Announces the Ph.D. Dissertation Defense of

Mohammed Alharbi

for the degree of Doctor of Philosophy (Ph.D.)

“A Probabilistic Checking Model for Effective Explainability Based on Personality Traits”

March 15, 2022, 10:30 a.m.
Virtual Dissertation

Zoom
Meeting ID: 317 180 4062
Passcode: T6.3QsH5

DEPARTMENT:
Electrical Engineering and Computer Science

ADVISOR:
Shihong Huang, Ph.D.

PH.D. SUPERVISORY COMMITTEE:
Shihong Huang, Ph.D., Chair
Imadeldin Mahgoub, Ph.D.
Hari Kalva, Ph.D.
Behnaz Ghoraani, Ph.D.

ABSTRACT OF DISSERTATION
It is becoming increasingly important for an autonomous system to be able to explain its actions to humans in order to improve trust and enhance human-machine collaboration. However, providing the most appropriate kind of explanations – in terms of length, format, and presentation mode of explanations at the proper time – is critical to enhancing their effectiveness. Moreover, since explanation entails costs, such as the time it takes to explain and for humans to comprehend and respond, the actual improvement in human-system tasks from explanations (if any) is not always obvious, particularly given various forms of uncertainty in knowledge about the human.

In this research, we describe an approach to address this issue. The key idea is to provide a structured framework that allows a system to model and reason about human personality traits as critical elements to guide proper explanation in human and system collaboration. In particular, we focus on the two concerns of modality and amount of explanation in order to optimize the explanation experience and improve overall system-human utility. Our models are based on probabilistic modeling and analysis (PRISM-games) to determine at run time what is the most effective explanation under uncertainty. To demonstrate our approach, we introduce a self-adaptive system called Grid – a virtual game – and the Stock Prediction Engine (SPE), which allow an automated system and a human to collaborate on a game and stock investments. Our evaluation of these exemplars, through simulation, demonstrates that a human subject’s performance and overall human-system utility is improved when considering the psychology of human personality traits in providing explanations.

Keywords— Explainability, Human System Co-adaptation, Self-adaptive Systems, Human-Computer Interaction (HCI), Affective Computing, Personality Traits, Human-in-the-Loop, Model Checking, Probabilistic Modelling.

BIOGRAPHICAL SKETCH
B.S., Taibah University, Medina, Saudi Arabia, 2011
M.S., Florida Atlantic University, Boca Raton, Florida, 2018
Ph.D., Florida Atlantic University, Boca Raton, Florida, 2022

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION

Time in Preparation: 2018 - 2022
Qualifying Examination Passed: Spring 2019

Published Papers:


