Decision Making for Affective Agents in Assistive Environments

Konstantinos Tsiakas

Heracleia Human Centered Computing Laboratory
Computer Science and Enginnering Department
Univeristy of Texas at Arlington, USA

Institute of of Informatics and Telecommunications
National Center for Scientific Research - NCSR 'Demokritos', Athens, Greece
konstantinos.tsiakas@mavs.uta.edu

Abstract. *In this paper, we discuss the Decision Making and Learning ability of Affective Agents to make human-like decisions. This work is in the context of Assistive Living Environments (ALE) applications, where an agent is capable of assisting a human in physical and cognitive rehabilitation through multimodal and adaptive interaction. The goal of this research is to investigate what role multimodality plays in producing a natural and effective interaction using Reinforcement Learning. We propose a hierarchical decision making framework for affective agents doing complex tasks. This framework incorporates an internal reward mechanism to make the learning more efficient.*

Keywords: Decision Making, Reinforcement Learning, Affective Agents, Multimodal Interaction, Adaptive Systems, Cognitive and Physical Rehabilitation Systems

I am a second year Ph.D. candidate and my research focuses on Multimodal Human Computer Interaction and specifically on the aspect of the Decision Making. I follow the Reinforcement Learning approach, where the agent receives multisensing data based on which the agent must take the most appropriate decision to make the interaction natural and efficient.

1 Introduction

During the past decade, researchers have investigated the decision making processes of autonomous agent systems. The aim has been to model system behavior that enables efficient and robust interaction with users. Recent advances in autonomous agents have shown that combining different modalities leads to a more natural interaction with humans. Using virtual agents or embodied conversational agents [1], and combining modalities such as, facial expressions, body motion, gesture recognition, gaze detection, haptic feedback, emotional state estimation and others, can lead to a more natural interaction, more like human to human communication where different modalities are used [2].

On the other hand, hardcoding rule-based decisions based on multisensing stimuli perceived by the agents, may lead to ineffective and time-consuming designs, since environmental changes or unseen events will not be modeled or incorporated into the design. Furthermore, since human factors play a major role in an interaction, any agent decision making system must be adaptive and robust to user preferences and needs, and in line with Human-Computer Interaction basic principles.

Reinforcement Learning has been applied extensively to problems where the system must learn behavior through *trial-and-error* interactions within a dynamic environment [3]. There have been many approaches applying Reinforcement Learning to *affective agent* decision making and learning. In [4], a motivational reward framework inspired by human learning and recent neuroscience was proposed. In this work, external rewards from the environment are combined with internal rewards of the agent cognitive and affective state. In [5], *Intrinsic Motivation for Reinforcement Learning* (IMRL) was introduced as a type of reinforcement learning that incorporates principles of intrinsic motivation to the traditional external rewards obtained from the environment. In [6], a computational model of emotions was presented as a mapping between RL primitives and emotion labels, where agent-based simulation experiments can replicate psychological and behavioral dynamics of emotion.

In my research, I explore Reinforcement Learning approaches applied to multimodal human-machine interaction. I study the process of decision making and learning in Assistive Environments applications, where the aim is to extract user behavioral and physical data as the subject performs physical and cognitive rehabilitation tasks. In particular, my plan is to design a computational framework that can be used to build two systems: (a) a Rehabilitation Session Manager and (b) an Embodied Conversational Agent for assessing Depression and Anxiety Disorders Screening described further below.

2 Research Plan

My research has three phases. The first one is the problem formulation and the data collection. For the two systems mentioned above, I will conduct many rounds of Wizard-of-Oz experiments to acquire multimodal data based on a specific scenario. In the second phase, I will use existing Reinforcement Learning frameworks to investigate if multisensing data and complex tasks can be handled effectively using traditional RL algorithms. In the third phase, I will introduce a new computational framework for affective agents that can handle different modalities and more complex tasks.

3 An Affective Agent for Physical and Cognitive Rehabilitation

We propose the use of affective agents for physical and cognitive rehabilitation tasks. The agent must interact in a natural and effective way to retrieve behavioral and cognitive data during its interaction with the user. In this section, I present Readapt;a Rehabilitation Session Manager and an Embodied Conversational Agent for assessing Depression and Anxiety Disorders Screening (DADS).

3.1 ReAdapt: A Multimodal Adaptive Rehabilitation Session Manager

ReAdapt is a multimodal adaptive rehabilitation session manager for monitoring remote exercising. The system is responsible to adapt the difficulty and the parameters of the rehabilitation session based on the subject's physical performance and condition. The problem is formulated as a Markov Decision Process (MDP) where the agent perceives multisensing data that represent the subject's physical performance and condition. Specifically, the state space includes the type and difficulty level of the current exercise, the time spent on this exercise, the subject's performance through body motion analysis, a pain detection system that extracts facial features and the user self-report of pain using speech. In response, the agent takes appropriate actions to adapt the exercise difficulty and to interact with the subject.

