Modalities, Styles and Strategies: An Interaction Framework for Human–Computer Co-Creativity

Anna Kantosalo[1], Prashanth Thattai Ravikumar[2], Kazjon Grace[3] and Tapio Takala[1]
[1]Department of Computer Science, Aalto University School of Science, P.O.Box 15400, FI-00076 AALTO, Finland
[2]Communications and New Media, National University of Singapore, 21, Lower Kent Ridge, Singapore 119077
[3]Design Lab, The University of Sydney School of Architecture, Design and Planning, 148 City Road, Darlington NSW 2008 Australia
anna.kantosalo@aalto.fi, prashanth.thattai@u.nus.edu, kazjon.grace@sydney.edu.au, tapio.takala@aalto.fi

Abstract

Human-computer co-creativity research is relatively new, and comparing how co-creative systems augment creativity is challenging even within the same creative domain. This paper proposes a framework to facilitate such comparisons, drawing on domain-agnostic concepts in interaction design research. It describes the different aspects of the interactions between one or more humans and their computationally creative collaborators at three levels: interaction modalities, interaction styles, and interaction strategies. Modalities are the channels by which information is exchanged, styles are the different behaviours and models that govern the system's actions over those channels, and strategies are the goals and plans that shape those behaviours. The paper ends with an analysis of nine co-creative systems, demonstrating how the framework makes the comparison of very different systems possible.

Introduction

Human–computer co-creativity is a subfield of computational creativity dedicated to the study and design of interactive, creative computational collaborators. These computationally creative collaborators interact with humans through a user interface, and together they form a co-creative system. Co-creative systems are distinguished by their shared locus of creative initiative. They represent a middle ground between autonomous creative systems, which are intended as the sole shepherds of their own creativity, and creativity support systems, which instead facilitate the creativity of their users. In recent years computationally creative collaborators have been designed and implemented for many creative domains, including drawing (Davis et al., 2015), poetry (Kantosalo et al., 2015), and music (Ravikumar and Wyse, 2019).

This interstitial and emerging research field has a number of open questions including how to design (Kantosalo et al., 2014), evaluate (Kantosalo et al., 2015; Karimi et al., 2018b) and describe (Kantosalo and Toivonen, 2016; Davis et al., 2015; Yannakakis, Liapis, and Alexopoulos, 2014) co-creative systems. Put simply: co-creativity is a goal as broad as creativity itself, and made yet more complex by the many and varied ways humans and computers can interact. As the field matures it will be necessary to cohere around a set of domain-general ways to describe the capabilities of co-creative systems. One attempt at finding a unified framing for co-creative systems has been to look at how users and systems behave while creating (e.g. Kantosalo et al., 2015; Karimi et al., 2018b; Long, Jacob, and Magerko, 2019). This paper strengthens that tradition with a framework for describing interactions with co-creative systems.

Interaction design—a human-computer interaction (HCI) discipline for designing interactions with products and services—has been suggested as one way to frame the design (Kantosalo et al., 2014) and evaluation of co-creative systems (Bown, 2014) in a behaviour-centric way. We extend this suggestion, drawing several terms from the interaction design literature to construct a framework for describing the behaviour of computers and users within human-computer co-creative systems. The goal of the framework is to equip co-creativity researchers with a domain-agnostic vocabulary to discuss the capabilities and shortcomings of existing and proposed interfaces for co-creation.

Our framework draws from three traditional interaction design concepts: interaction modalities, styles, and strategies. Together they form a layered description of interactions between a human and a computationally creative collaborator, from the most elemental level describing different qualities of their interaction channel, right through to the higher-level reasoning processes that govern how the computer selects and achieves its objectives.

We begin with a brief discussion on the background of these key interaction concepts. We then move on to defining our three interaction layers, discussing how they have been used in HCI as well as how they apply to co-creative systems research. We then show how this framework facilitates the analysis of nine co-creative systems. These systems are selected not as a comprehensive review of the current state of the art in human–computer co-creativity, but as a demonstration of the breadth of our framework. We end with a discussion of some new areas of co-creative systems research that our framework suggests are as-yet underexplored.

Background

As recently as the mid 1970s, the notion that computer systems should be in any way designed to be easy for their non-expert users to operate was largely treated with derision (Carroll, 1997). HCI as a research field and interaction design as its associated professional design domain have arisen

Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)
ISBN: 978-989-54160-2-8

in the four decades since. New kinds of design processes were needed for the digital era, as neither a focus on visual communication (as in graphic design) nor on physical consumer products (as in product/industrial design) were sufficient to tackle the design of digital products and services. Interaction design and HCI focus on the experiences evoked by digital products as well as their utility and aesthetics.

