The role of self-sustained dynamical activity in generalized neural nets

Claudius Gros

Institute for Theoretical Physics
J.W. Goethe University Frankfurt

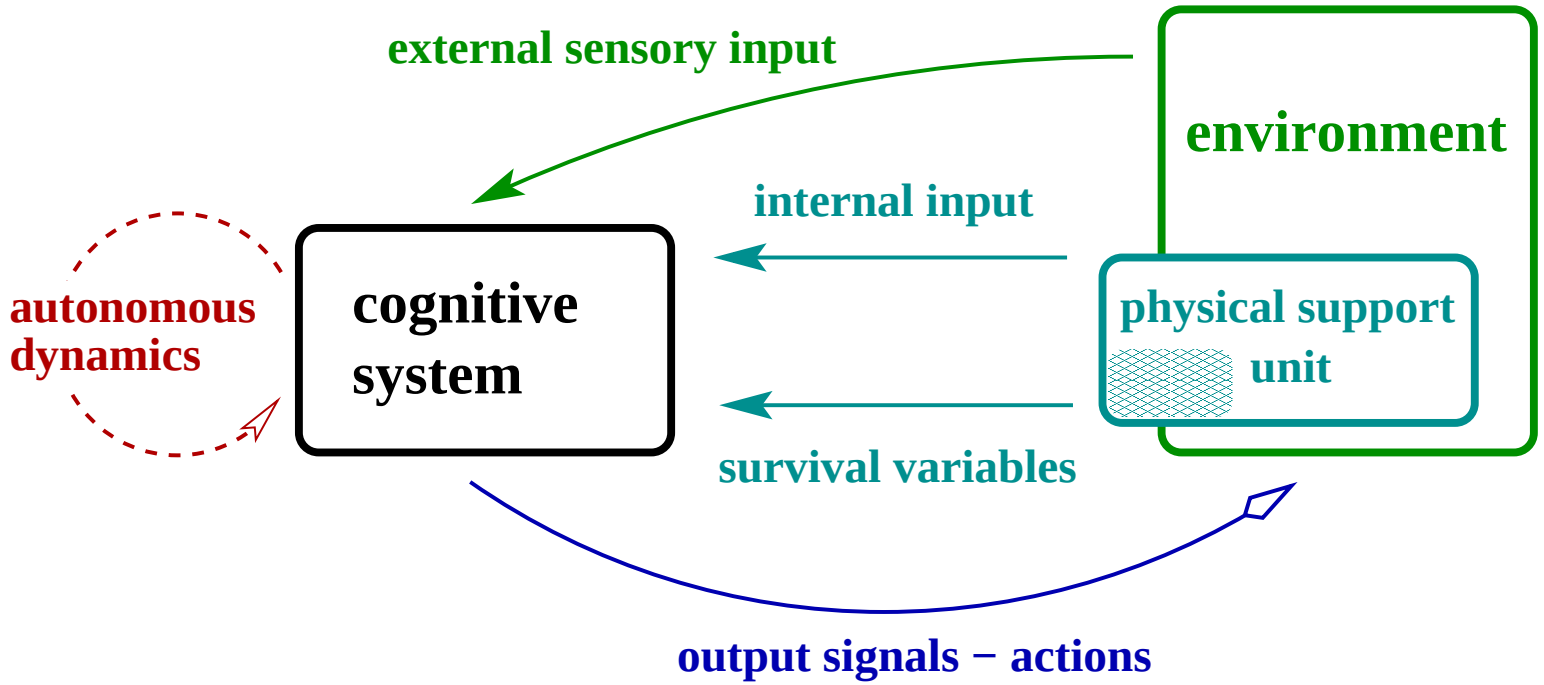
Outline

- elements of cognitive system theory
- principles of autonomous brain dynamics
  - ▷ self-sustained
  - ▷ neural competition
- neural nets with transient-state dynamics
  - ▷ internal dynamics vs. sensory data input stream
  - ▷ emergent cognitive capabilities
- outlook

what is a cognitive system?

complex and adaptive dynamical system

- living dynamical system
  - ▷ embedded
  - ▷ body for: action output
    sensory input



biologically inspired cognitive systems

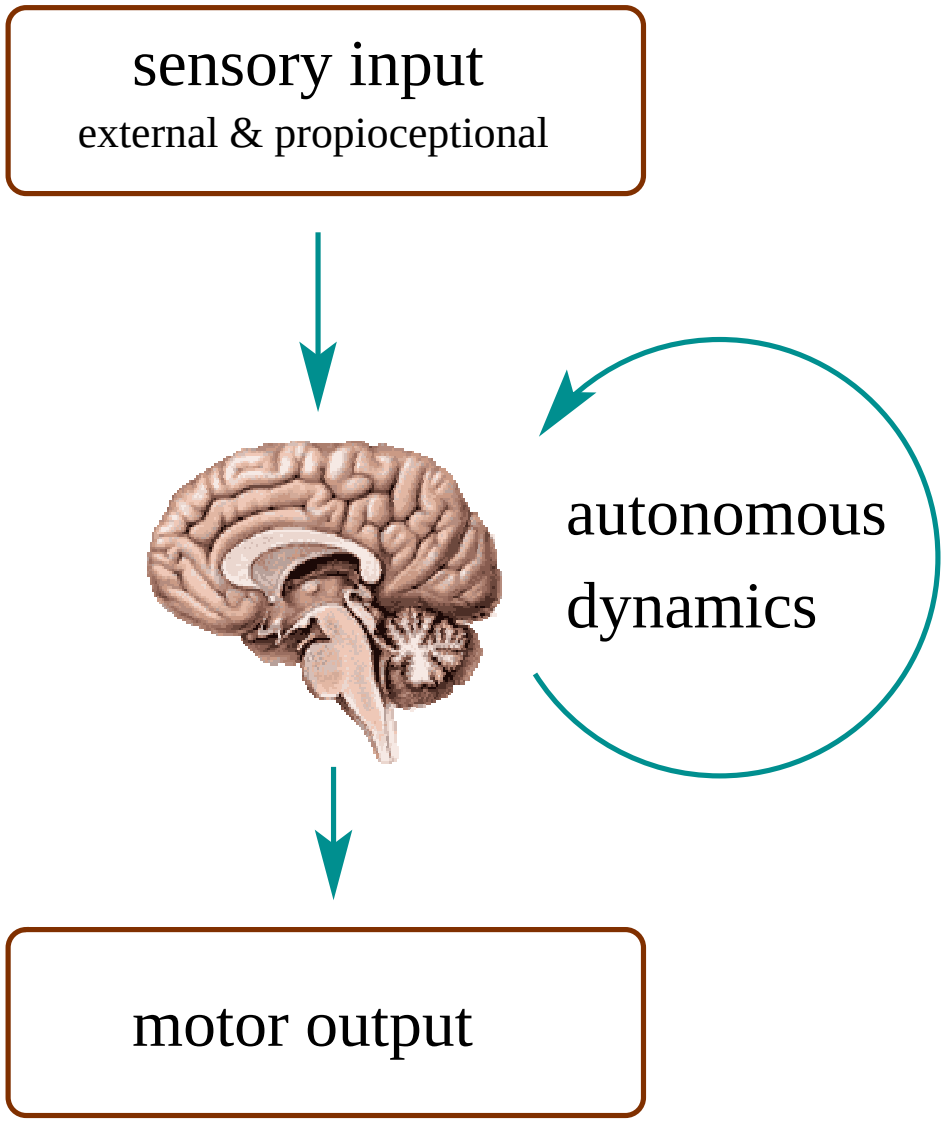
traditional AI (artificial intelligence)

