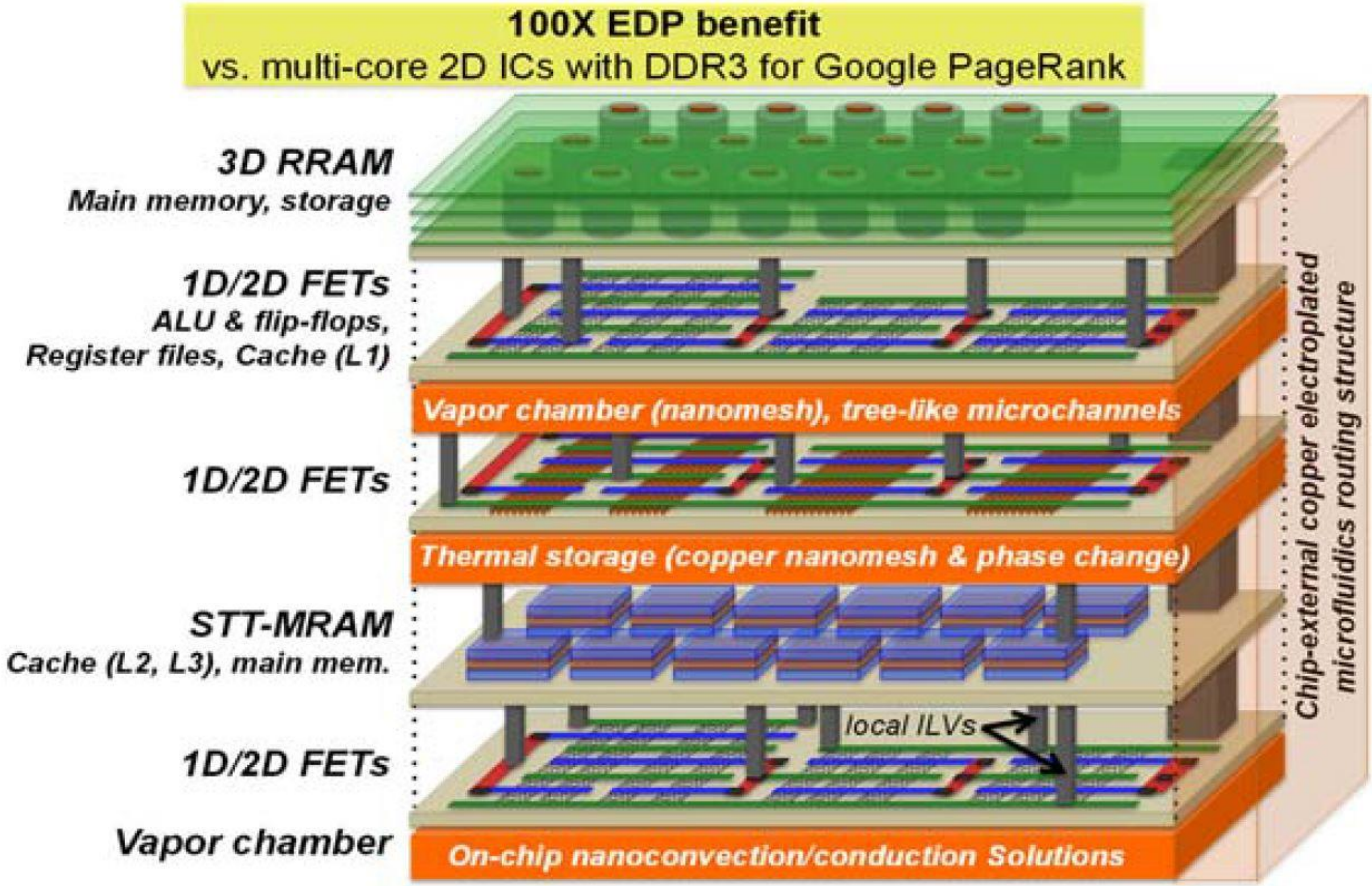


Large Gains for Accelerators

- Custom logic in a FPGA versus a 12-core chip multiprocessor



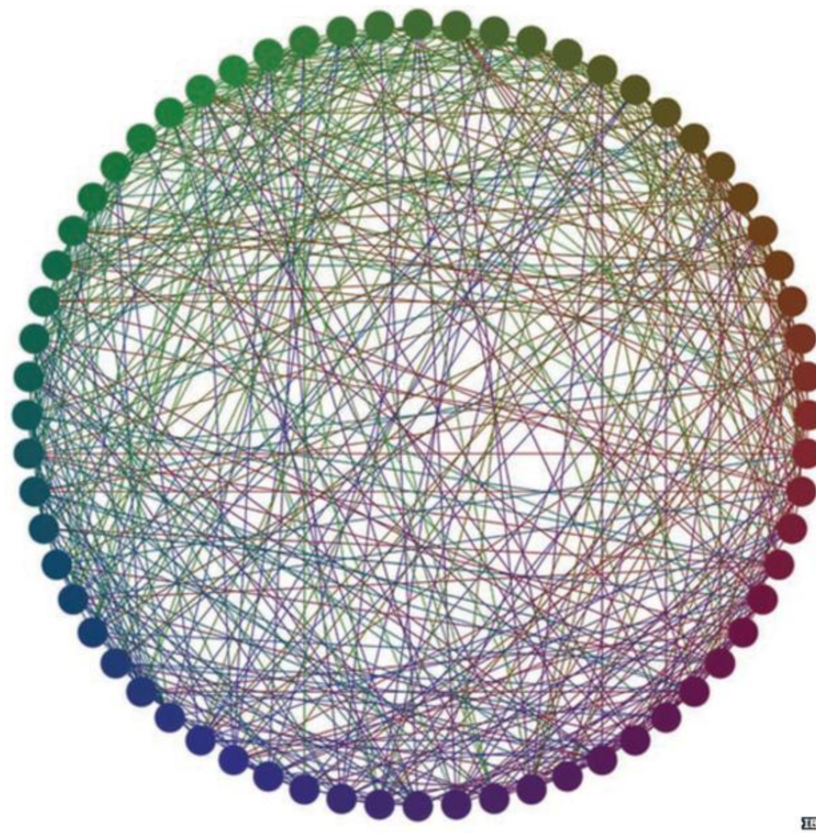
3D Stacking



100X EDP benefit
vs. multi-core 2D ICs with DDR3 for Google PageRank
3D RRAM
Main memory, storage
1D/2D FETs
ALU & flip-flops,
Register files, Cache (L1)
Vapor chamber (nanomesh), tree-like microchannels
1D/2D FETs
Thermal storage (copper nanomesh & phase change)
STT-MRAM
Cache (L2, L3), main mem.
