computational biology

Biological Systems are complex, thus, a combination of experimental and computational approaches are needed.

computational biology

- Computational Biology ≠ Bioinformatics
  - More than sequences, database searches, statistics or image analysis.
- A part of Computational Science
  - Using mathematical modeling, simulation and visualization
  - Complementing theory and experiment

simplest chemical reaction

$A \rightarrow B$

- irreversible, one-molecule reaction
- examples: all sorts of decay processes, e.g. radioactive, fluorescence, activated receptor returning to inactive state
- any metabolic pathway can be described by a combination of processes of this type (including reversible reactions and, in some respects, multi-molecule reactions)

simplest chemical reaction

A → B

various levels of description:

- homogeneous system, large numbers of molecules = ordinary differential equations, kinetics
- small numbers of molecules = probabilistic equations, stochastics
- spatial heterogeneity = partial differential equations, diffusion
- small number of heterogeneously distributed molecules = single-molecule tracking (e.g. cytoskeleton modelling)

kinetic description

- Imagine a box containing N molecules.

  How many will decay during time t? k N
- Imagine two boxes containing N/2 molecules each.

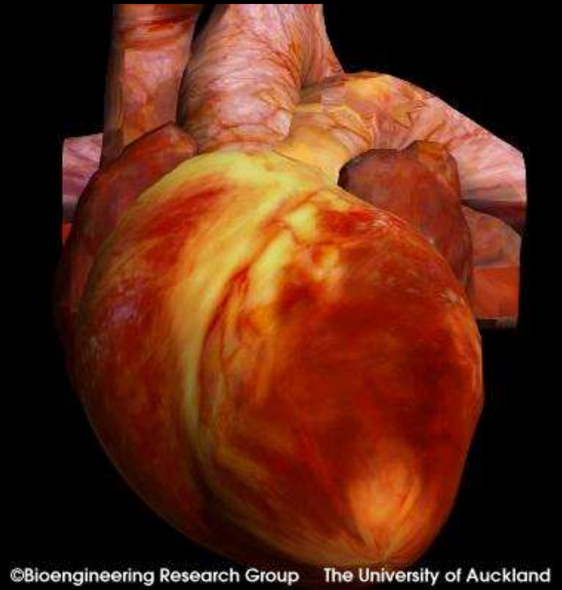
  How many decay? k N
- Imagine two boxes containing N molecules each.

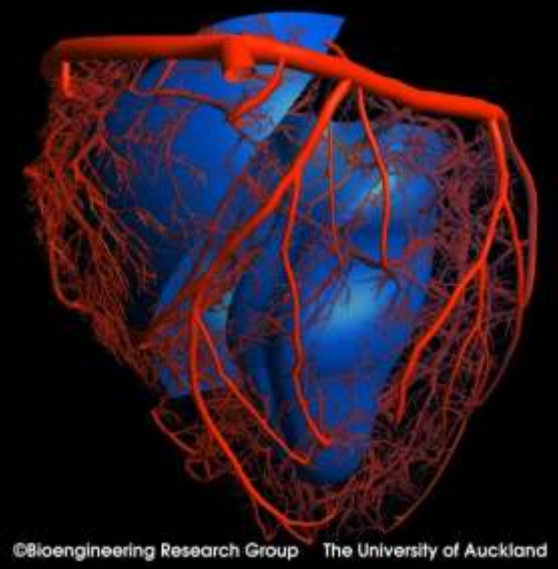
  How many decay? 2k N
- In general:

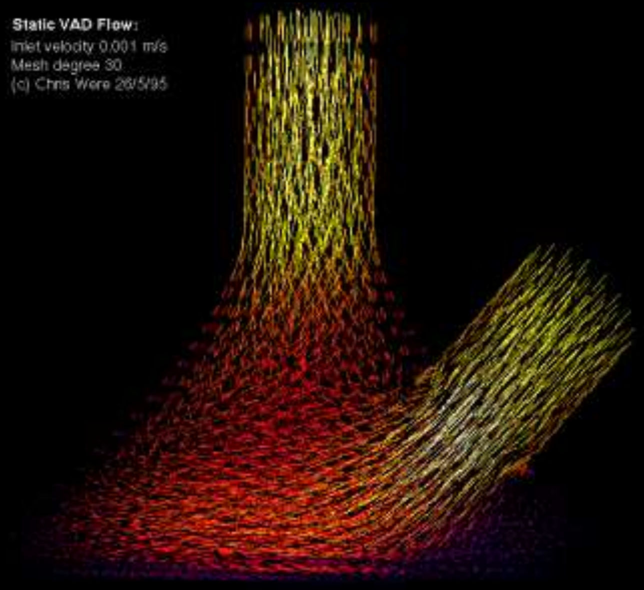
$$-\frac{dn(t)}{dt} = \lambda * n(t) \Leftrightarrow \quad n(t) = N_0 e^{-\lambda t}$$

virtual human

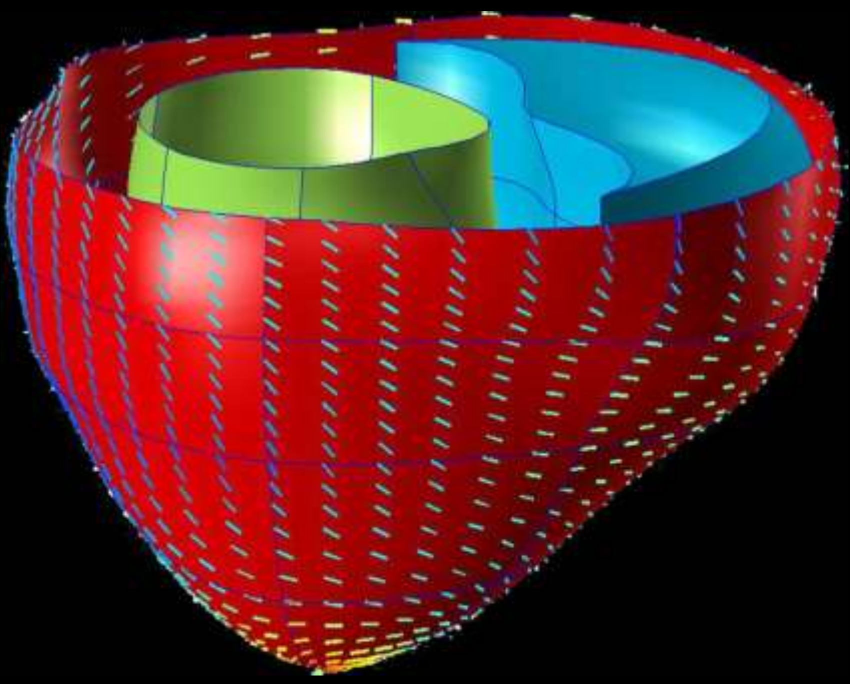
Simulation of complex models of cells, tissues and organs



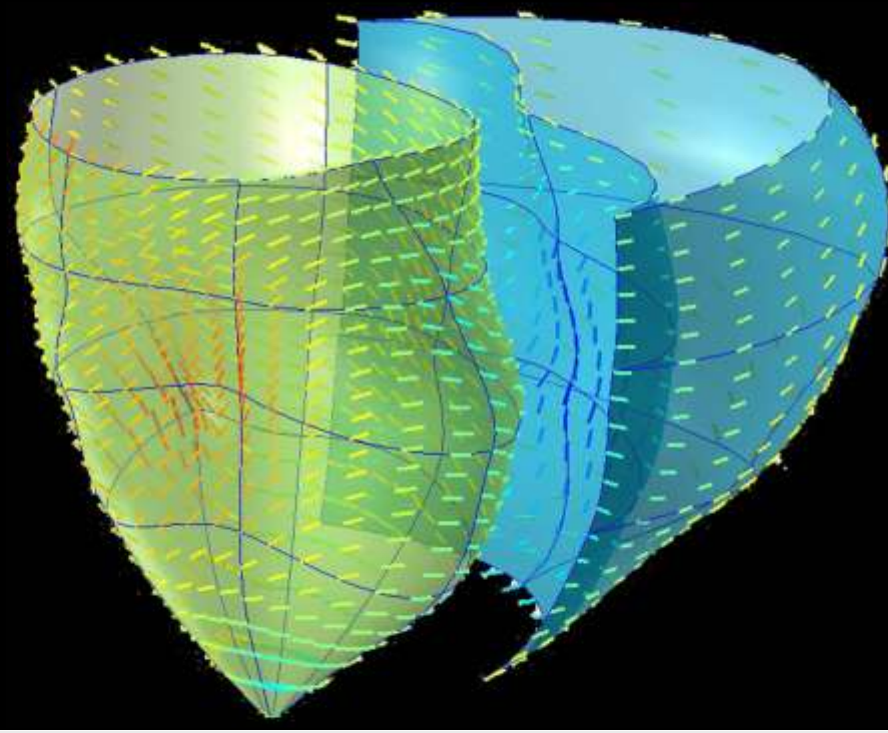




cardiac ventricular anatomy



Epicardial Fibers – FEM Model



Endocardial Fibers – FEM Model

www.ccmb.jhu.edu

physiome project

- "A worldwide effort to define the physiome by developing databases and models which will facilitate the understanding of the integrative functions of cells, organs and organisms."

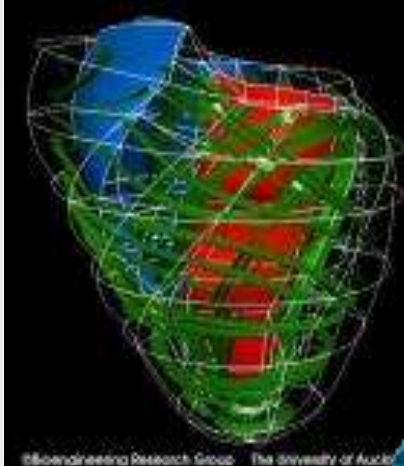
defenition

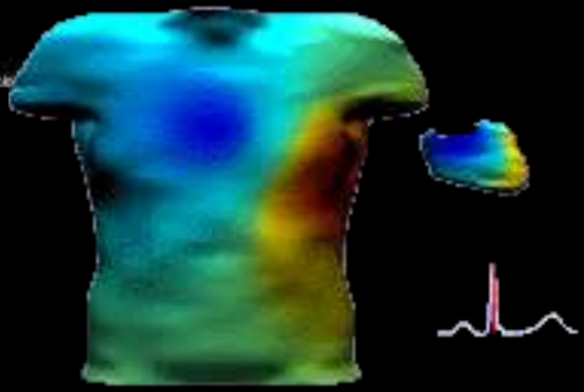
Physiome is the quantitative and integrated description of the functional behavior of the physiological state of an individual or species.

p h y s i o m e   p r o j e c t

main objective:

"… to understand and describe the human organism, its physiology and pathophysiology quantitatively, and to use this understanding to improve human health."





b i o i n f o r m a t i c s

what is bioinformatics?

- Methods to analyse, store, search, retrieve and represent biological data by computers /in computers
- massive amounts of data: databases
- extracting information and knowledge from "raw" data
- for most bioscientists, all they need in bioinformatics is sequence analysis

what is bioinformatics?

Bioinformatics – a Definition[1]

(Molecular) **bio** – **informatics:** bioinformatics is conceptualising biology in terms of molecules (in the sense of Physical chemistry) and applying **"informatics techniques"** (derived from disciplines such as applied maths, computer science and statistics) to **understand** and **organise** the **information** associated with these molecules, on a **large scale.** In short, bioinformatics is a management information system for molecular biology and has many **practical applications.**

[1]As submitted to the Oxford English Dictionary.

definitions of bioinformatics

some definitions

- use of computers to catalog and organize molecular life science information into meaningful entities.
- subset of computational biology

what does it do?

- bioinformatics is not just the storage of data in a computer.
- bioinformatics is the use of computers to test a biological hypothesis prior to performing the experiment in the laboratory.
- bioinformatics is the design of software programs that analyze data.

bioinformatics databases

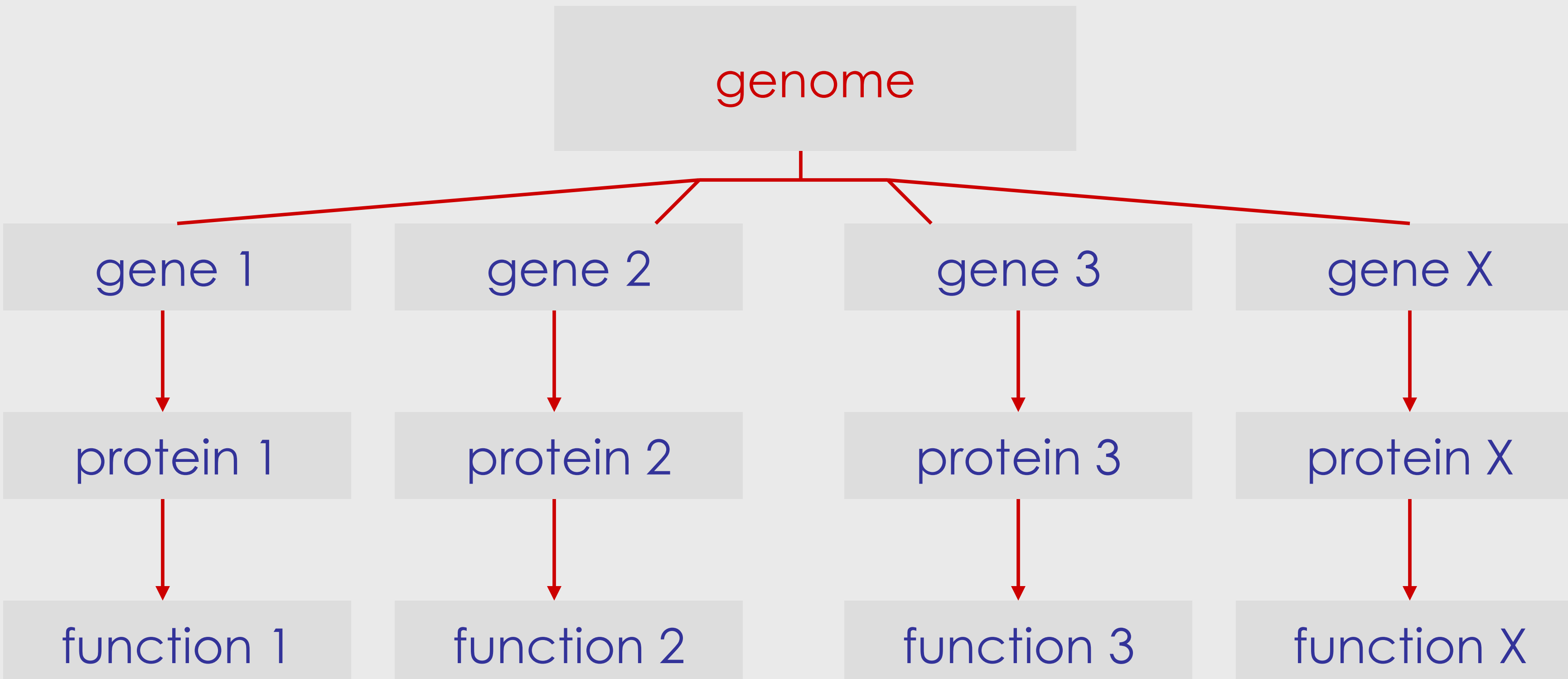
- nucleotide and protein sequences
- protein structures
- all sorts of functional data related to genes, proteins and their regulation, interactions etc.
- curated and non-curated databases

some goals

- sequence searching and sequence alignments
- looking at properties that can be analyzed/predicted from sequence data
- protein structures and their analysis
- structural classification
- visualisation of macromolecules

basis of molecular biology

- hierarchy of relationships:



genome size

FERN	160,000,000,000	
LUNGFISH	139,000,000,000	
SALAMANDER	81,300,000,000	
NEWT	20,600,000,000	
ONION	18,000,000,000	
GORILLA	3,523,200,000	
MOUSE	3,454,200,000	genes
HUMAN	3,400,000,000	31,000
Drosophila	137,000,000	13,500
C. Elegans	96,000,000	19,000
Yeast	12,000,000	6,315
E. Coli	5,000,000	5,361
smallest Genome	??????	

