

Complexity: Lessons learned from Networks

From static to dynamical networks

Markus Kirkilionis, University of Warwick

August 4, 2006

Overview

Graph Theory - Network Topologies

- Basic definitions

- Algebraic representations

- Motifs & Data

- Stochastic processes defined on network topologies

Attaching state spaces to graphs

- The state space choice

- Change of state

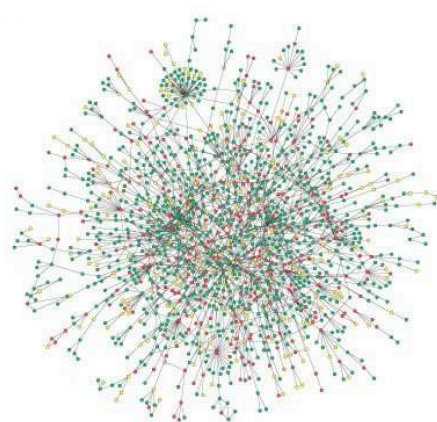
Time-continuous dynamical systems on networks

- Lotka–Volterra Systems

- Reaction schemes

- Analysis of reaction networks

- Representing reaction schemes by graphs



We start very simply by introducing the concept of network topologies, i.e. we exclude any temporal dynamics at the moment. Essentially we consider pure graph theory.

A graph G (or network) is a set of points or nodes (always called vertices in graph theory) V in space which are interconnected by a set E of lines or links (always called edges in graph theory), i.e. $G = (V, E)$.

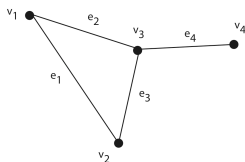


Figure: Example of an undirected graph G .

Here $G = (\{v_1, v_2, v_3, v_4\}, \{e_1, e_2, e_3, e_4\})$. Let $|S|$ denote the number of elements of a set S . If both $|V| < \infty$ and $|E| < \infty$ then the graph G is called finite. For any edge e joining the vertices v_i and v_j we set $e = (v_i, v_j)$. Because so far the vertices have no direction, we have also $e = (v_j, v_i)$. In this case G is called *undirected*.

An undirected graph describes a binary relationship between a set of vertices.

- ▶ If an edge e has v as an end-point, then e is called incident with v .

A very important concept:

Definition

The degree of a vertex v , written $d(v)$, is the number of edges incident with v .

An undirected graph describes a binary relationship between a set of vertices.

- ▶ If an edge e has v as an end-point, then e is called incident with v .
- ▶ If $(u, v) \in E$, u is said to be adjacent to v .

A very important concept:

Definition

The degree of a vertex v , written $d(v)$, is the number of edges incident with v .

An undirected graph describes a binary relationship between a set of vertices.

- ▶ If an edge e has v as an end-point, then e is called incident with v .
- ▶ If $(u, v) \in E$, u is said to be adjacent to v .
- ▶ Two edges are adjacent if they have a common end-point.

A very important concept:

Definition

The degree of a vertex v , written $d(v)$, is the number of edges incident with v .

Let $k := d(v)$ and $P(k)$ be the degree frequency (which can be interpreted as a degree probability distribution) in different types of graphs.

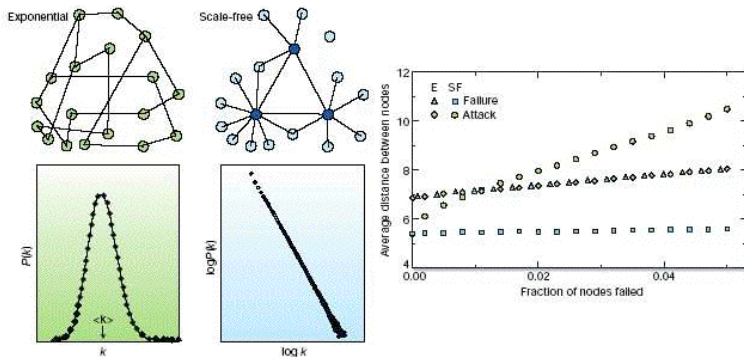


Figure: Examples of different types of degree distributions.