local ILVs
1D/2D FETs
Vapor chamber
On-chip nanoconvection/conduction Solutions
Chip-external copper electroplated
microfluidics routing structure

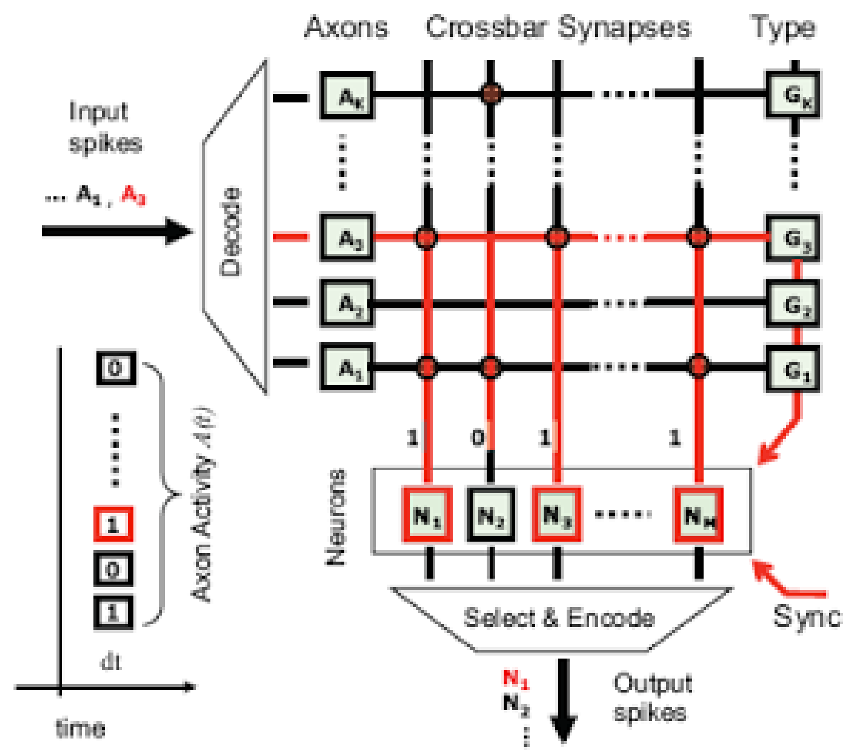
Neuromorphic

- Inspired by the human brain
  - Neurons (cells) and synapses (connections)
- IBM Truenorth consists of 4,096 neurosynaptic cores
  - On-chip mesh interconnect
  - 5.4 billion transistors
  - Yet only 70mW
  - Still digital computing!
- Architecture has neurons and synapses

