

What is Systems Biology? Where does it come from ? What are the challenges ahead?

Nicolas Le Novère, Babraham Institute, EMBL-EBI

RESEARCH ARTICLE

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,¹ John I. Glass,¹ Carole Lartigue,¹ Vladimir N. Noskov,¹ Ray-Yuan Chuang,¹ Mikkel A. Algire,¹ Gwynedd A. Benders,² Michael G. Montague,¹ Li Ma,¹ Monzia M. Moodie,¹ Chuck Merryman,¹ Sanjay Vashee,¹ Radha Krishnakumar,¹ Nacyra Assad-Garcia,¹ Cynthia Andrews-Pfannkoch,¹ Evgeniya A. Denisova,¹ Lei Young,¹ Zhi-Qing Qi,¹ Thomas H. Segall-Shapiro,¹ Christopher H. Calvey,¹ Prashanth P. Parmar,¹ Clyde A. Hutchison III,² Hamilton O. Smith,² J. Craig Venter^{1,2*}

2 JULY 2010 VOL 329 SCIENCE www.sciencemag.org

Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors

Cell

Kazutoshi Takahashi¹ and Shinya Yamanaka^{1,2,*}

¹Department of Stem Cell Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan

²CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan

*Contact: yamanaka@frontier.kyoto-u.ac.jp

DOI 10.1016/j.cell.2006.07.024

Cell 126, 663–676, August 25, 2006 ©2006 Elsevier Inc. 663



EXTREME GENETIC ENGINEERING

An Introduction to Synthetic Biology

January 2007

etc group

A synthetic oscillatory network of transcriptional regulators

Michael B. Elowitz & Stanislas Leibler

Departments of Molecular Biology and Physics, Princeton University, Princeton, New Jersey 08544, USA

NATURE | VOL 403 | 20 JANUARY 2000 | www.nature.com

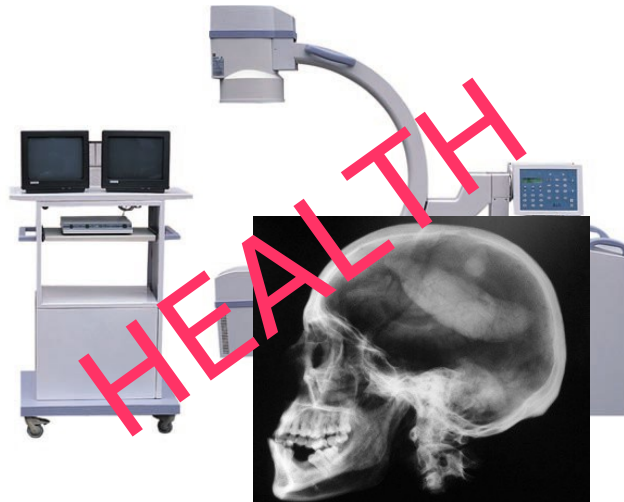
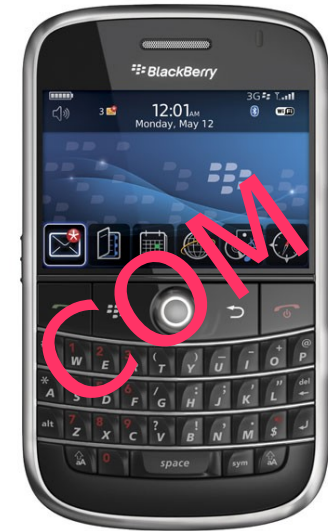


page discussion view source history teams

About

The **International Genetically Engineered Machine competition (iGEM)** is a Biology competition. Student teams are given a kit of biological parts at the beginning of the year. Working at their own schools over the summer, they use the

What did electrical engineering bring?



What will biological engineering bring?

HEALTH



FOOD



ENVIRONMENT

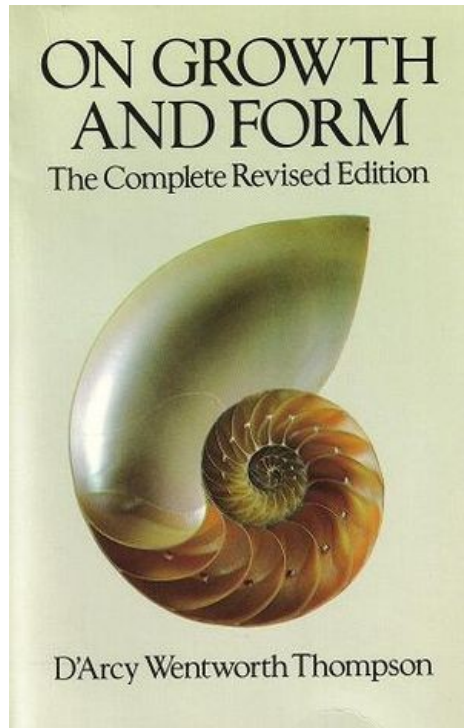


SECURITY



Why using mathematical models?

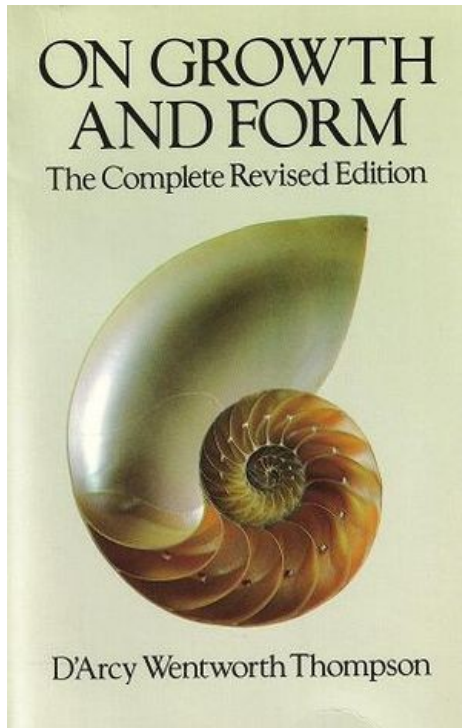
Describe



1917

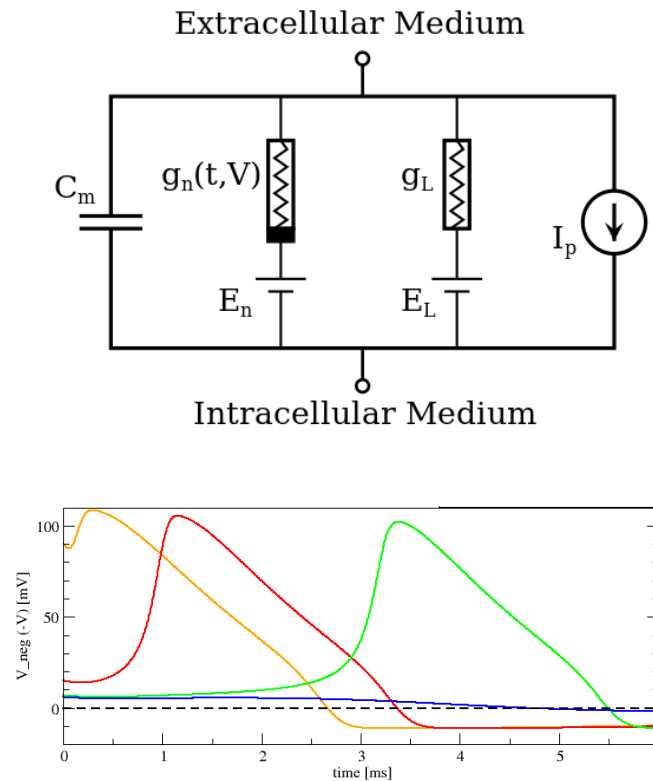
Why using mathematical models?

Describe



1917

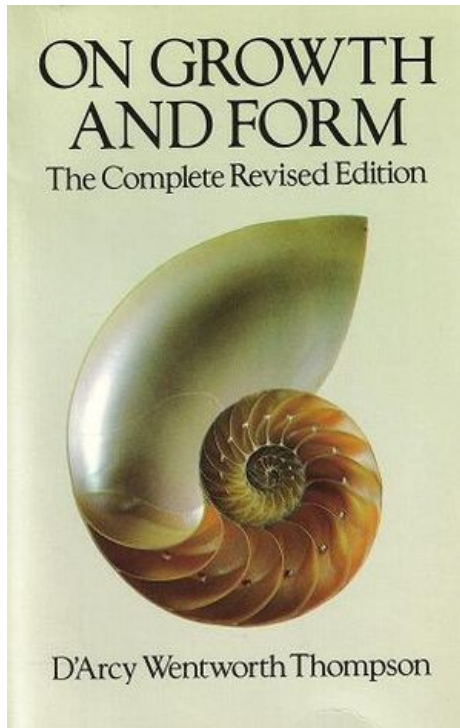
Explain



1952

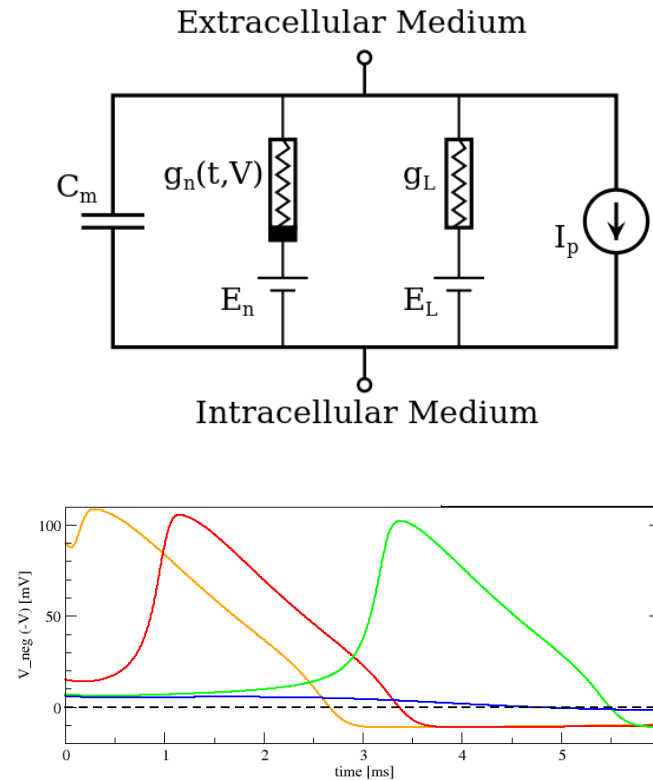
Why using mathematical models?

Describe



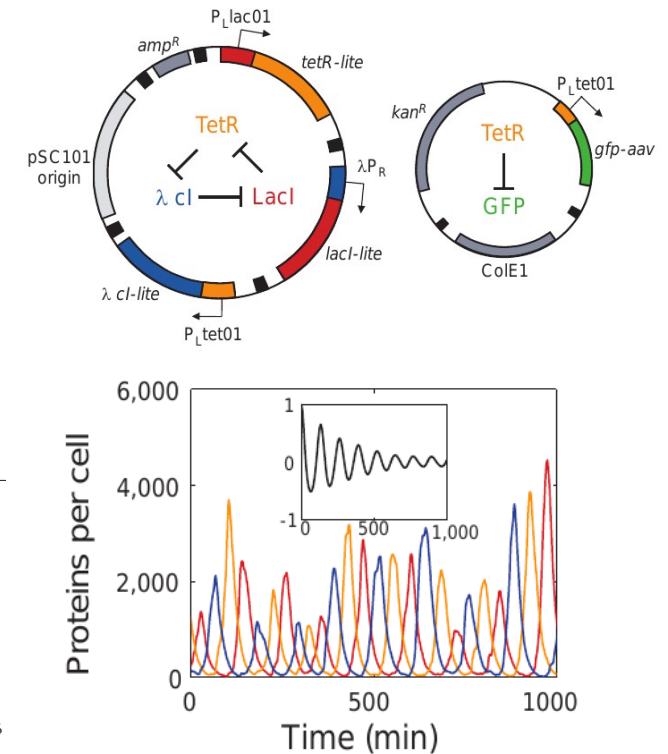
1917

Explain



1952

Predict



2000

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a **system** using **mathematical** concepts and language.”

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a **system** using **mathematical** concepts and language.”

variables

$[x]$

V_{\max}

K_d

EC_{50}

length

$t_{1/2}$

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a **system** using **mathematical** concepts and language.”

variables

$[x]$

V_{\max}

K_d

EC_{50}

length

$t_{1/2}$

relationships

$$K_d = \frac{[A] \cdot [B]}{[AB]}$$

$$d[X]/dt = k \cdot [Y]^2$$

$$\sum_i [X]_i - F(t) = 0$$

$$k(t) \sim N(k, \sigma^2)$$

If $\text{mass}_t > \text{threshold}$
then $\text{mass}_{t+\Delta t} = 0.5 \cdot \text{mass}$

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a **system** using **mathematical** concepts and language.”

variables

$[x]$

V_{\max}

K_d

EC_{50}

length

$t_{1/2}$

relationships

$$K_d = \frac{[A] \cdot [B]}{[AB]}$$

$$d[X]/dt = k \cdot [Y]^2$$

$$\sum_i [X]_i - F(t) = 0$$

$$k(t) \sim N(k, \sigma^2)$$

If $\text{mass}_t > \text{threshold}$
then $\text{mass}_{t+\Delta t} = 0.5 \cdot \text{mass}$

constraints

$$[x] > 0$$

Energy conservation

Boundary conditions
($v < \text{upper limit}$)

Objective functions
(maximise ATP)

Initial conditions

What is a mathematical model?

Wikipedia (April 17th 2013): “A mathematical model is a description of a **system** using **mathematical** concepts and language.”

variables

$[x]$

Vmax

Kd

EC₅₀

length

t_{1/2}

relationships

$$K_d = \frac{[A] \cdot [B]}{[AB]}$$

$$d[X]/dt = k \cdot [Y]^2$$

$$\sum_i [X]_i - F(t) = 0$$

$$k(t) \sim N(k, \sigma^2)$$

If mass_t > threshold
then mass_{t+Δt} = 0.5 · mass

constraints

$$[x] > 0$$

Energy conservation

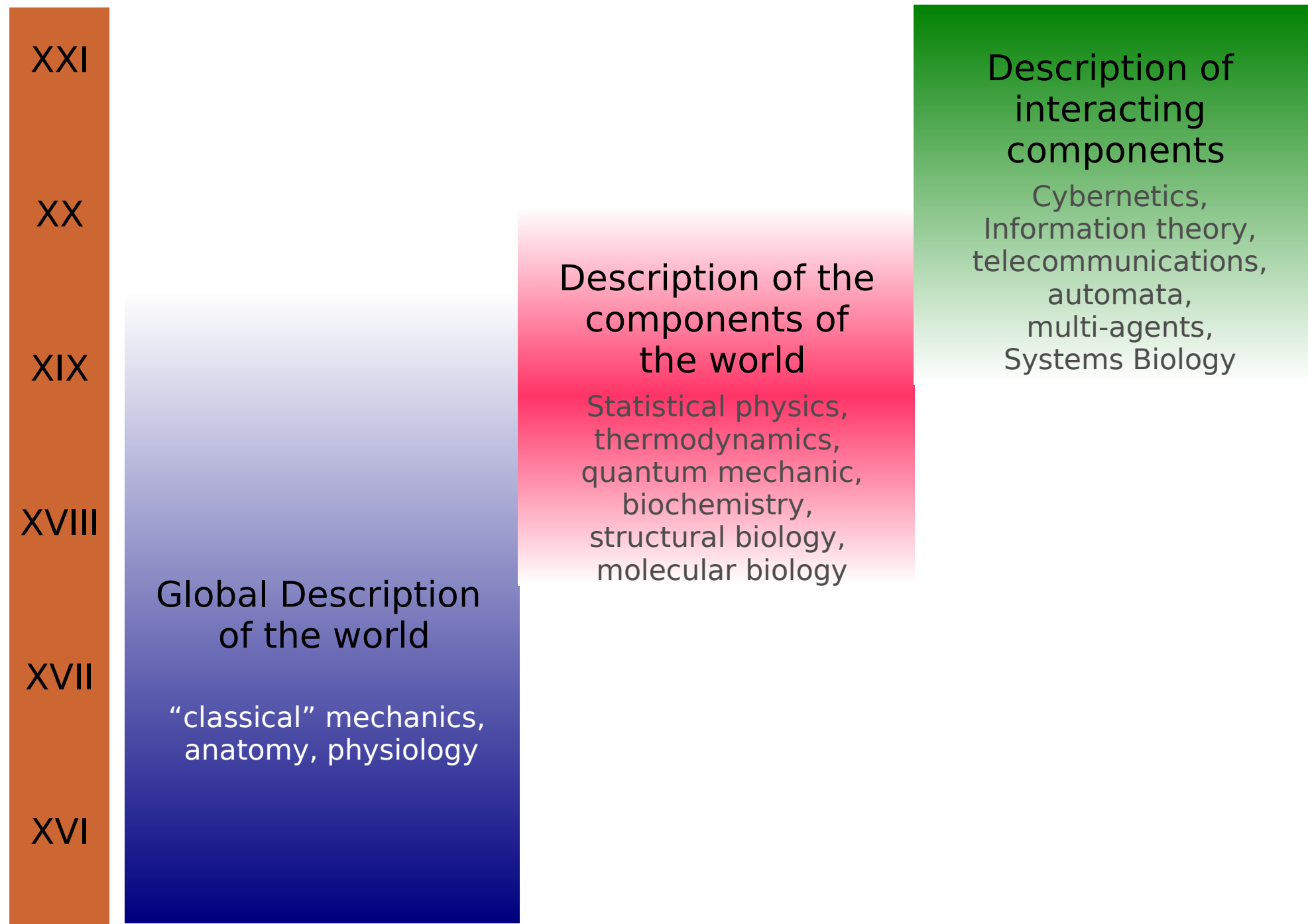
Boundary conditions
(v < upper limit)

Objective functions
(maximise ATP)

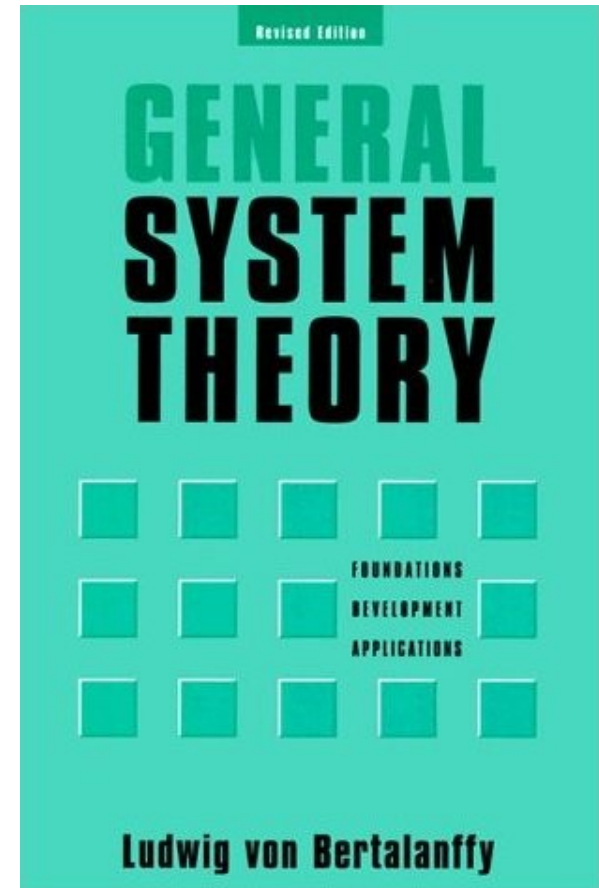
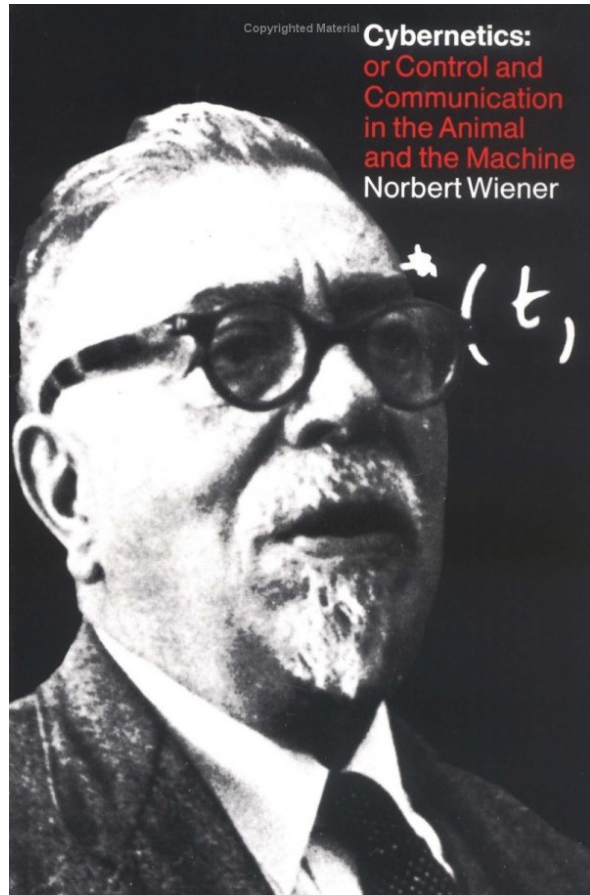
Initial conditions

Different types: Dynamical models, logical models, rule-based models, multi-agent models, statistical models, etc.

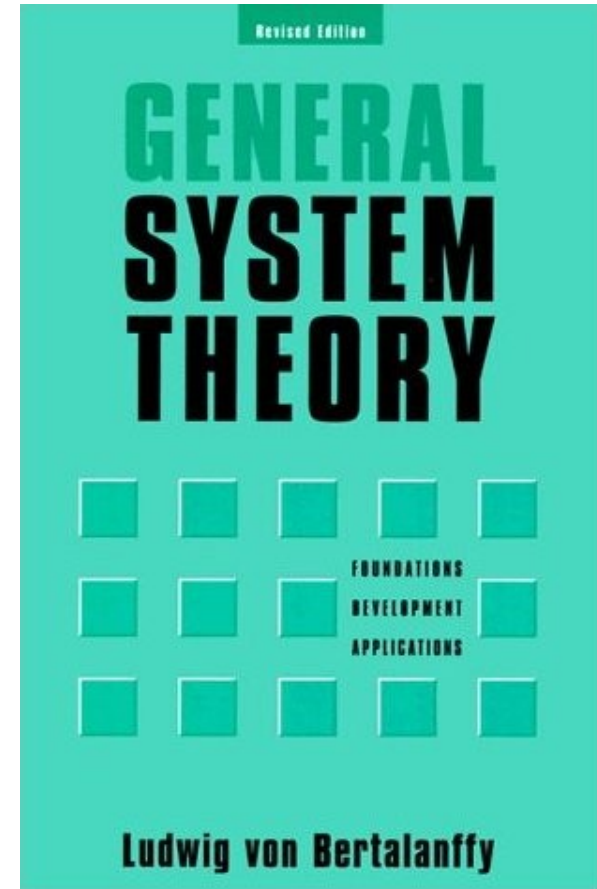
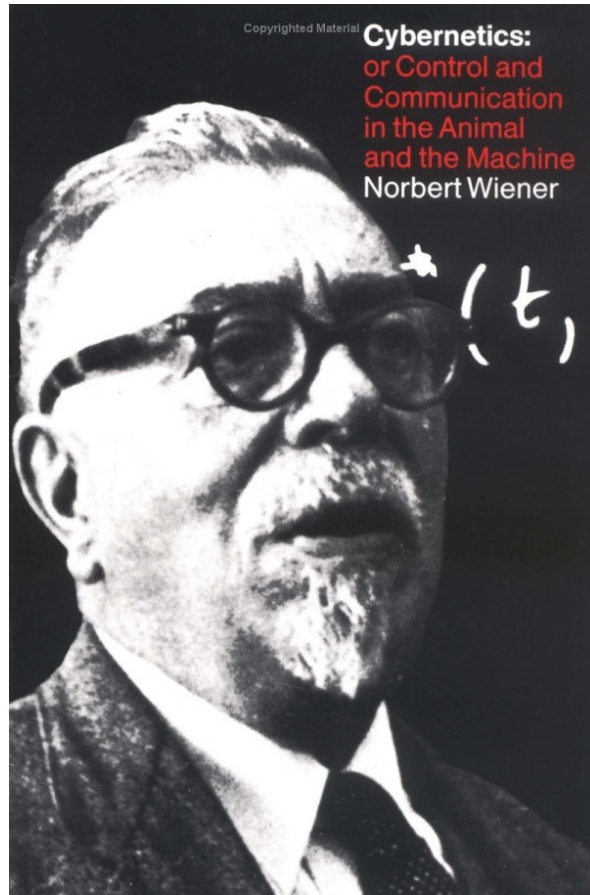
Emergence of the notion of system



Systems are formalised mid-XXth

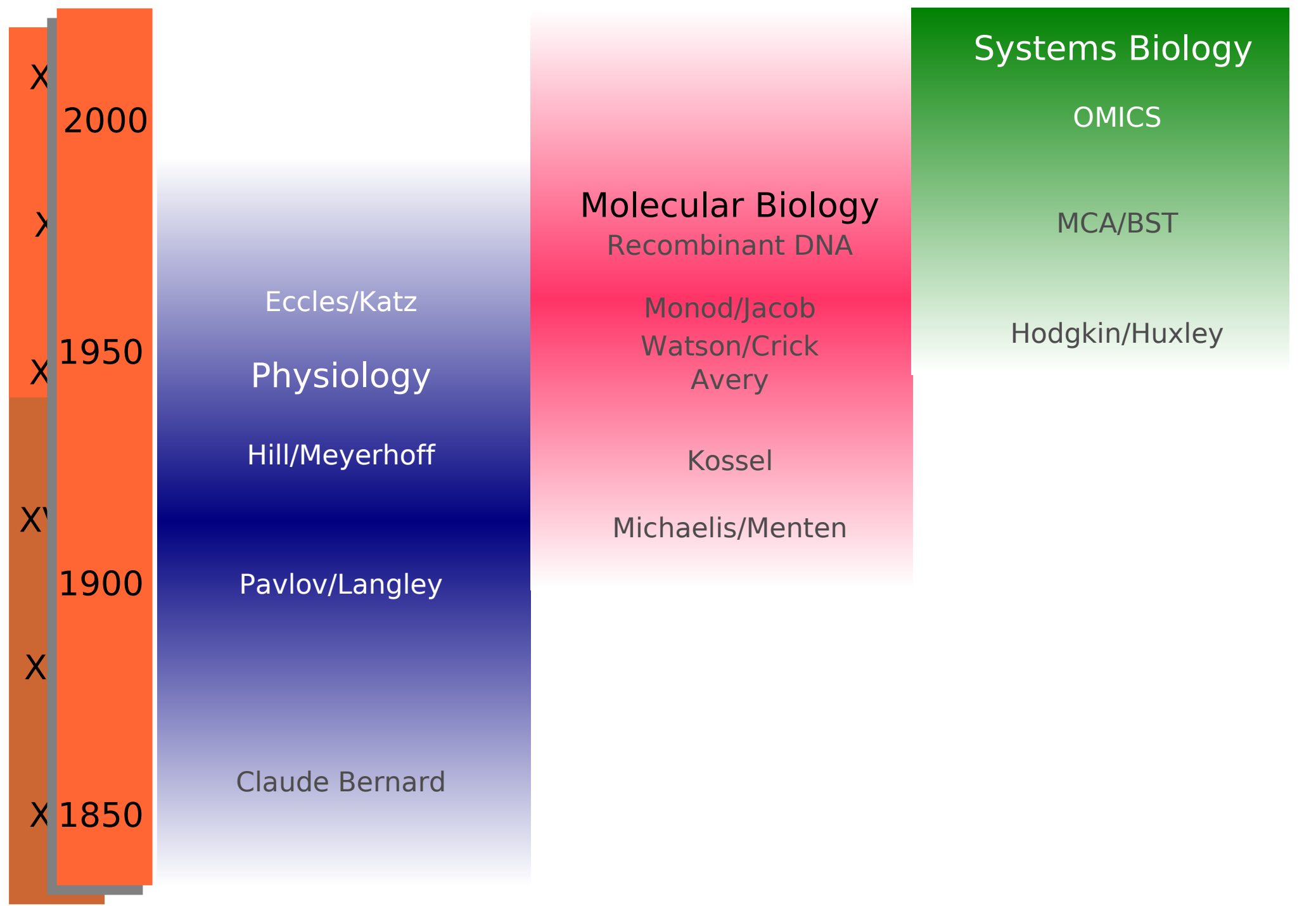


Systems are formalised mid-XXth



"[A system consists of] a dynamic order of parts and processes standing in mutual interaction. [...] The fundamental task of biology [is] the discovery of the laws of biological systems"
Ludwig von Bertalanffy, Kritische Theorie der Formbildung, 1928

The three paradigms of Biology



Computer simulations Vs. mathematical models

[37]

THE CHEMICAL BASIS OF MORPHOGENESIS

By A. M. TURING, F.R.S. *University of Manchester*

(Received 9 November 1951—Revised 15 March 1952)

It is suggested that a system of chemical substances, called morphogens, reacting together and diffusing through a tissue, is adequate to account for the main phenomena of morphogenesis. Such a system, although it may originally be quite homogeneous, may later develop a pattern or structure due to an instability of the homogeneous equilibrium, which is triggered off by random disturbances. Such reaction-diffusion systems are considered in some detail in the case of an isolated ring of cells, a mathematically convenient, though biologically unusual system. The investigation is chiefly concerned with the onset of instability. It is found that there are six essentially different forms which this may take. In the most interesting form stationary waves appear on the ring. It is suggested that this might account, for instance, for the tentacle patterns on *Hydra* and for whorled leaves. A system of reactions and diffusion on a sphere is also considered. Such a system appears to account for gastrulation. Another reaction system in two

Computer simulations Vs. mathematical models

[37]

THE CHEMICAL BASIS OF MORPHOGENESIS

By A. M. TURING, F.R.S. *University of Manchester*

(Received 9 November 1951—Revised 15 March 1952)

It is suggested that a system of chemical substances, called morphogens, reacting together and diffusing through a tissue, is adequate to account for the main phenomena of morphogenesis.

One would like to be able to follow this more general process mathematically also. The difficulties are, however, such that one cannot hope to have any very embracing theory of such processes, beyond the statement of the equations. It might be possible, however, to treat a few particular cases in detail with the aid of a digital computer. This method has the advantage that it is not so necessary to make simplifying assumptions as it is when doing a more theoretical type of analysis.

Birth of Computational Systems Biology

J. Physiol. (1952) 117, 500–544

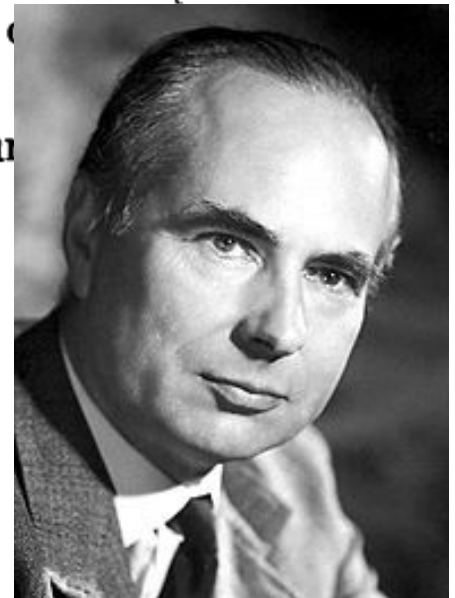
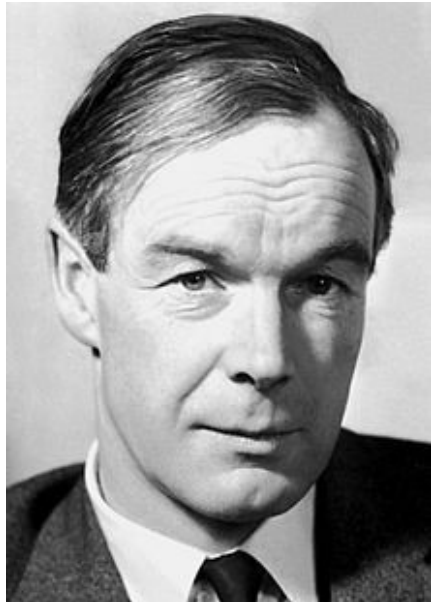
A QUANTITATIVE DESCRIPTION OF MEMBRANE CURRENT AND ITS APPLICATION TO CONDUCTION AND EXCITATION IN NERVE

BY A. L. HODGKIN AND A. F. HUXLEY

From the Physiological Laboratory, University of Cambridge

(Received 10 March 1952)

This article concludes a series of papers concerned with the flow of electric current through the surface membrane of a giant nerve fibre (Hodgkin, & Katz, 1952; Hodgkin & Huxley, 1952 *a–c*). Its general object is to put the results of the preceding papers (Part I), to put them in mathematical form (Part II) and to show that they will account for conduction and excitation in quantitative terms (Part III).



R.I.P. 30 May 2012
Cambridge

The Computational Systems Biology loop

“biological” model

mathematical model

$$I = C_M \frac{dV}{dt} + \bar{g}_K n^4 (V - V_K) + \bar{g}_{Na} m^3 h (V - V_{Na}) + \bar{g}_l (V - V_l),$$

$$\frac{dn}{dt} = \alpha_n (1 - n) - \beta_n n,$$

$$\frac{dm}{dt} = \alpha_m (1 - m) - \beta_m m,$$

$$\frac{dh}{dt} = \alpha_h (1 - h) - \beta_h h,$$

$$\alpha_n = 0.01 (V + 10) / \left(\exp \frac{V + 10}{10} - 1 \right),$$

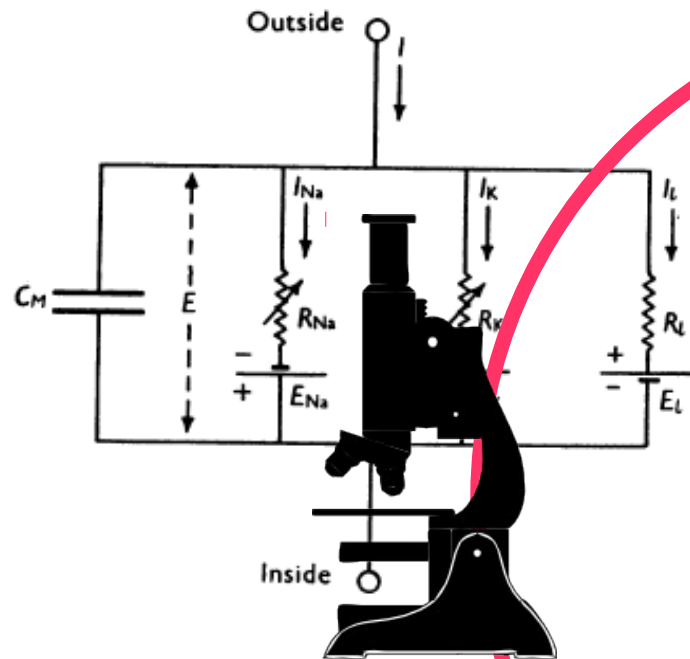
$$\beta_n = 0.125 \exp (V/80),$$

$$\alpha_m = 0.1 (V + 25) / \left(\exp \frac{V + 25}{10} - 1 \right),$$

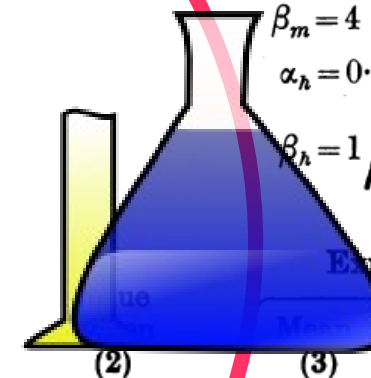
$$\beta_m = 4 \exp (V/18),$$

$$\alpha_h = 0.07 \exp (V/20),$$

$$\beta_h = 1 / \left(\exp \frac{V + 30}{10} + 1 \right).$$



validation



parameterisation

Constant
(1)

C_M ($\mu\text{F}/\text{cm}^2$)

V_{Na} (mV)

V_K (mV)

V_l (mV)

\bar{g}_{Na} (m.mho/cm²)

\bar{g}_K (m.mho/cm²)

\bar{g}_l (m.mho/cm²)

Range
(4)

0.8 to 1.5

-95 to -119

+9 to +14

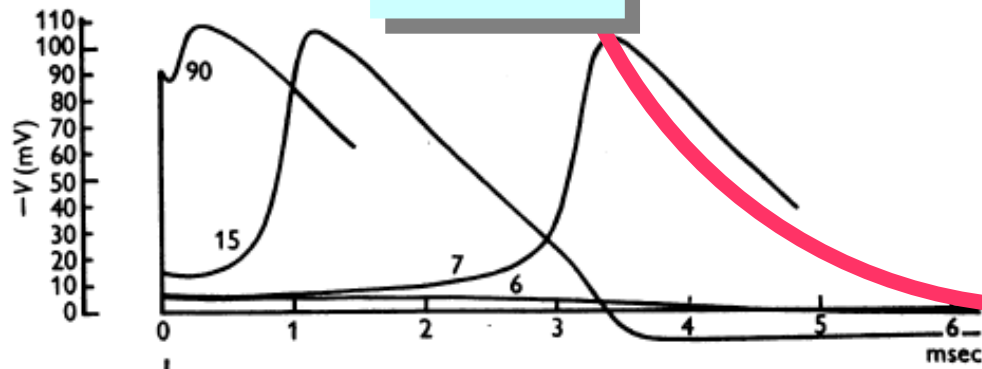
-4 to -22

65 to 90

120 to 260

26 to 49

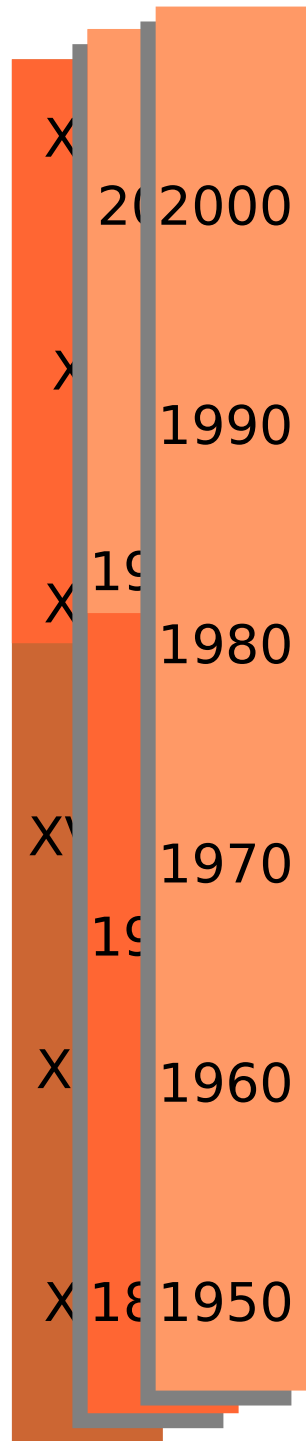
0.13 to 0.50



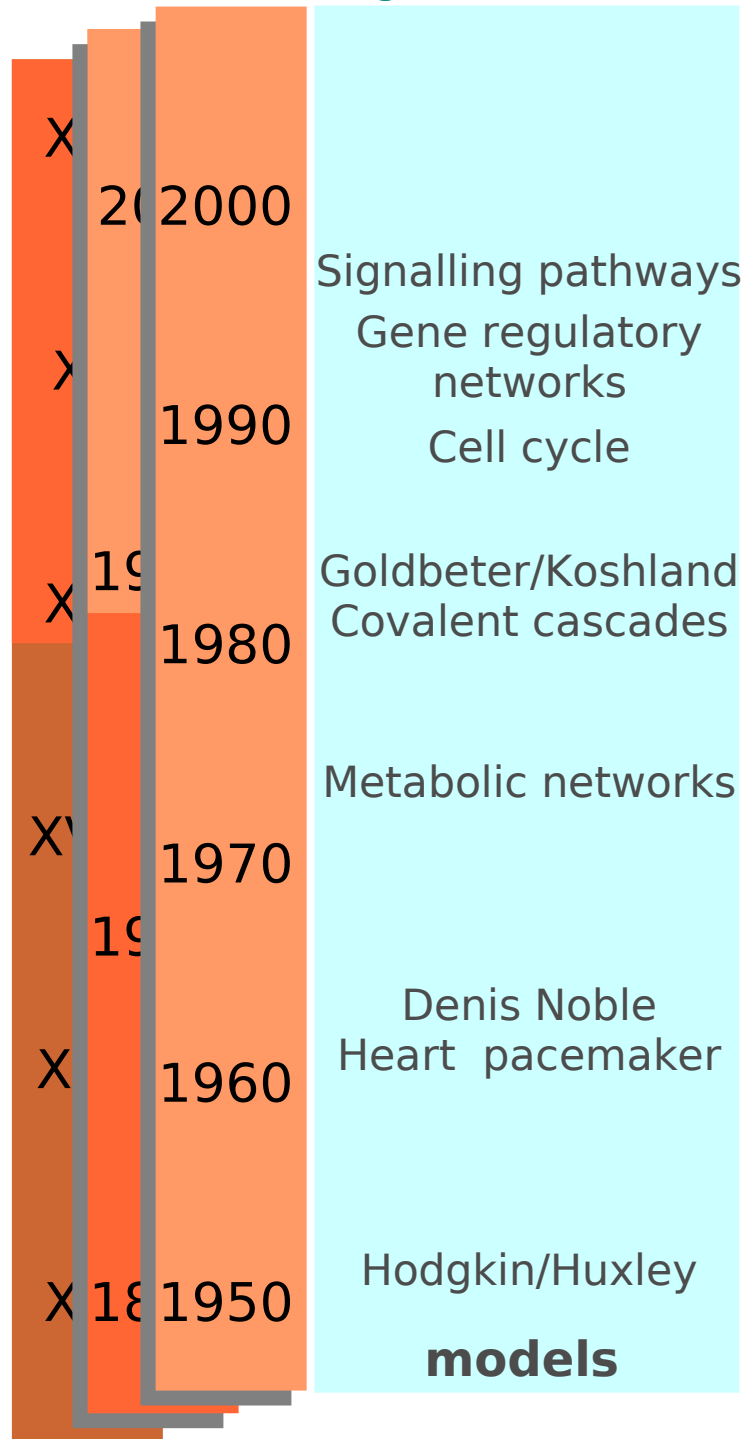
simulation

computational model

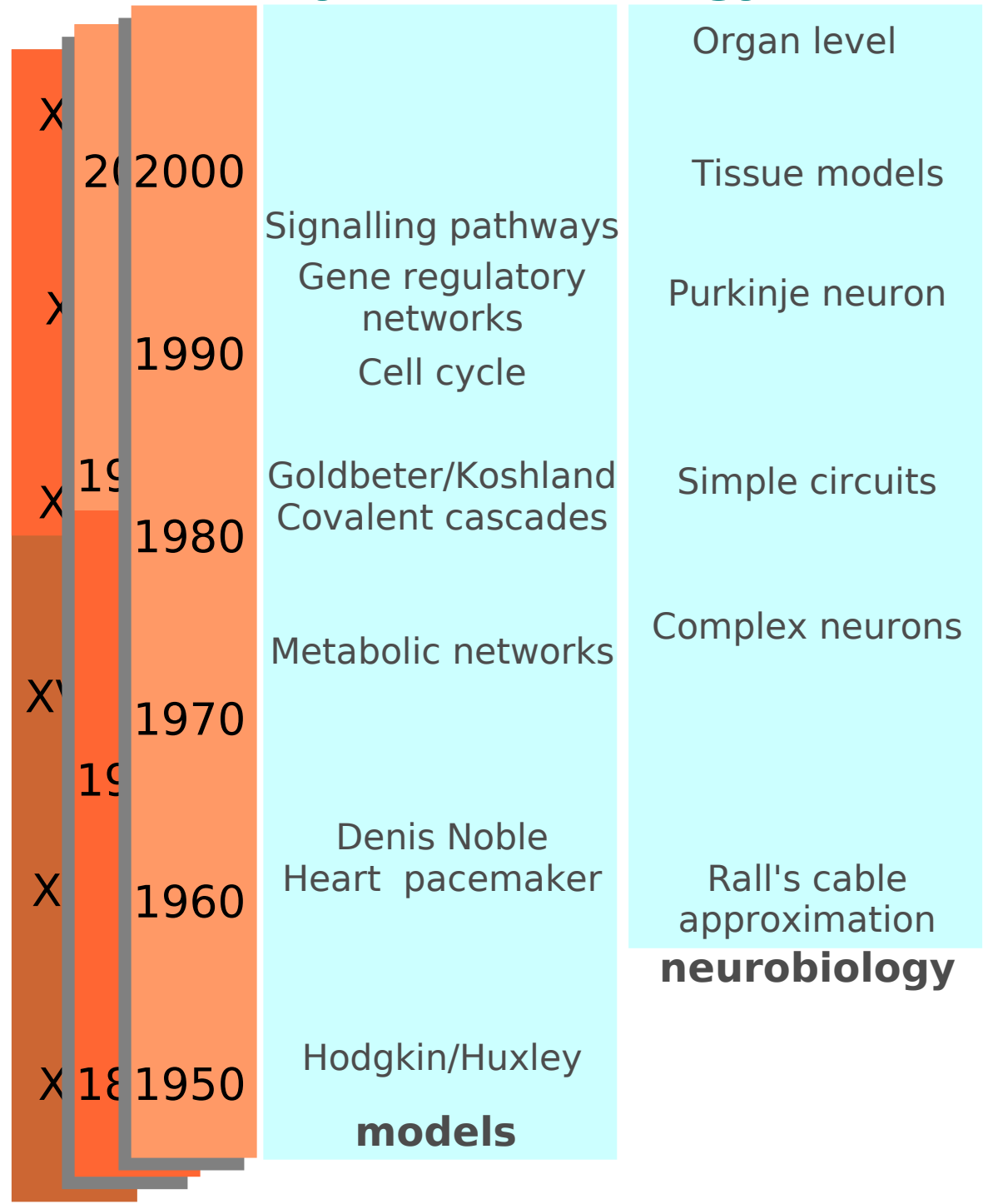
Towards Systems Biology



Towards Systems Biology



Towards Systems Biology

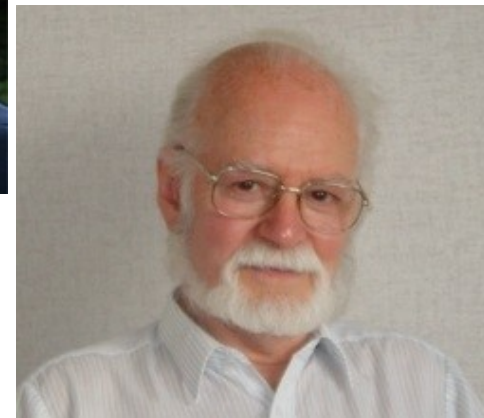


60s and 70s

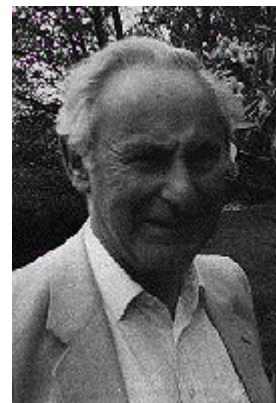
- Mihajlo Mesarovic: 1966 Symposium
“general systems theory and biology



