Why don't computers improvise with us?

Brian Magerko
Georgia Institute of Technology
Atlanta, GA USA
magerko@gatech.edu

Duri Long
Georgia Institute of Technology
Atlanta, GA USA
duri@gatech.edu

Abstract
Improvisation—or the collective individual and social process of interactively creating meaning together—can be a useful lens for designing engaging human-computer experiences. Improvisation has been explored in the AI/computational creativity communities, but it has yet to be incorporated in HCI more broadly. Based on our previous research studying and building improvisational systems, we propose a set of improvisational knowledge types to aid developers in incorporating improvisational techniques in human-computer experiences and discuss implications.

CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
DOI: https://doi.org/10.1145/3334480.XXXXXXX
*update the above block & DOI per your rightsreview confirmation (provided after acceptance)

Author Keywords
improvisation; creativity; design

CSS Concepts
• **Human-centered computing~Human computer interaction (HCI)~HCI theory, concepts and models**

Introduction
Improvisation is a fundamental technique we use to communicate, express ourselves, experiment, and make sense of the world around us. We improvise when we have conversations, when we play, and when we cook. However, as integral as improvisation is to social human interactions, our interactions with technology rarely follows suit—human-computer interactions often tend to be scripted, predictable, and driven by the user. Contemporary AI research has begun to explore how to use improvisation as a way of enabling computers to interact with humans in creative domains like dance [8], theater [7,12], drawing [4], and pretend play [3]. The question we propose in this paper is—***how can improvisational techniques be leveraged in HCI to provide more natural and dynamic user experiences?*** Or, in other words, why don't computers improvise with us?

Figure 1: *LuminAI* (formerly *ViewpointsAI*), an improvisational dance agent

Figure 2: *The Robot Improv Circus*, a VR space for human-AI prop improvisation

Figure 3*: Drawing Apprentice*, a co-creative drawing agent

Improvisation is described by different researchers as a social process with a variety of lenses (e.g. [9,11,14]). We summarize a large body of prior work on improvisation in the following definition: *improvisation is the collective individual and social process of interactively creating meaning together*. Prior work suggests that improvisation is a potentially useful metaphor for designing human-centered experiences with technology. There has been a push in the HCI community towards designing engaging human-computer experiences that promote curiosity, open-ended exploration, and diversion rather than performing a specific utilitarian function [13]. There is also a growing need for technologies that can easily adapt to novel, unforeseen circumstances, as computers are being introduced in ever-more complex and nuanced aspects of human life. Understanding and applying improvisational techniques to the development of AI agents has supported more open-ended, adaptable interactions (e.g. [4,8]), suggesting that these same techniques could also aid designers more broadly in designing engaging, exploratory human-computer experiences.

Improvisation Knowledge for HCI

We propose a minimal set of knowledge types that afford improvisational interaction between agents based on our previous research on improvisation and development of AI agents in theater, dance, music, and collaborative drawing (Fig 1-3). These knowledge types are: 1) *interactively-learned,* 2) *tacit*, 3) *interactional*, and 4) *transformational* knowledge.

Interactively-learned knowledge points to how authoring content or data for computing systems in open-ended improvisation is notoriously difficult, if not impossible. Current approaches look at how deep learning can learn models from available data sources (e.g. training a narrative improv agent on dialogue from literature [10]); however, these and previous strategies relying on hand-authoring ultimately have a difficult time capturing the kinds of social dynamics described above. We have explored how interactive learning can be employed to learn actions in open-ended spaces, such as an agent improvising with objects it has never seen before [3,7] or learning novel dance movements [8] (Fig 1, 2).

Tacit knowledge is the collection of behavioral conventions and formal terminologies involved in a particular domain. For instance, we have studied the tacit knowledge implicit in improvisational theater, which has led to a formal model of how actors (and agents) can improvise the beginning of a scene, how they find the 'point' of a scene, and how they establish characters over time in a scene [2,12]. This kind of knowledge enables an agent to perform according to the specific conventions that are needed for improvising within a specific domain.

Tacit knowledge can also come from existing formalizations within the relevant creative domain. For instance, Viewpoints and Laban movement theories are languages that have been used for decades to formally describe movement in dance and performance. Having a formal language to describe interactions in an open-ended space can provide an enormous bootstrap for interactive learning, for clustering approaches to learned moves, and for visualizing learned data [7].

Transformational knowledge refers to using tacit knowledge as the language to construct rules / policies

/ trained networks that can generate new moves. For example, a transform using the Viewpoints movement language could take a learned example of a human waving their hand goodbye and generate a new move with a different **tempo**, **duration**, and **shape** [8] that has structural similarities, but different temporal ones. Analogical reasoning has been found to be another effective method of transforming knowledge in improvisational systems, where knowledge is generated based on structural commonalities with authored or learned knowledge [16].

