Biologically Inspired Cognitive Architectures (2015) 13, 91–104

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bica

INVITED ARTICLE

NARLE: Neurocognitive architecture for the autonomous task recognition, learning, and execution

Yulia Sandamirskaya [a,*], Mikhail Burtsev [b]

[a] Institute of Neuroinformatics, University of Zurich and ETH Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[b] NBICS Centre, NRC ''Kurchatov Institute'', Kurchatov sq. 1, 123182 Moscow, Russian Federation

Received 19 May 2015; received in revised form 29 June 2015; accepted 30 June 2015

KEYWORDS
Neurocognitive architecture;
Dynamic Neural Fields;
Functional System Network;
Pervasive learning

Abstract
Robots controlled by the state of the art cognitive architectures are still far behind animals in their capabilities to learn complex skills and autonomously adapt to unexpected circumstances. The neurocognitive architecture proposed in this paper addresses the problem of learning and execution of hierarchical behaviors and complex skills. Learning is addressed both on the level of individual elementary behaviors and goal-directed sequences of actions. The proposed architecture comprises a Dynamic Neural Fields (DNFs) implementation of the low-level elementary behaviors and a Functional System Network (FSN) tying these behaviors in goal-directed sequences. The DNF framework enables a continuous, dynamical representation of perceptual features and motor parameters, which may be directly coupled to the robot's sensors and motors. Attractor states and instabilities of the DNFs account for segregation of cognitive states and mark behaviorally relevant events in the continuous flow of sensorimotor dynamics. The FSN, in its turn, comprises dynamical elements that can be arranged in a multilayered network by a learning process, in which new layers and elementary behaviors are added on demand. In our architecture, the FSN controls adaptation processes in the already acquired neural-dynamic elementary behaviors, as well as formation of new elementary behaviors. Combination of the DNF and FSN frameworks in a neurocognitive architecture NARLE enables pervasive learning both on the level of individual behaviors and goal-directed sequence, contributing to the progress towards more adaptive intelligent robotic systems, capable to learn new tasks and extend their behavioral repertoire in stochastic real-world environments.

* Corresponding author.
E-mail addresses: ysandamirskaya@ini.uzh.ch (Y. Sandamirskaya), burtsev.m@gmail.com (M. Burtsev).

http://dx.doi.org/10.1016/j.bica.2015.06.007

Introduction

Behavior of humans and higher animals is characterized by the ability to solve novel problems and to reuse acquired solutions in the future. This capacity for constant adaptation results in a continual refinement of skills and extension of the behavioral repertory. State of the art autonomous robots, to the contrary, only execute what they were programmed to do and thus accomplish only a limited number of predefined tasks in predefined environments. In an extreme example, an autonomous robot on the factory floor assembles predefined pieces picked up from predefined locations in a predefined order. Extending the use of robots into less well modeled environments, shared with naive human users, such as streets, homes, or hospitals, is facilitated if robots are adaptive and may autonomously acquire new behaviors. Indeed, such environments contain contingencies that are difficult to describe completely and specify in advance. It is similarly difficult to pre-train the perceptual system of a robot to recognize all objects that may be encountered in all possible configurations. On the motor side, new objects may require adaptation of the available skills, e.g. new movement parameters may be needed to insert a tool into a particular opening. The capacity to acquire new skills by combining available actions into meaningful sequences may help to accomplish missions under changed conditions. Clearly, new methods in robotics are needed to enable robots to autonomously learn on-site and deal with unexpected situations.

The classical approaches to autonomous robot action planning, such as dynamic programming, rely on *a priori* knowledge about the environment (Ingrand & Ghallab, 2014). This limits their effectiveness in real-world environments, which may change unexpectedly. Taking inspiration from biology, *behavior-based robotics*, starting with an early work of Rodney Brooks (Brooks, 1986), emphasized the idea that to function in a real-world situation, the robot needs a reactive controller. Such controller adjusts robot's behavior in accordance with the current environmental situation. This led to a modular, distributed, dynamical, and flexible behavior-based subsumption architecture, which was able to demonstrate complex behavior in a real-world environment. Behavior-based approaches range from purely reactive architectures (Maes, 1990) to methods integrating world knowledge and planning with the low-level control (Nicolescu & Mataric, 2003; Pirjanian, 1999). Since the stateless reactive approaches do not scale to more robust and complex behaviors (Bryson, 2003) hybrid architectures were developed that combine the strengths of the behavior-based and deliberative approaches (e.g. Bryson, 2003; Proetzsch, Luksch, & Berns, 2010). The behavioral modules in this framework may be subject to learning (Rylatt, Czarnecki, & Routen, 1998), although typically learning is not the main focus of behavior-based robotics. In a navigation scenario, reinforcement learning has been combined with principles of behavior-based control to enable autonomous learning of sequential behavior (Konidaris, 2005).

Developmental robotics (Cangelosi, Schlesinger, & Smith, 2015) offers an approach to making robot controllers flexible and adaptive. This approach exploits autonomous development and learning motivated by developmental theories of human cognition. Guided by an understanding of how biological systems interact with their environment during behavior and learning, developmental robotics achieved recently a considerable progress in implementing systems that learn to perceive and generate actions without the need to fully pre-program the involved sensory representations and motor controllers (Jamone et al., 2014; Saegusa, Metta, Sandini, & Natale, 2014). Such architectures may learn consequences of robot's actions (Natale, Rao, & Sandini, 2002), or learn to optimally process sensory inputs (Saeb, Weber, & Triesch, 2011). Exploration, required for learning, is achieved by different flavors of motor babbling, when exploratory motor commands are generated internally, or by demonstration. One of the promising approaches in this vain (Baldassarre & Mirolli, 2013) advances the idea that an intrinsic motivation and goal formation can be adopted for the control of explorative learning of skills.

In this work, we aim to go beyond adaptation of single skills, currently prevalent in developmental robotics, towards learning sequences of actions leading to a goal, while at the same time refining the elementary behaviors. Learning action sequences can be achieved in a robotic system by supervised, unsupervised, and reinforcement learning techniques. One of the most popular approaches here is *reinforcement learning* (RL) (Sutton & Barto, 1998) and its modifications. Traditionally, RL algorithms are used for problems with a single objective. However, robotic applications require hierarchical organization of behavior, where simple actions fit together in coherent skills and these combine to achieve a higher-level goal. Classical RL (Sutton & Barto, 1998) does not align with hierarchical structure of behavior, but there is an ongoing development to resolve this issue. Thus, using temporal abstraction, atomic sequences of state transitions, which correspond to different goals and skills, can be represented as *options* (Sutton, Precup, & Singh, 1999). An option generalizes primitive actions to include temporally extended courses of action. Options are defined by the initiation and termination sets of states, as well as by their policy. In the options framework, the *temporal-difference (TD) networks* are used for a general compositional specification of the goals of learning (Sutton, Rafols, & Koop, 2006).

When a sequence of actions is abstracted in the form of an option, the option itself can be used as an atomic unit for the composition of behavioral strategy at the higher level. This possibility is explored in the domain of *hierarchical RL* (Barto & Mahadevan, 2003; Botvinick, Niv, & Barto, 2009; Vigorito & Barto, 2010). Hierarchical RL is intended to produce gradual accumulation of behavioral modules that can be combined and reused for different tasks. The major challenge for the effective application of the hierarchical RL framework in robotics is the lack of a learning mechanisms, which could generate options, instead of options being explicitly coded by a designer. Other roadblocks on the way to applying RL in robotics were reviewed in detail recently (Kober, Bagnell, & Peters, 2013).

To progress in solving the challenge of *pervasive learning* both on the level of sequences and individual behaviors, we propose to implement principles of autonomous learning and adaptation, elaborated in the field of developmental

robotics, using the mathematical framework of neural dynamics, leading to the *Neurocognitive Architecture for the Autonomous Task Recognition, Learning and Execution* (NARLE). In this framework, the whole control architecture — from the low-level perception and motor control, to elementary behaviors and planning — will be formulated with attractor dynamics that are subject to adaptation and learning. The transparency and homogeneity of the architecture will allow us to side-step problems of integration of heterogeneous software modules, which many modern robotic systems face (Asfour et al., 2006; Krüger et al., 2011; Tenorth & Beetz, 2013). To this end, we aim to employ the mathematical framework of Dynamic Neural Fields (DNFs) (Sandamirskaya, 2013; Schöner, 2008) for implementation of a modular architecture inspired by work in behavioral robotics, supplemented with neural-dynamic learning and adaptation mechanisms. The resulting architecture will be integrated with *Functional Systems Network* (FSN) framework (Komarov, Osipov, & Burtsev, 2010) for learning of complex goal-directed sequences of behaviors. New behaviors may be added to the neural-dynamic architecture if no existing behavior is activated in a current situation. This mechanism is similar to the process of ''detection of unknown'', recently explored in relation to learning in robots and biological systems (Kushiro, Harada, & Takeno, 2013). The two levels of the NARLE architecture — the low-level sensorimotor level, built with continuous activation functions of DNFs and the hierarchical system to represent action sequences based on FSN — are tightly integrated, but nevertheless effectively segregate the system into a continuous, implicit representation of actions, which is linked to sensors and motors of the robotic agent, and a discrete, explicit representation of actions, in which planning of action sequence occurs. This segregation is similar to separation of procedural and declarative knowledge in human cognition and learning (Sun, Merrill, & Peterson, 1999; Sun, Mathews, & Lane, 2007). Our architecture uses the same principles of neural population coding as previous cognitive architectures for planning and decision making (e.g.Cisek, 2006). Coupling of these neural-dynamic mechanisms to the sensory-motor system of the agent is more detailed in our framework, which makes the architecture embodied and situated and enables learning based on the agent's experience.

