Nov 8: Eric Deeds, University of California at Los Angeles
"The evolution of cellular individuality"

Nov 15: Daniel Merkle, University of Southern Denmark
"Graph rewriting and chemistry"

Nov 22: Jean Krivine, IRIF, Université de Paris
"From molecules to systems: the problem of knowledge representation in molecular biology"

Nov 29: Eric Smith, Earth Life Sciences Institute, Tokyo
“Easy and Hard in the Origin of Life"

Dec 6: Massimiliano Esposito, University of Luxembourg
"Thermodynamics of Open Chemical Reaction Networks: Theory and Applications"

Dec 13: Yarden Katz, Harvard Medical School
"Cells as cognitive creatures"

Jan 17: Aleksandra Walczak, ENS Paris
"Prediction in immune repertoires"

Jan 24: Tommy Kirchhausen, Harvard Medical School
"Imaging sub-cellular dynamics from molecules to multicellular organisms"

Previous Lectures And Look-Ahead

1. The Topology of the Possible
   (La représentation de l'information biologique)

2. Propagation of Genetic, Phenotypic, and Molecular information
   (Limites de la transmission de l'information biologique)

3. Modeling cellular information processing the classical way
   (Modélisation 'classique' du traitement de l'information cellulaire)

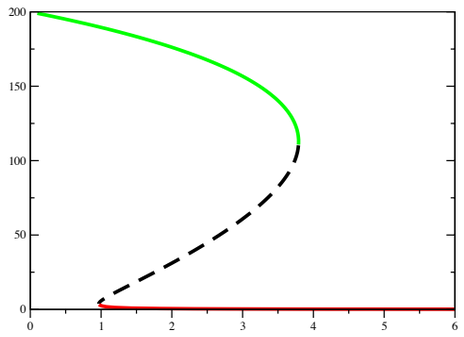
4. Modeling cellular information processing the rule-based way
   (Modélisation basé sur les règles; introduction)

5. Examples of rule-based models
   (Modélisation basé sur les règles; examples)

6. Causality in rule-based dynamics
   (Causalité)

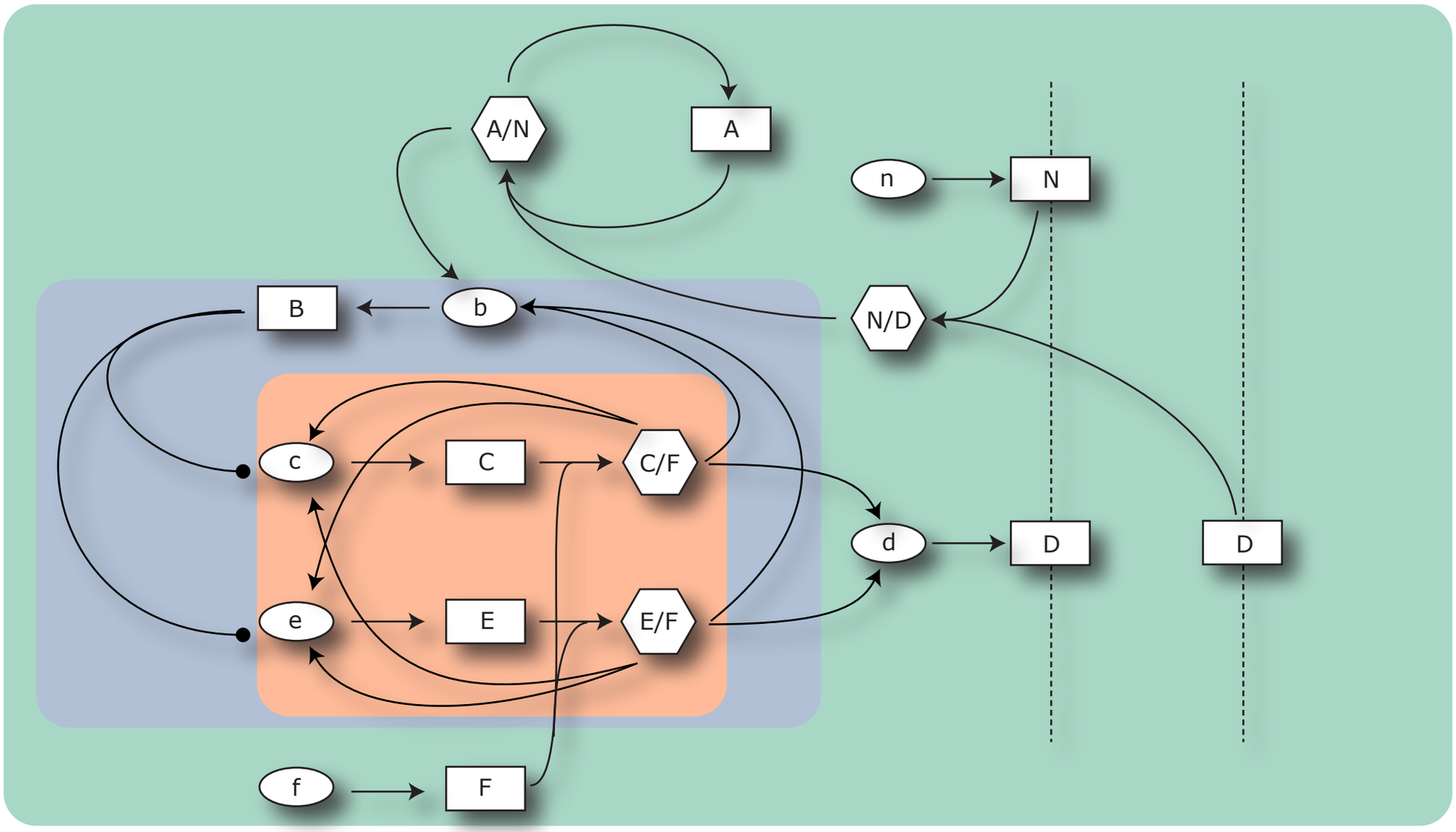
7. Combinatorial scaffolding
   (Echafaudage combinatoire)

8. Cellular learning?
   (Apprentissage cellulaire?)

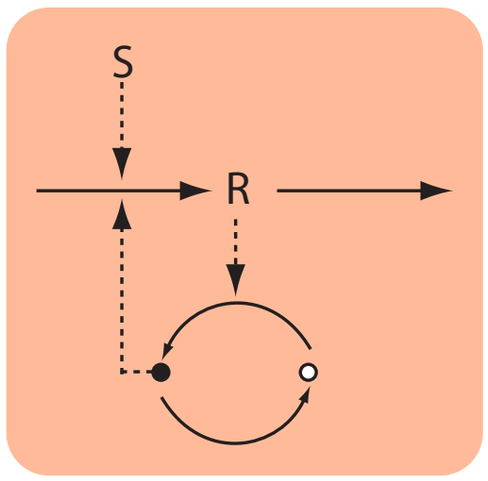


substrate S

Delta / Notch: Loops All The Way Down

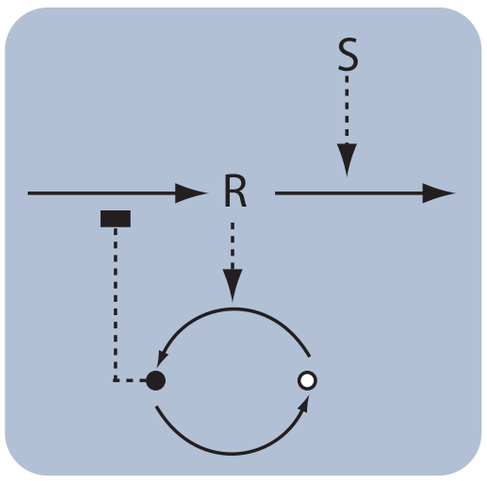


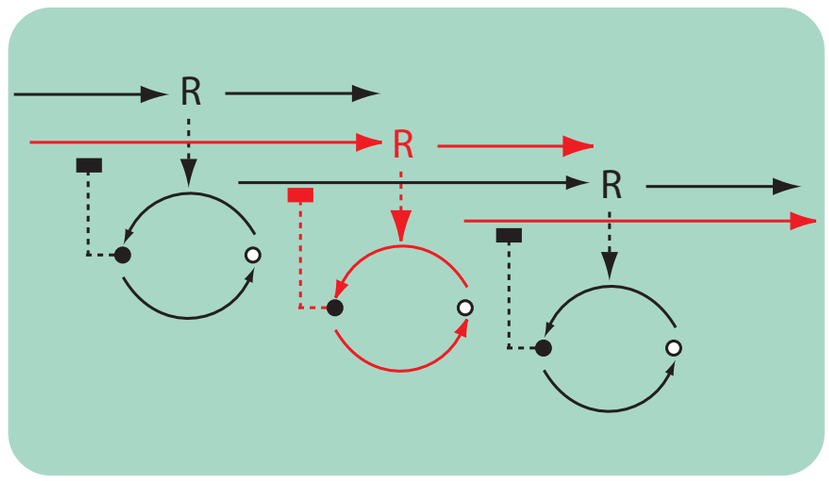
spatially interrupted "homeostat"



a bistable switch

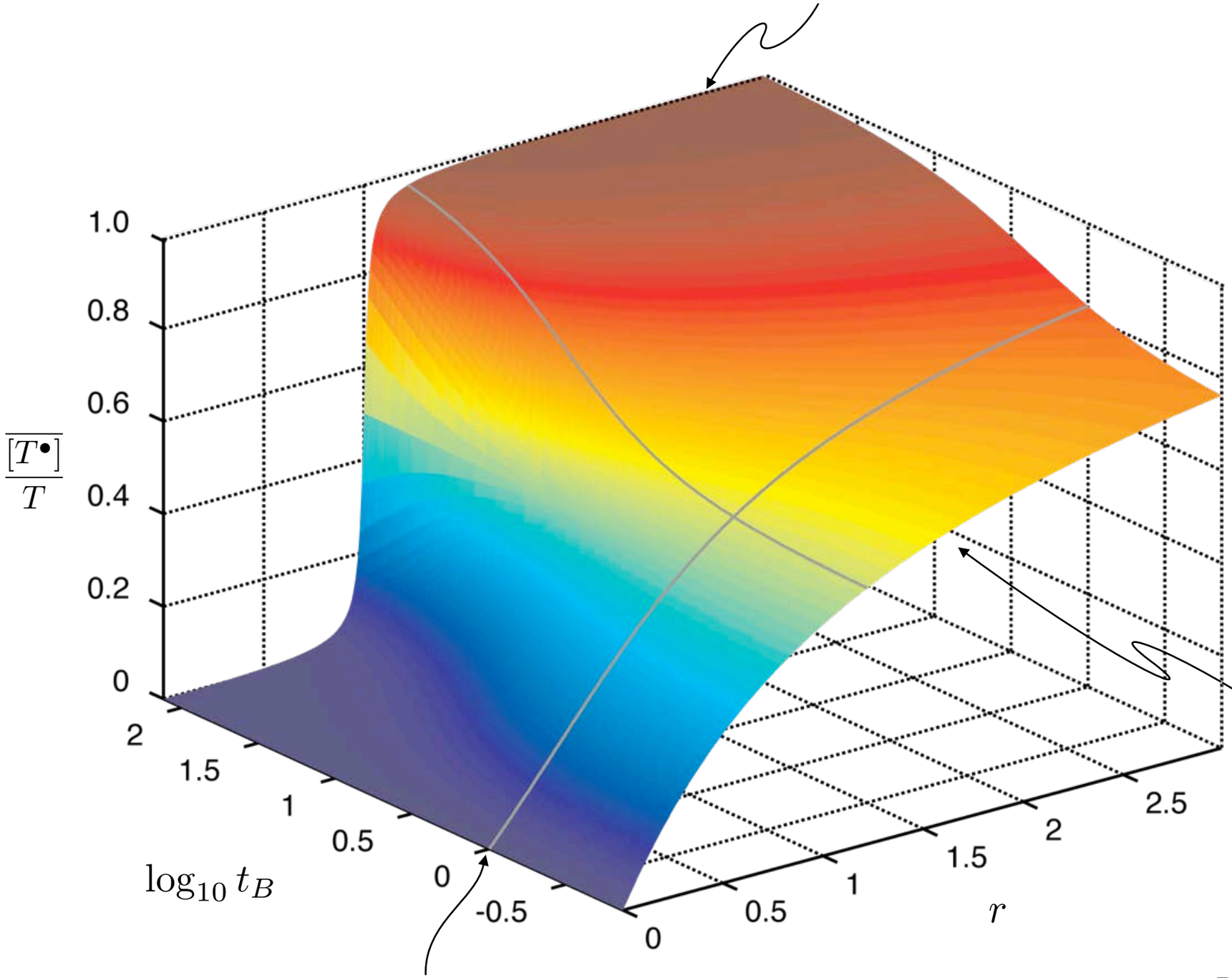
a homeostat



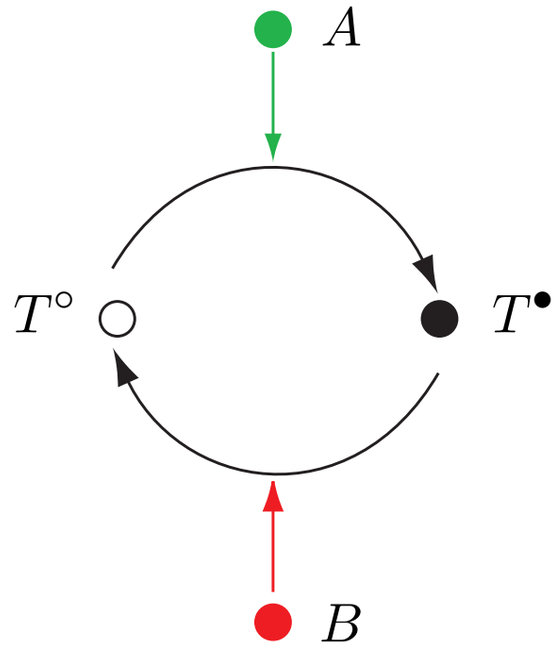


The Do-Undo Loop: An "Atom" Of Molecular Control

steady-state solution in saturated regime



$K_m^A = K_m^B = 0.74\ \mu M$



steady-state solution in linear regime

assuming $K = \frac{K_m^B}{K_m^A} = 1$

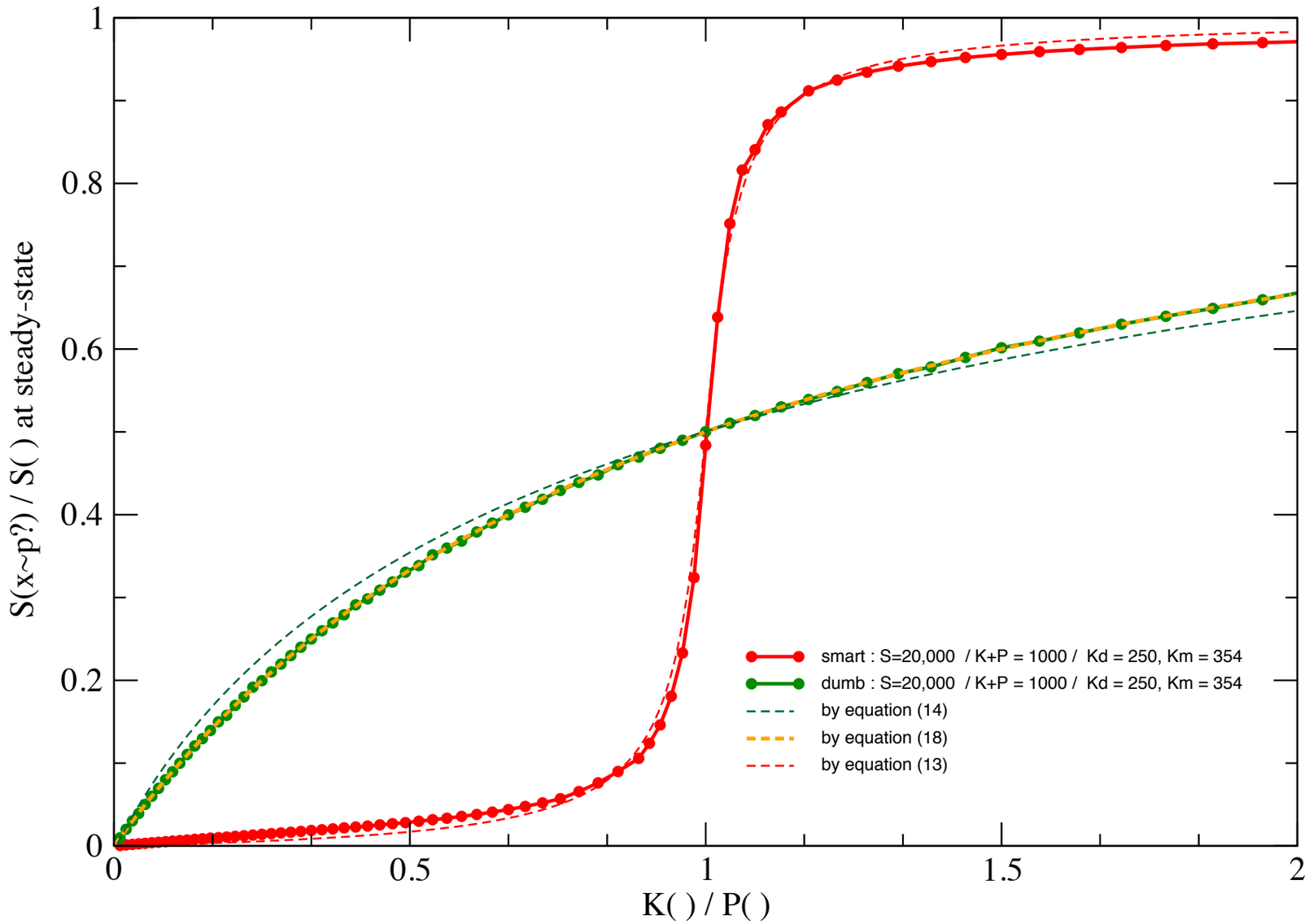
Details Matter

$$K + S^{\circ} \underset{k_{-1}^{K}}{\overset{k_{1}^{K}}{\rightleftharpoons}} KS^{\circ} \xrightarrow{k_{2}^{K}} K + S^{\bullet}$$

$$P + S^{\bullet} \underset{k_{-1}^{P}}{\overset{k_{1}^{P}}{\rightleftharpoons}} PS^{\bullet} \xrightarrow{k_{2}^{P}} P + S^{\circ}$$

$$K + S \underset{k_{-1}^{K}}{\overset{k_{1}^{K}}{\rightleftharpoons}} KS \qquad KS^{\circ} \xrightarrow{k_{2}^{K}} K + S^{\bullet}$$

$$P + S \underset{k_{-1}^{P}}{\overset{k_{1}^{P}}{\rightleftharpoons}} PS \qquad PS^{\bullet} \xrightarrow{k_{2}^{P}} P + S^{\circ}$$



Models Are...

...maps, not the territory

...about converting facts into knowledge

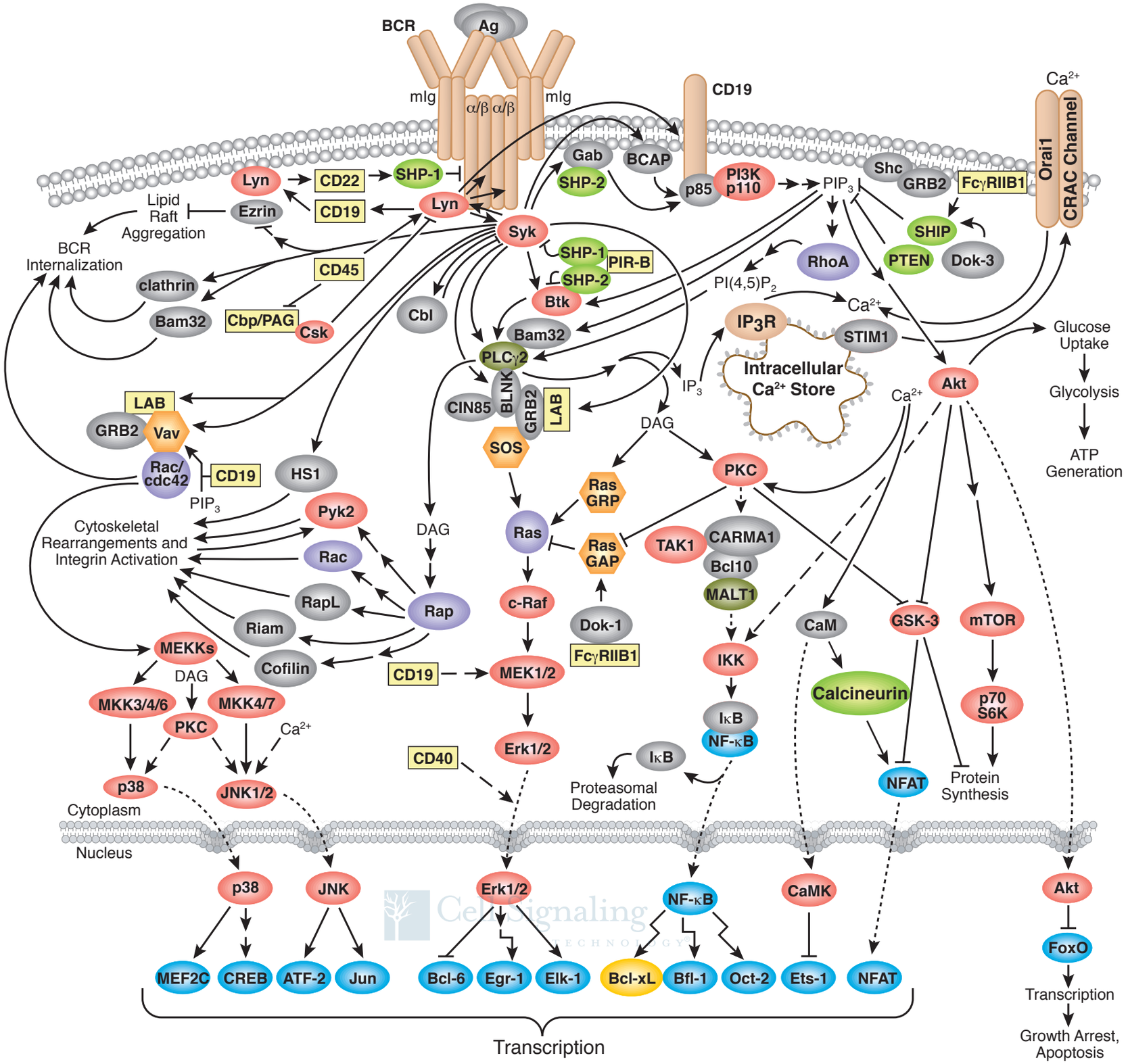
...never right or wrong, only useful or useless

Pablo Picasso: Don Quixote (1955)

4.

Modeling cellular information processing with rules (Kappa)

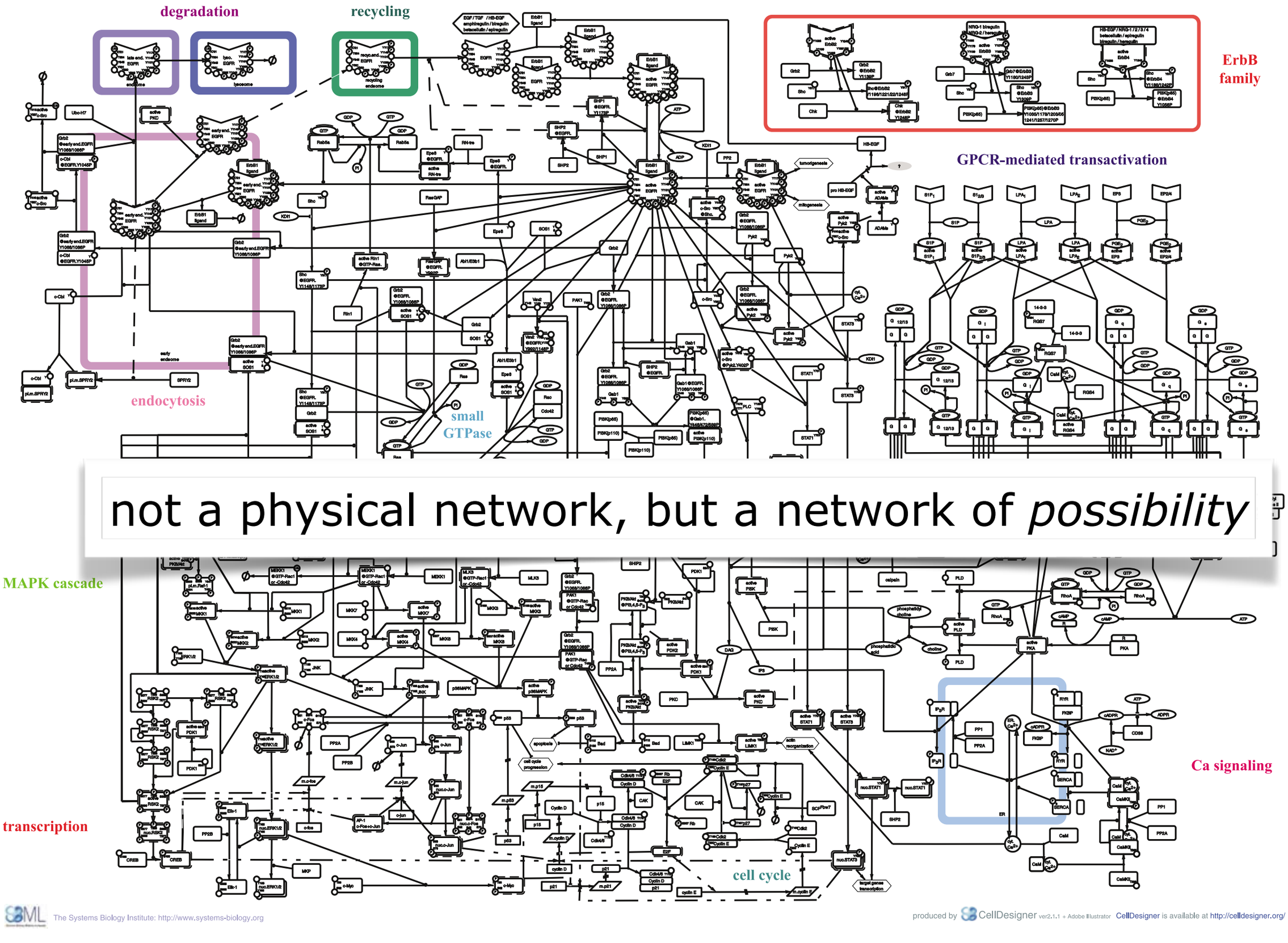
Signaling: A Circulatory System of Information



tasks:

cell fate,
division,
repair,
cell death,
motility,
morphology,
…

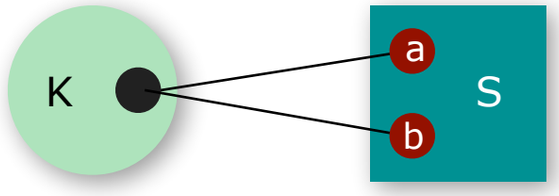
Signaling: A Circulatory System of Information



not a physical network, but a network of *possibility*

Extensional Models

$$\frac{d}{dt}x_1 = x_5 + x_9 + x_4 + x_3 + x_7 + x_8 + x_2 + x_6 - x_1x_{13} - x_1x_{13} - x_1x_{12} - x_1x_{11}$$
$$\frac{d}{dt}x_2 = x_3 - x_2$$
$$\frac{d}{dt}x_3 = x_1x_{11} - x_3 - x_3$$
$$\frac{d}{dt}x_4 = x_5 - x_4$$
$$\frac{d}{dt}x_5 = x_1x_{13} - x_5 - x_5$$
$$\frac{d}{dt}x_6 = x_7 - x_6$$
$$\frac{d}{dt}x_7 = x_1x_{12} - x_7 - x_7$$
$$\frac{d}{dt}x_8 = x_9 - x_8$$
$$\frac{d}{dt}x_9 = x_1x_{13} - x_9 - x_9$$
$$\frac{d}{dt}x_{10} = x_2 + x_6$$
$$\frac{d}{dt}x_{11} = x_3 + x_8 - x_1x_{11}$$
$$\frac{d}{dt}x_{12} = x_4 + x_7 - x_1x_{12}$$
$$\frac{d}{dt}x_{13} = x_5 + x_9 - x_1x_{13} - x_1x_{13}$$



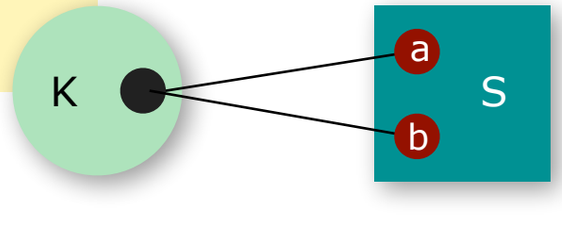
$$x_1 = [K]$$
$$x_2 = [K_aS_{ab}]$$
$$x_3 = [K_aS_b]$$
$$x_4 = [K_aS_a]$$
$$x_5 = [K_aS]$$
$$x_6 = [K_bS_{ab}]$$
$$x_7 = [K_bS_a]$$
$$x_8 = [K_bS_b]$$
$$x_9 = [K_bS]$$
$$x_{10} = [S_{ab}]$$
$$x_{11} = [S_b]$$
$$x_{12} = [S_a]$$
$$x_{13} = [S]$$

Extensional Models

$$\frac{d}{dt}[K] = \bullet + \bullet + \bullet + \bullet + \bullet + \bullet + \bullet + \bullet - \bullet - \bullet - \bullet - \bullet - \bullet$$
$$\frac{d}{dt}[K_a S_{ab}] = \bullet - \bullet$$
$$\frac{d}{dt}[K_a S_b] = \bullet - \bullet - \bullet$$
$$\frac{d}{dt}[K_a S_a] = \bullet - \bullet$$
$$\frac{d}{dt}[K_a S] = \bullet - \bullet - \bullet$$
$$\frac{d}{dt}[K_b S_{ab}] = \bullet - \bullet$$
$$\frac{d}{dt}[K_b S_a] = \bullet - \bullet - \bullet$$
$$\frac{d}{dt}[K_b S_b] = \bullet - \bullet$$
$$\frac{d}{dt}[K_b S] = \bullet - \bullet - \bullet$$
$$\frac{d}{dt}[S_{ab}] = \bullet + \bullet$$
$$\frac{d}{dt}[S_b] = \bullet + \bullet$$
$$\frac{d}{dt}[S_a] = \bullet + \bullet$$
$$\frac{d}{dt}[S] = \bullet + \bullet - \bullet - \bullet$$

13 equations

● kinase-related terms

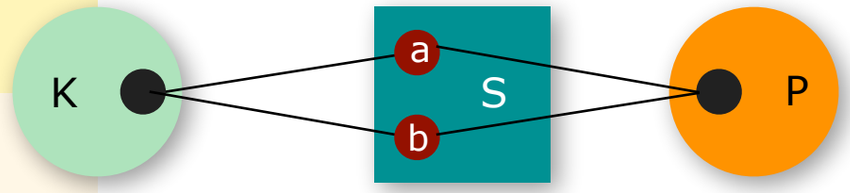


Extensional Models

$$\frac{d}{dt}[K] = \bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet-\bullet-\bullet-\bullet-\bullet-\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet+\bullet-\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_aS_{ab}] = \bullet-\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[K_aS_b] = \bullet-\bullet-\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[K_aS_a] = \bullet-\bullet+\bullet$$
$$\frac{d}{dt}[K_aS] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[K_bS_{ab}] = \bullet-\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[K_bS_a] = \bullet-\bullet-\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[K_bS_b] = \bullet-\bullet+\bullet$$
$$\frac{d}{dt}[K_bS] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[S_{ab}] = \bullet+\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[S_b] = \bullet+\bullet+\bullet+\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[S_a] = \bullet+\bullet+\bullet+\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[S] = \bullet+\bullet-\bullet-\bullet+\bullet+\bullet$$

13 equations, each changed !