ReAdapt has two goals: (a) to keep the subject engaged so that he completes the whole set of prescribed exercises and (b) that he spends the appropriate amount of time in each exercise, while preventing high levels of pain and ensuring safe rehabilitation. We adopted a user simulation model, where we used the *Dyna-Q* algorithm to evaluate the decision making and learning mechanisms. We evaluated our system by quantifying the system performance using a global reward for both goals stated above.

In future work, the exercises will be shown by a therapist avatar who is giving instructions on how to perform the exercises to human subjects. The preliminary experimental simulation and results are shown in Figure 1.

3.2 DADS: An Embodied Conversational Agent for Depressive and Anxiety Disorder Screening

In this section, I present ongoing work on *DADS*, which is an embodied conversational agent for Depressive and Anxiety Disorder Screening. The system is able to interact with the user using verbal and non-verbal communication. The main goal of the system is to be a self-assessment tool for Post-Traumatic Stress Disorder (PTSD) and related stress and anxiety disorders. DADS can make referrals for intervention and monitors progress. We focus on the Dialog Manager component, the module that is responsible for the decision making. During the dialogue interaction, the system takes into consideration verbal (speech) and non-verbal (facial expressions) input and adapts the interaction in order to elicit the required information needed while keeping the user in a 'calm' emotional state. The system architecture is shown in Figure 2

We model the dialogue using interconnected MDPs that describe different tasks. We split the dialogue into different sub-dialogues based on the user's replies and emotional state. Due to the complexity of the task, we will conduct

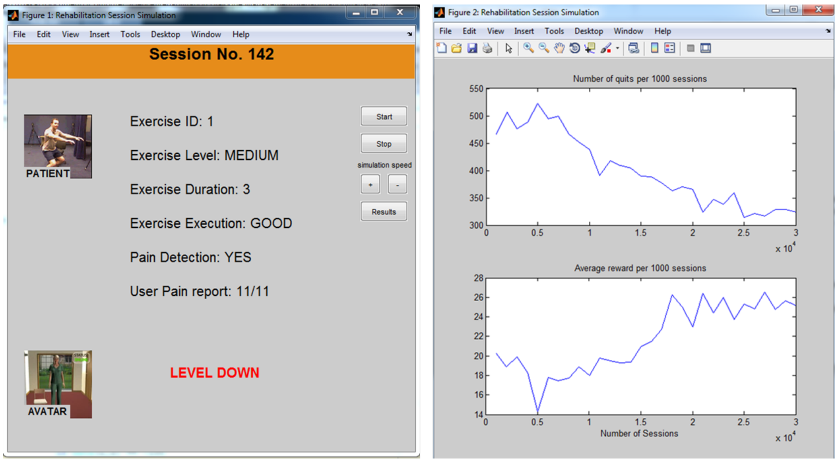

Fig. 1. Screenshot of ReAdapt: The left figure shows a simulation of the session manager showing the subject performance and a therapist avatar instructing the exercise and the selected action. In the right figure, we show that as the algorithm learns the optimal policy, the number of quits is decreasing, while the average reward for each episode/session increases.

a round of WoZ experiments in order to collect data for the user simulation model. We will use this data to feed the Reinforcement Learning algorithm for the system training. Then, we will evaluate the system with a second round of WoZ experiments receiving feedback from the users and a psychologist. The goal of this system is to adapt the order of questions based on the user responses, in order to elicit the required information for the assessment, while preventing the user from being in unwanted emotional states. As in previous task, we evaluate the system performance using a single reward that quantifies both desired goals.

4 Hierarchical Decisions for Affective Agents

Previously, we described two systems of affective agents for physical and cognitive rehabilitation. In order to utilize the multisensing data collected during the system interaction with the user, we incorporate both audiovisual and context data to the state space resulting to a large state space. Moreover, each available action is considered to be selectable from any possible state resulting to a large MDP whose solving is likely enough to be intractable or requires a large amount of iterations for an efficient learning. I will focus on two main approaches to deal with these problems.

1. **Hierarchical Decisions** A possible solution to the large state-action space problem is to reduce the state space using known methods, as function approximation. Another solution is to decompose the problem to lower-level problems, where each one will have its own state space and a subset of

