A Protein Substructure Based P System for Description and Analysis of Cell Signalling Networks

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

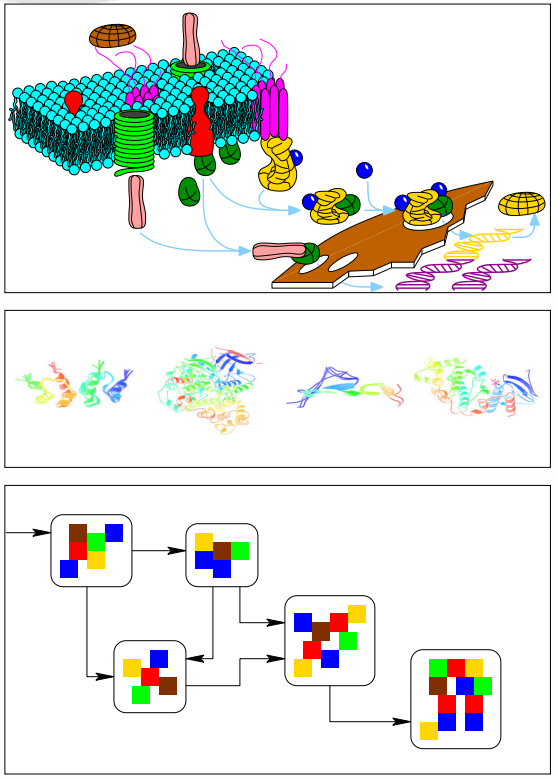
7th Workshop on
Membrane Computing

ESIGNET

Outline

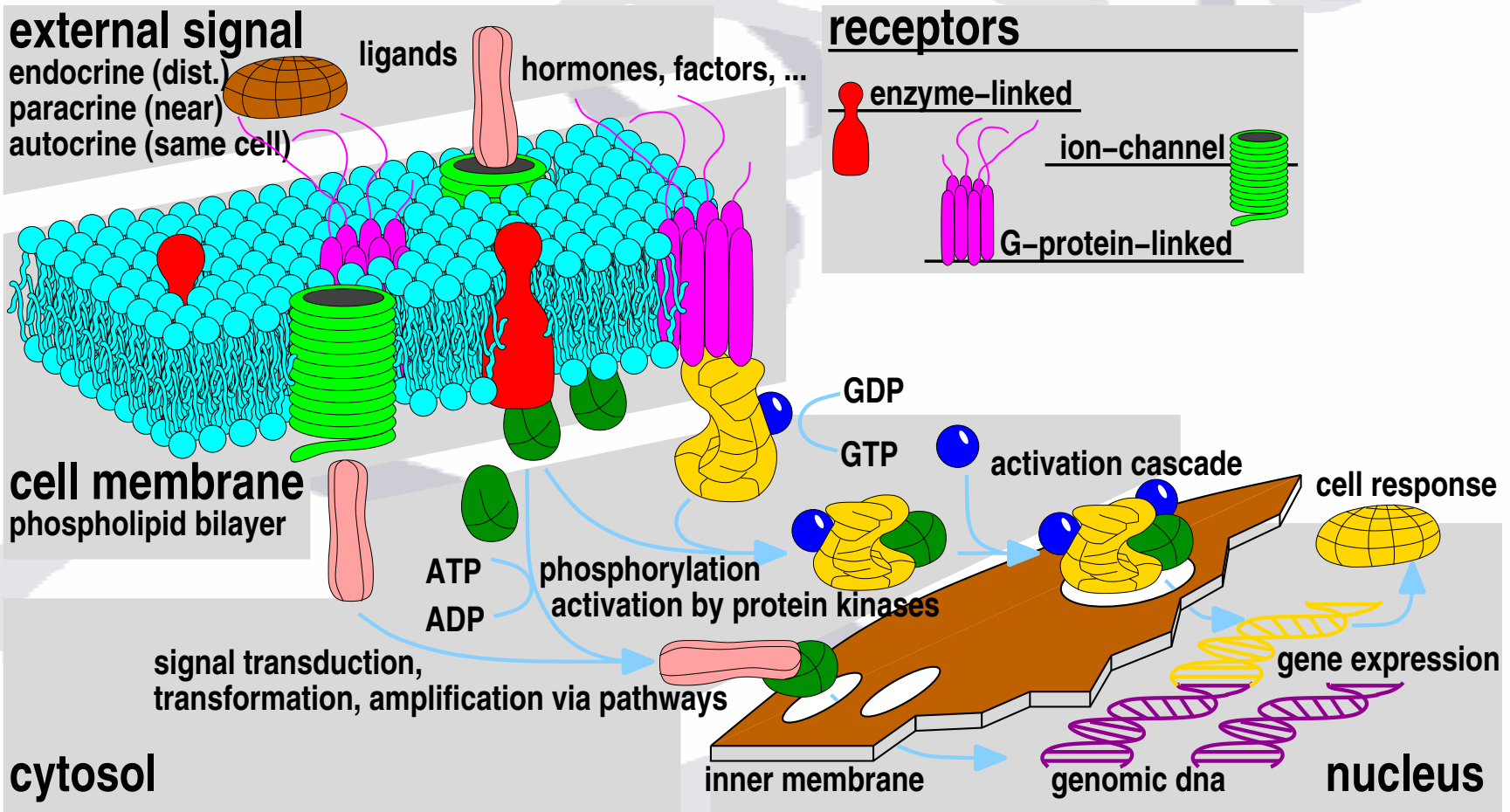
Bringing Membrane Systems Back to Biology