The computational framework of DNFs has been successfully applied in several behavioral scenarios and demonstrated a high potential for implementation in control architectures that may function and learn naturally in a human-centered environment (Bicho, Erlhagen, Louro, & Costa e Silva, 2011; Sandamirskaya, Zibner, Schneegans, & Schöner, 2013; Zibner, Faubel, Iossifidis, & Schöner, 2011). In particular, DNF architectures are (a) *robust*, i.e. capable to cope with uncertainties in sensory inputs and outcome of motor commands; (b) *adaptive*, i.e. capable to adjust to unforeseen changes in the environment and the body of the agent; and (c) *computationally efficient*, i.e. can cope with continuous parameter spaces of real-world environments. FSN is developed for goal-directed exploratory learning (Komarov et al., 2010) and features: (a) unsupervised goal-directed learning of action sequences; (b) solving multi-goal tasks in stochastic environments; (c) using alternative means for achievement of the same goal; (d) gradual accumulation of experience without forgetting. The FSN has been applied so far on abstract, discrete learning problems. DNFs will provide the low-level embodiment structure, enabling goal-directed sequence learning with FSN in a real-world environment, perceived and acted on by the robotic agent.

In this paper, we present the basics of the DNF and FSN frameworks and detail how they may be used to implement the NARLE system for a specific scenario, in which a humanoid robot NAO learns to assemble complex objects from colored blocks, distributed on the table-top. In particular, the systems for object and scene representation, action parsing, and sequence learning are detailed and their interplay in the envisioned task is described. Finally, the potential of this new biologically-inspired cognitive architecture is discussed.

Methodological background

Dynamic Neural Fields (DNFs)

Dynamic Neural Fields (DNFs) are mathematical models of neuronal activation, derived originally to understand formation of activity patterns in the brain (Amari, 1977; Grossberg, 1988; Wilson & Cowan, 1973). Later, DNFs were developed into a computational and conceptual framework to account for cognitive processes and development of cognitive functions in humans (Johnson, Spencer, & Schöner, 2008; Schöner, 2008), and recently also to control artificial cognitive systems, capable of adaptive behavior (Sandamirskaya, 2013; Sandamirskaya et al., 2013). These continuous in time and in space attractor dynamics allow to couple the controlling architecture to real physical sensors and motors, while providing an interface to discrete, or ''symbolic'', cognitive representations. The dynamics of a DNF is described by an integro-differential equation:

$$\tau\dot{u}(x,t) = -u(x,t) + h + \int f(u(x',t))\omega(|x'-x|)dx' + I(x,t). \tag{1}$$

In Eq. (1), $u(x,t)$ is the activation of the DNF at time t; x is a dimension that spans a behavioral parameter (e.g., color, location on the retina, postural state, or velocity); τ is the relaxation time-constant; h is the negative resting level; $f(\cdot)$ is the sigmoidal non-linearity shaping the output of the DNF; the lateral connections in the DNF are shaped by a Mexican hat lateral interaction kernel, $\omega(|x'-x|)$, with a short-range excitation and a long-range inhibition parts; $I(x,t)$ is the sum of the external inputs to the DNF. The lateral interactions bring about the existence and stability of a localized-peak solution that is the computational basis for cognitive processes of categorization, detection, selection, and memory in the DNF framework.

DNFs may be defined over spaces of different dimensionality: from a zero-dimensional, discrete dynamical node to a three-dimensional color-space DNF which describes, e.g. activation of neurons, sensitive to color and retinal space. DNFs of different dimensionality may be coupled through the external-input term, $I(x,t)$ in Eq. (1): the output of a

lower-dimensional field is expanded (copied along the not-shared dimension(s)) to be input to a higher-dimensional field, whereas output of a higher-dimensional field is ''projected'' (through summation or max-operation along the not-shared dimension(s)) to be input to a lower-dimensional DNF. The coupling strength between DNFs are subject to a simple, neurally-based (''Hebbian'') learning rule (Sandamirskaya, 2013):

$$\tau_l \dot{W}(x,y,t) = \epsilon(t)(-W(x,y,t) + f(u_1(x,t)) \times f(u_2(y,t))). \tag{2}$$

Here, $\epsilon(t)$ is a time-dependent learning signal, which marks a learning window and is controlled by a behavioral organization system (FSN here). $W(x,y,t)$ are coupling weights between DNFs $u_1(x,t)$ and $u_2(y,t)$, τ_l is the learning time-constant, $\times$ denotes the Kronnecker product (external product), so that each position (x,y) in the W function corresponds to a connection between the respective positions in the coupled DNFs. A version of the learning rule Eq. (2), in which only a single DNF is involved and the coupling weights form a layer, which feeds-back to this DNF, is known as memory trace, or preshape dynamics (Bastian, Schöner, & Riehle, 2003; Erlhagen & Schöner, 2002; Sandamirskaya, 2013; Wilimzig & Schöner, 2005) and has been used in the Dynamic Field Theory of cognitive development to model motor and perceptual long-term memory formation.

Functional Systems Network (FSN)

FSN architecture is inspired by the *Functional systems theory* proposed in 1935 by the soviet neurobiologist P.K. Anokhin (Anokhin, 1974). This theory considers how distributed morphological elements of an organism cooperate to achieve evolutionary adaptive results in the environment. Important conclusion of the theory is that an organism should have a self-organizing mechanism to coordinate diverse elements toward vital purposes. Such set of coordinated elements constitutes a system that has a function to achieve a goal, and a representation of the goal is key for *the functional system* (FS) integration.

A unit of the functional systems network is a functional system (FS). FS implements the following functionality: (1) recognition of the problem (context) input and subsequent activation of the output (i.e. execution of an action); (2) recognition of the goal input and subsequent deactivation of the output (i.e. termination of the action); (3) timing the period of the action execution and deactivation of the output if this period exceeds expected time for the action completion. Such set of coordinated elements constitutes a system that has a function to achieve a goal, and a representation of the goal is key for the functional system (FS) integration. A number of computational architectures utilizing functional systems theory were proposed and studied (Red'ko, Prokhorov, & Burtsev, 2004; Redko et al., 2007; Komarov et al., 2010; Lakhman, & Burtsev, 2014; Shirshova, & Burtsev, 2014; Vityaev, 2015). In NARLE we use a functional systems network model proposed in (Komarov et al., 2010).

This functionality can be implemented by the micro-circuit of three formal neurons (Fig. 1). The first ''problem'' neuron N^p detects situations when the action should be executed. The second ''action'' neuron N^a is activated by the ''problem'' neuron and has two self-referent collateral connections. The first collateral connection has a positive weight w^a and sustains activity of the ''action'' neuron N^a once it was kick-started by the ''problem'' neuron N^p. The second collateral connection has a negative weight w^f and its effect is delayed by the time τ expected for the successful action completion. Thus, the negative feedback shuts down N^a if it is active longer than usual. The third, ''goal'' neuron N^g detects situations when the action succeeded, i.e. the target state is detected. There is a connection with a negative weight w^g from N^g to N^a that turns off the latter. The micro-circuit, therefore, has two sets of inputs: activating with weights w^p (''problem'') and deactivating with weights w^g (''goal''), and one output U^a (''action'').

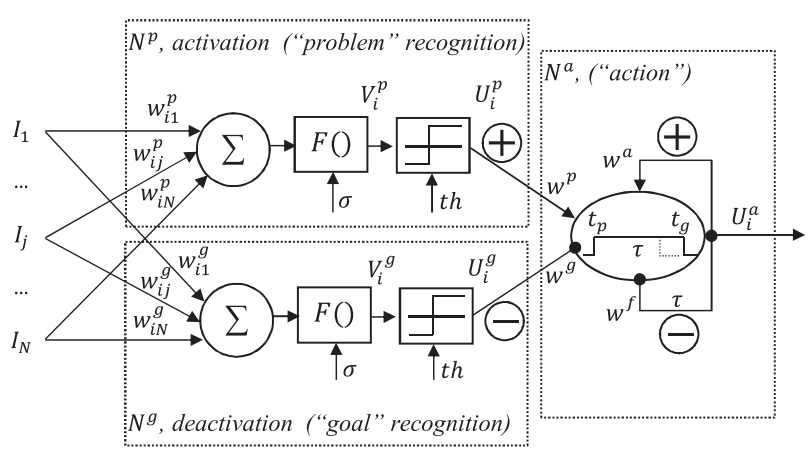

Fig. 1 A unit of the functional systems network can be implemented by a microcircuit of formal neurons. The neuron N^p recognizes contexts where activity of FS can contribute to the goal-directed behavior. This neuron activates action output neuron N^a, which stays active until target context is detected by the neuron N^g. Activation of N^g signals that the action was successful and inhibits output N^a. If a period of N^a activity exceeds expected time of the action completion τ, N^a is turned off and a ''failure'' signal is generated.

