UniversitätsSpital Zürich

Limitations of current experimental trends in biological research

Technical Journal Club

09.05.2017

Daniel Kirschenbaum

A fundamental question

- Research is moving towards large scale data.
- The *omes (here connectome), large-scale activity recordings in the brain, simulations.
  - (Although we are still limited in acquiring organism-wide complexity and in spatio-temporal resolution.)

- **Will only scaling up on data help to understand how biological systems work?**

Will scaling up help to understand how biological systems work?

A related question:

What would provide some sort of good positive control for our methods as a whole?

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Could a Neuroscientist Understand a Microprocessor?

Eric Jonas[1]*, Konrad Paul Kording[2,3]

They apply the standard neuroscience toolbox addressing:

- Connections,
- Lesions of individual components,
- Single-unit tuning curves,
- Joint statistics,
- Local activities,
- Estimated connections,
- Whole system recordings,

…to test if they can understand how a microprocessor works.

What does it mean to understand a system?

(Probably a lot to read about this in epistemiological works.)

Suggestions?

Two practical approaches:

1. ***If we can fix a broken implementation of a system.***
   Essentially, when the inputs, transformations and outputs can be replaced artifically.

2. ***David Marr's three complementary hierarchical levels.***

Almost never can a complex system of any kind be understood as a simple extrapolation from the properties of its elementary components. Consider, for example, some gas in a bottle. A description of thermodynamic effects—temperature, pressure, density, and the relationships among these factors—is not formulated by using a large set of equations, one for each of the particles involved. Such effects are described at their own level, that of an enormous collection of particles; the effort is to show that in principle the microscopic and macroscopic descriptions are consistent with one another. If one hopes to achieve a full understanding of a system as complicated as a nervous system, a developing embryo, a set of metabolic pathways, a bottle of gas, or even a large computer program, then one must be prepared to contemplate different kinds of explanation at different levels of description that are linked, at least in principle, into a cohesive whole,

David Marr's three complementary hierarchical levels.

Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

Figure 1–4. The three levels at which any machine carrying out an information-processing task must be understood.

Vision: generating a 3D model of the world which the individual can operate with.

Representation: E.g. arabic, digital and roman numerical system.

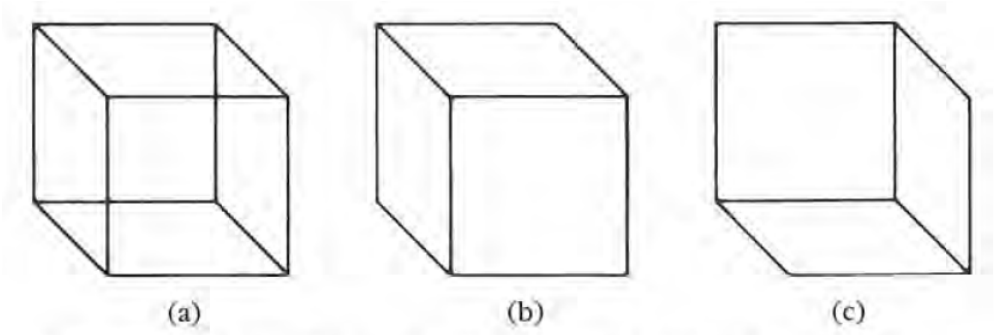
Algorithm: Addition

Transistors, neurons, synapses, ion channels.

David Marr's three complementary hierarchical levels.

The levels can sometimes be confused when trying to explain something:

- Seeing after images when staring into a light bulb or the fact that all colors are achieved by some mixture of three primary colors.→ Clearly hardware level phenomena.
- Necker's illusion:



Some bistable neuronal network in the brain. → Representational or algorithmic level.

PLOS | COMPUTATIONAL BIOLOGY

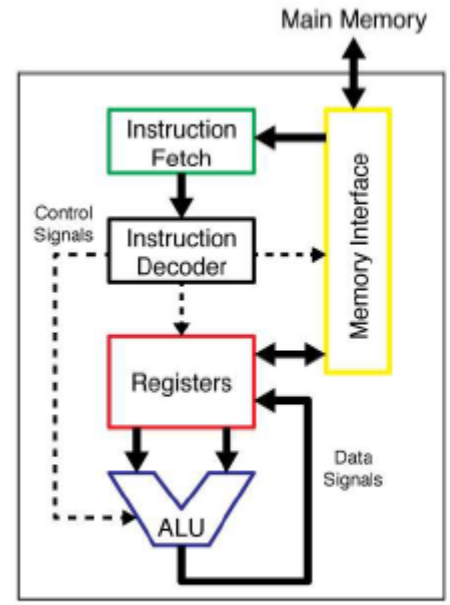
RESEARCH ARTICLE

Could a Neuroscientist Understand a Microprocessor?

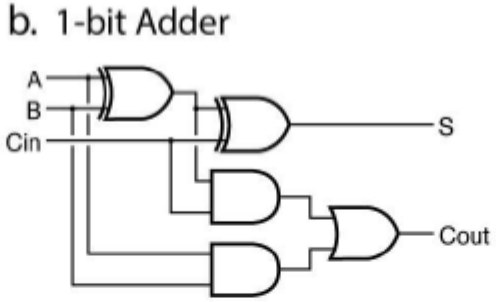
Eric Jonas[1]*, Konrad Paul Kording[2,3]

What is a satisfying understanding of a microprocessor?

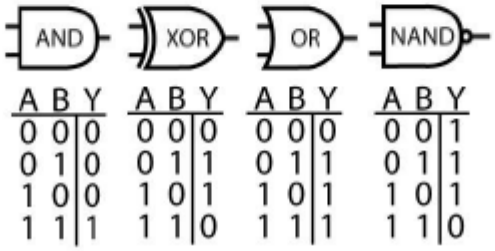
(serving as a ground truth)

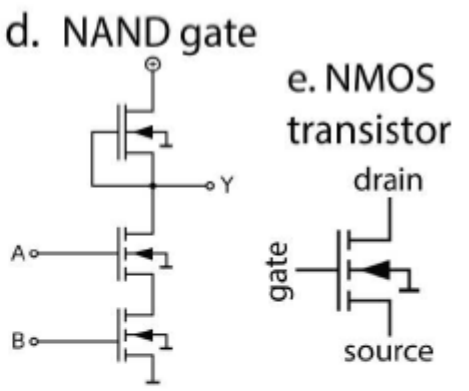


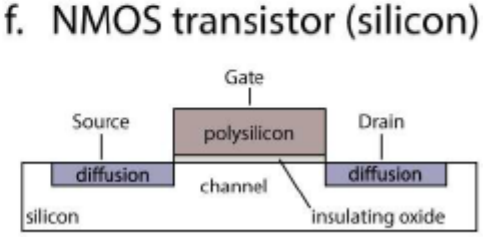
a.Processor Architecture



c. logic gate primitives

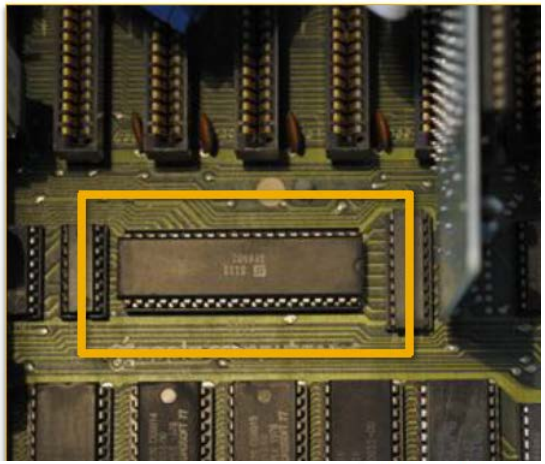




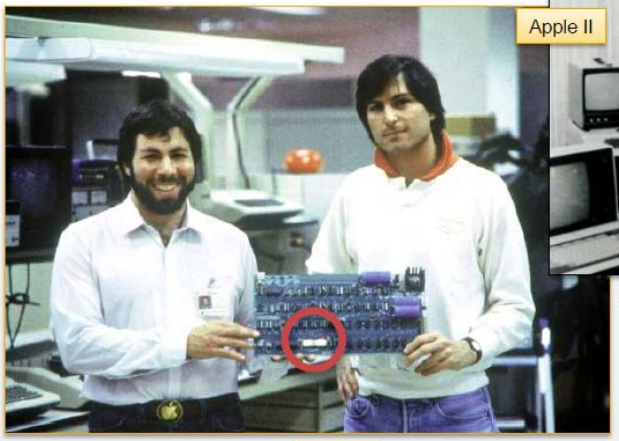


The model microprocessor

MOS Technology's
6502 CPU
Released 1975
Apple I, II
Commodore PET, C64
Atari 2600 (6507)
Atari 400, 800
Nintendo NES



