Safe Adaptation of Intelligent Interactive Systems

Principal Investigator: Prof. Fillia Makedon, Univ. of Texas at Arlington

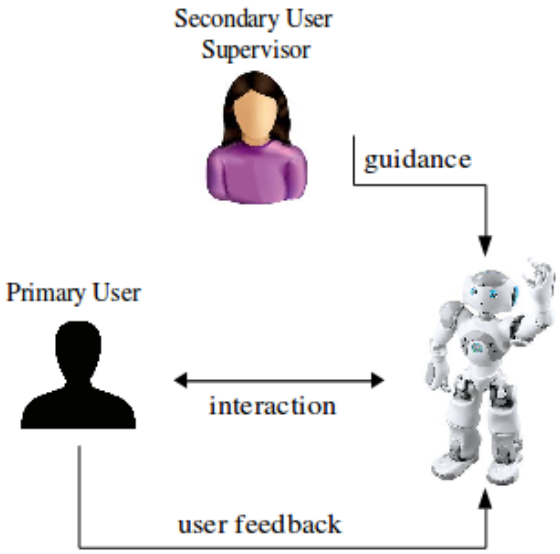
In this proposal, we present a computational framework and research methods that can facilitate the adaptation of Intelligent Interactive Systems to different users, ensuring a safe and efficient interaction. We follow the ***Reinforcement Learning*** framework that enables dynamical adaptation of a system to different users. We propose the use of Interactive Learning methods to exploit human knowledge and integrate it to the learning mechanism.

Need and Industrial Relevance

The need for Intelligent Interactive Systems (IIS) is essential in all applications that involve or depend on human-machine interaction. Building such systems to enable natural interaction is very hard because (a) they require intelligent technology that adapts to human abilities, limitations and preferences (b) users must be able to interact with such systems naturally but also in different modalities (voice, face, and other). Such systems need to interact with a variety of users. For an industrial product/service to be successful and usable, the interaction must be adaptive, tailored, safe and able to extract meaningful human interaction data for further improvement of its products.

Research Methods and Approach

We focus on application in the context of Assistive Environments, where a system interacts with a user to assist them perform a training task adapting their behavior towards each specific user [1]. Some possible applications are Rehabilitation Systems, Robot Aided Training, Vocational Therapy and Training, Human-Robot Collaboration Systems [2, 3]. Important attributes of such applications are *personalization* and *safety*. Human knowledge can be exploited and utilized to enhance system's learning ensuring a safe and tailored interaction [4] utilizing and combining Interactive Learning methods. The long term goal is a *lifelong learning* system that adapts to each specific user, as long as it interacts with them. In a real-world application, such a system should also include an expert user that supervises the interaction for safety issues. The proposed framework is shown in the next figure.



The primary user, who directly interacts with the system, can provide the system with implicit feedback during the interaction, that will enable the system adapt its behavior towards his/her specific abilities and preferences. The framework supports the participation of an external supervisor, as a secondary user, that supervises and observes the interaction, guiding the system during the early interaction steps, ensuring a safe and efficient interaction. This guidance also helps the system refine its specific goals based on the expert's decision making. Such a framework could have various applications to domains as Vocational Training and Rehabilitation, where the worker interacts with the system during a training task, as a primary user and the supervisor guides the system, as a secondary user. Moreover, human-robot collaboration systems can be successfully integrated with such a framework. The outcomes of this project will be an intelligent interactive system that has the ability to adapt safely to each specific user, integrating the knowledge of an expert that guides the early interaction steps and defines the desirable training goals.

References

1. Tsiakas, Konstantinos. *"Decision Making for Affective Agents in Assistive Environments."*, IVA 2015
2. Tsiakas, Konstantinos et al. "*A multimodal adaptive session manager for physical rehabilitation exercising*, ACM PETRA, 2015.
3. Tsiakas, Konstantinos, et al. "*A multimodal adaptive dialogue manager for depressive and anxiety disorder screening: a Wizard-of-Oz experiment*." ACM PETRA, 2015.
4. Tsiakas, Konstantinos. "Facilitating Safe Adaptation of Interactive Agents using Interactive Reinforcement Learning." Companion Publication of the 21st International Conference on Intelligent User Interfaces. ACM, 2016.