An Intelligent Interactive Learning and Adaptation Framework for Robot-Based Vocational Training

Konstantinos Tsiakas*, Maher Abujelala†, Alexandros Lioulemes‡ and Fillia Makedon§
Department of Computer Science and Engineering
The University of Texas at Arlington, USA
*konstantinos.tsiakas@mavs.uta.edu, †maher.abujelala@mavs.uta.edu,
‡alexandros.lioulemes@mavs.uta.edu, §makedon@uta.edu

***Abstract*—In this paper, we propose an Interactive Learning and Adaptation framework for Human-Robot Interaction in a vocational setting. We show how Interactive Reinforcement Learning (RL) techniques can be applied to such HRI applications in order to promote effective interaction. We present the framework by showing two different use cases in a vocational setting. In the first use case, the robot acts as a trainer, assisting the user while the user is solving the Towers of Hanoi problem. In the second use case, a robot and a human operator collaborate towards solving a synergistic construction or assembly task. We show how RL is used in the proposed framework and discuss its effectiveness in the two different vocational use cases, the Robot Assisted Training and the Human-Robot Collaboration case.**

I. Introduction

In the past few years, intelligent and autonomous robotic systems have been applied to vocational and manufacturing environments with a variety of applications, ranging from identifying training needs to risk prevention. Some of these robotic systems are fully-automated and placed in restricted area without interacting with humans for safety concerns. Nowadays, it became very essential to have humans and robots collaborate and work side by side, and this can be seen in vocational education, training and rehabilitation [1], [2], as well as in human-robot collaboration applications [3]. The development of such automated systems requires computationally advanced intelligent mechanisms in order to ensure safe and efficient interaction between the robotic system and the involved human worker, as the environment may involve heavy machinery [4], [5].

We can observe two main contributions of intelligent robotic systems in a vocational setting; (1) Robot-Based Vocational Training and (2) Human-Robot Collaboration. In the first case, robotic systems are deployed to train human workers in specific task, in a physical or simulated vocational setting, in order to enhance their abilities and thus their performance [6]. On the other hand, as more industrial robotic systems are being developed to operate and collaborate with human workers in the same physical space, there is increased need to advance technologies that enable safe, effective and efficient human-robot teamwork.

The deployment of robotic systems, either as robotic trainers or as teammates, raises the need of development of computational advancements. In particular, an important attribute of such robotic systems, is the development of agents that have behavior adaptability in enabling properties such as user modeling and user intention prediction, task planning, scene understanding, as well as understanding and adjusting to the abilities and preferences of the human user [7]. An important attribute of such robotic agents is the adaptability of their behavior towards the goal, as well as the abilities and preferences of the human user. Recent works discuss the aspect of *co-adaptation* in Human-Machine interaction [8], [9], referring to the process of adjustment of both the machine and the human operator during the interaction, considering both goal-oriented and human-centered approaches [10]. In this paper we consider the concept of ”co-adaptation” between a human operator and a machine interface and we formulate its application in the context of the two different scenarios mentioned earlier, the robot as a trainer and the robot as a team member.

In the first case of the Robot Assisted Training (RAT) systems, the robotic trainer must be able to adapt the training session to each specific user, in order to enhance specific and personalized attributes of each specific worker. In the second case, when a robot is employed as a team member, the robotic agent should be able to adjust its behavior towards the different set of abilities and preferences and the personalized style of each human member in the team. In both cases, the robot should also adjust its behavior according to the evolution of human skills and performance.

Considering the above, a dynamic adaptation mechanism is required to continuously adapt and adjust the robots policy towards each specific user's needs and abilities. In this paper, we present a unified computational framework that focuses on the adaptation mechanism of such robotic assistants under both aforementioned vocational settings. The paper is organized as follows: in Section II, we discuss our approach and methodology for modeling the robots behavior during the interaction with a human worker. In Section III, we present our proposed framework and its application to the different vocational settings. In Section IV, we present our use cases, illustrating our framework and we conclude, in Section V, presenting our future work.

