Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent

Jan Tekülve
Institute for Neural Computation
Ruhr-University Bochum
Bochum, Germany
jan.tekuelve@ini.rub.de

Gregor Schöner
Institute for Neural Computation
Ruhr-University Bochum
Bochum, Germany
gregor.schoener@ini.rub.de

***Abstract*—Intentionality is the capacity of mental states to be about the world, both in its "action" (world-to-mind) and its "perception" (mind-to-world) direction of fit. An intentional agent must be able to perceive, act, memorize, and plan. These psychological modes may be driven by desires and be informed by beliefs. We have previously proposed a neural process account of intentionality, in which intentional states are stabilized by interactions within populations of neurons that represent perceptual features and movement parameters. Instabilities in such neural dynamics activated the conditions of satisfaction of intentional states and induced sequences of intentional behavior. Here we explore the idea that the process organization of such intentional neural systems enables autonomous learning. We show how beliefs may be learned from single experiences, may be activated in new situations, and be used to guide behavior. Beliefs may also be dis-activated when their predictions do not match experience, leading to the learning of a new belief. We demonstrate the idea in a simple scenario in which a simulated agent autonomously explores an environment, directs action at objects and learns simple contingencies in this environment to form beliefs. The beliefs can be used to realize fixed desires of the agent.**

***Index Terms*—autonomous agent, fast learning, learning beliefs, neural dynamics, neural cognitive architecture**

I. Introduction

Although neurally inspired learning is increasingly used in robotics and computer vision, such work rarely addresses fast autonomous learning, that is, learning from experience as a system behaves and steers its own perception. Humans learn autonomously during lifelong development. They are able to learn and generalize from single instances of an experience. In fact, belief formation can be so fast as to be sometimes counter-productive, leading to superstitious behavior (even in pigeons [1]).

This ability to efficiently erect or derive beliefs about the environment provides enormous cognitive power [2]. Neurally inspired methods of learning are far from approximating this form of single shot learning. Fast learning has been demonstrated for object recognition based on exploiting prior knowledge about the visual appearance of objects [3] or on built-in knowledge about the transformations that enable generalization [4]. Early in development, contingency learning may be a way how infants break down the complexity of the world, which supports their social skills [5]. Our framing of belief acquisition approximates contingency learning.

We argue that neural models miss an important component for autonomous learning: The process structure to autonomously generate meaningful behavior and stable perceptual representations and thus, to generate experience. That process structure must support recognizing novelty, activating learned representations, and learning. We propose that the philosophical notion of *intentionality*, the capacity to generate internal states that are about the world, helps uncover the requisite process structure. The philosopher John Searle divides intentional states into two classes: The mind-to world direction of fit, which comprises "perceptual" states representing the world, and the world-to-mind direction of fit, which comprises "action" states representing desired world states [6].

We have previously analyzed the neural process requirements for intentional states of both directions of fit [7] (see [8] for a different analysis). Our analysis was based on Dynamic Field Theory (DFT) [9], a mathematical language that describes the neural dynamics in networks of neural populations. In particular, we exploited the notion that intentional states are stable patterns of neural activation, which may transition sequentially to other intentional states by inducing dynamic instabilities through a neural representation of the *condition of satisfaction* [10]. This mechanism gives neural dynamic architectures the potential to autonomously generate behavioral sequences (see [11] for a discussion of autonomy).

Our previous analysis led to a neural process account for four basis level psychological modes of intentionality (perception, memory, intention-in-action, and prior intention). In this paper we argue, that a fast form of learning captures the psychological mode of belief. We think of belief as a form of memory formation that generalizes beyond the specific instance by being categorical and propositional in nature. We are able to construct propositional content in a neural dynamic account again by organizing underlying processes through the condition of satisfaction [12]. We model the autonomous learning of beliefs which encompasses activation of existing beliefs, rejection of activated beliefs, formation of new beliefs, and integration of beliefs into the set of beliefs on which behavior is based.

The account is framed within a rudimentary toy scenario in which an agent is situated in a simple environment containing solely "bucket" and "canvas" objects of different color. The robotic agent explores the environment, moves towards objects and directs an effector to them to either pick-up paint from buckets or to dispense paint onto canvases, observing the resulting color. A network of neural dynamic fields is connected to the agent's sensory-motor surfaces and enables the agent to visually detect and select objects, build scene memories, generate sequences of actions to paint particular objects to achieve a particular color and ultimately to form and activate beliefs about which paint applied to which canvas generates which outcome.

II. Dynamic Field Theory

Dynamic Field Theory (DFT) [9] is a theoretical framework for understanding perception, motor behavior, and cognition based on neural principles. In DFT the activity of neural populations, tuned to metric dimensions, x, is modeled by activation fields, $u(x,t)$, described through the dynamics:

$$\tau\dot{u}(x,t) = -u(x,t)+h+s(x,t)+\int \omega(x-x')\sigma(u(x',t))dx'.$$

The time-continuous evolution of neural activation, $u(x)$, on the time scale τ relaxes to the stable solution, $h+s(x)$, defined by the field's resting level h and its localized inputs, $s(x)$, if the current activation $u(x)$ is below the sigmoidal threshold σ. Field locations with activation surpassing the threshold level engage in lateral interaction defined by the field's kernel, $\omega(x-x')$, which is excitatory locally, and inhibitory over longer distances, $x-x'$. This leads to the emergence of self-stabilized peaks of supra-threshold activation, which are the unit of representation in DFT (see Figure 1). Supra-threshold peaks arise as the sub-threshold state goes through the *detection instability*.

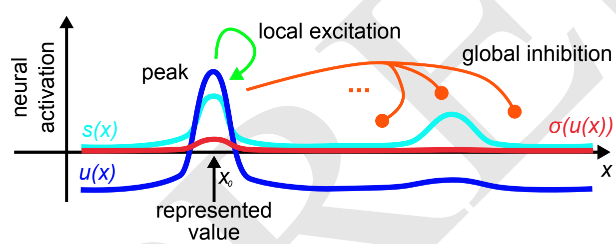

Fig. 1. A dynamic neural field spanned across the metric-dimension x representing value x_0 through a supra-threshold activation peak.

