Stochastic π-Calculus in Systems Biology

Marek Kwiatkowski

26 October 2006

Contents

Outline

1. Stochastic π
2. Modelling Biological Reality in Sπ
3. Related work

π-calculus: syntax

$$
\begin{aligned}
P ::= \quad & 0 \\
& \pi.P \\
& P + P \\
& P \mid P \\
& (\nu x)P
\end{aligned}
$$

Stochastic π-calculus: syntax

$$
\begin{aligned}
P ::= \quad & \\
& 0 \\
& (\pi, r).P \qquad r \in \mathbb{R}_+ \\
& P + P \\
& P \mid P \\
& (\nu x)P
\end{aligned}
$$

[Priami 1995]

Stochastic π-calculus: idea

Intuitively, all actions enabled in a process compete for execution and their relative chances of success are given by their rates. This is called *the race condition*.

Stochastic π-calculus: idea

Intuitively, all actions enabled in a process compete for execution and their relative chances of success are given by their rates. This is called *the race condition*.

More precisely: every action (π, r) experiences a delay which is a random variable with exponential distribution characterised by r. The action that drew the smallest delay is executed.

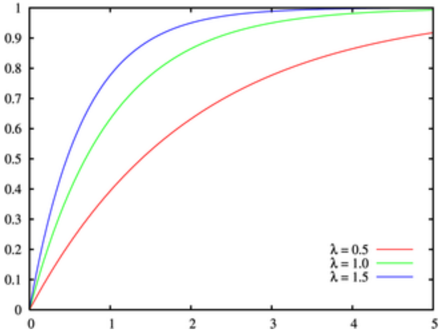
Stochastic π-calculus: idea

Intuitively, all actions enabled in a process compete for execution and their relative chances of success are given by their rates. This is called *the race condition*.

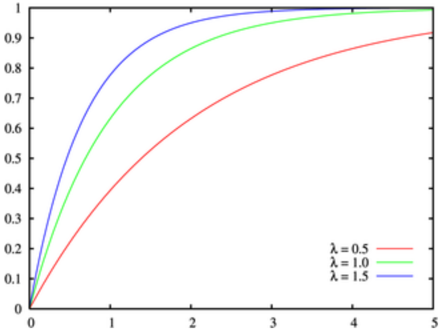
More precisely: every action (π, r) experiences a delay which is a random variable with exponential distribution characterised by r. The action that drew the smallest delay is executed.

The probability of two actions being assigned the same delay is zero.

Exponential distributions

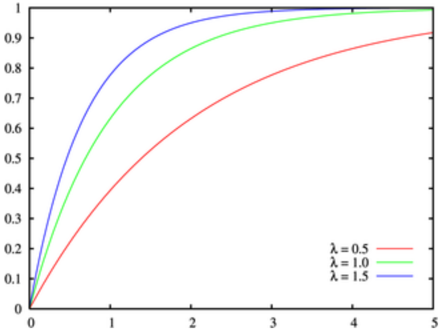


Exponential distributions



- Exponential distribution are *memoryless*: for any random variable X and real numbers r, s we have $P(X > t + s | X > t) = P(X > s)$

Exponential distributions



- Exponential distribution are *memoryless*: for any random variable X and real numbers r, s we have $P(X > t + s | X > t) = P(X > s)$
- The mean value is $1/\lambda$.

A sample process

Consider the process $P \stackrel{\text{df}}{=} ((a,3)+(b,5)) \mid (a,2)$.

A sample process

Consider the process $P \stackrel{\text{df}}{=} ((a,3)+(b,5)) \mid (a,2)$.
It can perform three transitions:

1. $P \xrightarrow{(a,3)} 0 \mid (a,2)$
2. $P \xrightarrow{(a,2)} ((a,3)+(b,4)) \mid 0$
3. $P \xrightarrow{(b,5)} 0 \mid (a,2)$

A sample process

Consider the process $P \stackrel{\text{df}}{=} ((a,3)+(b,5)) \mid (a,2)$.
It can perform three transitions:

1. $P \xrightarrow{(a,3)} 0 \mid (a,2)$
2. $P \xrightarrow{(a,2)} ((a,3)+(b,4)) \mid 0$
3. $P \xrightarrow{(b,5)} 0 \mid (a,2)$

- On average, the first transition will be performed 30% of the time, the second 20% of the time and the third 50% of the time.

A sample process

Consider the process $P \stackrel{\text{df}}{=} ((a,3)+(b,5)) \mid (a,2)$.
It can perform three transitions:

1. $P \xrightarrow{(a,3)} 0 \mid (a,2)$
2. $P \xrightarrow{(a,2)} ((a,3)+(b,4)) \mid 0$
3. $P \xrightarrow{(b,5)} 0 \mid (a,2)$

- On average, the first transition will be performed 30% of the time, the second 20% of the time and the third 50% of the time.
- An external observer sees a half of the time. This is captured by the notion of *the apparent rate*. The apparent rate of a in P is 5.

