Méthodes Formelles pour la Biologie des Systèmes

François Fages

Project-Team Lifeware

http://lifeware.inria.fr/

Institut National de Recherche en Informatique et Automatique

Inria Saclay – Ile de France

Two Lectures

Formal Methods for Systems Biology

1. Introduction to Chemical Reaction Networks (CRNs)
   TD enzymatic Lotka-Volterra oscillator
2. Continuous semantics by ordinary differential equations
   TD enzyme kinetics
3. Formal behaviours in temporal logics and model-checking
   TD robustness and parameter search

Systems Biology: How to build and validate CRN models ?

The Cell: a Chemical Analog Computer

1. Turing completeness of continuous CRNs
   TD synthesis of oscillators and switches
2. Logical circuits for diagnosis
   TD doctor in the cell

Synthetic Biology: How to program with CRNs ?

Biochemical Kinetics

Study of the concentration of chemical substances as a function of time.

Molecular species: A_1 ,…, A_m

|A|=Number of molecules A

[A]= |A| / Volume (e.g. unit ML^{-1}) Concentration of A in the solution

noted also A by abuse of notation

Molecular solutions: S, S', …:

multiset of molecules

linear expression with stoichiometric coefficients: $c_1 * A_1 + \ldots + c_n * A_n$

Reactions given with rate functions: f for S => S'

Mass Action Law Kinetics

Assumption: infinite diffusion speed, dilute solutions, low concentrations

Law: The number of reactions per time unit is proportional to the number of reactant molecules present in the solution

The rate of a reaction **A + B => C** is **k*[A]*[B]** for some reaction rate constant **k**

Continuous semantics: the time evolution of concentrations obeys the ODE

$$dA/dt = -k\ A\ B \qquad dB/dt = -k\ A\ B \qquad dC/dt = k\ A\ B$$

Stochastic semantics: SSA with same assumption of perfect diffusion

- infinite diffusion speed

Multi-agent semantics: simulation of diffusion

- finite diffusion speed for macromolecules

Interpretation of Rate Constants

Complexation rate constant: *probability of reaction upon collision*

specificity, affinity, position of matching surfaces and energy of bonds

Decomplexation rate constant: *total energy of the bonds*

Diffusion speeds:

small molecules > substrates > enzymes

cells are 10-100µm long, full of compartments

average travel in one random walk for one enzyme:

1 µm in 1s, 2µm in 4s, 10µm in 100s

500000 random collisions per second with a substrate concentration of 10^{-5}

50000 random collisions per second with a substrate concentration of 10^{-6}

Possibility of fast computations compared to DNA computation

Reaction Rate Functions

Mass action law kinetics

k*A for A => B

k*A*B for A+B => C

k*A^m*B^n for m*A + n*B => R

Guldberg and Waage, 1864

Henri-Michaelis-Menten kinetics

Vm*A/(Km+A) for A => B

Henry 1903, Michaelis and Menten 1913

Hill kinetics

Vm*A^n/(Km+A^n) for A => B

Archibald Hill 1910

Origin and justification of these other rate functions?

come from reductions of mass action law CRNs

Continuous Semantics of a CRN

To a set of reactions $\{ f_i \text{ for } S_i \Rightarrow S'_i \}_{i=1,\ldots,n}$ given with rate functions f_i

one associates the Ordinary Differential Equations (ODE) over $\{A_1, \ldots, A_k\}$

$$dA_k/dt = \Sigma^n_{i=1} (r_i(A_k) - l_i(A_k)) * f_i = \Sigma^n_{i=1} m_i(A_k) * f_i$$

where $l_i(A)$ is the stoichiometric coefficient of A in S_i

$r_i(A)$ is the stoichiometric coefficient of A in S'_i

$m_i = r_i - l_i$ is the net stoichiometric vector of reaction i

In matrix form: $\dot{x} = M.f(x)$

Solution x(t) by numerical integration

Steady states: concentrations x such that $M.f(x) = 0$ non-linear problem

Elementary modes: fluxes v such that $M.v = 0$ linear problem

Numerical Integration Methods

$dX/dt = f(X)$ with initial conditions X_0

Idea: discretize time t_0, $t_1=t_0+\Delta t$, $t_2=t_1+\Delta t$, … and compute a trace

