Towards Endowing Collaborative robots with fast learning for minimizing Tutors' demonstrations: What and When to do?[*]

Ana Cunha[1,3], Flora Ferreira[2], Wolfram Erlhagen[2], Emanuel Sousa[1], Luís Louro[3], Paulo Vicente[3], Sérgio Monteiro[3], and Estela Bicho[3] ✉

[1] Center for Computer Graphics, University of Minho, 4800-058 Guimaraes, Portugal
anaisabel.7@hotmail.com,
[2] Dept. of Mathematics and Applications, Center of Mathematics, University of Minho, 4800-058 Guimaraes, Portugal
flora.ferreira@gmail.com,
[3] Dept. Industrial Electronics, Algoritmi Center, University of Minho, 4800-058 Guimaraes, Portugal
estela.bicho@dei.uminho.pt

Abstract. Programming by demonstration allows non-experts in robot programming to train the robots in an intuitive manner. However, this learning paradigm requires multiple demonstrations of the same task, which can be time-consuming and annoying for the human tutor. To overcome this limitation, we propose a fast learning system – based on neural dynamics – that permits collaborative robots to memorize sequential information from single task demonstrations by a human-tutor. Important, the learning system allows not only to memorize long sequences of sub-goals in a task but also the time interval between them. We implement this learning system in Sawyer (a collaborative robot from Rethink Robotics) and test it in a construction task, where the robot observes several human-tutors with different preferences on the sequential order to perform the task and different behavioral time scales. After learning, memory recall (of what and when to do a sub-task) allows the robot to instruct inexperienced human workers, in a particular human-centered task scenario.

Keywords: Industrial Robotics, Assembly Tasks, Learning from Demonstration, Sequence Order and Timing, Rapid Learning, Dynamic Neural Fields

1 Introduction

One of the current challenges of the Industry 4.0 Era is the implantation of collaborative robots able to work side by side with (different) human-operators

---

in a large variety of tasks [11]. This implies that this new generation of robots (CoBots) have to master a wide variety of tasks and interaction scenarios that cannot be completely designed in advance by experts as in traditional industrial applications. For symbiotic and efficient human-robot collaboration in (sequential or supportive) tasks, it is fundamental to endow the robotic co-workers with cognitive and learning capabilities [14][3]. In this paper, we focus on the implementation of learning mechanisms that allow the robotic co-worker to acquire high-level knowledge about the sequential structure of multi-part assembly/disassembly tasks (which may have time constraints) without being explicitly programmed. We adopt the learning paradigm known as programming by demonstration/observation since it allows non-specialists in robot programming to train the robot in an intuitive and open-ended manner [9]. However, often this learning paradigm requires multiple demonstrations of the same task which can be time-consuming and annoying for the human tutor [10]. Thus, for user acceptance, it is crucial to make possible to the robot the acquisition of generalized task knowledge in very few demonstrations. With this in mind, we implement and test on the collaborative robot Sawyer (from Rethink Robotics) a computational model that integrates fast activation-based learning to robustly represent sequential information from single task demonstrations by a human-tutor. Important, this learning system allows not only to memorize long sequences of sub-goals in a task, but also the time interval between them. The integration of these two features – ordinal and temporal information – allows the robot to memorize in one shot 'what to do' and 'when to do' in a certain task scenario, which builds a fast memory mechanism that significantly reduces the number of demonstrations needed from a tutor. After learning, the autonomous reactivation of this memory can be used by the robot to instruct inexperienced workers, or to make decisions when it plays the role of an active assist/co-worker. To build this fast learning system, we apply the theoretical framework of dynamic neural fields (DNFs), that has been proven to provide key processing mechanisms for applications in cognitive robotics ([4],[13]) including robot learning ([16],[5]). As a specific task example we consider a thrusters/pipes assembly task. One or more human tutors show the robot Sawyer the assembly work consisting of a series of assembly steps necessary to construct the structure from its parts/thrusters (Fig. 1). Different tutors may have different preferences in the sequential order of assembly steps, and may act with different time scales. The remainder of the paper is structured as follows: section 2 describes the construction task and the robotic platform Sawyer; Section 3 contains the description of the DNFs based learning model; section 4 presents experimental results; and the paper ends with a discussion and future work in section 5.