II. Approach and Methodology

As already mentioned, we focus on two applications of robotic assistants in a vocational setting; Robot Assisted

Training and Human-Robot Collaboration. In both settings, the robot must be able to perceive and act based on its environment (task and human), adapting to environmental changes in order to ensure a safe and effective interaction. It is essential that these robotic agents follow a computational mechanism to encode how human users act towards the completion of a certain task.

To this end, we adopt the approach of *Shared Mental Models* as a computational model that captures the knowledge about the robot itself, the human user and task than needs to be performed. In [11], they represent the robot mental model as a Markov Decision Process (MDP). In the following sections, we present a basic introduction about MDP, as well as how we represent the mental models for the two use cases.

A. Markov Decision Processes

An MDP is described by a tuple $<$ *S, A, T, R* $>$ where:

- S is a finite set of states - state space
- A is the finite set of available actions – action space
- T is the transition model where T(s, a, s) denotes the probability of moving from state s to state s by performing action a
- R(s, a) is a reward function that gives a numerical reward of going to state s performing action a

The solution of an MDP results to an *optimal policy*. An optimal policy π is the mapping from states to actions that maximizes the total expected reward the agent receives, as it interacts with its environment. In order to solve an MDP, we follow the Reinforcement Learning framework [12].

However, finding an optimal policy often requires a large number of iterations, making it inappropriate for a real-time HRI application. Moreover, an optimal policy is learned based on specific environmental parameters and not to a dynamic environment, as when a human user is involved in the interaction. In our case, the environment is described by the task and the human user's abilities, preferences and intentions. As we already mentioned, an effective robotic assistant should be able to refine its policy based on each specific user, in order to adapt to the different user's set of abilities and skills. Taking these into consideration, we propose an Interactive Learning and Adaptation Framework [13], [14], that utilizes Interactive Reinforcement Learning methods to facilitate the safe adaptation of an agent to a different user.

III. Interactive Learning and Adaptation Framework

In this section, we present an interactive learning and adaptation framework that integrates Interactive Reinforcement Learning approaches to the adaptation mechanism. *Interactive Reinforcement Learning* (IRL) is a variation of RL that studies how a human can be included in the agent learning process. Human input can be either in the form of feedback or guidance. *Learning from Feedback* treats the human input as a reinforcement signal after the executed action [15], [16]. *Learning from Guidance* allows human intervention to the selected action before execution, proposing (corrective) actions [17]. To our knowledge, IRL methods have not been investigated for the adaptation of an agent to a new environment. Hence, we propose their integration to the adaptation mechanism, as *policy evaluation* metrics used to evaluate and modify a learned policy towards an optimal one, following proper transfer methods [18], [19]. Moreover, *Learning from Demonstration*, another IRL approach which studies how robots can learn a policy from observed or given example state-to-action mappings, can be applied to this framework, to provide the robot with prior knowledge about the task. A comprehensive survey of robot learning from demonstration can be found in [20]. In Figure 1, we present the algorithm for the proposed framework.

1: **procedure** Interactive Policy Adaptation(π, B)
2: $\quad s = start\_state$
3: $\quad N = 0$
4: $\quad$ **while** $s \neq goal\_state$ **do**
5: $\quad\quad$ Select action a based on s, π
6: $\quad\quad$ **if** $guidance$:
7: $\quad\quad\quad$ **if** $a \neq guidance$ **and** $N \leq B$
8: $\quad\quad\quad\quad a = guidance$
9: $\quad\quad\quad\quad N = N + 1$
10: $\quad\quad$ observe user action, next state s' and reward r
11: $\quad\quad Q(s,a)$ += $\alpha \cdot (r + \gamma \max_a Q(s',a) - Q(s,a))$
12: $\quad\quad$ **if** $feedback$:
13: $\quad\quad\quad Q(s,a) = Q(s,a) + \beta \cdot feedback$
14: $\quad\quad$ Record $[s, a, s', r, feedback, guidance]$
15: $\quad\quad s = s'$
16: $\quad$ **end while**
17: **end procedure**

