From Processes to ODEs by Chemistry

Luca Cardelli

Microsoft Research
Cambridge UK

2008-09-10

http://LucaCardelli.name

Outline

- Chemical Reactions
- The Chemical Ground Form (CGF)
- From CGF to ODEs
  - Discrete Chemistry
  - Continous Chemistry
  - ODEs
  - CGF to ODEs Examples
- The Chemical Parametric Form (CPF)
  - From CPF to CGF
  - CFP to ODEs Example
- Algebraic Laws by ODEs
- Conclusions

Chemical Reactions

Chemical Reactions

$A \to^r B_1 + \ldots + B_n$ $(n \geq 0)$	Unary Reaction	$d[A]/dt = -r[A]$	Exponential Decay
$A_1 + A_2 \to^r B_1 + \ldots + B_n$ $(n \geq 0)$	Hetero Reaction	$d[A_i]/dt = -r[A_1][A_2]$	Mass Action Law
$A + A \to^r B_1 + \ldots + B_n$ $(n \geq 0)$	Homeo Reaction	$d[A]/dt = -2r[A]^2$	Mass Action Law

(assuming $A \neq B_i \neq A_j$ for all i,j)

No other reactions!

JOURNAL OF CHEMICAL PHYSICS VOLUME 113, NUMBER 1

The chemical Langevin equation

Daniel T. Gillespie[a)]
Research Department, Code 4T4100D, Naval Air Warfare Center, China Lake, California 93555

Genuinely *trimolecular* reactions do not physically occur in dilute fluids with any appreciable frequency. *Apparently* trimolecular reactions in a fluid are usually the combined result of two bimolecular reactions and one monomolecular reaction, and involve an additional short-lived species.

Chapter IV: Chemical Kinetics
[David A. Reckhow , CEE 572 Course]

... reactions may be either elementary or non-elementary. Elementary reactions are those reactions that occur exactly as they are written, without any intermediate steps. These reactions almost always involve just one or two reactants. ... Non-elementary reactions involve a series of two or more elementary reactions. Many complex environmental reactions are non-elementary. In general, reactions with an overall reaction order greater than two, or reactions with some non-integer reaction order are non-elementary.

THE COLLISION THEORY OF REACTION RATES www.chemguide.co.uk

The chances of all this happening if your reaction needed a collision involving more than 2 particles are remote. All three (or more) particles would have to arrive at exactly the same point in space at the same time, with everything lined up exactly right, and having enough energy to react. That's not likely to happen very often!

Trimolecular reactions:

$A + B + C \to^r D$

the measured "r" is an (imperfect) aggregate of e.g.:

$A + B \leftrightarrow AB$

$AB + C \to D$

Enzymatic reactions:

$S \xrightarrow{E}{}^r P$

the "r" is given by Michaelis-Menten (approximated steady-state) laws:

$E + S \leftrightarrow ES$

$ES \to P + E$

The Chemical Ground Form (CGF)

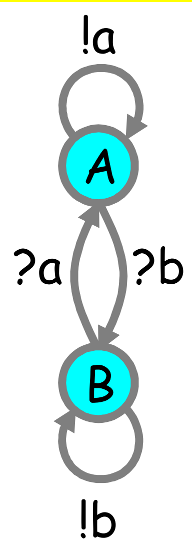
Chemical Ground Form (CGF)

A stochastic subset of CCS
(no values, no restriction)

E ::= 0 ⁝ X=M, E	Reagents
M ::= 0 ⁝ p;P ⊕ M	Molecules
P ::= 0 ⁝ X | P	Solutions
p ::= $\tau_{(r)}$ ⁝ $?a_{(r)}$ ⁝ $!a_{(r)}$	Actions (delay, input, output)
CGF ::= E,P	Reagents plus Initial Conditions

(To translate chemistry to processes we need a bit more than interacting automata: we may have "+" on the right of →, that is we may need "|" after p.)

⊕ is stochastic choice (vs. + for chemical reactions)
0 is the null solution (P|0 = 0|P = P)
and null molecule (M⊕0 = 0⊕M = M)
Each X in E is a distinct *species*
Each name a is assigned a fixed rate r: $a_{(r)}$

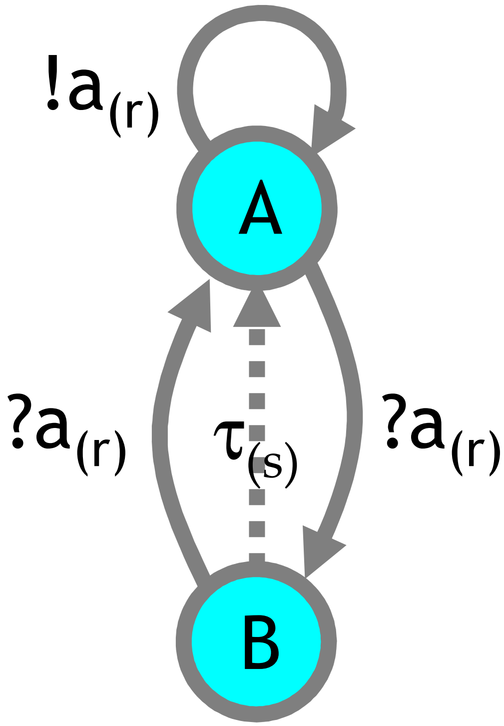


Ex: Interacting Automata
(= finite-control CGFs: they use "|" only in initial conditions):

A = !a;A ⊕ ?b;B — Automaton in state A
B = !b;B ⊕ ?a;A — Automaton in state B
A|A|B|B — Initial conditions: 2A and 2B

From CGF to Chemistry

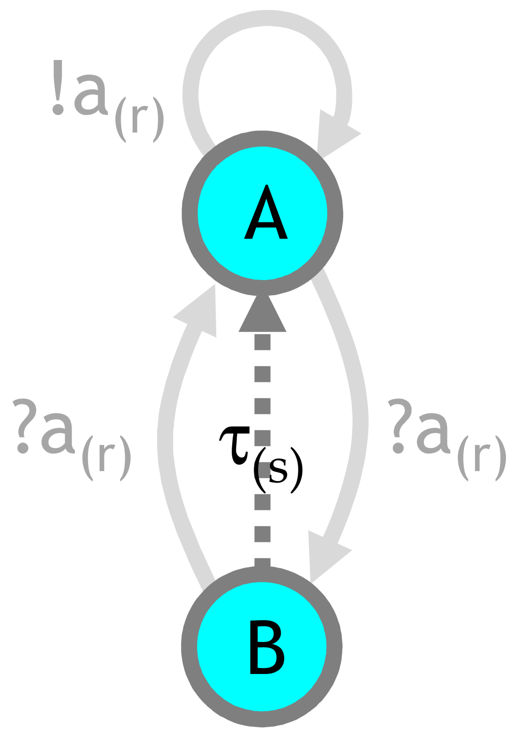
From CGF to Chemistry (by example)

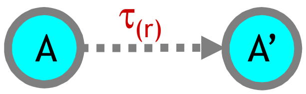


$A = !a_{(r)};A \oplus ?a_{(r)};B$

$B = ?a_{(r)};A \oplus \tau_{(s)};A$

From CGF to Chemistry (by example)





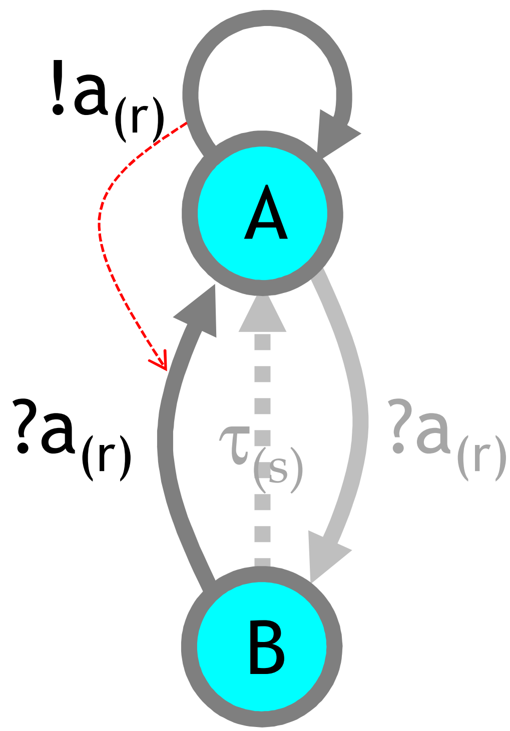
$A \rightarrow^{r} A'$

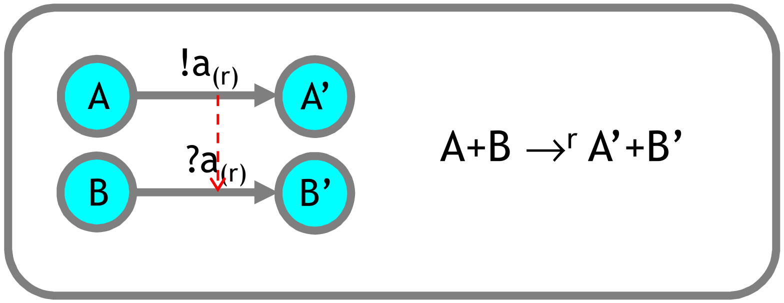
$B \rightarrow^{s} A$

$A = !a;A \oplus ?a;B$

$B = ?a;A \oplus \tau_{(s)};A$

From CGF to Chemistry (by example)