Other developments in Priami's original paper

- Immediate and passive actions, i.e. $r = \infty$ and $r = 0$

Other developments in Priami's original paper

- Immediate and passive actions, i.e. $r = \infty$ and $r = 0$
- Finitely branching, fully labelled (*proved*) semantics

Other developments in Priami's original paper

- Immediate and passive actions, i.e. $r = \infty$ and $r = 0$
- Finitely branching, fully labelled (*proved*) semantics
- Continuous Time Markov Chain derived from the transition system

Outline

Some (simple) biology

We are going to model protein interaction networks.

Some (simple) biology

We are going to model protein interaction networks.

- Proteins are complexes of domains or *motifs*

Some (simple) biology

We are going to model protein interaction networks.

- Proteins are complexes of domains or *motifs*
- Proteins interact with each other, possibly altering their potential for future interactions

Some (simple) biology

We are going to model protein interaction networks.

- Proteins are complexes of domains or *motifs*
- Proteins interact with each other, possibly altering their potential for future interactions
- Interaction can be limited by the presence of cellular compartments or be allowed only between proteins bound to each other

Some (simple) biology

We are going to model protein interaction networks.

- Proteins are complexes of domains or *motifs*
- Proteins interact with each other, possibly altering their potential for future interactions
- Interaction can be limited by the presence of cellular compartments or be allowed only between proteins bound to each other
- Quantitative aspects of a reaction depend on protein concentrations and the nature of the reaction.

Key correspondencies

Key correspondencies

- motifs/proteins $\sim$ processes

Key correspondencies

- motifs/proteins $\sim$ processes
- interaction $\sim$ communication

Key correspondencies

- motifs/proteins $\sim$ processes
- interaction $\sim$ communication
- complexes/compartments $\sim$ shared private channels

Key correspondencies

- motifs/proteins $\sim$ processes
- interaction $\sim$ communication
- complexes/compartments $\sim$ shared private channels
- basal rate $\sim$ rate of an activity

Key correspondencies

- motifs/proteins $\sim$ processes
- interaction $\sim$ communication
- complexes/compartments $\sim$ shared private channels
- basal rate $\sim$ rate of an activity

[Regev et al. 2001], [Priami, Regev et al. 2001]

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$
$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$$

A sample system

$$Sys = Gene\_A|Gene\_TF|Expression|Protein\_Deg$$

$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$
$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$

$Gene\_TF = (basal(), 4).(Gene\_TF|Protein\_TF) + (pA(), 40).(Gene\_TF|Protein\_TF)$
$Gene\_A = (basal(), 4).(Gene\_A|Protein\_A) + (pA(), 40).(Gene\_A|Protein\_A)$

$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$$

$$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$$
$$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site | Kinase)$$

A sample system

$$Sys = Gene\_A|Gene\_TF|Expression|Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF|Protein\_TF) + (pA(), 40).(Gene\_TF|Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A|Protein\_A) + (pA(), 40).(Gene\_A|Protein\_A)$$

$$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$$
$$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site|Kinase)$$

$$Binding\_Site = (\overline{bind}\, \langle bb1, bb2, bb3 \rangle\ , 0.1).Bound\_Site + (degp(), 0.1).(\overline{bb3}, \infty)$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$$

$$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$$
$$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site | Kinase)$$

$$Binding\_Site = (\overline{bind}\langle bb1, bb2, bb3\rangle, 0.1).Bound\_Site + (degp(), 0.1).(\overline{bb3}, \infty)$$
$$Kinase = (\overline{bb2}\langle ptail\rangle, 10).Kinase + (bb3(), \infty)$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$$

$$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$$
$$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site | Kinase)$$

$$Binding\_Site = (\overline{bind}\, \langle bb1, bb2, bb3 \rangle\, , 0.1).Bound\_Site + (degp(), 0.1).(\overline{bb3}, \infty)$$
$$Kinase = (\overline{bb2}\, \langle ptail \rangle\, , 10).Kinase + (bb3(), \infty)$$
$$Bound\_Site = (\overline{bb1}, 10).Binding\_Site + (degp(), 0.1).(\overline{bb3}, \infty).(\overline{bb3}, \infty)$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$$
$$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$$

$$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$$
$$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$$

$$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$$
$$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site | Kinase)$$

$$Binding\_Site = (\overline{bind}\langle bb1, bb2, bb3 \rangle, 0.1).Bound\_Site + (degp(), 0.1).(\overline{bb3}, \infty)$$
$$Kinase = (\overline{bb2}\langle ptail \rangle, 10).Kinase + (bb3(), \infty)$$
$$Bound\_Site = (\overline{bb1}, 10).Binding\_Site + (degp(), 0.1).(\overline{bb3}, \infty).(\overline{bb3}, \infty)$$

$$Bound\_TF =$$
$$(c\_bb1(), 10).Protein\_TF + (c\_bb3(), \infty) + (c\_bb2(tail), 10).((c\_bb1(), 10).Active\_TF(tail) + (c\_bb3(), \infty))$$

