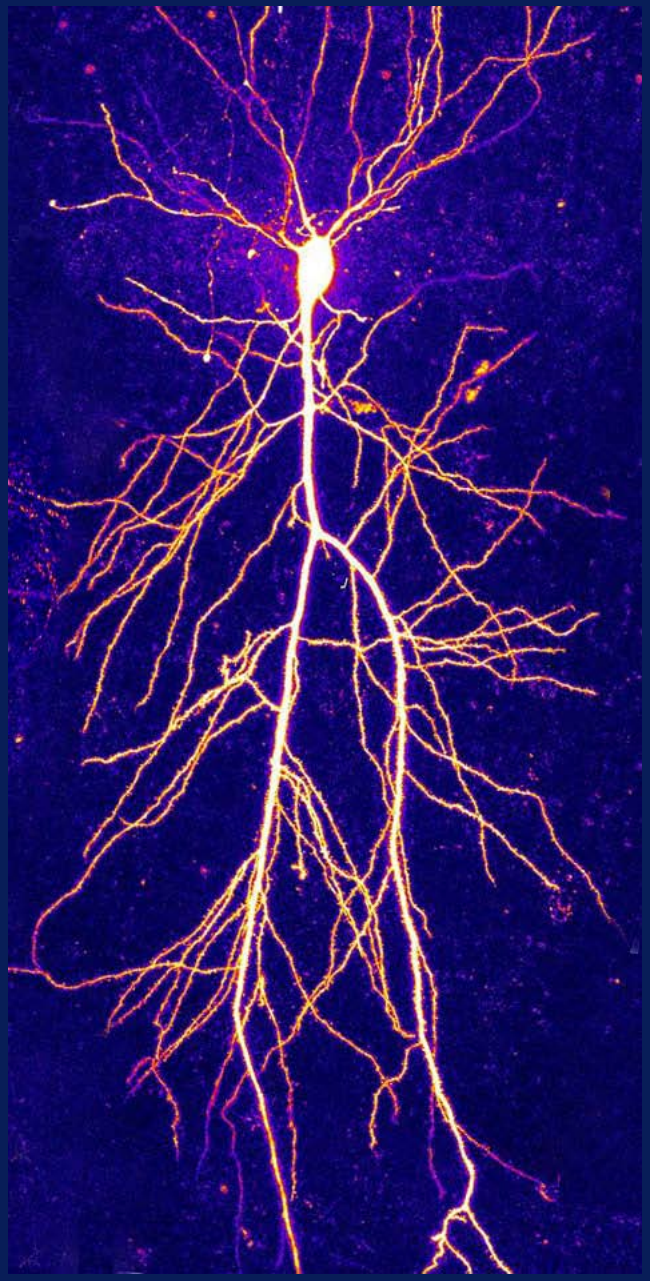


Physical Design of Biological Systems

Overview

- What is physical design?
- What is known about the design of biological systems
  - In particular, nervous systems
- Examine some problems where design of electronics and biology may overlap
- Convince you it’s not too early to think about the physical design of biological systems

What is physical design?

- Specify a physical object that performs a given function
- Predict how a given design will work
- Show the design does the desired function
- Understand the construction process
- Verify the design can be built

Build a CPU that implements the ARM instruction set

SPICE, delay calculation

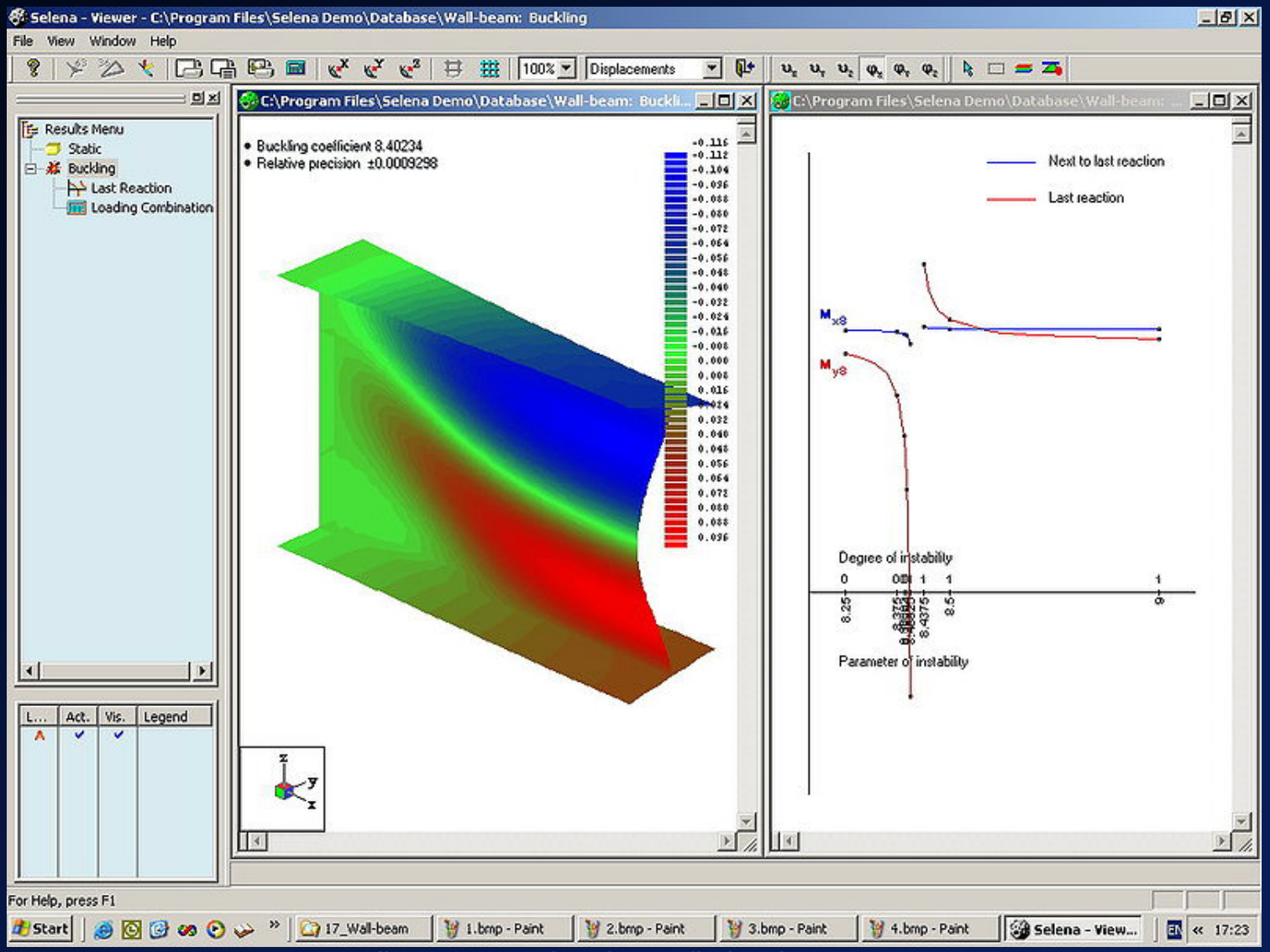
Logic simulation, formal verification, static timing

Design Rules, DRC, parasitics

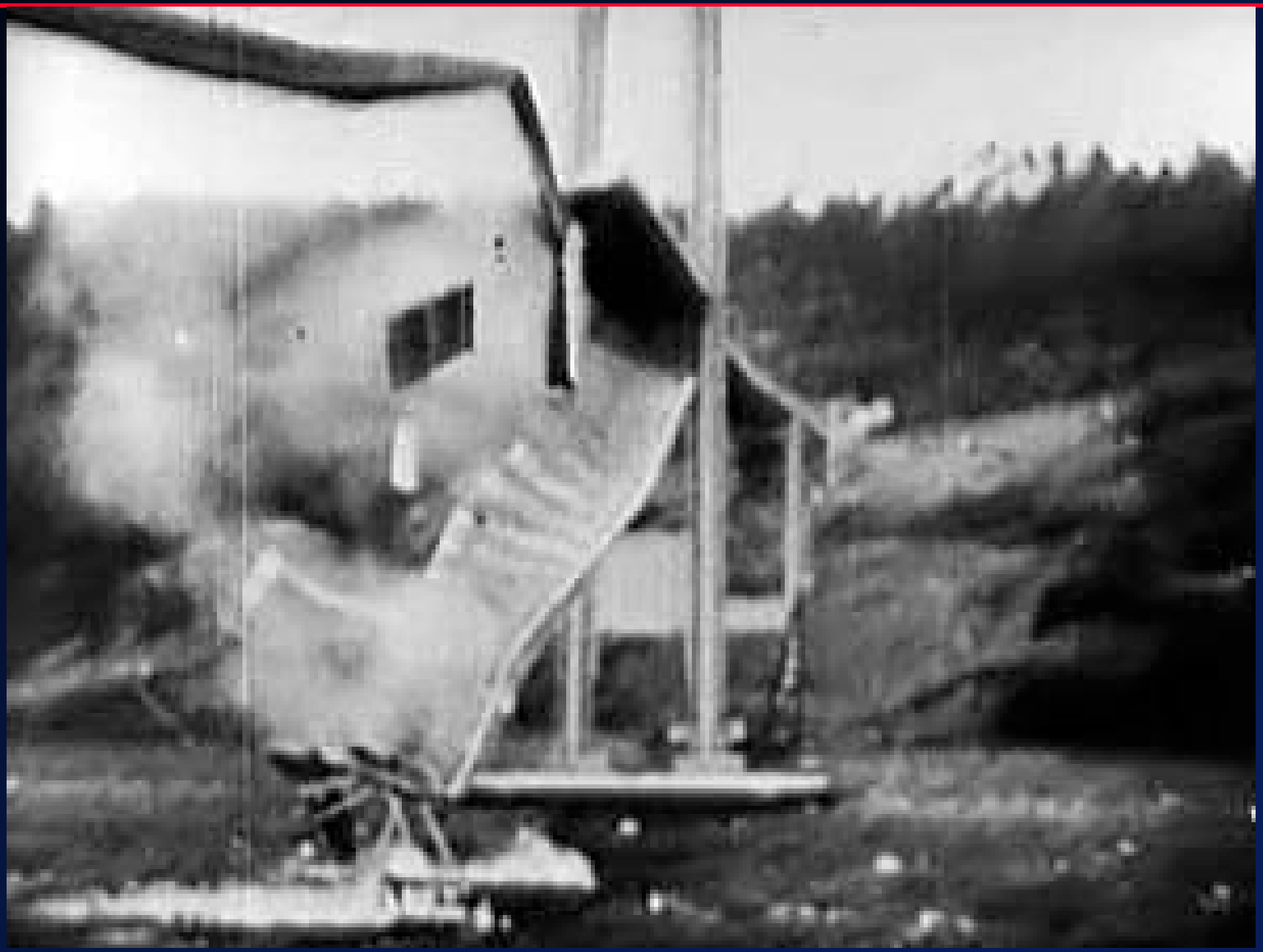
DRC, DFM tests

The same steps apply in any field

Understand how it works



Make sure it meets needs of application



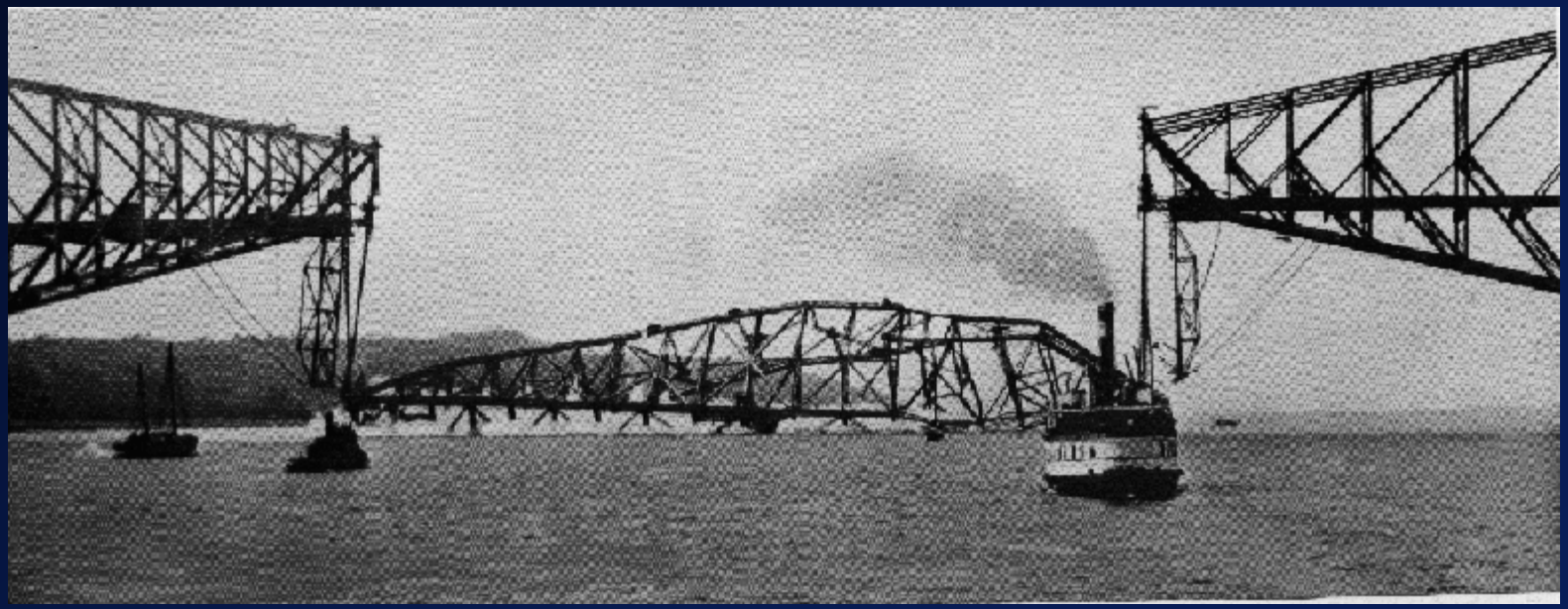
Janelia Farm, Howard Hughes Medical Institute

Know how it’s constructed

State Records NSW

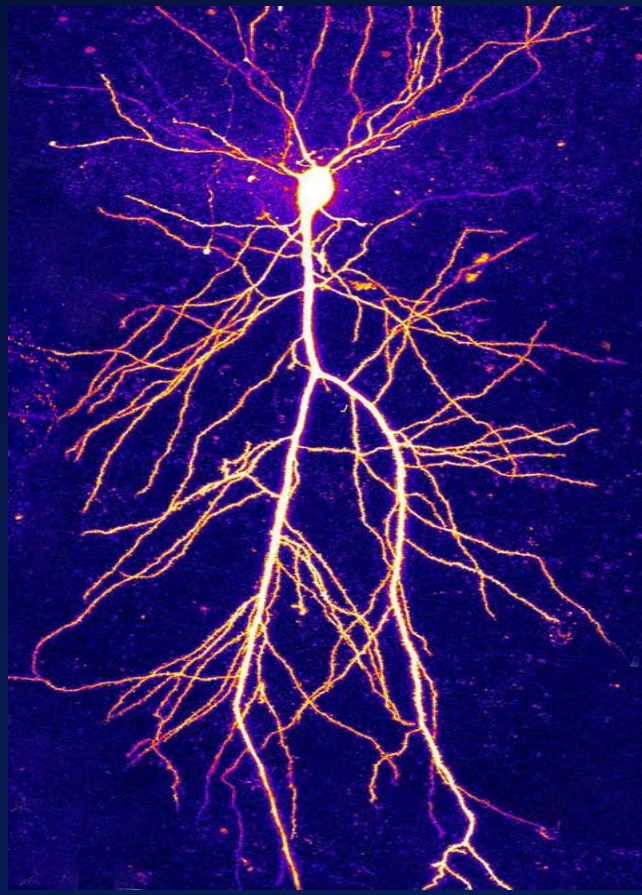
Make sure each step during construction is legal

- Quebec bridge



Do we know enough biology to get started on physical design?

- Bio systems operate in a very different way
  - Combination of chemical , mechanical, and electrical communication
- They are specified in a very different way
  - Grown and not constructed

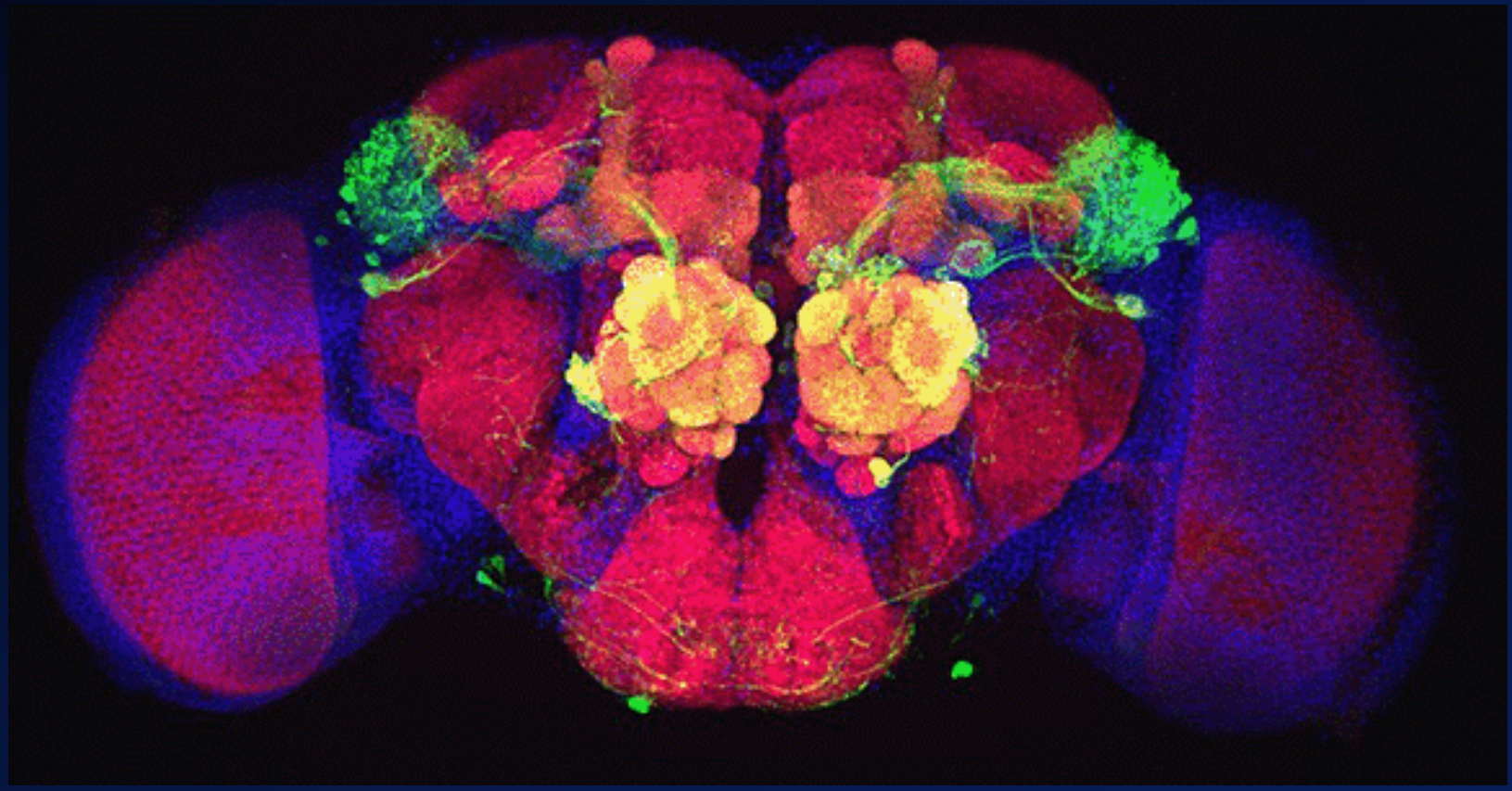


Do we understand how they work?

- Not completely, but lots of knowledge and lots of tools
  - Genetic tools are becoming very powerful
- Try to understand neural operation at many levels
  - High level of brain units
  - Medium level – construction of neural organs
  - Low level – operation of neurons and synapses
  - Very low level – molecules and chemicals

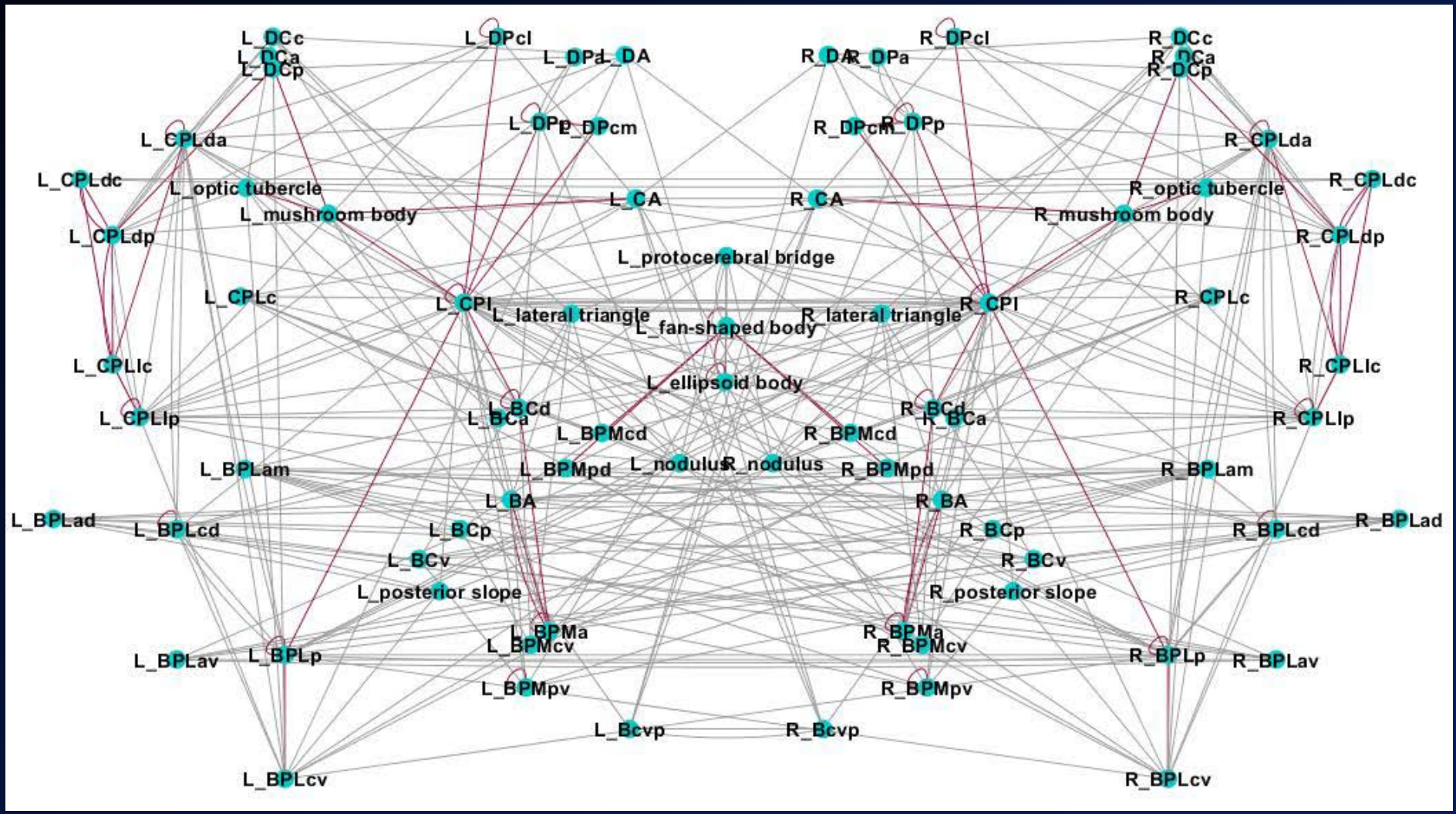
Example – the fruit fly

- Brain expressing colored markers



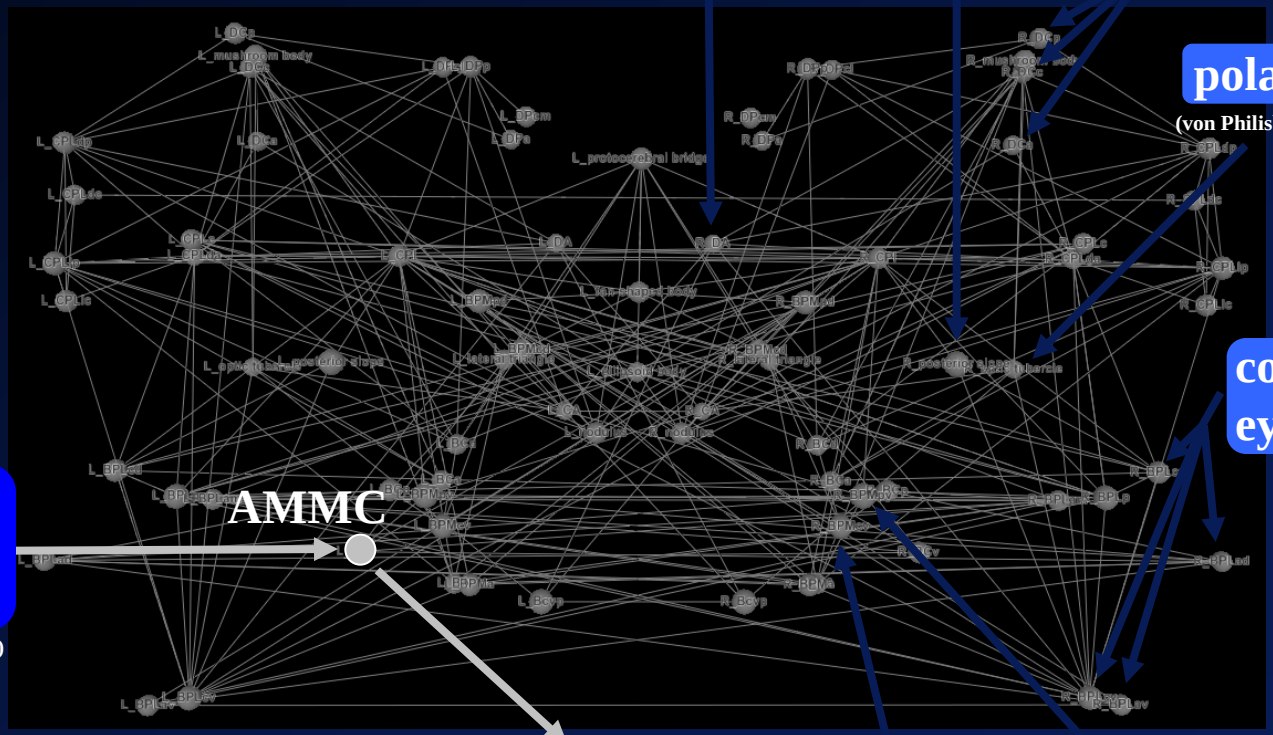
Janelia Farm, Howard Hughes Medical Institute Richard Axel, Columbia

How a fly works – large scale



Wayne Pereanu

We have a basic idea of what many of these parts do



ocelli
(Marcey & Stark, 1984)

gustatory
(Singh, 1997)

ascending
(Strausfeld & Hansen)

polarization
(von Philisborn & Labhart, 1990)

compound eye
(Strausfeld et al., 2007)

antenna mechanosensory stimuli
(Kamikouchi et al., 2006)

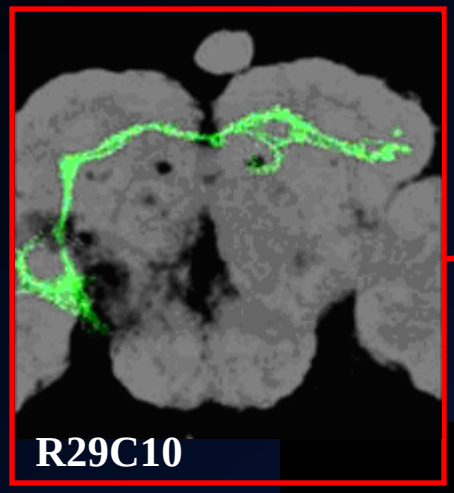
AMMC

nerve cord

olfactory
(Power, 1946)

hygrosensation
(Lu et al., 2007)

Map indicates ocelli may modulate response



ocelli

CPI

PS

AMM

C

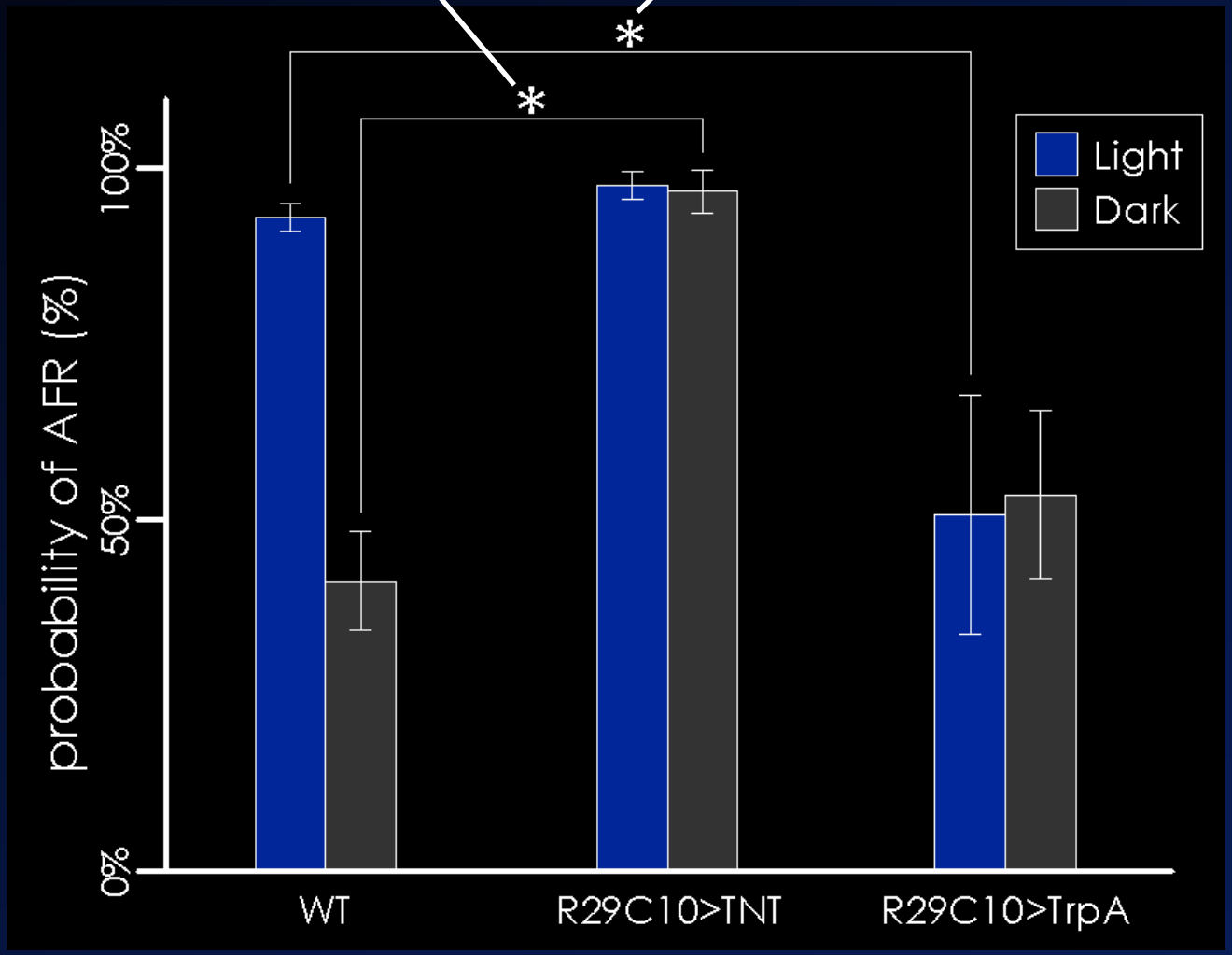
antenna mechanosensory stimuli

nerve cord

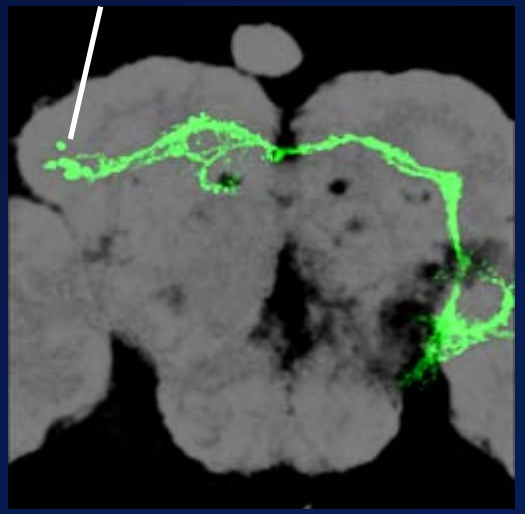
CPI>AMMC interneurons mediate this response

Inhibition results in light-like levels in the dark. Therefore, these neurons are necessary for the observed modulation.