Fig. 1. Interactive Policy Adaptation with Q-learning

The algorithm is a variation of Q-learning [21], modified for the interactive adaptation of a learned policy. We extend the algorithm by adding the two additional communication channels; $feedback$ and $guidance$. During the interaction, the agent chooses an action a based on its current state s and its learned policy π (Line 5). This action can be modified (Lines 6–9) following the *teaching on a budget* approach [22]. Based on this approach, the user can provide a limited amount of guidance (B), when needed, correcting the selected action. After the Q-value update, feedback provided by the user can be used to modify the policy, following the *Q-augmentation* combination technique (Line 13) [15]. After each interaction step, the algorithm records the interaction data (Line 14), that can be used for a model-based approach. In the next sections, we show how this interactive policy adaptation algorithm applies to the different settings.

1) Robot Assisted Training: In a Robot Assisted Training (RAT) task, a robot acts as a human trainer. Robot Assisted Training has been extensively applied to assist users during cognitive and physical tasks, aiming to enhance certain user abilities in a vocational setting [23], [24]. The robot must be able to perceive the user's state, providing them a personalized training session based on the specific set of abilities and skills. It is shown that personalized training is more effective than a generalized, maximizing user's performance and engagement.

In Figure 2, we outline our proposed framework in a RAT application, where a human user interacts with a robot during a training (physical or cognitive) task. In that case, the robot mental model captures all the required information to represent the task and user progress, formulating its state space S. Each state includes important task and user information (e.g., task difficulty, task duration, user performance, etc.), based on which it performs an appropriate action (e.g., task difficulty, hint type, encouragement, etc.). The transition model captures how the human user reacts under different robot actions. The optimal policy, the one that will maximize the expected total reward, will result to a personalized training session that will eventually assist the user to train specific skills and abilities.

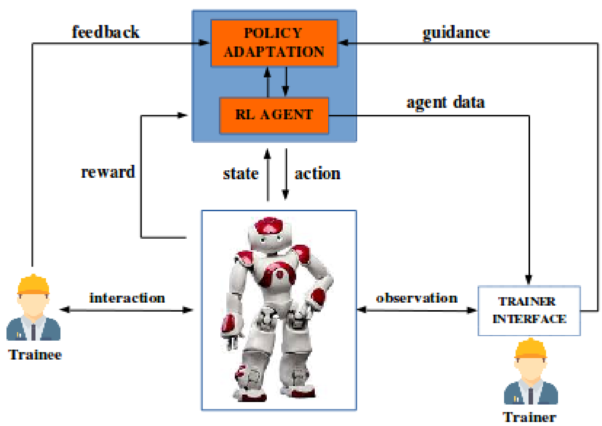

Fig. 2. Outline of the proposed framework in a Robot-Assisted Training task. We extend the RL framework by adding two additional communication channels; *feedback* and *guidance*. Their integration to the adaptation module can enable the agent to continuously adapt towards the current user, ensuring a safe and personalized interaction.

Interactive learning methods (learning from feedback and guidance) are integrated to the learning mechanism, facilitating the policy adaptation of the robotic agent to the current user. Feedback is considered to be a personalization factor, as it is provided implicitly by the primary user (trainee), in the form of facial expressions, speech, haptic or visual feedback, etc. User feedback is task-related and provides the agent with useful information about its current policy. To ensure an effective and safe training session, a secondary user (human trainer) can guide the early interaction steps, providing the system with suggested or corrective actions [22], [25]. In our preliminary results [13], we have shown that the integration of interactive learning methods facilitated the adaptation of a learned policy to a new user.

2) Human-Robot Collaboration: In a Human-Robot collaboration task, human users and robot are considered to work together towards a common task. Under this case, the robot should be able to recognize user's intentions, skills and coordinate with human workers in order to perform a task, in a synergistic manner. In order to achieve such a coordination, the robot is required to develop, maintain and keep track of the team's participants mental models including, among others, team member roles, intentions, intended goals, performed actions and so forth.