- best solution for given task
- robotics, chess, visual systems, ...
- driven by applications

cognitive system approach

- general principles - self motivation
- biologically inspired, neurobiology
- emergent cognitive capabilities

global brain dynamics

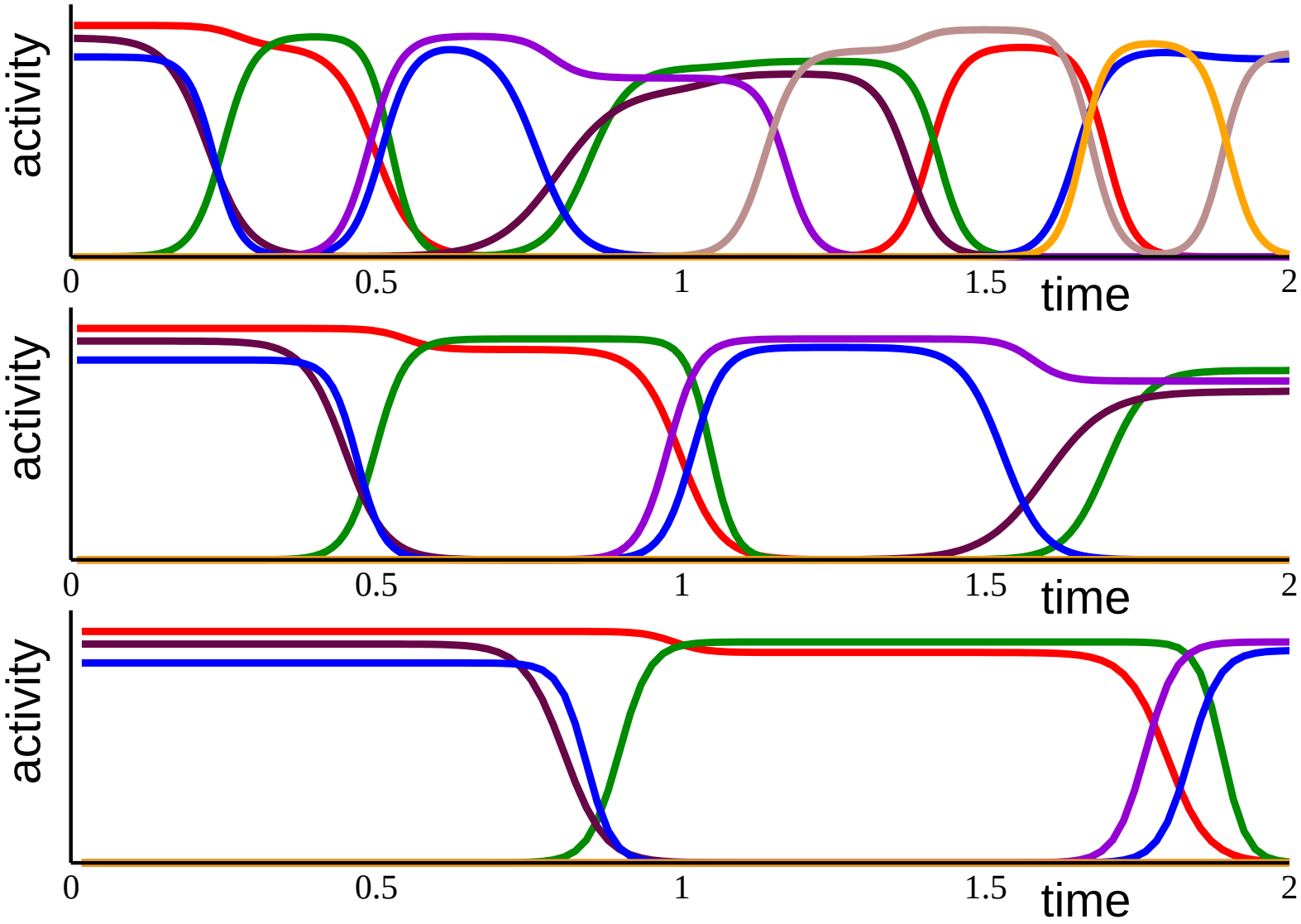


is the brain an input-output mapper?

- self-sustained neural activity
- relation to sensory input
  - ▷ driven or modulated?
  - ▷ epiphenomena or central functional role?
- states-of-the-mind [Edelmann & Tononi]
  - ▷ neural competition
  - ▷ transient neural states

transient state dynamics

transiently stable activity plateaus



multi winners-take-all networks

competitive global brain states

- “winning coalition” $\Leftrightarrow$ global states of mind
- time-sequence of winning coalitions $\Leftrightarrow$ associative thought process

clique encoding

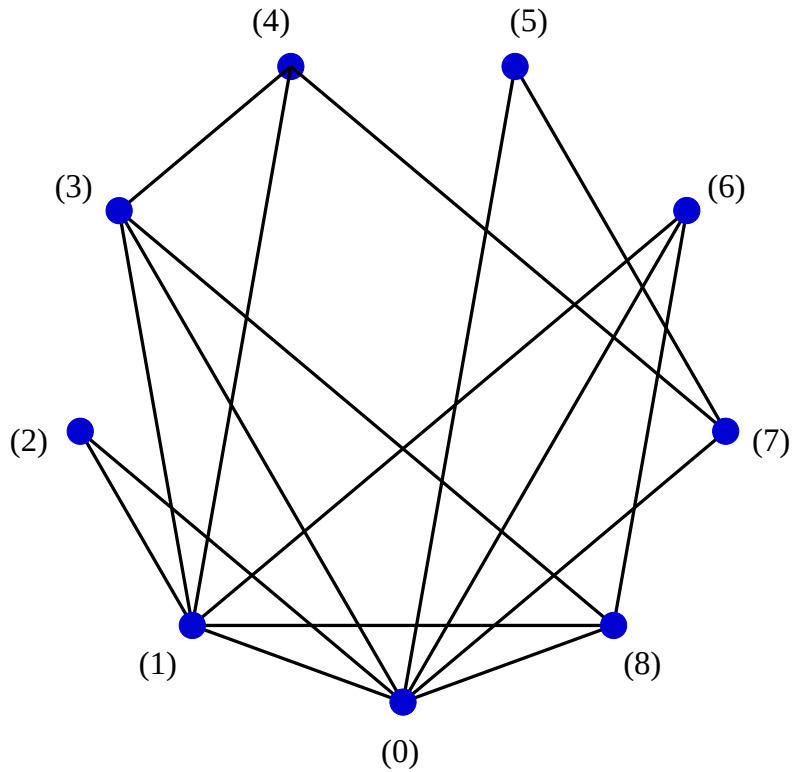
- clique $\equiv$ fully connected subgraph

cliques

(0,1,2)

(1,3,4)

(0,1,6,8)