Overactivation results in dark-like responses in the light. Therefore, these neurons are sufficient for the observed modulation

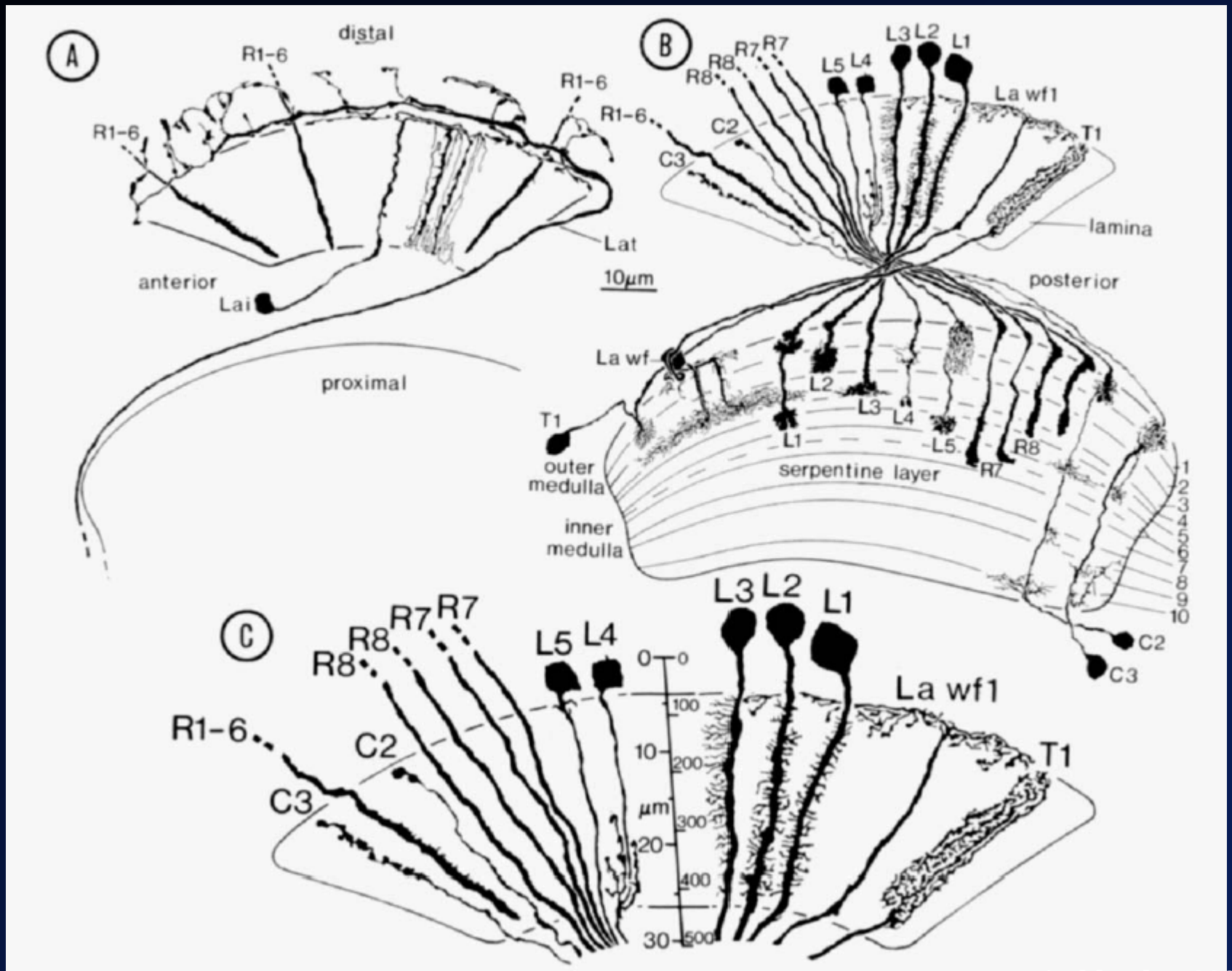


9 neurons from the BLVp2h1 lineage



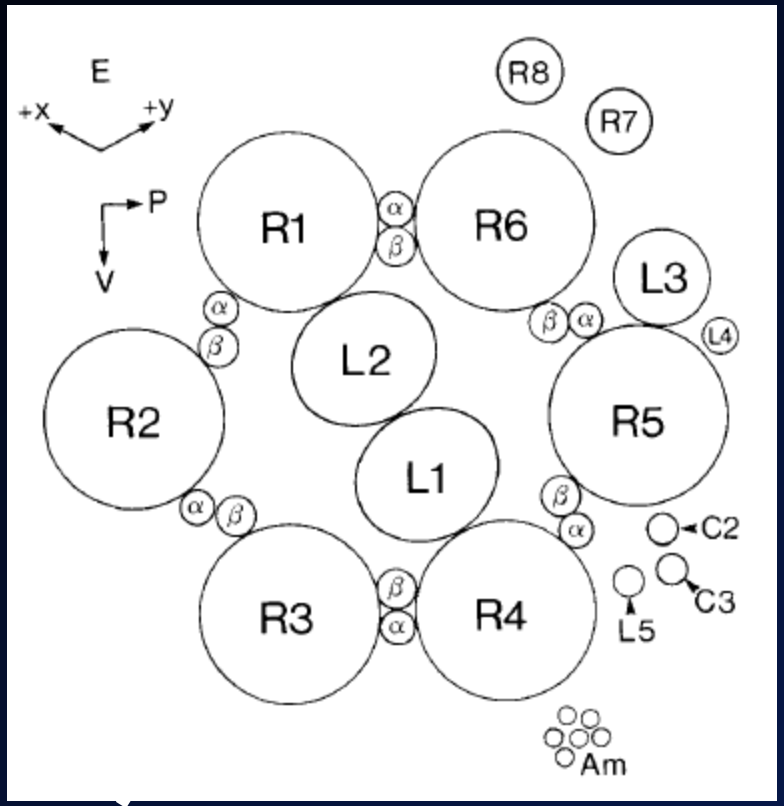
How it works – medium level

- For at least some of the computational units, we know how they are built down to the connections
  - Technology for getting the rest is under development

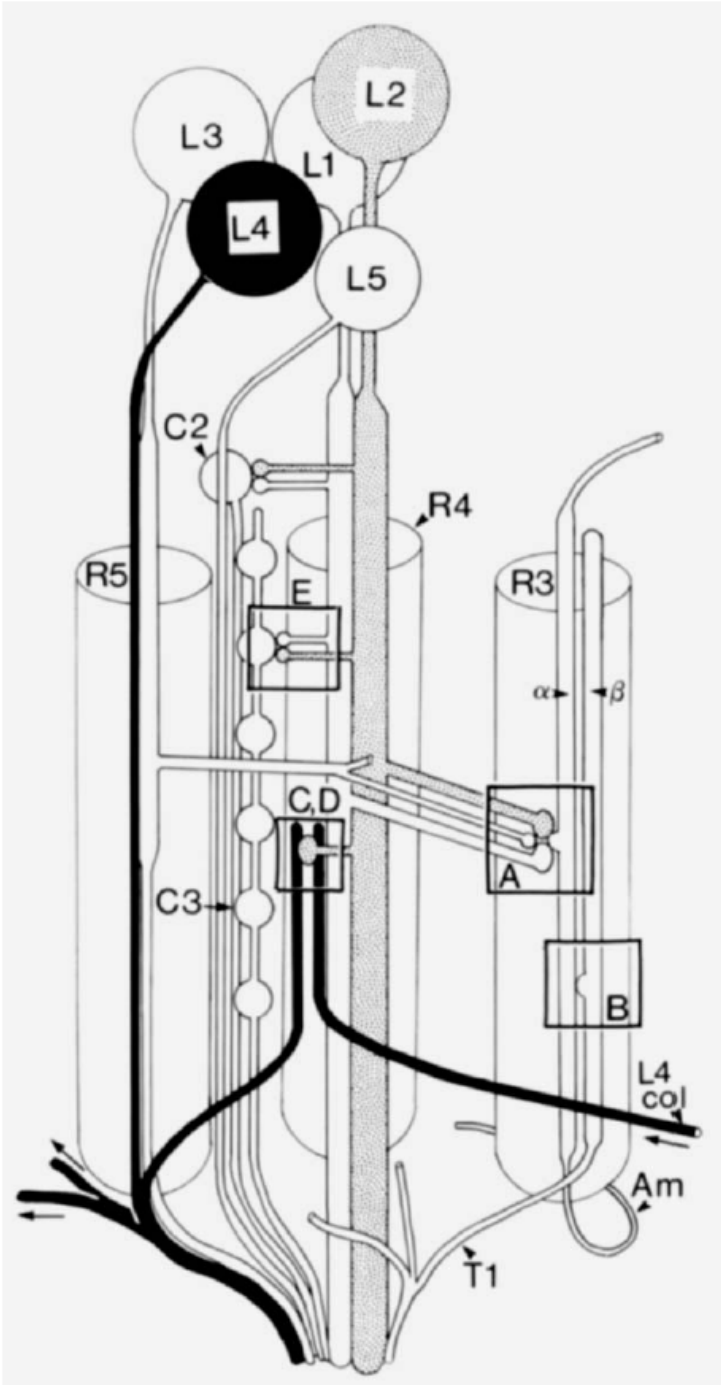


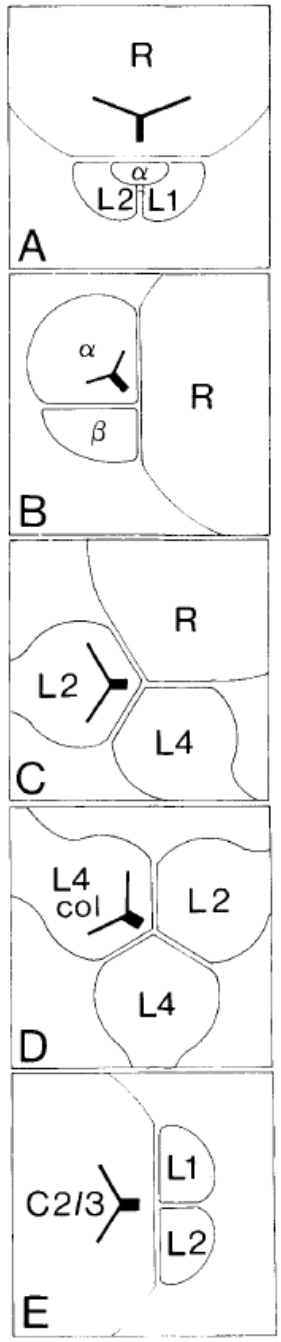
Fischbach, 1989

Lamina circuit



- Overhead view
- Side view





Meinertzhagen and O'Neil, 1991

We have the netlist

TABLE 1. *Drosophila melanogaster:* A Synaptic Inventory for Columnar Cells of the Wild-Type Lamina Cartridge[1]

Cell	Presynaptic upon elements	synapse	Fig.	Postsynaptic at elements	synapse	Fig.
R1–R6	L1, L2, α, α;			L2	dyad	23
	L1, L2, L3, α;		18, 19	L4	dyad, triad (m)	26, 27
	L1, L2, L3, gl;	tetrad	21, 22	α	dyad, triad (m)	31
	L1, L2, gl, gl.					
R7	—			—		
R8	—			—		
L1	—			R1–R6	tetrad	18, 19
				α	triad (1)	
				C2	dyad	37
				C3	dyad	38
L2	R, L4 (all three)		23	R1–6	tetrad	18, 19
	L4, L4y	dyad		L4 (all three)	dyad	24A, 25
	L4, L4x	dyad		L4 (all three)	triad (1)	26
	L4y, L4x	dyad	24B	α	dyad	
				α	triad (1)	
				C2	dyad	36, 37
				C3	dyad	38, 39
L3	—			R1–R6	tetrad	
				α	dyad	34
				α	triad (1)	
L4	L2, R	dyad		L2	dyad	23
	R(m), L2, other	triad	26, 27			
L4y	L2, R	dyad	24A, 25	L2	dyad	
	L2, L4x	dyad				
	R, R	dyad	28			
	R(m), L2, L4x	triad				
L4x	L2, L4	dyad		L2	dyad	
	R(m), L2, L4y	triad				
L5	—			α	dyad	35
				C2?	triad (1)?	
				?	triad	40
Am	R1–R6, β	dyad	31	R1–R6	tetrad	22
	R1–R6, L (3 or 2)	dyad		C2	dyad	
	β, L (3, 2 or 5)	dyad	34	C3	dyad	39
	R(m), β, L (1, 2 or 3)	triad				
	gnarl		32			
	glia					
T1	—			α	dyad	33, 34
					triad (1)	
C2	alpha, L (1, 2 or 3)	usu.		—		
	L2, L (1 or 3) or R	dyad	36–37			
C3	alpha, L (2 or 3)	dyad	39	—		
	L1, L2	dyad	38			
Ep-glia	—			R1–R6	tetrad	
				α	monad	32

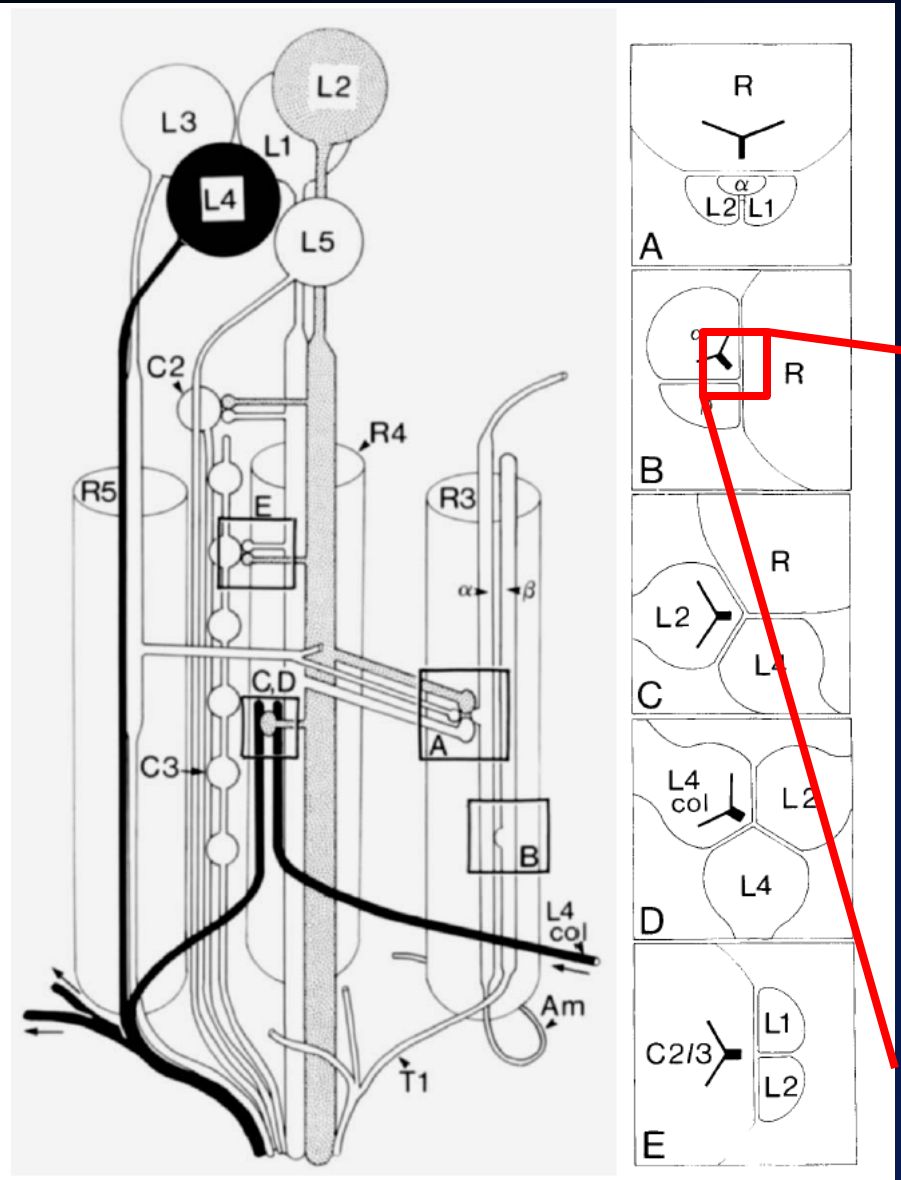
(m), (l): median and lateral elements, respectively, of a triad.
[1]cf Shaw ('81), his Table 1.

Meinertzhagen and O'Neil, 1991

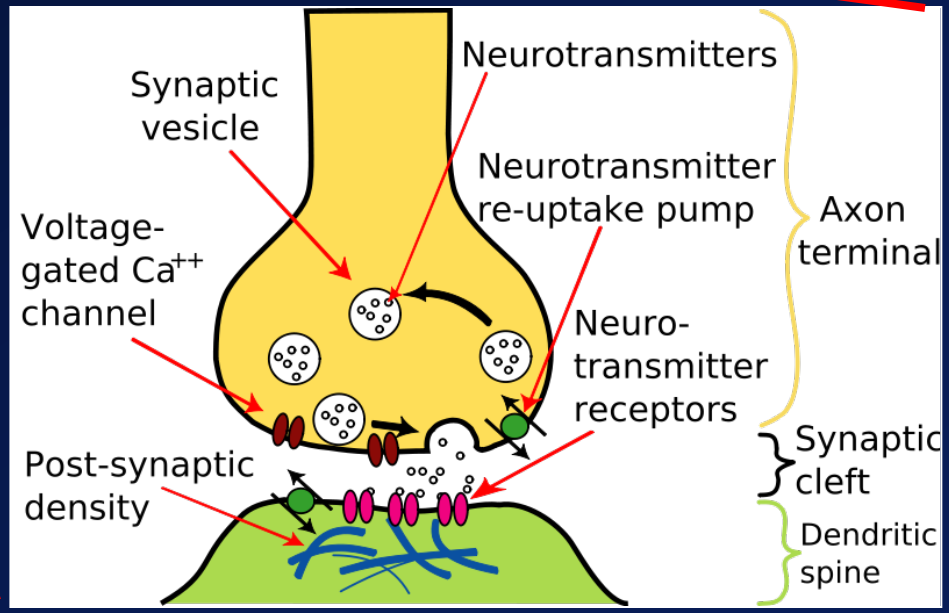
But we don’t quite have everything

- Signs of connections unknown
  - All receptors look the same on electron microscope pictures
  - Modern genetics (and the fact that all flies are identical here) will soon fix this…
- Neuromodulators
  - Support cells (glia) have incoming synapses, but no outgoing synapses

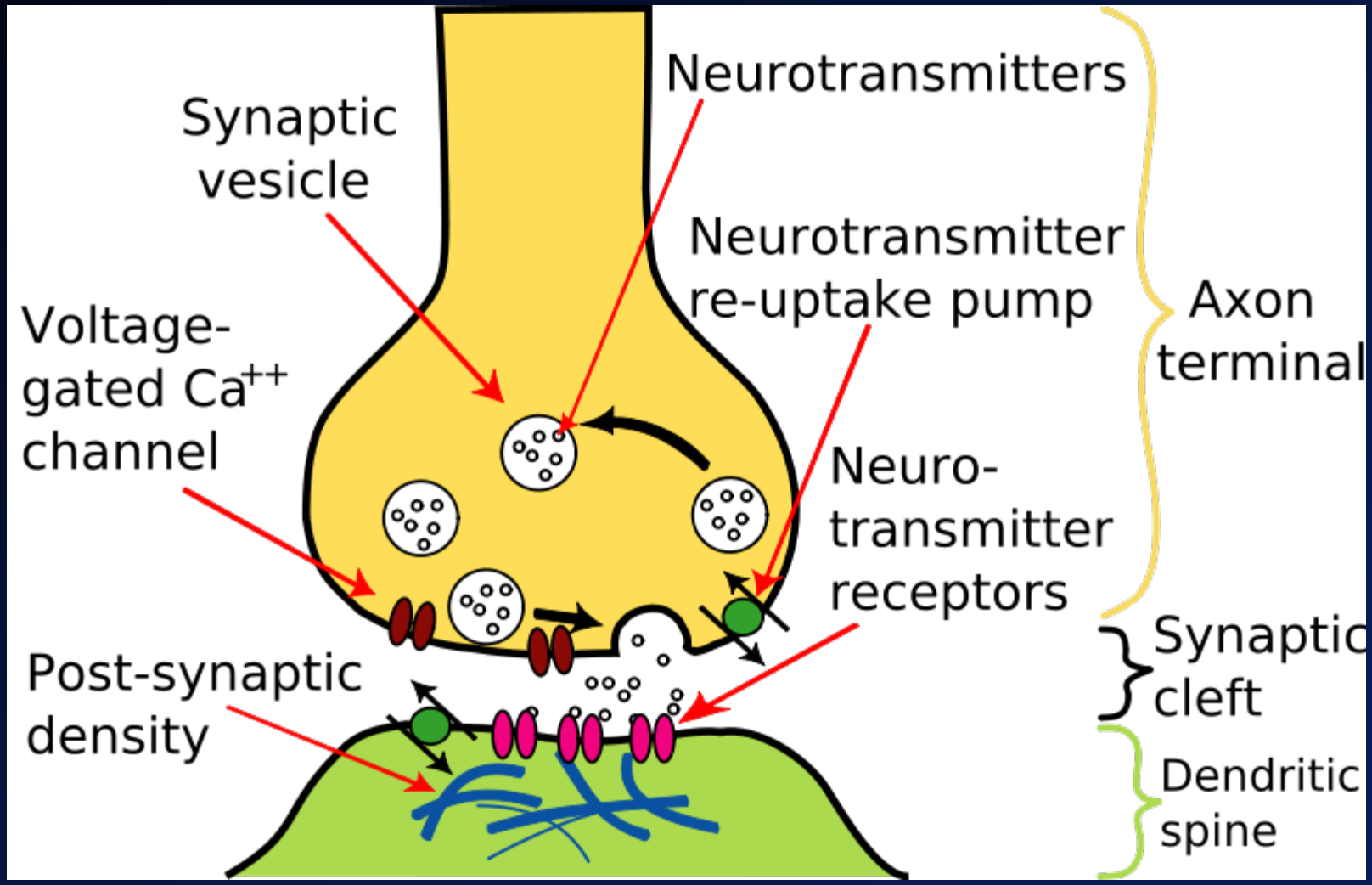
How it works – small scale



- Each synapse is itself a complex piece of machinery

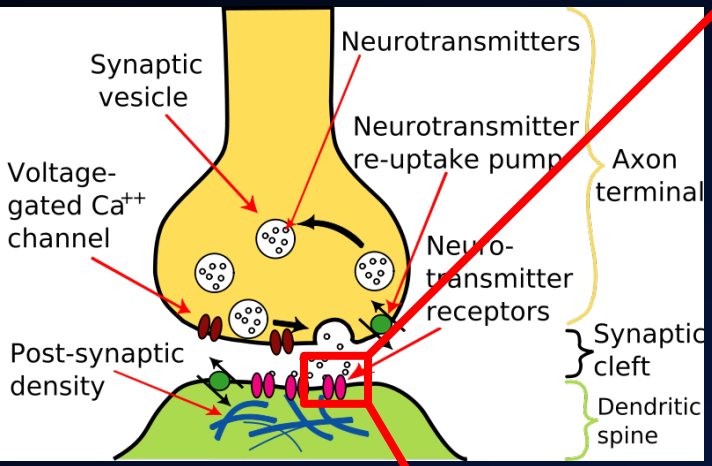


How a synapse works



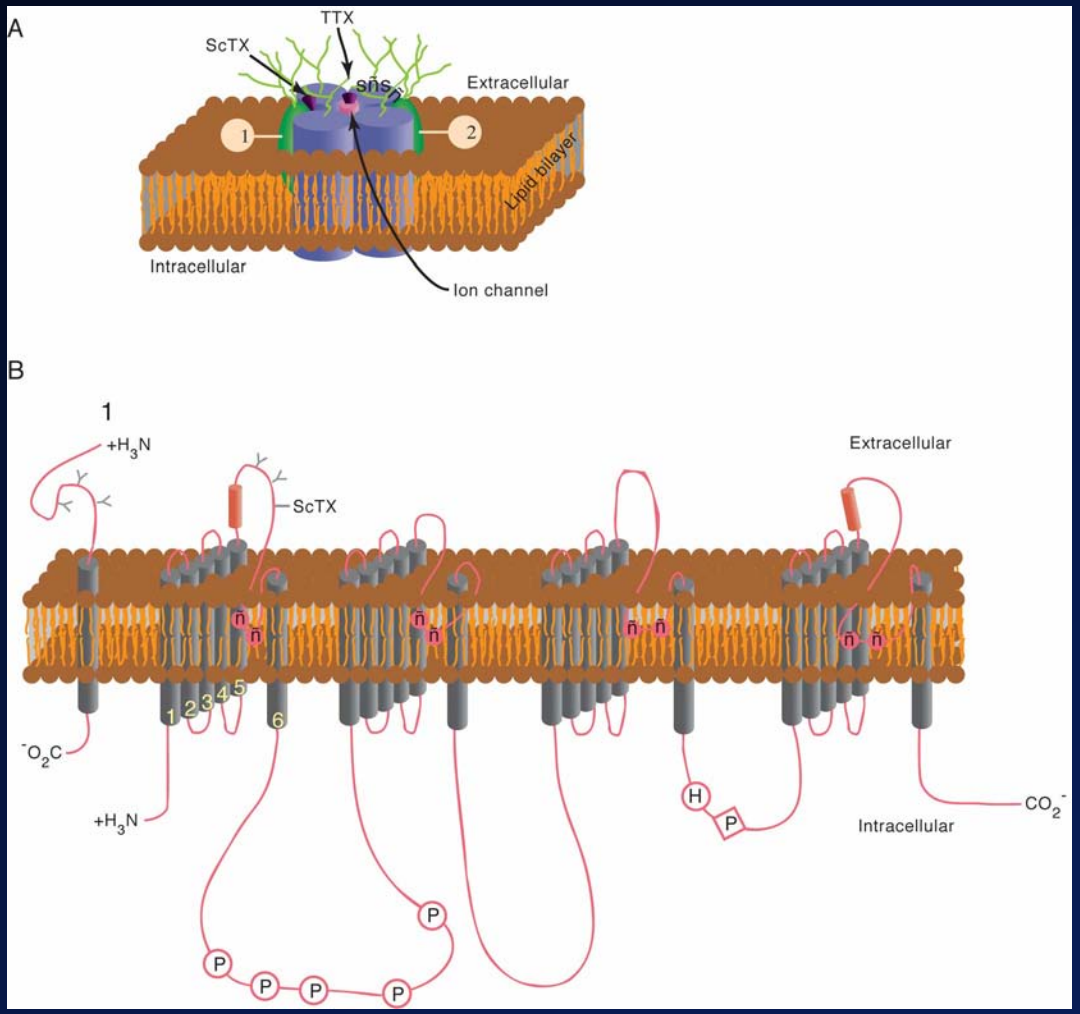
Wikipedia Synapse_Illustration2_tweaked.svg, 1 Nov 2009

How it works – tiny scale



- Operates on the scale of genes and molecules.
- Not needed for most analysis
  - Voltage vs current is enough
- But will need to be understood for learning and memory

Ion channels are complex structures



Jeff Magee

Janelia Farm, Howard Hughes Medical Institute

How creatures are built

- This is not as well understood as operation
- But is again yielding to modern genetics
- Typical model is a worm – one of the smallest creatures that has specific organs
- Link to movie

From the Goldstein lab at UNC

How it's built – large scale

"C. elegans develops from a single cell, the fertilized egg, to a 558-celled worm in about 14 hours. The worm that crawls out of its eggshell has a functioning feeding apparatus, gut, nervous system and muscles."

- Link to movie

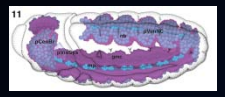
From the Goldstein lab at UNC

How it's built – medium scale

- Entire creature is built by successive division and specialization
  - Specialization driven by chemical environment and instructions from the parent
  - Interestingly, cell death is quite often used as well
- We know cell origins for each cell in worm
- Working on the same in flies

Proliferative activity of a typical neuroblast in *Drosophila*

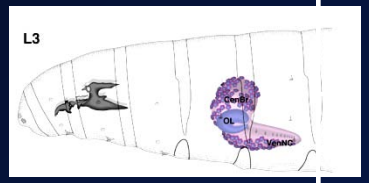
Embryo



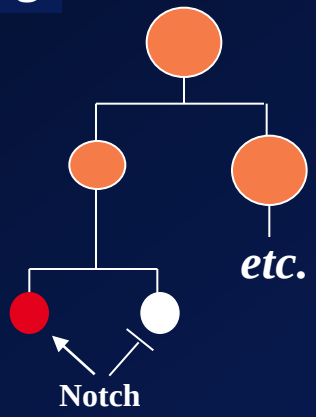
1st larval stage

2nd larval stage

3rd larval stage



hatching



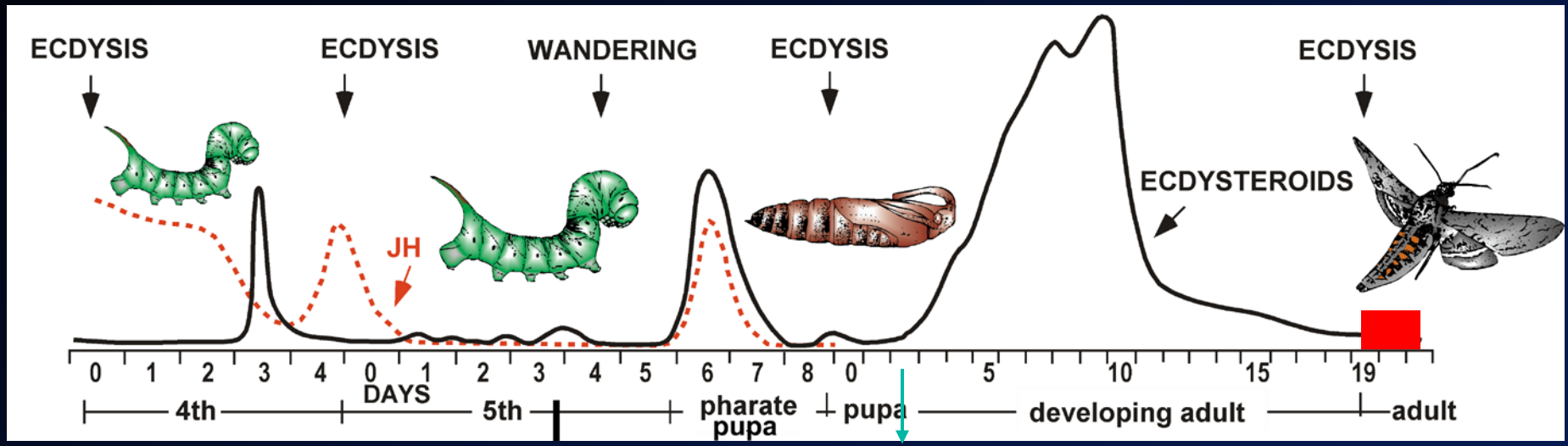
"Hemilineage" pattern for secondary neurons

primary

Secondary neurons

From Jim Truman's lab

Hormonal Control of Insect Development and Behavior



Pupal commitment

Adult commitment

Hormones

Ecdysone → molting

switching of programs at metamorphosis

Juvenile hormone (JH) → prevent switching

prevent premature differentiation

role in reproduction

From Lynn Rutherford's lab

How it’s built – small scale

- Passing state to daughter cells
- Chemical gradients
- Link to movie

OK, we sorta kinda somewhat understand biology. So what?

- Compare EDA today with classical physics in the late 1800s
- At this time, classical mechanics was almost complete
- Old stuff was and remains super useful
  - Newton's laws of motion
  - Euler angles
  - Gauss's method
- Still used most, and the first taught
- But new stuff adds a new level of concerns

Similarly, will keep using old stuff of EDA

- What we have done so far will be core technology from now on
  - Mapping and covering
  - FM and multi-level partitioning
  - Placement
  - Global and detailed routing
  - Logic and circuit simulation
- New problems for new areas of concern
  - Biology

How does the emerging bio field relate to EDA?

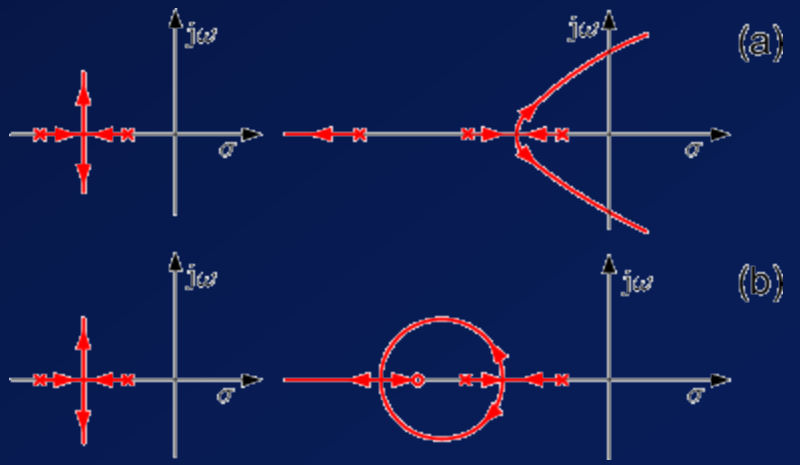
- Bio, especially nervous system bio, offers lots of interesting problems for:
  - Folks who enjoy math for problem solving
  - Folks who enjoy physical design
  - Folks who enjoy working on large software systems
- If you don't like working on at least one of these, why are you here?

Math part – interesting problems

- Linear vs non-linear systems
- ‘Small’ number of cycles
- Mixed A/D systems
- Statistical operation

Linear systems

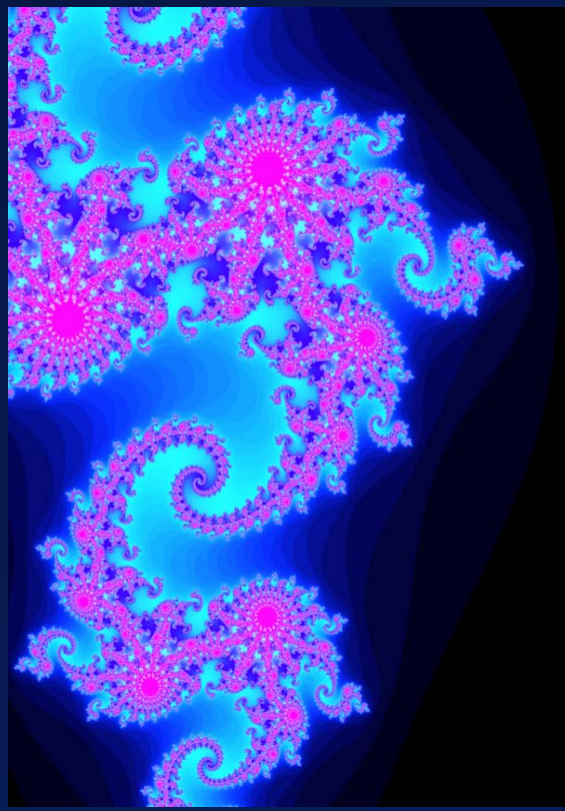
- Engineers like linear systems, and with good reason
- Many tools are available – examples
  - Given an N variable linear system
  - Can find min, max time constants
  - Sensitivities
  - Eigenvalues and vectors
  - Major and minor modes
  - Big box of tools



Non-linear systems

- In the general case, cannot say anything even for 3-4 variables
- But biology does not exploit the whole space of non-linear systems
  - Mostly monotonic, or just somewhat non-linear
  - Must be low sensitivity
- Need to discover the right approximations

One non-linear op!