(t_0,X_0), (t_1,X_1), …, (t_n,X_n)…

Forward *Euler*'s method: $t_{i+1}=t_i+\Delta t$ $\quad X_{i+1}=X_i+f(X_i)*\Delta t$

estimation error$(X_{i+1})=|f(X_i)-f(X_{i+1})|*\Delta t$

Midpoint method *(Runge-Kutta):* intermediate computation at $\Delta t/2$

Adaptive step size: $\Delta t_{i+1}=\Delta t_i/2$ while error>$error_{max}$, otherwise $\Delta t_{i+1}=2*\Delta t_i$

Implicit method *(Rosenbrock)*: solve $X_{i+1}=X_i+f(X_{i+1})*\Delta t$ by root finding

Biocham-3: Rosenbrock method implemented in Prolog

Biocham-4: GSL library with implicit method by default,

not as good as Biocham-3 → back to Biocham-3 implementation

+ Events: hybrid systems (both continuous ODE and discrete event transitions)

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \rightarrow^{c1} C \rightarrow^{c3} E+P$

$E+S \leftarrow^{c2} C$

Mass action law kinetics ODE:

$dE/dt =$

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \xrightarrow{c1} C \xrightarrow{c3} E+P$

$E+S \xleftarrow{c2} C$

Mass action law kinetics ODE:

$dE/dt = -c_1ES+(c_2+c_3)C$

$dS/dt =$

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \xrightarrow{c1} C \xrightarrow{c3} E+P$

$E+S \xleftarrow{c2} C$

Mass action law kinetics ODE:

$dE/dt = -c_1ES+(c_2+c_3)C$

$dS/dt = -c_1ES+c_2C$

$dC/dt =$

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \xrightarrow{c1} C \xrightarrow{c3} E+P$

$E+S \xleftarrow{c2} C$

Mass action law kinetics ODE:

$dE/dt = -c_1ES+(c_2+c_3)C$

$dS/dt = -c_1ES+c_2C$

$dC/dt = c_1ES-(c_3+c_2)C$

$dP/dt =$

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \xrightarrow{c1} C \xrightarrow{c3} E+P$

$E+S \xleftarrow{c2} C$

Mass action law kinetics ODE:

$dE/dt = -c_1ES+(c_2+c_3)C$

$dS/dt = -c_1ES+c_2C$

$dC/dt = c_1ES-(c_3+c_2)C$

$dP/dt = c_3C$

Two conservation laws (i.e. species s.t. $\Sigma^n_{i=1}Mi = constant$ since $\Sigma^n_{i=1} dMi/dt = 0$):

Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

$E+S \rightarrow^{c1} C \rightarrow^{c3} E+P$

$E+S \leftarrow^{c2} C$

Mass action law kinetics ODE:

$dE/dt = -c_1ES+(c_2+c_3)C$

$dS/dt = -c_1ES+c_2C$

$dC/dt = c_1ES-(c_3+c_2)C$

$dP/dt = c_3C$

Two conservation laws (i.e. species s.t. $\Sigma^n_{i=1}Mi$ = *constant* since $\Sigma^n_{i=1}\, dMi/dt = 0$):