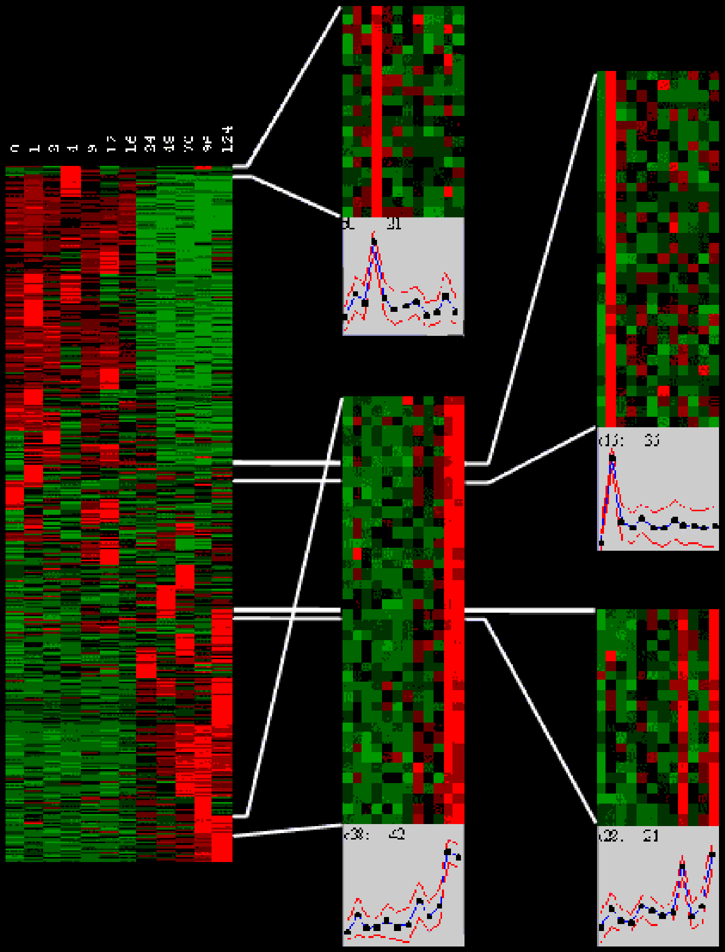
genomics

- comparative genomics
  - whole-genome analyses
  - evolution studies
  - analyses of components in a "complete" system
- functional genomics = inferring functions from data
  - expression patterns, gene regulation
  - sequence comparisons, homologue relationships
  - studies of gene variation, altered phenotypes

DNA microarrays

- massive data sets from simultaneous expression levels of thousands of genes
- impossible to grasp directly by the human mind
- methods are needed for finding meaningful results and patterns from the bulk of data

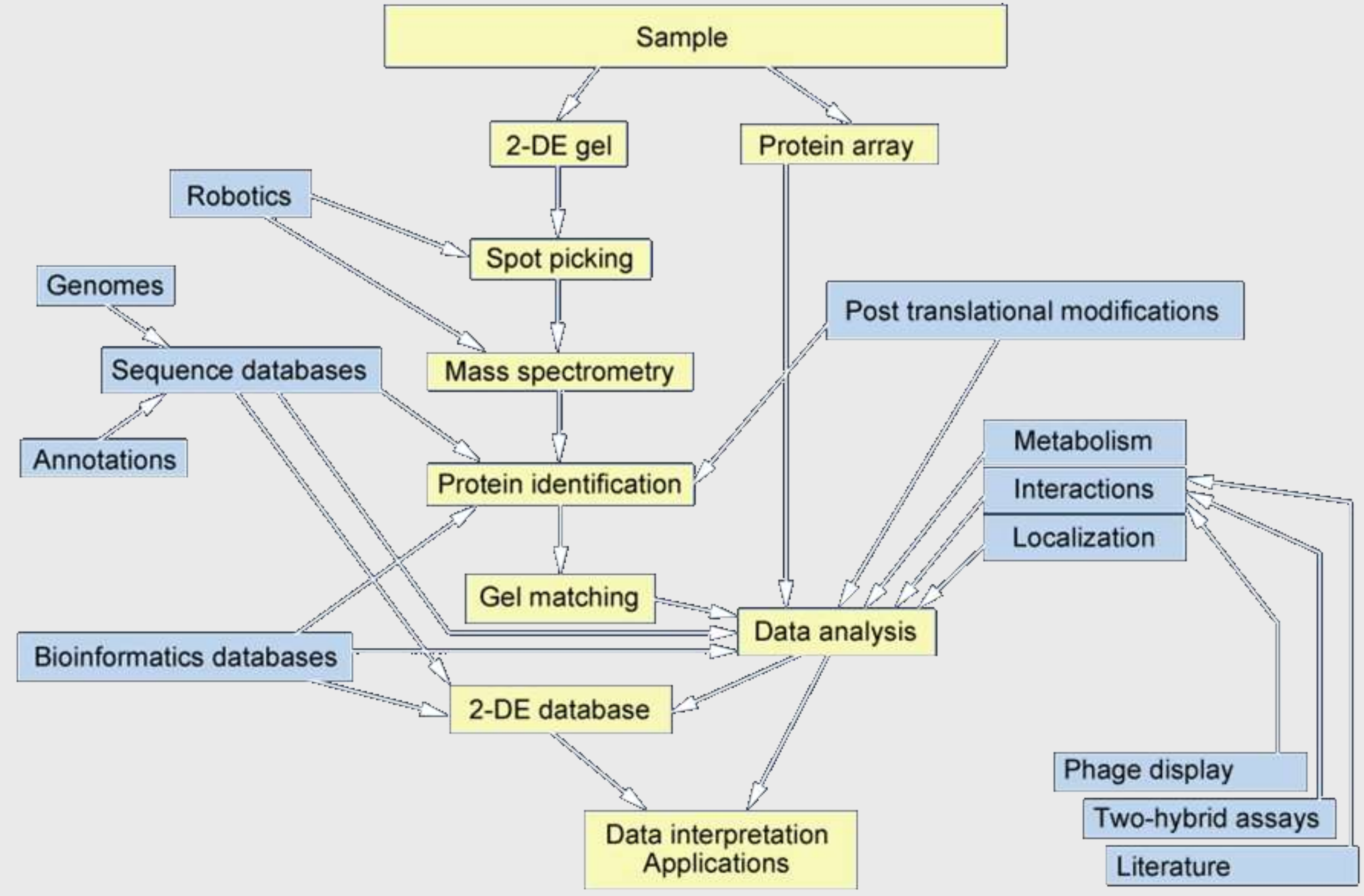
microarray data



p r o t e o m i c s

- expression proteomics = differential proteomics = 2D-PE + MS
- interaction proteomics
- functional proteomics = systematic perturbation or functi onal inactivation of proteins in a given environment
- structural proteomics

bioinformatics in proteomics



structural proteomics

- High-throughput determination of the 3D structure of proteins
- Goal: to be able to determine or predict the structure of every protein.
  - Direct determination - X-ray crystallography and nuclear magentic resonance (NMR).
  - Prediction
    - Comparative modeling -
    - Threading/Fold recognition
    - *Ab initio*

why structural proteomics?

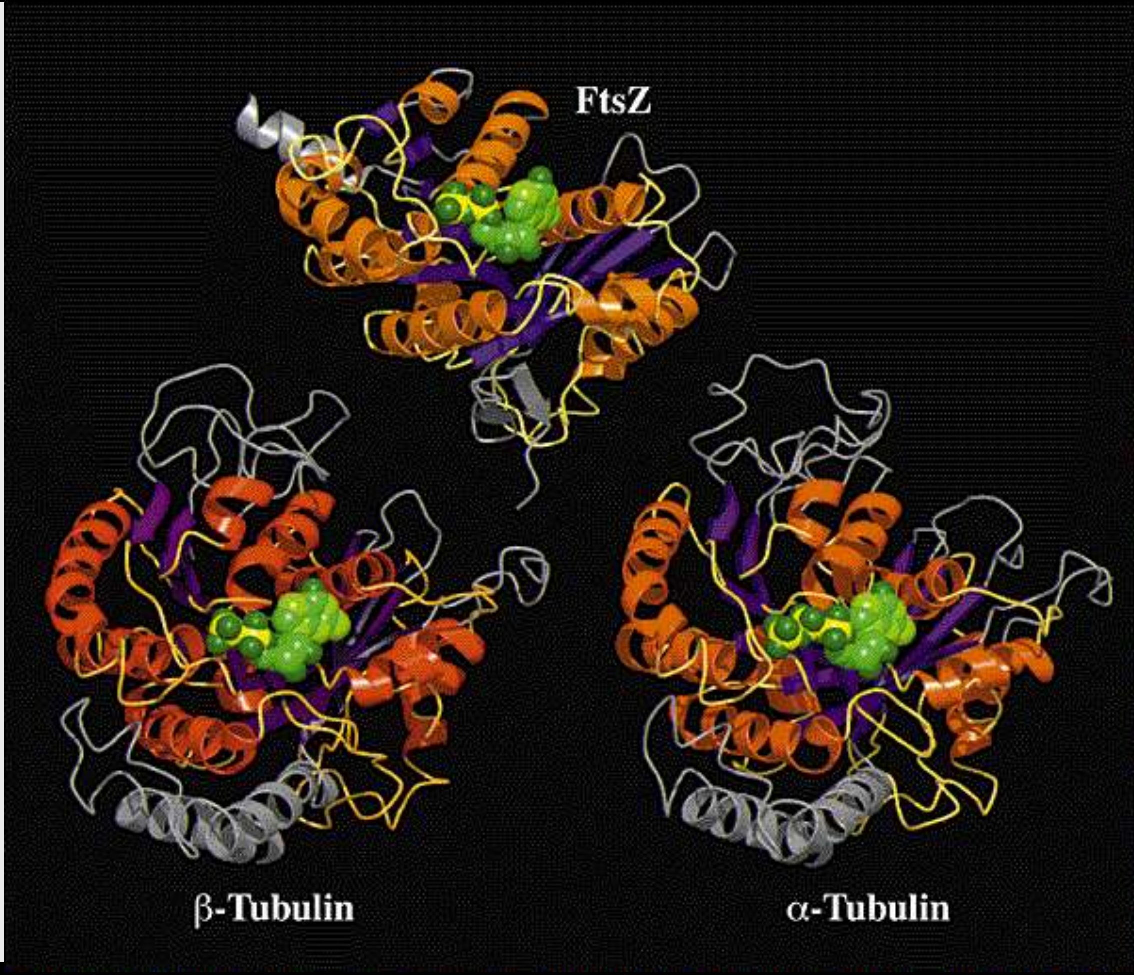
- To study proteins in their active conformation.
  - Study protein:drug interactions
  - Protein engineering
- Proteins that show little or no similarity at the primary sequence level can have strikingly similar structures.

an example

- FtsZ - protein required for cell division in prokaryotes, mitochondria, and chloroplasts.
- Tubulin - structural component of microtubules - important for intracellular trafficking and cell division.
- FtsZ and Tubulin have limited sequence similarity and would not be identified as homologous proteins by sequence analysis.

homologues

FtsZ and tubulin have little similarity at the amino acid sequence level



are FtsZ and tubulin homologues?

- Yes!

