Summary

Gregor Schöner
gregor.schoener@ini.rub.de

Cognition in the wild…

- attention/gaze
- active perception/working memory
- action plans/decisions/sequences
- goal orientation
- motor control
- background knowledge
- learning from experience

=> implied properties of the underlying neural processes

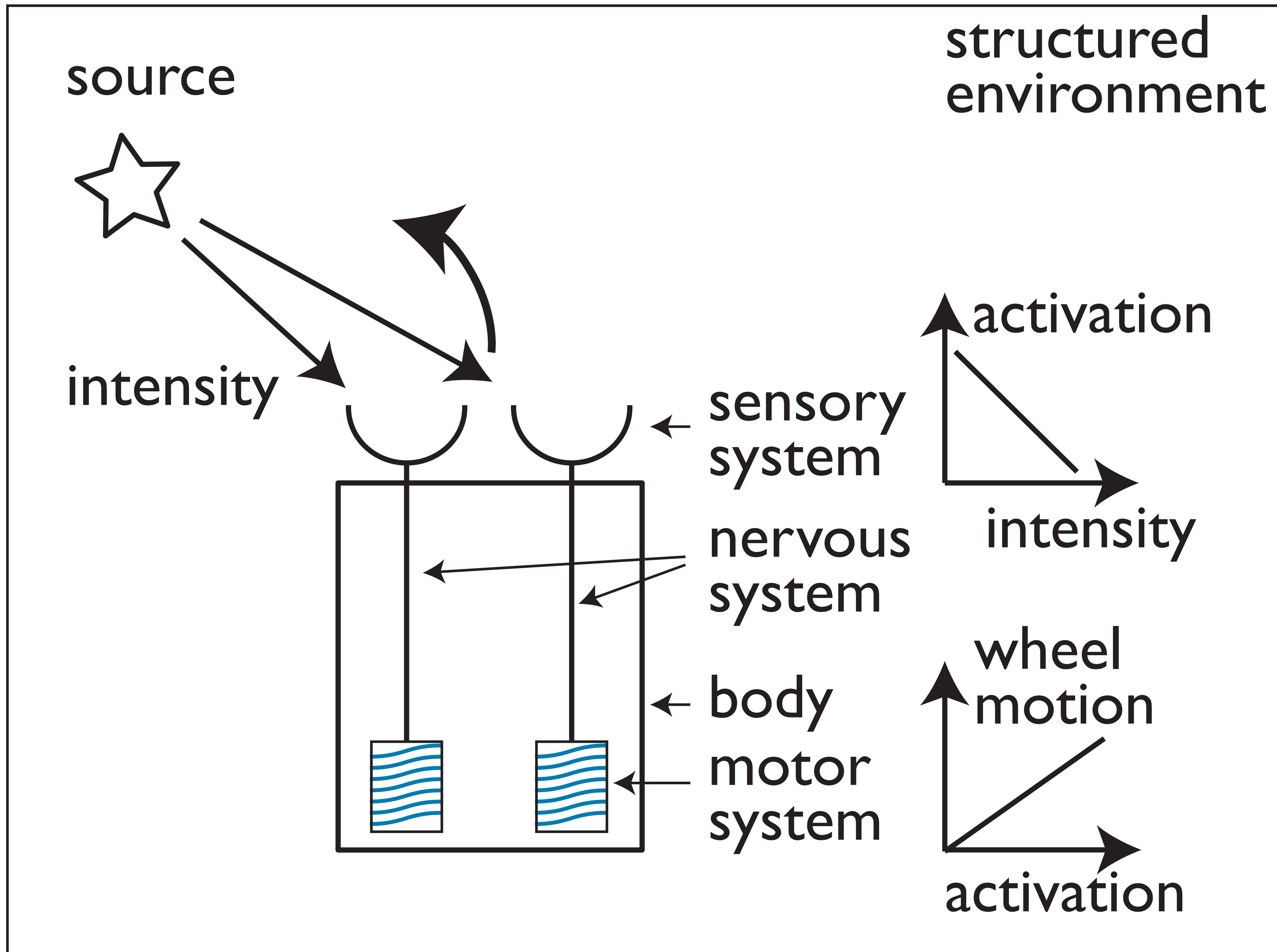
- graded state
- continuous time
- continuous/intermittent link to the sensory and motor surfaces
- from which discrete events and categorical behavior emerge
- in closed loop
- => states must be stable

Embodiment hypothesis

- all cognition is like soccer playing = has the properties of embodied cognition
- => there is no particular boundary up to which cognition is embodied and beyond which it is computational/symbolic

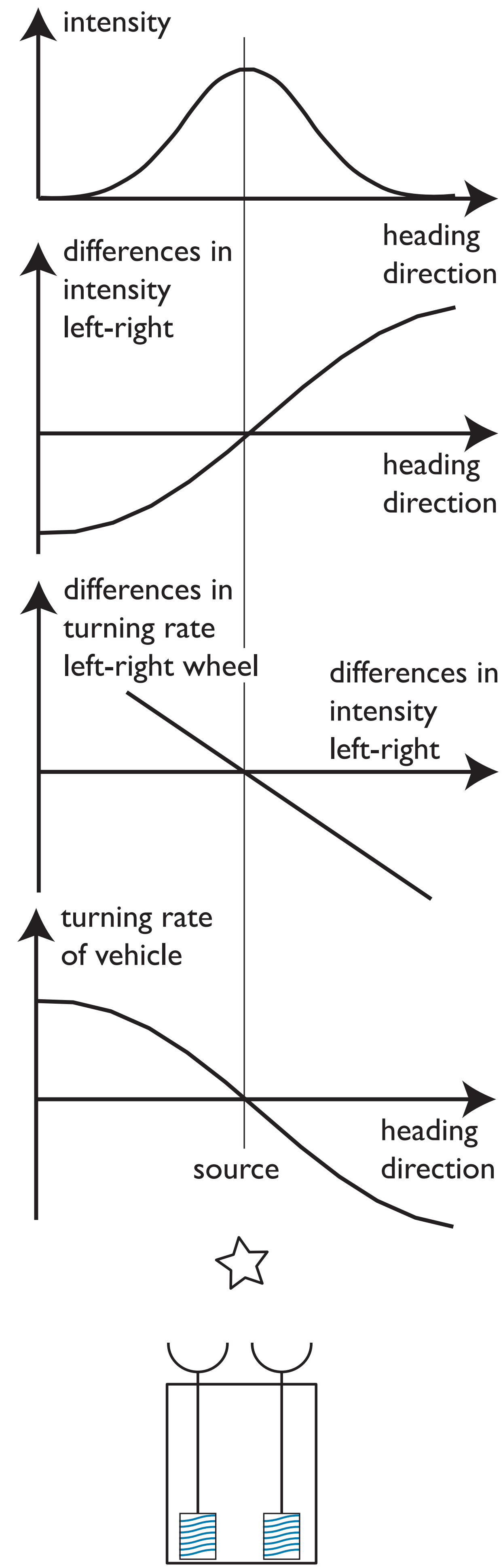
Five things needed to generate behavior

- sensors
- motors
- linked by a nervous system
- linked physically by a body
- an appropriately structured environment



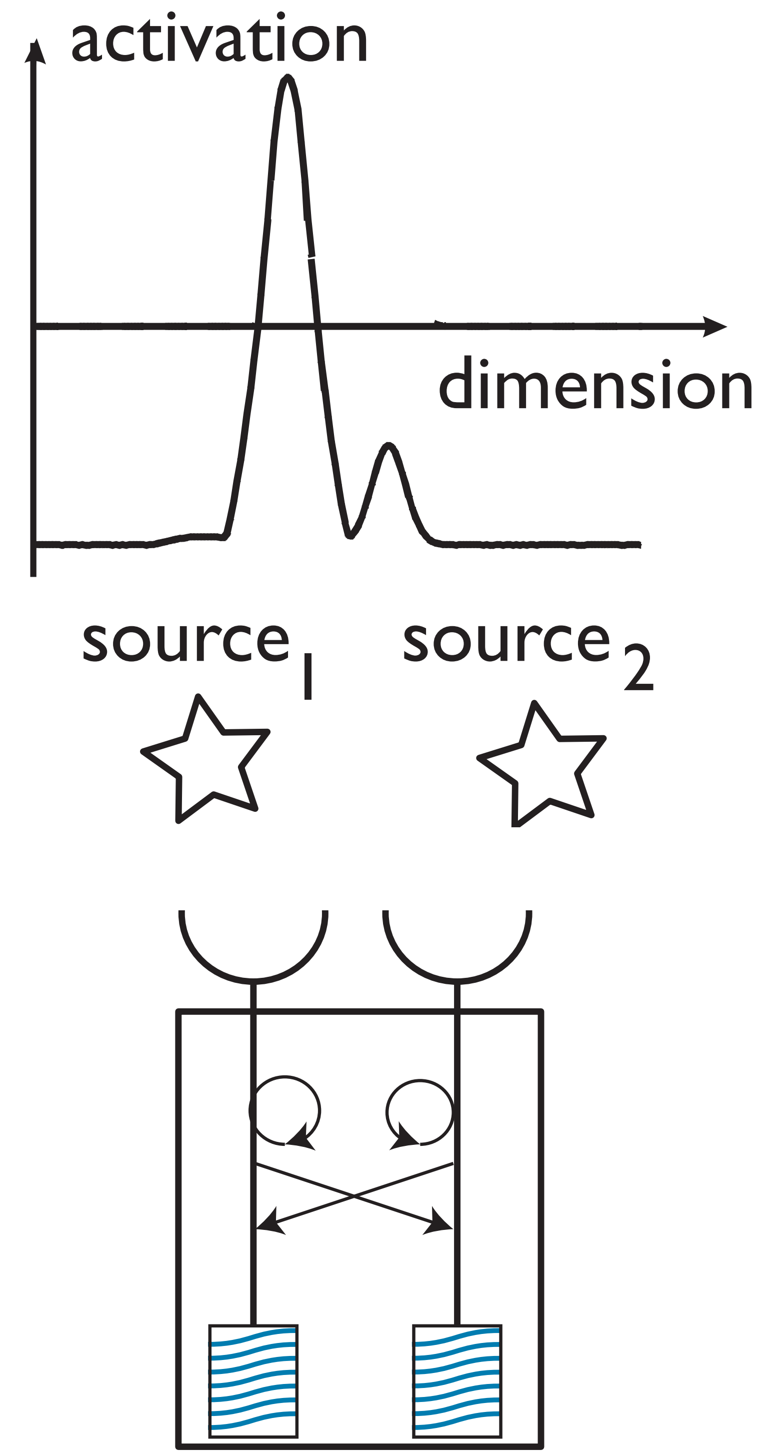
Emergent behavior: this is a dynamics

- feedforward nervous system
- + closed loop through environment
- => (behavioral) dynamics



Emergent cognition from neural dynamics

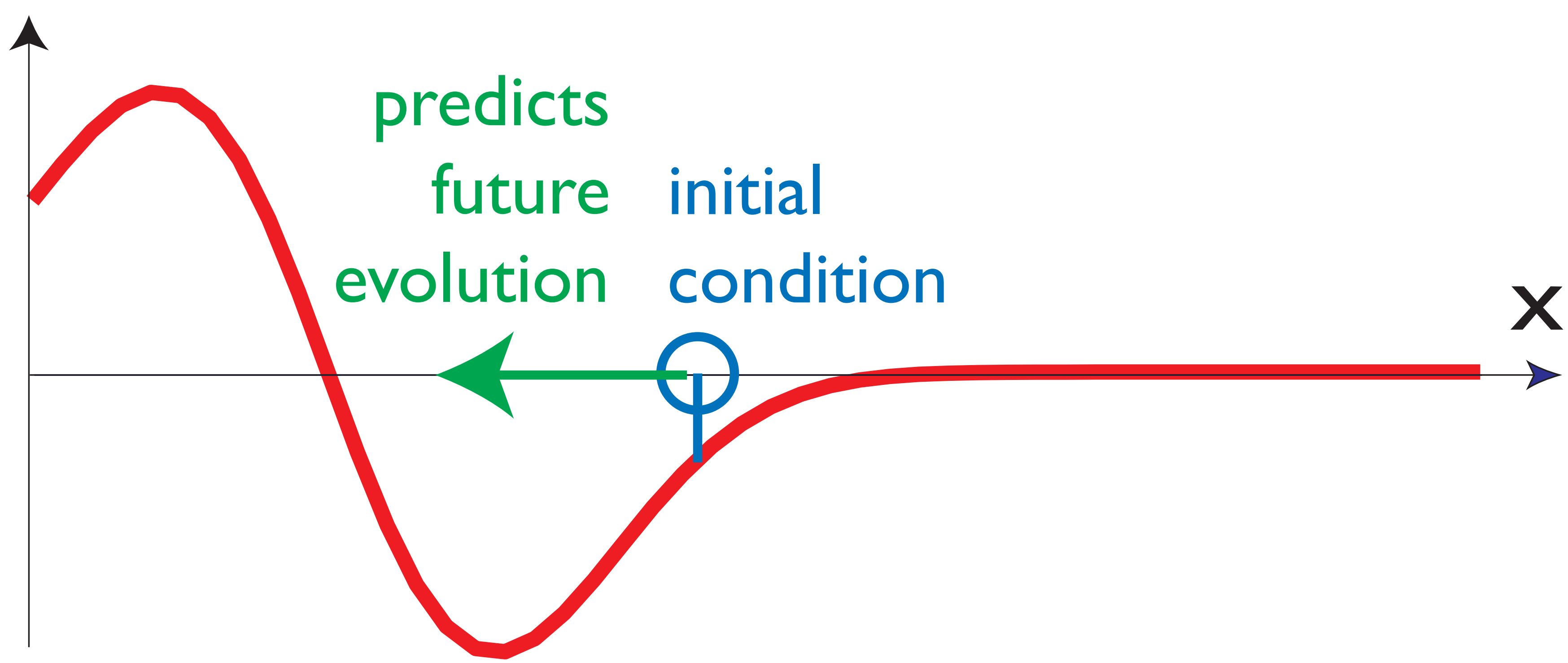
- mental decisions, working memory..



Dynamical system

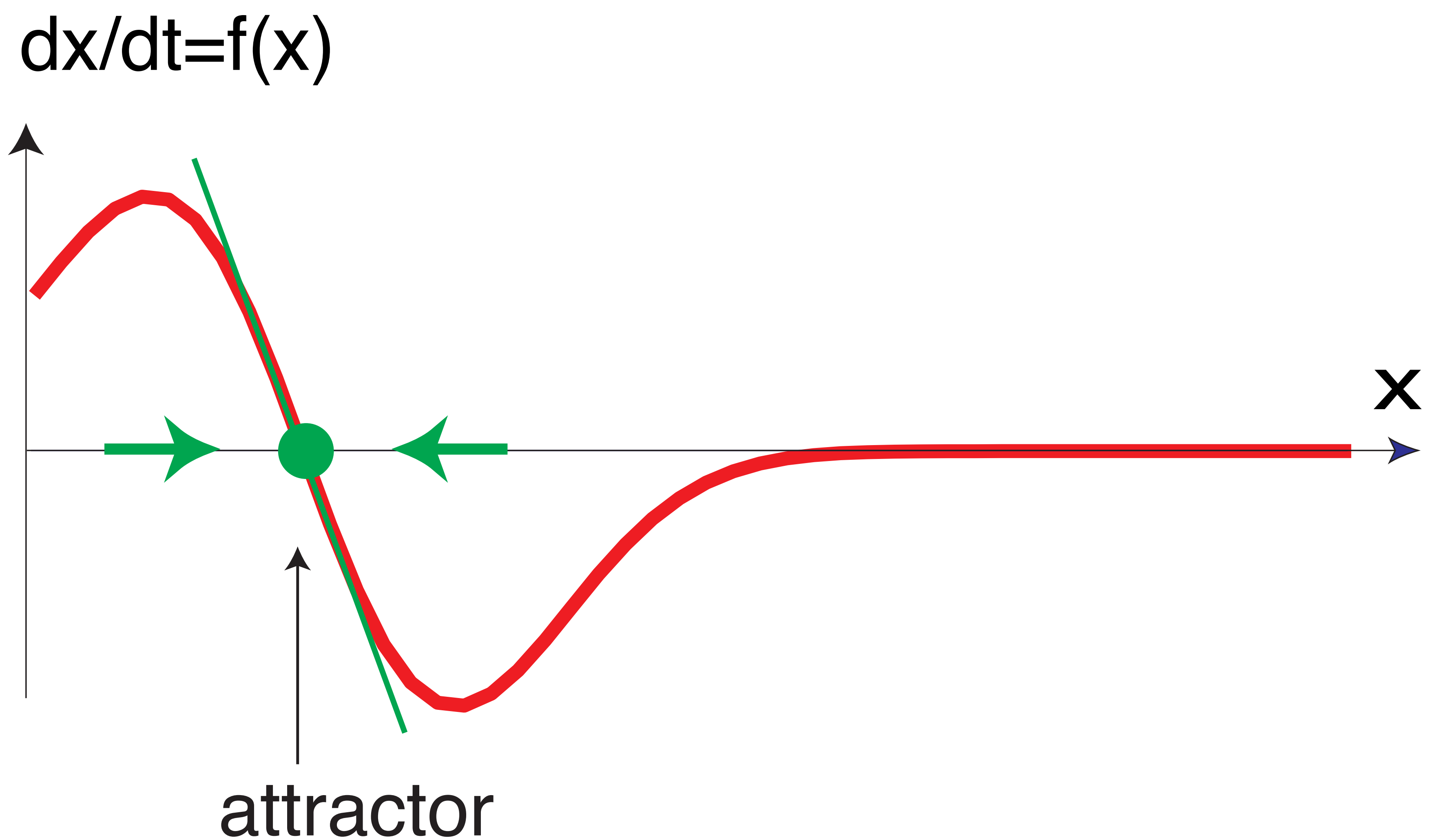
- present determines the future
  - given initial condition
  - predict evolution (or predict the past)

$dx/dt=f(x)$



attractor

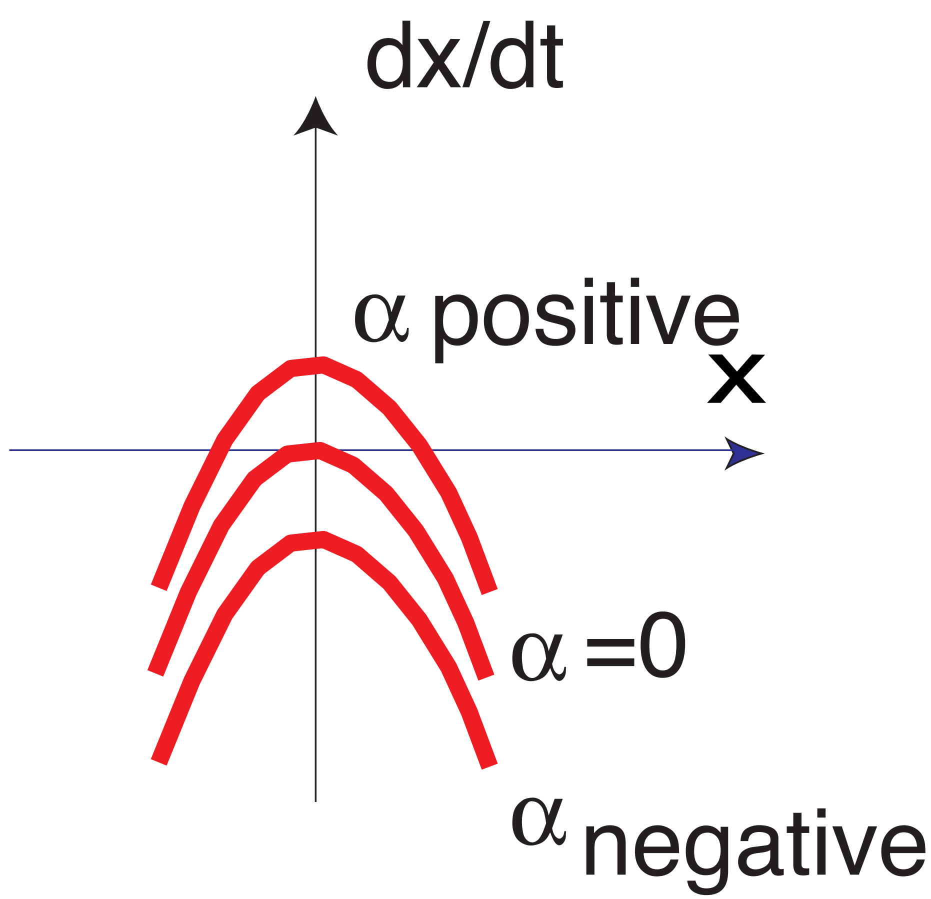
- fixed point, to which neighboring initial conditions converge = attractor