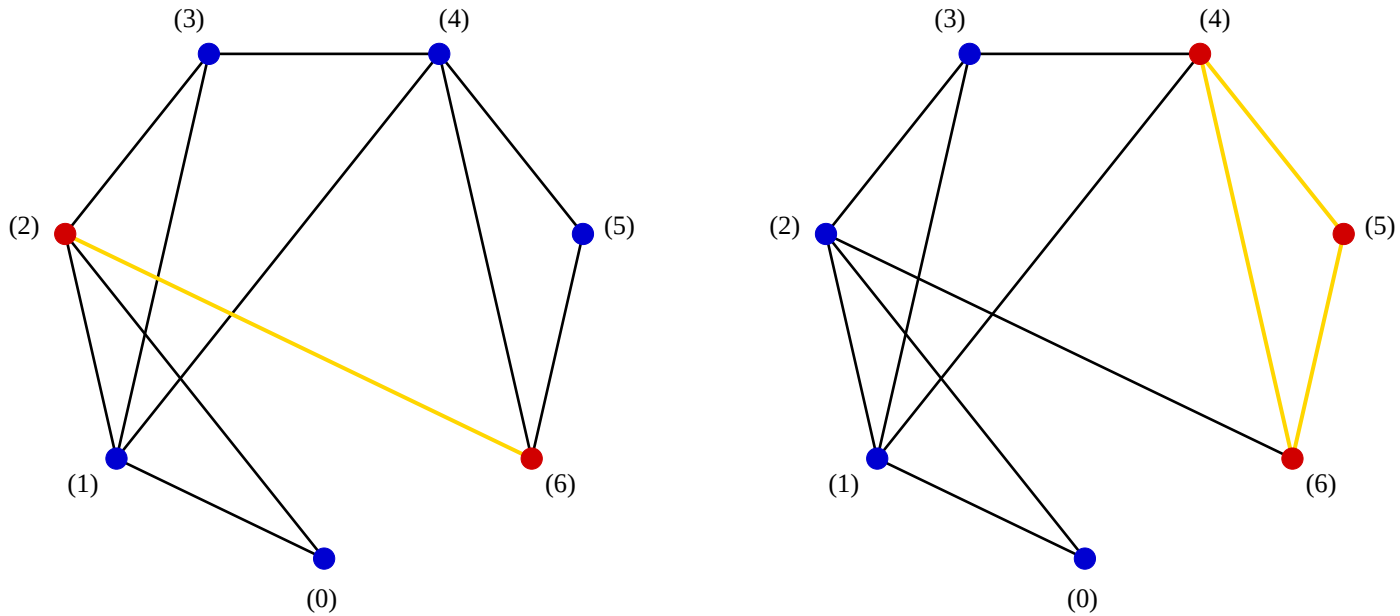


- clique $\equiv$ winning coalition

clique encoding

- lines: excitatory interactions (inhibitory background)
- stable memory states (transient attractors):

$$(0,1,2),\quad (1,2,3),\quad (1,3,4),\quad (4,5,6),\quad (2,6).$$



memory states $\Leftrightarrow$ fully connected clusters (cliques)

generalized Lotka-Volterra equations

- reservoir: $\varphi_i \in [0,1]$ $\qquad$ $x_i \in [0,1]$: activity levels

$$\begin{aligned} \dot{x}_i &= (1-x_i)\Theta(r_i)r_i + x_i[1-\Theta(r_i)]r_i \\ r_i &= g(\varphi_i)\sum_j w_{i,j}x_j + \sum_j z_{i,j}x_j f(\varphi_j) \end{aligned}$$

Γ_φ^- : depletion-rate
Γ_φ^+ : filling-rate

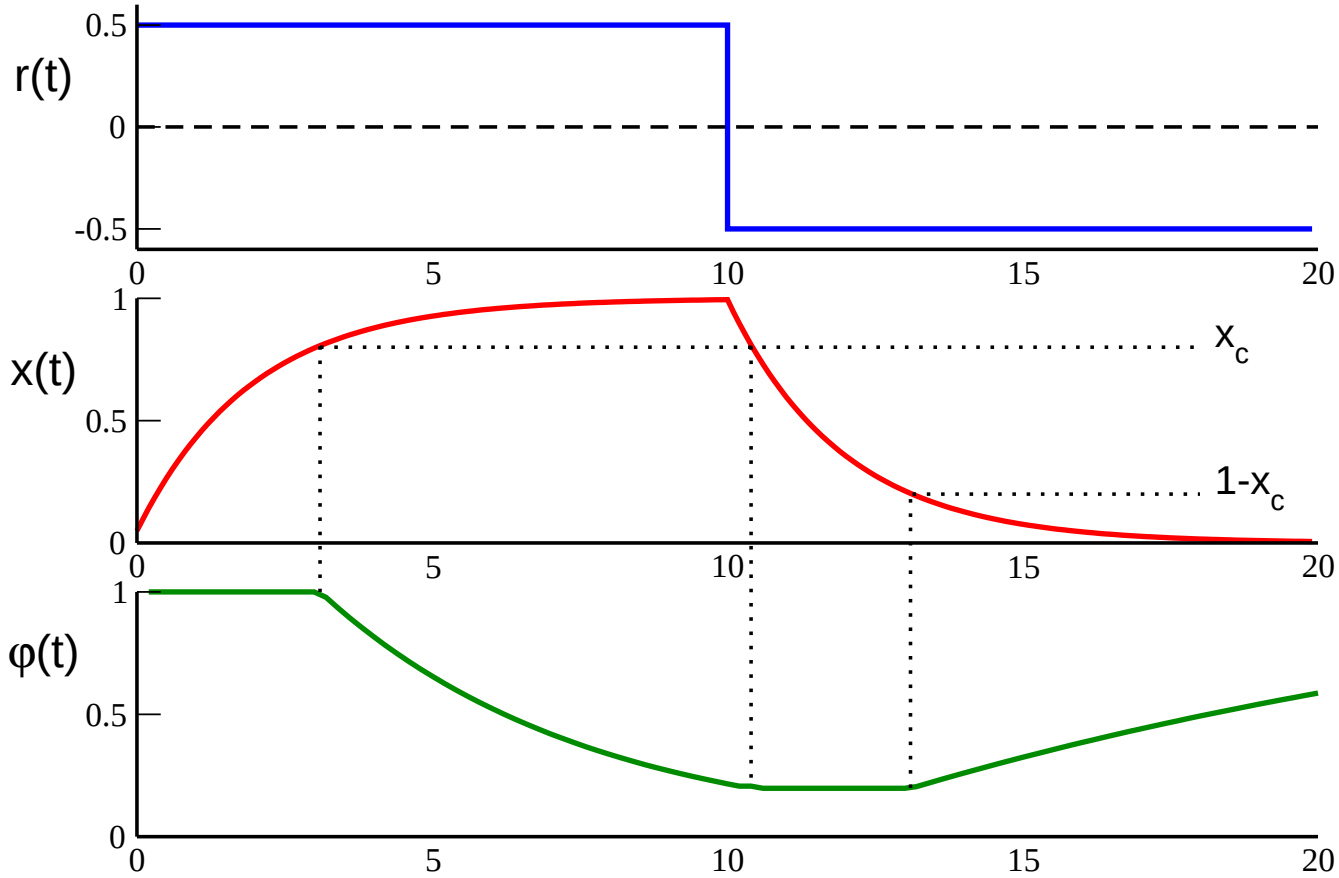
$\Gamma_\varphi^\pm \to 0$: transient states $\longrightarrow$ attractors

$$\dot{\varphi}_i = \Gamma_\varphi^+ \Theta((1-x_c)-x_i)(1-\varphi_i) - \Gamma_\varphi^- \Theta(x_i-x_c)\varphi_i$$

