Summary

Gregor Schöner
gregor.schoener@ini.rub.de
Introduction

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

Braitenberg
Emergent behavior: this is a dynamics

- feedforward nervous system
- + closed loop through environment
- => (behavioral) dynamics
Emergent cognition from neural dynamics

- mental decisions, working memory..
Neurons as input-output units

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

The diagram illustrates the threshold behavior of a neuron with a stimulus applied at different time points. The potential changes over time, showing the firing threshold for each shock applied. The graph on the right plots time in milliseconds against potential in millivolts (mV), with shocks applied at time points 1, 2, and 3.
temporal summation

- Neuron
- Schematic representation of neuron
- Graph showing potential over time with shocks and firing threshold
- Time (milliseconds) on the x-axis
- Potential in 10 mV on the y-axis
Neural dynamics

- replace spiking by a sigmoidal threshold function
- as an abstraction of the membrane potential
  => low levels of activation are not transmitted (to other neural systems, to motor systems)
  => high levels of activation are transmitted
  threshold at zero (by definition)
Neural dynamics

- Stationary state = fixed point = constant solution

- Stable fixed point: nearby solutions converge to the fixed point = attractor

\[
\frac{du(t)}{dt} = \dot{u}(t) = -u(t) + h \quad (h < 0)
\]
Neural dynamics

Attractor structures ensemble of solutions = flow

\[ \tau \dot{u}(t) = -u(t) + h \]

Graph showing neural dynamics with vector-field representation.
**Neuronal dynamics**

- Inputs = contributions to the rate of change
  - Positive: excitatory
  - Negative: inhibitory
- \( \tau \dot{u}(t) = -u(t) + h + \text{inputs}(t) \)
Neuronal dynamics with self-excitation

- single activation variable with self-excitation
- representing a small population with excitatory coupling

\[ \tau \dot{u}(t) = -u(t) + h + s(t) + c \sigma(u(t)) \]
Neuronal dynamics with self-excitation

\[ \tau \dot{u}(t) = -u(t) + h + S(t) + c\sigma(u(t)) \]
Stability from neural dynamics

- **autonomous activation from interaction**
- \( \dot{u}(t) = -u(t) + h + \text{input}(t) + \sigma(u(t)) \)
- **detection instability**
- **working memory**
- **reverse detection instability**
Neuronal dynamics with competition

- Two activation variables with reciprocal inhibitory coupling
- Representing two small populations that are inhibitorily coupled

\[
\begin{align*}
\tau \dot{u}_1(t) &= -u_1(t) + h + s_1(t) - c_{12} \sigma(u_2(t)) \\
\tau \dot{u}_2(t) &= -u_2(t) + h + s_2(t) - c_{21} \sigma(u_1(t))
\end{align*}
\]
Neuronal dynamics with competition

**Coupling**: the rate of change of one activation variable depends on the level of activation of the other activation variable.

\[
\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t)) \\
\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))
\]
Neuronal dynamics with competition

=> biased competition

before input is presented

after input is presented
Neural fields

self-excitation

mutual inhibition

self-excitation
Neural fields

... the same underlying math

coupling among continuously many activation variables

local excitatory coupling ("self-excitation")

global inhibitory coupling ("mutual inhibition")

\[ \tau \dot{u}(x, t) = -u(x, t) + h + s(x, t) + \int dx' w(x - x') \sigma(u(x', t)) \]

\[ w(x-x') \]

\[ x-x' \]
Neural fields

- Forward connectivity thus generates a map from sensory surface to feature dimension
- Neglect the sampling by individual neurons => activation fields

sensory signal, \( s(x) \)

activation field, \( u(y) \)
Example motion perception: space of possible percepts
Neural fields

gous notion for forward connectivity to motor surfaces...

(e.g., through neural oscillators and peripheral reflex loops)
Example: movement planning: space of possible actions
Dynamic of neural fields

- peaks as attractors
- detection instability
- working memory
- selection

$$\tau \dot{u}(x, t) = - u(x, t) + h + s(x, t) + \int dx' w(x - x') g(u(x', t))$$
Attractors and their instabilities

- input driven solution (sub-threshold)
- self-stabilized solution (peak, supra-threshold)
- selection / selection instability
- working memory / memory instability
- boost-driven detection instability

Noise is critical near instabilities
The detection instability stabilizes decisions

threshold piercing
detection instability

activation

stable state
The detection instability leads to the emergence of events.