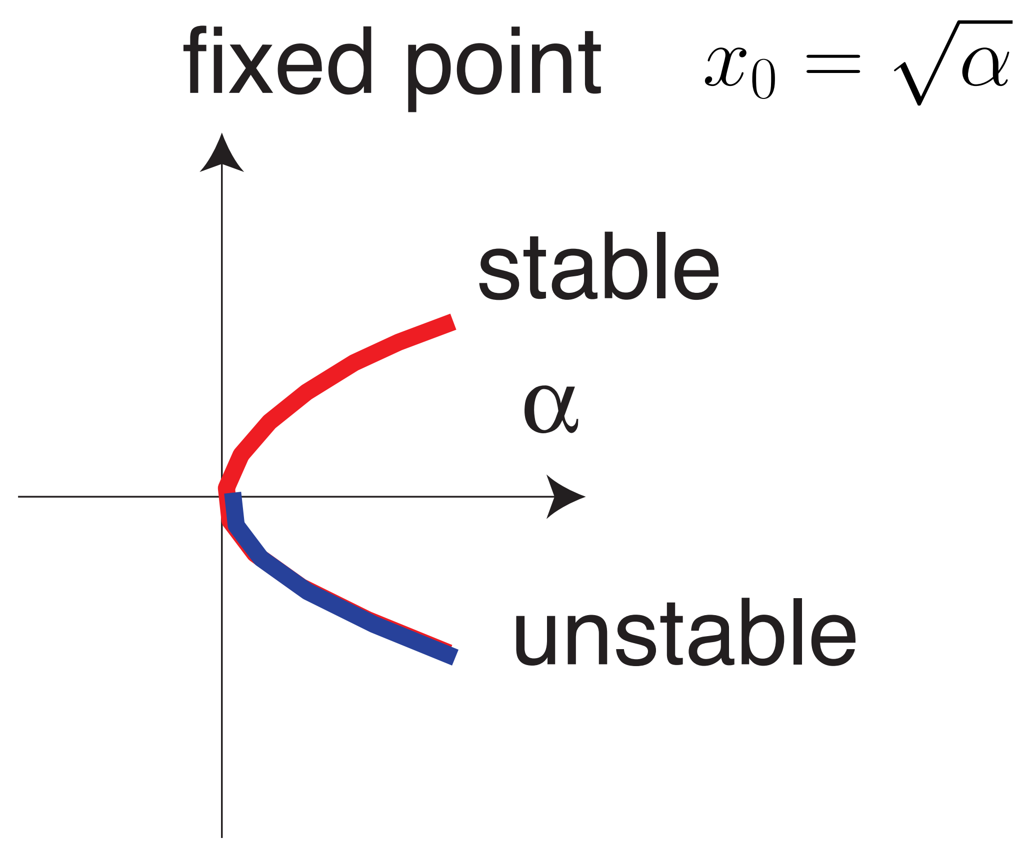


tangent bifurcation

- normal form of tangent bifurcation

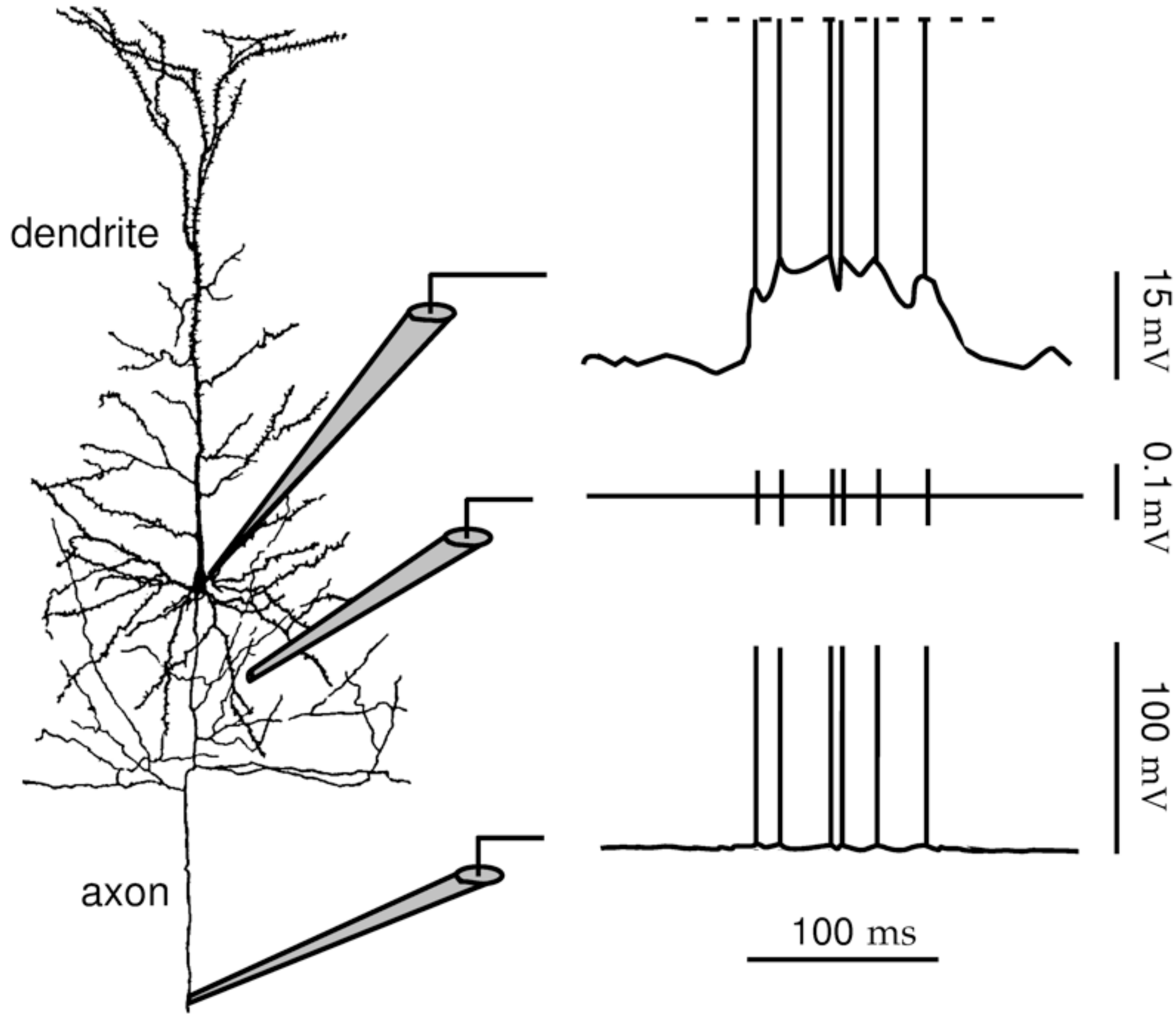
$$\dot{x} = \alpha - x^2$$



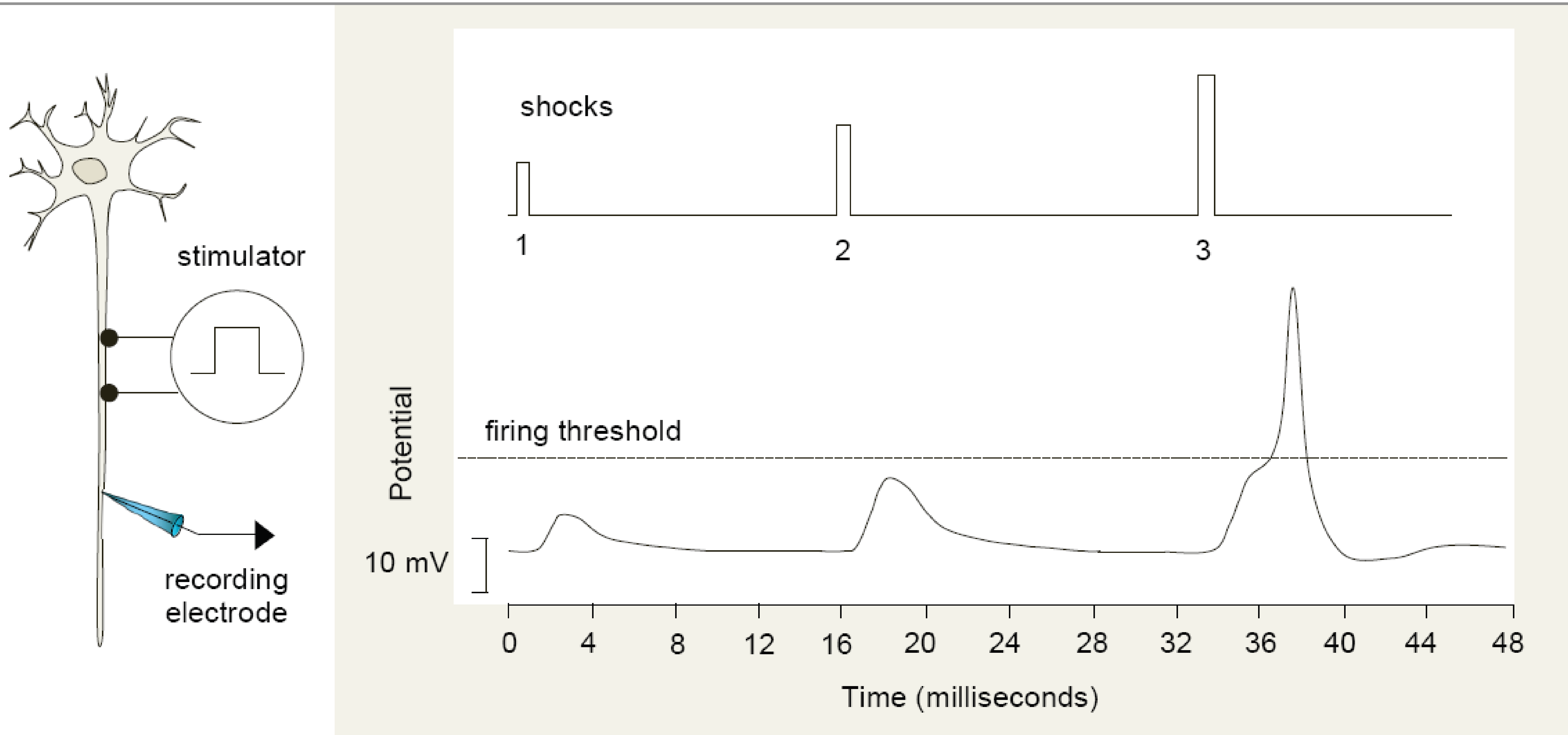


Neurons as input-output units

- inputs from dendrites
- spike formation at soma
- output at axon

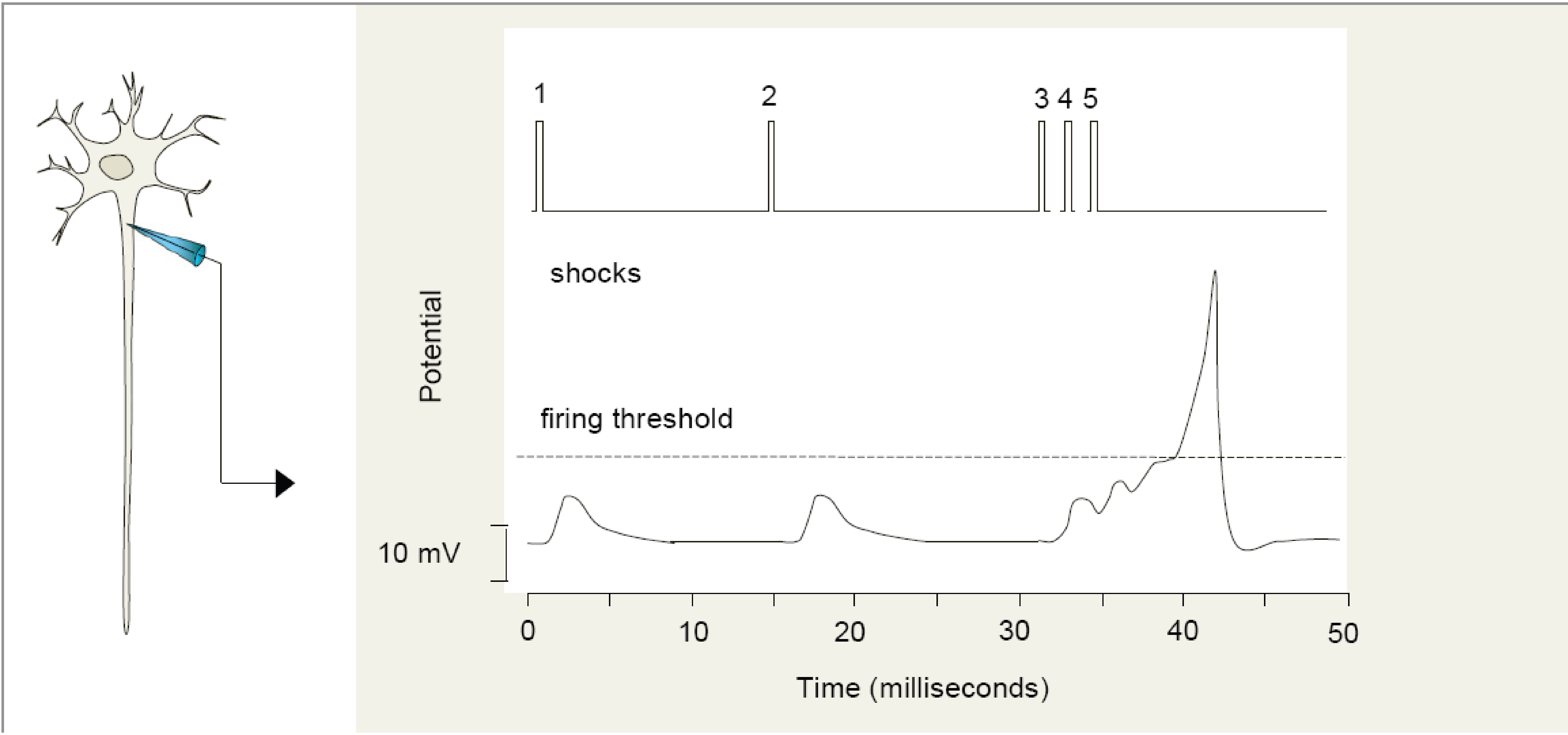


threshold behavior



$$\tau \dot{u} = -u + h + S(t) \qquad \tau \approx 5 - 10 \text{ ms}$$

temporal summation

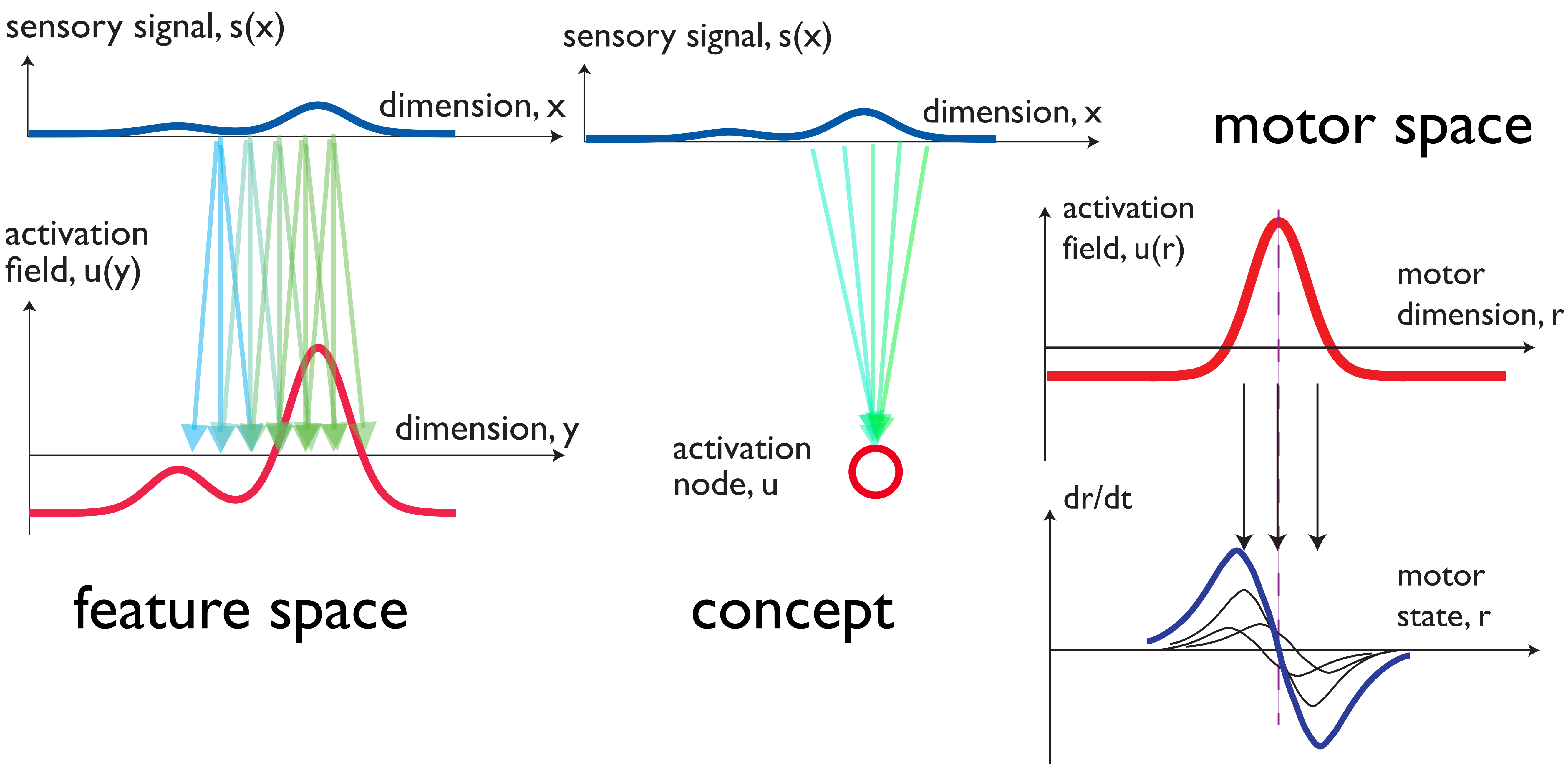


Foundations 1a: Space and time

- Space: fields
- Time:
  - neural dynamics
  - interaction
- Instabilities
- Simulating instabilities

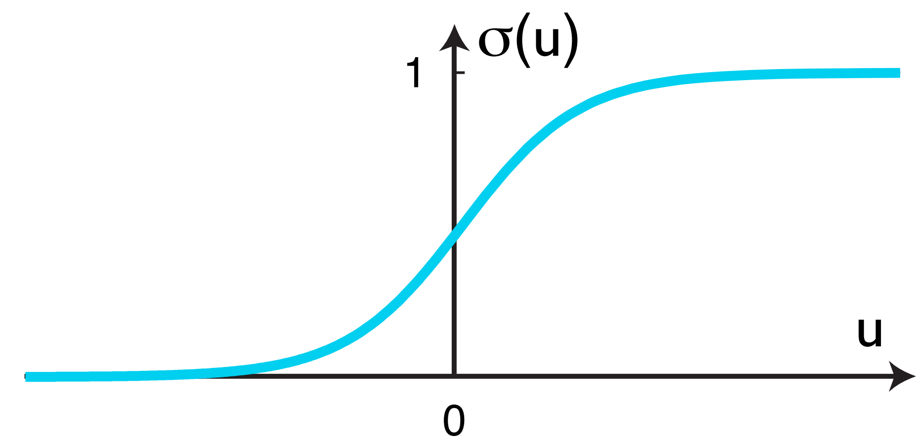
Where do the spaces come from?

- connectivity from sensory surfaces / to motor surfaces



Neural dynamics

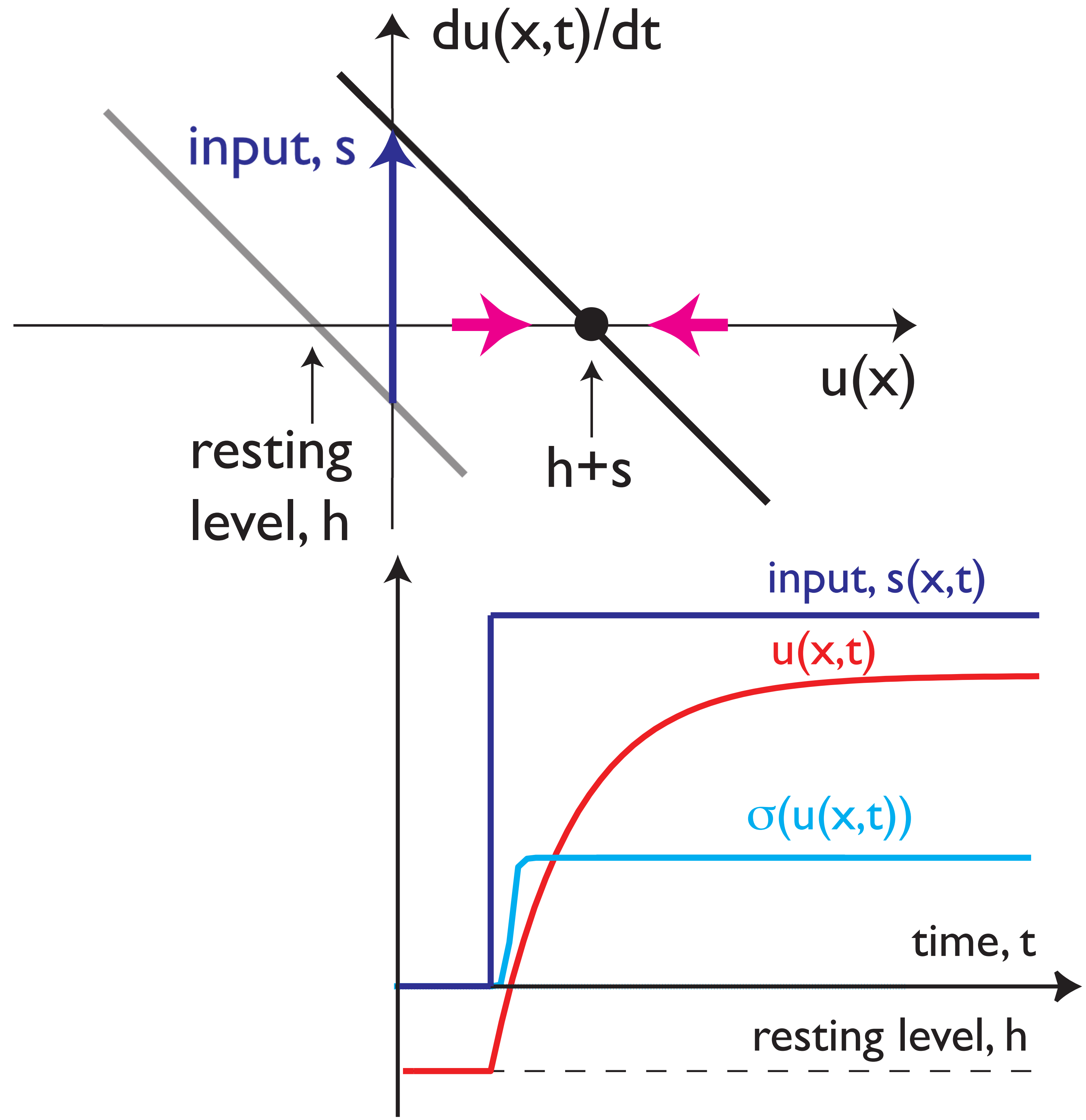
- activation u ~ population level membrane potential
- defined relative to sigmoid
  - above threshold: transmitted
  - below threshold: not transmitted



Neural dynamics

- activation dynamics = neural dynamics
- originates from membrane dynamics
- inputs as "forces"
    - positive: excitatory
    - negative: inhibitory

$$\tau \dot{u}(x,t) = -u(x,t) + h + s(x,t)$$



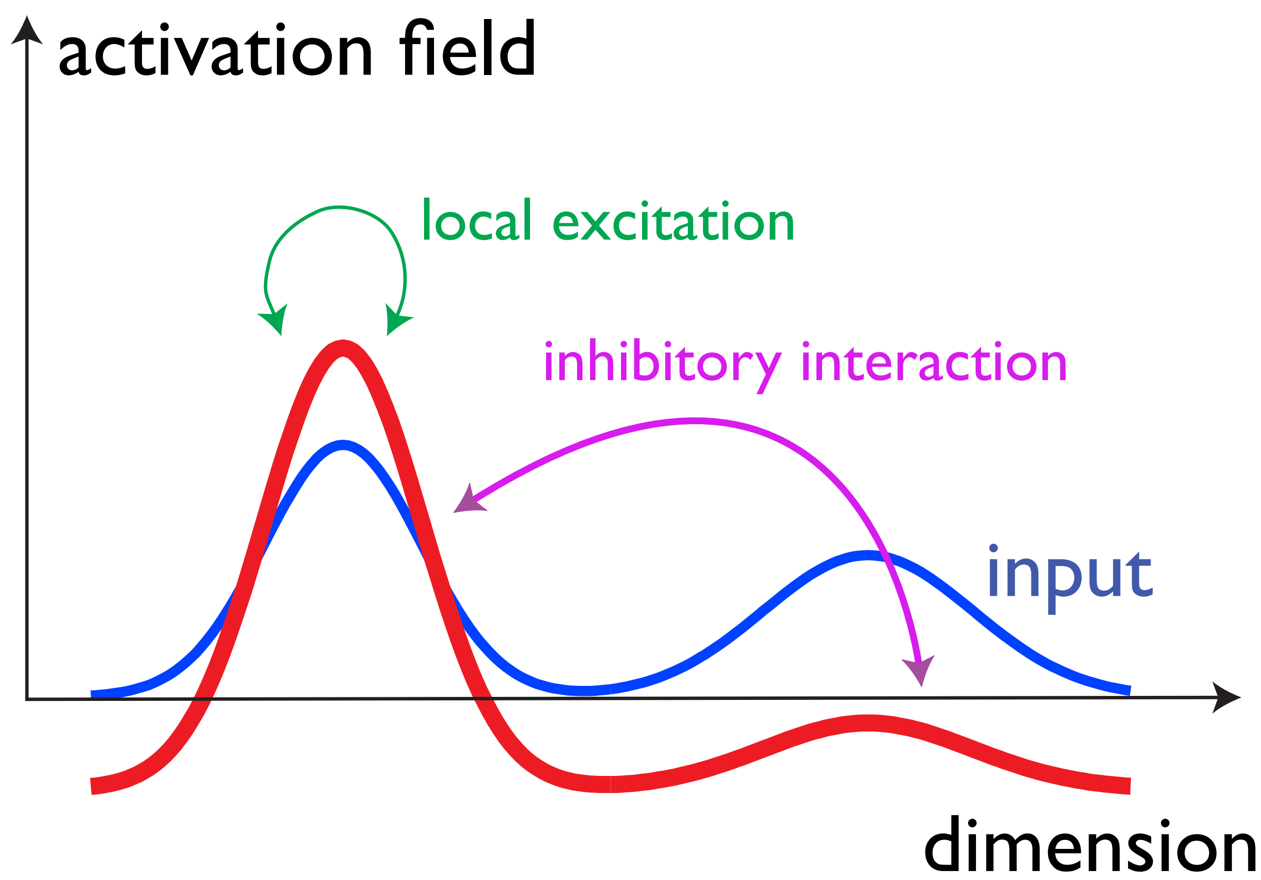
…beyond input driven activation

$$\tau \dot{u}(x,t) = -u(x,t) + h + s(x,t)$$

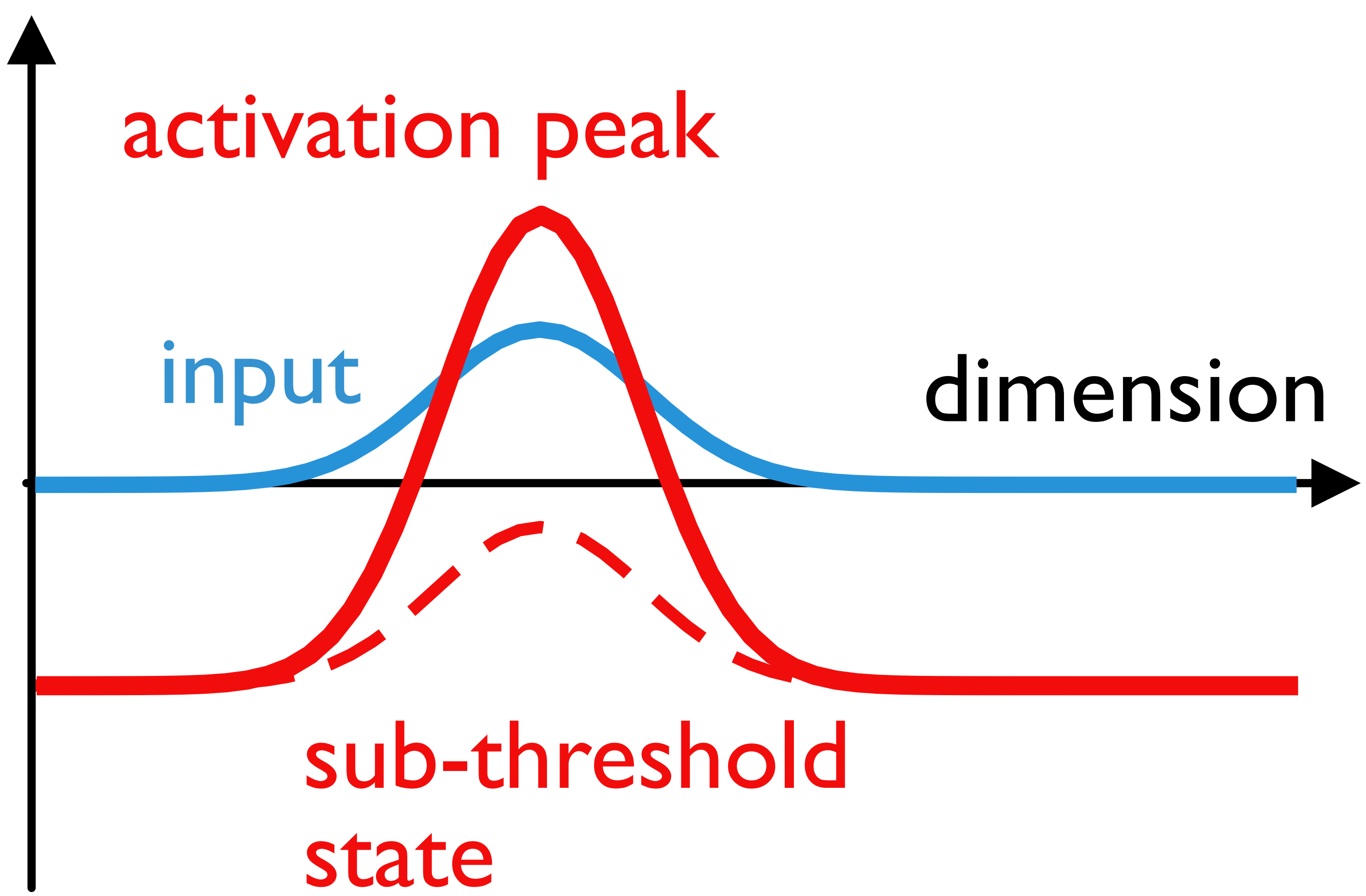
$$\boxed{+ \int w(x-x')\sigma(u(x',t))dx'}$$

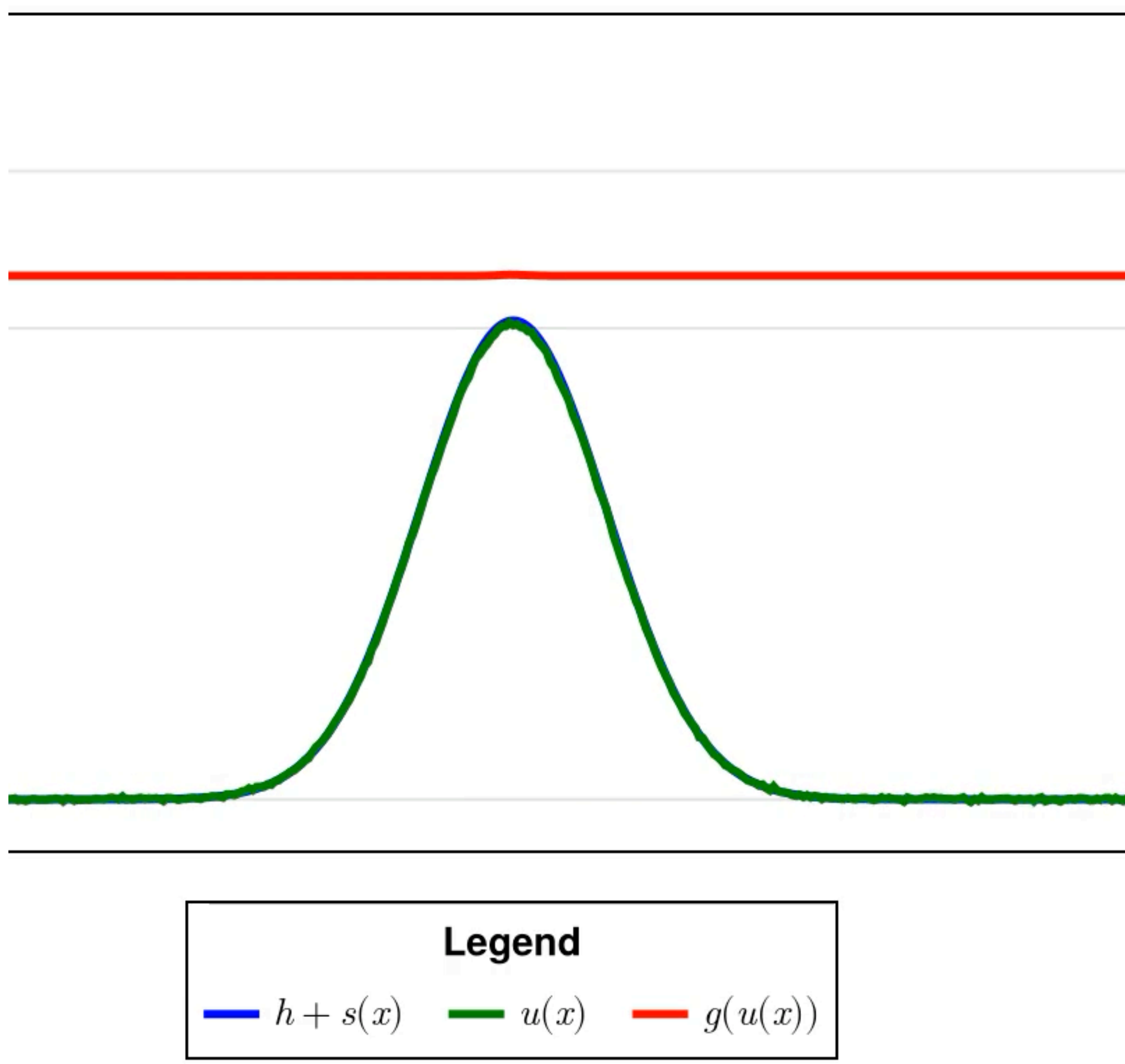
interaction

- strong recurrent connectivity within populations
- excitatory for neighbors in space
- inhibitory for activation at a spatial distance

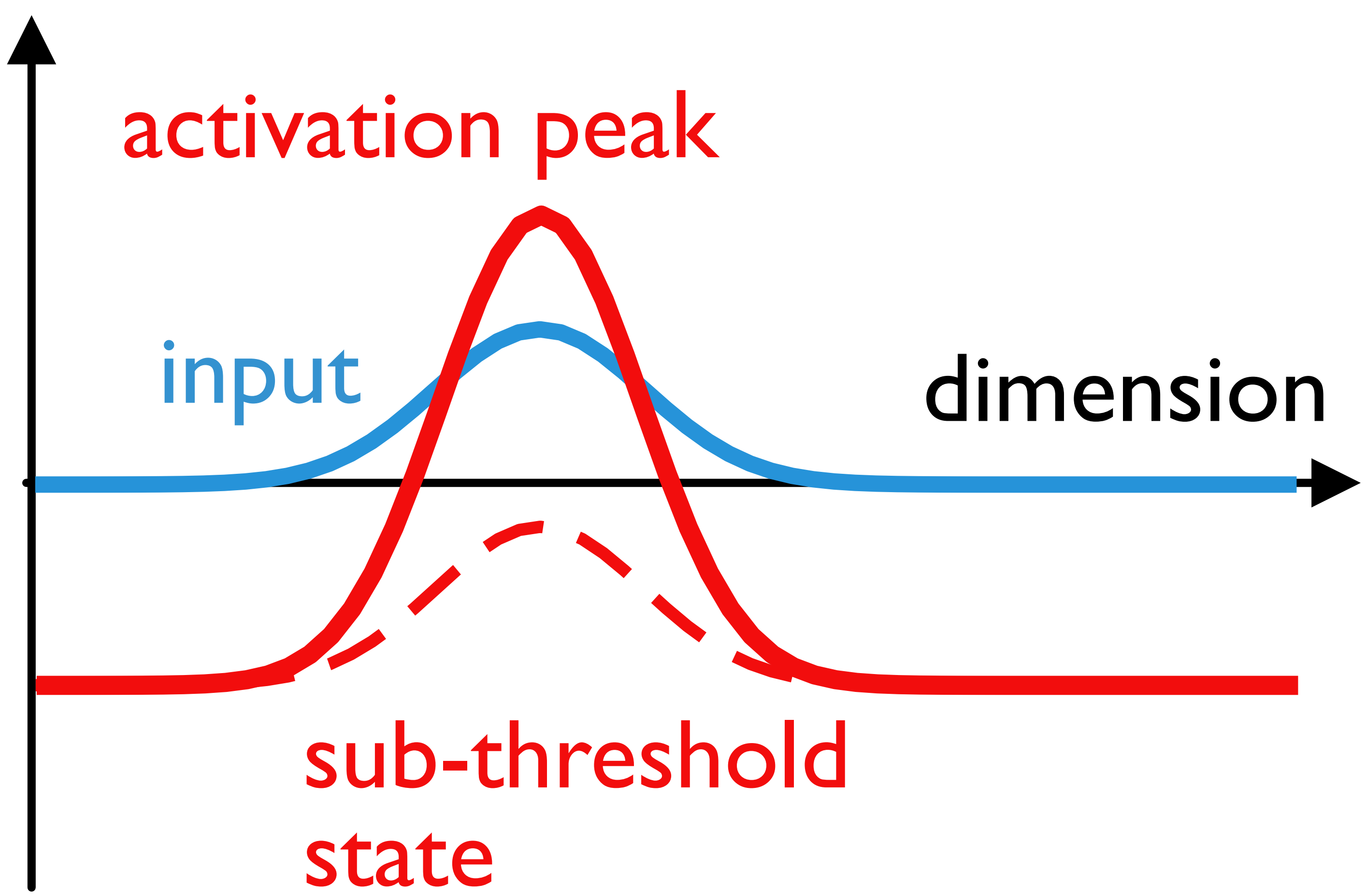


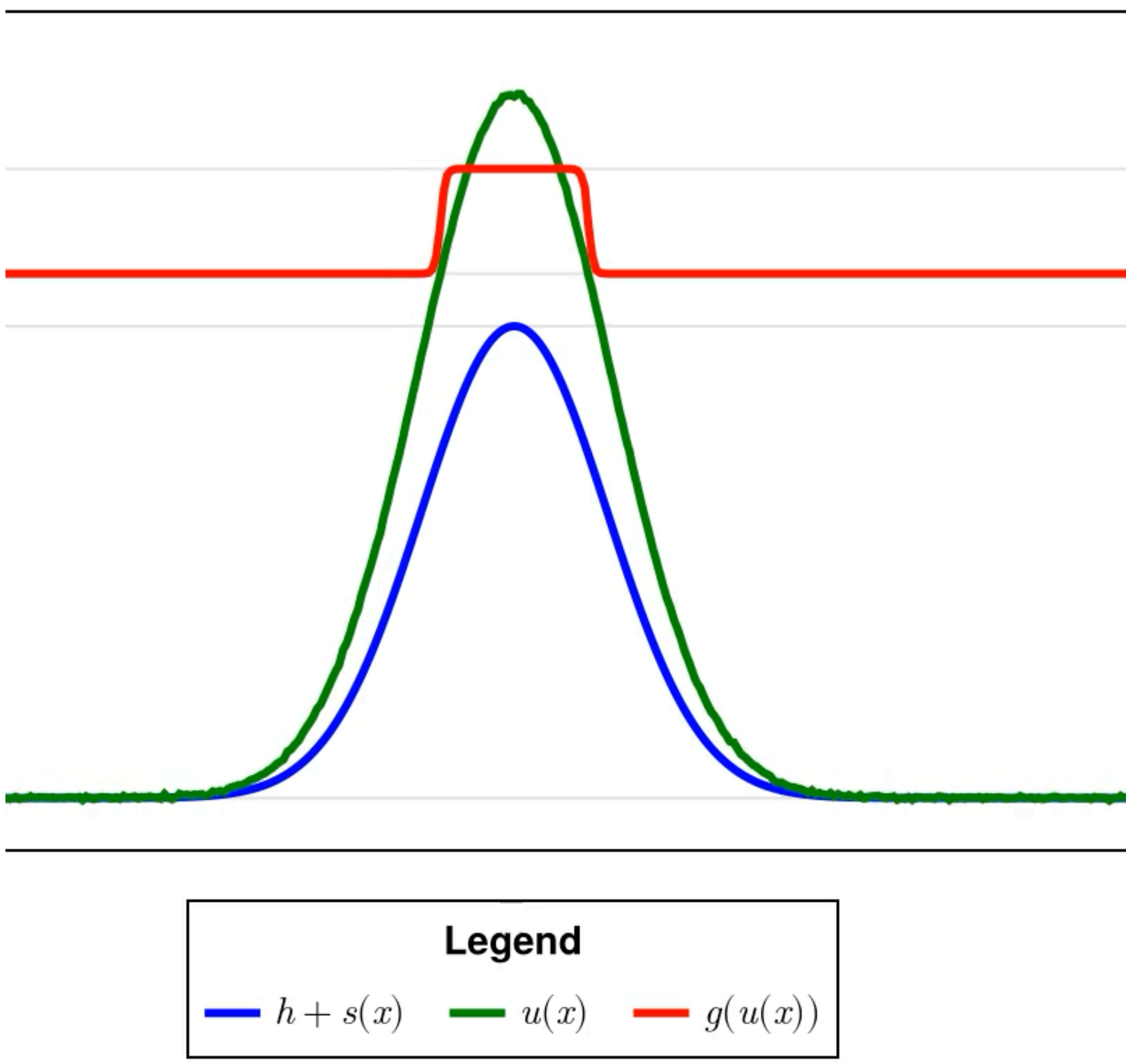
- detection instability of sub-threshold state=> switch to peak
- peak persists below detection instability => bistable





