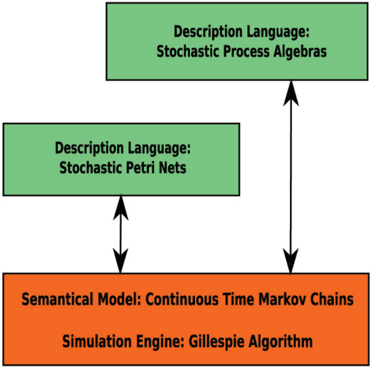
Modeling Biological Systems in Stochastic Concurrent Constraint Programming

Luca Bortolussi[1] Alberto Policriti[1]

[1]Department of Mathematics and Computer Science
University of Udine, Italy.

Workshop on Constraint Based Methods for Bioinformatics,
Nantes, 25th September 2006

Modeling Biological Systems with Stochastic Process Algebras



Pros

- Simple Language
- Compositionality

Cons

- Hard to encode general information
- Lacking computational extensibility

Constraints... why not?

Outline

1. Theory
   - Concurrent Constraint Programming
   - Continuous Time Markov Chains
   - Stochastic CCP
2. Bio-Modeling
   - Modeling Biochemical Reactions
   - Modeling Gene Regulatory Networks

Outline

1. Theory
   - Concurrent Constraint Programming
   - Continuous Time Markov Chains
   - Stochastic CCP

2. Bio-Modeling
   - Modeling Biochemical Reactions
   - Modeling Gene Regulatory Networks

Concurrent Constraint Programming

Constraint Store

- In this process algebra, the main object are **constraints**, which are *formulae over an interpreted first order language* (i.e. $X = 10$, $Y > X - 3$).
- Constraints can be added to a "pot", called the constraint store, but can never be removed.

Agents

Agents can perform two basic operations on this store:

- Add a constraint (**tell** ask)
- Ask if a certain relation is entailed by the current configuration (**ask** instruction)

Syntax of CCP

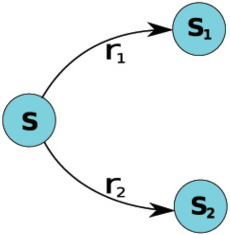
$$Program = Decl.A$$

$$D = \varepsilon \mid Decl.Decl \mid p(x) :- A$$

$$\begin{array}{ll} A = & \mathbf{0} \\ \mid & \mathrm{tell}(c).A \\ \mid & \mathrm{ask}(c_1).A_1 + \mathrm{ask}(c_2).A_2 \\ \mid & A_1 \parallel A_2 \mid \exists_x A \mid p(x) \end{array}$$

Continuous Time Markov Chains

A **Continuous Time Markov Chain** (CTMC) is a direct graph with edges labeled by a real number, called the rate of the transition (representing the speed or the frequency at which the transition occurs).



- In each state, we select the next state according to a *probability distribution* obtained normalizing rates (from S to S_1 with prob. $\frac{r_1}{r_1+r_2}$).
- The time spent in a state is given by an exponentially distributed random variable, with rate given by the *sum of outgoing transitions* from the actual node ($r_1 + r_2$).

Syntax of sCCP

Syntax of Stochastic CCP

$$Program = D.A$$

$$D = \varepsilon \mid D.D \mid p(\mathbf{x}) : -A$$

$$\pi = \mathrm{tell}_\lambda(c) \mid \mathrm{ask}_\lambda(c)$$
$$M = \pi.A \mid \pi.A.p(\mathbf{y}) \mid M + M$$
$$A = \mathbf{0} \mid \mathrm{tell}_\infty(c).A \mid \exists_x A \mid M \mid (A \parallel A)$$

Stochastic Rates

Each basic instruction (tell, ask, procedure call) has a rate attached to it. *Rates are functions from the constraint store* $\mathcal{C}$ *to positive reals*: $\lambda : \mathcal{C} \longrightarrow \mathbb{R}^+$.

sCCP soup

Operational Semantics

- There are *two transition relations*, one instantaneous (finite and confluent) and one stochastic.
- Traces are sequences of events with variable time delays among them.

Implementation

- We have an interpreter written in Prolog, using the *CLP engine of SICStus* to manage the constraint store.
- Efficiency issues.

Stream Variables

- *Quantities varying over time* can be represented in sCCP as unbounded lists.
- Hereafter: special meaning of $X = X + 1$.