The degree probability distribution is related to the famous 'Small World Property' (SWP)

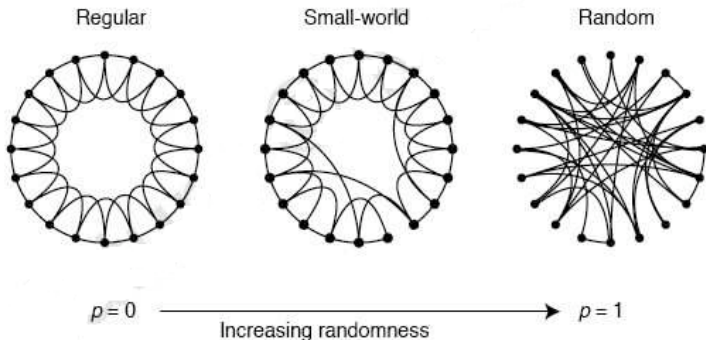
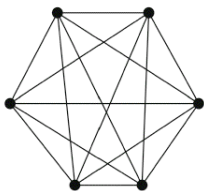


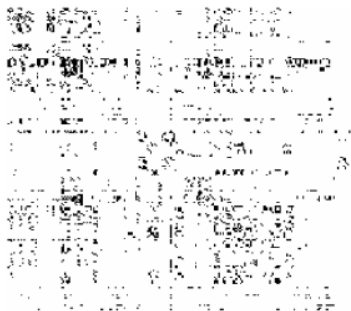
Figure: Degree distributions can be interpreted to be derived from random processes deleting and re-establishing edges.

Graphs can be analysed algebraically (a very fruitful concept!) by introducing adjacency and incidence matrices. In a graph $G = (V, E)$ define an $(n \times n)$ symmetric matrix $A = (a_{ij})$ by $a_{ij} = 1$ if $(v_i, v_j) \in E$ and zero otherwise. This adjacency matrix encodes all information on G . Powers of A can be used to calculate the number of paths between vertices. The coefficients of the characteristic polynomial of A encode information on the number of edges, triangles (the number of times the complete subgraph K_3 with 3 vertices occurs in G) etc.

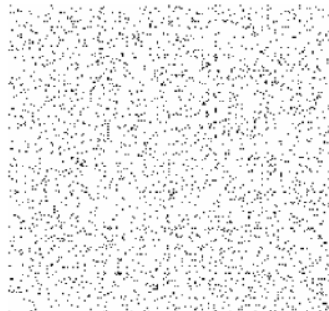


$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

The adjacency matrix can even be analysed visually: Comparison of a random graph with the *C. elegans* neuronal network.

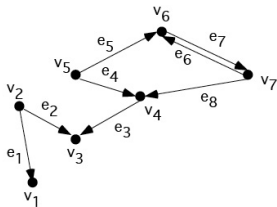


C. Elegans network



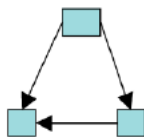
Random Network

After giving the graph $G = (V, E)$ an *orientation* (of the edges) the result is a directed graph \vec{G} . Instead of a binary relationship such a graph can model hierarchical relations, like 'x is more dominant than y', 'x is influencing y', 'x regulates y', etc. The adjacency matrix A of \vec{G} becomes in general non-symmetric. We can define an $(n \times m)$ incidence matrix $D = (d_{ij})$ by $d_{ij} = +1$ if v_i is the positive end of e_j , $d_{ij} = -1$ if v_i is the negative end of e_j , and zero otherwise. The incidence matrix D resulting from giving an arbitrary orientation to G has rank $n - c$, where c is the number of components of G .

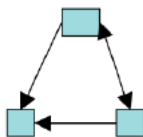


A motif is a specified subgraph contained in a graph G . The analysis of motif frequency has been a major route of investigation for all applications.

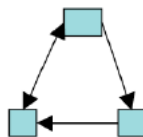
The *C. elegans* neuronal network is a simple directed network, i.e. can be based on the knowledge that synaptic transmission is unilateral. Shown are frequencies of certain configurations (from Johnson et al.).



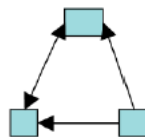
$N = 197$



$N = 1022$



$N = 92$



$N = 96$

Data for network structure:

- ▶ Genetics: For example knock-out experiments.

Data for network structure:

- ▶ Genetics: For example knock-out experiments.
- ▶ Metabolomics: 'binding', 'reacting' as binary relationships.