- Introduction
  - Biological Principles of Cell Signalling
  - ESIGNET – Research Project
  - Modelling Cell Signalling Networks
- A P System for Cell Signalling Networks
  - Motivation and Intention
  - System Definition
  - Matching and Matching Strategies
  - System Behaviour and Properties
- Example and Conclusion
  - Example: Yeast Pheromone Pathway
  - Conclusion and Future Work



Biological Principles of Cell Signalling

Information Processing in Living Cells



ESIGNET – Research Project

Evolving Cell Signalling Networks (CSNs) *in silico*

European interdisciplinary research project

- University of Birmingham (Computer Science)
- TU Eindhoven (Biomedical Engineering)
- Dublin City University (ALife Lab)
- University of Jena (Bio Systems Analysis)

SIXTH FRAMEWORK PROGRAMME

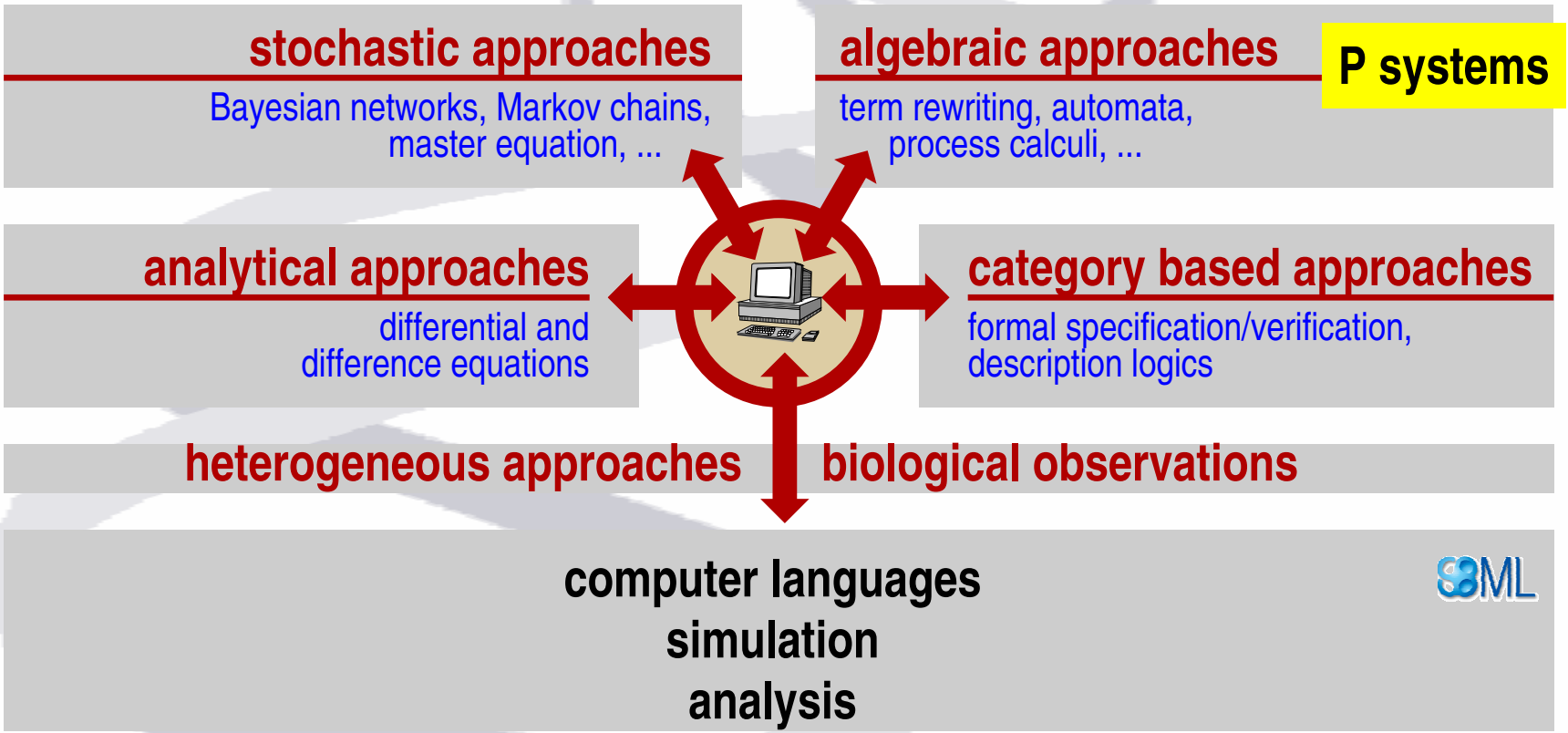
Objectives

- study the computational properties of CSNs
- developing new ways to model and predict real CSNs
- gain new theoretical perspectives on real CSNs

ESIGNET

Modelling Cell Signalling Networks

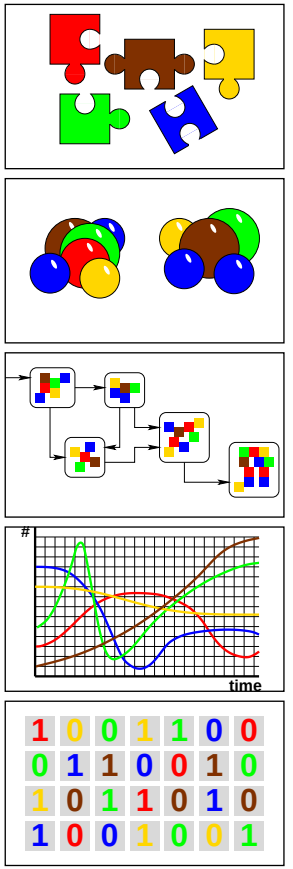
Phenotypic Description of CSNs – Structure, Behaviour, Function



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
  - 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

