On the Computational Complexity of Rate-Independent CRN/ODE Computations

François Fages

Joint work with Elisabeth Degrand and Sylvain Soliman CMSB 2020

INRIA Saclay

LIFEWARE group: *Computational Systems Biology and Optimization*

http://lifeware.inria.fr/

Analog Computations with Chemical Reaction Networks (CRN)

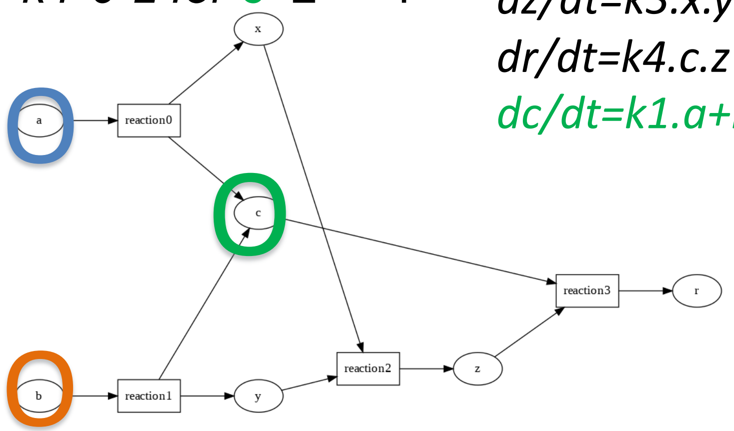
CRN:

*k1*a for* a => x+c

*k2*b for* b => y+c

*k3*x*y for* x+y => z

*k4*c*z for* c+z => r



ODE

da/dt=-k1.a

db/dt=-k2.b

dx/dt=k1.a-k3.x.y

dy/dt=k2.b-k3.x.y

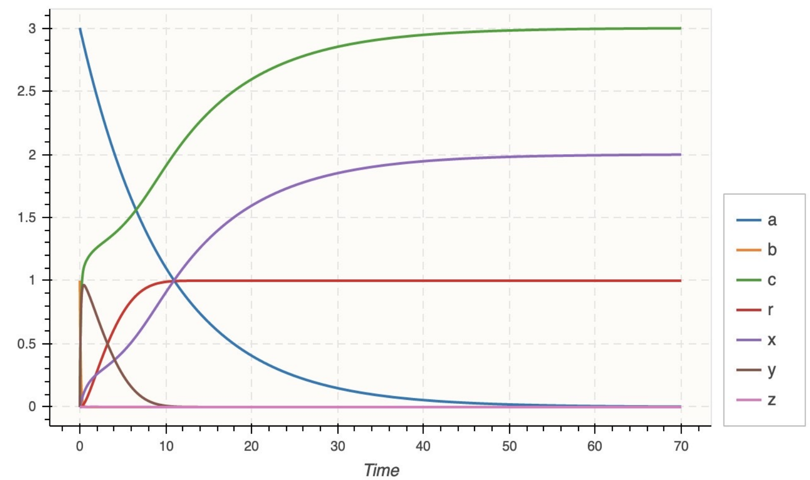
dz/dt=k3.x.y-k4.c.z

dr/dt=k4.c.z

dc/dt=k1.a+k2.b-k4.c.z

Input: a, b   Initialization: x=y=z=r=c=0

Output: c



Computed function at steady state:

c(∞)= max(a(0),b(0))

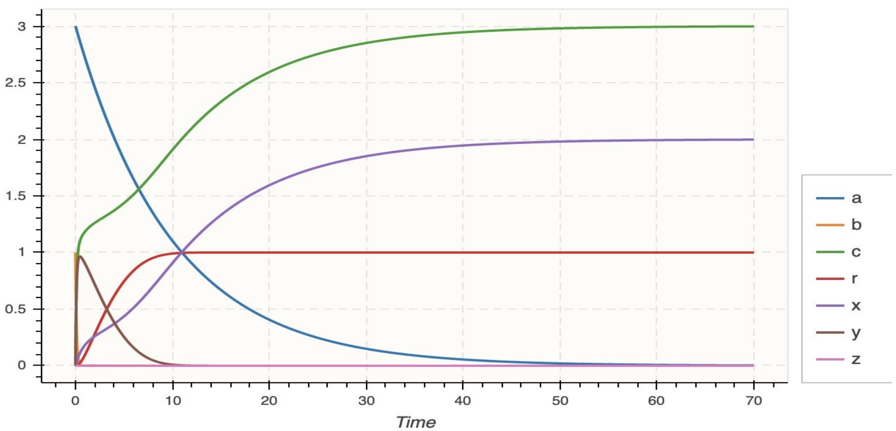
Theorem [F, Le Guludec, Bournez, Pouly CMSB 2017]

A real function is Turing-computable (in Ptime) if and only if it can be computed by a CRN over a finite set of molecular species (with polynomial length trajectories)

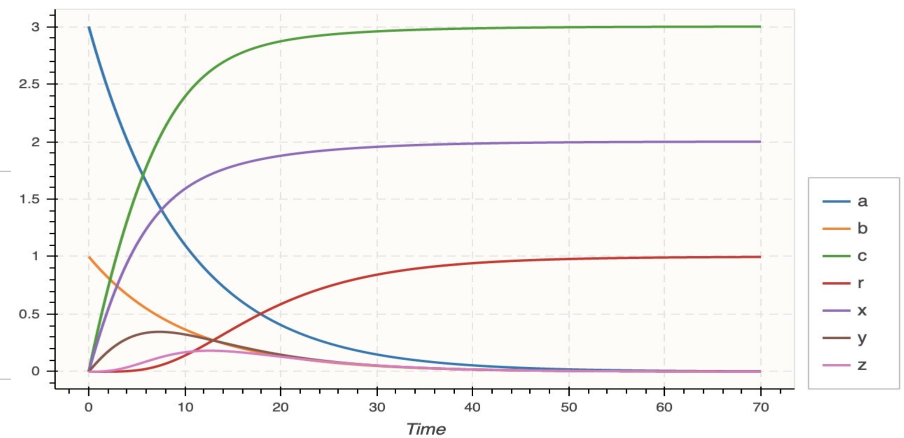
Rate-Independent CRN/ODE Computation

Input: a(0)=3 b(0)=1 **Result c*=3 independently of the reaction rates**

k1=0.1, k2=10.0, k3=1, k4=100.0:



k1=0.1, **k2=0.1, k3=0.1, k4=0.1**:



The I/O functions computed by that CRN are independent of the kinetics terms

a => x+c

b => y+c

x+y => z

c+z => r

$c^* = \max(a(0),b(0)) = a(0)+b(0)-\min(a(0),b(0))$

$x^* = \max(0,a(0)-b(0))$

$y^* = \max(0,b(0)-a(0))$

$r^* = \min(a(0),b(0))$

$z^* = 0$, $a^* = 0$, $b^* = 0$

Absolute robustness
Ideal circuit designs
for synthetic biology

Mathematical Characterization of the Functions Computed by Rate-Independent CRNs

Theorem [Chen Doty Soloveichik 2014 ITCS]
A real function is computable by a rate-independent CRN if and only if it is a positive continuous piecewise linear function with rational coefficients.

Theorem [Chalk Kornerup Reeves Soloveichik 2018 CMSB]
A real function is computable by a composable (i.e. not consuming its inputs) rate-independent CRN iff it is a superadditive (i.e. $f(x+y) \geq f(x)+f(y)$) positive continuous piecewise rational linear.

Does not help to show that a given CRN is rate-independent
nor to design a rate-independent CRN

Graphical conditions on the CRN
ensuring rate-independence ?

Monomolecular Rate-Independent CRN Structures

A => B

output B: computes the *identity function* B(∞)=A(0)+B(0) rate-independent !

output A: computes the *zero function* A(∞)=0 rate-independent !

A => C

B => C

Harmless join

output C: computes the *sum* C(∞)=A(0)+B(0)+C(0) rate-independent

output A: computes the *zero function* A(∞)=0 rate-independent !

C => A

C => B

Harmfull fork

output A: computes $A(\infty)=\frac{\alpha}{\beta}C(0)+A(0)$ not rate-independent !

C => A

C => B

B => C

Harmless fork with a circuit

output A: computes the *sum* A(∞)=C(0)+B(0)+A(0) rate-independent !