Depending on the individual parametrization of excitatory and inhibitory interaction strength, fields operate in different regimes. In the self-stabilized regime, supra-threshold peaks are stabilized against input noise. In the selective regime, lateral inhibition allows only a single peak at any point in time. In the self-sustained regime, peaks are retained after localized input is removed. Peaks in multi-dimensional fields represent conjunctions of feature dimensions. For instance, a peak in a two-dimensional field that spans across color and position represents a particular color seen at a particular position. Dynamic neural nodes are zero-dimensional fields that represent categorical states.

A field, u_{tar}, receives input from another field, u_{src}, if that field's output, $\sigma(u_{\text{src}})$, adds to the target field's rate of change, $\dot{u}_{\text{tar}}$, weighted with a homogeneous projection kernel $\omega_{\text{tar,src}}$. The source output might need to be contracted or expanded to match the target field's dimensionality [13]. Typically, contractions entail integrating over the excess dimension, while expansions provide input that is constant along the excess dimensions (e.g., ridges, tubes, or slices). *Concept nodes* are connected reciprocally to fields through a pattern of connectivity that encodes the feature representation of the concept. For instance, the concept node for "blue" is connected to an appropriate range of hue values in a hue feature field.

A. Networks of field form architectures

Networks of dynamic neural fields that connect to the sensory-motor surfaces of an agent define architecture from which complex behavior may emerge through autonomous transition between different stable macro states, each represented by peaks of supra-threshold activation. *Boost nodes* provide homogeneous input to a target field, and may induce such transitions by altering the dynamic regime of the target field, so that peaks may form from sub-threshold localized activation. Such boost nodes may effectively modulate the flow of activation within an architecture by enabling or disabling particular branches of the architecture to form peaks. Boost nodes may thus act as "gates", or also as "go" signals that trigger an action by activating a sub-network.

Pairs of fields, an excitatory "intention" field and an inhibitory "condition of satisfaction" (CoS) field, control the initiation and termination of actions or mental states [14] (see Figure 2). The intention field represents the desired end state of a particular action and activates a sub-network that ultimately realizes the desired action. The intention field pre-activates the CoS-field, in which a peak is formed when desired and perceived state overlap sufficiently. A peak in the CoS-field inhibits the intention field, destabilizing the peak there and deactivating the associated sub-network, which terminates the action. The CoS field inhibits any precondition node that prevented competing actions from becoming activated. This unlocks the next step in a sequence.

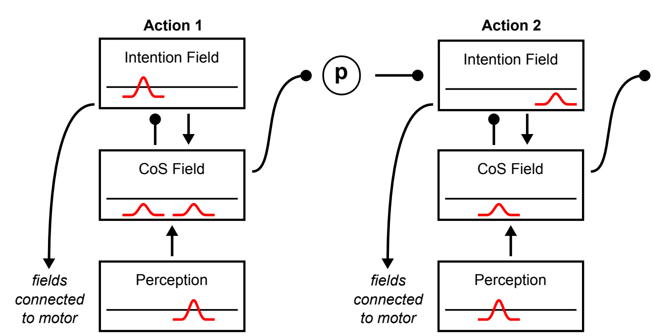

Fig. 2. Two consecutive elements of an action sequence each represented through a pair of intention and CoS field.

Other than through the CoS mechanism, autonomous transitions between macro states may also be induced by neural

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

representations of a “Condition of Dissatisfaction” (CoD) which detect failed actions or invalid perceptual states.

III. Model/Scenario

We illustrate the autonomous formation, utilization, and rejection of beliefs by a dynamic neural field architecture in a toy scenario, in which a robotic agent performs a simple painting task. The task requires the agent to collect paint from buckets (elongated cuboids) and to apply the paint to coat colored canvases (cubes) leading to a new resulting color of the canvas cube. The agent learns the relationship between the color of the paint, and the color of the canvas before and after the coating is applied. This simple toy scenario exemplifies semantic learning in a context that entails both mind-to-world intentionality (from perception to memory) and the world-to-mind intentionality (from prior-intentions to executing intentions-in-action, see Figure 3 for a schematic overview of the field architecture[1].

A. Architecture Capabilities

The simulated world consists of an array of cuboids of different height and color, which the robotic agent faces. The robot may move sideways along the array and reach with its arm towards objects in its current field of view to either collect paint from a tall cube (bucket) or to apply a coat of paint to a small cube (canvas). The architecture is capable of elementary acts of perceptions and of intention in action, which together enable the agent to perform the task. Perceptions may drive memory formation and intentions in action may be performed in sequences that realize prior intentions (see [7] for details).

1) Perceptions: Cuboids in the world, that are currently perceived, are represented as peaks of activation in the *retinal space/color perception* and *retinal space/height perception* fields. Peaks in this fields are induced sequentially through a mechanism of selective spatial attention (see [15] for details). A transient detector allows the agent to detect changes in the world (e.g. a sudden change of color of an object) and to direct spatial attention to the location of change. In addition to visual perception of the world, the agent perceives the current Cartesian position of its arm's end-effector, its own body position along line of cubes, and whether its painting tool contains a charge of paint.

All perceptions are represented as self-stabilized peaks of activation that may be sustained as working memory in absence of the inducing input in fields with strong excitatory kernels.

2) Memories: Memories of cuboid positions and features are stored as slowly decaying memory traces spanned across two-dimensional space/feature space fields. Each visual perception contributes to the build-up of a trace at the activated location and decay of activation at other locations, leading to memory that is subject to interference [16]. The memory trace generates sub-threshold activation in a *space/feature memory* field, which may receive additional inputs in the form of feature cues. Whenever feature cues overlap with the memory trace, a peak emerges in the memory field, recalling the memorized position and feature value of a cuboid.