Outline

1. Theory
   - Concurrent Constraint Programming
   - Continuous Time Markov Chains
   - Stochastic CCP

2. Bio-Modeling
   - Modeling Biochemical Reactions
   - Modeling Gene Regulatory Networks

General Principles

Measurable Entities	$\leftrightarrow$	Stream Variables
Logical Entities	$\leftrightarrow$	Processes (Control Variables)
Interactions	$\leftrightarrow$	Processes

Biochemical Arrows to sCCP processes

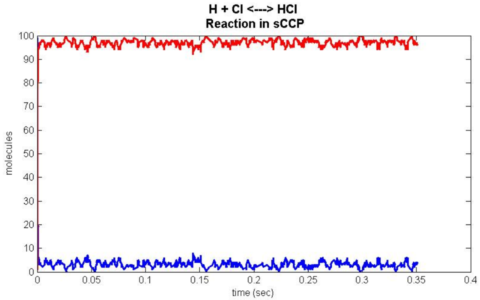
$R_1 + \ldots + R_n \rightarrow_k P_1 + \ldots + P_m$	$\text{reaction}(k, [R_1, \ldots, R_n], [P_1, \ldots, P_m]) :-$ $\text{ask}_{r_{MA}(k, R_1, \ldots, R_n)} \left(\bigwedge_{i=1}^{n} (R_i > 0) \right).$ $\left(\parallel_{i=i}^{n} \text{tell}_\infty (R_i = R_i - 1) \parallel_{j=1}^{m} \text{tell}_\infty (P_j = P_j + 1) \right).$ $\text{reaction}(k, [R_1, \ldots, R_n], [P_1, \ldots, P_m])$
$R_1 + \ldots + R_n \rightleftharpoons_{k_2}^{k_1} P_1 + \ldots + P_m$	$\text{reaction}(k_1, [R_1, \ldots, R_n], [P_1, \ldots, P_m]) \parallel$ $\text{reaction}(k_2, [P_1, \ldots, P_m], [R_1, \ldots, R_n])$
$S \mapsto_{K, V_0}^{E} P$	$\text{mm\_reaction}(K, V_0, S, P) :-$ $\text{ask}_{r_{MM}(K, V_0, S)}(S > 0).$ $(\text{tell}_\infty (S = S - 1) \parallel tell_\infty (P = P + 1)).$ $\text{mm\_reaction}(K, V_0, S, P)$
$S \mapsto_{K, V_0, h}^{E} P$	$\text{hill\_reaction}(K, V_0, h, S, P) :-$ $\text{ask}_{r_{Hill}(K, V_0, h, S)}(S > 0).$ $(\text{tell}_\infty (S = S - h) \parallel tell_\infty (P = P + h)).$ $\text{Hill\_reaction}(K, V_0, h, S, P)$

where $r_{MA}(k, X_1, \ldots, X_n) = k \cdot X_1 \cdots X_n$; $\quad r_{MM}(K, V_0, S) = \dfrac{V_0 S}{S + K}$; $\quad r_{Hill}(k, V_0, h, S) = \dfrac{V_0 S^h}{S^h + K^h}$

A simple reaction: H + Cl ⇋ HCl

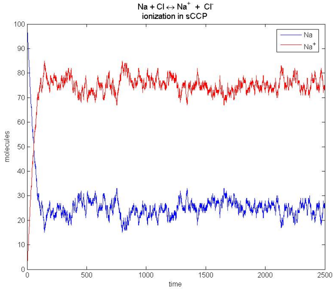
We have two reaction agents. The reagents and the products are stream variables of the constraint store (put down in the environment). *Independent on the number of molecules*.

$$\mathrm{reaction}(100, [H, CL], [HCL]) \parallel \mathrm{reaction}(10, [HCL], [H, CL])$$



Another reaction: $Na + Cl \leftrightharpoons Na^{+} + Cl^{-}$

$$\text{reaction}(100, [NA, CL], [NA+, CL-]) \parallel \text{reaction}(10, [NA+, CL-], [NA, CL])$$



Enzymatic reaction

$$S + E \rightleftharpoons^{k_1}_{k_{-1}} ES \rightarrow_{k_2} P + E$$

Mass Action Kinetics

enz_reaction($k_1, k_{-1}, k_2, S, E, ES, P$) :-
reaction(k_1, [S, E], [ES]) ||
reaction(k_{-1}, [ES], [E, S]) ||
reaction(k_2, [ES], [E, P])

Mass Action Equations

$$\frac{d[ES]}{dt} = k_1[S][E] - k_2[ES] - k_{-1}[ES]$$
$$\frac{d[E]}{dt} = -k_1[S][E] + k_2[ES] + k_{-1}[ES]$$
$$\frac{d[S]}{dt} = -k_1[S][E]$$
$$\frac{d[P]}{dt} = k_2[ES]$$

Michaelis-Menten Equations

$$\frac{d[P]}{dt} = \frac{V_0 S}{S+K}$$
$$V_0 = k_2[E_0]$$
$$K = \frac{k_2+k_{-1}}{k_1}$$