Bimolecular Rate-Independent CRN Structures

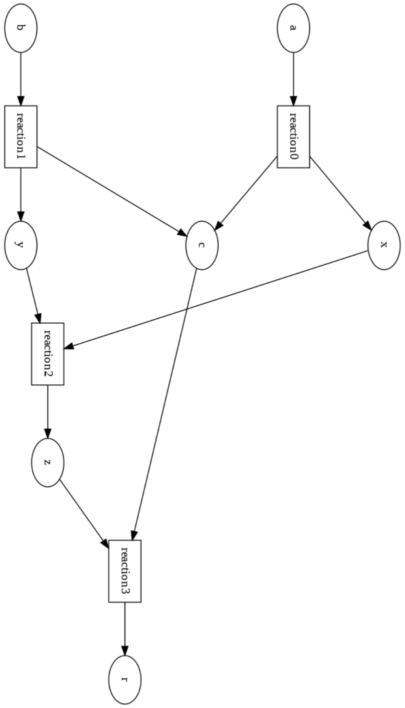
A+B => C

output C: computes C(∞)=*minimum*(A(0),B(0))+C(0) rate-independent !

output B: computes B(∞)=max(0,B(0)-A(0)) rate-independent !

C => A+B

output A: makes copies A(∞)=C(0)+A(0) rate-independent !



A => X+C

B => Y+C   Rate-independent on all species, why ?

X+Y => Z

C+Z => R

Definition *A* funnel *CRN is a CRN that is:*

- *fork-free on species nodes*
- *circuit-free*
- *synthesis-free*

Theorem [Degrand F Soliman CMSB 2020] A funnel CRN (ODE) is rate-independent for any output species.

Sufficient condition, not a necessary condition (e.g. harmless fork with circuit)

Global Rate-Independence Condition

Lemma *The structure of a funnel CRN* C *is a DAG with no reaction source node*

Lemma *All steady fluxes of a funnel CRN* C *are equal to* 0*.*

Proof: by induction on the topological order of the graph.

Definition *We shall denote* x^+_i *the total amount of species* x_i *available in an execution of the corresponding ODE system.*

$$x_i^+ = x_i^0 + \int_0^{+\infty} \frac{dx_i^+}{dt} = x_i^0 + \int_0^{+\infty} \sum_{P_j(x_i) > R_j(x_i)} (P_j(x_i) - R_j(x_i)) f_j$$

Theorem [Degrand F Soliman CMSB 2020] *The ODE system associated to* a funnel CRN *has a single steady state* x^* *that does not depend on the kinetic functions* f_i *of* C*.*

Corollary *A funnel CRN is globally rate-independent for all species.*

Theorem [Degrand F Soliman CMSB 2020] Any function computable by a rate-independent CRN/ODE is computable by a funnel CRN/ODE.

Proof: by Chen-Doty-Soloveichik's characterization and Ovchinnikov's max-min representation of continuous piecewise linear functions with rational coefficients.

Rate-Independence for « Persistent » Outputs

The harmless-fork-with-circuit CRN is rate-independent on outputs A, B, C

C => A   C => B   B => C

Def. A species x

- is a product of a CRN if it can only increase: $\forall i\ R_i(x) \leq P_i(x)$
- is structurally persistent if it is covered by a P-invariant S, $\forall i\ S.R_i = S.P_i$, and does not belong to a critical (gets empty) siphon (when empty remains empty)

Theorem [Degrand F Soliman CMSB 2020] Any CRN (ODE) is rate-independent on its structurally persistent products.

Proof: P-invariant covering ensures boundedness and convergence for products.

The species reaching 0 are localized in siphons and exclude persistent outputs.

Implemented in BIOCHAM using Constraint Logic Programming for computing P-invariants and siphons [Nabli, Martinez, F, Soliman 2016 *Constraints*]

Evaluation against BioModels repository

590 CRNs from SBML models (many not well-formed CRNs)

[F Gay Soliman 2011 *TCS*]

94 with rate-independent products

29 with non trivial rate-ind. products

2 globally rate-ind. CRNs

Size of those 29 models:

- 4-136 species
- 2-316 reactions

Constraint solving time:

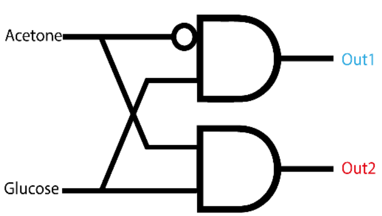
- between 0.07 and 151 seconds
- except 2 timeouts >240s

