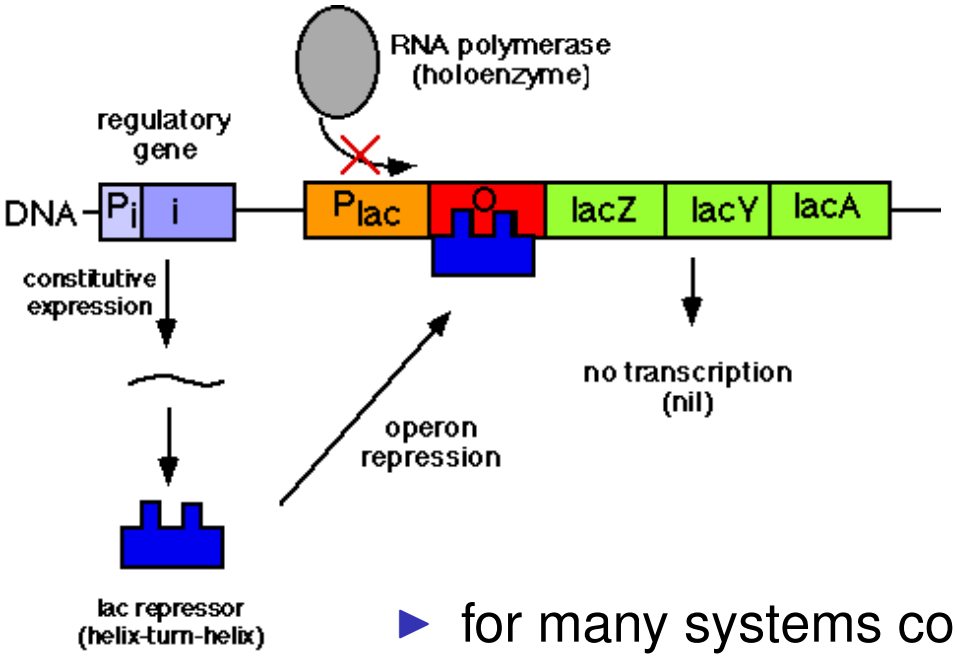
Cartoons and Text



"In the absence of lactose, the repressor protein encoded by the I gene binds to the lac operator and prevents transcription...."

- for many systems conceptual understanding only
  - lack of information on kinetic parameters, molecular concentrations, biochemical reaction mechanisms...
  - resulting systems of differential equations mostly not analytically solvable
- discrete modeling formalizes visual and verbal description and allows rigorous mathematical analysis

Discrete modeling

- system description by means of discrete functions
  - including structural information
  - capturing of interaction character and impact

  $\longrightarrow$ predicting/analyzing dynamics

Hypothesis: kinetic details of interactions less important than network organization

- degree of coarseness (Boolean, multivalued, hybrid)
- varied applications
  - derive structure and logic of networks from dynamics (reverse engineering)
  - modeling and analyzing (small) specific systems
  - studying properties of classes of networks

Boolean models

Ingredients:

- system with n components $v_1, \ldots, v_n$ interpreted as variable in $\{0,1\}$
- n functions $f_i : \{0,1\}^n \longrightarrow \{0,1\}$
- system description $f = (f_1, \ldots, f_n) : \{0,1\}^n \longrightarrow \{0,1\}^n$
  - f_i captures the rule to calculate the future value of v_i from the current values of its regulators

  $\longrightarrow$ f holds information on network structure

  $\longrightarrow$ f encodes system dynamics on state space $\{0,1\}^n$

Structure

Given $f : \{0,1\}^n \longrightarrow \{0,1\}^n$

Interaction graph $G(f)$ of f

- represents dependencies between components
- character of interactions
- possible context sensitivity (non-monotonicity) resulting in a multigraph