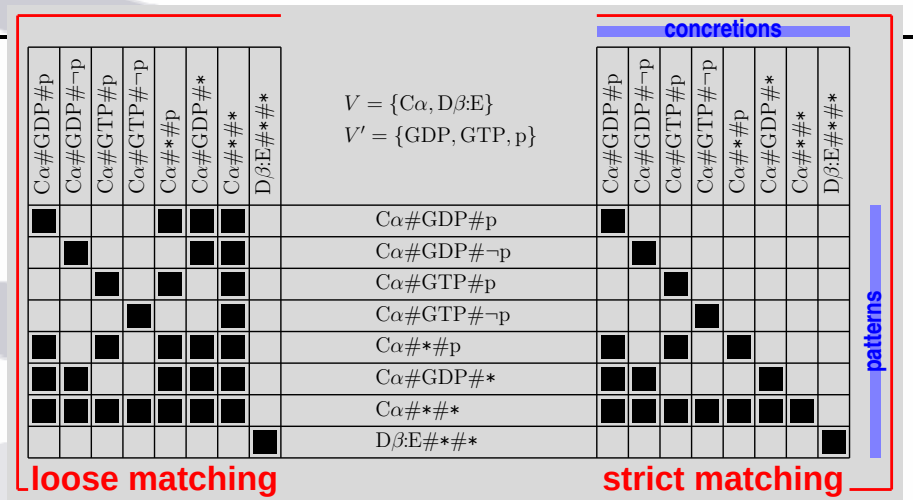
Matching and Matching Strategies

Comparison of String-Objects Describing Biomolecules

- String-based representation of proteins
  - **string-object** s: representation of a protein, its properties, substructure, binding domains, activation state, ligands
  - **structure**: $s \in V^+ \otimes (\{\#\} \otimes ((V')^+ \cup \{\neg\} \otimes (V')^+ \cup \{*\}))^*$
  - **meta-symbols**: **placeholder** (wild-card) $*$ $\longrightarrow$ appropriate or unknown substructure/property; **exclusion** $\neg$
  - test whether two string-objects identify same molecule
- Matching strategies

loose:

two string-objects match iff there is at least one common wild-card free representation