Small number of cycles

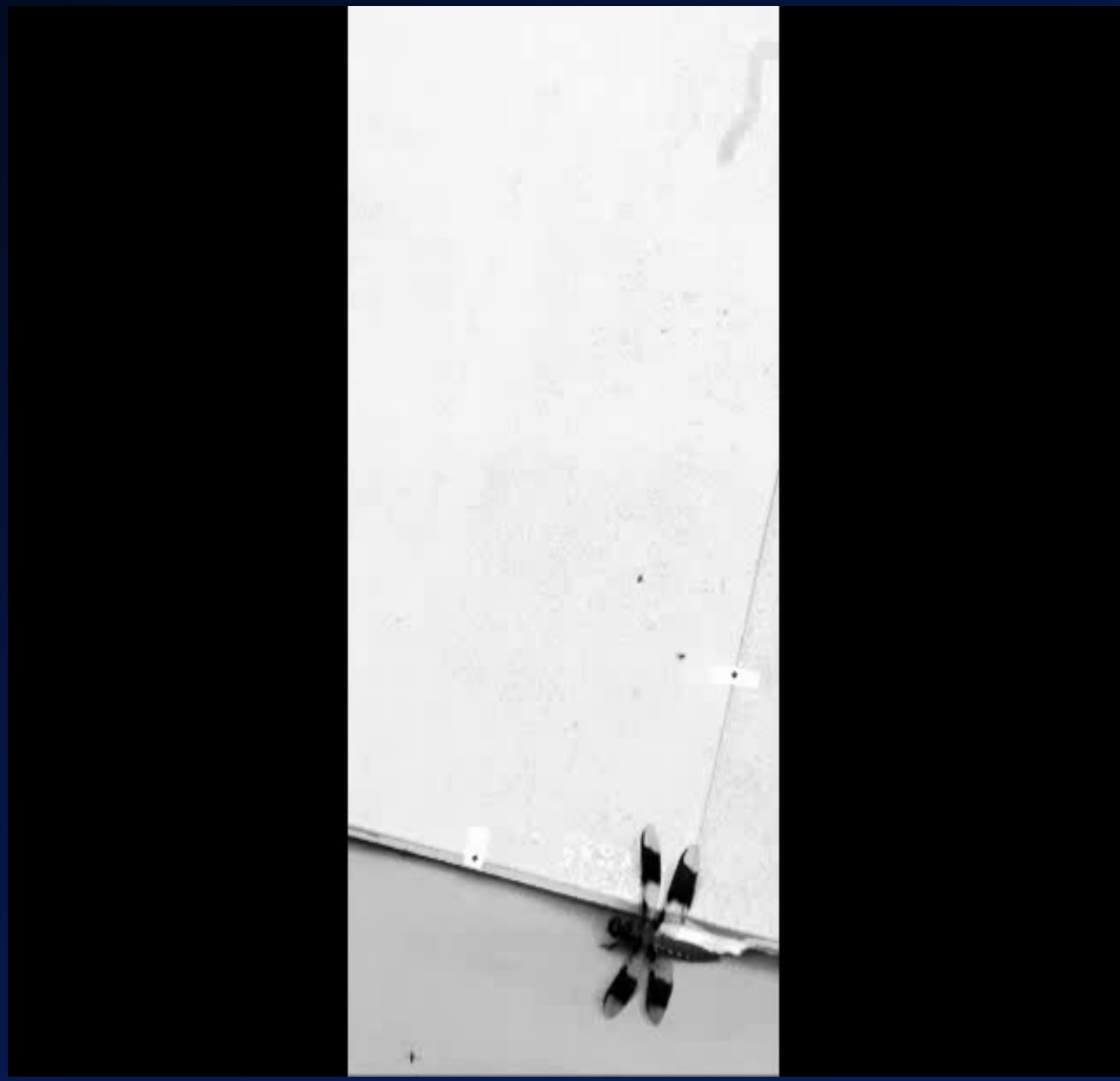
- EDA has built good tools for analog circuitry in two limiting cases
- Waveform itself is the objective (amplifiers, filtering, DSP)
- Modulation or perturbation on a large number of cycles is the objective.
- But nature operates in an intermediate regime…

Example

Dragonfly catching a fruit fly

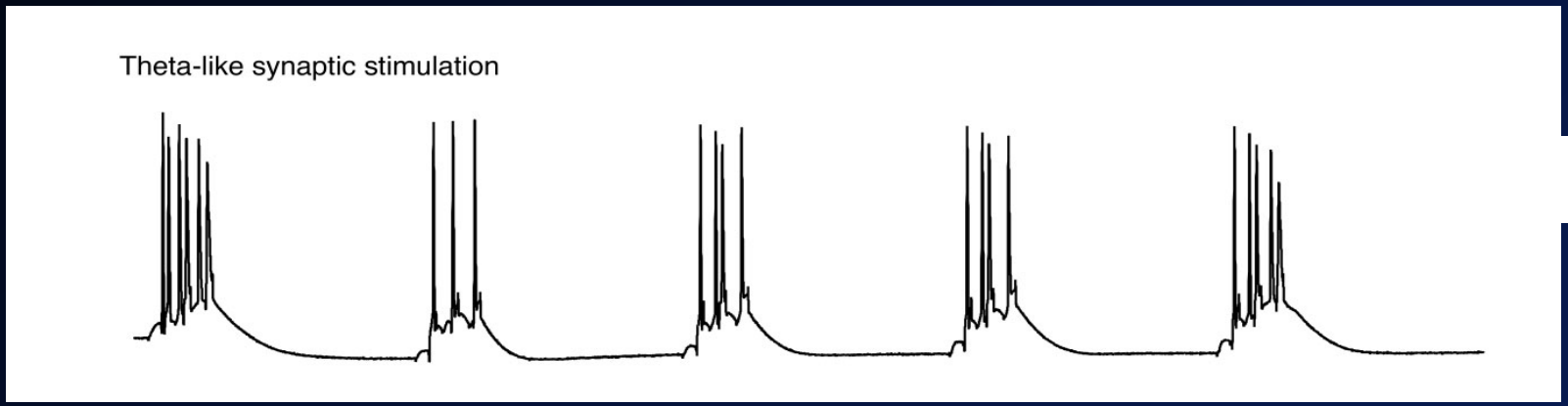
Or football player catching a pass

Or you aiming for coffee during the break….



From Anthony Leonardo's lab

Mixed A/D systems

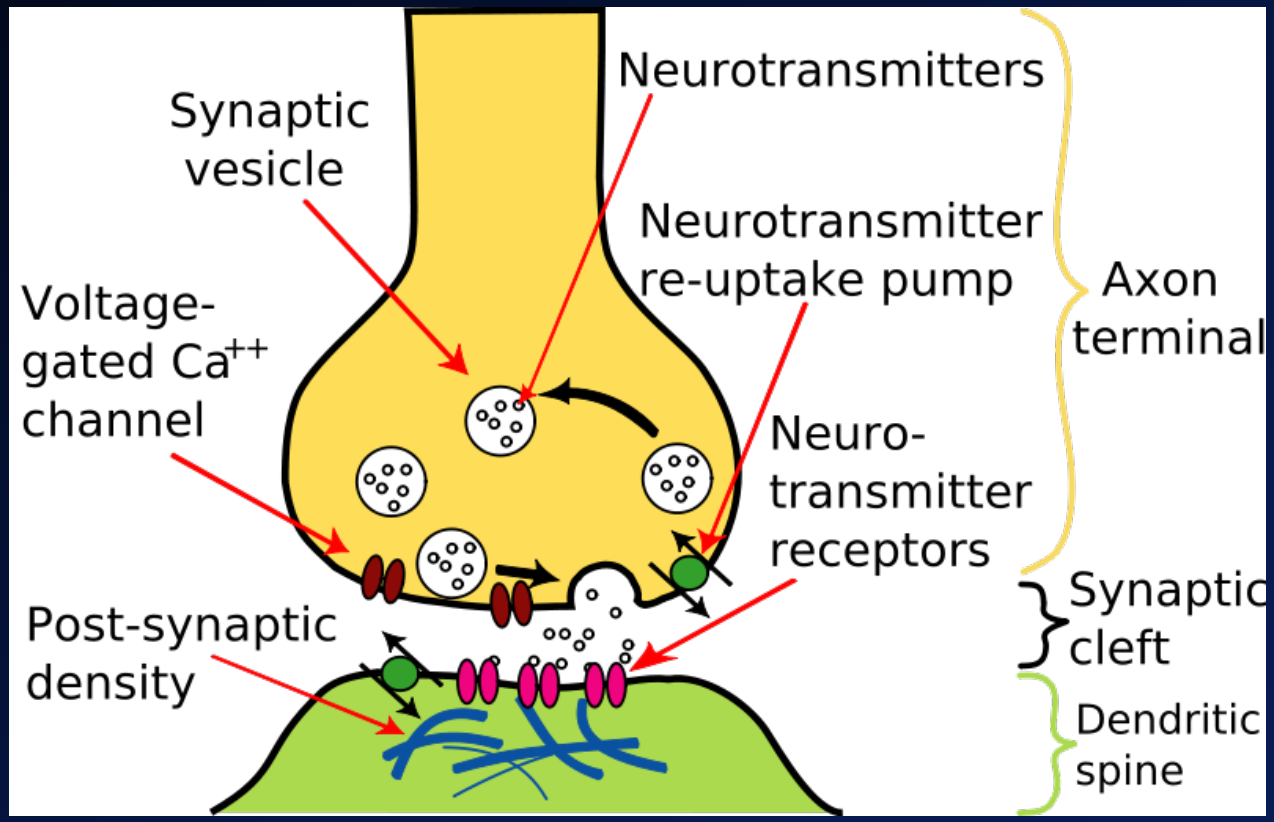


- How is information encoded?
  - Average amplitude?
  - Pulse timing?
    - Timing of first pulse?
    - Timing between pulses?
  - Pulse frequency?
  - Multi-neuron encoding?
- Answer is yes

 Waveform from Jeff McGee

Statistical analysis

Compared to neurons, statistical timing is just barely statistical



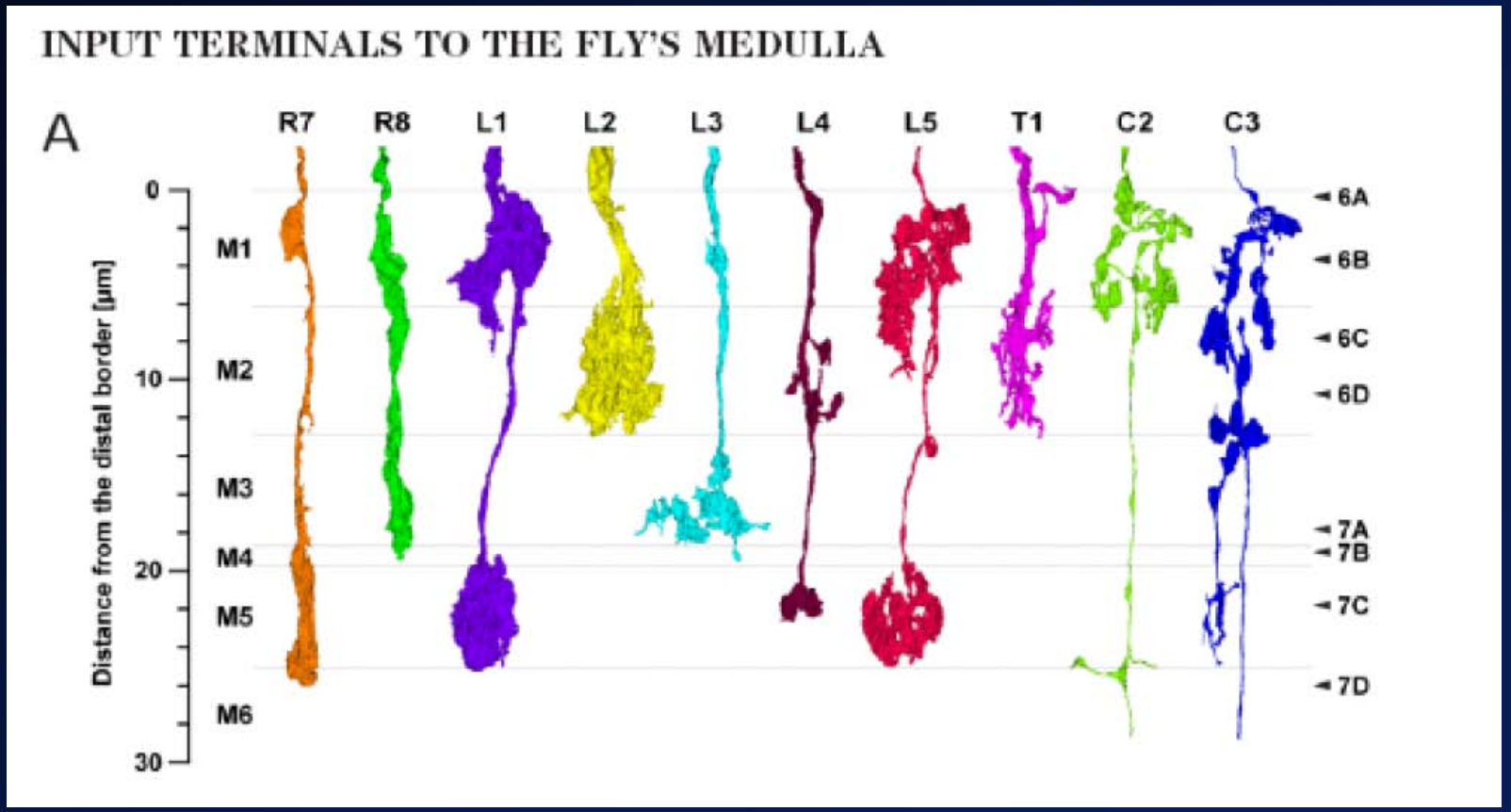
- Every stage of operation is statistical
- But behavior is robust

Physical design

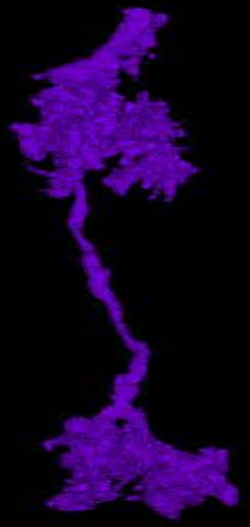
- Differs from IC design in several ways
  - Wires and gates of commensurate size
  - Fanouts are much larger, logic depths are less
  - Volume filling, not 2D filling
  - Design creation is very different

Wires and gates of near-equal size

- Neurons have connections to what would be wires in gates



Meinertzhagen and Takemura



Volume filling design

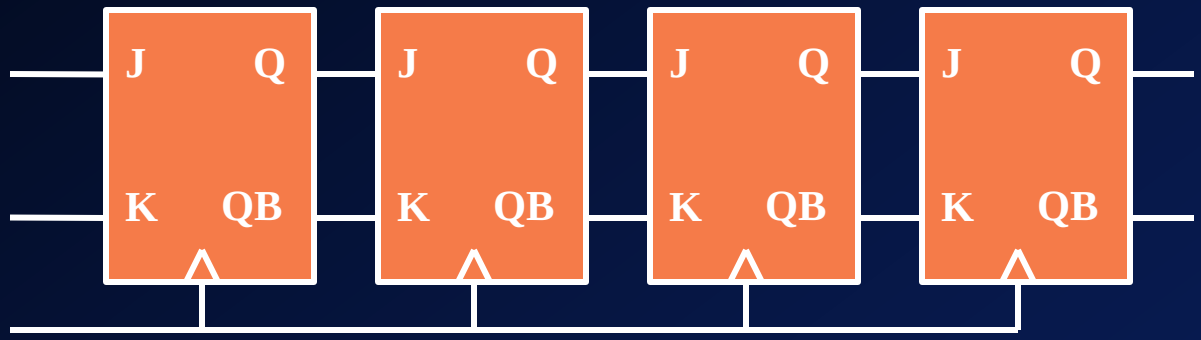
Small open spaces for food, oxygen, etc.

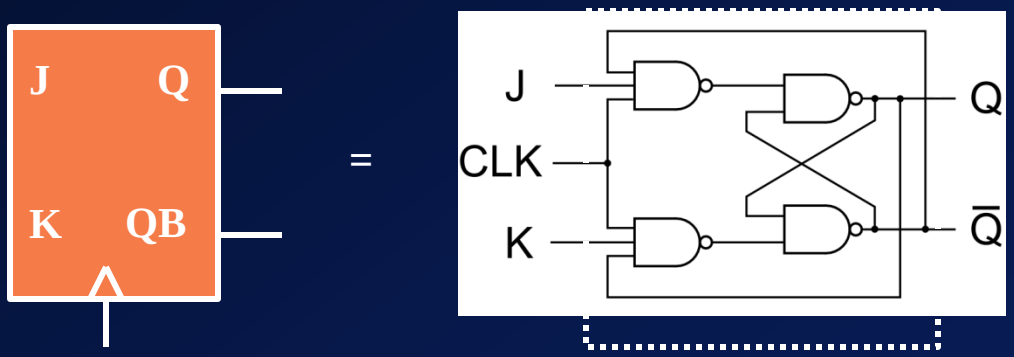
Mostly limited by size and need to make connections

Takemura & Meinertzhagen, 2010

Design creation

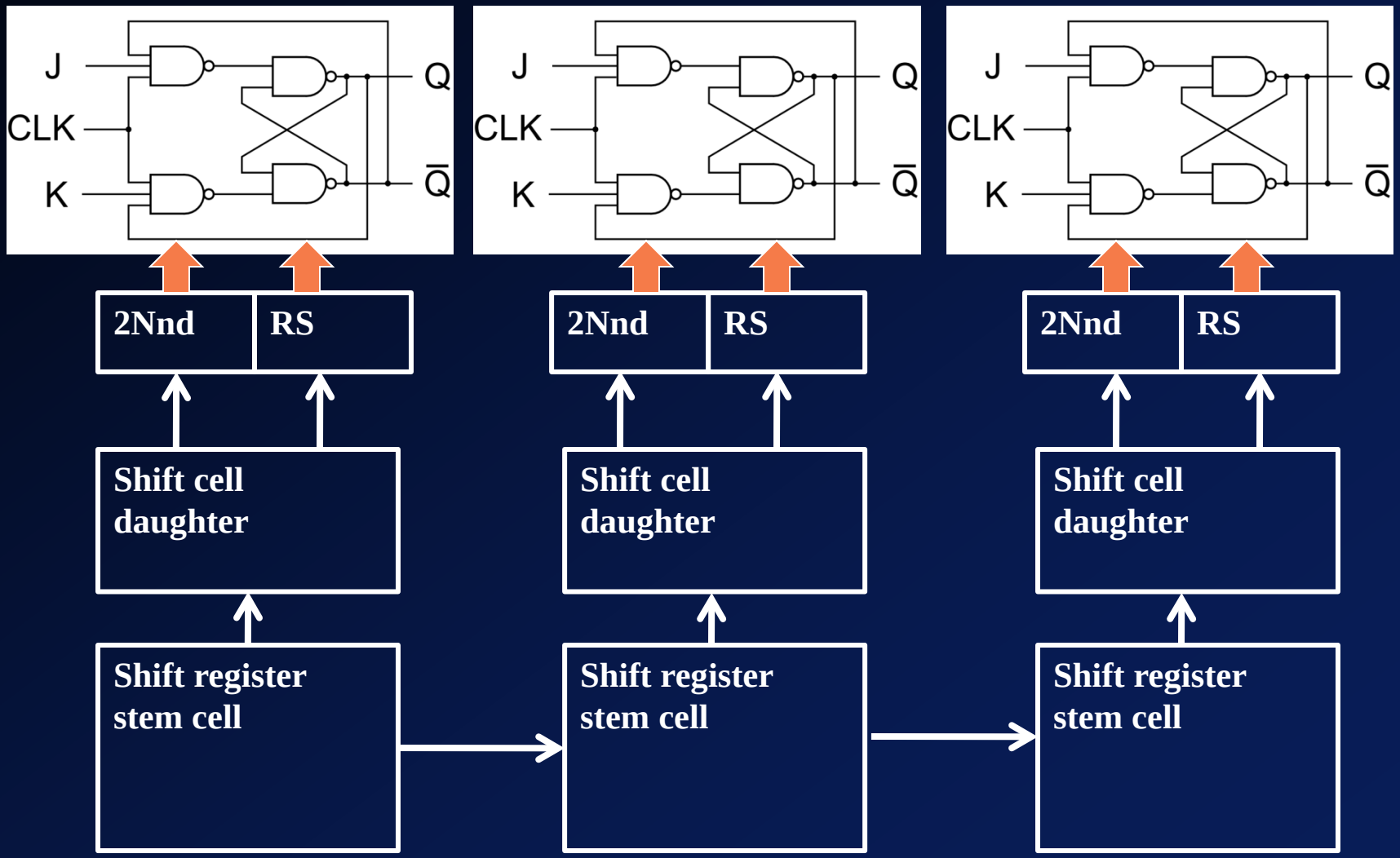
- How an engineer specifies a shift register





Design creation

- How biology might specify a shift register

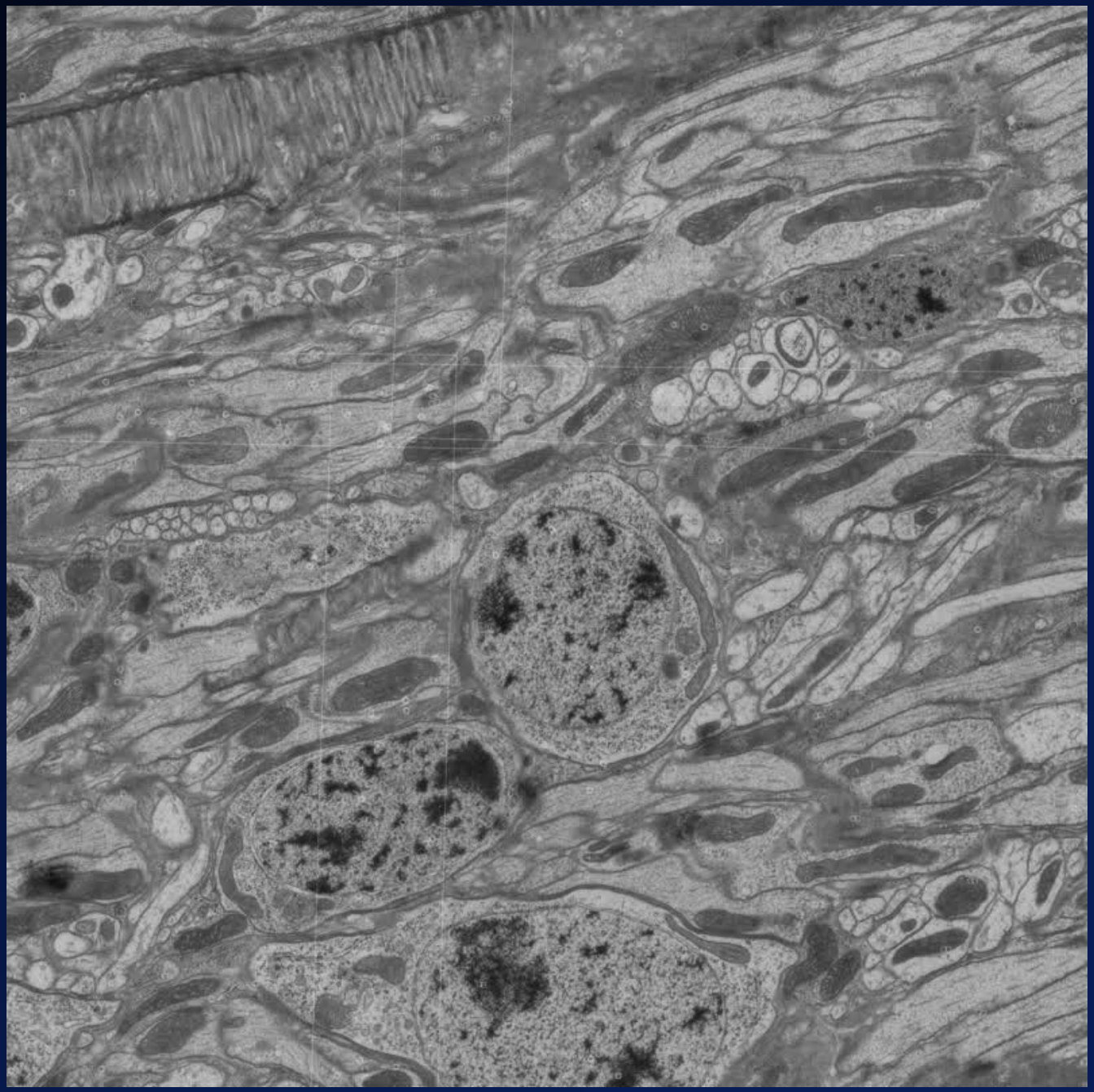


Large software systems

- Neural reconstruction
- Multi-site projects

Big software

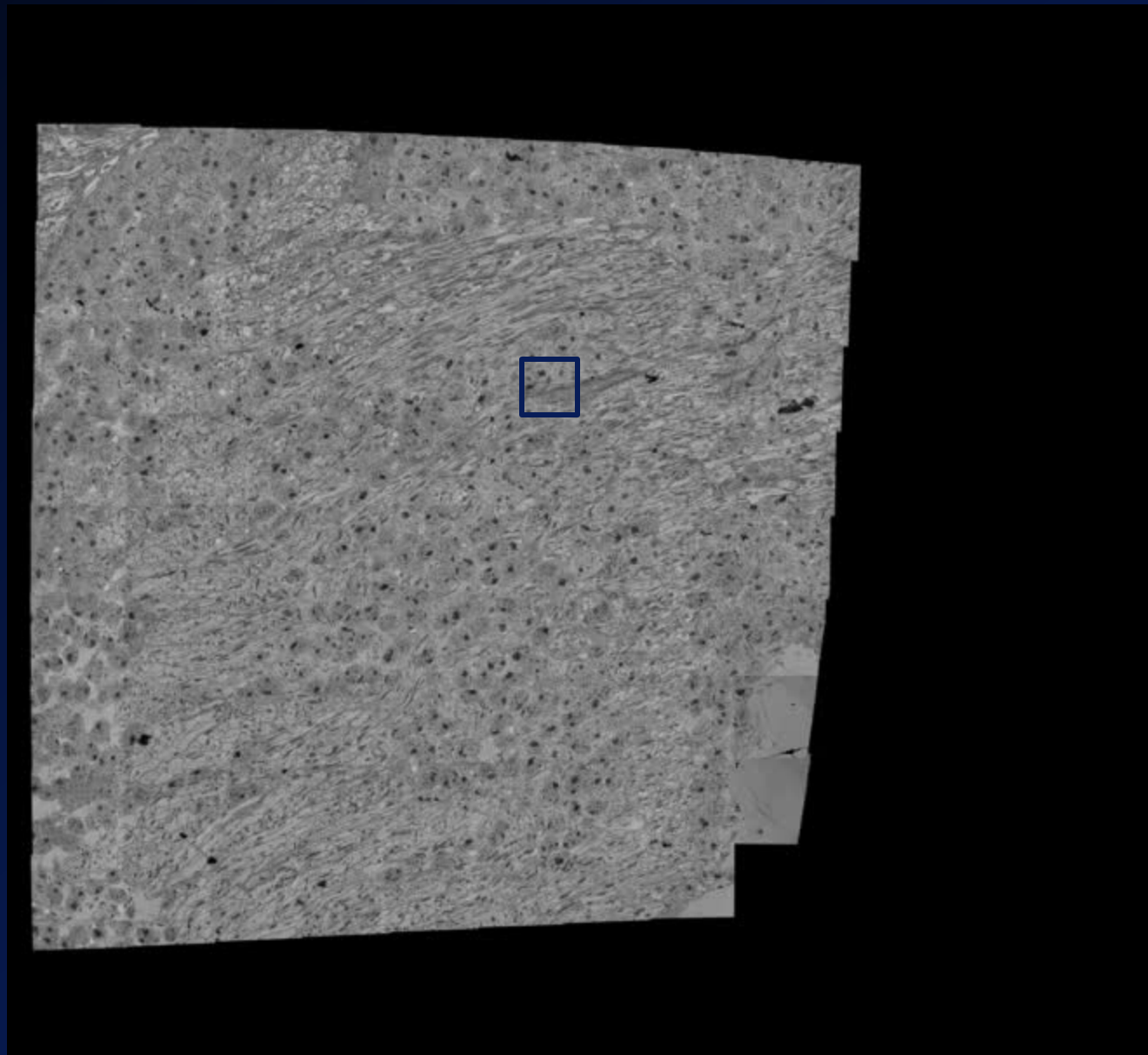
- Neural reconstruction
- 3-4 columns in the medulla
- One column took 2 person-years to correct
- 3Kx3K reduced to 500x500



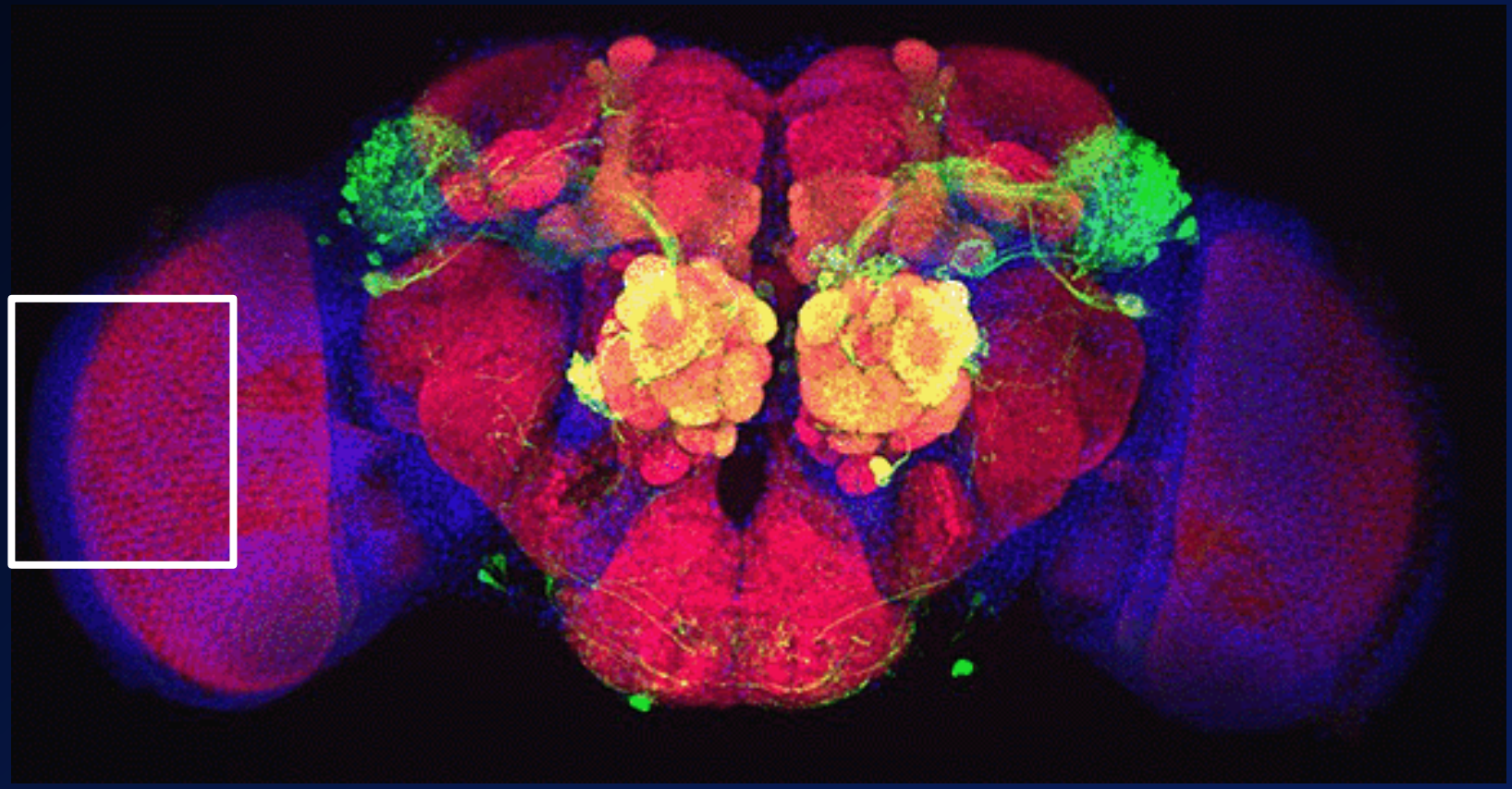
Just a small part

Now perhaps 250 columns of the medulla (out of 800 in total)

2 Terabytes of images
9 by 9 by 1700 stack

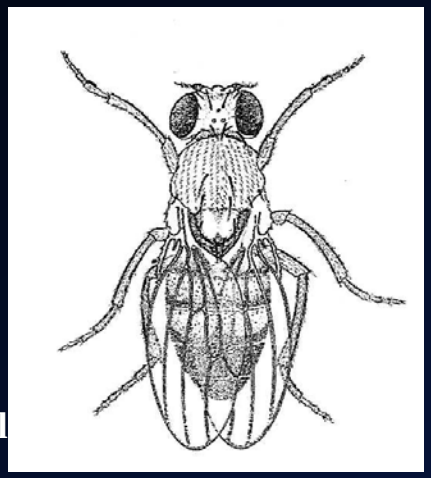


And this is only part of the brain



Janelia Farm, Howard Hughes Medical Institute

And this is a (small) fly

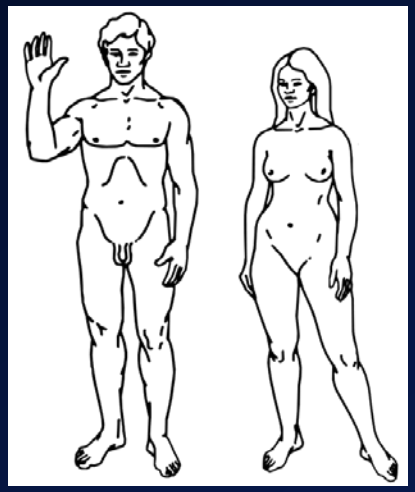


Perrimon l

100x

100x

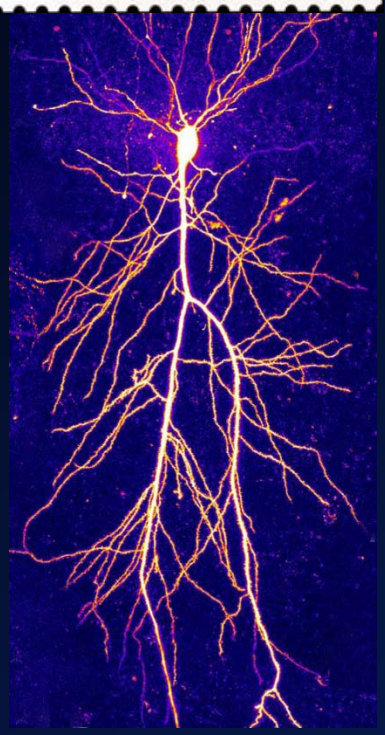
100x



Big software: Many sites and many methods

- http://elegans.swmed.edu/Worm_labs/ lists 232 groups working on C. Elegans
- http://flybase.org lists researchers from Aaronson to Zykov, 7613 in all
- Many methods; each method and each combination needs software
  - Optical
  - Electron Microscope (EM)
  - Genetic
  - Molecular
  - Cross method integration especially needed

Conclusions



- Biology presents many of the same problems as EDA
  - Not completely understood, but enough to start
  - Problems are similar in spirit though different in detail
    - How it works, how it's built, large complex systems
  - Crying need for tools and software
- A natural extension for current physical design community

Caveats

- I’m new to this field myself
- Information here is believed (by me anyway) to be reliable
- But treat it like Wikipedia
  - In general it should be OK
  - But before basing any important decisions (like a career change) on it, check the primary sources!

Why steal ideas from Electrical Engineering?

- Both EE and biology perform computation using large networks of tiny elements
- EE is 100 years old.