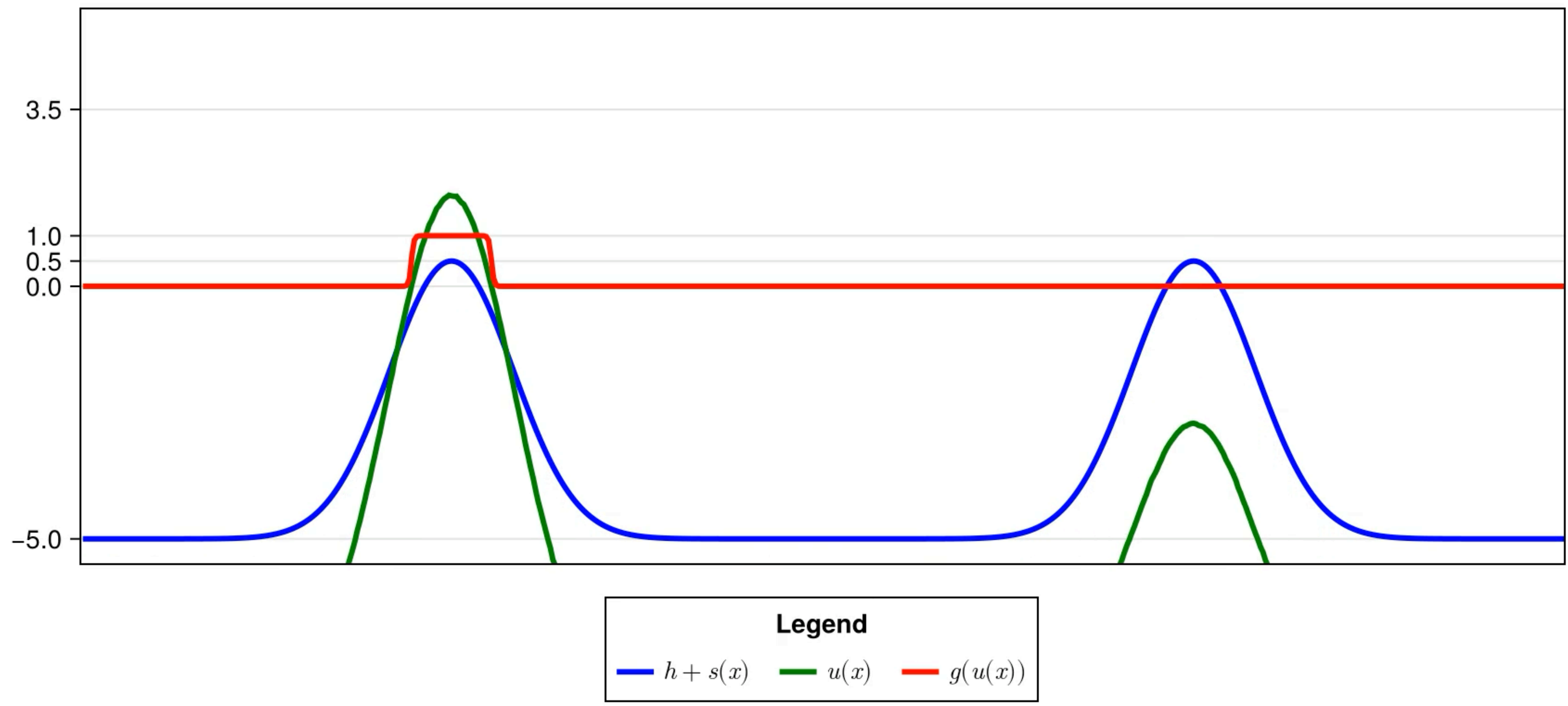
reverse detection instability of peak

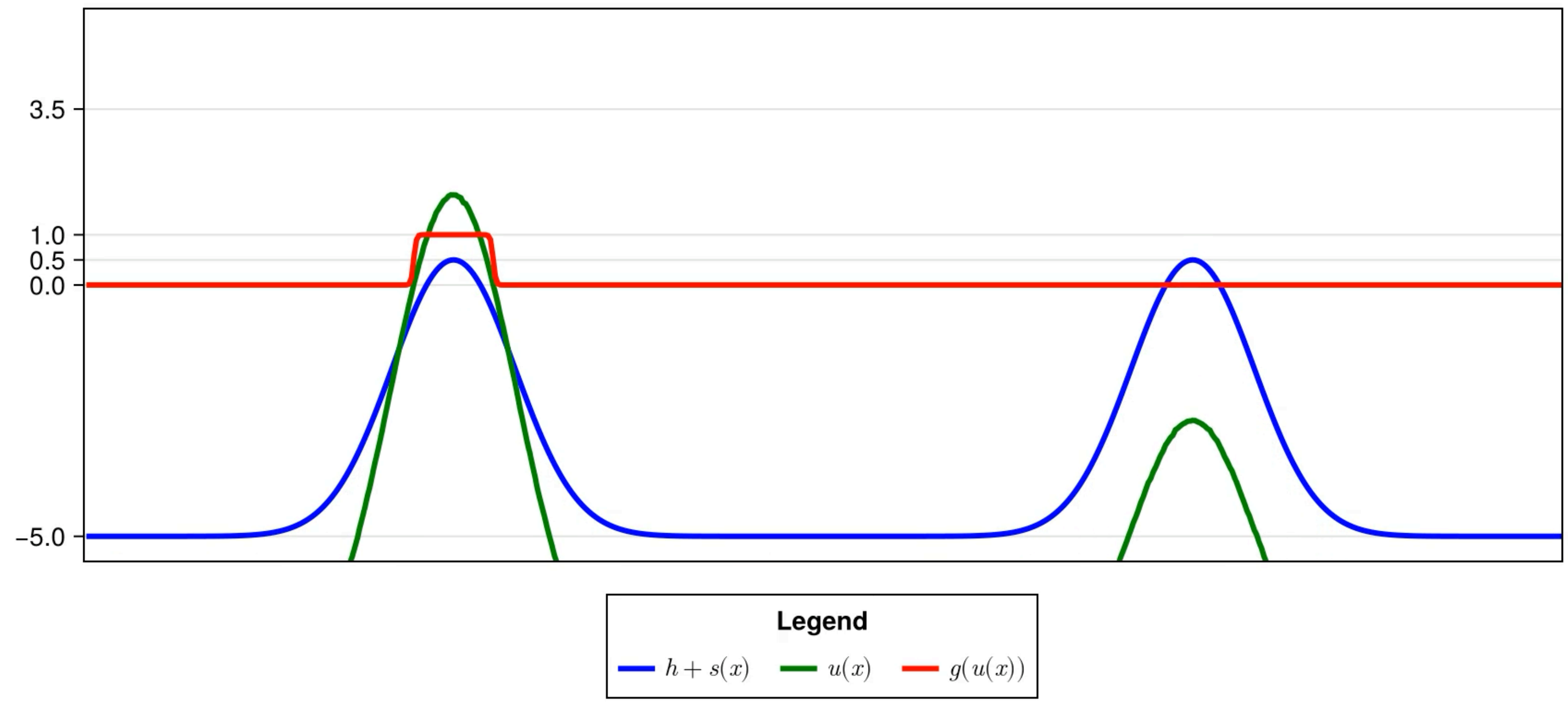




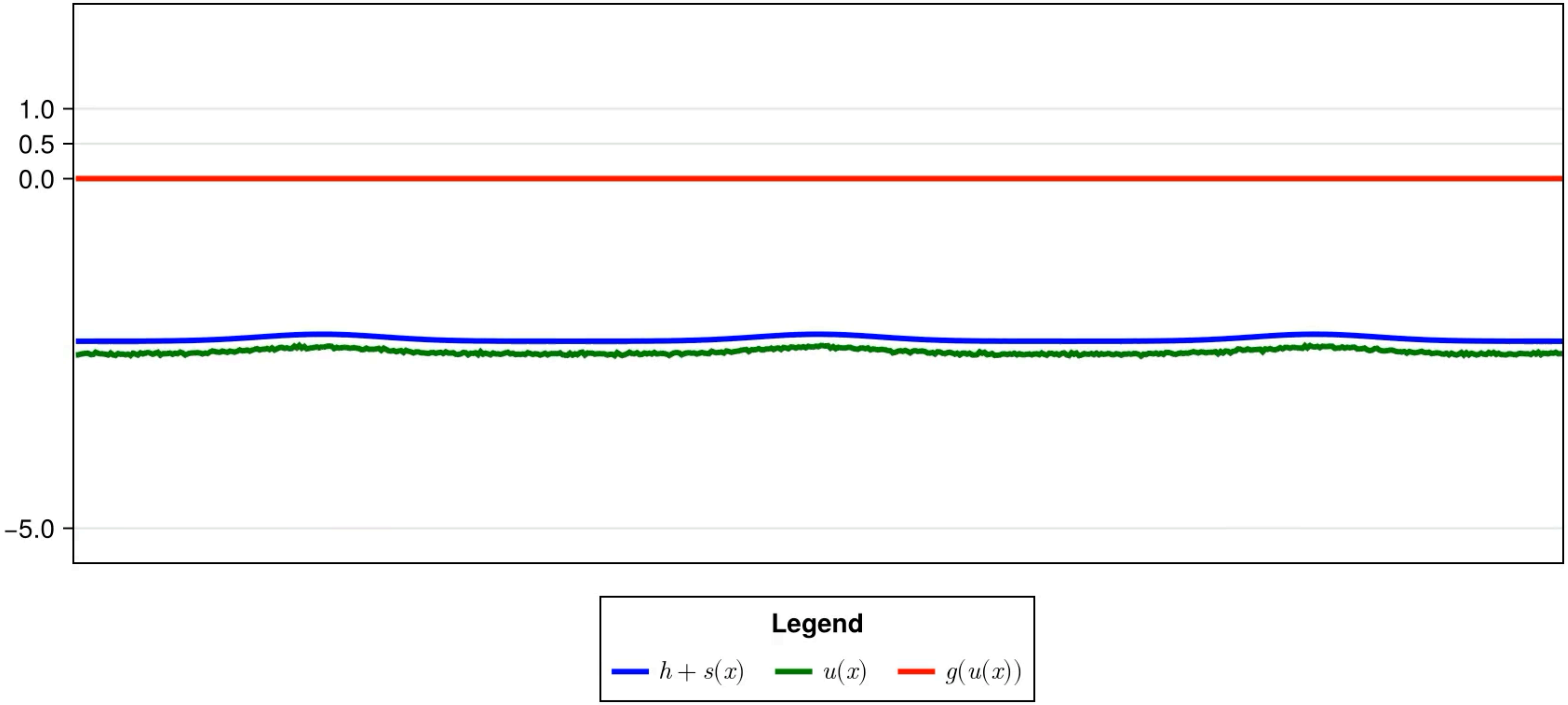
- sustained activation
- ~working memory

- selection
- selection instability





- detection and selection induced by homogeneous boost
- => amplify small inhomogeneities

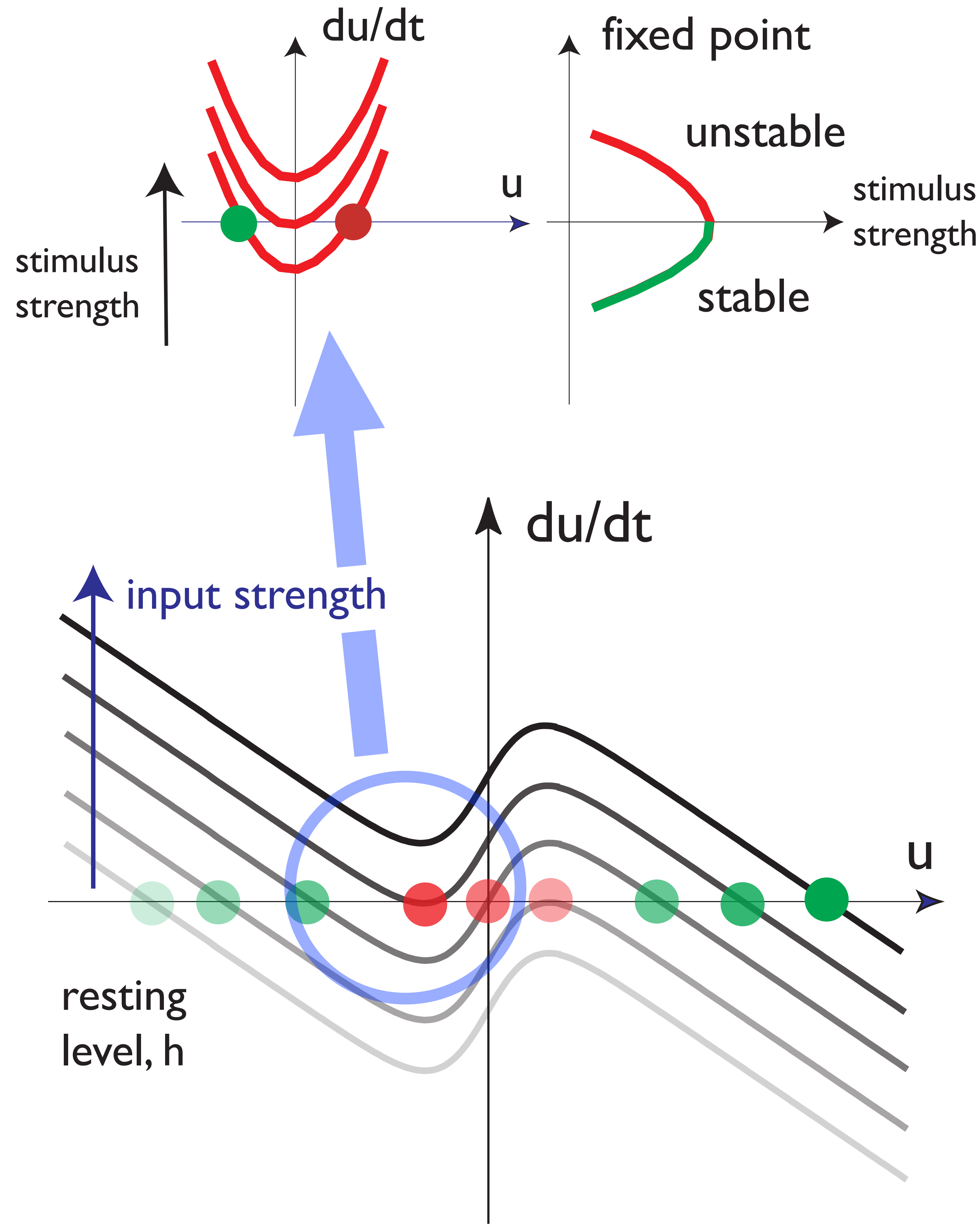


Foundations 1b: Mathematical analysis

- Discretization of fields
- Self-excitation
- Inhibitory interaction
- Mathematical formalization
- … beyond 1D fields

Neuronal dynamics with self-excitation

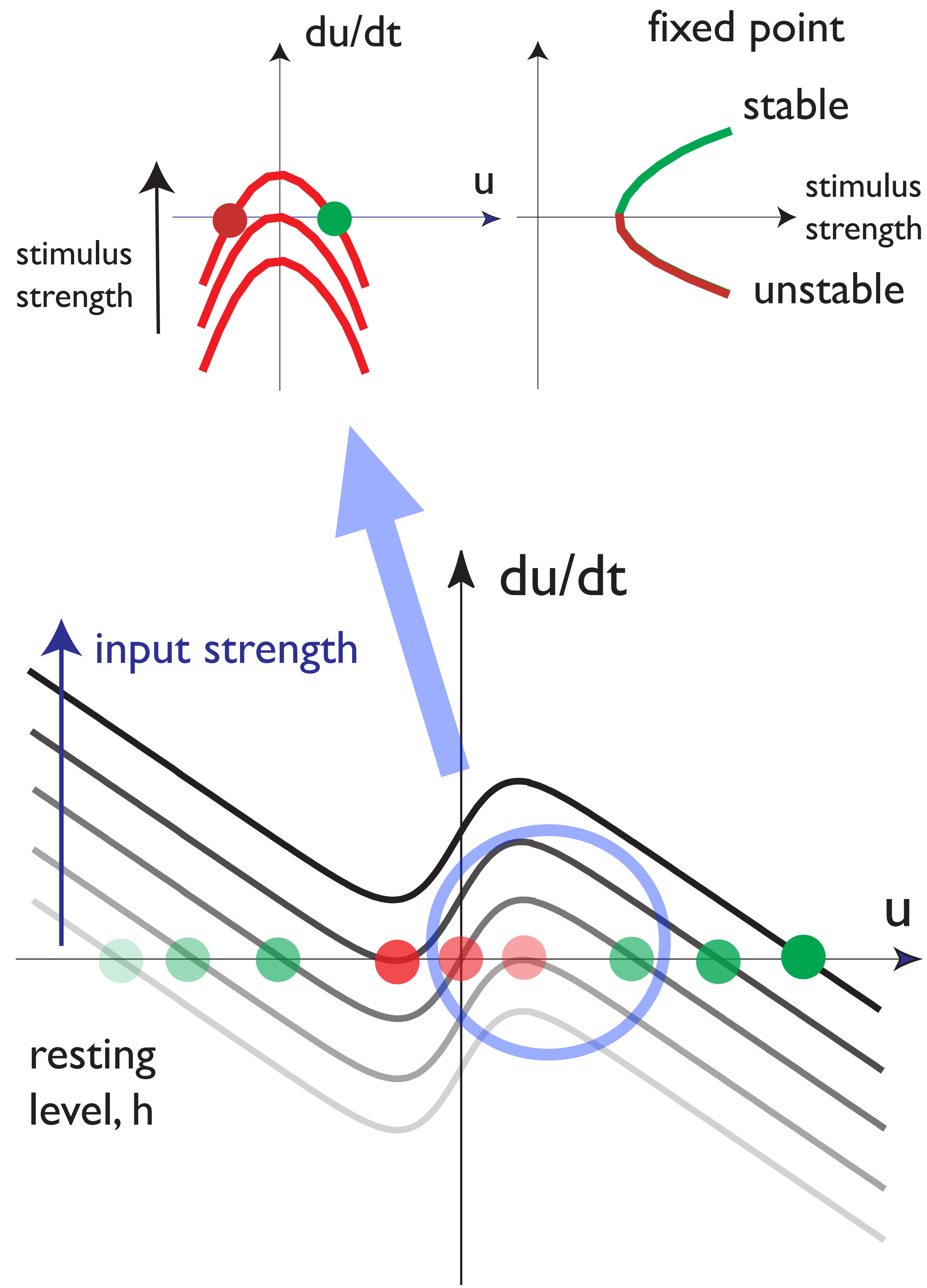
- increasing input strength => detection instability



$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

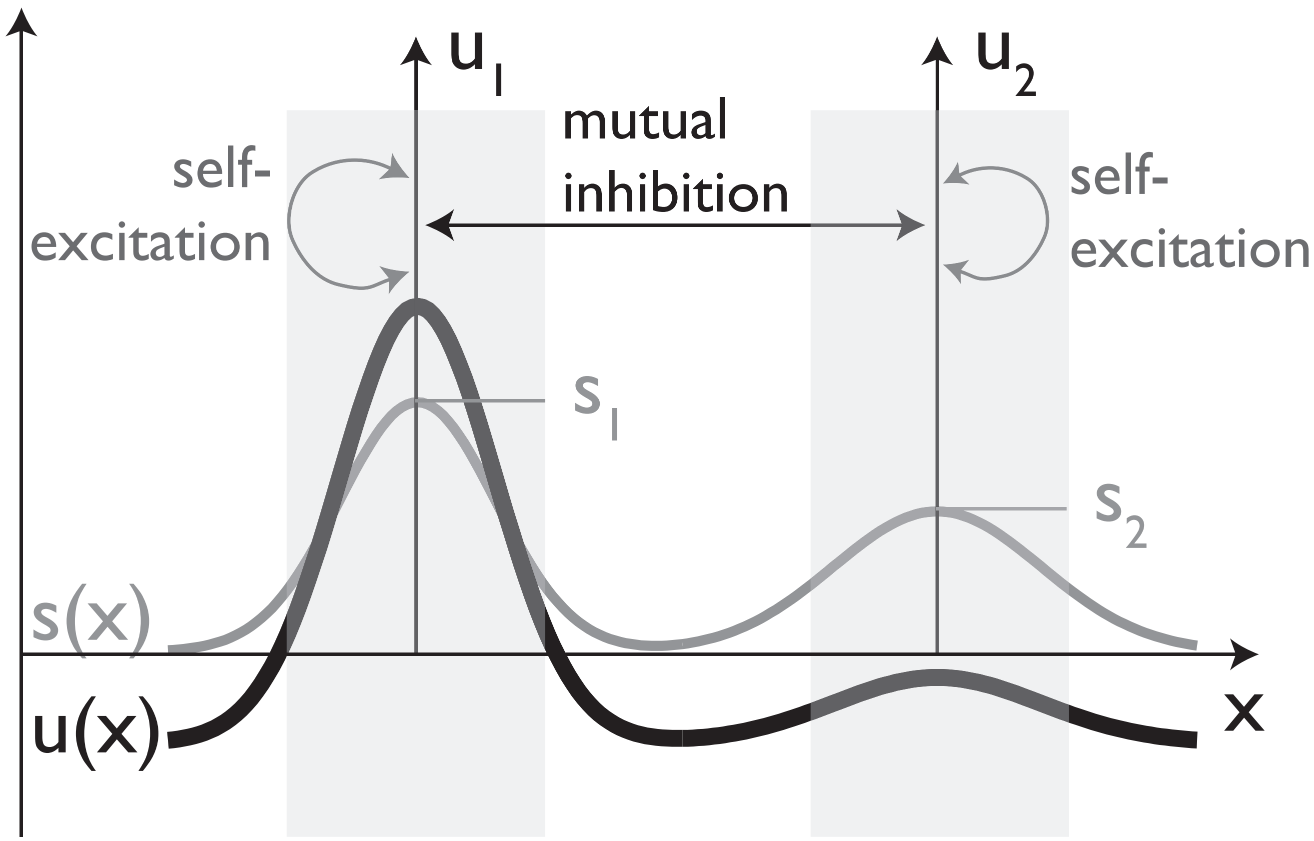
Neuronal dynamics with self-excitation

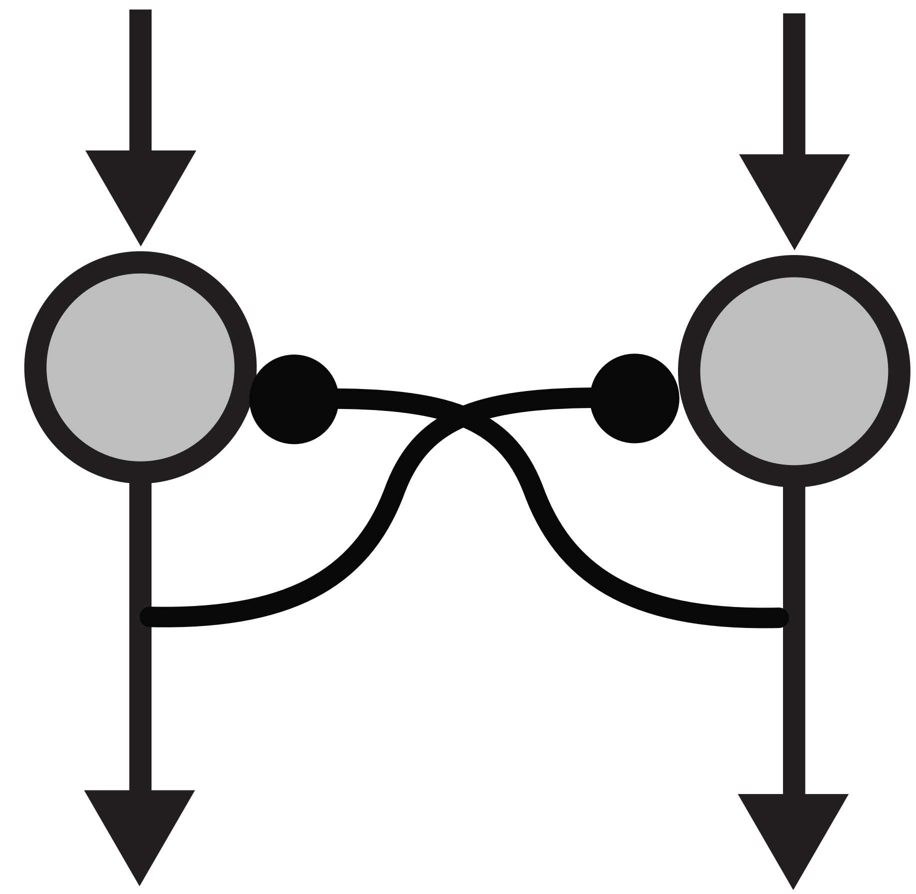
- decreasing input strength => reverse detection instability



$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Inhibitory interaction: inhibitory recurrent connectivity





coupling/interaction

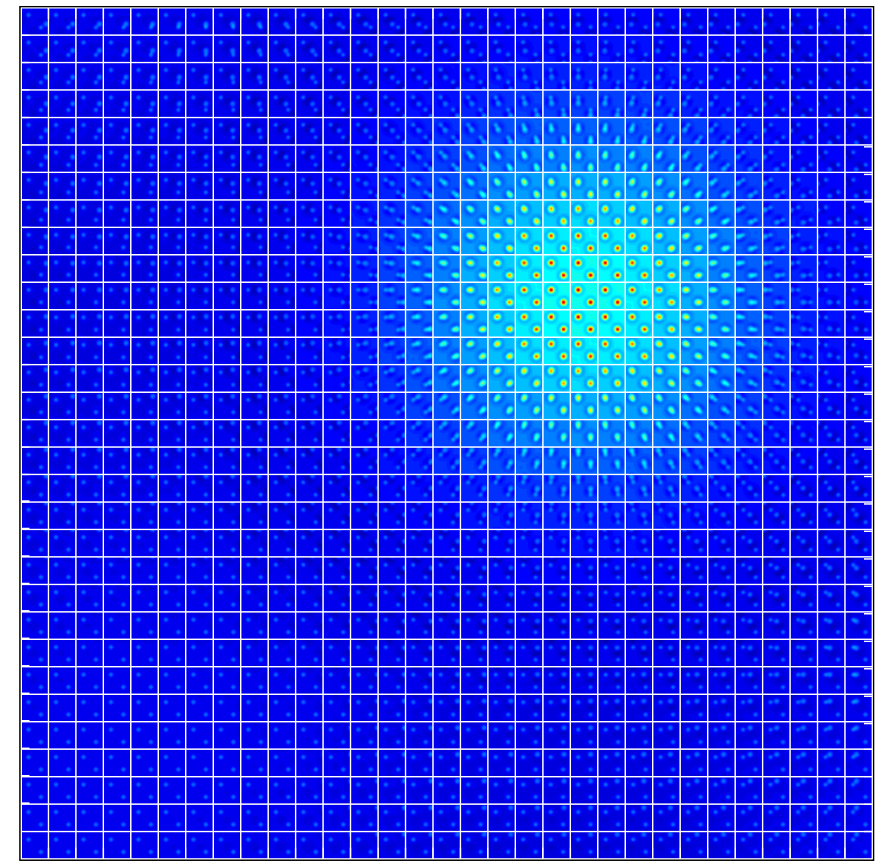
$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

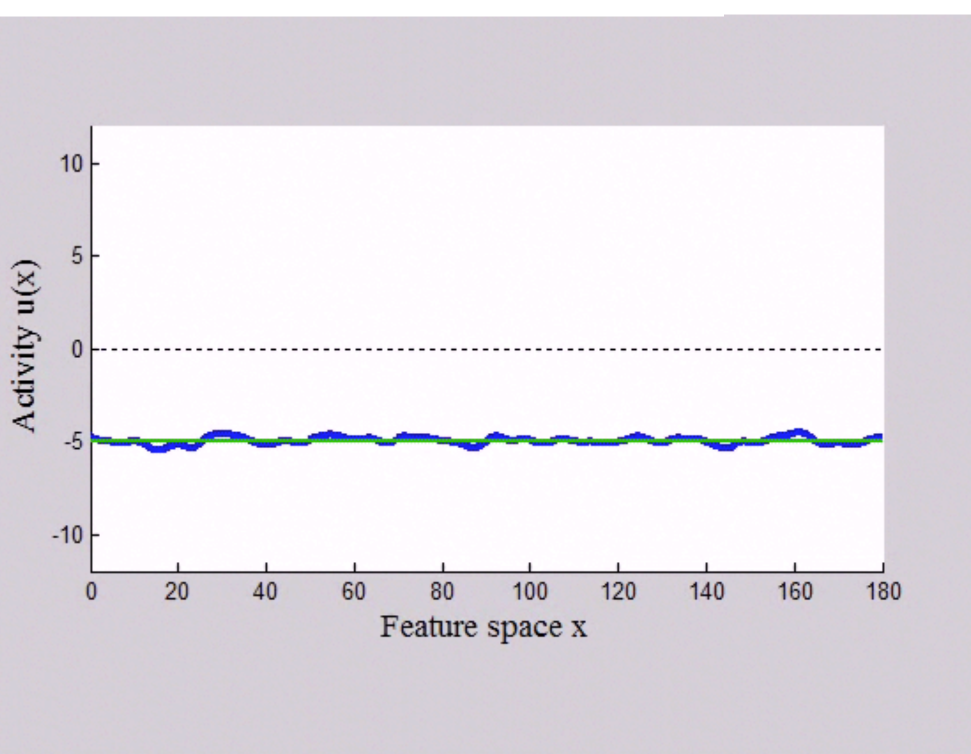
Field dynamics in different dimensions

- 1, 2, 3, 4… dimensions: peaks/ blobs as attractors

4-dimensional

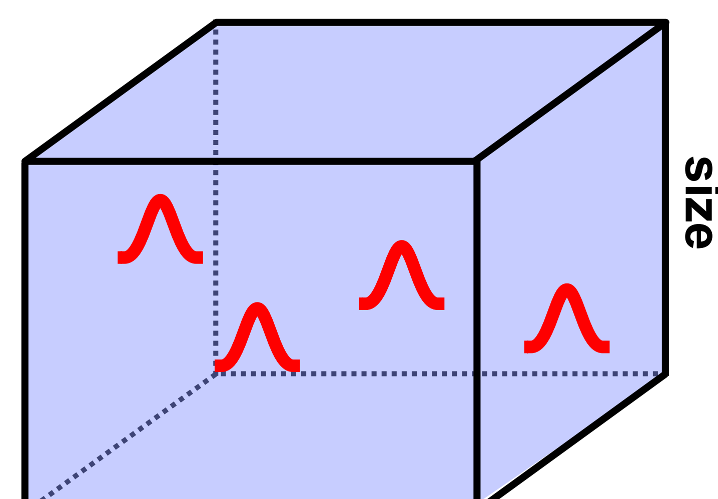


1-dimensional



2-dimensional

3-dimensional

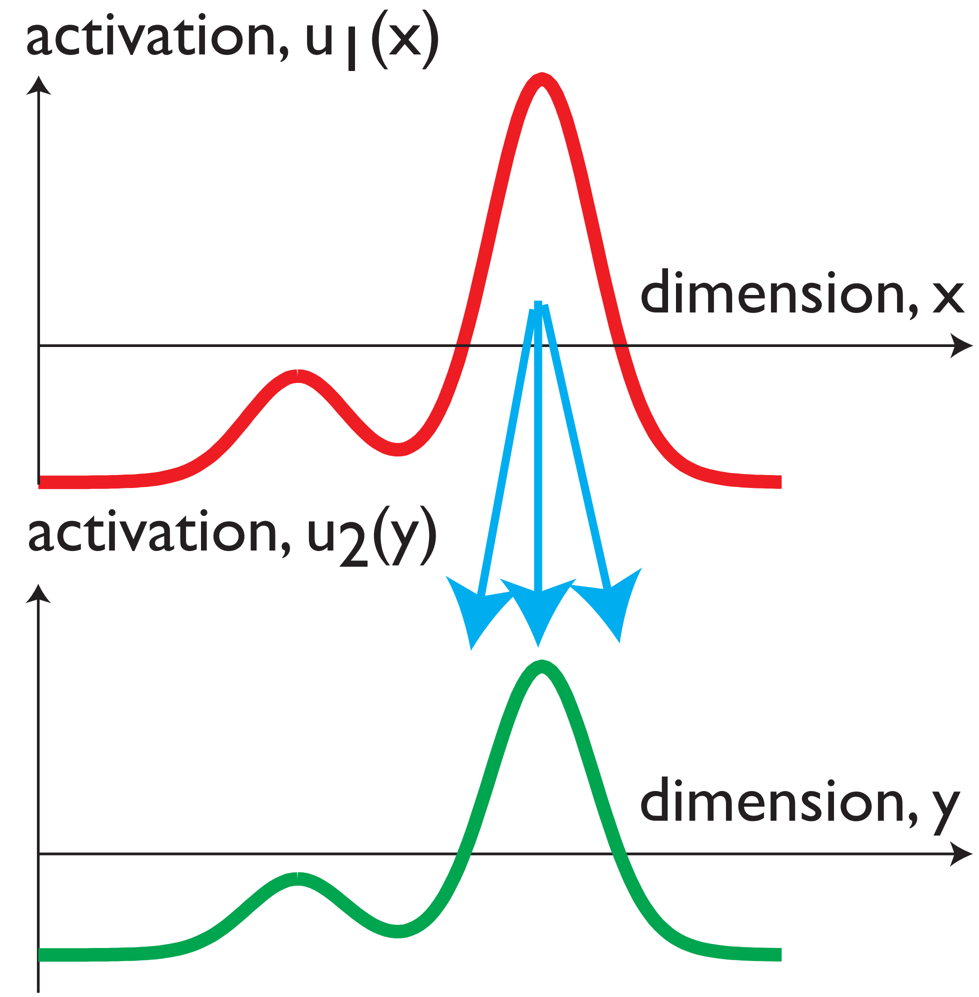


Learning

- Hebbian learning
- memory trace
- autonomous learning

Hebbian learning

- Hebbian learning of projections
    - among fields
    - forward from sensory input to fields
- interaction leads to localized rather than distributed representations (SOM)