$E+C=constant=E_0+C_0, \quad S+C+P=constant=S_0+C_0+P_0,$

we can eliminate E and P and get the equivalent parametric system

$dS/dt = -c_1(E_0+C_0-C)S+c_2C$

$dC/dt = c_1(E_0+C_0)S-(c_1S+c_2+c_3)C$

we shall further assume $C_0=0$, $P_0=0$

Slow/Fast Time Scales

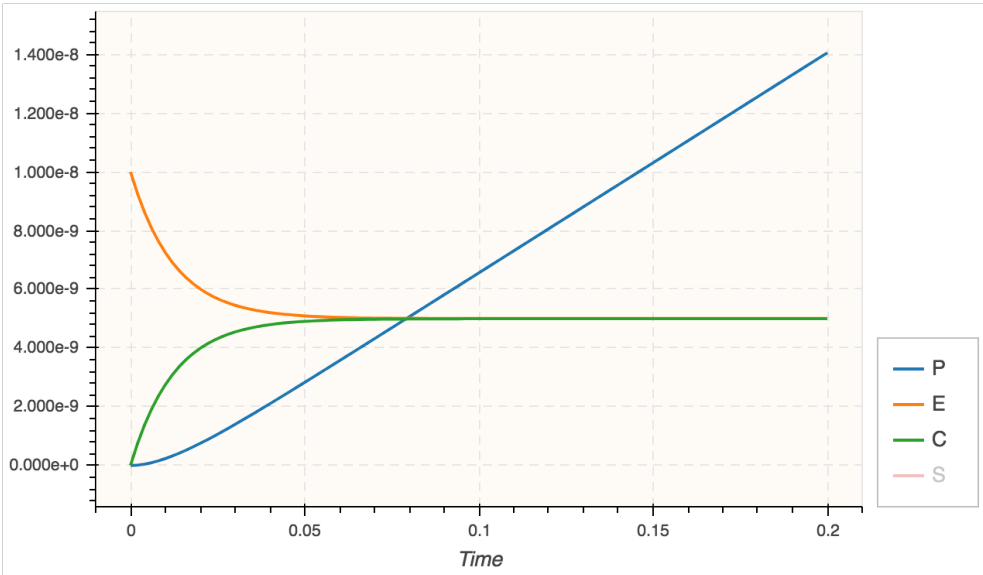
Hydrolysis of benzoyl-L-arginine ethyl ester by trypsin (protein of 247 amino acids)

present(E,1e-8). present(S,1e-5). *E << S*

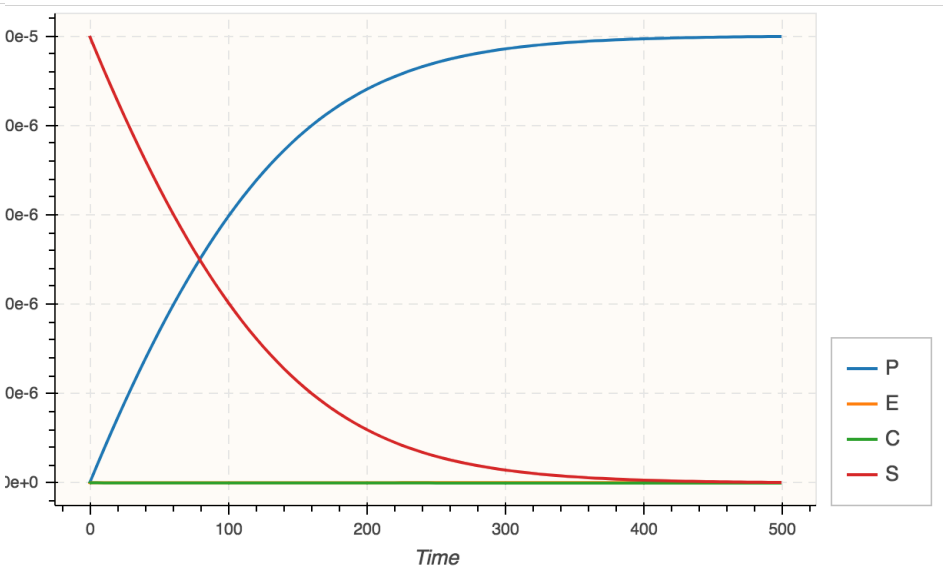
parameter(c1=4e6, c2=25, c3=15). *c1 >> c2, c3*

c1*E*S for E+S => C. c2*C for C => E+S . c3*C for C => E+P.