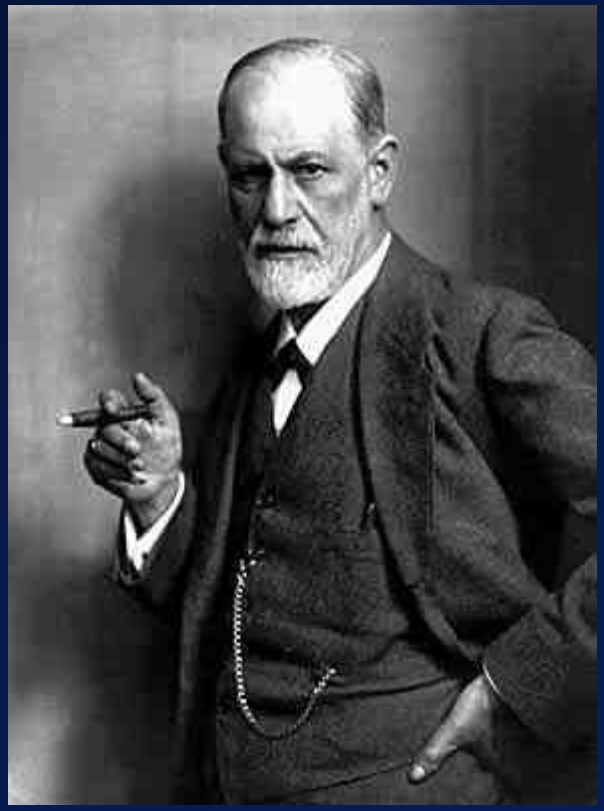
- It’s a $1,000,000,000,000 per year market
  - Very solid infrastructure of ideas and software

Our Goal: Understanding the Brain

- Many approaches are possible; almost all are being tried
  - Study the behavior of the organism and deduce brain function
  - Perturb the genetics and see how the function differs
  - Look at activity in areas of the brain
  - Statistical methods – look at large numbers of examples
- Each has limitations in terms of detailed understanding of function

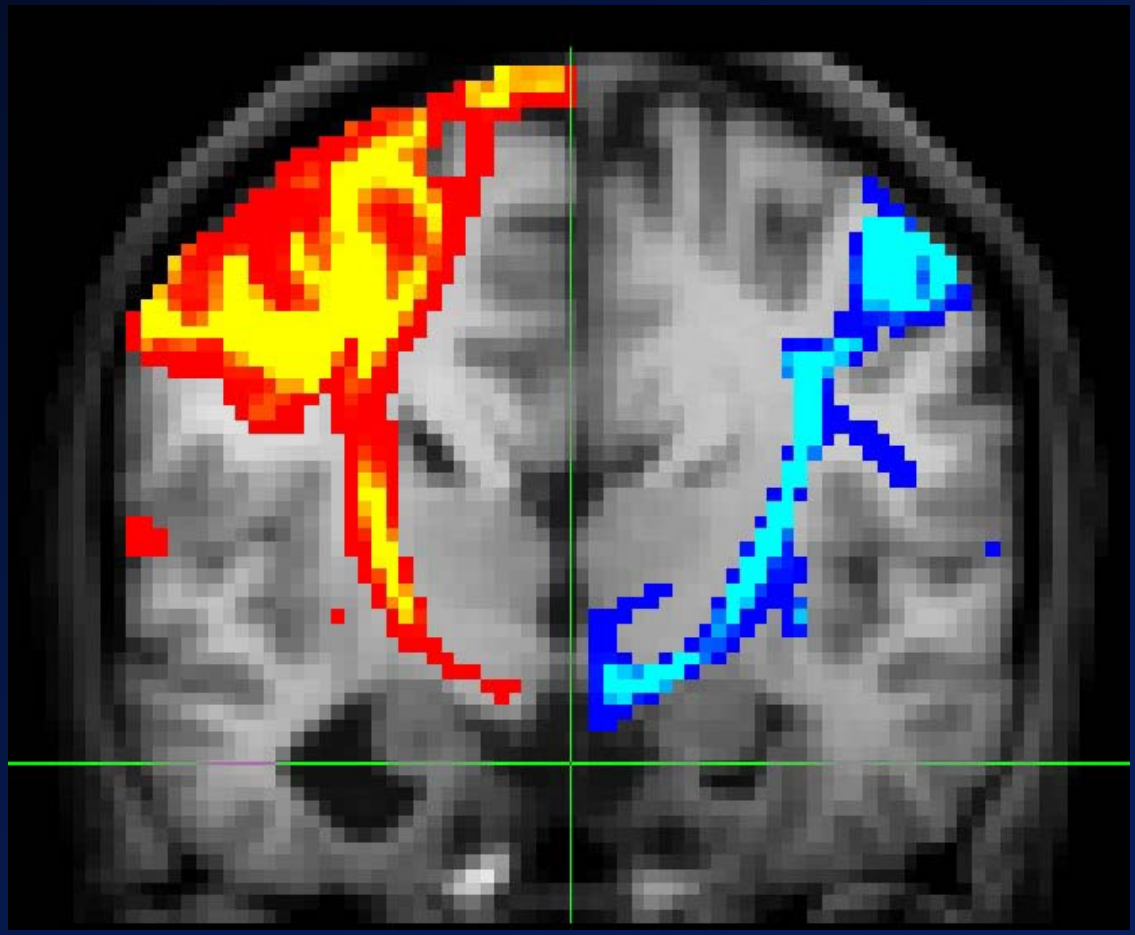
Trying to understand how the brain works from external behavior is hard

- After 150 years, people still debate Freud's theories
- In theory, results are at best ambiguous
  - Many structures can give precisely the same response



Techniques like functional MRI and PET scans look only at very large averages

- Like trying to figure out how planes work by looking at airport traffic



Janelia Farm, Howard Hughes Medical Institute

Genetic and statistical methods have limits

- Same genetic mechanism is often re-used in many places
- Function is a combination of many genes
  - Problems in finding genetic basis of diseases
- Statistical methods don’t reveal causes
- Evidence always admits of several possibilities
  - Does mercury in vaccines cause neural problems?

Alternative: take it apart to see how it works

- Idea is as old as engineering
  - Children are known for this approach
  - Patent system is a result of this method's success
  - Lots of historical examples

- Used in biology for more than 400 years
  - Starting with circulation of blood in the middle ages

But looking at brain structure is hard

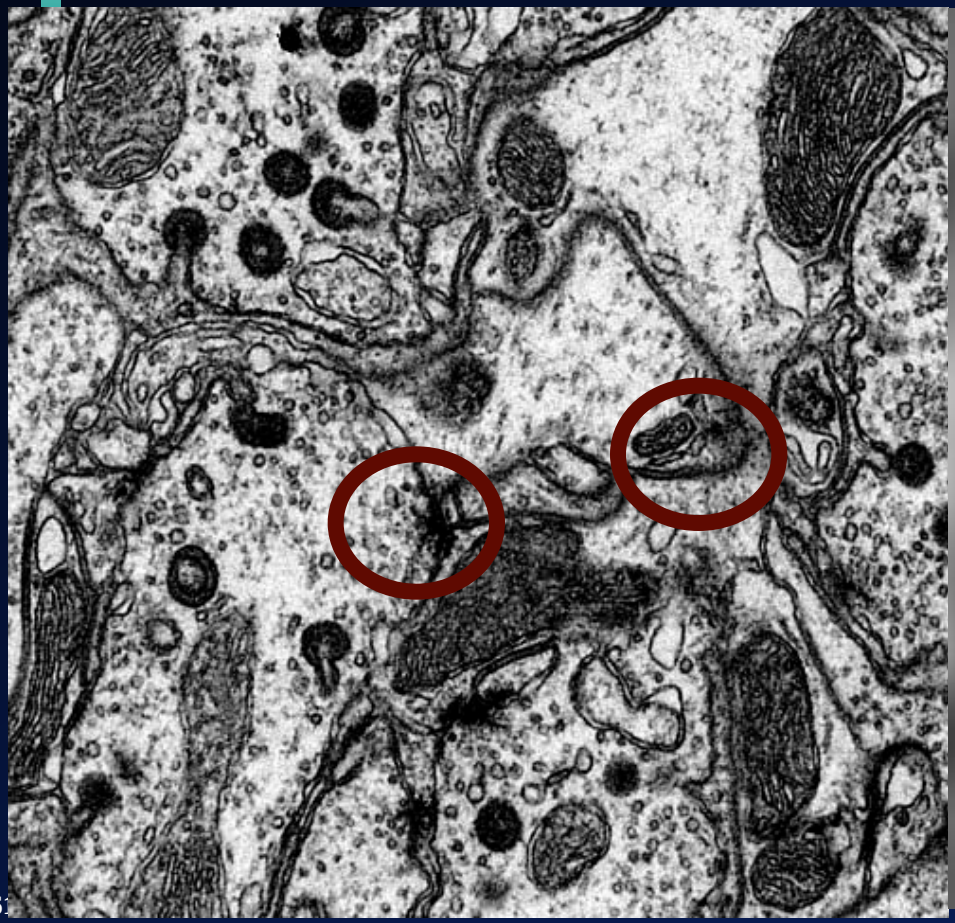
- Two main problems
  - Structures are very small
  - Network is very complex
- Until recently, only possible for very small animals with easy to resolve structure
  - C. Elegans, 302 brain cells, ~2K synapses
  - Took two decades and 10s of person-years

- Needed technical developments to make this feasible

Electron Microscopes make it possible

Electron microscope

Optical microscope



… But possible does not mean easy

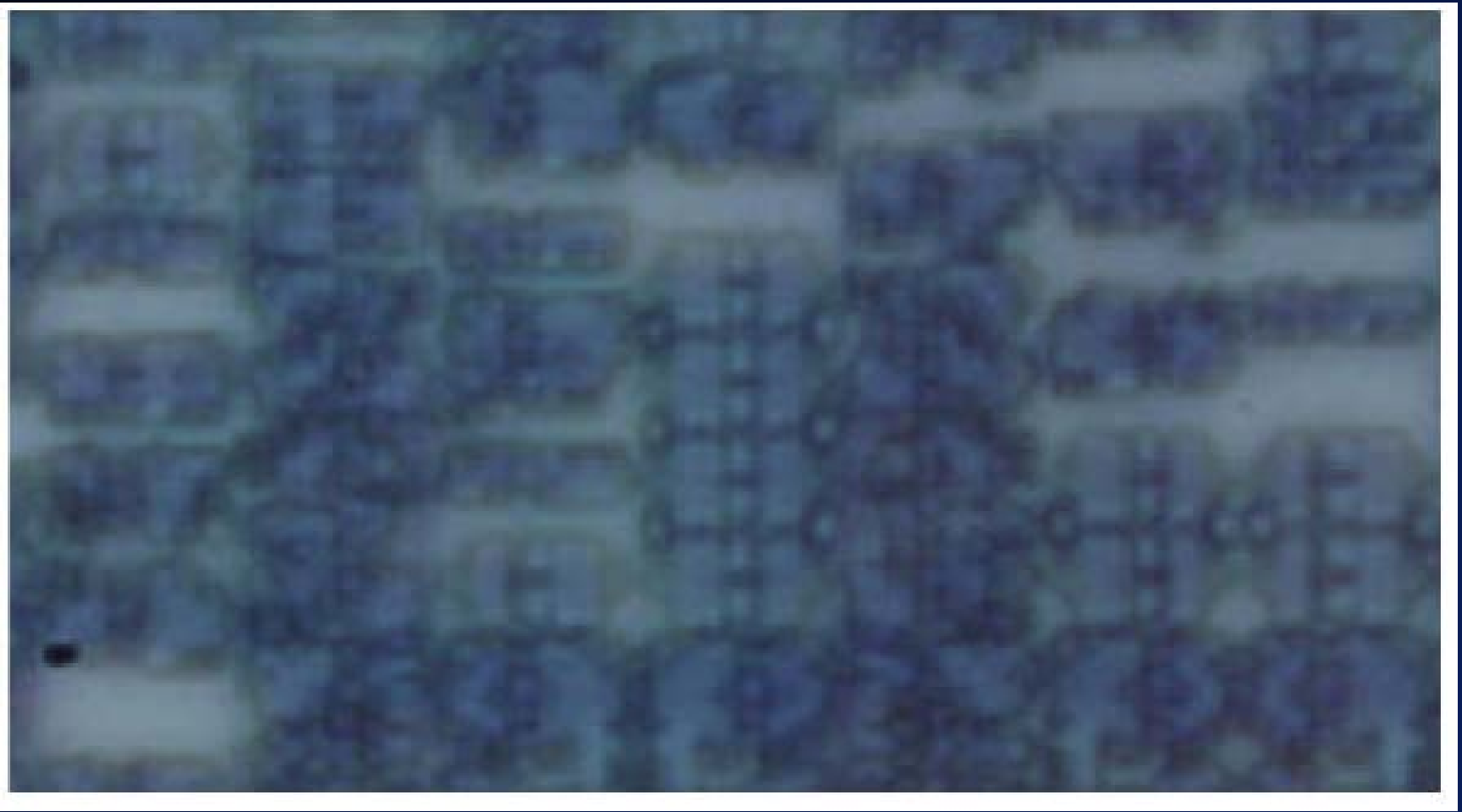
- Can do this manually now

Janelia Farm, Howard Hughes Medical Institute

There is another field with almost exactly the same problems

- Finding out exactly how a chip works from a physical example
- Needed because
  - Chip is out of production and need a replacement
  - Military intelligence
  - Competitive analysis
  - Legal enforcements of patents
- Similar technical problems of feature size and complexity

Optical microscopes can’t resolve chips anymore



“Reverse Engineering in the Semiconductor Industry”, Torrance and James

But electron microscopes give good images

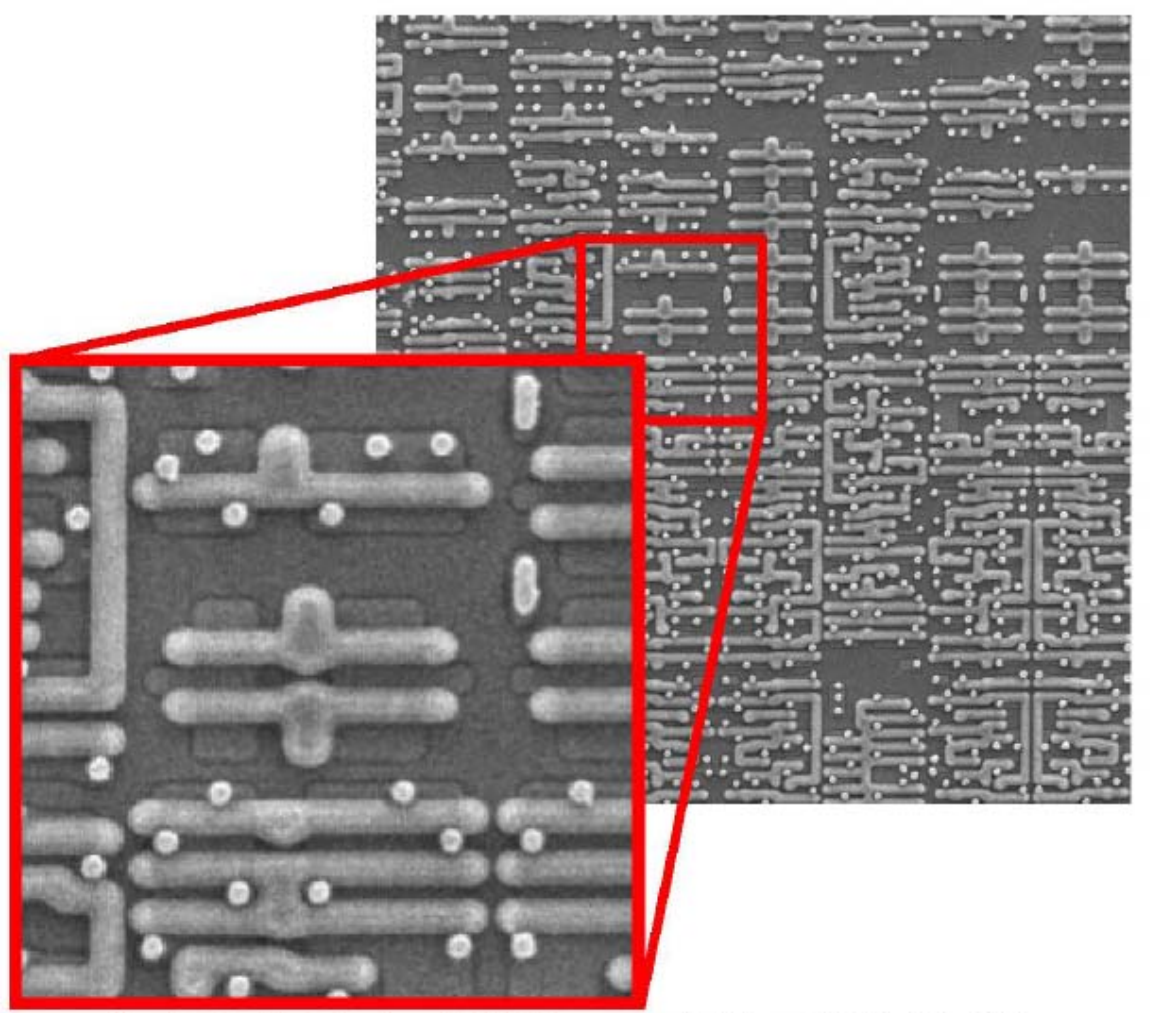


Fig. 8. Optical (top) and SEM images of 130-nm TI OMAP1510

"Reverse Engineering in the Semiconductor Industry", Torrance and James

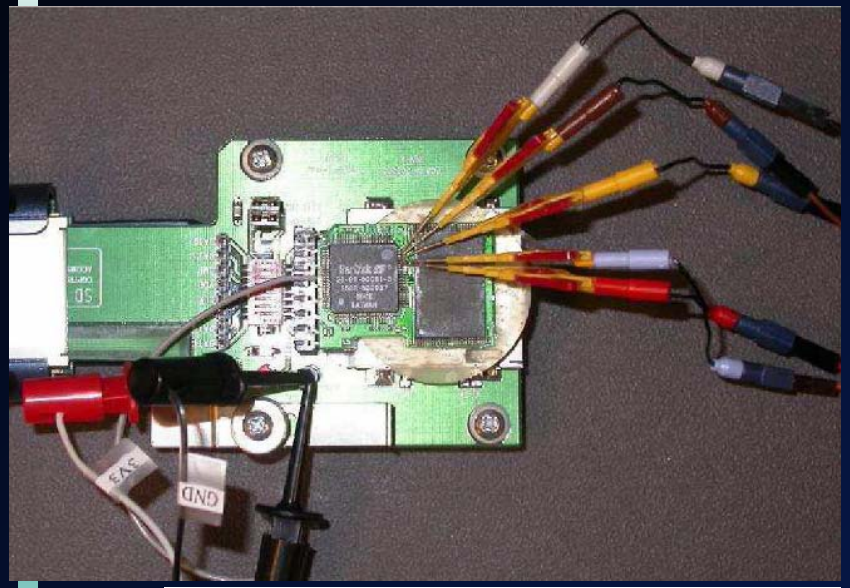
Results are large and hard to analyze

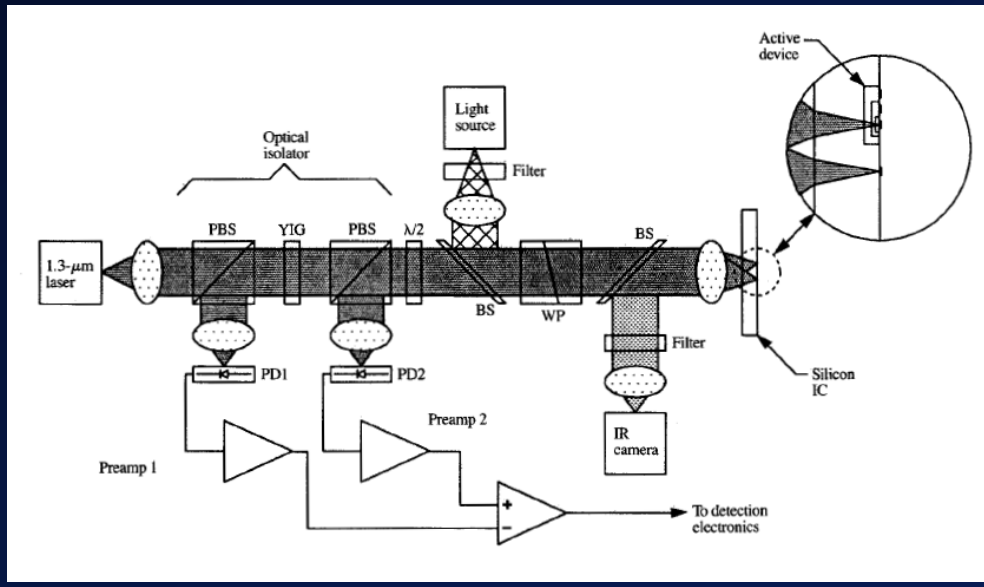
From “Chip Detectives”

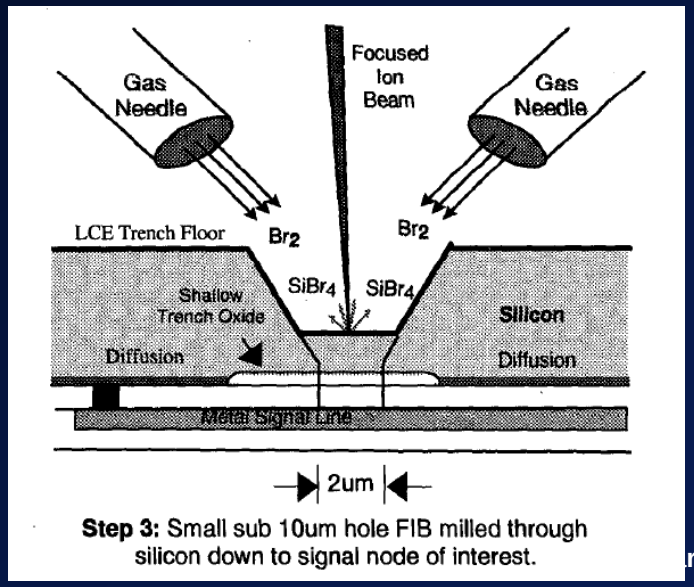
Fig. 5. As RE used to be done!

“Reverse Engineering in the Semiconductor Industry”, Torrance and James

Equivalent techniques in both fields



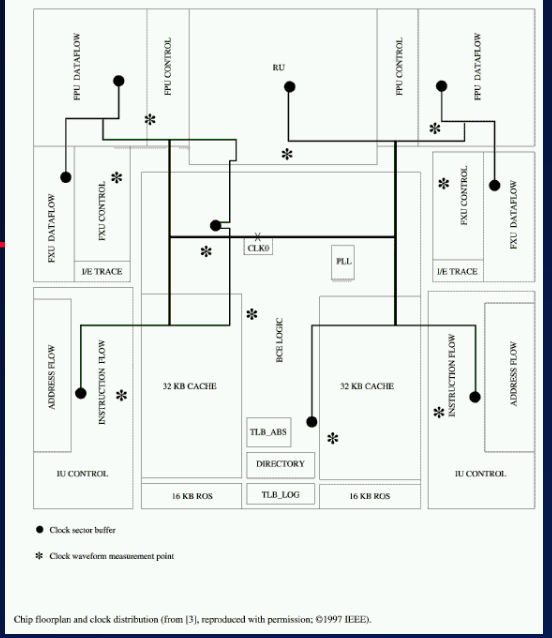




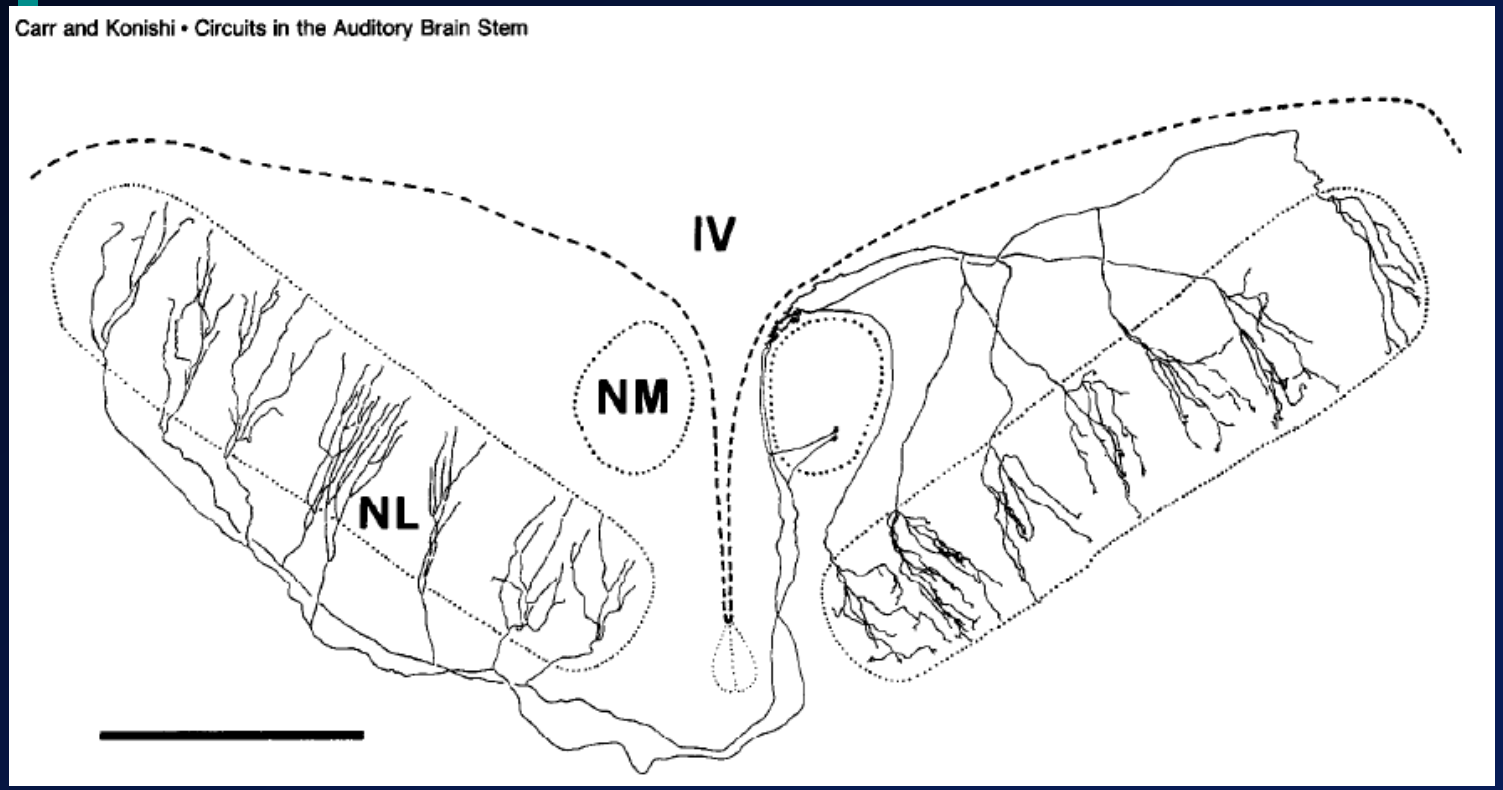
Step 3: Small sub 10um hole FIB milled through silicon down to signal node of interest.

Equivalent structures in both

- Clock tree on chip (IBM)
- Auditory circuits of barn owl.



Chip floorplan and clock distribution (from [3], reproduced with permission; ©1997 IEEE).



One big difference: Reverse engineering of chips is a well developed technology

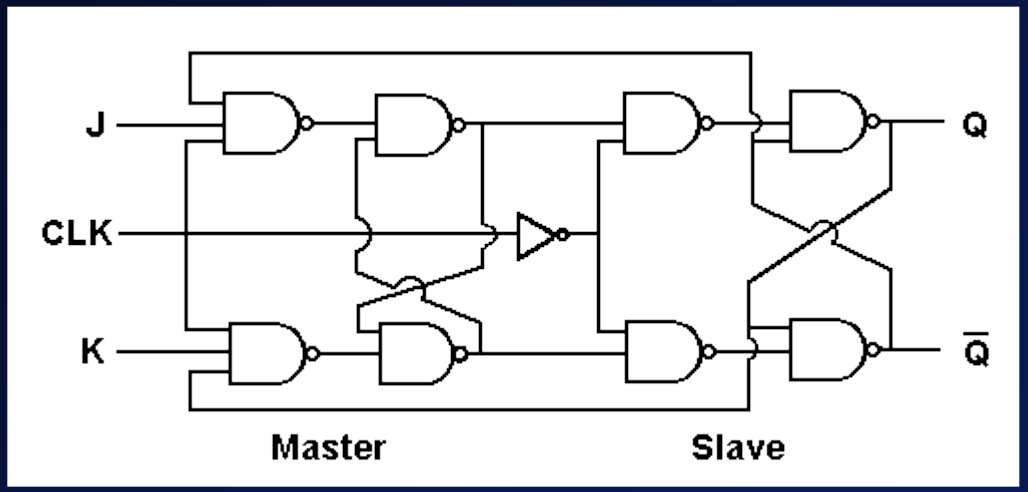
- Routinely done as a for-profit operation

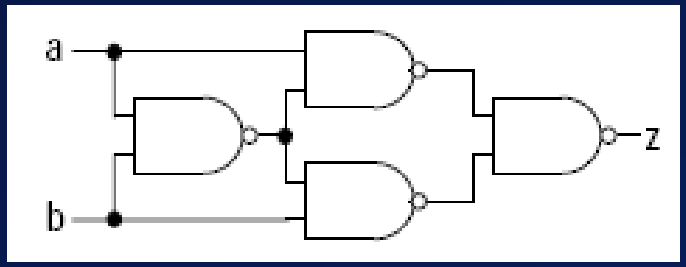
Possible techniques to borrow

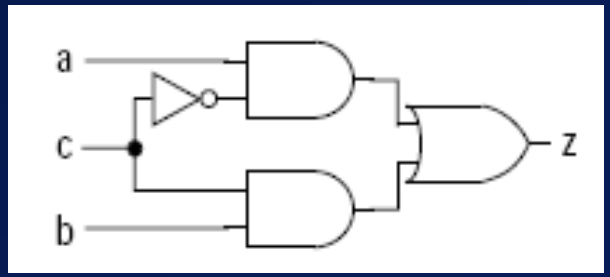
- Make automatic inferences more accurate by replacing hard decisions by probabalistic techniques
- Incorporate biological prior information in reconstruction
- Improve productivity using experience with similar graphical systems
- Attack up front the problems of a globally distributed, multi-group effort
- Plus many more speculative lines of attack

Use constraint that design uses known parts

- Chips are built from about 100 basic patterns
  - Three are shown below
- If you find something that is not one it's an error (usually) or a novel structure

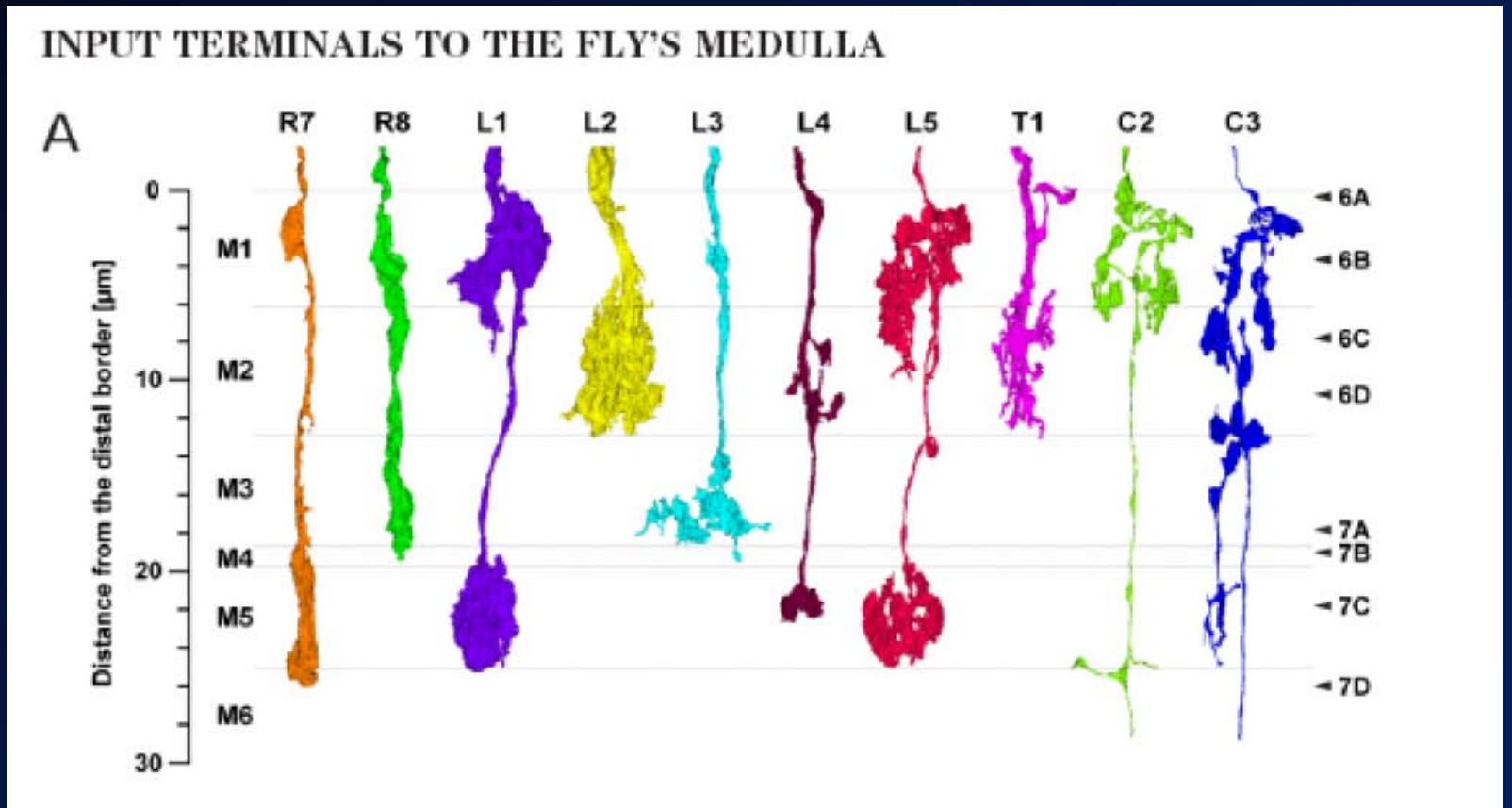






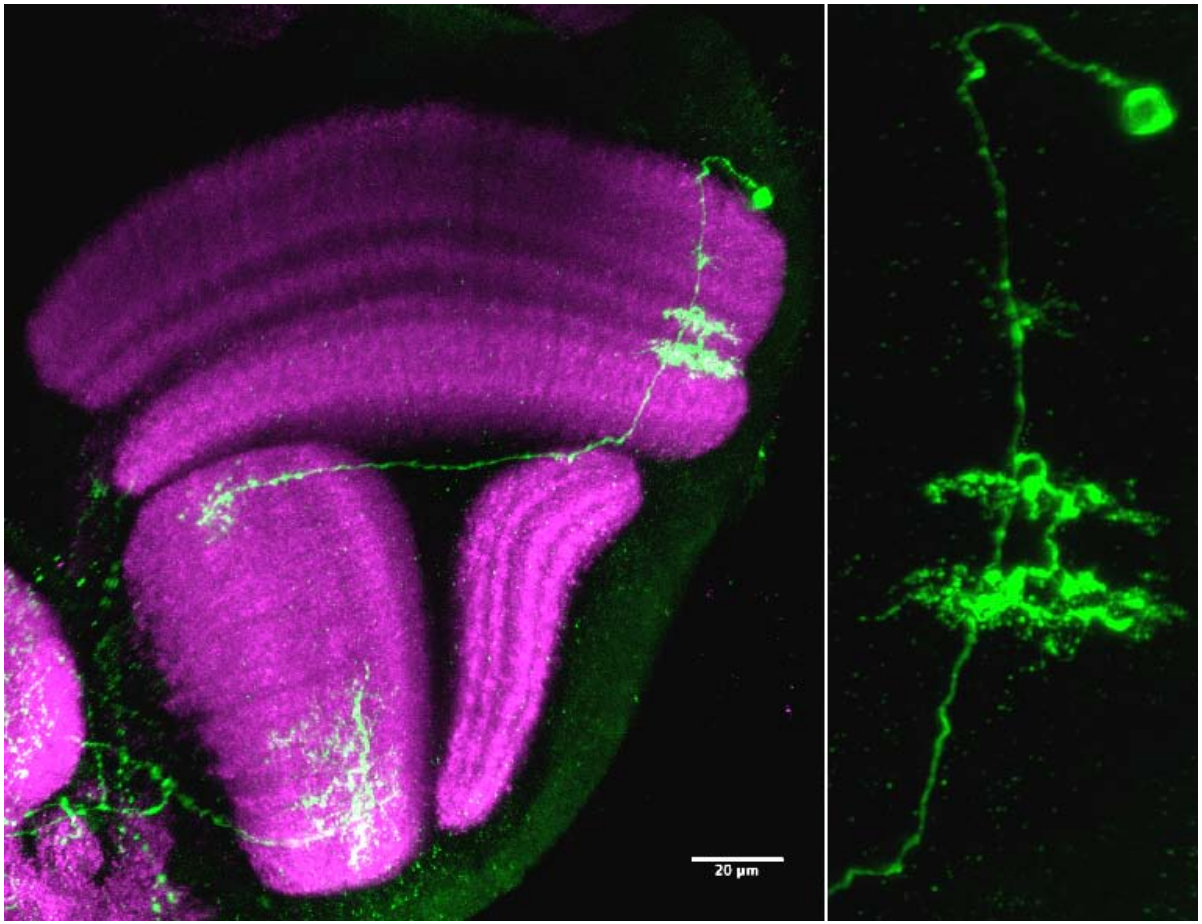
Use similar constraints from biology

- Genetics plus staining and optical techniques give us the library
  - Example – cells that go from the lamina to the medulla



Optical/genetic techniques give us the catalog

- Work of A. Nern here at Janelia



- Cannot show connections, but can show each type of component.
- Like a computer, millions of parts but only hundreds of types

Conclusions

- Brain analysis is a reverse engineering problem
- Reverse engineering of chips is a similar problem
  - Current chips about the scale of a fly's brain
- EEs have built lots of tools & software to aid in reverse engineering
  - At a minimum, can serve as a roadmap for what is needed in neuroscience
  - At best, maybe some of these tools can be used or adapted to aid in brain reconstruction and modeling