Interaction can have multiple purposes, such as communication of information, chit chat, refining common goals through discussion, planning a strategy, or giving commands (Barfield, 1993, pp. 208-209). It can also take many forms, including exchanging information with a system or direct manipulation of the properties of the system (Barfield, 1993, pp. 207-208). Research in the field of interaction design has sought to characterise the ways people perceive and interact with complex as well as creative (Shneiderman, 2007) tasks, providing literature which we draw on to construct our framework below.

Three terms emerge from interaction design literature for describing the interactions between humans and computers at different levels of abstraction. The first, modality, considers the physical properties of the interaction, focusing on information channels and nature of exchanged information. The second, interaction style, focuses on the design of the interface and how it supports different conceptual models and behaviours related to interaction. The final, interaction strategy, considers more abstract elements of interaction, such as how the different interaction styles evolve throughout interaction. These three concepts form the basis of our framework and we shortly discuss their use in interaction design literature below.

The HCI origins of interaction modalities

On a basic level, interactions between a human and a computer are limited by the affordances of the interaction participants, such as the input and output devices available to each party. Questions related to the selection of interaction modalities are studied especially by researchers interested in multi-modal interaction (Granström, House, and Karlsson, 2013, p.2). Different modalities can affect how additional or different information can be conveyed (e.g. Bernsen, 2002), as well as make interaction more natural and immersive (O'hara et al., 2013; Tham et al., 2018).

The notion of modality is a useful ground zero for describing co-creative systems. The key question here is what are the input and output channels of the human and computational collaborators? With the basic possibilities of a specific co-creative collaboration so established, it is then possible to examine their nature further.

The HCI origins of interaction styles

Interaction style is a somewhat loosely used term dealing with *how* communication happens between the system and its user. For example, it can be "conversational" or "direct manipulation" (Barfield, 1993, p.215). In graphical user interfaces interaction is based on "widgets" such as menus, forms, dialog boxes, or icon sets (Sutcliffe, 1988, pp.68-75). Other modalities, such as speech (natural language) and embodied motion (gestures) also make their own interaction styles. In this paper we adopt the definition by (Hix and Hartson, p. 57): "Interaction styles are a collection of interface objects and associated techniques from which an interaction designer can choose when designing the user interaction component of an interface. They provide a behavioral view of how the user communicates with the system."

Similar discussions of the ways, both broad and narrow, how interactive systems can behave during co-creation are almost entirely absent from this community's description of its systems. Where such descriptions do occur, such as in (Bray and Bown, 2016), they use bespoke terminology, making comparisons to other systems and other domains challenging. The field of computational creativity tends towards describing its systems in terms of their generative capacity and their evaluation metrics. Co-creativity researchers often add to that a description of what we would call their systems' interaction modalities, especially as regards the communication between user and agent. What is missing from these descriptions is a generalisable way of discussing interaction style: how do the user and the agent collaborate? And what impact does that choice have on creativity?

The HCI origins of interaction strategies

Where interaction styles describe the *way* a co-creative system interacts with its user to perform creative tasks, interaction strategies govern *why* it chooses one available action over another. At its simplest this is expressed through evaluation metrics such as value, quality, novelty, surprise, diversity, and so on. These cornerstones of computational creativity have extra importance in co-creativity research because of the need for collaboration: ignoring the user and searching for optimality does not make a good creative partner. It is therefore beneficial, in many cases, to adapt the search towards what the user appears to be doing. This has the potential to be more complex than adding additional evaluation metrics, and may include personalising or prioritising the evaluation process, or even a meta-search for new metrics or generative procedures (Wiggins, 2006).

The term "strategies" has been used to describe how interactive intelligent systems (which co-creative systems can be considered a kind of) adapt. It has been applied to describe how conversational agents might take different approaches based on past user behaviour (Schuller et al., 2006), and similarly to govern what educational content to deliver based on estimates of student mastery in intelligent tutoring systems (Al-Nakhal and Naser, 2017).

In our framework we define the term "interaction strategy" to refer to the system's evaluative metrics, its underlying goals, and the meta-reasoning process (Cox and Raja, 2007) by which it adjusts its metrics to best achieve those goals. Many systems do not implement all three parts of that definition, which we consider an underexplored and promising direction for future systems.

A Layered Framework for Interactions with Computational Collaborators

We have adapted the above concepts for human-computer co-creativity. As shown in Figure 1 they form a three layered