- coupling via reservoir functions $f(\varphi)$, $g(\varphi)$

reservoir dynamics

$$\begin{aligned} \dot{x} &= (1-x)\Theta(r)r + x[1-\Theta(r)]r \\ \dot{\varphi} &= \Gamma_{\varphi}^{+}\,\Theta((1-x_c)-x)(1-\varphi) - \Gamma_{\varphi}^{-}\,\Theta(x-x_c)\,\varphi \end{aligned}$$



stabilization of memory states

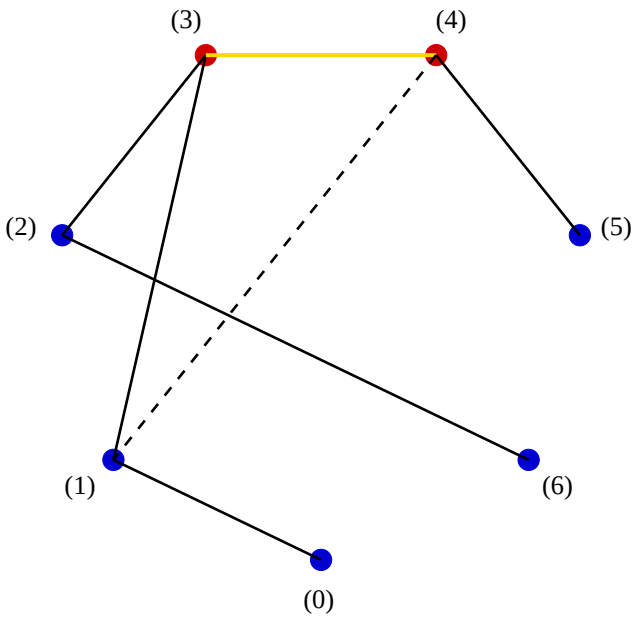
illustration:
homogeneous case

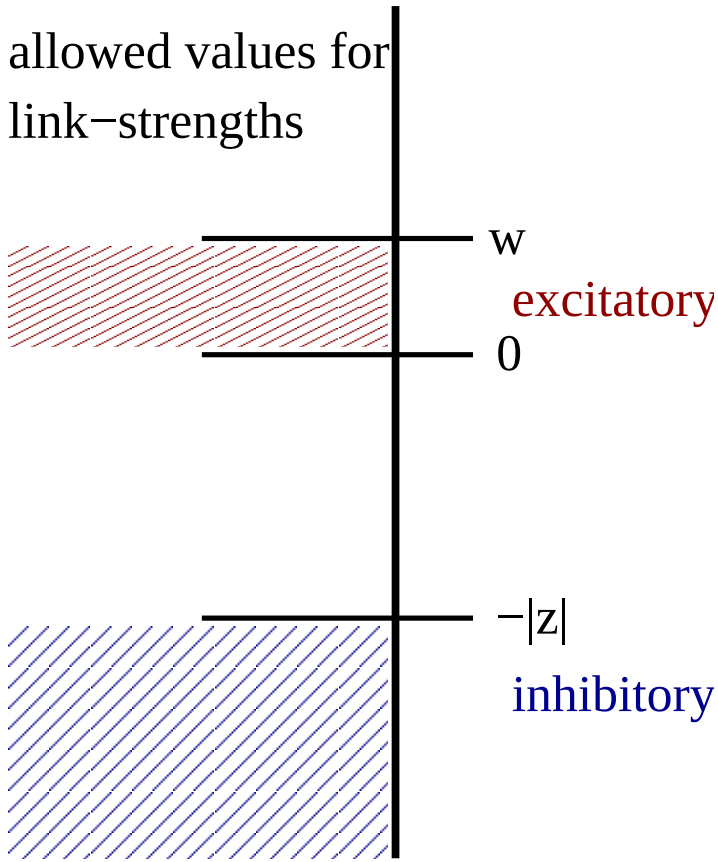
$$\begin{aligned} \dot{x}_i &= (1-x_i)\Theta(r_i)r_i + x_i[1-\Theta(r_i)]r_i \\ r_i &= g(\varphi_i)\sum_j w_{i,j}x_j + \sum_j z_{i,j}x_j f(\varphi_j) \end{aligned}$$

excitatory: $w_{i,j} = w > 0$

inhibitory: $z_{i,j} = z = -|z|$

▷ condition: $|z| > (Z-1)w$





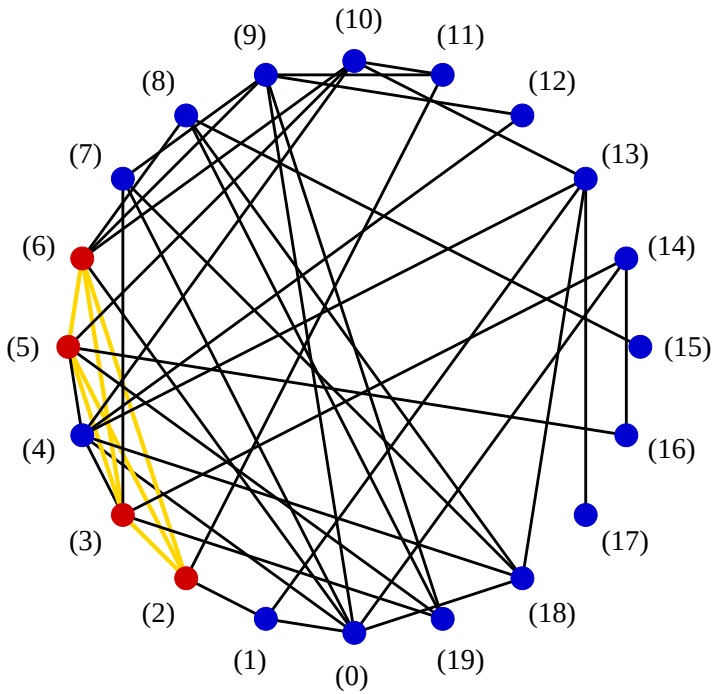
dense and homogeneous networks

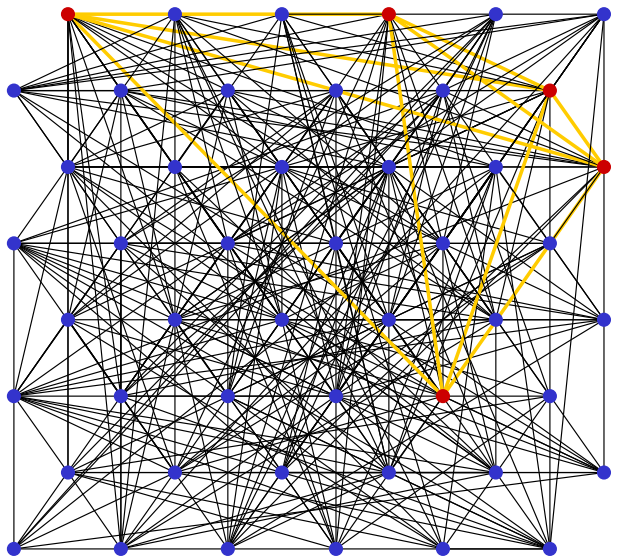
clique encoding

sparse coding $\Rightarrow$ high storage capacity

stable memory states:

	2-AC	3-AC	4-AC	5-AC	total
20-sites	19	10	1	0	30
48-sites	2	166	66	2	236



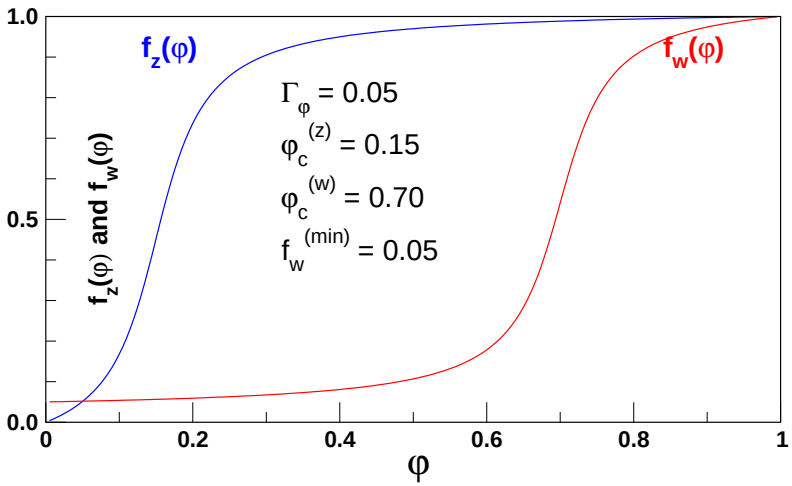


reservoir functions

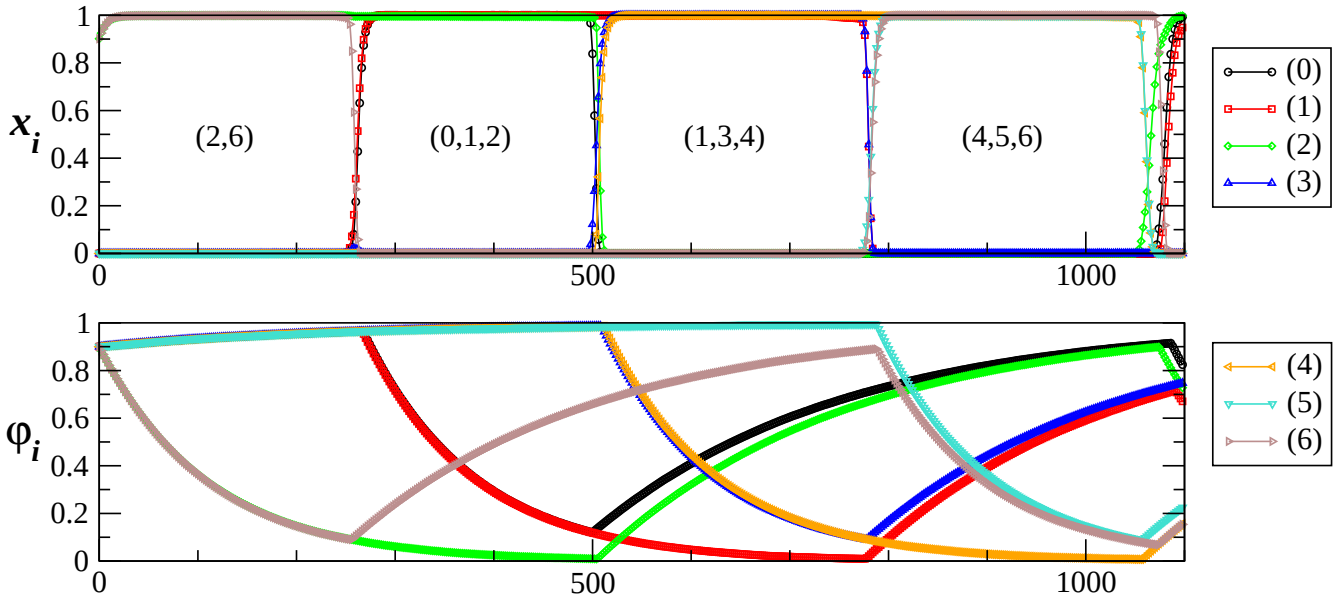
$$\begin{aligned} \dot{x}_i &= (1-x_i)\Theta(r_i)r_i + x_i[1-\Theta(r_i)]r_i \\ r_i &= g(\varphi_i)\sum_{j=1}^{N} w_{i,j}x_j + \sum_{j=1}^{N} z_{i,j}x_j f(\varphi_j) \end{aligned}$$

- ▷ neuron can inhibit other neurons if $\varphi > \varphi_c^{(f)} = 0.15$
- ▷ neuron can be activated by other neurons if $\varphi > \varphi_c^{(g)} = 0.8$

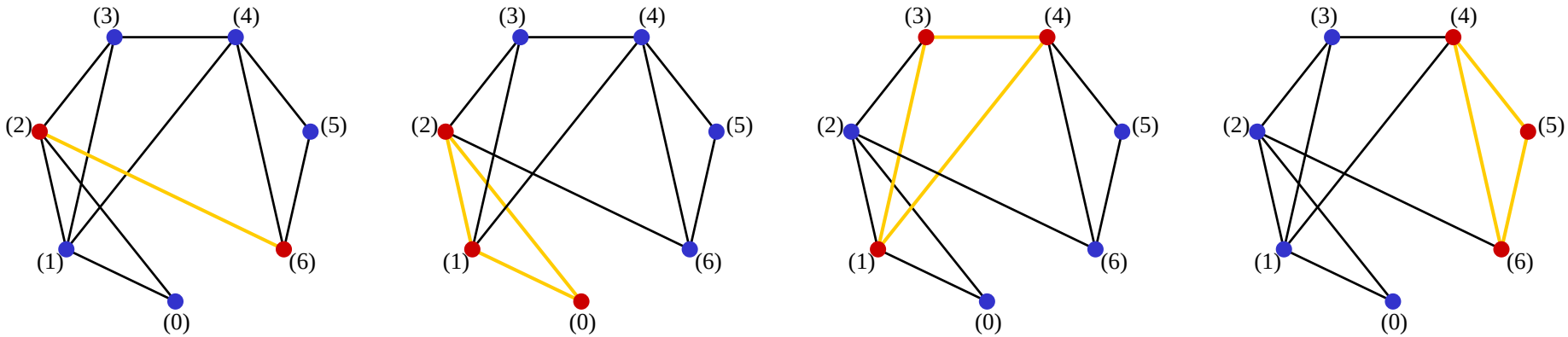
$$f(\varphi) = \Theta(\varphi - \varphi_c^{(f)}) \qquad g(\varphi) = \Theta(\varphi - \varphi_c^{(g)})$$