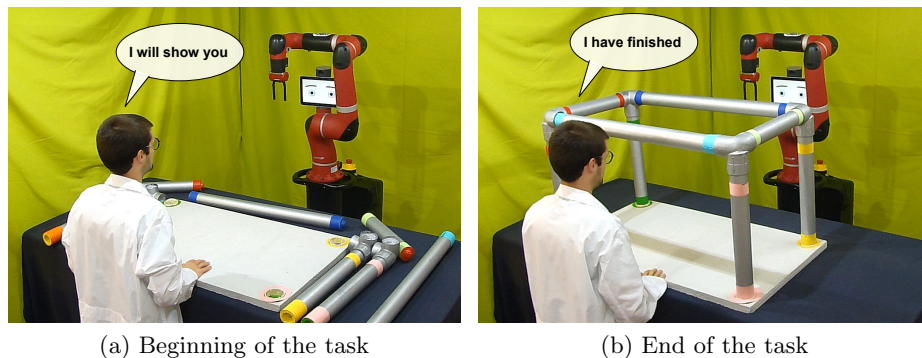(a) Beginning of the task   (b) End of the task

Fig. 1: Illustration of the construction task scenario: a tutor collaborates with the robotic co-worker in assembling a structure composed of eight colored thrusters

2   Experimental Setup

For the experiments we used the robot Sawyer (displayed in figure 2d) designed by the company "Rethink Robotics" to execute collaborative tasks [12]. Sawyer features a 7 degrees of freedom robot arm with 1.26 meters reach and its "head" is the LCD display that sits on top. Sawyer displays different eyes movements in a familiar way which contributes to its human-friendly design. The robot is also equipped by two cameras, one located in the head and another in the arm. The information about object color/type is provided by the head camera system. As a test scenario, we used a task of building a structure of eight thrusters/pipes, where the insertion of each one of the thrusters corresponds to one sub-task. We considered three different Layouts that imply three different assembly sequences and two tutors with different behavioral time scales. The disposal of the thrusters in each scenario can be seen in Fig. 2. A tutor demonstrates how to assemble the thrusters while the robotic co-worker observes and memorizes the serial order and timing of each assembly step. Later, the robot acts as a tutor and recalls the memorized task to a different operator, respecting the order and time interval of each step. Depending on how the thrusters are disposed on the workplace, different sequences can be used to assemble all the parts, which will require the robotic platform to learn multiple possible sequences to build the structure. Moreover, different tutors will assemble the sequence with different time intervals: for example, an older tutor may take longer to reach and insert all the thrusters/parts than a younger one. The speech synthesizer allows the robot to communicate the result of its decision process to the human co-worker.
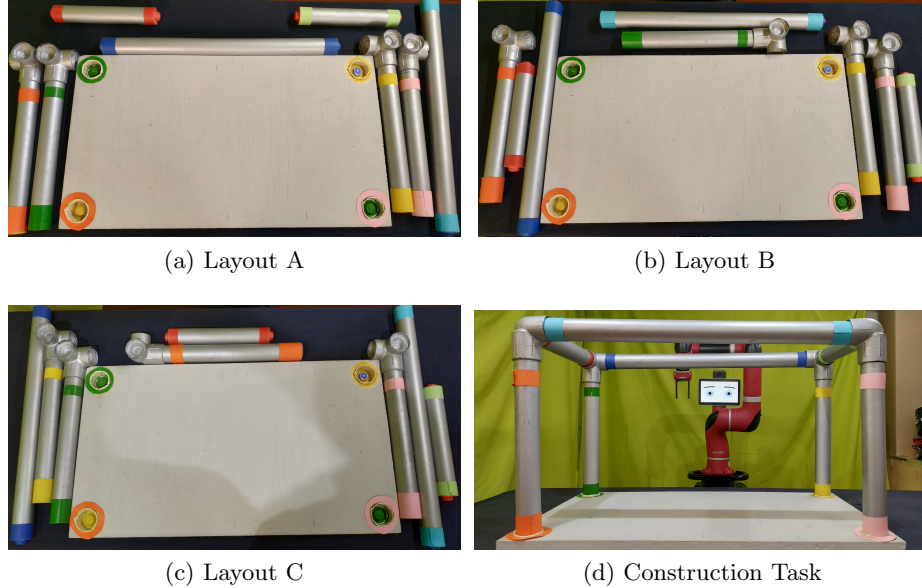
(a) Layout A

