Cell Signalling Networks: Biological Computers with Universal Computational Capacity

Thomas Hinze Thorsten Lenser Peter Dittrich

Bio Systems Analysis Group
Friedrich Schiller University Jena

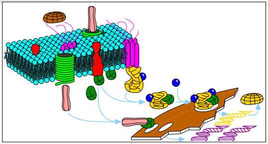
{hinze,thlenser,dittrich}@cs.uni-jena.de
www.minet.uni-jena.de/csb

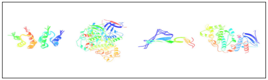
16. Theorietag
Automaten und Formale Sprachen

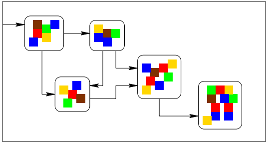
ESIGNET

Outline

Synthetic Biology: Applications for Theoretical Computer Science



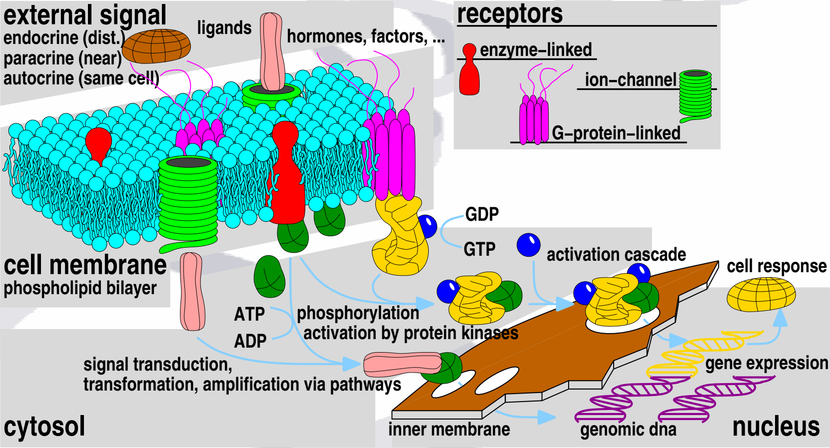




ESIGNET

Biological Principles of Cell Signalling

Information Processing in Living Cells



ESIGNET

Modelling Cell Signalling Networks

Phenotypic Description of CSNs – Structure, Behaviour, Function

stochastic approaches
Bayesian networks, Markov chains, master equation, ...

algebraic approaches — **P systems**
term rewriting, automata, process calculi, ...

analytical approaches
differential and difference equations

category based approaches
formal specification/verification, description logics

heterogeneous approaches

biological observations

computer languages
simulation
analysis

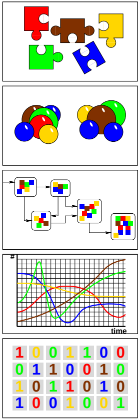
SBML

ESIGNET

Motivation and Intention

Adopting Advantages of P Systems to Project Objectives

- Capture significant aspects of cellular signalling:
  - components, topology, modularity
  - protein substructures and properties
  - dynamical behaviour
  - discrete characteristics
  - signal coding and transduction
  - robustness to perturbations
  - computational capacity
- Provide a general framework
- P systems feature different levels of abstraction
- Keep formalism tractable
- Balance detailedness with computational needs
- Facilitate system modification, recombination, and construction ab initio



ESIGNET

System Definition

Identification of Components Based on Biological Model

- System $\Pi_{\mathrm{CSN}} = (V, V', E, M, n)$
  - V: alphabet of protein identifiers
  - V': alphabet of protein substructure/property identifiers
  - M: **modules** $\longrightarrow$ functional reaction units
  - E: graph $\longrightarrow$ transduction **channels** between modules
  - n: number of modules (degree of the P system)
- Modules $M_i = (R_{i1}, \ldots, R_{ir_i}, \mathrm{f}_{i1}, \ldots \mathrm{f}_{ir_i}, A_i) \in M$
  - R_{ij}: reaction rule $\longrightarrow$ multisets of educts and products may contain meta-symbols $\longrightarrow$ **matching** required
  - f_{ij}: function corresponding to **kinetics** of R_{ij}, number of educt objects taken from module within one reaction step
  - A_i: multiset of axioms $\longrightarrow$ initial contents of M_i
- Channels $e_{ij} = (i, j, I_{ij}, \mathrm{d}_{ij}) \in E$
  - weighted directed channel from module i to module j
  - I_{ij}: filter interface (**receptor pattern** and **conc. gradient**)
  - d_{ij}: time **delay** (number of system **steps**) for passage

