

**On observational learning of hierarchies in sequential tasks:
A dynamic neural field model**

Emanuel Sousa

*Department of Industrial Electronics, University of Minho, Guimarães, Portugal
E-mail: esousa@dei.uminho.pt*

Wolfram Erlhagen

*Centre of Mathematics, University of Minho, Guimarães, Portugal;
Donders Institute for Cognition Brain and Behavior, Radboud University Nijmegen,
The Netherlands
E-mail: wolfram.erlhagen@math.uminho.pt*

Estela Bicho

*Department of Industrial Electronics, University of Minho, Guimarães, Portugal
E-mail: estela.bicho@dei.uminho.pt*

Many of the tasks we perform during our everyday lives are achieved through sequential execution of a set of goal-directed actions. Quite often these actions are organized hierarchically, corresponding to a nested set of goals and sub-goals. Several computational models address the hierarchical execution of goal-directed actions by humans. However, the neural learning mechanisms supporting the temporal clustering of goal-directed actions in a hierarchical structure remain to a large extent unexplained. In this paper we investigate in simulations, of a dynamic neural field (DNF) model, biologically-based learning and adaptation mechanisms that can provide insight into the development of hierarchically organized internal representations of naturalistic tasks. In line with recent experimental evidence from observational learning studies, the DNF model implements the idea that prediction errors play a crucial role for grouping fine-grained events into larger units. Our ultimate goal is to use the model to endow the humanoid robot ARoS with the capability to learn hierarchies in sequential tasks, and to use that knowledge to enable efficient collaborative joint tasks with human partners. For testing the ability of the system to deal with the real-time constraints of a learning-by-demonstration paradigm we use the same assembly task from our previous work on human-robot collaboration. The model provides some insights on how hierarchically structured task representations can be learned and on how prediction errors made by the robot and signaled by the demonstrator can be used to control such process.

1. Introduction

In our everyday activities such as cleaning the kitchen, preparing a meal or serving a coffee a complex stream of object-directed actions has to be organized in real time in order to fulfill the task. Although these actions unfold in time as a linear sequence, goal-directed behavior is very often organized hierarchically, corresponding to a nested set of goals and subgoals.¹⁻³ Groups of goal-related actions appear to be linked together at multiple timescales through their relations to a common cause and not primarily through their temporal and causal relations to each other. Such a hierarchical organization of behavior is believed to support efficient action planning since it allows the actor to select individual actions and action clusters based on their predicted consequences in the near and distal future. It also promotes flexibility in task execution since the temporal order of subgoals and their associated subroutines may be easily adapted to changing environmental constraints. Think of the coffee serving example. The ingredients coffee, milk and sugar can in principle be added in an any order, but all subgoals have to be completed before the cup can be handed over.

Several existing computational accounts explain important aspects of hierarchical action production in humans (e.g. Refs. 4 and 5). However, the neural learning mechanisms that give rise to the temporal clustering of goal-directed actions in a hierarchical structure remains to a large extent unexplained. Developmental studies with children reveal that infants at a surprisingly early age of 24-36 months are able to imitate the goal-subgoal structure of novel action sequences composed of several separate, but jointly necessary means steps.^{6,7} In other words, the toddlers are not simply copying the chain of events demonstrated by the adult but imitate on a hierarchical level. In these studies, the children had little explicit event knowledge that might help them to structure the task from the beginning on, suggesting the operation of bottom-up mechanisms in learning the hierarchical structure.

In the present paper, we investigate in simulations of a dynamic neural field (DNF) model biologically-based learning and adaptation mechanisms that can provide insight into the development of hierarchically organized internal representations of naturalistic tasks. Specifically, we explore the idea that prediction errors may play a fundamental role in grouping fine-grained events into larger units.⁸ Ultimately, we are interested in using the DNF model to endow the humanoid robot ARoS developed by our group with a basic hierarchy imitation competence. It is thus important that the model is able to capture the real-time constraints of a learning by demonstration