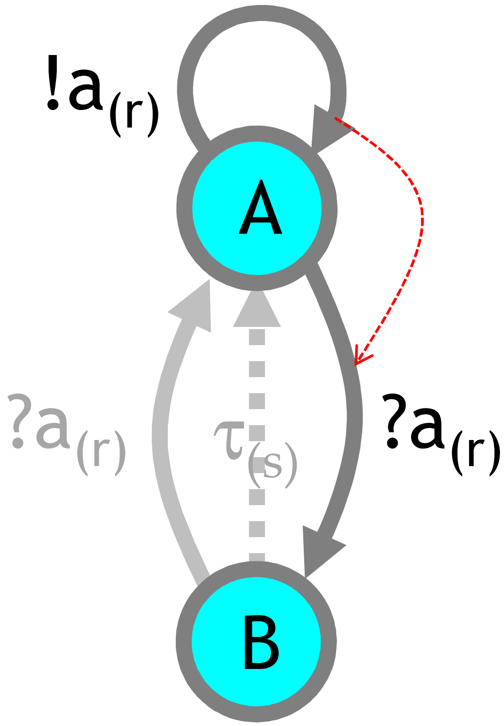
$B \rightarrow^{s} A$

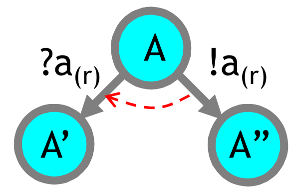
$A+B \rightarrow^{r} A+A$

$A = !a;A \oplus ?a;B$

$B = ?a;A \oplus \tau_{(s)};A$

From CGF to Chemistry (by example)





$A+A \rightarrow^{2r} A'+A''$

$B \rightarrow^{s} A$

$A+B \rightarrow^{r} A+A$

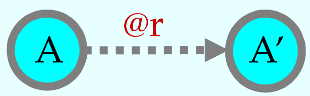
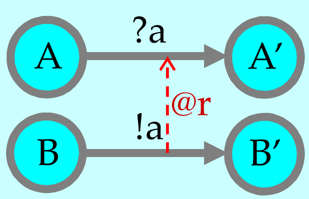
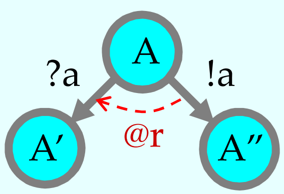
$A+A \rightarrow^{2r} A+B$

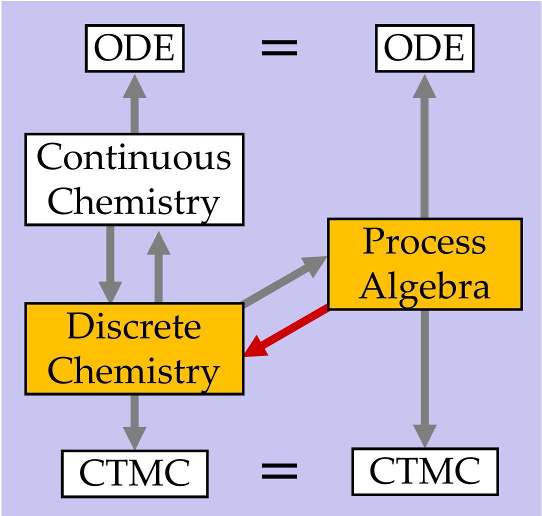
Double rate for homeo reactions

$A = !a;A \oplus ?a;B$

$B = ?a;A \oplus \tau_{(s)};A$

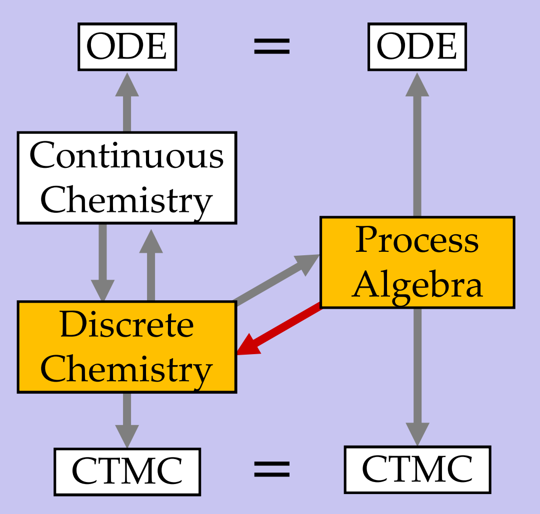
From CGF to Chemistry (by example)

Interacting Automata →	Discrete Chemistry
initial states A | A | … | A	initial quantities $\#A_0$
	$A \xrightarrow{r} A'$
	$A+B \xrightarrow{r} A'+B'$
	$A+A \xrightarrow{2r} A'+A''$



From CGF to Chemistry: Ch(E)

E ::= 0 ∶ X=M, E	Reagents
M ::= 0 ∶ p;P ⊕ M	Molecules
P ::= 0 ∶ X | P	Solutions
p ::= $\tau_{(r)}$ ∶ $?a_{(r)}$ ∶ $!a_{(r)}$	Interactions
CGF ::= E,P	Reagents plus Initial Conditions



Chemical reactions for E,P: (N.B.: <...> are reaction tags to obtain multiplicity of reactions, and P is P with all the | changed to +)

Ch(E) :=

{(<{X.i}>: X $\to^{r}$ P) *s.t.* E.X.i = $\tau_{(r)}$;P} ∪

{(<{X.i,Y.j}>: X + Y $\to^{r}$ P + Q) *s.t.* X≠Y, E.X.i = $?a_{(r)}$;P, E.Y.j = $!a_{(r)}$;Q} ∪

{(<{X.i,X.j}>: X + X $\to^{2r}$ P + Q) *s.t.* E.X.i = $?a_{(r)}$;P, E.X.j = $!a_{(r)}$;Q}

Initial conditions for P:

Ch(P) := P

From Discrete to Continuous Chemistry

The “Type System” of Chemistry

The International System of Units (SI) defines the following physical units, with related derived units and constants; note that *amount of substance* is a base unit in SI, like length and time:

mol	(a base unit)	mole, unit of *amount of substance*
m	(a base unit)	meter, unit of *length*
s	(a base unit)	second, unit of *time*
$L = 0.001 \cdot m^3$		liter (volume)
$M = mol \cdot L^{-1}$		molarity (concentration of substance)
$N_A : mol^{-1} \cong 6.022 \times 10^{23}$		Avogadro’s number (number of particles per amount of substance)

For a substance $X{:}mol$, we write $[X]{:}M$ for the concentration of X, and $[X]^{\bullet}{:}M{\cdot}s^{-1}$ for the time derivative of the concentration.

A continuous chemical system (C,V) is a system of chemical reactions C plus a vector of initial concentrations V_X: M, one for each species X.

The rates of unary reactions have dimension s^{-1}.

The rates of binary reactions have dimension $M^{-1}s^{-1}$.

(because in both cases the rhs of an ODE should have dimension $M{\cdot}s^{-1}$).

For a given volume of solution V, the volumetric factor γ of dimension M^{-1} is:

$\gamma : M^{-1} = N_A V$ where $N_A{:}mol^{-1}$ and V:L

$\#X / \gamma : M =$ concentration of X molecules

$\gamma{\cdot}[X] : 1 =$ total number of X molecules (rounded to an integer).

The Gillespie Conversion

Discrete Chemistry	Continuous Chemistry $\gamma = N_A V$ $:M^{-1}$
initial quantities $\#A_0$	initial concentrations $[A]_0$ with $[A]_0 = \#A_0/\gamma$
$A \overset{r}{\dashrightarrow} A'$	$A \rightarrow^{k} A'$ with $k = r$ $:s^{-1}$
$A+B \overset{r}{\dashrightarrow} A'+B'$	$A+B \rightarrow^{k} A'+B'$ with $k = r\gamma$ $:M^{-1}s^{-1}$
$A+A \overset{r}{\dashrightarrow} A'+A''$	$A+A \rightarrow^{k} A'+A''$ with $k = r\gamma/2$ $:M^{-1}s^{-1}$

V = interaction volume

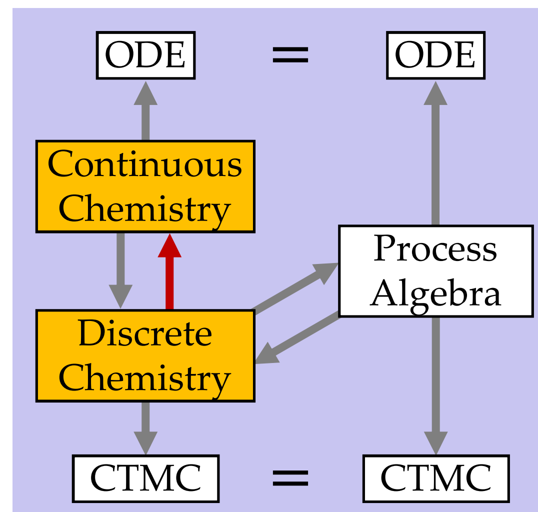
N_A = Avogadro's number

Think $\gamma = 1$

i.e. $V = 1/N_A$

$M = mol \cdot L^{-1}$

molarity (concentration)