- kinase-related terms
- phosphatase-related terms



adding a phosphatase

$$\frac{d}{dt}[K_aP_bS_{ab}] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_aP_bS_b] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_aP_bS_a] = \bullet+\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_aP_bS] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_bP_aS_{ab}] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_bP_aS_b] = \bullet+\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_bP_aS_a] = \bullet+\bullet-\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[K_bP_aS] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P] = \bullet+\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P_aP_bS_{ab}] = \bullet+\bullet-\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P_aP_bS_b] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P_aP_bS_a] = \bullet+\bullet-\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P_aP_bS] = \bullet+\bullet-\bullet-\bullet$$
$$\frac{d}{dt}[P_aS_{ab}] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_aS_b] = \bullet-\bullet+\bullet+\bullet$$
$$\frac{d}{dt}[P_aS_a] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_aS] = \bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_bS_{ab}] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_bS_b] = \bullet-\bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_bS_a] = \bullet-\bullet+\bullet$$
$$\frac{d}{dt}[P_bS] = \bullet-\bullet+\bullet$$

+ 21 new equations

Limitations of the ODE Framework

the formalism of differential equations does not represent agents, only their concentration

- need to know in advance all molecular species that can occur
- vulnerable to combinatorial explosions
- cumbersome to build and modify large models

the formalism of differential equations entangles kinetics and causation

- physical time is inadequate for describing distributed systems
- hard to reason about “mechanisms of action”

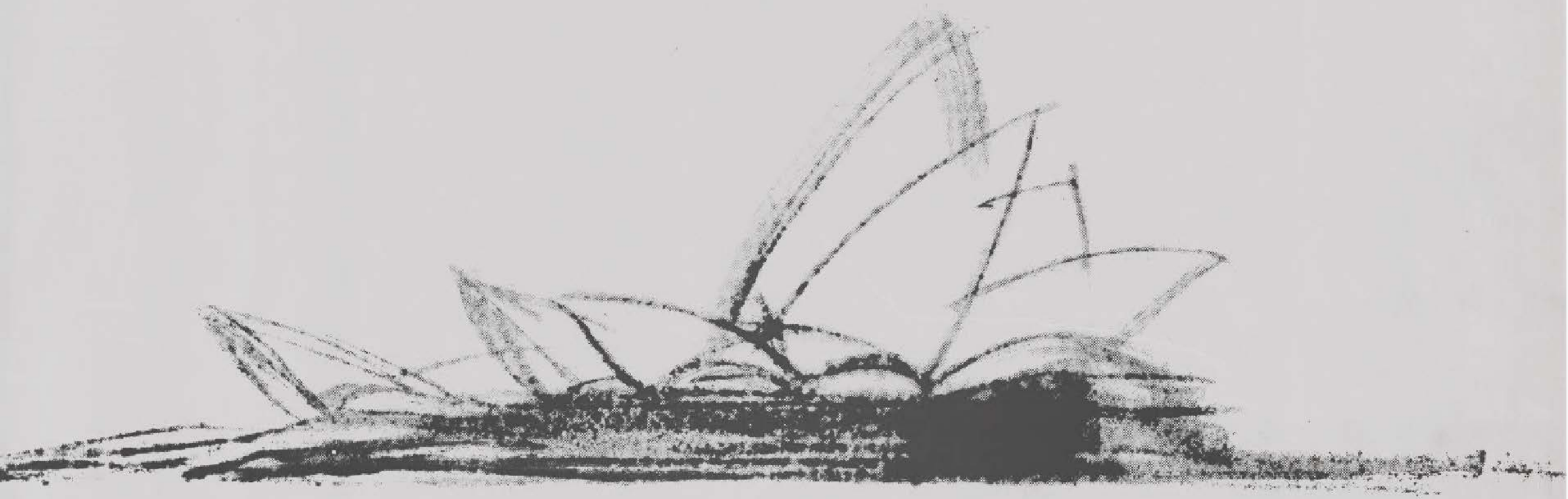
the formalism of differential equations separates model and knowledge

- “I forgot why I put this term in my model...”
- models are not self-documenting and don’t organize knowledge effectively

Modeling in Physics (Approximately)

understanding precedes modeling

Jørn Utzon, 1958



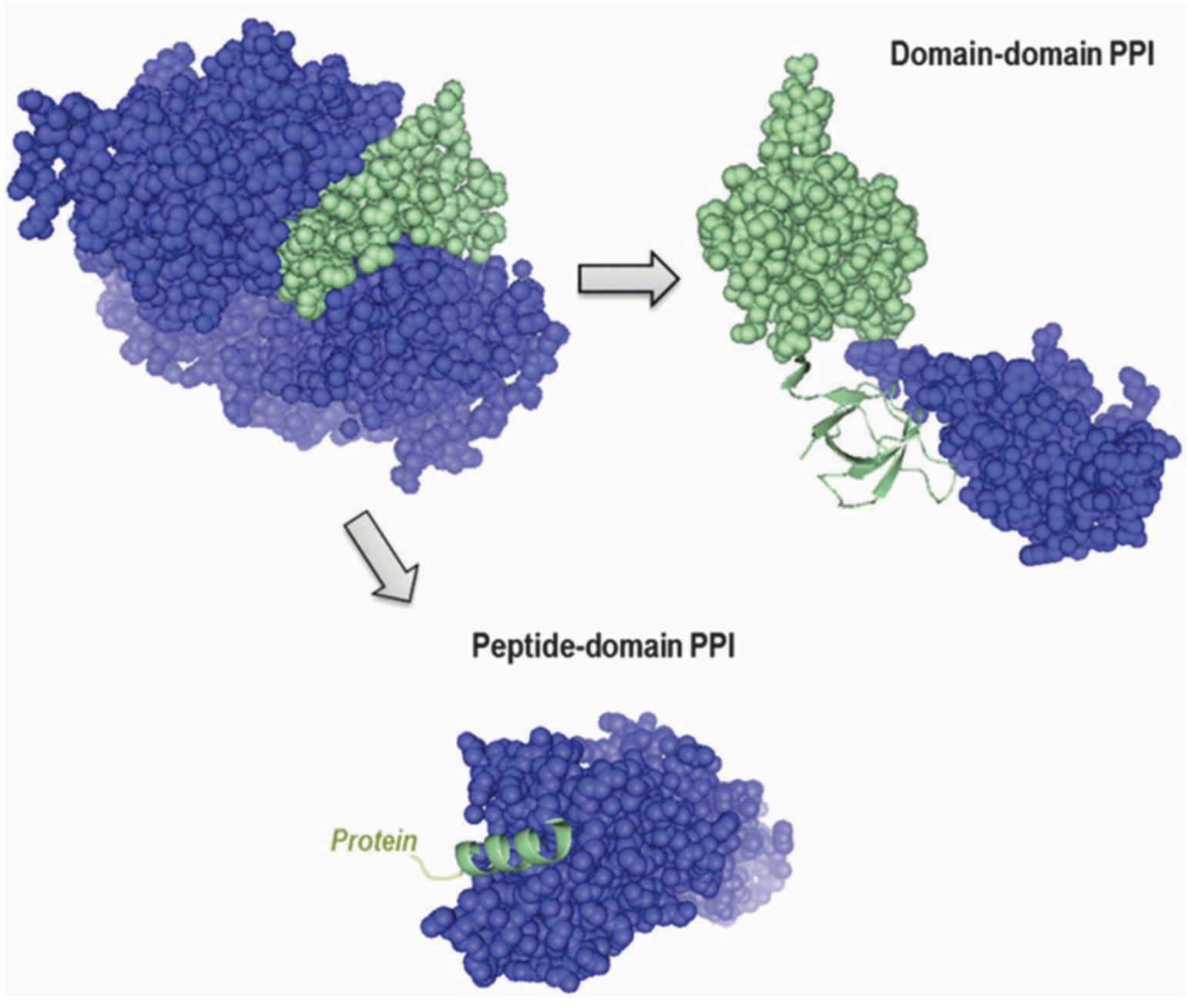
Jørn Utzon, 1958

Modeling in Biology ?

modeling precedes understanding

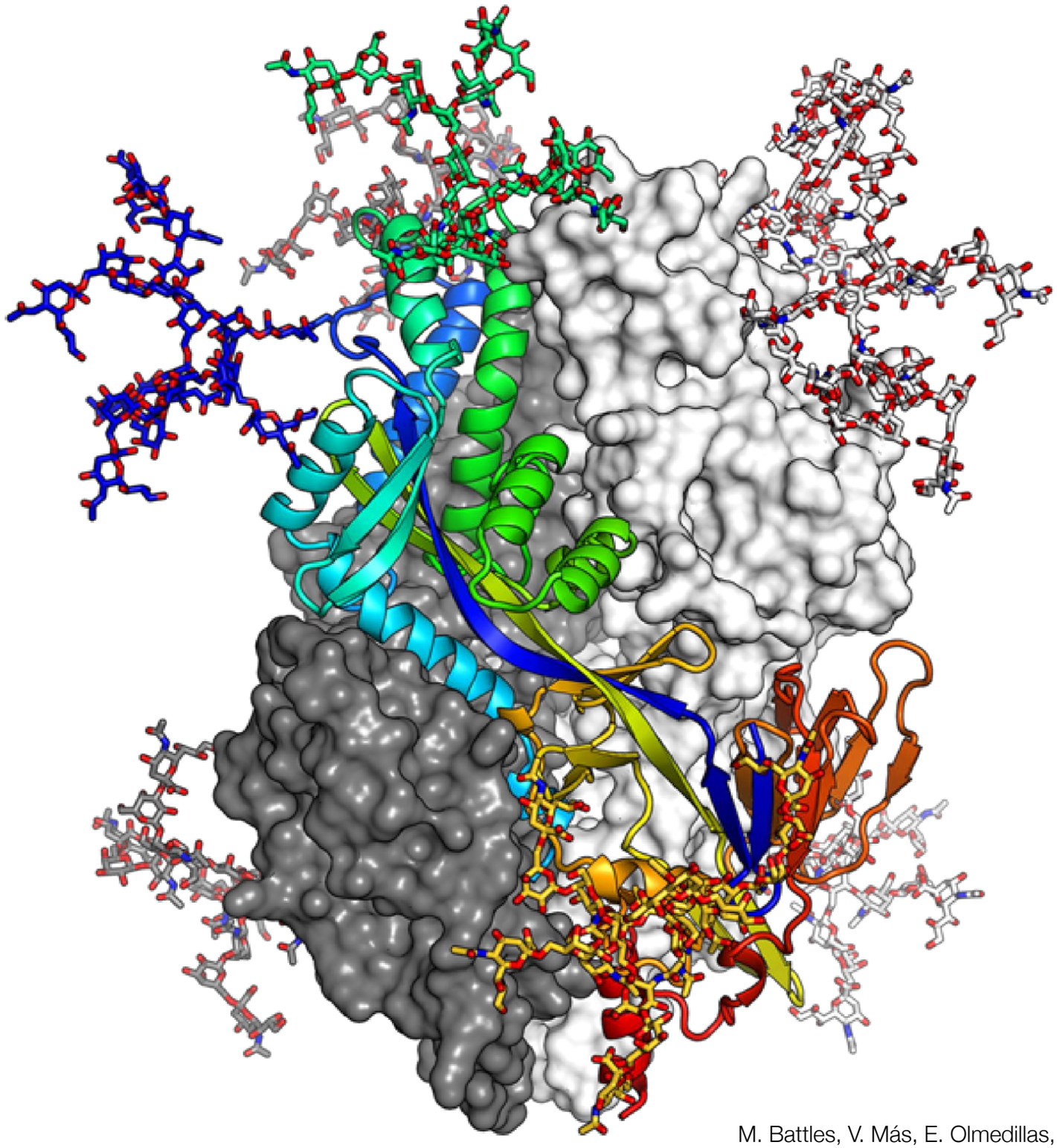
A model should be a formal and executable representation of the facts it rests upon.

The Scope: Proteins as "Agents"



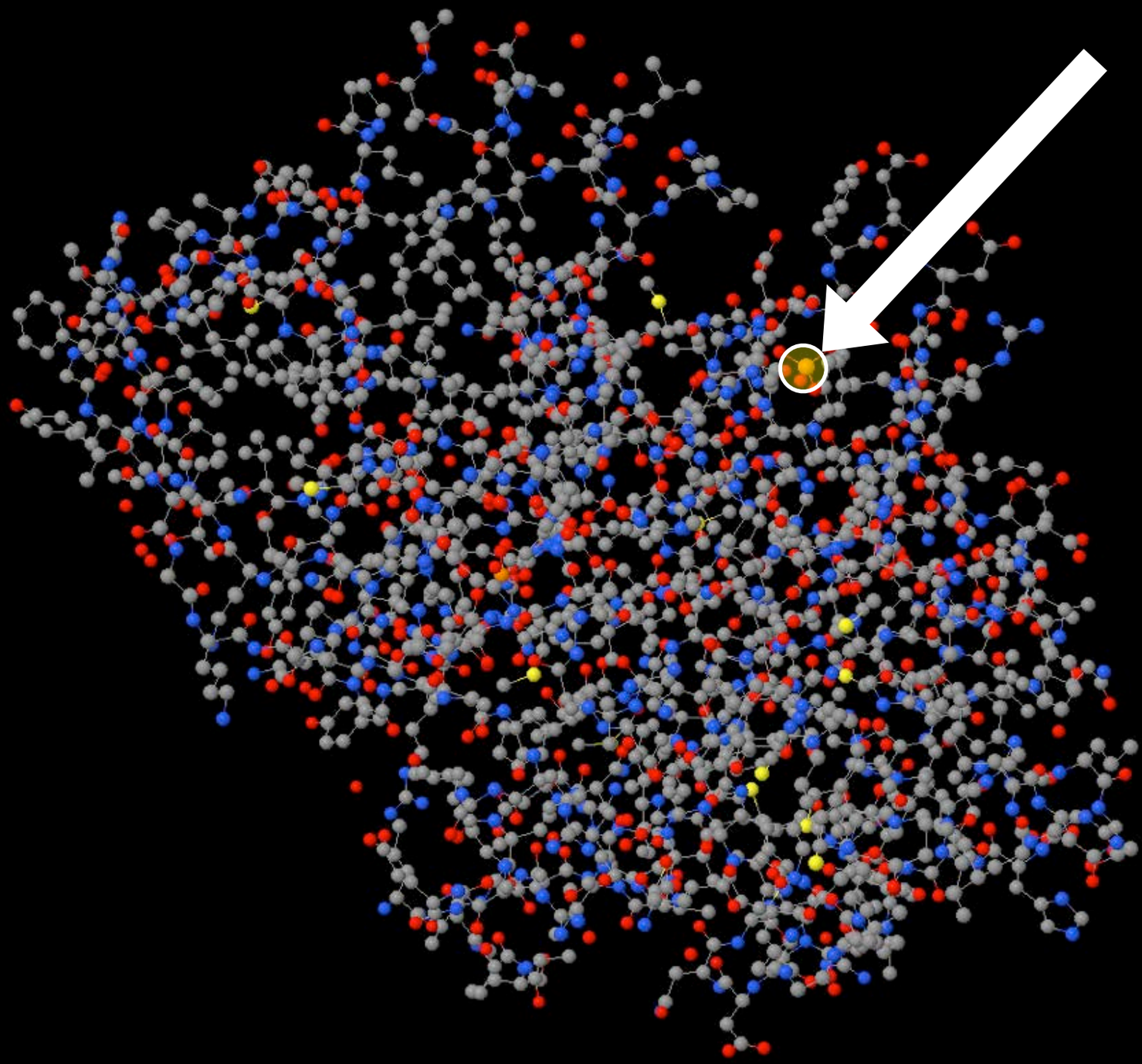
Laura Nevola and Ernest Giralt, Chem. Commun., 2015, 51, 3302-3315

The Scope: Proteins as "Agents"

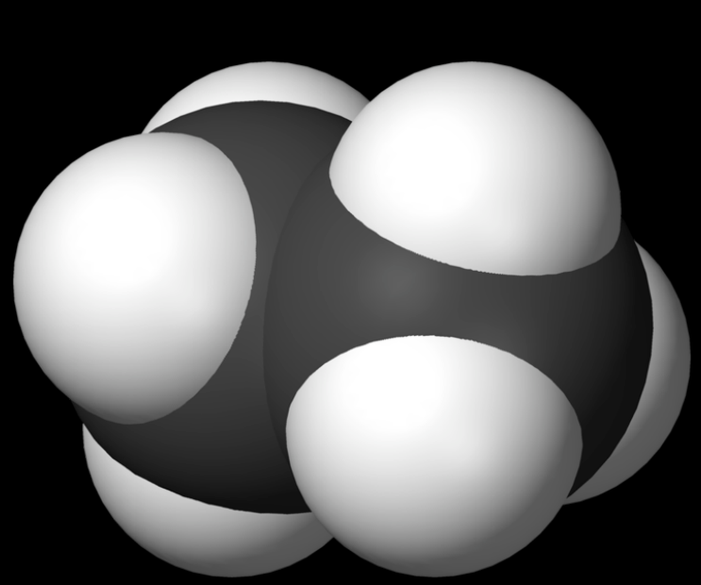


M. Battles, V. Más, E. Olmedillas, O. Cano, M. Vázquez, L. Rodríguez, J. Melero, & J. McLellan. "Structure and immunogenicity of pre-fusion-stabilized human metapneumovirus F glycoprotein," Nat Commun 8, 1528 (2017)