paradigm.⁹ To this end, we adopt for the model simulations an assembly task from our previous DNF approach to natural human-robot interactions^{10,11} for a socially guided learning paradigm.¹² One or more human teachers first demonstrate possible linear sequences of assembly steps that result in the construction of a toy vehicle from its parts. During observation, the dynamic field model establishes synaptic links between sufficiently active neural populations encoding the perceived consequences of subsequent assembly steps using Hebbian learning principles. The robot then tries to reproduce the sequential task and makes at every stage of the assembly work a prediction about the subgoal that has to be achieved next. In the socially guided learning paradigm, the teacher gives immediate positive or negative feedback about this prediction. If a prediction failure occurs, the error signal is used to lower the threshold for associative learning during new task demonstrations.¹³ As a result, time-dependent population activity representing the memory of already accomplished subgoals remains above threshold for a longer time period. The system is thus able to develop longer-term sequential dependencies that reduce the likelihood of prediction failure. In addition, the error signal becomes associated with the segmentation of the assembly work in independent subtasks.^{8,14,15} A neural population, which is initially driven by the error signal, establishes connections to the cluster of all subgoal representations defining a certain subtask.

Learning of task hierarchies thus necessarily means that the activity of the higher-level representation has to be maintained above the learning threshold during the course of the sequential activation of the subgoal representations. Dynamic neural field theory provides a rigorous mathematical framework to explain the existence of self-stabilized activity patterns of neural populations representing such a working memory function.^{16,17}

2. Dynamic Neural Field Model

The model is based on the theoretical framework of Dynamic Neural Fields (DNFs) that was originally proposed as a simplified mathematical model to explain firing patterns of neuronal populations in cortical tissue.^{16,18} The architecture of this model family reflects the hypothesis that strong recurrent excitatory and inhibitory interactions in local populations of neurons form a basic mechanism for cortical information processing. The recurrent interactions cause non-trivial dynamic behaviour in neural assemblies. Most importantly, population activity which is initiated by time-dependent external signals may become self-sustained in the absence of any external

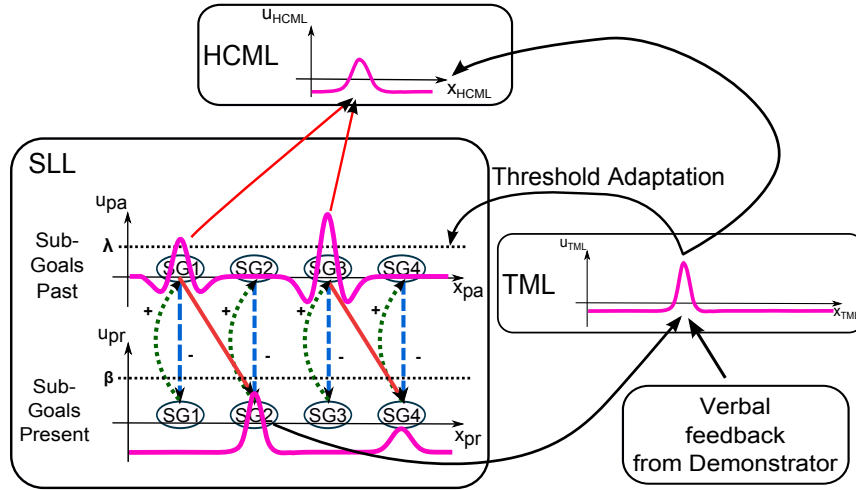


Fig. 1. Schematic of the model architecture with three connected layers implementing sequence learning (SLL), task monitoring (TML), and higher level cognitive memory (HCML).

input. Such attractor states of the population dynamics may be used by the nervous system to guide complex goal-directed behavior that goes beyond simple input-output mappings. DNF models have been used in the past to model cognitive capacities like working memory, decision making, prediction and anticipation^{19,20} and to synthesize these functionalities in artificial agents (for a review see Ref. 17).

Figure 1 presents a sketch of the model architecture for learning and representing hierarchical structure in sequential tasks. It consists of three layers with neural populations that receive input from connected neural pools and external sources (vision, speech). In the sequence learning layer (SLL), a Hebbian mechanism is applied to store the temporal order(s) of subgoals as demonstrated by the teacher. The task monitoring layer (TML) contains populations that become active when during task execution an erroneous prediction about a subsequent assembly step is detected. Suprathreshold population activity in TML triggers in turn the development of population activity in the higher cognitive memory layer (HCML) that signals the accomplishment of an entire subtask composed of several subgoals represented in SLL.