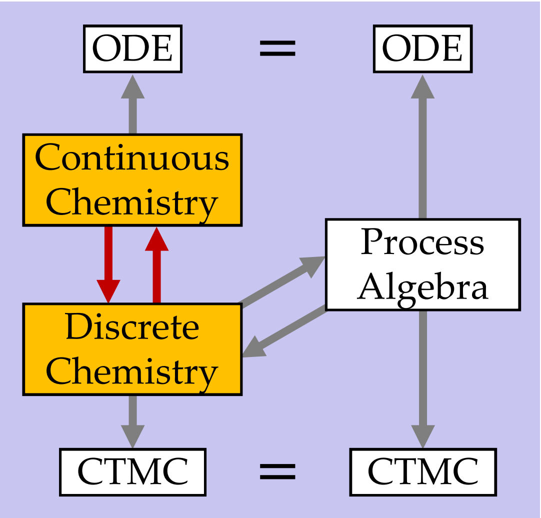
$Cont_\gamma$ and $Disc_\gamma$

4.2–3 Definition: $Cont_\gamma$ and $Disc_\gamma$

For a volumetric factor $\gamma{:}M^{-1}$, we define a translation $Cont_\gamma$ from a discrete chemical systems (C,P), with species X and initial molecule count $\#X_0 = \#X(P)$, to a continuous chemical systems (C,V) with initial concentration $[X]_0 = V_X$. The translation $Disc_\gamma$ is its inverse, up to a rounding error $\lceil\gamma[X]_0\rceil$ in converting concentrations to molecule counts. Since γ is a global conversion constant, we later usually omit it as a subscript.

$Cont_\gamma(X \rightarrow^r P)$	$= X \rightarrow^k P$	with $k = r$,	$r{:}s^{-1}$	$k{:}s^{-1}$
$Cont_\gamma(X+Y \rightarrow^r P)$	$= X+Y \rightarrow^k P$	with $k = r\gamma$	$r{:}s^{-1}$	$k{:}M^{-1}s^{-1}$
$Cont_\gamma(X+X \rightarrow^r P)$	$= X+X \rightarrow^k P$	with $k = r\gamma/2$	$r{:}s^{-1}$	$k{:}M^{-1}s^{-1}$
$Cont_\gamma(\#X_0)$	$= [X]_0$	with $[X]_0 = \#X_0/\gamma$	$X_0{:}mol$	$[X]_0{:}M$
$Disc_\gamma(X \rightarrow^k P)$	$= X \rightarrow^r P$	with $r = k$,	$k{:}s^{-1}$	$r{:}s^{-1}$
$Disc_\gamma(X+Y \rightarrow^k P)$	$= X+Y \rightarrow^r P$	with $r = k/\gamma$	$k{:}M^{-1}s^{-1}$	$r{:}s^{-1}$
$Disc_\gamma(X+X \rightarrow^k P)$	$= X+X \rightarrow^r P$	with $r = 2k/\gamma$	$k{:}M^{-1}s^{-1}$	$r{:}s^{-1}$
$Disc_\gamma([X]_0)$	$= \#X_0$	with $\#X_0 = \lceil\gamma[X]_0\rceil$	$[X]_0{:}M$	$X_0{:}mol$

$$Ch_\gamma := Cont_\gamma \circ Ch$$



From Continuous Chemistry to ODEs

From Reactions to ODEs (Law of Mass Action)

v_1: $A+B \rightarrow_{k_1} C+C$
v_2: $A+C \rightarrow_{k_2} D$
v_3: $C \rightarrow_{k_3} E+F$
v_4: $F+F \rightarrow_{k_4} B$

Write the coefficients by columns

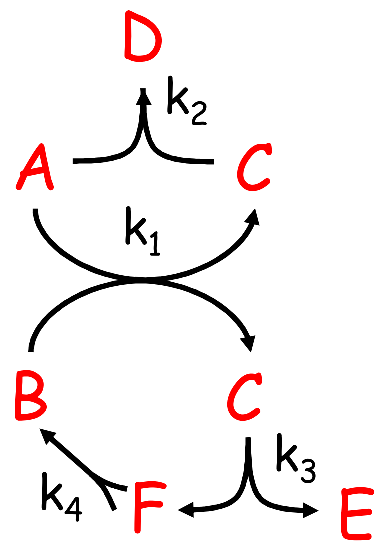
Stoichiometric Matrix

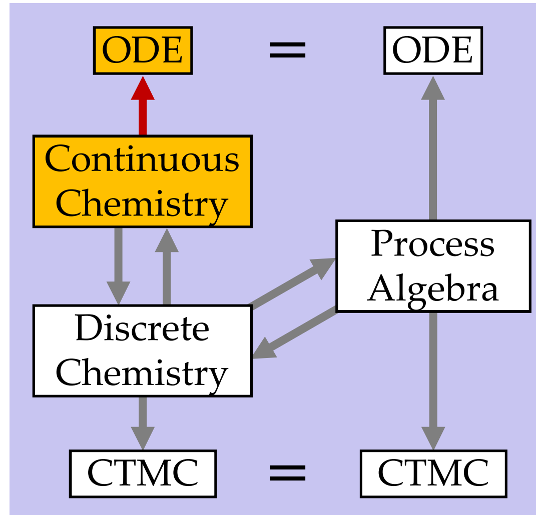
reactions

species

N	v_1	v_2	v_3	v_4
A	-1	-1		
B	-1			1
C	2	-1	-1	
D		1		
E			1	
F			1	-2

X





Quantity changes

Stoichiometric matrix

Rate laws

$d[X]/dt = N \cdot l$

$d[A]/dt = -l_1 - l_2$
$d[B]/dt = -l_1 + l_4$
$d[C]/dt = 2l_1 - l_2 - l_3$
$d[D]/dt = l_2$
$d[E]/dt = l_3$
$d[F]/dt = l_3 - 2l_4$

Read the concentration changes from the rows

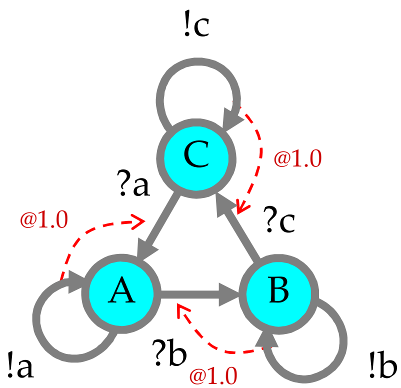
E.g. $d[A]/dt = -k_1[A][B] - k_2[A][C]$

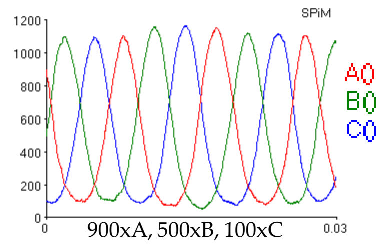
Set a rate law for each reaction
(Degradation/Hetero/Homeo)

	l
l_1	$k_1[A][B]$
l_2	$k_2[A][C]$
l_3	$k_3[C]$
l_4	$k_4[F]^2$

X: chemical species
[-]: quantity of molecules
l: rate laws
k: kinetic parameters
N: stoichiometric matrix

From Processes to ODEs via Chemistry!



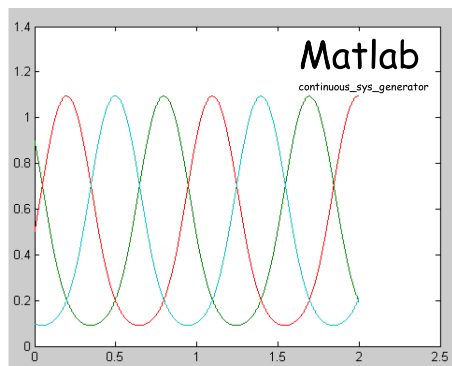



```
directive sample 0.03 1000
directive plot A(); B(); C()

new a@1.0:chan new b@1.0:chan new c@1.0:chan
let A() = do !a;A() or ?b; B()
and B() = do !b;B() or ?c; C()
and C() = do !c;C() or ?a; A()

run (900 of A() | 500 of B() | 100 of C())
```