Consistency of f and $G(f)$

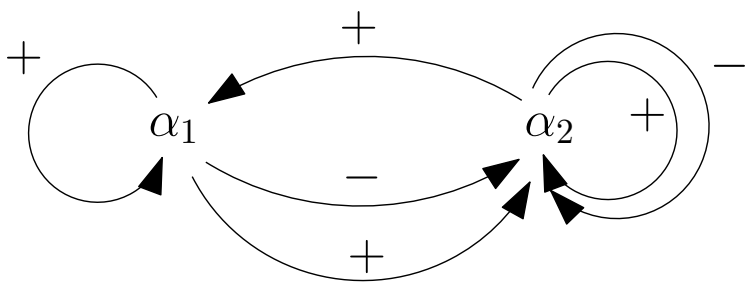
Structure

Given $f : \{0,1\}^n \longrightarrow \{0,1\}^n$

Interaction graph $G(f)$ of f

- represents dependencies between components
- character of interactions
- possible context sensitivity (non-monotonicity) resulting in a multigraph

$$f : \{0,1\}^2 \to \{0,1\}^2,$$
$$(x_1, x_2) \mapsto (x_1 \wedge x_2,\ (\overline{x_1} \wedge \overline{x_2}) \vee (x_1 \wedge x_2))$$



Consistency of f and $G(f)$

Consistency of f and $G(f)$

- function f should be consistent with the interaction graph
  - $f_i(x)$ only depends on x_j if α_j is predecessor of α_i
  - functionality of edges and sign consistency:
    $$\alpha_i \to \alpha_j: \quad \exists\, s \in \mathcal{B}^n \,:\, f_j(s) \neq f_j(\overline{s}^i) \quad \text{and} \quad \alpha_i \xrightarrow{+} \alpha_j \Leftrightarrow f_j(s) = s_i$$

Definition For $x \in \{0,1\}^n$ let $G(x)$ be the graph with vertices $\alpha_1, \ldots, \alpha_n$ and an edge $\alpha_j \to \alpha_i$ if $f_i(x_1, \ldots, x_j, \ldots, x_n) \neq f_i(x_1, \ldots, 1 - x_j, \ldots, x_n)$, with positive sign if $x_j = f_i(x)$ and negative otherwise. $G(x)$ is called *local interaction graph in* x. We call $G(f) := \bigcup_{x \in \{0,1\}^n} G(x)$ *global interaction graph of* f.

Consistency of f and $G(f)$

- function f should be consistent with the interaction graph
  - $f_i(x)$ only depends on x_j if α_j is predecessor of α_i
  - functionality of edges and sign consistency:
    $$\alpha_i \to \alpha_j: \quad \exists\, s \in \mathcal{B}^n \,:\, f_j(s) \neq f_j(\overline{s}^i) \quad \text{and} \quad \alpha_i \xrightarrow{+} \alpha_j \Leftrightarrow f_j(s) = s_i$$

Definition For $x \in \{0,1\}^n$ let $G(x)$ be the graph with vertices $\alpha_1, \ldots, \alpha_n$ and an edge $\alpha_j \to \alpha_i$ if $f_i(x_1, \ldots, x_j, \ldots, x_n) \neq f_i(x_1, \ldots, 1 - x_j, \ldots, x_n)$, with positive sign if $x_j = f_i(x)$ and negative otherwise. $G(x)$ is called *local interaction graph in* x. We call $G(f) := \bigcup_{x \in \{0,1\}^n} G(x)$ *global interaction graph of* f.

- $G(x)$ is a graph, $G(f)$ may be a multi-graph
- $G(x)$ is a graph representation of the discrete Jacobian matrix of f in x
- consider behavior in $B_{d_H}(x,1)$ with $d_H(x,y) := \sum_{i=1}^n |x_i - y_i|$ (Hamming distance)
- $G(f)$ represents the *functional* and sign consistent network topology of f

Modeling

Modeling specific systems often starts with structural information:

- translate data on biochemical interactions into directed, signed (multi)graphs

Definition An *interaction (multi-)graph* (or wiring diagram) I is a labeled directed multigraph with vertex set $V := \{\alpha_1, \ldots, \alpha_n\}$, $n \in \mathbb{N}$, and edge set $E \subseteq V \times V \times \{+, -\}$. In Boolean models, vertices are understood as $\{0,1\}$-variables.

vertex: component (genes, proteins, chemical complexes,...), set of components (similar function, identification {gene, RNA, protein},...), signal,...

edge: inhibiting/activating activity (TFs, enzymes,...), complex forming, information flow,...

value: activity status, concentration, configuration,...