The model takes inspiration from recent findings in cognitive neuroscience. Neural populations in areas of the prefrontal cortex (PFC) are known to be concerned with cognitive aspects of behavioral planning.²¹

They seem to encode the end result of object-directed behavior (e.g., a change in the state of the target object) rather than the intended movements required to perform the task. We conceptualize this finding in our model by assuming that distinct neural populations encode the desired end-states or goals of individual assembly steps. For successful execution of a sequential task, the working memory of already completed steps has to be separated from steps that are still to be accomplished. Firing patterns of neural populations in PFC suggest that previous and future goals are indeed represented by separate but interacting neuronal assemblies.²² The DNF-model reflects this finding by assuming the existence of a “past” layer, u_{pa} , and a “present” layer, u_{pr} , each with representations of end-states of all assembly steps. Populations encoding the same subgoal are coupled by pre-defined excitatory and inhibitory connections (see Fig. 1, dotted lines in SLL). During task demonstration, vision input indicating a change in the state of the target object initiates an interaction dynamics that results in an automatic updating of the working memory in u_{pa} . Subsequently, a Hebbian learning mechanism establishes connections to populations in u_{pr} representing newly demonstrated subgoals (see Fig. 1, solid arrows in SLL) as long as the population activity in u_{pa} remains above the learning threshold²³ (for a discussion of the threshold concept see Ref. 13). Figure 2 compares the time course of population activity in u_{pr} and u_{pa} when at time t_0 the population in u_{pr} receives input from the vision system. The input triggers a supra-threshold activity pattern that in turn starts to drive through the pre-defined excitatory connections the corresponding population in u_{pa} . The activity in u_{pa} continues to grow due to the excitation from u_{pr} , the vision input and the recurrent interactions within the population. The inhibitory feedback connections cause in turn a decrease of the activity in u_{pr} to resting level, resulting in a transient activity pattern in layer u_{pr} . Also the activity in u_{pa} decreases to some extent over time. However, the recurrent interactions within the populations are strong enough to self-stabilize the activity well above resting level, thus establishing a memory function. The learning threshold λ defines the time window $[t_3, t_5]$ of high activation in which new associations can be established to sufficiently active populations ($u_{pr} > \beta$) representing newly demonstrated subgoals.

For executing the sequential task, the robot has to switch from an observational learning mode to an active mode in which at each stage of the construction the robot generates itself a prediction about possible subgoals. To realize this switch, we exploit the finding in neurophysiological and computational studies showing that a task-dependent change in baseline acti-

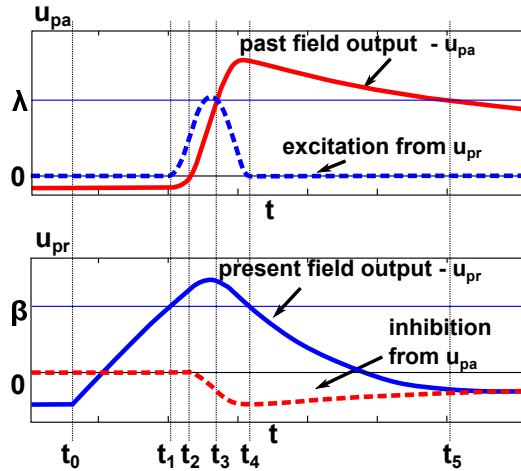


Fig. 2. Time course of activity (solid lines) and input (dashed lines) of two interconnected populations in u_{pr} and u_{pa} encoding the same subgoal.

vity may have a profound influence on the integrative properties of neuronal populations.^{24,25} A higher resting level in the active mode makes it possible that the visual input alone drives the populations in u_{pa} to a persistent state. Through the newly established connections, populations in layer u_{pr} become subsequently activated above resting level. Competition between different subgoal representations mediated by lateral inhibition ensures that only one population reaches a suprathreshold activation level at the same time. In the real-world robotics implementations, the robot verbally communicates this prediction about the next subgoal and the human teacher gives immediate positive or negative feedback. Negative feedback signalling a prediction error (e.g., a premature execution of a certain subgoal) shapes the learning process in two ways. First, the verbal input together with input from the active population in u_{pr} creates a self-stabilized activity pattern of a specific population in the task monitoring layer (TML). We generalize here the notion of a comparison between population representations relevant for correct task execution implemented in our previous dynamic field model of natural human-robot interactions.¹¹ A “generic” error detection system seems to be sensible not only to internal but also to external sources such as feedback.²⁶ Suprathreshold activity in TML causes an adaptation of the gating threshold λ for Hebbian learning during new demonstrations of the task.¹³ Functionally, a lower threshold increases the time window for learning in u_{pa} , allowing the formation of associations to several subsequent

subgoal representations in u_{pr} .