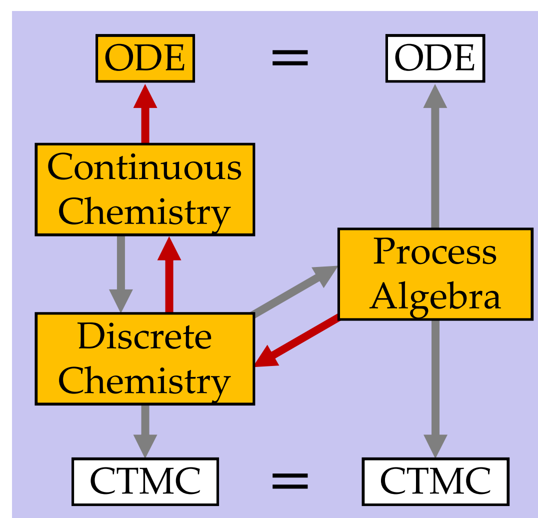
$$A = !a_{(s)};A \oplus ?b_{(s)};B$$
$$B = !b_{(s)};B \oplus ?c_{(s)};C$$
$$C = !c_{(s)};C \oplus ?a_{(s)};A$$

$$A+B \rightarrow^{s} B+B$$
$$B+C \rightarrow^{s} C+C$$
$$C+A \rightarrow^{s} A+A$$

$$d[A]/dt = -s[A][B]+s[C][A]$$
$$d[B]/dt = -s[B][C]+s[A][B]$$
$$d[C]/dt = -s[C][A]+s[B][C]$$




```
interval/step [0:0.001:20.0]
(A)    dx1/dt = - x1*x2 + x3*x1    0.9
(B)    dx2/dt = - x2*x3 + x1*x2    0.5
(C)    dx3/dt = - x3*x1 + x2*x3    0.1
```



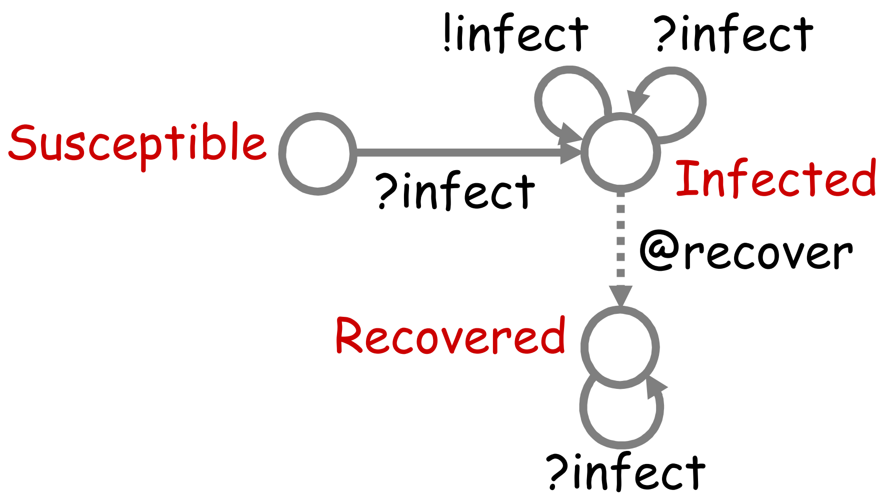
Epidemics

Beyond Chemical Interactions
to Population Interactions

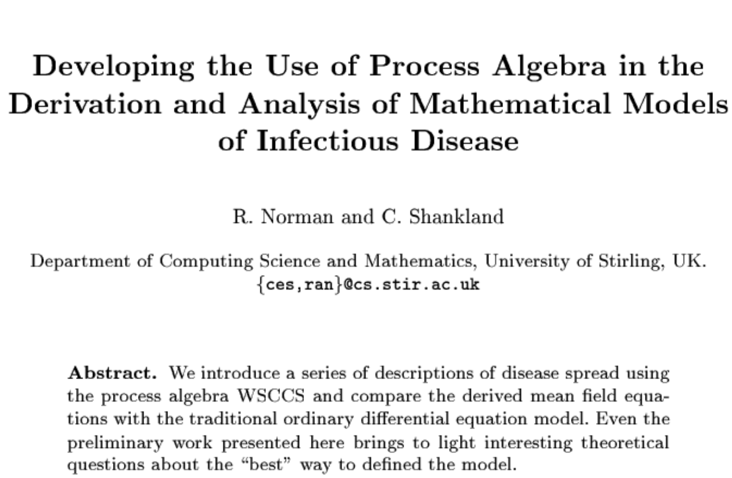
Kermack, W. O. and McKendrick, A. G. "**A Contribution to the Mathematical Theory of Epidemics.**" *Proc. Roy. Soc. Lond. A* **115**, 700-721, 1927.

http://mathworld.wolfram.com/Kermack-McKendrickModel.html

Epidemics




```
directive sample 500.0 1000
directive plot Recovered(); Susceptible(); Infected()

new infect @0.001:chan()
val recover = 0.03

let Recovered() =
  ?infect; Recovered()

and Susceptible() =
  ?infect; Infected()

and Infected() =
  do !infect; Infected()
  or ?infect; Infected()
  or delay@recover; Recovered()

run (200 of Susceptible() | 2 of Infected())
```

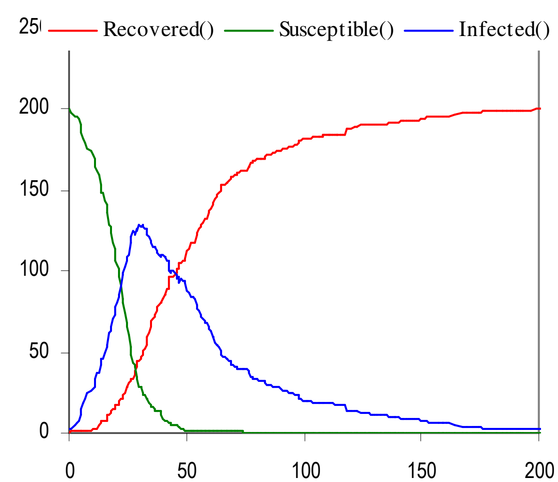


Developing the Use of Process Algebra in the Derivation and Analysis of Mathematical Models of Infectious Disease

R. Norman and C. Shankland

Department of Computing Science and Mathematics, University of Stirling, UK.
{ces,ran}@cs.stir.ac.uk

Abstract. We introduce a series of descriptions of disease spread using the process algebra WSCCS and compare the derived mean field equations with the traditional ordinary differential equation model. Even the preliminary work presented here brings to light interesting theoretical questions about the "best" way to defined the model.



ODE

$S = ?i_{(t)};I$

$I = !i_{(t)};I \oplus ?i_{(t)};I \oplus \tau_r;R$

$R = ?i_{(t)};R$

Differentiating Processes!

$S + I \rightarrow^{t\gamma} I + I$

$I + I \rightarrow^{t\gamma} I + I$

$I \rightarrow^{r} R$

$R + I \rightarrow^{t\gamma} R + I$

"useless" reactions

$[S]^\bullet = -t\gamma[S][I]$

$[I]^\bullet = t\gamma[S][I]-r[I]$

$[R]^\bullet = r[I]$

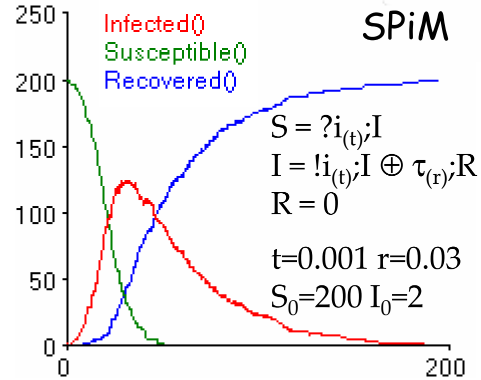
Automata produce the standard ODEs!

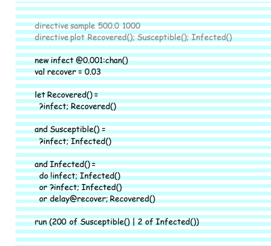
$$\frac{dS}{dt} = -aIS$$

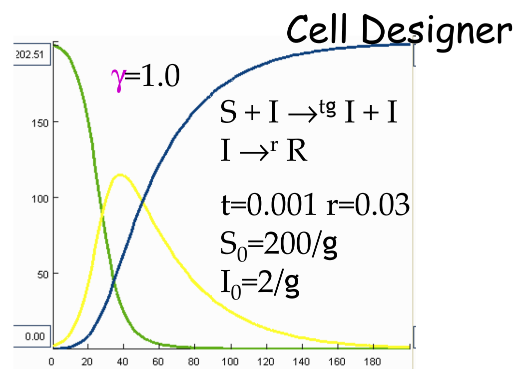
$$\frac{dI}{dt} = aIS - bI$$

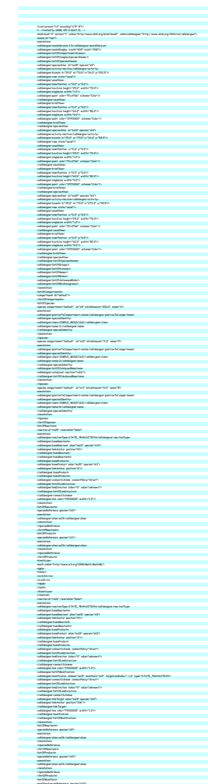
$$\frac{dR}{dt} = bI$$

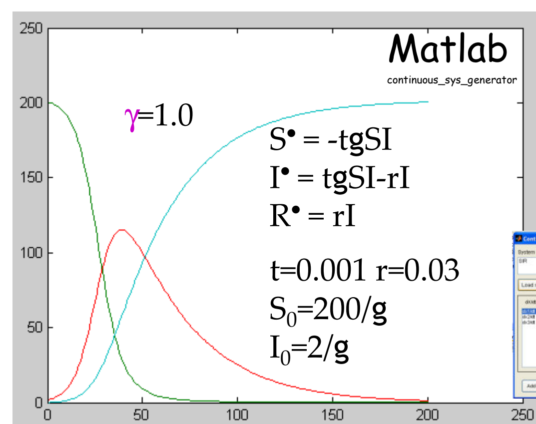
(the Kermack-McKendrick, or SIR model)