3) Intentions in Action (IiA): Each action of the robot is specified by a pair of intention and CoS-fields connected to various sub-networks. Reaching and driving to a position are realized by sub-networks simplified from [17], which generate velocity profiles for either the joint angles of the arm or the wheels of the vehicle. The simulated actions to collect and apply pain manifest themselves through a change of the fill status of the painting device. In addition to driving to a specified position, the robot may also explore its environment by moving in a one of the two directions along the line of cuboids, until a previously unattended cuboid is detected. The latter three intentions in action are categorical in nature and are represented by neural nodes.

The actions “visual search”, “recall”, and “activate belief” do not directly induce motor actions, but are aimed at particular perceptual, memory or belief states in which a given feature cue matches visual information, memory, or a learned belief.

4) Prior Intentions: Sequences of actions (or rather, of the entailed intentions in action (IiAs) are represented by intention fields that project excitatory onto all component IiA-fields and onto precondition nodes between them, that implicitly impose a serial order of activation. The IiA that is not inhibited by a precondition node is activated first.

To coat a particular cube with a particular paint, the agent must first collect paint of the specified color from a bucket (tall cuboid) and then apply the paint to a canvas (small cube). Both actions comprise a sequence of actions: Collecting paint comprises locating a tall cuboid, reaching for it, and picking-up the paint. Applying a coat of paint comprises locating the small cube, reaching for it, and dispensing the paint. Locating a cuboid of particular height and color entails the sequence of recalling the cuboid's location, driving to the location, and visually searching for the cuboid within the visual array after reaching the location.

Sequences may also entail alternative action plans, that may be activated once a particular action terminates in failure. This is mediated by activation of its “condition of dissatisfaction” (CoD). For instance, activation of the CoD node of “recall” or “visual search” destabilizes the precondition node of the “explore” intention, which will then activate and move the robot to a new position.

What kinds of paint buckets and canvases are actually sought is determined by the architecture's state of its goals and belief representation.

B. Belief Model

Beliefs differ from memories in that they are combinations of activated concepts rather then perceptual or motor experiences, that would be directly expressed in metric feature spaces such as space, color or height. Any individual belief activates a subset of concepts, but not necessarily the associated feature values. How is such set of concepts formed into a belief

[1]The full architecture containing all parameter values is available for download under: https://www.ini.rub.de/the_institute/people/jan-tekulve/

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

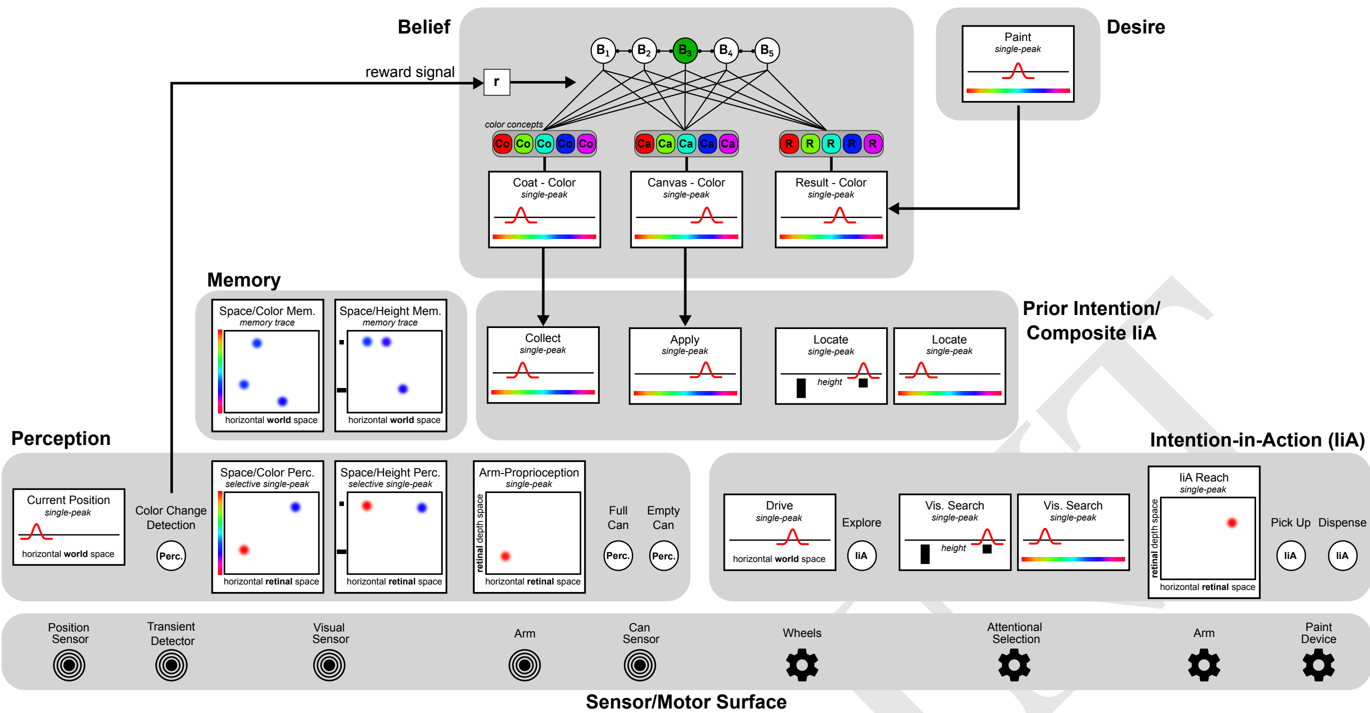

Fig. 3. Schematic overview of the dynamic fields and nodes representing the agent's intentional states grouped according to their psychological modes. For clarity's sake, we show only the most relevant connections to the belief sub-network, which is depicted in more detail in Figure 4.

and how may an individual belief become activated based on matches of its concepts? In Figure 4, we illustrate a neural-dynamic architecture that is inspired by Carpenter's and Grossberg's Adaptive Resonance Theory (ART) [18] and associates a triple of color concepts (paint, canvas and post-coating canvas) represented by concept nodes with a neural node representing a belief. During the execution of a painting action, a concept node for each role (coat, canvas, result) activates autonomously. At the end of the painting sequence, a belief node becomes connected to these nodes through a reward-modulated Hebbian learning rule. Formed beliefs may become activated to paint a particular canvas with a particular paint or to achieve a particular result. They may become deactivated if they do not correctly predict the observed outcome. We step through the belief system illustrated in Figure 4 from bottom up.