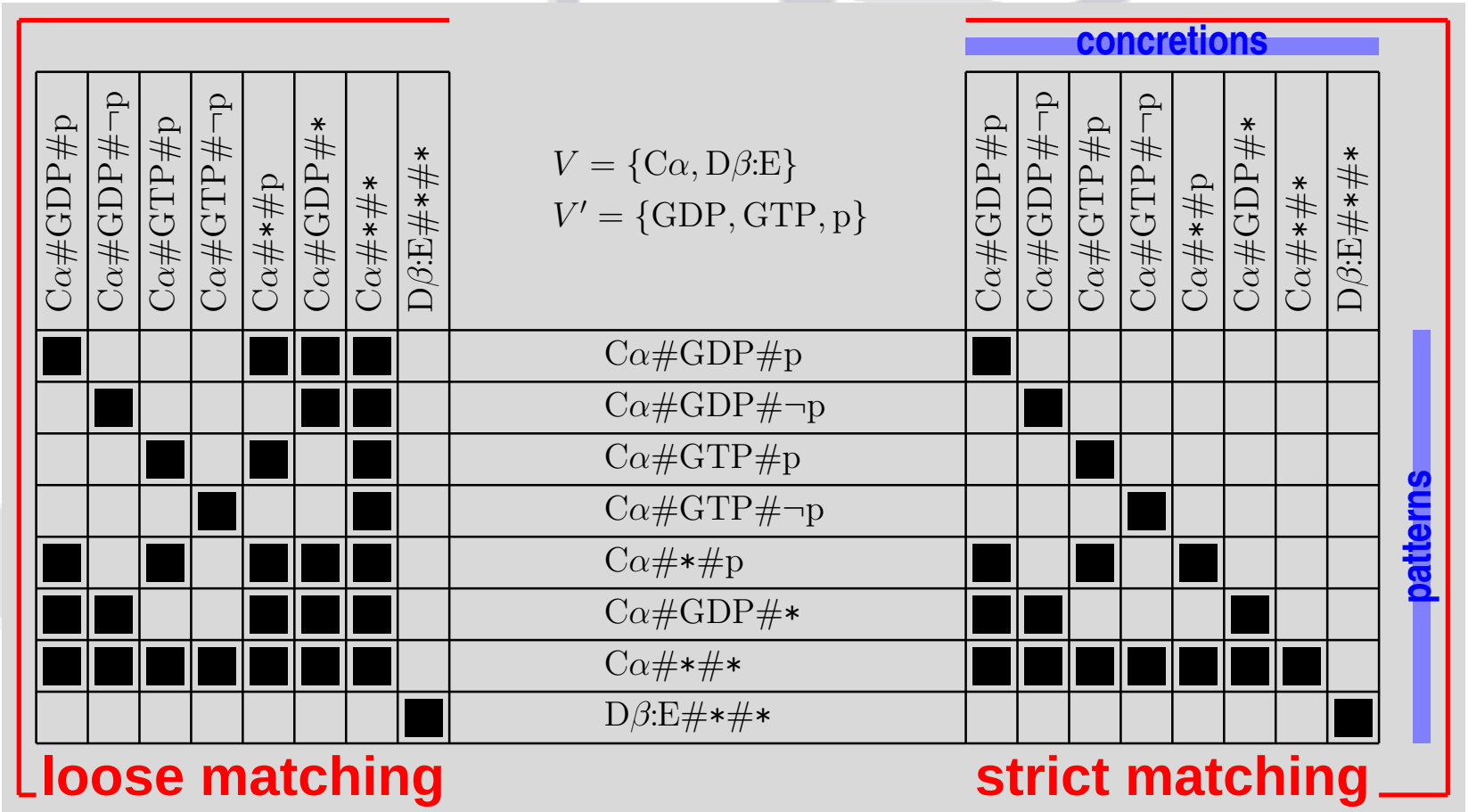


strict:

participating string-objects interpreted as a pattern and a candidate (concretion of the pattern)

Matching and Matching Strategies

Comparison of String-Objects Describing Biomolecules



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:

$$\begin{aligned} L_i(0) &= A_i \\ L_i'(t) &= L_i(t) \ominus \mathit{Educts}_i(t) \uplus \mathit{Products}_i(t) \\ L_i(t+1) &= L_i'(t) \ominus \mathit{Outgoing}_i(t) \uplus \mathit{Incoming}_i(t) \end{aligned}$$