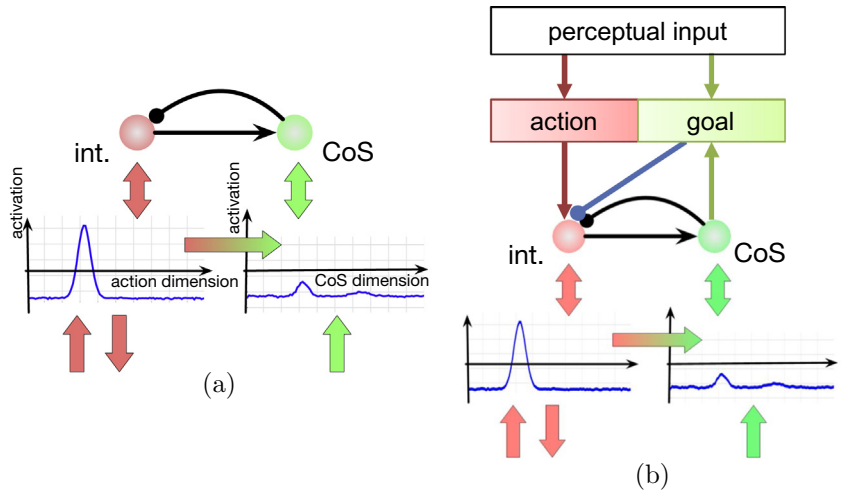

Fig. 2 (a) A simple elementary behavior (EB) with one 1D intention and 1D condition of satisfaction (CoS) DNFs. The neural-dynamic nodes (red and green circles on top of the DNFs) signal activation of the respective DNF. The CoS node inhibits the intention node when the behavior is complete. (b) Coupling of DNFs with the functional system (FS) for control of the elementary behavior. Functional system recognizes the state of the task and activates the elementary behavior that was associated with this state earlier during the learning. Output of the FS activates the intention node of the corresponding EB and then supervises EB's execution by monitoring timing of the behavior via connection from CoS node to the goal input of the FS. If the outcome of EB was not achieved in an expected time-frame, the FS inhibits the EB and initiates switching to an alternative EB or exploratory learning.

Output for every neuron in the FS microcircuit is calculated as a nonlinear transformation of a weighted sum of inputs:

$$V_i = F\left(\sum_{j\in I} w_{ij} I_j\right) + \sigma, \tag{3}$$

$$U_i = \begin{cases} V_i & : V_i > \text{th} \\ 0 & : V_i \leqslant \text{th} \end{cases} \tag{4}$$

In Eqs. (3) and (4), w_{ij} is the weight, $F(x)$ is a nonlinear function (typically, a sigmoid), σ is a random value, U_i is output of the ith neuron.

The twofold structure of an FS coincides with the structure of *Elementary Behavior* (EB), recently developed in the DNF framework to represent behaviors, which have to be initiated when appropriate, executed as long as needed, and terminated when the behavioral goal is reached (Richter, Sandamirskaya, & Schöner, 2012; Sandamirskaya, Richter, & Schöner, 2011). Fig. 2b shows a neural-dynamic EB and how it may be linked to the FS: the ''action'' output of the FS activates the *intention* node of the EB, the activated *condition of satisfaction (CoS)* node of the EB provides excitatory input to the ''goal'' neuron of the FS, which, in its turn, inhibits the intention node. The DNFs in the lower part of the EB figure encode the continuous parameters of the intention and CoS of the elementary behavior. The FSN decides when an EB is activated or adapted, and when a new EB is created. Both the FS and EB receive directly sensory input from the environment, but only EBs control actions of the agent.

Next, we present essential elements of the NARLE architecture in an exemplary robotic scenario, demonstrating their functionality and reviewing some preliminary results.

Description of NARLE

Outline of the task

We will describe a concrete implementation of the NARLE architecture for a particular robotic task. In this robotic scenario, a humanoid robot (NAO) is presented with a pictorial representation of a composition of small colored objects to be assembled (''target''). The robot is then placed in front of a tabletop on which colored objects are scattered, as depicted in Fig. 3. The objects are light-weight foam cubes that may be arranged in different configurations. The task for the robot is to build the figure presented in the beginning of the trial. Similar tasks have been used to study decision making, attention control, and action generation in humans subjects (Hayhoe & Ballard, 2005).

The neurocognitive architecture, introduced next, will control the robot in the envisioned task by implementing the following functionality: (1) perceive and parse the task diagram, (2) create an internal representation of a goal state in memory, (3) perceive and parse the tabletop, (4) compare current and goal states, (5) motivate, plan, and execute actions if current and goal states are different, (6) learn how to achieve the goal state if direct planning is not possible.

Perception and representation of objects in NARLE

Objects in our architecture are represented with DNFs using an object recognition system inspired by cognitive perception in humans, which relies on low-dimensional, ''wholistic'' features (Faubel & Schöner, 2008, 2010; Zibner, Faubel, & Schöner, 2011). The DNF object recognition

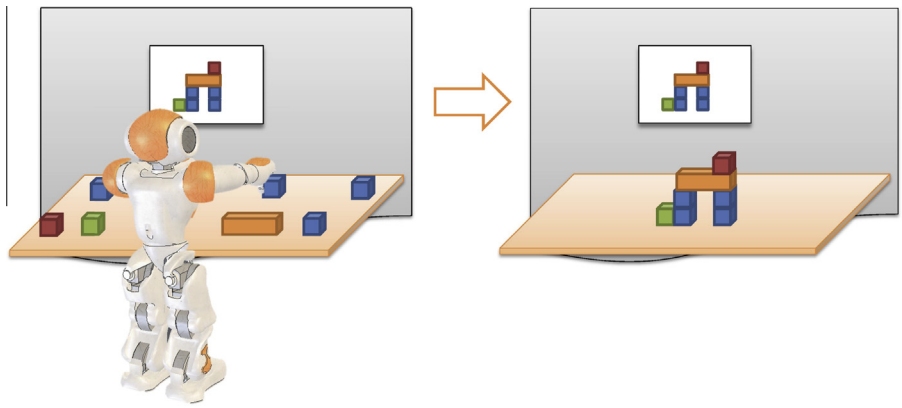

Fig. 3 The task setting. Left: NAO is presented with a picture of the objects' arrangement to build. Right: The robot builds the required objects' arrangement out of the objects scattered on the table.

architecture reaches a considerable performance, being able to differentiate 30 everyday objects in nine different poses after observing each object in a canonical pose. An improved architecture, which introduces distributed, histogram-based features, achieves an even better performance (Zibner, Faubel, Iossifidis et al., 2011). Recently, this architecture has been combined with a template-based pose recognition system (Knips, Zibner, Reimann, Popova, & Schöner, 2014) and a key-point based recognition system (Lomp, Terzić, Faubel, du Buf, & Schöner, 2014; Terzić & du Buf, 2014) to enable object recognition and grasp parameters extraction in a robotic table-top task.

In our scenario, objects are discriminated by two types of features (see Fig. 4), color and shape:

Color. In the framework of Dynamic Neural Fields, the state of the system is described by activation of the DNFs, Eq. (1), defined over different behavioral spaces. Color (hue value) is one such behavioral space, which is continuous, circular, and typically spans the range [0, 180]. From a color image of a conventional camera, the hue value is extracted for each pixel (after transforming the image into hue-saturation-value color space). The resulting hue-matrix is input to a DNF, which is defined over the dimensions of color and space. The dynamics of this perceptual DNF (Sandamirskaya, 2013; Sandamirskaya et al., 2013; Schöner, 2008) produces one or several narrow peaks of activity representing color localization over dimensions of position and hue. Such activity peaks correspond to colors and locations of objects in the image of the perceived scene. Through a scene-representation dynamics (Zibner, Faubel, Iossifidis et al., 2011), peaks for different objects are created sequentially and the color of each object is extracted. The memory traces of the color DNF activity (Sandamirskaya, 2013) can be retained to create memory