ESIGNET

System Behaviour and Properties

Controlled Interplay of Components at Global Level

- Definition of dynamical system behaviour
  - contents of module M_i at global time $t \in \mathbf{N}$: multiset $L_i(t)$
  - system step by module M_i:

$$
\begin{aligned}
L_i(0) &= A_i \\
L_i'(t) &= L_i(t) \ominus Educts_i(t) \uplus Products_i(t) \\
L_i(t+1) &= L_i'(t) \ominus Outgoing_i(t) \uplus Incoming_i(t)
\end{aligned}
$$

1. Determine multiset of educts using $L_i(t)$, $R_{i1}, \ldots, R_{ir_i}, \mathrm{f}_{i1}, \ldots \mathrm{f}_{ir_i}$; involves matching
2. Remove educt objects from module contents
3. Determine and add multiset of reaction products, obtain $L_i'(t)$
4. Determine and separate objects leaving host module evaluate $L_i'(t)$ and I for each outgoing channel, matching
5. Add objects received from incoming channels, consider d

- System properties
  - modularity – static system topology – ability to identify objects/substructures – flexibility in level of abstraction
  - determinism – computational tractability – **universality**

Random Access Machine (RAM)

Established Turing-Complete Model for Computation

- Syntactical denotation of components
  $RAM = (R, L, P, l_1)$
  - l_1: jump label of first instruction
  - P: program (finite set of instructions)
  - L: finite set of jump labels $L = \{l_1, \ldots, l_c\}$
  - R: finite set of registers $R = \{r_1, \ldots, r_m\}, \quad r_k \in \mathbf{N}$
- Available instructions
  - l_i : $\mathtt{INCR}(r_k), l_j$ — increment register r_k, jump to l_j
  - l_i : $\mathtt{DECR}(r_k), l_j$ — decrement register r_k, jump to l_j
  - l_i : $r_k \neq 0, l_j, l_p$ — if $r_k \neq 0$ jump to l_j else jump to l_p
  - l_i : HALT — terminate program and output
- Useful assumptions
  - consecutive indexing of jump labels and registers
  - determinism
  - initialisation of registers at start
  - output of all m registers when HALT

ESIGNET

Random Access Machine (RAM)

Example for a Computation and Subsequent Simulation by Π_{CSN}

- Consider for simulation and transformation

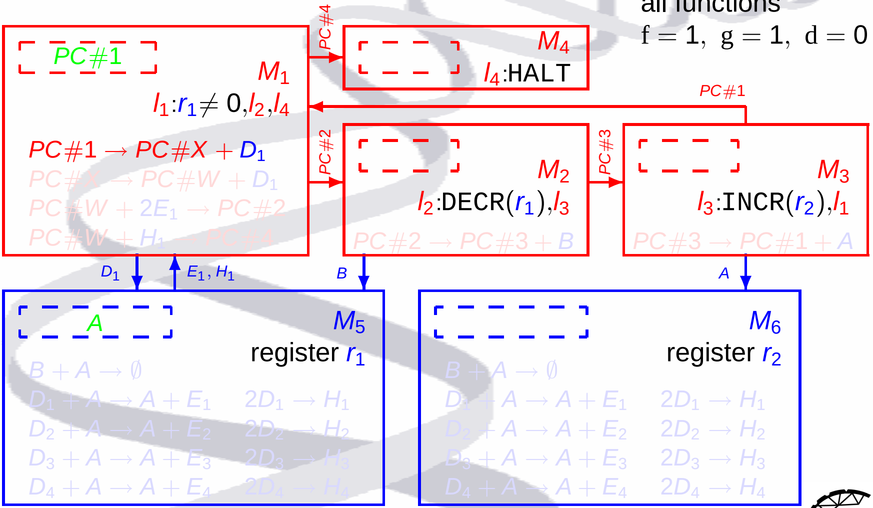
$$
\begin{aligned}
RAM &= (R, L, P, l_1) \\
R &= \{r_1, r_2\} \\
L &= \{l_1, l_2, l_3, l_4\} \\
P &= \{l_1 : r_1 \neq 0, l_2, l_4, \\
&\quad\; l_2 : \mathtt{DECR}(r_1), l_3, \\
&\quad\; l_3 : \mathtt{INCR}(r_2), l_1, \\
&\quad\; l_4 : \mathtt{HALT}\}
\end{aligned}
$$