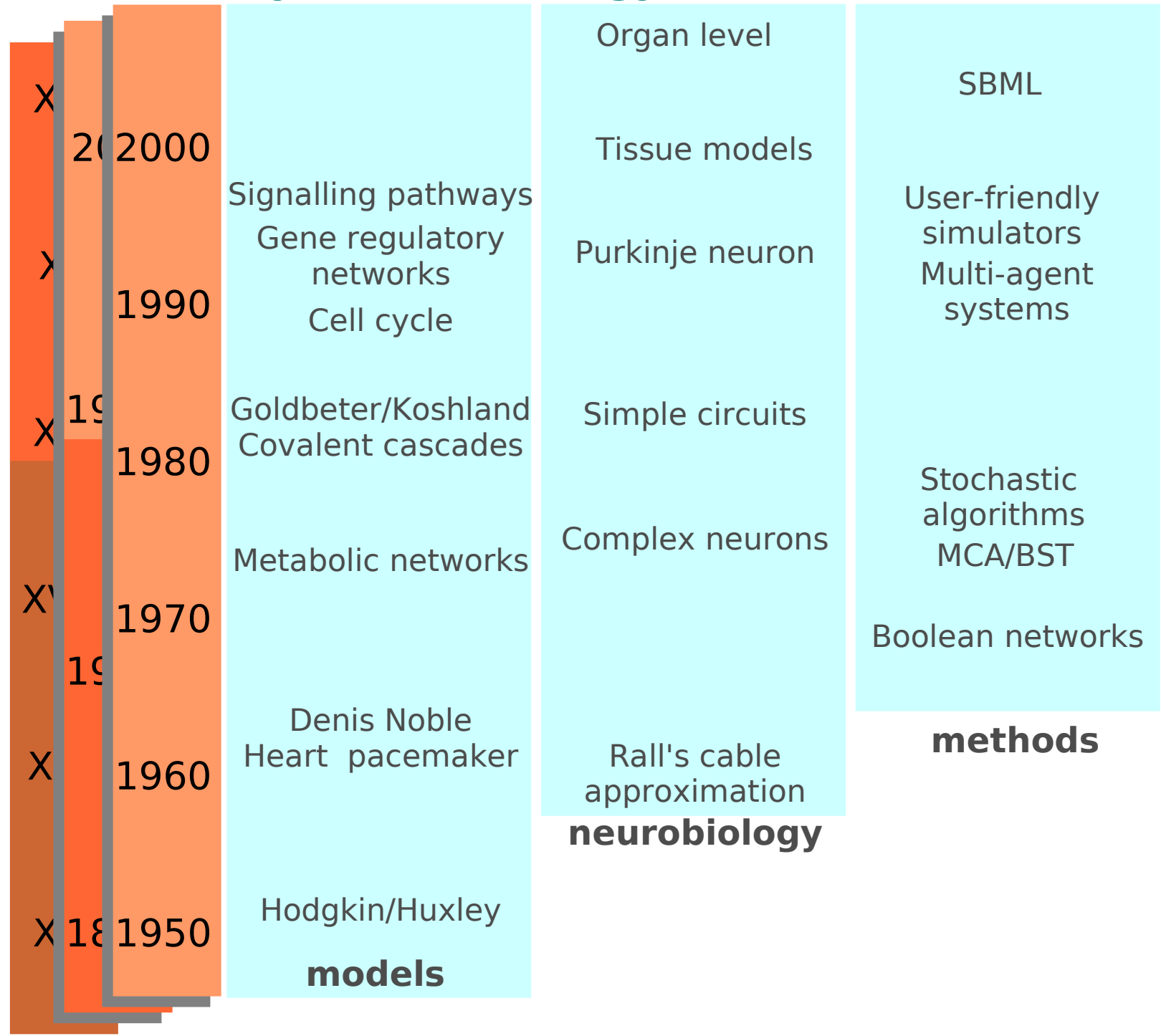
- Stuart Kaufmann,
Rene Thomas: 1969-73
boolean networks for
gene regulation



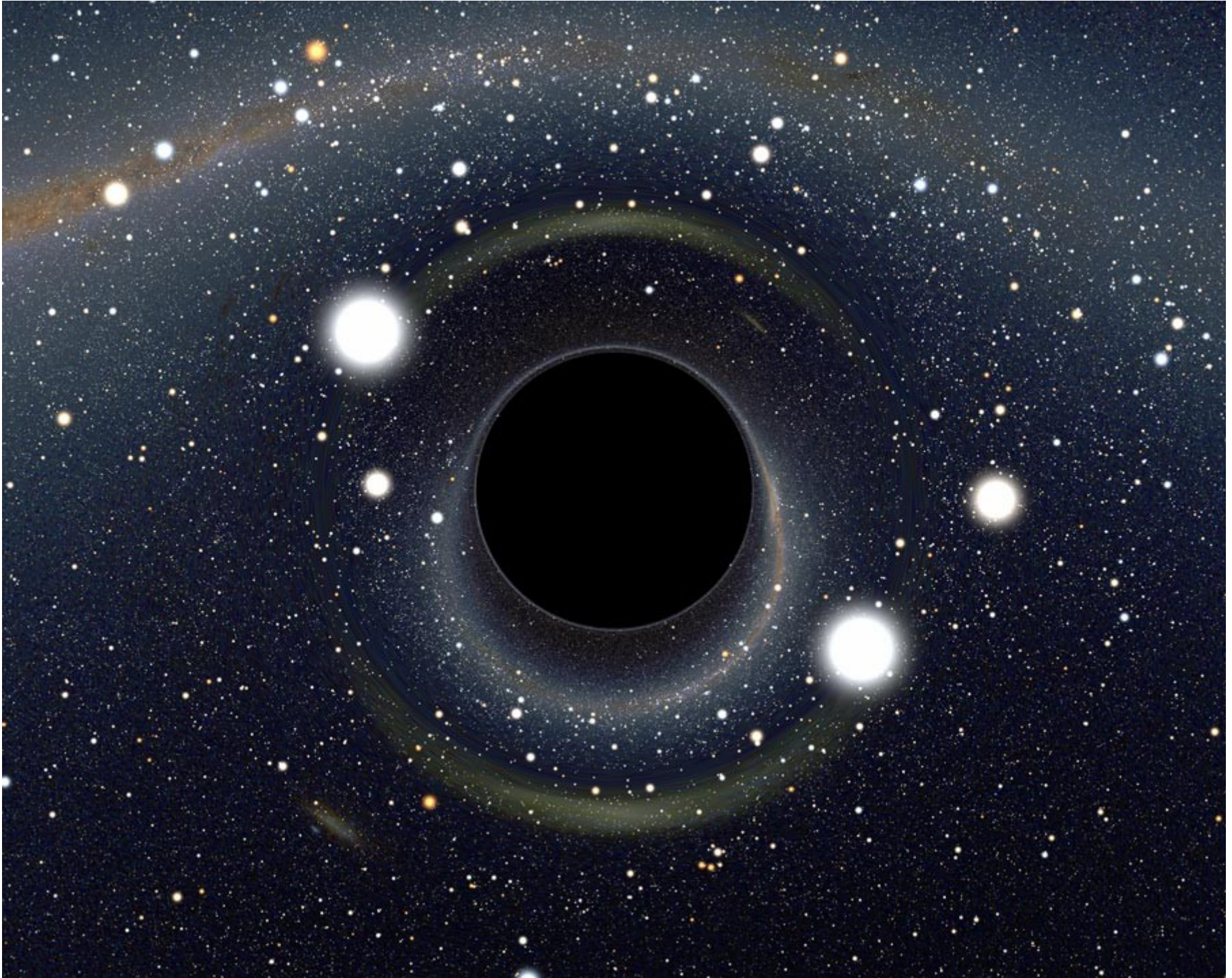
- Henri Kacser:
Metabolic control analysis,
Michel Savageau:
Biochemical Systems Theory



Towards Systems Biology



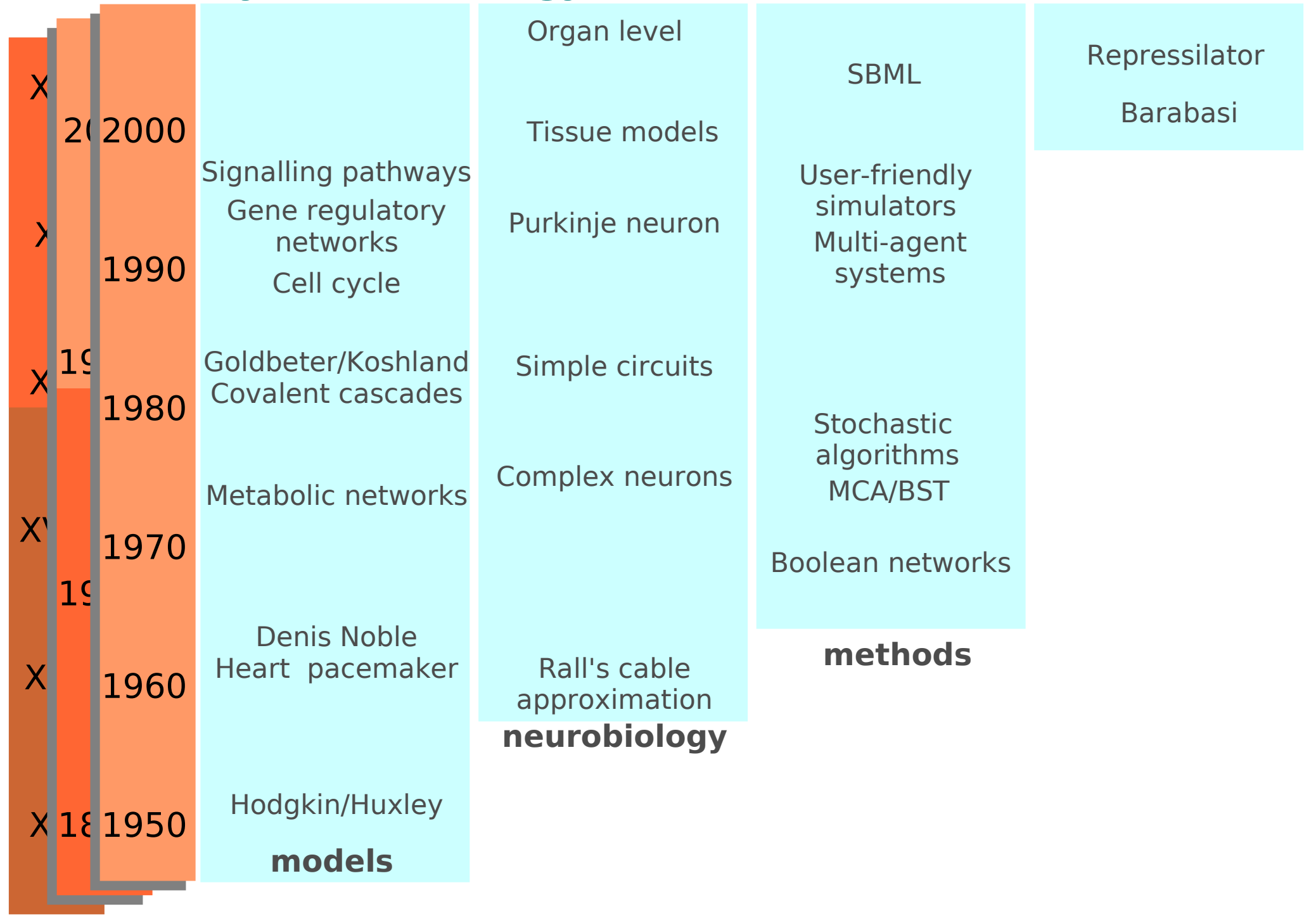
80s: The reign of Molecular Biology



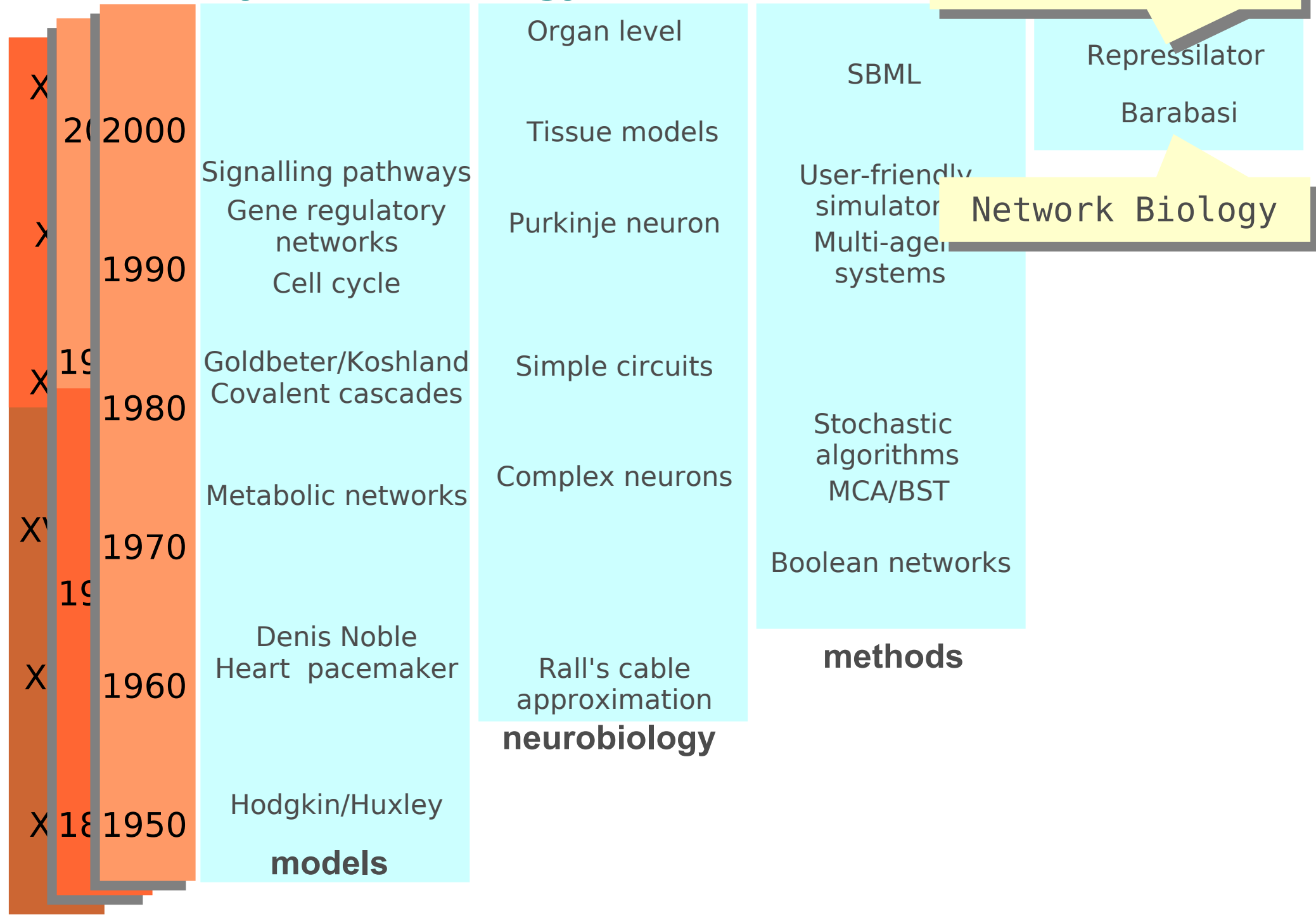
90s: maturation of the community

- Publication of modelling work in high visibility journals, e.g.:
 - Tyson. modeling the cell-division cycle - cdc2 and cyclin interactions. *PNAS* 1991; McAdams and Shapiro. Circuit simulation of genetic networks. *Science* 1995; Barkai and Leibler. Robustness in simple biochemical networks. *Nature* 1997; Neuman et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 1998; Yue et al. Genomic cis-regulatory logic: Experimental and computational analysis of a sea urchin gene . *Science* 1998; Bray et al. Receptor clustering as a cellular mechanism to control sensitivity. *Nature* 1998; Bhalla ad Iyengar. Emergent properties of signaling pathways. *Science* 1998
- Structuring of the community modelling metabolism
- Large-scale modelling and simulation projects
 - E-Cell project 1996; The Virtual Cell 1998
- Availability of high-throughput data on parts and interactions
 - Two-hybrids (1989); microarrays (1995) etc.
- Large-scale funding for wet+dry studies of biological systems
 - Alliance For Cellular Signalling (<http://www.afcs.org/>). First of the NIH “glue grants”. 1998

Towards Systems Biology



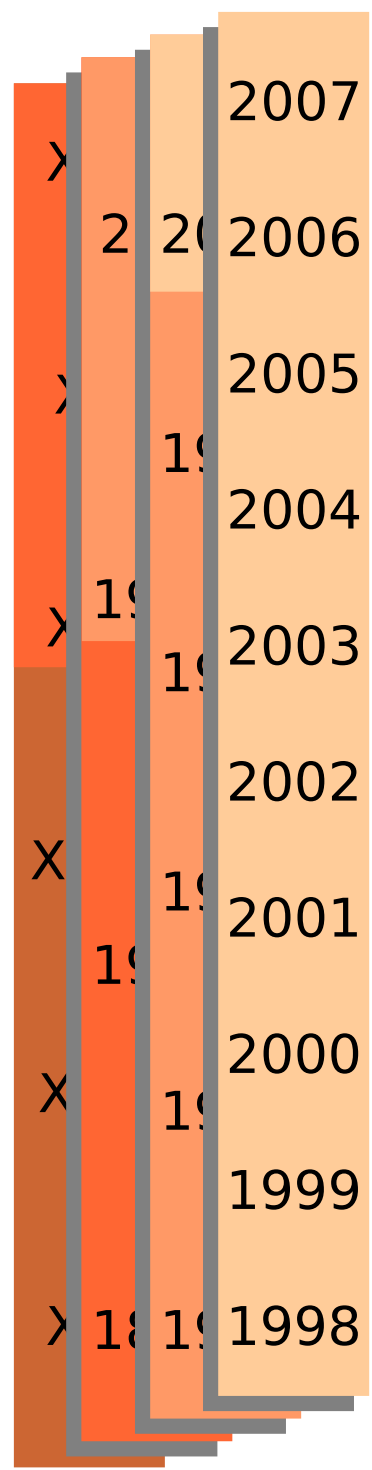
Towards Systems Biology



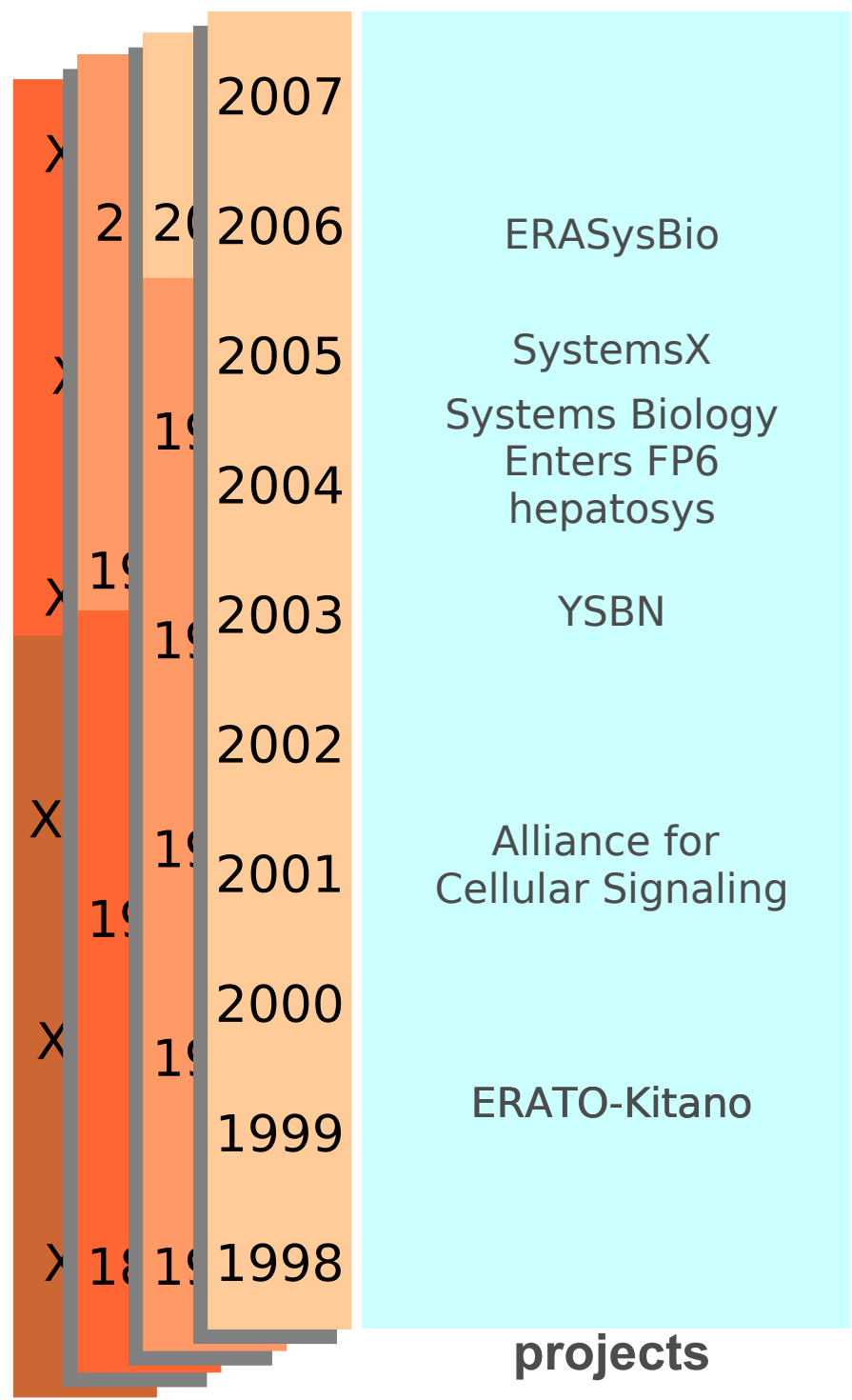
Formal revival of Systems Biology

- “Modelling” Systems Biology
 - 1998 - Hiroaki Kitano founds the Systems Biology Institute in Tokyo
 - First appearance: Kyoda, Kitano. Virtual Drosophila project: Simulation of drosophila leg formation. *Genome Informatics Series* (1998)
 - Kitano, H. Perspectives on systems biology. *New Generation Computing* Volume 18, Issue 3, 2000, Pages 199-216
- “Network” Systems Biology
 - First appearance: Leroy Hood. Systems biology: new opportunities arising from genomics, proteomics and beyond. *Experimental Hematology*. Volume 26, Issue 8, 1998, Page 681
 - Schwikowski B, Uetz P, Fields S. A network of protein-protein interactions in yeast. *Nat Biotechnol*. 2000 Dec;18(12):1257-61.
 - 2000 - Leroy Hood founds the Systems Biology Institute in Seattle