Steve Wozniak, Steve Jobs



Apple I

[Klein] CCL 1.3

Paul Allen, Bill Gates

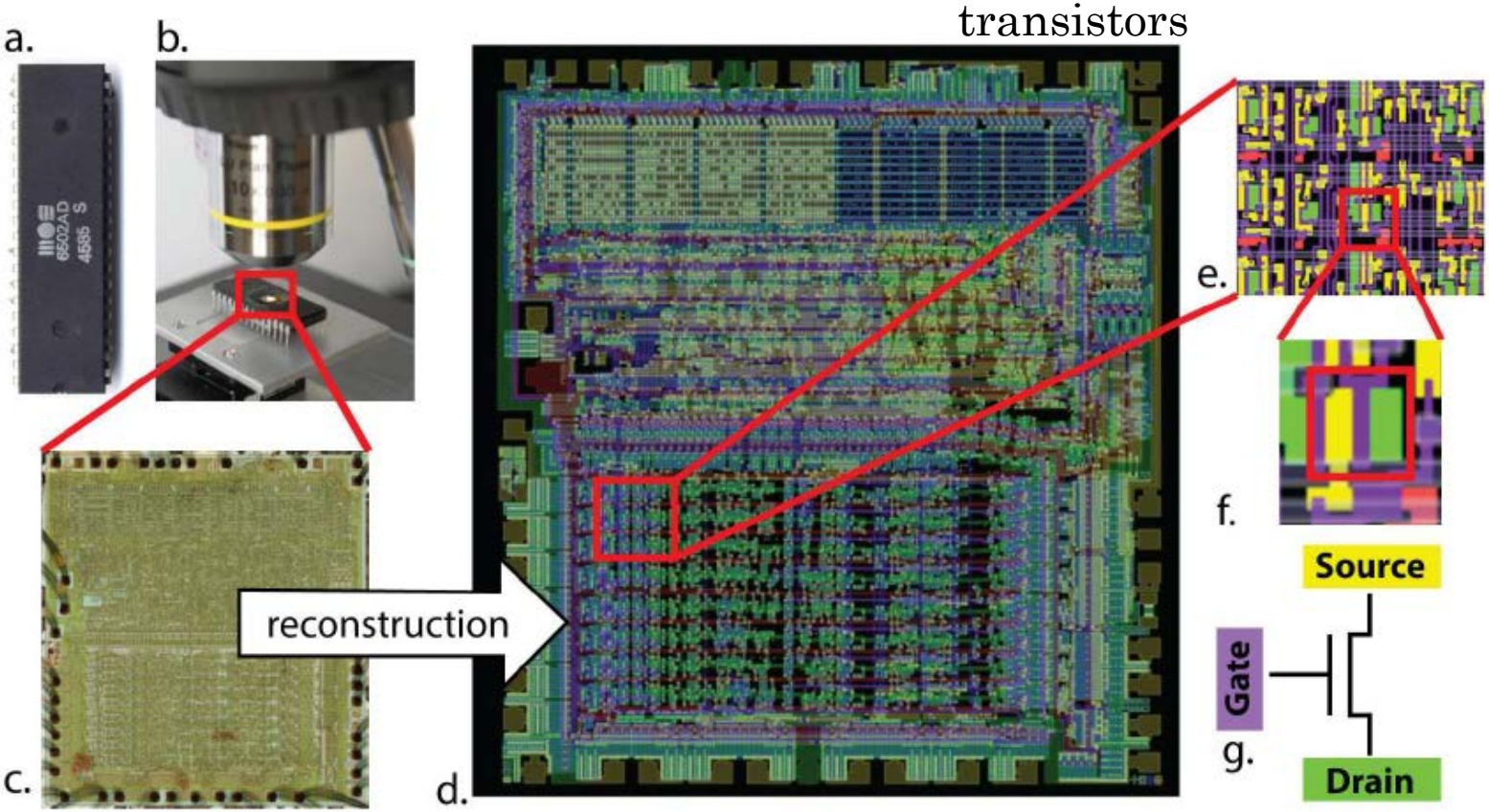
[Klein] CCL 1.3

Apple II

Commodore PET

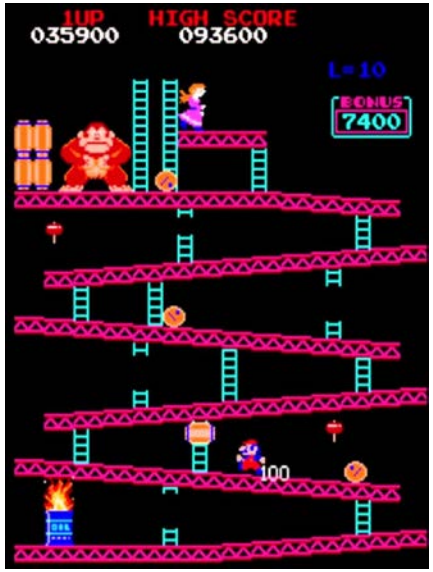
SIGGRAPH 2010

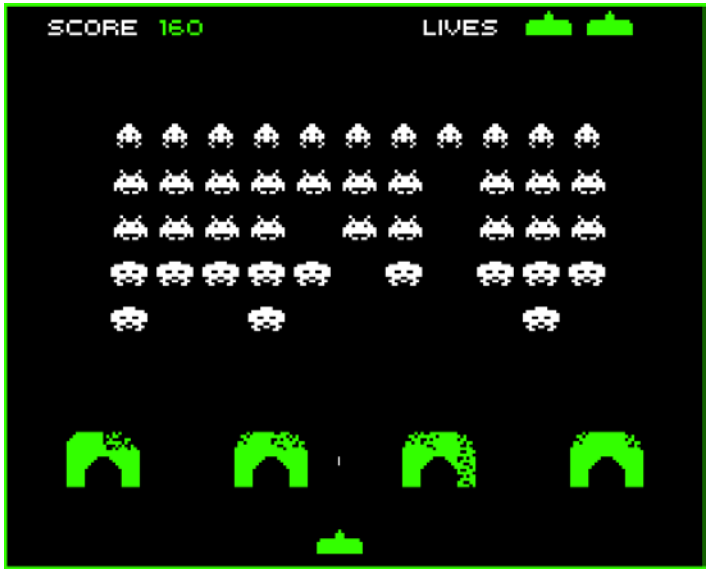
- Reverse engineering the chip
  - No digital representation
  - Designs lost



Experimental paradigm

- The neuroscientific methods should assess the chip while it «behaves».
- The authors chose the booting of three video games as «behaviour».
  - Video games have well defined inputs and outputs.
  - Dynamic input dependent output.

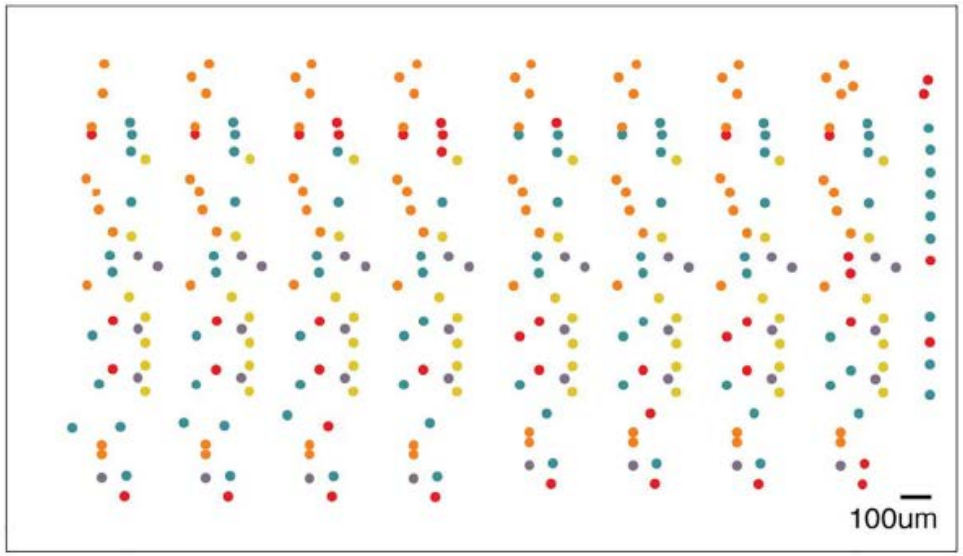




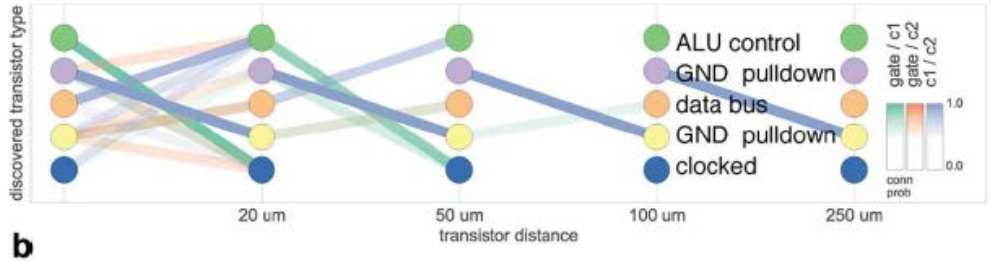
Donkey Kong (1981) Space Invaders (1981) Pitfall (1978)

Connectomics

- All connections of the model are known.
- Graph analysis methods for identifying cell types and circuit motifs.