In a synergistic task, such models are mostly referred as *shared mental models* [26], [11], representing the knowledge of an individual (human or robot) about how team members should coordinate towards a common goal. This requires the ability of a robot to adapt to each different teammate, in terms of preferences, skills and intentions. A robot that works with a person according to another user's preferences and skills is highly likely to be ineffective. Moreover, even the same person's intentions and abilities are subject to change over time, or even within the task, as human skills, preferences and performance may evolve (co-adaptation). Based on these, we argue that an online adaptation mechanism is needed, to enable the robotic team member to continuously adapt to the current member's expertise and preferences. In Figure 3, we illustrate our proposed framework in a synergistic task.

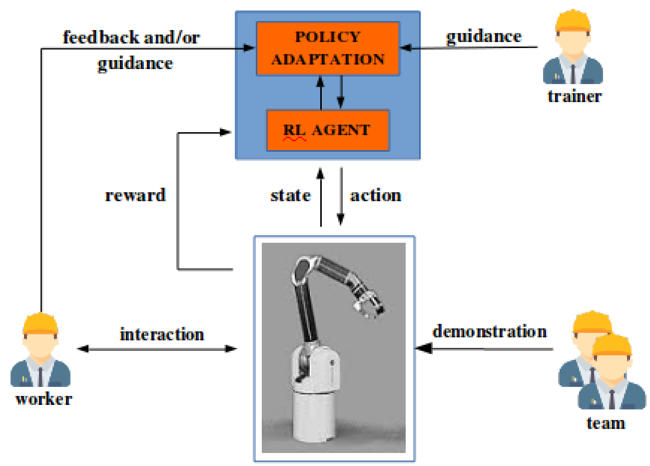

Fig. 3. Illustration of the proposed framework in a human-robot collaboration task. The robot can learn an initial policy by observing two or more user during a collaborative task. Feedback from the primary user is used to dynamically adjust the learned policy towards the needs and abilities of the current user. Moreover, a human trainer can dynamically define a training protocol by providing guidance directly to the system.

An initial policy can be learned by (or provided to) the robot from observed or demonstrated examples between two or more human workers. This enables the robot to acquire prior knowledge of the task and the human selection mechanism during a collaborative task. This initial policy needs to be refined as the robot collaborates with a new worker to meet the specific individual set of abilities and work plan. Even if the robot is used to substitute a member of a known team (from the observed team), a refinement of the model is needed to ensure a safe and efficient coordination. IRL methods can be applied to enable the human team member(s) provide the robot with feedback (after a robot selected action) or with guidance (prior to robot action execution).

IV. Use Cases

In this section, we present two use cases, covering the aforementioned scenarios and the application of the proposed framework for each one.

A. Robot-Assisted Training: The Tower of Hanoi task

In this section, we illustrate our proposed framework with a use case in Robot Assisted Training. We follow a specific scenario, where the user is asked to perform the 'Tower of Hanoi' task[1]. The Tower of Hanoi has been used in many studies to determine the user planning and problem solving abilities [27], [28]. Similarly, in this framework, we use it to identify, appraise and evaluate an individuals level of functioning for employment decision making in industries and to test their problem solving skills. A NAO robot is employed to assist the user by providing hints, when needed, as shown in the experimental setup in Figure 4.

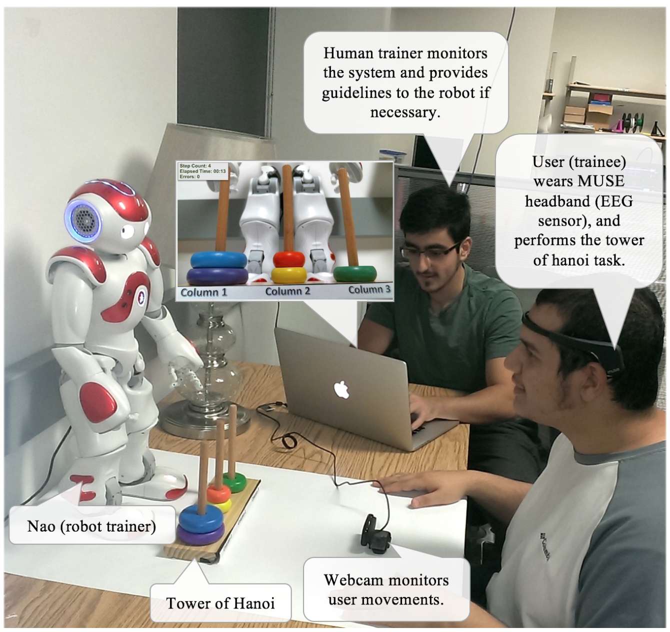

Fig. 4. The experimental setup for the 'Towers of Hanoi' training session. The NAO robot provides hints to the user during the interaction. A secondary user is able to control the robot's action selection mechanism (learning from guidance). We plan to measure user's engagement and integrate it as feedback to the learning mechanism (learning from feedback).