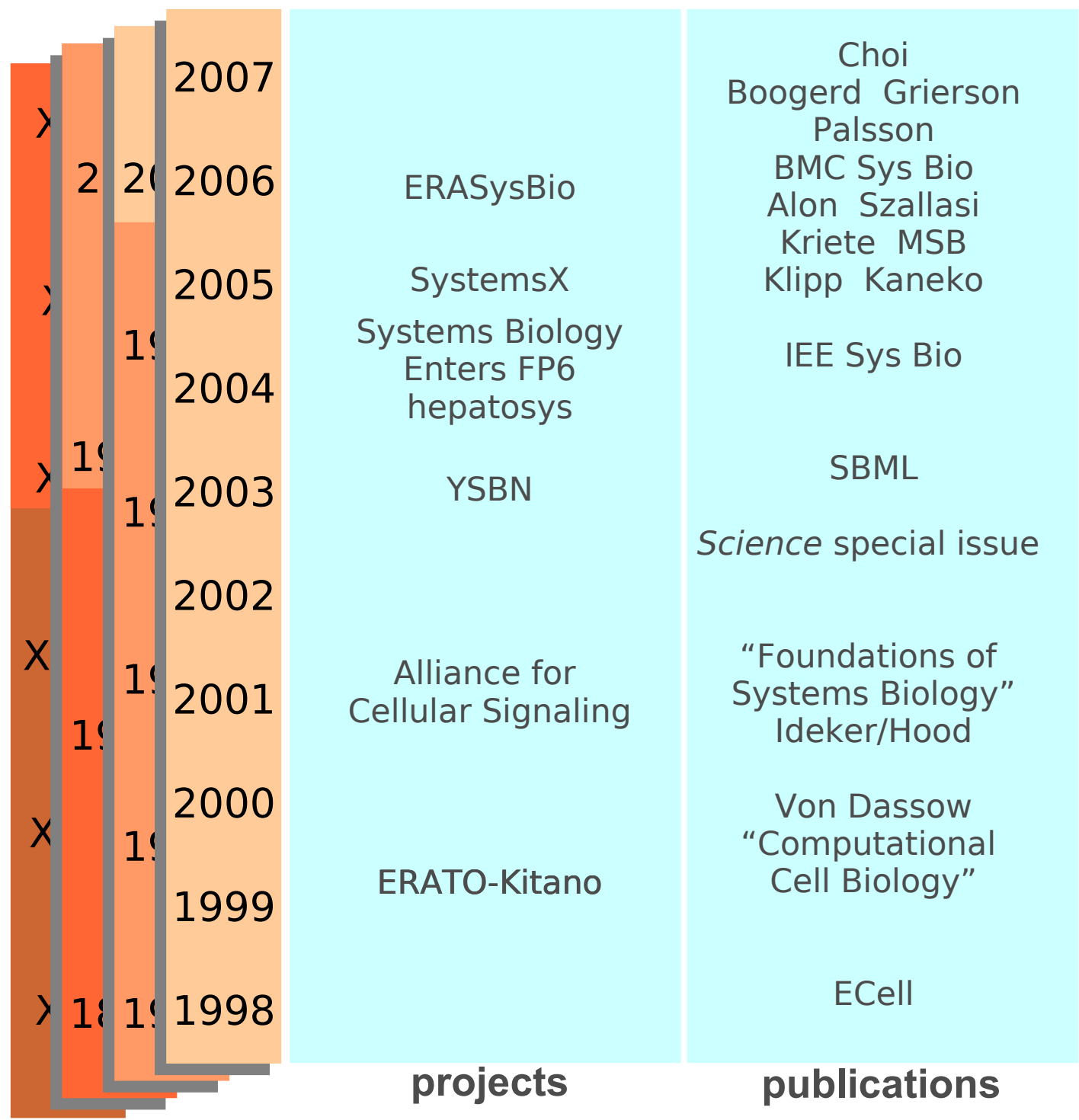
Rise of Systems Biology as a paradigm



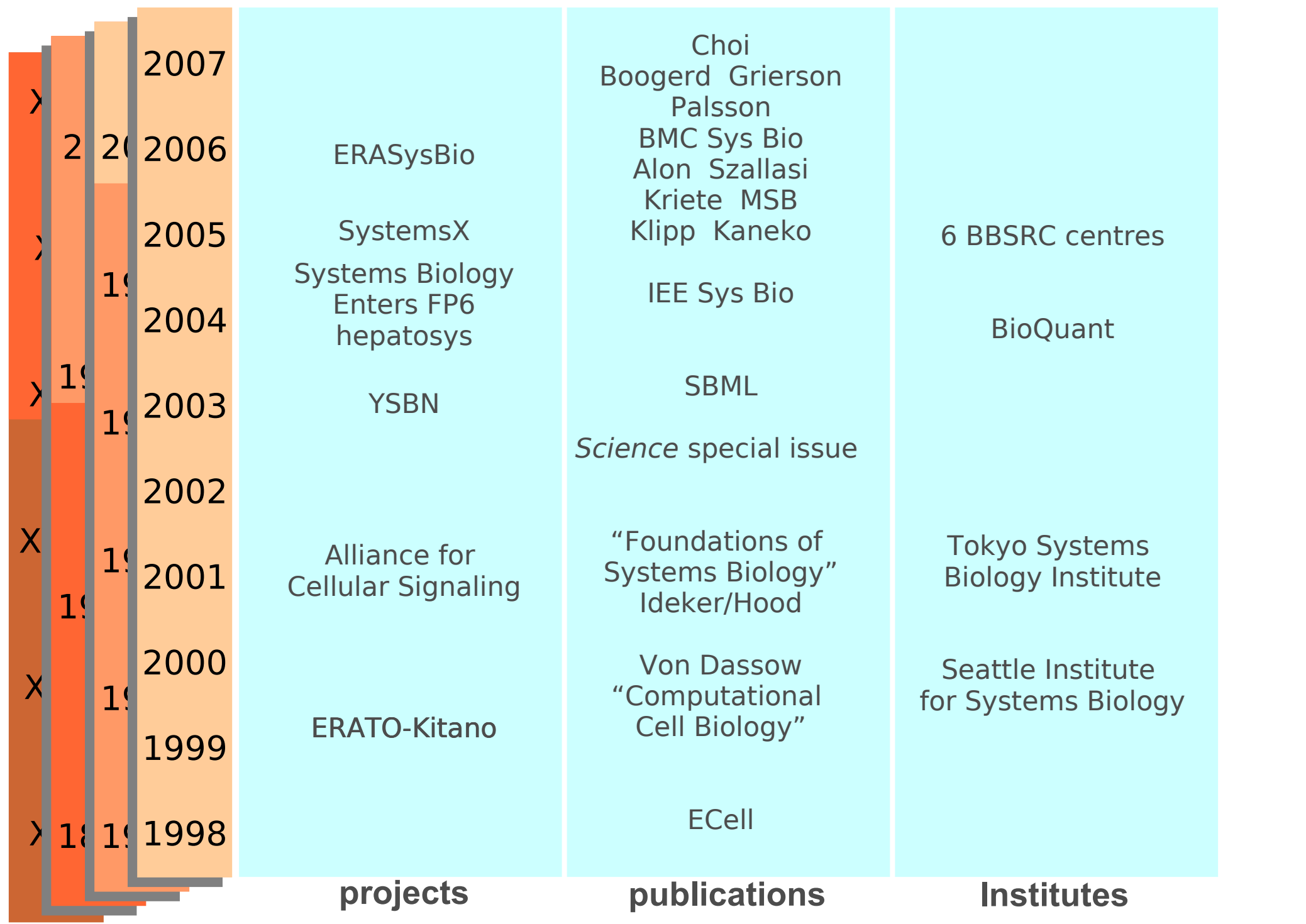
Rise of Systems Biology as a paradigm



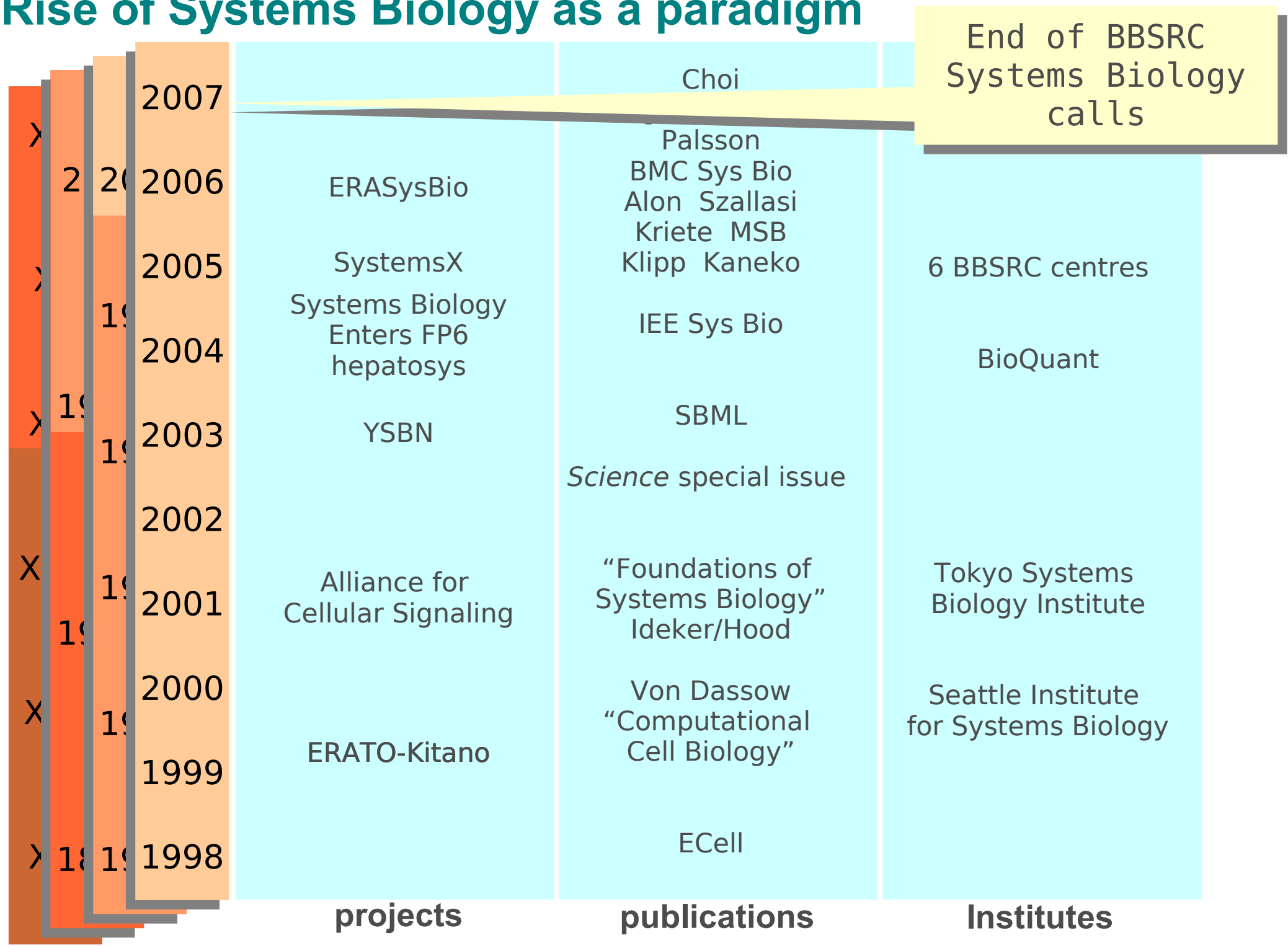
Rise of Systems Biology as a paradigm



Rise of Systems Biology as a paradigm



Rise of Systems Biology as a paradigm



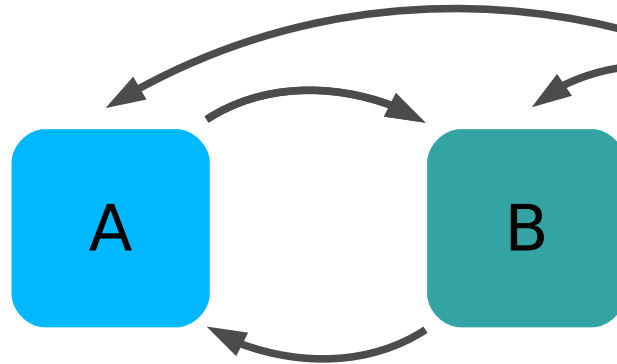
But what is Systems Biology???