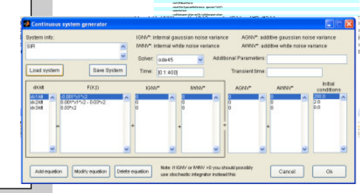




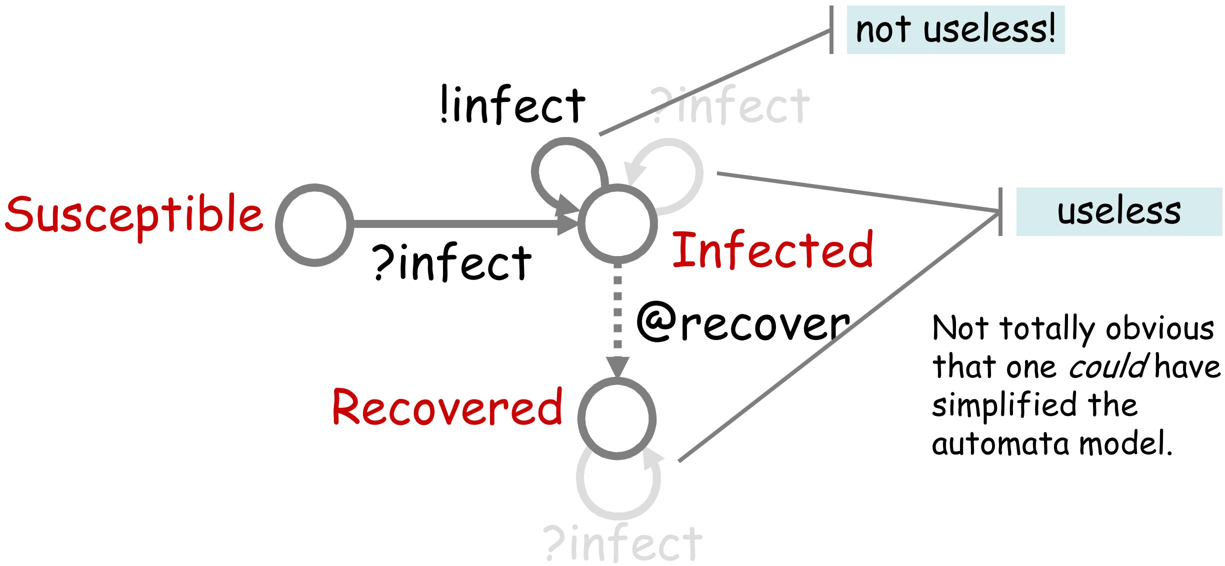








Simplified Model



$S = ?i_{(t)};I$

$I = !i_{(t)};I \oplus \tau_r;R$

$R = 0$

$S + I \rightarrow^{t\gamma} I + I$

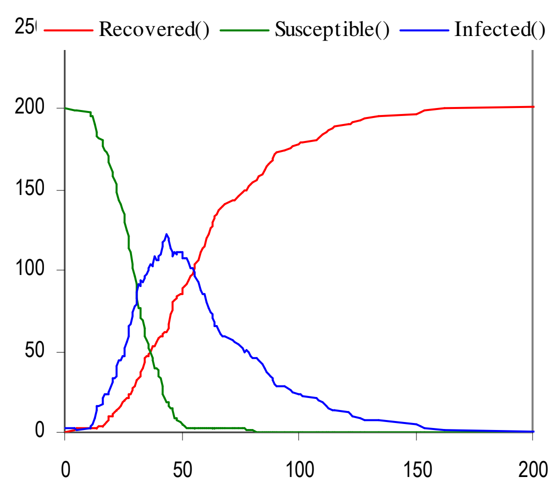
$I \rightarrow^{r} R$

$[S]^\bullet = -t\gamma[S][I]$

$[I]^\bullet = t\gamma[S][I]-r[I]$

$[R]^\bullet = r[I]$

```
directive sample 500.0 1000
directive plot Recovered(); Susceptible(); Infected()

new infect @0.001:chan()
val recover = 0.03

let Recovered() =
  ()

and Susceptible() =
  ?infect; Infected()

and Infected() =
  do !infect; Infected()
  or delay@recover; Recovered()

run (200 of Susceptible() | 2 of Infected())
```

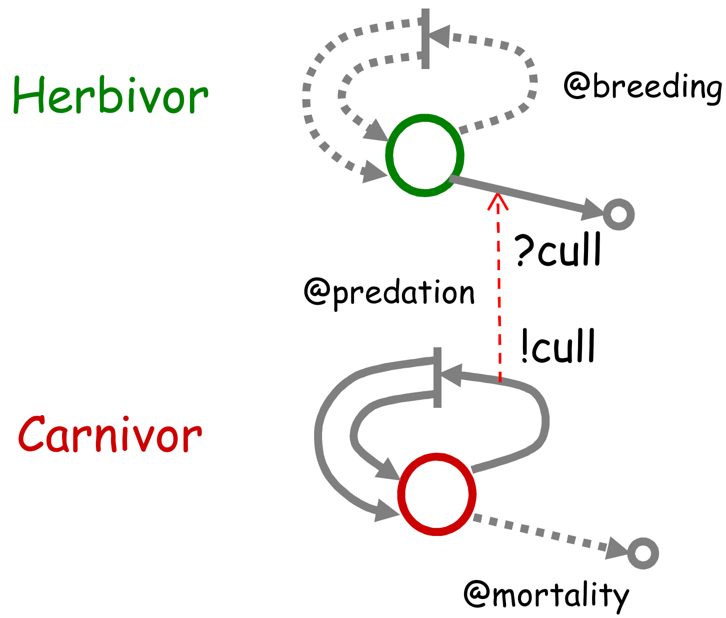


Same ODE, hence equivalent automata models.

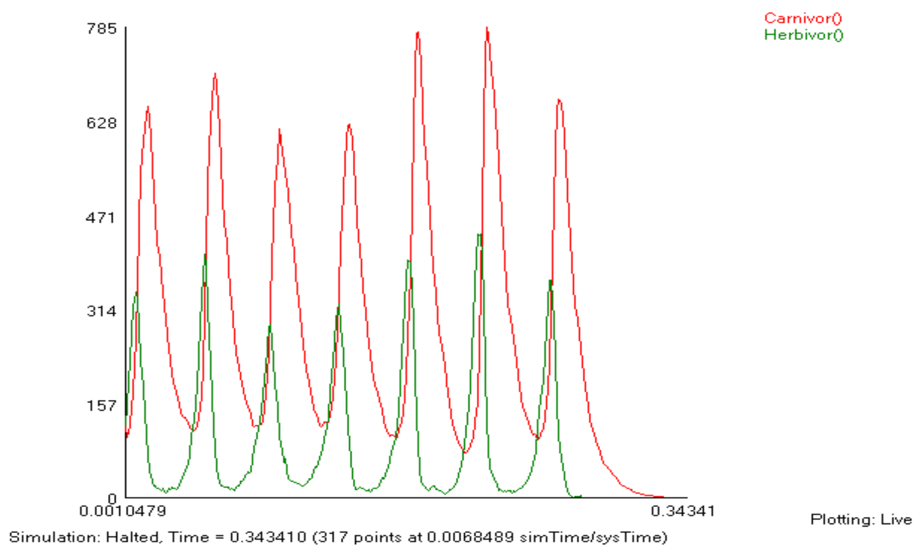
Lotka-Volterra

Beyond Simple Automata
to Unbounded State

Predator-Prey




```
directive sample 1.0 1000
directive plot Carnivor(); Herbivor()

val mortality = 100.0
val breeding = 300.0
val predation = 1.0
new cull @predation:chan()

let Herbivor() =
  do delay@breeding; (Herbivor() | Herbivor())
  or ?cull; ()

and Carnivor() =
  do delay@mortality; ()
  or !cull; (Carnivor() | Carnivor())

run 100 of Herbivor()
run 100 of Carnivor()
```



An unbounded state system!