  Proteins that have conserved secondary structure can be derived from a common ancestor even if the primary sequence has diverged to the point that no similarity is detected.

bioMEMS

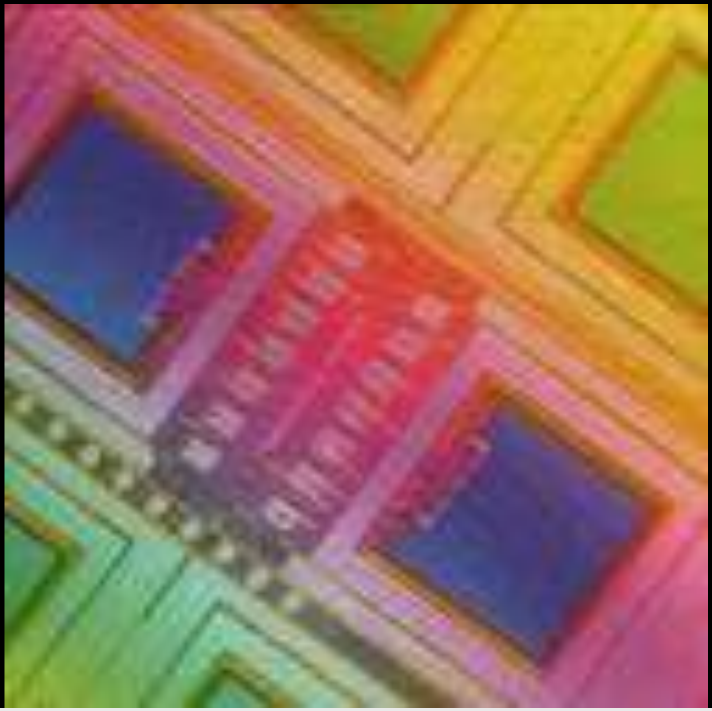
MEMS-Micro Electro Mechanical Systems

- definition
  the integration of mechanical elements, sensors, actuators, and electronics on a common silicon substrate through microfabrication technology" (memsnet.org)

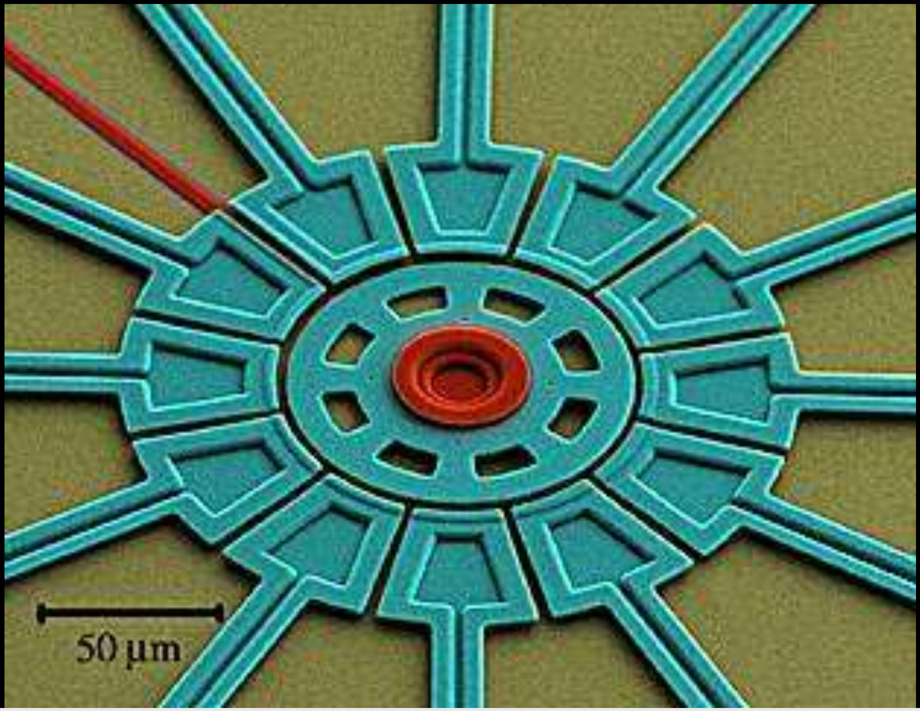
- micromachining
  the fabrication of devices with at least some of their dimensions in the micrometer range. (Madou, 2002)

- broad interpretation
  Microfabrication, micromanufacturing, microsystems technology, MEMS, BIOMEMS, MOEMS

MEMS-Micro Electro Mechanical Systems



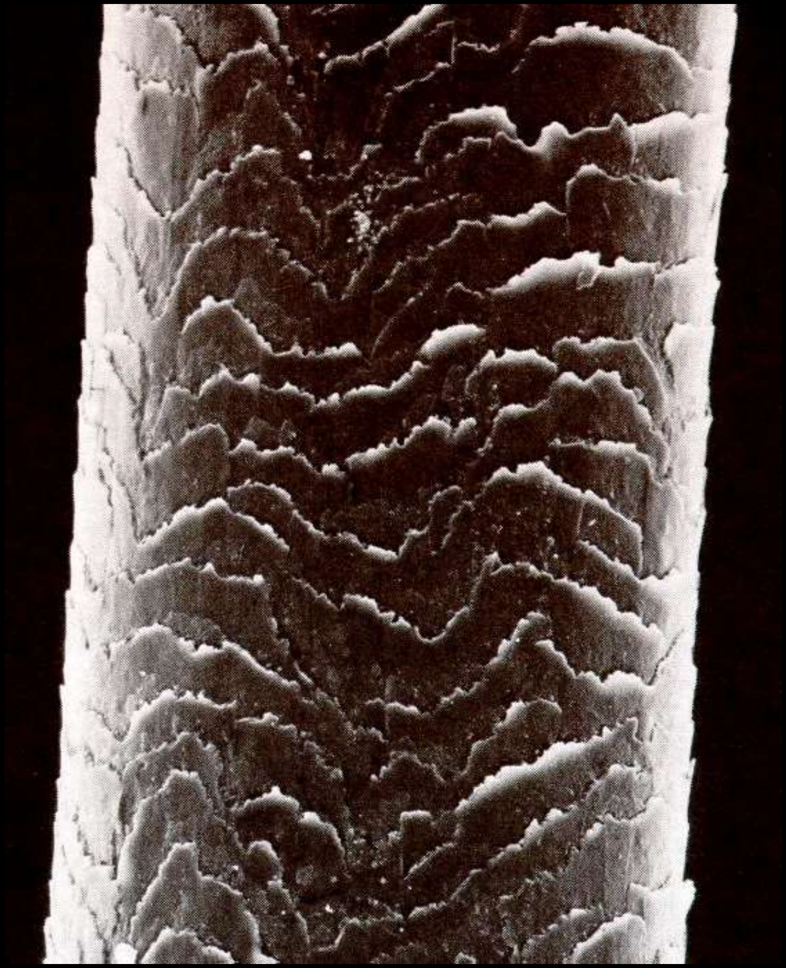
accelerometer



working micro-motor smaller than a human hair

sense of scale, 100 μm

sense of scale, 100 μm



DNA analysis chip

front

back

Photos courtesy of Nanogen, Inc.

