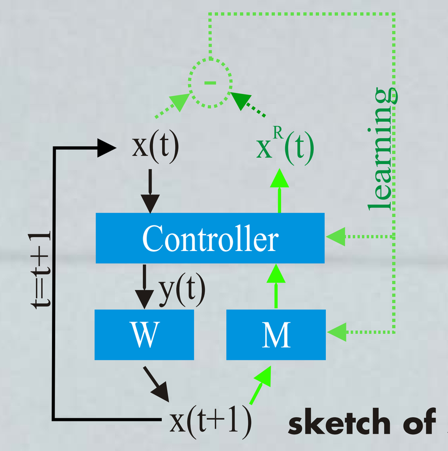
Self-organized exploration and automatic sensor integration from the homeokinetic principle

UNIVERSITÄT LEIPZIG

Institut für Informatik, POB 920, D 04009 Leipzig

Ralf Der, Frank Hesse, René Liebscher {der|fhesse|liebschr}@informatik.uni-leipzig.de

Sensorimotor loop



sketch of sensorimotor loop

— sensorimotor loop
— time loop
··· time loop error used as learning signal

World and world model

M world model: $x^M(t+1)=a(x)y(t)$

W world: $x(t+1)=a(x)y(t)+\xi(t)$

$x(t)$-sensor value

a(t)-response factor

$y(t)$-controller ouput (motor command)

$\xi(t)$- model error

Time loop error

-Difference between true x(t) and reconstructed sensor value $x^R(t)=C^{-1}(M^{-1}(x(t+1)))$ with C as Controller

- Minimizing this difference leads to stabilization backward in time, this means destabilization forward in time (Nonlinearities in the system confine unlimited increase)

A closed loop velocity control of a robot realized by a leaky-integrator neuron: The membrane potential z of the leaky-integrator neuron is calculated as $\tau^{-1}\Delta z = -z + c_i x_i + H$

c_i - synaptic strength of channel i

x_i - input of channel i

H - bias

The neuron output y is $y = g(z) = \tanh(z)$

The homeocinetic principle (gradient descending the time loop error) in some approximation produces the following dynamics for the synaptic strength c_i and the bias (internal state variable) H.

$$\Delta c_i = \mu a_i - 2\mu z x_i - \gamma\mu c_i$$

$$\Delta H = -2\mu z \quad \text{where} \quad \mu = \varepsilon g'\xi^2$$

μ - synaptic gain control (modified update rate)