(b) Layout B

(c) Layout C

(d) Construction Task

Fig. 2: Construction task and Layout scenarios used during the experiments

3   Model Description

The model presented in this paper is based on previous research on natural human-robot interaction [2], [16], [17] based on Dynamic Neural Fields (DNFs). DNFs provide a rigorous theoretical framework to implement neural computations that endow a robot with crucial cognitive functions such as working memory, prediction and decision making [15]. DNFs are formalized by nonlinear integro-differential equations in which activity of neurons is summarized into the activity function $u(x,t)$, which can be used to reduced computational complexity and can be mathematically analyzed. The concept behind dynamic field models is that task-relevant information is expressed by supra-threshold bumps of neural populations where each bump represents a specific action or sub-task. Input from external sources, such as vision, causes activation in the correspondent populations that remain active with no further external input due to recurrent excitatory and inhibitory interactions within the populations. Those interactions are able to hold an auto-sustained multi-bump pattern which can be turned into a memory mechanism for order and time interval of sequential processes ([6], [8], [17]).

Fig. 3 presents the learning by demonstration of several sequence memory fields. Each Sequence Memory field (u_{SM}) stores a sequence of stimulus events as a multi-bump pattern. A bump represents an event triggered through excitatory sensory input. The strength of each memory representation reflects the time elapsed since stimulus presentation, resulting in an activation gradient from the
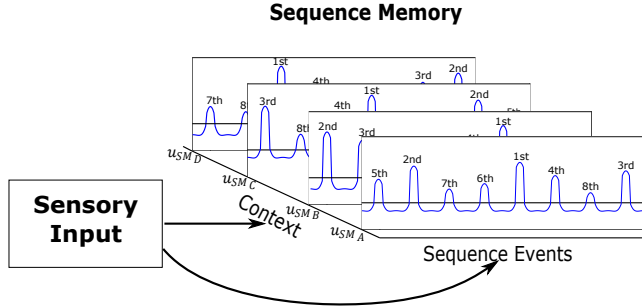
Sequence Memory

Fig. 3: Sketch of sequence learning process.

first to the last event [6]. According to the context (e.g. position of the objects in the workplace, characteristics of the co-worker), several sequences with different order and timing can be memorized.

Fig. 4 depicts an overview of the sequence recall process. Taking into account a specific context, a Sequence Memory field (u_{SM}) is chosen from the set of stored fields. The Sequence Recall field (u_{SR}) receives the multi-bump pattern of u_{SM} as subthreshold input. During sequence recall, the continuous increase of the baseline activity in u_{SR} brings all subpopulations closer to the threshold for the evolution of self-stabilized bumps. When the currently most active population reaches this threshold, the corresponding sensory output is triggered. At the same time, the excitatory-inhibitory connections between associated populations in u_{SR} and the Past Events field (u_{PE}) guarantee that the suprathreshold activity representing the latest sequence event becomes first stored in u_{PE} and subsequently suppressed.
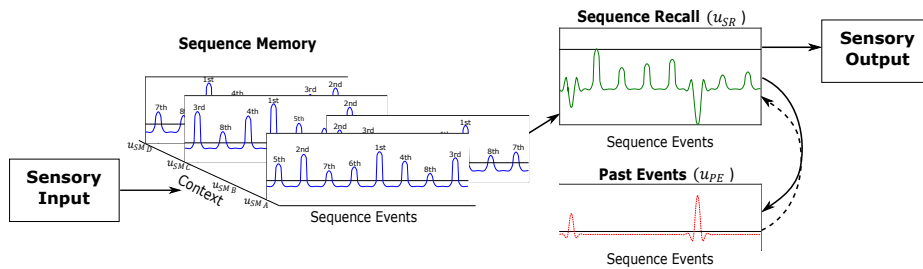

Fig. 4: Sketch of sequence recall process. Dashed lines indicate inhibitory connections, solid lines excitatory connections.