Michaelis-Menten Kinetics

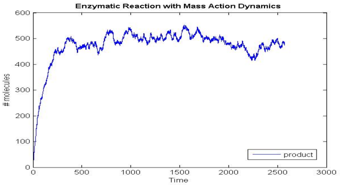
$$\text{mm\_reaction}\left(\frac{k_2 + k_{-1}}{k_1}, k_2 \cdot E, S, P\right)$$

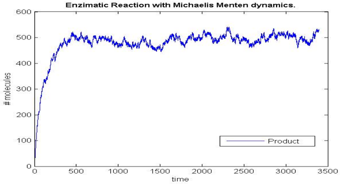
Enzymatic reaction

$$S + E \rightleftharpoons^{k_1}_{k_{-1}} ES \rightarrow_{k_2} P + E$$

Mass Action Kinetics

enz_reaction($k_1, k_{-1}, k_2, S, E, ES, P$) :-
  reaction(k_1, [S, E], [ES]) ||
  reaction(k_{-1}, [ES], [E, S]) ||
  reaction(k_2, [ES], [E, P])

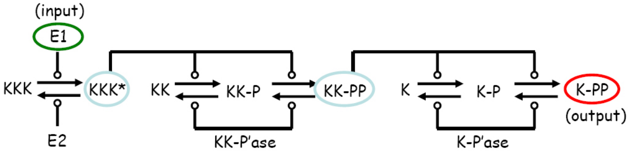




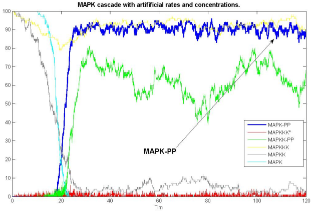
Michaelis-Menten Kinetics

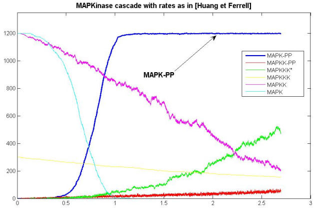
$$\text{mm\_reaction}\left(\frac{k_2 + k_{-1}}{k_1}, k_2 \cdot E, S, P\right)$$

MAP-Kinase cascade

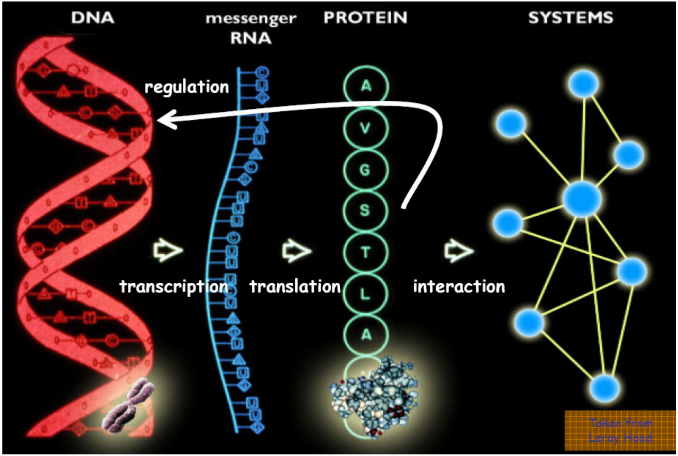


$\text{enz\_reaction}(k_a, k_d, k_r, KKK, E1, KKKE1, KKKS) \parallel \text{enz\_reaction}(k_a, k_d, k_r, KKKS, E2, KKKSE2, KKK) \parallel$
$\text{enz\_reaction}(k_a, k_d, k_r, KK, KKKS, KKKKKS, KKP) \parallel \text{enz\_reaction}(k_a, k_d, k_r, KKP, KKP1, KKPKKP1, KK) \parallel$
$\text{enz\_reaction}(k_a, k_d, k_r, KKP, KKKS, KKPKKKS, KKPP) \parallel \text{enz\_reaction}(k_a, k_d, k_r, KP, KP1, KPKP1, K) \parallel$
$\text{enz\_reaction}(k_a, k_d, k_r, K, KKPP, KKKPP, KP) \parallel \text{enz\_reaction}(k_a, k_d, k_r, KKPP, KKP1, KKPPKKP1, KKP) \parallel$
$\text{enz\_reaction}(k_a, k_d, k_r, KP, KKPP, KPKKPP, KPP) \parallel \text{enz\_reaction}(k_a, k_d, k_r, KPP, KP1, KPPKP1, KP)$