Data for network structure:

- ▶ Genetics: For example knock-out experiments.
- ▶ Metabolomics: 'binding', 'reacting' as binary relationships.
- ▶ Neuronal networks: imaging ('Geometric network')

Data for network structure:

- ▶ Genetics: For example knock-out experiments.
- ▶ Metabolomics: 'binding', 'reacting' as binary relationships.
- ▶ Neuronal networks: imaging ('Geometric network')
- ▶ Ecology: trophic relations, patch connectivity etc.

Data for network structure:

- ▶ Genetics: For example knock-out experiments.
- ▶ Metabolomics: 'binding', 'reacting' as binary relationships.
- ▶ Neuronal networks: imaging ('Geometric network')
- ▶ Ecology: trophic relations, patch connectivity etc.
- ▶ Economy: information, resources

From a mathematical perspective the stochastic processes defined on graphs and leading to different degree distributions must have the following ingredients:

- ▶ A selection rule to create one event or a chain of (conditional) events (vertex selection, edge selection, growing/static/shrinking number of vertices/edges)

Here is an example of a probability with which after the selection of a vertex v_i in a graph G with n vertices (at time t) $(n - 1)$ new edges $e_j = (v_i, v_j)$, $j = 1, \dots, n - 1$, (possibly after relabeling) are created from v_i to every other vertex in the graph. Let $k_j = d(v_j)$.

$$P_{ij} = \frac{(1 - p)k_j + p}{\sum_{j=1}^{n-1} (1 - p)k_j + p}, \quad 1 \leq p \leq 1.$$

From a mathematical perspective the stochastic processes defined on graphs and leading to different degree distributions must have the following ingredients:

- ▶ A selection rule to create one event or a chain of (conditional) events (vertex selection, edge selection, growing/static/shrinking number of vertices/edges)
- ▶ A probability with which each of these events occur.

Here is an example of a probability with which after the selection of a vertex v_i in a graph G with n vertices (at time t) $(n - 1)$ new edges $e_j = (v_i, v_j)$, $j = 1, \dots, n - 1$, (possibly after relabeling) are created from v_i to every other vertex in the graph. Let $k_j = d(v_j)$.

$$P_{ij} = \frac{(1 - p)k_j + p}{\sum_{j=1}^{n-1} (1 - p)k_j + p}, \quad 1 \leq p \leq 1.$$

The next mathematical possibility is to attach state spaces to the (static or evolving) graph, either to the vertices or to the edges. Let X be a finite set (i.e. X is either the set of vertices V or edges E) and $f : X \rightarrow K$ with $K = \mathbb{N}, \mathbb{R}, \mathbb{C}$. Usually the set of all such functions f form a vector space, i.e. the length of roads (the traveling salesman) usually adds up etc. But this depends on the model. The general idea of introducing state spaces is to attach numerical values to the relationships, or to make these relationships depending on some quantitative measure of the node (like size, richness, status etc.)

Generalisation: Vector-valued or function state spaces defined on K .

Data: ... (much more difficult, quantification of relationships)

Change of discrete state spaces via probabilistic models

There are two types of possible evolutions (changes in time) on networks with state spaces.

- ▶ The evolution (or construction) of the graph structure, i.e. the network topology. This network evolution can now depend on the edge and vertex states.



Change of discrete state spaces via probabilistic models

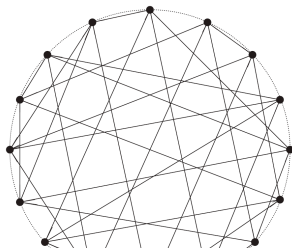
There are two types of possible evolutions (changes in time) on networks with state spaces.

- ▶ The evolution (or construction) of the graph structure, i.e. the network topology. This network evolution can now depend on the edge and vertex states.
- ▶ The evolution of the state spaces.



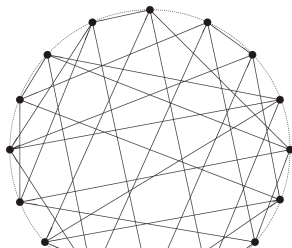
As an example of state space evolution we take a simple 'opinion' *YNY* model (equivalent to an *SIS* model of the spread of an infectious disease) on a fixed graph (created as a homogeneous random graph).

- ▶ Each vertex (individual, agent) can have two states, either to say 'yes (Y)' or 'no (N)'.