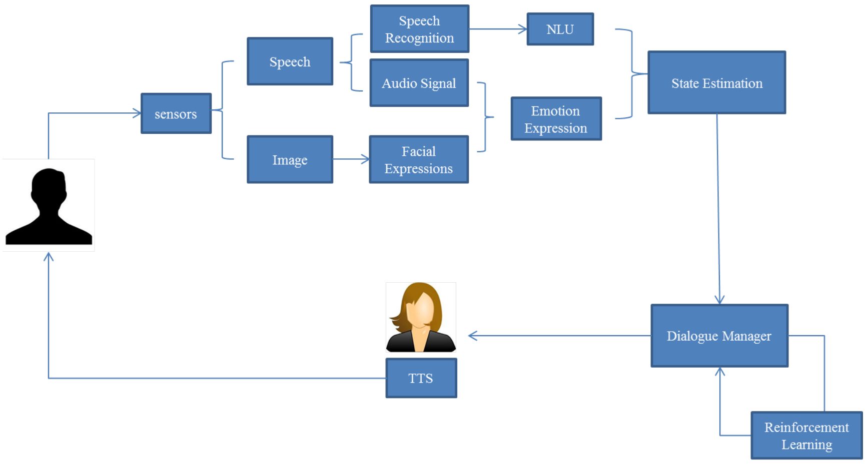

Fig. 2. DADS Architecture: The system receives verbal and non-verbal data to extract information needed for the assesment. The manager decides for the next action expressed by the avatar and a Text-to-Speech (TTS) system.

selectable actions. In this way, each lower-level decision takes into consideration specific state attributes relative to the sub-goal resulting to a reduced state-action space. Moreover, complex actions can be composed of weighted decisions of lower-level primitive actions.

2. **External and Internal Rewards** In this section we discuss ways to address the problem with the large number of iterations needed for the learning, mentioned earlier. In the preliminary work described here, we quantify the performance combining all sub-goals to a single reward. Integrating internal rewards for each one of the lower-level decisions with the overall external environmental reward will make the decision making more efficient and the learning faster. It has been shown that using internal rewards serves as motivation for the affective agent to expedite the learning and make it more efficient.

The contibution of thsi research will be a computational framework for hierarchical decision making integrating internal rewards. The proposed framework is shown in Fig. 3, where each decision can be composed from lower-level decisions. Based, on the current state, the agent has to learn which sub-decision must be activated to select the most appropriate action. Each decision has its own perception model of the environment according to the task it serves. Adapting the interaction based on multisensing data (facial expressions, audiovisual emotion, etc.) could be a lower level decision, while completing the task) is the overal (high-level) decision. The evaluation of this framework will focus on both learning convergence rates and interaction efficiency, following basic HCI evaluation methodologies.

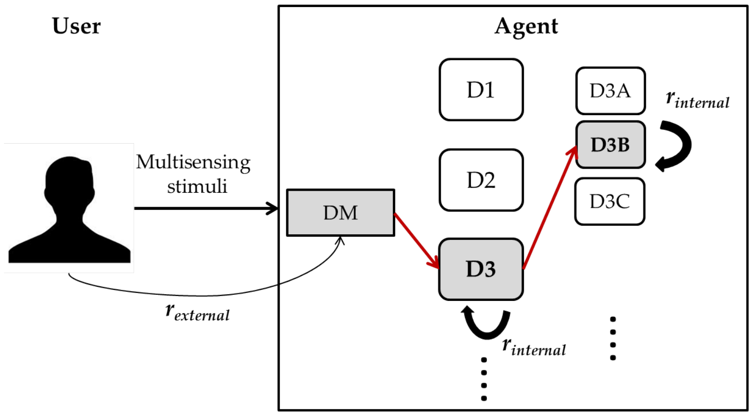

Fig. 3. In this proposed framework, each decision making module is composed of lower-level decisions, serving different tasks. Higer-level decisions activate lower-level decisions based the internal reward given the required task. In this way, complex action such as 'What and how to say it' can be made by a conversational agent or a companion robot.

Acknowledgments. This paper is partially supported by the project "Robots in assisted living environments: Unobtrusive, efficient, reliable and modular solutions for independent ageing - RADIO", which has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 643892. For more details, please see http://www.radio-project.eu

References

1. Poppe, R., Bck, R., Bonin, F., Campbell, N., de Kok, I., Traum, D. (2014). From multimodal analysis to real-time interactions with virtual agents. Journal on Multimodal User Interfaces, 8(1), 1-3.
2. Cassell J, Prevost S, Sullivan J, Churchill E (2000) Embodied Conversational Agents. MIT Press, Cambridge
3. Sutton, R. S., Barto, A. G. (1998). Introduction to reinforcement learning. MIT Press.
4. Ahn, H., Picard, R. W. (2005). Affective-cognitive learning and decision making: A motivational reward framework for affective agents. In Affective Computing and Intelligent Interaction (pp. 866-873). Springer Berlin Heidelberg.
5. Barto, A. G. (2013). Intrinsic motivation and reinforcement learning. In Intrinsically motivated learning in natural and artificial systems (pp. 17-47). Springer Berlin Heidelberg.
6. Jacobs, E., Broekens, J., Jonker, C. (2014). Emergent dynamics of joy, distress, hope and fear in reinforcement learning agents. In Adaptive Learning Agents workshop at AAMAS2014.