A second important role of the prediction failure at the beginning of the learning is that it introduces a breakpoint in the linear processing of the sequential task. It has been hypothesized that this breakpoint might be exploited by the learning system to “chunk” groups of subgoals together to form a higher-level subtask representation.^{8,14,27} Such a strategy might provide the basis of hierarchical processing. We take inspiration from models of cognitive control that attribute different levels of temporal abstraction in sequential actions to different areas in the frontal lobe.^{3,28} In the model, the population activity in TML enables the spontaneous creation and maintenance of a suprathreshold activation pattern of a population in the higher cognitive memory layer (HCML) by providing homogeneous excitatory input to this population. The pattern becomes meaningful during the learning process since the Hebbian principle establishes connections to all subgoal representations defining the subtask. The process of creating the higher level cognitive memory stops when the sustaining homogeneous input from TML disappears. This happens to occur when positive verbal feedback during task execution destabilizes the population representation of the initial prediction error.

2.1. *Mathematical details*

Each layer of the model is formalized by a DNF. The temporal evolution of activity, $u(x_i, t)$, of a neuron at position x in field i is governed by a particular form of a dynamic neural field:

$$\tau_i \frac{\partial u(x_i, t)}{\partial t} = -u(x_i, t) + \int w_i(x_i - x') f_0[u(x', t)] dx' + h_i + S(x_i, t) \quad (1)$$

where $\tau_i > 0$ is the time constant of the dynamics and $h_i < 0$ defines the baseline level to which field activity relaxes without external input. The integral term describes the intra-field interactions to which only sufficiently active neurons contribute. The non-linear transfer function f_0 is chosen as the Heaviside step function with threshold 0.

The input $S(x_{pa}, t)$ to field $u(x_{pa}, t)$ is given by:

$$S(x_{pa}, t) = C_{pa} I(x, t) + \int f_\beta[u(x_{pr}, t)] G(x_{pr}, x_{pa}) dx_{pr} \quad (2)$$

where $I(x, t)$ represents Gaussian input provided by vision system with a strength parameter C_{pa} . The integral term represents the summed input from $u(x_{pr}, t)$ mediated by pre-defined excitatory connections $G(x_{pr}, x_{pa})$

between corresponding populations. The threshold function $f_\beta[u(x_{pr}, t)]$ ensures that only activity above β is propagated. For the field $u(x_{pr}, t)$ the input $S(x_{pr}, t)$ is given by:

$$S(x_{pr}, t) = C_{pr}I(x, t) + \int f_0[u(x_{pa}, t)] J(x_{pa}, x_{pr}) dx_{pa} + \int f_0[u(x_{pa}, t)] a(x_{pa}, x_{pr}, t) dx_{pa} \quad (3)$$

where $S(x_{pr}, t)$ consists of three components: the input from the vision system $I(x, t)$, the input from $u(x_{pa}, t)$ mediated by the pre-defined inhibitory connections $J(x_{pa}, x_{pr})$ and the summed activity from populations in $u(x_{pa}, t)$ mediated by the adaptive excitatory connections $a(x_{pa}, x_{pr}, t)$ established during the learning process. The mathematical formulation of the learning rule for setting the synaptic connections $a(x_{pa}, x_{pr}, t)$, between any two sufficiently active neurons $x_{pa} > \lambda$ and $x_{pr} > \beta$ is defined by:

$$\tau_a \frac{\partial a(x_{pa}, x_{pr}, t)}{\partial t} = f_\beta[u(x_{pr}, t)] f_\lambda[u(x_{pa}, t)] e(x_{pr}, t) \quad (4)$$

It represents a supervised Hebbian Learning rule (see Ref. 29) where the parameter τ_a defines the time scale of learning and the error term $e(x_{pr}, t)$ describes the difference between the desired active state of a neuron in $u(x_{pr}, t)$ and the summed input from all connected neurons in $u(x_{pa}, t)$:

$$e(x_{pr}, t) = f_\beta[u(x_{pr}, t)] - \int f_\lambda[u(x_{pa}, t)] a(x_{pa}, x_{pr}, t) dx_{pa} \quad (5)$$