1) Belief Formation: The belief learning sub-network is coupled to the rest of the architecture through three working memory color role fields, u_{role}, where role $\in \{\text{coat}, \text{canvas}, \text{result}\}$. The role fields are connected to the color perception field each via a different gating field that becomes activated when pick-up, dispense, and color change detection nodes are on, respectively. In each case, a single peak is formed in a role field that reflects the used or experienced colors during the painting process.

The role fields are reciprocally connected to color-concept nodes, $u^{\text{role}}_{\text{color}}$, that discretize the continuous hue-space in a simple form of abstraction (color $\in \{\text{yellow}, \text{green}, \text{orange}, \text{cyan}, \text{blue}, \text{purple}, \text{pink}, \text{red}\}$, and role as above) and are governed by this neural dynamics

$$\tau \dot{u}^{\text{role}}_{\text{color}} = - u^{\text{role}}_{\text{color}} + h_{\text{con}} + \sum_i l^{\text{role}}_{i,\text{color}} \; \sigma(b_i) + w_{\text{rcl}} \; \sigma(u_{\text{rcl}}) + \int w_{\text{color}}(x) \; \sigma(u_{\text{role}}(x)) dx. \tag{1}$$

Here, $w_{\text{rcl}} \; \sigma(u_{\text{rcl}})$ describes the resting level boost during recall, $w_{\text{color}}(x)$ the connection pattern that describes the particular color representation in the role field, $u_{\text{role}}(x)$, and $l^{\text{role}}_{i,\text{color}}$ the learned connection strength to a belief node b_i. Incoming connection strengths and resting level, h_{con}, are set-up such that a concept node may either be activated through a peak in the role field, u_{role}, or the combination of and activated recall boost, u_{rcl}, and activated associated belief, b_i.

Each *belief* node, b_i, is connected to all concept nodes via the plastic connections, $l^{\text{role}}_{i,color}$, initialized to zero. Belief nodes follows the dynamics:

$$\begin{aligned}\tau \dot{b}_i = & - b_i + h_b + w_b \sigma(b_i) + w_{\text{com}} \sigma(u_{\text{com}}) - w_c \sigma(c_i) \\ & - w_{\text{inh}} \sum_{j \neq i} \sigma(b_j) - w_{\text{cod}} \sigma(u_{\text{cod}}) \\ & + \sum_{\text{role}} \sum_k \left[l^{\text{role}}_{i,k} \sigma(u^{\text{role}}_k) - w_{\text{inc}} \sigma(u^{\text{role}}_k) \right]\end{aligned} \tag{2}$$

where w_b describes the self-excitation, w_{inh} the inhibitory connections from other belief nodes, w_{com} the input from the commit node, w_c the inhibition from the corresponding commit state node, c_i, and w_{cod} the inhibition from the condition of dissatisfaction node. The connection pattern between the

