Chair of Cognitive Modeling Department of Computer Science & Department of Psychology Faculty of Science University of Tübingen



An Integrative, Event-Predictive Theory of Cognition

Behavioral Evidence and Artificial Neural Network Models



Example 1: The Disappearing Hand Trick (closely related to the Rubber Hand Illusion)



by Roger Newport, Helen Gilpin, & Catherine Preston



General Explanation of the Disappearing Hand Trick

Our brain attempts to

continuously maintain a

consistent postural "image" of

our own body's current state.

(body posture as well as its position and orientation relative to the outside environment)



Example 2: Levesque's Winograd Schema Challenge

(Levesque, 2011), The Winograd Schema Challenge

• Consider the following sentence:

The ball fits into the suitcase, because it is **large**.

Grammatical regularities lead to (slight) garden path effect:

> Pronoun is more likely to reference the subject, i.e., "the ball".

• Which leads to **surprise** and a revision of the pronoun referent model.



Our brain attempts to

continuously construct a

consistent semantic "image" of

the described state of affairs.



5 | Martin V. Butz | Event-Predictive Cognition

General Explanation for Both Examples

- Our brain attempts to continuously construct (on the fly) a consistent postural / semantic "image" of the perceived state of affairs.
- It predicts next sensory inputs (e.g. words, visual information about the hand) and is surprised when certain predictions are violated, attempting to revise the "image" accordingly.

In the examples:

- Reaching for the other hand yields empty space leading to the disappearing hand illusion and then quickly the revision of the postural image.
- 2. Binding of pronoun *it* yields a "**semantic model**" that indicates a reasoning inconsistency, thus requiring a revision.



- What is a semantic image ?
- > How is it constructed by and within our brain ?
- > Which encodings are involved ?
- How are the involved encodings dynamically (& selectively) activated ?
- How can these encodings be learned ?



Brain as a Generative / Predictive Model

(cf. e.g. Moshe Bar, Horace Barlow, Lawrence Barsalou, Karl Friston, Peter König, and many others...)

- The brain develops a **generative, predictive model** of the encountered environment in order to interact with it in an effective, goal-directed manner.
- The **free energy principle** allows the formalization of this development including the involved behavior but ...
 - ... the formalism is overly general.
 - > Structural, inductive biases –

that is, tendencies to develop certain encoding structures – must shape cognitive development further!



Types of Predictive Encodings & Inductive Biases

- At least 3 fundamental types of predictive encodings are distinguishable:
 - Spatial
 - E.g.: relative spatial mappings of entities and frames of reference
 - Top-down
 - E.g.: Gestalt-related perceptual expectations
 - Temporal
 - E.g.: Sensorimotor codes but also force-dependent codes
- Significant free energy signal changes lead to the encoding of
 - Events
 - Active set of P.E.s
 - Predictive attractors with low residual error
 - Event boundary / event transition encodings
 Predictions across event attractors

- Event-schemata can be formed, which encode
 - **Preconditions** necessary to initiate an event;
 - Final consequences that are typically reached when the event ends;
 - The event code itself.
- Hierarchically organized event encodings enable
 - hierarchical, conceptual, goal-directed planning / reinforcement learning and behavioral control;
 - conceptual and hypothetical reasoning.



Nested Event Schemata Yield Hierarchical Structure

Event Boundary



• Events

- Predictive (probabilistic) encoding **network** of how (sensorimotor, -force, & more abstract) dynamics unfold over time.
- Event Boundaries
 - Conditional predictive encodings of event transitions.
- Event Hierarchy
 - Hierarchical / taxonomical structure of events, event boundaries and their typical interactions.



Unfolding Predictive & Active Inferences



Butz, M. V. (2016). Towards a Unified Sub-Symbolic Computational Theory of Cognition. *Frontiers in Psychology*, *7*, 10.3389/fpsyg.2016.00925. Butz, M. V. & Kutter, E. F. (2017). How the Mind Comes Into Being: Introducing Cognitive Science from a

Functional and Computational Perspective. Oxford, UK: Oxford University Press.

"The ball fits into the suitcase, because it is large." **The ball ...**







"The ball fits into the suitcase, because it is large."

... fits into ...









"The ball fits into the suitcase, because it is large." ... the suitcase









"The ball ... fits into ... the suitcase."