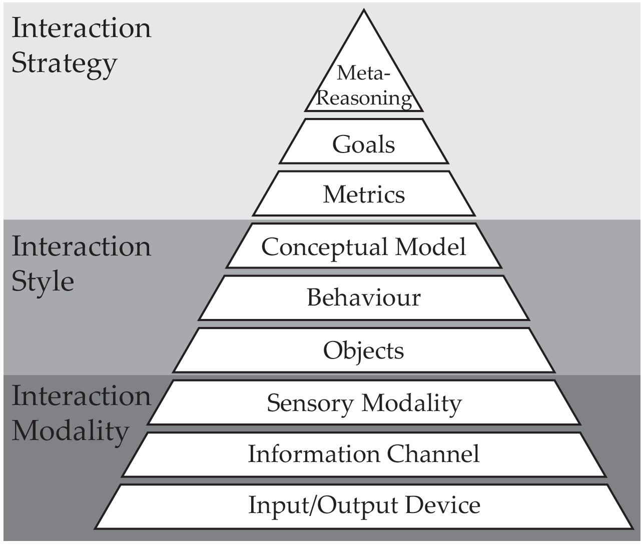

Figure 1: Interaction layers

framework for describing interactions with a computationally creative colleague. The first layer, *interaction modality*, describes the channels and media of the interactions. The second layer, *interaction style*, builds on the modality, focusing on conceptual interaction and behaviours. The final layer, *interaction strategy*, allows for designing and describing systems with more elaborate goals and interaction plans, than traditional productivity systems.

Interaction Modalities

Definition: Interaction modality describes the medium of communication implemented through one or more sensory channels. There may be specifications related to each channel, such as restrictions or constraints on the flow of information through each channel.

Interaction modalities describe the fundamental elements of sending and receiving information between a human and a computer. We consider them as unique combinations of different input and output devices, information channels, and sensory modalities. While interaction can be unimodal, using only one modality, interactions are usually multimodal, combining different modalities (Bernsen, 2002). Multimodality is used to increase the naturalness of human–computer interaction, to provide different types of information (Granström, House, and Karlsson, 2013, p.1) and to reduce errors in productivity systems (Jaimes and Sebe, 2007), but it also plays a role in creativity support. For example the modalities draw, write, talk and gesture have all been found important for supporting collaborative design (Eris, Martelaro, and Badke-Schaub, 2014).

Sensory modalities relate interaction modalities to human senses, such as hearing, smell, sight, touch, taste (Jaimes and Sebe, 2007; Schuller et al., 2006), the sense of balance (Schuller et al., 2006) or more obscure senses like the senses of temperature, kinesthetics, or vibrations.

Information channels are often associated with sensory modalities: The sense of sight typically relates to the visual channel, sense of hearing to the auditive channel, etc. (Schuller et al., 2006). The same sensory channel can be used to carry different kinds of information, and some information is suitable for multiple channels. For example the visual channel can be used to carry images or linguistic information, but linguistic information could also be carried over an auditory channel. Information can also be mapped or transposed to different channels as part of various representations, such as in a musical stave or a choreographer's notes. Channels may be simplex or duplex, and there is no requirement for the channels used in a co-creative system to be symmetrical between human and computer.

Input and output devices are used to implement different information channels. Example input devices, conveying information from the human to the computer, include keyboard, touch screen, mouse, joystick, hand writing and speech recognition, and output devices for conveying information from the computer to the human include displays, data projectors, sound, speech, printers, plotters, and haptic devices (Benyon, Turner, and Turner, 2005, pp.149–157).

Information from one agent may go through preprocessing when it travels through a channel or device to another agent. This may significantly change the format of the information. For example, the Drawing Apprentice (Davis et al., 2015), takes human input given via a tactile channel with a pointing device, but interprets it visually as pen strokes.

Interaction Styles

Definition: Interaction styles describe how interaction is structured between the human and the computational collaborator. They describe the inter- and intra-part relationships which distinguish different periods of interaction from one another. Multiple styles can be manifested in a single system. Styles build on the symbolic representation of objects, linking them with different conceptual models of interaction and behaviours of the computational collaborators.

Interaction styles are design paradigms describing how users interact with a product. They can be used as tools for thinking how different conceptual models of interaction could be realised through different interaction modalities (Rogers, Sharp, and Preece, 2011, p.206). With co-creative systems, we consider that an interaction style emerges from the interplay of a conceptual model of interaction and the behaviours of a system. The interaction between the two is mediated by objects that provide a shared context for the system and the user.

Objects act as a conduit between interaction modality and the rest of the interaction style. They are representations of important shared concepts and interface control elements that can be manipulated or interacted with during co-creation. They are often multimodal and may include an interpretation of the meaning of the object in addition to its physical representations in a medium[1]. For example the creative artefact under construction may be represented as an

[1]Medium is at times used as an aspect of modality describing the physical realisation of information at the human–computer interface (see Bernsen (2002)). Due to the possibilities of mixing the term medium with the artistic medium, we have chosen the word

object. Such objects often have a specific iconic representation like the notation used in sheet music, or the layout of a poem on a piece of paper. The full understanding of these visual objects requires knowledge of the representation style and possibly a translation to another modality, such as audio.

Behaviours describe the actions of the human and the computational collaborator during interaction. These actions are linked to specific objects on the interface. They can involve concrete manipulation of the objects, be initiated through an object, or they can relate objects to each other. For example a human could directly manipulate a word object in a co-creative poetry system by editing the word, then request validation of the rhyme from the computer by clicking a button object, and finally the computer could return a validation of the word object, with respect to the context provided by other word objects in the poem object.