Probabalistic techniques

- We need to make billions of decisions
- Some of them will be wrong
- Need to correct based on constraints among decisions
- Example from EE: Low Density Parity Check codes
  - 10,000 decisions
  - 7.5% of them wrong
  - 10,000 constraints
  - Theoretical limit is 11% for this example

Credit: Prof. David J.C. MacKay, Cambridge

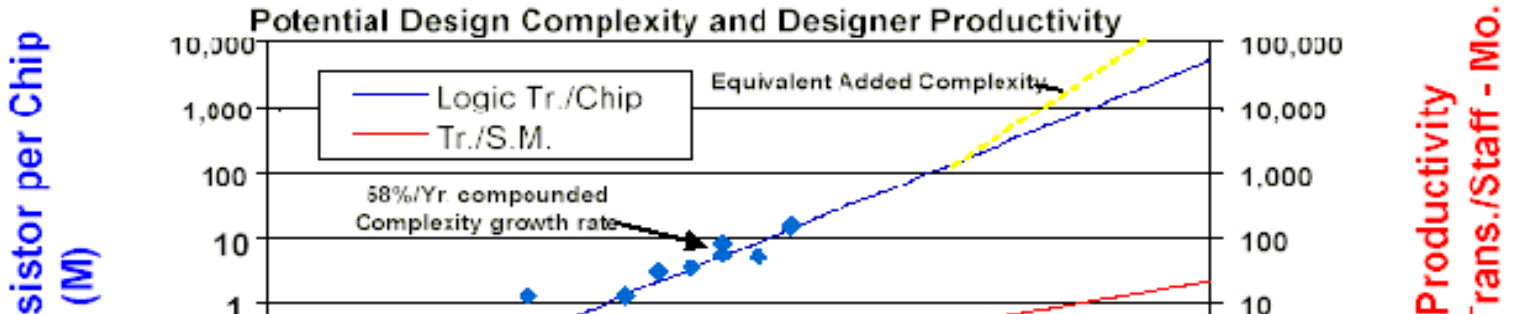
Tools to manage multi-site, multi-person efforts

- Large data sets 9x9x1900 = 153K images, 3 TB

- How to divide up the work?
- How/where is the data stored? In what formats?
- Network bandwidth needs
- Updates on other's work
- Software versions and compatibility
- Naming conventions
- Ergonomics

Janelia Farm, Howard Hughes Medical Institute

Imitate graphics/languages/tools that helped productivity increases in EE



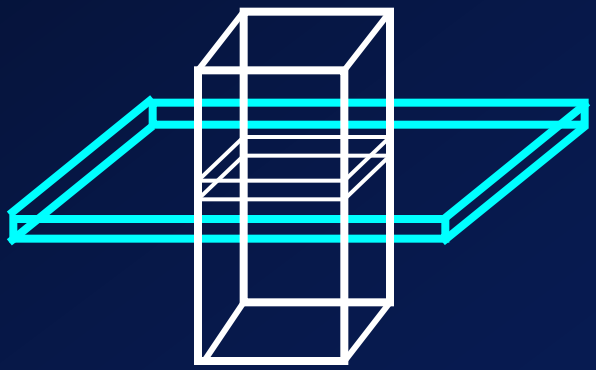
Year	Technology	Chip Complexity	Frequency	3 Yr. Design Staff	Staff Cost*
1997	250 nm	13 M Tr.	400	210	90 M
1998	250 nm	20 M Tr.	500	270	120 M
1999	180 nm	32 M Tr.	600	360	160 M
2002	130 nm	130 M Tr.	800	800	360 M

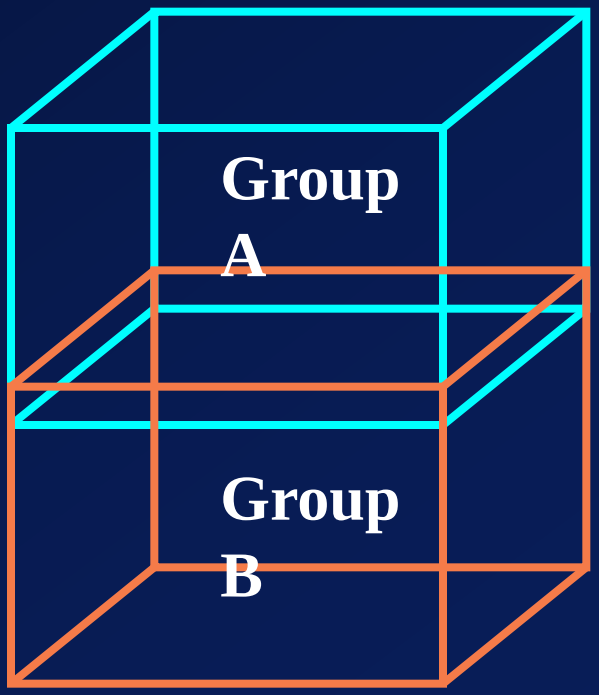
* @ $150K / Staff Yr. (In 1997 Dollars)

Allow changes in the middle of projects

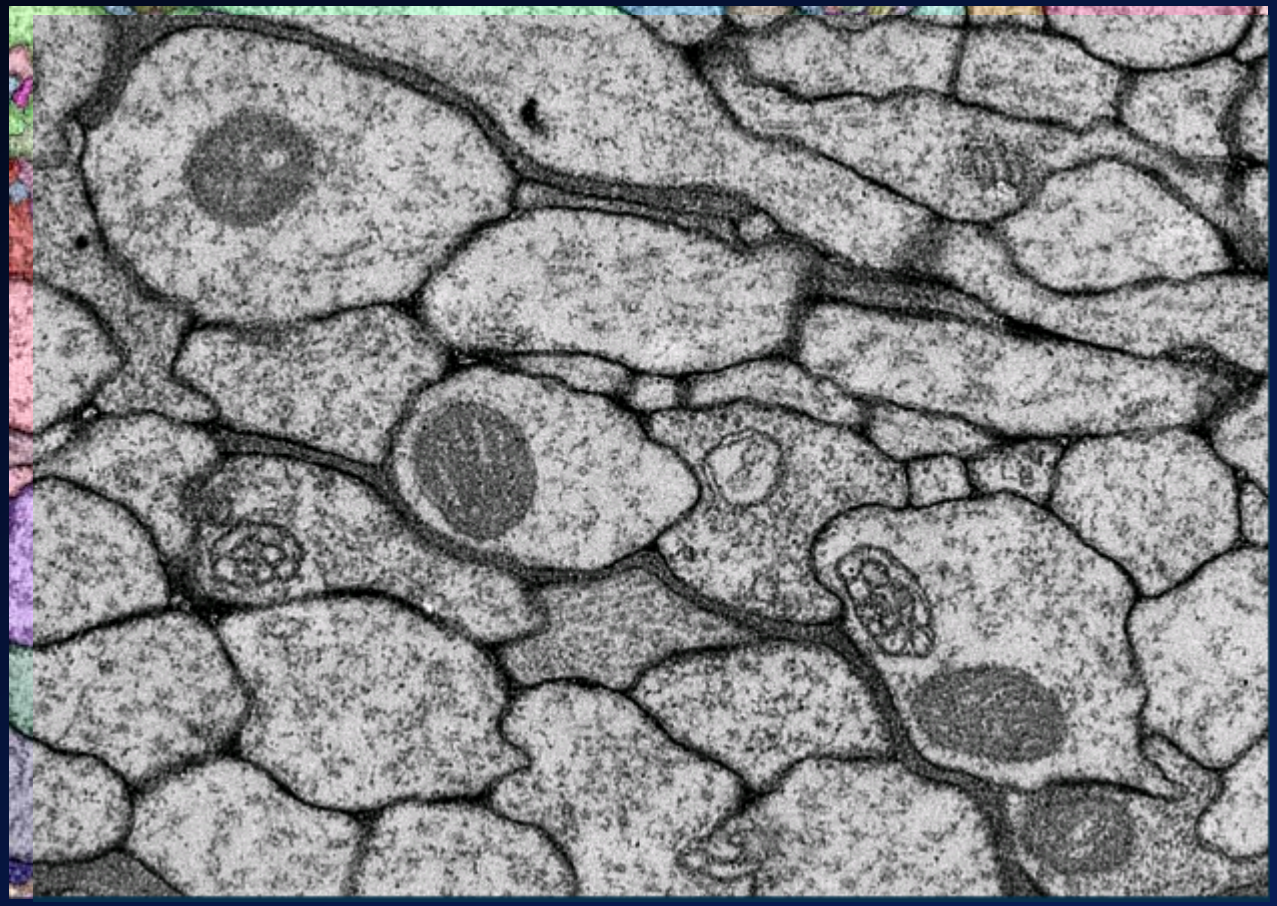
- Known as ECO – Engineering Change Order
  - Better algorithms
  - Enlarge the region of interest
  - Merge/split efforts

Image Segmentation





Replace hard decisions



EM Images

Image segmentation

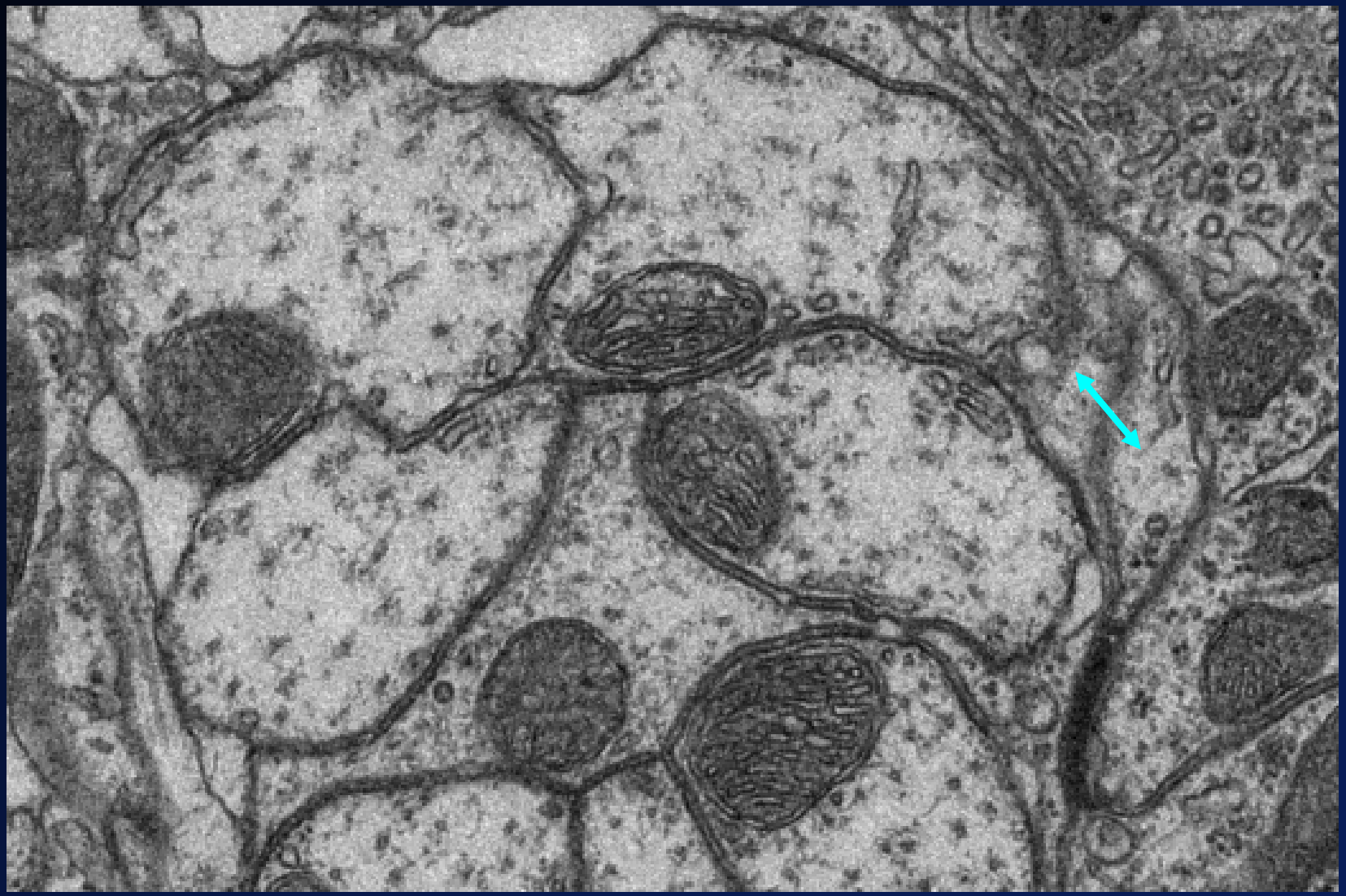
Link neurons in 3D

Manual Proofreading

Identify synapses

Circuit diagram for analysis

Data is ambiguous

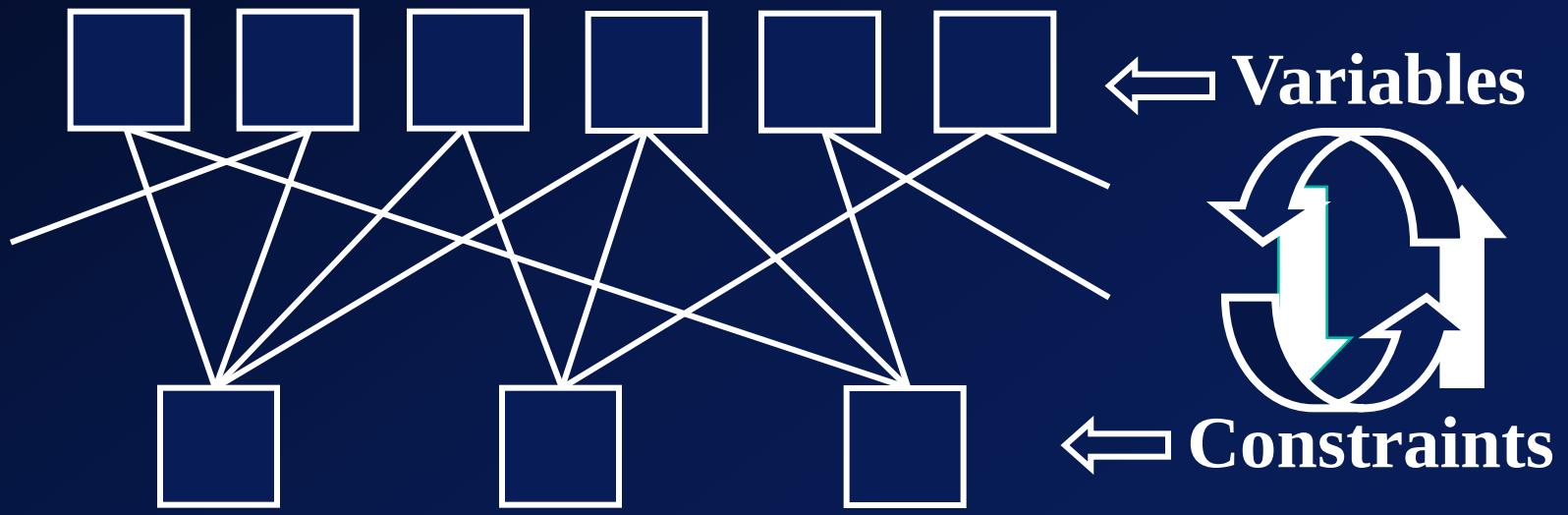


Replace early hard decisions with overall optimization

- ***When it is not necessary to make a decision, it is necessary not to make a decision."*** *Lord Falkland* (1610-1643)
- Express decisions in terms of probabilities
- Find best overall explanation for the data
  - Requires revisiting local decisions
- Many techniques already exist
  - Perform near the theoretical limit (of Shannon)
- Easy to add human input to the mix
- Algorithms are efficient: could be made real-time reaction during proofreading

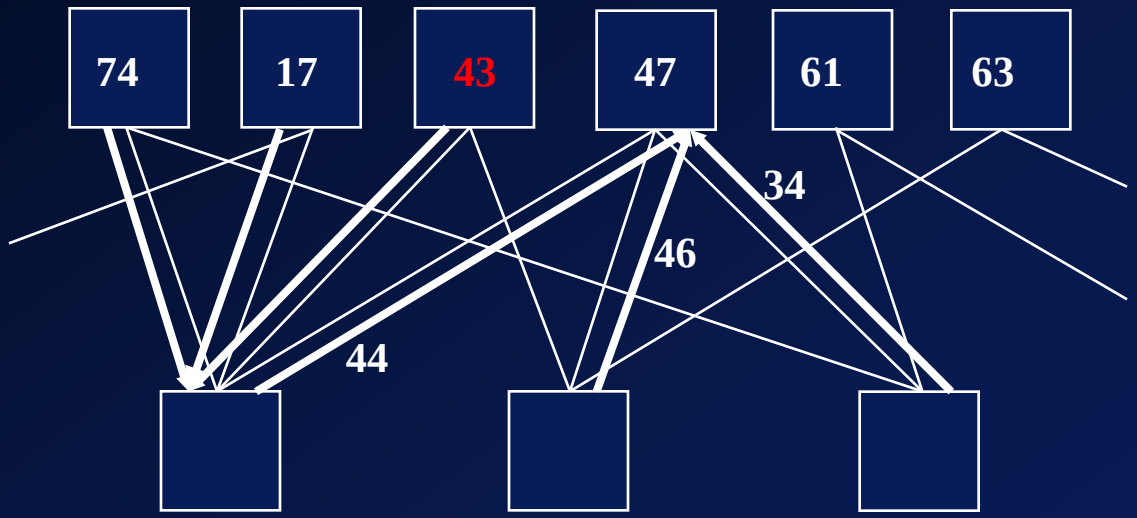
Belief propagation

- Used to solve many types of problems in engineering
  - Decoding of noisy signals
  - Finding clusters in data
  - Solving sets of boolean equations



Belief propagation

Data	1	0	1	0	1	1
Noise	+1.19	-1.20	-0.02	+1.50	+0.16	+0.25
Rcvd	2.19	-1.20	0.98	1.50	1.16	1.25
% a 1	84	15	62	73	65	68



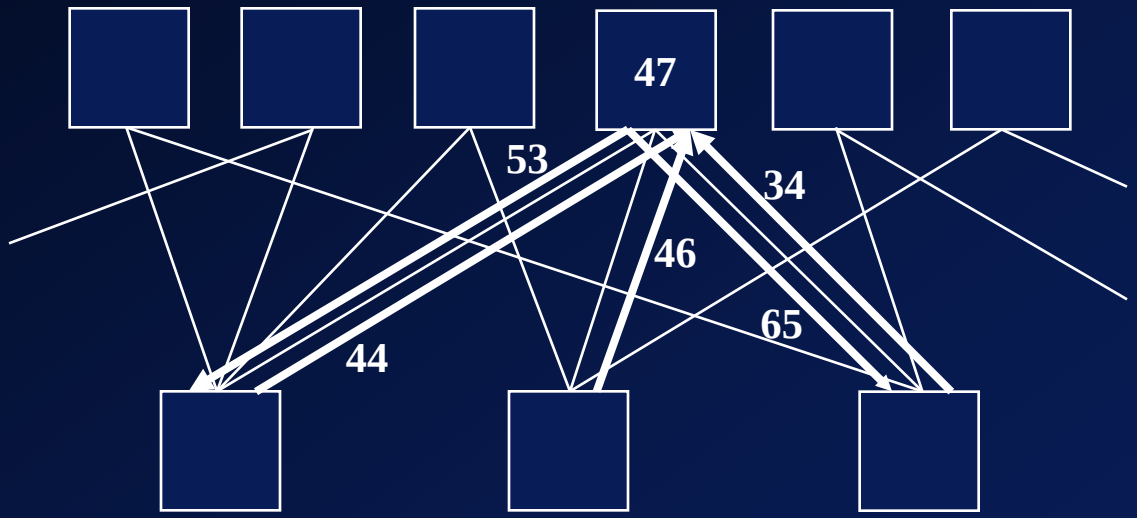
Variables

Constraints (must be even)

Belief propagation

Data	1	0	1	0	1	1
% of 1	84	15	62	73	65	68

Pass 2 %:	**74**	**17**	**42**	**47**	**61**	**63**
Pass 3 %:	**80**	**15**	**60**	**56**	**67**	**68**
Pass 4 %:	**76**	**16**	**56**	**48**	**62**	**67**

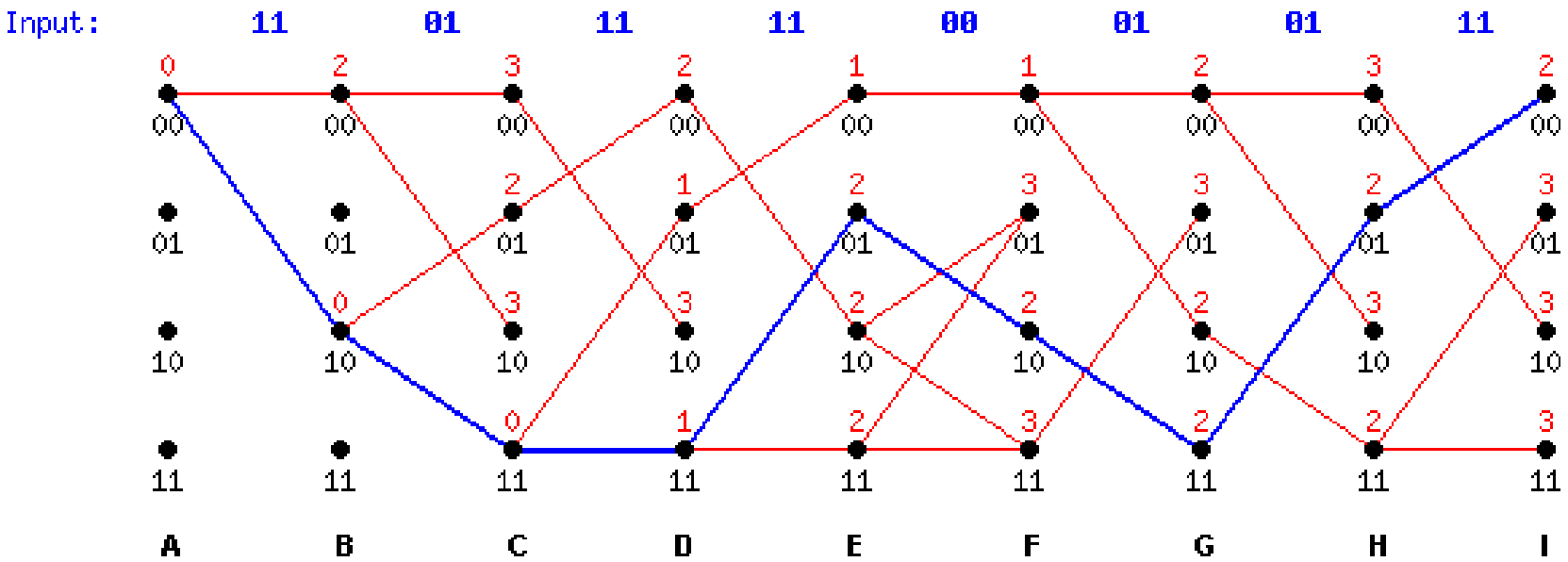


Variables

Constraints

Tracing consequences through many layers

- Viterbi decoding

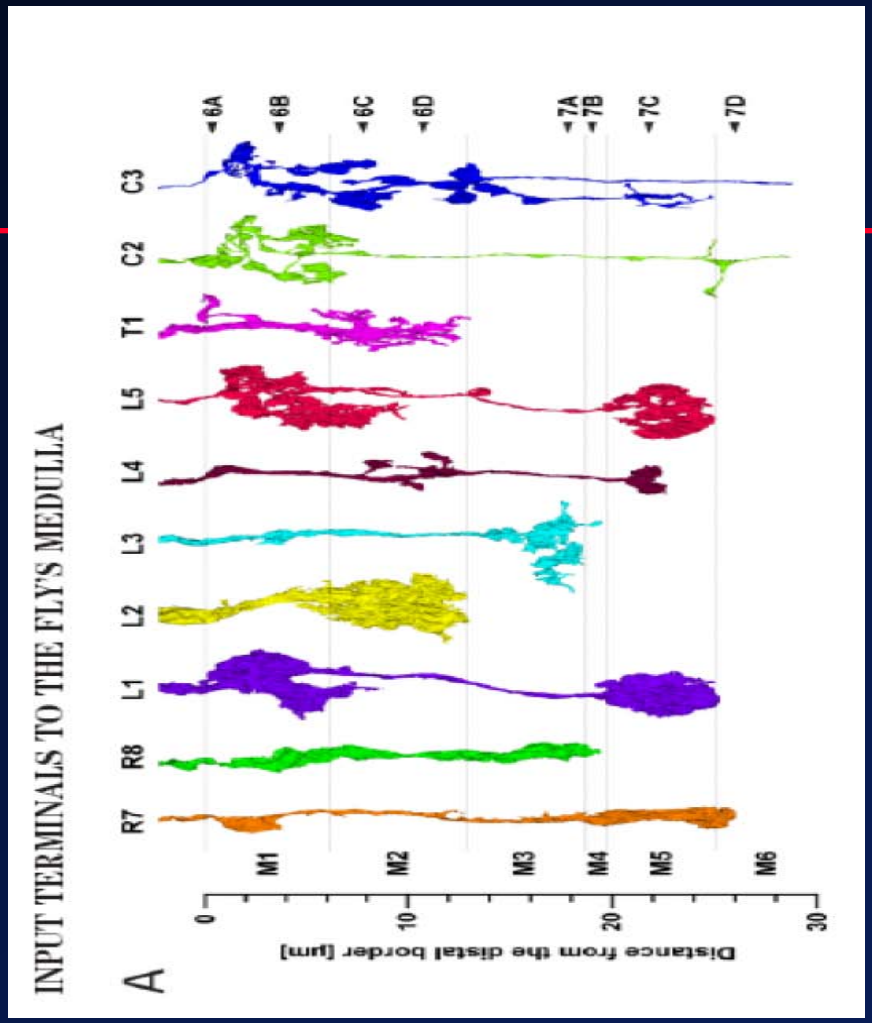


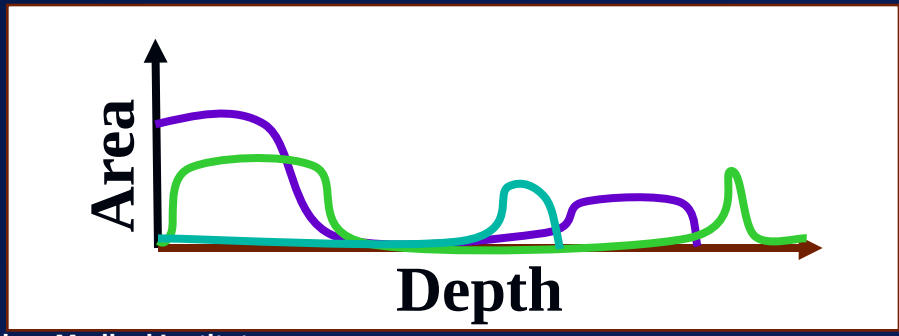
Incorporate biological priors

- We have lots of biological prior knowledge
  - Neuron types from previous work
  - Entry/exit knowledge
  - General biology (each cell has a nucleus, mitochondria are contained within cells, etc.)
- We can use this to help reconstruction
- Same idea used extensively in EE for the same reasons

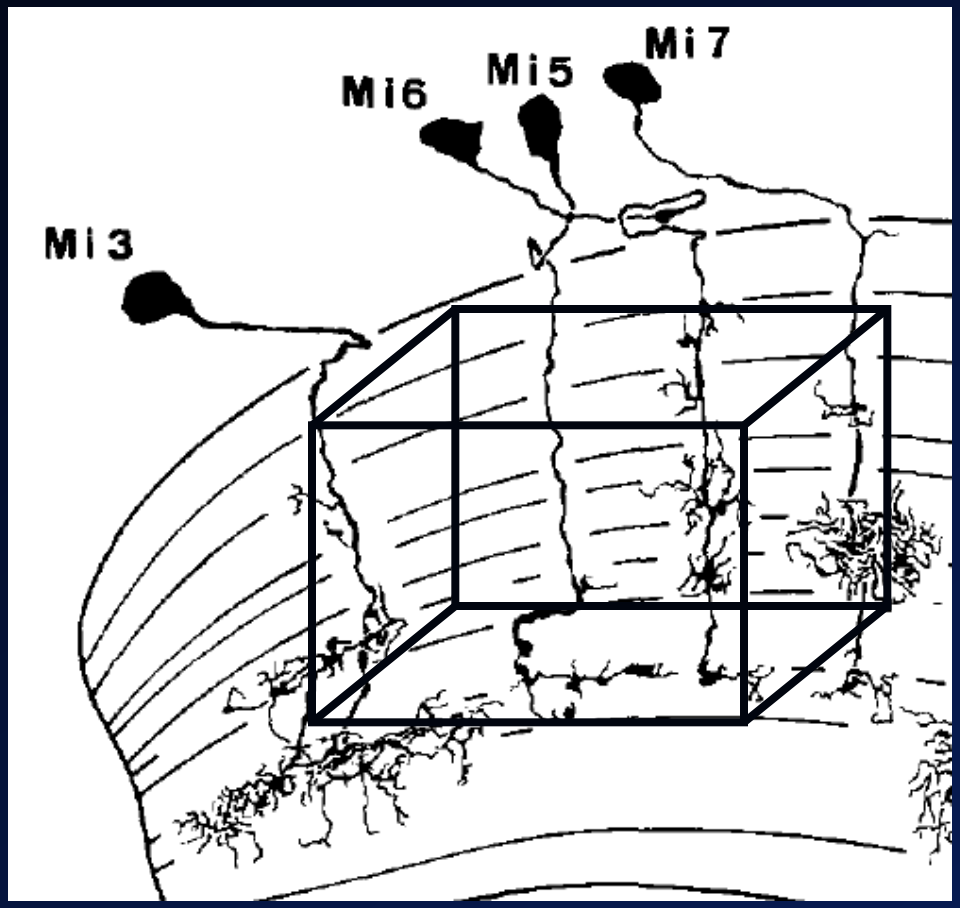
Use EE experience

- Need data in machine readable form
- Try to match each neuron
- If you can, use the info to improve work.
- If no match, either
  - Error in reconstruction
  - New type





Example 2 – input/output constraints



We know every neuron must connect to a face of the volume

Example: 5Kx5Kx10 layers

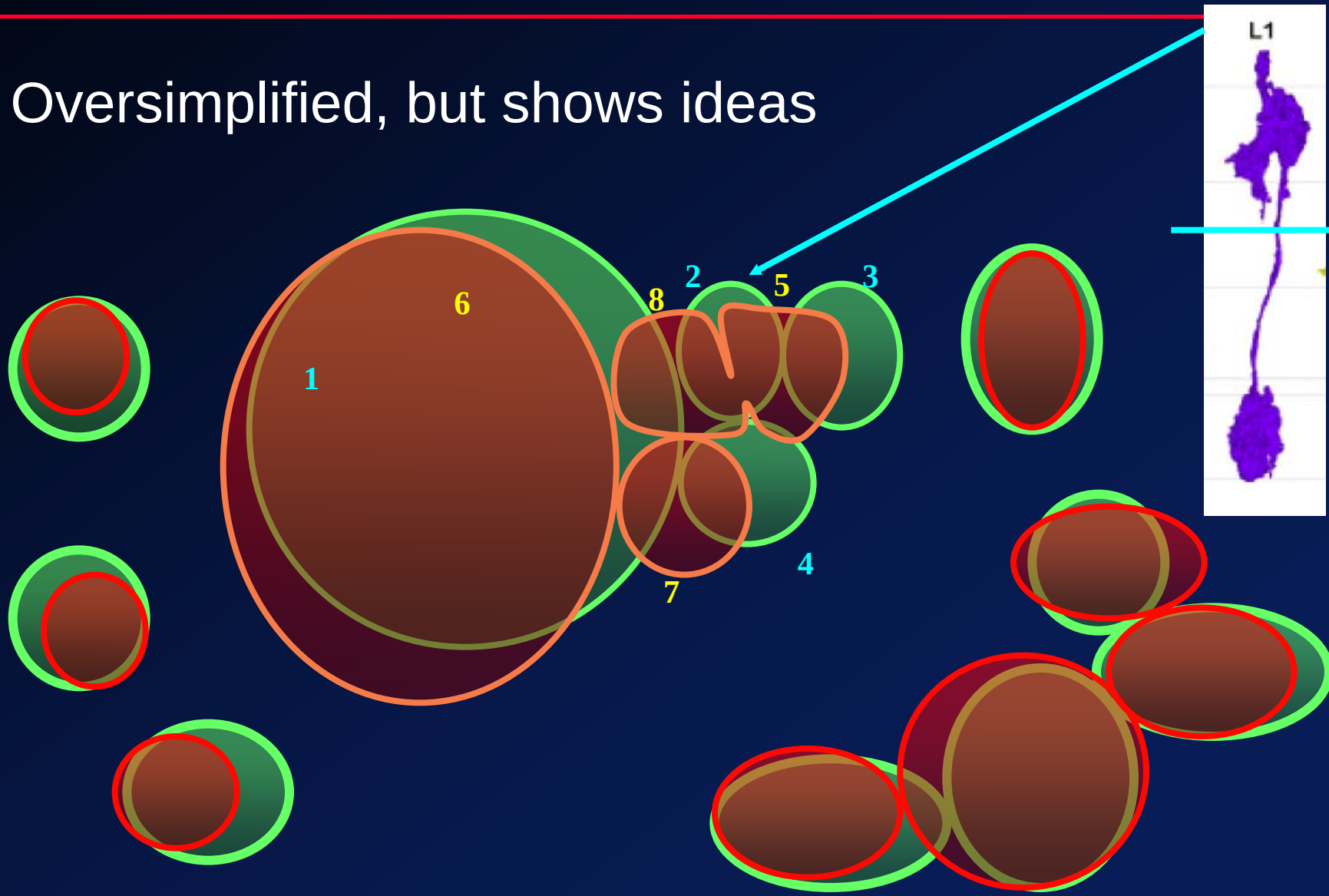
262K segments (sections of neurons)

18K connect to the face of the volume

➔ Lots of errors

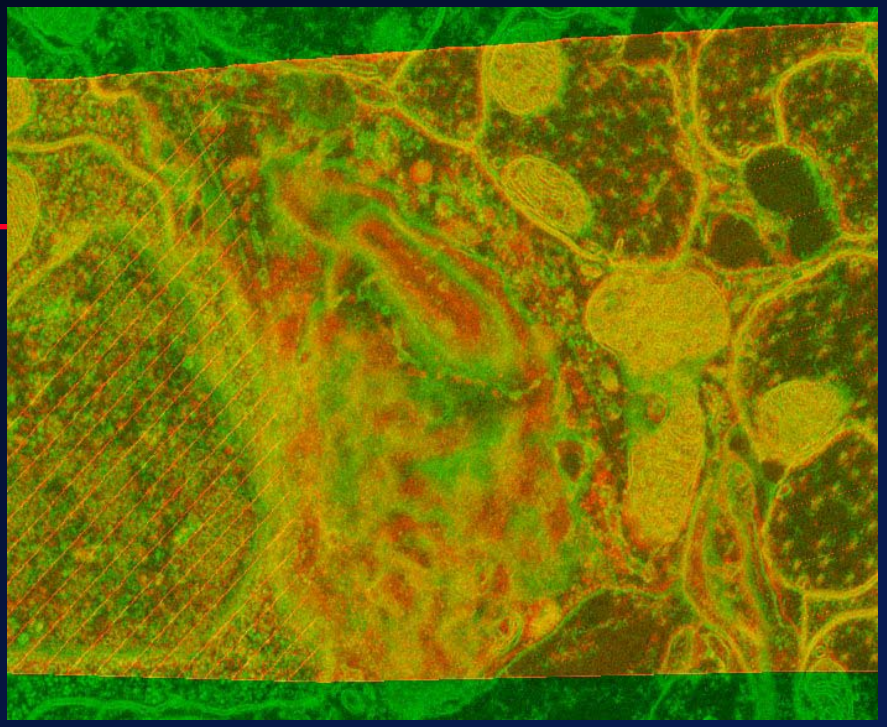
Example of biological prior and probability

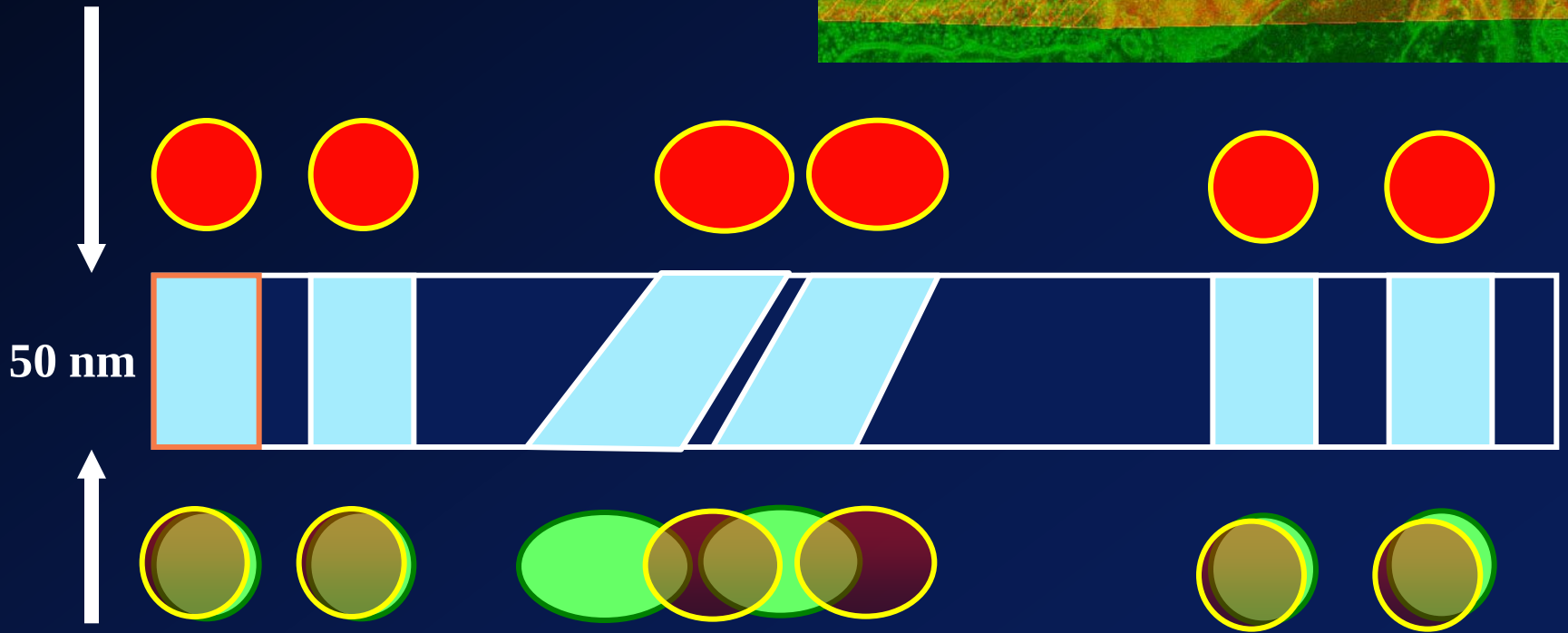
- Oversimplified, but shows ideas



Why not just align better?