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

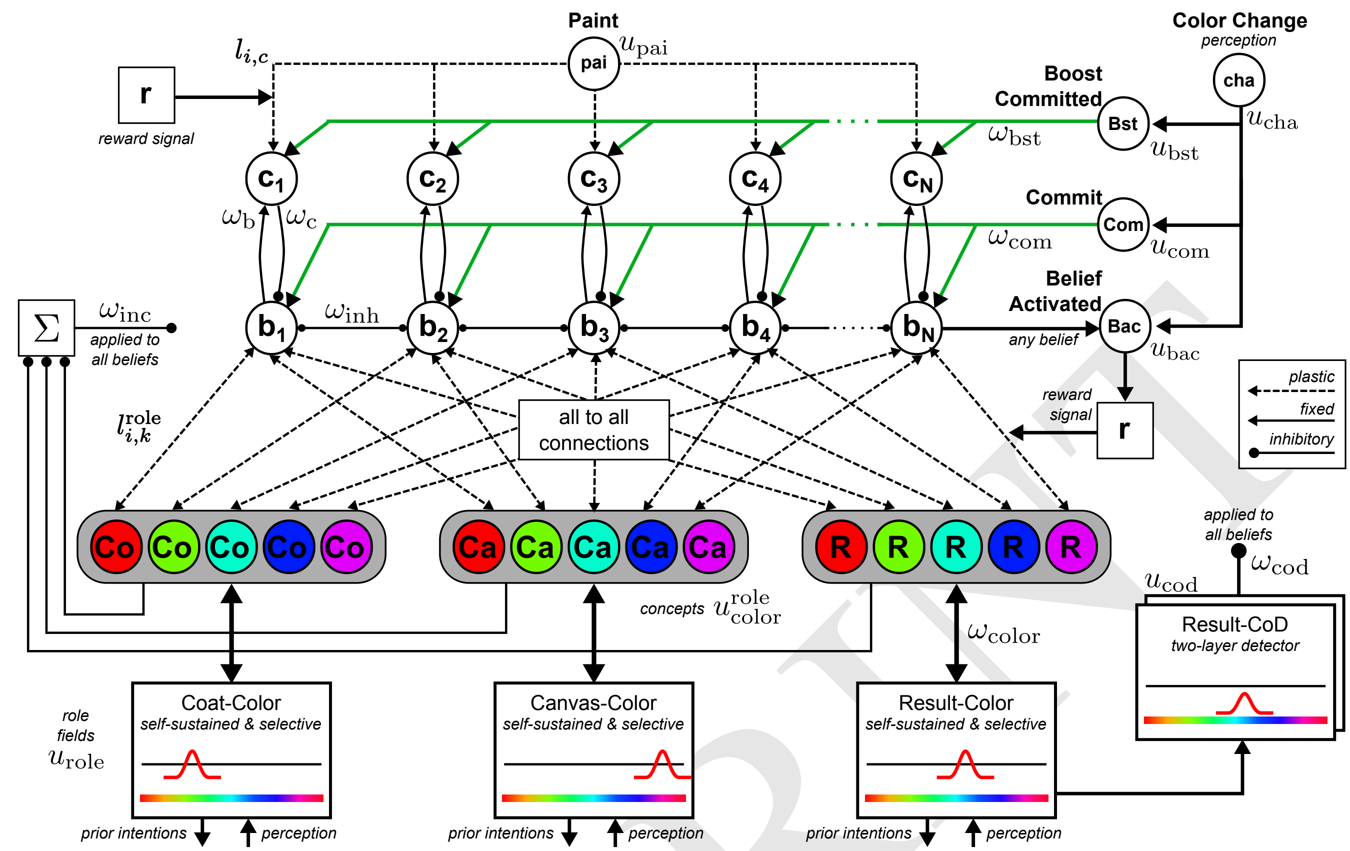

Fig. 4. The sub-network responsible for belief formation, activation and rejection.

concept nodes and the belief nodes is set up such that supra-threshold activation in any of the concept nodes, u_k^{role}, leads to an inhibition in all belief nodes with fixed strength, w_{inc}. Any node, b_i, that has already learned a connection, $l_{i,k}^{\mathrm{role}} \neq 0$, receives excitatory input that cancels out the inhibitory effect, w_{inc}, of that concept node. Thus, when n concept nodes are active, only a belief node that matches all n concepts can become activated.

An change of color in the scene leads to activation of the change detection node, u_{cha}. If a belief node, b_i, is active at the same time, then the node, u_{bac} (*belief activated*), is pushed through the detection instability and a transient reward signal, $r(t)$, is generated (see [19] for the details of that). This strengthens the connections, $l_{i,k}^{\mathrm{role}}$, according to the following reward based Hebbian learning rule:

$$\dot{l}_{i,k}^{\mathrm{role}} = -\eta\; r(t)\; \sigma(b_i)\; (l_{i,k}^{\mathrm{role}} - \sigma(u_k^{\mathrm{role}})) \tag{3}$$

where the learning rate, η, is sufficient to learn a perceived rule in a single presentation (one-shot learning) (see [20] for a review of neural dynamic learning rules). The stability of the learned state, $l_{i,k}^{\mathrm{role}} = \sigma(u_k^{\mathrm{role}})$, makes that later presentations of the same belief cause little to no weight change.

If no belief node is active at the time at which a color change is detected, then the current set of concepts nodes has not previously been learned. The processes of selecting a new belief to learning this belief is then initiated by activation of the *commit*, u_{com}, and the *boost committed*, u_{bst}, nodes, both activated through u_{cha}. The commit node, u_{com}, boosts all belief nodes. Lateral inhibition among belief nodes result in the activation of a single belief node for that purpose. Only previously uncommitted belief nodes may join the competition.

This is implemented by the *commit state nodes*, c_i, representing commitment of the belief node, b_i, to a concept configuration. The boost committed node u_{bst} activates all commit state nodes, c_i, with non-zero weights. This happens before the belief nodes, because the dynamics of u_{bst} evolves on a faster timescale than the dynamics of u_{com}. The outgoing inhibition of b_i from c_i is strong enough to prevent activation of b_i when b_i is below threshold. If b_i is already active, however, inhibition from c_i does not induce a reverse detection instability, and does not, therefore, deactivate b_i. The commit state nodes, c_i, are governed by the following dynamics:

$$\tau\dot{c}_i = -\,c_i + h_c + w_b\sigma(b_i) + w_{\mathrm{bst}}\sigma(u_{\mathrm{bst}}) + l_{i,c}\sigma(u_{\mathrm{pai}}), \tag{4}$$

where the resting level, h_c, is set up such that at least b_i or the combination of u_{bst} and u_{pai} need to be active to push c_i through the detection instability.

Commitment is encoded in plastic connections, $l_{i,c}$, from the paint task node, u_{pai}, to each of the c_i nodes based on a

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

similar reward-modulated Hebbian rule as in Eq. 3:

$$\dot{l}_{i,c} = -\eta \; r(t) \; \sigma(u_{\text{pai}}) \; (l_{i,c} - \sigma(c_i)). \tag{5}$$

2) Belief Activation: The learned reciprocal connections between concept and belief nodes allow for the activation of beliefs through a (partial) concept node cue. For example, if input from a desire node activates a particular result color concept node, $u_{k_0}^{\text{result}}$, this inhibits all belief nodes through w_{inc}, and simultaneously excites all nodes, b_i, with a previously learned connection to that result color concept, $l_{i,k_0}^{\text{result}} \neq 0$. Of these, only one becomes activated due to lateral inhibition among beliefs. A homogeneous boost, u_{rcl}, across all concepts leads to activation of coat and canvas concept nodes defined through the connections, l_{i,k_1}^{coat} and $l_{i,k_2}^{\text{canvas}}$, of the activated belief. Activation of these concept nodes leads to the formation of self-sustained working memory peaks in the role fields, which are projected onto the intention fields of "collect" and "apply".

3) Belief Rejection: If environmental conditions change, beliefs might become outdated and an update of an already established belief might be necessary. In the toy scenario, this may be because a color mixing rule changes such that coat color, a, and canvas color, b, which previously resulted in color, c, now result in color, d, instead. A belief update would occur if the agent tried to produce color, c, based on the belief a, b, c, but instead observed color, d.