Conceptual models structure the interaction between the human and the computer. Example behaviours of co-creative collaborators include shadowing, mirroring, coupling and negotiating interactions (Young and Bown, 2010), operation based, request based and ambient interaction (Bown and Brown, 2018) and additive and iterative interaction styles (Clark et al., 2018). General terms like highly encapsulated systems, direct manipulation and programmable interfaces can be used for describing different levels of control and visibility in human-computer co-creative systems (Bray, Bown, and Carey, 2017).

For the purpose of analysing interaction between system and the user, conceptual models are classified into contribution structures, and interaction models. In additive interactions, the human and the system combine their contributions to build a shared object, whereas in iterative interactions, the human and the computational system evaluate their contributions to successively refine a shared object (Clark et al., 2018). Interaction models are characterized as operation based, request based and ambient interaction models of interaction by Bown and Brown (2018). In operation-based interaction styles, human users directly manipulate a generative algorithm to produce outcomes. In systems that use a request-based interaction style, human users submit requests to a system that returns results. In ambient interaction systems, the system listens to the user's actions and generates its responses through autonomous meta-creative processes that run in the background.

Interaction Strategies

Definition: A co-creative system's interaction strategy governs what it values at any given moment, how those values change, and why. We consider an interaction strategy to consist of metrics (how a system values creative artefacts), goals (the motivational drives by which a system alters its metrics), and metareasoning (the way a system derives metrics from its goals).

The notion of an interaction strategy gives us a way to examine different goals of a computational collaborator, going beyond the discussion of its evaluation metrics. We intend the term to subsume both its criteria for determining what is good and its reason for adopting those criteria. We want to declare an open bias in constructing our framework this way: we want the discussion around co-creative systems to more seriously entertain adaptive evaluation metrics that adapt to user behaviour. Separating goals and metrics allows us to describe systems that adapt their generative capacity to the user and/or their creations. Such systems might, for example, switch between different metrics (or even generate whole new metrics) in order to encourage their human collaborators in one direction or another.

Metrics are a core component of research in both autonomous and collaborative creative systems (Jordanous, 2012). They give creative systems a way to discern valuable creative behaviours. Much of the formative discussion in the first ten years of ICCC revolved around what metrics were necessary and sufficient for evaluating creativity. These questions also apply to co-creative contexts, although with some nuance due to the presence of a human collaborator. We see metrics in co-creative systems as falling into three broad categories:

1. Metrics of *value*: evaluation of the quality of the resulting artefact, including appropriateness, utility, and aesthetics.
2. Metrics of *novelty*: evaluation of the divergence of resulting artefacts, including surprise and diversity.
3. Metrics of *interaction*: evaluation of the user or the way the user and the system are behaving together, rather than evaluations of the artefact directly.

Many co-creative systems have only value metrics, relying on stochastic processes and human curation to inject meaningful divergence into creative artefacts. Such systems can suggest novel artefacts to humans, but not reason or meta-reason about novelty themselves, as they have no way to detect it.

Goals—as a seminal topic in AI research—can play many roles in creative systems. Here we focus on strategic goals, defined as a desired state that the computer aims to bring about by varying its evaluation metrics. A system with fixed evaluation metrics has only *implicit* goals: they are hard-coded into its construction. By contrast, goals may be *explicitly* represented within an agent (and therefore potentially reasoned about). Goals can also be *interactive*: able to be manipulated by the user through some interface, giving control over the interaction strategy to the human.

Meta-reasoning in an interaction strategy refers to the process by which a co-creative system arrives at its current evaluation metrics given its current goals. Meta-reasoning describes how an agent reasons about (and affects) its own reasoning processes (Cox and Raja, 2007). In our framework, the system reasons about how to adapt its evaluation metrics to achieve its goals. Framing metareasoning as part of an interaction strategy adapts research in creative meta-search (Wiggins, 2006) to a co-creative context. In the co-creative case meta-search does not occur in isolation, because there are multiple creative agents (at least one human and one artificial) that are attempting creative search in parallel. This means that an agent may need to adapt its search space not only to the progress of its own search but to the

object and use it broadly combining representation and content.

behaviour of its human collaborator(s) (see e.g. Kantosalo and Toivonen (2016)). There are several possible "levels" of metareasoning for co-creative systems:

- *No metareasoning* occurs without explicit goals.
- *Metric prioritisation* occurs when a system weights (or otherwise chooses between) its evaluation metrics given the current context.
- *Metric formulation* occurs when a system creates or modifies an evaluation function, like through meta-search.
- *Goal reasoning* occurs when a system uses higher-level goals or drives to adapt its goals, which in turn allow it to adapt its metrics by one of the above methods.