- Because we can't

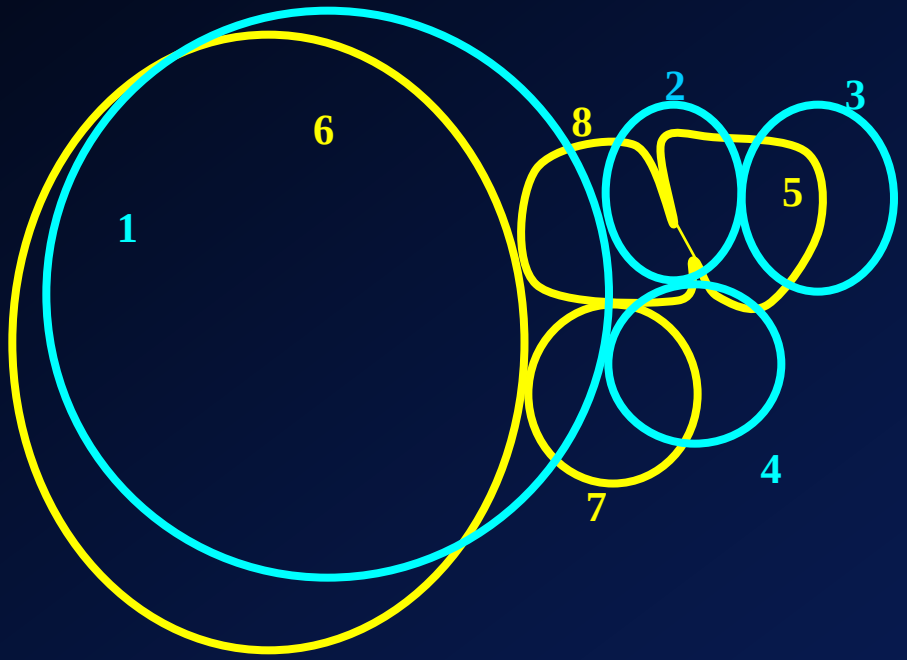


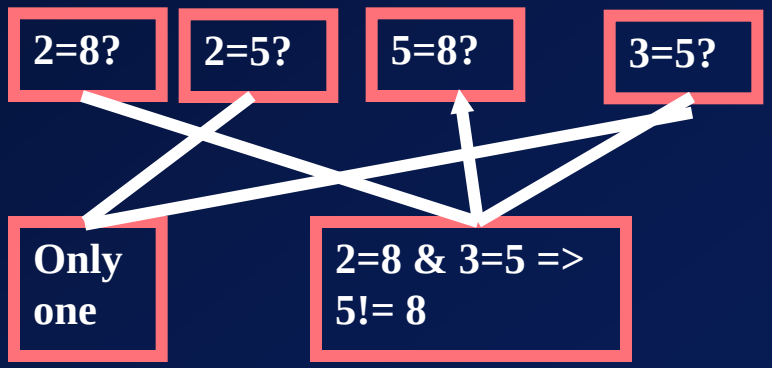


Janelia Farm, Howard Hughes Medical Institute

Example of biological prior and probability

Maps directly to belief propagation problem

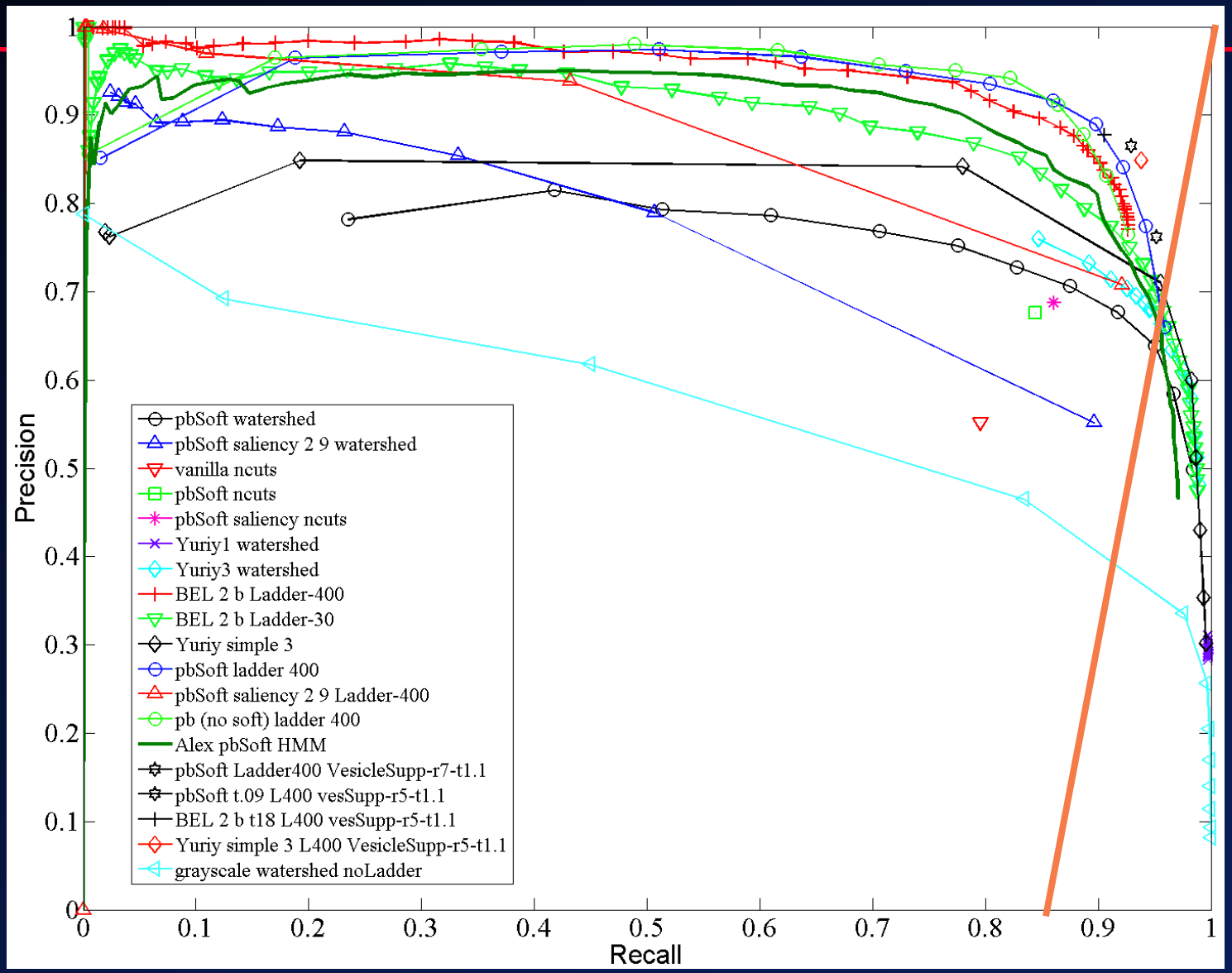




Using chip design experience to improve the reconstruction process

- Technically, task is like chip reverse engineering
- Operationally, more like chip design
- Lots of EE experience is potentially helpful
  - Organizing a multi-site effort on common data
  - Productivity enhancements
  - Accommodating changes during processing

Many of these techniques can be used in reconstruction

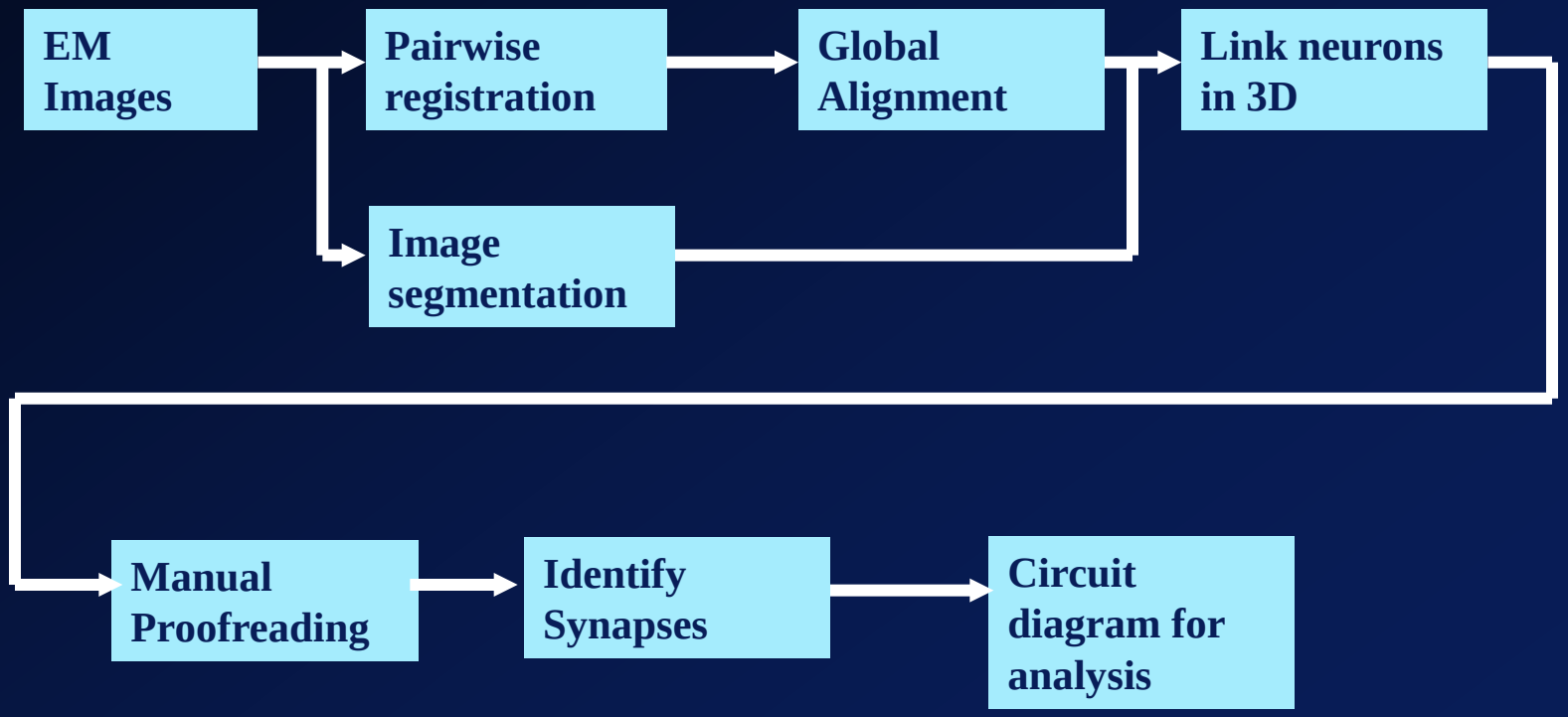


Fraction of detected boundaries that are correct

Fraction of ground truth boundaries that are detected

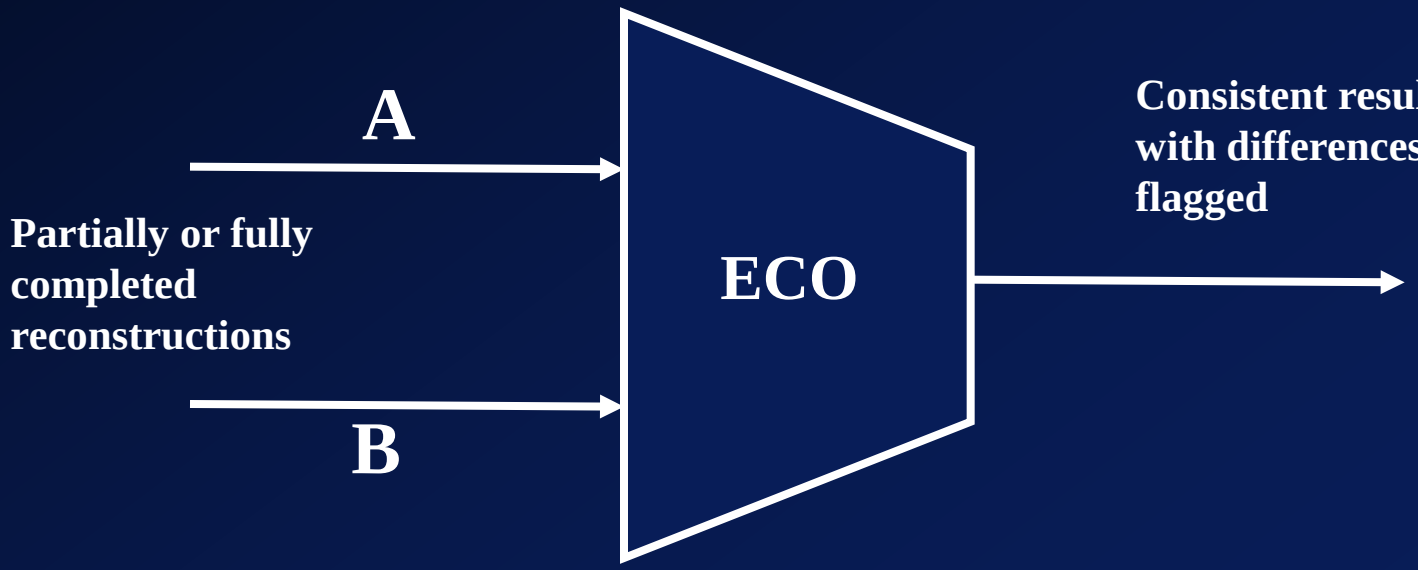
Changes during processing

- In theory, a step by step linear process
- Can start again if there is any problem
- But this kills both productivity and morale



One tool that's needed badly

- A common problem in engineering
  - In EE, called an Engineering Change Order, or ECO
  - In software engineering, a version merge
  - Need to update to a new design, keeping as much as possible of the old



Mechanics of ECO from chip experience

- Find a mapping from old<->new that includes as much information as possible
  - Need heuristics since this is NP complete
  - Many heuristics exist for corresponding EE problem.
- Construct the new problem
  - Easier to keep consistency
- Copy the old data where possible
- Report where this creates conflict (merge conflict)
- Mark appropriate regions for human attention

Other research possibilities

- What are these circuits doing, and how?
  - Positive and negative feedback, AGC, oscillators, parallel comp, etc.
- Better probe/instrumentation electronics
  - Cleaner signals/smaller electrodes
  - Digital circuits for near-analog timing
- Techniques to get more detailed data
  - Combined optical/EM on the same samples
  - Tip/tilt imaging of slices
  - Etc.

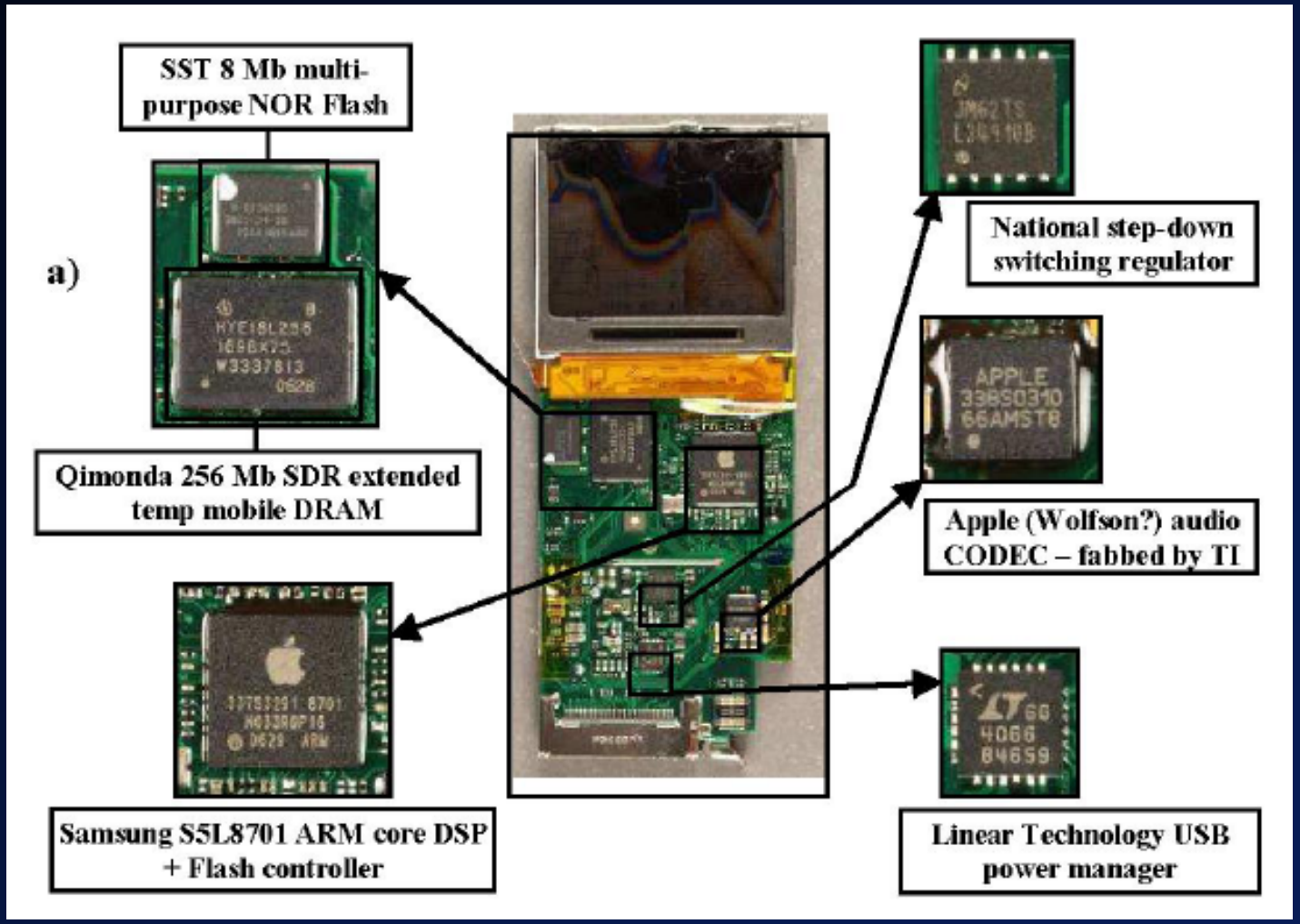
Conclusions

- Need reconstructions, but other techniques too
  - Reconstructions make them more efficient
- Lots of useful techniques/knowledge from chip design
  - Probabalistic techniques
  - Incorporation of prior knowledge
  - User interface and software engineering
  - Applications of EE techniques to biology & instrumentation

How does this relate to Dmitri's work?

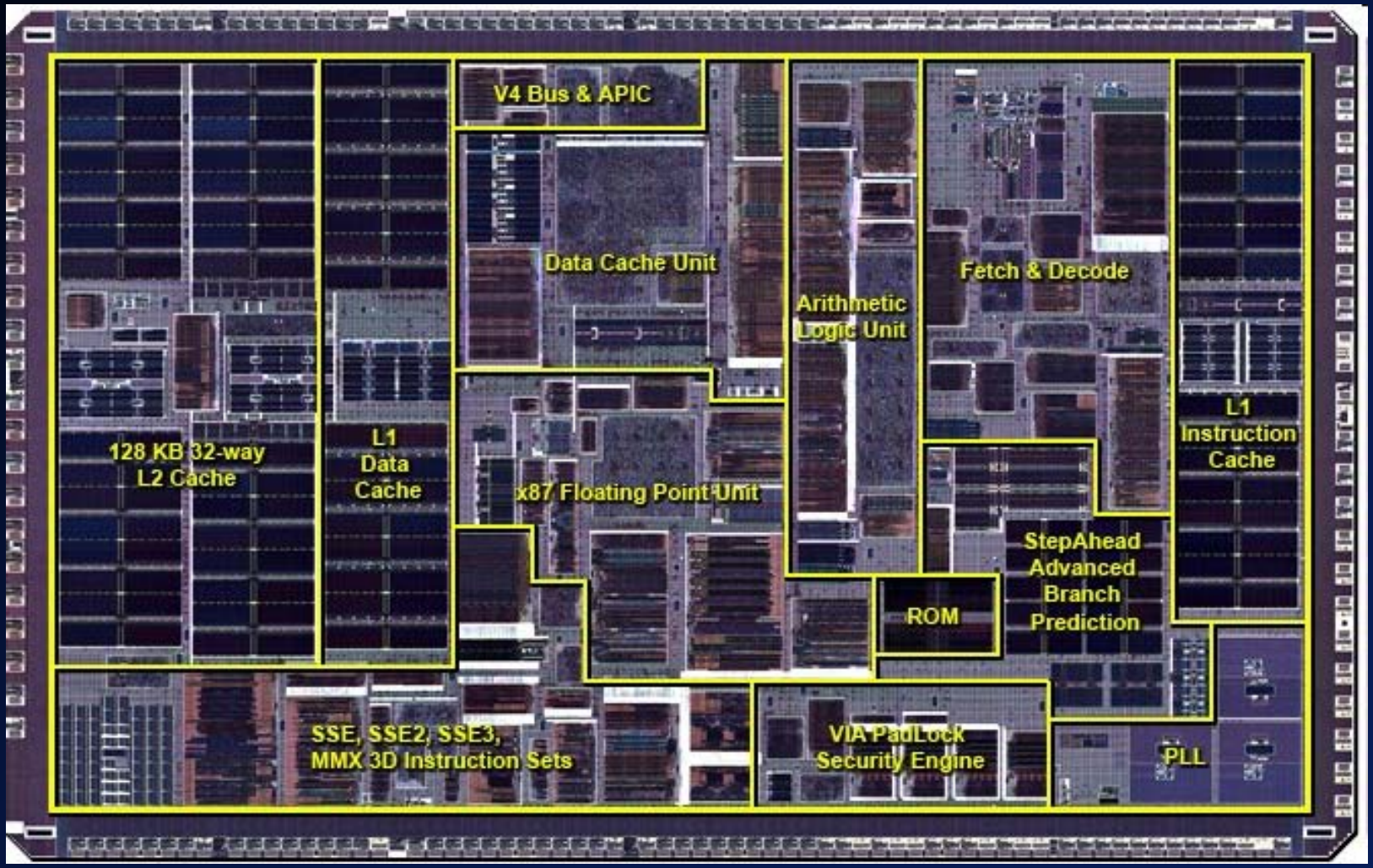
- Basically an addition to it

Example – Reverse Engineering a Cell Phone



"Reverse Engineering in the Semiconductor Industry", Torrance and James

Each chip has a number of functional units



C7 die layout from linuxdevices.com

Support Multi-group, multi-site work

- Technical process of EM re-construction is similar to the process used on chips
  - See previous talks
- But from an operational point of view, more similar to designing a chip
  - Need more than just algorithms
  - Work divided among teams (and probably locations)
  - Work needed not known exactly until job is finished
  - Software tools and data storage must support integration of results

How do EEs do this?

- Start with a goal (chip that does XYZ by next Christmas for Z dollars)
- Decide an overall chip architecture and floorplan
- Create sub-problems
  - Defined area, interface, cost, timing etc.
- Multiple teams start to work
  - Figure out their status, report on goals
  - Negotiations among groups
- Final integration

General observations

- Need a common interchange format
- Much easier if everyone uses the same software
  - Not strictly needed, but
  - Training time is minimized
  - Can trade results, people, etc.
- Interfaces naturally defined in the fly
- In EE, these are conserved.
  - Example – MP3, JPEG. Readers and writers change, but interface remains
  - Maybe also true in biology

Step 1 – prepare the samples

From “Chip Detectives”

General observations

- Groups need to know what other groups are doing
  - Partial results, confidence levels
- Groups need to figure out needs early
  - Computers and people
- Need ability to handle exceptions in an organized way

Getting a feel early

- Do whole result to low accuracy
- Pick a portion (the hardest portion if you know it) and do it to full accuracy
- Form estimate of how hard to do the full job to full accuracy.
- In reconstruction
  - Do the whole network fully automatically (will contain errors, but sizes the problem)
  - Do a portion fully

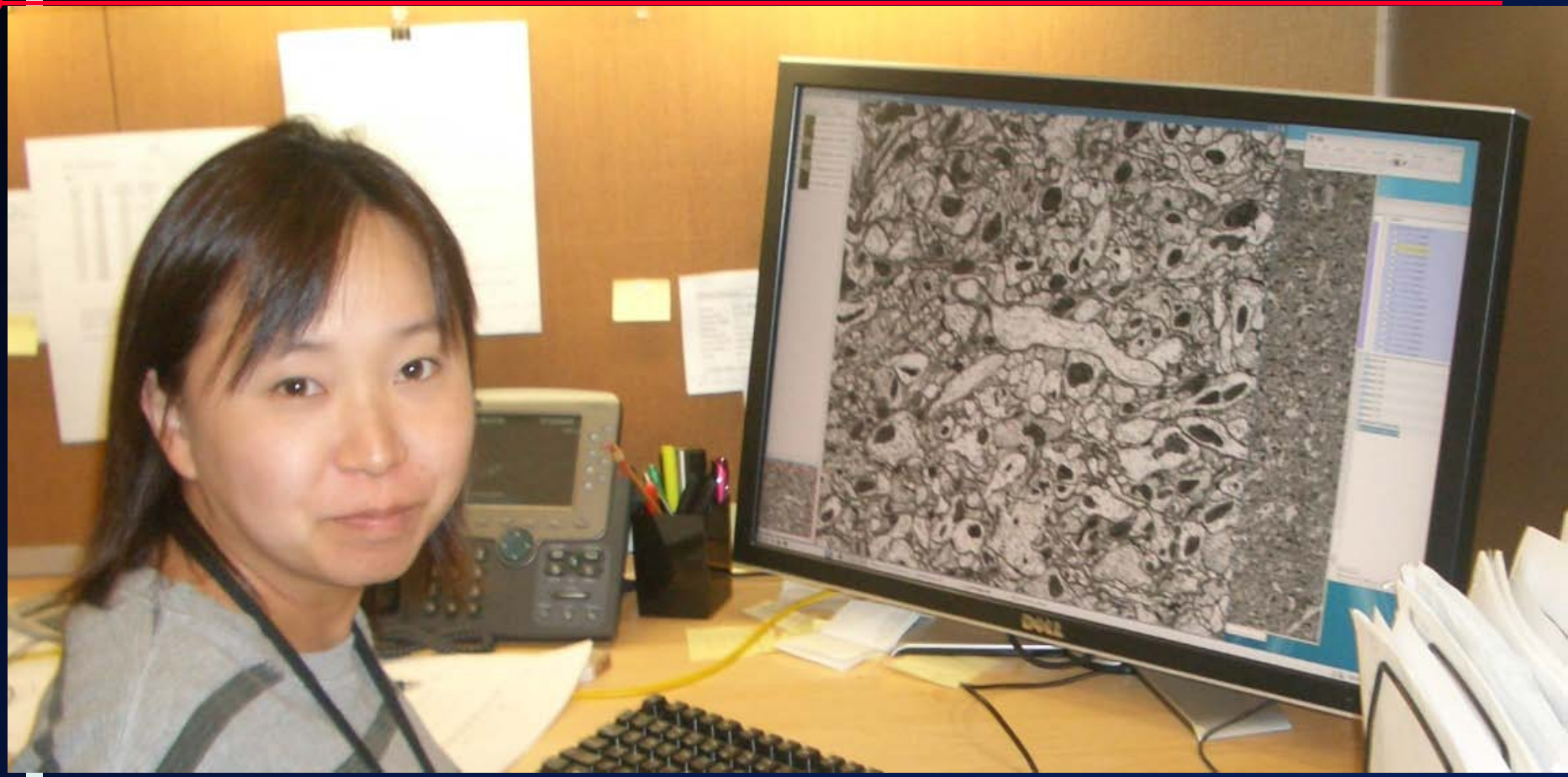
Other projects

- Generalize registration techniques
- Use parallels with EE design to try to understand networks of neurons
- Better probe design
- Digital electronics for other projects

Example RE Company

- Picked this one since there are good descriptions on their web site:
  - “It all starts with high quality reverse engineering that effectively decaps and delayers devices cleanly to enable automated SEM mosiacs of 1000's of high magnification images to be taken and stitched together to calibrated standards. “
- Very similar to the first steps in biology

This is where we are today in brain RE



- Still tracing wiring by hand!

Chipworks – Next level of tools

- Provide software tools to help make sense of the images (ICsurveyor) – from their web site:
  - “Seamlessly navigate massive high magnification SEM-based image mosaics of your competitors’ chips”
  - “Accurately measure key features, using the ruler tool, for use in simulation, costing, and yield estimating”
  - “View multiple metal layers at one time when tracing signals”
  - “Annotate blocks for sharing information within the team and across the organization”

Viewing multiple superimposed layers

From “Chip Detectives”

Viewing multiple registered layers

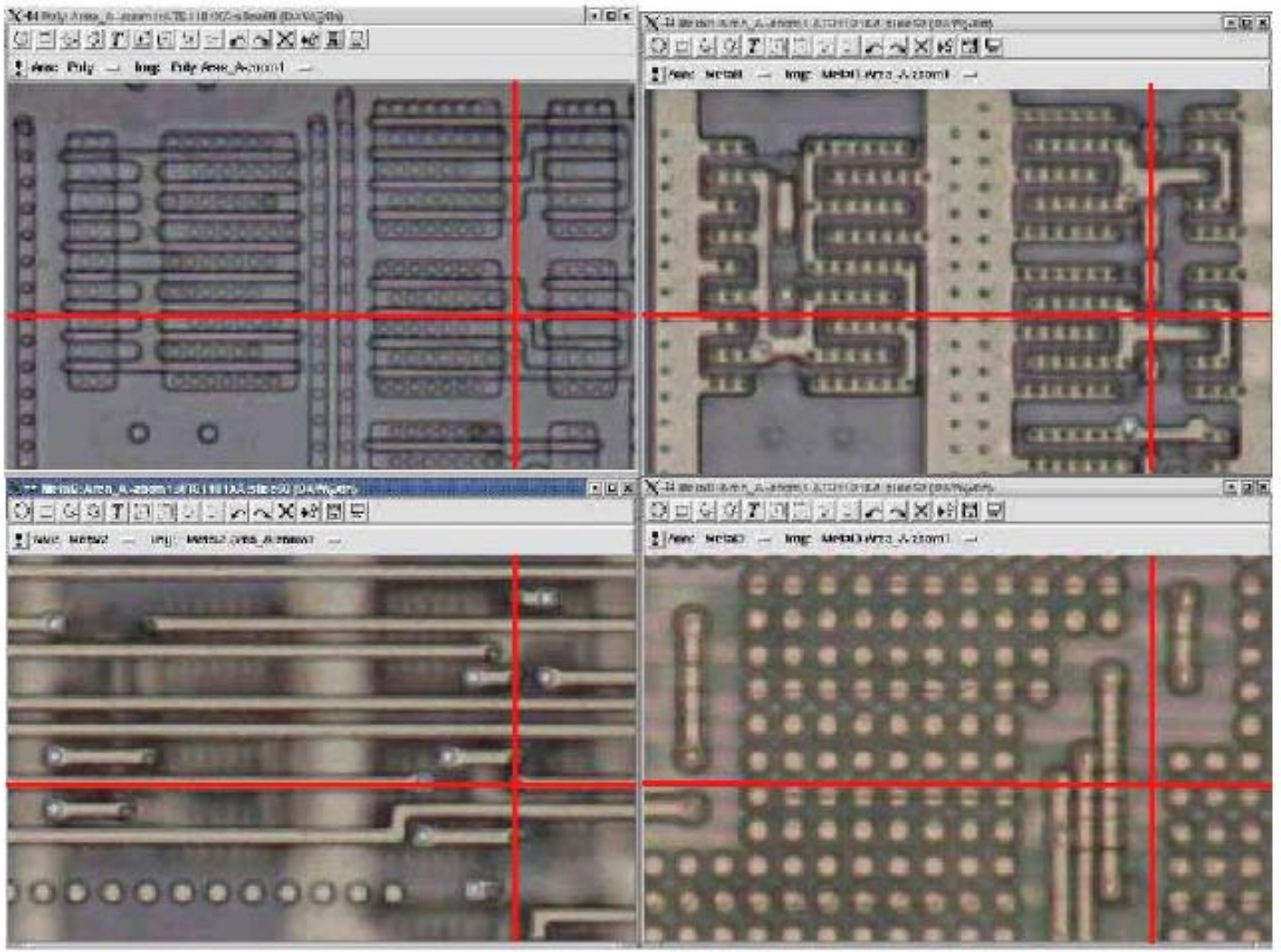


Fig. 9. Simultaneous windows showing images of three metal layers and polysilicon layer

"Reverse Engineering in the Semiconductor Industry", Torrance and James

Analysis of extracted circuits - Underlined parts biologists would like to have

"Report Contents:

- Package and die overview including a package x-ray.
- Annotated die photograph **identifying the major functional blocks**.
- Complete set of **hierarchical schematics** to capture the operation of the device.
- Top-level schematic overview to capture the entire function of the circuit.
- **Schematics for each analyzed block**.