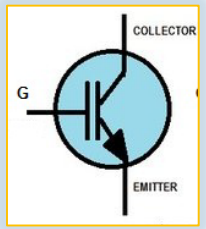


a



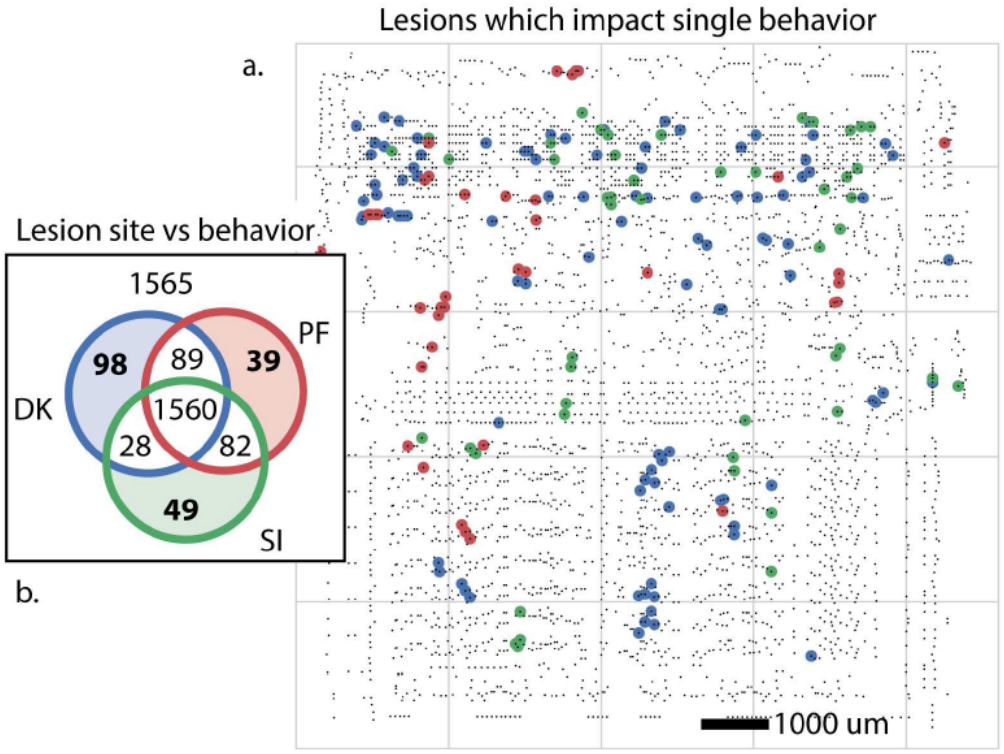
b

- Not showing any pattern of the known modularity of the chip.
- In reality there is one transistor type.



Lesions of single transistors

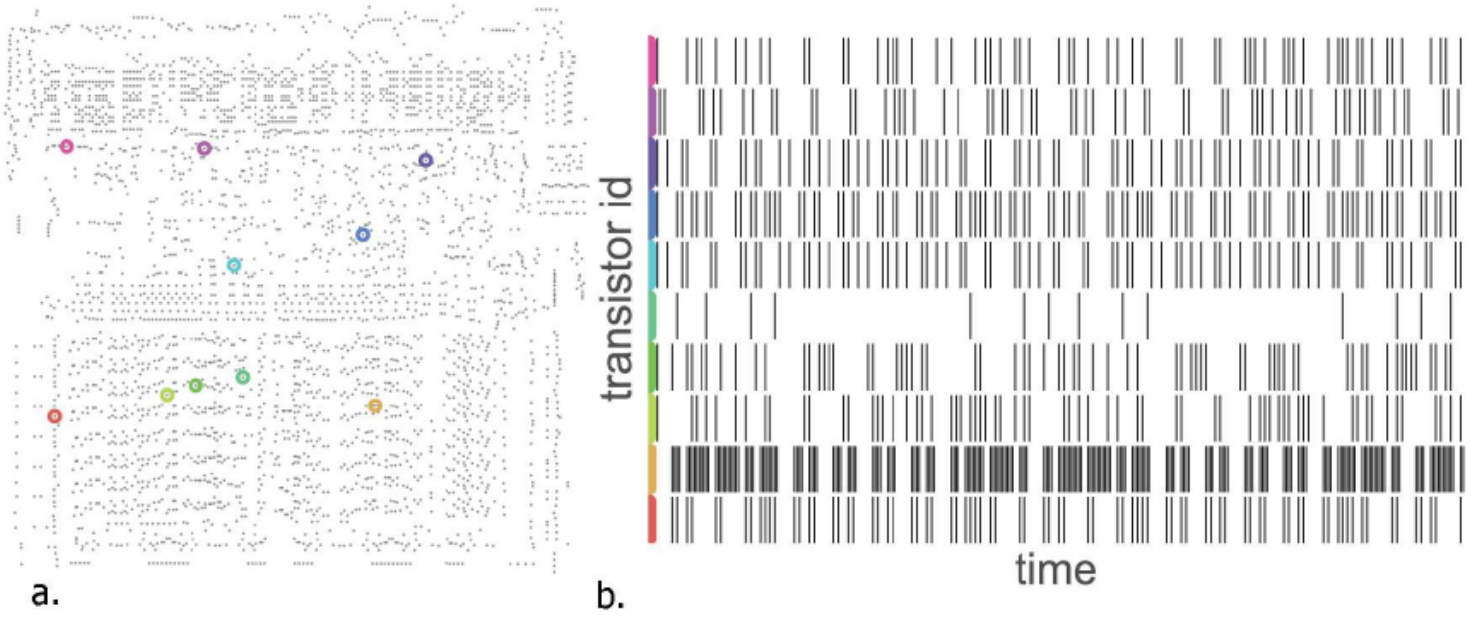
- Which transistor is necessary for which behaviour?



- Grossly misleading: one cannot attribute a transistor to Donkey Kong.
- It could be meaningful if one could narrow down the analysis to very simple behaviours, like arithmetic tasks.
  →Very difficult to achieve such a simplification in neuroscience.
- Even if a lesion abolishes a function it is difficult to interpret in terms of general computation.

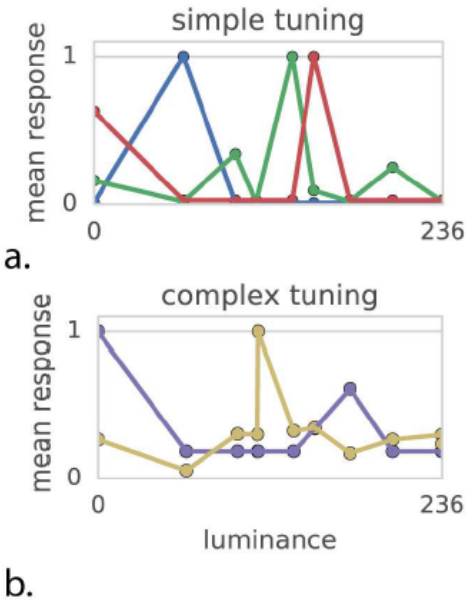
Tuning properties of individual transistors

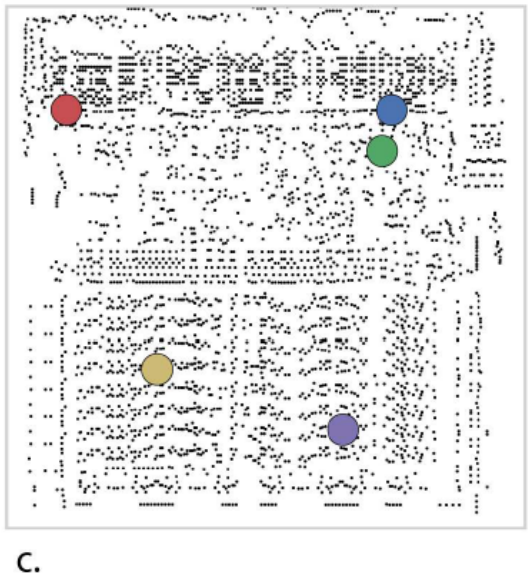
- Very broadly used for cells, circuits and brain areas.
- When a transistor switches it is considered as a spike.
- Correlation of spike rate to the luminance of recently displayed pixels.



Tuning properties of individual transistors

- Very frequently used for cells, circuits and brain areas.
- When a transistor switches it is considered as a spike.
- Correlation of spike rate to the luminance of recently displayed pixels.





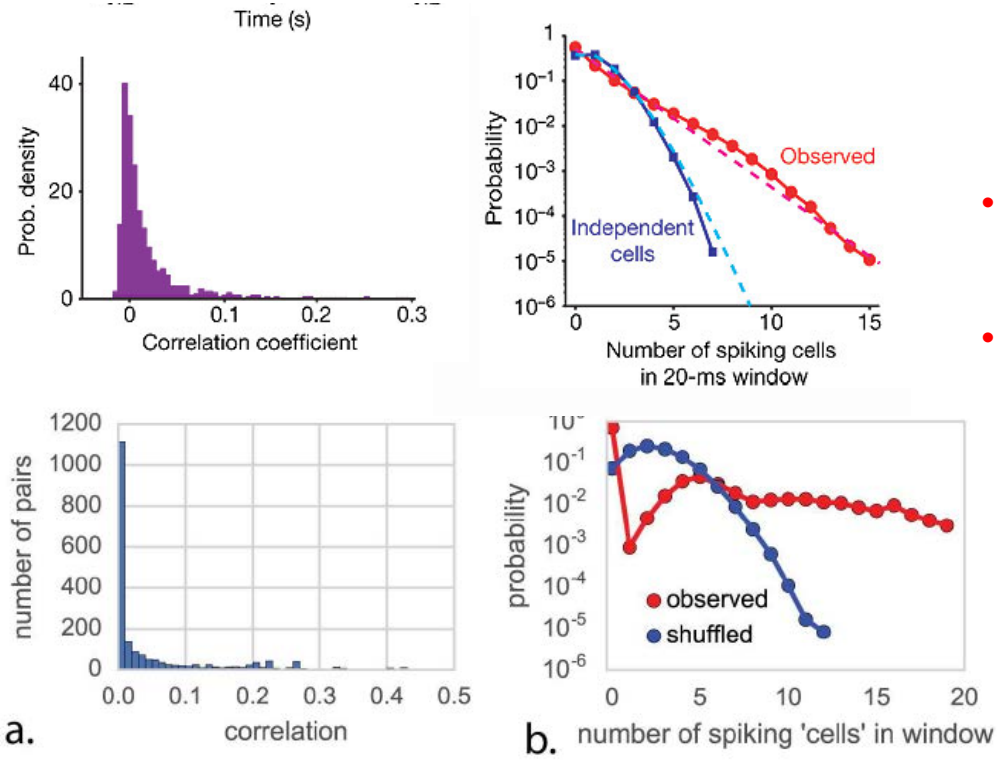
- Albeit strong tuning we know that there is no direct relation between individual transistor activity and pixel luminance.
- The tuning is highly non-linear and as such non-conclusive.
- Neurons can be involded in, or be upstream or downstream to a function. Infering from this to their exact function is very difficult.