Those last two categories of metareasoning are speculative: we do not know of any systems that are best described in that way. They are not, however, unimaginable: goal reasoning could be governed by more abstract motivations like curiosity (Grace and Maher, 2016), autonomy (Jennings, 2010), goal-awareness (Linkola et al., 2017), or intent (Cook and Colton, 2011). Such discussions are a regular feature in autonomous creative systems, and this framework provides a place for them in co-creativity.

Analysing Co-Creative Interactions

To examine how interaction modalities, styles and strategies are utilised by human-computer co-creative systems, we analysed nine systems from different application areas of human-computer co-creativity. Three of the systems, Masse (Ravikumar and Wyse, 2019), Shimon (Weinberg et al., 2009) and MiMi (François, Chew, and Thurmond, 2007) are all interactive improvisation systems focusing on different types of music. While Masse and Mimi synthesise music, Shimon robotically plays the instrument marimba. Three systems, the Poetry Machine (Kantosalo et al., 2015), LyriSys (Watanabe et al., 2017), and an unnamed story generation system that we hereafter refer to as "Clark et al.'s story system" (Clark et al., 2018) represent different interfaces in the domain of linguistic co-creativity. The first focuses on poetry, the second on lyrics, and the final on the collaborative writing of prose. The final three systems are drawn from other domains of human-computer co-creativity, including drawing (Creative Sketching Apprentice (Karimi et al., 2018a)), improvised dance (ViewpointsAI (Jacob et al., 2013)) and game level design (Sentient Sketchbook (Liapis, Yannakakis, and Togelius, 2013)). The systems and their domains are summarised in Table 1. To demonstrate the capacity of our framework to describe and compare co-creative systems, we examine the modalities, styles and strategies used by these tools below.

Modalities in the Sample Systems

The modalities used by the different systems can be divided into modalities focused on conveying domain specific information and modalities conveying other information. For example, the MASSE system only conveys music specific, intra-musical information through a full-duplex auditory channel and a tactile input channel consisting of a

System	Reported in	Domain
MASSE	Ravikumar and Wyse (2019)	Music
Shimon	Weinberg et al. (2009)	Music
MiMi	François, Chew, and Thurmond (2007)	Music
Poetry Machine	Kantosalo et al. (2015)	Linguistic
Unnamed Story Generation System	Clark et al. (2018)	Linguistic
LyriSys	Watanabe et al. (2017)	Linguistic
Sketching Apprentice	Karimi et al. (2018a)	Drawing
Viewpoints AI	Jacob et al. (2013)	Dance
Sentient Sketchbook	Liapis, Yannakakis, and Togelius (2013)	Game level design

Table 1: Analysed co-creative systems and their domains.

Midi-keyboard. Similarly, the linguistic system Poetry Machine focuses on intra-linguistic information and uses text suggestions as an important communication medium.

Most other systems seem to employ both domain specific and general modalities in important roles: For example the two other musical systems, Mimi and Shimon both use sound and sight. Shimon adds a headbobbing movement to represent rythm, for example. Clark et al.'s story system uses extra-linguistic static visual prompts for inspiring the user to start writing a story and the Lyrisys lyrics writing system visualises extra-linguistic cues on top of the co-created lyrics such as the syllable structure, syllable count, and story. The co-creative sketching tool Sketching Apprentice provides a verbal description of the sketch to draw to the user, reversing the use of an image to prompt text in the unnamed story writing system described above.

However the dualism for intra- and extra-domain information does not fit all domains of co-creativity perfectly: The interactive dance performance system ViewpointsAI and the game level design system Sentient sketchbook are interesting contrasts to the previous systems, as dance and games are themselves multimodal experiences that require the integration of several modalities and temporal effects. They therefore use representations, such as animations of dance movements or visual representations of game levels to convey important information.

The timing factors related to the use of different channels is also interesting. The musical system Shimon's modalities are synchronous: it nods its head in time with the beat This tight coupling of visual modalities with human input is similar to the ViewpointsAI dance system. The Mimi system offers a counter example: it uses the auditory channel to produce music that is played presently, and the visual channel to indicate music that will played in the future. Some systems also restrict the use of different channels following a specific form: For example the story system by Clark et al. allows the human and the computer only to input one line at a time into to the story, but several other systems allow for the human to continuously modify the shared artefact and ask feedback for it. These systems include the linguistic systems Poetry Machine and Lyrisys, as well as the Sketching Apprentice and Sentient Sketchbook.

Interaction Styles in the Sample Systems

There are several different conceptual models; ambient, request-based and operation-based interaction, among the examined systems. These connect to different objects and behaviours within the systems.