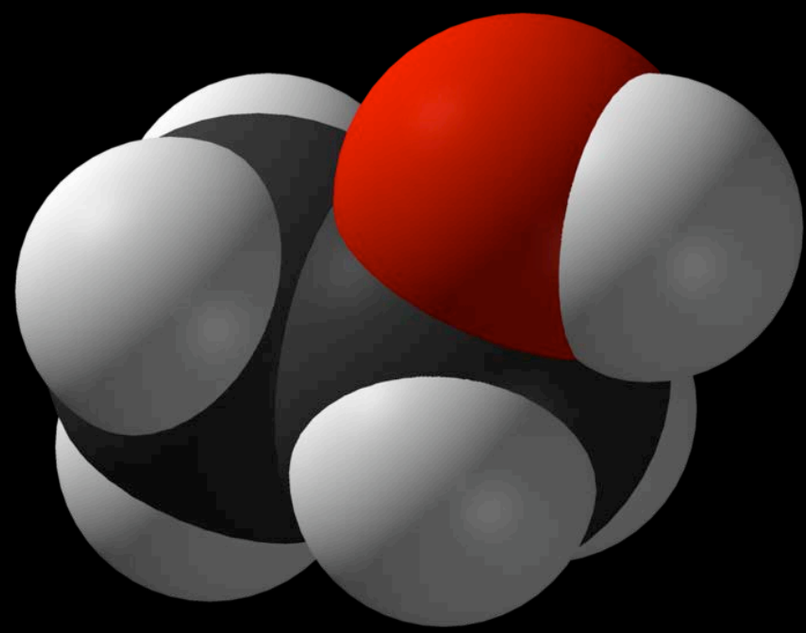
The Same Protein In a Different State



NOT The Same Molecule In a Different State

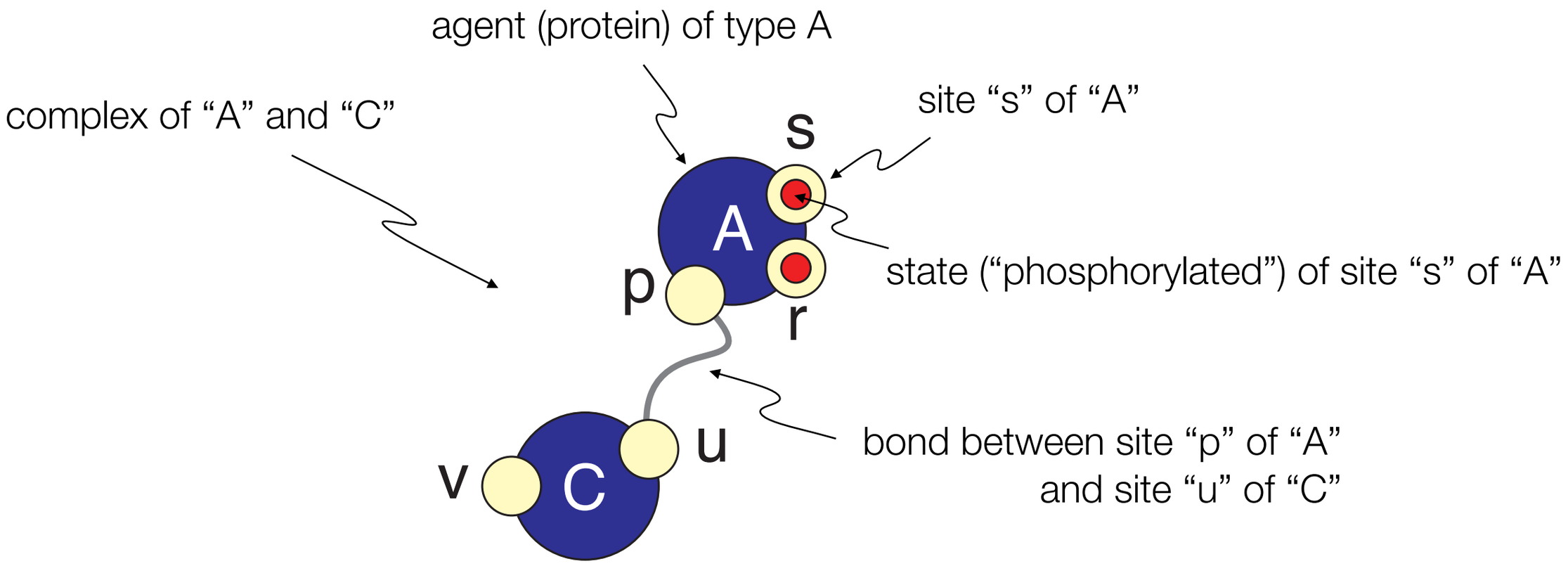


ethane



ethanol

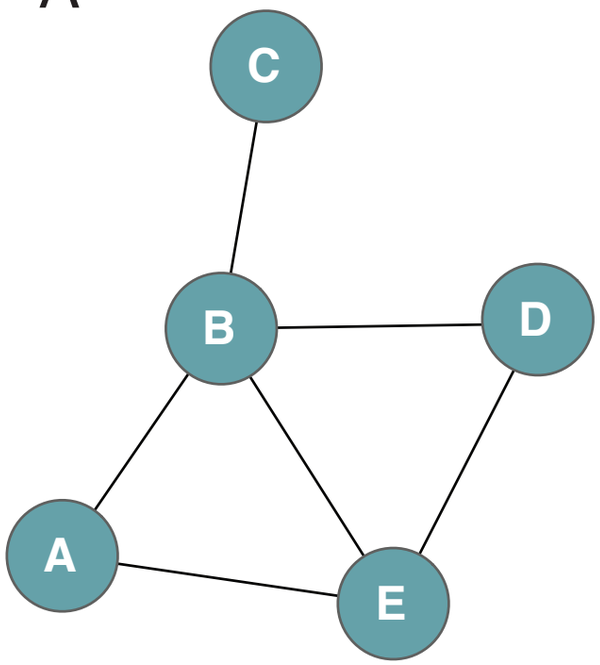
Abstraction Level



A(s{p}[.], r{p}[.], p[1]), C(v[.],u[1])

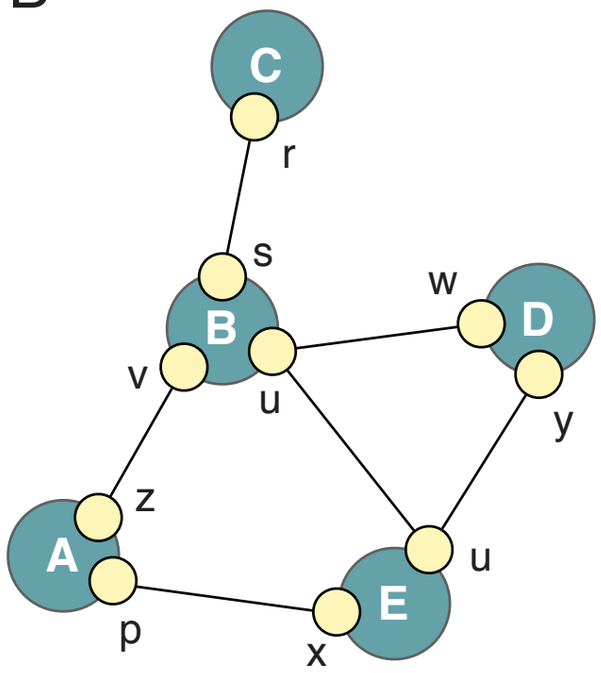
Graphs And Site Graphs

A



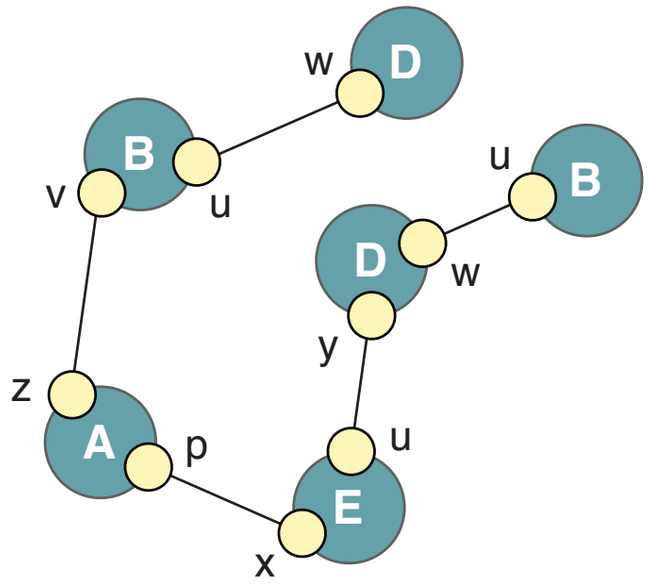
plain graph

B



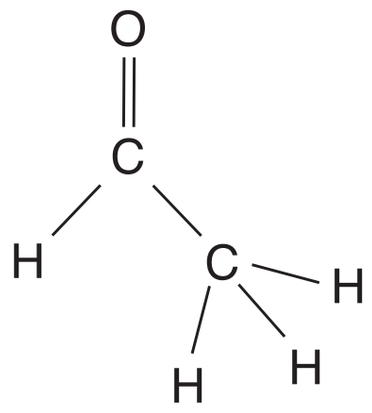
site graph (contact graph)

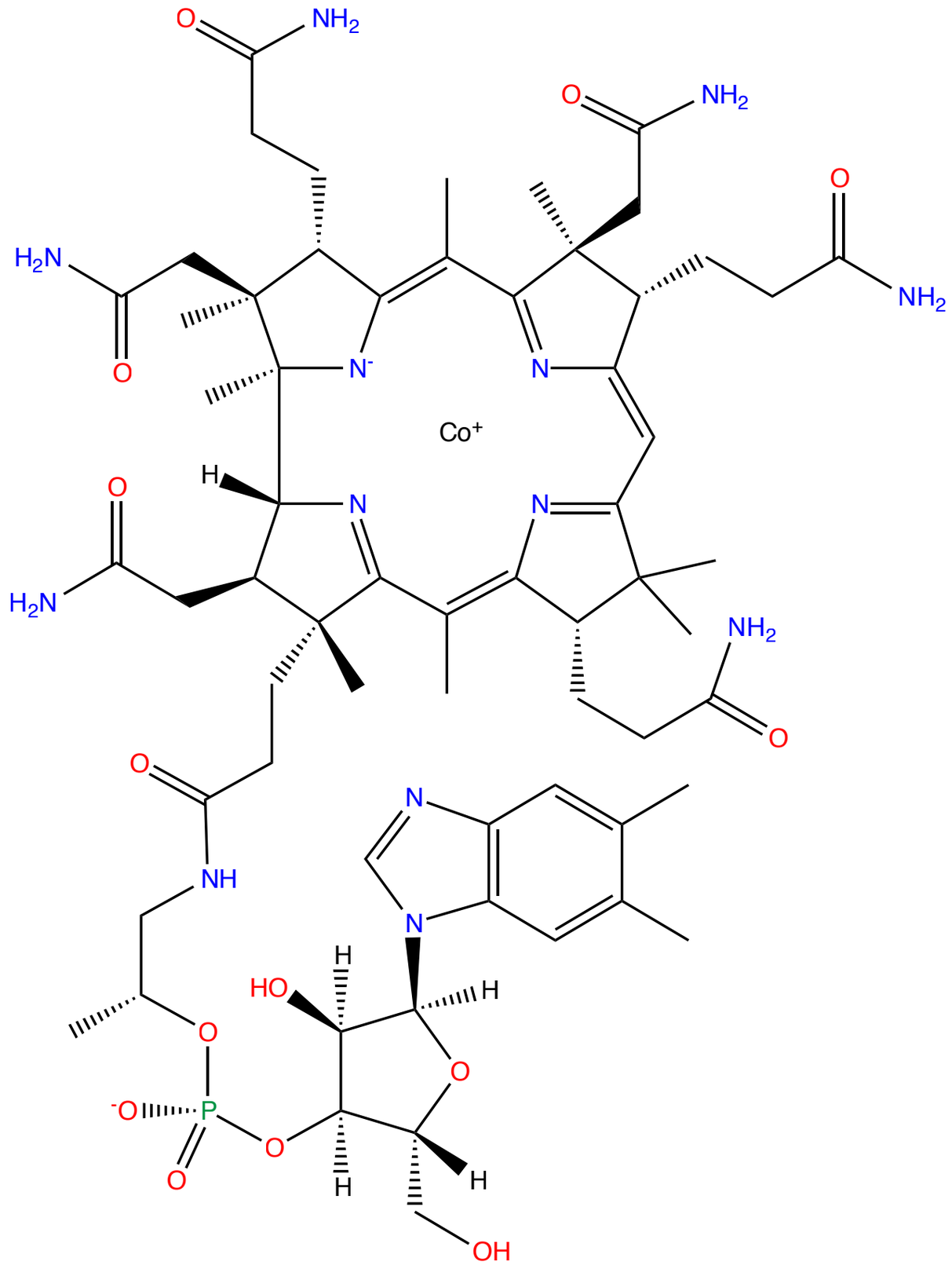
C



site graph (pattern)