- translate behavioral rules into Boolean function $f : \{0,1\}^n \longrightarrow \{0,1\}^n$
  - for each component decide the impact of its predecessors in a given state on its value $\longrightarrow$ choice of parameters

Consistency: $I = G(f)$

Exploring the Structure

Consider interaction graph $I = (V, E)$

Use graph theoretical characteristics and measures

- quantify organizational features of I
  - importance of nodes
  - reachability among nodes
  - homogeneity/heterogeinity w.r.t. a given property
- relate to biological features
  - robustness, sensitivity, control,...
  - identify modules with characteristic function

Degree distribution and clustering

- **degree** of a node: # of edges originating (**outdegree**) or ending (**indegree**) in the node
- **hubs:** highest degree nodes
- **degree distribution** $P(k)$: fraction of nodes with degree k (indegree/outdegree distribution)
  [cellular networks are often scale-free]
- **neighborhood** of a node v: set of nodes $\neq v$ adjacent to v (in/out-neighborhood)
- **clique**: completely connected subgraph
- **clustering coefficient** of a node: ratio of # of edges in neighborhood and # of edges if neighborhood were a clique
  [large average clustering coefficients indicate redundancy, cohesiveness; observed in protein-protein interaction and metabolic networks]

Paths and connectivity

- **distance** between two nodes: shortest path length connecting the nodes
- **small world**: average shortest path length of large networks stays small
  [facilitates rapid spread of information in response to input; signal transduction, protein interaction, metabolic networks]
- **path redundancy** [robustness]
- **betweenness centrality** of node v: ratio of # shortest paths from s to t through v and total # of shortest st-paths
  [importance of a node in flow from sources to sinks]
- **connectivity** of the network: existence of paths between every pair of nodes (distinguish directed/undirected graphs)
  - **strongly connected** directed graphs: all node pairs connected in both directions
  - **strongly connected components:** maximal subgraphs that are strongly connected $\longrightarrow$ acyclic scc-graph with initial and terminal components
    [modularity of signaling networks]

Modules and motifs

- **modules** of a network: subnetworks distinguishable by dense intra-module and sparse inter-module connectivity
  - identification should include biological characteristics: physical location, function, evolutionary conservation
  - difficulties: cross-talk, overlap, hierarchical modularity
- **motifs** of a network: significant, small-subgraphs of well-defined topology
  - e.g. feedback loops, feedforward loops, cascades,...
  - statistical importance (U. Alon et al.)
  - classification and comparison of networks

Modules and motifs

- **modules** of a network: subnetworks distinguishable by dense intra-module and sparse inter-module connectivity
  - identification should include biological characteristics: physical location, function, evolutionary conservation
  - difficulties: cross-talk, overlap, hierarchical modularity
- **motifs** of a network: significant, small-subgraphs of well-defined topology
  - e.g. feedback loops, feedforward loops, cascades,...
  - statistical importance (U. Alon et al.)
  - classification and comparison of networks

Remark: structural analysis yields also dynamical information, but is generally not sufficient for understanding corresponding dynamical systems

Dynamics

Given $f : \{0,1\}^n \to \{0,1\}^n$

State transition graph $S(f)$ of f

- vertex set $\{0,1\}^n$ (state space)
- edge set $\{(x, f(x)) \mid x \in \{0,1\}^n\}$ – synchronous update