Ambient interaction is used by MASSE, Mimi, Shimon, Viewpoints AI and the Sketching Apprentice. The first two musical systems, MASSE and Mimi, use ambient interaction models with additive contributions that are simultaneously played with the input. Both systems listen to a predetermined number of musical bars before they start to play along with the musician. The rythmic objects generated and listened by MASSE are 2-bars in length, while the audio input and visual music notation output of Mimi are arbitrary in length.

Shimon, Viewpoints AI, and the Sketching Apprentice, use an ambient interaction model to produce additive contributions in a turn-taking fashion. The objects used by the systems are domain specific: Shimon listens to musical objects, the Viewpoints AI gathers stylized human movements, and the Sketching Apprentice uses visual line drawings as interaction objects. Corresponding objects are used by the systems to listen for interaction events to decide when their turn begins. Viewpoints AI and the Sketching Apprentice are similar in their interaction behaviors and contribution types, producing interaction behaviours through transformation functions, while Shimon predominantly attempts to imitate the style of the musician. Then again, both Shimon and Viewpoints AI produce contributions that combine with the user input and are additive, while the Sketching Apprentice system generates suggestions to refine a creative artefact, showing an example of iterative contribution.

A request-based model of interaction is used by Poetry Machine, Clark et al.'s story system, and the Sentient Sketchbook. The users of these systems modify a shared artefact, and send requests to the system for feedback or additional materials. The two linguistic systems allow the user to directly manipulate object representations through a WYSWYG editor, whereas the Sentient Sketchbook provides a tile-based interface for the user to make design changes to the game world. All systems offer means for direct user manipulation. In the Poetry Machine system, the user directly makes changes to poem text and sends a request to the system to validate the rhyme, or request additional materials. In response, the system returns a validation of the poem, or materials suitable for a specific context. In the story generation system, the user directly manipulates the sentences that are automatically generated for every line and adopts them into the story. The Sentient Sketchbook system uses similar methods and also allows the users to directly manipulate the game-level, send requests, and adopt the suggestions generated by the system.

The final system, Lyrisys, is an example of a system that uses an operation-based model with direct manipulation of the interface objects. The Lyrisys system provides interface objects as parameters (e.g., syllable count, lyric and story) that the user manipulates to generate the lyric. The user changes parameters and pushes a button to generate the lyrics that are consistent with other context. In contrast to the previous systems, the user makes no changes in the content of the artefact but observes the behaviours of the system in response to parametric changes.

Interaction Strategies in the Sample Systems

The musical systems Shimon and Mimi both use a static interaction strategy to select their musical responses: they have implicit goals and no meta-reasoning. Mimi does provide an interface that allows the performer to evaluate the contributions of the system and change the preparatory musical material used by the computational collaborator, but its metrics remain unchanged by this configuration. Mimi has no metrics, being purely stochastic, while Shimon uses an interaction metric: the degree to which it is musically conforming with its collaborators. The MASSE system, by contrast, has an explicit goal to maintain a target level of stability and togetherness in the performance. It senses deviations from the goal state and sets targets for its, a kind of metric prioritisation. It has two metrics: stability (a value metric), and togetherness (an interaction metric).

The Poetry Machine and Clark et al's story system both have value metrics: rhyme, meter and alliteration for the former, and word-level conditional probability for the latter. Both can generate novel artefacts stochastically, like Mimi and Shimon, but do not have novelty metrics. Both also have implicit goals and no meta-reasoning. Lyrisis also has value metrics (meter, rhyme, and topicality) but differs in that the human can interactively select a goal topic.

Viewpoints AI has a variety of response modes that it switches between randomly: repeating the user, picking a novel gesture from its library, or transforming it (by reflecting a limb, switching limbs, or duplicating movements to other limbs). In various modes it applies a tree-based gesture similarity metric to either conform with (an interaction metric) or diverge from (a novelty metric) the user. Each mode is effectively a separate generative system, and its random switching between them means it has no explicit meta-reasoning. The Sketching Apprentice has an evaluation model based on visual and conceptual similarity. In one version of the system it maximises visual similarity (a value metric) while minimising conceptual similarity (a novelty metric). In another version the target levels for both similarities are user-configurable, meaning it has interactive goals. The Sentient Sketchbook also has interactive goals, providing several sliders and check-boxes for the user to control the weights of its wide variety of value metrics (playability, safety, resource availability, explorability, symmetry, and several forms of balance) plus novelty.

None of the nine systems used a more complex meta-reasoning strategy than metric prioritisation, suggesting that creative meta-search and other forms of goal-based reasoning have yet to be broadly applied in co-creativity research.

Discussion

We have suggested a framework for describing interactions within human-computer co-creative systems at three levels

- interaction modalities, interaction styles, and interaction strategies. We have used the framework to analyse nine co-creative systems across four domains – music, poetry, sketching, and dance – and facilitate comparisons in the way they represent, evaluate, reason about and communicate artefacts. The framework illuminates similarities and differences among the systems, and highlights directions for future co-creative systems research.