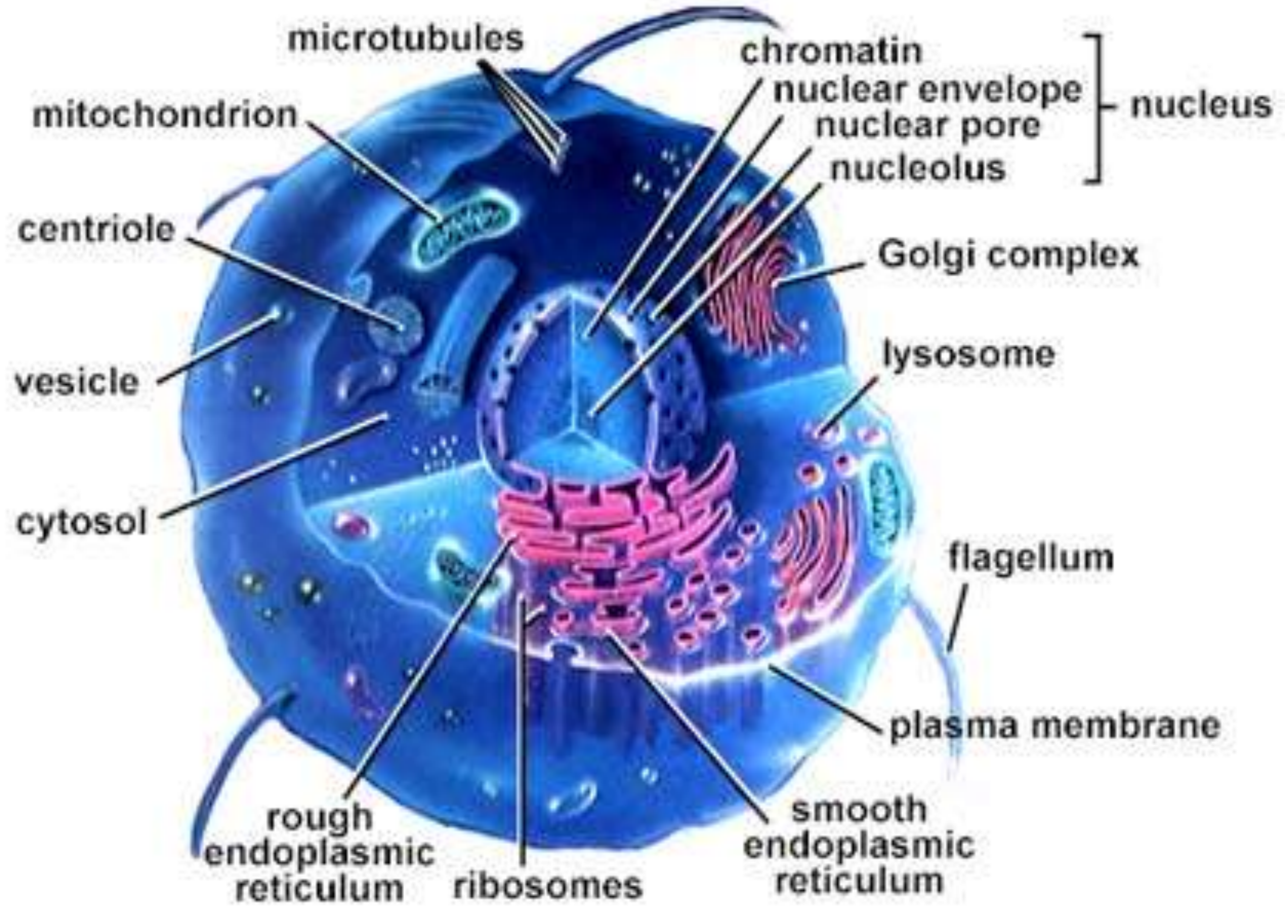
c e l l s   a n d   e - c e l l s

c e l l s

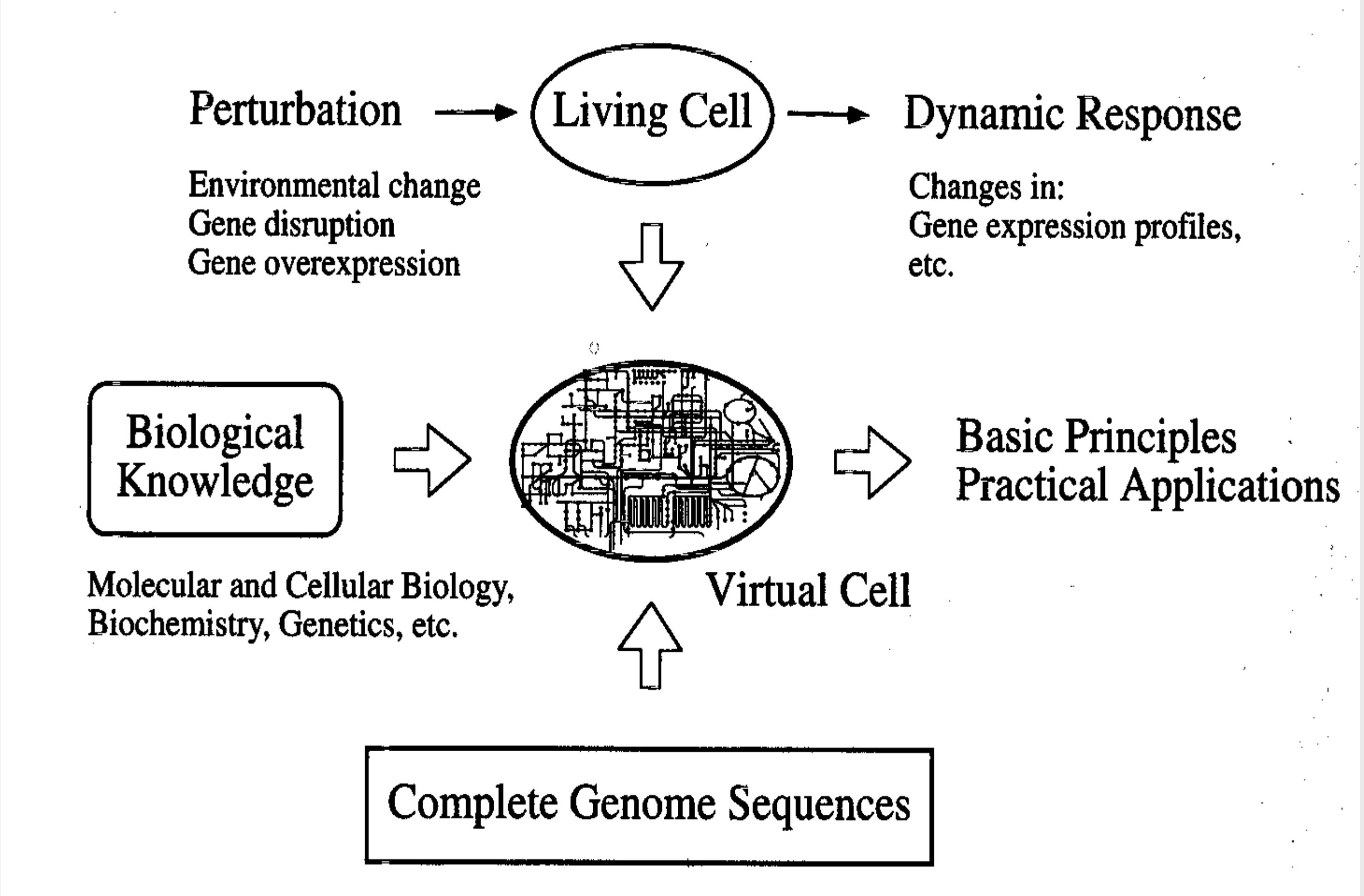
Humans

60 trillion cells

320 cell types



virtual cells

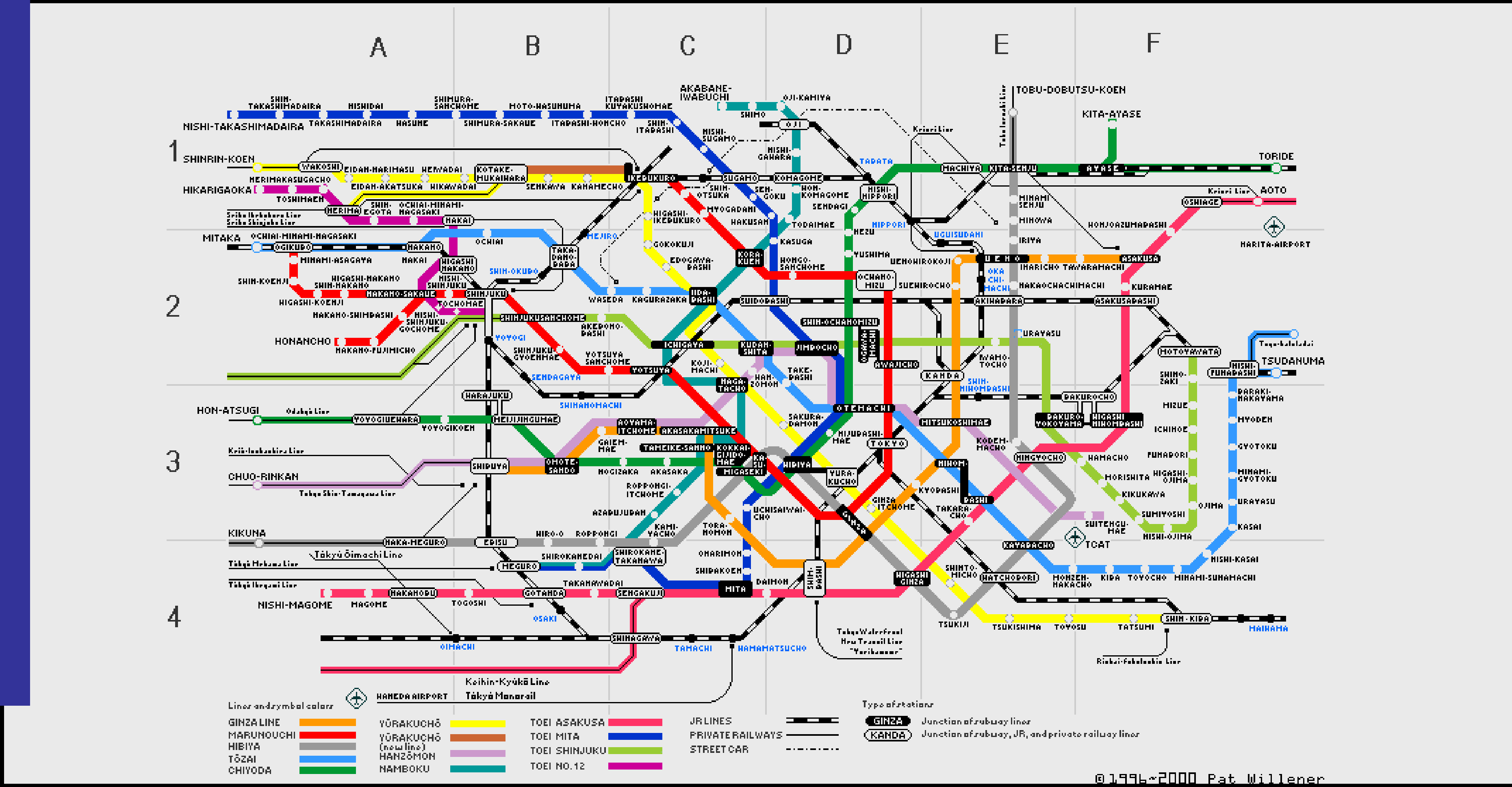


biological computers

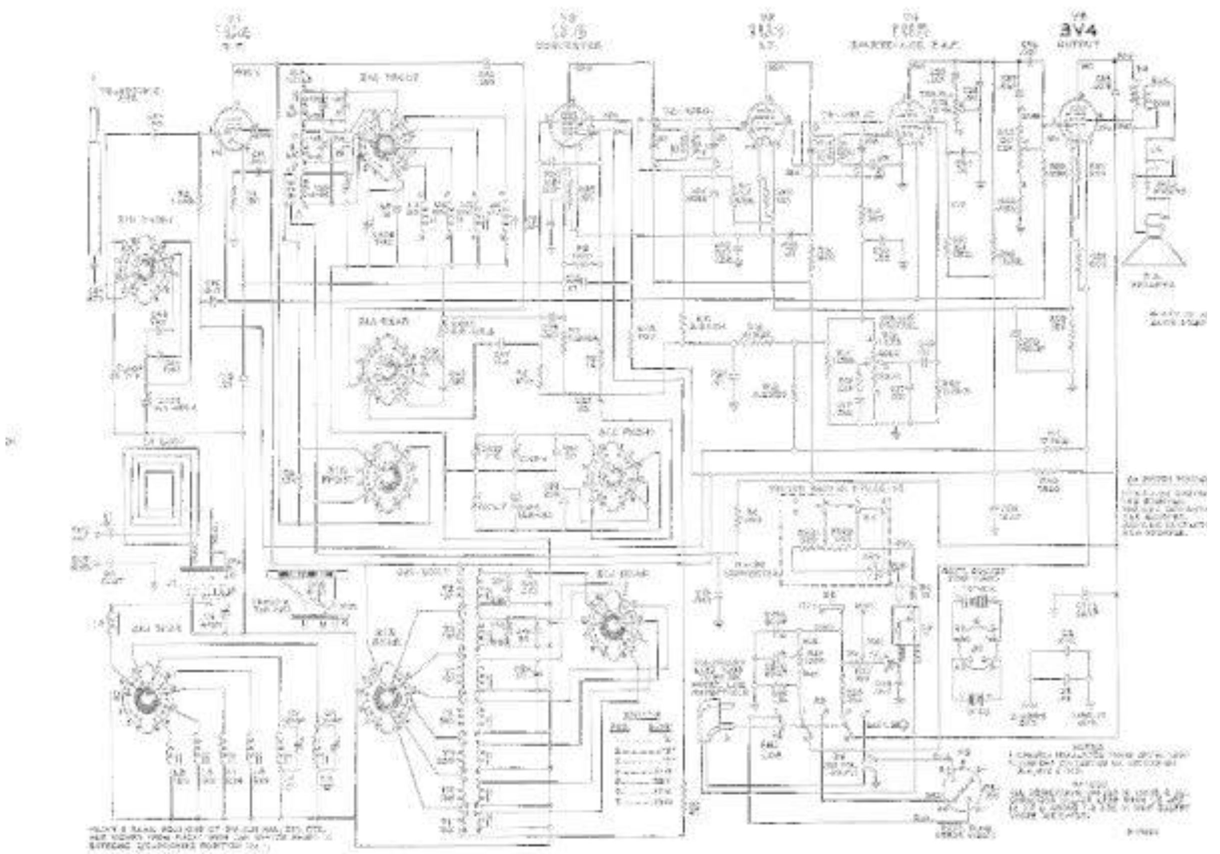
part 2

transcription and regulation

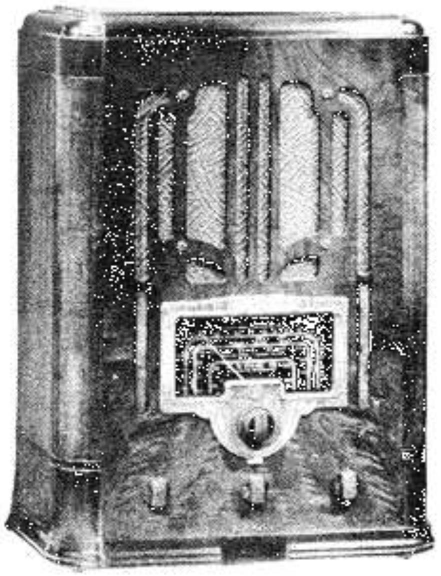
tokyo subway



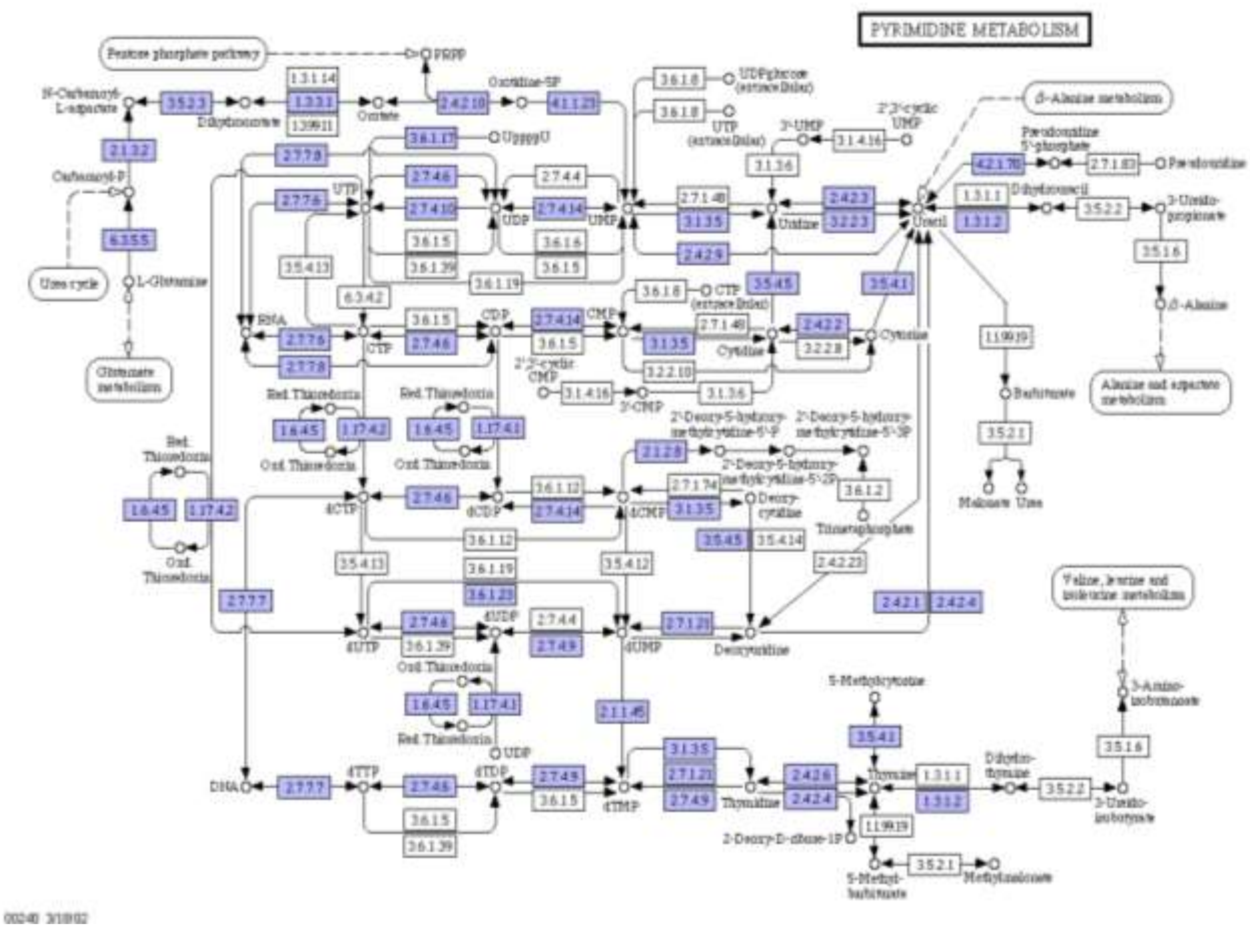
electronic pathway

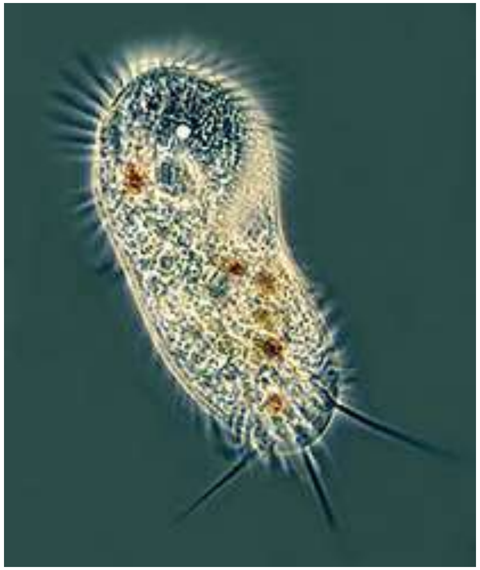


Schematic Diagram—Chassis No. RC-1125



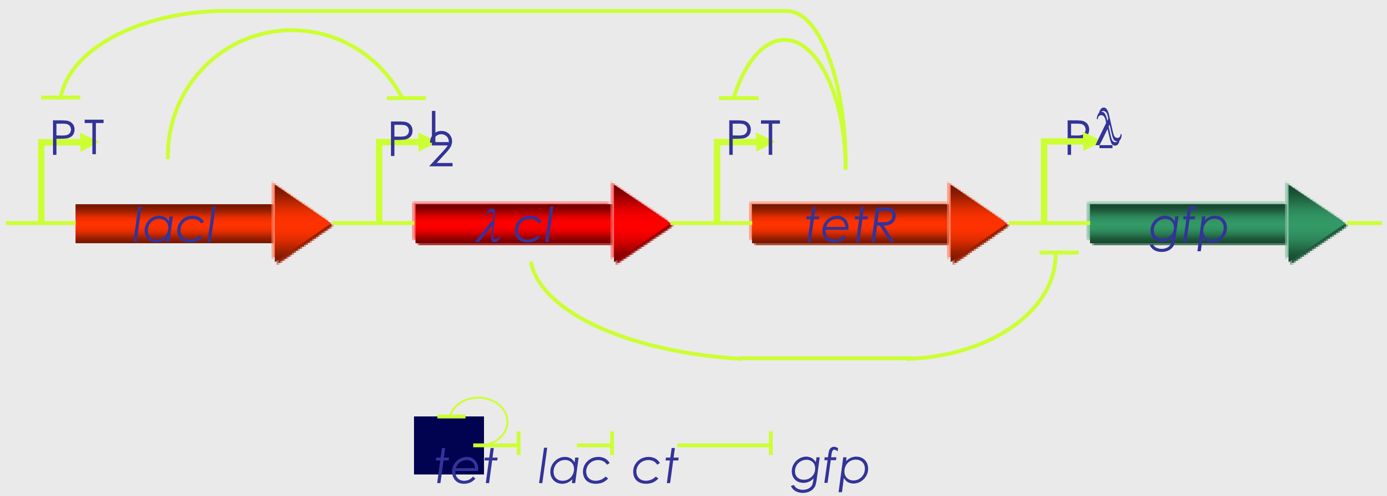
pyrimidine pathway





transcription regulators

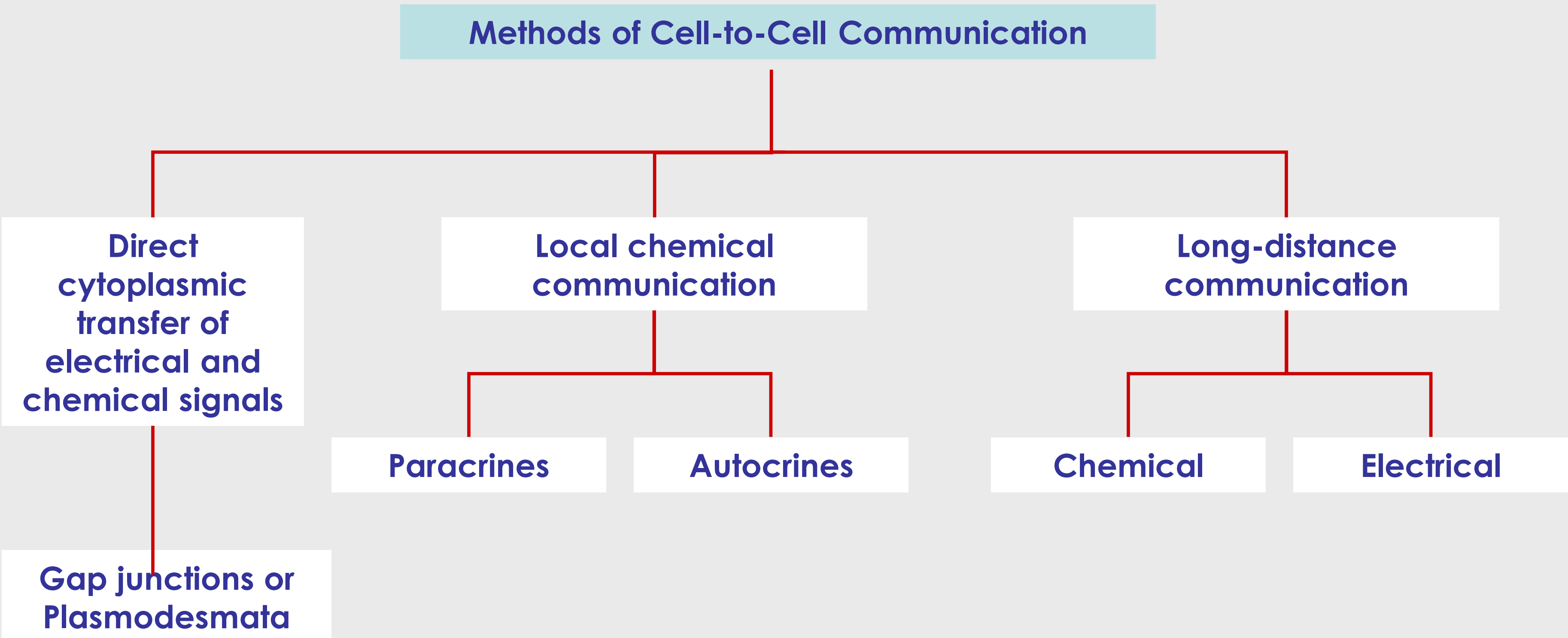
- lac⁻ strain CMW101
- three promoter genes: *lacI, λ cI, tetR*
- the binding state of lacI and tetR can be changed with IPTG (isopropyl β-D-thiogalactopyranoside) and aTc (anhydro-tetracycline).
- only signal when aTc but no IPTG