Molecular and Cellular Biology

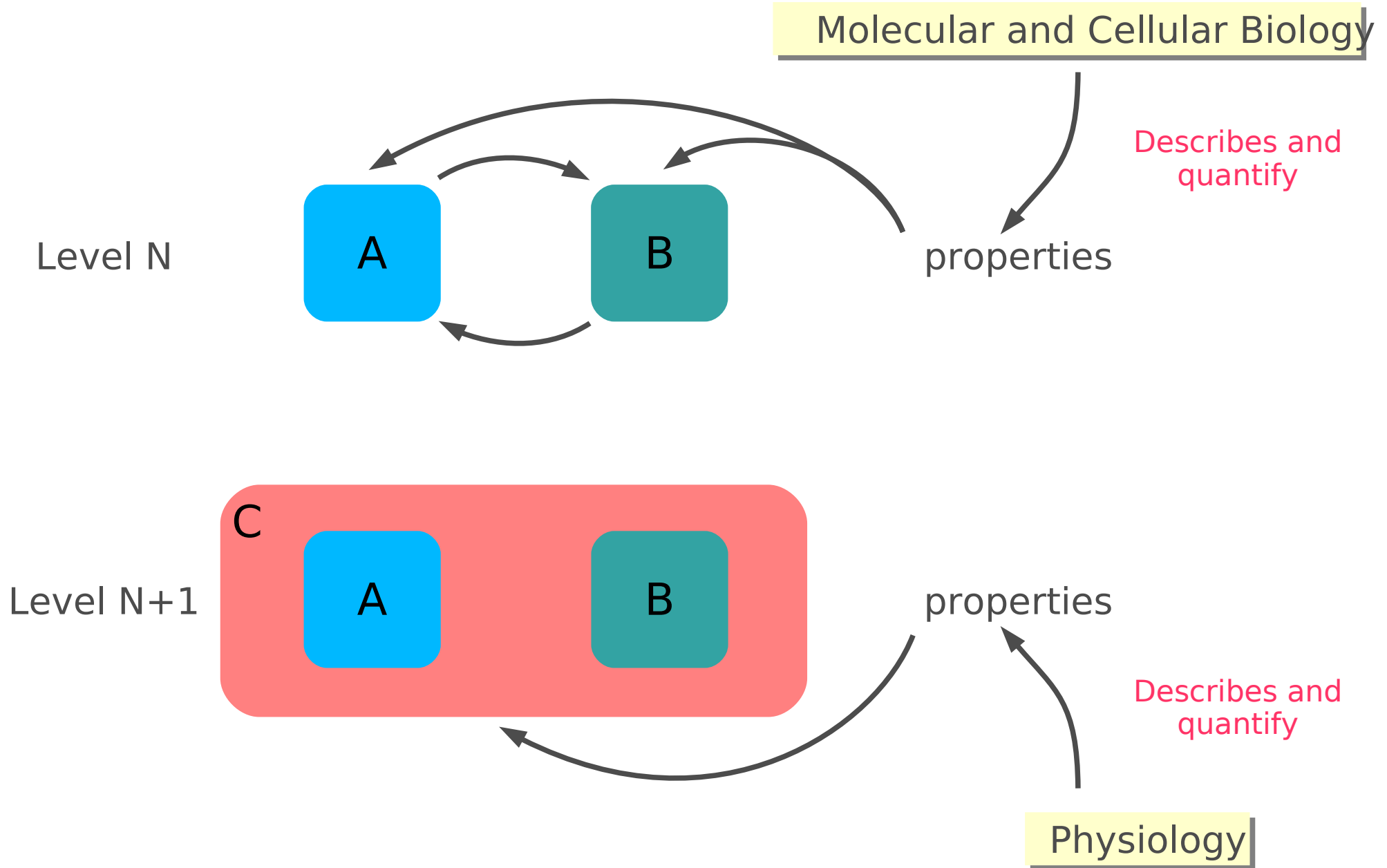
Describes and
quantify

properties

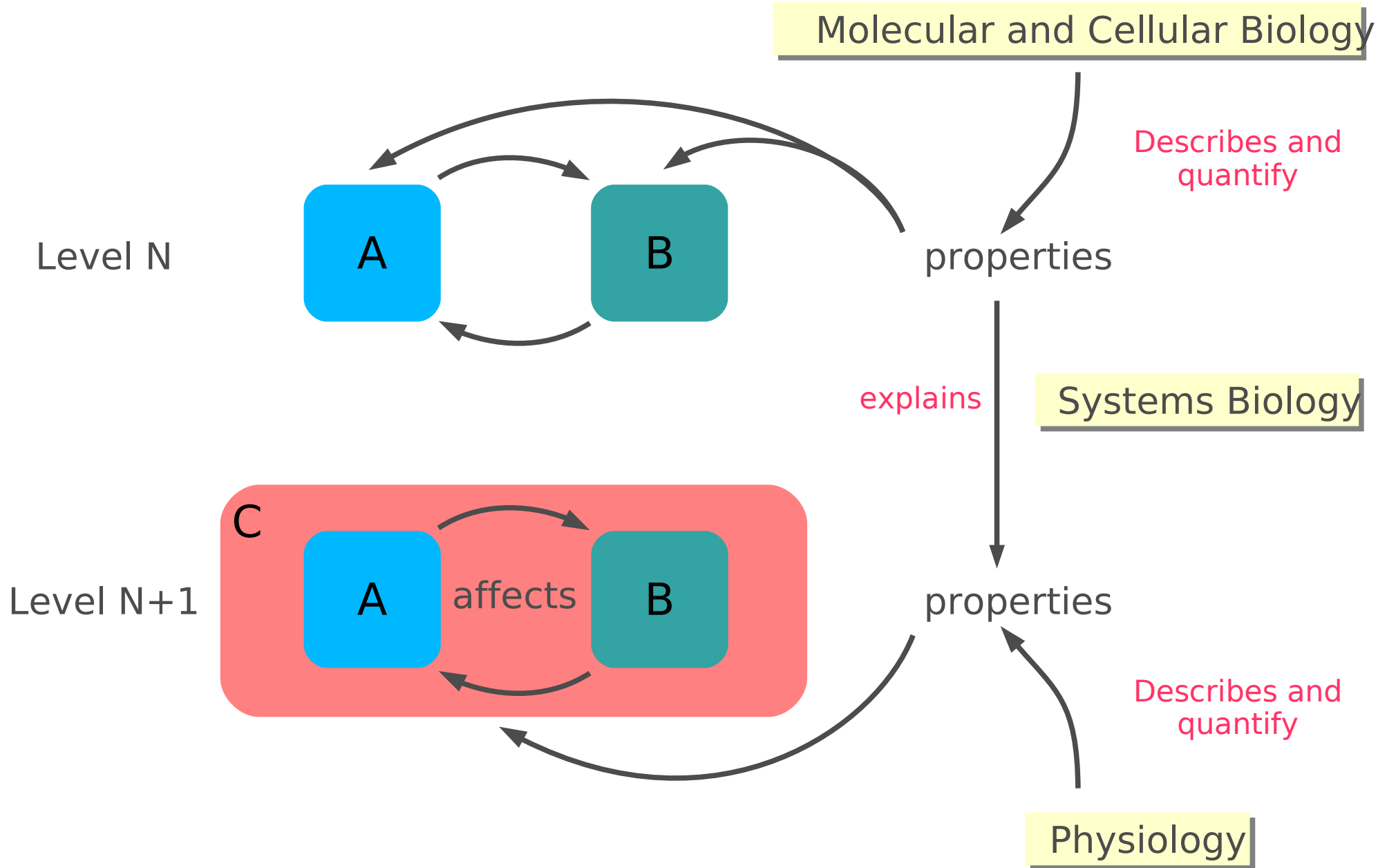
Level N



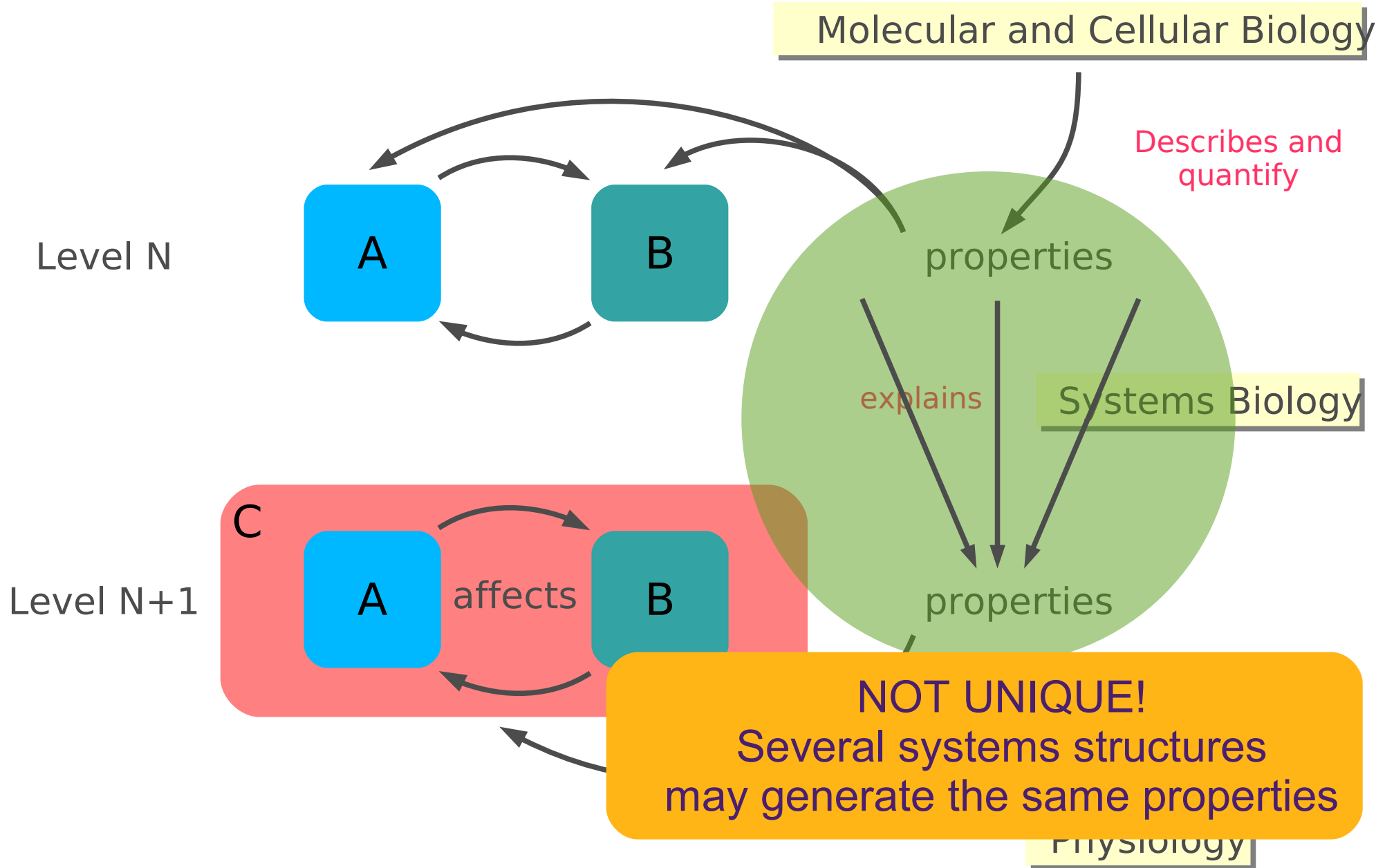
But what is Systems Biology???



But what is Systems Biology???



But what is Systems Biology???



What Systems Biology is not!

Purely theoretical: Most systems biologists are actually experimental biologists

What Systems Biology is not!

Purely theoretical: Most systems biologists are actually experimental biologists



Based on large datasets: in a system of two enzymes, the behaviour of both reactions is different than the ones observed in isolation

What Systems Biology is not!

Purely theoretical: Most systems biologists are actually experimental biologists

Based on large datasets: in a system of two enzymes, the behaviour of both reactions is different than the ones observed in isolation

Focused on biomolecular systems: systems biology is scale-free, and a biological system can be made up of molecules, cells, organs or individuals

What Systems Biology is not!

Purely theoretical: Most systems biologists are actually experimental biologists

Based on large datasets: in a system of two enzymes, the behaviour of both reactions is different than the ones observed in isolation

Focused on biomolecular systems: systems biology is scale-free, and a biological system can be made up of molecules, cells, organs or individuals

Systems Biology is the study of the *emerging* properties of a biological system, taking into account all the *necessary* constituents, their *relationships* and their *dynamics*

Two kinds of Systems Biology?

Systems-wide analysis (omics)

- Born: 1990s
- Technologies: high-throughput, statistics
- People's background: molecular biologists, mathematicians
- Key lesson: the selection of a phenotype is done at the level of the system, not of the component (gene expression puzzle)

Application of systems-theory

- Born: 1960s
- Technologies: quantitative measurements, modelling
- People's background: biochemists, engineers
- Key lesson: the properties at a certain level are emerging from the dynamic interaction of components at a lower level

Procedure does not depend on directionality

Bottom-up

Top-down

Build the system

literature

“pathway” database

Put numbers

biochemistry

“omics”

Parametrise

parameter search

hmm....

Analyse

Simulation

structural analysis,
steady-state analysis

perturb

Inhibition, stimulation,
suppression, overexpression

The challenges ahead

Types of representation

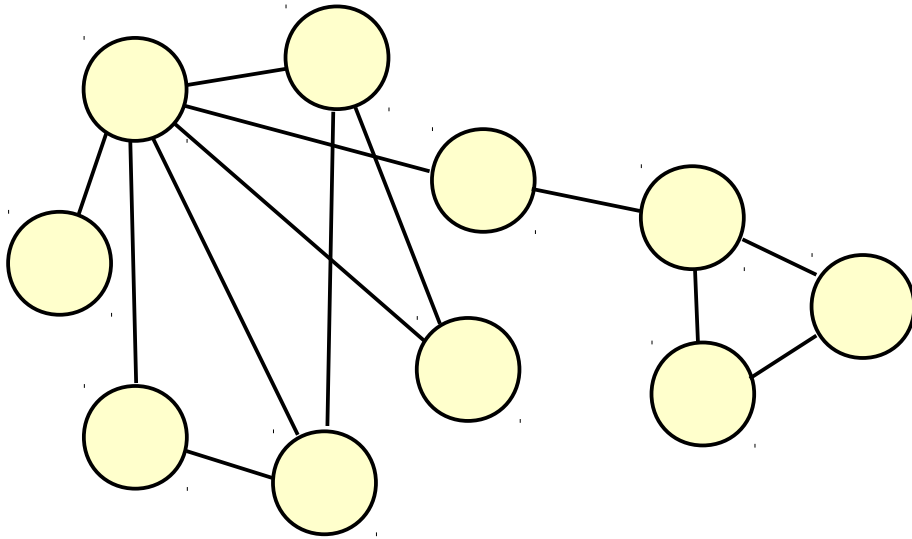
Scales and the mesoscopic gap

Genotype-system-phenotype problem

Drug discovery models
Vs systems modelling

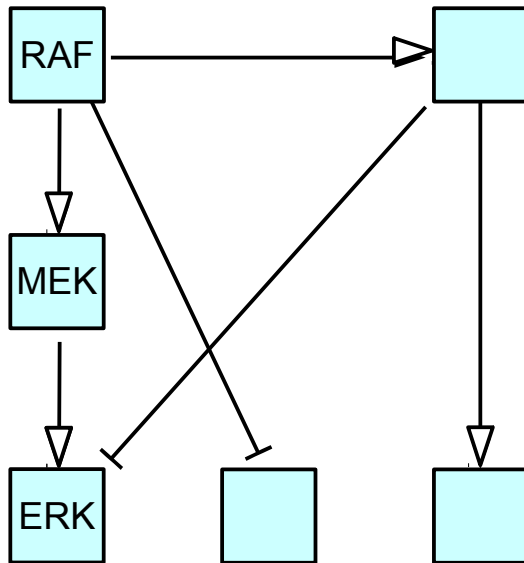
Drug discovery models
Vs “omics” data

Interaction networks



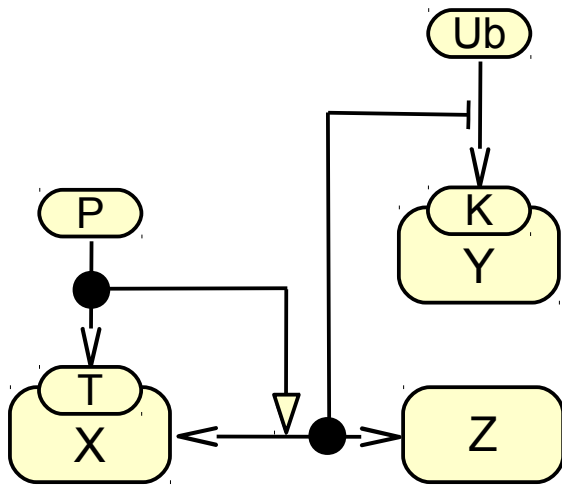
- Non-directional
- Non-sequential
- Non-mechanistic
- Statistical modelling
- Functional genomics
- IntAct, DIP, String

Activity-Flows



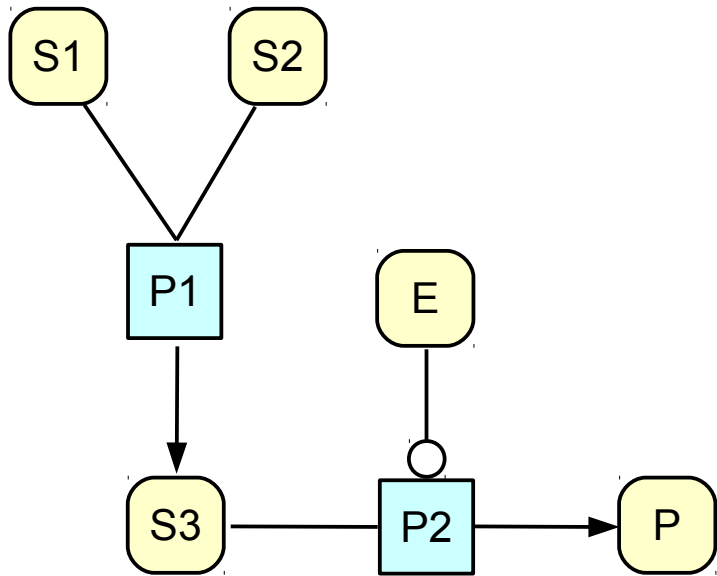
- Directional
- Sequential
- Non-mechanistic
- Logical modelling
- Signalling pathways, gene regulatory networks
- KEGG, STKEs

Entity Relationships

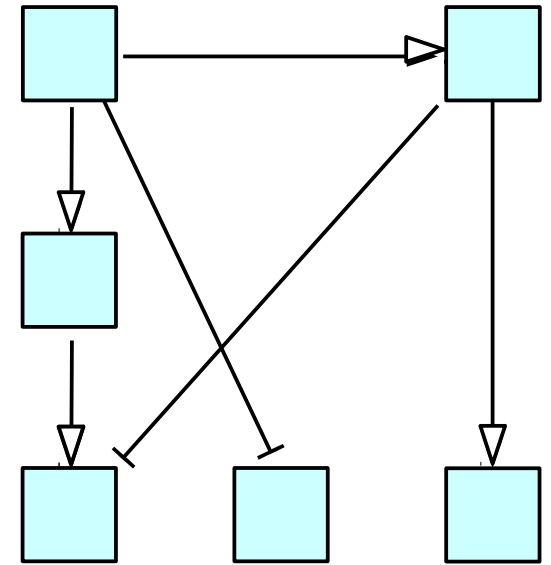
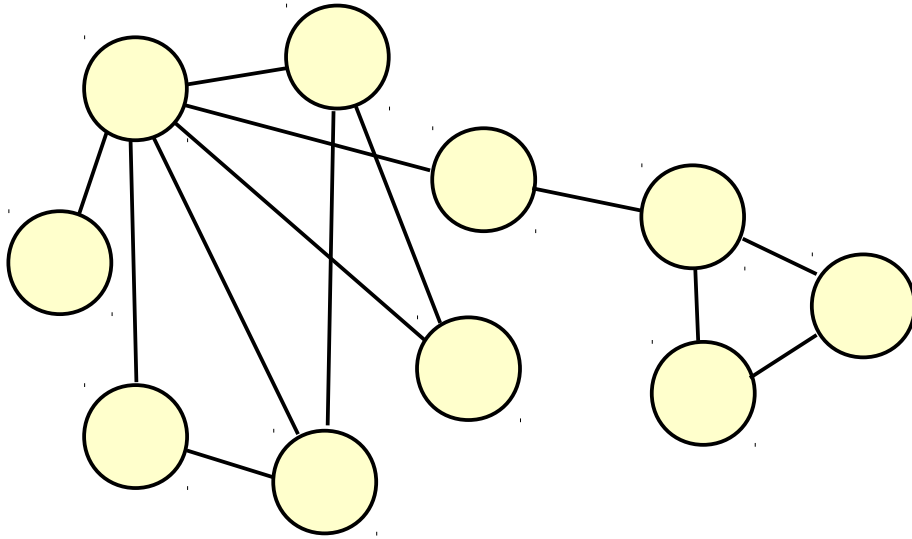


- Directional
- Non-sequential
- Mechanistic
- Independent rules: no explosion
- Rule-based modelling
- Molecular Biology
- MIM

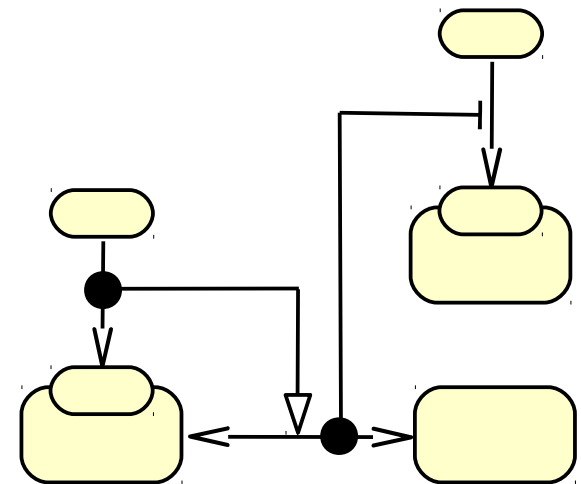
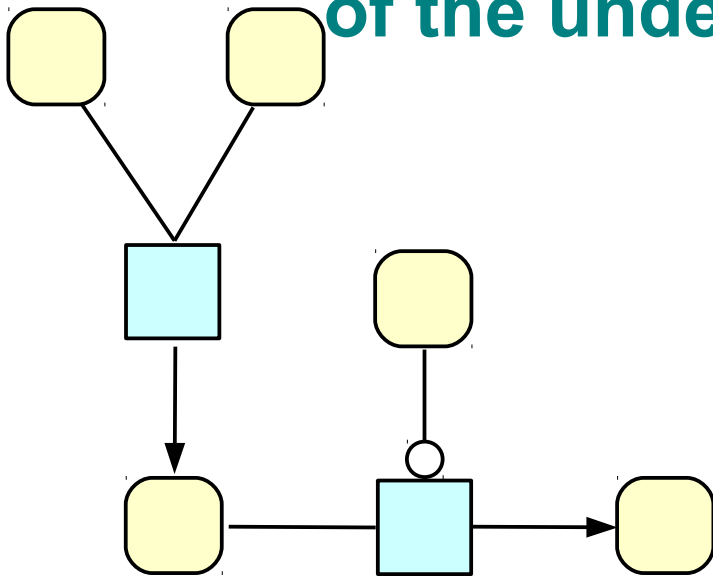
Process Descriptions

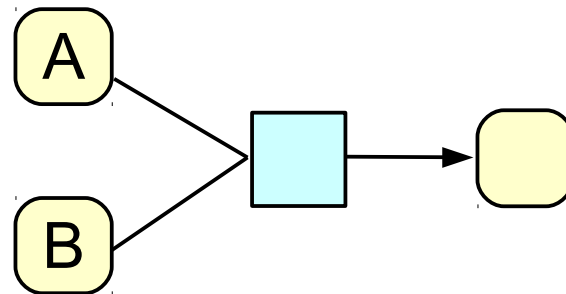
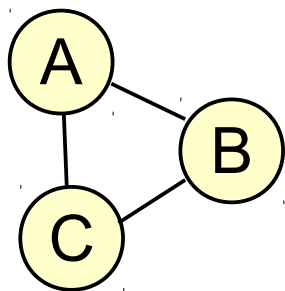


- Directional
- Sequential
- Mechanistic
- Subjected to combinatorial explosion
- **Process modelling**
- **Biochemistry**, Metabolic networks
- KEGG, Reactome

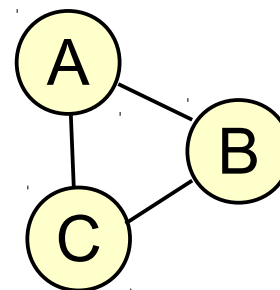
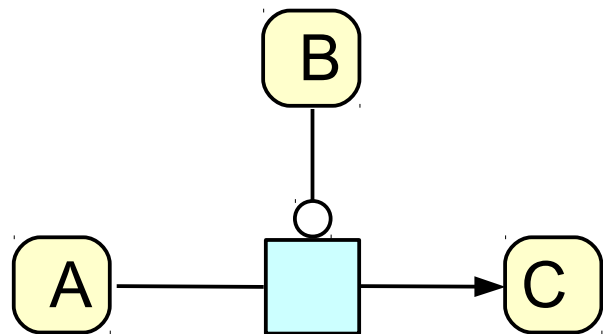


The four views are orthogonal projections of the underlying biological phenomena





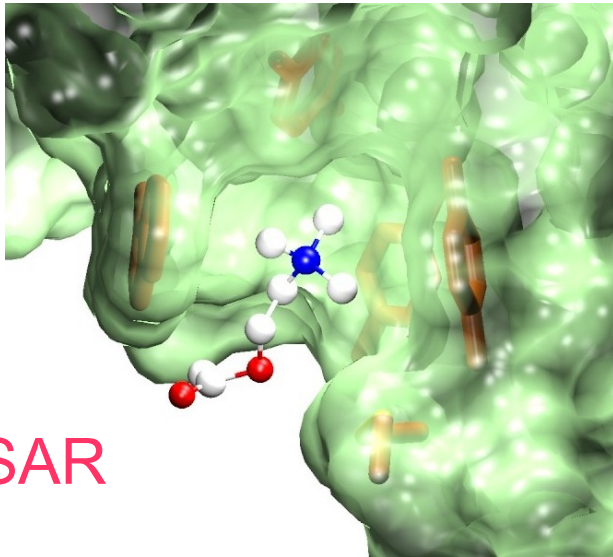
Think scaffolding proteins



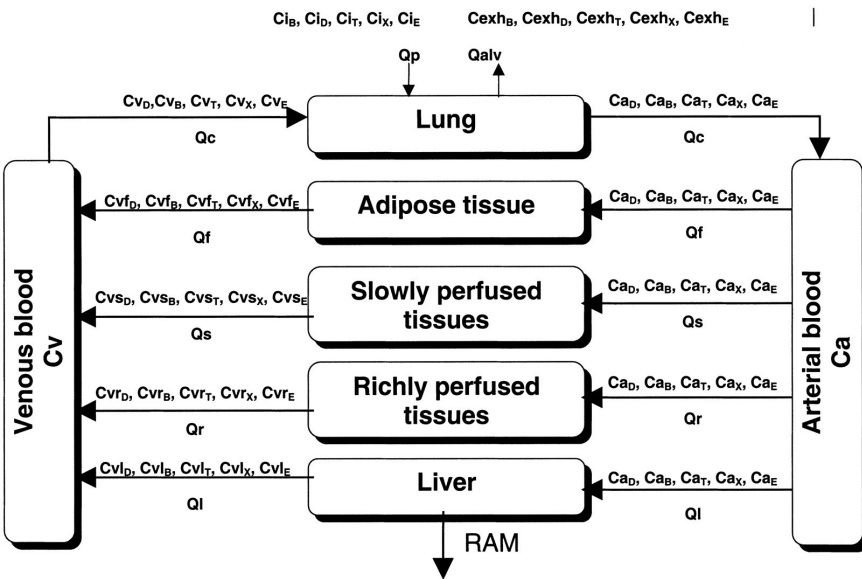
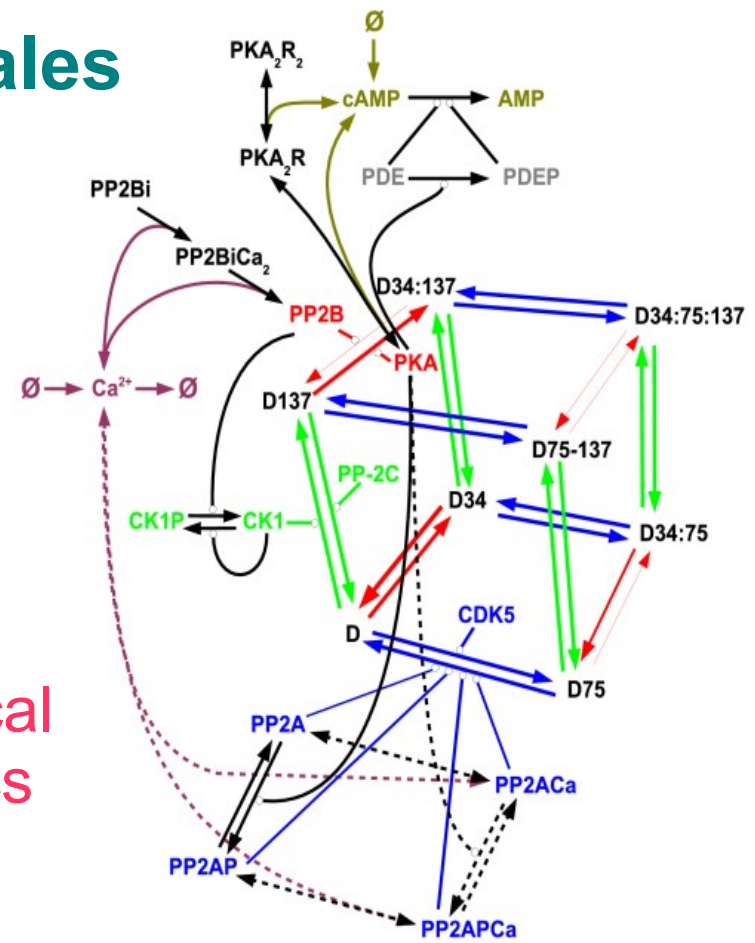
Think phosphorylated signals