3. Modeling Results

The main objective of the present modeling work is to show that the neuro-plausible learning mechanisms in a social teaching context are in principle sufficient to endow the robot ARoS with the capacity to learn and represent generalized task knowledge in a hierarchical organization. For our learning by demonstration approach it is important to stress that ARoS already has the necessary perceptual and motor capacities (for details see Refs. 10 and 11): the different reach-grasp-place sequences to achieve each individual assembly step are in its motor repertoire, the vision system is trained to recognize the outcome of each assembly step, and a speech synthesizer/recognizer system endows the robot with elementary verbal communication skills.

ARoS has to learn the construction of a toy vehicle consisting of a bottom platform (BP) with an axle on which two wheels (LW, RW) have first

to be mounted and then fixed with nuts (LN, RN). Four columns (C1, C2, C3, C4) have to be inserted into specific holes on the bottom platform before a top platform (TP) can be attached (Fig. 3). The human teachers are instructed to show in each demonstration the whole sequential task, but are allowed to vary the sequential order. As a concrete example, we focus in the present modeling work on learning a specific subtask: C1 to C4 may be inserted in any order, but TP requires the presence of all columns.

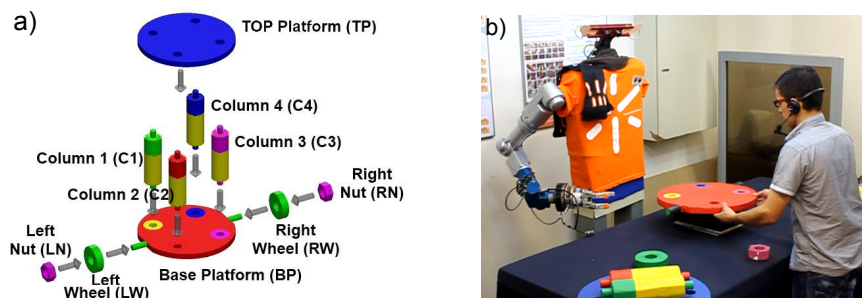


Fig. 3. (a) Toy vehicle (b) Robot ARoS interacting with a human teacher

3.1. Demonstrating of two sequential orders

In the first experiment, two teachers demonstrate the assembly task using two sequences of subgoals with the ordinal position of C3 and C4 interchanged. Each of the two sequences were demonstrated 15 times in alternating order. The recall trials start with the presentation of BP by the teacher and the vision input activates the respective population in u_{pa} . In line with the training sequences, the model predicts C1 as the next subgoal. After having inserted C2, the robot will predict C3 and C4 as equally likely future assembly steps. Small differences in the learned weights and the population activity due to noise in the system gives preference to one of the options. In the example of Fig. 4, the population representing C4 becomes suprathreshold. In any case, the robot always predicts the last missing column as new subgoal before trying to place the top platform. This successful imitation of the demonstrated sequences does not mean, however, that the robot has already understood the causal relationship between the columns and the top platform. The following experiment shows this.

3.2. Demonstration of a new sequential order

Now the new subsequence $BP \rightarrow C4 \rightarrow C3 \rightarrow C1 \rightarrow C2 \rightarrow TP$ is demonstrated in alternation with the two previously demonstrated sequential orders (15 presentations for each order). During task execution, the robot follows the new sequential order. However, after having correctly predicted

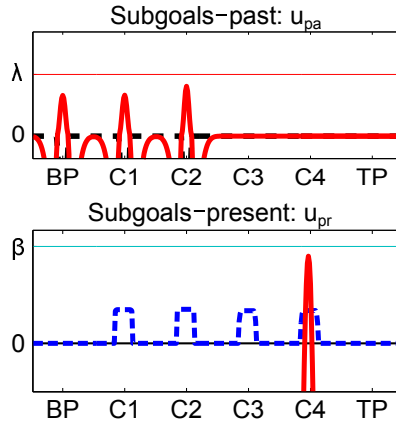


Fig. 4. Population activity (solid lines) in the past layer u_{pa} and in the present layer u_{pr} . The dashed line in u_{pr} represents the input from active populations in u_{pa} through learned connections $a(x_{pa}, x_{pr}, t)$. The robot makes the prediction to insert C4 after having finished subgoals BP, C1 and C2.