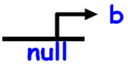




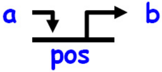
The gene machine



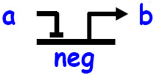
The instruction set



$$\text{null\_gate}(k_p, X) :- \text{tell}_{k_p}(X = X+1).\text{null\_gate}(k_p, X)$$



$$\text{pos\_gate}(k_p, k_e, k_f, X, Y) :- \text{tell}_{k_p}(X = X+1).\text{pos\_gate}(k_p, k_e, k_f, X, Y) + \text{ask}_{r(k_e, Y)}(\mathit{true}).\text{tell}_{k_e}(X = X+1).\text{pos\_gate}(k_p, k_e, k_f, X, Y)$$

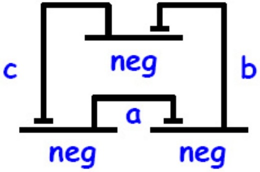


$$\text{neg\_gate}(k_p, k_i, k_d, X, Y) :- \text{tell}_{k_p}(X = X+1).\text{neg\_gate}(k_p, k_i, k_d, X, Y) + \text{ask}_{r(k_i, Y)}(\mathit{true}).\text{ask}_{k_d}(\mathit{true}).\text{neg\_gate}(k_p, k_i, k_d, X, Y)$$

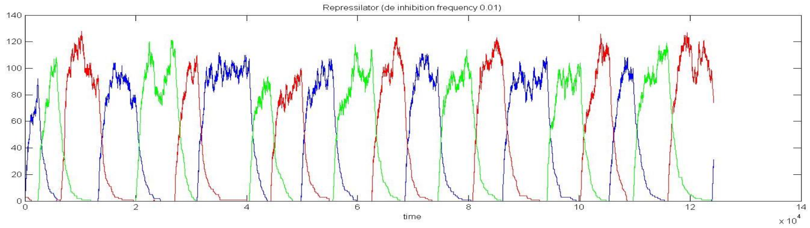
where $r(k, Y) = k \cdot Y$.

L. Cardelli, A. Phillips, 2005.

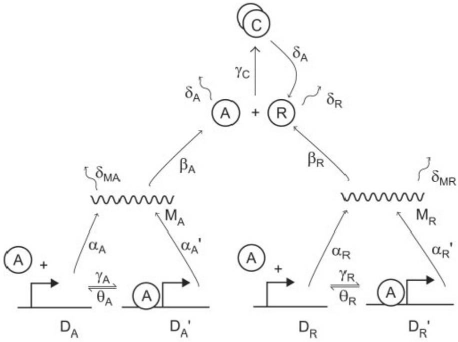
Repressilator



$$\text{neg\_gate}(0.1, 1, 0.0001, A, C) \parallel$$
$$\text{reaction}(0.0001, [A], []) \parallel$$
$$\text{neg\_gate}(0.1, 1, 0.0001, B, A) \parallel$$
$$\text{reaction}(0.0001, [B], []) \parallel$$
$$\text{neg\_gate}(0.1, 1, 0.0001, C, B) \parallel$$
$$\text{reaction}(0.0001, [C], [])$$

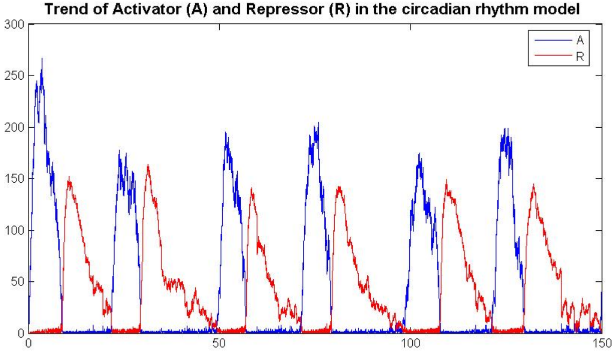


Circadian Clock



Circadian Clock

$$\text{pos\_gate}(\alpha_A, \alpha'_A, \gamma_A, \theta_A, M_A, A) \parallel \text{pos\_gate}(\alpha_R, \alpha'_R, \gamma_R, \theta_R, M_R, A) \parallel$$
$$\text{reaction}(\beta_A, [M_A], [A]) \parallel \text{reaction}(\delta_{MA}, [M_A], []) \parallel$$
$$\text{reaction}(\beta_R, [M_R], [R]) \parallel \text{reaction}(\delta_{MR}, [M_R], []) \parallel$$
$$\text{reaction}(\gamma_C, [A, R], [AR]) \parallel \text{reaction}(\delta_A, [AR], [R]) \parallel$$
$$\text{reaction}(\delta_A, [A], []) \parallel \text{reaction}(\delta_R, [R], [])$$



Conclusions

- We have introduced a stochastic version of CCP, with functional rates.
- We showed that sCCP may be used for modeling biological systems, defining *libraries* for biochemical reactions and gene regulatory networks.
- We showed that non-constant rates allow to use more complex chemical kinetics than mass action one.

The End

THANKS FOR THE ATTENTION!

QUESTIONS?