- Initialisation: $r_1 = 1$ and $r_2 = 0$ (arbitrarily chosen)
- Program moves contents of r_1 cumulatively to r_2
- RAM consists of $m = 2$ registers and $c = 4$ instructions
- Each register and each instruction forms separate module of P system $\Pi_{\mathrm{CSN}} = (V, V', E, M, c + m)$

ESIGNET

Simulation of RAM by P System Π_{CSN}

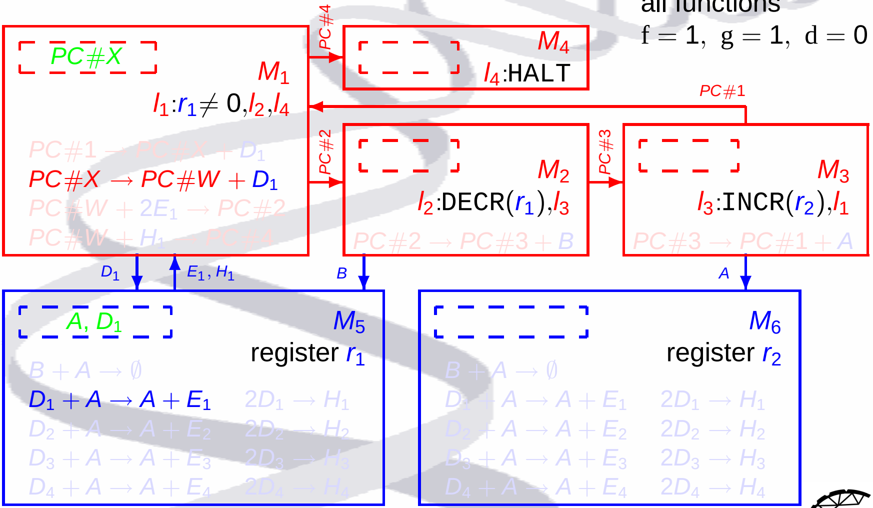
all functions
$\mathrm{f} = 1, \ \mathrm{g} = 1, \ \mathrm{d} = 0$



ESIGNET

Simulation of RAM by P System Π_{CSN}

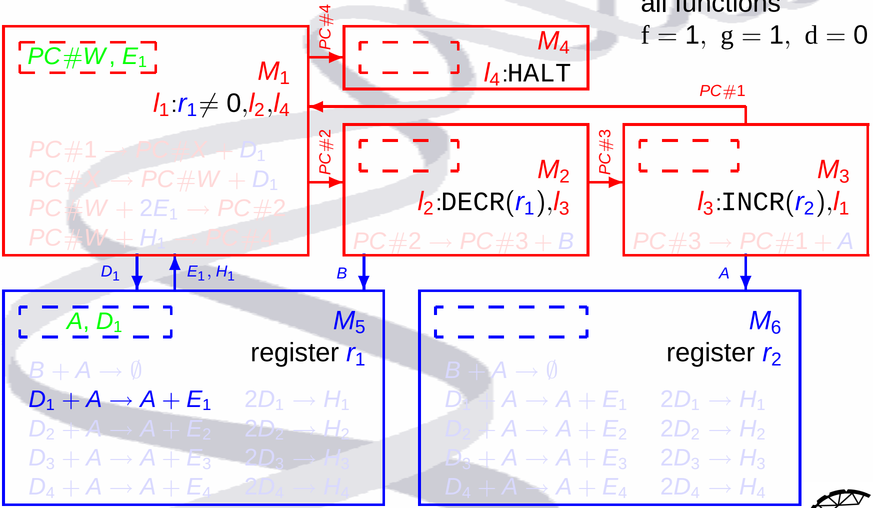
all functions $f = 1, \; g = 1, \; d = 0$