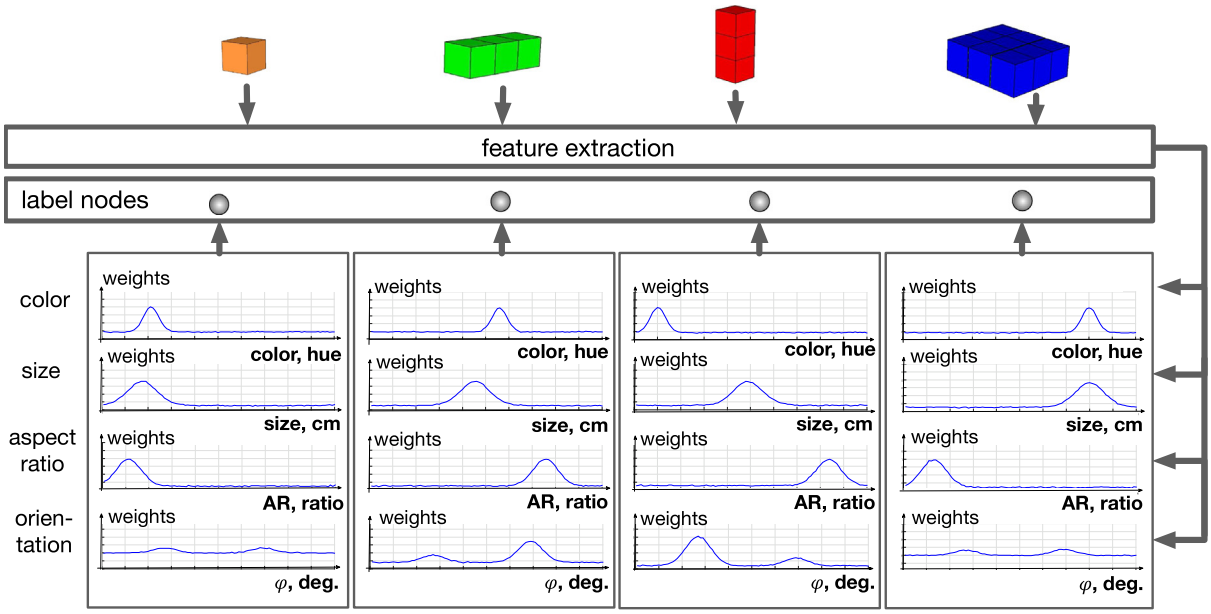

Fig. 4 Representation of the basic objects in the NARLE: four exemplar basic objects are shown with different combinations of the four features (color, size, aspect ratio, and prevalent edge orientation). The vision-based feature extraction routines generate input to four 1D DNFs, one for each feature dimension. The diagram shows connection weights (memory traces) formed during object-learning session between a label node and the four DNFs for each of the object.

of color-based representations of objects. These memory traces, by their width in the hue dimension, reflect the variability of each object's color, experienced by the robot during its behavior.

Shape. Several shape factors may be used to represent objects: (1) aspect ratio of the object's view projection onto the table surface (the system can discriminate round/square object from objects of different lengths), (2) distribution of edge orientations, and (3) size. According to their shape, the objects may be categorized into ''small graspable object'', ''medium graspable object'', ''long graspable object'', ''lying object'', ''standing object'', or ''large flat object''. Different low-level shape parameters — size, aspect ratio, and orientation — can be calculated with standard methods of computer vision. The estimated distribution of these parameters at each moment in time is input to the corresponding DNFs (see Fig. 4). Categorization of the shape parameters for novel objects may be learned in a supervised fashion through the user labeling and creation of memory traces in the DNFs' shape dimensions.

Parsing spatial relations (analyzing the task image)

To recognize a task, NARLE analyzes the task image in terms of the present objects and spatial relations between them. The features of the objects are extracted and stored in memory traces of the perceptual DNFs representing color, aspect ratio, size, and orientation. Each object is processed in a sequence in accordance with the DNF scene representation architecture (Zibner, Faubel, Iossifidis et al., 2011). The spatial relations architecture, developed in the DNF framework (Lipinski, Sandamirskaya, & Schöner, 2009; Lipinski, Schneegans, Sandamirskaya, Spencer, & Schöner, 2012; Richter, Lins, Schneegans, Sandamirskaya, & Schöner, 2014) is used to represent the spatial arrangement of objects in the scene relative to each other.

Fig. 5 shows an exemplary snapshot of the scene understanding process. The presented image is shown on the top of the figure and consists of two objects — a ''high'' green object and a ''small, square'' blue object to the right from the green one. Four plots in the two stacks hold the perceptual representation of the two objects, each consisting of four preshapes over four perceptual dimensions: color, size, aspect ratio, and orientation. An activity bump in each dimension specifies the estimated value of each of the four perceptual parameters for each object. Each bundle of four preshapes is linked to a label node, which represents the two objects, which are, in this case, known to the system. In the lower part of the figure, the spatial language system is shown, which consists of (1) two nodes that mark whether the selected object is a target or a reference of the spatial description, (2) a set of spatial-language templates (one for each relational term, ''left'', ''right'', ''below'', and ''above'' are shown here), and (3) the spatial term nodes, which signal the estimated spatial relation for each object in the currently parsed scene. The resulting distribution of the activation profiles in a number of preshapes and dynamical nodes represents the goal, or the task state, for the learning robot.

Action vocabulary (basic elementary behaviors)

When learning complex tasks, the state-action space available to the robot becomes very large. We address this issue by splitting learning in several stages. On the first stage, a number of basic *elementary behaviors* (EBs) can be designed by hand or learned through exploratory learning. On the second stage, sequences of these EBs for composite skills are acquired with FSN learning. Pervasive learning is achieved by addition of new EBs with variation of parameters of the existing EBs and establishing new intention-to-condition of satisfaction (CoS) associations. Feasibility of this approach has been demonstrated recently (Luciw, Kazerounian, Sandamirskaya, Schöner, & Schmidhuber, 2014). The basic EBs can be, for instance, ''find object x'' (where ''x'' is a parameter, provided externally, such as color or shape), ''close/open gripper'', ''move hand towards location y'', or ''turn hand'' (Richter & Sandamirskaya, 2012).

Designing a basic EB means to create the structure, schematically shown in Fig. 2a: (1) a DNF representation of the intention of the action spans parameters of the action, e.g. color of the object to find, size of the object being grasped, location to which the arm should be moved, orientation, at which the hand should be put, or target position for the docking maneuver; (2) a DNF representation of the final condition of actions (condition of satisfaction, or CoS), signaling, e.g. that the object of the sought color is centered in the visual field, the hand is closed over the object (contact sensed + opening of the hand), or that the hand has reached the intended orientation or location. The intention and the CoS DNFs are coupled through a predictive mapping, in which the anticipation about the final state of the current action is represented. The intention of the EBs is linked to the motor system of the robot and sets attractors for the involved dynamical systems (controllers) and to the sensory system of the robot, setting biases for the involved features. During action observation, the intention DNF is driven by the perceptual input, so that the intention for the particular behavior is activated if this behavior is observed.

Action selection

To complete a task, the robot should produce a sequence of elementary behaviors. This sequential execution is controlled by the FSN subsystem of our architecture (Fig. 6). At the task initiation, a goal is set to the robot, in particular, an image of the figure to be assembled is presented. Motivational FS, which is associated with this image is activated by corresponding ''problem'' inputs (extracted through scene parsing, described above) and remains in the active state until the task completion. The motivational FS induces subthreshold activation for all learned FSs related to the task. Every task-related FS is turned on at its position in the behavioral sequence by recognition of the memorized state of the environment. An active FS initiates the connected elementary behavior and supports it until goal state of the environment, which is specific for this EB in this task, is obtained. Results of EB activity are controlled by the ''goal'' inputs of the FS. When the current