The dynamics of each Sequence Memory field u_{SM}, the Sequence Recall field u_{SR} and the Past Events field u_{PE} are governed by the following equations,

respectively [1] [6]:

$$\tau_{SM}\frac{\partial u_{SM}(x,t)}{\partial t} = -u_{SM}(x,t) + \int w(x-y)f\left(u_{SM}(y,t)\right)\mathrm{d}y \qquad (1)$$
$$+s(x,t) + h_{SM}(x,t)$$

$$\tau_{SR}\frac{\partial u_{SR}(x,t)}{\partial t} = -u_{SR}(x,t) + \int w(x-y)f\left(u_{SR}(y,t)\right)\mathrm{d}y \qquad (2)$$
$$-\int w(x-y)f\left(u_{PE}(y,t)\right)\mathrm{d}y + u_{SM}(x) + h_{SR}(t)$$

$$\tau_{PE}\frac{\partial u_{PE}(x,t)}{\partial t} = -u_{PE}(x,t) + \int w(x-y)f\left(u_{SM}(y,t)\right)\mathrm{d}y \qquad (3)$$
$$+u_{SR}(x,t)f(u_{SR}(x,t)) + h_{PE}$$

where $u_{SM}(x,t)$, $u_{SR}(x,t)$ and $u_{PE}(x,t)$ represents the activity at time t of a neuron tuned to the feature value x. The parameters $\tau_{SM},\tau_{SR},\tau_{PE} > 0$ define the time scale of each field. The connection function $w(x)$ determines the coupling between neurons within the field and to enable multi-bump solutions is use the following function [7]:

$$w(x) = Ae^{-b|x|}\left(b\sin|\alpha x| + \cos(\alpha x)\right), \qquad (4)$$

where $b > 0$ determines the rate at which the oscillations in w decay with distance and $A > 0$ and $0 < \alpha \le 1$ control the amplitude and the spatial phase of w, respectively. The function $s(x,t)$ represents the time dependent localized input at site x from the sensory input ($s(x,t) > 0$ when the encoded variable has an excitatory input, and $s(x,t) = 0$ otherwise). The strength of individual memory representations in u_{SM} is controlled by the baseline dynamics $h_{SM}(x,t)$:

$$\frac{\partial h_{SM}(x,t)}{\partial t} = (1 - f(u_{SM}(x,t)))\left(-h_{SM}(x,t) + h_{SM_0}\right) + \frac{1}{\tau_{h_{SM}}}f(u_{SM}(x,t)) \quad (5)$$

where $h_{SM_0} < 0$ defines the level to which h_{SM} converges without suprathreshold activity at position x and $\tau_{h_{SM}}$ measures the growth rate when it is present. The baseline activity $h_{SR}(t)$ evolves continuously in time described by the equation:

$$\frac{\partial h_{SR}(t)}{\partial t} = \frac{1}{\tau_{h_{SR}}}, h_{SR}(t_0) = h_{SR_0} < 0 \qquad (6)$$

where $\tau_{h_{SR}}$ controls the growth rate of h_{SR}. The baseline activity $h_{PE} < 0$ is constant. $f(x)$ is the output function of the neuron and is taken as the Heaviside step function with threshold 0.

4   Experimental Results

In this Section, we describe the experimental results of learning the sequential order and timing of assembling the structure displayed in Fig. 2d, using the implemented model described in Section 3. Two tutors with different behavioral time scales, who already had experience with the construction task, were asked to build the structure starting from three different Layouts (A, B and C, Fig. 2). At the same time, the robotic co-worker Sawyer pays attention to the tutor and stores the sequential order and time interval of each step, for all demonstrations. Afterward, the robotic platform acts as a tutor and teaches two other inexperienced workers with no knowledge of the construction task. From the memorized sequences, the system selects the most suitable one according to the characteristics of the new worker and the distribution of the pieces in the table. The robot recalls the selected sequence, verbalizing the color of the piece that should be inserted, according to the order and time stored in the Sequence Memory field during the demonstration trials.