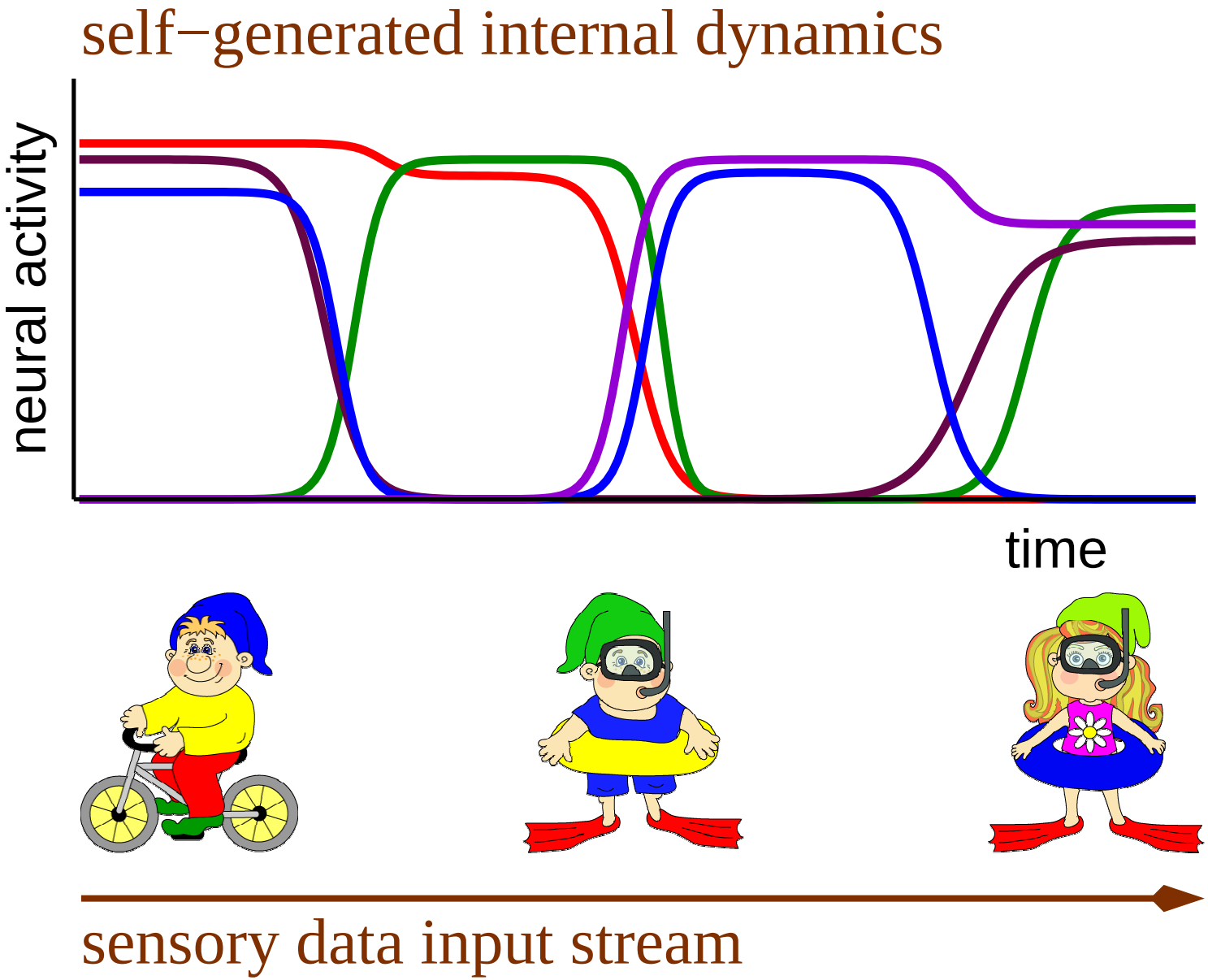
a simple associative thought process



thought process: $(2,6) \rightarrow (0,1,2) \rightarrow (1,3,4) \rightarrow (4,5,6)$



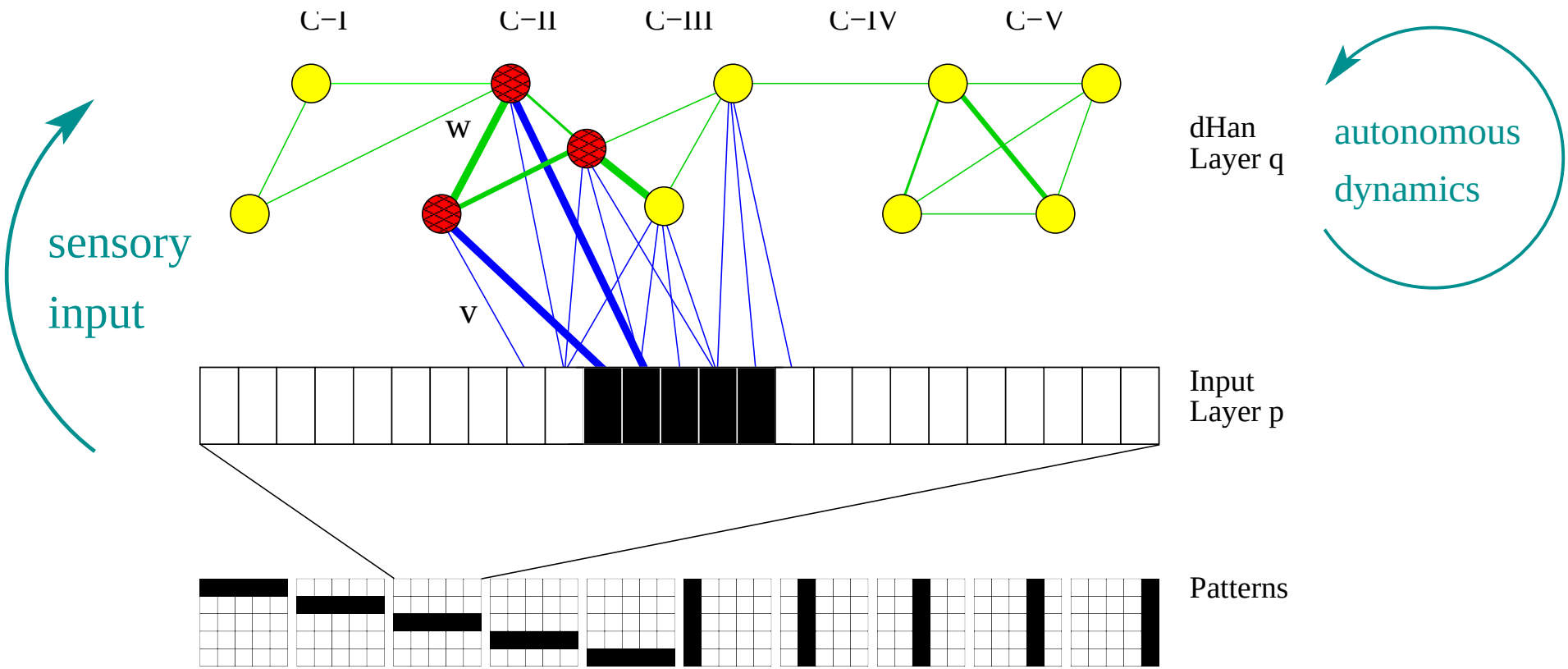
correlations: internal vs. external?



coupling to sensory input

self-generated internal dynamics

- dHan (dense homogeneous associative network)