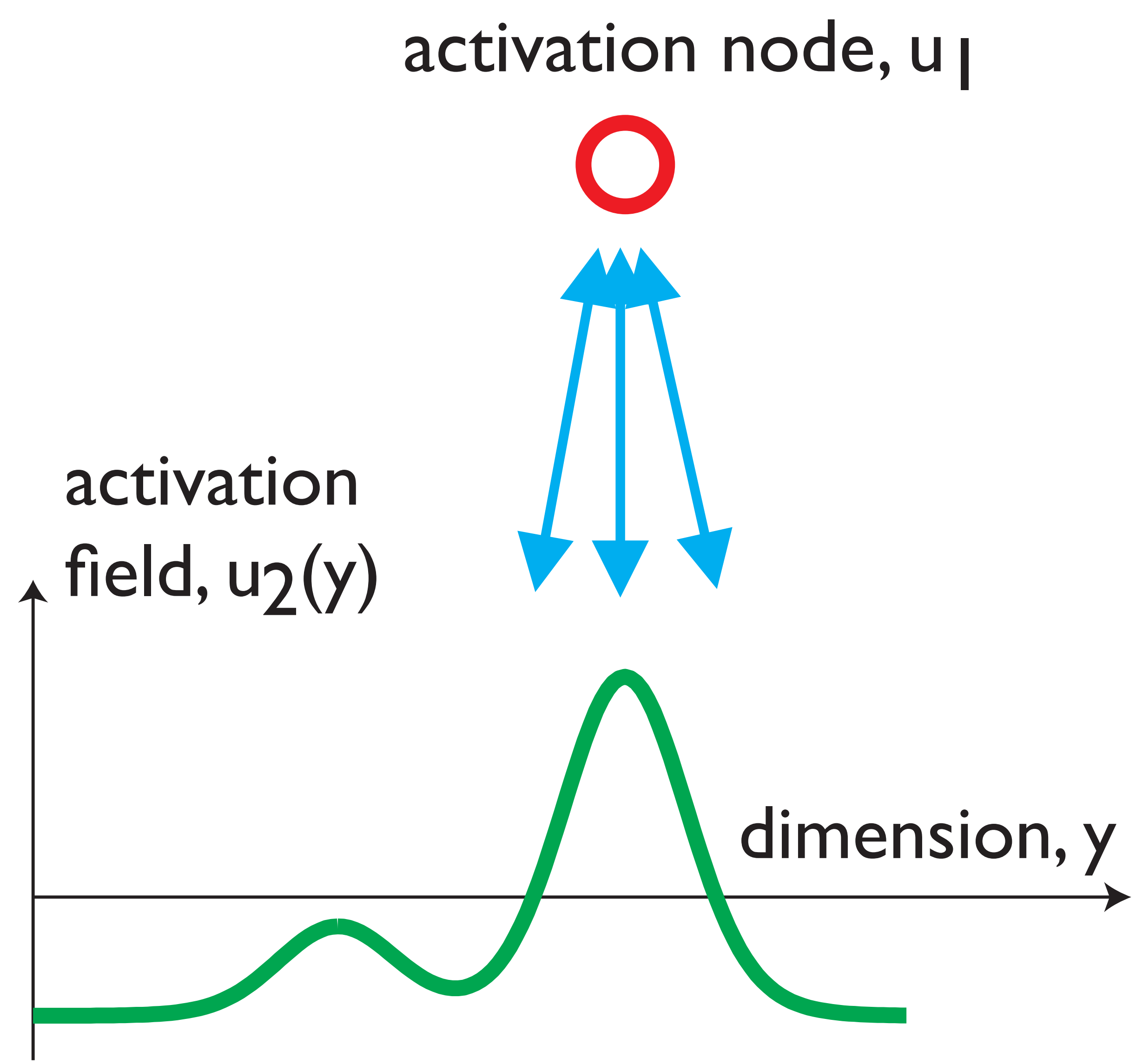


$$\tau \dot{W}(x, y, t) = \epsilon(t)\Big(- W(x, y, t) + f(u_1(x, t)) \times f(u_2(y, t))\Big)$$

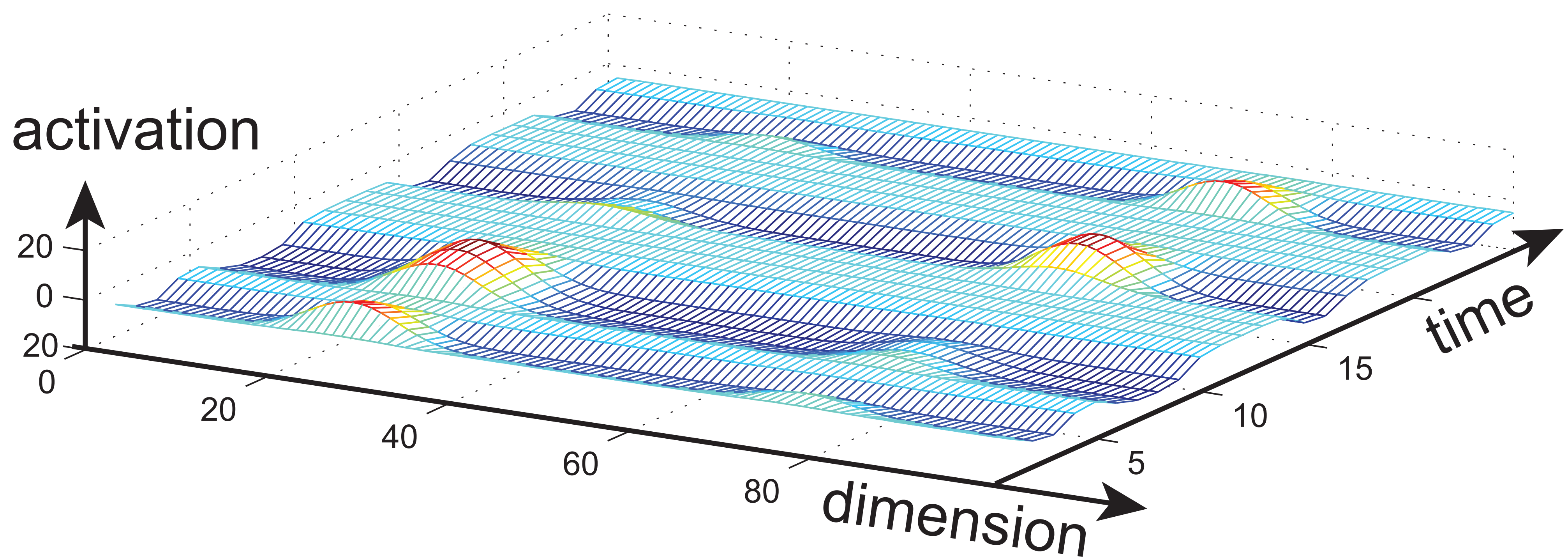
[Sandamirskaya, *Frontiers Neurosci* 2014]

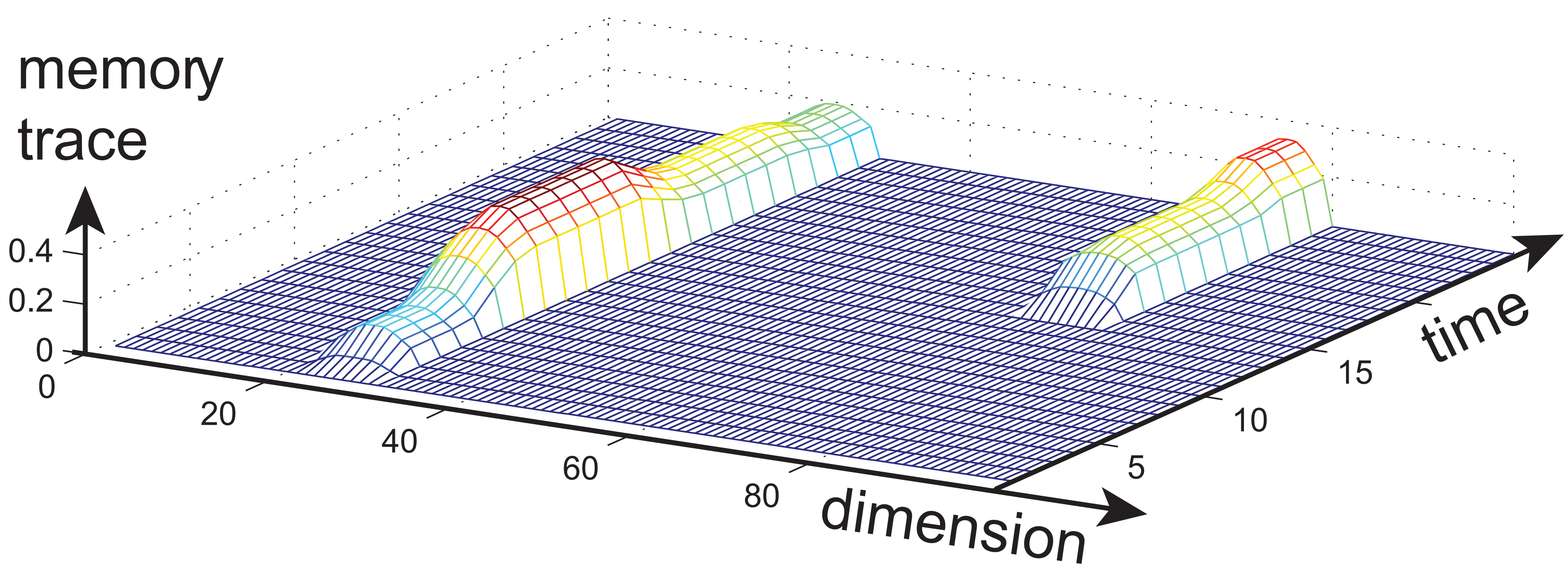
Hebbian learning

- learning reciprocal connections between zero-dimensional nodes and fields
- => grounded concepts
- analogous to the output layer of DNN
- => ensembles of such nodes coupled inhibitorily form the basis for conceptual thinking…



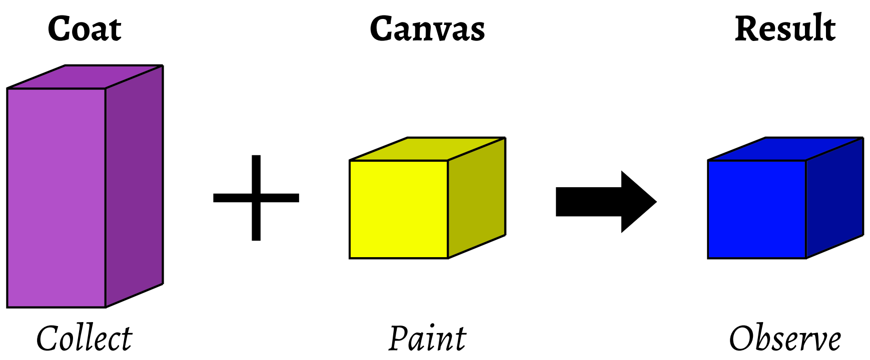
=> the memory trace reflects the history of detection decisions

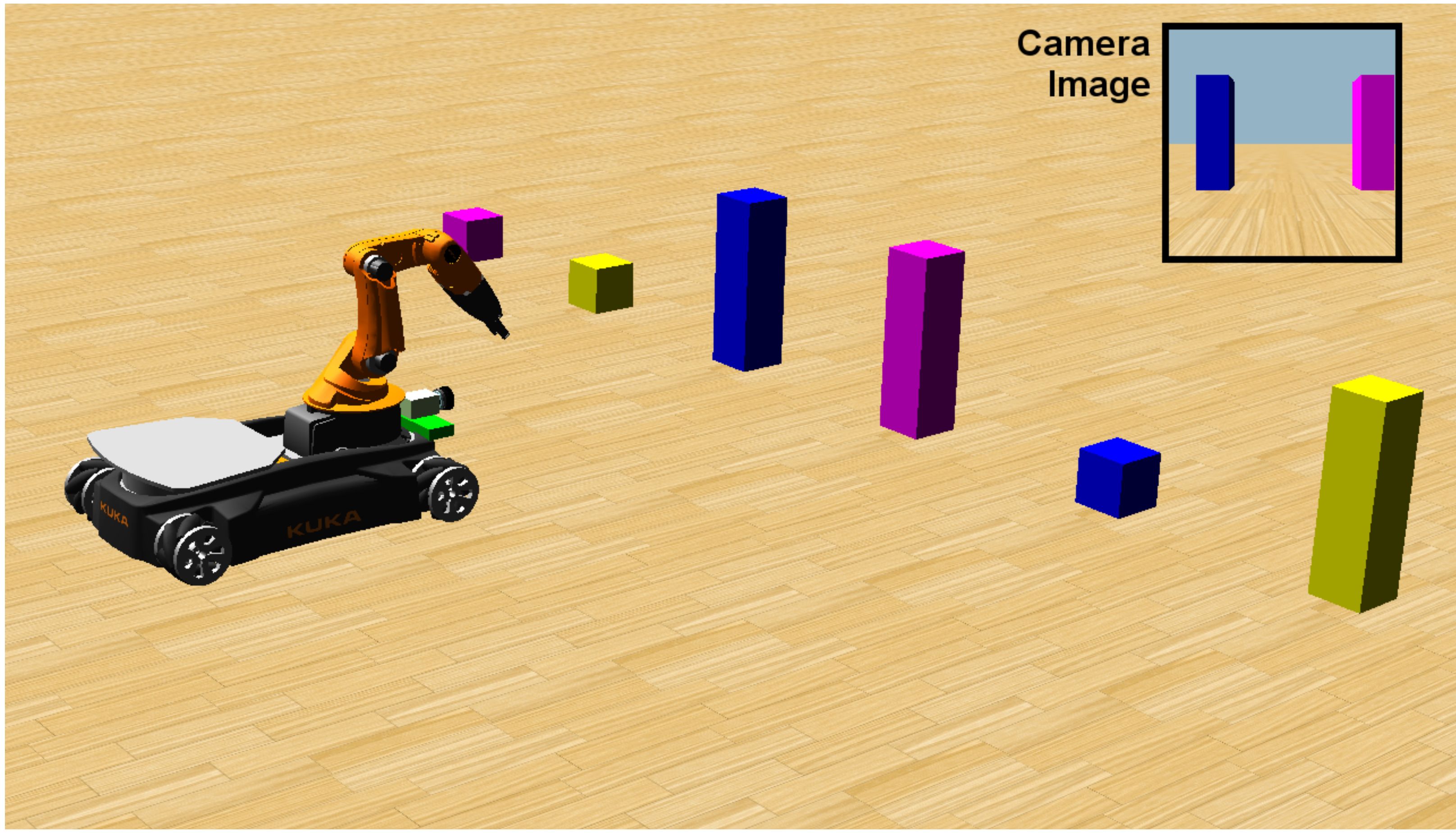




Autonomous learning

- learning regularities in the world (contingencies) by acting on the world

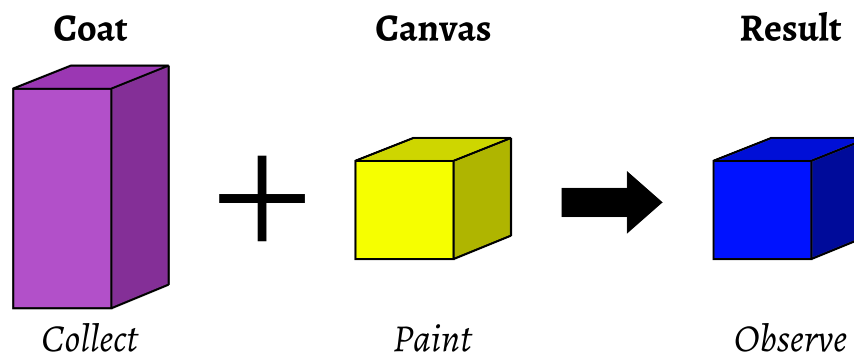


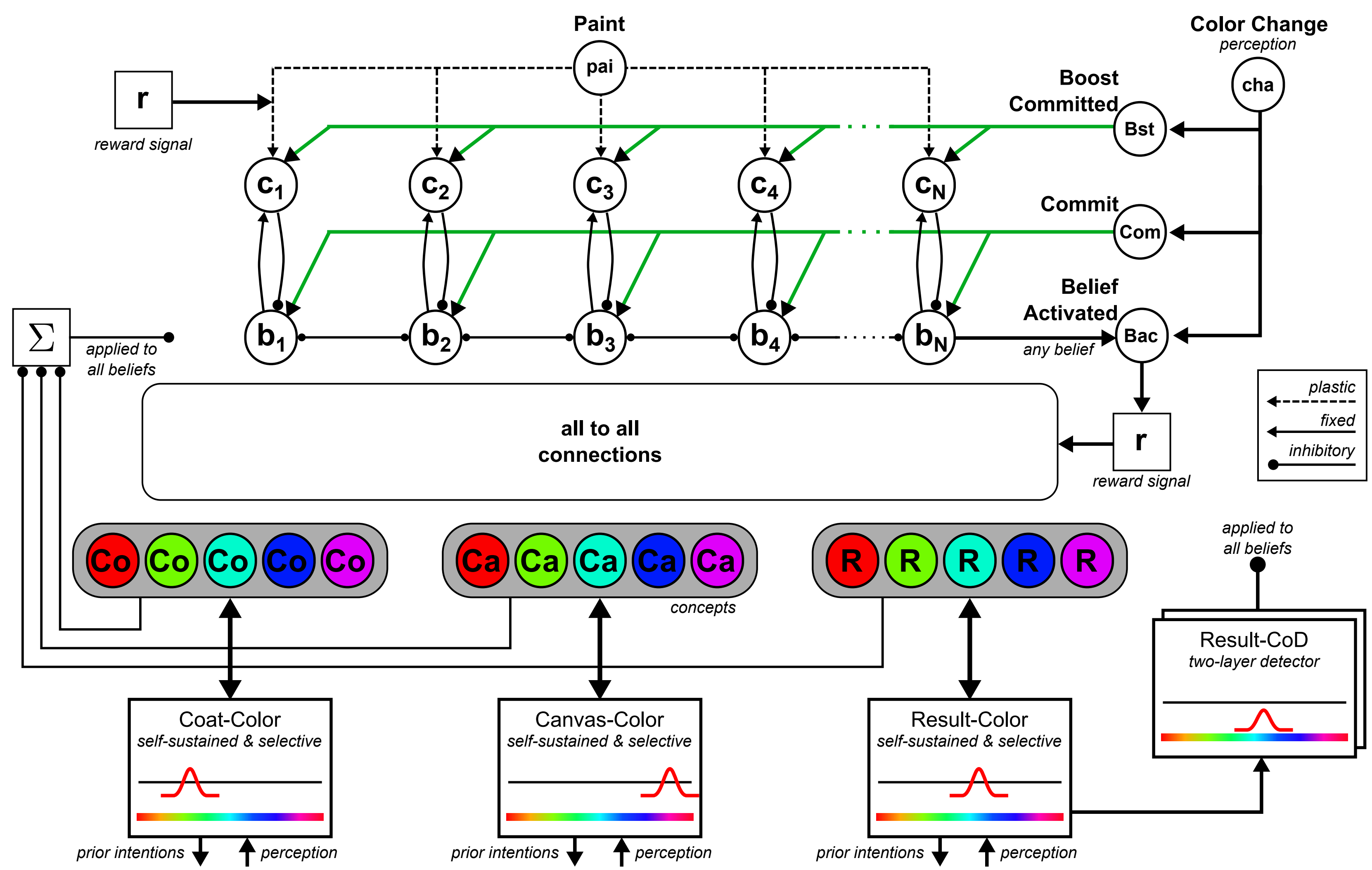


[Tekülve, Schöner, IEEE Trans Cog Dev Sys 2022;
Tekülve, Schöner Cog Science, in press (2024)]

Autonomous learning

- contingency represented in a "belief network
- learning rate modulated: learning "event" (opportunity)





[Tekülve, Schöner, *IEEE Trans Cog Dev Sys* 2022;
Tekülve, Schöner *Cog Science* (2024)]

Experimental signatures

- metric effects: distances between potential states matter
- effects of timing: time matters, spatio-temporal co-variation
- instabilities: it matters how far a state is from becoming unstable…

Foundations 2: Space-time coupling

- Background: different notions of binding
- Joint representations and coupling patterns
- Binding through space/ordinal dimension
- Coordinate transforms

Intuition for "binding"

red cutter horizontally aligned

Binding

- features, category, and location are all “bound” together..

Binding is flexible

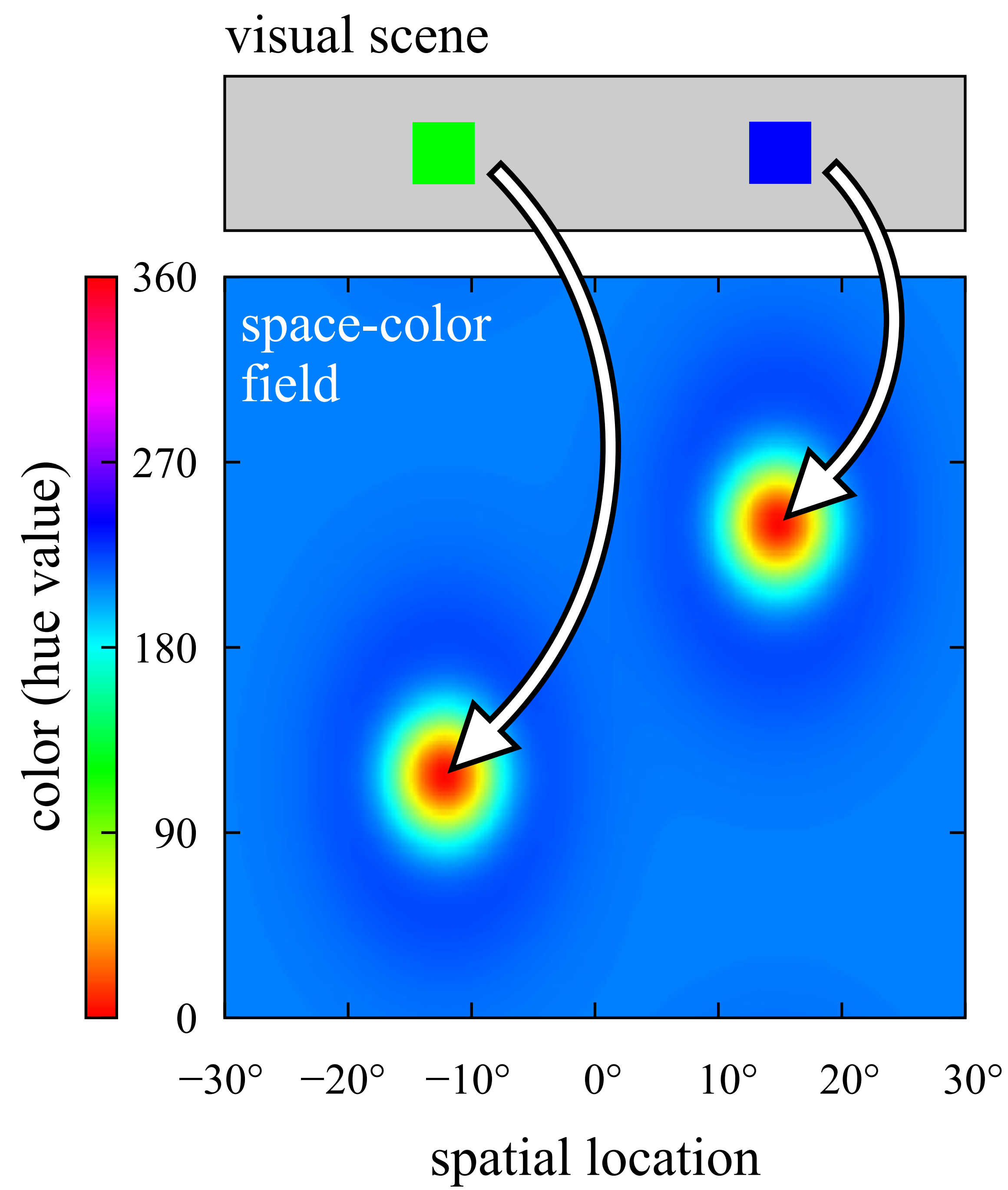
- feature combinations never seen before may be bound
- mis-bindings may occur in "illusory conjunctions" (e.g. Treisman, 98)

the "S" is green

[Chapter 5, DFT book 2016]

Joint space-feature representation

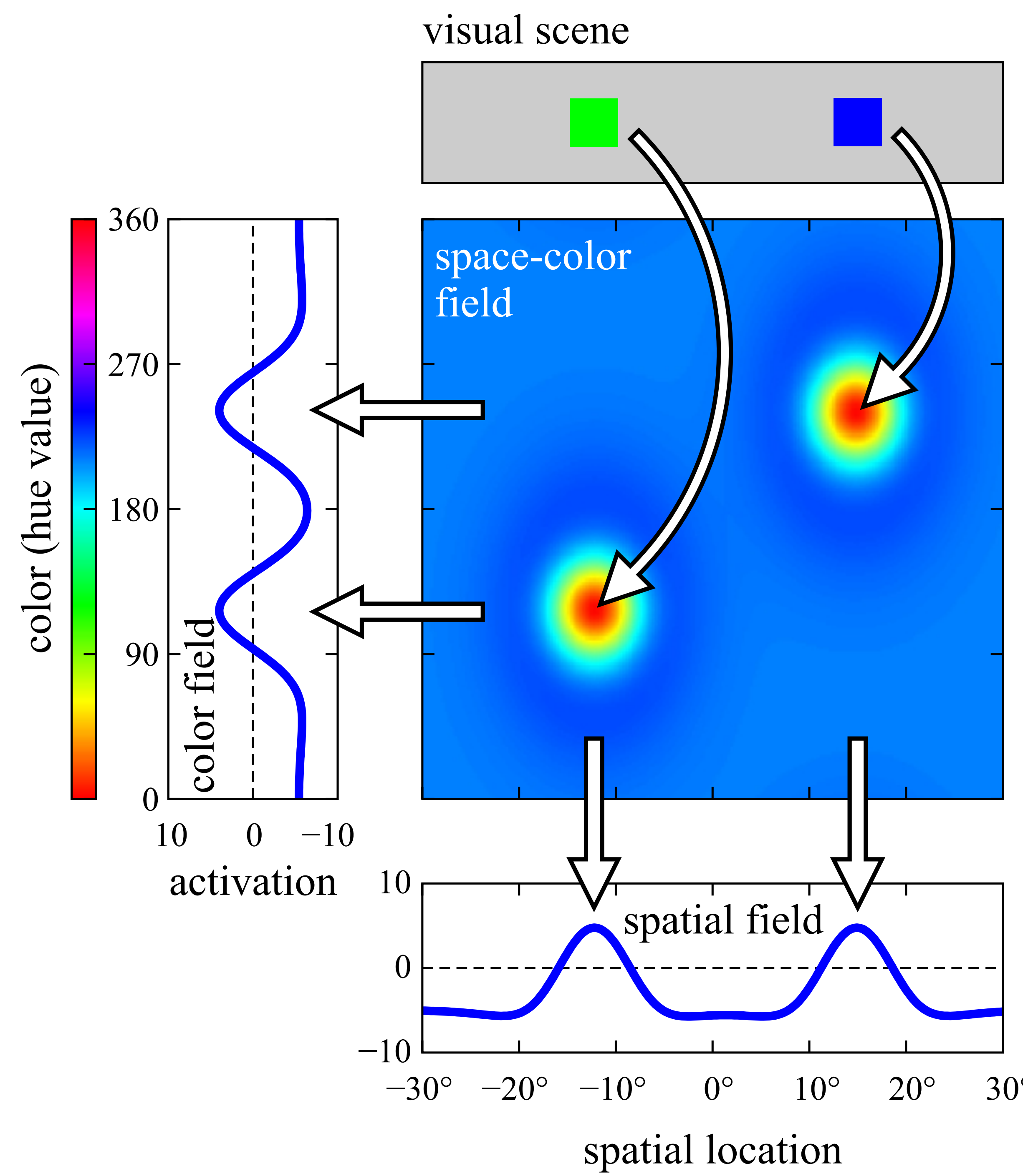
- in a joint representation, localized peaks represent instances in which the different features dimensions are "anatomical bound"
- fixed: need the neural substrate every possible bound state



[Schneegans et al.,Ch 5 of *DFT Primer,* 2016]

Extract features: unbinding

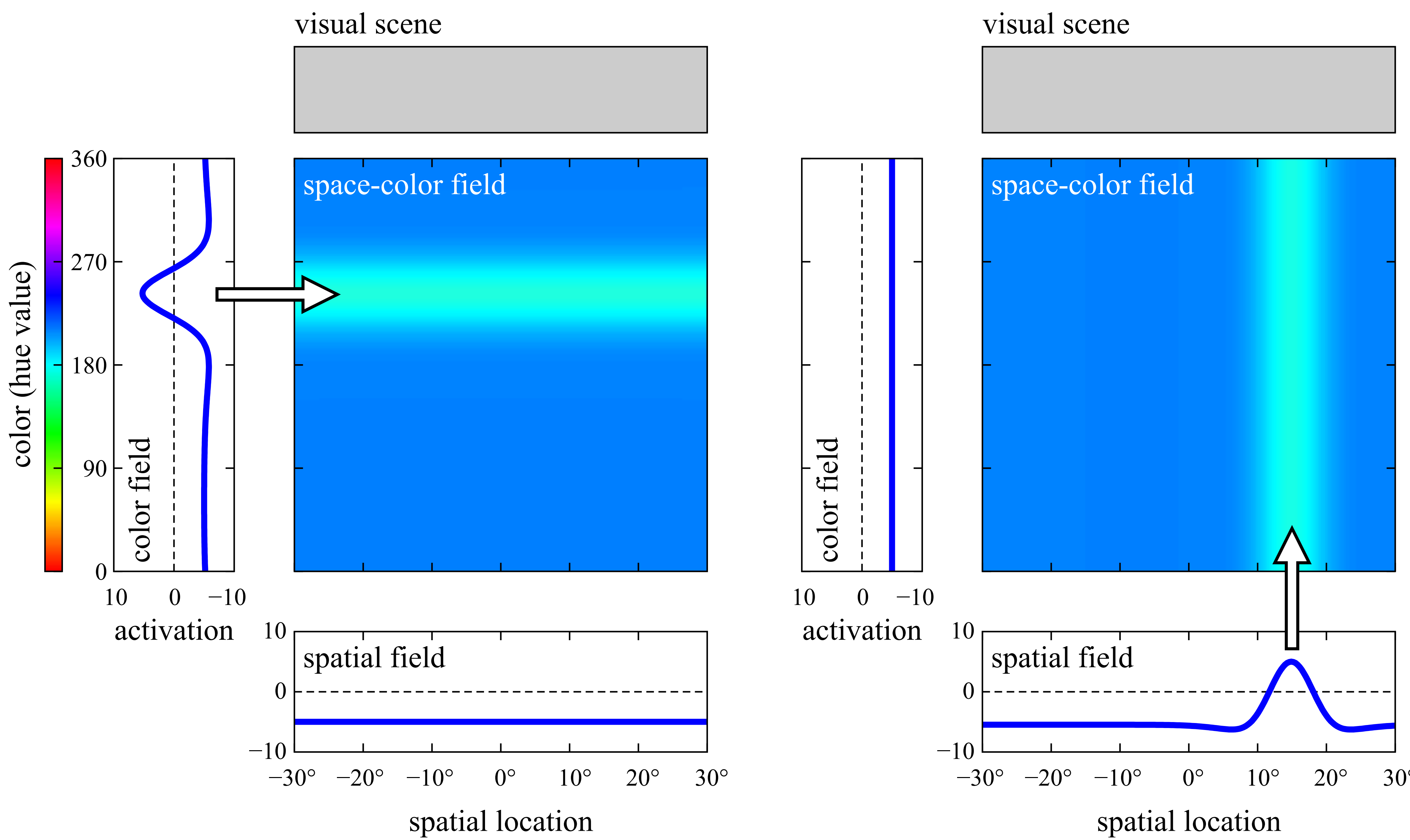
- projecting to lower-dimensional fields by summing along the marginalized dimensions
- contraction mapping