During the execution of the paint sequence, the color role fields contain peaks at the locations a, b, and c based on the active belief. Once a color change is detected through the transient-detector, the observed new color, d, leads to the formation of a new peak of activation in the result role field, which will override the old representation of c due to strong lateral inhibition. The activation pattern change in the role field is detected through a transient detector that activates the CoD node, u_{cod}, which inhibits all beliefs via w_{cod} (see [21] for details on the transient detector). This leads to the destabilization of the active belief and the activation of a new uncommitted belief node if no previously committed belief matches the newly, active concept node pattern. Once a belief is activated, the reward signal is generated and attributes the new pattern to the active belief node.

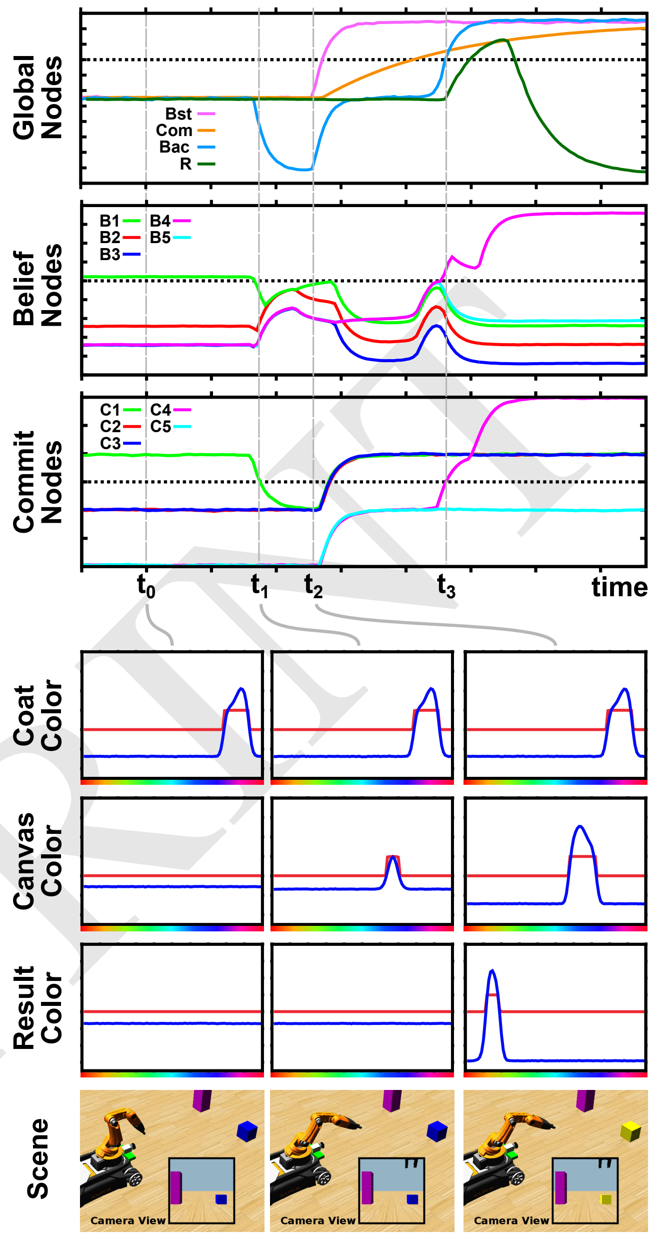

Fig. 5. Activation time course of selected nodes during the formation of a new belief. The first row displays activation of u_{bst}, u_{com}, and u_{bac}, while the second and third rows display the activation of five selected belief nodes, b_i, and their commit state nodes, c_i. The bottom half of the figure shows activation snapshots of the role fields at three different points in time.

IV. Results

We demonstrate how activation of the belief sub-network develops over time during belief formation, belief activation, and belief rejection.

A. Belief Formation

Figure 5 depicts an activation time course of the belief sub-network during the final portion of a painting sequence.

At the point, t_0, in time, purple color has already been collected from a tall, purple cuboid, and that caused the formation of a working memory peak in the coat color role field. This leads to activation of the previously learned belief, $B1$ (Coat:Purple, Canvas:Purple, Result:Yellow) that matches the coat color. The corresponding commit node, $C1$, is also activated.

Once the agent begins dispensing paint onto the blue cube at point, t_1, a working memory peak emerges in the canvas color role field. The blue canvas color representation causes a destabilization of the non-matching belief, $B1$, which will in turn deactivate $C1$ and lower the activation level of the belief activated node, u_{bac}.

At point, t_2, the cube changes its color from blue to yellow leading to an activation of the color change node, u_{cha}, that boosts three nodes: "boost committed", u_{bst}, "commit", u_{com},

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

and "belief activated" u_{bac}. The observed changed color forms a peak in the the result color role field, which excites matching beliefs, such as $B1$. No belief matches all three color roles, however. Activation of u_{bst} activates the commit nodes of all previously committed beliefs lowering their activation level.

Once the slower u_{com} passes the detection threshold at t_3, all beliefs receive a boost, which activates the previously uncommitted belief, $B4$. This activates u_{bac}, and generates a transient reward signal that connects the represented color roles with the belief, $B4$. The weights of $B1$ and $B4$ after t_3 are depicted in Figure 6.

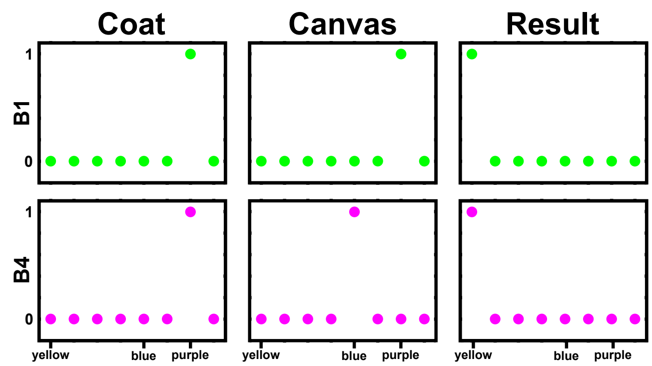

Fig. 6. Connections $l_{i,k}^{\mathrm{role}}$ of $B1$ and $B4$ after the learning episode depicted in Figure 5.