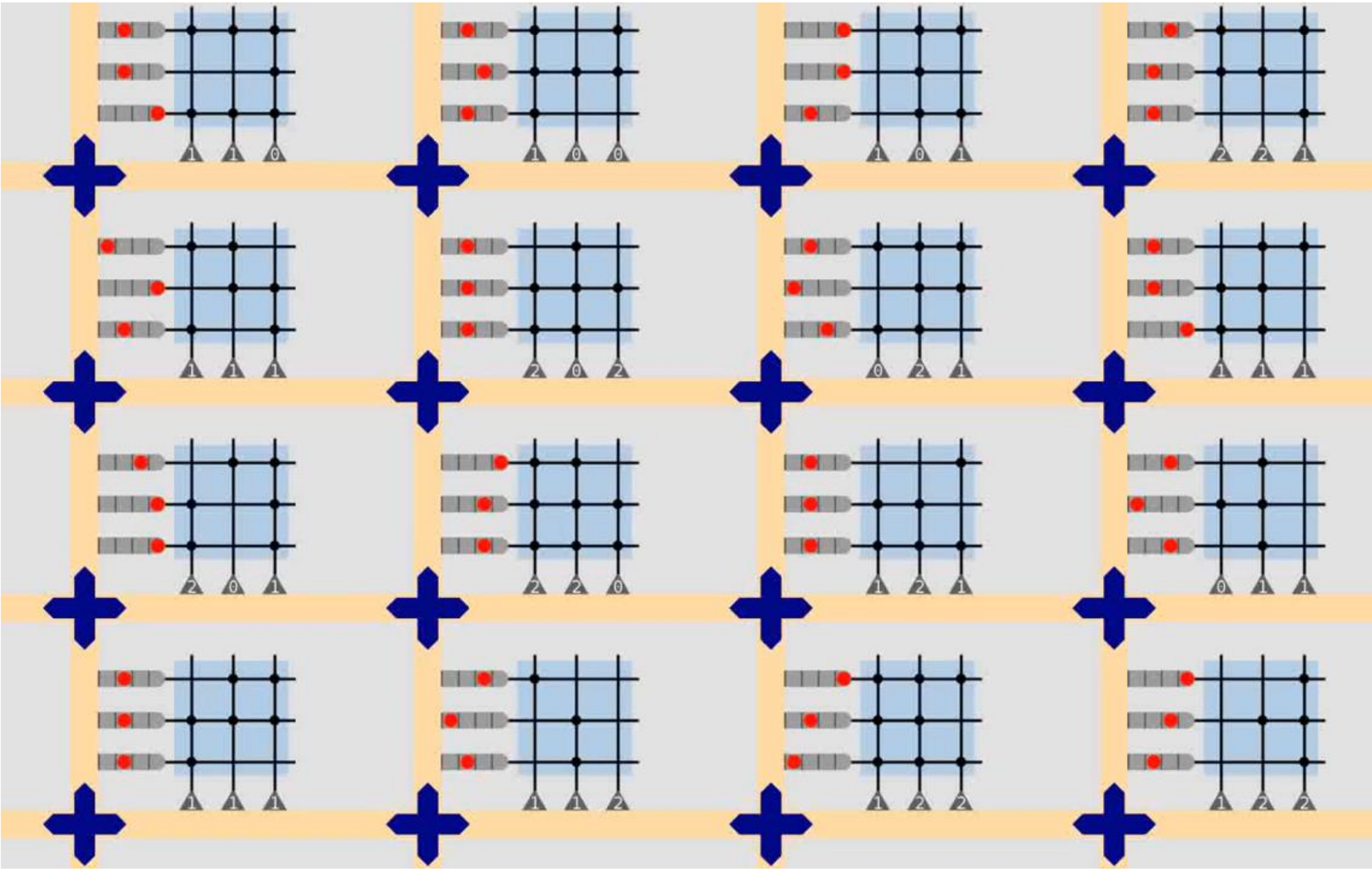


Neurosynaptic Core

- No clock
  - Event-driven operation
- 256 inputs (axons), 256 outputs (neurons), a bank of SRAM to store data for each neuron, and a router
- When a neuron receives the right number of spikes at the right times to match a pre-learned pattern, it sends a spike of its own
  - This spike adds to the pattern seen by other neurons

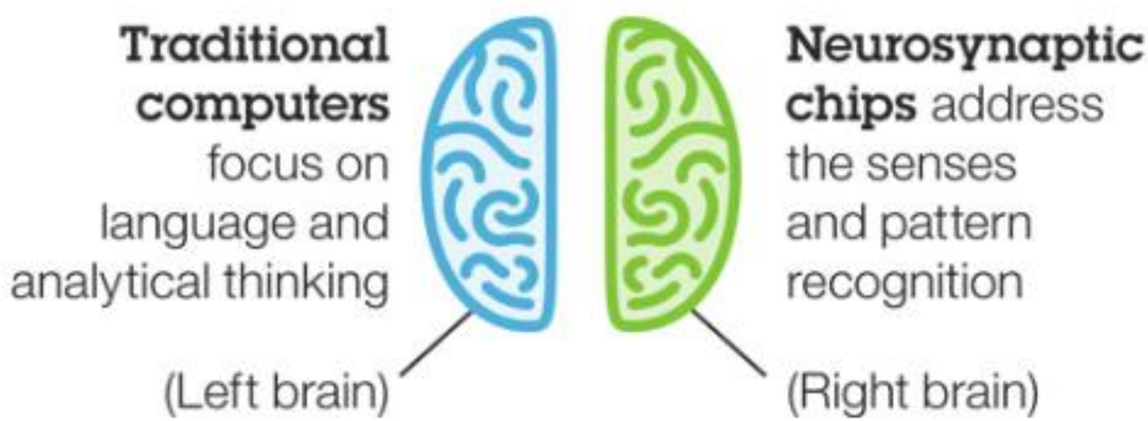


Flow of Data



Training and Applications

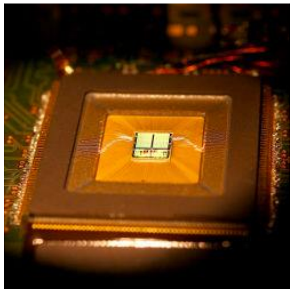
- One downside is training
  - Chips must be trained in advance in a process that can take multiple days in a supercomputer
  - "Synapse university" trains humans how to train neuromorphic chips