As interaction modalities are the component of our framework most closely tied to the actual artefacts and their representations, it is challenging to construct a fully domain-agnostic way of describing them. Analysing the interaction modalities used in the example co-creative systems shows a division between those using solely intra-domain information channels and those that supplement with extra-domain information. This allows us to describe systems as being uni-modal or multi-modal in their interaction, although that distinction is less useful in creative domains like music where artefacts are often represented using different modalities. We also used the concept of interaction modalities to identify aspects such as timing differences between systems, as well as restrictions on system feedback.

We analysed the interaction styles used in the sample co-creative systems based on conceptual models, structure of interaction, and structure of contributions. The nine systems were characterised as using ambient, request-based and operation-based models of interaction, as well as being structured to use either simultaneous interaction or turn-taking. This was supplemented with one more level of analysis that characterised the structure of human and computer contributions as additive or iterative. These categories can be easily applied to systems across many creative domains, and suggest the possibility of comparative efficacy studies and the establishment of design patterns and other best practices using this terminology.

Finally, we analysed the co-creative systems for interaction strategies based on metrics, goals, and meta-reasoning capabilities. All the evaluation metrics we found in the example systems were either about artefact value, artefact novelty, or system/user interaction. This taxonomy held across all the creative domains we studied, suggesting its utility as a way of describing co-creative evaluation more generally. The systems we analysed could also be described as following either a static strategy (where their metrics never change) or a dynamic one (where metrics were adapted based on meta-reasoning). The majority had static strategies, and all the dynamic strategies involved some form of metric prioritisation. None of the systems we looked at were able to adapt their evaluation metrics more substantially, nor create new situation-specific metrics based on some higher-level reasoning. This suggests that co-creative systems have yet to leverage some of the research in autonomous creative systems surrounding meta-search and transformational creativity. Co-creative systems that are capable of adapting their evaluation metrics could lead to more intentional and interaction-aware collaboration, linking human–computer co-creativity with the broader discussion on autonomy and meta-creativity (Linkola et al., 2017).

To summarise our discussion, studies of interaction typically ask what is the most effective way to exchange information between a system (a co-creative agent) and its user (its human collaborator) (Bernsen, 2002). In the co-creative context that question is hard to answer, as we lack a rich and general terminology for how concepts, objects, artefacts and more are exchanged. Deconstructing co-creative interaction into modalities, styles, and strategies allows a finer-grained discussion of information flow during human-computer co-creativity. We hope this terminology serves as a useful design tool for exploring different ways that new co-creative systems can be realised. With the framework designers could envision and experiment with different modalities and styles.

Acknowledgments

This research has been funded by the Academy of Finland (decision #311090, Digital Aura) and the Australian Research Council (DP200101059).

References

Al-Nakhal, M. A., and Naser, S. S. A. 2017. Adaptive intelligent tutoring system for learning computer theory. *European Academic Research* 4(10).

Barfield, L. 1993. *The User Interface Concepts & Design*. Addison-Wesley Publishing Company.

Benyon, D.; Turner, P.; and Turner, S. 2005. *Designing interactive systems: People, activities, contexts, technologies*. Mateu Cromo, Spain: Pearson Education Ltd.

Bernsen, N. O. 2002. Multimodality in language and speech systems—from theory to design support tool. In *Multimodality in language and speech systems*. Springer. 93–148.

Bown, O., and Brown, A. R. 2018. Interaction design for metacreative systems. In *New Directions in Third Wave Human-Computer Interaction: Volume 1 - Technologies*. Cham: Springer International Publishing. 67–87.

Bown, O. 2014. Empirically grounding the evaluation of creative systems: Incorporating interaction design. In *Proc. ICCC*, 112–119.

Bray, L., and Bown, O. 2016. Applying core interaction design principles to computational creativity. In *Proc. ICCC*, 93–97.

Bray, L.; Bown, O.; and Carey, B. 2017. How can we deal with the design principle of visibility in highly encapsulated computationally creative systems ? In *Proc. ICCC*, 65–71. ACC.

Carroll, J. M. 1997. Human-computer interaction: Psychology as a science of design. *Annual Review of Psychology* 48(1):61–83. PMID: 15012476.

Clark, E.; Ross, A. S.; Tan, C.; Ji, Y.; and Smith, N. A. 2018. Creative writing with a machine in the loop: Case studies on slogans and stories. In *Proc. IUI'18, Tokyo, Japan*, 329–340. ACM.

Cook, M., and Colton, S. 2011. Automated collage generation-with more intent. In *Proc. ICCC*, 1–3. Citeseer.

Cox, M., and Raja, A. 2007. Metareasoning: A manifesto. *BBN Technical*.

Davis, N.; Hsiao, C.-P.; Popova, Y.; and Magerko, B. 2015. An enactive model of creativity for computational collaboration and co-creation. In *Creativity in the Digital Age*. Springer. 109–133.