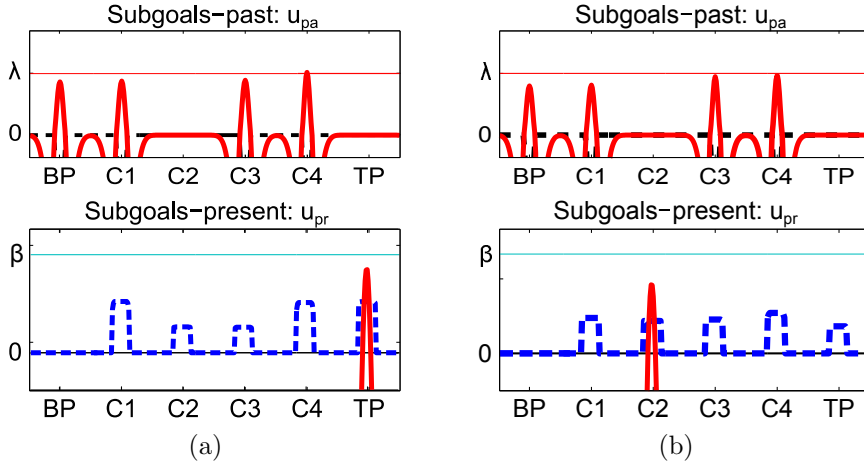


Fig. 5. (a) Wrong prediction of TP given that C2 is still missing. (b) Correct prediction of C2 following the learning of longer-term dependencies during new demonstrations.

and inserted C4, C3 and C1, the robot makes the prediction to attach TP instead of inserting C2 (Fig. 5(a)). This premature choice is a direct consequence of the simple associative chaining that the model has created during task demonstration. The TP population gets direct input from the population C4 and C3 (sequences 1 and 2) whereas C2 is supported only by population C1 (sequence 3). In response to the negative verbal feedback, a self-sustained activation peak in TML evolves (not shown) representing the prediction error in the self-generated sequence. This activation is assumed to lower the adaptive learning threshold λ in u_{pa} by a fixed amount,¹³ thus increasing the time window for associative learning. As a consequence, connections to several subsequent subgoal representations may be learned during new demonstration trials. The threshold adaptation in response to negative feedback may continue for several demonstration-execution iterations until positive feedback indicates a successful execution of the whole subsequence. This is shown in Fig. 5(b) where now C2 follows C1, C3 and C4. Note that also TP continues to get input from active populations in u_{pa} , which is however weaker compared to the input to C2. The suprathreshold activity in TML also triggers and sustains the activation of a pool of neurons in HCML (compare the time course in Fig. 6(a) on top). During execution trials, subgoal representations in the past layer u_{pa} become associated with this highly active population. When the prediction error is corrected, that is, the activity in TML has decayed to resting level, all populations defining the subtask of finishing the base with 4 columns (TB plus C1 to C4) are active and thus drive the population in HCML. This is shown in Fig. 6(b) when at the time of placing C2, suprathreshold activity in HCML evolves. Note the delay of temporal evolution compared to the time course of suprathreshold activity of population C2. This model prediction is in line with experimental findings in sequence learning studies showing that it takes some extra time to mentally mark the boundaries of the higher level subtask when the associated last subgoal has been achieved.²⁷

4. Discussion

Sequentially organized human activity, such as assembling an object, can be identified at a range of different time-scales, from fine-grained to coarse-grained. The segmentation on different temporal grains reflects a hierarchical organization of behavior where goals tend to be satisfied by the recursive satisfaction of subgoals. Recent experimental evidence from observational learning studies suggests that being able to encode the hierarchical structure of observed activity strongly promotes the successful imitation of new