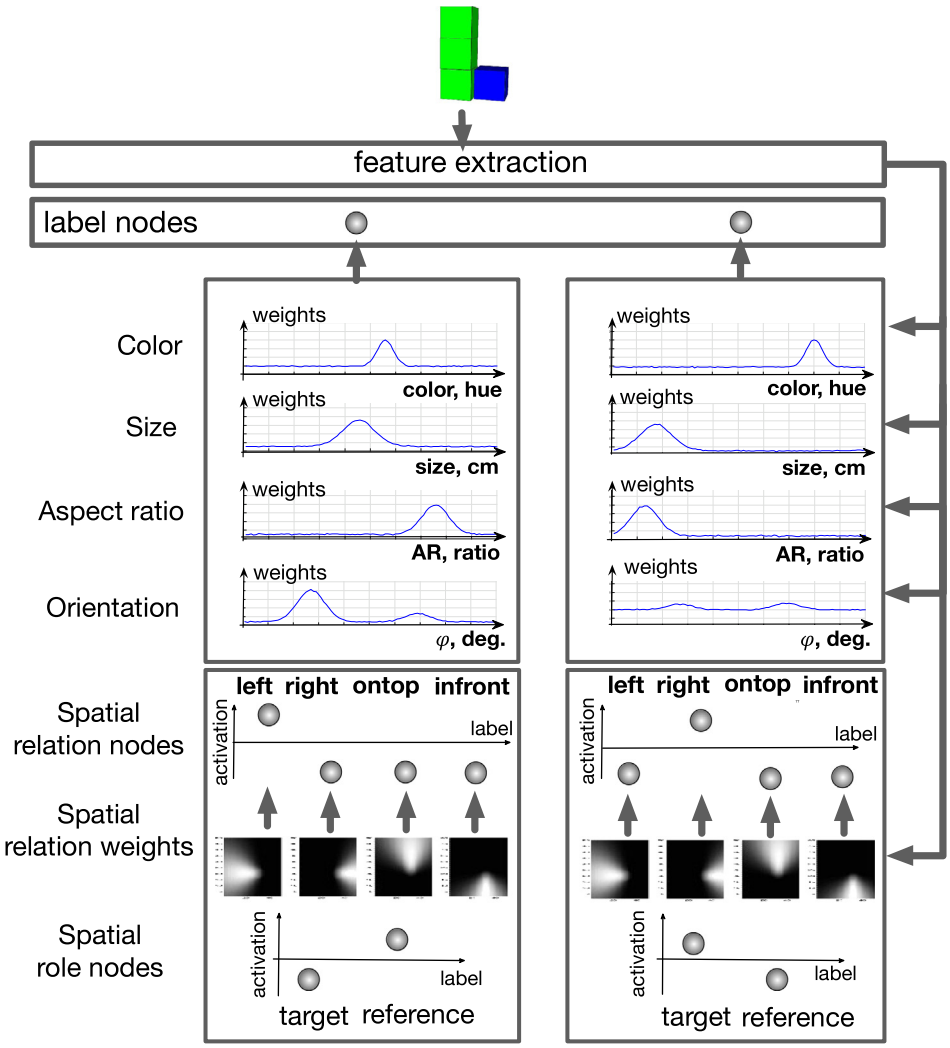

Fig. 5 Understanding state of the task: each object in a scene is processed sequentially by the dynamics of autonomous scene representation from Zibner, Faubel, Iossifidis et al. (2011). Visual features of each object are extracted and stored in a neural-dynamic object representation system (Faubel & Schöner, 2008). For the known object the memory trace is retrieved. A spatial language system (Lipinski et al., 2009, 2012; Richter et al., 2014) is used to extract and represent the arrangement of objects relative to each other (the spatial relations between them). According to the size or salience one of the objects is selected to be the reference and the spatial relations of other objects are represented relative to this reference.

step of the sequence is finished, the FS deactivates the EB and itself. The next task-related FS pops up in the same manner and remains active until activation of ''goal'' inputs of the current motivational FS is achieved. This event signals that the task as a whole is completed and the motivational FS is deactivated.

Learning elementary behaviors by action observation

For learning elementary behaviors by observation, the robot has to infer the actions performed by the human demonstrator. In contrast to kinesthetic teaching usually applied for imitation learning (e.g. Calinon, Guenter, & Billard, 2007), NARLE relies on visual perception. To accomplish this, a neural-dynamic action parser, presented recently (Lobato, Sandamirskaya, Richter, & Schöner, 2015), is extended with FSN-based memory. The temporal sequence of the demonstrated actions are interpreted by this system in terms of the involved actions (e.g., reaching, grasping, dropping, or turning) and the objects, at which these actions are directed. The objects are represented in the same way as objects during scene analysis (see Fig. 4). The goal is to interpret and memorize human actions as an activation distribution, which is equivalent to, e.g. such phrases as, ''pick-up red object'', ''put red object on top of the blue one'', ''dock green cylinder onto the red object'', etc. Fig. 7 shows the information flow between sensory and cognitive systems of the architecture during action parsing. The perceptual DNFs extract static features (such as color and shape descriptors of objects), as well as motion features (approach of the hand to and movement away from objects), which activate EBs, associating them to the specific action parameters and features of objects, at which the

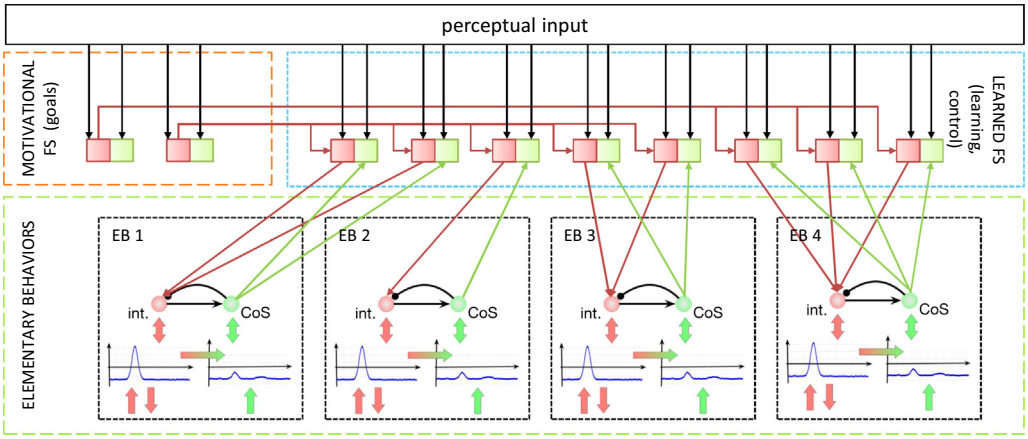

Fig. 6 Action selection. Every particular task is associated with a single motivational FS (represented by two (red and green) squares in the figure). Each FS has memory about parameters of the outcome required for the task, stored in the weighted connection between the FS's nodes and DNFs of the EB. A motivational FS turns on when the task description is presented to the robot and supports activity of all task-related FSs until the behavioral goal is reached. Every FS, pre-activated by the current motivation, monitors the current state of environment and steps in on its turn. An active learned FS initiates and controls the progress of the connected elementary behavior.

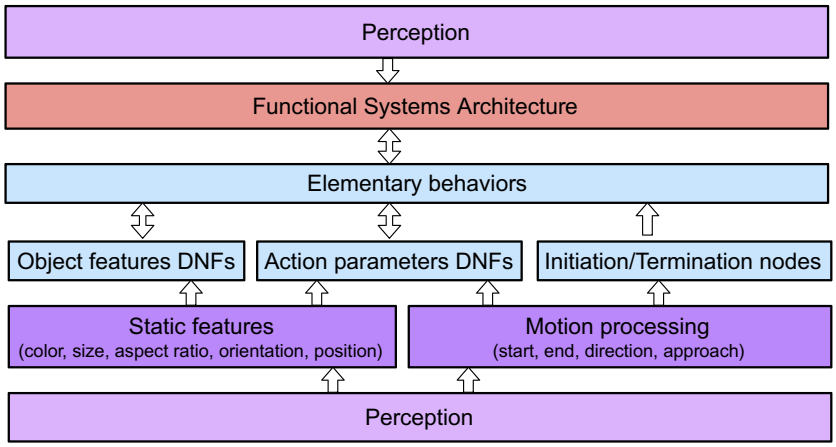

Fig. 7 The scheme for learning elementary behaviors by imitation. Low-level perceptual DNFs extract static features (such as color and shape descriptors of objects), as well as motion features (approach of the hand to and movement away from objects), which activate EBs, associating them to specific action and object features. Initiation and termination nodes detect temporal boundaries between behaviors. The FS architecture stores the observed sequences and associates it to the current motivational state.

action is directed. Initiation and termination nodes detect temporal boundaries between behaviors (see Lobato et al., 2015 for the details of the action parsing network). The FS architecture stores the observed sequence and associates it to the current motivational state.

Sequence learning with FSN

The functional systems network learns by an extension of the existing network with new functional systems. In the most simple case, there is only one motivational FS in the network at the beginning of learning. ''Goal'' inputs of this network are defined to detect the target state. The learning starts when only the motivational FS is active but not any other FS in the network (if there are any other FSs). In this case, a tentative FS is created to store parameters of the current state and motivation in its ''problem'' inputs. Since there is no learned FS to control actions, the random action is executed. If the results of an action mismatch the ''goal'' inputs of an active motivational FS or the ''problem'' inputs of any learned FS, the tentative FS cannot be integrated with the existing experience of the robot, and it is consequently discarded. A new tentative FS is generated in this case, as before, to continue learning. In the case when a random action leads to the expected task goal (i.e. the ''goal'' inputs of a motivational FS are activated) or activation of an already learned FS (i.e. the ''problem'' inputs of a learned FS are activated), a connection from the tentative FS to the executed action is created, as well as its ''goal'' inputs are set to store the state produced by

selected action. This FS is then added to the network and considered *learned*. As a result, every learned FS associates a particular motivation and a state of the environment with an action to be executed. It is also able to control action execution with its ''goal'' inputs.

This simple learning algorithm makes it possible to learn behavioral trajectories connecting different starting states with the goal state. It is important to note, that the FSN not only generates a sequence of goal-directed actions, but also controls success of every step in a distributed manner. It is very important for the robotic applications, where the environment can be very stochastic. If the prediction of an active learned FS has failed and no other FS has been activated, this FS is deactivated. Absence of active FSs initiates learning, and an alternative behavior is acquired. Thus, after learning over typical environmental conditions, the FSN can switch quickly between memorized courses of actions if some of them fail.