The problem of scales

QSAR



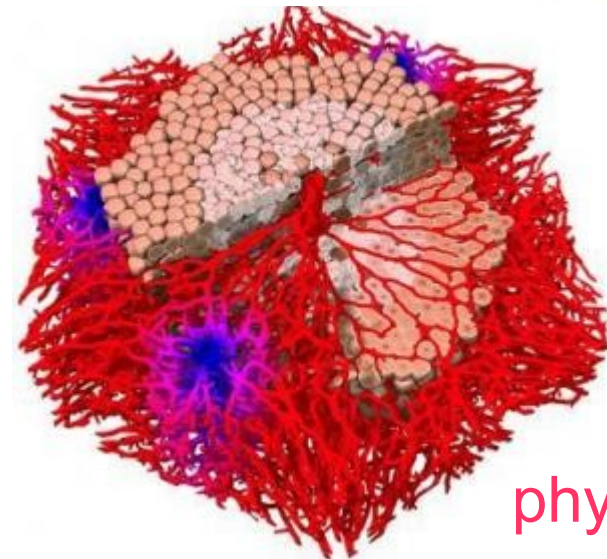
chemical kinetics



$$RAM_D = \frac{V_{max_D} \cdot C_{vD}}{K_{mD} \left(1 + \frac{C_{vD}}{K_{iD}} + \frac{C_{vR}}{K_{iR}} + \frac{C_{vL}}{K_{iL}} + \frac{C_{vE}}{K_{iE}} \right) + C_{vD}}$$

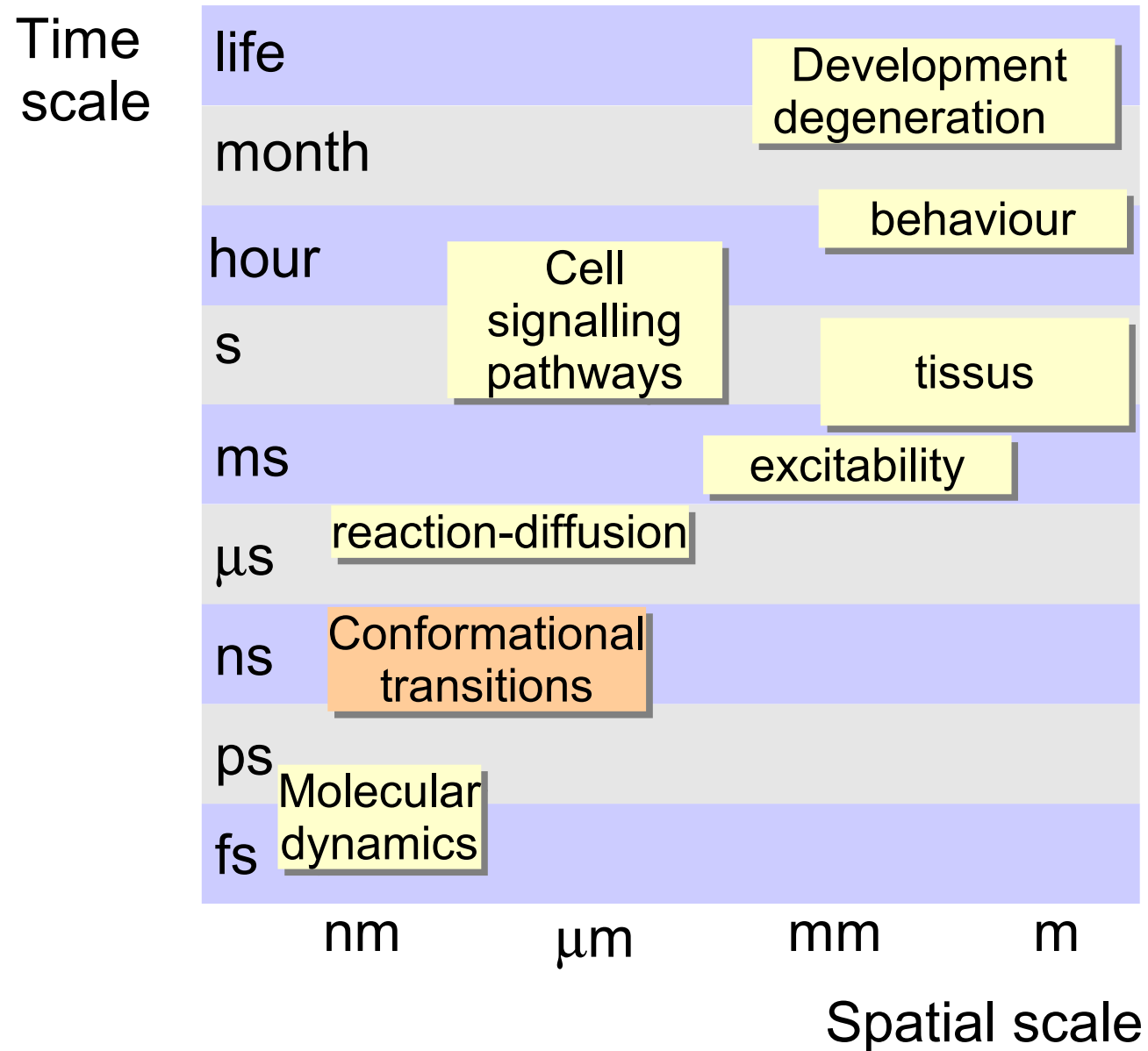
$$RAM_S = \frac{V_{max_S} \cdot C_{vS}}{K_{mS} \left(1 + \frac{C_{vD}}{K_{iD}} + \frac{C_{vR}}{K_{iR}} + \frac{C_{vL}}{K_{iL}} + \frac{C_{vE}}{K_{iE}} \right) + C_{vS}}$$

pharmacometrics

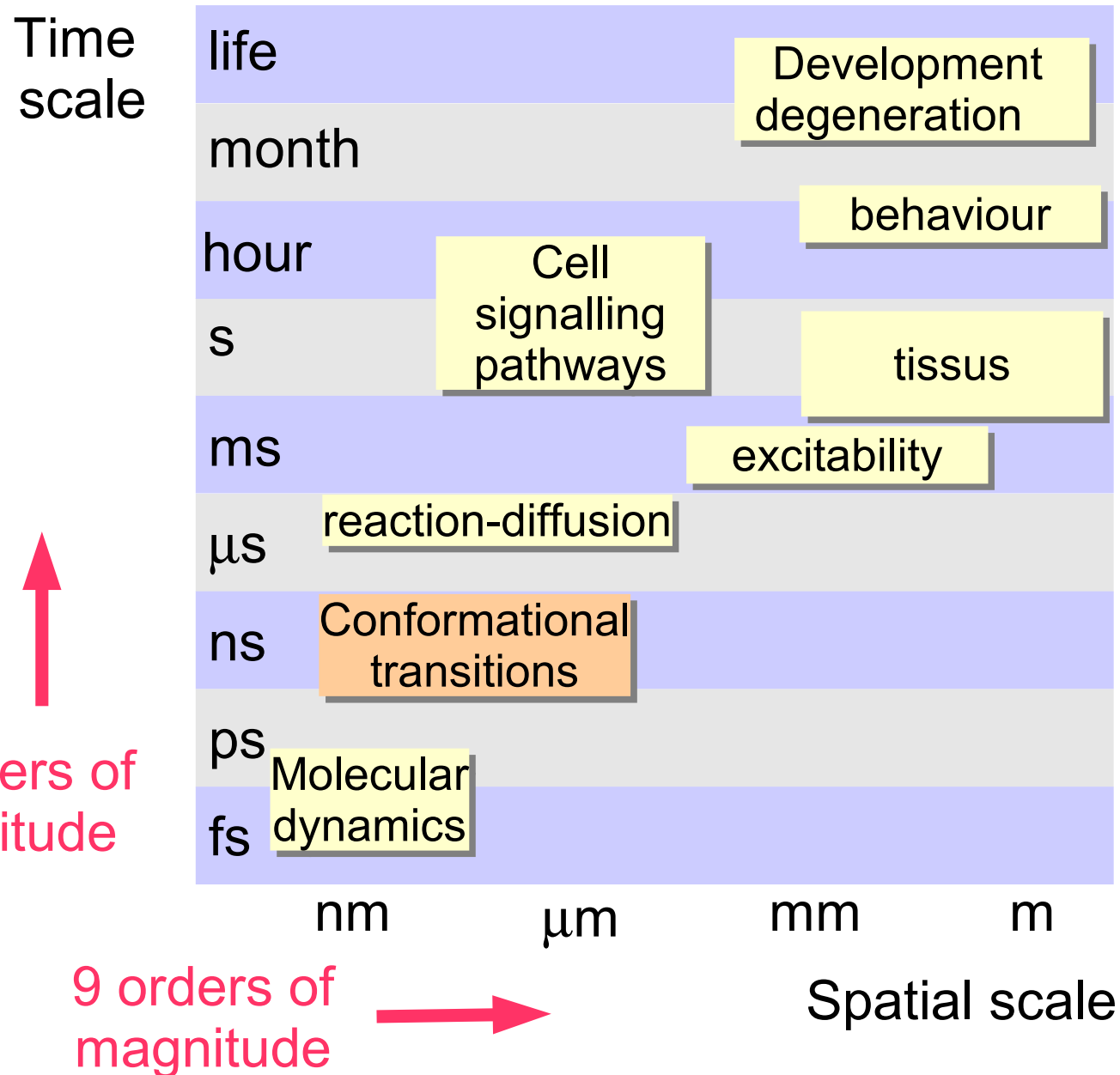


physiology

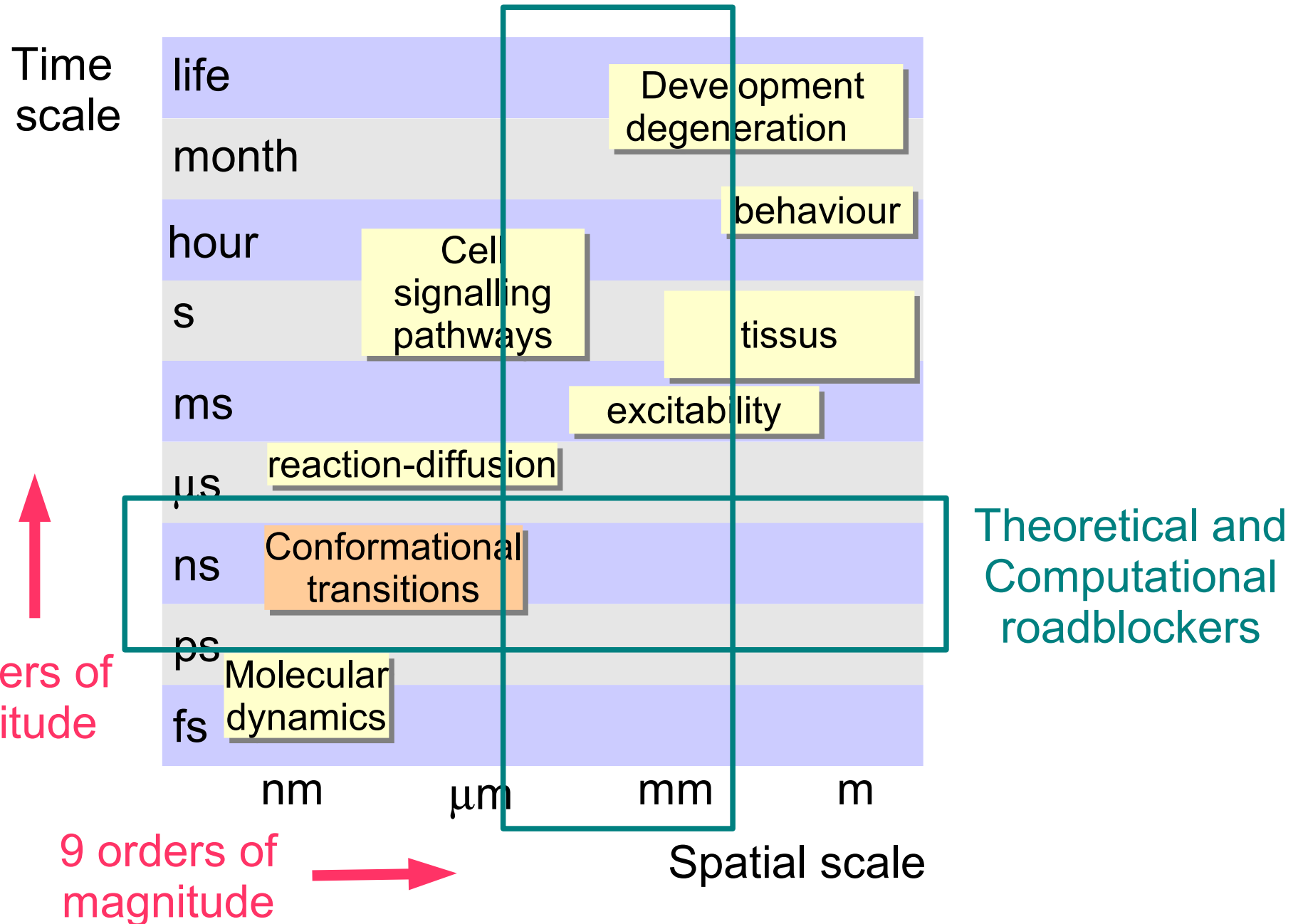
The mesoscopic gap



The mesoscopic gap



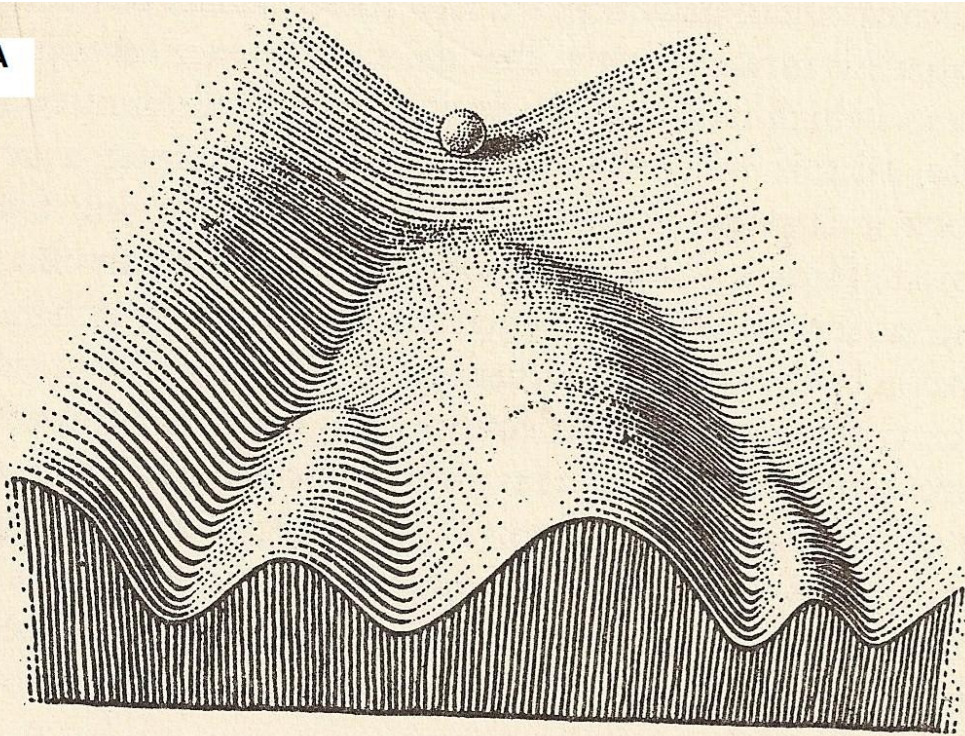
The mesoscopic gap



Multi-scale does not mean fine scale plus massive power

- Computational issues
 - In 2008, largest simulations were 1 million atoms for 50 ns (using > 30 years CPU) and 10000 atoms for 0.5 ms.
 - E-coli ~ 1 million million atoms ...
 - We need at least 10^{14} more power to simulate E coli during 1 s
- Theoretical issues
 - Molecular dynamics does not scale linearly with number of interactions
 - We do not know how to simulate mesoscopic phenomena: conformational transitions, secondary structure movements etc.

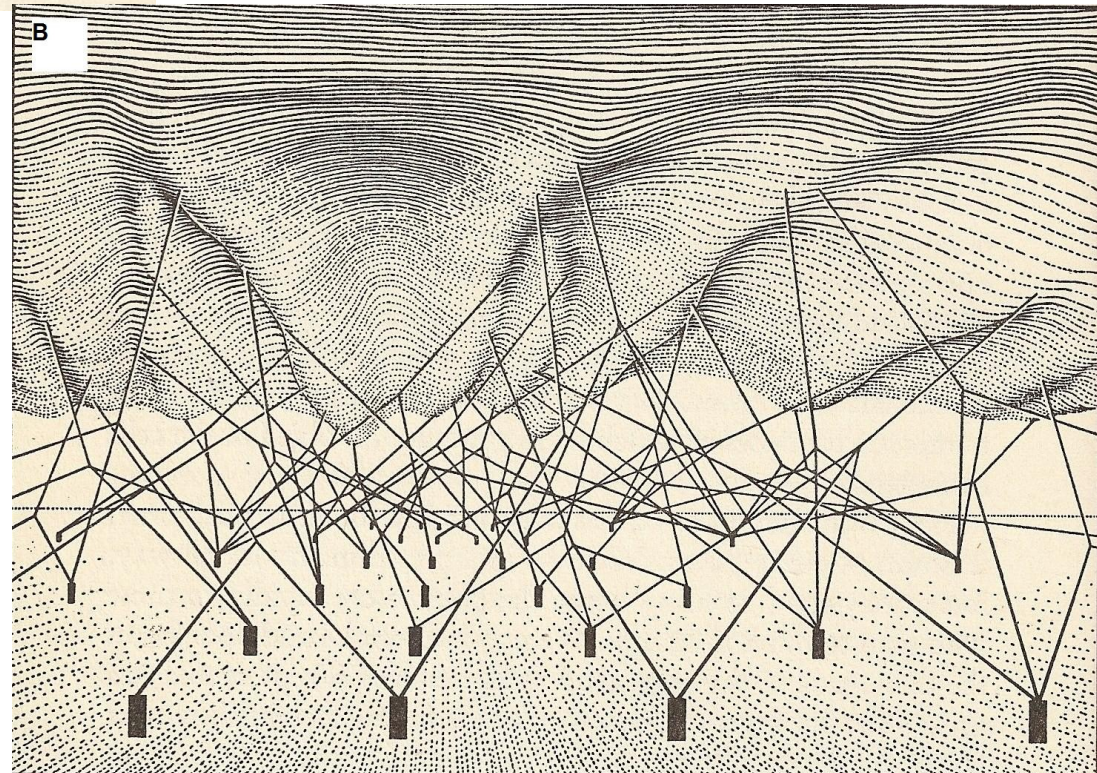
A



Emergent properties and the gene-system-phenotype puzzle

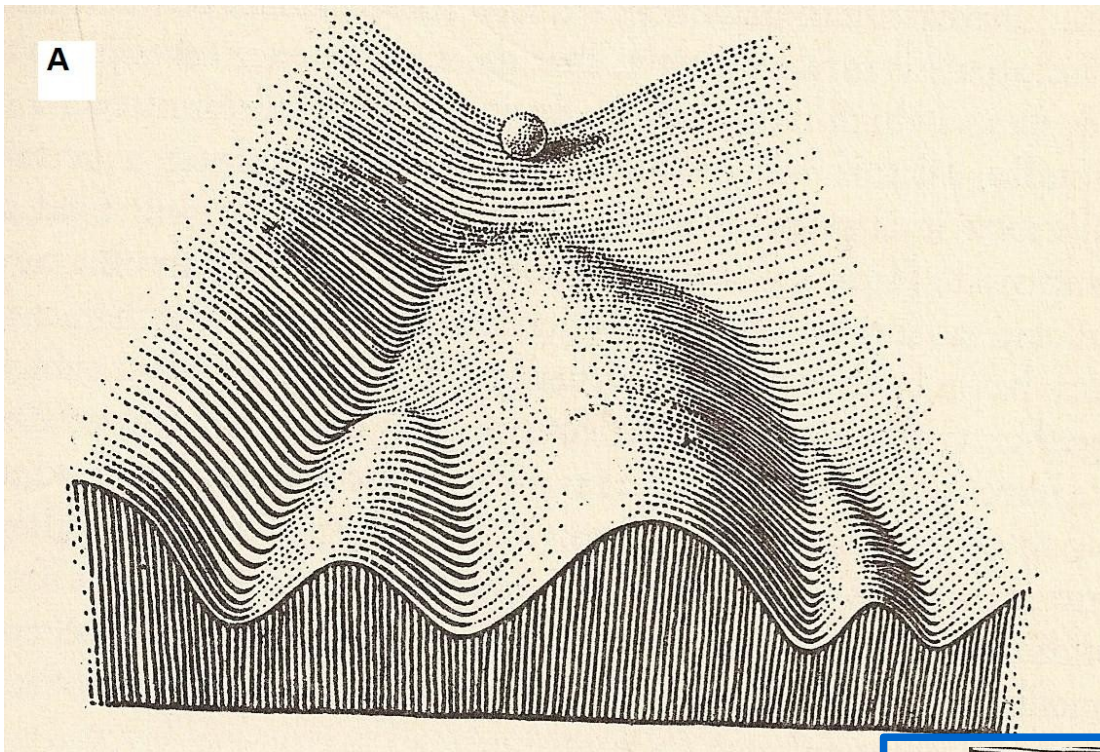
Waddington C.H., Kacser H (1957)
The Strategy of the Genes:
A Discussion of Some Aspects of
Theoretical Biology.
George Allen & Unwin