From Guet *et al.*, Science 24 May 2002

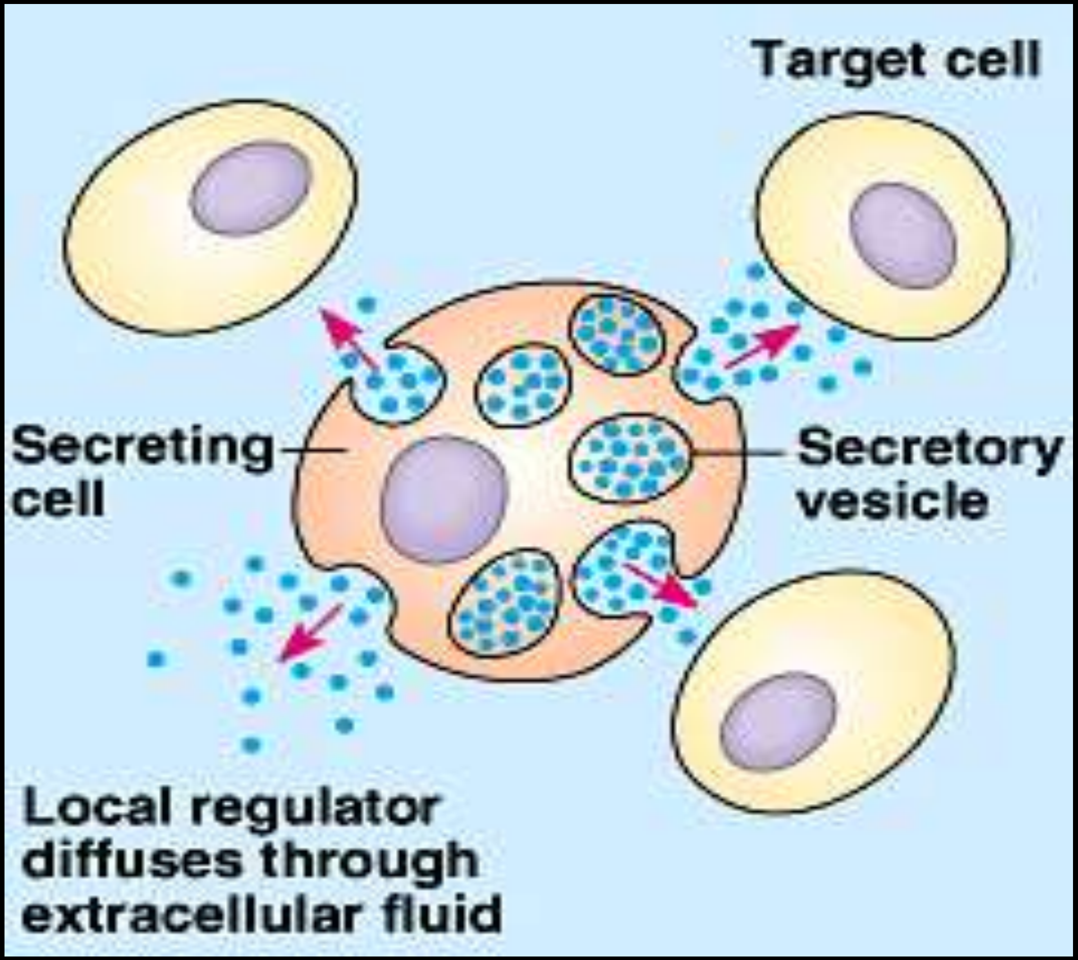
cell communication

cell-cell communication

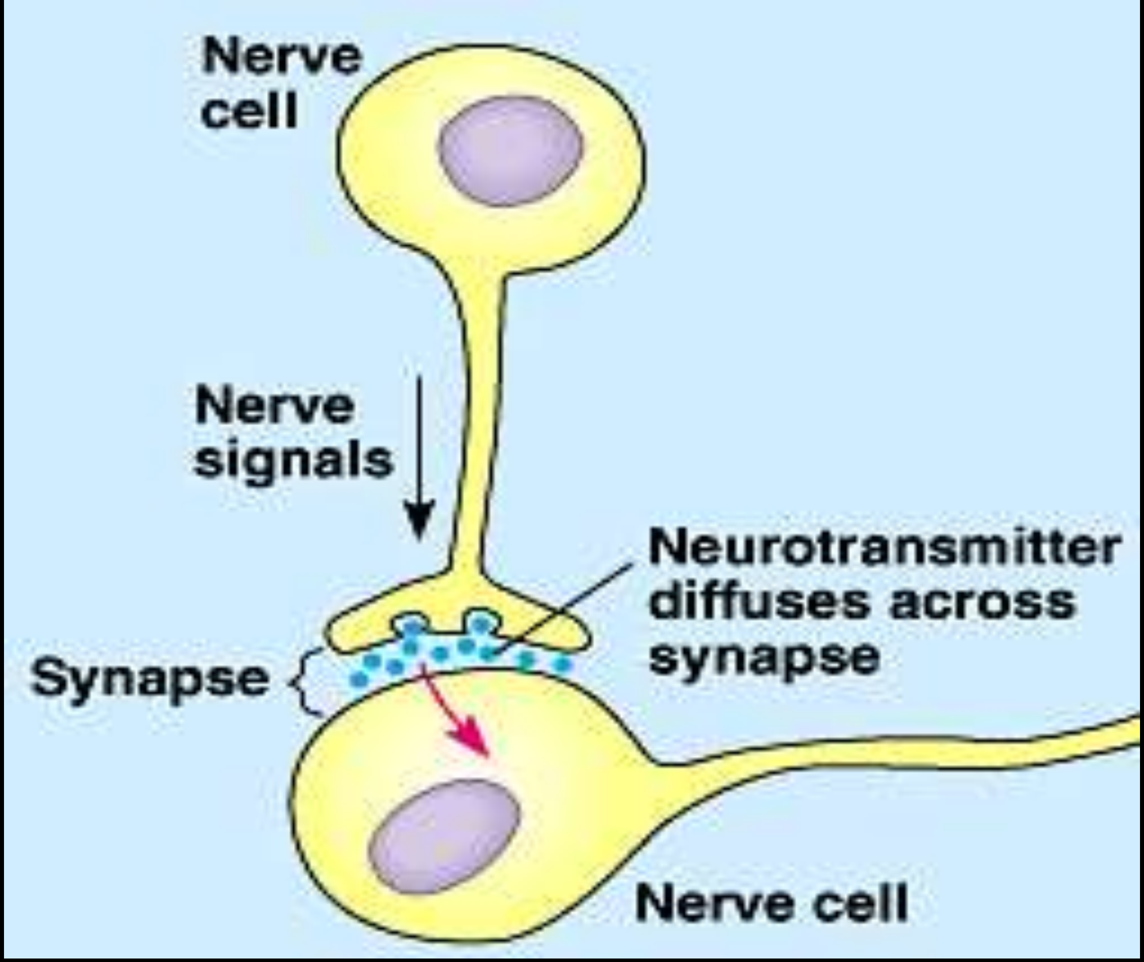


A given molecule can function as a signal by more than one method

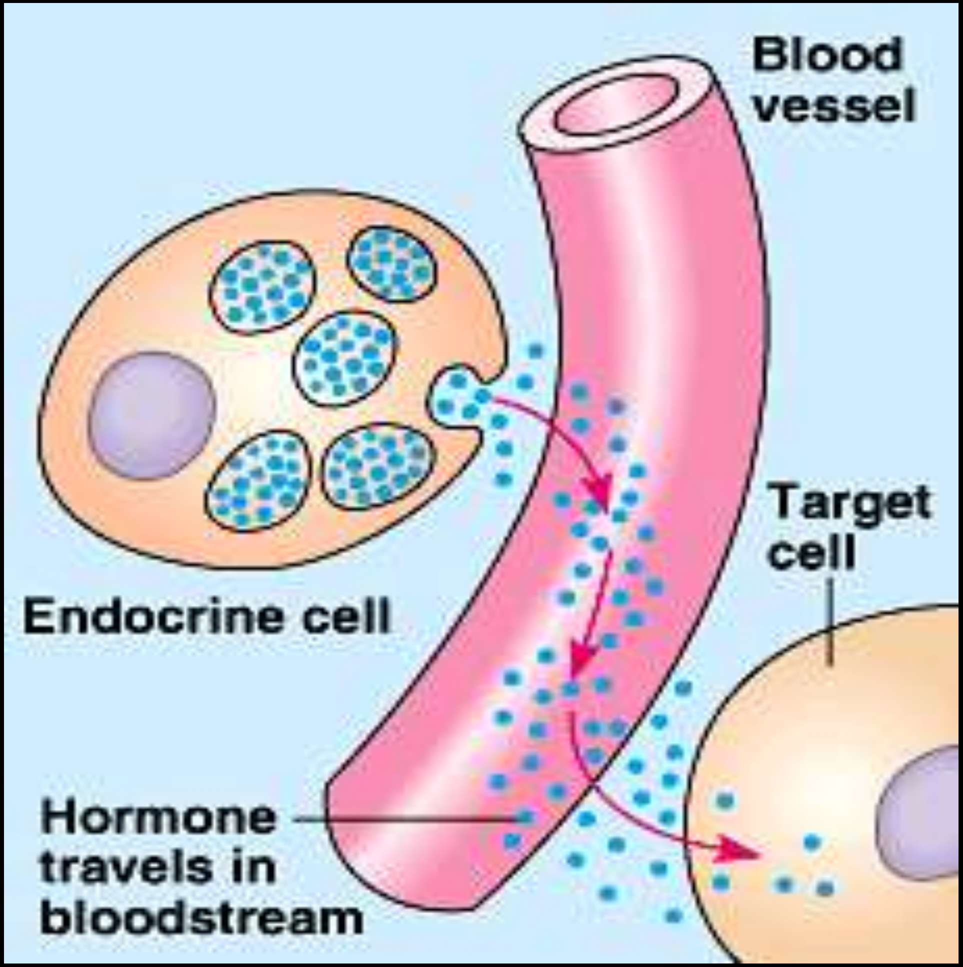
paracrine signaling



synaptic signaling



long distance (hormonal) signaling



q u o r u m   s e n s i n g

quorum sensing

- the ability of bacteria to sense and respond to environmental stimuli such as pH, temperature, the presence of nutrients, etc has been long recognized as essential for their continued survival
- it is now apparent that many bacteria can also sense and respond to signals expressed by other bacteria
- quorum sensing is the regulation of gene expression in response to cell density and is used by Gram positive and Gram negative bacteria to regulate a variety of physiological functions
- it involves the production and detection of extracellular signaling molecules called auto-inducers

quorum sensing

Vibrio fischeri is a specific bacterial symbiont with the squid *Euprymna scolopes* and grows in its light organ

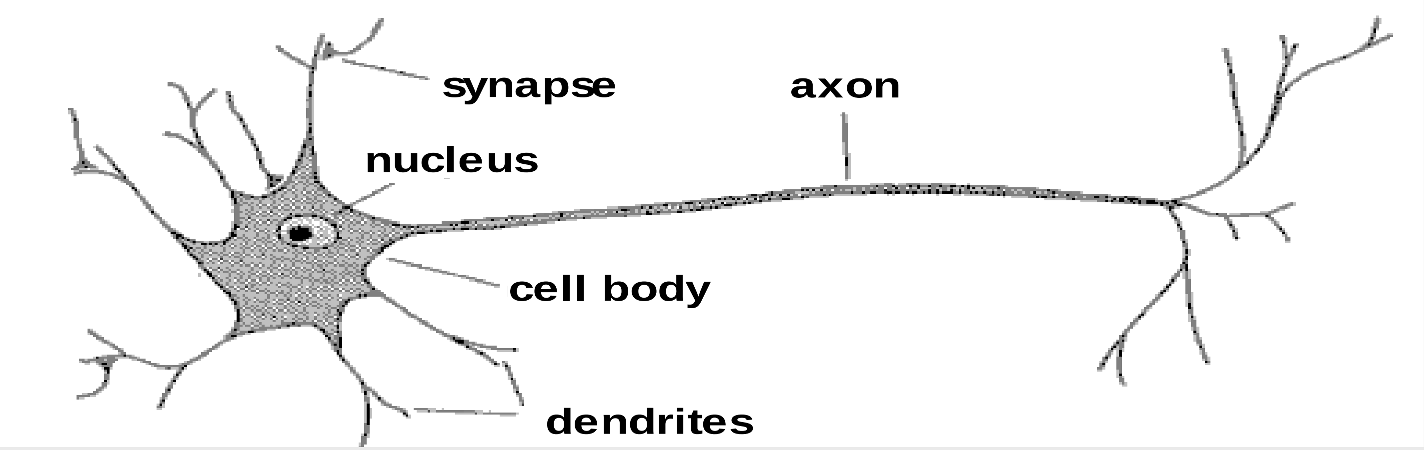
quorum sensing

- the squid cultivates a high density of cells in its light organ, thus allowing the auto-inducer to accumulate to a threshold concentration
- at this point, auto-inducer combines with the gene product luxR to stimulate the expression of the genes for luciferase, triggering maximal light production
- studies have shown that hatchling squid fail to enlarge the pouches that become the fully developed organ when raised in sterile seawater

quorum sensing

neural networks

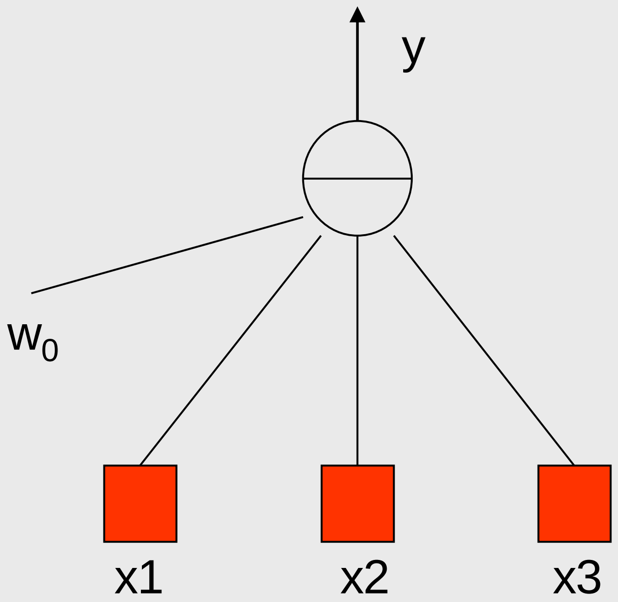
biological neuron



- A neuron has
  - A branching input (dendrites)
  - A branching output (the axon)
- The information circulates from the dendrites to the axon via the cell body
- Axon connects to dendrites via synapses
  - Synapses vary in strength
  - Synapses may be excitatory or inhibitory