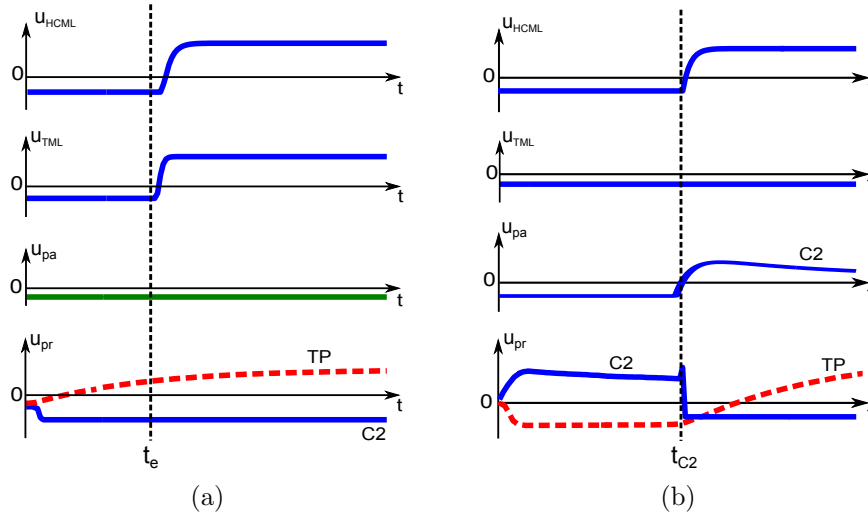


Fig. 6. Time course of activity in the various model layers (a) Error trial: Population TP in u_{pr} becomes active before the last column (C2) has been inserted. At time t_e , the human gives negative feedback to this prediction, triggering the suprathreshold activity first in TML and subsequently in HCML. (b) Correct trial: After learning, the system correctly recalls C2 before TP. C2 is inserted at time t_{C2} , TP is predicted and the HCML population indicating the completion of the subtask becomes activated.

action sequences (e.g. Refs. 7 and 27). Learning by observation and imitation has also attracted a lot of attention over the last couple of years in the robotics community since it represents an intuitive and efficient way how ordinary people could teach a robot new skills and tasks.^{9,30} In this paper we have addressed the crucial question, in a learning by observation paradigm, of how an agent with no prior task knowledge (e.g., a robot or a child) may learn to segment the continuous flow of sequentially structured behavior into a goal-subgoal hierarchy. In line with recent experimental evidence, the dynamic field model implements the idea that prediction errors play a crucial role for grouping fine-grained events (e.g., individual assembly steps) into larger units (e.g., the subtask of inserting 4 columns). An adaptation of the learning threshold triggered by the prediction error updates the long-term sequential dependencies (stored in the learned inter-field connections) necessary to achieve accurate prediction. The mechanism is similar to the gating signal that Reynolds et al.¹⁵ have used in their connectionist network model of event segmentation to learn and update internal context information reflecting event knowledge. They applied their segmentation model to perform one-step predictions of human movement patterns. This

suggests that prediction errors might be sufficient to learn part-subpart structures simultaneously on different levels of the action hierarchy. In the social learning situation of the robotics experiments, the human teacher gives immediate binary feedback when the robot predicts the next subgoal of the assembly task. This greatly improves the efficiency of the learning compared to reinforcement-based mechanisms since the credit-assignment problem in complex sequences is avoided. Interestingly, integrating the notion of temporal abstract actions associated with (pseudo-)rewarded subgoals into the reinforcement learning framework can significantly speed up learning (for discussion see Ref. 3).

A unique feature of the model is that the time window for Hebbian learning is defined by the time course of the population activity above learning threshold. This means that only the initial stage of working memory (WM) maintenance of already accomplished subtasks contributes to successful long-term memory (LTM) formation of sequential order. This model prediction is in line with recent findings in combined behavioral and brain imaging studies of brain circuits commonly believed to support the transition from WM and LTM.³¹

The capacity to form hierarchies of sequential behavior is not only important for successful imitation. Compact representations of temporally extended activity are also beneficial to plan future actions and to coordinate with others in joint action tasks.¹¹ In the present implementation, the population activity in HCML encodes that a certain subtask has been completed. We are currently working on an extended version of the learning model which basically copies the two layered structure of the subgoal level to allow a sequential activation of future subtask representations.

Acknowledgements

Work supported by: (1) Portuguese FCT Grant SFRH/BD/48529/2008, financed by POPH-QREN-Type 4.1-Advanced Training, co-funded by the European Social Fund and national funds from MEC; (2) FEDER Funds through Competitiveness Factors Operational Program - COMPETE and National Funds by FCT Portuguese Science and Technology Foundation under the Project FCOMP-01-0124-FEDER-022674. (2) Project NETT: Neural Engineering Transformative Technologies, EU-FP7 ITN proj. nr. 289146

References

1. K. Lashley, The problem of serial order in behavior, in *Cerebral Mechanisms in Behaviour*, ed. L. A. Jeffres (Wiley, New York, 1951) pp. 112–142.