B



Emergent properties and the gene-system-phenotype puzzle

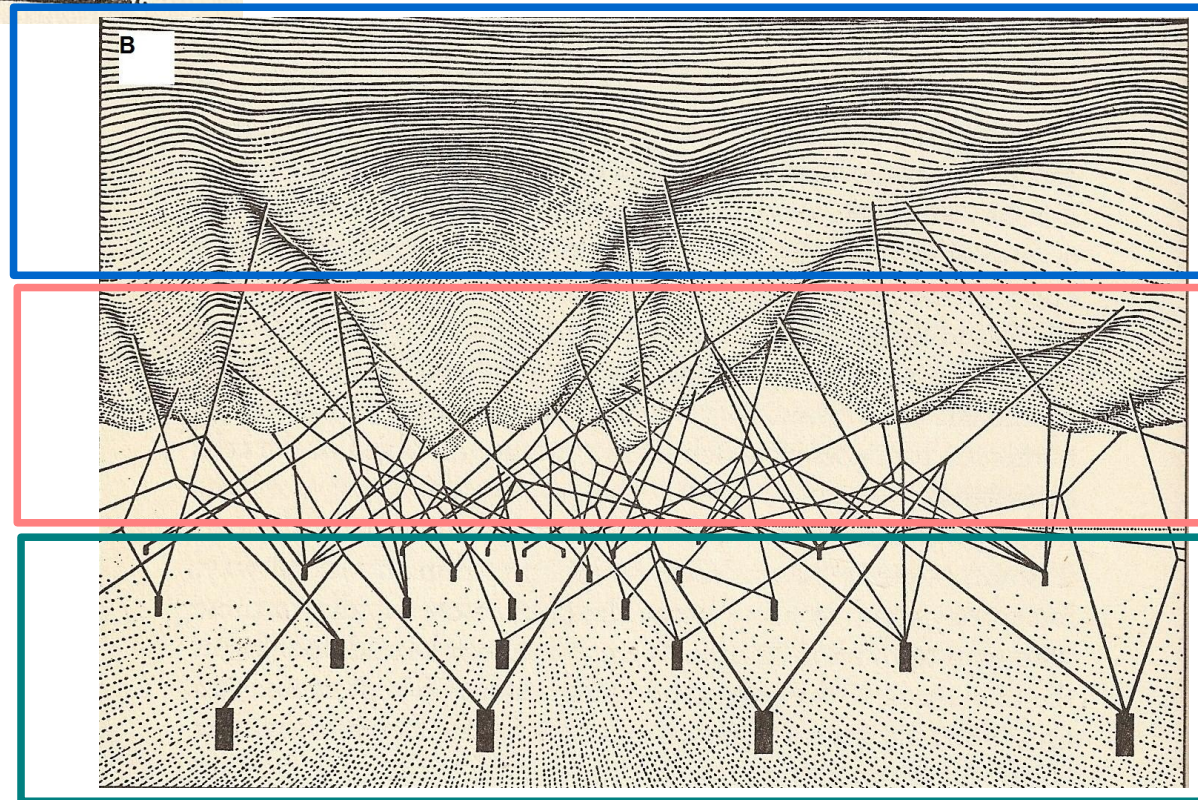
Waddington C.H., Kacser H (1957)
The Strategy of the Genes:
A Discussion of Some Aspects of
Theoretical Biology.
George Allen & Unwin



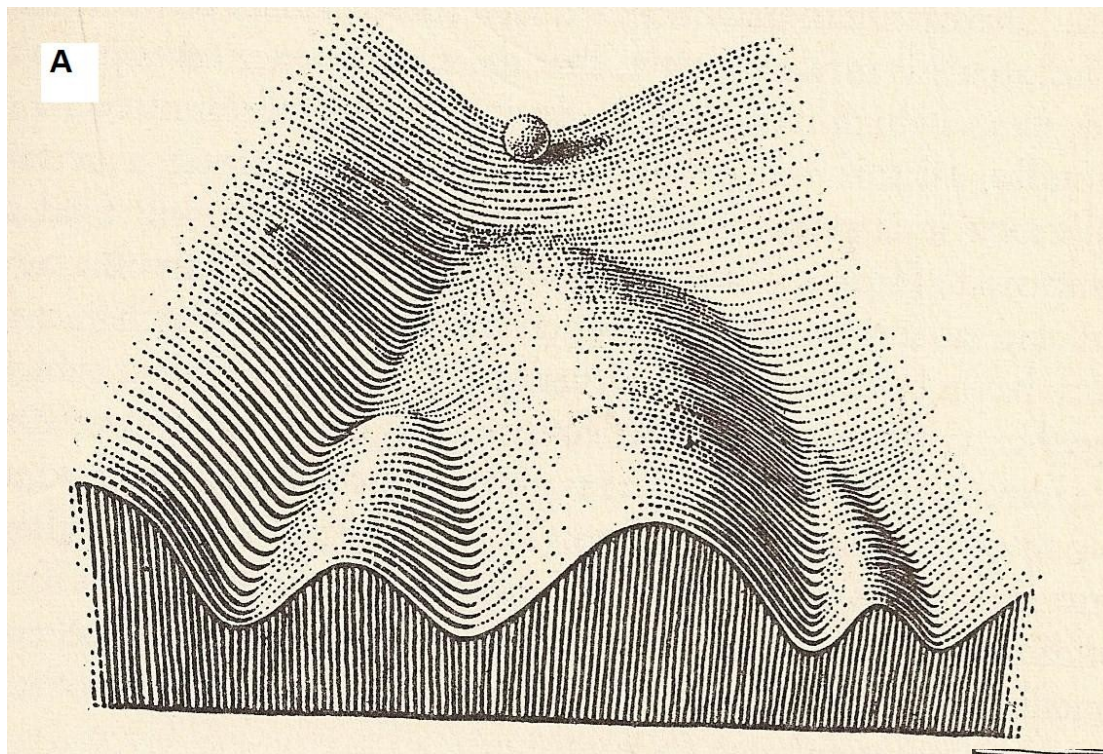
physiology ← phenotype

systems biology ← system

genetics ← genotype

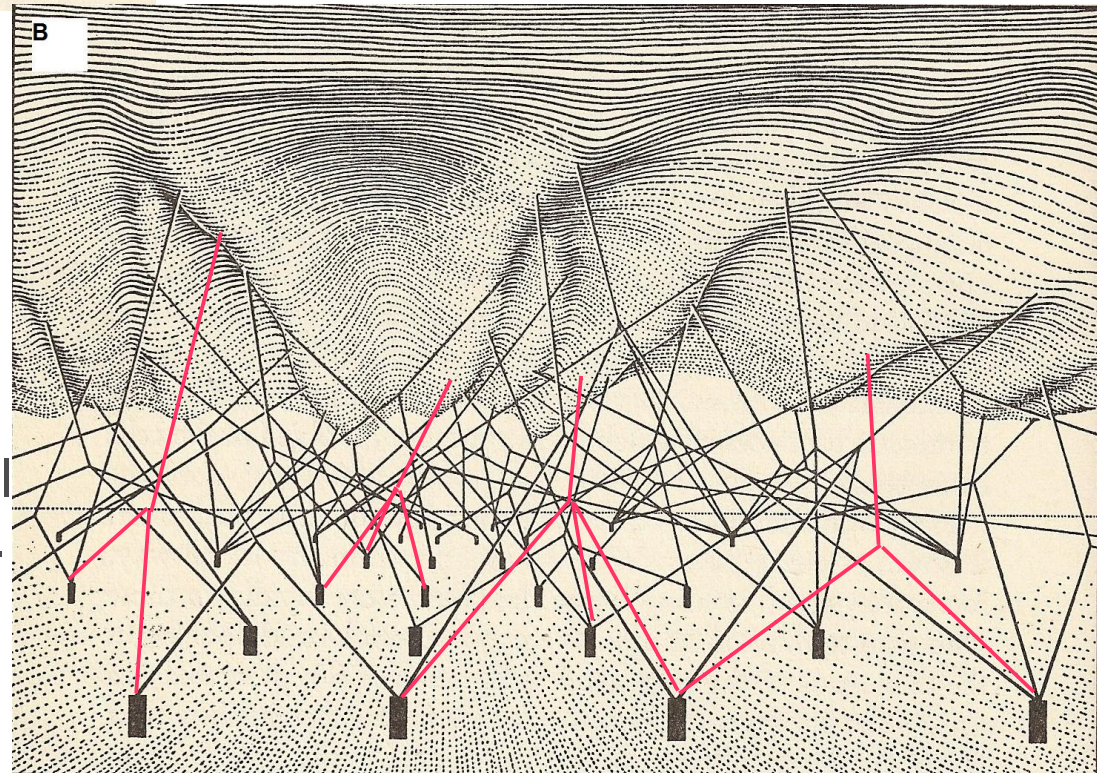


Emergent properties and the gene-system-phenotype puzzle



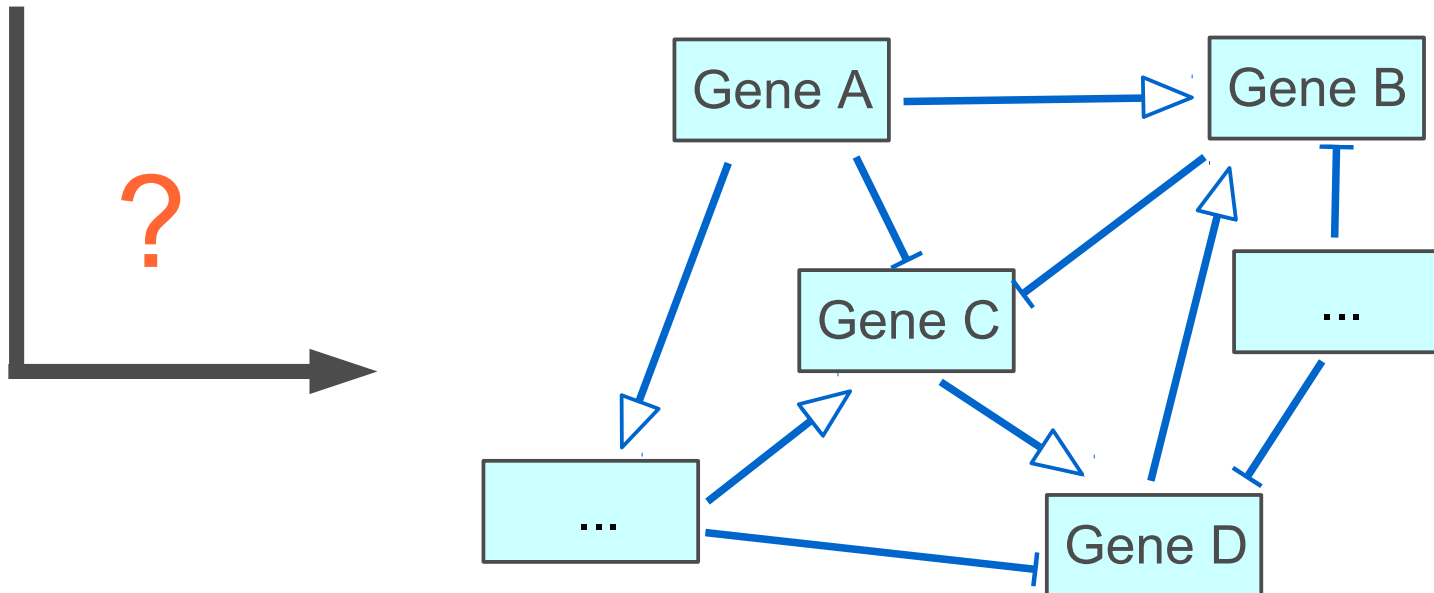
Waddington C.H., Kacser H (1957)
The Strategy of the Genes:
A Discussion of Some Aspects of
Theoretical Biology.
George Allen & Unwin

Many networks can theoretically generate the same phenotype, and this happens, in a synchronous (sister cells with same phenotype but different transcript/prote/metabol/omes) and diachronous manner (*omes of a cell changes over time but same phenotype).

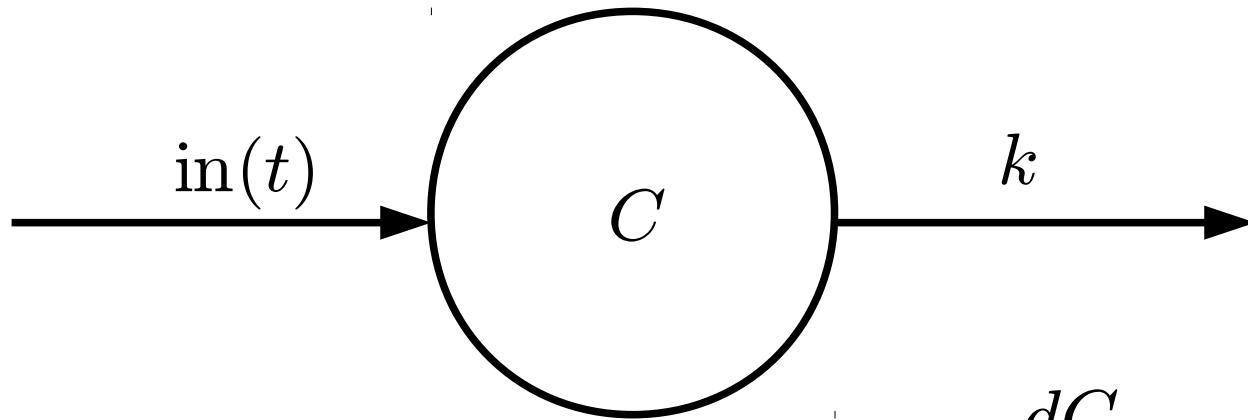


Reverse engineering is hard ...

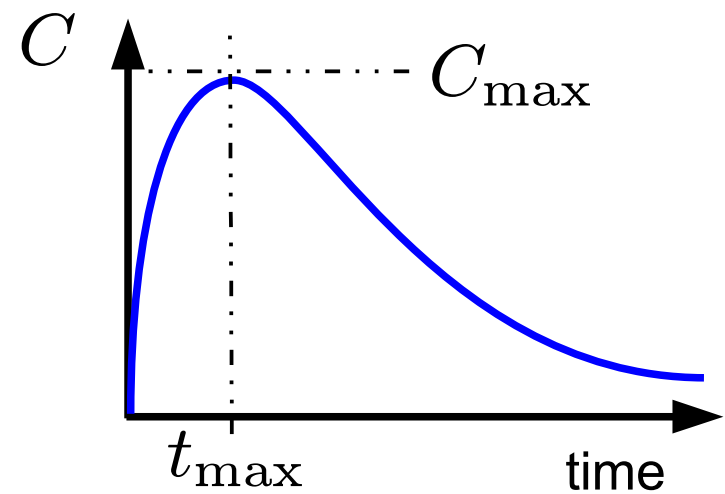
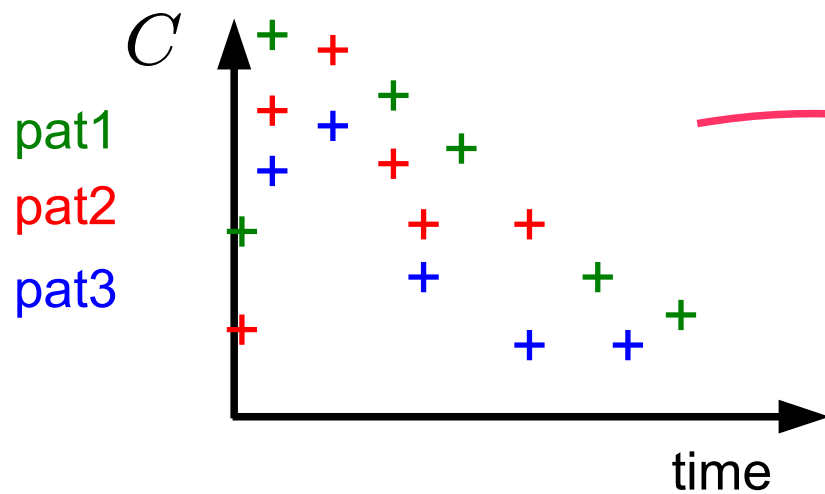
	Gene A	Gene B	Gene C	Gene D	...
Phenotype X	✓	✗	✓	✗	
Phenotype Y	✓	✗	✗	✓	
Phenotype Z	✗	✓	✓	✗	
...					



Drug discovery modelling: pharmacometrics

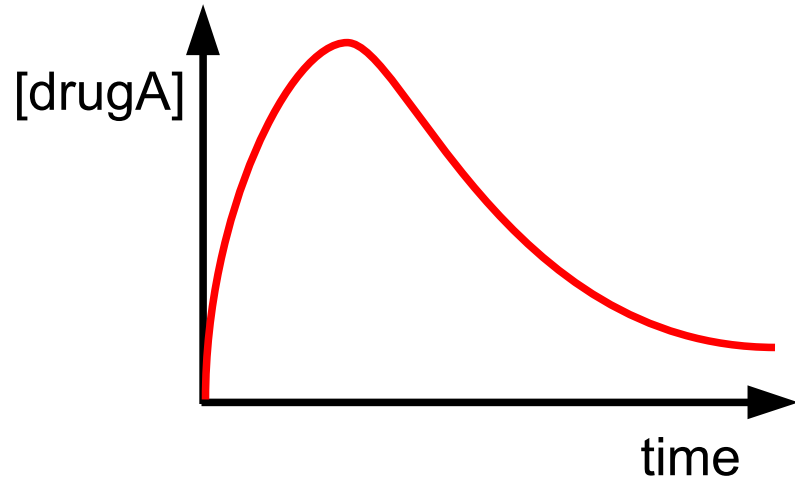


$$\frac{dC}{dt} = \text{in}(t) - k \times C$$

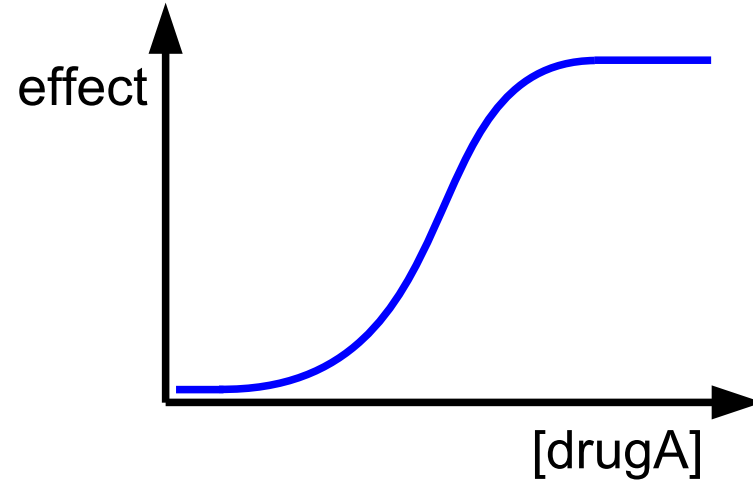


Drug discovery modelling: pharmacometrics

PK

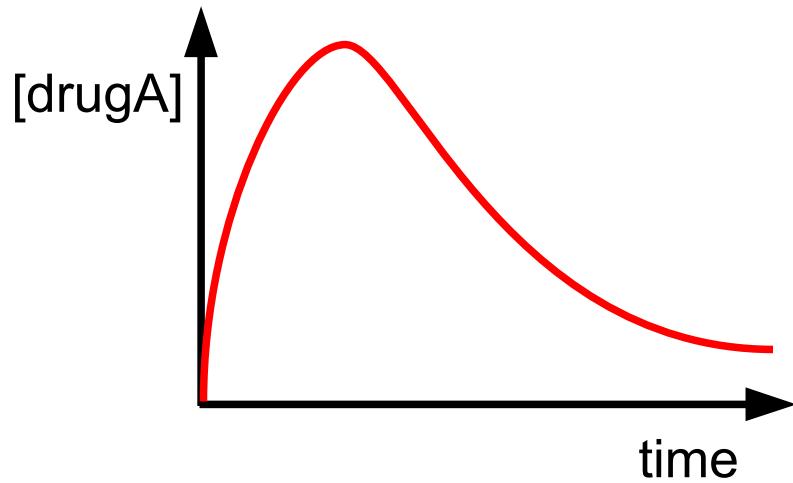


PD

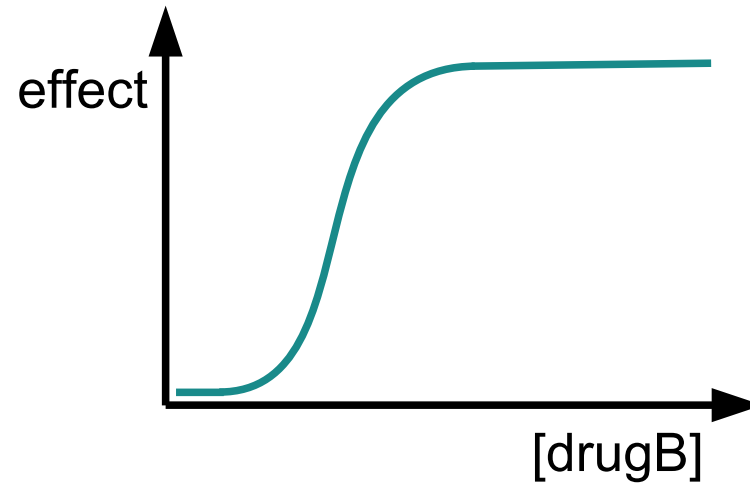
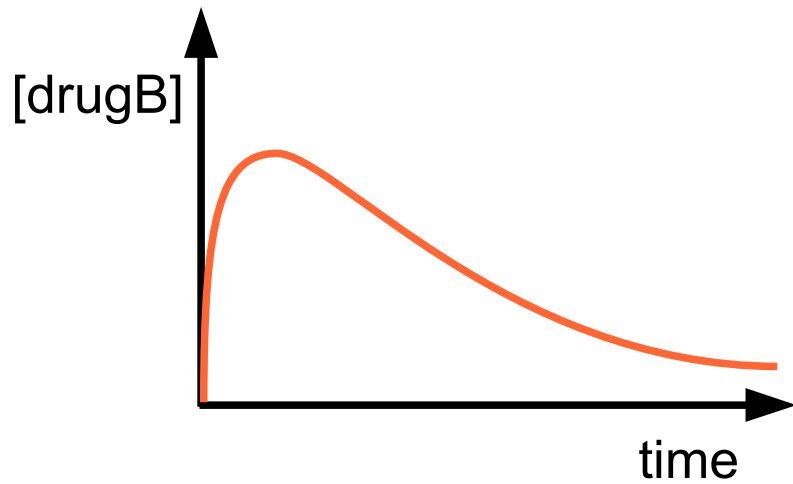
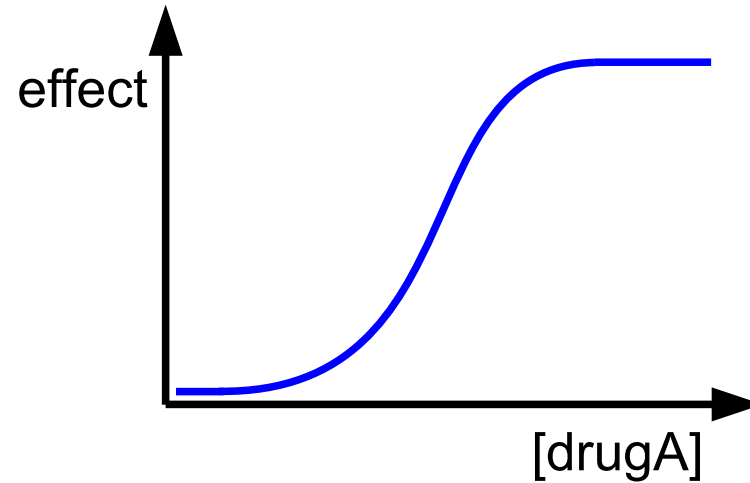