Here, we propose to connect ''goal'' inputs of the FSN to the CoS DNFs and outputs to the intention DNFs of elementary behaviors (as in Fig. 6). The FSN will then be able to learn action sequences to solve tasks posed to the robot by observing execution behaviors and their success and failure.

Learning EBs by exploration

The FSN framework allows exploratory learning of action sequences towards the defined outcome, consequently, it can also be used to acquire elementary behaviors out of sensorimotor activation patterns. For example, to learn EB for reaching an object, a motivational FS with ''goal'' inputs associated with correct reaching outcome is created. A tentative FS is created with ''problem'' inputs corresponding to the currently active peaks in the perceptual DNFs, becoming, effectively, an intention node of the new EB (Fig. 2b). The CoS DNFs are activated at random and condition of reaching is probed by the motivational FS. If movement is successful, the tentative FS receives activating connection from the motivational FS and sets connection to induce CoS for the new EB. If the movement fails, the tentative FS is discarded. Our work on using T-learning and shaping to facilitate exploration in the neural-dynamic reinforcement learning architecture lays foundation for this work (Luciw et al., 2014), whereas using FSN allows to probe and add new behaviors to the network.

Discussion

We have described a novel cognitive neural-dynamics architecture (NARLE), capable of pervasive learning in a behaving agent. In this section, we will discuss the merits of NARLE and position it with respect to related work in robotics and cognitive science.

Generalisation of the approach

How would the approach presented in this paper scale beyond the simple scenario we considered? Given the modularity of the approach, achieving scaling of perceptual features poses only a limited computational challenge: each new feature space adds a low-dimensional DNF to the architecture, which may be connected to the system of EBs and FSN through learning. In terms of motor control, on the other hand, the elementary behaviors of NARLE compete to control the shared degrees of freedom of the robot. The more degrees of freedom are available, the less conflict between different EBs arise for their arbitration, which then can be controlled by the task demands through the FSN. Therefore, there is no theoretical problem to scaling the architecture to more advanced hardware, although the sequence learning task may become more complex. This is related to the third dimension of scaling — the complexity of the task. Ultimately, autonomous learning must remain limited in scope so that enough experience can be gained with different aspects of the task to be learned. The analogy to human development can be insightful here: at different developmental stages, only a subset of tasks are open to learning. Perhaps, in analogy, autonomous learning in real-world robot applications will need to be restrained to limit the number of dimensions of a task, on which learning takes place at any moment in time. This will be an interesting topic for future research.

Related work

The NARLE architecture tackles the problem of learning to organize behaviors in time to achieve an overall goal. A pletora of cognitive architectures has been suggested, which offer different ways to address the problem of behavioral organization both in the field of understanding human cognition and for building autonomous intelligent robots.

In the field of modeling human cognition, a debate exists between the proponents of a structured, symbolic approach, in which behaviors are represented explicitly and symbolic calculations (e.g. search) on graph-like representations lead to planning of goal-directed sequences (Cooper & Shallice, 2006b), and those who believe that a system which organizes complex behaviors may develop through learning from a homogeneous substrate of a neural network, leading to a distributed representation of behaviors and sequences (Botvinick & Plaut, 2004; Elman, 1991). Whereas first approaches suffer from fixed behaviors, which cannot be easily learned or adjusted to a given situation, the latter, distributed neural network approaches, do not scale well to complex hierarchical action sequences (Cooper & Shallice, 2006a). The learning process in such systems is supervised, non-autonomous, and offline. The question of the nature of modularity in biological neuronal system, related to the symbolic vs. subsymbolic debate, is yet to be resolved (Bullinaria, 2007). NARLE, in its structure, lies between the symbolic and non-symbolic approaches. It offers a subsymbolic, continuous substrate of Dynamic Neural Fields to represent parameters of actions and features of objects at which the actions are directed, at the same time providing a flexible link to discrete-nodes, ''symbolic'' representations of the behaviors. In other words, NARLE postulates more structure on the neuronal substrate which underlies behavior than typical neural network approaches (Elman, 1991), retaining, nevertheless, a possibility for adaptation and learning.

A shared limitation of many cognitive models is their disembodiment: e.g., action schemas (Cooper & Shallice, 2006b; Dominey, Hoen, & Inui, 2006) provide action descriptions, but are not yet actionable behaviors — they lack stabilisation of the activated action, required to ensure that the respective behavior may in fact be realized in the physical environment, as well as the means to detect that the action has been successfully accomplished, or has failed. Simple attractor networks that represent movement sequences (Stringer, Rolls, & Taylor, 2007), which use a similar associative learning mechanism as NARLE to represent a sequence of movements by associating a high-level movement selector with a set of movement selector output patterns (each activating a movement primitive), may also not easily be embodied. The neural-dynamic structure of an EB in NARLE offers a substrate that may be directly linked to sensors and effectors, due to stabilising properties of DNFs and the EB structure, and control behavior of a robotic agent.

The embodiment problem has to be solved, obviously, in behavior organization architectures that control robotic hardware. Many cognitive architectures were specifically designed to control robots and increase their flexibility and robustness in situations, which require learning and adaptation (Asfour et al., 2008; Pardowitz, Knoop, Dillmann, & Zöllner, 2007; Uchibe, Asada, & Hosoda, 1996; Tenorth & Beetz, 2013). These architectures typically do not focus on learning and adaptation of the behavioral modules, which is exactly the challenge which NARLE aims to tackle. Learning from demonstration (Argall, Chernova, Veloso, & Browning, 2009) and imitation learning of movement primitives (Ijspert, Nakanishi, & Schaal, 2002; Mochizuki, Nishide, Okuno, & Ogata, 2013) are among the most successful attempts to implement robotic learning systems. Using insights from work on human—robot cooperation (Dominey, Mallet, & Yoshida, 2009; Lallee, Madden, Hoen, & Dominey, 2010; Madden, Hoen, & Dominey, 2010) and segmentation of observed action sequences (Guerra-Filho & Aloimonos, 2006; Wörgötter et al., 2013), in our work, we aim at more ''cognitive'' actions, which not only specify the desired trajectory of the robot's effectors, but also the properties of the target object and its desired final state (Cuijpers, Schie, Koppen, Erlhagen, & Bekkering, 2006). Learning such goal-directed behaviors and learning sequence of such behaviors simultaneously is an open problem.

The approach we pursue is related to other work on representation and learning of action sequences using Dynamic Neural Fields (Bicho, Louro, & Erlhagen, 2010; Erlhagen et al., 2006; Ferreira, Erlhagen, & Bicho, 2011) and more general neural dynamics (Cisek, 2006, 2007; Cisek & Kalaska, 2010). An example of a similarly comprehensive robotic learning project was developed early on, but lacks experimental validation (Grupen, 2003). The EBs that compose the NARLE architecture are similar to cognitive simulators, proposed by Barsalou (2008) as processes, which ground cognition in sensorimotor interactions. The learning mechanisms of NARLE implement in a computational framework the theoretical principles of ''ongoing emergence'', which is thought to be at the core of epigenetic robotics (Prince, Helder, & Hollich, 2003).

Overall, the NARLE architecture builds on a body of work in modeling human cognition, in particular representation and planning of action sequence, as well as in cognitive and developmental robotics.

Conclusions

In this paper, we propose to combine several functional components based on Dynamic Neural Fields, which were previously validated in robotic experiments, with Functional Systems Network validated in machine learning scenarios. Here, we demonstrated the main principles and architectural decisions leading to a new neural-dynamic cognitive architecture — NARLE — which enables pervasive learning, in the sense that both the elementary behaviors and goal-directed sequences of actions can be adapted according to the demands of the current environment and task. The NARLE, apart from respecting the principles of modularity of sensorimotor controllers, offers the following advantageous properties for development of extendable and flexible robotic learning architectures:

- Homogeneous substrate: In NARLE, the behavioral modules, including the representation of features and labels, spatial terms and spatial arrangements, memory, motor parameters, and goals, all are formulated with Dynamic Neural Fields — activation functions, defined over different dimensions and following a particular, generic and powerful attractor dynamics. DNFs may be coupled to each other through connecting weight functions, which are subject to simple, neurally-based, learning rules (shunted Hebbian learning). Augmented with memory trace dynamics, the DNFs offer computational substrate, in which all behaviors of the agent may be described — including new EBs, created during behavior based on the successes and failures of the learning agent.
- A bridge between the continuous sensorimotor representations and discrete ''cognitive'' representations: The dynamic instabilities, characteristic for DNFs, create localized activity peaks as units of representation. These localized peaks may drive discrete dynamical nodes, which then will stand for the particular range of features, corresponding to the selected color, position, or movement parameter. Thus, the DNF dynamics bridge the continuous sensorimotor representations and the discrete representations that may be used as symbols in an action planning process.
- Continuous dynamics and segregation of behaviorally-relevant events in time: The EBs and DNFs of NARLE architecture continuously integrate sensory inputs, inputs from other parts of the architecture, and self-excitatory inputs, leading to temporally distinct instabilities at transitions between an active and an inactive state of each DNF. These instabilities mark behaviorally relevant events, effectively chunking behavior in time in a flexible, context-dependent manner.
- Pervasive learning: Every functional system has a built-in feedback loop to control effectiveness of its own action. As a result, the FS network simultaneously generates commands for actions and a distributed prediction of the consequences of these actions. This allows to detect

and localize elements of the network related to failures in goal-directed behaviors and to initiate learning to solve the problem.