As an example of state space evolution we take a simple 'opinion' YNY model (equivalent to an SIS model of the spread of an infectious disease) on a fixed graph (created as a homogeneous random graph).

- ▶ Each vertex (individual, agent) can have two states, either to say 'yes (Y)' or 'no (N)'.
- ▶ The events are defined as $Y \rightarrow N$ and $N \rightarrow Y$ and respectively occur with rates δ and $k\lambda$, where k is the number of neighbours having opinion 'Y'.



The rate of change of an average quantity ϕ (such as the fraction of sites in a particular state) is described as

$$\dot{\phi} = \sum_{v \in V} \sum_{e_v \in E_v} r(e_v)(\phi_{e_v} - \phi)$$

where V is the set of all vertices (agents), and E_v represents the set of all events that can occur at v . A particular event e_v changes the average from ϕ to ϕ_{e_v} and occurs at rate $r(e_v)$.

Under a 'mean-field hypothesis' (i.e. every agent feels the influence of opinion of an average neighbour) we can establish ODE describing the time-evolution of the fraction of agents with opinion 'Y' (described by ρ_1) and opinion 'N' (described by ρ_0). We have

$$\dot{\rho}_1 = -\delta\rho_1 + \lambda k\rho_1\rho_0,$$

where of course $\rho_0 = 1 - \rho_1$.

This 'mean-field hypothesis' is of course invalid, and the approximation of $\dot{\rho}_1$ can be improved by looking at the fraction of neighbouring pairs of agents which have the same or different opinion. Let for example ρ_{11} be the fraction of randomly chosen edges such that both incident vertices (agents) are in configuration 'YY'. Then

$$\dot{\rho}_1 = -\delta\rho_1 + \lambda k\rho_{10}\rho_{0|1}$$

and

$$\dot{\rho}_{11} = -2\rho_{11} + 2\lambda\rho_{10}\rho_{0|1} + 2\lambda(k-1)\rho_{1|01}\rho_{10}$$

Here the $\rho_{0|1}$ is conditional probability (can be interpreted as a local density) that if an agent is of state 'N' there is a neighbouring agent in state 'Y', etc.

Data rate-driven models:

- ▶ Genetics: time-scales of transcription and degradation.

Data rate-driven models:

- ▶ Genetics: time-scales of transcription and degradation.
- ▶ Metabolomics: reaction probabilities.

Data rate-driven models:

- ▶ Genetics: time-scales of transcription and degradation.
- ▶ Metabolomics: reaction probabilities.
- ▶ Neuronal networks: spike frequencies.

Data rate-driven models:

- ▶ Genetics: time-scales of transcription and degradation.
- ▶ Metabolomics: reaction probabilities.
- ▶ Neuronal networks: spike frequencies.
- ▶ Ecology: trophic encounters, patch transition

Data rate-driven models:

- ▶ Genetics: time-scales of transcription and degradation.
- ▶ Metabolomics: reaction probabilities.
- ▶ Neuronal networks: spike frequencies.
- ▶ Ecology: trophic encounters, patch transition
- ▶ Economy: information, resource distribution rates.

The next (and last possibility discussed) is that we consider a graph G (assumed static like in the previous examples) with continuous state spaces attached to the vertices. We assume this state space is modelling a density of some entity, like the density of people of a certain type occurring in a country etc. Furthermore we assume this density changes deterministically and continuously in time (by assuming a law of large numbers). For our example this means there are continuously transitions (collisions, changes of type) between a large number of people of these different types. This leads to a large system (depending on the number of different types) of coupled ODE. The best studied prototype for n such types is

$$\dot{x}_i = (r_i + \sum_{j=1}^n a_{ij}x_j)x_i.$$

We interpret $A = (a_{ij})$ as a weighted adjacency matrix.

We can apply all the framework developed for time-continuous dynamical systems, i.e. determining attractors, invariant sets, change of attractors (bifurcation theory), etc. This type of model leads to a number of interesting questions:

- ▶ What is the connection between the structure of the underlying graph (network topology) and the type of attractors of the network?