Lotka-Volterra in Matlab

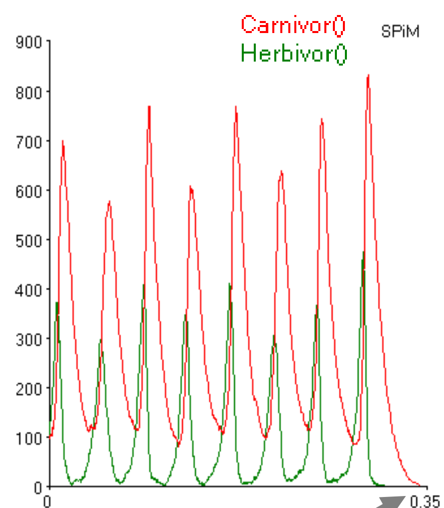
$H = \tau_b; (H|H) \oplus ?c_{(p)};0$
$C = \tau_m;0 \oplus !c_{(p)};(C|C)$
$\#H_0, \#C_0$

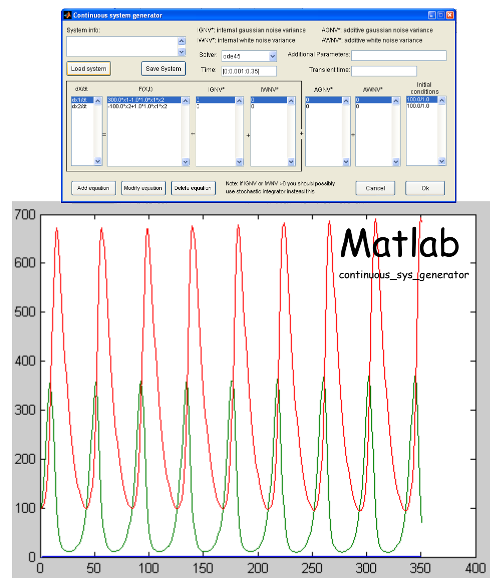
$H \to^b H + H$
$C \to^m 0$
$H + C \to^{pg} C + C$
$[H]_0 = \#H_0/g$
$[C]_0 = \#C_0/g$

$[H]^\bullet = b[H]-pg[H][C]$
$[C]^\bullet = -m[C]+pg[H][C]$
$[H]_0 = \#H_0/g$
$[C]_0 = \#C_0/g$

$m=100.0$
$b=300.0$
$p=1.0$
$g=1.0$
$\#H_0 = 100$
$\#C_0 = 100$

```
directive sample 0.35 1000
directive plot Carnivor(); Herbivor()

val mortality = 100.0
val breeding = 300.0
val predation = 1.0
new cull @predation:chan()

let Herbivor() =
  do delay@breeding; (Herbivor() | Herbivor())
  or ?cull; ()

and Carnivor() =
  do delay@mortality; ()
  or !cull; (Carnivor() | Carnivor())

run 100 of Herbivor()
run 100 of Carnivor()
```





Extinction

No extinction

Which one is the "right prediction"?

The Chemical Parametric Form (CPF)

Chemical Parametric Form (CPF)

E	$::= X_1(\mathbf{p}_1)=M_1, \ldots, X_n(\mathbf{p}_n)=M_n$	Reagents	$(n \geq 0)$
M	$::= p_1;P_1 \oplus \ldots \oplus p_n;P_n$	Molecules	$(n \geq 0)$
P	$::= X_1(\mathbf{p}_1) \mid \ldots \mid X_n(\mathbf{p}_n)$	Solutions	$(n \geq 0)$
p	$::= \tau_r \quad ?n(\mathbf{p}) \quad !n(\mathbf{p})$	Interactions	
CPF	$::= E,P$	with initial conditions	

$\oplus$ is stochastic choice (vs. + for chemical reactions)
0 is the null solution ($P|0 = 0|P = P$)
and null molecule ($M\oplus 0 = 0\oplus M = M$) ($t_0;P = 0$)
X_i are distinct in E, **p** are vectors of names
p are vectors of distinct names when in binding position
Each free name n in E is assigned a fixed rate r:
written either $n_{(r)}$, or $r_{CPF}(n)=r$.

Not bounded-state systems.
Not finite-control systems.
But still finite-species systems.

A translation from CPF to CGF exists
(expanding all possible instantiation of parameters from the initial conditions)

An incremental translation algorithm exists
(expanding on demand from initial conditions)

From CPF to CGF

CPF to CGF: Handling Parameters

Consider first the CPF subset with no communication (pure ?n, !n).

$E ::= X_1(\mathbf{p}_1)=M_1, \ldots, X_n(\mathbf{p}_n)=M_n$
$M ::= p_1;P_1 \oplus \ldots \oplus p_n;P_n$
$P ::= X_1(\mathbf{p}_1) \mid \ldots \mid X_n(\mathbf{p}_n)$
$p ::= \tau_r \quad ?n \quad !n$

Grounding (replace parameters with constants)

where X/**p** is a name in bijection with <X,**p**>
(each X/**p** is seen as a separate *species*)

$$/(p_1;P_1 \oplus \ldots \oplus p_n;P_n) =_{def} p_1;/(P_1) \oplus \ldots \oplus p_n;/(P_n)$$
$$/(X_1(\mathbf{p}_1) \mid \ldots \mid X_n(\mathbf{p}_n)) =_{def} X_1/\mathbf{p}_1 \mid \ldots \mid X_n/\mathbf{p}_n$$

Let N be the set of free names occurring in E.

E_G is the **Parametric Explosion** of E (still a finite species system)
computed by replacing parameters with all combinations of free names in E

$$E_G := \{(X/\mathbf{q} = /(M\{\mathbf{p}\leftarrow\mathbf{q}\}))\ \text{s.t.}\ (X(\mathbf{p}) = M) \in E \text{ and } \mathbf{q} \in N^{\#\mathbf{p}}\}$$
$$P_G := /P \qquad \text{(simply ground the given initial conditions once)}$$

E_G is a CGF! To obtain the chemical reactions $Cp_\gamma(E)$, just compute $Ch_\gamma(E_G)$

$$Cp_\gamma(E,P) = Ch_\gamma(E_G,P_G) \qquad \text{the chemical system of a CPF}$$

CPF to CGF: Handling Communication

Grounding (replace parameters with constants)

just one main change: now also convert each input parameter into a ground choice of all possible inputs

N is the set of free names in E,P

#**p** is the length of **p**

n/**p** is a name in bijection with <n,**p**>

X/**p** is a name in bijection with <X,**p**>

(each X/**p** is seen as a separate *species*)

$E ::= X_1(\mathbf{p}_1)=M_1, \ldots, X_n(\mathbf{p}_n)=M_n$

$M ::= p_1;P_1 \oplus \ldots \oplus p_n;P_n$

$P ::= X_1(\mathbf{p}_1) \mid \ldots \mid X_n(\mathbf{p}_n)$

$p ::= \tau_r \quad ?n(\mathbf{p}) \quad !n(\mathbf{p})$

$/_N(\tau_r;P) = \tau_r;\ /_N(P)$

$/_N(!n_{(r)}(\mathbf{p});P) = !n/\mathbf{p}_{(r)};\ /_N(P)$

$/_N(?n_{(r)}(\mathbf{p});P) = \oplus(\mathbf{q}\in N^{\#\mathbf{p}}) \text{ of } ?n/\mathbf{q}_{(r)};\ /_N(P\{\mathbf{p}\leftarrow\mathbf{q}\})$

$/_N(p_1;P_1 \oplus \ldots \oplus p_n;P_n) = /_N(p_1;P_1) \oplus \ldots \oplus /_N(p_n;P_n)$

$/_N(X_1(\mathbf{p}_1) \mid \ldots \mid X_n(\mathbf{p}_n)) = X_1/\mathbf{p}_1 \mid \ldots \mid X_n/\mathbf{p}_n$

E_G is again the **Parametric Explosion** of E

$E_G := \{(X/\mathbf{q} = /_N(M\{\mathbf{p}\neg\mathbf{q}\}))\ \text{s.t.}\ (X(\mathbf{p}) = M) \in E \text{ and } \mathbf{q} \in N^{\#\mathbf{p}}\}$

$P_G := /_N(P)$ (simply ground the given initial conditions once)

E_G is a again a CGF!