Complex formation 5e-9 in 0.1s



Product formation 1e-5 in 400s



Michaelis Menten Reduction

Assume dC/dt=0…

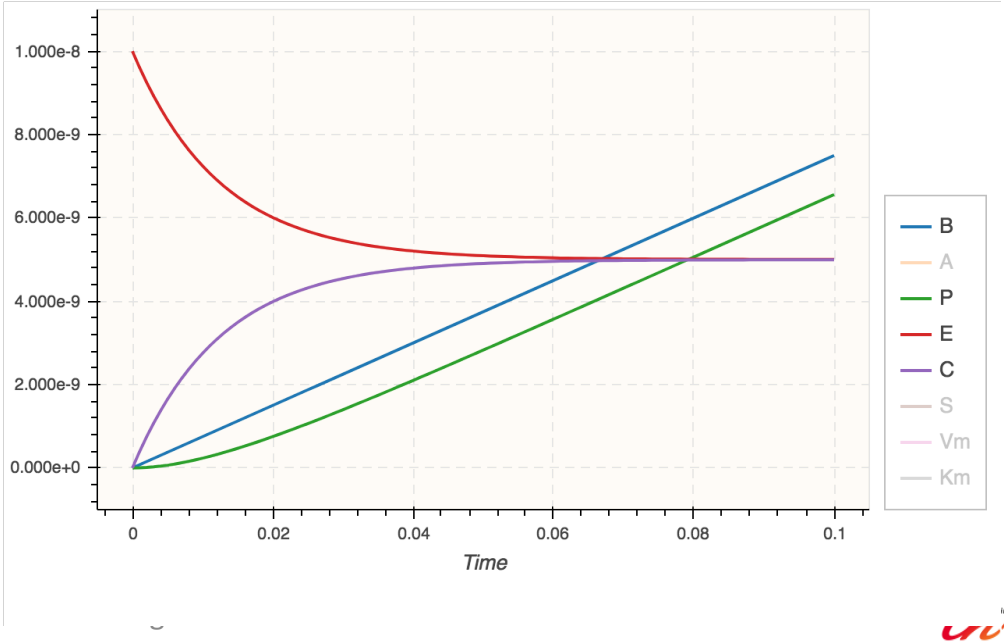
Vm*S/(Km+S) for S => P.

parameter(E0=1e-8).

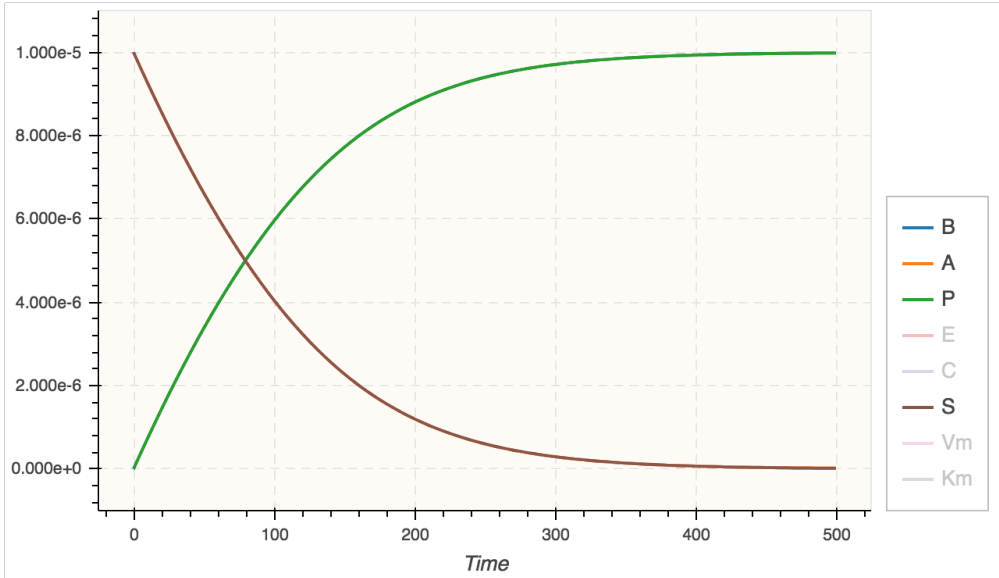
function(Vm=c2*E0).

function(Km= (c2+c3)/c1).