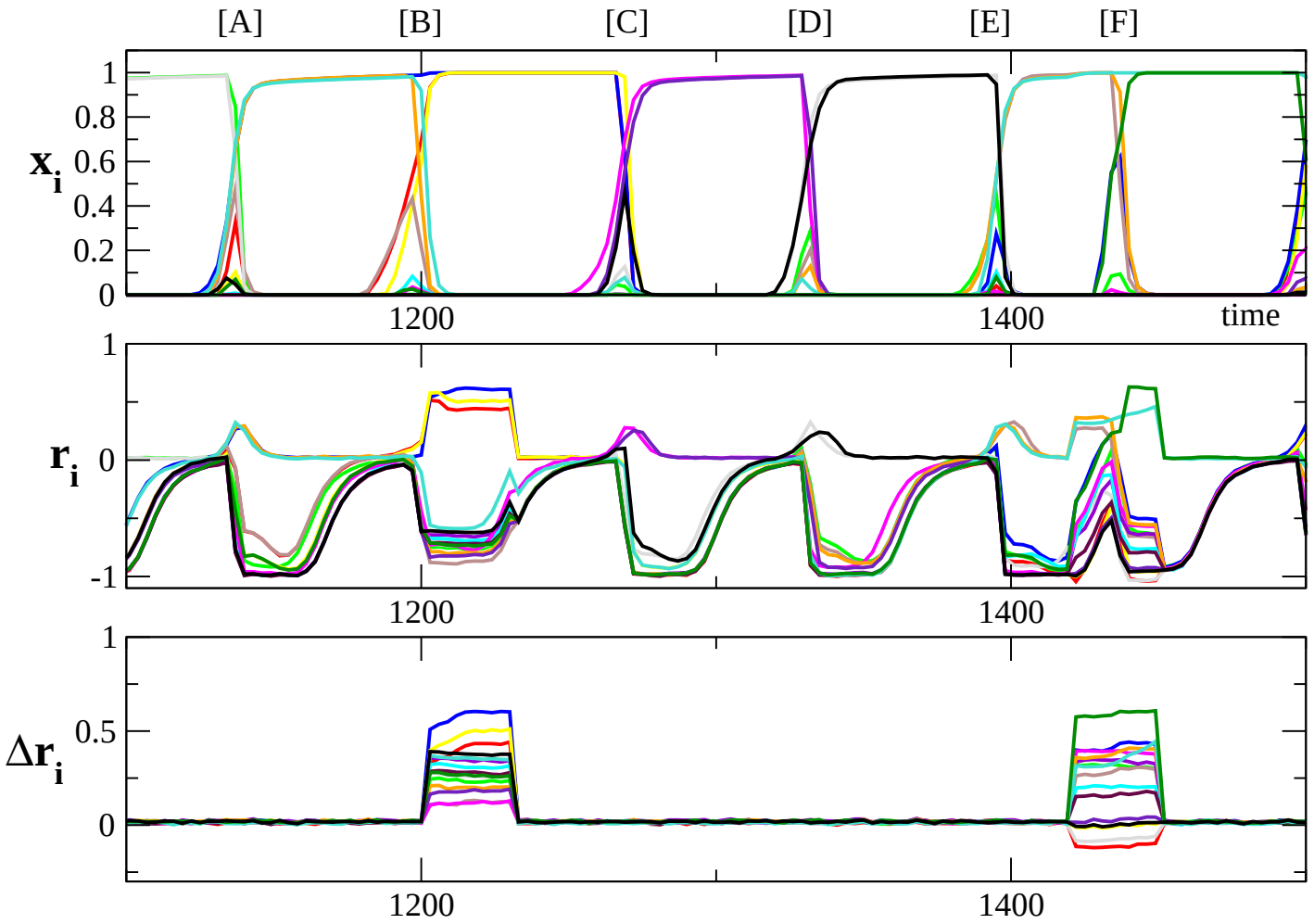
input data stream

- unrelated to internal dynamics

learning during sensible periods

competition

internal dynamics ⇔ sensory input



activities

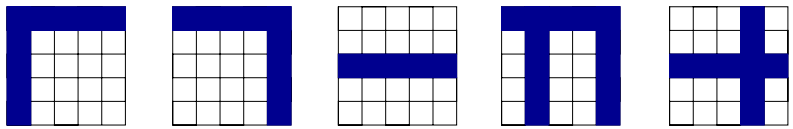
growth rates

sensory input

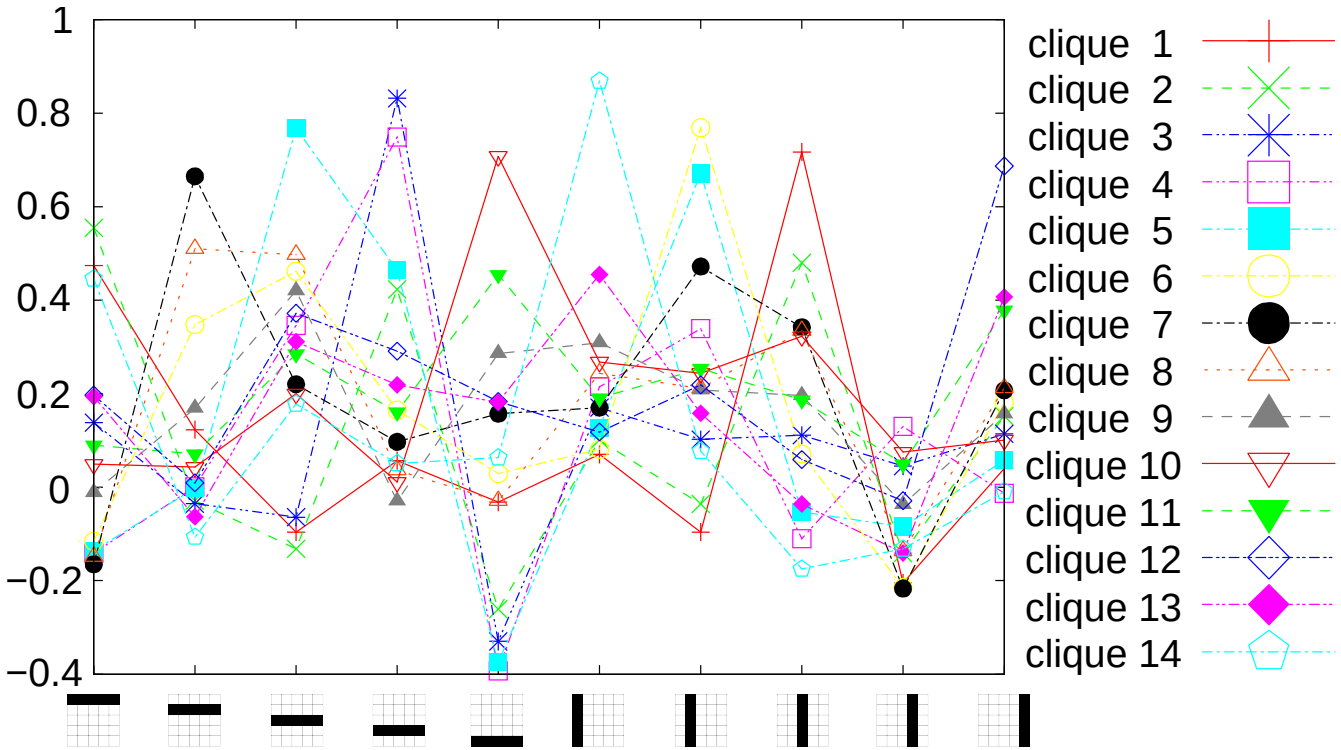
bars problem

random superposition

- horizontal & vertical bars



independent component analysis



[Gros & Kaczor ‘08]

emerging cognitive capabilities

working principles

- ▷ self-sustained dynamical activity
- ▷ transient states - competitive dynamics

coupling to input signals

- ▷ competitive dynamics - learning in sensitive periods
- ▷ emergent - not directly encoded