4.1   Learning the Sequence Order and Timing of the Assembly Task

During the demonstration period, when the system detects that one of the colored thrusters is going to be inserted, an input stimulus is generated in the location of the population of neurons encoding the respective colored thruster, which leads to activation in the u_{SM} field, forming a bell-shaped bump that grows gradually as a function of time. Fig. 5 pictures the demonstration experiment performed, where both tutors show to the robotic co-worker how to build the structure. In the scenario illustrated in Fig. 5a and 5b, the colored thrusters are distributed in the work table according to Layout A. A video of this trial can be found in the following link: https://youtu.be/YTZTDJzzGYw. Next, Fig. 5c and 5d are snapshots of the Sequence Memory fields stored during the demonstration, where the amplitude of the bumps encodes the serial order of the inserted thrusters, with the highest peak (orange) being the color of the first inserted piece and the shortest one (light blue) being the last. Although the serial order used by both tutors is the same, the intervals of time between each assembly step in both trials are considerably different, which will be demonstrated in Section 4.2.
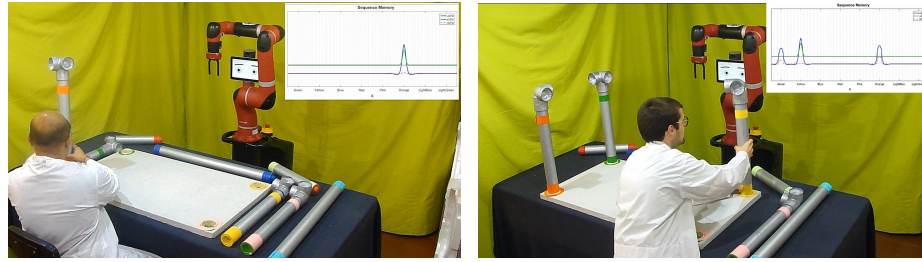
(a) **Older tutor** inserting the Orange thruster (1$^{\text{st}}$ piece)

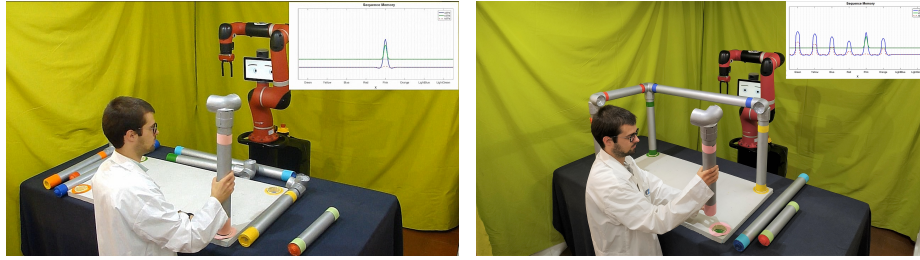(b) **Younger tutor** inserting the Yellow thruster (3$^{\text{rd}}$ piece)
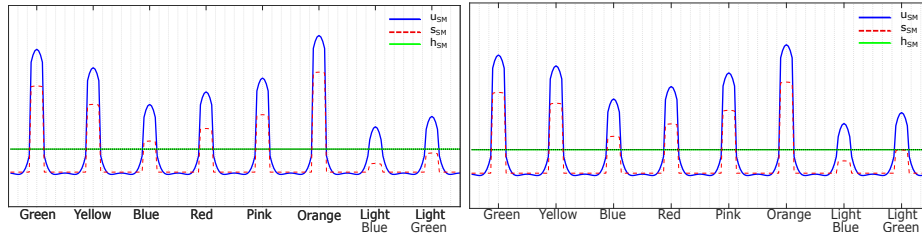
(c) Sequence Memory Field A demonstrated by the **Older tutor**
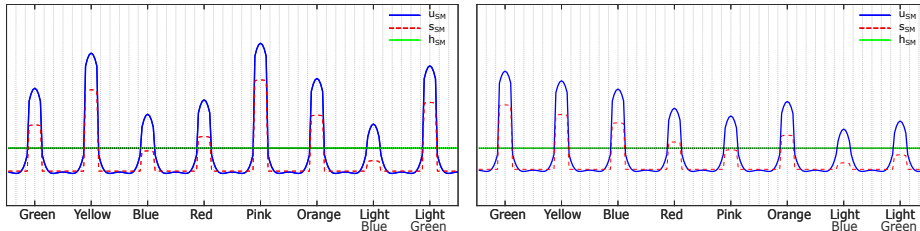