16 | Martin V. Butz | Event-Predictive Cognition





"The ball fits into the suitcase "



"... because it is large."

This implies:

"Was it small, then the first part of the sentence would be false."

Counterfactual reasoning.

- "... because **the ball** is large?"
- > "Was the ball small, the ball would still fit.





"The ball fits into the suitcase "



"... because it is large."

This implies:

"Was it small, then the first part of the sentence would be false."

Counterfactual reasoning.

"... because the **suitcase** is large?"

"Was the suitcase small, the ball would NOT fit."





18 | Martin V. Butz | Event-Predictive Cognition

Neuro-Computational Cognitive Models

... and some psychological experiments.

- 1. Event Dynamics
- 2. Inference of and within Events
- 3. Event Boundary Anticipation
- 4. Event Schemata: Segmentation and Inference







Learning to Infer Biological Motion Events

A Perspective Taking, Gestalt Inference, & Mirror Neuron Model

INFERENCE OF BIOLOGICAL MOTION

Schrodt, F. (2018). PhD dissertation.

Schrodt, F., & Butz, M. V. (2016). Just Imagine! Learning to Emulate and Infer Actions with a Stochastic Generative Architecture. *Frontiers in Robotics and AI*. doi:10.3389/frobt.2016.00005



20 | Martin V. Butz | Event-Predictive Cognition

An ANN of Biological Motion Inference



- "Intention module" clusters types of motion via a supervised temporal autoencoder
- Spatiotemporal autoencoders predict cluster-interdependencies and dynamics over time
- Top-down predictive autoencoders yield biological motion clusters
- Separation into posture, motion direction, & magnitude Gestalt encodings
- Perspective- & focus-relative encodings



Generative ANN Model Evaluation



Exemplar Model Capabilities



Stimulus and FORsvs.Model ExpectationStimulus and FORsvs.Model Expectation(1/5 speed)(1/5 speed)(1/5 speed)





Sensorimotor forward model used for control NEURAL ACTIVE-INFERENCE-BASED IMAGINING AND CONTROL

Sebastian Otte, Theresa Schmitt, Karl Friston, & Martin V. Butz (2017). Inferring adaptive goaldirected behavior within recurrent neural networks. *ICANN 2017*.



The RNN Model

- Recurrent neural network architecture (LSTM) learns a sensorimotor forward model.
- Scenario: A "rocket" with simulated gravity and two thrust motors

"Rocket" scenario



- Input to the LSTM:
 - Current location of rocket and current motor activity
- Output of the LSTM:
 - Resulting velocity and next location of rocket
- During training:
 - Training samples of somewhat systematic random explorations
- During testing:
 - Active inference-based motor control derivation



Controlling a Glider (Otte, Schmitt, Friston, & Butz, ICANN 2017)

- Learning a sensorimotor forward model with an RNN (LSTM)
- Control via active inference:
 - Attempting to reach goal state M_G
 - Inferring policy π by maximizing expected Q_{τ} over expected future model states M_{τ} and observations o_{τ} .

$$Q_{\tau}(\pi, s_{t}) = -D_{KL}[P(\boldsymbol{o}_{\tau}|\boldsymbol{M}_{\tau})||P(\boldsymbol{o}_{\tau}|\boldsymbol{M}_{G})] - E_{\pi}[H(P(\boldsymbol{o}_{\tau}|\boldsymbol{M}_{\tau}))]$$





Sebastian Otte, Theresa Schmitt, Karl Friston, & Martin V. Butz (2017). Inferring adaptive goal-directed behavior within recurrent neural networks. *ICANN 2017*.



Concurrent Event Inference and Goal-Directed Control REPRISE: A RETROSPECTIVE & PROSPECTIVE INFERENCE SCHEME

Martin V. Butz, David Bilkey, Alistair Knott, & Sebastian Otte (in press). REPRISE: A Retrospective and Prospective Inference Scheme. CogSci 2018.