Example Application

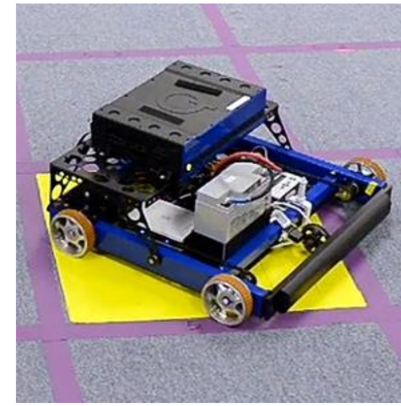
A single chip can process color video in real-time while consuming 176,000 times less energy than a current Intel chip performing the exact same analysis. Note the Intel chip can *not* do this analysis in real-time and is in fact 300 times slower!

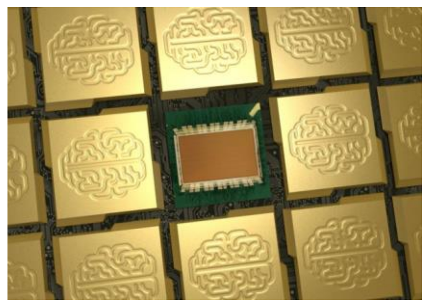
Neuromorphic Computing



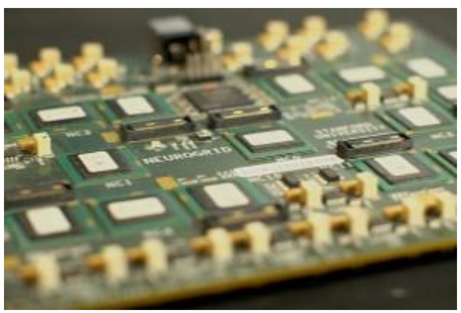
"Spikey" from Electronic Visions group in Heidelberg

Qualcomm's NPU's for robots.



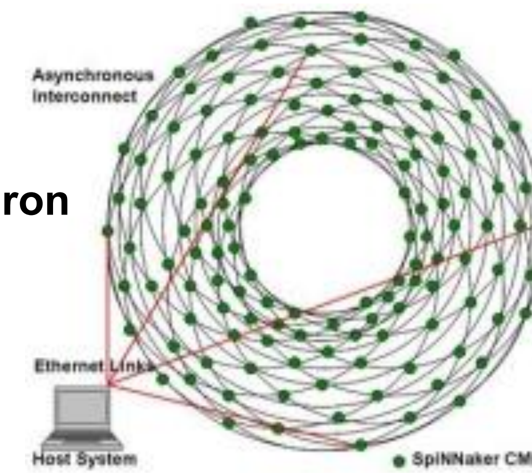


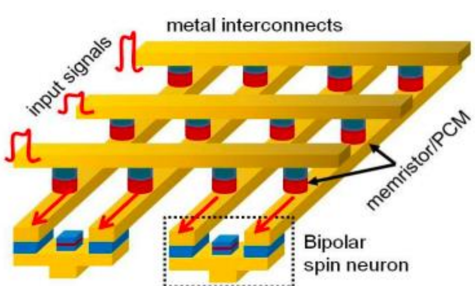
IBM's TrueNorth



Stanford's Neurogrid

SpiNNaker's 1B neuron machine

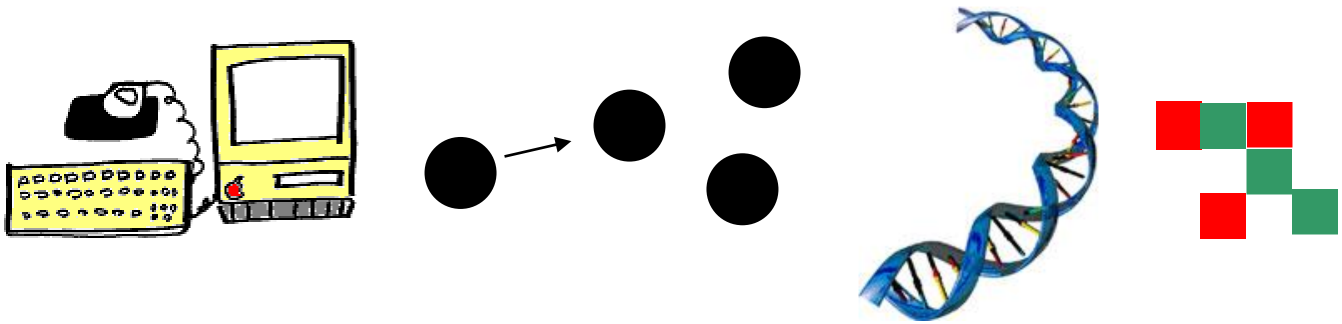




(Peter Nugent, LBNL)

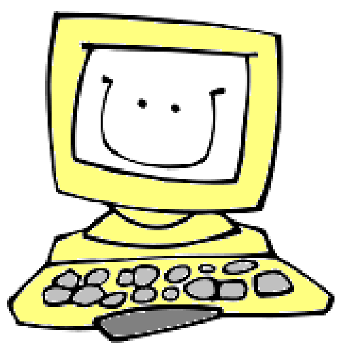
What Is Quantum Computation?