- Accumulation of experience (''open-ended learning''): the FSN algorithm stores new experiences by adding new FS elements to the network. That has a number of advantages. First, the network structure for the old behavior is kept intact after learning. Consequently, even if connections acquired during learning are erased, the old behaviors remain undisturbed. In this way, the agent can learn a number of alternative behaviors for the same problem and try them one by one, which is an effective solution in many real-world situations. As the structure of FSN is not fixed, but grows on demand, open-ended learning is possible in the system.

This cognitive architecture, we believe, will go beyond current state of the art in reinforcement learning applied to robotics, since it can learn on the interface between the continuous parameters of the sensorimotor states and the discrete, labeled action representations, used for planning. Thus, learning can unfold on continuous spaces that characterize actions and states, can be one-shot, and does not suffer from catastrophic forgetting (unsuccessful actions are decoupled from the current sequence, but their representation remains in the FSN). Multi-goal, hierarchical learning of goal-directed sequences is possible in this new framework. The DNF+FSN framework NARLE allows not only to learn sequences of behaviors, but also to adapt the behaviors to the changing environmental demands and add new behaviors as needed. Thus, NARLE goes beyond behavior-based robotics, adding pervasive learning to its principles of modular control. We hope that implementation and further development of NARLE architecture will make a significant contribution to the field of developmental and adaptive robotics, implementing learning in a complex behaving architecture that can perform and chain a multitude of behaviors.

Aknowledgements

The authors gratefully acknowledge the financial support of the EU FP7-ICT-2009-6 Grant NeuralDynamics (270247) and the EU ERC Grant neuroP (257219).

References

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics, 27*, 77—87.

Anokhin, P. (1974). *Biology and neurophysiology of the conditioned reflex and its role in adaptive behavior*. Oxford: Pergamon.

Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems, 57*(5), 469—483.

Asfour, T., Azad, P., Vahrenkamp, N., Regenstein, K., Bierbaum, A., Welke, K., et al (2008). Toward humanoid manipulation in human-centred environments. *Robotics and Autonomous Systems, 56*(1), 54—65.

Asfour, T., Regenstein, K., Azad, P., Schroder, J., Bierbaum, A., Vahrenkamp, N., et al (2006). Armar-iii: An integrated humanoid platform for sensory-motor control. In *2006 6th IEEE-RAS international conference on humanoid robots* (pp. 169—175). IEEE.

Baldassarre, G., & Mirolli, M. (2013). *Intrinsically motivated learning systems: An overview*. Berlin, Heidelberg: Springer.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 1—41.

Barto, A. G., & Mahadevan, S. (2003). Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems, 13*(4), 41—77.

Bastian, A., Schöner, G., & Riehle, A. (2003). Preshaping and continuous evolution of motor cortical representations during movement preparation. *European Journal of Neuroscience, 18*, 2047—2058.

Bicho, E., Erlhagen, W., Louro, L., & Costa e Silva, E. (2011). Neurocognitive mechanisms of decision making in joint action: A human—robot interaction study. *Human Movement Science, 30*(5), 846—868.

Bicho, E., Louro, L., & Erlhagen, W. (2010). Integrating verbal and nonverbal communication in a dynamic neural field architecture for human robot interaction. *Frontiers in Neurorobotics, 4*(May), 13.

Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition, 113*(3), 262—280.

Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review, 111*(2), 395—429.

Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation, 2*(1), 14—23.

Bryson, J. (2003). The behavior-oriented design of modular agent intelligence. In *Agent technologies*.

Bullinaria, J. A. (2007). Understanding the emergence of modularity in neural systems. *Cognitive Science, 31*(4), 673—695.

Calinon, S., Guenter, F., & Billard, A. (2007). On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society, 37*(2), 286—298.

Cangelosi, A., Schlesinger, M., & Smith, L. B. (2015). *Developmental robotics: From babies to robots*. MIT Press.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience, 26*(38), 9761—9770.

Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1485), 1585—1599.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience, 33*, 269—298.

Cooper, R., & Shallice, T. (2006a). Structured representations in the control of behavior cannot be so easily dismissed: A reply to Botvinick and Plaut (2006). *Psychological Review, 113*(4), 929—931.

Cooper, R. P., & Shallice, T. (2006b). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review, 113*(4), 887—916, discussion 917—31.

Cuijpers, R. H., Schie, H. T. V., Koppen, M., Erlhagen, W., & Bekkering, H. (2006). Goals and means in action observation: A computational approach. *Neural Networks, 19*, 311—322.

Dominey, P. F., Hoen, M., & Inui, T. (2006). A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience, 18*(12), 2088—2107.

Dominey, P. F., Mallet, A., & Yoshida, E. (2009). Real-time spoken-language programming for cooperative interaction with a humanoid apprentice. *International Journal of Humanoid Robotics, 6*(2), 147—171.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning, 7*, 195—225.

Erlhagen, W., Mukovskiy, A., Bicho, E., Panin, G., Kiss, C., Knoll, A., et al (2006). Goal-directed imitation for robots: A bio-inspired approach to action understanding and skill learning. *Robotics and Autonomous Systems, 54*(5), 353—360.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review, 109*, 545—572.

Faubel, C.& Schöner, G. (2010). Learning objects on the fly — Object recognition for the here and now. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*.

Faubel, C., & Schöner, G. (2008). Learning to recognize objects on the fly: A neurally based dynamic field approach. *Neural Networks, 21*, 562—576.

Ferreira, F., Erlhagen, W., & Bicho, E. (2011). A dynamic field model of ordinal and timing properties of sequential events. *Lecture Notes in Computer Science, 6792*, 325—332.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks, 1*, 17—61.

Grupen, R. A. (2003). A developmental organization for robot behavior. In *Proceedings of the third international workshop on epigenetic robotics: Modeling cognitive development in robotic systems*. Lund, Sweden: Lund University Cognitive Studies.

Guerra-Filho, G., & Aloimonos, Y. (2006). Understanding visuo-motor primitives for motion synthesis and analysis. *Computer Animation and Virtual Worlds, 17*, 207—217.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188—194.

Ijspert, A., Nakanishi, J., & Schaal, S. (2002). Learning rhythmic movements by demonstration using nonlinear oscillators. In *IEEE/RSJ international conference on intelligent robots and systems, Vol. 1(October)*, pp. 958—963.

Ingrand, F., & Ghallab, M. (2014). Deliberation for autonomous robots: A survey. *Artificial Intelligence*.

Jamone, L., Brandao, M., Natale, L., Hashimoto, K., Sandini, G., & Takanishi, A. (2014). Autonomous online generation of a motor representation of the workspace for intelligent whole-body reaching. *Robotics and Autonomous Systems, 62*(4), 556—567.

Johnson, J. S., Spencer, J. P., & Schöner, G. (2008). Moving to higher ground: The dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology, 26*, 227—251.

Knips, G., Zibner, S. K. U., Reimann, H., Popova, I., & Schöner, G. (2014). Reaching and grasping novel objects: Using neural dynamics to integrate and organize scene and object perception with movement generation. In *Proceedings of the joint IEEE International Conference on Development and Learning & Epigenetic Robotics (ICDL-EPIROB)*.

Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research, 32*(11), 1238—1274.

Komarov, M., Osipov, G., & Burtsev, M. (2010). Adaptive functional systems: Learning with chaos. *Chaos, 20*(4), 045119.

Konidaris, G. D. (2005). An architecture for behavior-based reinforcement learning. *Adaptive Behavior, 13*(1), 5—32.

Krüger, N., Geib, C., Piater, J., Petrick, R., Steedman, M., Wörgötter, F., et al (2011). Object—action complexes: Grounded abstractions of sensory—motor processes. *Robotics and Autonomous Systems, 59*(10), 740—757.

Kushiro, K., Harada, Y., & Takeno, J. (2013). Robot uses emotions to detect and learn the unknown. *Biologically Inspired Cognitive Architectures, 4*, 69—78.

Lakhman, K., & Burtsev, M. (2014). Evolution, development and learning with predictor neural networks. In *ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems* (pp. 457—464). The MIT Press. http://dx.doi.org/10.7551/978-0-262-32621-6-ch075.

Lallee, S., Madden, C., Hoen, M., & Dominey, P. F. (2010). Linking language with embodied and teleological representations of action for humanoid cognition. *Frontiers in Neurorobotics, 4*(June), 8.