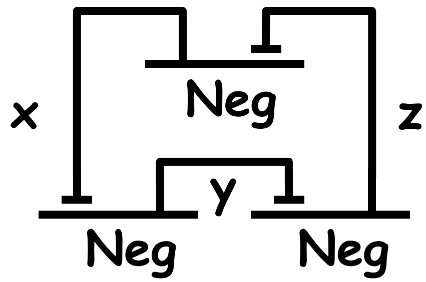
$Cp_\gamma(E,P) = Ch_\gamma(E_G,P_G)$ the chemical system of a CPF

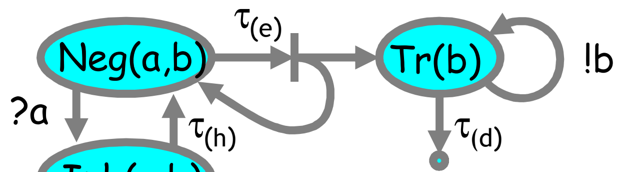
CPF Example:
Gene Networks

And Yet It Moves

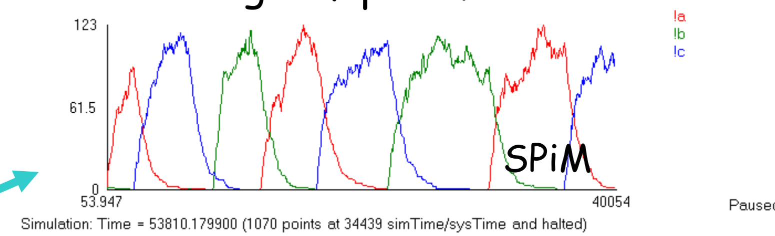
R.Blossey, L.Cardelli, A.Phillips: **Compositionality, Stochasticity and Cooperativity in Dynamic Models of Gene Regulation** (HFSP Journal)

The Repressilator





A fine stochastic oscillator over a wide range of parameters.




```
directive sample 50000.0 1000
directive plot !a; !b; !c

val dk = 0.0[illegible]    (* Decay rate *)
val in[illegible] 0.001    (* Inhibition rate *)
val cs[illegible] = 0.[illegible]    ([illegible] Constitutive rate *)

let [illegible](p:c[illegible]an()) [illegible] do !p; tr(p) or delay@dk

let neg(a:chan(), [illegible]b:chan()) =
 do ?a; del[illegible]y@i[illegible]h; neg(a,b)
 or de[illegible]y@cst; ([illegible]r(b) | neg(a,b))

val b[illegible]d = 1.0    ([illegible]* Protein binding rate *)
new [illegible]@bnd:chan[illegible]) new b@bnd:chan() new c@b[illegible]d:chan()
run (neg(c,a) | n[illegible]g(a,b) | neg(b,c))
```

Parametric representation

$\mathrm{Neg}(a,b) = ?a;\ \mathrm{Inh}(a,b) \oplus \tau_e;\ (\mathrm{Tr}(b) \mid \mathrm{Neg}(a,b))$

$\mathrm{Inh}(a,b) = \tau_h;\ \mathrm{Neg}(a,b)$

$\mathrm{Tr}(b) = !b;\ \mathrm{Tr}(b) \oplus \tau_g;\ 0$

$\mathrm{Neg}(x_{(r)},y_{(r)}) \mid \mathrm{Neg}(y_{(r)},z_{(r)}) \mid \mathrm{Neg}(z_{(r)},x_{(r)})$

$Neg/x,y \to^{e} Tr/y + Neg/x,y$

$Neg/y,z \to^{e} Tr/z + Neg/y,z$

$Neg/z,x \to^{e} Tr/x + Neg/z,x$

$Tr/x + Neg/x,y \to^{r} Tr/x + Inh/x,y$

$Tr/y + Neg/y,z \to^{r} Tr/y + Inh/y,z$

$Tr/z + Neg/z,x \to^{r} Tr/z + Inh/z,x$

$Inh/x,y \to^{h} Neg/x,y$

$Inh/y,z \to^{h} Neg/y,z$

$Inh/z,x \to^{h} Neg/z,x$

$Tr/x \to^{g} 0$

$Tr/y \to^{g} 0$

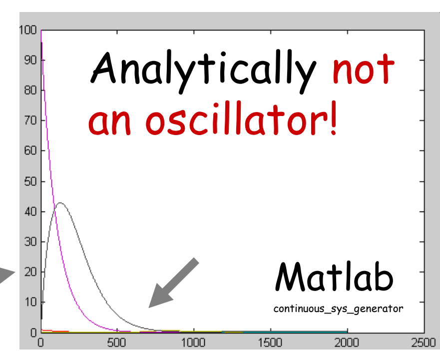
$Tr/z \to^{g} 0$

$Neg/x,y + Neg/y,z + Neg/z,x$

d[Neg/x,y]/dt = -r[Tr/x][Neg/x,y] + h[Inh/x,y]
d[Neg/y,z]/dt = -r[Tr/y][Neg/y,z] + h[Inh/y,z]
d[Neg/z,x]/dt = -r[Tr/z][Neg/z,x] + h[Inh/z,x]
d[Inh/x,y]/dt = r[Tr/x][Neg/x,y] - h[Inh/x,y]
d[Inh/y,z]/dt = r[Tr/y][Neg/y,z] - h[Inh/y,z]
d[Inh/z,x]/dt = r[Tr/z][Neg/z,x] - h[Inh/z,x]
d[Tr/x]/dt = e[Neg/z,x] - g[Tr/x]
d[Tr/y]/dt = e[Neg/x,y] - g[Tr/y]
d[Tr/z]/dt = e[Neg/y,z] - g[Tr/z]

simplifying (N is the quantity of each of the 3 gates)

d[Neg/x,y]/dt = hN – (h+r[Tr/x])[Neg/x,y]
d[Neg/y,z]/dt = hN – (h+r[Tr/y])[Neg/y,z]
d[Neg/z,x]/dt = hN – (h+r[Tr/z])[Neg/z,x]
d[Tr/x]/dt = e[Neg/z,x] - g[Tr/x]
d[Tr/y]/dt = e[Neg/x,y] - g[Tr/y]
d[Tr/z]/dt = e[Neg/y,z] - g[Tr/z]



interval/step [0:10:20000]   N=1, r=1.0, e=0.1, h=0.001, g=0.001

(Neg/x,y)	dx1/dt = 0.001 – (0.001 + x4)*x1	1.0
(Neg/x,y)	dx2/dt = 0.001 - (0.001 + x5)*x2	1.0
(Neg/x,y)	dx3/dt = 0.001 - (0.001 + x6)*x3	1.0
(Tr/x)	dx4/dt = 0.1*x3 - 0.001*x4	100.0
(Tr/y)	dx5/dt = 0.1*x1 - 0.001*x5	0
(Tr/z)	dx6/dt = 0.1*x2 - 0.001*x6	0

CPF to CGF Iterative Algorithm. Ex: Neg(x,x)

E =

$\text{Neg}(a,b) = ?a;\ \text{Inh}(a,b) \oplus t_e;\ (\text{Tr}(b) \mid \text{Neg}(a,b))$
$\text{Inh}(a,b) = t_h;\ \text{Neg}(a,b)$
$\text{Tr}(b) = !b;\ \text{Tr}(b) \oplus t_d;\ 0$
$\text{Neg}(x,x)$

----- initialization -----

$E_c := \{\text{Neg}/x,x = ?x;\ \text{Inh}/x,x \oplus t_e;\ (\text{Tr}/x \mid \text{Neg}/x,x)\}$

----- iteration 1 -----

$C := \{\text{Neg}/x,x \rightarrow^e \text{Tr}/x + \text{Neg}/x,x\ \}$

$E_c := \{\text{Neg}/x,x = ?x;\ \text{Inh}/x,x \oplus t_e;\ (\text{Tr}/x \mid \text{Neg}/x,x)$
$\quad \text{Tr}/x = !x;\ \text{Tr}/x \oplus t_d;\ 0\}$

----- iteration 2 -----

$C := \{\ \text{Neg}/x,x \rightarrow^e \text{Tr}/x + \text{Neg}/x,x$
$\quad \text{Tr}/x \rightarrow^d 0$
$\quad \text{Tr}/x + \text{Neg}/x,x \rightarrow^{r(x)} \text{Tr}/x + \text{Inh}/x,x\ \}$

$E_c := \{\text{Neg}/x,x = ?x;\ \text{Inh}/x,x \oplus t_e;\ (\text{Tr}/x \mid \text{Neg}/x,x)$
$\quad \text{Tr}/x = !x;\ \text{Tr}/x \oplus t_d;\ 0$
$\quad \text{Inh}/x,x = t_h;\ \text{Neg}/x,x\}$

----- iteration 3 -----

$C := \{\ \text{Neg}/x,x \rightarrow^e \text{Tr}/x + \text{Neg}/x,x$
$\quad \text{Tr}/x \rightarrow^d 0$
$\quad \text{Tr}/x + \text{Neg}/x,x \rightarrow^{r(x)} \text{Tr}/x + \text{Inh}/x,x$
$\quad \text{Inh}/x,x \rightarrow^h \text{Neg}/x,x\}$

$E_c :=$ no change

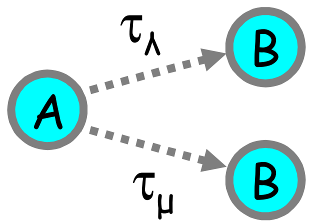
----- termination -----

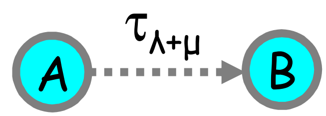
$\text{Neg}/x,x \rightarrow^e \text{Tr}/x + \text{Neg}/x,x$
$\text{Tr}/x \rightarrow^d 0$
$\text{Tr}/x + \text{Neg}/x,x \rightarrow^{r(x)} \text{Tr}/x + \text{Inh}/x,x$
$\text{Inh}/x,x \rightarrow^h \text{Neg}/x,x$
$\text{Neg}/x,x$