Slightly different early trajectory in 0.1s



Same trajectory in 400 s



TD2: Enzyme Kinetics

BIOCHAM-4: version online http://lifeware.inria.fr/biocham4/online/

Michaelis-Menten enzymatic reaction CRN

- CRN of 3 reactions with mass action law kinetics
- Real parameter values for the hydrolysis of benzoyl-L-arginine ethyl ester by trypsin (protein of 247 amino acids)

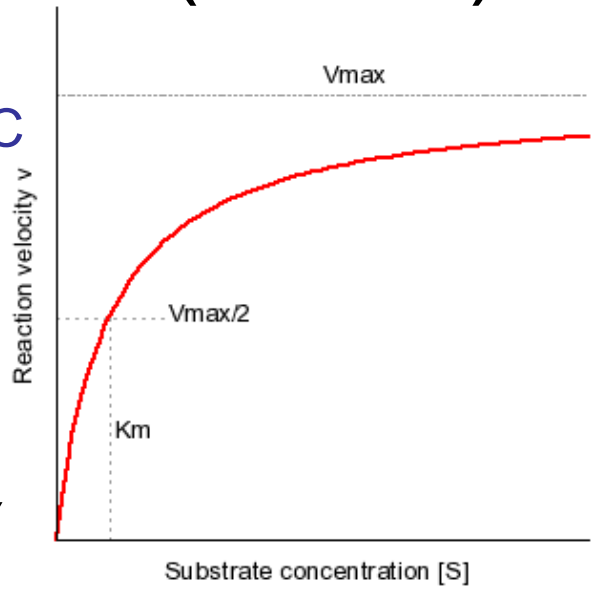
In [1]:
```
present(E,z). parameter(z=1e-8).
present(S,s). parameter(s=1e-5).
absent(C).
absent(P).
```

In [2]:
```
parameter(k1=4e6, k2=25, k3=15).
```

In [3]:
```
k1*E*S for E+S => C.
k2*C for C => E+S.
k3*C for C => E+P.
```

Quasi-Steady State Approximation (QSSA)

