

EXPLAINABLE AI FOR DESIGNERS:

A HUMAN-CENTERED PERSPECTIVE

ON MIXED-INITIATIVE CO-CREATION

Jichen Zhu (Drexel University, USA)

Antonios Liapis (University of Malta, Malta)

Sebastian Risi (IT University of Copenhagen, Denmark)

Rafael Bidarra (Delft University of Technology, The Netherlands)

G. Michael Youngblood (Palo Alto Research Center, USA)

INTRODUCTION

Growing interest in eXplainable Artificial Intelligence (XAI) aims to make AI and machine learning (ML) more understandable to human users. However, most existing work focuses on new algorithms, instead of usability, practical interpretability and efficacy on real users. We propose a new research area of **eXplainable AI for Designers (XAIID)**, specifically for game designers.

By focusing on a specific user group, and on their needs and tasks, we propose a **human-centered approach** for facilitating game designers to co-create with AI/ML techniques through XAIID.

RELATED WORK

Existing XAI Approaches

Black-Box AI

Existing work's focuses:

- visualize features
- elucidate the relationship between neurons

White-Box AI

Existing work's focuses:

- explanation of evidence
- explanation of the model
- explanation of the reasoning

Mixed-Initiative Co-Creative Systems

Numerous mixed-initiative tools have focused on visualizing properties of their specific artifacts. However, there is little research in explaining the creative process (rather than the final artifact) in co-creative tools; most attempts have focused on visualizations rather than on NLP generating the explanation.

eXplainable AI for Designers

The goal of XAIID includes investigating the actual **usability** of XAI in terms of how it supports game designers in specific design tasks.

XAIID can enhance game designers' capabilities to co-create playable experiences with AI, including but not limited to ML, agent control, procedural content generation, and planning.

EXPLAINING EXPLAINABILITY

1 AXIOM 1:
Explanation without introspection is not explanation

2 AXIOM 2:
Understanding through external probing is observation

3 AXIOM 3:
Reactive elements that have always been reactive from inception are not explainable, but may be observable

4 AXIOM 4:
The explanation of a reactive model lies in the deliberative process that created it

MAPPING THE SPACE OF XAIID

Spectrum of Explainability

Ranges from

- a) explanations that provide introspection into the operation of AI techniques, to
- b) observations that offer insights of the input-output pattern.

Spectrum of Initiative

Ranges from

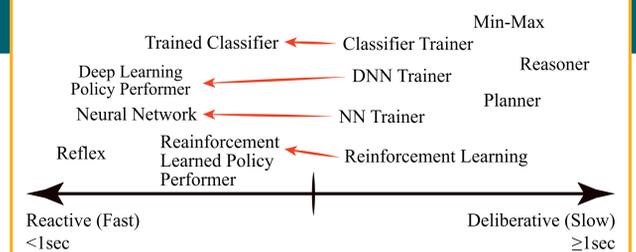
- a) passively providing on-demand explanations, to
- b) autonomously taking on activities that require an understanding of the design processes and goals

Spectrum of Domain Overlap

Ranges from

- a) AI and human designer making use of the same tools for the same task (on-task co-creativity), to
- b) a designer works on an aspect of the game while the AI handles another that is tangentially affected by the designer's input (off-task co-creativity).

Mapping deliberative AI/ML techniques to reactive processes



THREE USE CASES

1: White-Box PCG System

A level designer is using a computer-aided design tool to create the overworld map for a car-racing game. The tool generate the entire level, using grammars developed by a tool programmer.

XAIID can present the designer with compelling and intuitive narrative:

- summary of highlights of the generative process;
- sequentially in the order that the system makes decisions;
- non-sequentially, summarizing the explanation from the most important points.

2: Black-Box PCG System

A game designer is using a black-box PCG to design a city layout. The PCG assistant uses a DNN to recognize ideal topology for road placement, which is then placed by a set of construction rules.

In order to build common ground, the PCG assistant will interactively show the human designer how the provided land and topology are perceived and how its prior examples are used to generate roads based on the language provided.

3: Black-Box NPC Behavior System

A game designer has a new game level and wants to see if the enemy NPC, controlled by a trained DNN, will behaves as intended. The XAIID system can help the designer to better understand the NPC AI by:

- given the layout of the level, showing the likely distribution of actions the NPC will take;
- given a particular NPC action, showing all possible situations that can lead to this action and draw the designers attention to the salient ones.

OPEN CHALLENGES

White-Box Systems

- how to compact their process into clustered, consistent and meaningful activity information only?
- can a computational model of a designer help determine which data is relevant?

Black-Box Systems

- how to grow from simple interpretations to full explanations of their inner workings?
- how to provide meaningful, hierarchical abstractions of the transformation of training data into learned models?