Laws by ODEs

Choice Law by ODEs

$$\tau_\lambda;B \oplus \tau_\mu;B = \tau_{\lambda+\mu};B$$





$$A = \tau_\lambda;B \oplus \tau_\mu;B$$

↓

$$A \rightarrow^{\lambda} B$$
$$A \rightarrow^{\mu} B$$

↓

$$[A]^\bullet = -\lambda[A] - \mu[A]$$
$$[B]^\bullet = \lambda[A] + \mu[A]$$

=

$$A = \tau_{\lambda+\mu};B$$

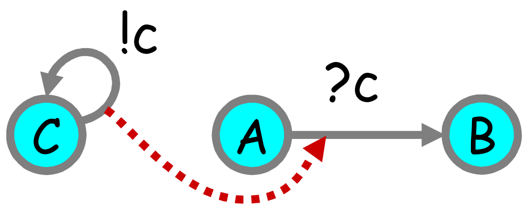
↓

$$A \rightarrow^{\lambda+\mu} B$$

↓

$$[A]^\bullet = -(\lambda+\mu)[A]$$
$$[B]^\bullet = (\lambda+\mu)[A]$$

Idle Interaction Law by ODEs



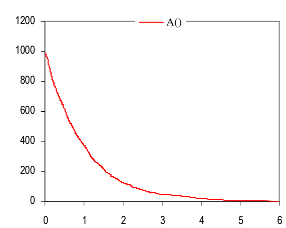
$A = ?c;B$

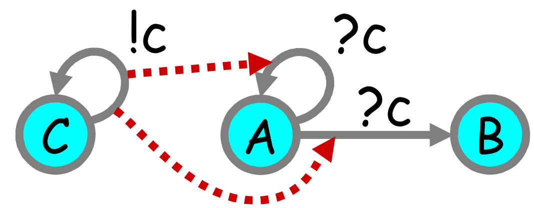
$C = !c;C$

$A+C \rightarrow^{r} B+C$

```
directive sample 6.0 1000
directive plot A()

new c@1.0:chan

let A() = ?c; B()
and B() = ()
and C() = !c; C()

run (C() | 1000 of A())
```

$[A]^{\bullet} = -r[A][C]$
$[B]^{\bullet} = r[A][C]$
$[C]^{\bullet} = 0$





It may seem like A should decrease half as fast, but NO! Two ways to explain:

- State A is *memoryless* of any past idling.
- Activity on c is double

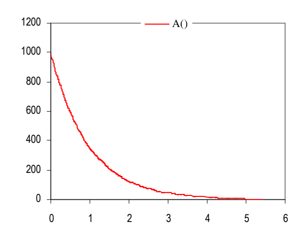
$A = ?c;A \oplus ?c;B$

$C = !c;C$

$A+C \rightarrow^{r} A+C$
$A+C \rightarrow^{r} B+C$

```
directive sample 6.0 1000
directive plot A()

new c@1.0:chan

let A() = do ?c; B() or ?c; A()
and B() = ()
and C() = !c; C()

run (C() | 1000 of A())
```

$[A]^{\bullet} = -r[A][C]$
$[B]^{\bullet} = r[A][C]$
$[C]^{\bullet} = 0$

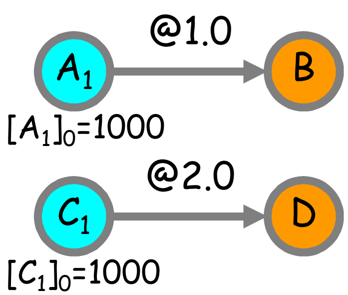


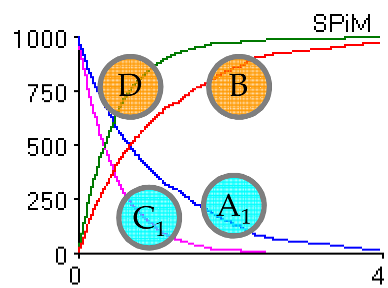
Stochastic Interleaving

Hermanns: Interactive Markov Chains. Sec 4.1.2

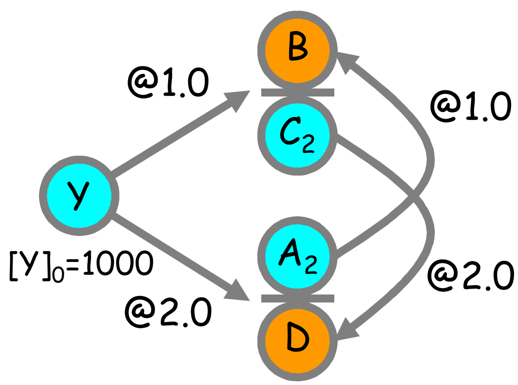
$$\tau_\lambda;B \mid \tau_\mu;D \;=\; \tau_\lambda;(B \mid \tau_\mu;D) \oplus \tau_\mu;(\tau_\lambda;B \mid D)$$

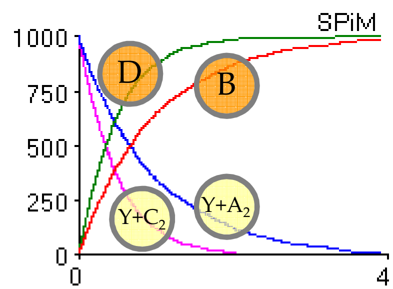
Ex: Λ=1.0, μ=2.0






```
directive sample 4.0 10000
directive plot A(); B(); C(); D()

let A() = delay@1.0; B()
and B() = ()

let C() = delay@2.0; D()
and D() = ()

run 1000 of (A() | C())
```






```
directive sample 4.0 10000
directive plot
   ?YA; B(); ?YC; D(); Y(); A(); C()
new YA@1.0:chan new YC@1.0:chan

let A() = do delay@1.0; B() or ?YA
and B() = ()

let C() = do delay@2.0; D() or ?YC
and D() = ()

let Y() =
   do delay@1.0; (B() | C())
   or delay@2.0; (A() | D())
   or ?YA or ?YC

run 1000 of Y()
```

Amazingly, the B's and the D's from the two branches sum up to exponential distributions

Stochastic Interleaving Law by ODEs

$\tau_\lambda;B \mid \tau_\mu;D \;=\; \tau_\lambda;(B \mid \tau_\mu;D) \oplus \tau_\mu;(\tau_\lambda;B \mid D)$

Want to show that B and D on both sides have the "same behavior" (equal quantities of B and D produced at all times)

$A_1 = \tau_\lambda;B$
$C_1 = \tau_\mu;D$
$n x A_1 \mid n x C_1$

$A_1 \to^\lambda B$
$C_1 \to^\mu D$
$[A_1]_0 = [C_1]_0 = n/\gamma$

$[A_1]^\bullet = -\lambda[A_1]$
$[B]^\bullet = \lambda[A_1]$
$[C_1]^\bullet = -\mu[C_1]$
$[D]^\bullet = \mu[C_1]$

=?

$Y = \tau_\lambda;(B \mid C_2) \oplus \tau_\mu;(A_2 \mid D)$
$C_2 = \tau_\mu;D$
$A_2 = \tau_\lambda;B$
$n x Y$

$Y \to^\lambda B + C_2$
$Y \to^\mu A_2 + D$
$C_2 \to^\mu D$
$A_2 \to^\lambda B$
$[Y]_0 = n/\gamma$

$[Y]^\bullet = -\lambda[Y]-\mu[Y]$
$[A_2]^\bullet = \mu[Y]-\lambda[A_2]$
$[B]^\bullet = \lambda[Y]+\lambda[A_2]$
$[C_2]^\bullet = \lambda[Y]-\mu[C_2]$
$[D]^\bullet = \mu[Y]+\mu[C_2]$

0

$[Y+A_2]^\bullet = -\lambda[Y+A_2]$
$[B]^\bullet = \lambda[Y+A_2]$
$[Y+C_2]^\bullet = -\mu[Y+C_2]$
$[D]^\bullet = \mu[Y+C_2]$

$[Y+A_2]^\bullet = [Y]^\bullet+[A_2]^\bullet$
$= -\lambda[Y]-\mu[Y]+\mu[Y]-\lambda[A_2]$
$= -\lambda[Y]-\lambda[A_2]$
$= -\lambda[Y+A_2]$ — $[Y+A_2]$ decays exponentially!

$[B]$ and $[D]$ have equal time evolutions on the two sides provided that $[A_1]=[Y+A_2]$ and $[C_1]=[Y+C_2]$.

Moreover we have $[A_1]_0 = [C_1]_0 = [Y]_0 = n/\gamma$ and $[A_2]_0 = [C_2]_0 = 0$ since only Y is present on the r.h.s., so that $[A_1]_0 = [Y+A_2]_0$ and $[C_1]_0 = [Y+C_2]_0$. Similarly $[B]_0 = [D]_0 = 0$.

Therefore the final ODEs have the same initial conditions for all variables, and have the same relationships between variables, and in particular between $[B]$ and $[D]$.

So, for example, if we run a stochastic simulation of the left hand side with 1000*A1 and 1000*C1, we obtain the same curves for B and D than a stochastic simulation of the right hand side with 1000*Y.

Conclusions

Conclusions

- From Processes to ODEs
  - A way of relating automata-like models to classical ODE-based models
  - A way of investigating difference between discrete (stochastic) and continuous semantics
  - A compositional language (CPF) for describing systems of ODEs