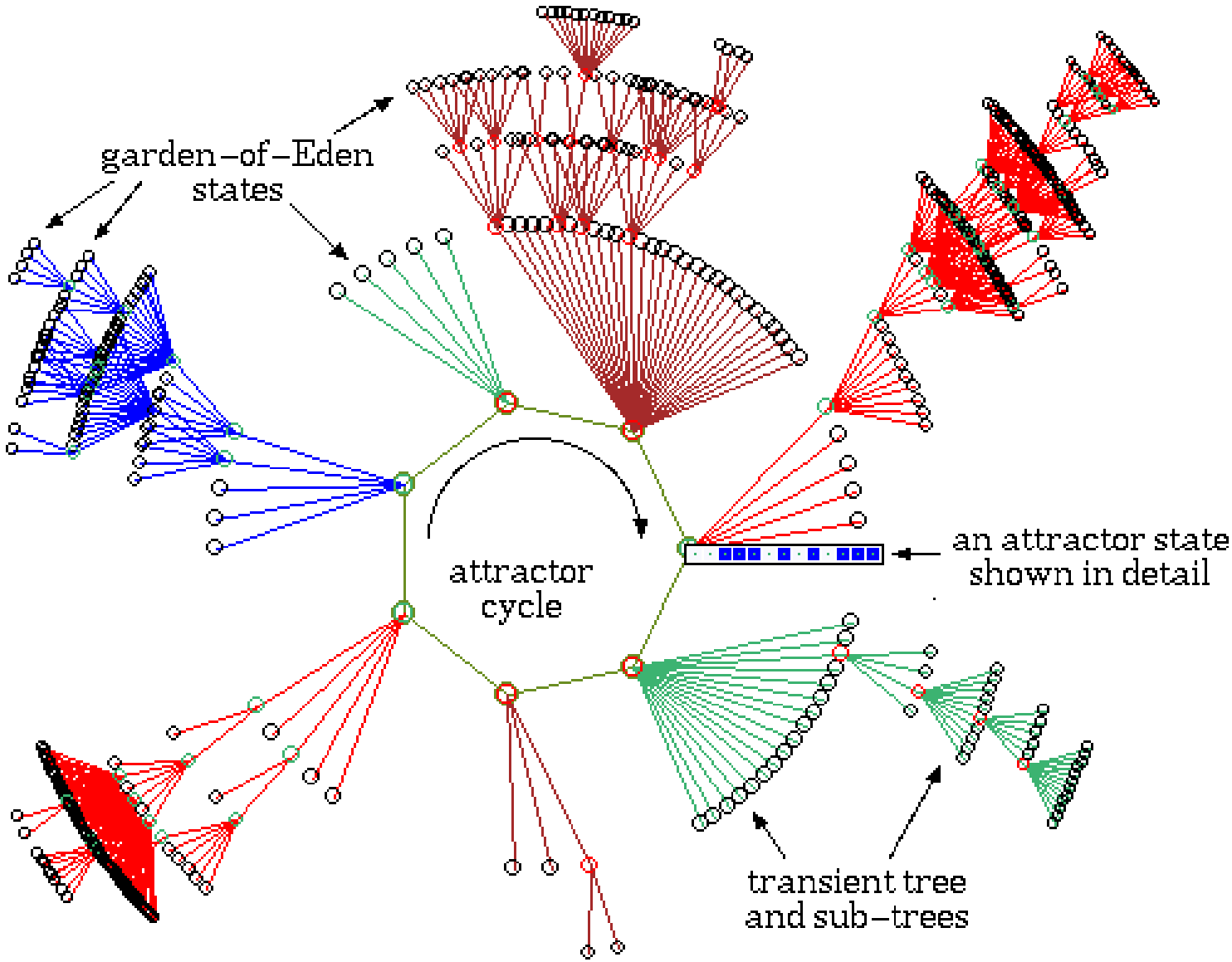
Trajectories: infinite paths $(x(0), x(1), \ldots)$ in $S(f)$ (simulation)

Note: conceptual differences to ODE/PLDE description

- explicit description of trajectories
- trajectories can merge

Consequences of synchronous update and finite state space

- deterministic behavior
- each trajectory ends in a cycle
- components of $S(f)$ consist of single cycle and attached trees



Andy Wuensche, www.ddlab.com

Attractors

Given state transition graph $S(f)$

Definition A set A of vertices (states) of $S(f)$ is called trap set, if no trajectory starting in A can leave A. If in addition A is strongly connected, then A is called attractor.

- attractors are terminal strongly connected components
- attractors are fixed points and periodic points
- every trajectory leads to an attractor (basins of attraction)
- distinct attractors are disjoint
- asymptotical behavior (biological meaningful)