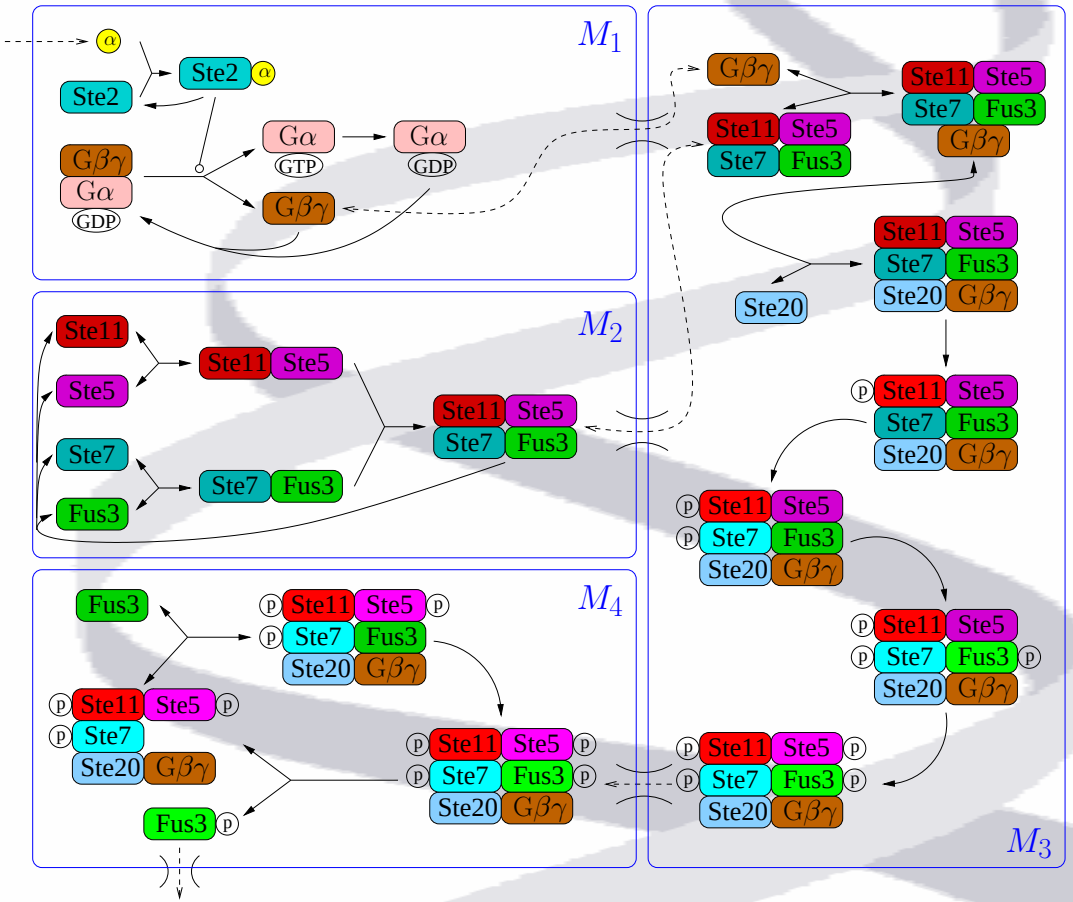
    1. Determine multiset of educts using $L_i(t)$, $R_{i1}, \ldots, R_{ir_i}, \mathrm{f}_{i1}, \ldots \mathrm{f}_{ir_i}$; involves matching
    2. Remove educt objects from module contents
    3. Determine and add multiset of reaction products, obtain $L_i'(t)$
    4. Determine and separate objects leaving 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

ESIGNET

Example: Yeast Pheromone Pathway

Signal Transduction in *Saccharomyces cerevisiae*

A well-known signalling system controlling mating behaviour



$$\Pi = (V, V', E, M, 4)$$

$$V = \{\mathrm{Ste2}, \alpha, \mathrm{G}\beta\gamma, \mathrm{G}\alpha, \ldots\}$$

$$V' = \{\mathrm{a}, \mathrm{GDP}, \mathrm{GTP}, \mathrm{p}\}$$

$$M = \{M_1, M_2, M_3, M_4\}$$

$$M_1 = (R_{11}, R_{12}, R_{13}, R_{14}, R_{15}, \mathrm{f}_{11}, \mathrm{f}_{12}, \mathrm{f}_{13}, \mathrm{f}_{14}, \mathrm{f}_{15}, A_1)$$

$$R_{11} = \mathrm{Ste2}\#\neg\mathrm{a} + \alpha \rightarrow \mathrm{Ste2}\#\mathrm{a}$$

$$R_{12} = \mathrm{Ste2}\#\mathrm{a} \rightarrow \mathrm{Ste2}\#\neg\mathrm{a}$$

$$\mathrm{f}_{11} = \lfloor \mathrm{k}_{11}[\mathrm{Ste2}\#\neg\mathrm{a}][\alpha]/\mathrm{V}_1^2 \rfloor$$

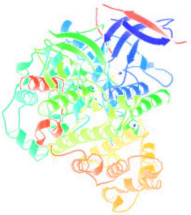
$$\vdots$$

ESIGNET

Conclusion and Future Work

Summary

- P systems provide a suitable framework for modelling CSNs
- Valuable insight can be gained from combining previously isolated ideas
- Our system takes first steps in this direction



Outlook

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

ESIGNET