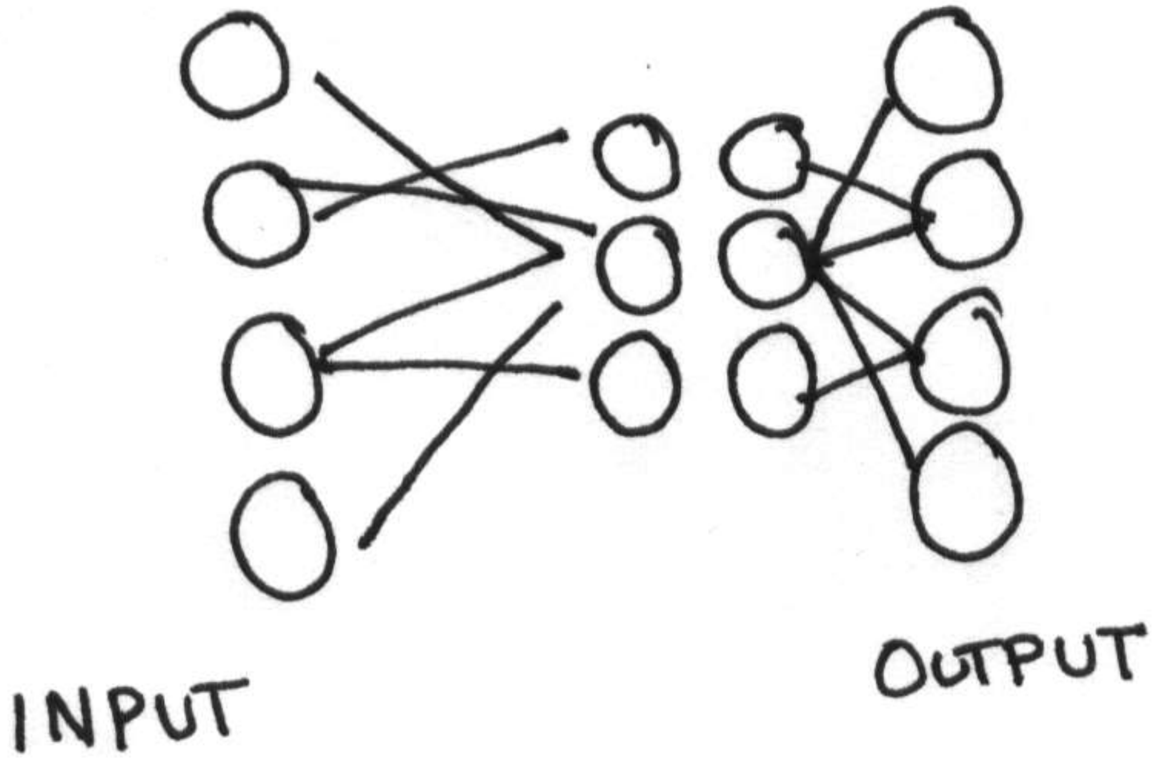
artificial neuron

- Definition: non linear, parameterized function with restricted output range



$$y = f\left(w_0 + \sum_{i=1}^{n-1} w_i x_i \right)$$

artificial neural networks



computer verus brain

a computer

Clock freq. - ~ Gigahertz (10^9 per s)

Memory - ~ Gigabytes (10^{10} bits)

Sync. and sharing problems

Very strong with formal problems,
Very weak in informal problems

One 'heart' – the CPU

our brain

Switching rate – 1000 per sec.

Number of neurons - ~ 10^{13}

Connectivity - ~10^{4-5}

Image recognition - ~ 0.1 sec.

Very parallel

DNA computing

what is DNA computing?

- The field of DNA computing is concerned with the possibility of performing computations using biological molecules.

- It is also concerned with understanding how complex biological molecules process information in an attempt to gain insight into new models of computation.

- Cells and nature compute by reading and rewriting DNA by processes that modify sequence at the DNA or RNA level. DNA computing is interested in applying computer science methods and models to understand such biological phenomena and gain insight into early molecular evolution and the original of biological information processing.

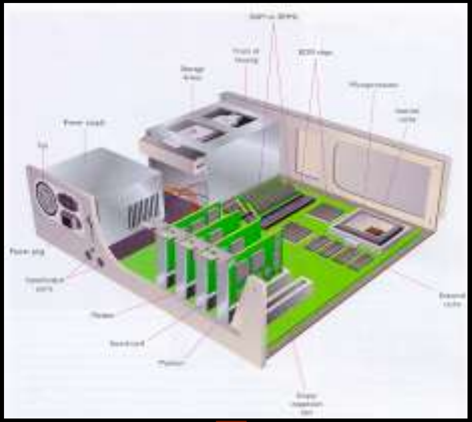
what is DNA computing?

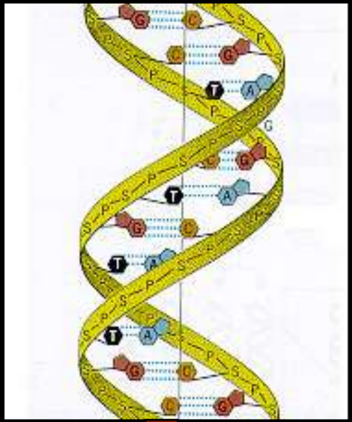
Molecular electronics, Theoretical biology, Evolutionary biology, Emergent computation, Brain sciences, Organic chemistry, Biomimetic engineering, Parallel processing, Distributed computing, Behavioural ecology, Cytology, Discrete mathematics, Optimisation theory, Artificial Intelligence, Cognitive science, Botany, Psychology, Algorithmics, Clinical engineering, Biophysics, Connectionism, Integrative physiology, Technology transfer, Selectionism, Immunology, Automata theory, Evolutionary computation, Simulation of computational systems, Histology, Ethology, Medical computing, Signal transduction and processing, Cellular automata, Electronic engineering, Vision, Object oriented design, Philosophy of science, VLSI, Non-linear dynamical systems, Game theory, Communication, Bioengineering, Self-organisation, Biochemistry, Pattern recognition, Information theory, Machine learning, Biosystem simulation, Genetics, Mathematical biology, Microbiology, Zoology, Science education, Physiology, Systems theory, Biosensors, Analogue devices and sensors, Microtechnology, Robotics ...

what is DNA computing?

Molecular electronics, Theoretical biology, Evolutionary biology, Emergent computation, Brain sciences, Organic chemistry, Biomimetic engineering, Parallel processing, Distributed computing, Behavioural ecology, Cytology, Discrete mathematics, Optimisation theory, Artificial Intelligence, Cognitive science, Botany, Psychology, Algorithmics, Clinical engineering, Biophysics, Connectionism, Integrative physiology, Technology transfer, Selectionism, Immunology, Automata theory, Evolutionary computation, Simulation of computational systems, Histology, Ethology, Medical computing, Signal transduction and processing, Cellular automata, Electronic engineering, Vision, Object oriented design, Philosophy of science, VLSI, Non-linear dynamical systems, Game theory, Communication, Bioengineering, Self-organisation, Biochemistry, Pattern recognition, Information theory, Machine learning, Biosystem simulation, Genetics, Mathematical biology, Microbiology, Zoology, Science education, Physiology, Systems theory, Biosensors, Analogue devices and sensors, Microtechnology, Robotics ...

what is DNA computing?





011001101010001

ATGCTCGAAGCT

what is DNA computing?

- not to confuse with bio-computing which applies biological laws (evolution, selection) to computer algorithm design.

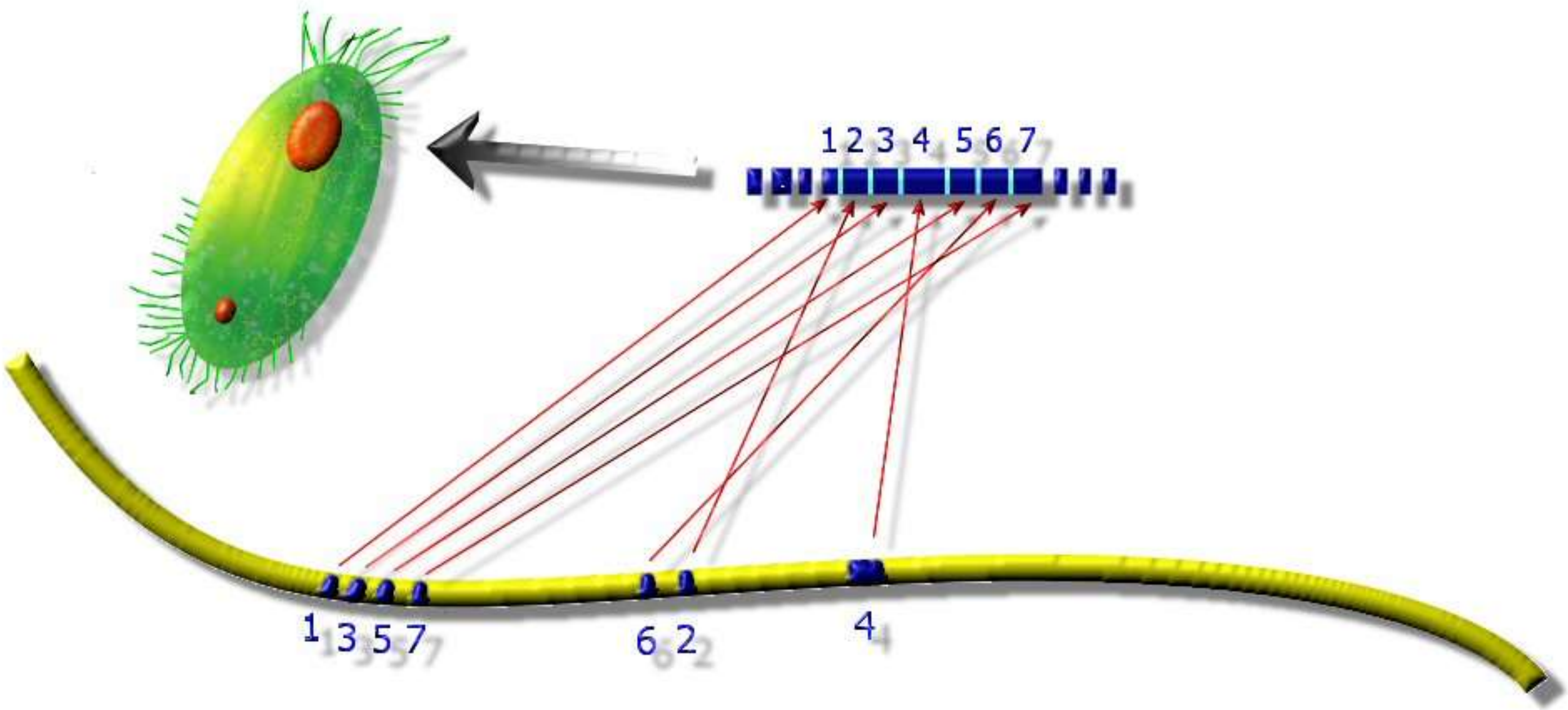
DNA vs. conventional computers

DNA-based computers	Microchip-based computers
slow at individual operations	fast at individual operations
can do billions of operations simultaneously	can do substantially fewer operations simultaneously
can provide huge memory in small space	smaller memory
setting up a problem may involve considerable preparations	setting up only requires keyboard input
DNA is sensitive to chemical deterioration	electronic data are vulnerable but can be backed up easily