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Bind features

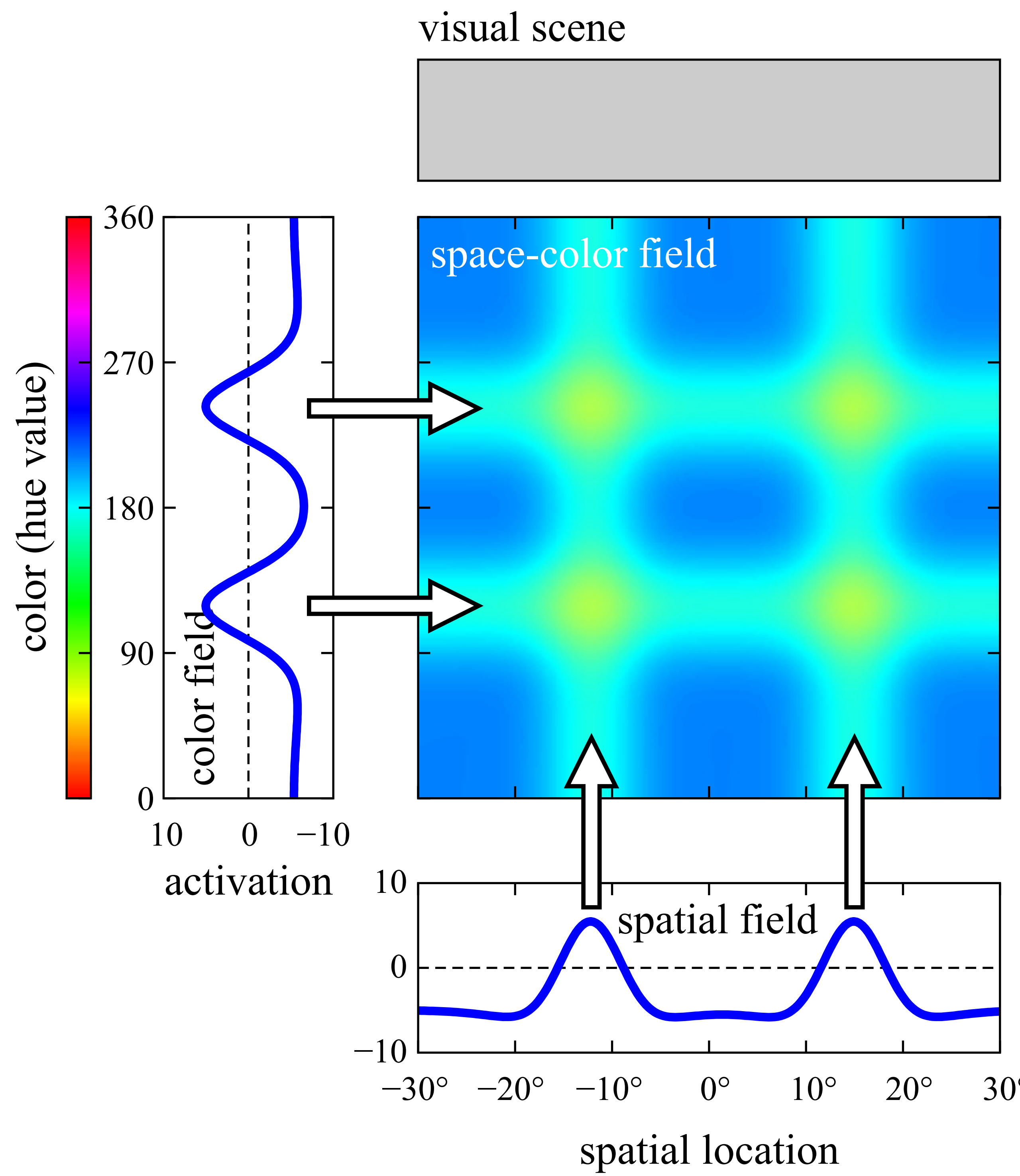
- project lower-dimension field onto higher-dimensional field: expansion mapping



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Binding problem

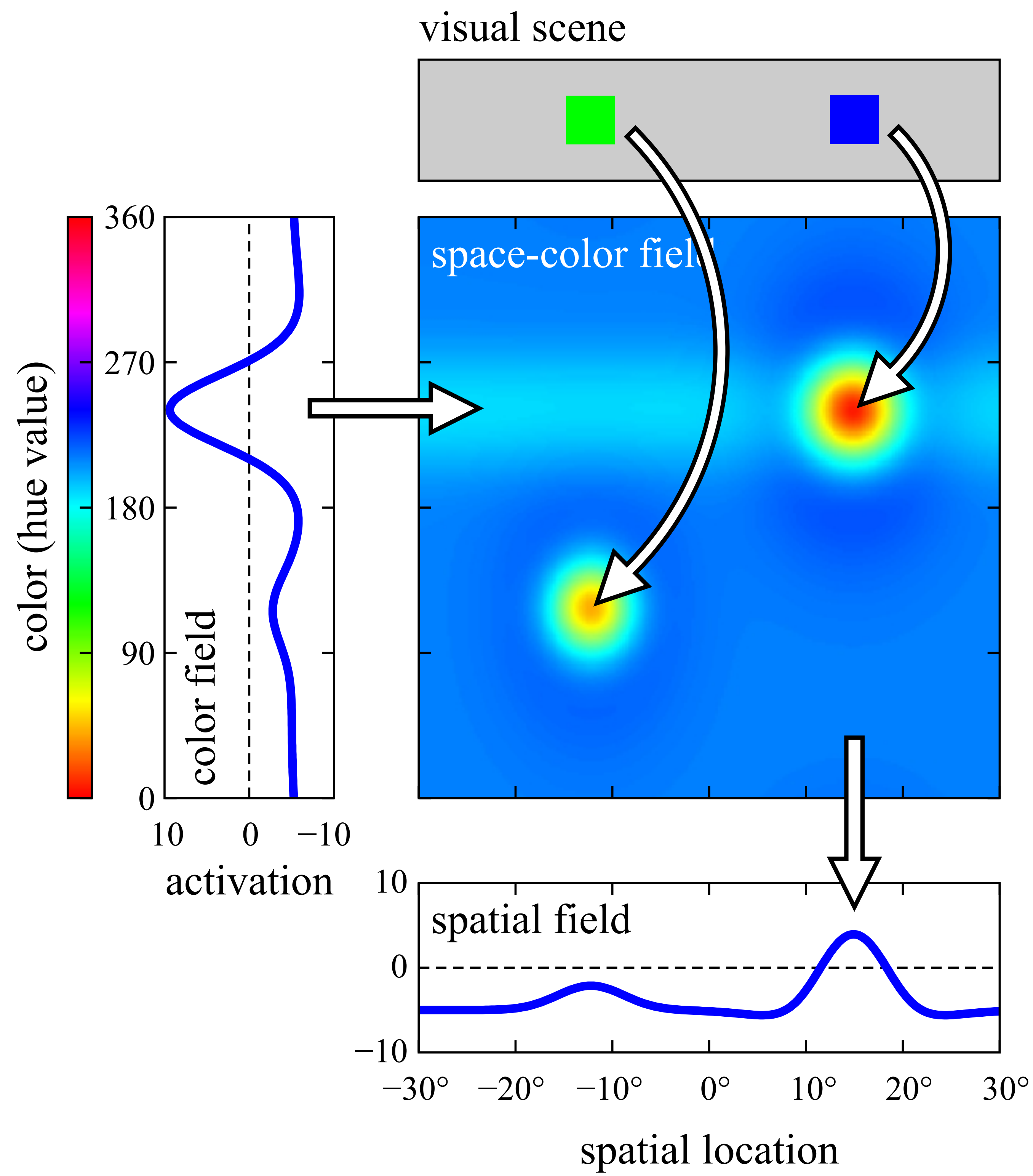
- this binding operation runs into the binding problem
- solution: bind one object at a time
- => attentional bottleneck



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Cued selection

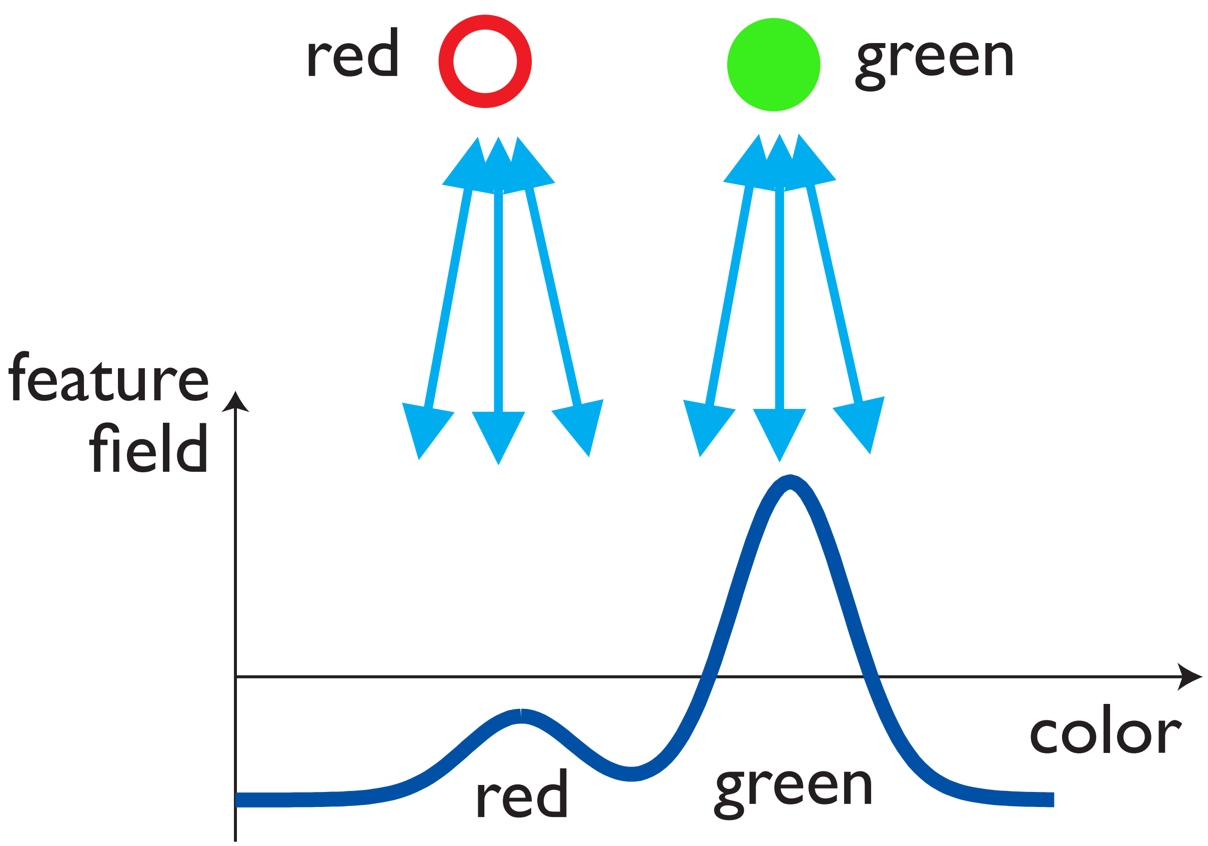
- an operation that uses joint and individual representations
- combining expansion and contraction



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Role-filler binding

- color concepts… grounded in feature fields

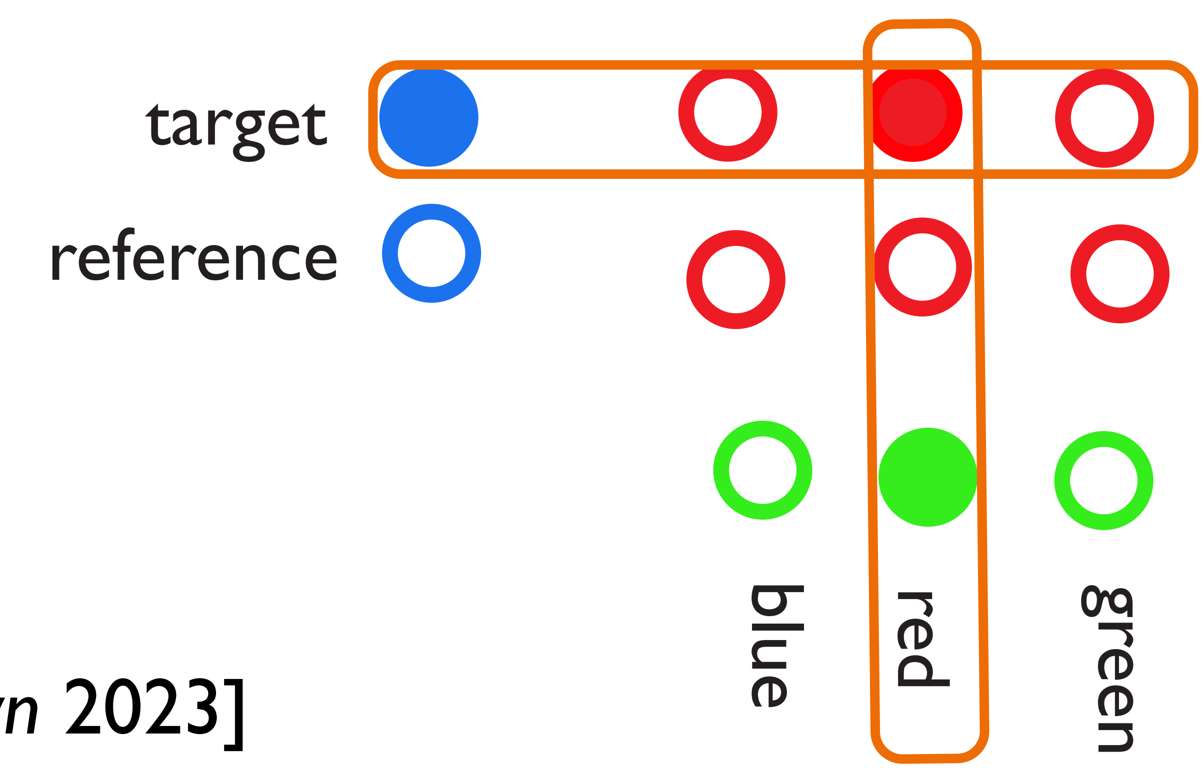


- roles: reference, target, agent, tool, …

"green to the right of red"

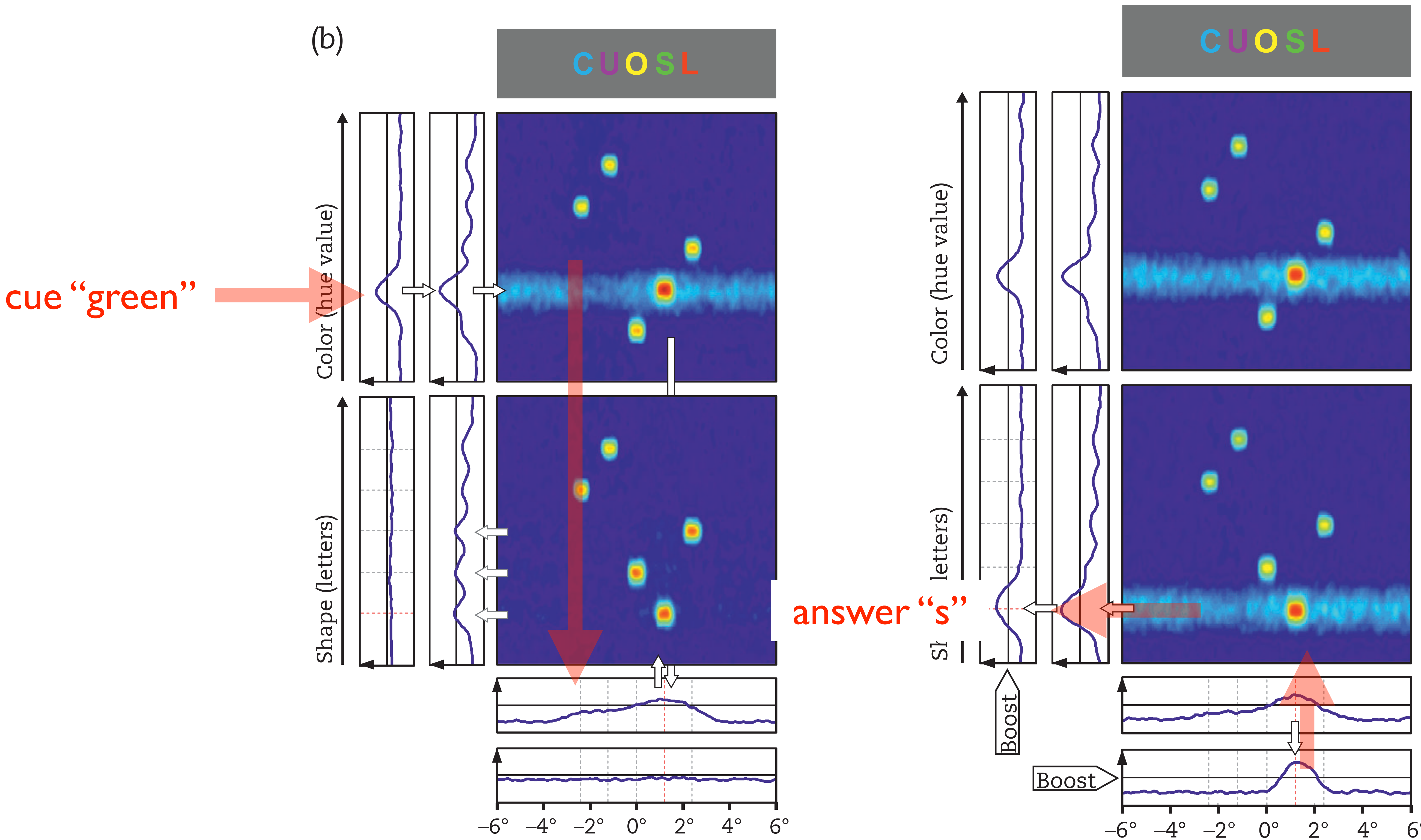
target reference

- joint representation of roles and concepts



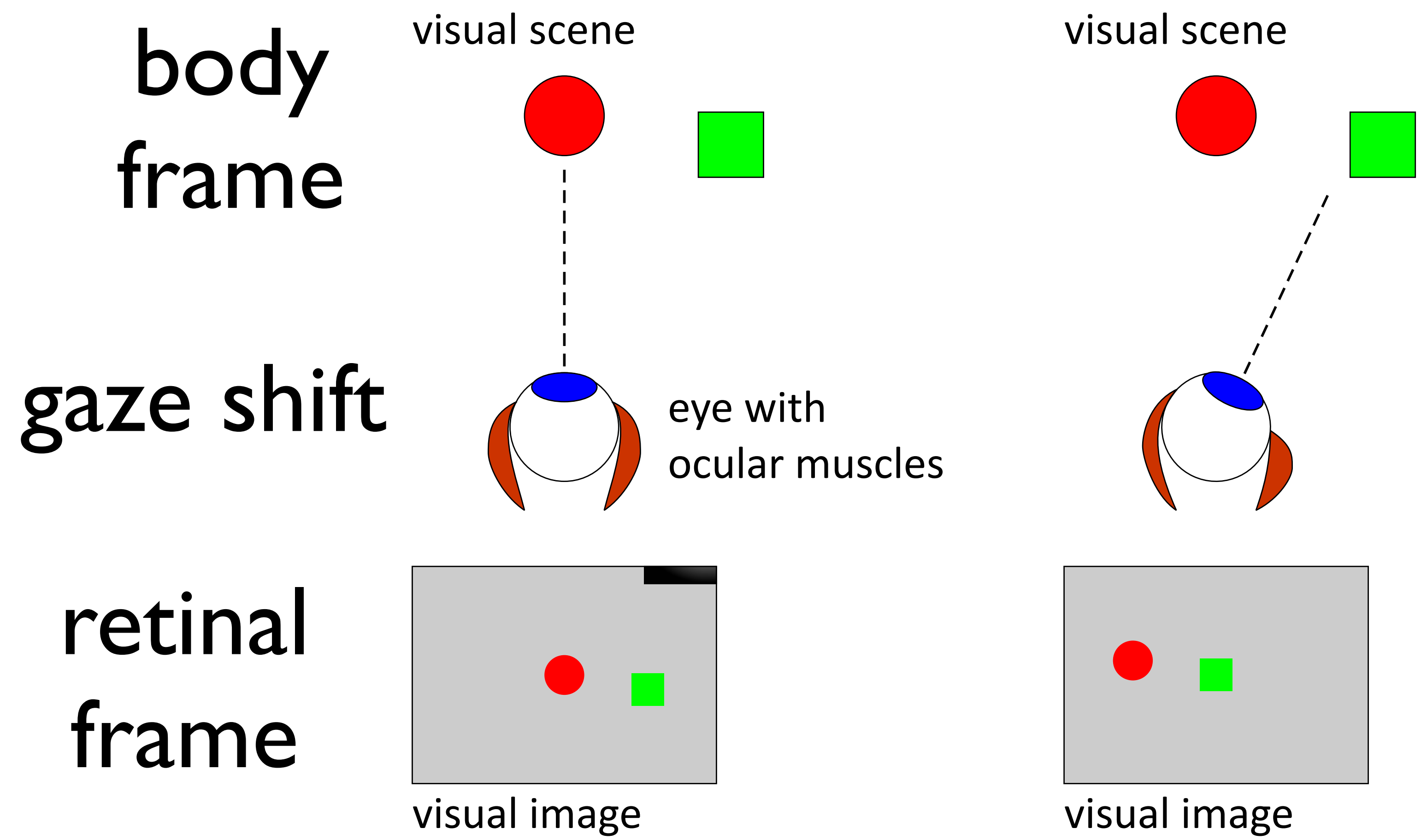
[Sabinasz, Richter, Schöner: *Cog Neurodyn* 2023]

Binding through space

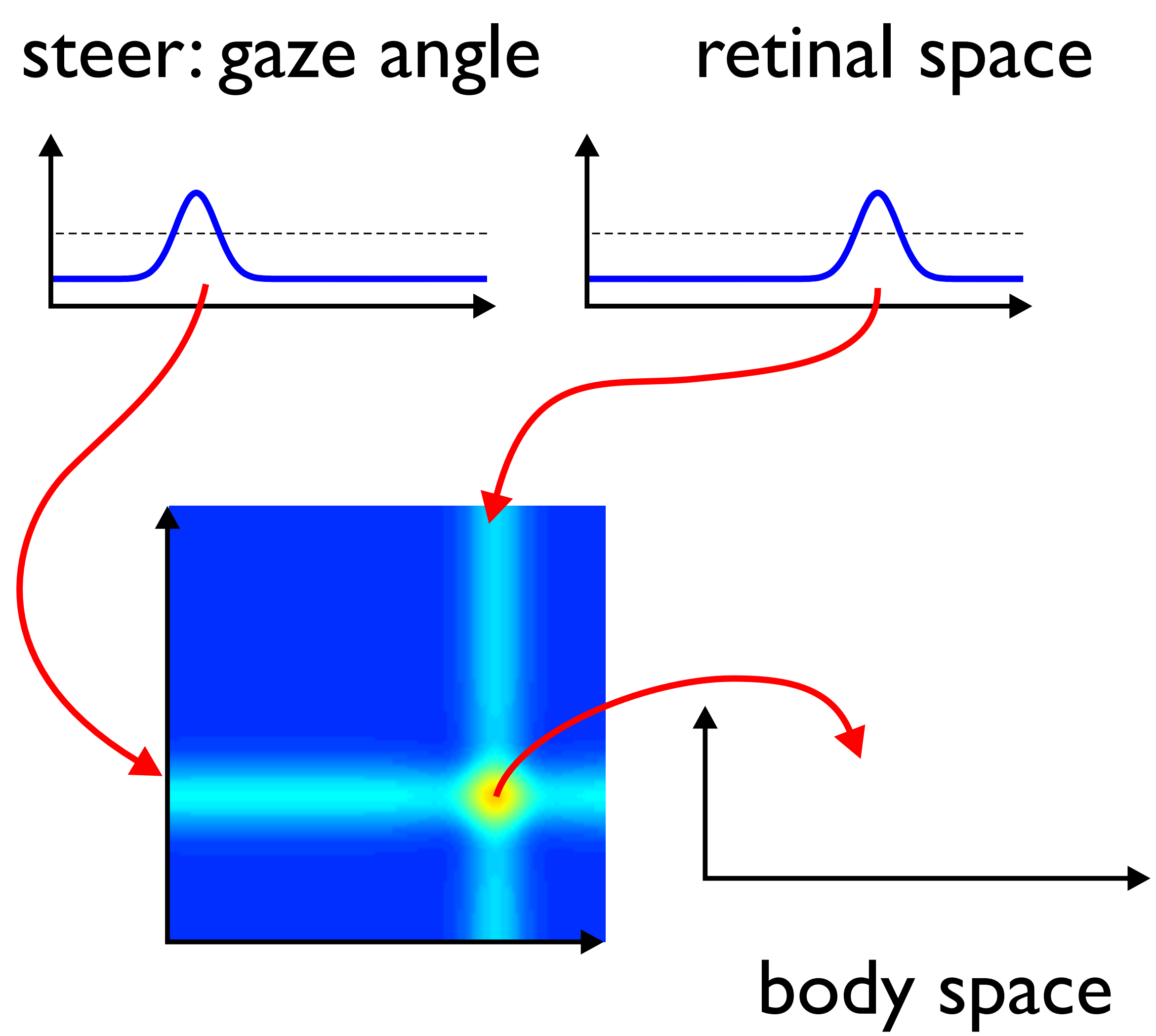


[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Coordinate transforms

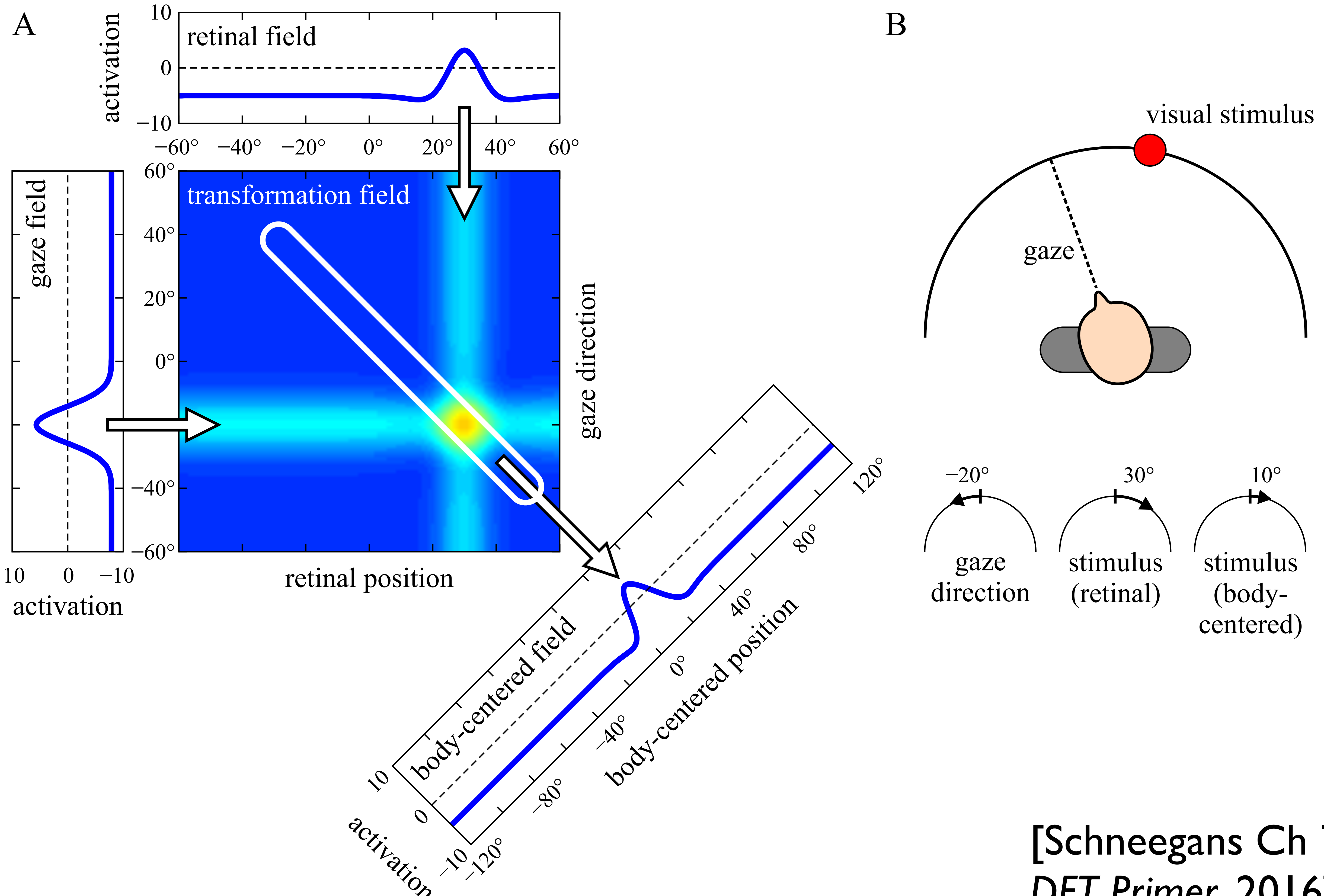


- bind neural representations of
  - retinal space
  - gaze angle
- into a joint representation
  - (gain field ~Andersen/Pouget
- then contract to body space



[Schneegans Ch 7, *DFT Primer*, 2016]

Coordinate transform



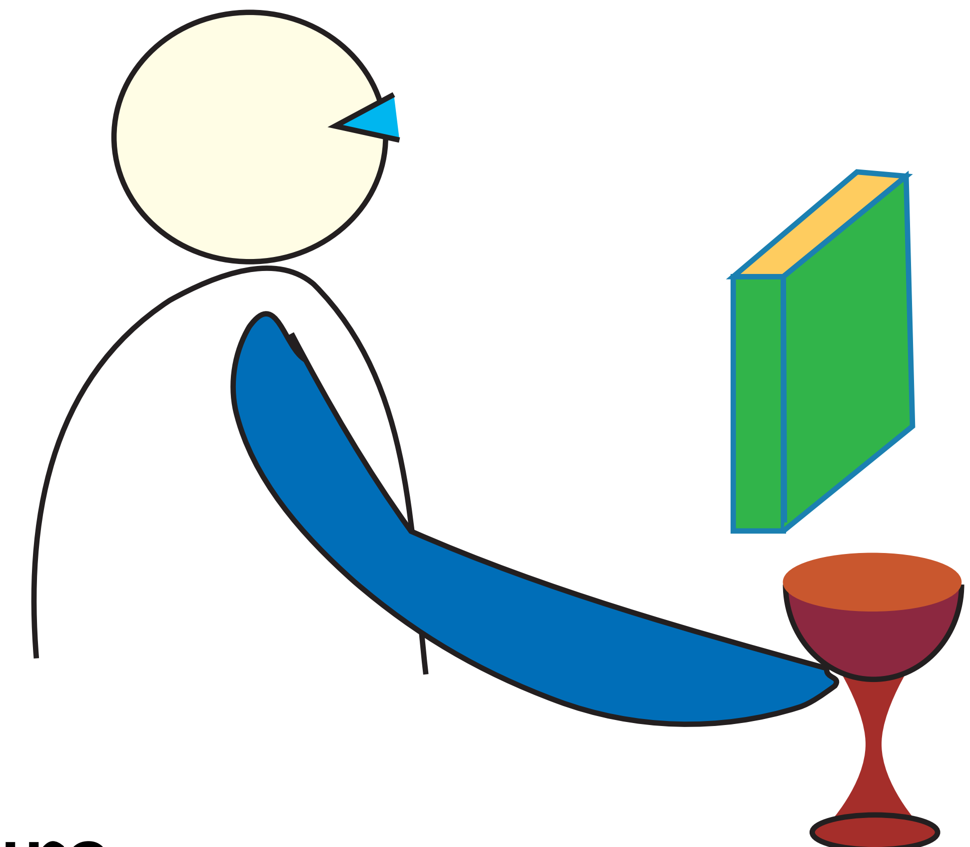
[Schneegans Ch 7 of *DFT Primer*, 2016]