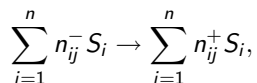
We can apply all the framework developed for time-continuous dynamical systems, i.e. determining attractors, invariant sets, change of attractors (bifurcation theory), etc. This type of model leads to a number of interesting questions:

- ▶ What is the connection between the structure of the underlying graph (network topology) and the type of attractors of the network?
- ▶ What is the robustness of the different types in the network?

We can apply all the framework developed for time-continuous dynamical systems, i.e. determining attractors, invariant sets, change of attractors (bifurcation theory), etc. This type of model leads to a number of interesting questions:

- ▶ What is the connection between the structure of the underlying graph (network topology) and the type of attractors of the network?
- ▶ What is the robustness of the different types in the network?
- ▶ etc.

Let there be r reactions $\mathbf{R} = (R_1, \dots, R_r)$, and different types (species) $\mathbf{S} = \{S_1, \dots, S_n\}$. A single *reaction* is defined by



with n_{ij}^- and n_{ij}^+ being the stoichiometric coefficients of the forward and backward reactions. The coefficients

$$n_{ij} := n_{ij}^+ - n_{ij}^-$$

define the stoichiometric matrix N . The law-of-mass-action induces the function

$$v_j : \mathbb{R}_+^n \times P \rightarrow \mathbb{R}_+$$

which is the reaction velocity depending on concentrations $\mathbf{x} = (x_1, \dots, x_n)$ and the parameter $p = (k_1, \dots, k_n)$. The triple $T = (\mathbf{S}, \mathbf{R}, v)$ is called a *reaction scheme*. Dynamics:

$$\dot{\mathbf{x}} = Nv(\mathbf{x}, p)$$

The molecularity of the type S_i in reaction (transition) R_j is encoded in the *kinetic exponent* κ_{ij} . All these exponents are assembled in the *kinetic matrix* κ . These exponents can be defined by

$$\kappa_{ij} = \frac{\partial \log(v_j(x_0, k_j))}{\partial \log(x_{0,i})}.$$

With this definition in most cases the n_{ij}^- equal the κ_{ij} . Each reaction rate can be written as

$$v_j(x, k_j) = k_j \prod_{i=1}^m x^{k_{ij}}.$$

All these monomials form the reaction vector v we have already encountered.

Is there a dissection of the nonlinear network behaviour? This must be a central idea of any complexity theory based on networks. Transition schemes: consider the stationary states $Nv = 0$. Transformation into reaction coordinates yields the following expression for the Jacobian J :

$$J = N \operatorname{diag}(v) \kappa^t \operatorname{diag}(h_0)$$

with κ being the kinetic matrix and $h_0 = x_0^{-1}$ is the inverse of stationary type concentrations.

The intersection of the set of stationary reaction rates with the kernel of N induces a convex polyhedral cone.

$$K_v = \{v \in \mathbb{R}^r \mid Nv = 0, v \geq 0\} = (\ker(N)) \cap \mathbb{R}_+^r \\ = \sum_{i=1}^t j_i E_i, j_i > 0 \forall i\}$$

The minimal number of generating vectors E_i are called *extremal currents*. All equilibrium reactions can be written in the form $v(j) = \sum_{i=1}^t j_i E_i$.

Substitution of the extremal currents into the Jacobian yields

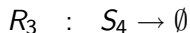
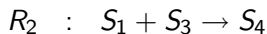
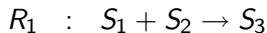
$$\tilde{J}(v) = N \operatorname{diag}\left(\sum_{i=1}^t j_i E_i\right) \kappa^t \operatorname{diag}(h_0).$$

With the implicit definition $J(v) = \tilde{J}(j) \operatorname{diag}(h_0)$ the local asymptotic stability of every equilibrium can be obtained by summing up the stability properties of the different extremal currents:

$$\tilde{J}(j) \operatorname{diag}(h_0) = j_1 \tilde{J}(E_1) \operatorname{diag}(h_0) + \dots + j_t \tilde{J}(E_t) \operatorname{diag}(h_0).$$

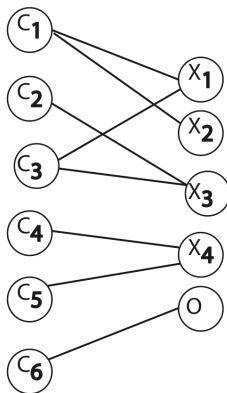
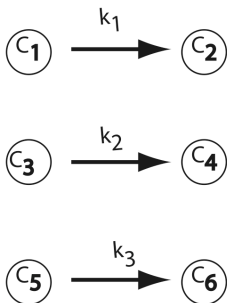
Extremal currents can be stable, unstable, or mixing stable (Clarke, definition based on local Lyapunov functions). A stable and mixing stable extremal current is called positive loop. Unstable or not mixing stable extremal currents are called stoichiometric generators. They lead to complex dynamical behaviour.