We follow the approach of an adaptive Socially Assistive Robotic (SAR) system ([29], [30]), where the robot follows a personalized strategy for each individual. Each trainee has to complete the task within a predefined amount of time. After the end of the session, the user receives a score which is a function of the number of movements, and the elapsed time, and the number of errors (rules violations)/extra steps. The user's goal is to maximize their score during each session by finishing the task in a shorter time while making less errors and steps. The state space includes the user's last move $(correct, wrong)$, task completion (yes or no), the elapsed time (in time units). Based on this information, the robot must learn when to provide the user with a hint. The robot can follow three different training strategies:

1) No hint,
2) Inform the user if the last move was correct or not
3) Provide the user with the correct action to perform

The core RL agent will receive a positive reward under a successful task completion, relative to the user's score. Individuals of different abilities and task expertise require different strategies, so as to maximize their score and their engagement to the task.

In the initial development stage, we have implemented a simple and robust vision-based algorithm that can track user movements during the task. This algorithm analyses the Hue-Saturation-Value (HSV) color space of the input video from an external web-cam facing the Tower of Hanoi. The algorithm determines the position of each ring and detects whether the user is interacting with the system. Based on the current position of the rings, the process provides the trainer (human and robot) with real-time feedback whether the previous move was correct and finds the best next movement.

Apart from the objective user's performance (game performance), we propose to measure psychological data to estimate user's engagement during the task. Psychological data may be very essential in monitoring users state to optimize their performance [31]. Therefore, to make the interaction between the robot (trainer) and the user (trainee) more adaptive, we propose adding a Brain-Computer Interface (BCI) that can provide the user's engagement as implicit feedback to the adaptation mechanism. In our previous work [32], we have developed a method for evaluating user engaged enjoyment, using a commercially available EEG tool (Muse[2]). Our method is able to measure brain activities, reflecting user enjoyment in a given task, which allows for task comparison, in terms of enjoyment. This mental feedback is essential to assist the robot to refine its policy to follow when the user (trainee) makes a mistake or takes a longer time to perform a single move. Task engagement can be derived from the MUSE EEG output as a function of the alpha, beta and theta bands [33]. We define the task engagement levels as three states: *bored, engaged* and *stressed*. This implicit user feedback can be exploited as a *policy evaluation* metric, considering the correlation between task difficulty and user expertise [34], [35]. For example, if an expert user is continuously provided with hints, their engagement may decrease. The received feedback will help the system refine this non-optimal policy and adapt it to the specific user. In order to ensure more reliable readings, we will follow self-calibrating methods to enhance the sensor output data fidelity [36].

On the other hand, in order to facilitate the policy adaptation and the effectiveness of the training session, our framework supports the participation of a secondary user, who observes the interaction and is able to control the action selection mechanism by intervening with corrective or suggesting alternate

[1] https://en.wikipedia.org/wiki/Tower_of_Hanoi

[2] http://www.choosemuse.com

actions, until the robot acts effectively in an autonomous way. Feedback and guidance can be used as an adaptation parameters following the algorithm showed in Figure 1. Considering the form and the amount of guidance, we investigate methods that enhance expert's decision making, while minimizing their workload, by making the learning process transparent [37], [17].

B. Human-Robot Collaboration: Lego Construction task

Human-robot coordination is considered a challenging task, in terms of safety and efficiency, especially when human and robot need to cooperate in the same physical space [38]. Moreover, collaborative robots need to act in a *contextually-rich* environment, dealing with various objects during the synergistic task [39]. In this work, we focus on the collaborative task modeling, to enable the robot adapt to the different set of preferences and abilities of each human coworker, following our proposed framework.