(d) Sequence Memory Field A demonstrated by the **Younger tutor**

Fig. 5: Two different tutors assembling **Sequence A**: Orange → Green → Yellow → Pink → Red → Blue → Light Green → Light Blue

Subsequently, Fig. 6 portraits the demonstration of two other sequences, this time by the younger tutor. In Fig. 6a, the parts are organized in the scenario according to Layout B, so the tutor starts the construction by inserting the pink thruster, opting for a different assembly sequence. A complete video of this trial can be found at https://youtu.be/0F8T_d_y2xs. Similarly, in the scenario displayed in Fig. 6b, the tutor was asked to assemble the structure starting from Layout C, which resulted in a new sequence. After the demonstration of each trial, the information acquired in both fields - Sequence Memory B (Fig. 6c) and Sequence Memory C (Fig. 6d) - is stored in the model, so later the robotic platform can use the memorized information to instruct inexperienced workers, taking into account the Layout of the parts that constitute the sequence.

(a) Younger tutor assembling **Sequence B**, starting from the Pink thruster (1$^{\text{st}}$ piece).

(b) Younger tutor assembling **Sequence C**, inserting the Pink thruster (6$^{\text{th}}$ piece)


(c) **Sequence Memory Field B:** Pink → Yellow → Light Green → Orange → Green → Red → Blue → Light Blue

(d) **Sequence Memory Field C**: Green → Yellow → Blue → Orange → Red → Pink → Light Green → Light Blue

Fig. 6: Younger tutor assembling **Sequence B and C**

4.2 Recalling the Memorized Sequences

In order to verify if the robotic co-worker was able to memorize not only the sequence but also the time interval between each construction step, two workers with no previous knowledge of the task were asked to collaborate with Sawyer and follow its instructions to learn the assembly steps of the sequence A, which was previously demonstrated by the two previous tutors, with different time intervals. The system was programmed to verbalize each construction step, taking into consideration the sequence and timing between each insertion.

Fig. 7a and 7b picture the robotic system as a tutor, guiding two different workers through the construction sub-tasks of the structure, while Fig. 7c and 7d illustrate the time course of the maximal activation of each sub-neuronal population when the sequence was assembled by each worker, according to the instructions given by the robot. A video example can be found at https://youtu.be/Vn0_1raKq4I.
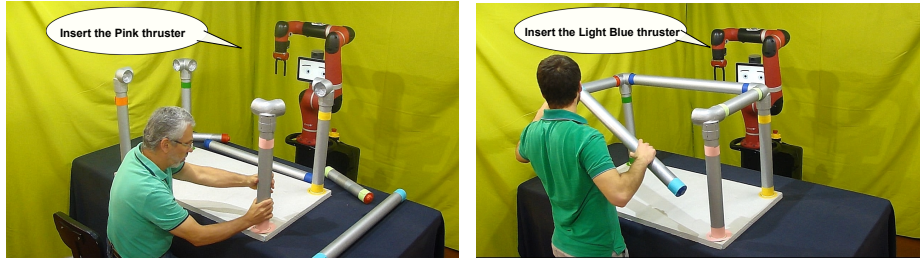
(a) **Older worker** following Sawyer's instruction to assemble Sequence A