Chemical Abstraction For Comparison

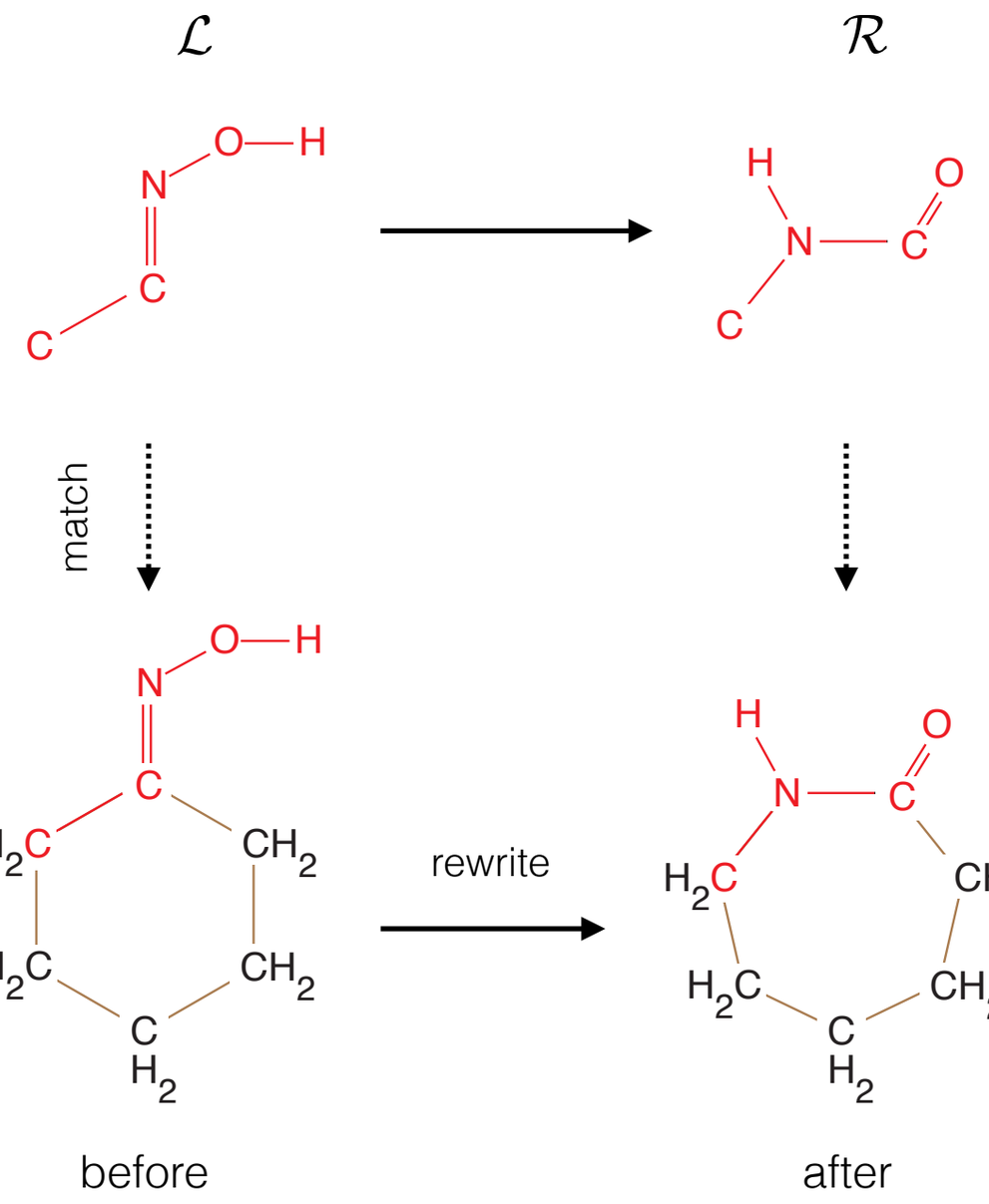


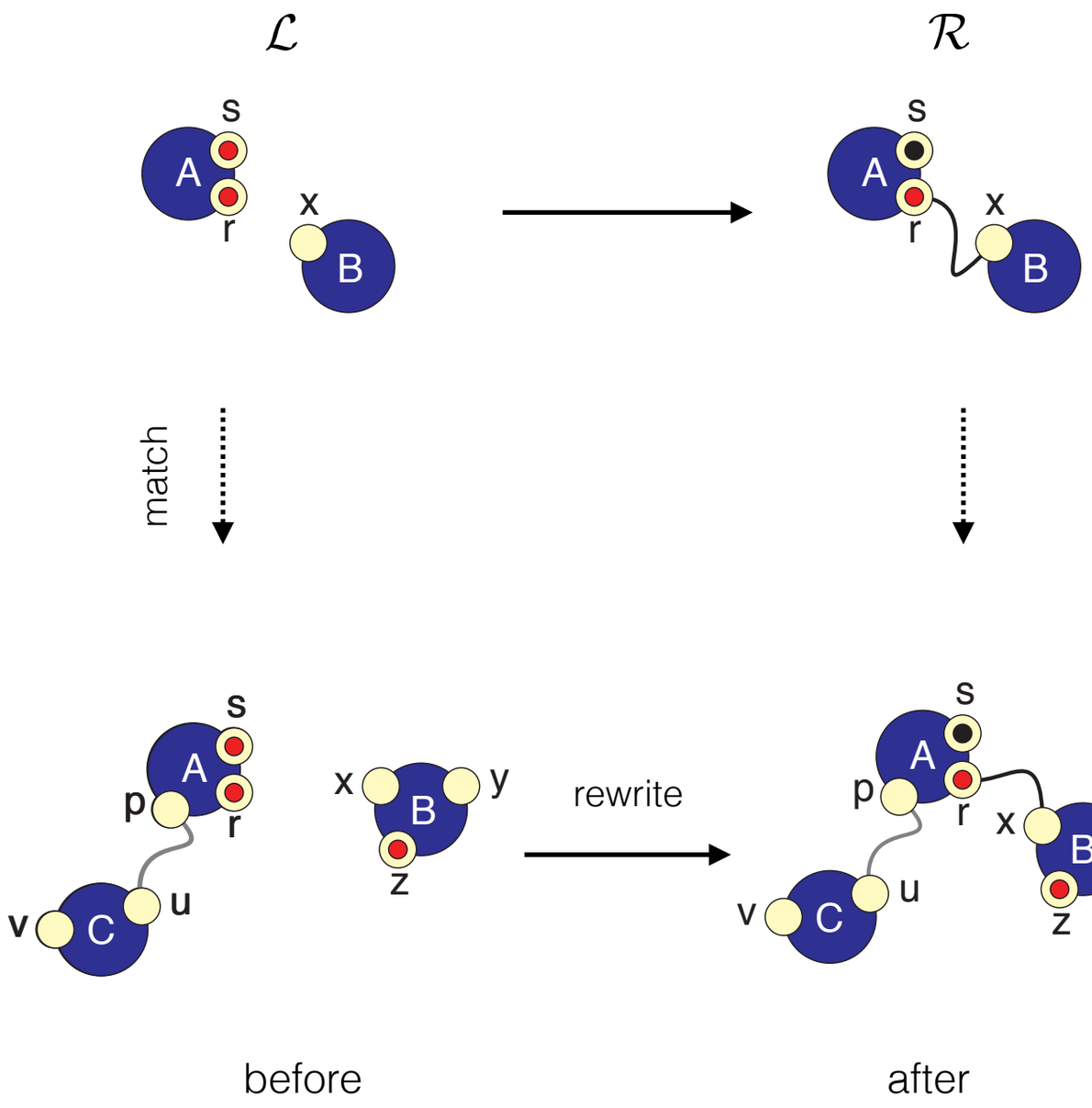


Graph Rewriting

chemistry

“molecular biology”

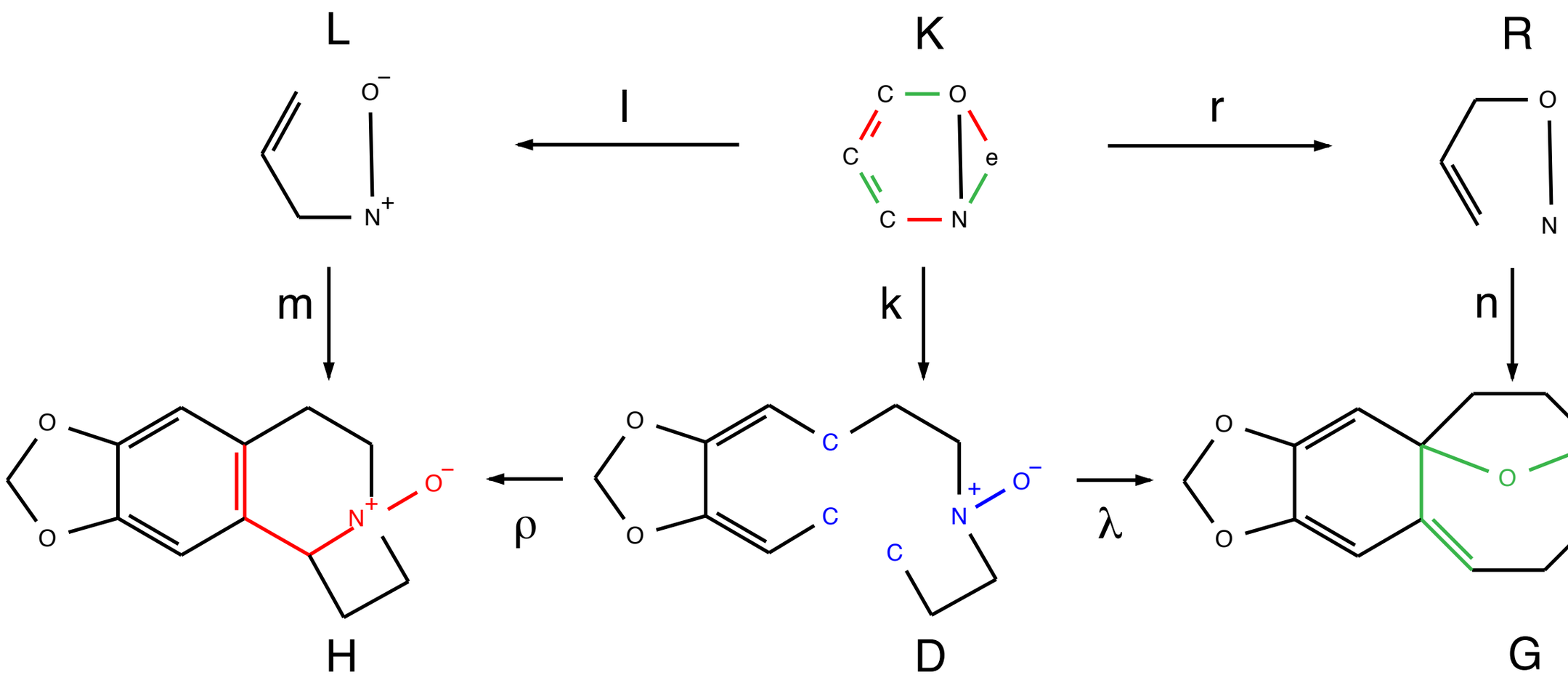




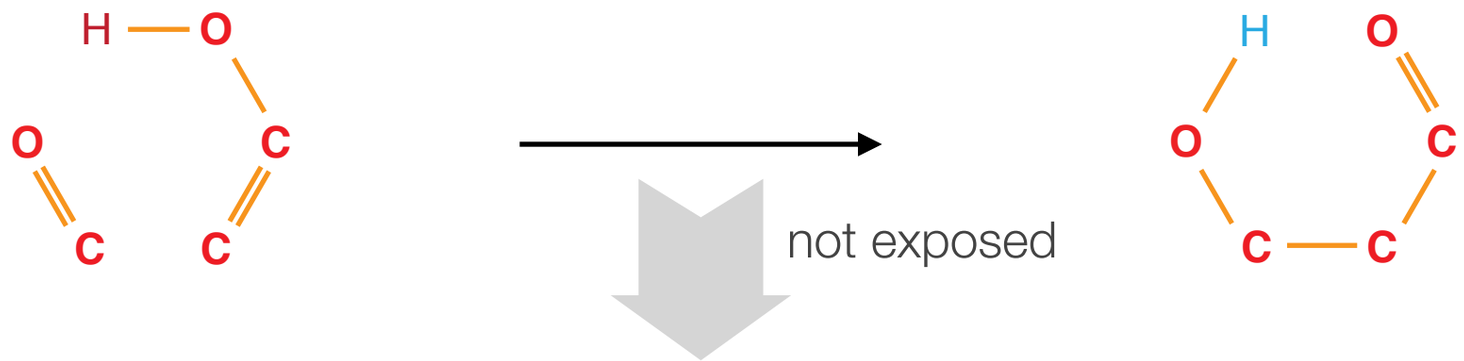
implementation: Mød

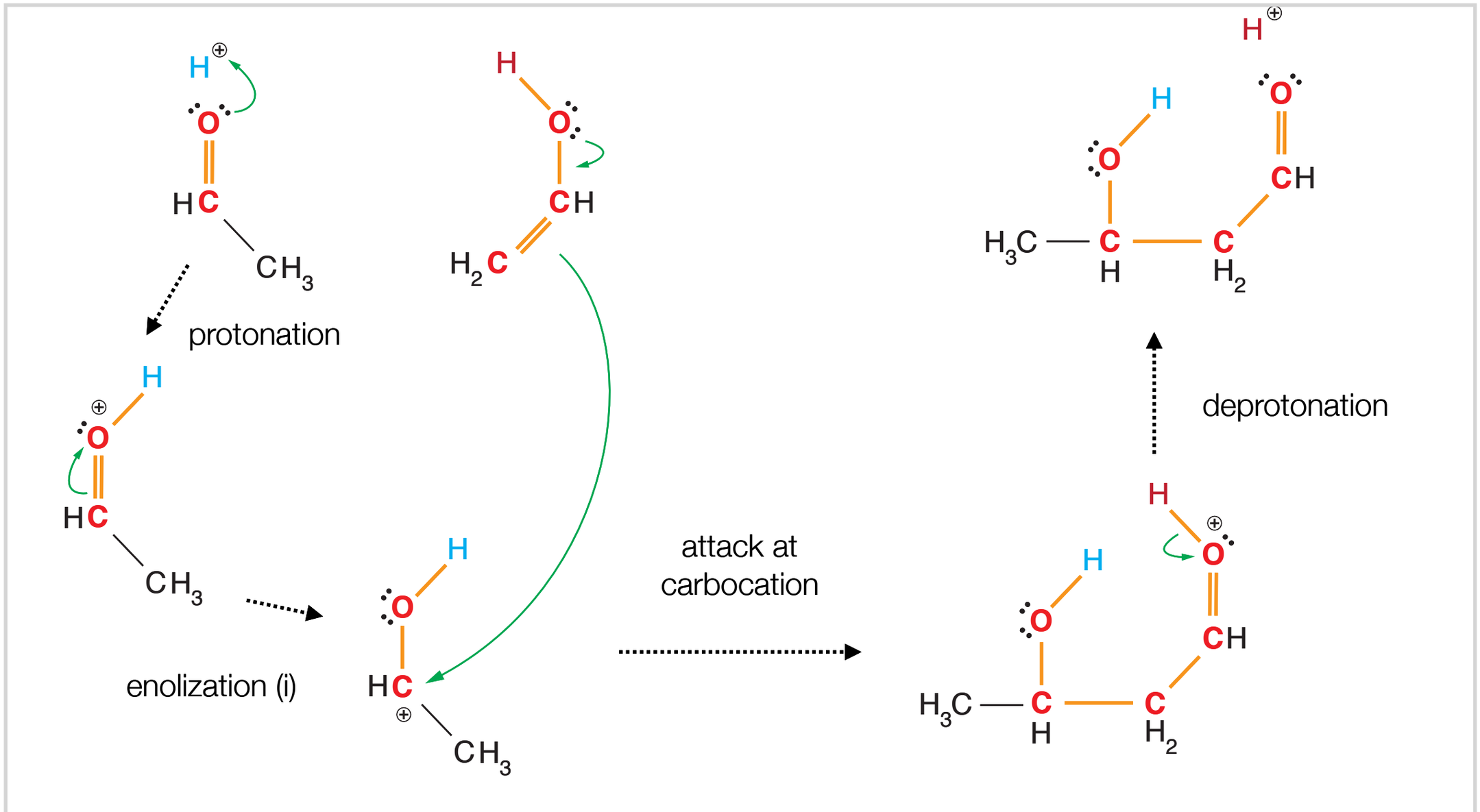
implementation: Kappa

DPO Graph Rewriting in Chemistry



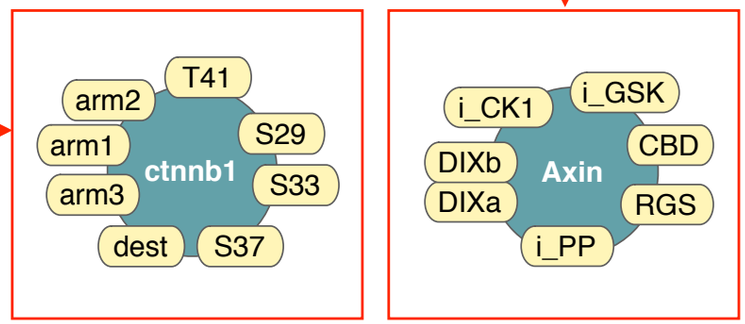
Unboxing of Rules



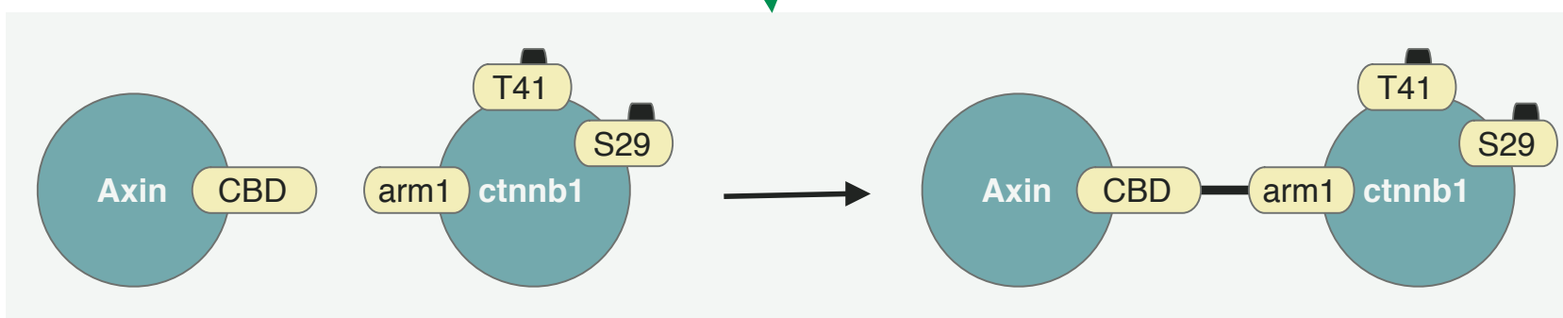


Fusing Modeling and Knowledge Representation

"Axin binds a region in the armadillo repeat of β-catenin, if β-catenin is un-phosphorylated at T41 and S29."