Transformational knowledge also includes a way of discerning 'good' generated moves from 'bad ones' [9]—and without any gold standard of ideal behavior to compare against. For example, an agent ideally should be able to recognize how to perform interesting and enjoyable dance moves based on its learned data. We contend that this can be solved with a heuristic about what a 'creative move' is, which we and others have previously defined computationally as the value, surprise, and novelty of a thing [7]. Furthermore, one can consider a 'creative arc', which is how the different elements of creativity (value, surprise, and novelty) ideally modulate over time (e.g. generating moves with low surprise early in an experience and increasing over time to highly surprising moves) [7].

Interactional knowledge encapsulates the social decision making processes involved in improvisation. This ultimately dictates what moves are most appropriate at any given time. We have found across multiple expressive domains that *turn-taking* is one central component to this process [4,8]. Turn-taking refers to the give and take of leadership in an exchange (e.g. how two dancers dynamically negotiate leader/follower roles through movement). This ebb and flow of control is evident in the ways in which we converse, dance, and design together. It is critical to understand how to identify, participate in, and transition between different roles related to control in an improvisational system.

Interactional knowledge also encompasses the participatory sensemaking process we undergo together while co-creating [4]. The ways in which we negotiate meaning in a shared space is crucial to how we improvise and create together in real-time. Improvisational moves between both creative experts and novices improvising in everyday interactions have their roots in basic interactional patterns like "mimicry" or "call and response" [1,14].

Finally, the ways in which we negotiate a shared understanding of the creative world that we are improvising in is crucial in collaboration. The construction of *shared mental models* over time is one lens to examine this process. The shared mental model (c.f. [5]) and common ground literatures [15] point to a wealth of processes for how humans naturally negotiate shared understanding during conversation and common interactions. We argue that sensemaking processes like these are crucial to having more naturalistic interactions between humans and computers.

Discussion

We have presented the kinds of formal knowledge we have seen a need for in our prior work developing improvisational AI systems. How could these findings translate more broadly to human-computer interaction? There are already inklings of improvisational behaviors

in creativity support tool research, where interfaces augment the user's creative process and in intelligent user interface research, where user modeling is has often been employed to react to the perceived user's mental state. Co-creativity is also increasingly becoming a core topic in the computational creativity community as well (c.f. [6]). We hope that by formalizing these requirements for improvisational human computing we can bolster further exploration of improvisation in the HCI community.

References

[1] Margaret A Boden. 2004. *The Creative Mind: Myths and Mechanisms*. Routledge, New York, NY, USA.

[2] António Brisson, Brian Magerko, and Ana Paiva. 2011. Tilt Riders: Improvisational Agents Who Know What the Scene is About. In *Intelligent Virtual Agents 10th International Conference, IVA 2011, Reykjavik, Iceland, September 15-17, 2011. Proceedings*.

[3] Nicholas Davis, Margeaux Comerford, Mikhail Jacob, Chih-Pin Hsiao, and Brian Magerko. 2015. An enactive characterization of pretend play. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, 275–284.

[4] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative sense-making: Quantifying interaction dynamics in co-creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*, 356–366.

[5] Daniel Fuller and Brian Magerko. 2010. Shared mental models in improvisational performance. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, 15.

[6] Matthew Guzdial and Mark Riedl. 2019. An interaction framework for studying co-creative ai.

[7] Mikhail Jacob. 2019. Improvisational Artificial Intelligence for Embodied Co-creativity. Georgia Institute of Technology.

[8] Mikhail Jacob, Gaëtan Coisne, Akshay Gupta, Ivan Sysoev, Gaurav Verma, and Brian Magerko. 2013. Viewpoints AI. In *Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE '13)*.

[9] Philip N Johnson-Laird. 2002. How jazz musicians improvise. *Music Perception: An Interdisciplinary Journal* 19, 3: 415–442.

[10] Lara J Martin, Brent Harrison, and Mark O Riedl. 2016. Improvisational computational storytelling in open worlds. In *International Conference on Interactive Digital Storytelling*, 73–84.

[11] David Mendonça and William A Wallace. 2004. Cognition in jazz improvisation: An exploratory study. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

[12] Brian O'Neill, Andreya Piplica, Daniel Fuller, and Brian Magerko. 2011. A Knowledge-Based Framework for the Collaborative Improvisation of Scene Introductions. In *ICIDS*, 85–96.

[13] Yvonne Rogers. 2006. Moving on from Weiser's Vision of Calm Computing: Engaging Ubicomp Experiences. In *International conference on Ubiquitous computing*, 404–421.

[14] Keith Sawyer. 1993. The pragmatics of play: interactional strategies during children's pretend play. *Pragmatics* 3, 3: 259–282.

[15] Robert Stalnaker. 2002. Common ground. *Linguistics and philosophy* 25, 5/6: 701–721.

[16] Tony Veale and Guofu Li. Specifying Viewpoint and Information Need with Affective Metaphors.