Assume quasi-steady state $dC/dt \simeq 0 \simeq c_1E_0S-(c_1S+c_2+c_3)C$



Then $C = c_1E_0S/(c_1S+c_2+c_3)$

$= E_0S/(K_m+S)$

where $K_m=(c_2+c_3)/c_1$

substrate concentration with half maximum velocity

We get $dP/dt = -dS/dt = -c_1(E_0+C_0-C)S+c_2C$

$= V_mS / (K_m+S)$ where $V_m= c_3E_0$

maximum velocity at saturing substrate concentration

Michaelis-Menten kinetics: **$V_mS /(K_m+S)$ for S => P**

C and E are eliminated but sometimes E is re-injected as a variable…

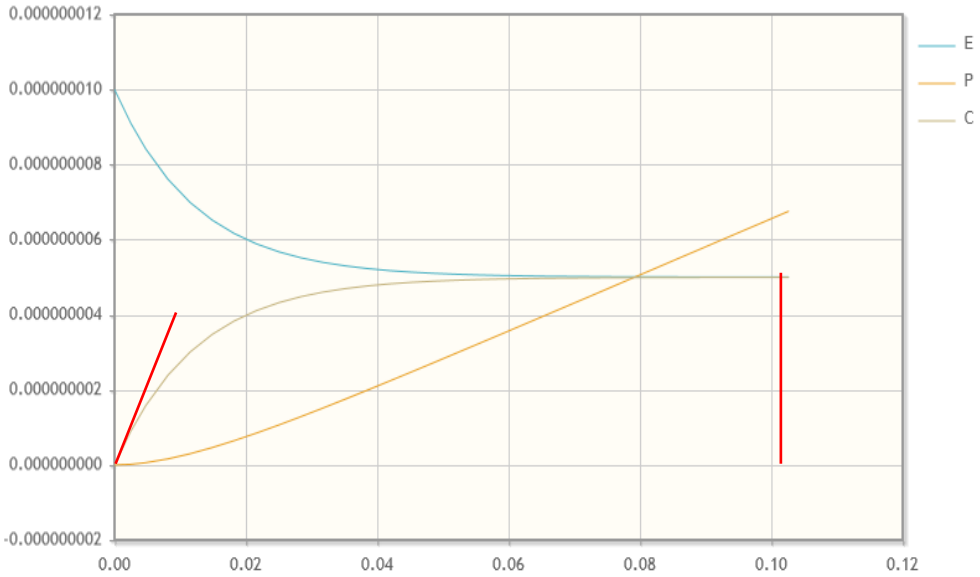
c_3*E*S /(K_m+S) for S+E => E+P

(Weak) Justification by Preservation of Time Scales

"Time taken for a significant change"

Time scale of f(t) ≈ $\frac{fmax - fmin}{|\frac{df}{dt}|max}$

Time scale of C(t) ≈ $\frac{5e{-}8}{4e{-}8/0.01} = \frac{1}{80}$



Formally, suppose $S(t)=S_0$ we get

$dC/dt = c_1(E_0+C_0)S_0 - (c_1S_0+c_2+c_3)\,C$

$C(t) = (C_0-\overline{C})\,\mathbf{e^{-kt}} + \overline{C}$ where $k= c_1S_0+c_2+c_3$ and $\overline{C} = (E_0+C_0)S_0/(K_m+S_0)$

Taking k^{-1} as time scale of e^{-kt} (i.e. decrease of $e^{-1}\approx 1/3$ in k^{-1} time)

gives in the Trypsin example $1/(10^{-5}.4.10^6+25+15)=1/80 = 0.0125s$

Validity Condition of QSSA

When the time scale of S is much longer than the time scale of C…

S varies from S_0 to 0

| $dS/dt = -c_1(E_0-C)S+ c_2C$ | is maximal when $S=S_0$, $C=C_0$

Time scale of S ≈ $1/(c_1E_0)$

In the Trypsin example $1/(4.10^6. 10^{-8})$≈25s

Validity condition: $c_1E_0 << k = c_1S_0+c_2+c_3$ (e.g. Trypsin: $4.10^{-2}<<80$)

i.e. QSSA valid when $E_0 << S_0+K_m$

in particular when $E_0<<S_0$ [Briggs and Haldane 1925]

Better justification by approximation in all time points [Tikhonov theorem]

Quasi-Equilibrium Approximation (QE)

Assume reaction equilibrium $c_1 ES \simeq c_2 C$ *(fast complexation/decomplexation cycle)*

From $E=E_0-C$ we get $c_1 E_0 S - c_1 CS = c_2 C$

$C = c_1 E_0 S/(c_2+c_1 S)$

$C = E_0 S/(K_d+S)$

where $K_d=c_2/c_1$

substrate concentration with half maximum velocity

$E+S \rightarrow^{c1} C \rightarrow^{c3} E+P$ $\quad c_3 << c_2$

$E+S \leftarrow^{c2} C$

Hence $dP/dt = -dS/dt = V_m S / (K_d+S)$

where $V_m = c_3 E_0$

maximum velocity at saturing substrate concentration

Michaelis-Menten quasi-equilibrium kinetics: $\mathbf{V_m S / (K_d+S)}$ **for** $\mathbf{S \Rightarrow P}$

justified when complex equilibrium reached on a fast time scale $c_3/c_2 << 1$

Conclusion

Michaelis-Menten kinetics, Hill kinetics of order *n* and more general kinetics come from reductions of elementary CRNs with Mass Action law kinetics

QSS approximation: projection on slow dynamics variables

- fast dynamics species E, C act as slaves of slow species S

QE approximation: elimination of fast reaction equilibria

The slow/fast separation of the CRN dynamics may change over time

- resulting in a hybrid automaton of piece-wise reduced CRNs
- helps to understand the CRN dynamics

Using Michaelis-Menten kinetics in a CRN may be not justified (and wrong)