Grounding language

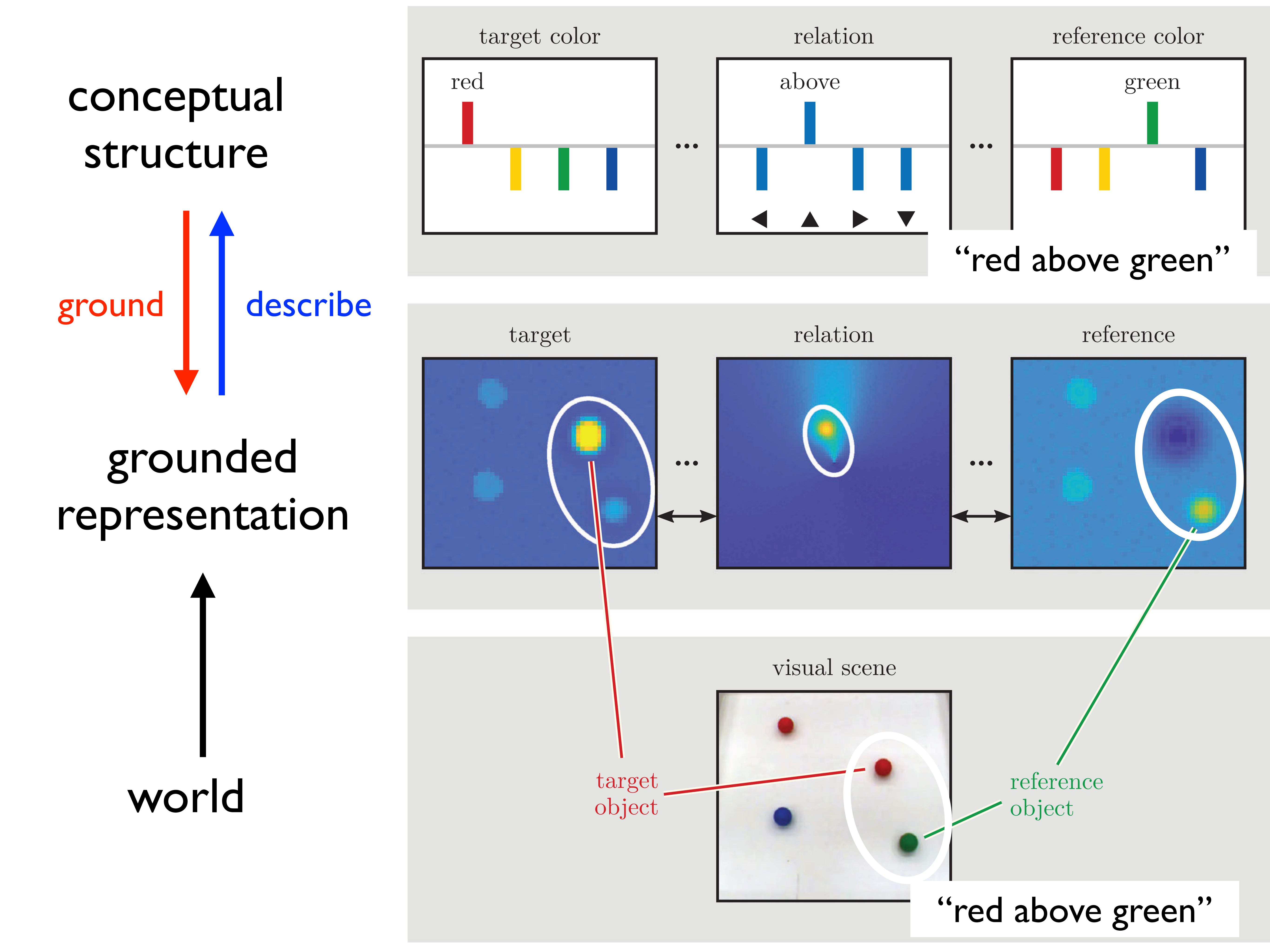
- Grounding language
- Representing conceptual structure
- Grounding relational concepts
- Mental mapping
- Grounding nested phrases

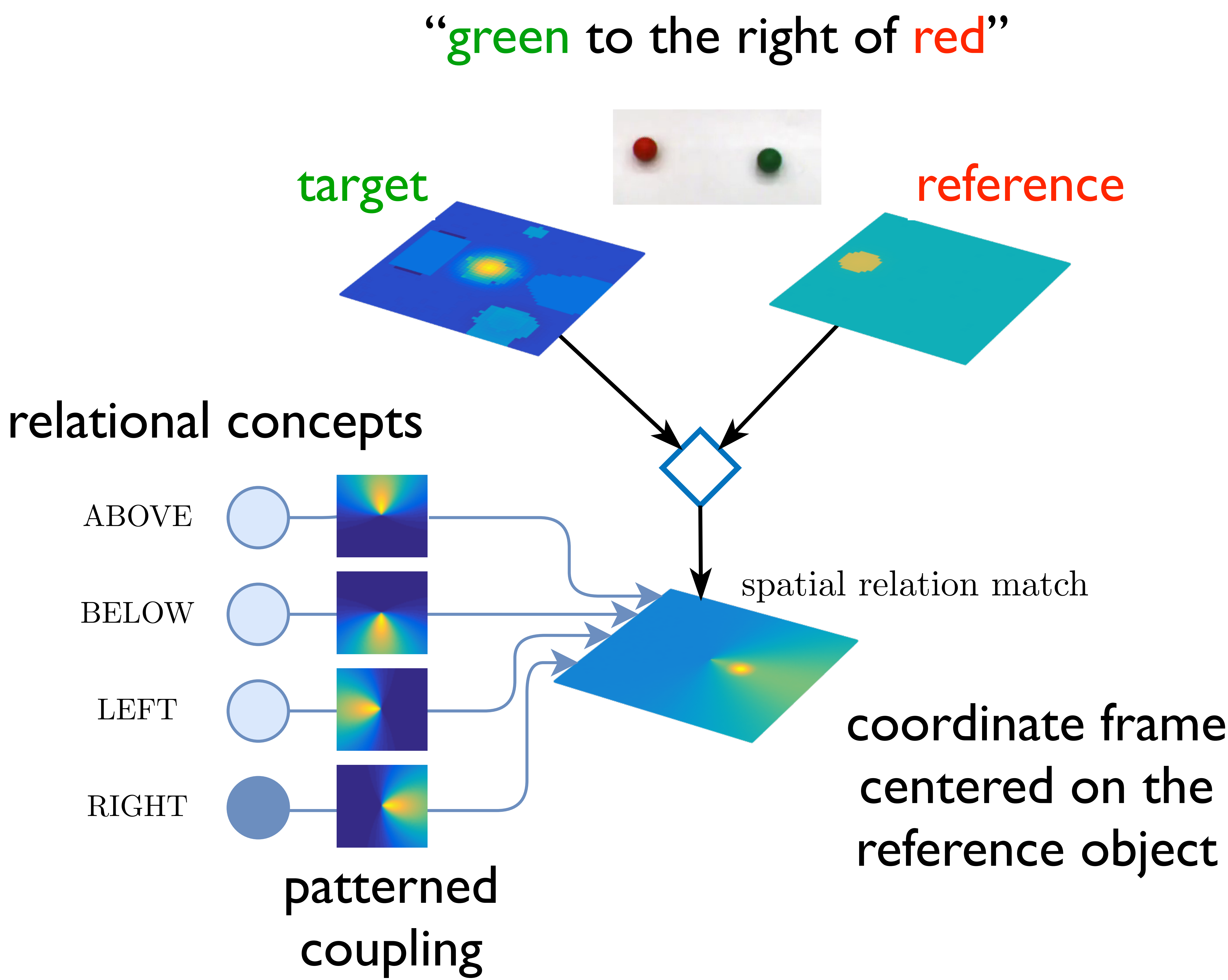
Perceptual grounding

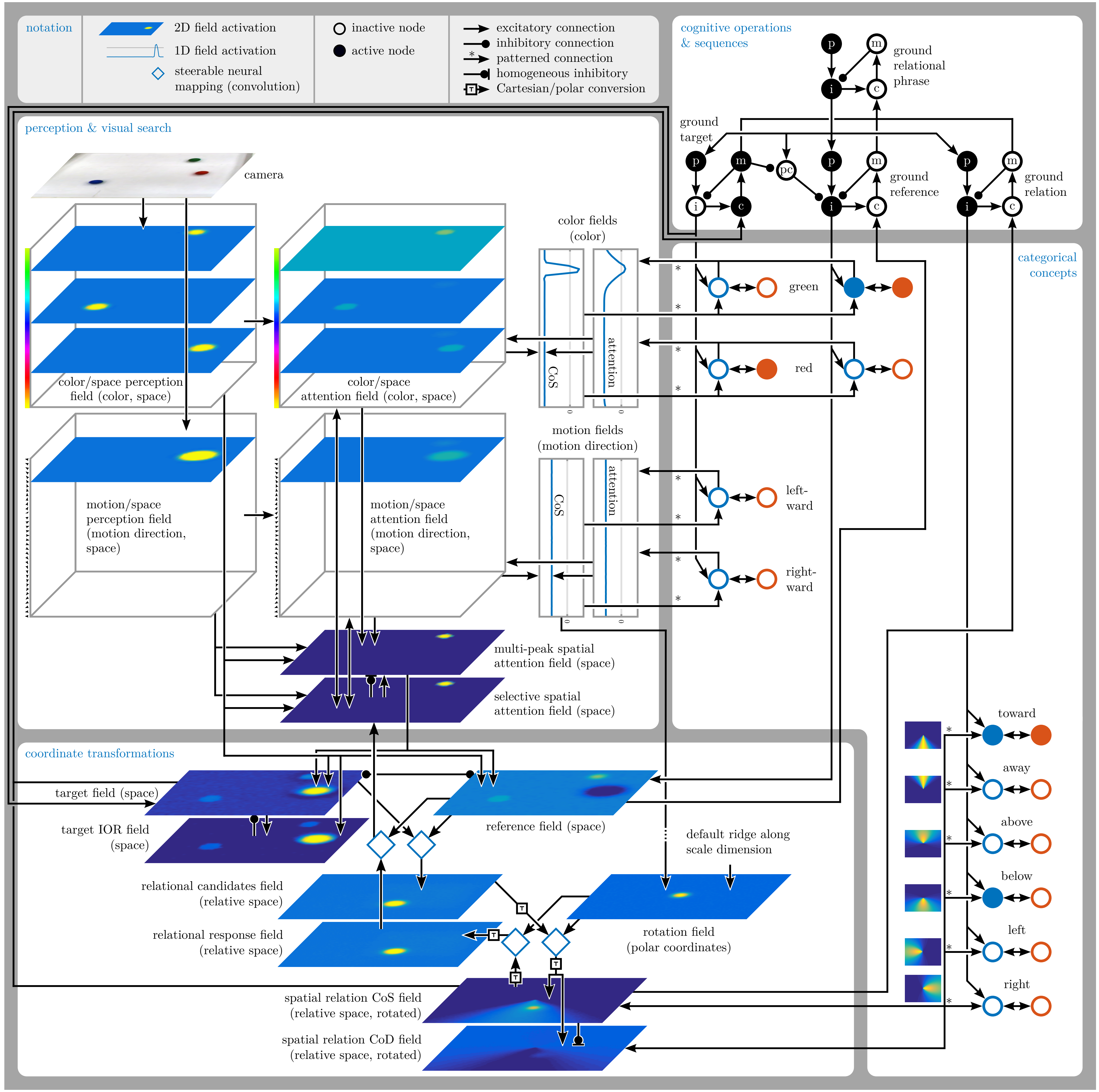
- “the cup to the right of the green book”
- “he reaches for the cup”



- grammatical/semantic structure
  - nouns vs relations/actions
  - roles: target object is related to a reference object

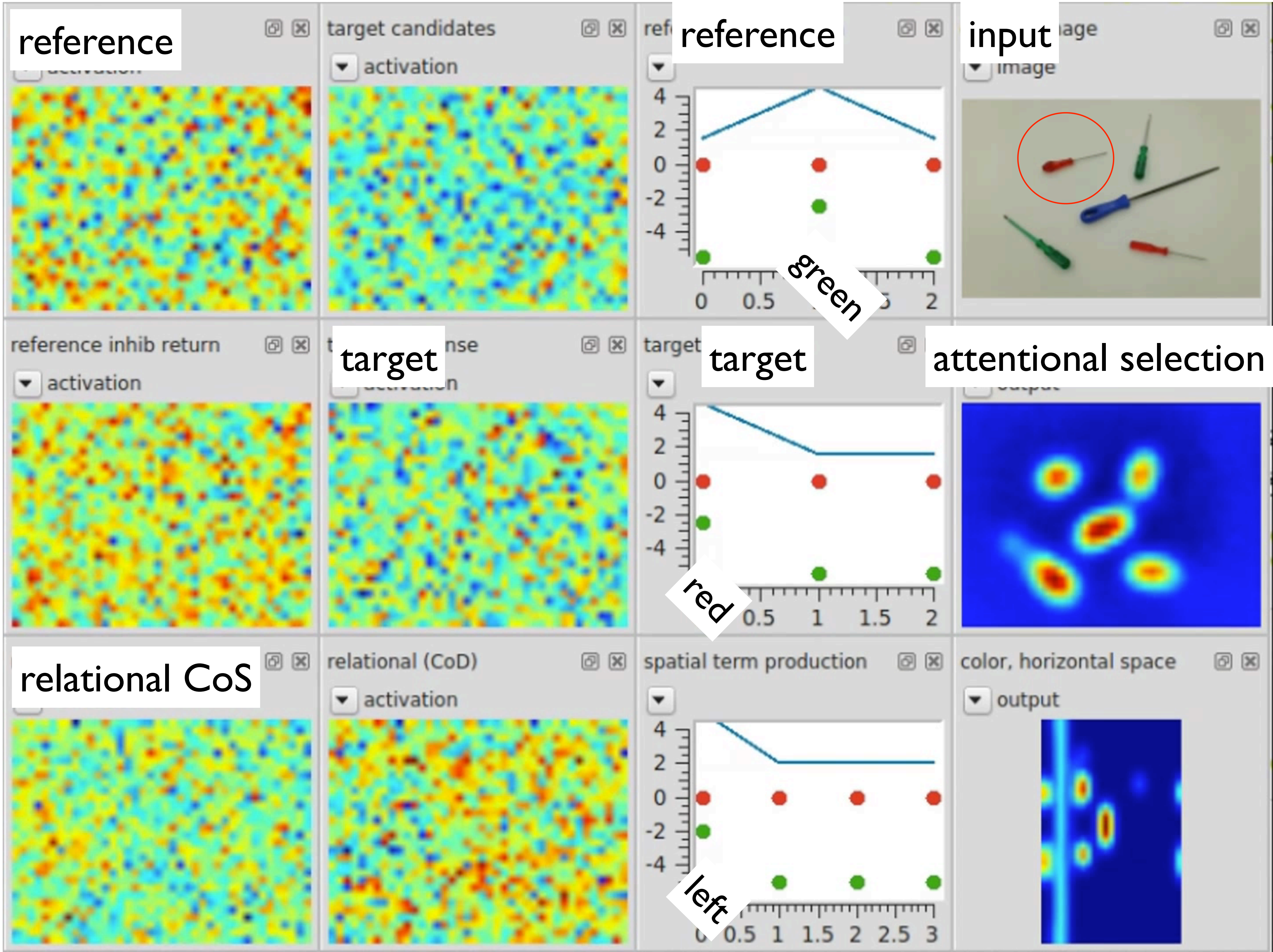

conceptual structure
ground
describe
grounded representation
world
target color
red
relation
above
reference color
green
…
"red above green"
target
relation
reference
visual scene
target object
reference object
"red above green"


"green to the right of red"
target
reference
relational concepts
ABOVE
BELOW
LEFT
RIGHT
spatial relation match
coordinate frame centered on the reference object
patterned coupling



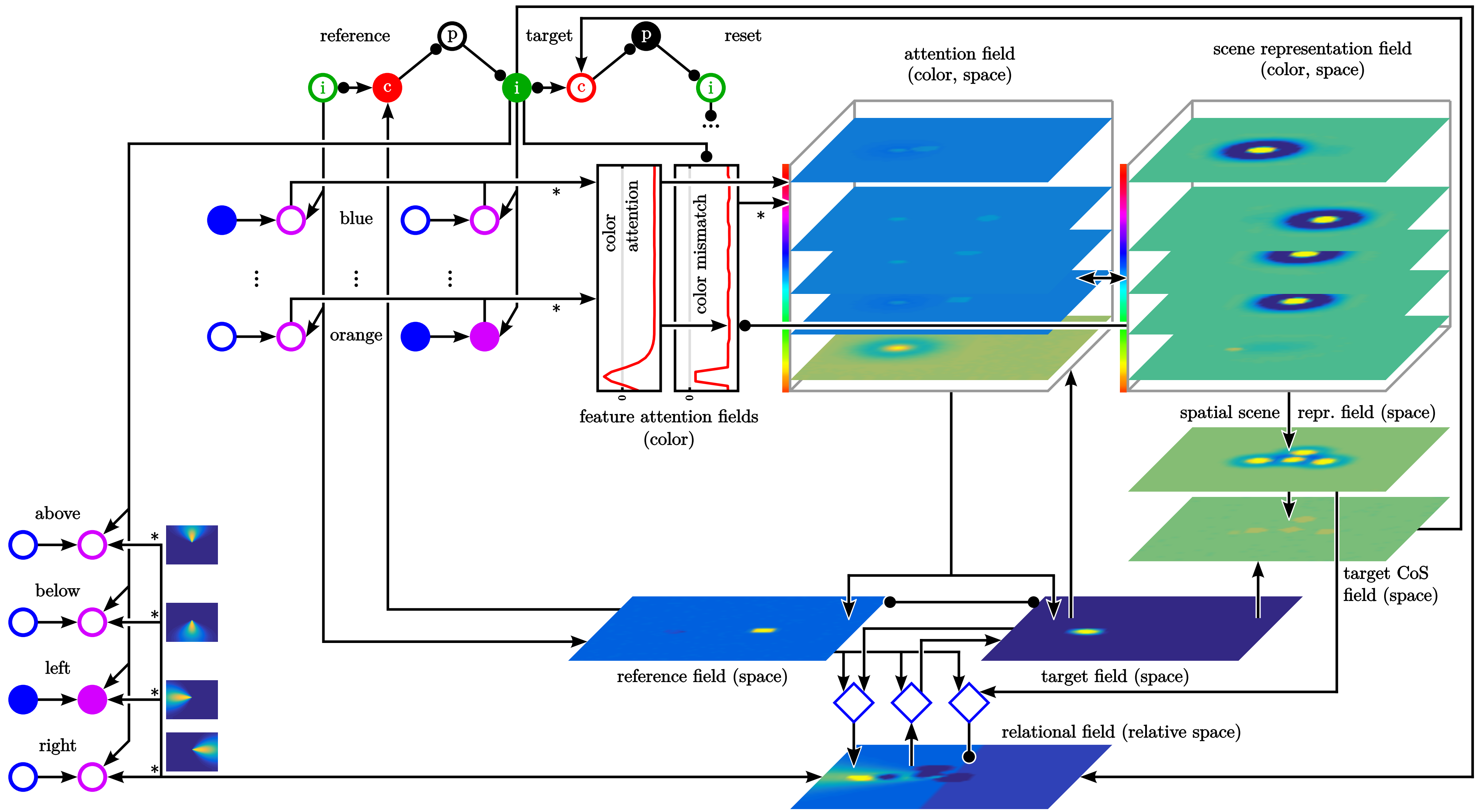
[Richer, Lins, Schöner, *Cognitive Science* 2021]

“red to the left of green?” conceptual structure



Mental mapping

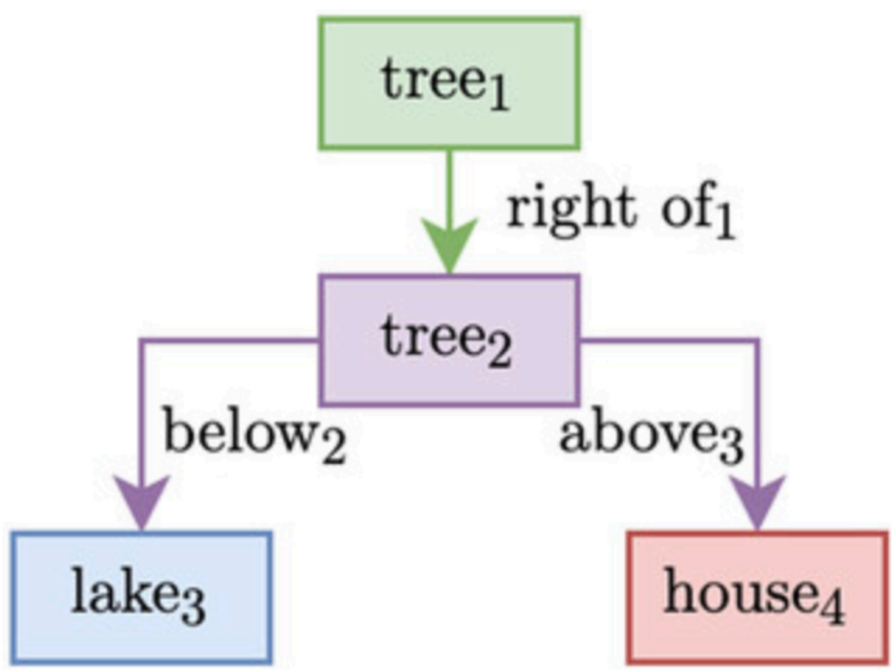
- propositions
  - “There is a cyan object above a green object.”
  - “There is a red object to the left of the green object.” …
- inference
  - “Where is the blue object relative to the red object?”

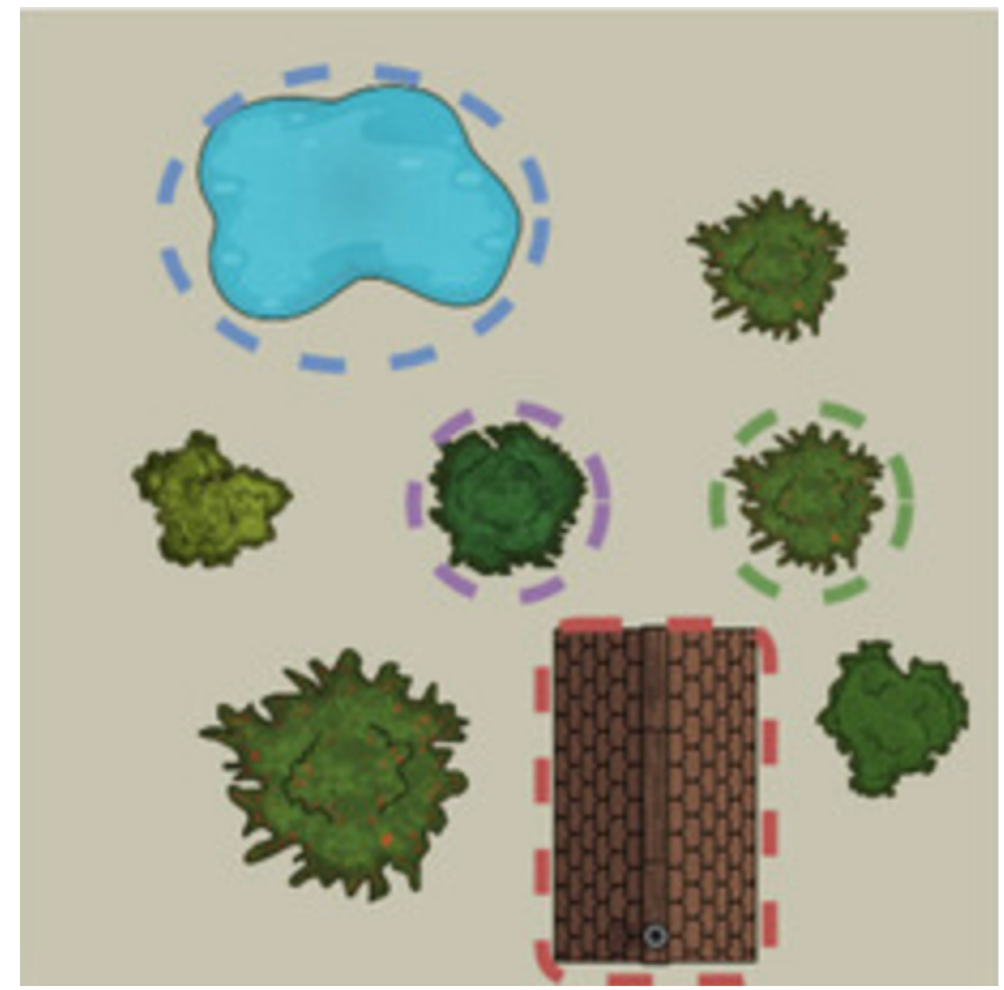


[Kounatidou, Richter, Schöner, *CogSci* 2018]

Grounding nested phrases

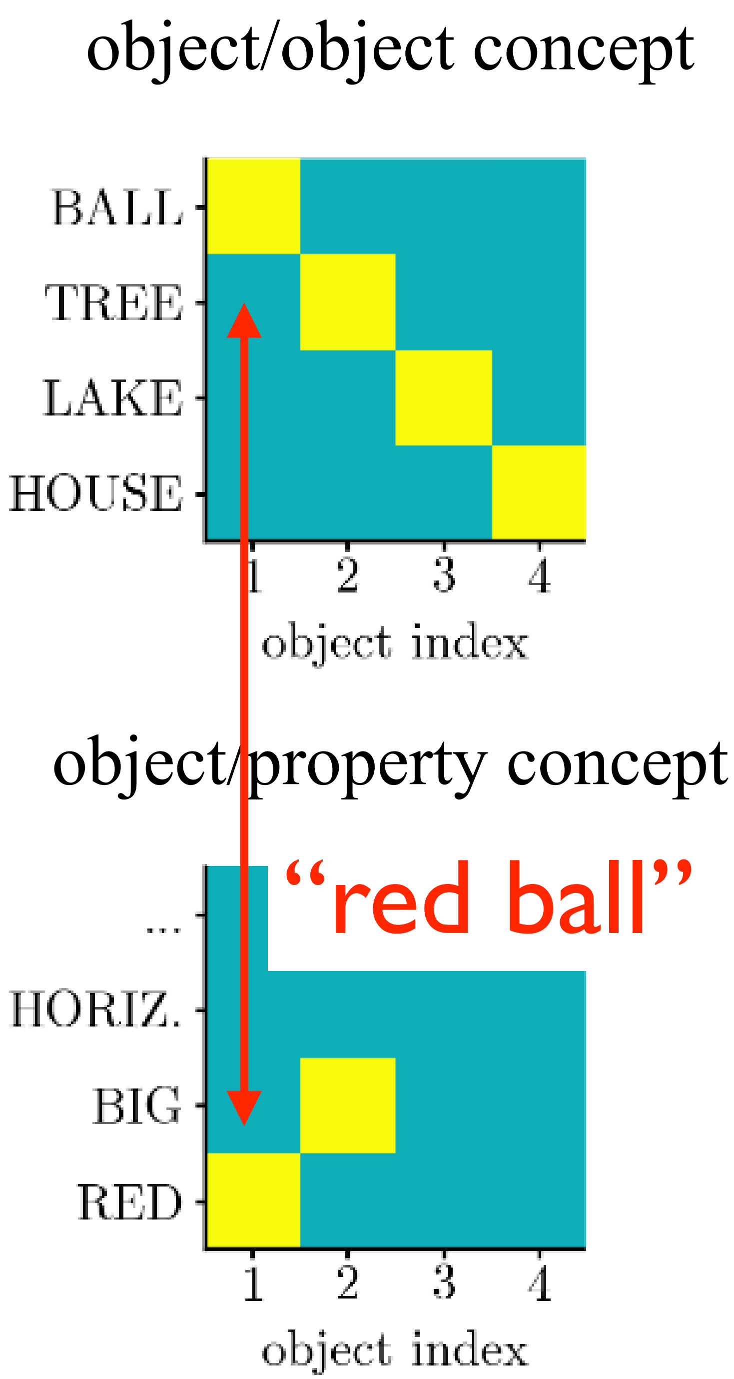
"the tree to the right of the tree that is below the lake and above the house"

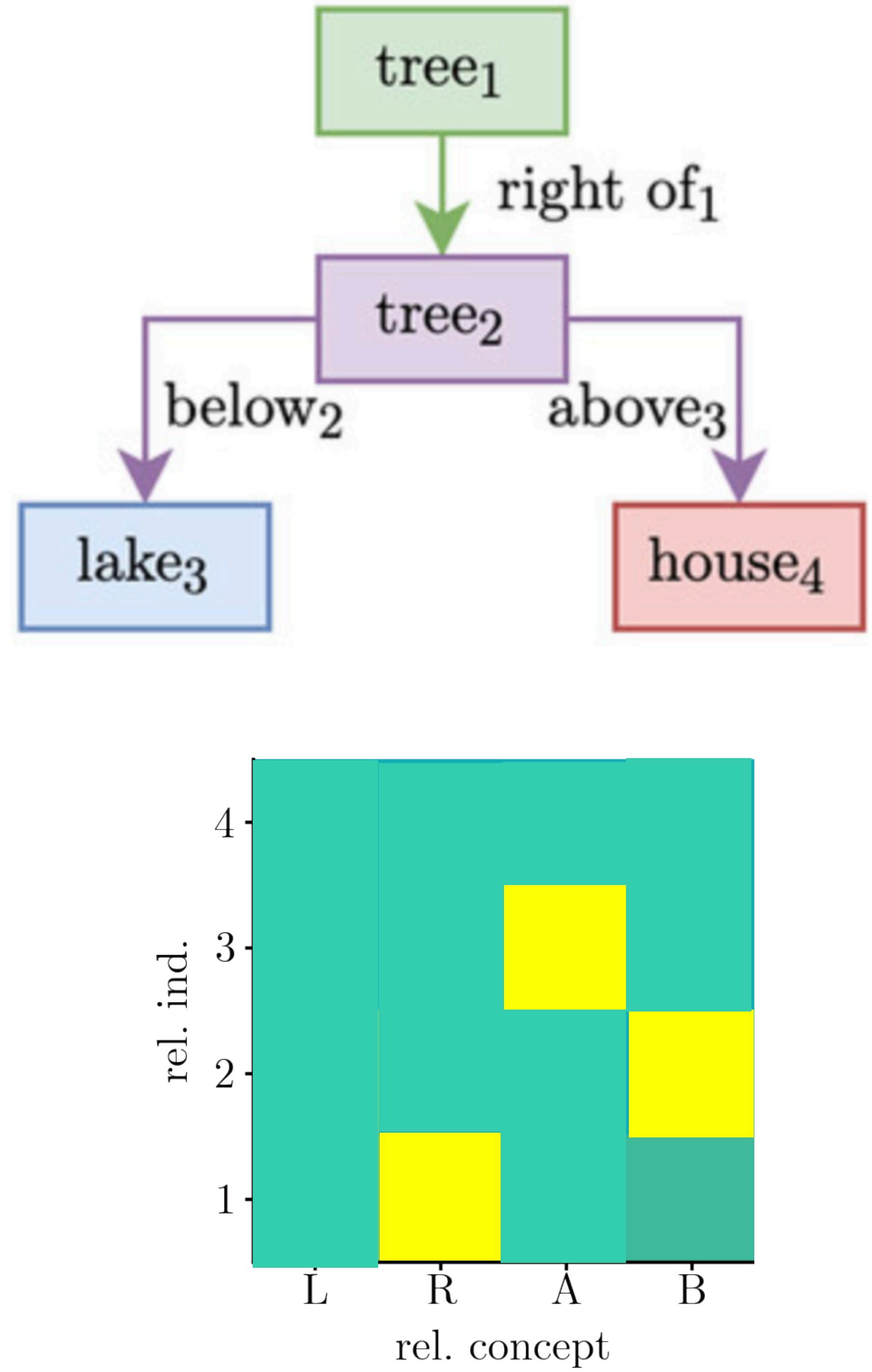




[Sabinasz, Schöner, TopiCS 2023; Sabinasz, Richter, Schöner Cog Neurodyn 2023]