In our use case for a collaborative task, a robot and a human need to work together to assemble a specific LEGO® construction (Figure. 5) using the available parts; red and blue parts and screws (white parts). The role of both teammates (human and robot) is interchangeable, i.e. they both have the same set of available actions. The challenge in such a task is to enable the robot learn a personalized strategy according to each teammate. Moreover, the robot can learn a training strategy, by altering its policy, to train the human user how to perform the task in different ways, or under a task-switching [40] or cross-training environment [11].

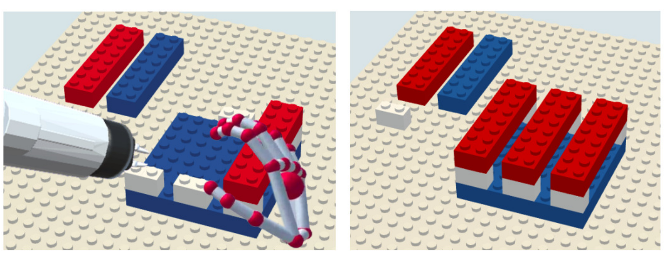

Fig. 5. Lego assembly task simulation using the Oculus Rift and a custom 3D structural game. In the left image, we show the virtual environment, where the user and the robot perform an assembly task. In the right photo, we show the final construction – the goal state.

For example, one worker prefers placing the blue parts first, then assign to the robot to place (drill) the screws and then the human places the red parts. In a different case, a worker may need the robot to hold a specific part as they place a screw, or another part. Consequently, there is a variety of policies that the robot can learn and apply. We argue that our framework enables the robot to update its current policy towards each specific user in an online fashion, since each user provides the robot with direct feedback and guidance during the interaction that matches their working profile. The policy is updated following the proposed algorithm (Figure 1). Feedback and guidance can be provided to the robot, in real time, via predefined speech commands.

The action space includes the available actions (by both human and robot) defined by the tuple $[action, object]$, where $action = [place, pick, drill]$ and $object = [red, blue, white]$. Moreover, no_action is included. The state space includes information about the task completion (yes or no) and the current human action e.g., $[places, red]$. The core RL algorithm will receive a positive reward for the task completion. Based on the state and its current policy, the robot performs an action. The human trainee performs and the user can either guide the robot (by suggesting an alternate action) or by providing feedback (positive or negative) for the executed action. In the case of human training in a collaborative task, an expert human trainer can define their own training strategies and goals by altering robot's policy through guidance.

Our initial experimental setup (before the physical experiments) is on a simulation environment. To capture the users' hand, we are using the Kinect's skeleton tracker and the Leap Motion's fingers tracking module. Also we are using the Barrett WAM Arm manipulator because of its backdrivability and its high sensitivity features. Finally, we employ the Oculus Rift and a custom 3D structural game at the Unity platform to boost the user's immersion for these simulated cases.

V. Conclusion and Future Work

In this work, we outlined an Interactive Learning and Adaptation framework that facilitates the adaptation of a robotic agent towards the current user, in a Robot-Based vocational training and collaboration setting. We discussed the need of an online adaptation mechanism and we illustrated our framework in two use cases, covering different applications in a vocational setting.

Our future work includes data collection from users interacting with the two presented applications, under the proposed framework. These data will be used to develop user models for further investigation and improvement of the presented algorithm, in the form of interaction data ($[s, a, s', r, feedback, guidance]$). For the Tower of Hanoi, these data include physiological data from the MUSE sensor, the guidance from the secondary user, given the current state, as well as visual data from the webcam that estimate the current state of the task and the action to be taken. During the assembly task, our data collection plan includes acquisition of the state-action trajectories, as well as the provided feedback and guidance in the form of simple voice commands, that will be used for the RL policy evaluation. Moreover, user studies will be conducted on both use cases, to receive evaluative feedback for both the proposed framework and the user experience during the interaction.