ESIGNET

Simulation of RAM by P System Π_{CSN}

all functions
$f = 1,\ g = 1,\ d = 0$

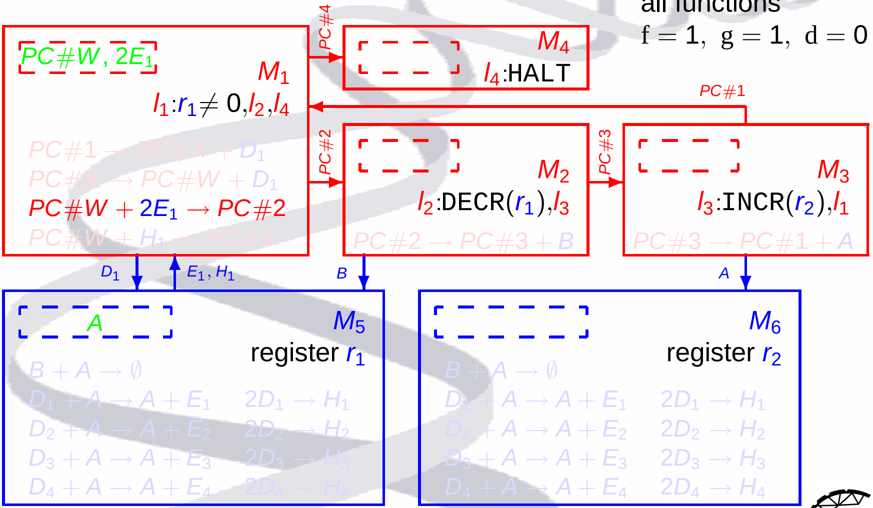


ESIGNET

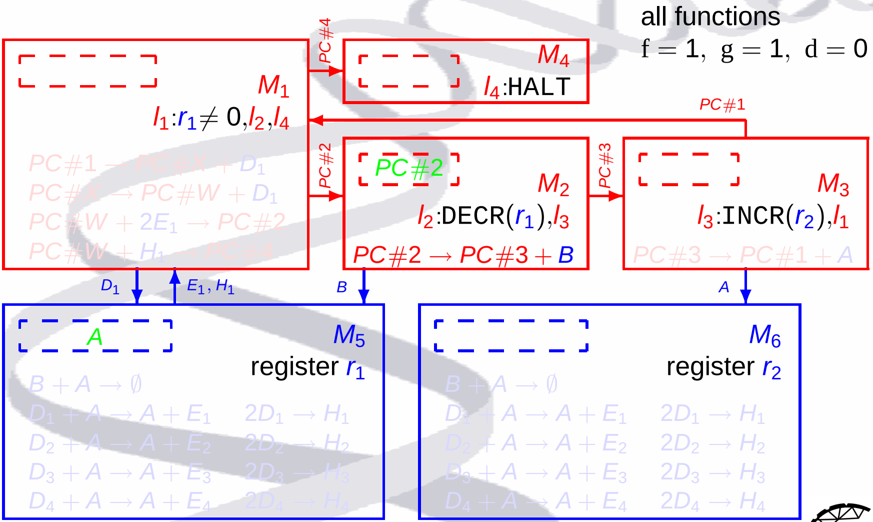
Simulation of RAM by P System Π_{CSN}

all functions