2. C. A. Kurby and J. M. Zacks, *Trends in Cognitive Sciences* **12**, 72 (2008).
3. M. M. Botvinick, *Trends in Cognitive Sciences* **12**, 201 (2008).
4. R. P. Cooper and T. Shallice, *Psychological Review* **113**, 887 (2006).
5. M. M. Botvinick and D. C. Plaut, *Psychological Review* **111**, 395 (2004).
6. L. L. Travis, Goal-based organization of event memory in toddlers, in *Developmental spans in event comprehension and representation: Bridging fictional and actual events*, (Lawrence Erlbaum Associates, Inc, 1997) pp. 111–138.
7. A. Whiten, E. Flynn, K. Brown and T. Lee, *Developmental Science* **9**, 574 (2006).
8. J. M. Zacks, C. A. Kurby, M. L. Eisenberg and N. Haroutunian, *Journal of Cognitive Neuroscience* **23**, 4057 (2011).
9. W. Erhlagen, A. Mukovskiy, E. Bicho, G. Panin, C. Kiss, A. Knoll, H. T. van Schie and H. Bekkering, *Robotics and Autonomous Systems* **54**, 353 (2006).
10. E. Bicho, W. Erhlagen, L. Louro and E. C. e Silva, *Human Movement Science* **30**, 846 (2011).
11. E. Bicho, W. Erhlagen, L. Louro, E. C. e. Silva, R. M. Silva and N. Hipólito, A dynamic field approach to goal inference, error detection and anticipatory action selection in human-robot collaboration, in *New Frontiers in Human-Robot Interaction*, eds. K. Dautenhahn and J. Sanders (John Benjamins Publishing Company, 2011) pp. 135–164.
12. A. Thomaz and C. Breazeal, *Artificial Intelligence* **172**, 716 (2008).
13. A. R. Seitz and H. R. Dinse, *Current Opinion in Neurobiology* **17**, 148 (2007).
14. K. Sakai, K. Kitaguchi and O. Hikosaka, *Experimental Brain Research* **152**, 229 (2003).
15. J. R. Reynolds, J. M. Zacks and T. S. Braver, *Cognitive Science* **31**, 613 (2007).
16. S.-i. Amari, *Biological Cybernetics* **27**, 77 (1977).
17. W. Erhlagen and E. Bicho, *Journal of Neural Engineering* **3**, R36 (2006).
18. H. R. Wilson and J. D. Cowan, *Kybernetik* **13**, 55 (1973).
19. G. Schöner, *Cambridge Handbook of Computational Cognitive Modeling*, 101 (2007).
20. W. Erhlagen and G. Schöner, *Psychological Review* **109**, 545 (2002).
21. J. Tanji, K. Shima and H. Mushiake, *Trends in cognitive sciences* **11**, 528 (2007).
22. A. Genovesio, P. J. Brasted and S. P. Wise, *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **26**, 7305 (2006).
23. D. O. Hebb, *The organization of Behavior* (Lawrence Erlbaum Associates Inc, Cambridge, MA, USA, 1949).
24. E. Salinas, *Neural Computation* **15**, 1439 (2003).
25. W. Asaad, G. Rainer and E. Miller, *Journal of Neurophysiology* **84**, 451 (2000).
26. W. H. R. Miltner, C. H. Braun and M. G. H. Coles, Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a generic neural system for error detection (1997).
27. B. M. Hard, S. C. Lozano and B. Tversky, *Journal of Experimental Psychol-*

- ogy. General* **135**, 588 (2006).
28. J. Fuster, *International Journal of Psychophysiology* **60**, 125 (2006).
 29. P. Dayan and L. Abbott, Plasticity and learning, in *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, (MIT Press, Cambridge, MA, USA, 2001)
 30. C. L. Dautenhahn, Kerstin and Nehaniv (ed.), *Imitation in Animals and Artifacts* (MIT Press, Cambridge, MA, USA, 2002).
 31. C. Ranganath, M. X. Cohen and C. J. Brozinsky, *Journal of Cognitive Neuroscience* **17**, 994 (2005).