A sample system

$$Sys = Gene\_A | Gene\_TF | Expression | Protein\_Deg$$

$Protein\_Deg = (\overline{degp}, 1).Protein\_Deg$
$Expression = (\overline{basal}, 4).Expression + (ptail(), 100).(\overline{pA}, 40).Expression$

$Gene\_TF = (basal(), 4).(Gene\_TF | Protein\_TF) + (pA(), 40).(Gene\_TF | Protein\_TF)$
$Gene\_A = (basal(), 4).(Gene\_A | Protein\_A) + (pA(), 40).(Gene\_A | Protein\_A)$

$Protein\_TF = (bind(c\_bb1, c\_bb2, c\_bb3), 0.1).Bound\_TF + (degp(), 0.1)$
$Protein\_A = (\nu\, bb1, bb2, bb3)(Binding\_Site | Kinase)$

$Binding\_Site = (\overline{bind}\langle bb1, bb2, bb3\rangle, 0.1).Bound\_Site + (degp(), 0.1).(\overline{bb3}, \infty)$
$Kinase = (\overline{bb2}\langle ptail\rangle, 10).Kinase + (bb3(), \infty)$
$Bound\_Site = (\overline{bb1}, 10).Binding\_Site + (degp(), 0.1).(\overline{bb3}, \infty).(\overline{bb3}, \infty)$

$Bound\_TF =$
$(c\_bb1(), 10).Protein\_TF + (c\_bb3(), \infty) + (c\_bb2(tail), 10).((c\_bb1(), 10).Active\_TF(tail) + (c\_bb3(), \infty))$

$Active\_TF(tail) = (\overline{tail}, 100).Active\_TF(tail) + (degp(), 0.1)$

Technicalities

- Priami's original semantics allow many transitions with the same action between the same two processes.

Technicalities

- Priami's original semantics allow many transitions with the same action between the same two processes.
- Bio–Sπ semantics have to take concentrations (i.e. quantities) into account.

$$\frac{P \xrightarrow{(x, r_b \cdot r_0 \cdot r_1)} P'}{Q|P \xrightarrow{(x, r_b \cdot r'_0 \cdot r'_1)} Q|P'}, \quad \begin{cases} r'_0 = r_0 + In_x(Q) \\ r'_1 = r_1 + Out_x(Q) \end{cases}$$

Outline

1. Stochastic π
2. Modelling Biological Reality in Sπ
3. Related work

Extensions and refinements

- BioAmbients ([Regev et al. 2004])

Extensions and refinements

- BioAmbients ([Regev et al. 2004])
- Beta Binders ([Priami, Quaglia 2005])

Extensions and refinements

- BioAmbients ([Regev et al. 2004])
- Beta Binders ([Priami, Quaglia 2005])
- Brane calculi ([Cardelli 2005])

Extensions and refinements

- BioAmbients ([Regev et al. 2004])
- Beta Binders ([Priami, Quaglia 2005])
- Brane calculi ([Cardelli 2005])

Automation

- BioSPi – the original system of Regev, Shapiro and Silverman. Newest version released 25 Mar 2006.
  `http://www.weizmann.ac.il/home/biospi/aspic-release.tar.gz`
- SPiM (Stochastic Pi Machine) – Andrew Phillips, MS Research. Version 0.042 released 2 Sep 2006.
  `http://research.microsoft.com/~aphillip/spim/`

Bibliography

Core papers:

- [Priami 1995] C. Priami *Stochastic π-calculus*, The Computer Journal 38(6), 1995.
- [Priami, Regev et al. 2001] C. Priami, A. Regev, E.Shapiro, W. Silverman *Application of a stochastic name-passing calculus to representation and simulation of molecular processes*, Information Processing Letters 80, 2001

Other referenced papers:

- [Regev et al. 2001] A. Regev, W. Silverman, E. Shapiro *Representation and simulation of biochemical processes using the π-calculus process algebra*, Proceedings of the Pacific Symposium of Biocomputing 2001,
- [Regev et al. 2004] A. Regev, E. M. Panina, W. Silverman, L. Cardelli, E. Shapiro *BioAmbients: an abstraction for biological compartments*, TCS 325(1), 2005.
- [Priami, Quaglia 2005] C. Priami, P. Quaglia *Beta Binders for Biologcal Interactions*, LNCS 3082, 2005.
- [Cardelli 2005] L. Cardelli *Brane calculi. Interactions of biological membranes* LNCS 3082, 2005.

Other relevant papers:

- A. Regev *Computational systems biology: a calculus for biomolecular knowledge.* PhD thesis, Tel-Aviv University 2002
- A. Regev, E. Shapiro *The π-calculus as an abstraction for biomolecular systems*, chapter in G. Ciobanu, G. Rozenberg (eds.) *Modelling in Molecular Biology*, Natural Computing Series, Springer 2004