Perturbations

Minimal perturbation (noise): transiently flipping the value of a component

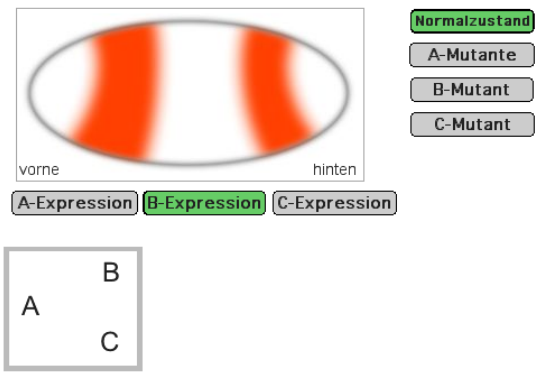
- comparison of different initial conditions
- how does the change cascade through the network?
- change in basin of attraction/attractor

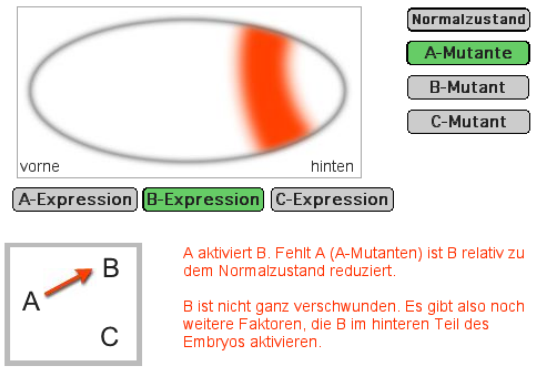
Structural perturbation (mutation): permanently changing a coordinate function f_i

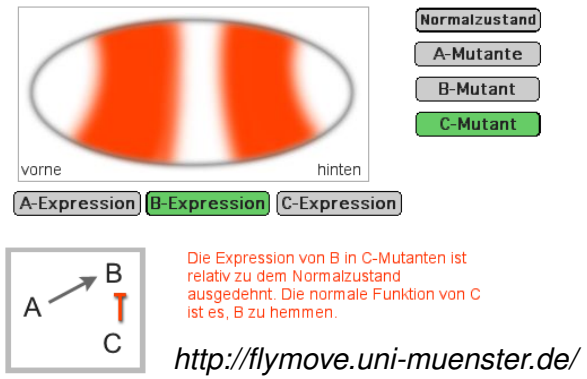
- comparison of two different networks
- attractors, basins of attraction, stability,...

Network Inference - Reverse Engineering

- analyzing binding sites and mutants







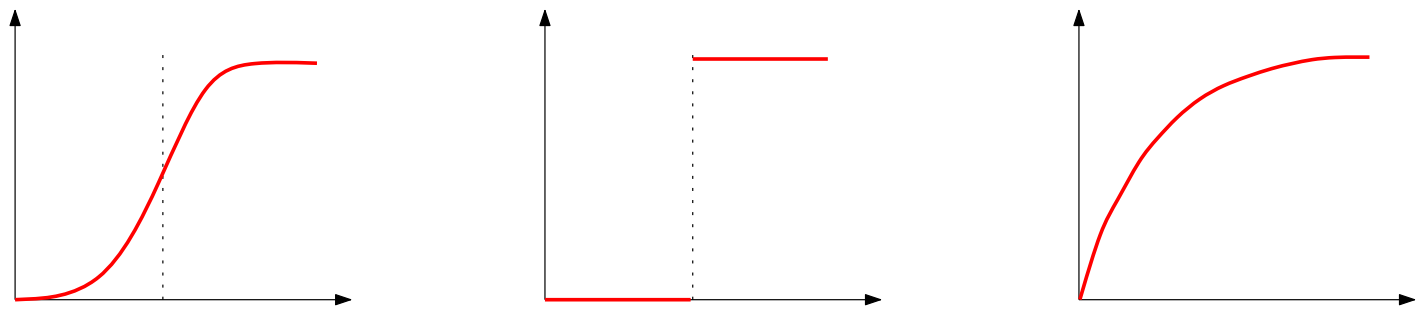
http://flymove.uni-muenster.de/

- time series data
  - data discretization
  - often many admissible models

Inferring interaction graphs: Given a function $f : \{0,1\}^n \longrightarrow \{0,1\}^n$, we can derive an interactions graph consisting of functional edges in agreement with the dynamics determined by f by using the previously introduced formulas describing functionality of edges and sign consistency.

Being Aware of the Level of Abstraction

- omitting components, simplifying processes
- logical idealization of regulatory interactions



- all or nothing functionality
- ignoring spatial and temporal data