Lipinski, J., Sandamirskaya, Y., & Schöner, G. (2009). Swing it to the left, swing it to the right: Enacting flexible spatial language using a neurodynamic framework. *Cognitive Neurodynamics, 3*(4).

Lipinski, J., Schneegans, S., Sandamirskaya, Y., Spencer, J., & Schöner, G. (2012). A neuro-behavioral model of flexible spatial language behaviors. *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*(6), 1490—1511.

Lobato, D., Sandamirskaya, Y., Richter, M., & Schöner, G. (2015). Parsing of action sequences: A neural dynamics approach. *Journal of Behavioral Robotics (Paladyn)*.

Lomp, O., Terzić, K., Faubel, C., du Buf, J. M. H., & Schöner, G. (2014). Instance-based object recognition with simultaneous pose estimation using keypoint maps and neural dynamics. In *International Conference for Artificial Neural Networks, ICANN*, pp. 451—458.

Luciw, M., Kazerounian, S., Sandamirskaya, Y., Schöner, G., & Schmidhuber, J. (2014). Reinforcement-driven shaping of sequence learning in neural dynamics. In *Simulation of Adaptive Behavior, SAB*.

Madden, C., Hoen, M., & Dominey, P. F. (2010). A cognitive neuroscience perspective on embodied language for human—robot cooperation. *Brain and Language, 112*(3), 180—188.

Maes, P. (1990). Situated agents can have goals. *Designing autonomous agents: Theory and practice from biology to engineering and back*. Cambridge, MA: MIT Press (pp. 49—70). Cambridge, MA: MIT Press, chapter.

Mochizuki, K., Nishide, S., Okuno, H. G., & Ogata, T. (2013). Developmental human—robot imitation learning of drawing with a neuro dynamical system. In *2013 IEEE international conference on systems, man, and cybernetics*, pp. 2336—2341.

Natale, L., Rao, S., & Sandini, G. (2002). Learning to act on objects. In H. Bülthoff, C. Wallraven, S.-W. Lee, T. Poggio, L. Natale, & S. Rao, et al. (Eds.). *Biologically motivated computer vision* (Vol. 2525, pp. 89—100). Berlin/Heidelberg: Springer.

Nicolescu, M. & Mataric, M. (2003). Linking perception and action in a control architecture for human—robot domains. In *2003 Proceedings of the 36th annual Hawaii international conference on system sciences*, p. 10.

Pardowitz, M., Knoop, S., Dillmann, R., & Zöllner, R. D. (2007). Incremental learning of tasks from user demonstrations, past experiences, and vocal comments. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society, 37*(2), 322—332.

Pirjanian, P. (1999). Behavior coordination mechanisms-state-of-the-art. Technical report, Institute for Robotics and Intelligent Systems, University of Southern California, IRIS-99-375.

Prince, C. G., Helder, N. A., & Hollich, G. J. (2003). Ongoing emergence: A core concept in epigenetic robotics. In *Proceedings of the fifth international workshop on epigenetic robotics: Modeling cognitive development in robotic systems* (pp. 63—70). Lund University Cognitive Studies.

Proetzsch, M., Luksch, T., & Berns, K. (2010). Development of complex robotic systems using the behavior-based control architecture iB2C. *Robotics and Autonomous Systems, 58*(1), 46—67.

Red'ko, V. G., Prokhorov, D. V., & Burtsev, M. S. (2004). Theory of functional systems, adaptive critics and neural networks. In *Proceedings. 2004 IEEE International Joint Conference on Neural Networks, 2004* (Vol. 3, pp. 1787—1792). IEEE.

Redko, V. G., Anokhin, K. V., Burtsev, M. S., Manolov, A. I., Mosalov, O. P., Nepomnyashchikh, V. A., et al., (2007). Project animat brain: Designing the animat control system on the basis

of the functional systems theory. In M. V. Butz, O. Sigaud, G. Pezzulo, & G. Baldassarre (Eds.)., *Anticipatory behavior in adaptive learning systems, Vol. 4520 of Lecture Notes in Computer Science* (pp. 94—107). Springer Berlin Heidelberg.

Richter, M. & Sandamirskaya, Y. (2012). Neural dynamics for behavioral organization of an embodied agent. In *16th International Conference on Cognitive and Neural Systems (ICCNS)*.

Richter, M., Sandamirskaya, Y., & Schöner, G. (2012). A robotic architecture for action selection and behavioral organization inspired by human cognition. In *IEEE/RSJ international conference on Intelligent Robots and Systems, IROS*.

Richter, M., Lins, J., Schneegans, S., Sandamirskaya, Y., & Schöner, G. (2014). Autonomous neural dynamics to test hypotheses in a model of spatial language. In *The annual meeting of the cognitive science society, CogSci*.

Rylatt, M., Czarnecki, C., & Routen, T. O. M. (1998). Connectionist learning in behaviour-based mobile robots: A survey. *Artificial Intelligence Review, 12*(6), 445—468.

Saeb, S., Weber, C., & Triesch, J. (2011). Learning the optimal control of coordinated eye and head movements. *PLoS Computational Biology, 7*(11), e1002253.

Saegusa, R., Metta, G., Sandini, G., & Natale, L. (2014). Developmental perception of the self and action. *IEEE Transactions on Neural Networks and Learning Systems, 25*(1), 183—202.

Sandamirskaya, Y. (2013). Dynamic neural fields as a step towards cognitive neuromorphic architectures. *Frontiers in Neuroscience, 7*, 276.

Sandamirskaya, Y., Richter, M., & Schöner, G. (2011). A neural-dynamic architecture for behavioral organization of an embodied agent. In *IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL EPIROB 2011)*.

Sandamirskaya, Y., Zibner, S. K., Schneegans, S., & Schöner, G. (2013). Using dynamic field theory to extend the embodiment stance toward higher cognition. *New Ideas in Psychology, 31*(3), 322—339.

Schöner, G. (2008). Dynamical systems approaches to cognition. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modeling* (pp. 101—126). Cambridge, UK: Cambridge University Press.

Shirshova, M. P., & Burtsev, M. S. (2014). Evolution of event and delay controlled neuronal network for locomotion. In *Intl. Conf. Genetic and Evolutionary Methods GEM14* (pp. 41—47).

Stringer, S. M., Rolls, E. T., & Taylor, P. (2007). Learning movement sequences with a delayed reward signal in a hierarchical model of motor function. *Neural Networks, 20*, 172—181.

Sun, R., Mathews, R. C., & Lane, S. M. (2007). Implicit and explicit processes in the development of cognitive skills: A theoretical interpretation with some practical implications for science education. *Educational Psychology Research Focus*, 1—26.

Sun, R., Merrill, E., & Peterson, T. (1999). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. MIT Press.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence, 112*(1), 181211.

Sutton, R., Rafols, E., & Koop, A. (2006). Temporal abstraction in temporal-difference networks. *Advances in Neural Information Processing Systems, 18*, 1313.

Tenorth, M., & Beetz, M. (2013). KnowRob: A knowledge processing infrastructure for cognition-enabled robots. *The International Journal of Robotics Research, 32*(5), 566—590.

Terzić, K. & du Buf, J. (2014). An efficient naive Bayes approach to category-level object detection. In *IEEE International Conference on Image Processing (ICIP)*, pp. 1658—1662.

Uchibe, E., Asada, M., & Hosoda, K. (1996). Behavior coordination for a mobile robot using modular reinforcement learning. In *Proceedings of the 1996 IEEE/RSJ international conference on Intelligent Robots and Systems' 96*, IROS 96, pp. 1329—1336.

Vigorito, C. M., & Barto, A. G. (2010). Intrinsically motivated hierarchical skill learning in structured environments. *IEEE Transactions on Autonomous Mental Development, 2*(2), 132—143.

Vityaev, E. E. (2015). Purposefulness as a principle of brain activity. In M. Nadin (Ed.), *Anticipation: Learning from the past, chapter Purposeful*, pp. 231—254.

Wilimzig, C., & Schöner, G. (2005). The Emergence of stimulus-response associations from neural activation fields: Dynamic field theory. In *In proceedings of the twenty-seventh annual cognitive science society* (pp. 2359—2364). Stresa, Italy: Cognitive Science Society.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik, 13*, 55—80.

Wörgötter, F., Aksoy, E. E., Krüger, N., Piater, J., Ude, A., & Tamosiunaite, M. (2013). A simple ontology of manipulation actions based on hand-object relations. *IEEE Transactions on Autonomous Mental Development, 5*, 117—134.

Zibner, S. K. U., Faubel, C., & Schöner, G. (2011). Making a robotic scene representation accessible to feature and label queries. *In Proceedings of the first joint IEEE International Conference on Development and Learning and on Epigentic Robotics, ICDL-EPIROB*.

Zibner, S. K. U., Faubel, C., Iossifidis, I., & Schöner, G. (2011). Dynamic neural fields as building blocks for a cortex-inspired architecture of robotic scene representation. *IEEE Transactions on Autonomous Mental Development, 3*(1), 74—91.