Biomodel#	#species	#reactions	#products	#RI	#NTRI	NTRI-product species	Time (s)
037	12	12	2	2	2	Yi, Pi	0.950
104	6	2	3	3	1	species_4	0.074
105	39	94	11	3	1	AggP_Proteasome	63.366
143	20	20	4	1	1	MLTH_c	3.333
178	6	4	1	1	1	lytic	0.139
227	60	57	2	1	1	s194	17.299
259	17	29	1	1	1	s10	2.308
260	17	29	1	1	1	s10	2.310
261	17	29	1	1	1	s10	2.297
267	4	3	1	1	1	lytic	0.086
283	4	3	1	1	1	Q	0.053
293	136	316	14	4	3	aggE3, aggParkin, AggP_Proteasome	>240
313	16	16	4	2	1	IL13_DecoyR	2.071
336	18	26	1	1	1	IIa	4.148
344	54	80	7	2	1	AggP_Proteasome	>240
357	9	12	1	1	1	T	0.561
358	12	9	4	2	1	Xa_ATIII	0.892
363	4	4	1	1	1	IIa	0.067
366	12	9	4	2	1	Xa_ATIII	0.901
415	10	5	7	7	7	s10, s11, s12, s13, s14, s9, s15	0.894
437	61	40	22	8	1	T	16.109
464	14	10	6	3	1	s12	2.282
465	16	14	5	5	1	s23	59.554
525	18	19	8	3	1	p18inactive	33.479
526	18	19	8	3	1	p18inactive	33.858
540	22	11	12	11	8	s14, s15, s16, s17, s18, s19, s20, s21	56.134
541	37	32	13	9	7	s14, s15, s16, s17, s18, s19, s21	31.573
559	90	136	18	2	2	s493, s502	150.954
575	76	58	9	1	1	DA_GSH	66.806

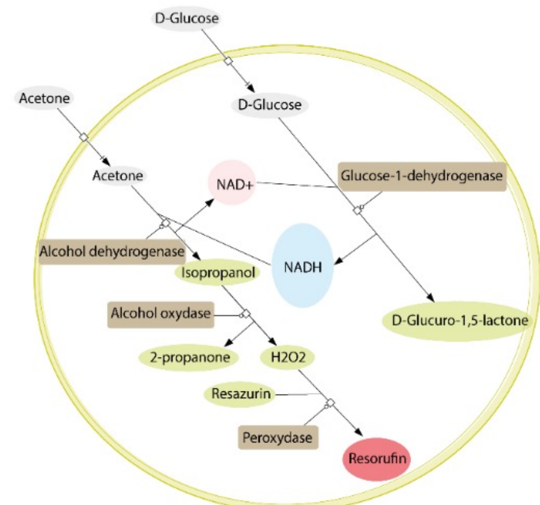
Relevance to Synthetic Biology

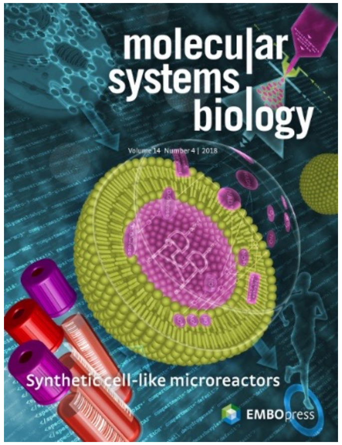
- Absolute robustness through rate-independence by design
  - Graphical constraints for CRN design
  - Constraint-based synthesis method

- Concrete chemical implementation with « morally » rate-independent CRNs
  - Rate-independent CRN kernel
  - Plus reverse reactions breaking formal rate-independence (limited robustness)
  - Boolean function CRNs for coma diagnosis [Courbet Amar F Renard Molina 2018 MSB]



Glucose	Acetone	Out2	Out1
0	0	0	0
1	0	0	1
0	1	0	0
1	1	1	0





Time Complexity of the Functions Computed by Rate-independent CRN/ODE

Proposition. The input-output function computed by a rate-independent CRN/ODE with rational kinetics and rational inputs is computable in time $O(n \log n)$ and log space on a multitape Turing machine where n is the number of bits to encode the rational inputs with exact precision.

Proof. By Chen Doty Soloveichik 2014 ICTS the function computed by a rate independent CRN is a continuous positive piecewise linear function

$y = \sum_{i=1}^{k} a_i x_i = \sum_{i=1}^{k} a_i x_i$ where

- The fixed k rational coefficients a_i's can be encoded with by pairs of integers
- the x_i's are the input rational numbers encodable by pairs of integers of n bits
- and y is the output rational number similarly encoded with a pair of integers

By Harvey and Van der Hoeven 2021 multiplication is in $O(n \log n)$ on a multitape Turing machine (required for the output with common denominator),

addition and substraction are in $O(n)$ as piece-wise discrimination.