(b) **Younger worker** following Sawyer's instruction to assemble Sequence A
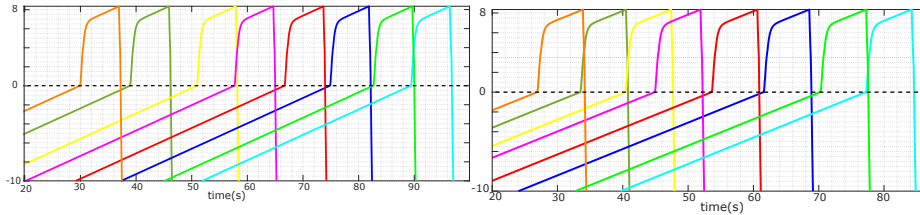
(c) Older worker: Total time of 67s **(slower)**

(d) Younger worker: Total time of 58s **(faster)**
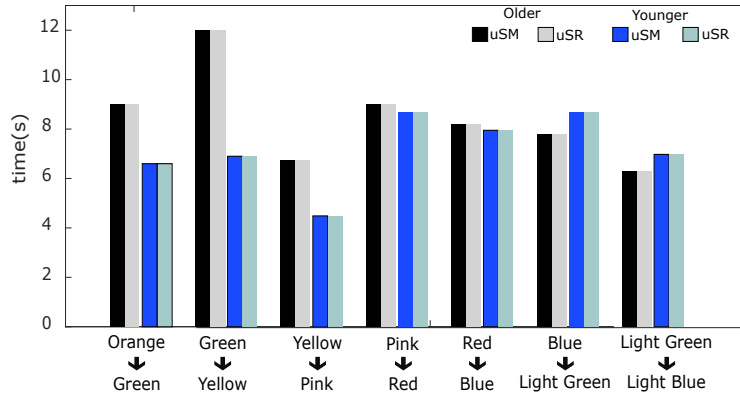
(e) Contrast between the time intervals of consecutive assembly steps during the construction of the sequence A, performed by an older (slower) and a younger (faster) worker

Fig. 7: Co-worker Sawyer recalls the memorized sequence, respecting the time interval of each demonstration

Each instruction is verbalized when each sub-neuronal population encoding the corresponding colored thruster in the Sequence Recall field (u_{SR}) reaches the threshold level, as stated in Section 3. As can be observed in both figures, all sub-neuronal populations seem to have a pre-activation strength that respects

the temporal order of the sequential task, learned during the demonstration trials. The first thrusters (orange) are inserted at t=30s (older tutor trial) and t=27s (younger tutor trial), while the last ones (light blue) are placed at t=97s (older tutor trial) and t=85s (younger tutor trial). By comparing both trials (Fig. 7e), we can observe that the older worker was slower than the younger in the majority of the steps and the younger worker took less time to perform the complete task, as it should be expected since the workers are following the order and time memorized in the previous demonstrations (Fig. 5).

5   Discussion

In this paper, we have proposed and tested, on a collaborative robot, a rapid learning system that allows the robot to memorize knowledge about ordinal and temporal aspects of sequential tasks in a learning by demonstration paradigm. One benefit of this learning system is that a single demonstration is sufficient, thus minimizing the efforts of the human tutor to train the robot. We have shown that after learning, the recall of the memorized information can be used by the robot to instruct inexperienced human operators, in the same context. This instructional process was, however, performed in open-loop. This fast learning system offers other benefits that will be explored in future work. For example, the recall of the stored information can be used as input to a long-term learning mechanism that allows the robot also to learn the connections between the several sub-tasks [16], thus extrapolating task knowledge, i.e. that a task can be performed in many different ways. As future work, the collaborative robot Sawyer can also contribute by using its robotic arm to interact with its co-worker and jointly build the construction task, reducing the workload of the human partner thus increasing the efficiency of the process.

Endowing a collaborative robot with the capacity to predict not only the ordinal sequence structure but also the time interval between successive events is central for efficient coordination of actions and decisions in space and time, in human-robot joint action tasks. It allows the robot to anticipate what the human operator will need, or will do, and when it should start an adequate complementary behavior in the service of the joint task.