Eris, O.; Martelaro, N.; and Badke-Schaub, P. 2014. A comparative analysis of multimodal communication during design sketching in co-located and distributed environments. *Design Studies* 35(6):559 – 592.

François, A. R.; Chew, E.; and Thurmond, D. 2007. Mimi-a musical improvisation system that provides visual feedback to the performer. Technical Report 07-889, University of Southern California CS Department.

Grace, K., and Maher, M. L. 2016. Surprise-triggered reformulation of design goals. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Granström, B.; House, D.; and Karlsson, I. 2013. *Multimodality in language and speech systems*, volume 19. Springer Science & Business Media.

Hix, D., and Hartson, R. *Developing User Interfaces: Ensuring Usability Through Product and Process*.

Jacob, M.; Coisne, G.; Gupta, A.; Sysoev, I.; Verma, G. G.; and Magerko, B. 2013. Viewpoints ai. In *Proc. AiIDE*.

Jaimes, A., and Sebe, N. 2007. Multimodal human–computer interaction: A survey. *Computer vision and image understanding* 108(1-2):116–134.

Jennings, K. E. 2010. Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines* 20(4):489–501.

Jordanous, A. 2012. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation* 4(3):246–279.

Kantosalo, A., and Toivonen, H. 2016. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proc. ICCC*, 77–84.

Kantosalo, A.; Toivanen, J. M.; Xiao, P.; and Toivonen, H. 2014. From isolation to involvement: Adapting machine creativity software to support human-computer co-creation. In *Proc. ICCC*, 1–7.

Kantosalo, A.; Toivanen, J. M.; Toivonen, H.; et al. 2015. Interaction evaluation for human-computer co-creativity: A case study. In *Proc. ICCC*, 276–283.

Karimi, P.; Grace, K.; Davis, N.; and Maher, M. L. 2018a. Creative sketching apprentice: Supporting conceptual shifts in sketch ideation. In *International Conference on-Design Computing and Cognition*, 721–738. Springer.

Karimi, P.; Grace, K.; Maher, M. L.; and Davis, N. 2018b. Evaluating creativity in computational co-creative systems. In *Proc. ICCC*, 104–111. ACC.

Liapis, A.; Yannakakis, G.; and Togelius, J. 2013. Sentient sketchbook: Computer-aided game level authoring. In *Proc. ACM FDG*.

Linkola, S.; Kantosalo, A.; Männistö, T.; and Toivonen, H. 2017. Aspects of self-awareness: An anatomy of metacreative systems. In *Proc. ICCC*, 189–196.

Long, D.; Jacob, M.; and Magerko, B. 2019. Designing co-creative ai for public spaces. In *Proc. ACM Creativity & Cognition*. 271–284.

O'hara, K.; Harper, R.; Mentis, H.; Sellen, A.; and Taylor, A. 2013. On the naturalness of touchless: Putting the "interaction" back into nui. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20(1):1–25.

Ravikumar, P. T., and Wyse, L. 2019. Masse: A system for music action selection through state evaluation. In *Proc. ICCC*, 132–139.

Rogers, Y.; Sharp, H.; and Preece, J. 2011. *Interaction Design, Beyond human-computer interaction*. John Wiley Sons Ltd. 3rd edition.

Schuller, B.; Ablaßmeier, M.; Müller, R.; Reifinger, S.; Poitschke, T.; and Rigoll, G. 2006. *Speech Communication and Multimodal Interfaces*. Springer Berlin Heidelberg. 141–190.

Shneiderman, B. 2007. Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM* 50(12):20–32.

Sutcliffe, A. 1988. *Human-Computer Interface Design*. MacMillan Education Ltd. Reprinted in 1990.

Tham, J.; Duin, A. H.; Gee, L.; Ernst, N.; Abdelqader, B.; and McGrath, M. 2018. Understanding virtual reality: Presence, embodiment, and professional practice. *IEEE Transactions on Professional Communication* 61(2):178–195.

Watanabe, K.; Matsubayashi, Y.; Inui, K.; Nakano, T.; Fukayama, S.; and Goto, M. 2017. Lyrisys: An interactive support system for writing lyrics based on topic transition. In *Proc. IUI*, IUI '17, 559–563. New York, NY, USA: ACM.

Weinberg, G.; Blosser, B.; Mallikarjuna, T.; and Raman, A. 2009. The creation of a multi-human, multi-robot interactive jam session. In *NIME*, 70–73.

Wiggins, G. A. 2006. Searching for computational creativity. *New Generation Computing* 24(3):209–222.

Yannakakis, G. N.; Liapis, A.; and Alexopoulos, C. 2014. Mixed-initiative co-creativity.

Young, M., and Bown, O. 2010. Clap-along: A negotiation strategy for creative musical interaction with computational systems. In *Proc. ICCC*, 215–222.