Consider some simple reaction scheme like



The six complexes are $C_1 = (1, 1, 0, 0)^t$, $C_2 = (0, 0, 1, 0)^t$, etc.

The reaction network can be represented by two graphs, a weighted directed graph to describe the reactions, and a bipartite undirected graph to describe the reactants. In our example we get:



For the directed graph the respective adjacency and incidence matrices are of importance. The first one, I_a contains the information whether the complex is the initial (entry -1) or the end vertex of an edge (entry 1). This entry distinguishes reactant complexes from product complexes. The second one, I_k contains nonzero entries only for initial vertices, i.e. for reactant complexes. The entries are the weight of the corresponding edge, which is the rate constant k_l .

The reaction scheme $\dot{x} = Nv$ can now be rewritten in the form

$$\dot{x} = YI_a I_k \Psi(x),$$

so $N = YI_\alpha$ and $v = I_k \Psi(x)$. In other words the nonlinearities of the dynamical system defining the reaction scheme can be investigated in terms of an incidence and adjacency matrix, and a vector of monomials.

The modularisation of the nonlinear dynamical system allows for an intelligent bifurcation analysis:

- ▶ The set of stationary solutions depends on a certain parametrization. This is a set of the kinetic reaction constants k_i , if one considers the stationary concentrations, or it is a set of convex coordinates j_i if one considers the stationary reaction rates, or it is a mixing of both sets, if one considers the curve coordinates defined by the toric variety. The stability results are not affected by the coordinate system, neither by the parametrization.

The modularisation of the nonlinear dynamical system allows for an intelligent bifurcation analysis:

- ▶ The set of stationary solutions depends on a certain parametrization. This is a set of the kinetic reaction constants k_i , if one considers the stationary concentrations, or it is a set of convex coordinates j_i if one considers the stationary reaction rates, or it is a mixing of both sets, if one considers the curve coordinates defined by the toric variety. The stability results are not affected by the coordinate system, neither by the parametrization.
- ▶ A bifurcation is the reaction of a system from one dynamical state to another. This may be the reaction from a hyperbolic equilibria to an oscillatory motion or it may be a change in the number of steady states.

Data for dynamical networks:

- ▶ Besides all event rates additional knowledge on non-linearities is needed.

Data for dynamical networks:

- ▶ Besides all event rates additional knowledge on non-linearities is needed.
- ▶ Example: 'integrate&fire'-dynamics as an approximation of Hodgkin-Huxley.

Data for dynamical networks:

- ▶ Besides all event rates additional knowledge on non-linearities is needed.
- ▶ Example: 'integrate&fire'-dynamics as an approximation of Hodgkin-Huxley.
- ▶ etc.

Discussion and Extensions

- ▶ There are various extensions...(dynamical systems on graph evolutions..., stochasticity of reactions)

Discussion and Extensions

- ▶ There are various extensions...(dynamical systems on graph evolutions..., stochasticity of reactions)
- ▶ Methods for dynamical networks still under-developed. Special algebraic structures often used.

Discussion and Extensions

- ▶ There are various extensions...(dynamical systems on graph evolutions..., stochasticity of reactions)
- ▶ Methods for dynamical networks still under-developed. Special algebraic structures often used.
- ▶ On top of the dynamical networks we have to define strategy sets, utility and cost functions influencing the graph evolution and the reaction rates. This allows us to define equilibria with a game theoretic interpretation.

Discussion and Extensions

- ▶ There are various extensions...(dynamical systems on graph evolutions..., stochasticity of reactions)
- ▶ Methods for dynamical networks still under-developed. Special algebraic structures often used.
- ▶ On top of the dynamical networks we have to define strategy sets, utility and cost functions influencing the graph evolution and the reaction rates. This allows us to define equilibria with a game theoretic interpretation.
- ▶ The role of stochasticity neglected for the 'reactions'.