$f = 1, \; g = 1, \; d = 0$

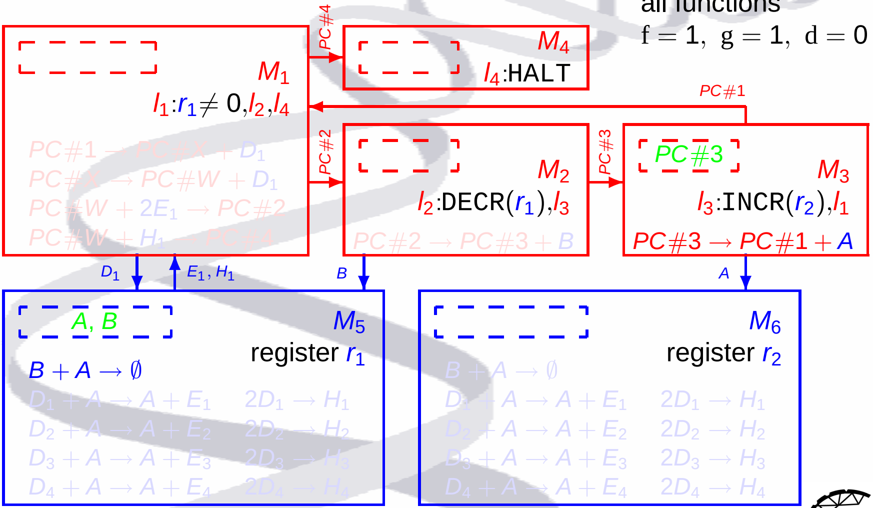


Simulation of RAM by P System Π_{CSN}



Simulation of RAM by P System Π_{CSN}

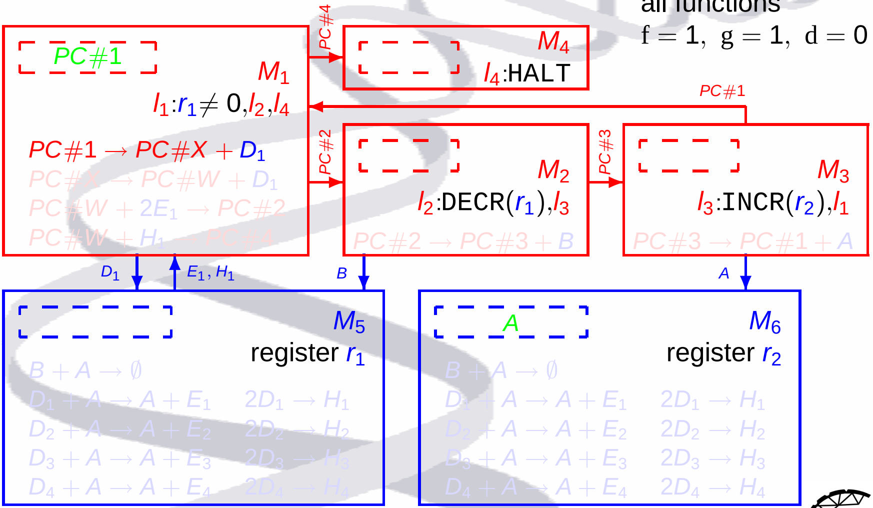
all functions
$\mathrm{f} = 1,\ \mathrm{g} = 1,\ \mathrm{d} = 0$