Hence the function computed by a rate-independent CRN/ODE is computable in $O(n \log n)$ time and log space.

Analog Characterization of Low Complexity Class?

Rational Polynomials can also be evaluated in O(n log n) time and log space on a multitape Turing machine with binary encoding of the inputs.

There is no notion of O(n log n) nor O(n) complete problems.

Characterizing L the Logspace complexity class for decision problems ?

- Turing machine with 3 tapes: R input n, RW work $\log n$, W output $P(n)$
- L$= DSPACE(O(\log n)) \subseteq DTIME\left(2^{O(\log n)}\right) = DTIME\left(n^{O(1)}\right) =$ P
- L⊆ NL⊆ P but it is unknown wheter L=NL, NL=P, even L=NP

- 2-SAT and ST-connectivity are NL-complete under log-space reduction
- ST-connectivity is also NL-complete under First-Order reduction
- L = languages expressible in FO + transitive closure predicate (data complexity)

Logspace reduction $\leq_L$

Defined on languages $L_1 \leq_L L_2$ such that $L_2 \in$ L implies $L_1 \in$ L
$L_1 \leq_L L_2$ if $x \in L_1 \leftrightarrow f(x) \in L_2$ where f is computable in LOGSPACE.

Space complexity for computing a function:

- Turing machine with 3 tapes: R input n, RW work $\log n$, W output $P(n)$
- Problem f(x) may be large, so consider TM computing f(x) bitwise on the fly

f ∈ SPACE(S(n)) if both languages L' and L" are in SPACE(S(n))

L'= {(x, i) : f_i(x) = 1} and

L" = {(x, i) : i ≤ |f(x)|}

This gives

- Transitivity: $L_1 \leq_L L_2$ and $L_2 \leq_L L_3$ implies $L_1 \leq_L L_3$
- The composition of two L computable functions is L computable
- Yet every decision problem in L is L-complete under $\leq_L$ (useful w.r.t. P not L)
- Stronger notion of reduction for meaningful L-completeness, e.g. $\leq_{FO}$

Conclusion

- Funnel CRN/ODE structure: fork-free, circuit-free, synthesis-free
  - Theorem. Any funnel CRN/ODE is rate-independent on all outputs
  - Theorem [Soloveichik et al. 2014]. Any function computed by a rate independent CRN/ODE is a continuous piecewise linear rational function.
  - Theorem. Any function computed by a rate independent CRN/ODE can be computed by a funnel CRN/ODE

- Relevance to low analog computational complexity classes
  - Functions computed by rate-independent CRN/ODE are in $O(n \log n)$ time and logspace on a multitape Turing machine with binary representation of the inputs
  - Investigate Logspace completeness with an appropriate notion of reduction? $\leq_{QFFO}$?

- Meaning for rate-independent difference equations?

Lifeware team at INRIA Saclay

François Fages (Research Scientist, Prof X)
Mathieu Hemery (Engineer)
Anna Niaraki (Assoc. Prof. Univ. Evry)
Sylvain Soliman (Research Scientist)
Sahar Aghakhani (PhD Univ. Evry)
Marine Collery (PhD IBM)
Elisabeth Degrand (PhD Inria)
Eléa Greugny (PhD Johnson&Johnson)
Jeremy Grignard (PhD Servier)
Julien Martinelli (PhD INSERM-Inria)

Contract collaborations:

- O. Bournez, Ecole Polytechnique, Palaiseau, A. Pouly, CNRS Univ. Paris
  - ANR δifference project: Computing with Discrete Differential Equations
- IBM on rule learning from temporal data, CIFRE Johnson&Johnson, CIFRE Servier
- F. Molina CNRS Sys2diag, Montpellier, J.H. Jiang, NTU, Taiwan
  - ANR-MOST BIOPSY project: Biochemical Programming Systems
- A. Weber, Univ. Bonn, T. Sturm, Inria Nancy, O. Radulescu, Montpellier, S. Walcher
  - ANR-DFG SYMBIONT project: Symbolic Methods for Biological Networks
- A. Ballesta, Institut Curie, INSERM Paris