The detection instability explains how a time-continuous neuronal dynamics may create macroscopic events at discrete moments in time.
Selection decisions are stable

[Wilimzig, Schöner, 2006]
reaction time (RT) paradigm

imperative signal = go signal

response

task set

time

RT
metric effect

- Predict faster response times for metrically close choices than for metrically far choices.

[from Schöner, Kopecz, Erlhagen, 1997]
**The memory trace**

- activation leaves a trace that may influence the activation dynamics later… in a simplest form of learning, the “bias” term of NN.

- powerful in DFT because the detection instability may amplify the induced into peaks of activation.
Tuning of neurons

Neural grounding

Bastian, Riehle, Schöner, 2003
Distribution of Population Activation (DPA) <=> neural field

Distribution of population activation = \[\sum_{\text{neurons}} \text{tuning curve} \times \text{current firing rate}\]

**Note:** neurons are not localized within DPA!

[Bastian, Riehle, Schöner, 2003]
Decision making in DPA

Evidence that the nervous system can simultaneously represent multiple potential actions suggests a straightforward interpretation of the finding, described above, that early responses in many premotor and parietal regions first appear to encode information about relevant stimuli and later change to encode motor variables. Perhaps the early activity, time-locked to stimulus appearance, does not encode the stimuli themselves but rather the set of potential actions that are most strongly associated with those stimuli (Wise et al. 1996), such as actions with high stimulus-response compatibility (Crammond & Kalaska 1994). This would imply that the functional role of this activity does not change in time from sensory to motor encoding but simply reflects the arrival of selection influences from slower but more sophisticated mechanisms for deciding which action is most appropriate.

Recent computational models have proposed that whenever multiple potential targets are available, representations of potential actions emerge within several frontoparietal neural populations, each composed of a continuum of cells with different preferences for the potential parameters of movement (Cisek 2006, Erlhagen & Schöner 2002, Tipper et al. 2000). In each population, cells with similar preferences mutually excite each other (even if they...

[Cisek, Kalaska 2005]
Joint representations

- "anatomical" binding
- example: a joint representation of color and visual space "binder" these two dimensions

[Schneegans et al., Ch 5 of *DFT Primer*, 2016]
Extract the bound features

- project to lower-dimensional fields
- by summing along the marginalized dimensions
- (or by taking the softmax)

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

- project lower-dimension field onto higher-dimensional field as “ridge input”

[Schneegans et al., Ch 5 of DFT Primer, 2016]
Assemble bound representations

Ridge Intersections

-30° −20° −10° 0° 10° 20° 30°

Activation color (hue value)

Space-color field

Visual scene

Spatial field

[Schneegans et al., Ch 5 of DFT Primer, 2016]
Assemble bound representations

- **binding problem:** multiple ridges along lower-dimensional space lead to a correspondence problem

- => assemble one object at a time...

- => sequentiality bottleneck!

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

- Ridge input along one dimension extracts from bound representation matching objects
- Other dimensions of those objects can then be extracted
- E.g. visual search

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

[Slides adapted from Sebastian Schneegans, see Schneegans, Chapter 7 of Dynamic Field Theory-A Primer, OUP, 2015]
Perceptual grounding of a relation: bringing the target object into the attentional foreground

“red to the left of green”

[Lipinski, Sandamirskaya, Schöner 2009 … Richter, Lins, Schöner, Topics 2017]
into the reference and target field and enable these fields to track moving objects even if spatial attention is currently focused elsewhere.

3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a "T." Steerable neural mappings are denoted as diamonds.
“red to the left of green”
Sequential behaviors or mental acts

- behaviors/mental states are attractors
- that resist change…
- to induce change in sequential behavior/thinking: induce an instability
Sequence generation

- the CoS organizes the transition away from an ongoing behavior/mental state

- based on a signal from perception or from an inner state of a neural architecture that is predicted to be indicative of successful completion of the behavior/mental act
What skills do you learn?

- **academic skills**
  - read and understand scientific texts
  - write technical texts, using mathematical concepts and illustrations
What skills do you learn?

- 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 do you learn?

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