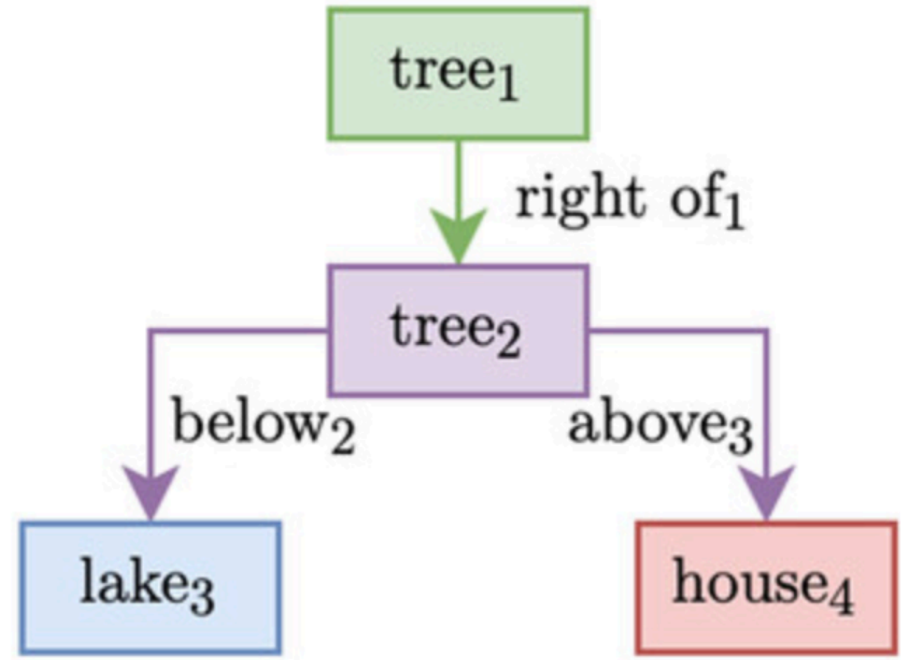
Binding through ordinal position



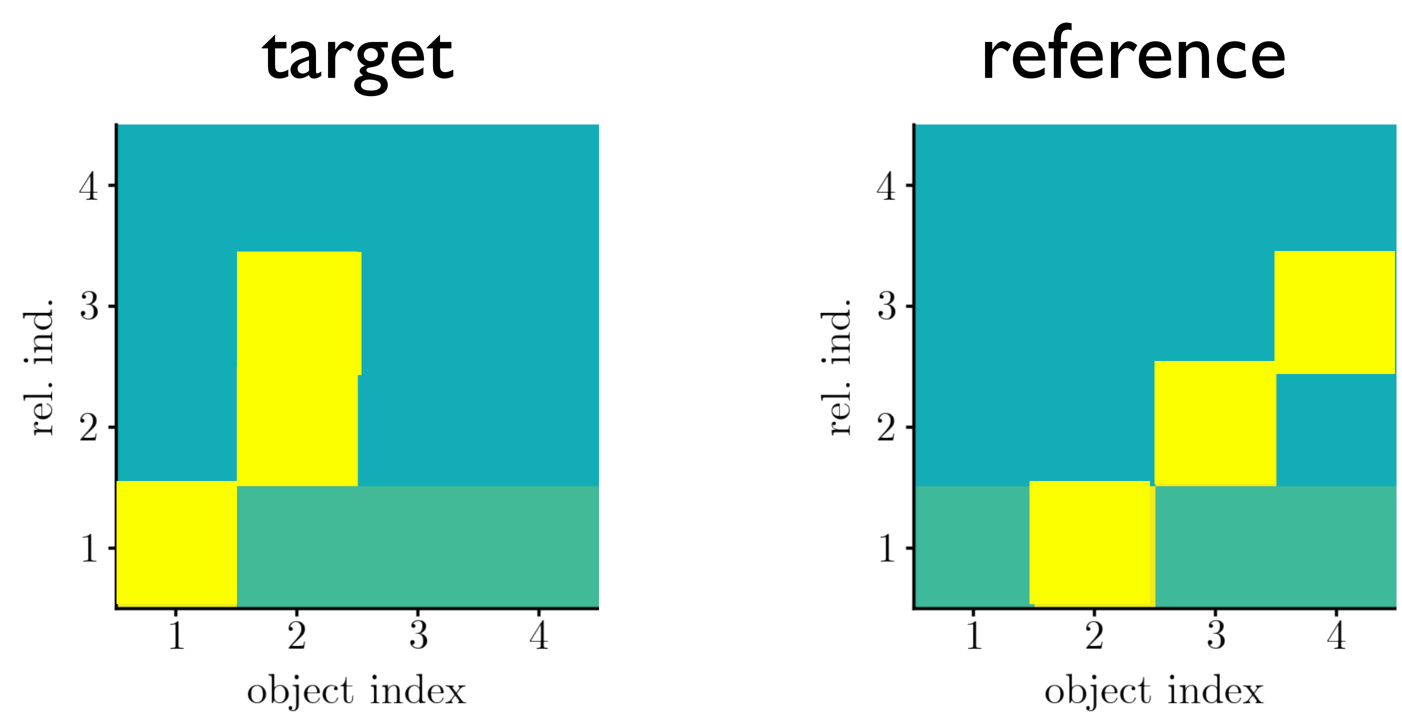


[Sabinasz, Richter, Schöner, *Cog. Neurodyn*. 2023]

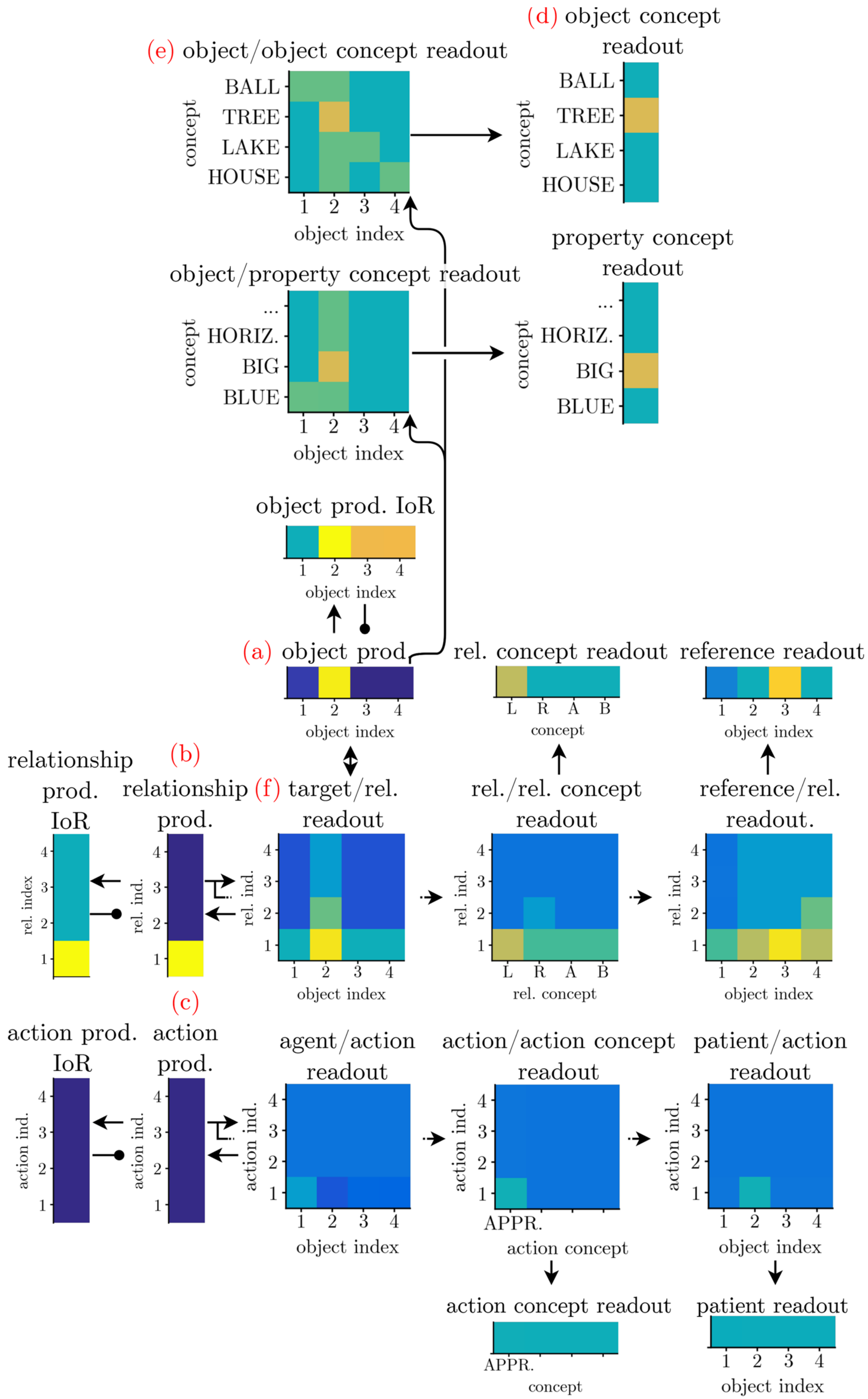
Neural representation of conceptual structure

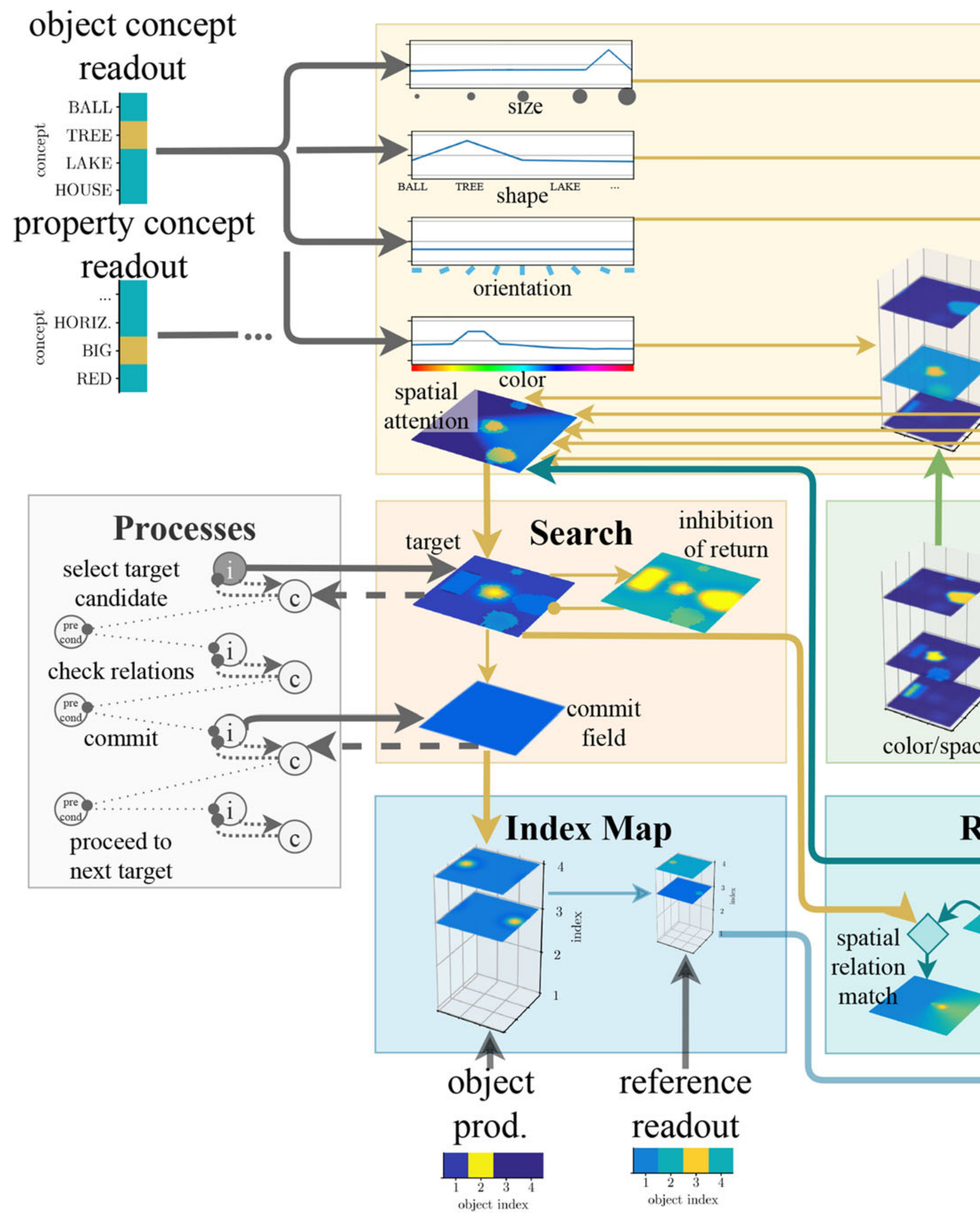


- binding arguments in particular roles to relations through the index dimensions

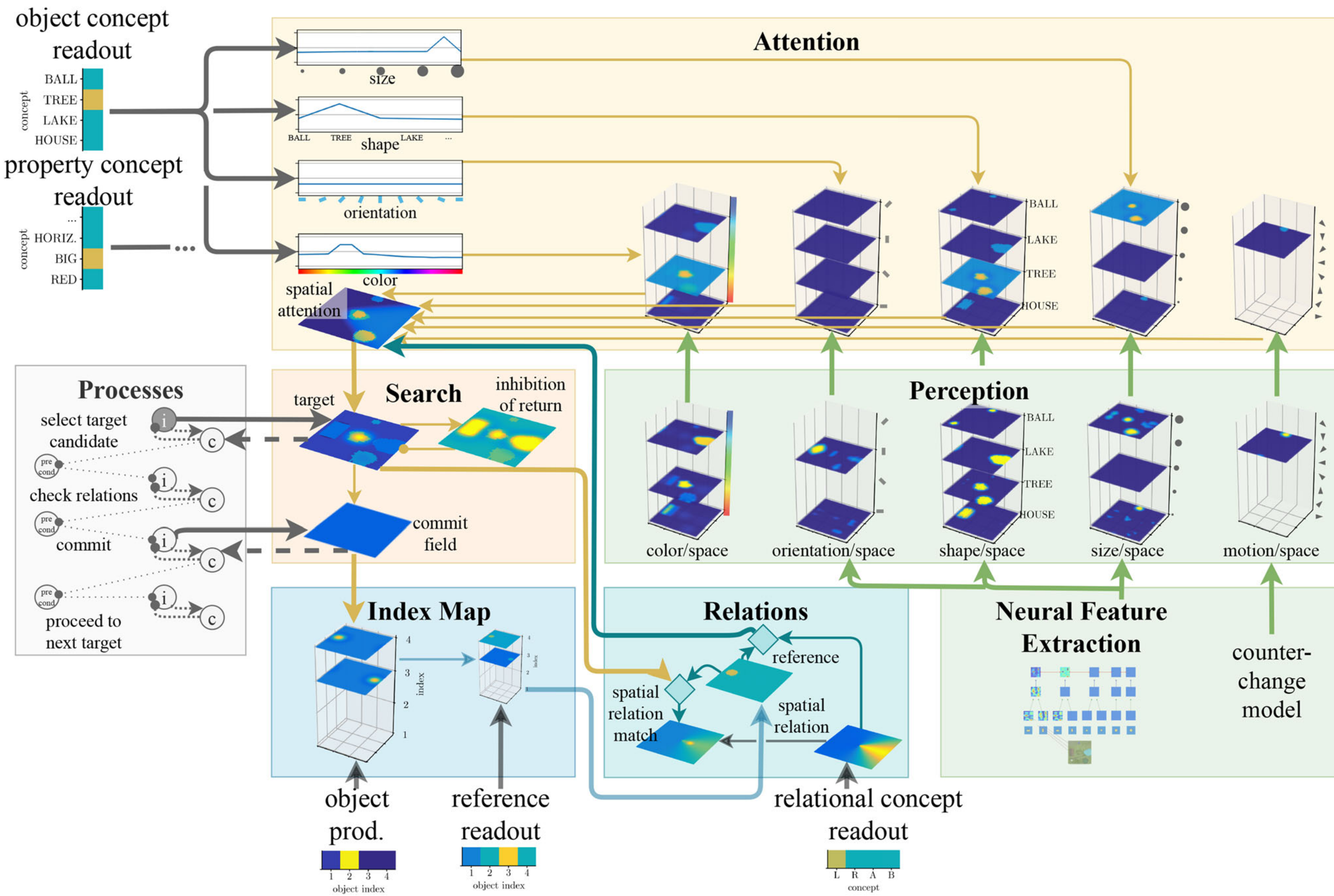


[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023]

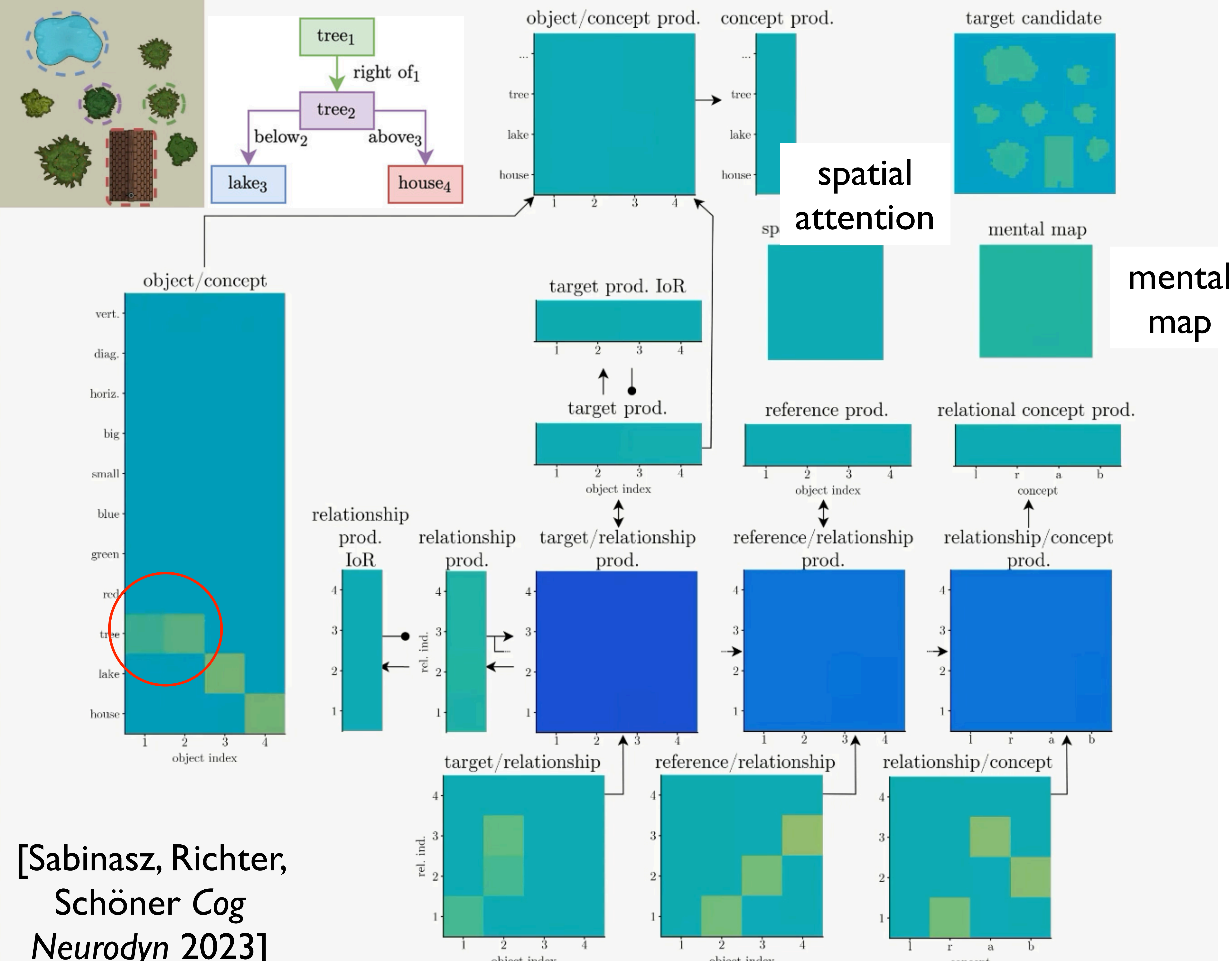




[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023]



[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023]



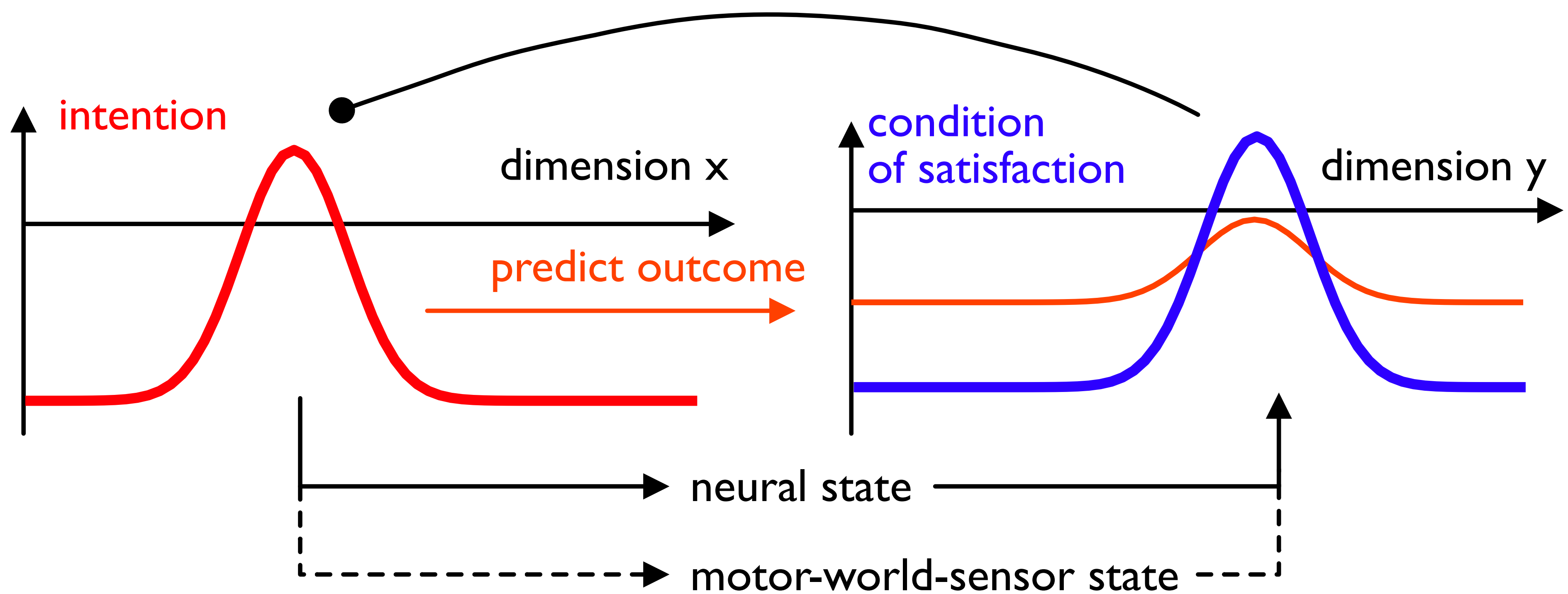
[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023]

Sequence generation

- the Condition of Satisfaction
- global view of sequence generation
- what state next?
- what if the CoS fails?

Neural dynamic principle

- the current neural attractor state = intention
- predicts its condition of satisfaction (CoS)
- input matching prediction: CoS activated
- CoS inhibits intention…

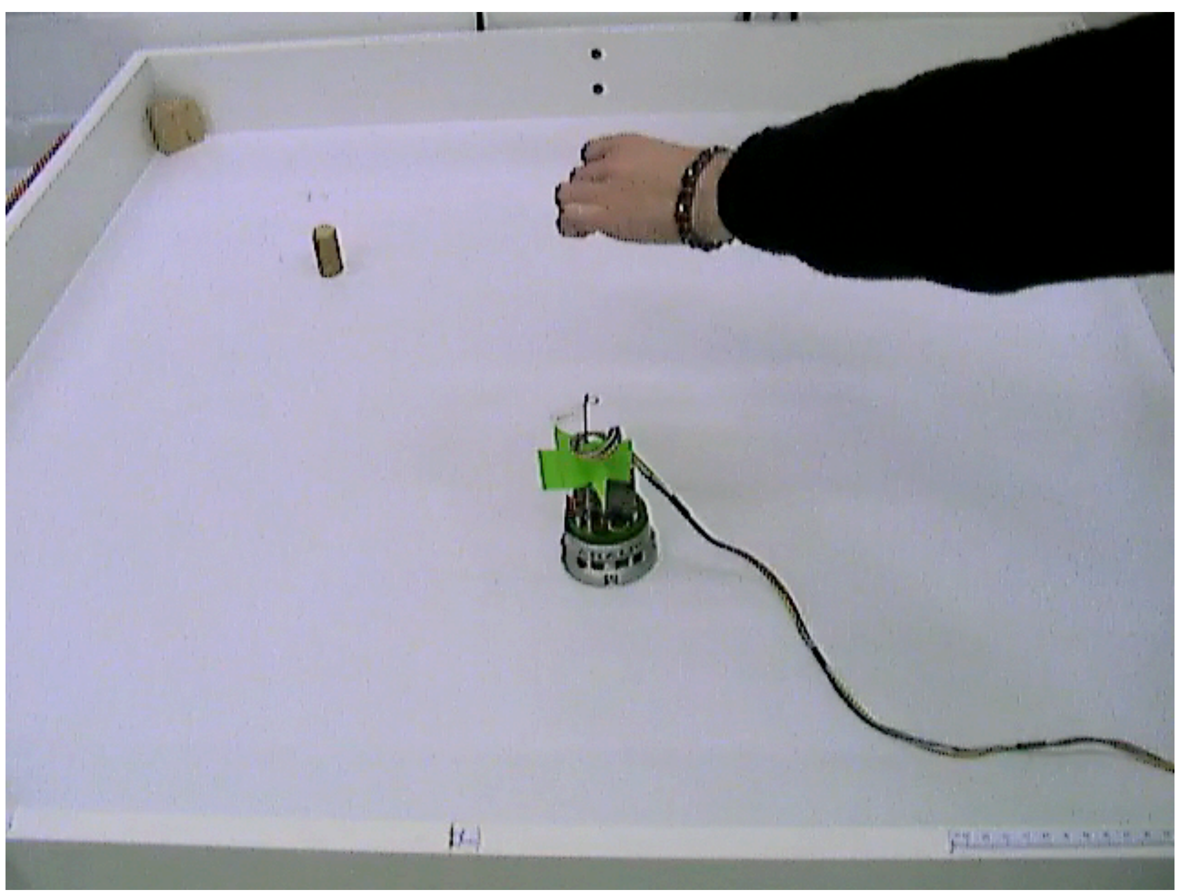


[Sandamirskaya, Schöner: *Neural Networks* 2010; Sandamirskaya *DFT primer* 2016]

Serial order task

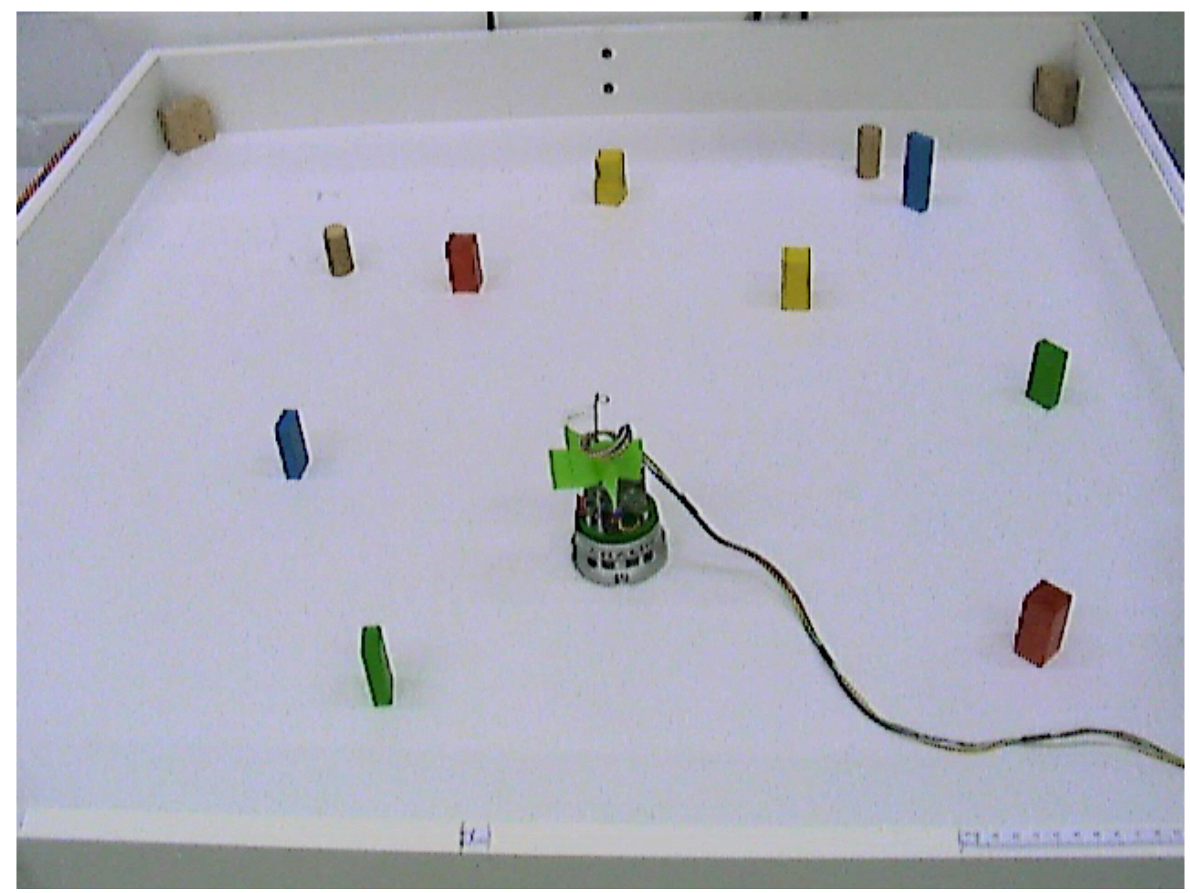
- learn a serially ordered sequence from a single demonstration

yellow-red-green-blue-red



- perform the serially ordered sequence with new timing

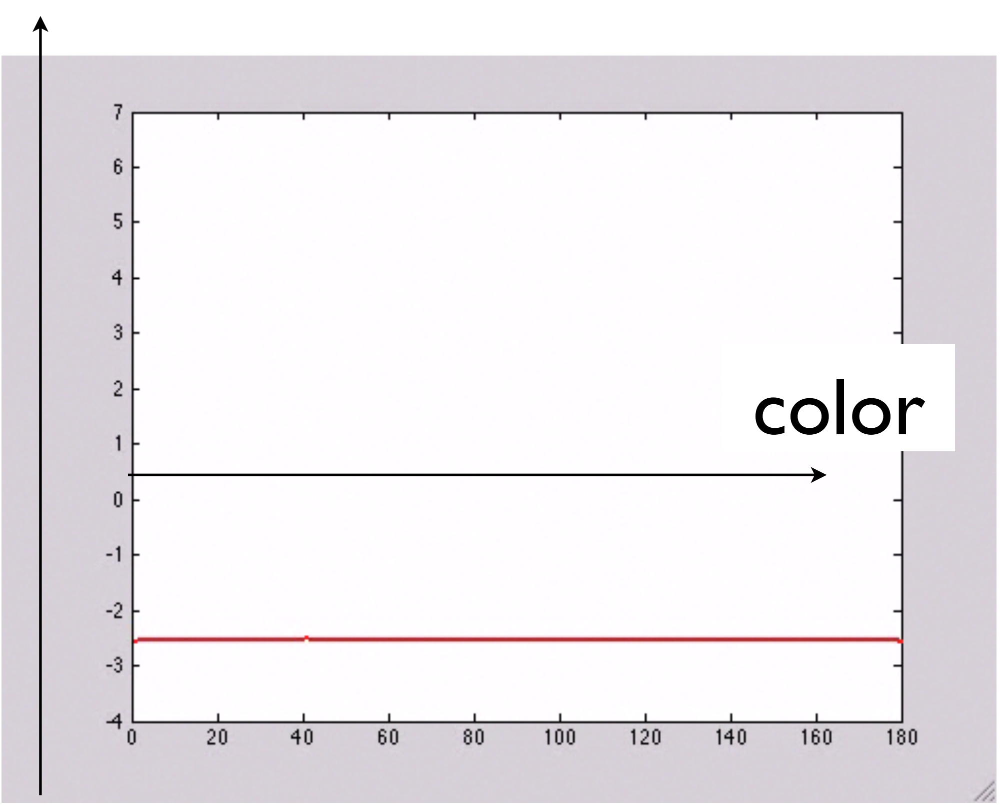
yellow-red-green-blue-red



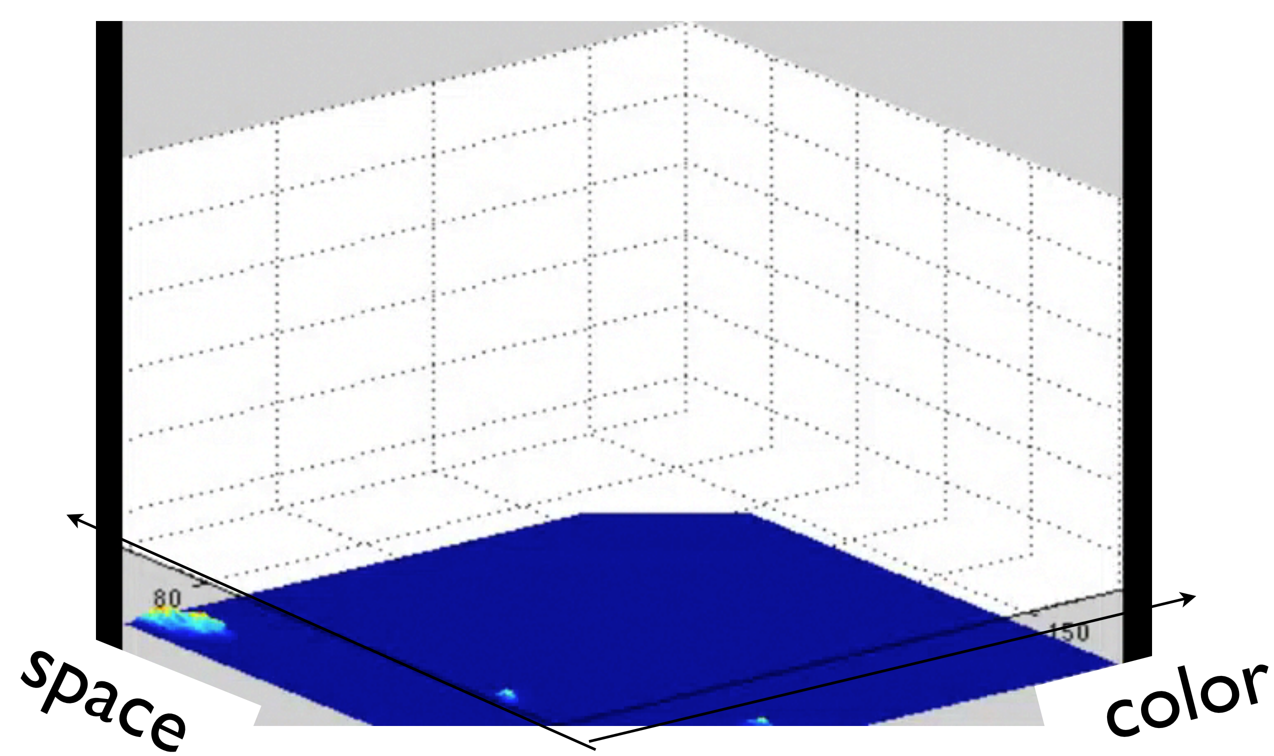
[Sandamirskaya, Schöner: *Neural Networks* 23:1163 (2010)] ordinal stack