ESIGNET

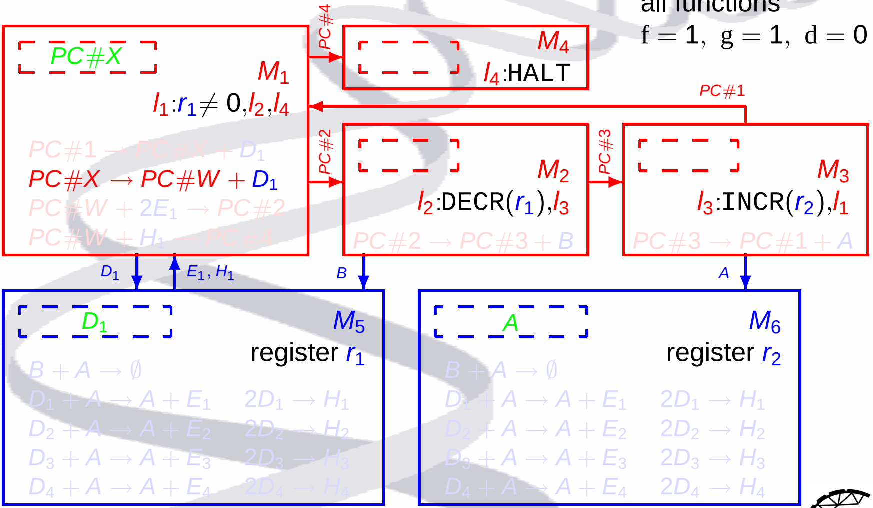
Simulation of RAM by P System Π_{CSN}

all functions
$f = 1, \; g = 1, \; d = 0$

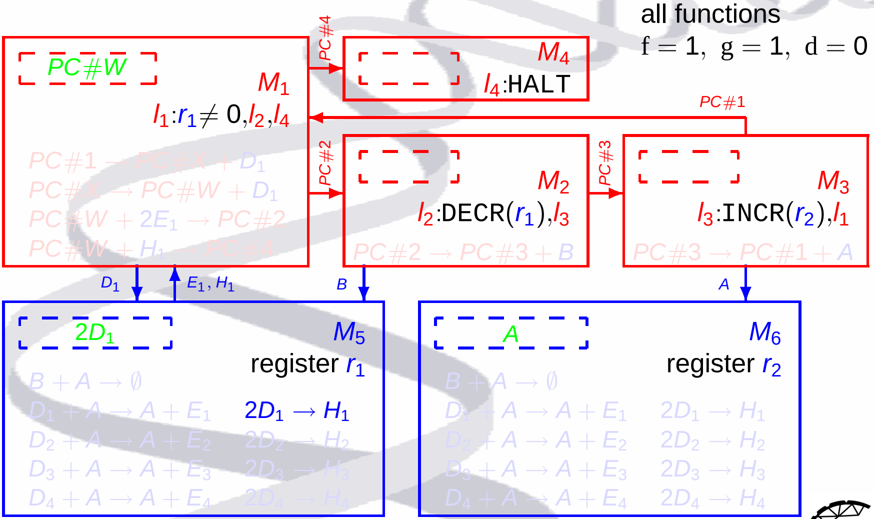


Simulation of RAM by P System Π_{CSN}

all functions
$f = 1, \ g = 1, \ d = 0$

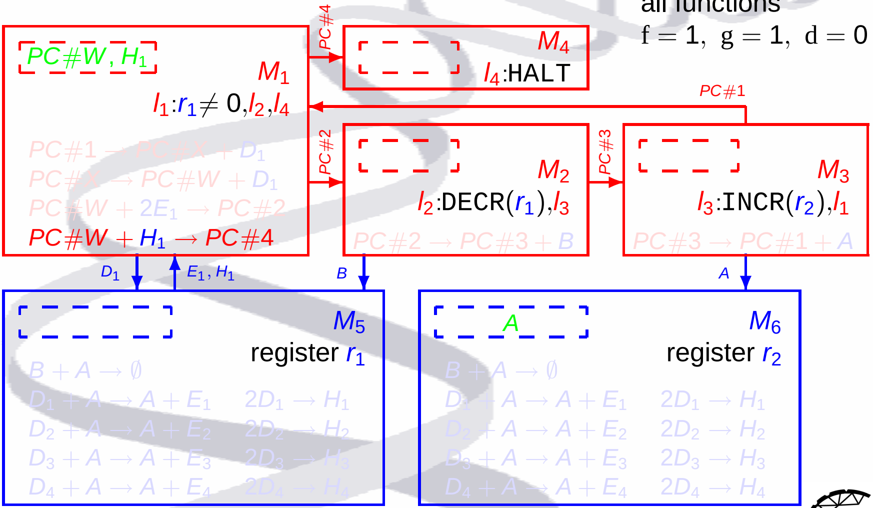


Simulation of RAM by P System Π_{CSN}



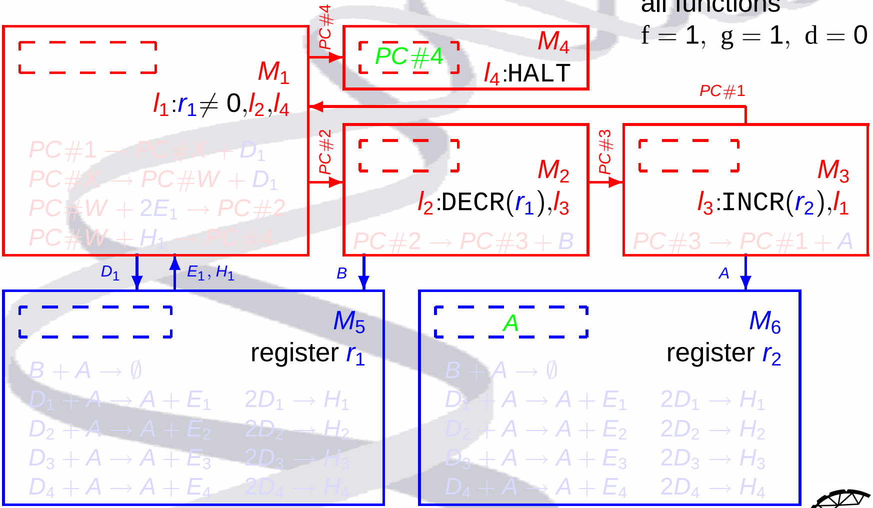
Simulation of RAM by P System Π_{CSN}

all functions
$f = 1,\ g = 1,\ d = 0$



Simulation of RAM by P System Π_{CSN}

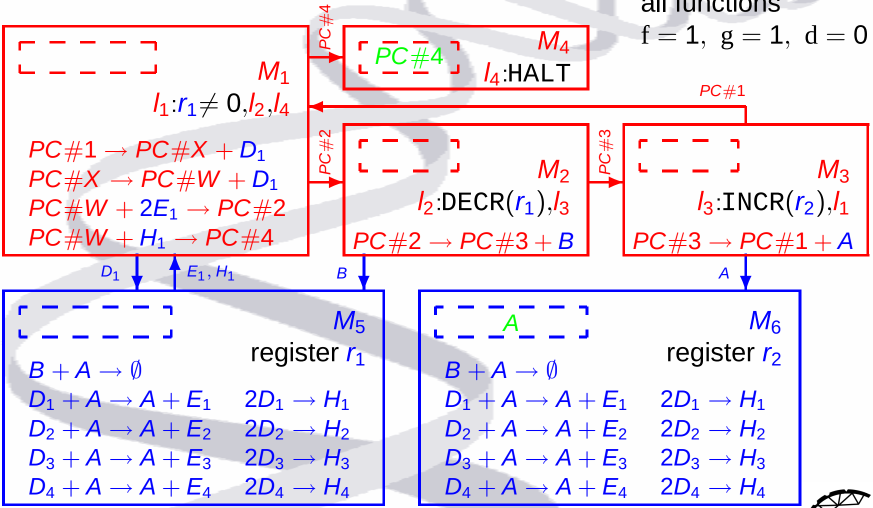
all functions
$\mathrm{f} = 1, \ \mathrm{g} = 1, \ \mathrm{d} = 0$



ESIGNET

Simulation of RAM by P System Π_{CSN}

all functions
$\mathrm{f} = 1, \ \mathrm{g} = 1, \ \mathrm{d} = 0$

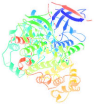


ESIGNET

Conclusion and Future Work

Summary

- P systems provide a suitable framework for modelling cell signalling networks
- Valuable insight can be gained from combining previously isolated ideas
- Capabilities for computing in vivo

Outlook

- Software simulation to prove practicability of the approach
- Interface to biological databases / experimental knowledge
- Investigation into the evolution of CSNs