ICT PhD

SEMINAR Announcement

Wednesday 17th March 2021, Time:16.00-17.00
Microsoft Teams - https://bit.ly/3qGpPWm

Professor Joost Broekens

University of Leiden
http://www.joostbroekens.com/
Email: d.j.broekens@liacs.leidenuniv.nl

Emotions in Reinforcement Learning Agents

Abstract: Emotions are tied to appraisal of personal relevance, motivation and adaptation of behavior. Many animals show signs of emotion in their behavior. Therefore, emotions must be related to mechanisms that aid survival, and emotions must be evolutionary continuous phenomena. I propose that emotions are manifestations of Temporal Difference Reinforcement Learning (TDRL) error assessment. The TD error reflects the estimated gain or loss of utility – well-being – resulting from new evidence.

In this talk I will give an overview of my work on emotion modelling in reinforcement learning agents. I will introduce the Temporal Difference Reinforcement Learning (TDRL) Theory of Emotion. Then I will mostly focus on two topics: the kinds of emotions that can be modelled with this approach, highlighting some of my own work, and, how such emotions might be used for expression of emotion and robot/agent transparency.

CV: I study Artificial Intelligence, in particular Affective Computing and the interaction between humans and socially interactive agents. I am the President Elect of the Association for the Advancement of Affective Computing (AAAC), which is the main international organisation in my research field. I am an assistant professor at the Leiden Institute of Advanced Computer Science (LIACS) of Leiden University. I am head of the Affective Computing and Human Robot Interaction group at LIACS. Finally, I am also co-founder and CTO of Interactive Robotics, enabling students from any age to learn with and from social robots.

My research interests include computational modelling of emotions in reinforcement learning, computational models of cognitive appraisal, emotion psychology, emotions in computer games, explainability of AI and transparency, human perception and effects of emotions expressed by virtual agents and robots, emotional and affective self-report, human-robot and human-agent interaction, and educational humanoid robots. My current research focuses on human-robot interaction, and Reinforcement Learning as formal model for emotional appraisal.

Prof. Silvia Rossi – silvia.rossi@unina.it

Università degli Studi di Napoli Federico II