B. Belief Activation

The learned connections are utilized in belief recall, activating beliefs that match a role cue. In Figure 7, a peak representing yellow has emerged in the result color role field due to the agent's desire to paint a yellow cube. This leads to an increase in activation for all beliefs that match the yellow result color ($B1$ and $B4$) and a decrease for all non-matching beliefs. Belief, $B1$, passes the activation threshold first by chance, leading to inhibition of all other belief nodes and activation of the connected concept nodes. This causes the emergence of peaks in the connected coat and canvas role fields, which guide the subsequent painting sequence.

C. Belief Rejection

Once a belief has been activated by recall, it may be acted upon during the painting task. In Figure 8, the evolution of activation of different belief nodes is shown, while the agent acts upon belief, $B2$, which should result in a blue-colored cuboid. However, the properties of the simulated world were changed such that the combination of coat and canvas color used in $B2$ will now result in the color cyan.

At point t_0, the agent is engaging in the dispense color action while belief, $B2$, has been recalled and is active, which leads to the formation of a peak representing the anticipated result color "blue". At t_1, the change to a cyan color occurs in the scene and the observed change color is projected into the result color role field. Due to the stronger connection strength from the perceptual fields and selective inhibitory coupling, the peak at the cyan position destabilizes the blue peak. The

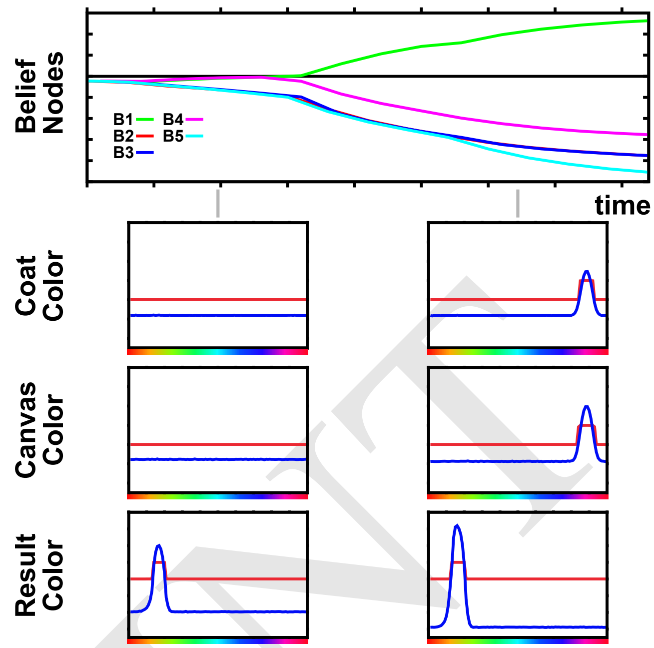

Fig. 7. Recall of a belief with a yellow result color. The learned connections are the same as shown in Figure 6

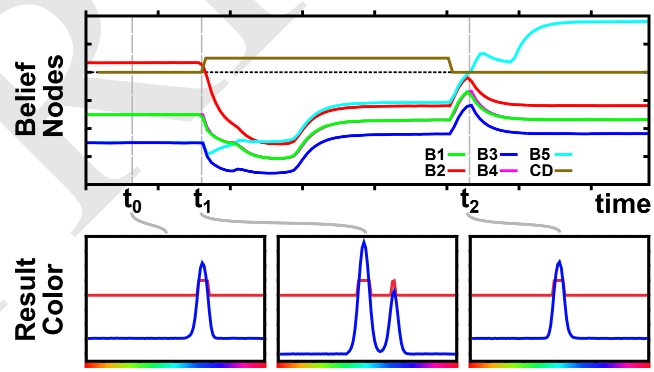

Fig. 8. Belief rejection: Belief, $B2$, predicts result color "blue", but instead "cyan" was observed. The detected change in the result field leads to the deactivation of $B2$, and the eventual commitment of $B5$.

color change in the result field is detected by the CoD transient detector, which, in turn, inhibits all beliefs for a brief period of time, counteracting the excitatory boost from the commit node, u_{com}.

Once inhibition from the CoD vanishes at point, t_2, the activation boost from u_{com} is strong enough to activate the previously uncommitted belief, $B5$, which can then be associated with the new color mixing rule.

V. Discussion

We have presented a network of neural dynamic fields that endows a robotic agent with the capability to form, activate, and reject beliefs in a simulated task environment. During belief learning, activated concept nodes become associated with a neural-dynamic belief node through a reward-modulated Hebbian learning rule. Activating beliefs is achieved by a

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019

neural match operation that is similar to the resonance principle of ART [18], combining the learned reciprocal connections with homogeneous global inhibition. Because the learned associations reside at the level of concepts, they afford generalization of the learned contingencies to slightly different environments. The rejection of a candidate belief occurs autonomously through the neural representation of a condition of dissatisfaction (CoD). That representation is triggered when a mismatch between perceived and predicted sensory representations of concepts is detected. The CoD inhibits the candidate belief and thus frees neural support for an alternative belief from inhibition which then proceeds to learn the new belief.

The neural-dynamic belief system is embedded in a larger network of neural dynamic fields that controls a robotic agent. That network generates stable representations of intentional states of four elementary psychological modes (perception, memory, intention-in-action, and prior intention). Transitions between intentional states occur through instabilities induced by the neural CoS or CoD representations. It is from such transitions between different intentional states that the agent's behavior emerges. The stable representation of actions and perceptions support working memory that provides an interface to the belief network. Using beliefs in action and learning beliefs both are insensitive to the duration of a behavioral episode, providing robustness.