References

1. Amari, S.i.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics **27**(2), 77–87 (1977). https://doi.org/10.1007/BF00337259
2. Bicho, E., Erlhagen, W., Louro, L., Costa e Silva, E.: Neuro-cognitive mechanisms of decision making in joint action: A human-robot interaction study. Human Movement Science **30**(5), 846–868 (2011). https://doi.org/10.1016/j.humov.2010.08.012
3. El Zaatari, S., Marei, M., Li, W., Usman, Z.: Cobot programming for collaborative industrial tasks: An overview. Robotics and Autonomous Systems **116**, 162–180 (2019). https://doi.org/10.1016/j.robot.2019.03.003

4. Erlhagen, W., Bicho, E.: The dynamic neural field approach to cognitive robotics. Journal of Neural Engineering **3**(3), R36–R54 (2006). https://doi.org/10.1088/1741-2560/3/3/R02

5. Erlhagen, W., Mukovskiy, A., Bicho, E., Panin, G., Kiss, C., Knoll, A., Schie, H., Bekkering, H.: Goal-directed imitation for robots: A bio-inspired approach to action understanding and skill learning. Robotics and Autonomous Systems **54**, 353–360 (2006). https://doi.org/10.1016/j.robot.2006.01.004

6. Ferreira, F., Erlhagen, W., Bicho, E.: A dynamic field model of ordinal and timing properties of sequential events. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2011). https://doi.org/10.1007/978-3-642-21738-8_42

7. Ferreira, F., Erlhagen, W., Bicho, E.: Multi-bump solutions in a neural field model with external inputs. Physica D: Nonlinear Phenomena **326**, 32–51 (2016). https://doi.org/10.1016/j.physd.2016.01.009

8. Ferreira, F., Erlhagen, W., Sousa, E., Louro, L., Bicho, E.: Learning a musical sequence by observation: A robotics implementation of a dynamic neural field model. IEEE ICDL-EPIROB 2014 - 4th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics pp. 157–162 (2014). https://doi.org/10.1109/DEVLRN.2014.6982973

9. Kyrarini, M., Haseeb, M.A., Ristić-Durrant, D., Gräser, A.: Robot learning of industrial assembly task via human demonstrations. Autonomous Robots **43**(1), 239–257 (2019). https://doi.org/10.1007/s10514-018-9725-6

10. Orendt, E., Fichtner, M., Henrich, D.: Robot programming by non-experts: Intuitiveness and robustness of one-shot robot programming. pp. 192–199 (2016). https://doi.org/10.1109/ROMAN.2016.7745110

11. Papanastasiou, S., Kousi, N., Karagiannis, P., Gkournelos, C., Papavasileiou, A., Dimoulas, K., Baris, K., Koukas, S., Michalos, G., Makris, S.: Towards seamless human robot collaboration: integrating multimodal interaction. The International Journal of Advanced Manufacturing Technology pp. 1–17 (2019). https://doi.org/10.1007/s00170-019-03790-3

12. Robotics, R.: Sawyer collaborative robot (2018), http://www.rethinkrobotics.com/sawyer/

13. Sandamirskaya, Y., Zibner, S., Schneegans, S., Schöner, G.: Using dynamic field theory to extend the embodiment stance toward higher cognition. New Ideas in Psychology **31**, 322–339 (2013). https://doi.org/10.1016/j.newideapsych.2013.01.002

14. Schaal, S.: The new robotics-towards human-centered machines. HFSP journal **1**, 115–26 (08 2007). https://doi.org/10.2976/1.2748612

15. Schöner, G.: Dynamical Systems Approaches to Cognition. No. January (2012). https://doi.org/10.1017/cbo9780511816772.007

16. Sousa, E., Erlhagen, W., Ferreira, F., Bicho, E.: Off-line simulation inspires insight: A neurodynamics approach to efficient robot task learning. Neural Networks **72**, 123–139 (2015). https://doi.org/10.1016/j.neunet.2015.09.002

17. Wojtak, W., Ferreira, F., Louro, L., Bicho, E., Erlhagen, W.: Towards temporal cognition for robots: A neurodynamics approach. 7th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, ICDL-EpiRob 2017 pp. 407–412 (2018). https://doi.org/10.1109/DEVLRN.2017.8329836