Credit: Neil Shenvi



Conventional computers, no matter how exotic, all obey the laws of

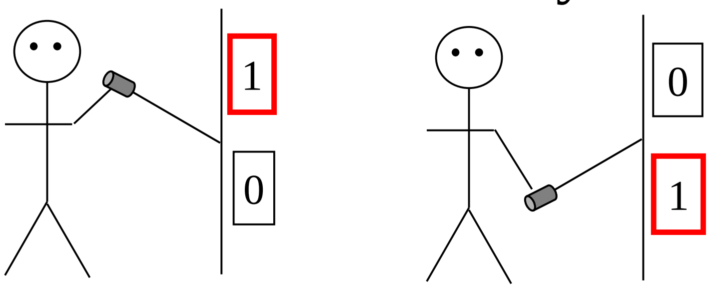
classical physics.



On the other hand, a quantum computer obeys the laws of quantum physics.

The Bit

The basic component of a classical computer is the bit, a single binary variable of value 0 or 1.



At any given time, the value of a bit is either '0' or '1'.

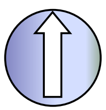
The state of a classical computer is described by some long bit string of 0s and 1s.

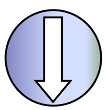
0001010110110101000100110101110110...

The Qubit

A quantum bit, or qubit, is a two-state system which obeys the laws of quantum mechanics.

Spin-½ particle

 $=|1\rangle$

 $=|0\rangle$

Valid qubit states:

$|\psi\rangle = |0\rangle$

$|\psi\rangle = |1\rangle$

$|\psi\rangle = (|0\rangle - e^{i\pi/4}|1\rangle)/\sqrt{2}$

$|\psi\rangle = (2|0\rangle - 3e^{i5\pi/6}|1\rangle)/\sqrt{13}$

The state of a qubit $|\psi\rangle$ can be thought of as a vector in a two-dimensional Hilbert Space, $\mathcal{H}_2$, spanned by the Basis vectors $|0\rangle$ and $|1\rangle$.

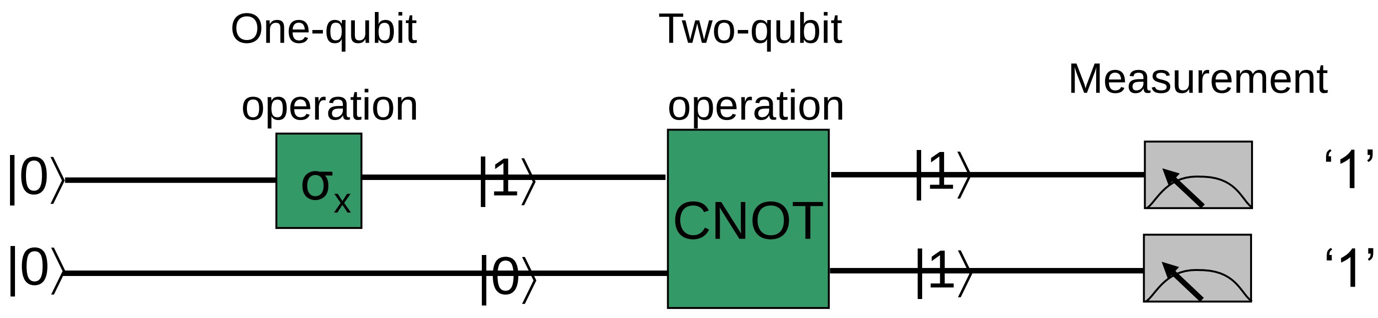
Computation with Qubits

How does the use of qubits affect computation?

Classical Computation		Quantum Computation	
Measurement: deterministic		**Measurement: stochastic**	
State	Result of measurement	State	Result of measurement
x = '0'	'0'	$\lvert\psi\rangle = \lvert 0\rangle$	'0'
x = '1'	'1'	$\lvert\psi\rangle = \lvert 1\rangle$	'1'
		$\lvert\psi\rangle = \frac{\lvert 0\rangle - \lvert 1\rangle}{\sqrt{2}}$	'0' 50% '1' 50%