Correlational structures of individual transistors

Nature **440**, 1007-1012 (20 April 2006) | doi:10.1038/nature04701; Received 22 September 2005; Accepted 6 March 2006; Published online 9 April 2006

Weak pairwise correlations imply strongly correlated network states in a neural population

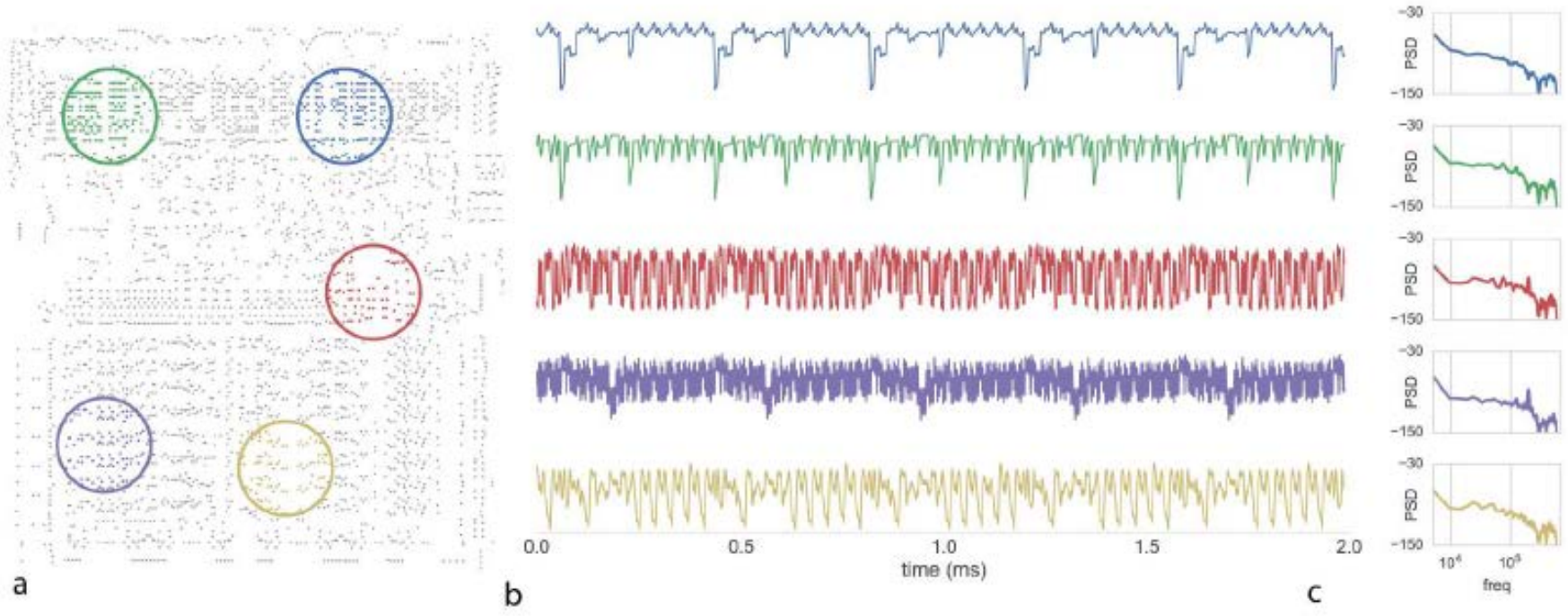
Elad Schneidman[1,2,3], Michael J. Berry, II[2], Ronen Segev[2] & William Bialek[1,3]



- In case of the processor there is no higher organisational principle explaining this.
- This interpretation may be questionable in the brain too.

Analyzing local field potentials

- Analogous to LFP or the BOLD signal in fMRI



- The PCU shows a rough power-law like distribution of frequencies. This is usually interpreted as a strong marker for self organized criticality. Very unlikely here.

- In neuroscience, frequency distributions and oscillations are thought to be indicative of certain tasks/regions. Here they are only epiphenomena.

Granger causality

- Granger causality aims to describe functional connectivity. It is extensively used.

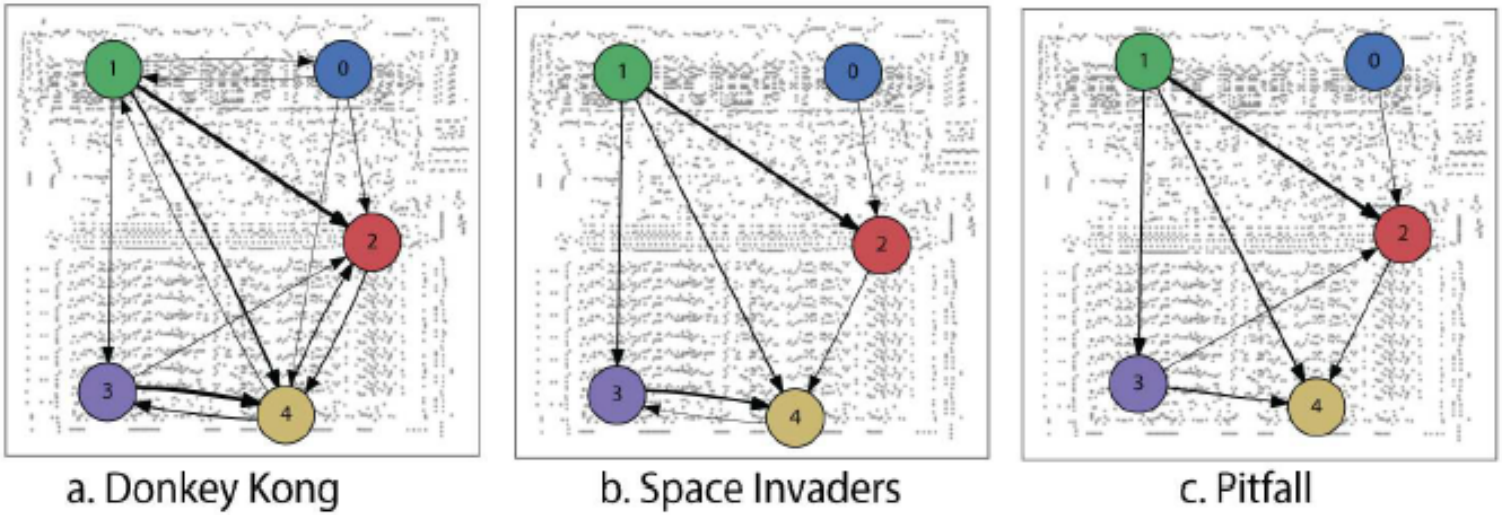
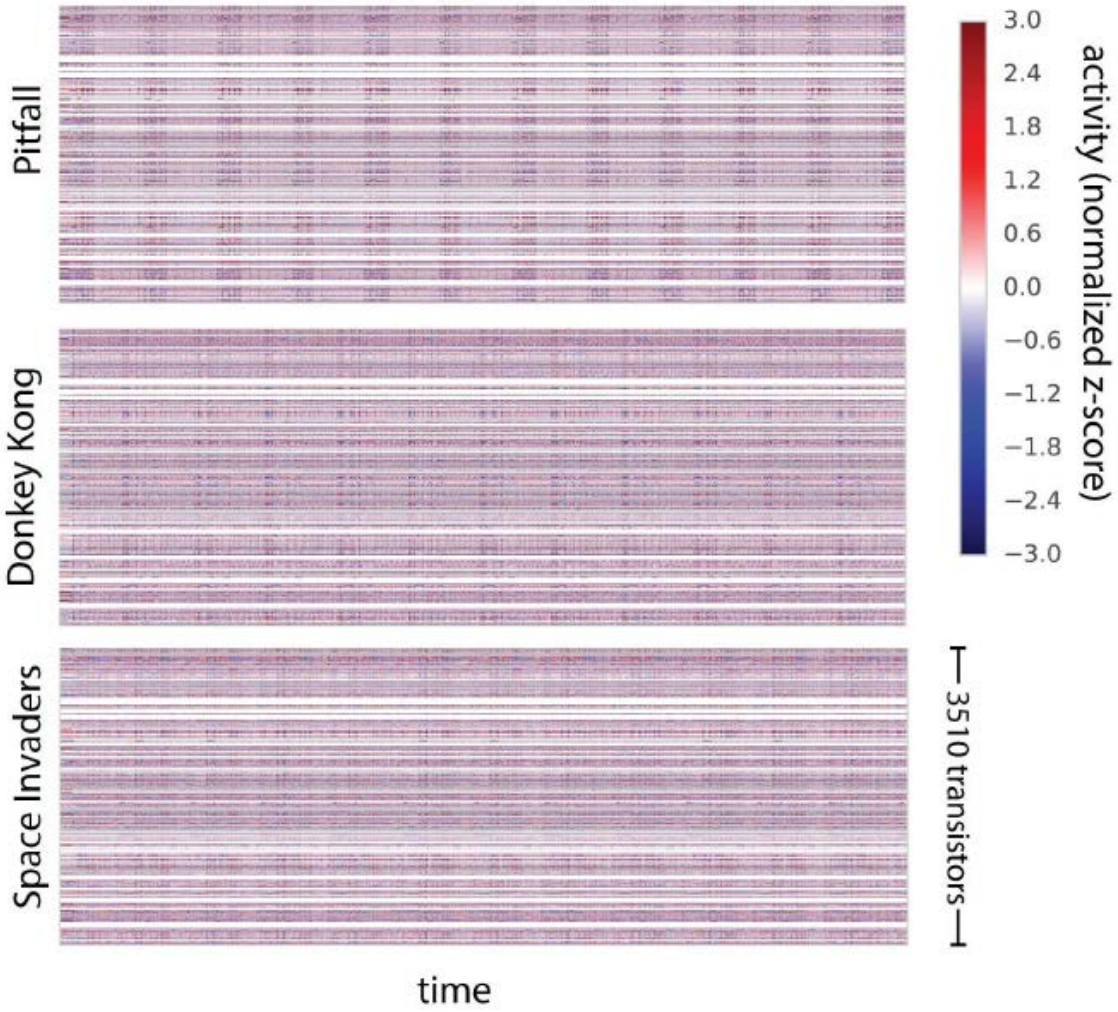