In this ongoing work, we investigate different techniques of how feedback and guidance can be integrated to the adaptation mechanism. We will compare our results with other existing frameworks [25], [29], [11], aiming to minimize human workload as the agent learns. We are conducting user studies, following these use cases, to investigate how human users provide an interactive agent with feedback and guidance and we will refine both our framework and the use cases. The long-term goal of this research is to develop progressively-autonomous systems that learn as long as they interact with real users, maximizing their efficiency and effectiveness in vocational settings.

ACKNOWLEDGMENT

This material is based upon work supported by NSF under award numbers CNS 1338118, 1035913. The authors would like to acknowledge the work of the following REU and Ph.D. students on the use cases: Ashwin Ramesh Babu, Akilesh Rajavenkatanarayanan, Joseph Tompkins and Dylan Ebert.

REFERENCES

[1] J. L. Schuyler and R. M. Mahoney, "Assessing human-robotic performance for vocational placement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 3, pp. 394–404, 2000.

[2] F. M. López-Rodríguez and F. Cuesta, "Andruino-a1: low-cost educational mobile robot based on android and arduino," *Journal of Intelligent & Robotic Systems*, vol. 81, no. 1, pp. 63–76, 2016.

[3] A. Bauer, D. Wollherr, and M. Buss, "Human–robot collaboration: a survey," *International Journal of Humanoid Robotics*, vol. 5, no. 01, pp. 47–66, 2008.

[4] R. Alami, A. Albu-Schäffer, A. Bicchi, R. Bischoff, R. Chatila, A. De Luca, A. De Santis, G. Giralt, J. Guiochet, G. Hirzinger *et al.*, "Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 1–16.

[5] C. Heyer, "Human-robot interaction and future industrial robotics applications," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 4749–4754.

[6] K. Tsiakas and F. Makedon. (2016) Safe adaptation of intelligent interactive systems. [Online]. Available: http://iperform.uta.edu/IAB_spring16/project_summaries/projsummary_mar2016_Safe_Adaptation_of_Intelligent_Interactive_Systems.pdf

[7] B. Scassellati and B. Hayes, "Human-robot collaboration," *AI Matters*, vol. 1, no. 2, pp. 22–23, 2014.

[8] P. Gallina, N. Bellotto, and M. Di Luca, "Progressive co-adaptation in human-machine interaction," in *Informatics in Control, Automation and Robotics*. IEEE, 2015.

[9] S. Chandramohan, M. Geist, F. Lefevre, and O. Pietquin, "Co-adaptation in spoken dialogue systems," in *Natural Interaction with Robots, Knowbots and Smartphones*. Springer, 2014, pp. 343–353.

[10] S. Schaal, "The new robotics – towards human-centered machines," *HFSP journal*, vol. 1, no. 2, pp. 115–126, 2007.

[11] S. Nikolaidis and J. Shah, "Human-robot teaming using shared mental models," *ACM/IEEE HRI*, 2012.

[12] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. MIT Press Cambridge, 1998, vol. 135.

[13] K. Tsiakas, M. Dagioglou, V. Karkaletsis, and F. Makedon, "Adaptive robot assisted therapy using interactive reinforcement learning," in *International Conference on Social Robotics*. Springer, 2016.

[14] K. Tsiakas, "Facilitating safe adaptation of interactive agents using interactive reinforcement learning," in *Companion Publication of the 21st International Conference on Intelligent User Interfaces*. ACM, 2016, pp. 106–109.

[15] W. B. Knox and P. Stone, "Reinforcement learning from simultaneous human and mdp reward," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 2012, pp. 475–482.

[16] J. Broekens, "Emotion and reinforcement: affective facial expressions facilitate robot learning," in *Artifical Intelligence for Human Computing*. Springer, 2007, pp. 113–132.

[17] S. Chernova and M. Veloso, "Interactive policy learning through confidence-based autonomy," *Journal of Artificial Intelligence Research*, vol. 34, no. 1, p. 1, 2009.

[18] L. Torrey, "Lightweight adaptation in model-based reinforcement learning." in *Lifelong Learning*, 2011.