agent abstraction

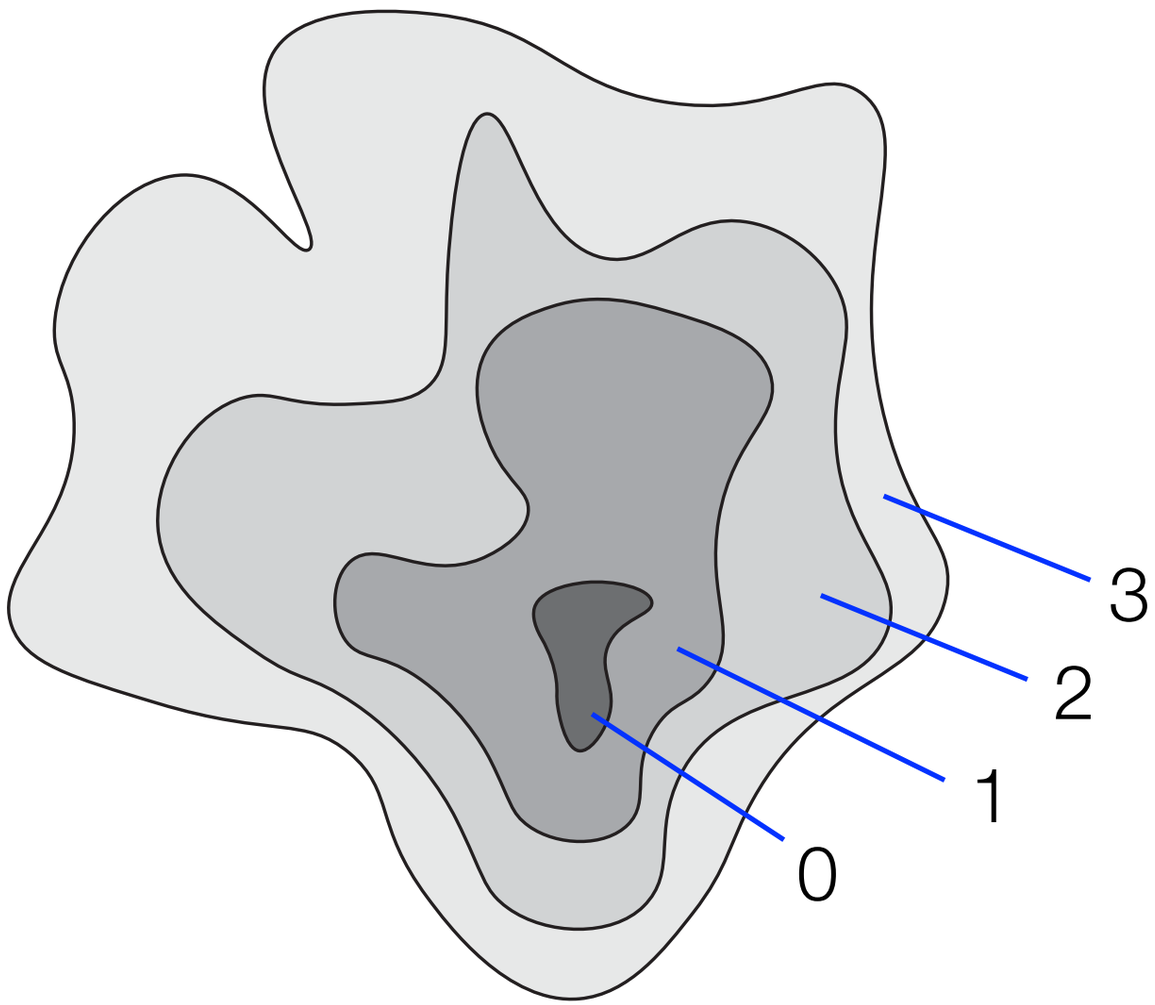


an interaction rule

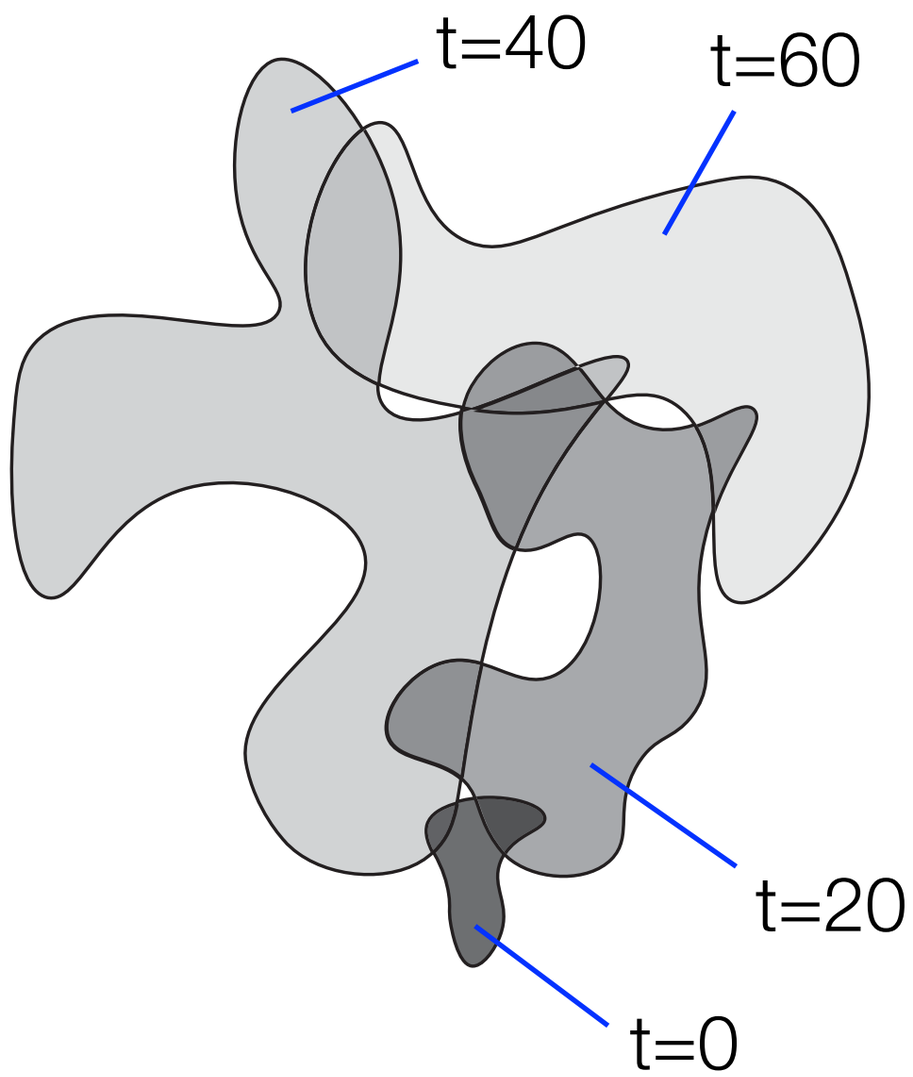
Statics and Dynamics

static expansion

$$\mathcal{A}_{i+1} = (\mathcal{A}_i \circ \mathcal{A}_i) \cup \mathcal{A}_i$$

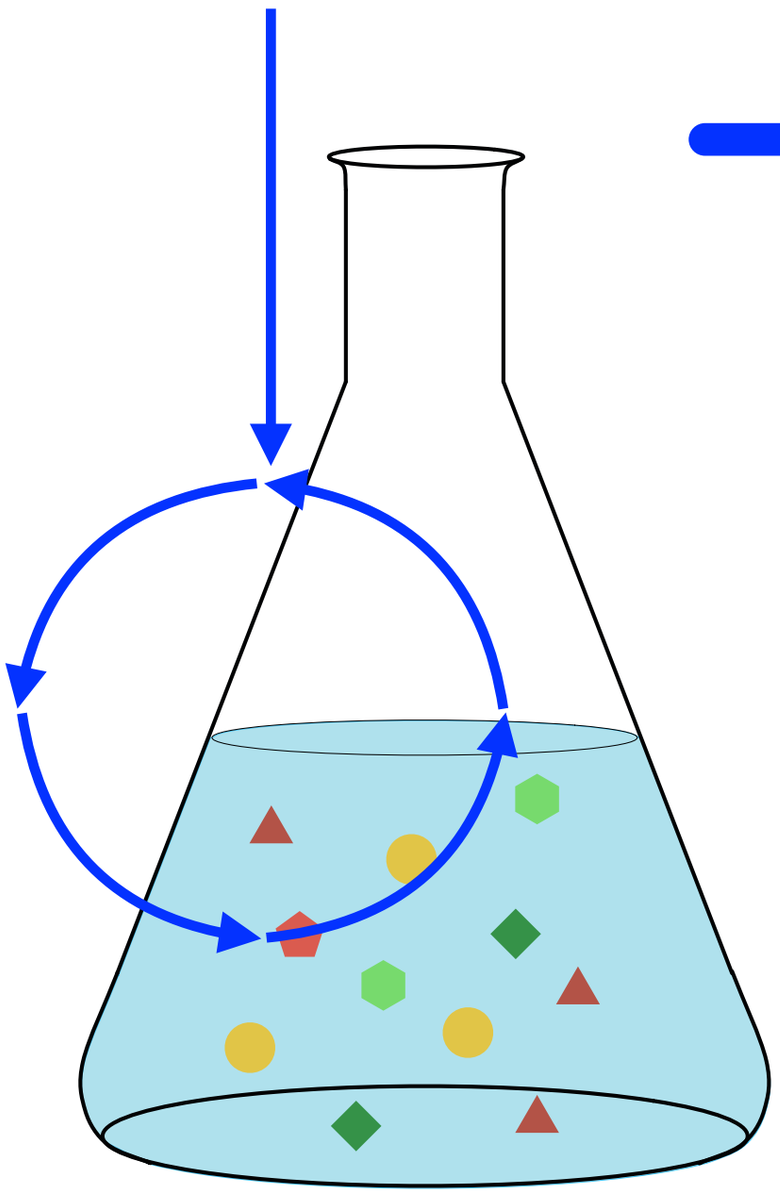


dynamic evolution



Dynamics

dynamic evolution



choose initial “soup” of molecules

for each rule

identify all its matchings in the soup;
compute its probability to fire, based on the # of matchings

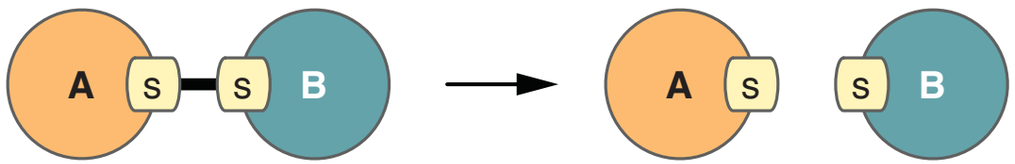
determine the time of the next event

choose a rule and a matching according to probability

apply the rule to the matched molecules

The Rate Constant

$\mathcal{L}_r$ $\mathcal{R}_r$

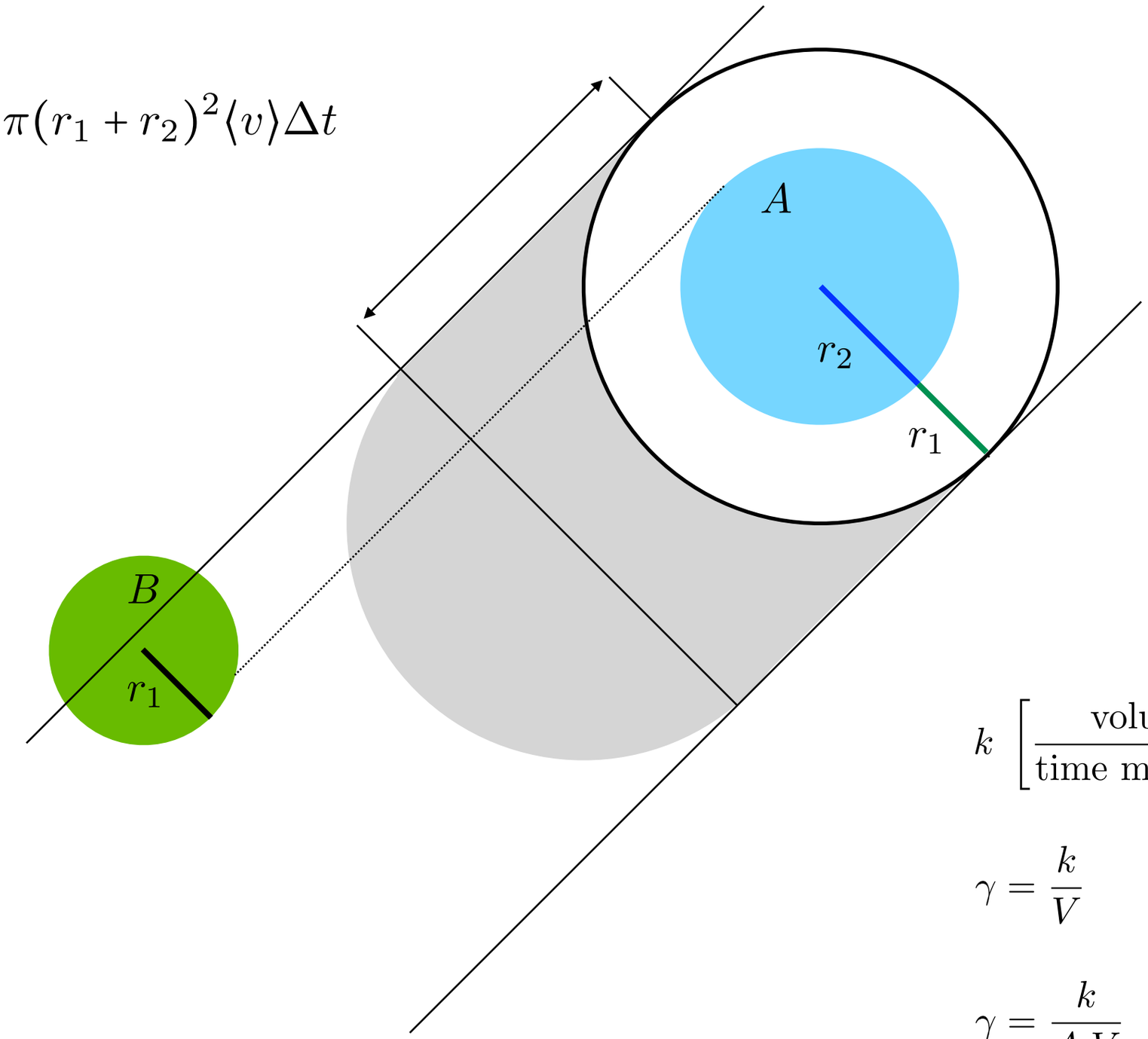


$@ \quad \gamma$

The Stochastic Rate Constant

volume V in which the reaction occurs

$$\Delta V = \pi(r_1 + r_2)^2 \langle v \rangle \Delta t$$



$$k \left[\frac{\text{volume}}{\text{time molecule}} \right] \qquad M^{-1}s^{-1}$$

$$\gamma = \frac{k}{V} \qquad mol^{-1}s^{-1}$$

$$\gamma = \frac{k}{A\,V} \qquad molecule^{-1}s^{-1}$$

The Stochastic Rate Constant

$\Delta V = \pi(r_1 + r_2)^2 \langle v \rangle \Delta t$ volume swept by collision cross-section by 1 A

$\pi d^2 \langle v \rangle \Delta t [B]$ opportunities for that A to hit some B

$\pi d^2 \langle v \rangle \Delta t V [A][B]$ opportunities for some A to hit some B

$\pi d^2 \langle v \rangle [A][B]$ rate per time and volume $k = \pi d^2 \langle v \rangle$ $M^{-1}s^{-1}$

$$\frac{\pi d^2 \langle v \rangle \Delta t}{V} n_B n_A = \frac{\Delta V}{V} n_B n_A$$