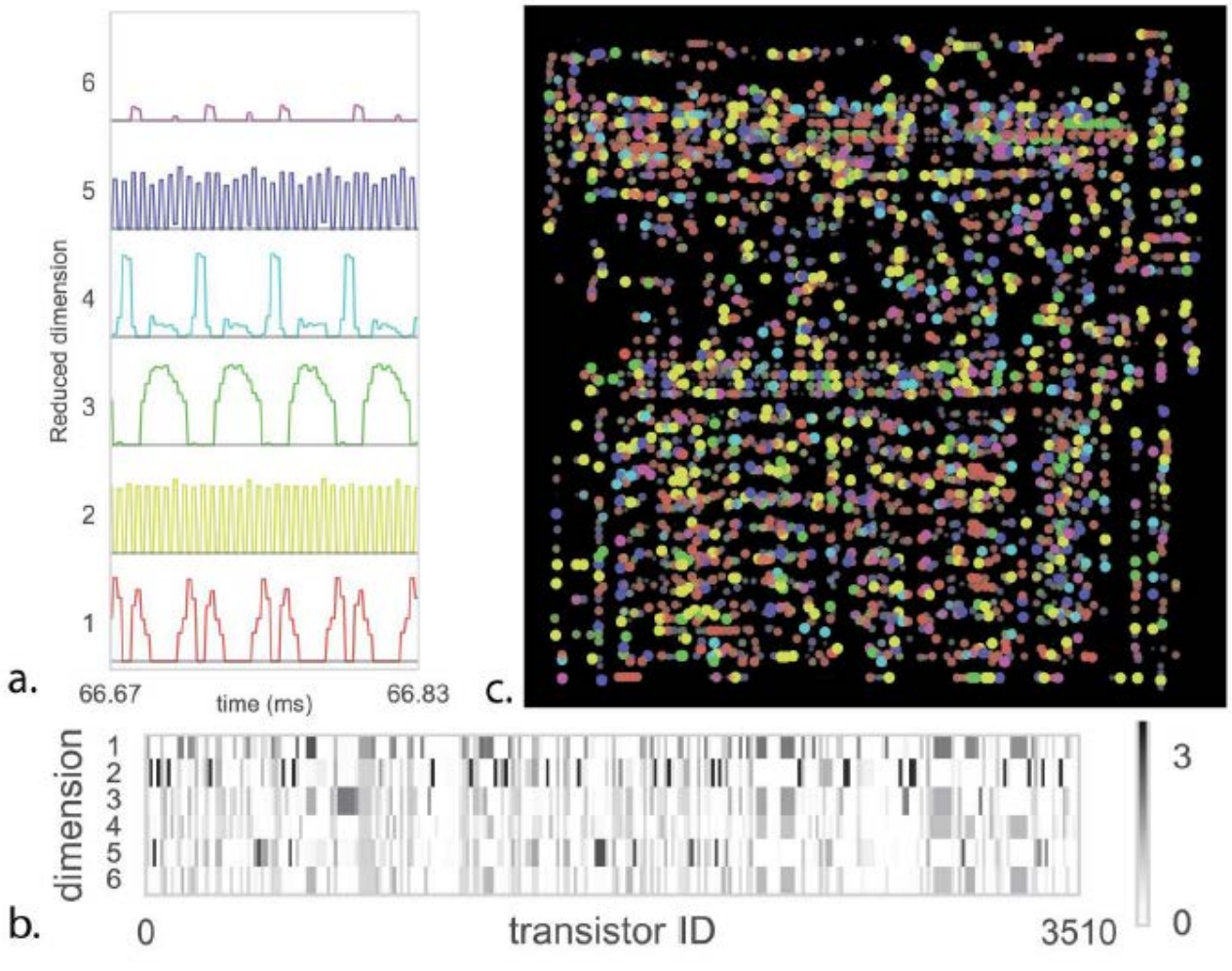
Fig 9. Analyzing conditional Granger causality to understand functional connectivity. Each of the recordings come from a well defined functional subcircuit. Green and blue are two parts of the decoder circuit. Red includes the status bits. Violet are part of the registers and yellow includes parts of the accumulator. We estimated for each behavioral state from LFP sites indicated in Fig 8. Arrows indicate direction of Granger-causal relationship, arrow thickness indicates effect magnitude.

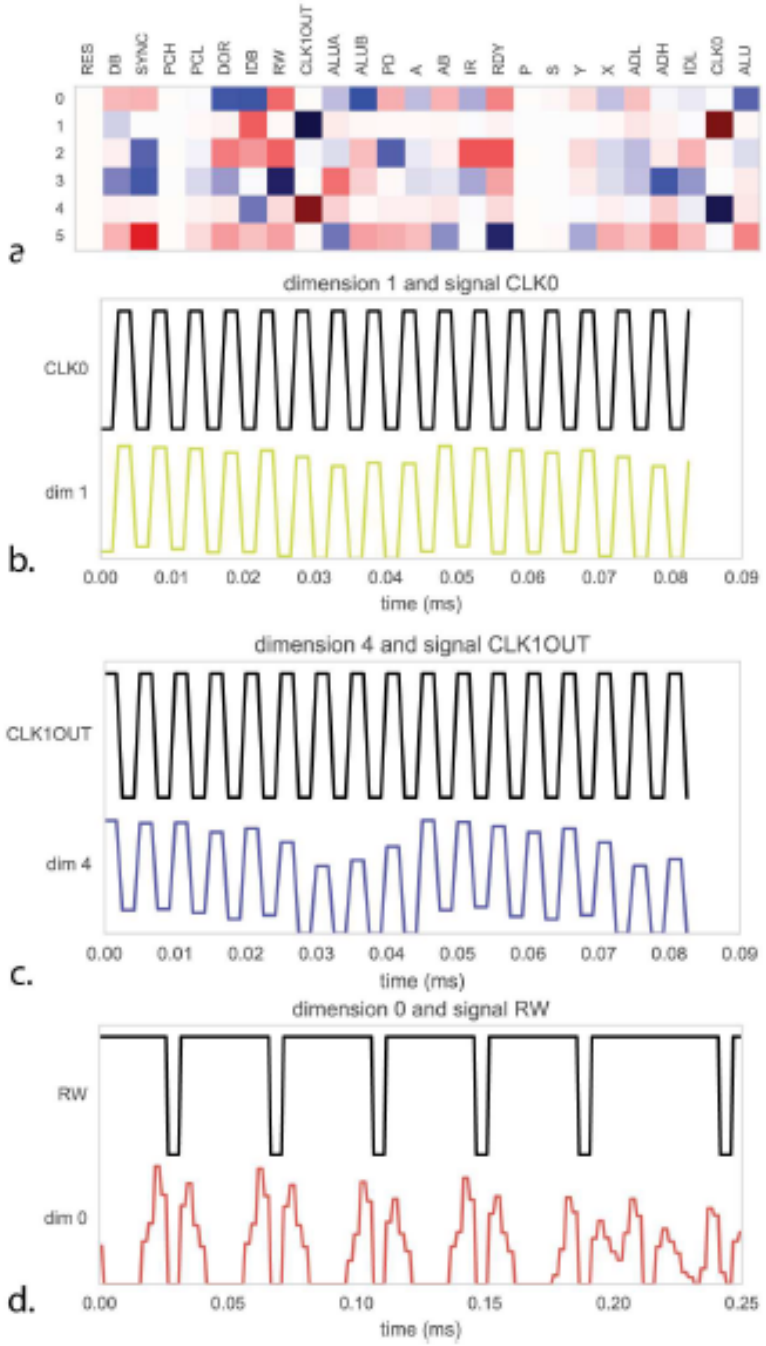
- ”Granger causality tells us how activity in the past are predictive of activity in the future, and the link from there to causal interactions is tentative at best.”

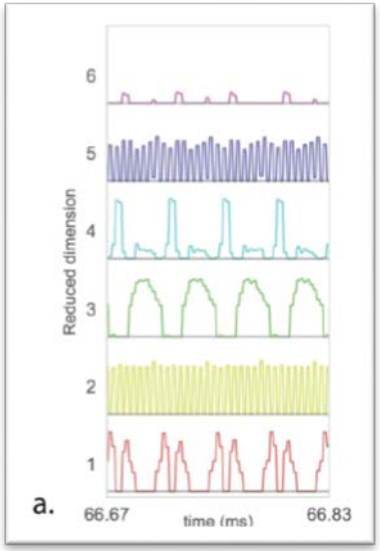
Dimensionality reduction



Dimensionality reduction







- Such correlations are seen in neuroscience and give rise to important papers.