27 | Martin V. Butz | Event-Predictive Cognition

The Setup

- Recurrent neural network architecture (LSTM) learns a sensorimotor forward model.
 - > With dedicated, stable hidden state event indicator neurons
- 3 "Event" Scenario:
 - 1. "Rocket" simulation (2 thrust directions, inertia gravity)
 - 2. Stepper simulation (4 combinable directional step motions, no inertia)
 - 3. Glider simulation (4 thrust directions, inertia)
- Input to the LSTM:
 - Current location of the "vehicle" and current motor activity
- Output of the LSTM:
 - Resulting motion of rocket
- During training:
 - Switching between three vehicles.
 - Switch triggers
 - 1. Retrospective (stable) hidden state inference
 - 2. BPTT learning
- During testing:
 - Retrospection: hidden state inference
 - Prospection: Active inference-based motor control derivation



"Stepper" event



"Glider" event





Martin V. Butz, David Bilkey, Alistair Knott, & Sebastian Otte (in press). REPRISE: A Retrospective and Prospective Inference Scheme. CogSci 2018.

REPRISE: Retro- and Prospective Inference Dynamics





29 | Martin V. Butz | Event-Predictive Cognition

Martin V. Butz, David Bilkey, Alistair Knott, & Sebastian Otte (in press). REPRISE: A Retrospective and Prospective Inference Scheme. CogSci 2018.

REPRISE Performance





Goal-Oriented Object Interactions

ANTICIPATING UPCOMING EVENT BOUNDARIES

Belardinelli, A. & Butz, M. V. (2013). Gaze strategies in object identification and manipulation. *Annual Conference on Cognitive Science (CogSci 2013)*, 1875-1880.

Belardinelli, A., Herbort, O., & Butz, M. V. (2015). Goal-oriented gaze strategies afforded by object interaction. *Vision Research, 106*, 47–57.

Belardinelli, A., Stepper, M. Y., & Butz, M. V. (2016). It's in the eyes: Planning precise manual actions before execution. *Journal of Vision*, *16(1)*, 18. doi:10.1167/16.1.18

Belardinelli, A., Barabas, M., Himmelbach, M., & Butz, M. V. (2016). Anticipatory eye fixations reveal tool knowledge for tool interaction. *Experimental Brain Research*, 234, 2415–2431.



Videos of Typical Trial Interactions

21em 8em 95cm 13.5 6cm 5.5cm

3.5cm

<u>Task:</u> Grasp & *Drink* from object

<u>Task:</u> Grasp & *Handover* object





32 | Martin V. Butz | Event-Predictive Cognition

Time Course of the Average Trial

• Note the normalized time window from the first mapped fixation until grasp onset.





33 | Martin V. Butz | Event-Predictive Cognition

Heat maps until object grasp: Objects are fixated in the light of the final event

Task: Grasp & drink from object

Grasp and hand over object



Task:

34 | Martin V. Butz | Event-Predictive Cognition



Anticipatory Eye Behavior – Time Course





1.	Event Dynamics
2.	Inference of and within Events
3.	b) Event Boundary Anticipation (Body)
4.	Event Schemata: Segmentation and Inference

Manual peripersonal space maps into the future

ANTICIPATING BODILY STATES AT FUTURE EVENT BOUNDARY

Belardinelli, A., Lohmann, J., Farnè, A., & Butz, M. V. (2018). Mental space maps into the future. *Cognition*, *176*, 65–73. doi:10.1016/j.cognition.2018.03.007



Event-Predictive Encodings in Action

- Focus:
 - Manual interactions with objects
 - Peripersonal space around the hand
 - > Known to integrate multiple sensory information sources.
 - Surrounds hand / face and other body parts.
 - \succ Can be modified via tool usage.
- Dual task paradigm with
 - 1. Expected cross-modal congruency
 - Does a light flash close to FUTURE FINGER POSITIONS influence the detection of vibrotactile stimulations of the respective fingers?
 - 2. Task-oriented object interaction
 - Transport an object to the right and place it upright



Event-Predictive Encodings in Action – Trial Schedule





38 | Martin V. Butz | Event-Predictive Cognition

Belardinelli, A., Lohmann, J., Farnè, A., & Butz, M. V. (2018). Mental space maps into the future. *Cognition*, 176, 65–73. doi:10.1016/j.cognition.2018.03.007

Cross-Modal Congruency Effect

Vibration detection response time is influenced by the visual distractor in anticipation of future hand posture (finger locations). (data shown is averaged over all three SOAs.)