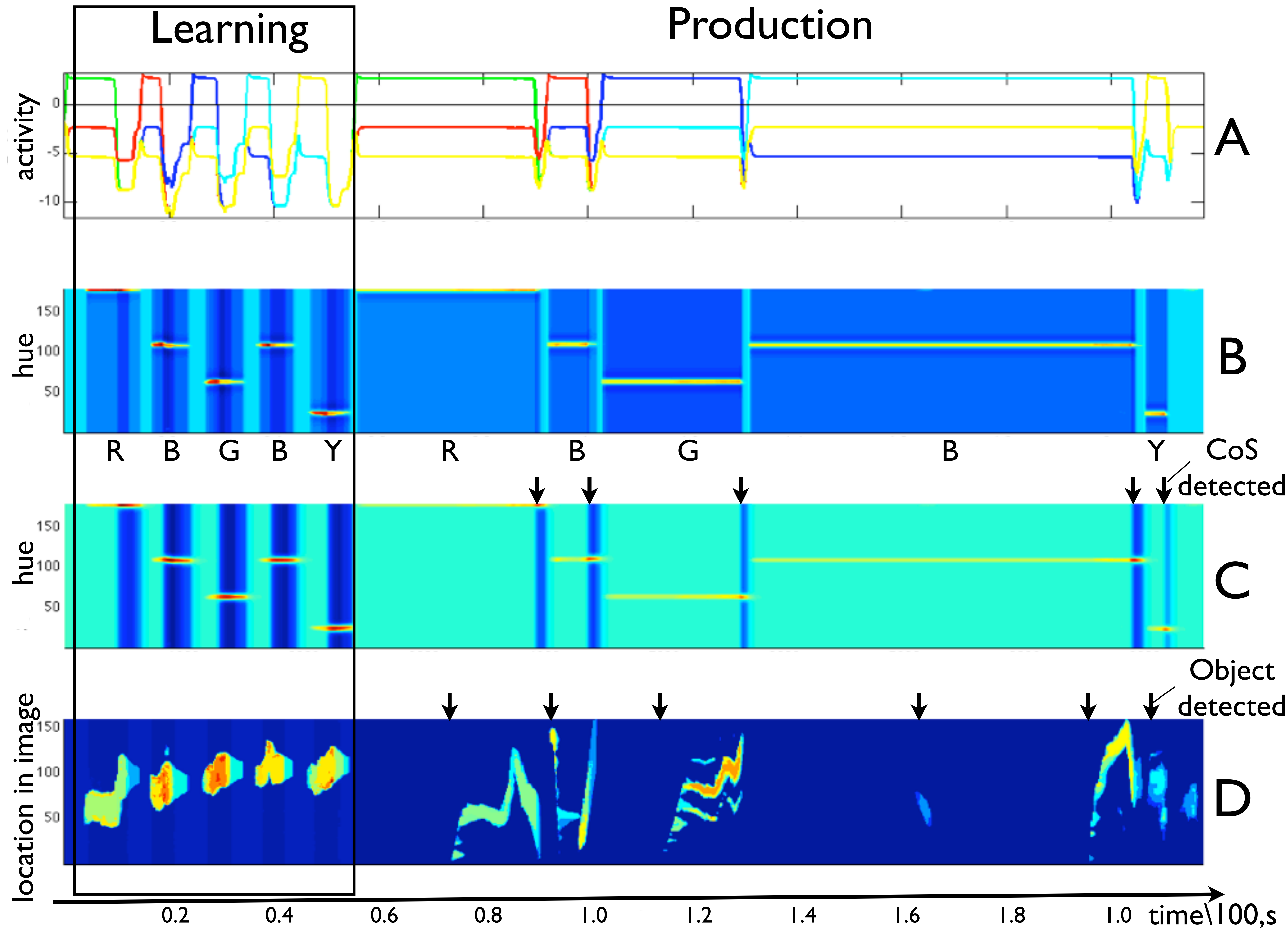
condition of satisfaction (CoS)

intentional state



2D color-space field

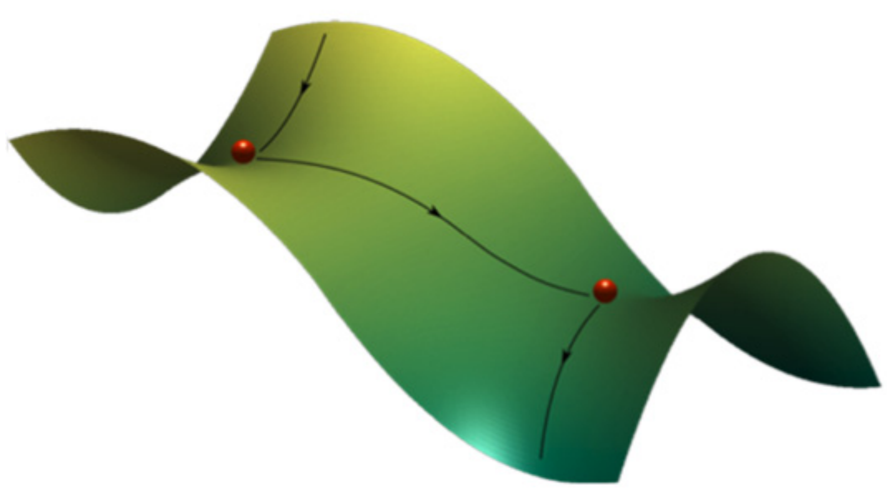


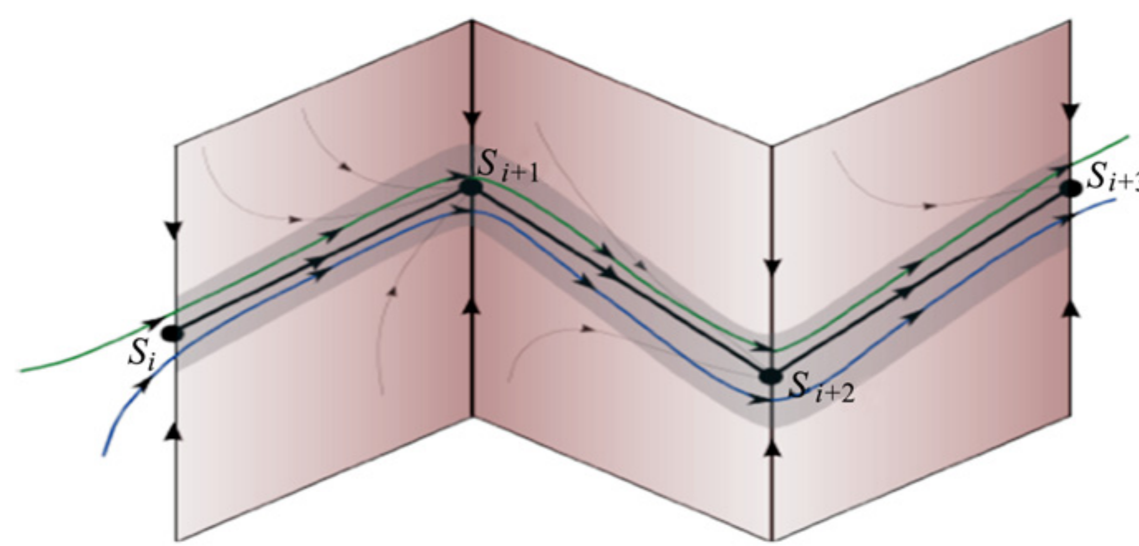


Learning
Production
activity
0
-5
-10
A
hue
150
100
50
B
R
B
G
B
Y
R
B
G
B
Y
CoS detected
hue
150
100
50
C
Object detected
location in image
150
100
50
D
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
1.0
time\100,s

Global view of sequence generation

- Rabinovich’s heteroclinic chain: many more dimensions are stable then unstable…
- the stability of neural attractors is the organizing principle!

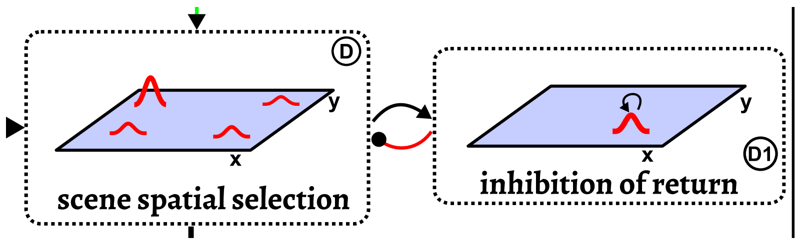


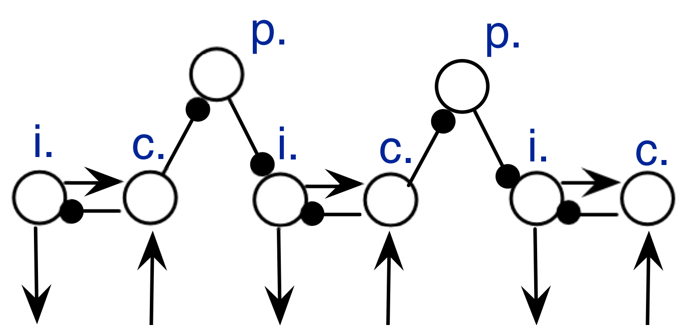


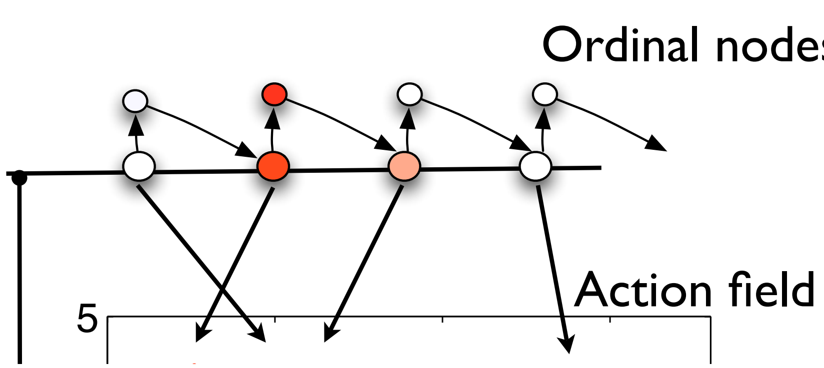
[Rabinovich et al., *Physics of Life Reviews* 2011]

What state next?

- 1 gradient-based selection
- 2 chaining
- 3 positional representation



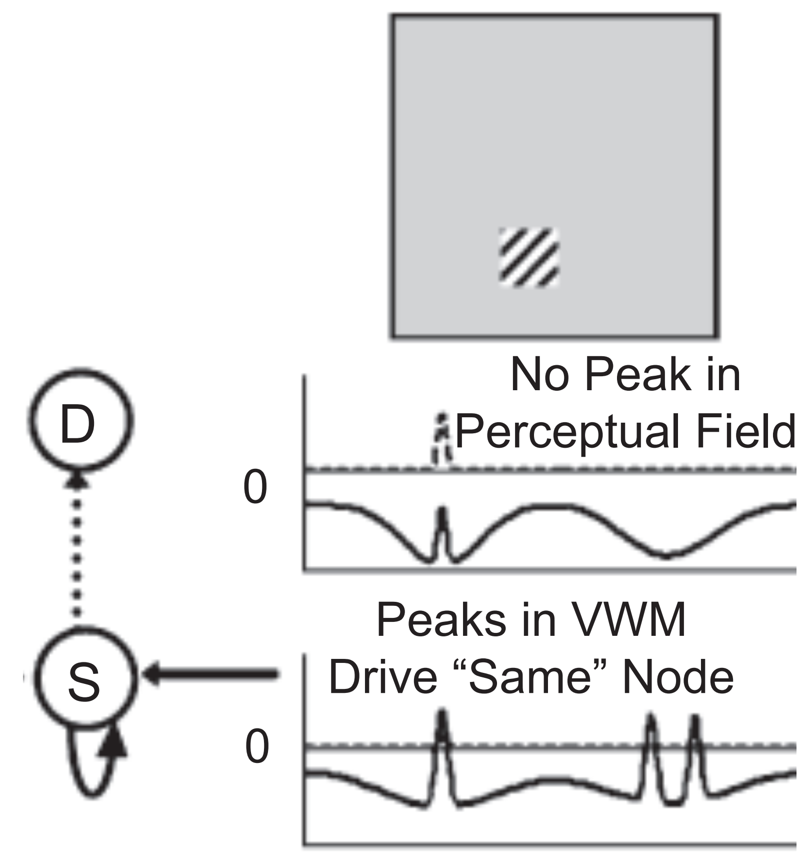




[Henson Burgess 1997]

What if CoS is not reached?

- Condition of Dissatisfaction (CoD) is the default answer after a time out of a “clock”
- may be interrupted by the CoS



[Johnson, et al. 2009]

Intentional agent

- intentionality
- direction of fit: "action" (W2M) and "perception" (M2W)
- psychological mode
- thinking is "action"

Intentionality

- *Intentionality* = the capacity of organisms and their nervous systems to generate mental states that are *about things in the world*
    - *things* may include an organism's own body
    - *things* may ultimately also includes the nervous system's own states

Two *directions of fit* of intentional states (Searle)

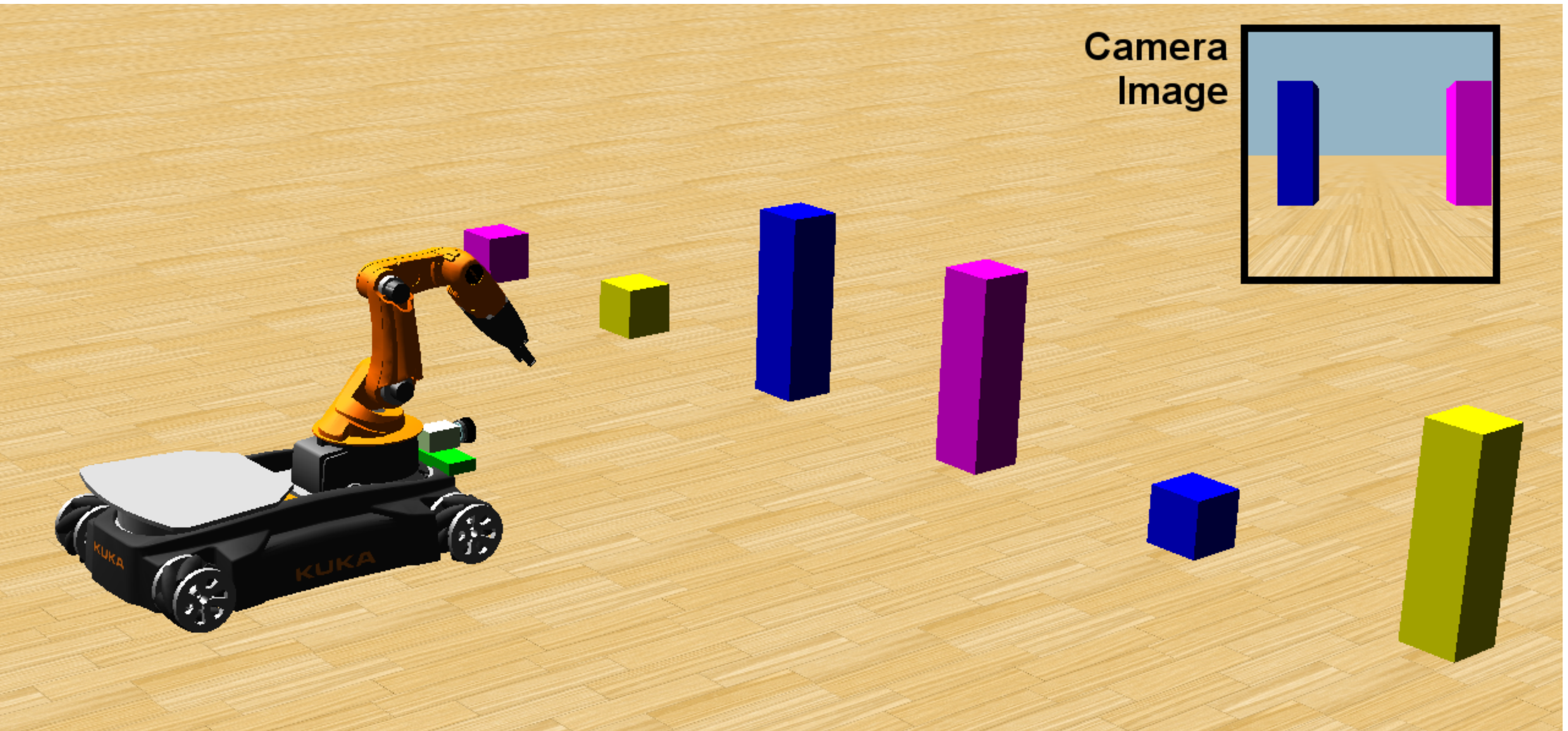
- *world-to-mind*: the world must match the intentional state to fulfill that state's *condition-of-satisfaction* (CoS) => the "motor" flavor of intentionality
- *mind-to-world*: the intentional state must match the state of the world to fulfill the CoS => "perceptual" flavor of intentionality

Six *psychological modes* of intentional states (Searle)

- *mind-to-world*
  - *perception*
  - *memory*
  - *belief ~ knowledge*
- *world-to-mind*
  - *intention-in-action ~ action*
  - *prior intention ~ plans*
  - *desire ~ goals*

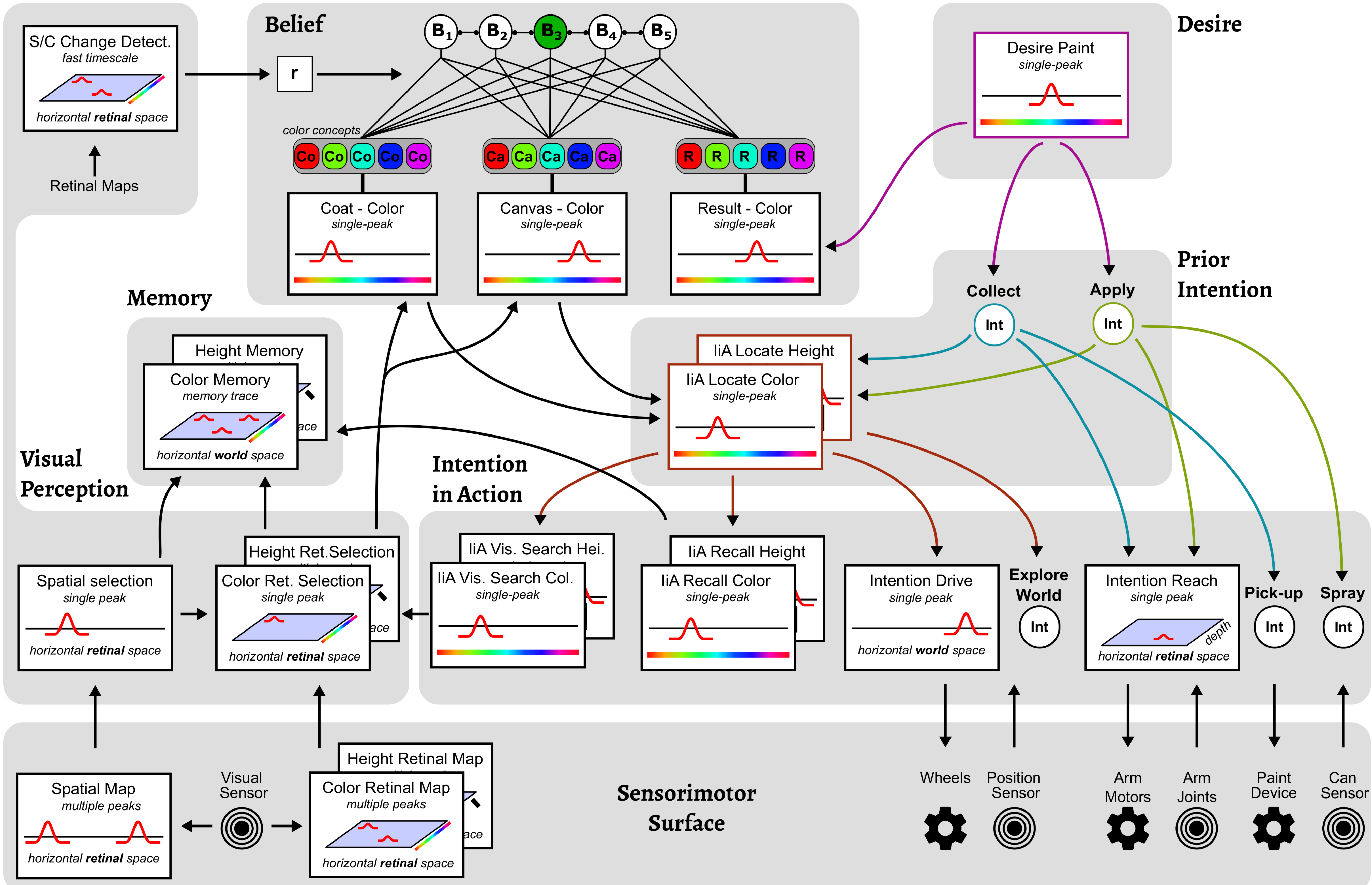
An intentional agent

[Tekülve, Schöner *IEEE Trans Cog Dev Sys* (2022);
Tekülve, Schöner *Cognitive Science* (2024)]



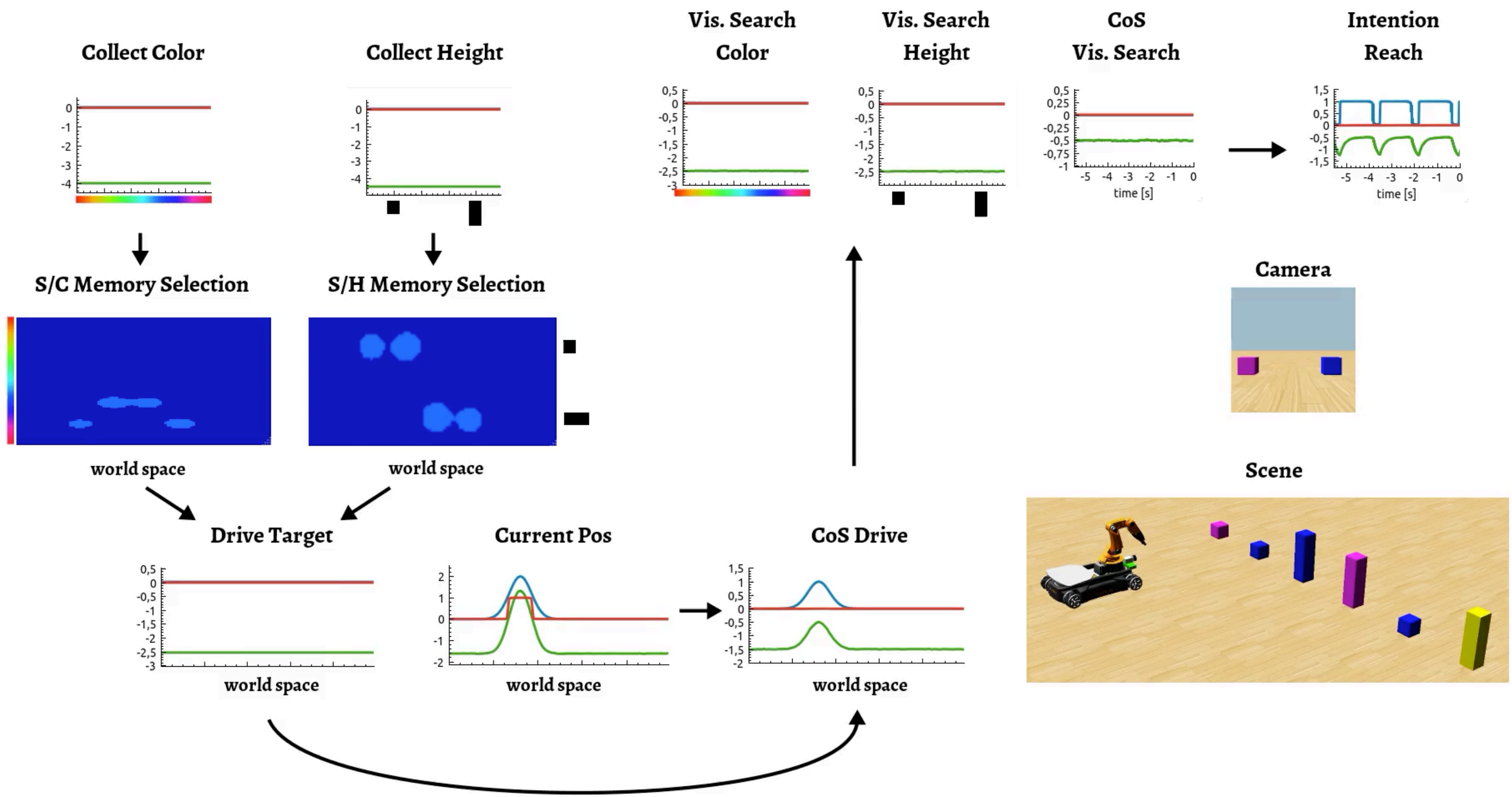
- agent
- environment
- perception
- intention in action
- memory
- prior intentions
- beliefs (propositional)
- desires

Neural dynamic architecture



Recall-visual search-drive

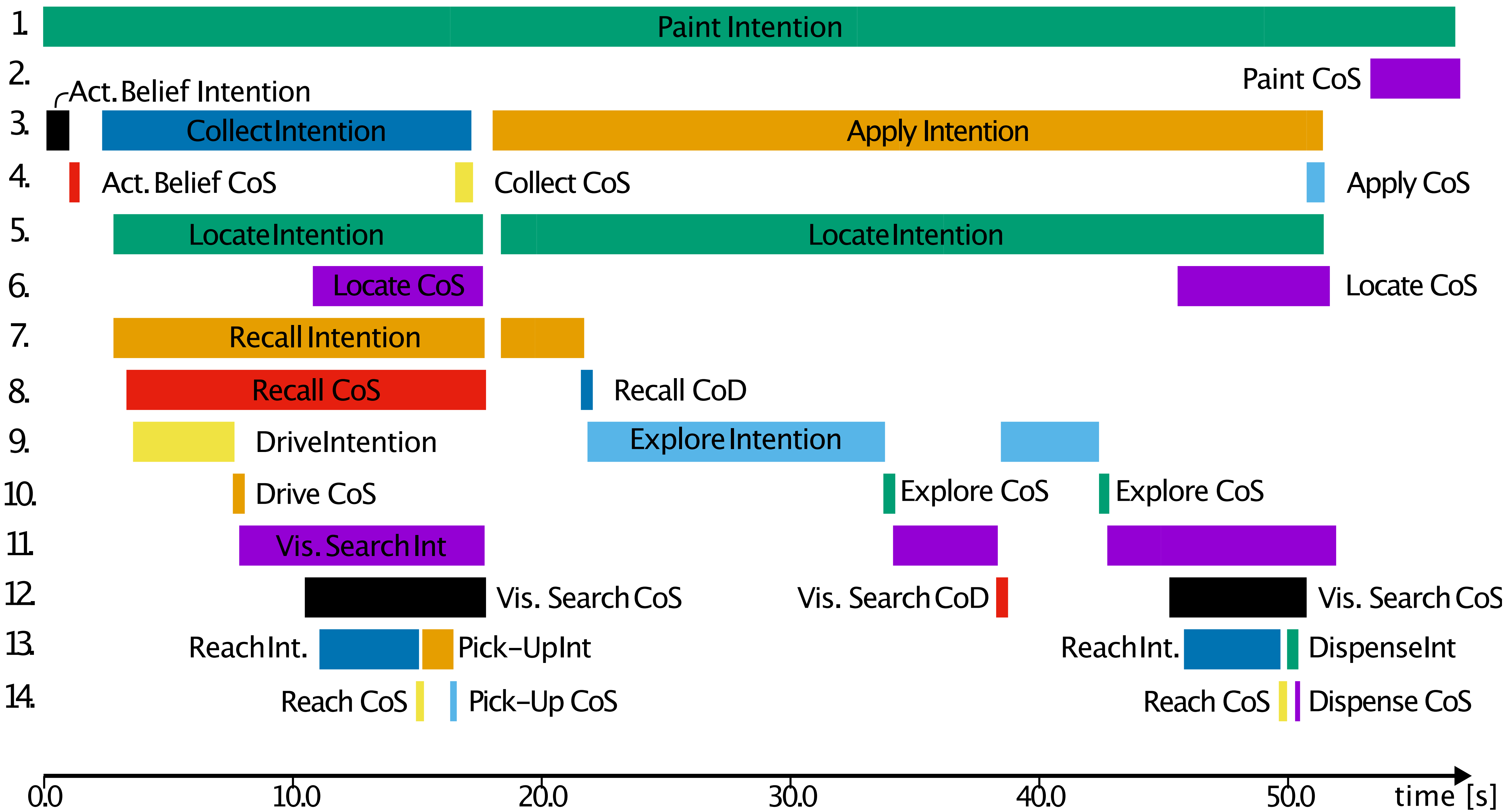
[based on a desire and an activated belief, looking for a tall pink object, which is in memory]



Instantiating a “perception” intention is an “action” intention

- e.g. cued recall from memory, visual search, opportunistic activation of knowledge
- => time structure: intention active until its CoS becomes active, then de-activated
- most thinking is of that nature !

Achieving a desire



Key principles

- 1 activation = state of neural networks
- 2 sigmoidal threshold function
- 3 functional significance of activation comes from its connectivity …
- 1, 2, 3 = connectionism

Key principles

- 4 autonomy
  - conceptually, connectionist networks are input driven (M2W)
  - thought and action are driven by the inner state of the mind = autonomy (W2M)
  - => neural dynamics with strong interaction

Key principles

- 5 higher cognition: binding
  - joint representations
  - expansion/contraction coupling
  - binding through space/index
  - coordinate transforms

What skills have you learned?

- academic skills
    - read and understand scientific texts
    - write technical texts, using mathematical concepts and illustrations

What skills have you learned?

- mathematical skills
  - conceptual understanding of dynamical systems
  - capacity to read differential equations and illustrate them
  - perform “mental simulation” of differential equations
  - use numerical simulation to test ideas about an equation

What skills have learned?

- interdisciplinary skills
  - handle concepts from a different discipline
  - handle things that you don’t understand
  - sharpen sense of what you understand and what not