Quantum Circuit Model

Example Circuit

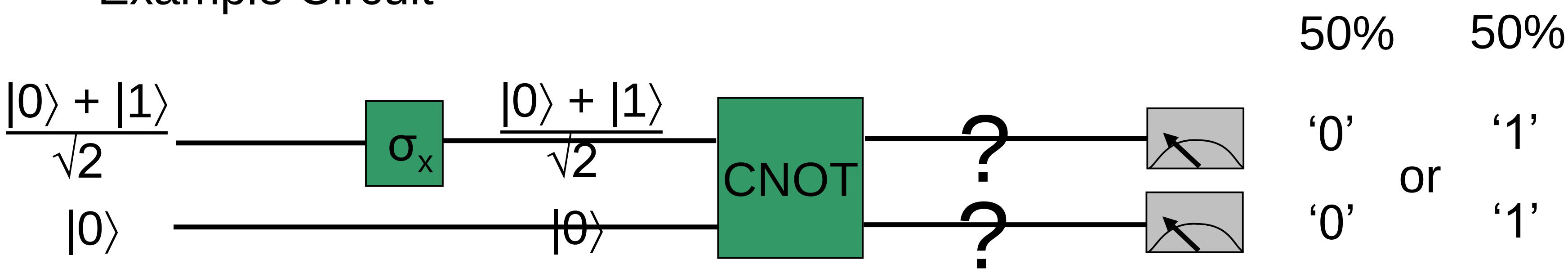


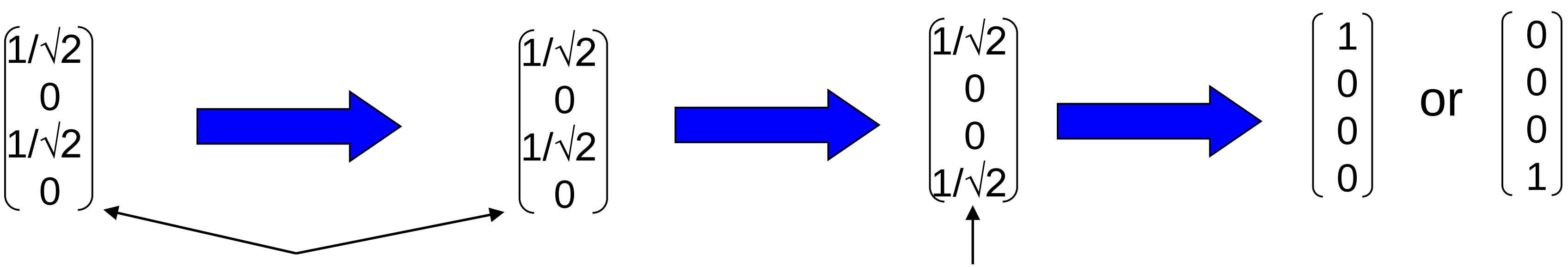
$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \rightarrow \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \rightarrow \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \rightarrow \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

$$\sigma_x \otimes I = \begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix} \qquad \text{CNOT} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}$$

Quantum Circuit Model

Example Circuit





Separable state:

can be written as

tensor product

$|\Psi\rangle = |\phi\rangle \otimes |\chi\rangle$

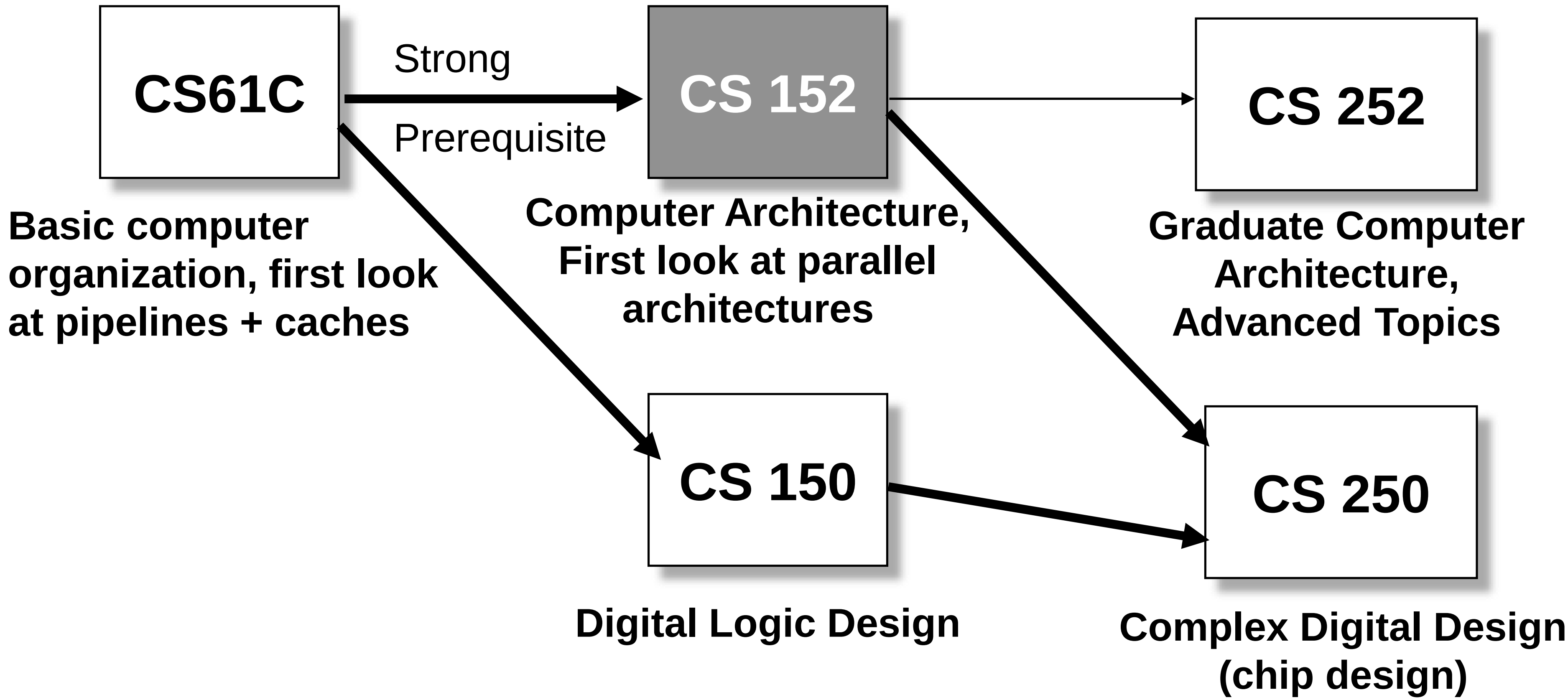
Entangled state:

cannot be written

as tensor product

$|\Psi\rangle \neq |\phi\rangle \otimes |\chi\rangle$

Related Courses



Advice: Get involved in research

E.g.,

- ASPIRE – specialized processors
- AMP Lab – algorithms, machines, people, cloud computing
- LBNL – computer architecture laboratory

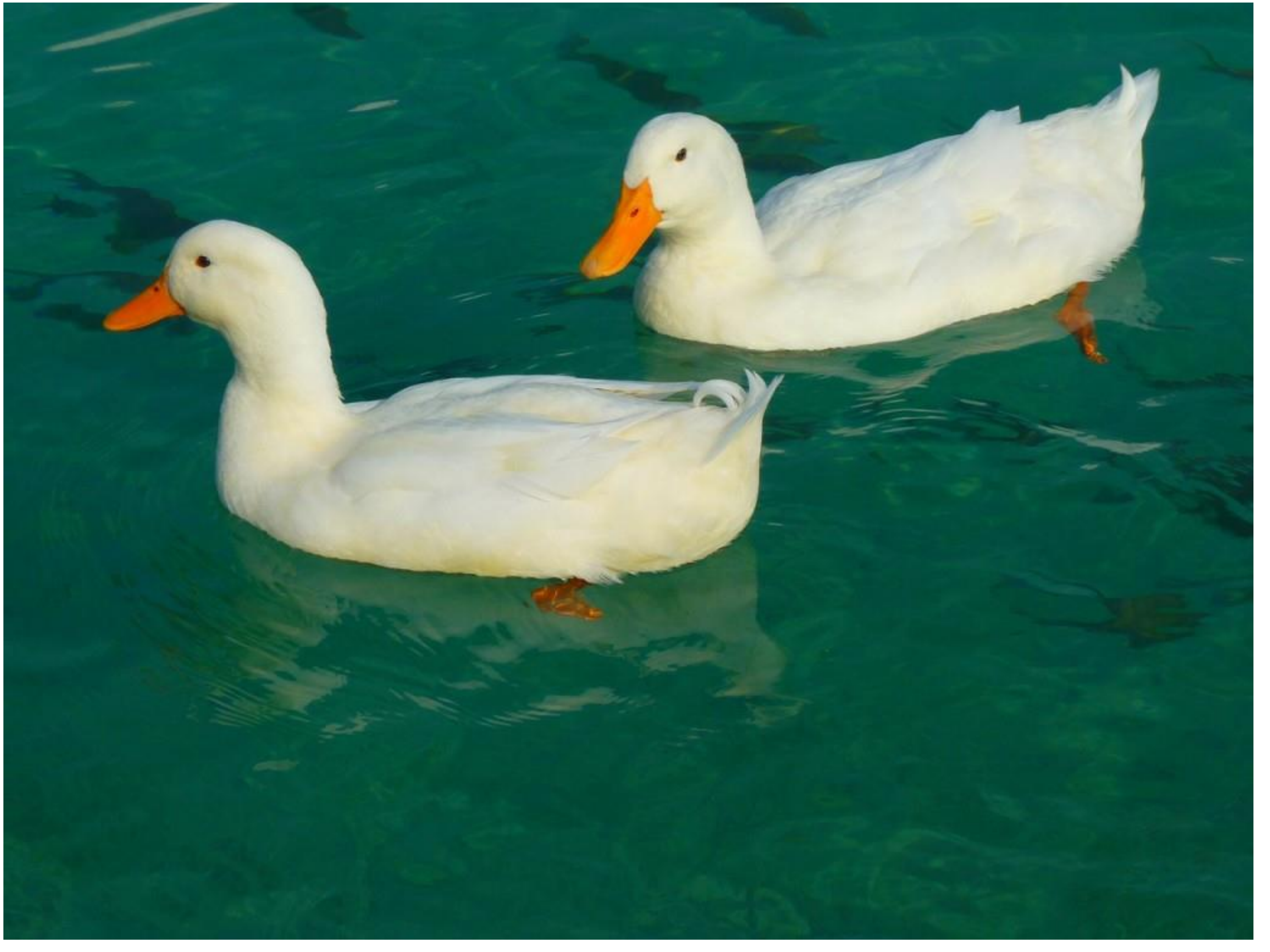
- Undergrad research experience is the most important part of application to top grad schools, and fun too
  - Also builds connections

Some Parting Thoughts

Sometimes We Think or Say

- “Everybody else is smart, I’m struggling to keep up”
- “Others just understand the material. I have to study so much harder”
- “I keep scoring below average”