Spatially Distributed Local Fields in the Hippocampus Encode Rat Position

Gautam Agarwal[1], Ian H. Stevenson[1,*], Antal Berényi[2,3], Kenji Mizuseki[2,†], György Buzsáki[2,††], Friedrich T. Sommer[1,††]

+ See all authors and affiliations

Science 09 May 2014:

- But even in a human designed CPU where we know what to look for, we are not at all able to describe how information flows.

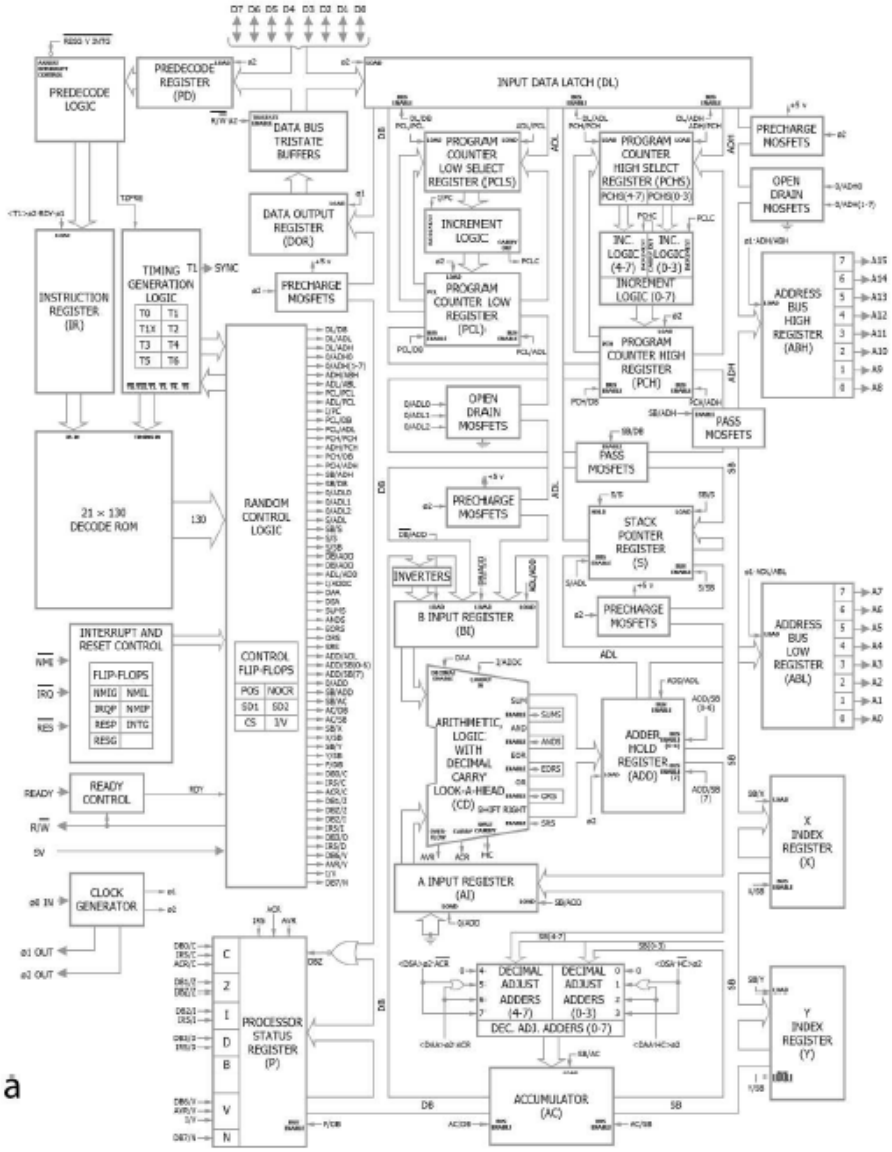
Summary and notes

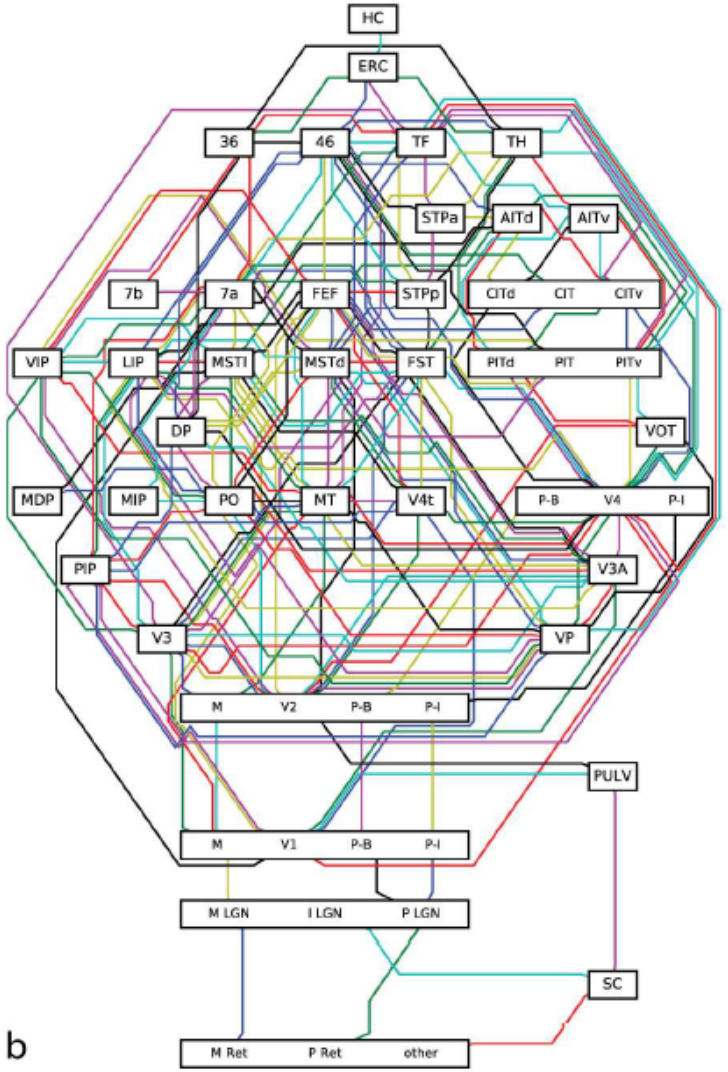
- A set of current data acquisiton and interpretation tools in neuroscience were tested.
- It was not possible with these tools to get any rough understanding of how a human-designed microprocessor works.

Summary and notes

	CPU	Brain
Signal	Digital	Analog
Biophysical complexity	Low	High
Temporal scale	Fast	Slow
Parallelity	Low	High
Inputs	Limited	Many
Development	Design	Evolution

- Nevertheless, there is no reason to think that our methods would be more meaningful when testing the brain.


D7 D6 D5 D4 D3 D2 D1 D0
PREDECODE LOGIC
PREDECODE REGISTER (PD)
INPUT DATA LATCH (DL)
DATA BUS TRISTATE BUFFERS
PRECHARGE MOSFETS
OPEN DRAIN MOSFETS
PROGRAM COUNTER LOW SELECT REGISTER (PCLS)
PROGRAM COUNTER HIGH SELECT REGISTER (PCHS)
DATA OUTPUT REGISTER (DOR)
INCREMENT LOGIC
INC. LOGIC (4-7)
INC. LOGIC (0-3)
INCREMENT LOGIC (0-7)
INSTRUCTION REGISTER (IR)
TIMING GENERATION LOGIC
SYNC
PROGRAM COUNTER LOW REGISTER (PCL)
PROGRAM COUNTER HIGH REGISTER (PCH)
ADDRESS BUS HIGH REGISTER (ABH)
21 × 130 DECODE ROM
RANDOM CONTROL LOGIC
PASS MOSFETS
STACK POINTER REGISTER (S)
INVERTERS
B INPUT REGISTER (BI)
ADDRESS BUS LOW REGISTER (ABL)
INTERRUPT AND RESET CONTROL
FLIP-FLOPS
CONTROL FLIP-FLOPS
ARITHMETIC, LOGIC WITH DECIMAL CARRY LOOK-A-HEAD (CD)
ADDER HOLD REGISTER (ADD)
READY CONTROL
X INDEX REGISTER (X)
A INPUT REGISTER (AI)
CLOCK GENERATOR
DECIMAL ADJUST ADDERS (4-7)
DECIMAL ADJUST ADDERS (0-3)
DEC. ADJ. ADDERS (0-7)
Y INDEX REGISTER (Y)
PROCESSOR STATUS REGISTER (P)
ACCUMULATOR (AC)
a


HC
ERC
36
46
TF
TH
STPa
AITd
AITv
7b
7a
FEF
STPp
CITd
CIT
CITv
VIP
LIP
MSTI
MSTd
FST
PITd
PIT
PITv
DP
VOT
MDP
MIP
PO
MT
V4t
P-B
V4
P-I
PIP
V3A
V3
VP
M
V2
P-B
P-I
PULV
M
V1
P-B
P-I
M LGN
I LGN
P LGN
SC
M Ret
P Ret
other
b

Summary and notes

- The authors propose for the CPU problem:
  - More hierarchical analytic tools.
  - Running one simple specific code.

What is the case for molecular biology?