a_i - response factor of channel i

ξ^2 - average model error

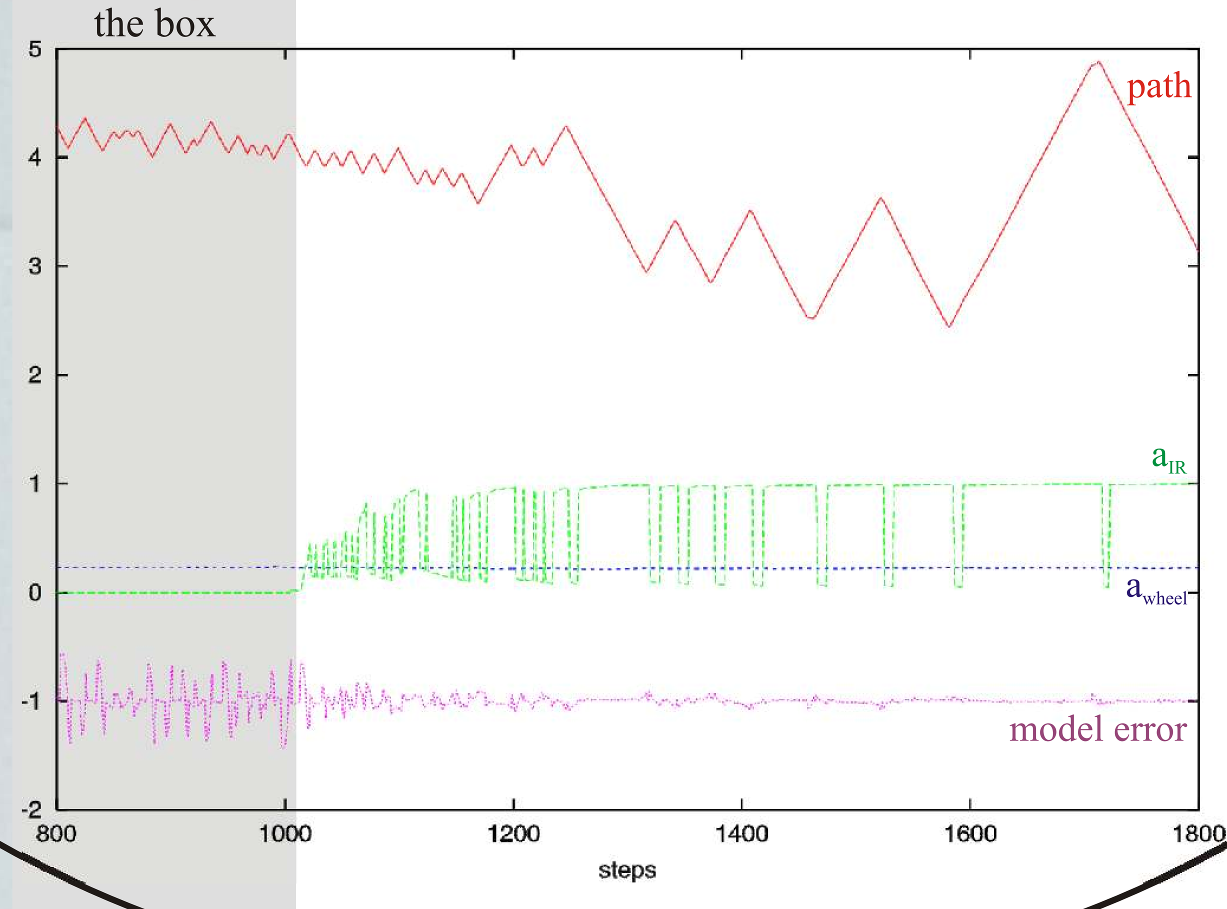
γ - term for small decay of weights

$$g'(z) = \tanh'(z) = 1 - \tanh^2(z)$$

Experimental results

large modeling error, robot does not push the box

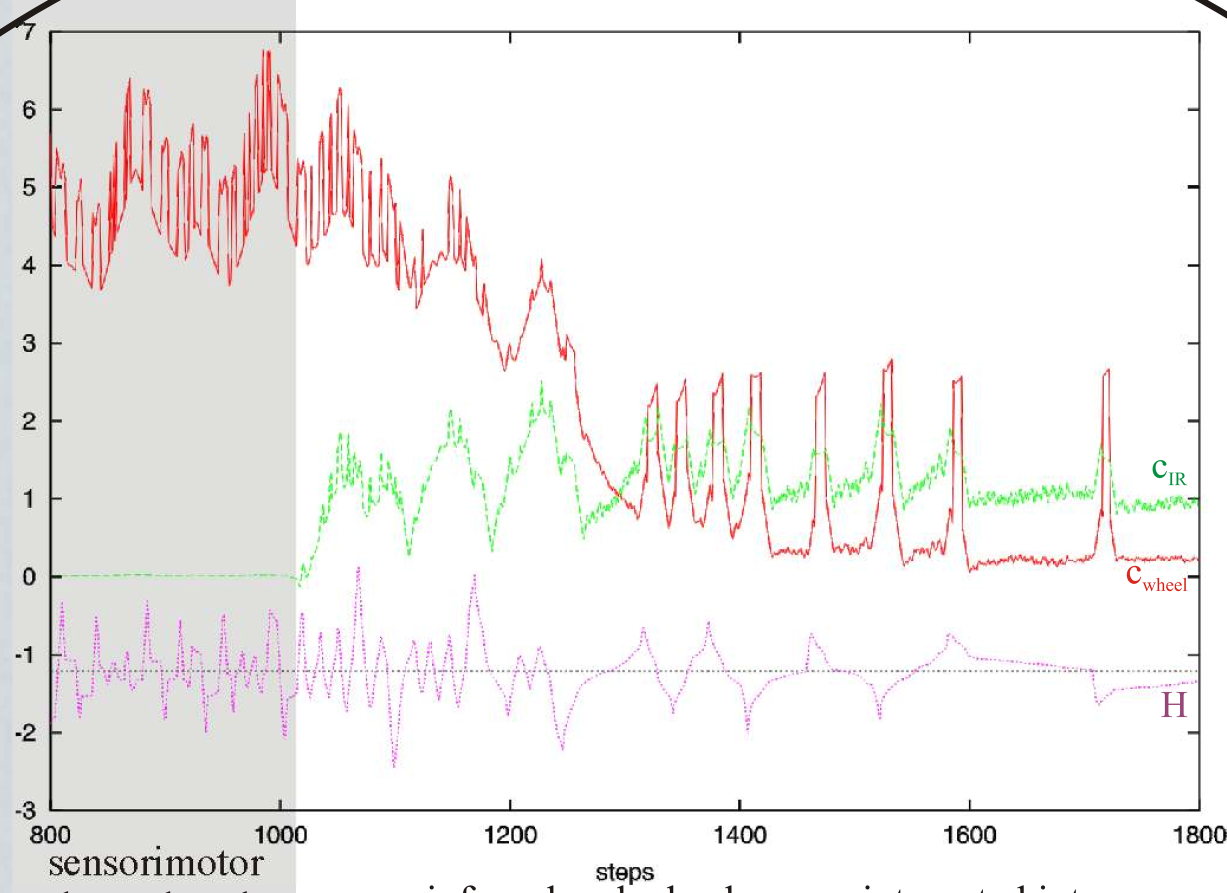
small modeling error, robot moves the box around and explores the whole range of the arena