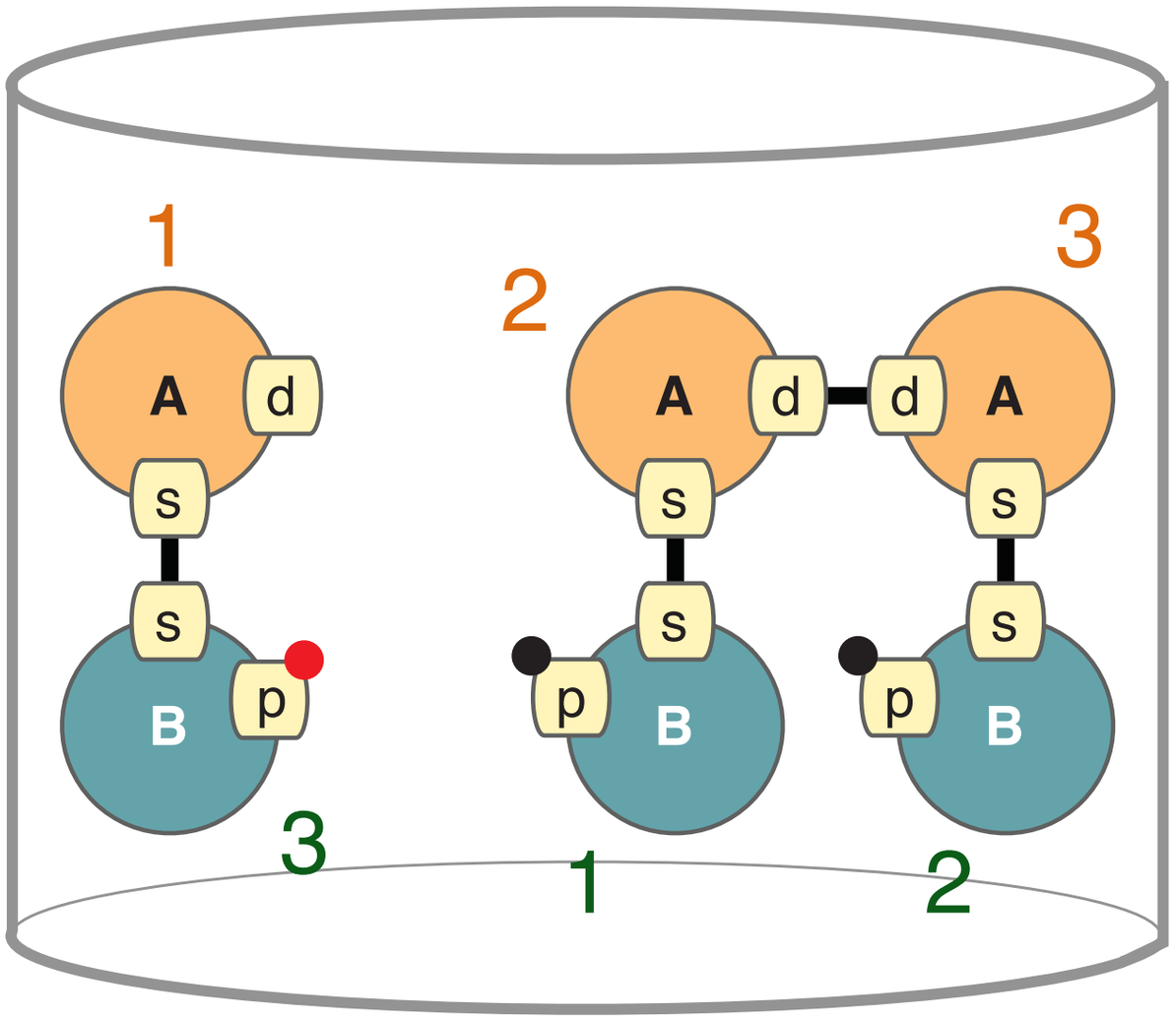
probability of encountering a B

$$\gamma = \frac{\pi d^2 \langle v \rangle}{V} = \frac{k}{V}$$

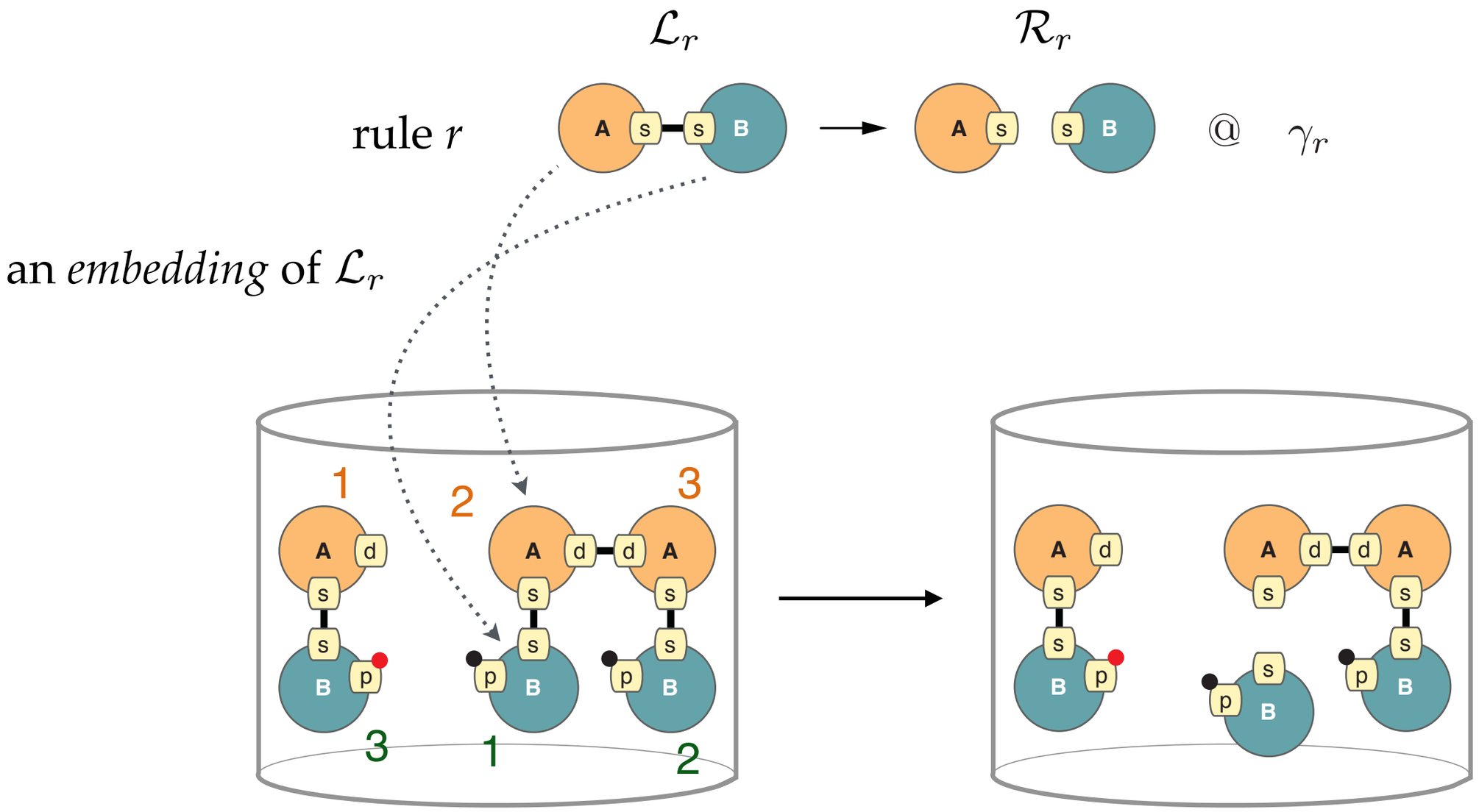
$$\gamma = \frac{k}{\mathcal{A}V} \quad \text{molecule}^{-1} s^{-1}$$

$\mathcal{A} = 6.022\,10^{23}\ \text{mol}^{-1}$

The Mixture



Rule Application



rule activity $\alpha_r = \gamma_r \cdot [\mathcal{L}_r; \mathcal{S}]$ (will modify…)

Memoryless Process

$a \cdot d\tau$ ... probability density of catching a fish in the time interval $[\tau, \tau + d\tau]$

$p(t)dt = ?$

$$P_0(t + \Delta t) = P_0(t)(1 - a\Delta t)$$

$$\Delta t \to 0 \quad \frac{dP_0(t)}{dt} = -a\, P_0(t)$$

no dinner up to time t $\quad P_0(t) = e^{-at}$

dinner within time t $\quad P(t) = 1 - e^{-at}$

probability density of dinner at time t $\quad p(t) = \frac{dP(t)}{dt} = ae^{-a\,t}$

The "First" Event In A Reaction Channel

$$R: \qquad A + B \xrightarrow{k} \text{something} \qquad\qquad v_R = k[A][B]$$

for an individual reaction event $\quad p(t) = \gamma e^{-\gamma t}$ $\quad$ with $\quad \gamma = \dfrac{k}{A\,V} \quad molecule^{-1}s^{-1}$

that many possible events $\quad n_A n_B$

for the reaction $\quad p(t) = \alpha\, e^{-\alpha t}$ $\quad$ with $\quad \alpha = \gamma n_A n_B$

$$p(t) = n_A n_B p(t)[1 - P(t)]^{n_A n_B - 1} = n_A n_B \gamma e^{-\gamma t} e^{-(n_A n_B - 1)\gamma t} = n_A n_B \gamma e^{-n_A n_B \gamma t}$$

CTMC

Given: n reaction types $i = 1, \ldots, n$ with activities α_i

Want to know: $p(i,t)$ probability that the first reaction happens at time *t* and it is reaction *i*

(1) probability that reaction *i* happens at time *t* and it is the first: $p_i(t)$

$$p_i(t) = \alpha_i e^{-\alpha_i\, t} \prod_{j \neq i}^{n} p(t_j > t) = \alpha_i e^{-\alpha_i\, t} \prod_{j \neq i}^{n} e^{-\alpha_j\, t} = \alpha_i e^{-\lambda\, t}$$

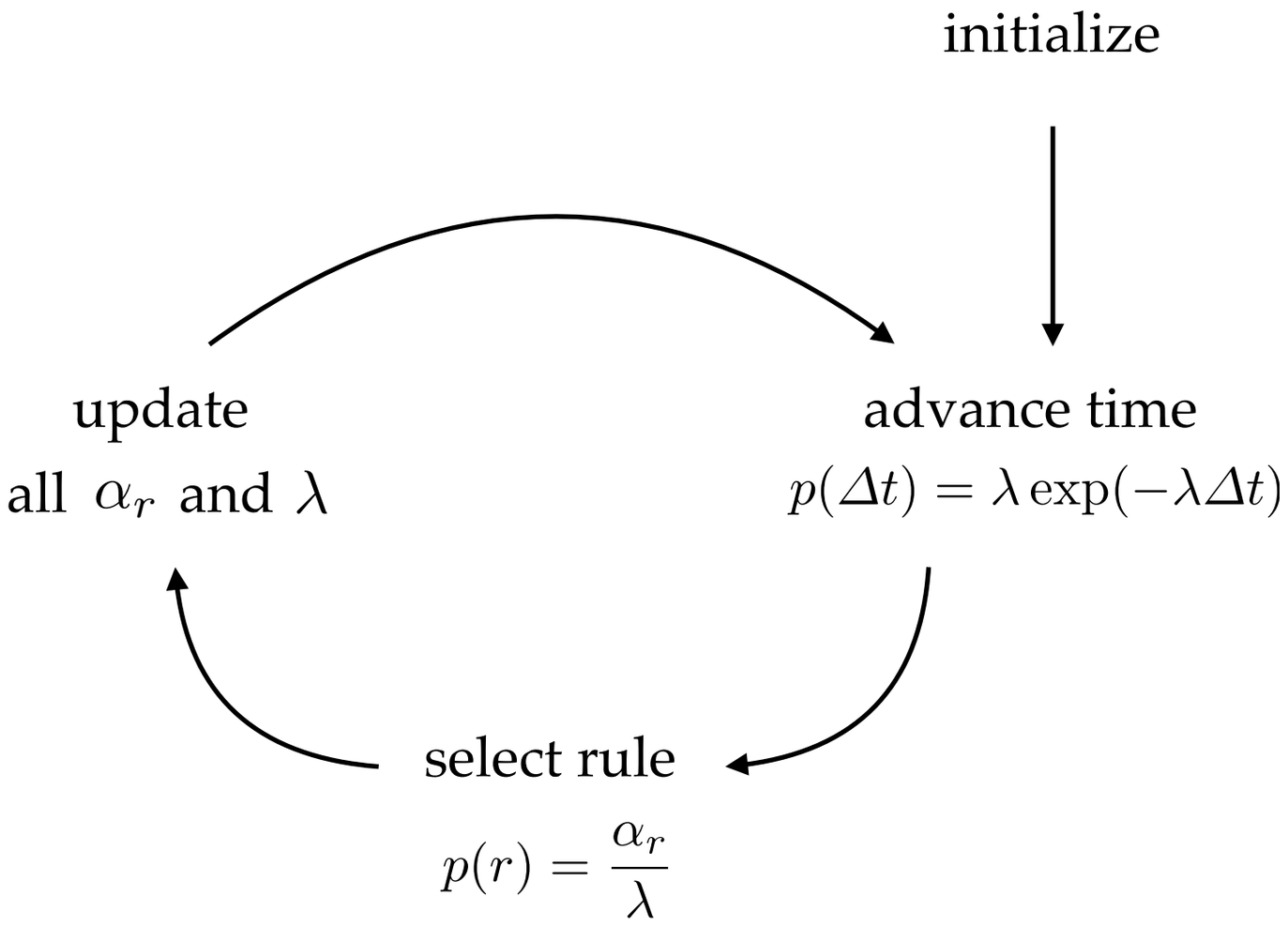
with $\lambda = \sum_{j=1}^{n} \alpha_j$

(2) $p(i,t) = p(i \,|\, t)p(t)$

$$p(t) = \sum_{i=1}^{n} p_i(t) = \sum_{i=1}^{n} \alpha_i e^{-\lambda t} = \boxed{\lambda e^{-\lambda t}}$$

$$p(i \,|, t) = \frac{p_i(t)}{\sum_{i=1}^{n} p_i(t)} = \frac{p_i(t)}{p(t)} = \frac{\alpha_i e^{-\lambda t}}{\lambda e^{-\lambda t}} = \boxed{\frac{\alpha_i}{\lambda}}$$

Basic CTMC Loop



rules

$$r: \quad \mathcal{L}_r \rightarrow \mathcal{R}_r \quad @ \, \gamma_r$$

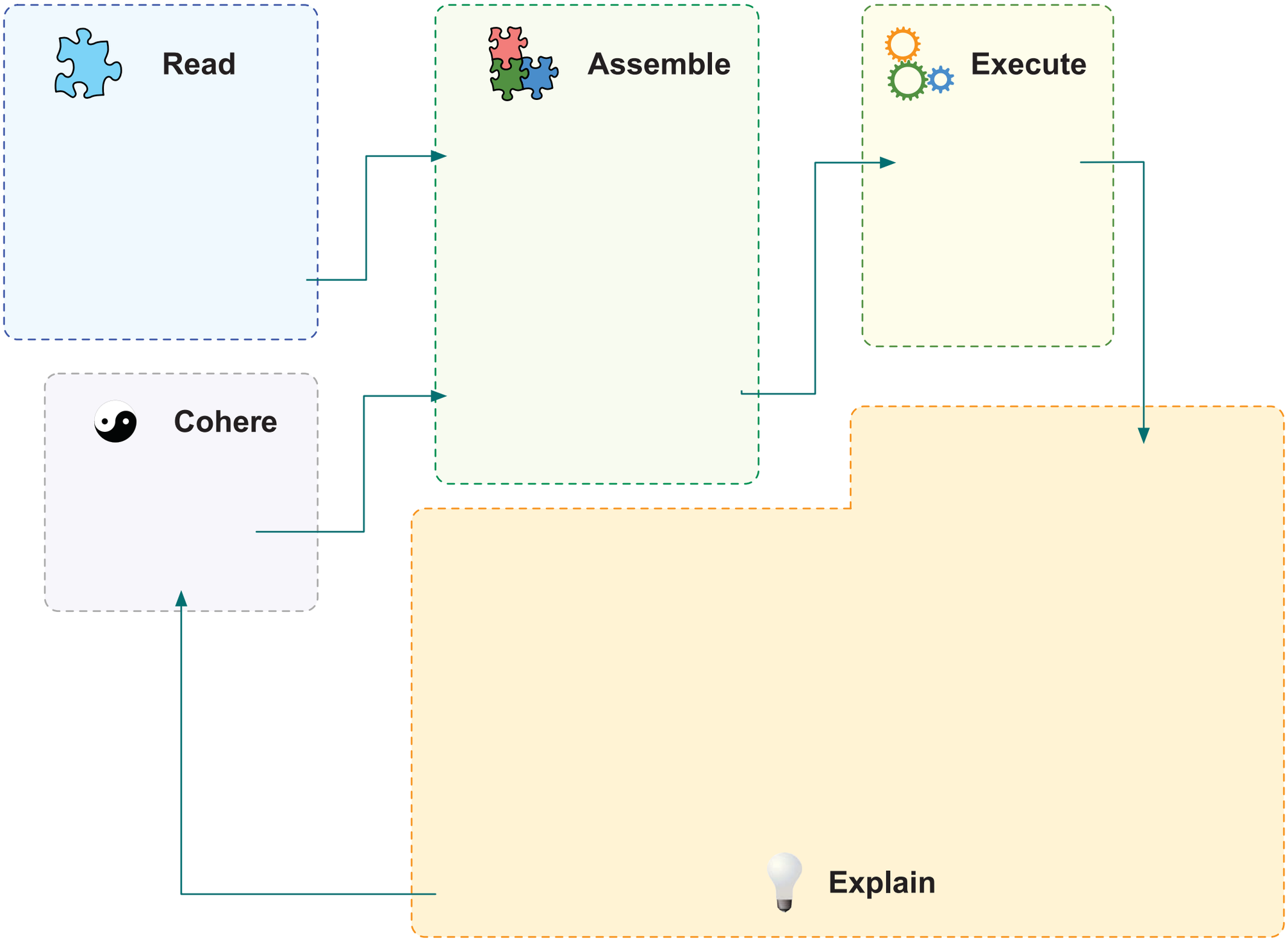
rule activities

$$\alpha_r = \gamma_r \cdot [\mathcal{L}_r; \mathcal{S}] \cdot \frac{1}{\sigma_r}$$