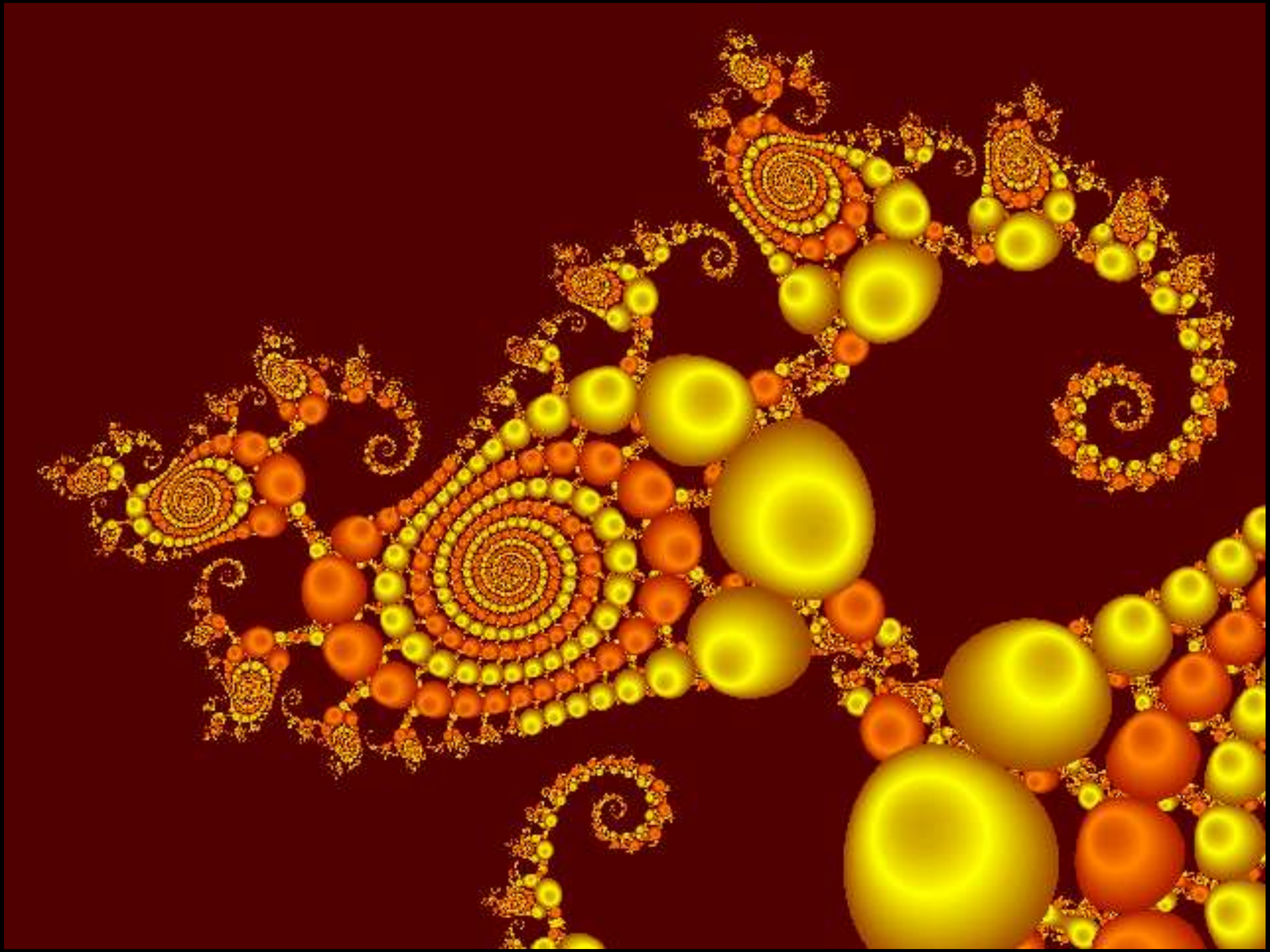
computing in biology

- Cells and nature compute by reading and rewriting DNA by processes that modify sequence at the DNA or RNA level. DNA computing is interested in applying computer science methods and models to understand biological phenomena and gain insight into early molecular evolution and the origin of biological information processing.

computing in biology



fractals and patterns



where are fractals found?

- star distribution in the galaxy
- rings of Saturn
- weather patterns
- trees
- geological formations
- vascular networks
- bioelectrical activity
- dendritic branching patterns

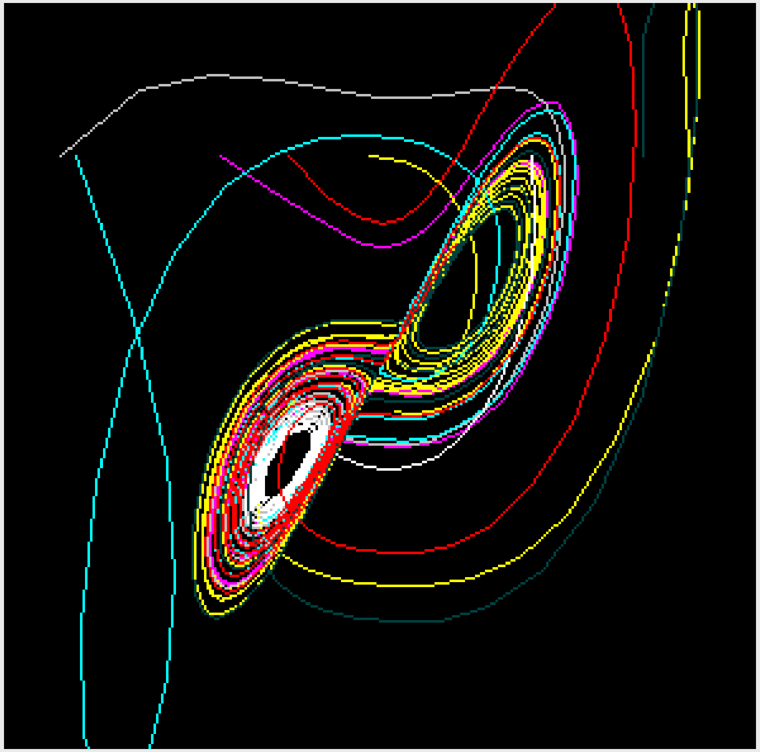
rings of saturn

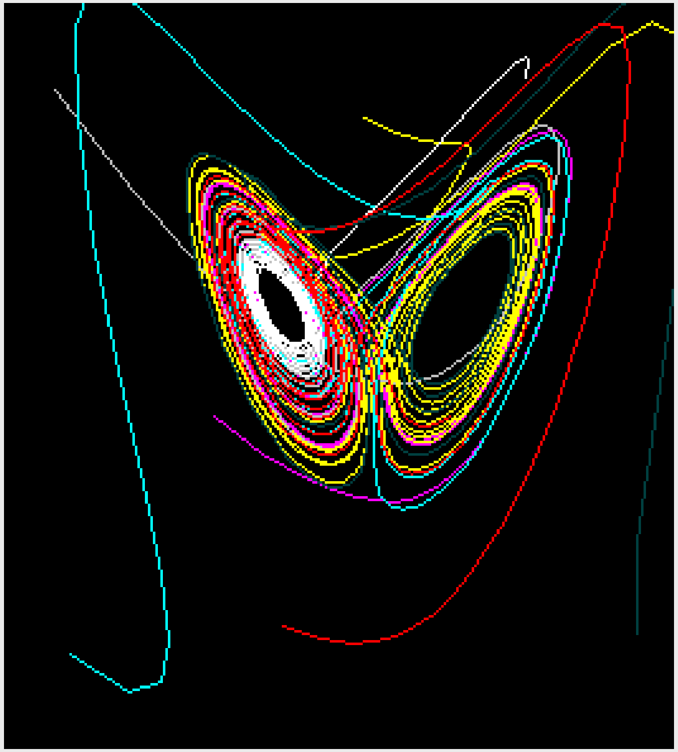
fractals in biology

- Time series - EMG, ECG, EEG
- Codes - DNA
- Population distribution / Urban expansion
- Gait analysis
- Vessel distribution
  - Diabetic retinopathy, lung bronchioli
- Pathology
- Classification of images
- Neurophysiology

what is a fractal?

- A snapshot of a chaotic process
  - The Lorenz Attractor

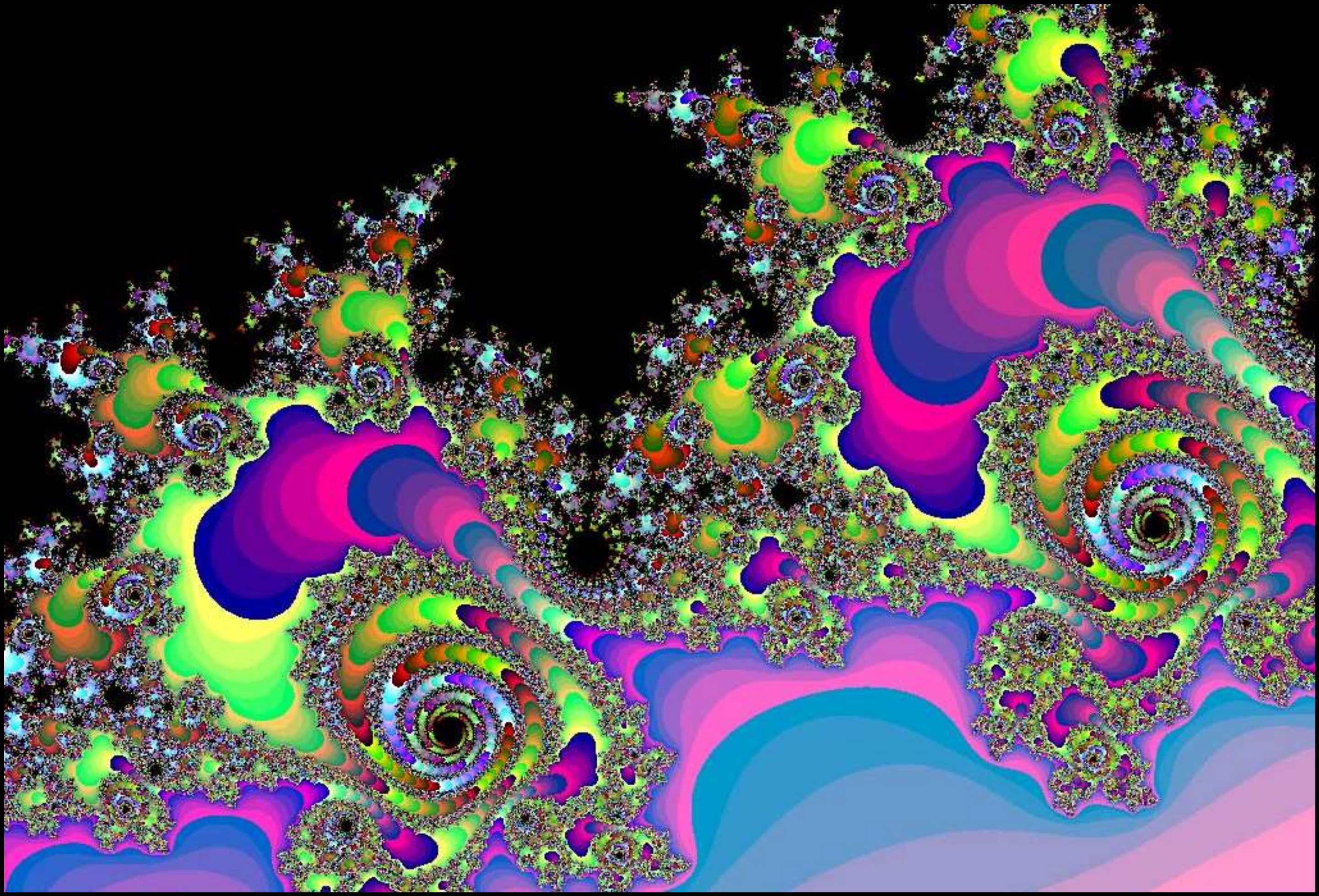




definition of fractals

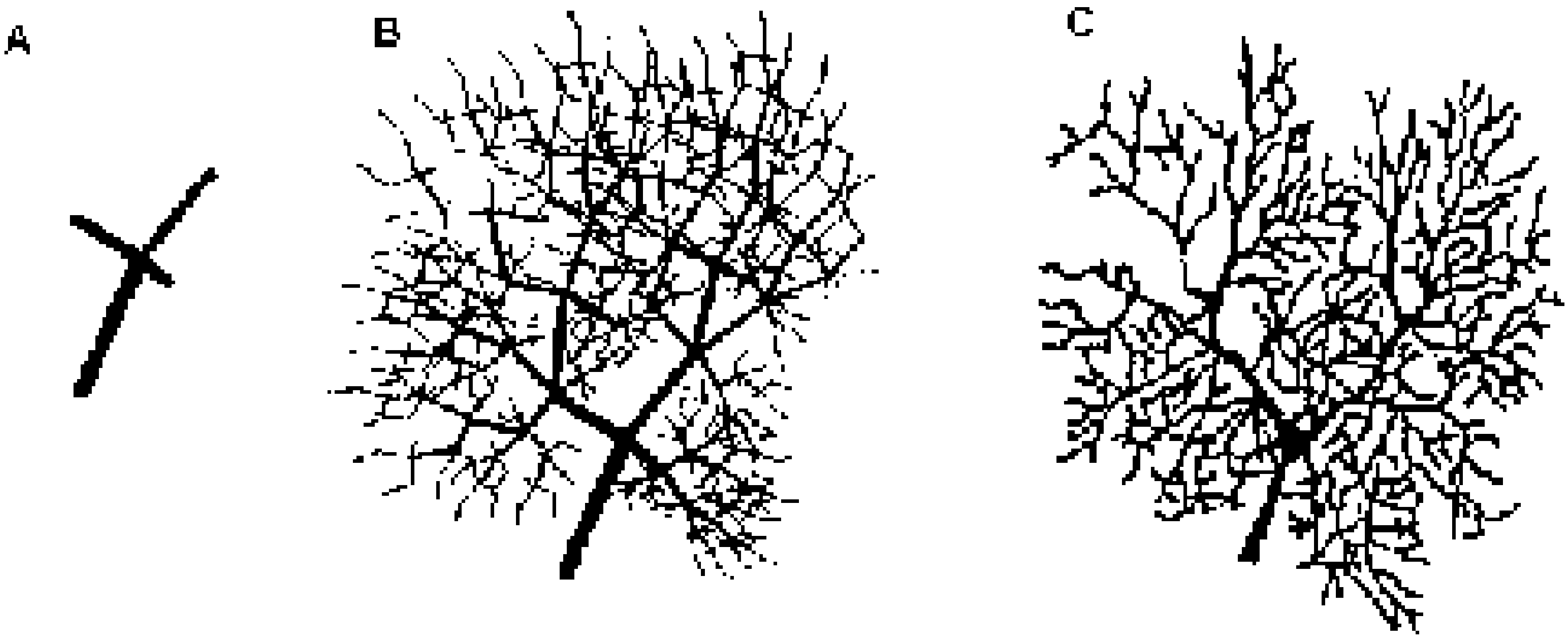
- Mandelbrot's definition
  - A fractal set is a metric space for which the Hausdorff-Besicovitch dimension D is greater than the topological dimension D_T
- Fractals for the layperson
  - Non Euclidean forms that are not easily described by Euclidean geometry
  - A fractal set is characterised by an unlimited process of repeated transformations of an invariant geometrical form.