travelled path, response factors and model error during experiment

no model learning a_{IR} set to 0

model learning activated a_{IR} adapted



synaptic strength and bias (internal state variable) during experiment

sensorimotor loop closed only over wheel channel

infrared and wheel sensor integrated into the sensorimotor loop, according to their response strength

Result of learning: $c_i = \alpha a_i$ for all i

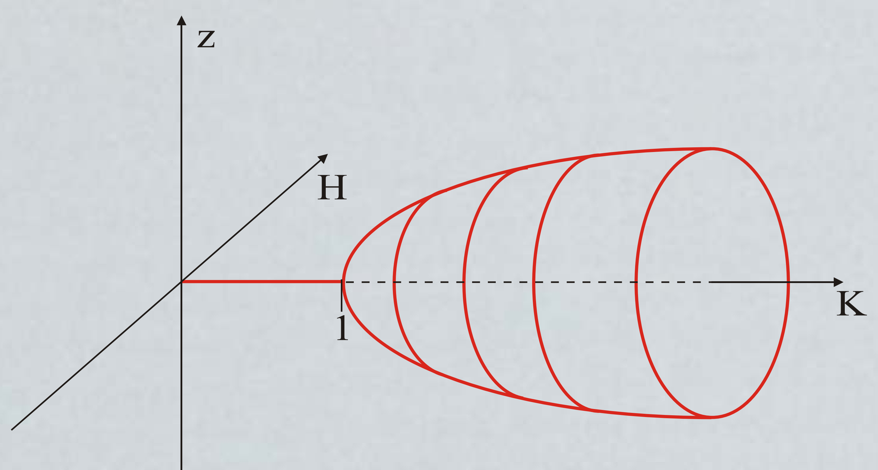
Automatic sensor integration according to response strength

Properties of a system with one channel

-Dynamics of z and H lead to a limit cycle in the (z,H) space

-With increasing K we get a Hopf-bifurcation (K=ca is the feed back strength of the system)

-c self-regulating to a slightly supercritical value of K of approx. 1.2



sketch of Hopf-bifurcation

-Frequency of the limit cycle oscillation determined by synaptic gain control

-Large modeling error ξ leads to high frequency

-Self-regulating exploration rate:

·Regions with large modeling error explored more intensely

·Information input for model learning enhanced

Experimental conditions:

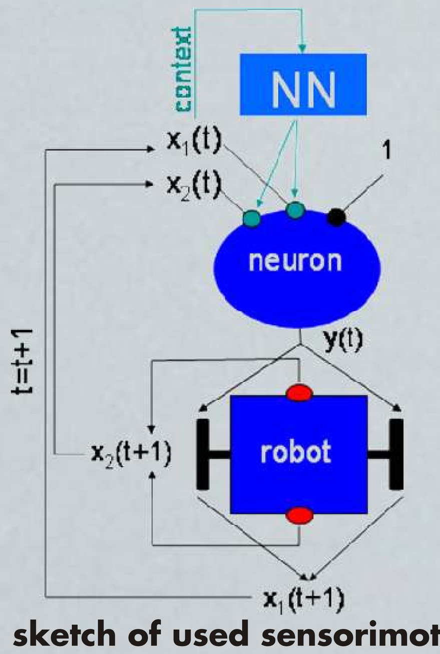
- Khepera robot inside moveable box
- Fixed walls as borders of the test area
- Sensors
  - Velocity of wheels x_1
  - Pseudo-infrared sensor triggered by physical infrared sensors r

$$x_2 = \begin{cases} b y_{t-1} : \max_{i=0}^{7} r_i > r_{min} \\ 0 : else \end{cases}$$

- Long time memory (NN) with context m

$$m = \begin{cases} 1 : \max_{i=0}^{7} r_i > r_{min} \\ 0 : else \end{cases}$$

- a_2, c_1, c_2 learned with rules given above



sketch of used sensorimotor loop

background picture shows experimental conditions

printed at the computer center of Leipzig University