[19] M. E. Taylor and P. Stone, "Transfer learning for reinforcement learning domains: A survey," *The Journal of Machine Learning Research*, 2009.

[20] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Rob.Aut.Sys*, vol. 57, no. 5, pp. 469–483, 2009.

[21] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[22] L. Torrey and M. Taylor, "Teaching on a budget: Agents advising agents in reinforcement learning," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 1053–1060.

[23] P. S. Lum, C. G. Burgar, and M. Van Der Loos, "The use of a robotic device for post-stroke movement therapy," in *Proceedings of the International Conference on Rehabilitation Robotics*, 1997.

[24] D. De Santis, J. Zenzeri, M. Casadio, L. Masia, A. Riva, P. Morasso, and V. Squeri, "Robot-assisted training of the kinesthetic sense: enhancing proprioception after stroke," *Frontiers in human neuroscience*, vol. 8, p. 1037, 2015.

[25] E. Senft, P. Baxter, J. Kennedy, and T. Belpaeme, "Sparc: Supervised progressively autonomous robot competencies," in *International Conference on Social Robotics*. Springer, 2015, pp. 603–612.

[26] J. E. Mathieu, T. S. Heffner, G. F. Goodwin, E. Salas, and J. A. Cannon-Bowers, "The influence of shared mental models on team process and performance." *Journal of applied psychology*, vol. 85, no. 2, p. 273, 2000.

[27] G. Glosser and H. Goodglass, "Disorders in executive control functions among aphasic and other brain-damaged patients," *Journal of Clinical and experimental Neuropsychology*, vol. 12, no. 4, pp. 485–501, 1990.

[28] S. V. Borys, H. H. Spitz, and B. A. Dorans, "Tower of hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state," *Journal of Experimental Child Psychology*, vol. 33, no. 1, pp. 87–110, 1982.

[29] A. Tapus, C. Tapus, and M. Mataric, "The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia," in *2009 IEEE International Conference on Rehabilitation Robotics*, 2009.

[30] G. Gordon, S. Spaulding, J. K. Westlund, J. J. Lee, L. Plummer, M. Martinez, M. Das, and C. Breazeal, "Affective personalization of a social robot tutor for childrens second language skills," in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[31] S. H. Fairclough, L. J. Moores, K. C. Ewing, and J. Roberts, "Measuring task engagement as an input to physiological computing," *VOLUME II*, p. 33, 2009.

[32] M. Abujelala, C. Abellanoza, A. Sharma, and F. Makedon, "Brain-ee: Brain enjoyment evaluation using commercial eeg headband," in *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2016.

[33] P. McCullagh, G. Lightbody, L. Galway, C. Brennan, and D. Trainor, "Assessment of task engagement using brain computer interface technology," in *Workshop Proceedings of the 11th International Conference on Intelligent Environments*, vol. 19. IOS Press, 2015, p. 244.

[34] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun, "Boredom, engagement and anxiety as indicators for adaptation to difficulty in games," in *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*. ACM, 2008, pp. 13–17.

[35] A. Plotnikov *et al.*, "Exploiting real-time eeg analysis for assessing flow in games," in *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 688–689.

[36] T. Karydis *et al.*, "Performance characterization of self-calibrating protocols for wearable eeg applications," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2015, p. 38.

[37] C. Chao, M. Cakmak, and A. L. Thomaz, "Transparent active learning for robots," in *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, 2010, pp. 317–324.

[38] M. C. Gombolay, R. J. Wilcox, A. Diaz, F. Yu, and J. A. Shah, "Towards successful coordination of human and robotic work using automated scheduling tools: An initial pilot study," in *Proc. Robotics: Science and Systems (RSS) Human-Robot Collaboration Workshop (HRC)*, 2013.

[39] H. S. Koppula, A. Jain, and A. Saxena, "Anticipatory planning for human-robot teams," in *Experimental Robotics*. Springer, 2016, pp. 453–470.

[40] W. Sheng, A. Thobbi, and Y. Gu, "An integrated framework for human–robot collaborative manipulation," *IEEE transactions on cybernetics*, vol. 45, no. 10, pp. 2030–2041, 2015.