The Duck Syndrome

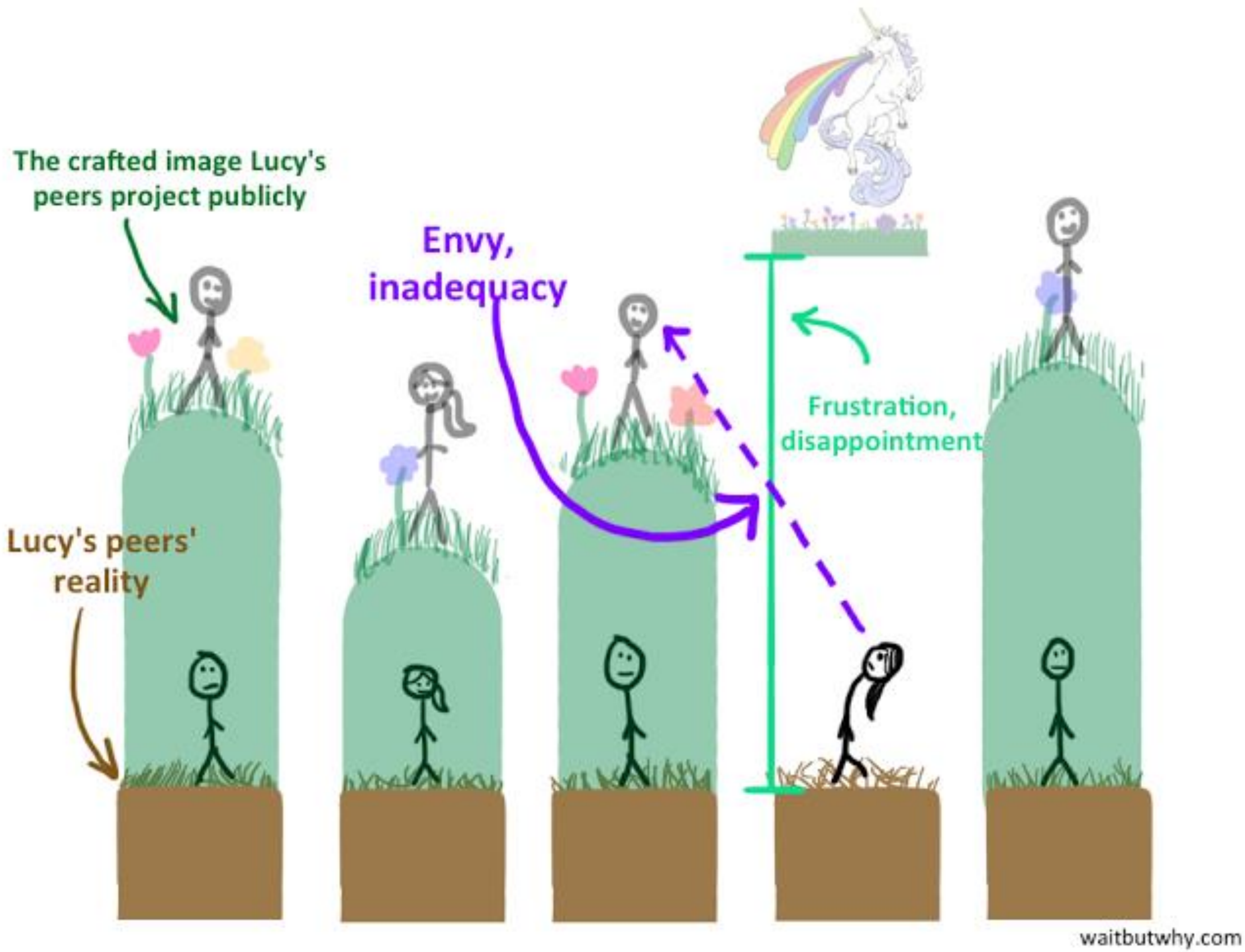


The Duck Syndrome

The Duck Syndrome: In Humans

Causes Unhappiness in Modern Times

- http://waitbutwhy.com/2013/09/why-generation-y-yuppies-are-unhappy.html



Different Kinds of Smart

- There are different kinds of smart (refers to how one learns)
- Some people prefer visual/audio inputs, others need to process the information, others need to work on the material
  - Look at the end result
- It'll help to identify what kind of smart you are
- Others are not smarter, they may be smart differently

"I Scored (or Feel) Below Average"

- Typical scenario: student transfers or gets admitted to top-tier university from a "less prestigious" school
- Student was top student in his/her home institution
- Now it's hard just to be average

Don’t Forget Where You Are

- If you are average in Berkeley, you aren’t average country-wide
- If you want to do great things you’ll be around smart people
- Would you rather be the big fish in a small pond, or a fish in a big ocean?

Importance of Grades

- Grades or early feedback did not stop people who are determined
  - They may only impose momentary setbacks

- 2014 UC Berkeley top student was told that she would never read at college level
  - http://abc7news.com/education/cals-top-student-was-once-told-she-had-a-low-iq/68443/

- Grades in a few classes won't make a difference in the grand scheme of things
  - Especially grades in a few midterms

End of CS152

- Stay for class evaluation
- Thanks for taking the class!

End of CS152

- Stay for class evaluation
- Thanks for taking the class!