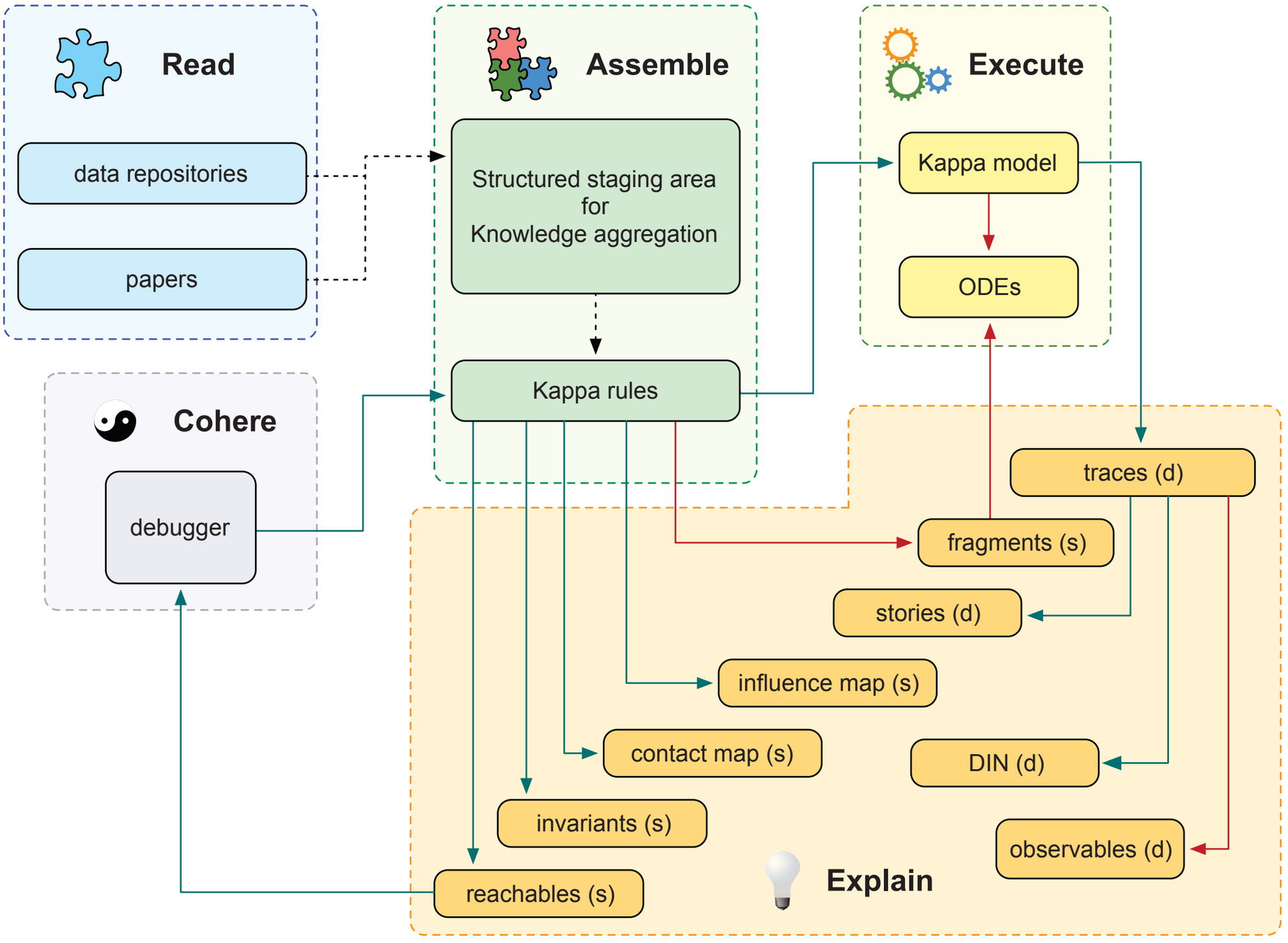
system activity

$$\lambda = \sum_s \alpha_s$$

Models as "Executable Knowledge"



Models as "Executable Knowledge"

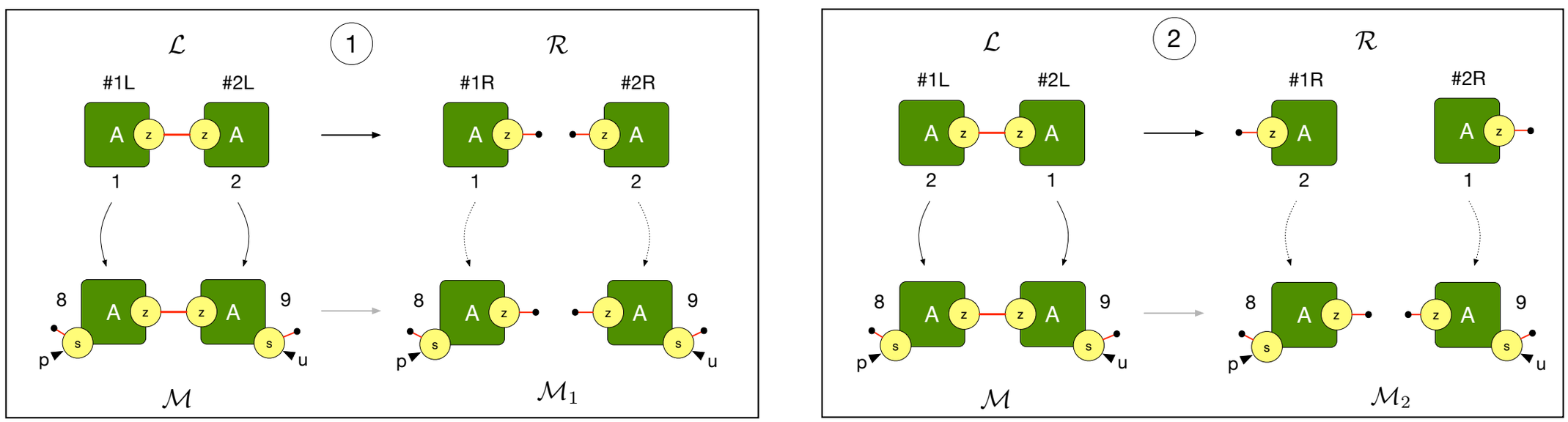


kappalanguage.org

(includes a detailed language manual)

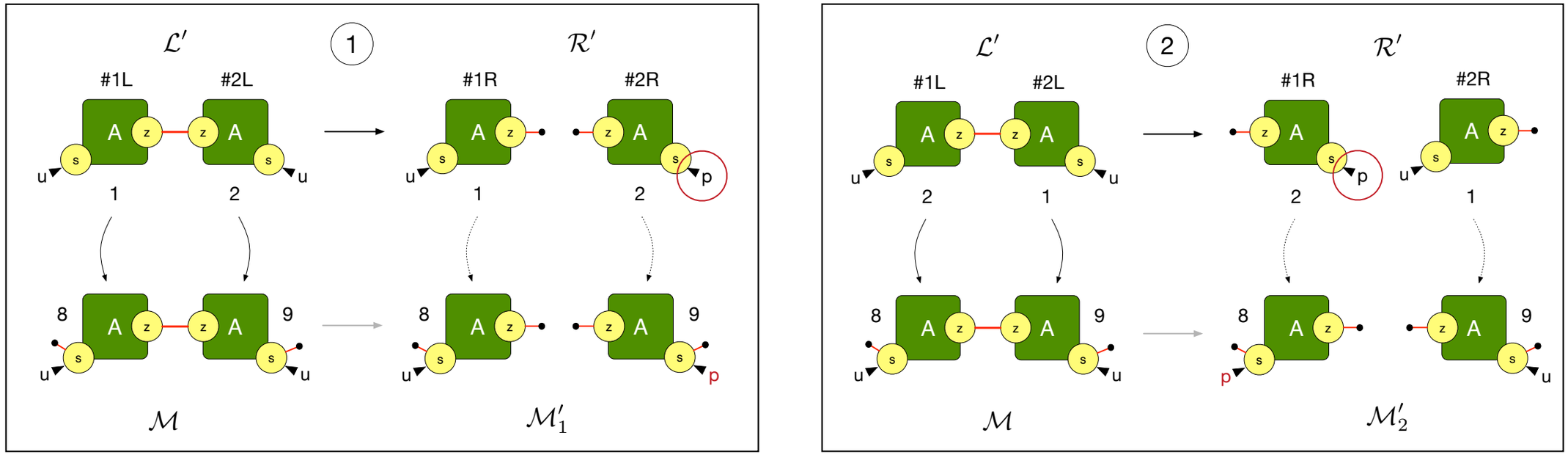
Rule Symmetry

Case A



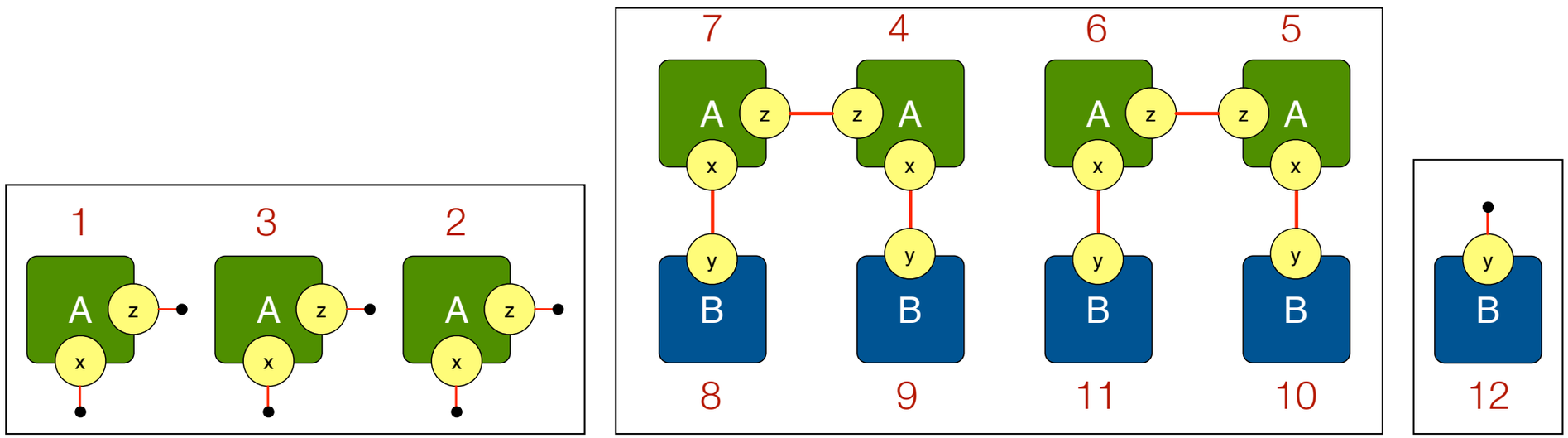
indistinguishable microstates; two embeddings represent the same physical event

Case B



distinguishable microstates; two embeddings represent different physical events

Symmetry



number of distinct classes of connected components

total symmetries in pattern

$$\omega_{\mathcal{P}} = \prod_{c=1}^{C(\mathcal{P})} n_c! \prod_{i=1}^{C(\mathcal{P})} (\omega_c)^{n_c}$$

$$3!\ 2!\ 1!\ 1^3\ 2^2\ 1^1 = 48$$

number of isomorphic instances in component c

number of symmetries of component c

Your Options?

rule $\mathcal{L}_r \longrightarrow \mathcal{R}_r \ @ \ \gamma_r$

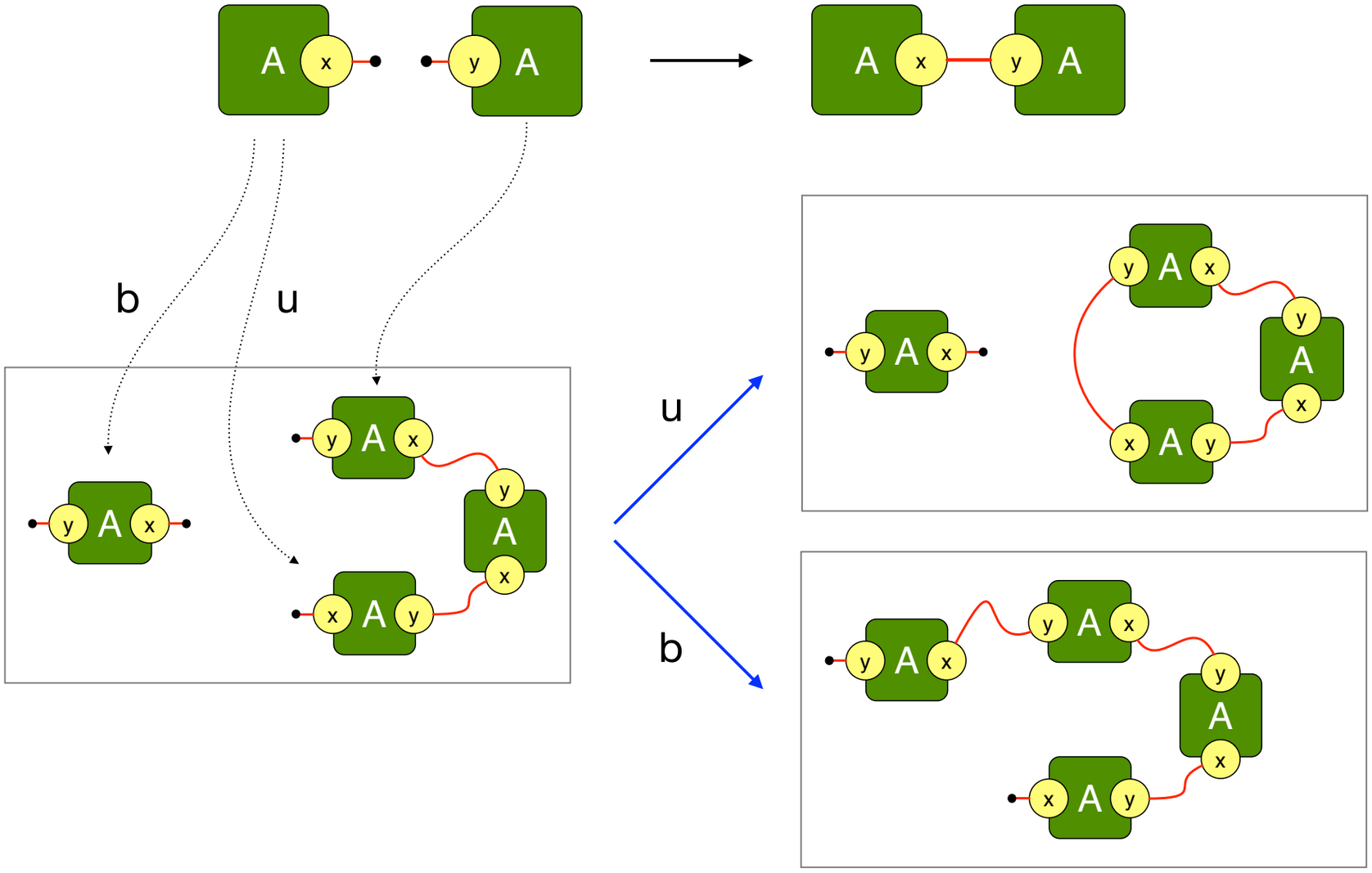
activity $\alpha_r = \gamma_r[\mathcal{L}_r; \mathcal{M}]\Omega_r$

(1) deterministic view $\Omega_r = \dfrac{1}{\#\,\mathrm{auto}(\mathcal{L}_r \to \mathcal{R}_r)}$

(2) nondeterministic view $\Omega_r = \dfrac{1}{\omega_{\mathcal{L}_r}}$

(3) “mathematical” view $\Omega_r = 1$

Molecular Ambiguity



$$\mathcal{L}_r \longrightarrow \mathcal{R}_r \;\; @ \;\; \gamma_{r,2} \;\; (\gamma_{r,1})$$