- Chipworks' schematics include **individual device size measurements**, **signal descriptions**, **cross references**, and a **signal reference list** that summarizes all signal sources and destinations."

Chipworks: Software tools available to help

- “The Chipworks' ICInside Browser is a key part of a Chipworks’ circuit analysis deliverable that enables you to view, analyze and interact with device schematics and associated layout images. It offers **bi-directional cross probing capability** between the two views so that once you identify a specific area of interest, you can **follow a signal path**, drill down, **toggle views** and **discover exactly what’s going on in terms of design, structure and functionality**.”

Automated tracing through multiple layers

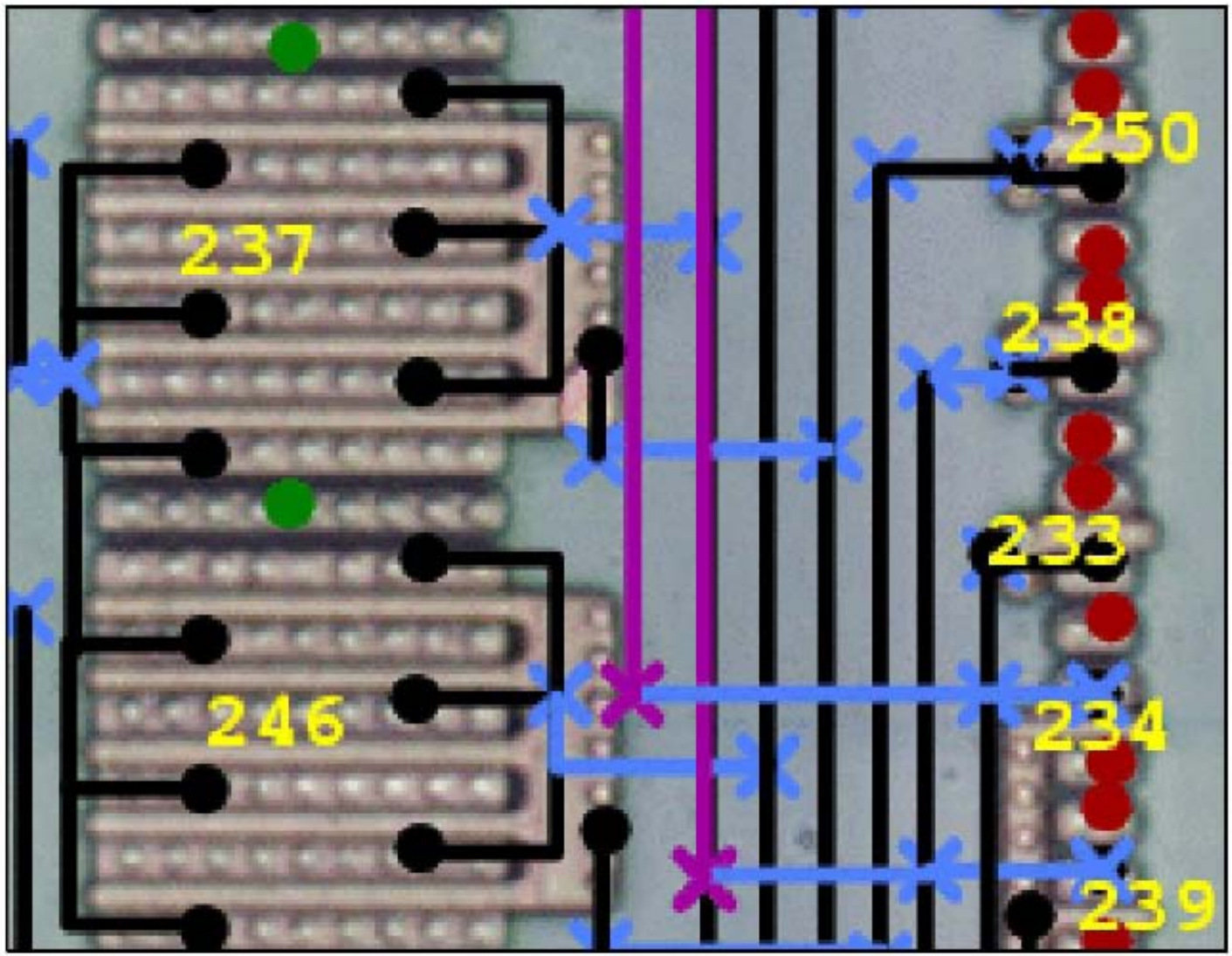


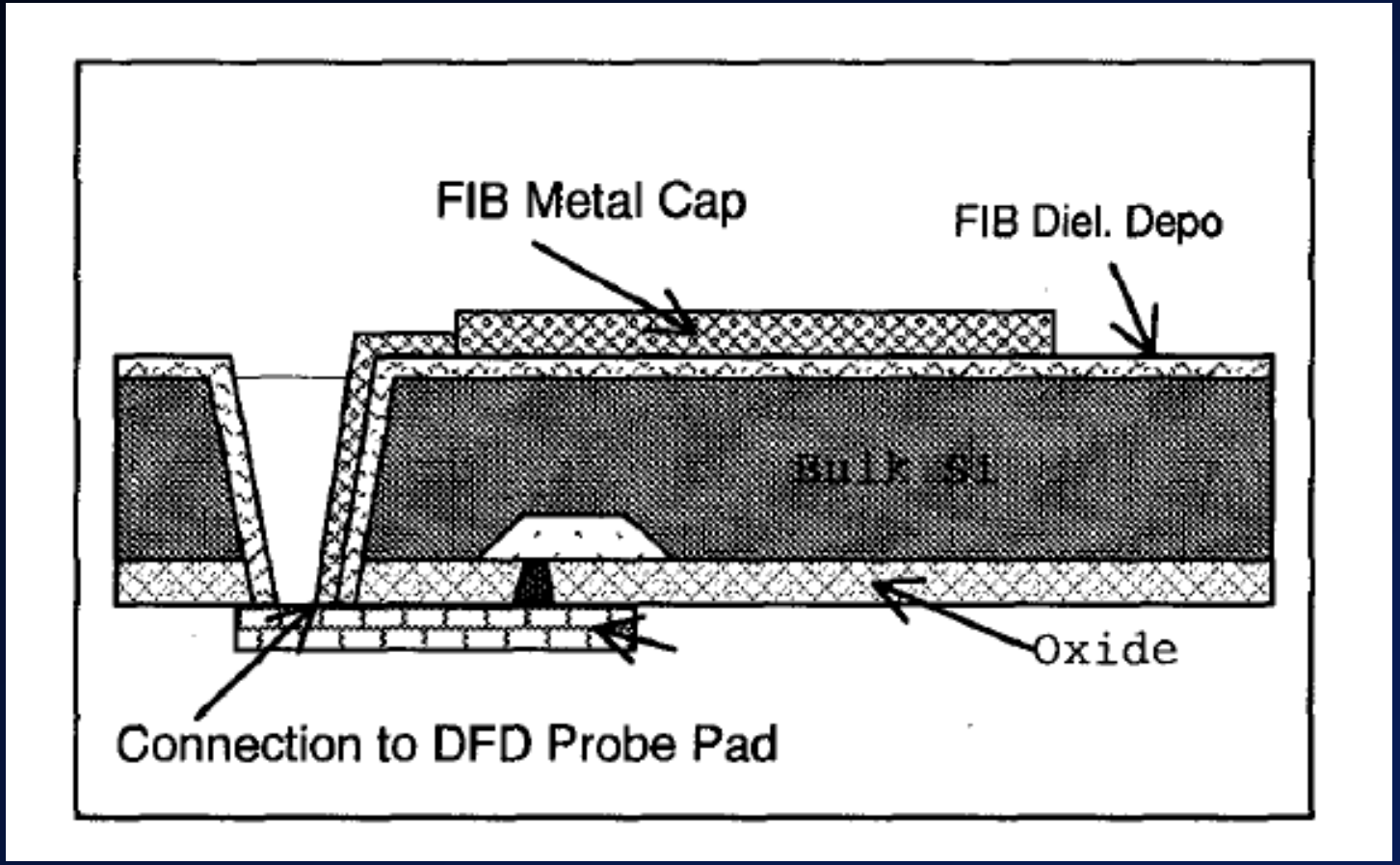
Fig. 10. Image annotated using DAW workstation

"Reverse Engineering in the Semiconductor Industry", Torrance and James

Other ways of probing signals

- Atomic force microscopy
- E-beam (watch energy of scattered electrons)
- Watching the back of the wafer (hot electrons recombine and generate IR)
- Drilling a hole and attaching to a signal

FIB = Focused Ion Beam



DESIGN FOR (PHYSICAL) DEBUG FOR SILICON MICROSURGERY AND PROBING OF FLIP-CHIP PACKAGED INTEGRATED CIRCUITS

Richard H. Livengood,
Donna Medeiros
Intel Corporation, Santa Clara, CA.

Can also modify chips, measure changes

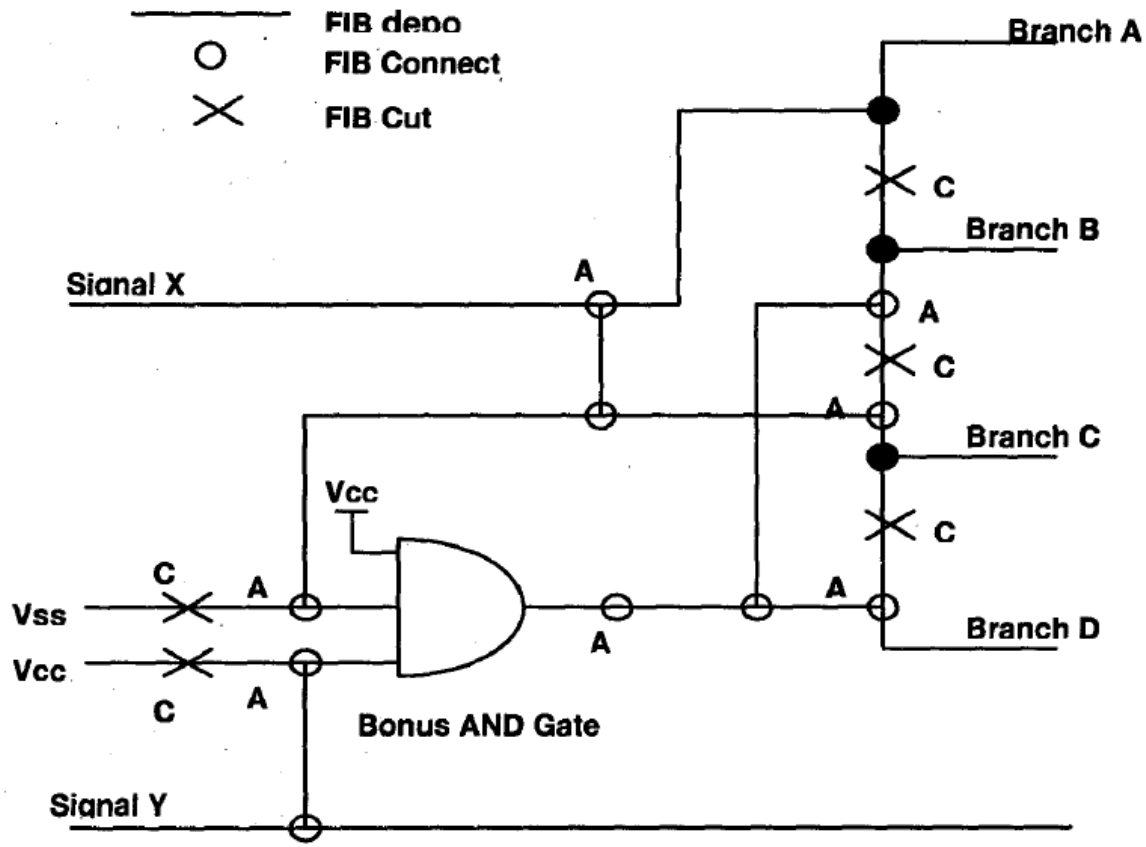


Figure 7: Schematic of complex logic change using spare logic gate with designed in FIB connection and cut constructs.

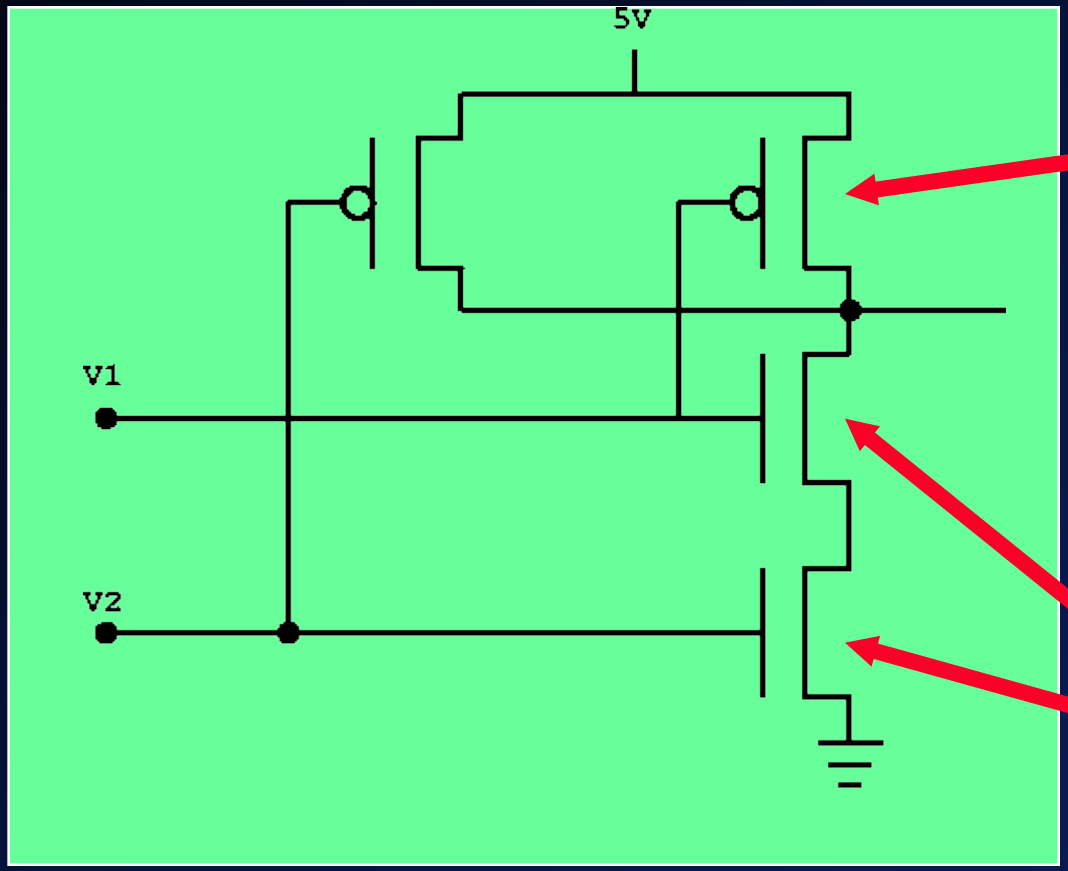
- Cut wires and add new wires

Design for (physical) debug for silicon microsurgery and probing of flip-chip packaged integrated circuits, Livengood R.H. and Medeiros, D.

Examining features that cannot be seen

- Two options
  - Deduce them from what can be seen
  - Make them visible in some way
- Examples:
  - Transistor types in EE – NMOS or PMOS. These differ by doping levels, which cannot be seen in an EM image
  - Synapse type in biology – inhibit or excite. Cannot be seen in a EM image

Deduce transistor type from the connections



These transistors must be P-type (High input level inhibits function)

These transistors must be N-type (High input level enhances function)

Option 2 – make it visible

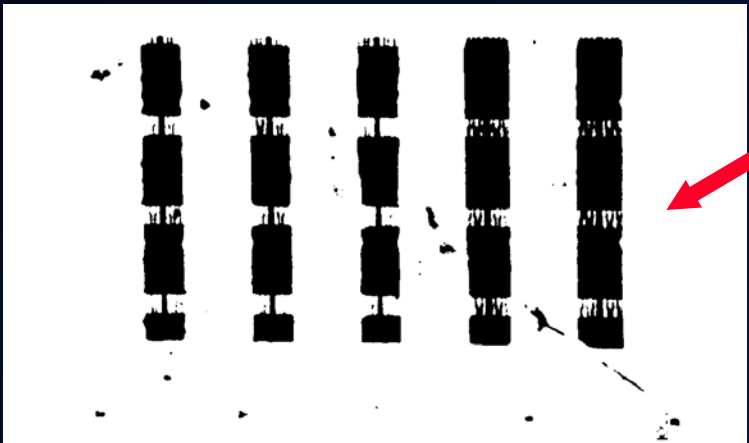


Fig. 4. Schottky barrier EBIC image of CMOS chip using Pd metallization.

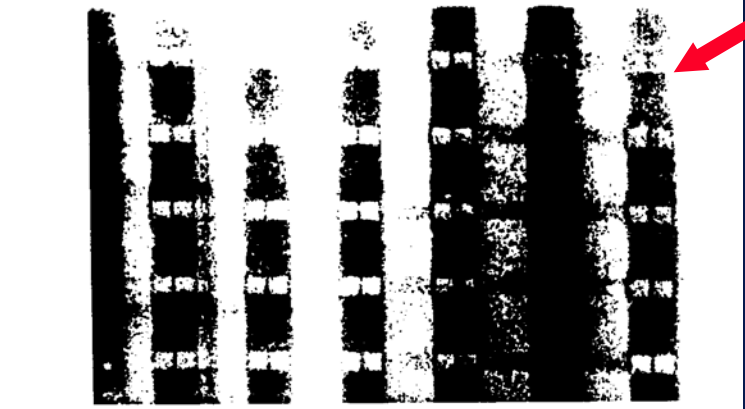


Fig. 5. Schottky barrier EBIC image of CMOS chip using Al metallization.

Coat substrate with a metal of one work function, take a picture

Strip first metal, coat substrate with a metal of another work function, take another picture

Layout Reconstruction of Complex Silicon Chips, by Simon Blythe, Beatrice Fraboni, Sanjay Lall, Haroon Ahmed, and Ugo de Riu

Why are the EE and biology problems so different in practice?

- The two problems are roughly the same scale
- So why is one an industrial process, common enough to require defensive measures,
- And the other is a research project?

Why is EE reverse engineering easy?

- All instances of a chip are identical
  - Especially helpful for probing vs reconstruction
- Few and well defined layers (< 20)
- Constructed of parallel lines and simple geometries
- Small underlying library
- Designed to be regular
- Substrate is stable and strong
- Layers are well registered

Same list as a to-do list for brain RE

- All instances of a chip are identical
  - Learn to get all slices from one brain
- Few and well defined layers (< 20)
  - Must handle 6000 layers
- Constructed of parallel lines and simple geometry
  - Structures must be biologically plausible
- Small underlying library
  - Discover the underlying library of the brain
- Designed to be regular
  - Find the regularities
- Substrate is stable and strong; layers are well registered
  - Perform registration by features

What steps do EEs perform that are missing in brain RE?

- Automated multi-layer path tracing

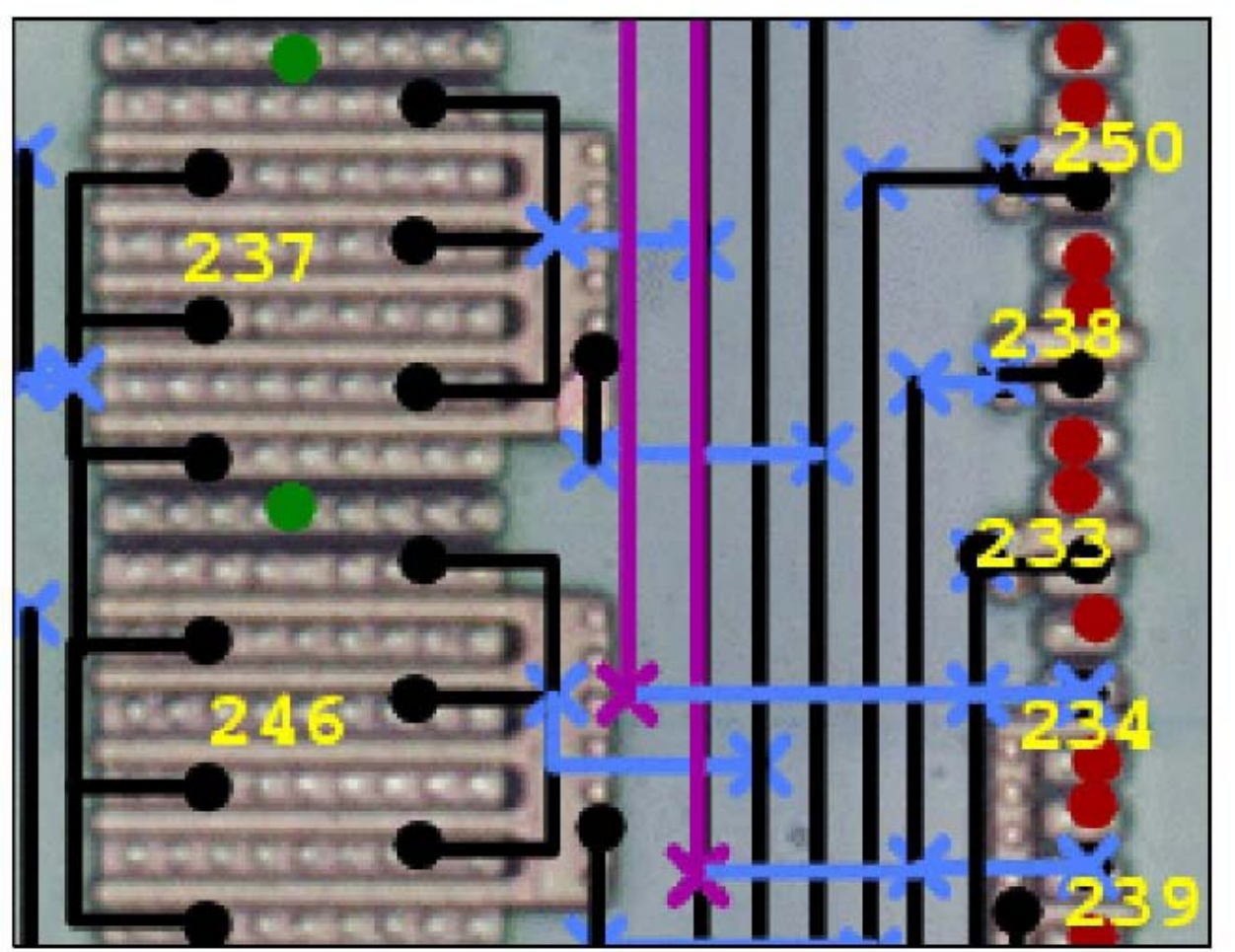


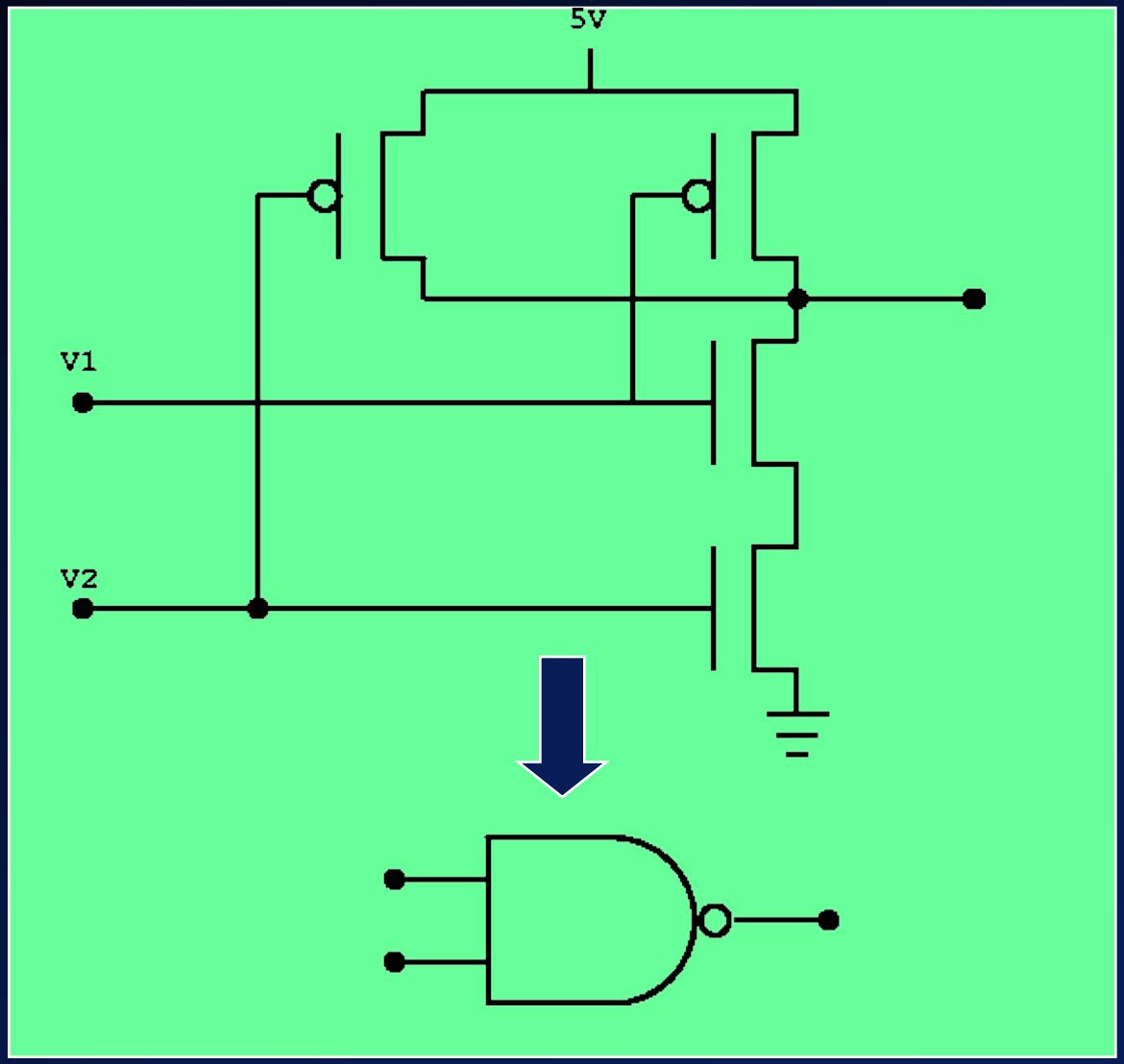
Fig. 10. Image annotated using DAW workstation

"Reverse Engineering in the Semiconductor Industry", Torrance and James

Also missing – equivalent of transistors to gates

- Basic chip RE gives a transistor netlist
- Gate level netlist is much more useful
  - Easier to understand
  - Much smaller (factor of at least 4)
  - Simulation much faster
  - More tools available
    - Automatic test pattern generation
    - Fault simulation
    - Static timing

Convert transistors to gates



Many possible methods – here are four:

- Rule based reduction
  - User specifies a set of rules; can add more later
- Boolean algebra
  - Assume each input is 0 or 1, find output
  - Substitute gates that do the same thing (maybe different structure)
- 4 valued analysis
  - Add X (unknown) and DC (Don’t care)
- Sub-graph isomorphism
  - Pattern matching

What if you have only the structure

- Images can only show parts and how they are connected
- Flow of information does not show up in pictures
  - Some can be deduced from network connections
  - Optic nerve carries information from the eye to the visual system (we think…)
- Operation of components may not be known
  - For example, synapse type (inhibit or excite)

Finding directions of information flow

- Based on the assumption that circuit is doing something useful.
- Information flow of some parts is uni-directional, even if type is unknown
  - Transistor in EE: Gate affects the channel, though polarity may be unknown
  - Synapse in biology: For a chemical synapse, pre-synaptic neuron affects post-synaptic, but not the other way
- Use known flows to find flows in other parts of the graph.

What if the gate function is unknown?

- Can determine function of circuit without knowing the functions of the parts
- Clues to overall function
  - Components with known patterns
  - Repeated sub-structures
  - Global structure
- “Another boost to the reverse-engineering process is the tendency of designers to follow published or textbook designs.”
  - Biology equivalent is using related organisms

Example: The ISCAS 1985 benchmarks

- Industrial examples for physical design problems
- Took real circuits, and removed all functions
- Left with 1 input boxes, 2 input boxes, etc.
  - 1 input can be one of 2 types, buffer or inverter
  - 2 input gate can be 16 possible types, including AND, OR, NAND, NOR, XOR, XNOR
  - 3 input gate can be one of 256 gate types, including AND, OR, NAND, NOR, MUX, FF, etc,
  - 4 inputs one of 65536 types, and so on

A Neutral Netlist of 10 Combinational Benchmark Circuits and a Target Translator into FORTRAN, F Brglez, H Fujiwara . IEEE International Symp. on Circuits and Systems (ISCAS), Kyoto, …, 1985

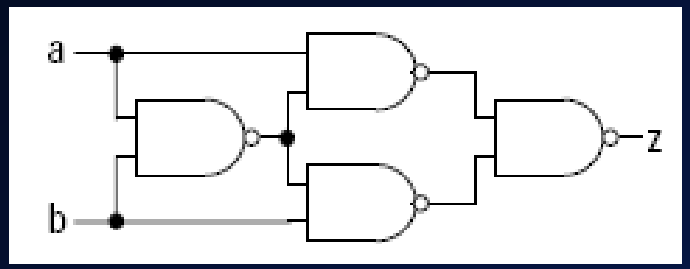
But people determined the exact function anyway!

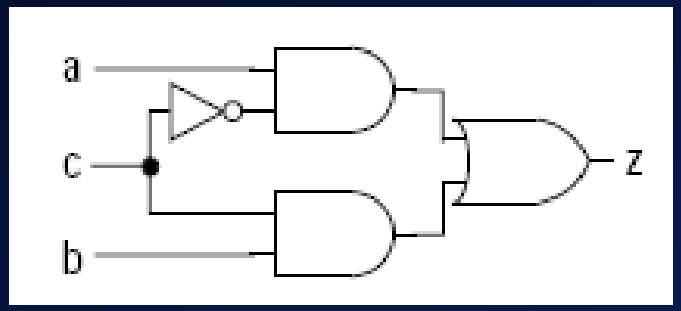
- SEC = Single Error Correct, etc.
- Much easier to understand, and much smaller

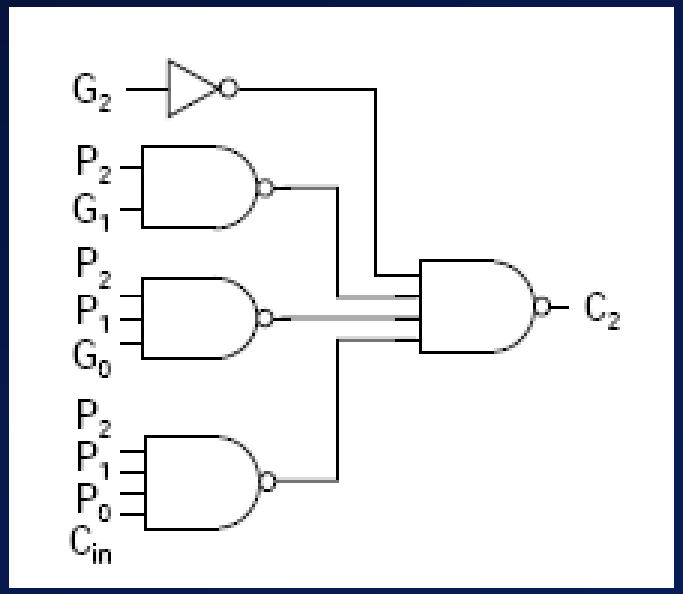
Circuit	Function	No. of logic gates	No. of major functional blocks
c432	27-channel interrupt controller	160	5
c499	32-bit SEC circuit	202	2
c880	8-bit ALU	383	7
c1355	32-bit SEC circuit	546	2
c1908	16-bit SEC/DED circuit	880	6
c2670	12-bit ALU and controller	1,193	7
c3540	8-bit ALU	1,669	11
c5315	9-bit ALU	2,307	10
c6288	16 × 16 multiplier	2,406	240
c7552	32-bit adder/comparator	3,512	8

How does this work?