The complete architecture was demonstrated in a simple toy scenario, in which an agent explores different color mixing combinations, acquires beliefs about color mixing rules, activates these beliefs to achieve desired resulting colors, and rejects beliefs if the resulting color do not match prediction. The scenario was minimalistic to promote conceptual clarity, the "concepts" certainly being trivial, their power of generalization unimpressive, and the possible actions being trivialized. Quantitative evaluation of the learning process may not be worthwhile at this stage.

Our emphasis was primarily to demonstrate that appropriate process infrastructure enables fast and robust belief learning. It is plausible that neural infrastructure of this kind exists. For instance, a model of cortical and basal-ganglionic processes for learning serially ordered behaviors has similar prior structure [22]. Schrodt and Butz [23] have explored rule learning in a scenario similar to ours and argue for its neural plausibility. We believe that the problem of autonomously learning beliefs, rules, or contingencies from experience is best framed as the problem of how the underlying architectures of neural processes are structured (rather than as a problem of finding special learning rules).

Apart from the obvious scaling question, there are many fronts on which we would like to see this work expanded. An important one is to address more complex desires, perhaps even abstract ones like the desire to learn new things about the world [24].

REFERENCES

[1] B. F. Skinner, ""Superstition" in the pigeon," *Journal of Experimental Psychology*, vol. 38, pp. 168–272, 1948.

[2] N. J. Nilsson, *Understanding Beliefs*. The MIT Press, 2014.

[3] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories," *Computer Vision and Image Understanding*, vol. 106, no. 1, pp. 59–70, 2007.

[4] O. Lomp, C. Faubel, and G. Schöner, "A Neural-Dynamic Architecture for Concurrent Estimation of Object Pose and Identity," *Frontiers in Neurorobotics*, vol. 11, no. April, p. 23, 2017.

[5] J. B. Northrup, "Contingency detection in a complex world," *International Journal of Behavioral Development*, vol. 41, no. 6, pp. 723–734, 2017.

[6] J. R. Searle, "The intentionality of intention and action*," *Cognitive Science*, vol. 4, no. 1, pp. 47–70, 1980.

[7] J. Tekülve and G. Schöner, "Neural dynamic concepts for intentional systems," in *41th Annual Conference of the Cognitive Science Society (CogSci 2019)*, 2019.

[8] R. Kozma, "Intentional systems: Review of neurodynamics, modeling, and robotics implementation," *Physics of Life Reviews*, vol. 5, no. 1, pp. 1–21, 2008.

[9] G. Schöner, J. P. Spencer, and T. DFT Research Group, *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press, 2015.

[10] Y. Sandamirskaya and G. Schöner, "An embodied account of serial order: How instabilities drive sequence generation," *Neural Networks*, vol. 23, no. 10, pp. 1164–1179, 2010.

[11] K. Thórisson and H. Helgasson, "Cognitive Architectures and Autonomy: A Comparative Review," *Journal of Artificial General Intelligence*, vol. 3, no. 2, pp. 1–30, 2012.

[12] M. Richter, J. Lins, and G. Schöner, "A Neural Dynamic Model Generates Descriptions of Object-Oriented Actions," *Topics in Cognitive Science*, vol. 9, pp. 35–47, 2017.

[13] S. K. U. Zibner and C. Faubel, "Dynamic scene representations and autonomous robotics," in *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press, 2015, pp. 227–246.

[14] M. Richter, Y. Sandamirskaya, and G. Schöner, "A robotic architecture for action selection and behavioral organization inspired by human cognition," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 2457–2464.

[15] R. Grieben, J. Tekülve, S. K. U. Zibner, S. Schneegans, and G. Schöner, "Sequences of discrete attentional shifts emerge from a neural dynamic architecture for conjunctive visual search that operates in continuous time," in *CogSci 2018*, J. Z. Chuck Kalish, Martina Rau and T. Rogers, Eds., July 2018, pp. 427–432.

[16] W. Erlhagen and G. Schöner, "Dynamic field theory of movement preparation." *Psychological Review*, vol. 109, no. 3, pp. 545–572, 2002.

[17] G. Schöner, J. Tekülve, and S. Zibner, "Reaching for objects: A neural process account in a developmental perspective," in *Reach-to-grasp Behavior: Brain, Behavior, and Modelling Across the Life Span*. Routledge, 2018, pp. 297–334.

[18] G. A. Carpenter and S. Grossberg, *Adaptive resonance theory*. Springer, 2016.

[19] J. Tekülve, S. K. Zibner, and G. Schöner, "A neural process model of learning to sequentially organize and activate pre-reaches," in *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2016, pp. 318–325.

[20] Y. Sandamirskaya, "Dynamic neural fields as a step toward cognitive neuromorphic architectures," *Frontiers in Neuroscience*, vol. 7, no. 276, pp. 1–13, 2014.

[21] M. Berger, C. Faubel, J. Norman, H. Hock, and G. Schöner, "The counter-change model of motion perception: an account based on dynamic field theory," in *International Conference on Artificial Neural Networks*. Springer, 2012, pp. 579–586.

[22] D. G. Beiser and J. C. Houk, "Model of cortical-basal ganglionic processing: encoding the serial order of sensory events," *Journal of Neurophysiology*, vol. 79, no. 6, pp. 3168–3188, 1998.

[23] F. Schrodt, J. Kneissler, S. Ehrenfeld, and M. V. Butz, "Mario becomes cognitive," *Topics in cognitive science*, vol. 9, no. 2, pp. 343–373, 2017.

[24] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE transactions on evolutionary computation*, vol. 11, no. 2, pp. 265–286, 2007.

Published as: Jan Tekülve and Gregor Schöner. Autonomously learning beliefs is facilitated by a neural dynamic network driving an intentional agent. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2019 Joint IEEE International Conferences on*, pages 143–150, 2019