Drug discovery modelling: pharmacometrics

PK

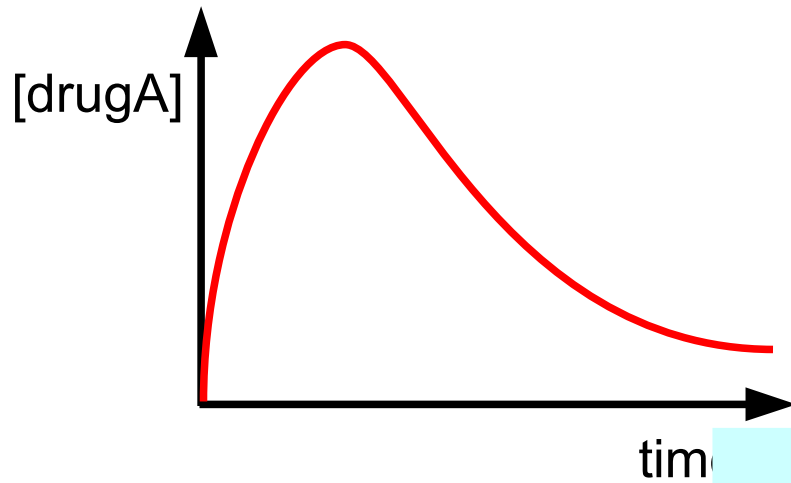


PD

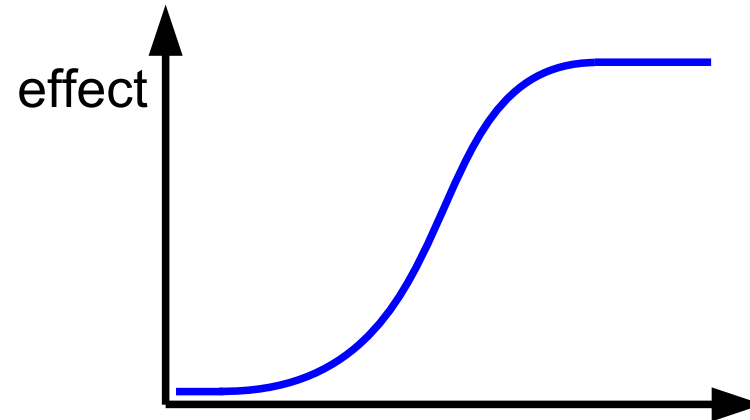


Drug discovery modelling: pharmacometrics

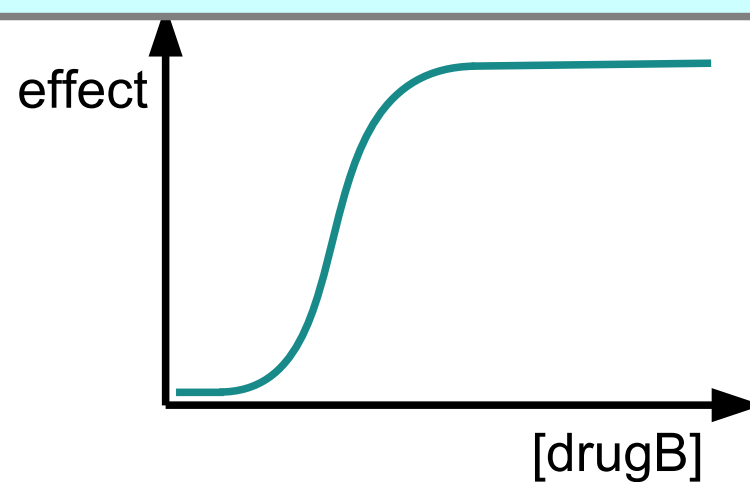
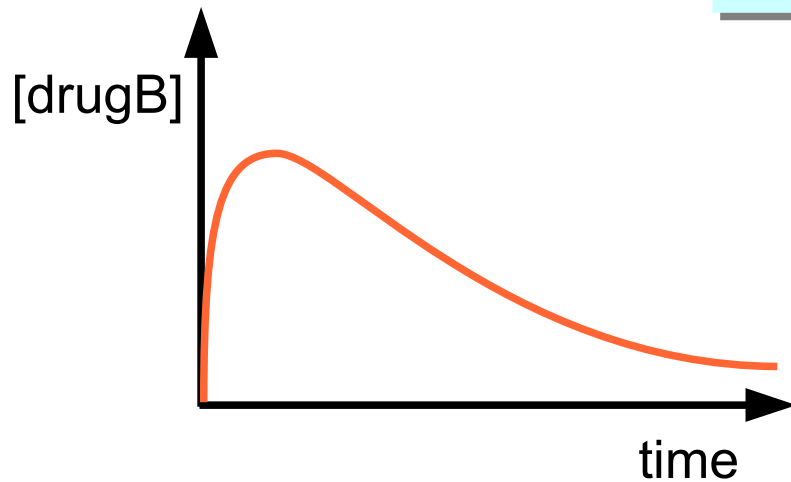
PK



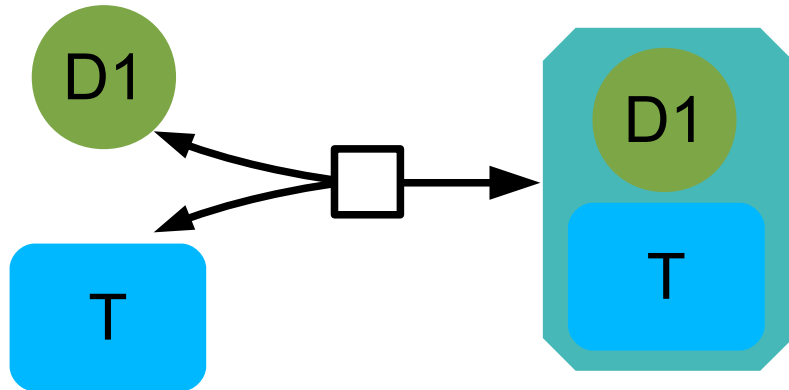
PD



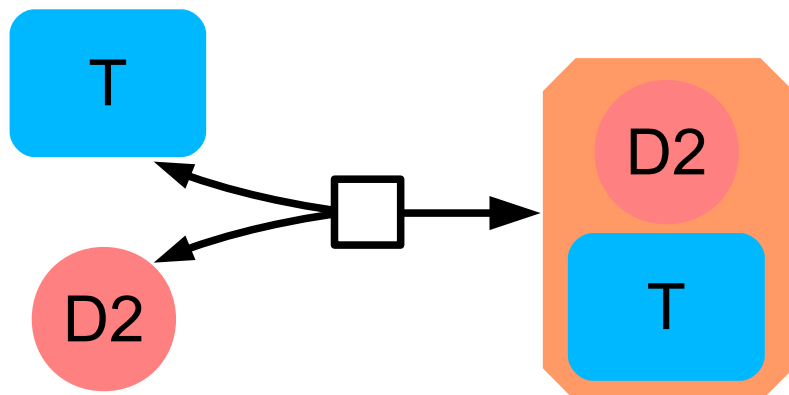
Effect of A+B?



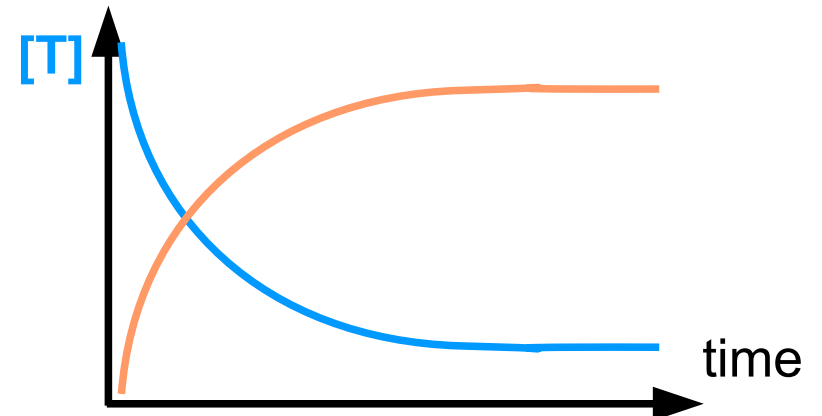
Systems modelling



$$\frac{d[T]}{dt} = -k_1 \times [T] \times [D1]$$

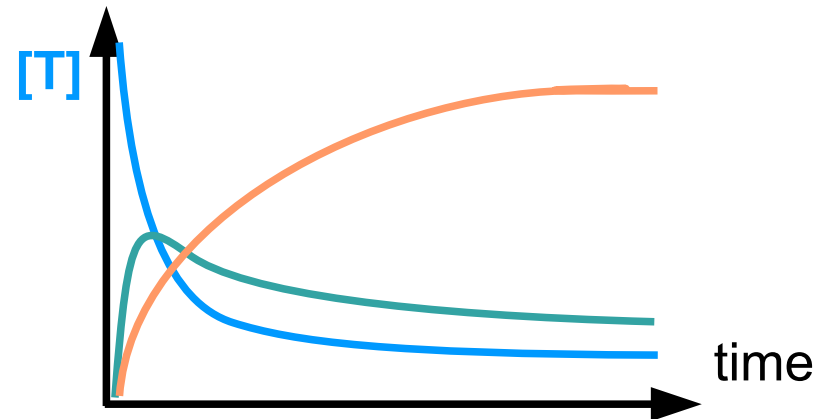
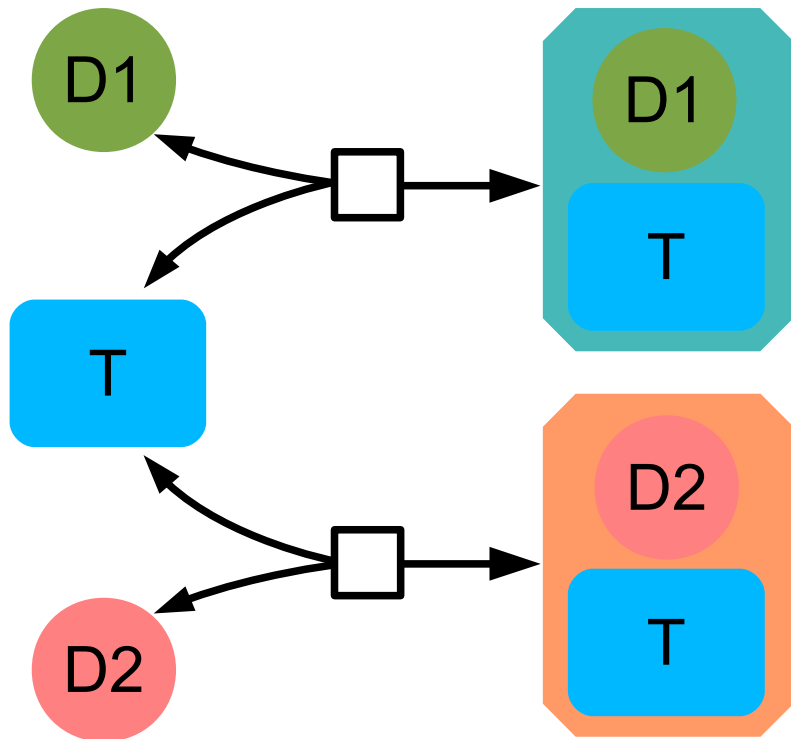


$$\frac{d[T]}{dt} = -k_2 \times [T] \times [D2]$$

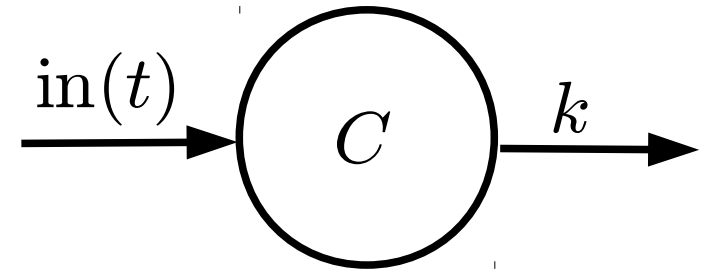
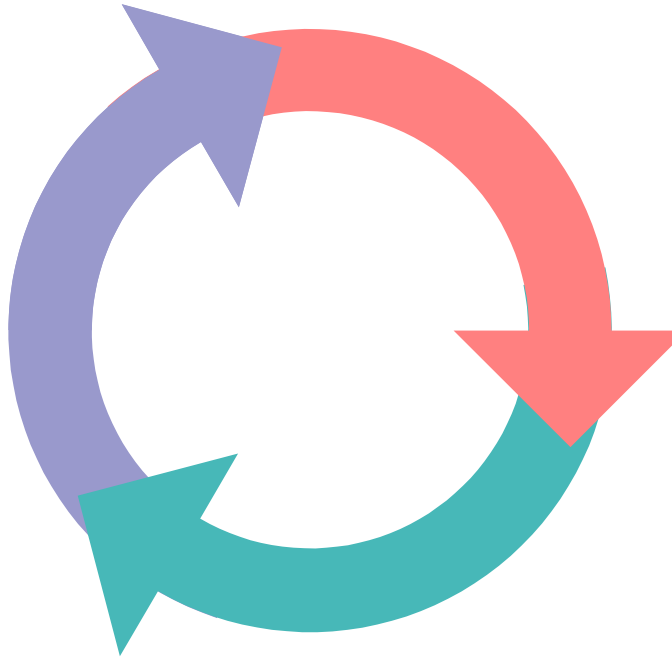
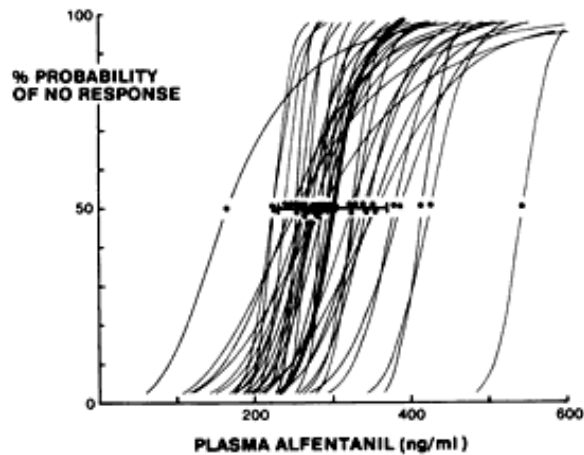
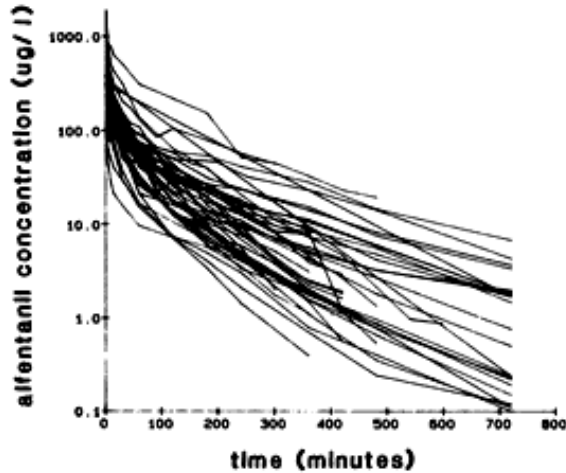


Systems modelling

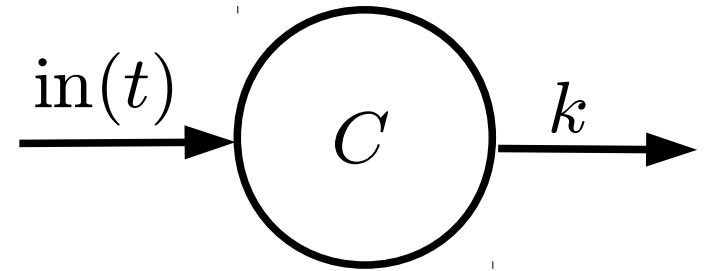
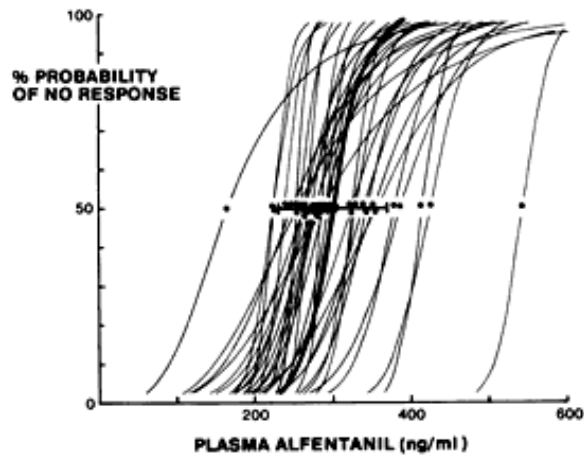
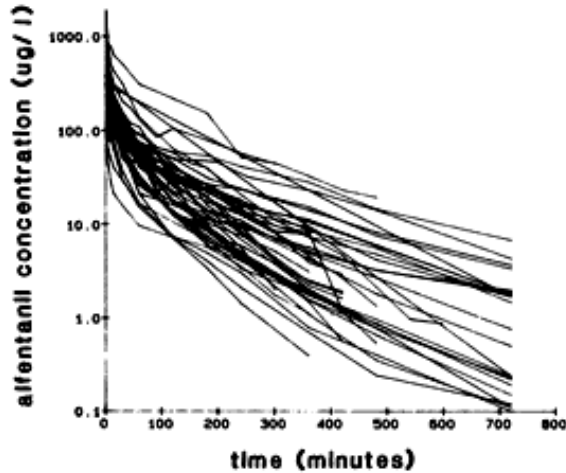
$$\frac{d[T]}{dt} = -k1 \times [T] \times [D1] - k2 \times [T] \times [D2]$$



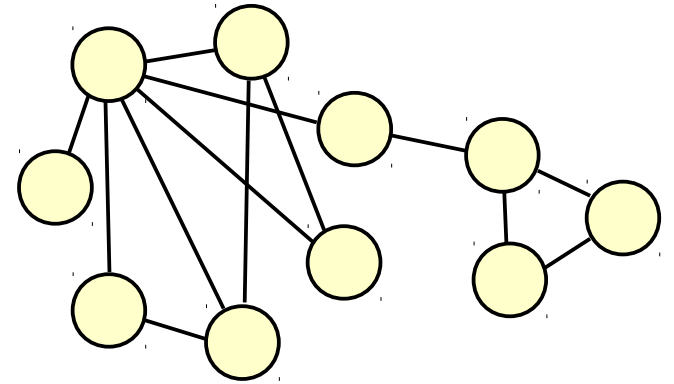
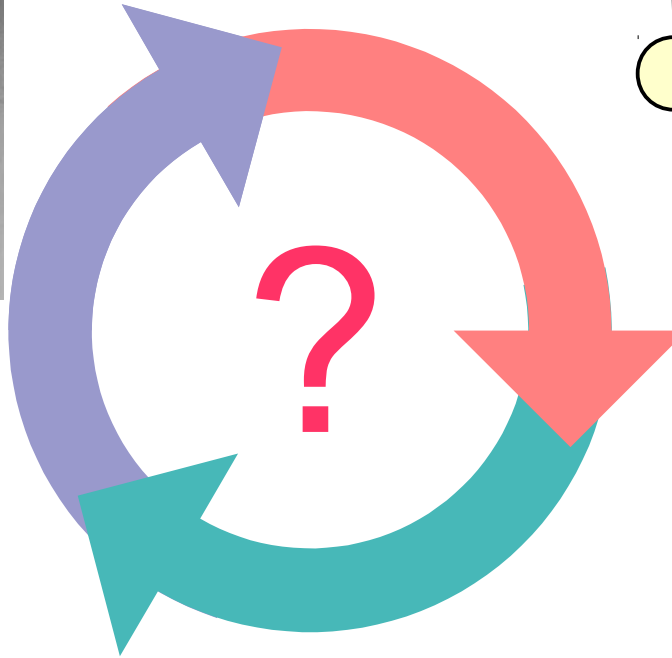
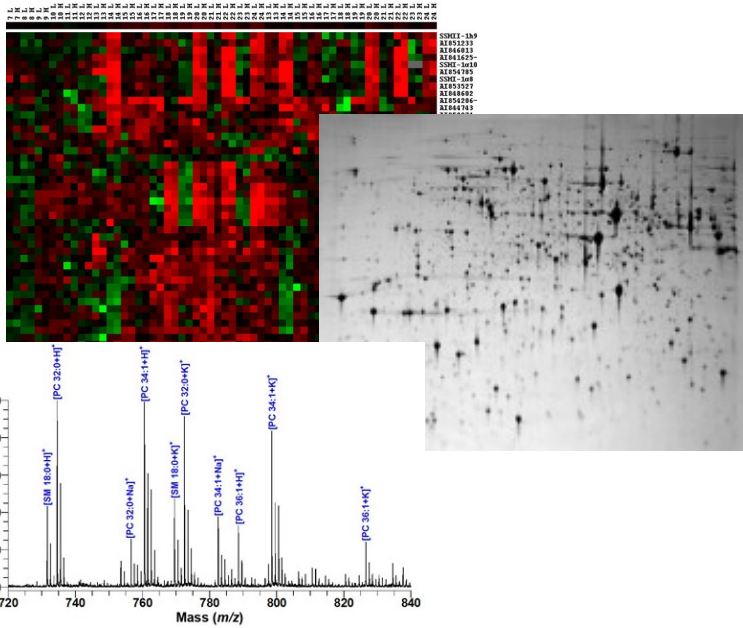
Drug discovery and pharmacometrics models



Drug discovery and pharmacometrics models



Drug discovery and omics



Systems Biology

Edda Klipp, Wolfram Liebermeister, Christoph Wierling,
Axel Kowald, Hans Lehrach, and Ralf Herwig

A Textbook

 WILEY-
BLACKWELL

DYNAMIC MODELS

 IN BIOLOGY

STEPHEN P. ELLNER AND
JOHN GUCKENHEIMER