m a n d l e b r o t



m a n d l e b r o t

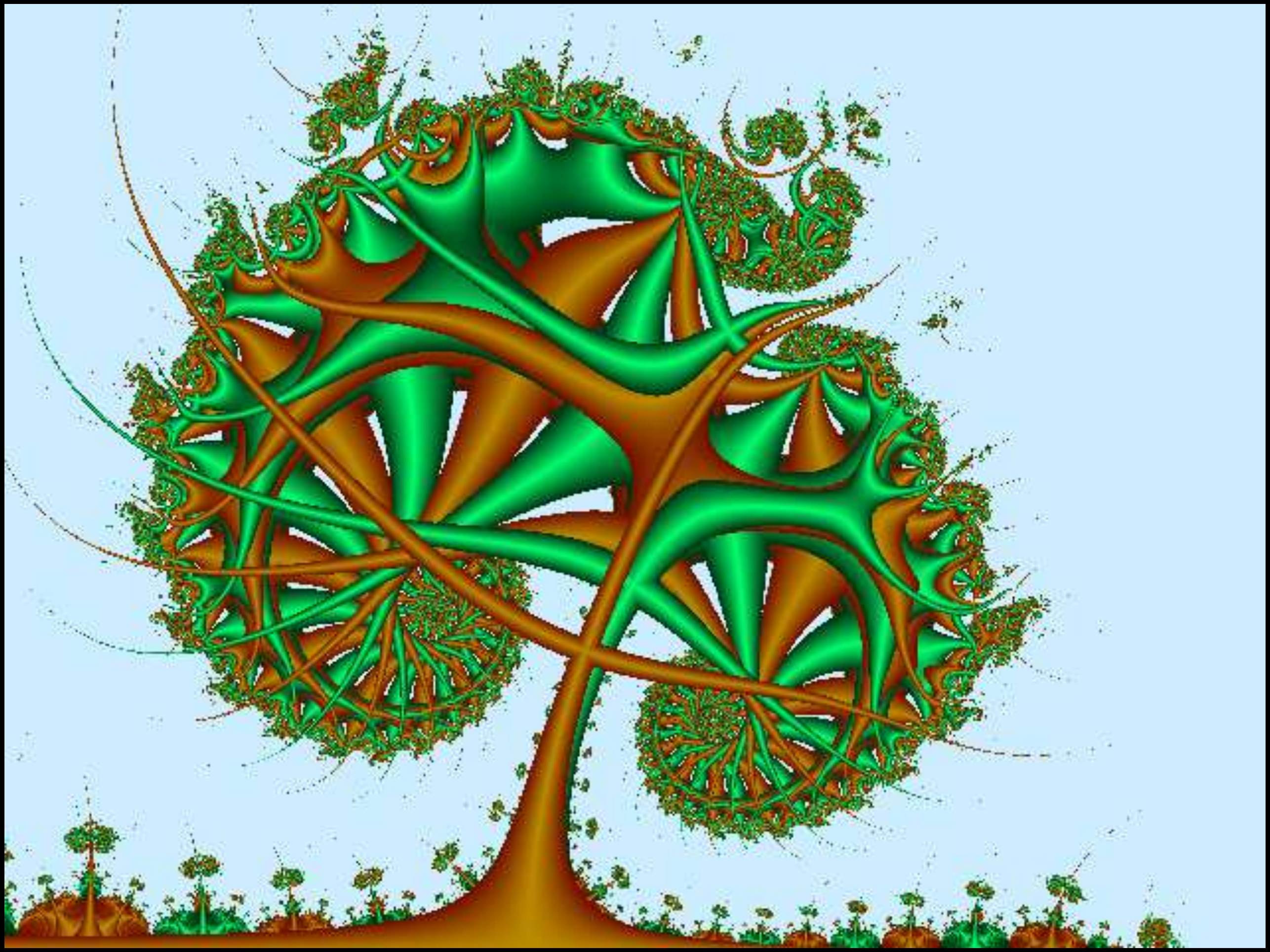
f r a c t a l   n e u r o n



Purkinje cell construction using recursive growth rule with fractal scaling

self similarity

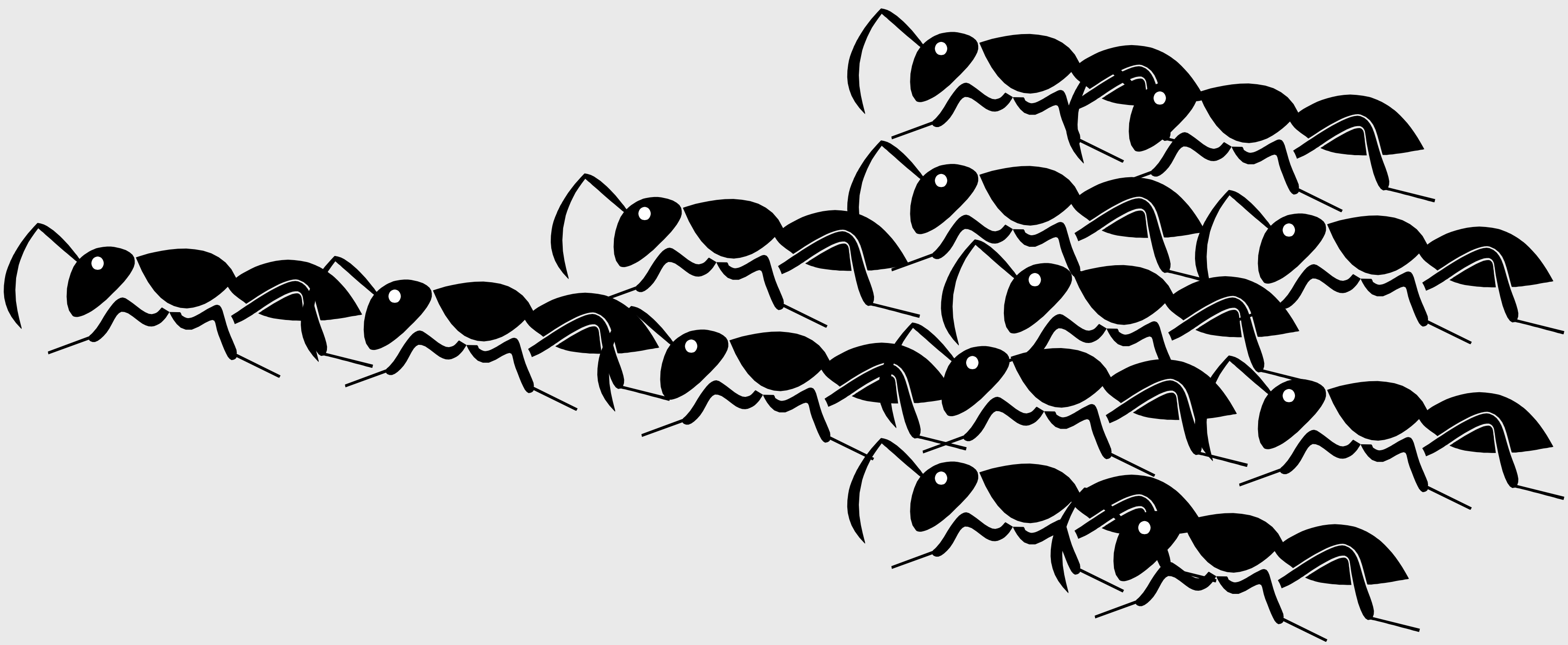
fractals in biology



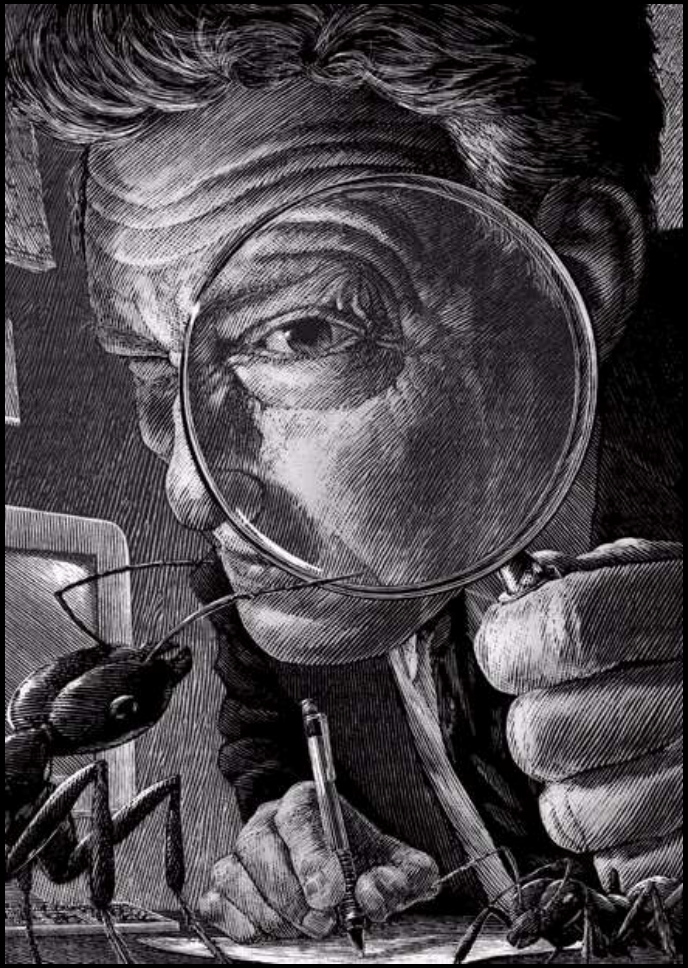
the birds and the bees

.... and ants

do ants follow rules?



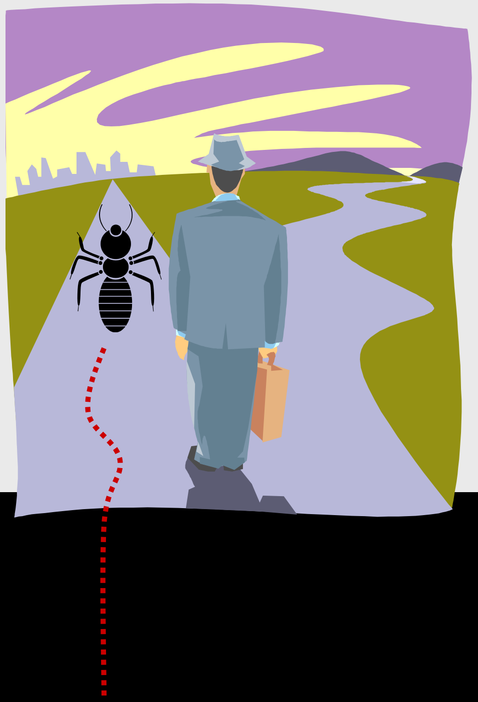
ants were closely studied…



ants find the shortest path

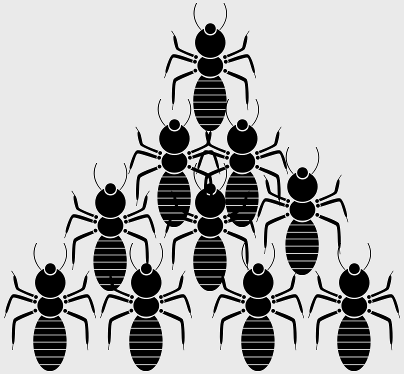
Ants find the most efficient routes to a food source.

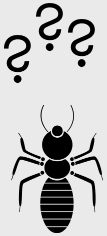
- ants finding the shorter path return sooner.
- ants following the shorter path reinforce the odor – pheromones – on the shorter path
- ants communicate even shorter paths to groups
- or individuals in the ant colony, as the pheromones dissipate.



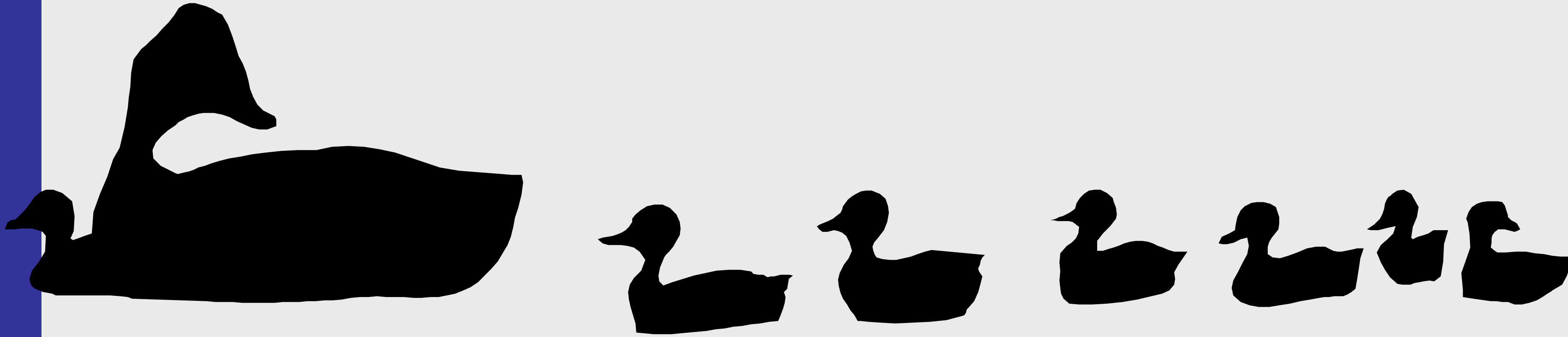
b u t   s i n g l e   a n t s ?

- a single ant acts randomly...
- a colony of ants provides sustenance and defensive protection for the entire population...
  - the ants combine to produce an effect that is much more than the sum of the parts.
  - again global effects from local interactions.

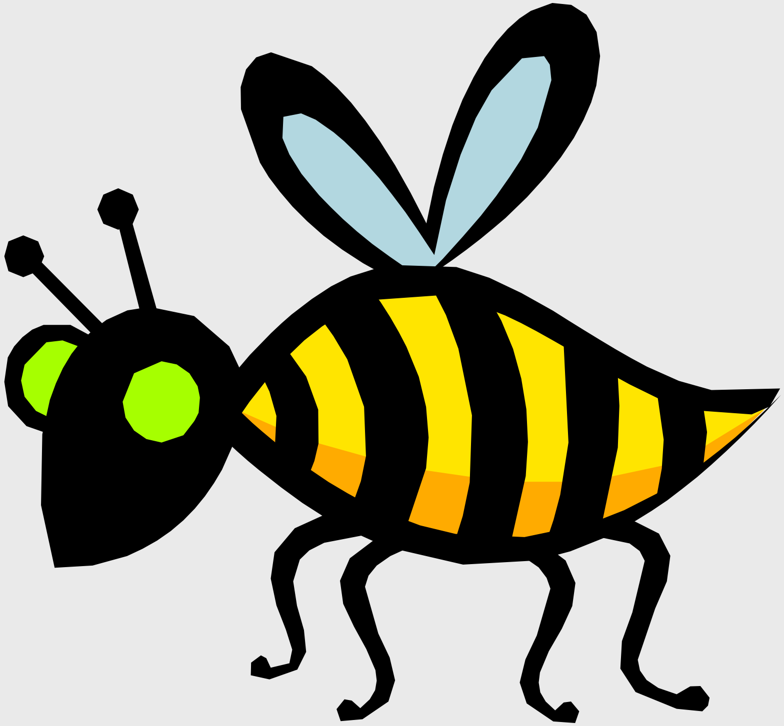




how about the birds?



…and the bees?



see you next time