- There are certain economical ways to build gates
- These are unique even without gate information







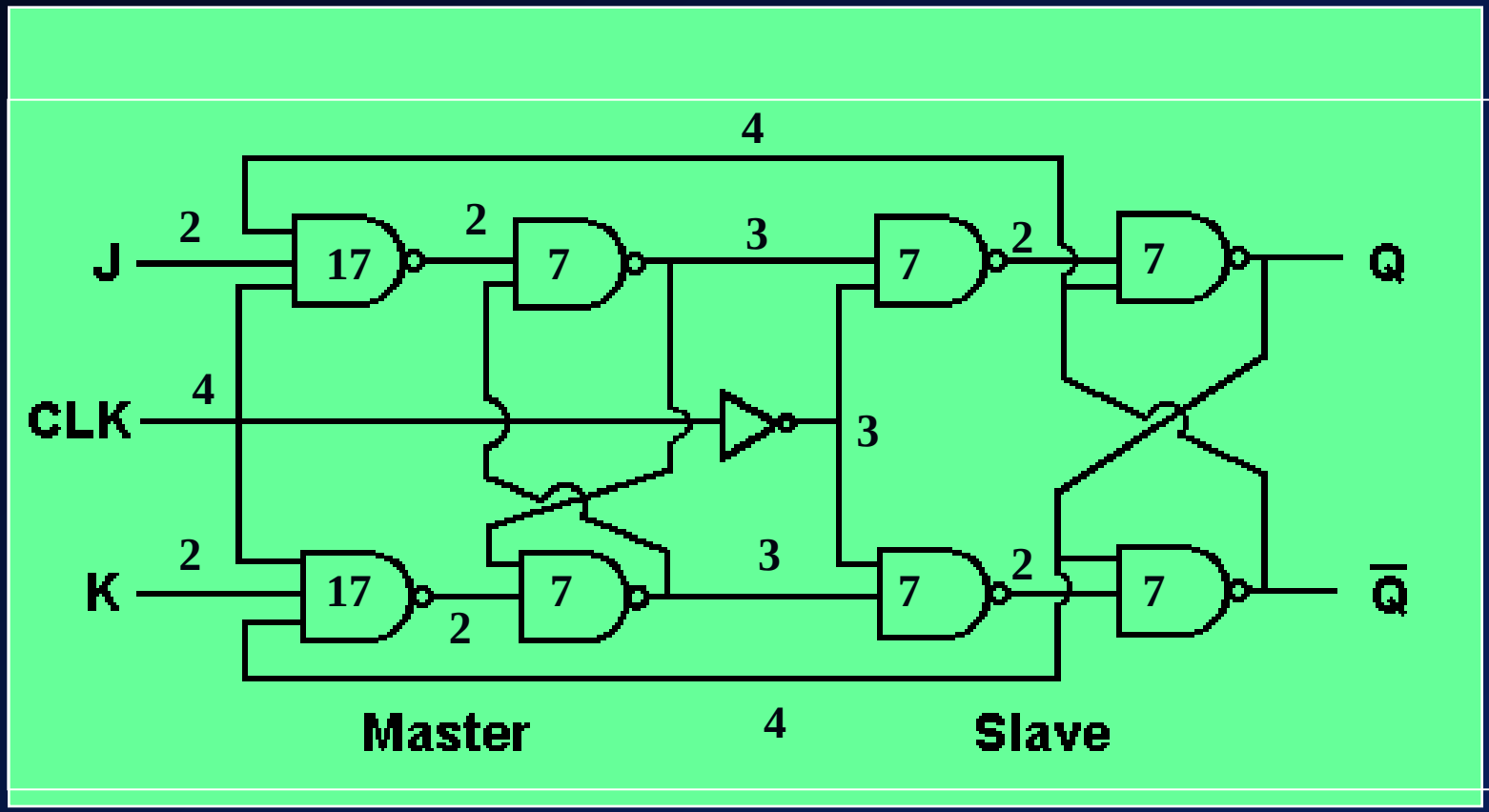
Unveiling the ISCAS-85 benchmarks: a case study in reverse engineering. Hansen, M.C.; Yalcin, H.; Hayes, J.P.; Design & Test of Computers, IEEE Volume 16, Issue 3, July-Sept. 1999 Page(s):72 - 80

Then there are ways to find these in graphs

- Sub-graph isomorphism problem
- NP-complete problem, but lots of heuristic solutions in EE and computer science
  - SubGemini: identifying subcircuits using a fast subgraph isomorphism algorithm, Ohlrich, M. and Ebeling, C. and Ginting, E. and Sather, L., Proceedings of the 30th international conference on Design automation, pages=31--37, year=1993

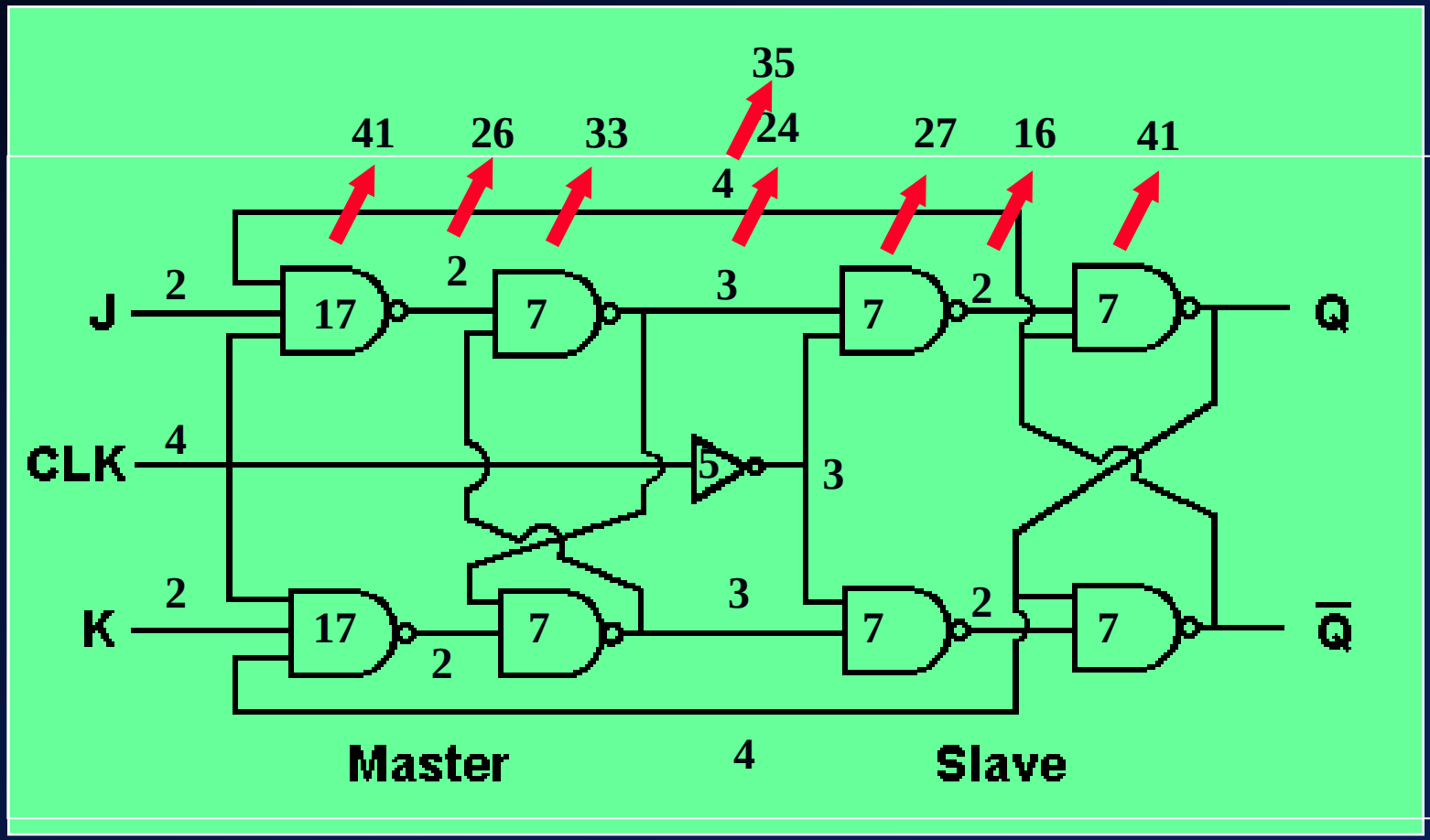
How does this work?

Label each part by type, each net by # of connections



How does this work?

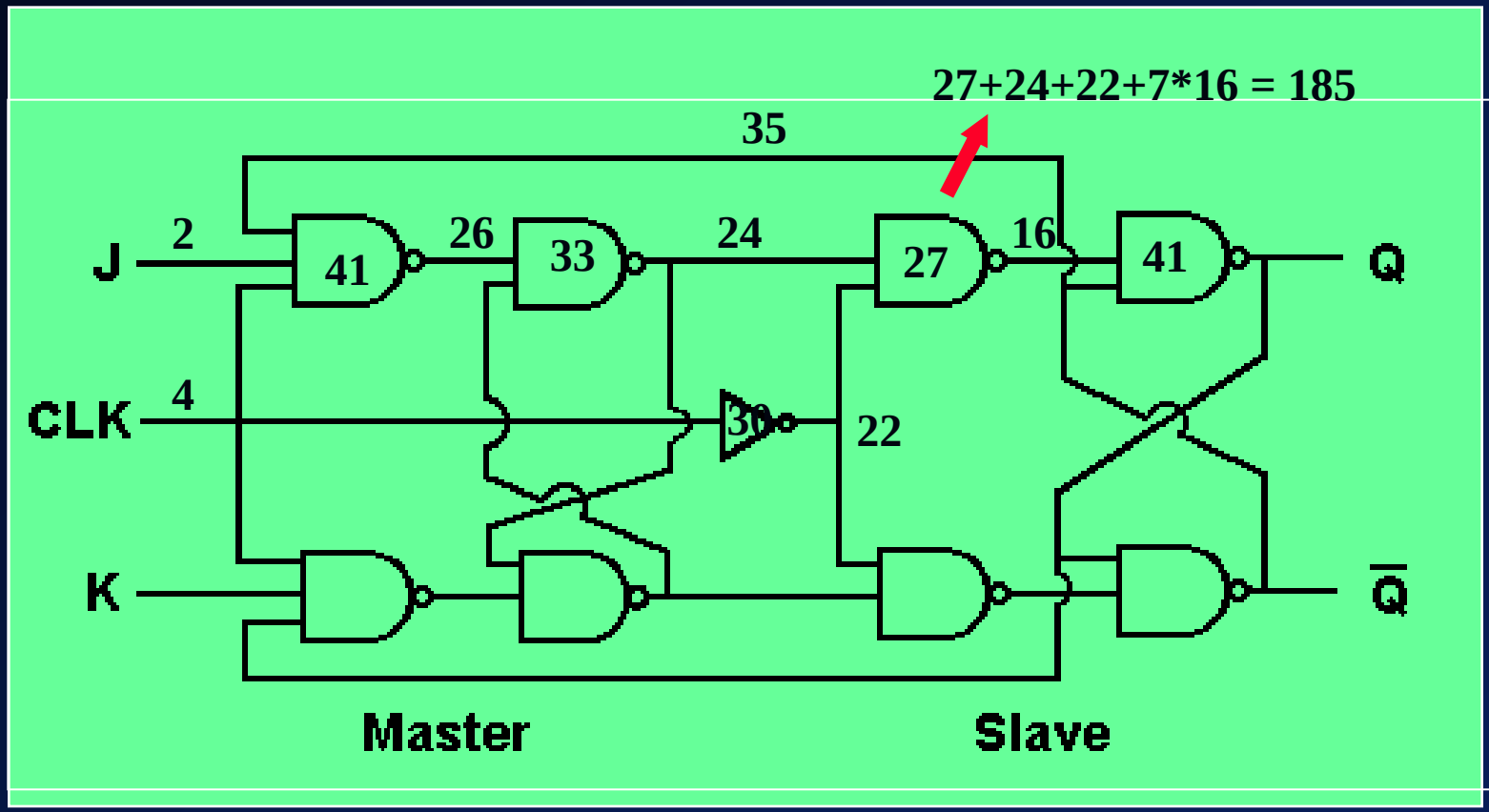
Now, nets add connected parts; parts add sum of inputs + 7*output (example only)



How does this work?

Repeat as many times as needed (graph radius of subcircuit)

Only gates with identical numbers are potential matches



Variations on the theme of sub-graphs

- Extensions that can cope with errors
  - A new algorithm for error-tolerant subgraph isomorphism detection, Messmer, BT and Bunke, H., IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 20, issue 5, pp. 493—504, 1998.
- Discover sub-graphs that occur often
  - Frequent subgraph discovery, Kuramochi, M. and Karypis, G., Proceedings of the 2001 IEEE International Conference on Data Mining, pp. 313—320.
- Biologists would like a combo (near identical graphs that occur often)

Once you have the connections, what can you do with them?

- EE has produced a large number of algorithms to analyze netlists
  - Analog simulation (level of differential equations)
    - Flat, hierarchical, mixed with digital, etc.
  - Analog response modeling and macromodels
    - Faster, simpler models when the variable of interest is firing rate, average voltage, etc. and not the detailed waveform
  - Digital simulation
    - Emulated, compiled code, hardware assist
  - Digital analysis
    - Observability, controllability, formal analysis, fault response, etc.

Finding directions of signal flow

- D. T. Blaauw, D. G. Saab, J. Long, and J. A. Abraham, “Derivation of signal flow for switch-Level simulation,” *EDAC*, 1990, 301-305
- Derivation of Signal Flow Direction in MOS VLSI, Jouppi, N. This paper appears in: **Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on** Publication Date: May 1987 Volume: 6, Issue: 3 On page(s): 480- 490
- An integrated system for assigning signal flow directions to CMOStransistors
- Kuen-Jong Lee Chih-Nan Wang Gupta, R. Breuer, M.A. Dept. of Electr. Eng., Nat. Cheng Kung Univ., Tainan;
- This paper appears in: **Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on**
  Publication Date: Dec 1995
  Volume: 14, Issue: 12
  On page(s): 1445-1458

A few more methods

- [**1]** G. Ditlow, **W.** Donath and A. Ruehli, “Logic equations for MOSFET circuits”, IEEE International Symposium on Circuits and Systems, pp. **752-755,** May **1983**
- **[2] Z.** Barzilai, L. Huisman, G. M. Silberman, D. T. Tang and L. ***S.*** Woo, “Simulating pass transistor circuits using logic simulation machines”, Design Automation Conference, pp. **157-163,** June **1983**
- **[3]** R. E. Bryant, “Boolean analysis of MOS circuits”, IEEE Transactions in Computer Aided Design, vol. **6,** pp. **634-649,** July **1987**
- **[4]** D. T. Blaauw, D. G. Saab, P. Banerjee and J. Abraham, “Functional abstraction of logic gates for switch level simulation”, European Conference on Design Automation, pp. **329-333,** Feb-

Example papers on this process

- **[5]** R. E. Bryant, “Extraction of gate level models from transistor circuits by four valued symbolic analysis”, International Conference in Computer-Aided Design, pp. **350-353,** November **1991**
- **[6]** R. E. Bryant, D. Beatty and **K.** Brace, “COSMOS: A compiled code simulator for MOS circuits”, Design Automation Conference, pp. **9-16,1987**
- **[7]** A. Kuehlmann, D.I.Cheng, A. Srinivasan and **D.** P. Lapotin, “Error diagnosis for transistor level verification”, Design Automation Conference, pp. **218-224,** June **1994**
- **[8] GateMaker: a transistor to gate level model extractor for simulation, automatic test pattern generation and verification** Kundu, S. Test Conference, 1998. Proceedings., International Date: 18-23 Oct 1998, Pages: 372 - 381

More references

- Layout Reconstruction of Complex Silicon Chips, by Simon Blythe, Beatrice Fraboni, Sanjay Lall, Haroon Ahmed, and Ugo de Riu

Five reasons for detailed connections

- Theory
- Example from Computer Science
- Cost and consumer electronics
- Examples from defense
- Industrial reverse engineering of chips

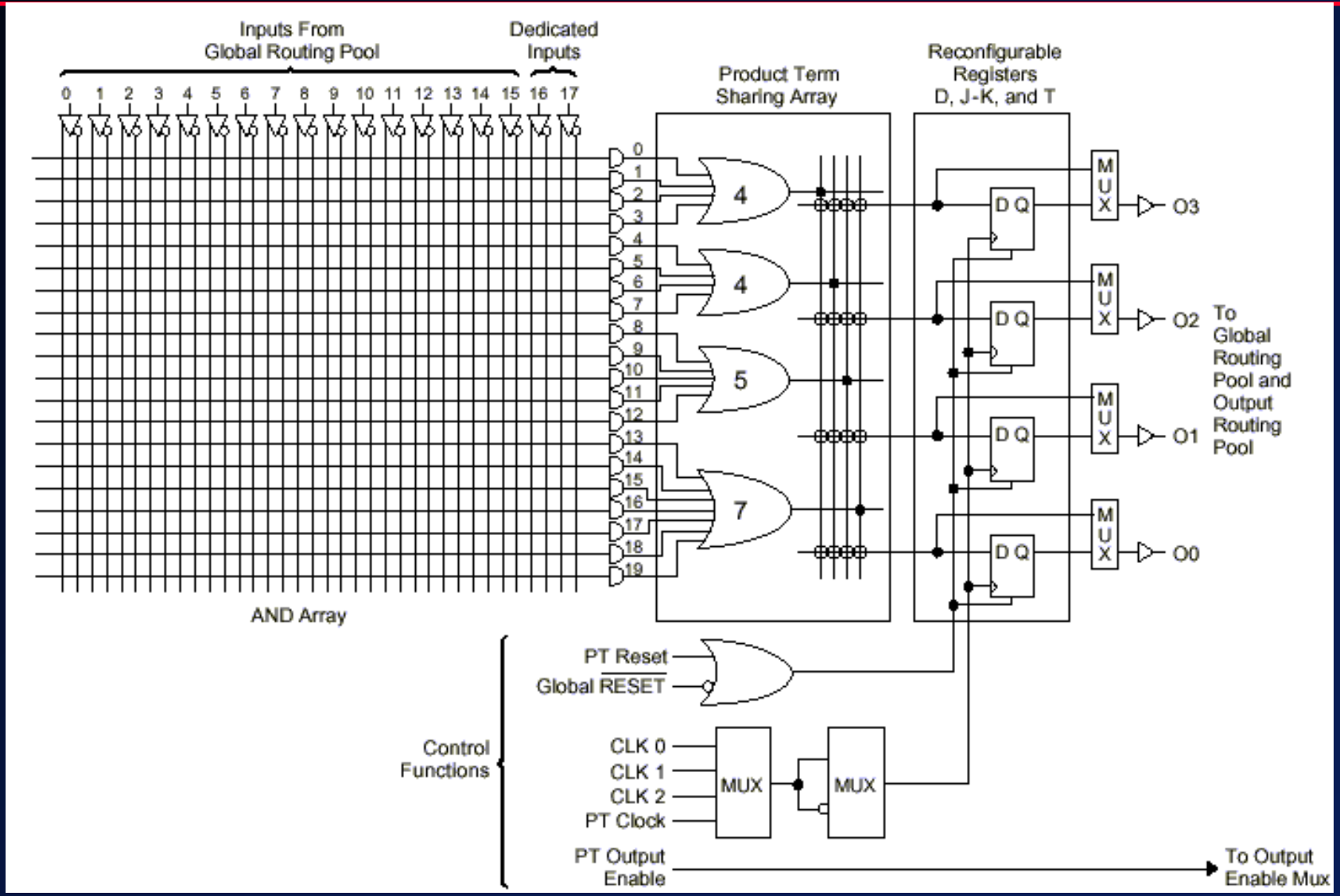
Software

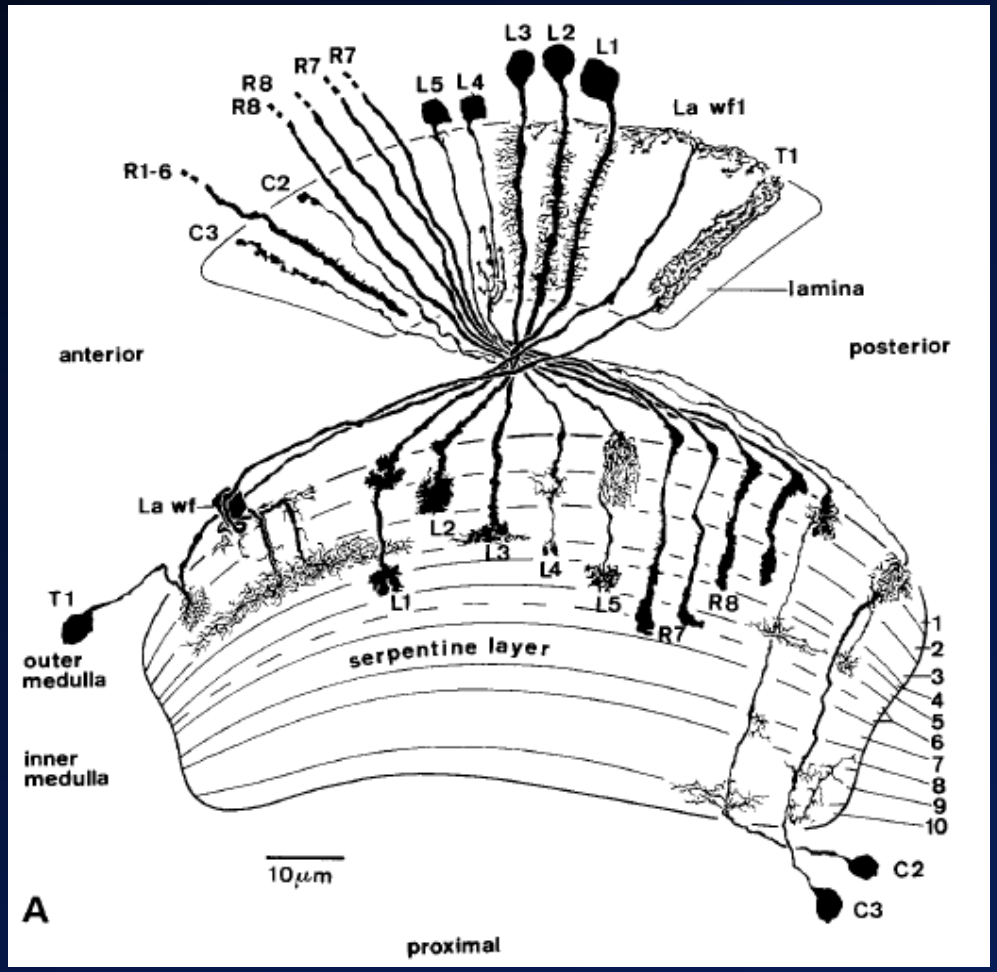
- Take your knowledge of the program
- Make sure it works on a number of cases
- Infer it will work on cases you have not tried

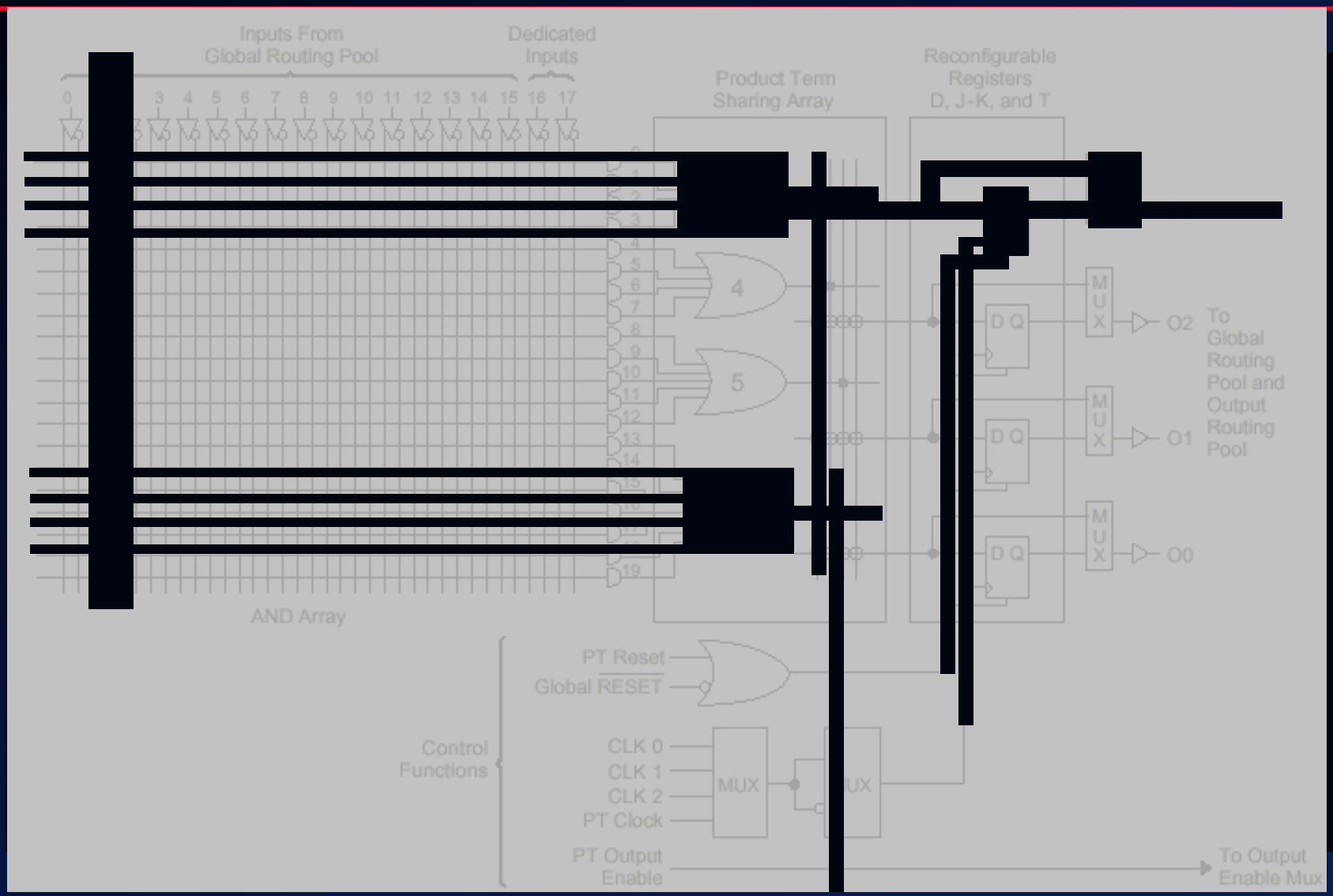
Question comes up in EE quite often

- How do you know a given chip does what you want it to?
  - For chips in production, present some finite number of patterns, want to know others will be OK
- Conversely, how can you show it's not doing anything else?
  - For a military chip, check to see if it's doing the right thing with some finite number of patterns, want to make sure it won't do anything else
- Both require detailed connections

Here's a PLD (each user can customize)




L3 L2
L1
R7 R7
R8
R8
L5 L4
La wf1
T1
R1-6
C2
C3
lamina
anterior
posterior
La wf
L2
L3
L4
T1
L1
L5
R8
R7
serpentine layer
outer medulla
inner medulla
1
2
3
4
5
6
7
8
9
10
10μm
C2
C3
A
proximal


Inputs From
Global Routing Pool
Dedicated
Inputs
0 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Product Term
Sharing Array
Reconfigurable
Registers
D, J-K, and T
4
5
M
U
X
D Q
O2
O1
O0
To
Global
Routing
Pool and
Output
Routing
Pool
AND Array
PT Reset
Global RESET
Control
Functions
CLK 0
CLK 1
CLK 2
PT Clock
MUX
PT Output
Enable
To Output
Enable Mux

Can you make any inference about function?

- No, this configuration is picked because it is ‘universal’
- Any logical function can be implemented by changing only the ‘synapses’

Motivation

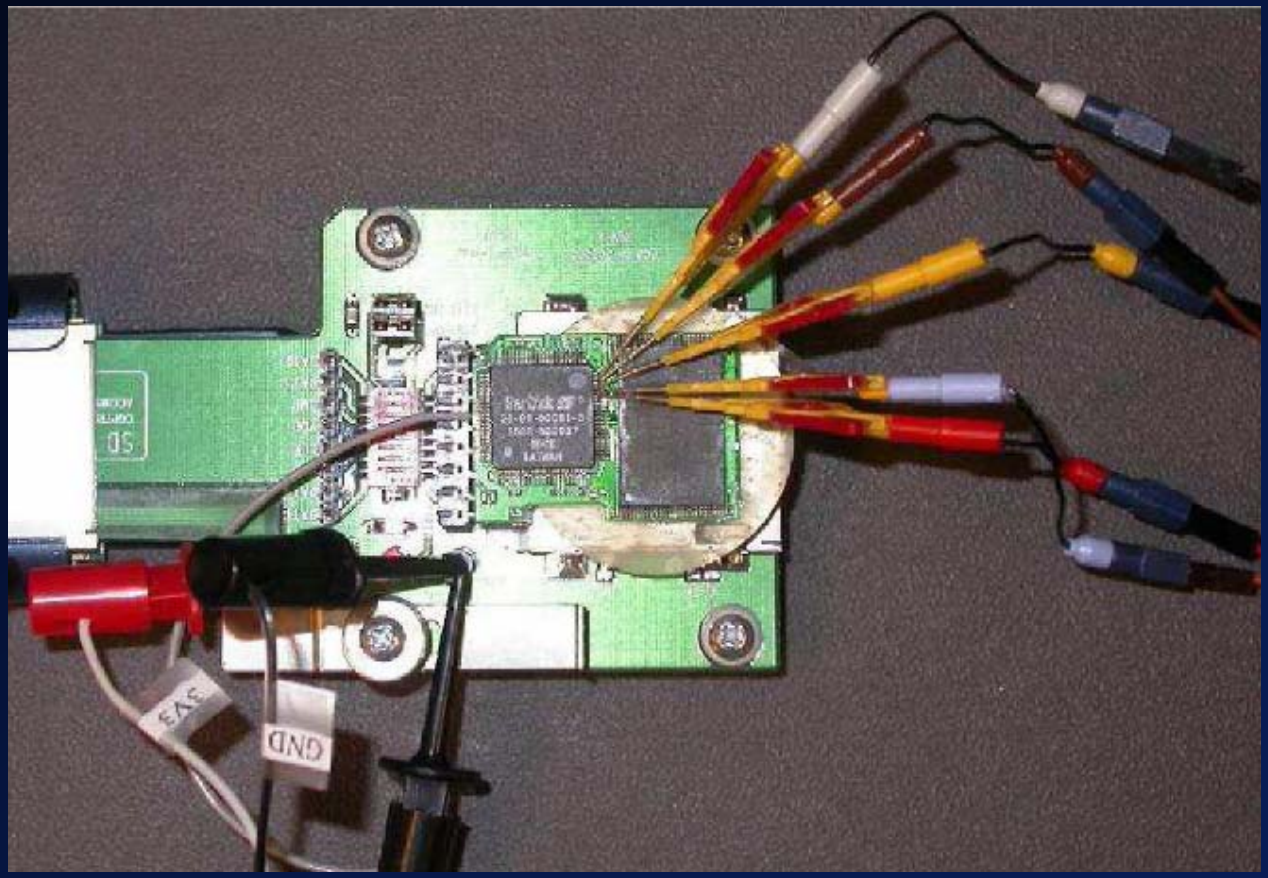
- Do we really need the detailed connections?
- Cannot we infer how the brain works from some easier/cheaper/faster/higher-level technique?
- Will results really aid in understanding?
- Engineering view: We understand a system when we can predict the results for experiments that have not been tried

Three other lines of reasoning

- Even consumer electronics adds extra costs to make sure each ‘synapse’ is working
- Defense department does a lot of work to make sure ‘synapses’ are what the expect
- When reverse engineering chips, engineers could deduce function operation etc., from probing, examining IOs, etc.
  - But they don’t – they slice it up and get the circuits

EEs can experiment with behavior

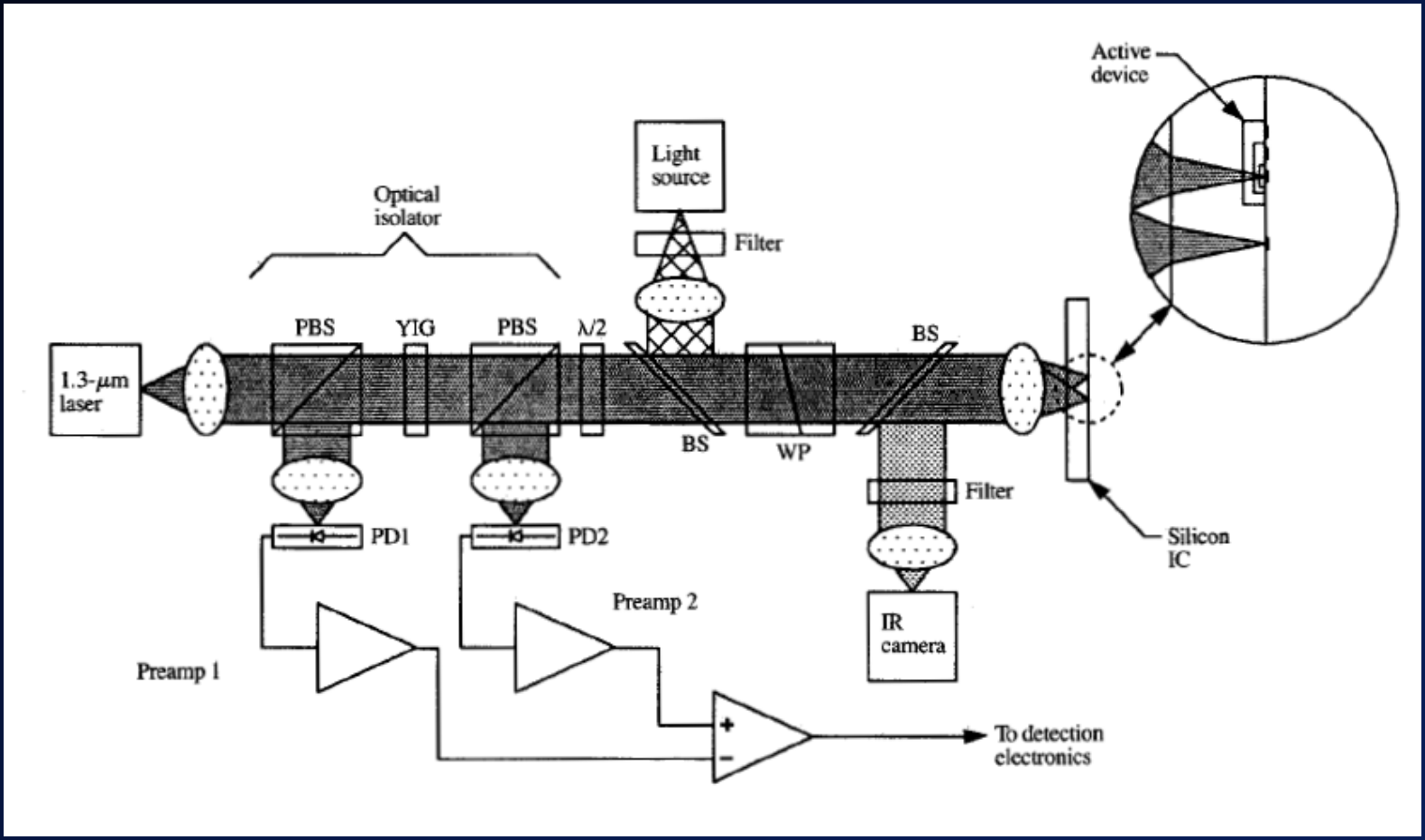
- External pins are easy to probe



"Reverse Engineering in the Semiconductor Industry", Torrance and James

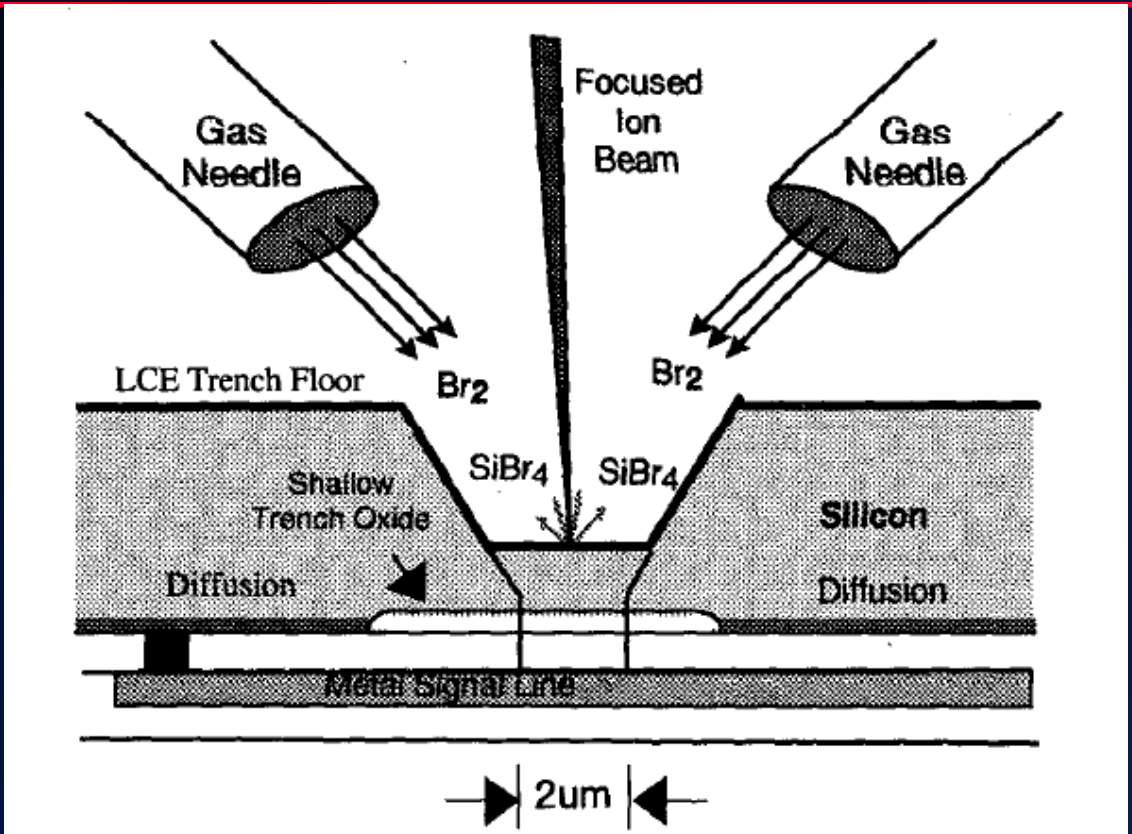
Internal signals are harder, but possible

- Using change in optical properties w/voltage



Picosecond Noninvasive optical detection of internal Electrical signals in flip-chip-mounted silicon integrated circuits, by H. K. Heinrich IBM J. of R&D

Equivalent of knock-out and knock-in



Step 3: Small sub 10um hole FIB milled through silicon down to signal node of interest.

Design for (physical) debug for silicon microsurgery and probing of flip-chip packaged integrated circuits, Livengood, R.H. and Medeiros, D.

But they end up doing reconstruction

From “Chip Detectives”

Janelia Farm, Howard Hughes Medical Institute