39 | Martin V. Butz | Event-Predictive Cognition

Belardinelli, A., Lohmann, J., Farnè, A., & Butz, M. V. (2018). Mental space maps into the future. *Cognition*, 176, 65–73. doi:10.1016/j.cognition.2018.03.007



A computational model of event & event transition detection, abstraction, and planning via free-energy based (active) inference **LEARNING EVENT TAXONOMIES FROM SENSORIMOTOR EXPERIENCES**

Gumbsch, C., Kneissler, J., & Butz, M. V. (2016). Learning behavior-grounded rvent segmentations. Proceedings of the 38th Annual Meeting of the Cognitive Science Society (pp. 1787–1792). Austin, TX: Cognitive Science Society. Gumbsch, C., Otte, S., & Butz, M. V. (2017). A computational model for the dynamical learning of event taxonomies. Proceedings of the 39th Annual Meeting of the Cognitive Science Society (pp. 452–457). London, UK: Cognitive Science Society.

Gumbsch, C. (2018). Master dissertation.



Computational Model of Event Processing and Inference

Predictive processing & learning



Hierarchical, active-inference-based goal-directed behavior



- Event Models are sensorimotor forward models
- Event Boundary Models are multivariate Gaussians
- **Predictive Perceptual Space** yields information fusion and imaginations
- Motivations selectively activate EBMs to reach anticipated reward
- Motor system executes motor commands and provides efference copies
- **Observations** yield relative locations of hand, mouth, and object, as well as object properties.

Gumbsch, C., Otte, S., & Butz, M. V. (2017). A computational model for the dynamical learning of event taxonomies. Proceedings of the 39th Annual Meeting of the Cognitive Science Society (pp. 452–457). London, UK: Cognitive Science Society.



Learned Behavioral Events

• Starting with Differential Extrinsic Plasticity control

(Ralf Der & Georg Martius)





42 | Martin V. Butz | Event-Predictive Cognition

Goal-Directed Event-Predictive Control

Reaching Goals

Collecting & Transporting Objects







43 | Martin V. Butz | Event-Predictive Cognition

Gumbsch, C. (2018). Master dissertation.

Hexapod Inferring Event Boundaries and Event Schemata









A Mario clone to explore the learning of event schema rules and the link to language.

EVENT-ORIENTED ABSTRACTIONS TOWARDS LANGUAGE COMPREHENSION



Mario's Brain – SEMLINCS Architecture





Schrodt, F., Kneissler, J., Ehrenfeld, S., & Butz, M. V. (2017). Mario becomes cognitive. *Topics in Cognitive Science*, *9*(2), 343–373. doi:10.1111/tops.12252

Main Features of SEMLINCS Architecture



- Speech interface enables
 - querying and manipulating:
 - Event schema knowledge
 - Goal selection
 - Motivational system state
 - Current motor commands.
 - autonomously agreeing on (sequential) joint action plans.

- Motivational system
 - Based on homeostatic variables;
 - Allows the autonomous selection of goals.
- Event schema knowledge in the form of condition-action-effect rules...
 - ... is learned from surprising event signals (e.g. disappearing object);
 - ... allows temporal forward predictions and inverse planning on a conceptual level.
- Planning:
 - Sensorimotor planning is currently hard-coded (A*) relying on game engine (simulator)
 - Schematic planning relies on event schema rule knowledge (i.e. production rules)



Schrodt, F., Kneissler, J., Ehrenfeld, S., & Butz, M. V. (2017). Mario becomes cognitive. *Topics in Cognitive Science*, *9*(*2*), 343–373. doi:10.1111/tops.12252

Example 1: Learning from Object Interaction Events and Observations





48 | Martin V. Butz | Event-Predictive Cognition

Schrodt, F., Kneissler, J., Ehrenfeld, S., & Butz, M. V. (2017). Mario becomes cognitive. *Topics in Cognitive Science*, *9*(2), 343–373. doi:10.1111/tops.12252 Mario Becomes Social!" video available online on YouTube.

Example 2: Coordinating Plans

Mario







Schrodt, F., Röhm, Y., & Butz, M. V. (2017). An Event-Schematic, Cooperative, Cognitive Architecture Plays Super Mario. In *Proceedings of EUCognition* 2016: *Cognitive Robot Architectures* (pp. 10–15).



(Many) Open Challenges

For ANN-based neuro-cognitive models...

- More complex
- More real
- Optimizing motion primitives
- Episodic memory
- Emergent linkage to linguistic structure
- Conceptual abstractions beyond sensorimotor behaviors