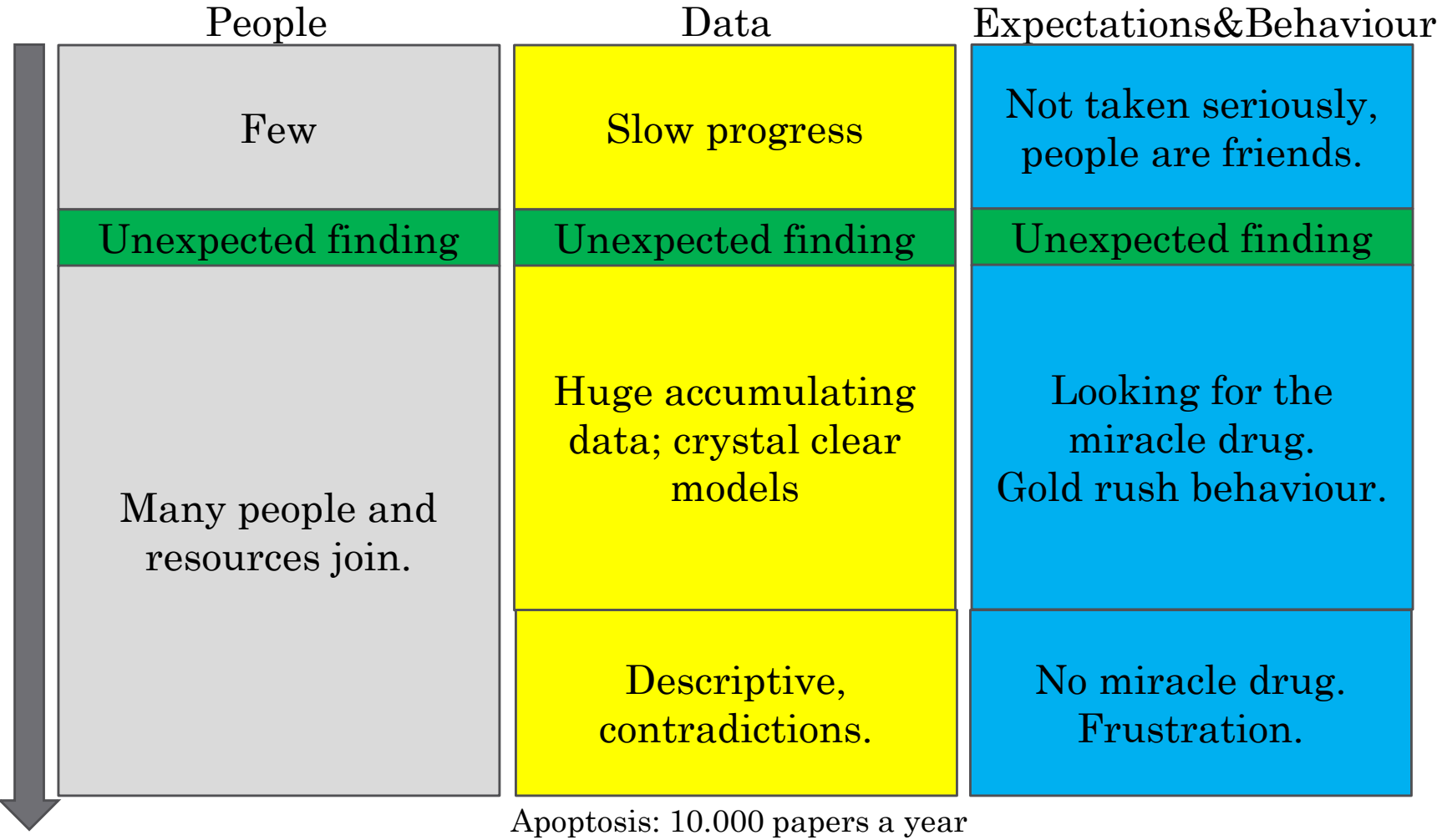
Cancer Cell

CORRESPONDENCE

Can a biologist fix a radio?—Or, what I learned while studying apoptosis

Yuri Lazebnik

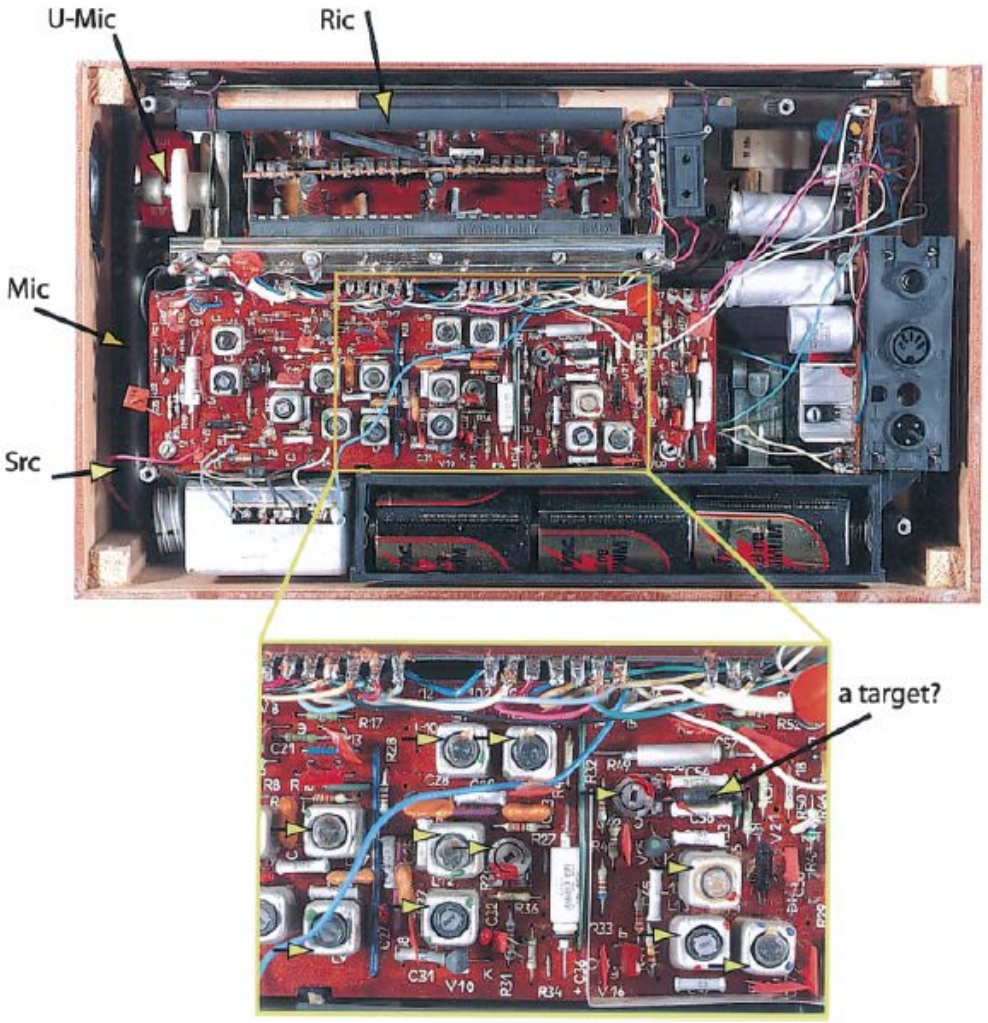
Personal observation: How do fields develop.



Apoptosis: 10.000 papers a year
P53: >85000 papers

His argument is that there is a problem with the methodological approach in general.

What would happen if a biologist tries to fix a broken radio?



A Radio converts signal, has hundreds of components ~ like a complex signal transduction pathway.

For the start the biologist sees it is a box that plays music.

Buy many functioning radios.

Take them apart and describe: many object with various shapes, colors, legs.

Categorizing them into families.

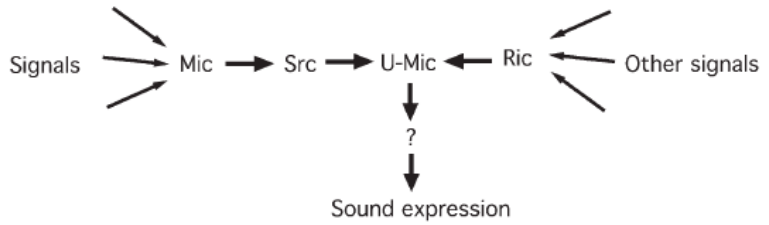
Does changing these properties influence function?

Cut single connections.

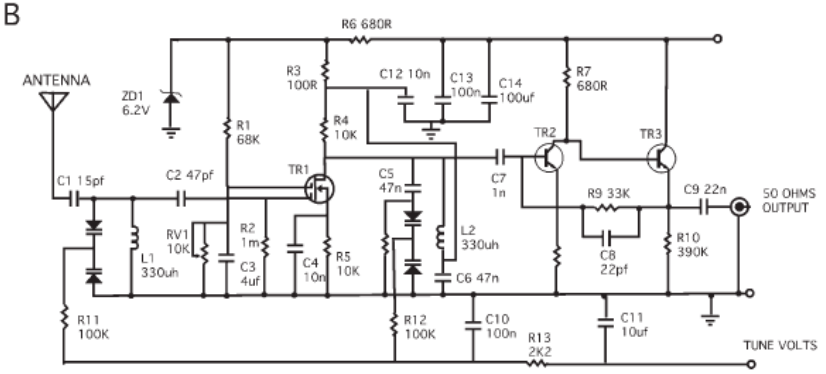
Crush the radio and check which pieces stay together.

Remove single components or shoot the radio at certain regions.

A



B



If we have cataloged all components, connections and knockouts: can we fix the radio?

- If a certain component is broken: yes. (pharma industry's «give me a target»)

- Otherwise no, because the components are tuneable. It is likely that multiple components are tuned incorrectly.