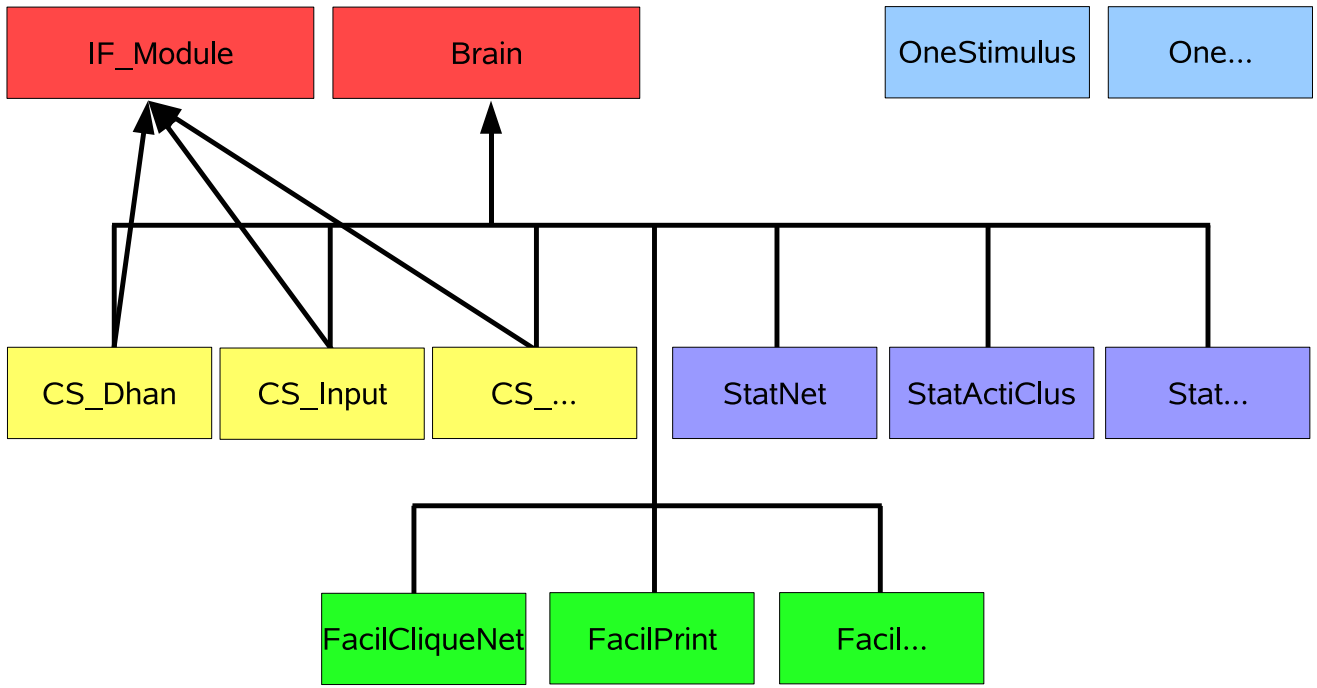
extraction of statistically independent components

- ▷ rudimentary object recognition capabilities

modular approach

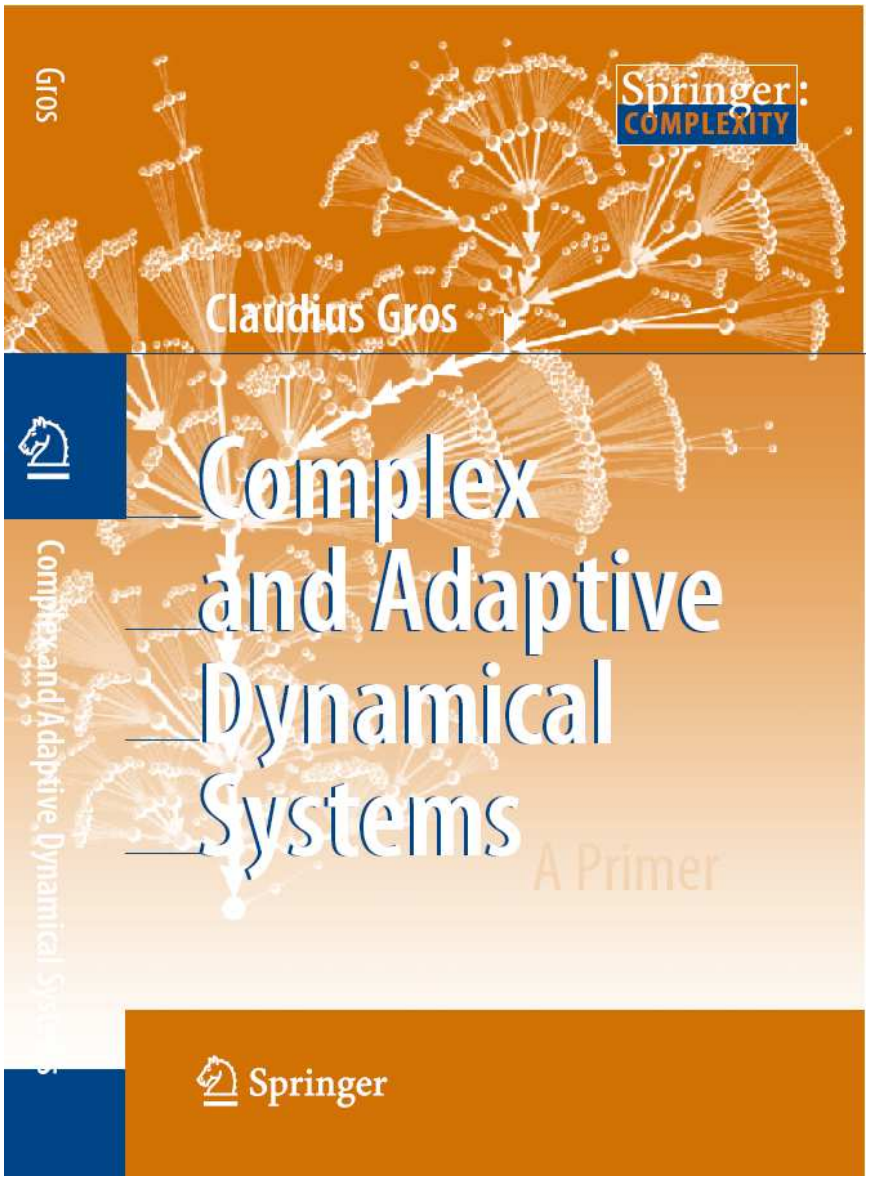
meta network of neural networks

- JAVA platform: class diagram



- full flexibility: full on-the-fly architectural reconfiguration
- status: several working modules

graduate-level textbook



- The small world phenomenon in social and scale-free networks
- Phase transitions and self-organized criticality in adaptive systems
- Life at the edge of chaos and coevolutionary avalanches resulting from the unfolding of all living
- Living dynamical systems and emotional diffusive control within cognitive system theory

(Springer, 2008)