Cancer Cell

CORRESPONDENCE

Can a biologist fix a radio?—Or, what I learned while studying apoptosis

Yuri Lazebnik

The authors argument is that these processes can only be modelled if we apply formal language, like engineers do. Doing additional experiments will not simplify the problem to basic arithmetic operations or simple relations.

What does this mean practically?

A systems biology approach.

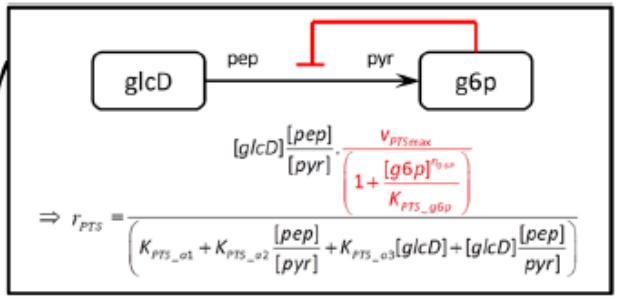
→But we tend to skip an important level of knowledge:
«…biochemistry disappeared in the same year as communism. (Sydney Brenner 1995)»
Kinetic properties of proteins are reported to much smaller extents as desirable. Modelling without information about kinetics is very difficult.

Integrating Kinetic Model of *E. coli* with Genome Scale Metabolic Fluxes Overcomes Its Open System Problem and Reveals Bistability in Central Metabolism

Ahmad A. Mannan[1*], Yoshihiro Toya[2,3], Kazuyuki Shimizu[3,4], Johnjoe Mc
Andrzej M. Kierzek[1‡*], Andrea Rocco[1‡*]

[1] Kinetic Model (KM) Construction

- Literature search for reaction enzyme mechanisms and kinetic parameters.
- Writing reaction formulae and ODEs.



[2] Model Parameterization

- Use multi-omics steady state data – Fluxome, metabolome and proteome → Keio multi-omics dataset.

[3] Integrating GSMN Model with KM

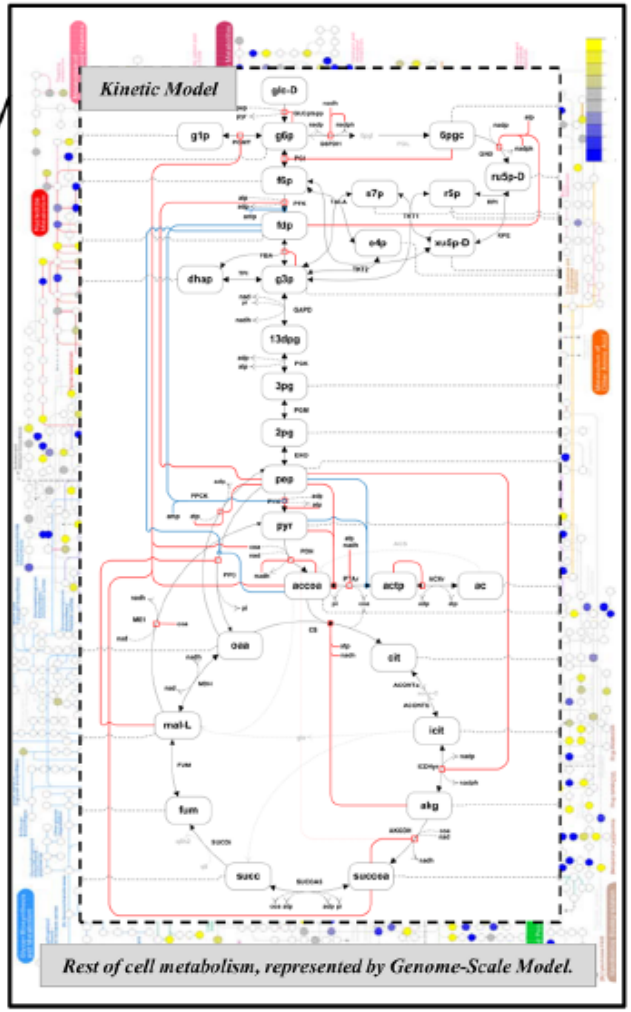
- Parameterize genome-scale metabolic network (GSMN) model with multi-omics data → generate steady state fluxes.
- Determine fluxes of KM reactions and connecting reactions from GSMN model.
- Parameterize KM with both GSMN model fluxes and multi-omics data.

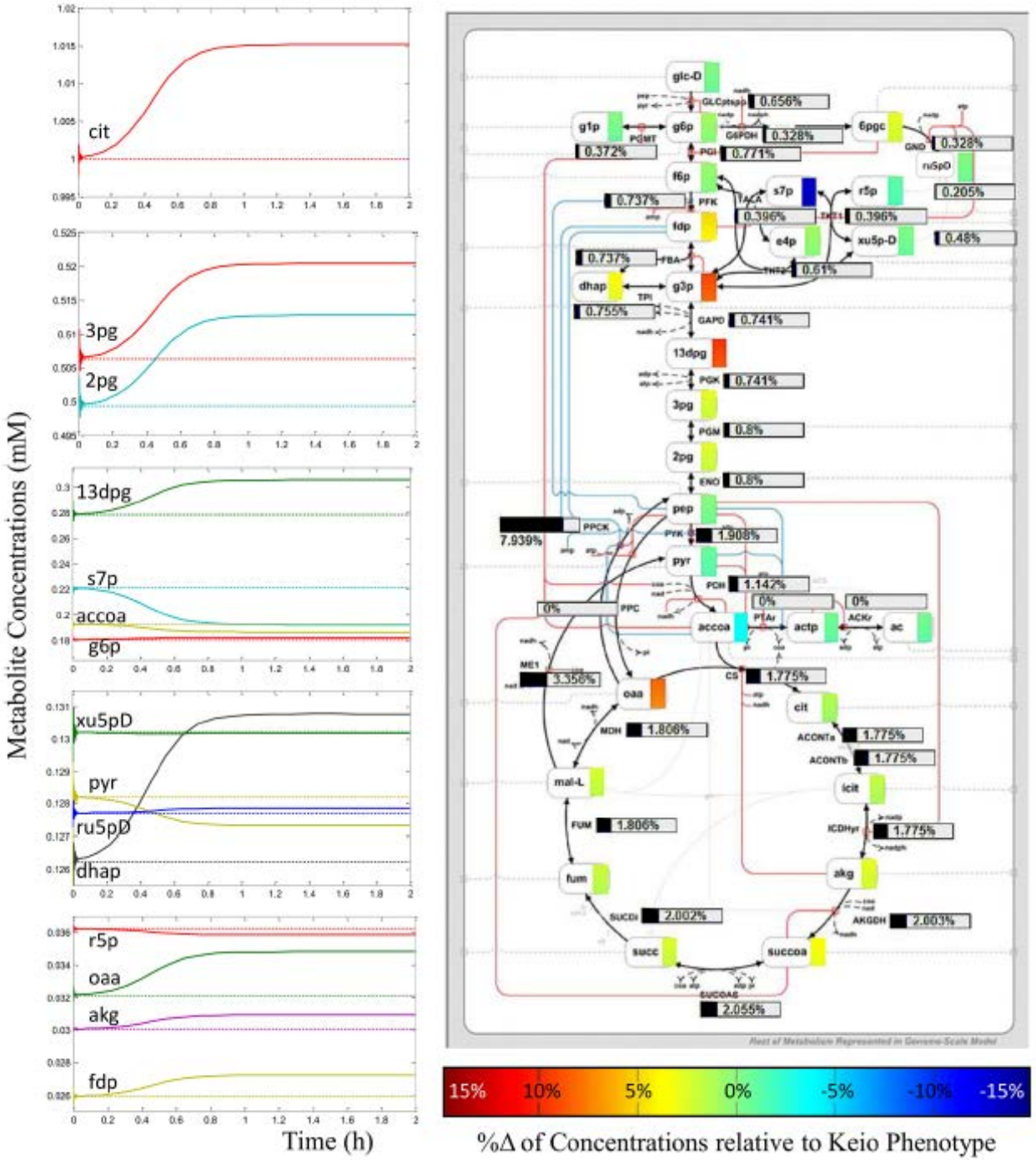
[4] Stabilizing Keio Steady State

- Optimize to minimally adjust KM reaction parameters to yield a stable steady state.

[5] Finding System Steady States

- Fix constant growth and media conditions → Set $d[X]/dt = d[glcD_{ex}]/dt = 0$.
- Run simulation from 1000's of random initial values and determine steady states.
- Assess their stability – Eigenvalues.
- Stable steady state = Metabolic steady state.




Metabolite Concentrations (mM)
cit
3pg
2pg
13dpg
s7p
accoa
g6p
xu5pD
pyr
ru5pD
dhap
r5p
oaa
akg
fdp
Time (h)
glc-D
g1p
g6p
6pgc
ru5pD
f6p
s7p
r5p
fdp
e4p
xu5p-D
dhap
g3p
13dpg
3pg
2pg
pep
pyr
accoa
actp
ac
oaa
cit
mal-L
icit
fum
akg
succ
succoa
15%
10%
5%
0%
-5%
-10%
-15%
%Δ of Concentrations relative to Keio Phenotype

Summary

- More data is needed about the kinetic properties of individual components in a signal transduction pathway.
- The systems biology approach is difficult, but probably the only way.
- Keeping in mind David Marr's proposition.
  - In order to construct a cohesive whole picture all 3 levels should be congruent:
    - Goal is e.g. bistable state of proliferation, health/cancer.
    - The representation is the set of proteins. The algorithm we only have very roughly because kinetic data is lacking (hardware level).
    - Hardware level: Also the proteins, but actually their structural and biochemical (kinetic) properties.