Critique as Creativity: Towards Developing Computational Commentators of Creative Works

Douglas H. Fisher Computer Science
Haerin Shin English; Media Studies
Vanderbilt University
Nashville, TN 37235
douglas.h.fisher@vanderbilt.edu haerin.shin@vanderbilt.edu

Abstract

If beauty is in the eye of the beholder, then creativity can be sought in computational commentators on arts and artifacts across diverse form and media. While most research in computational creativity focuses on the generative nature of creativity, we address how interpretative aspects of creativity can be manifest by and implemented in computational agents as well.

Introduction

Human commentators, both professional and amateur, are acting creatively by making connections between their own perspectives and social context, with an artifact. Their interpretation, comparisons, and criticisms of artistic, literary, and engineering creations are colored by their experience and aesthetic. For example, a song that elicits a rich inner animation by the listener represents a highly personal remediation of the music and lyrics, and illustrates that creativity is both generative and interpretative.

This paper focuses on interpretative processes that yield narrative critique, criticism, and commentary, which we informally regard as synonymous. Stiny and Gips (1978) were early proponents of this (a reviewer pointed us to that book), though penetration of it into CC is limited, and a deeper comparison than we have made here is warranted.

Smith (1991) suggests that these interpretive artifacts lie on a continuum between commentaries on specific subject texts (or other artifacts) to works of secondary literature that go well beyond the subject work that inspired the commentary. To Smith's long list of commentary types we add *(peer) review*, which is at the specific-subject end of the continuum, and a good candidate as an initial computational commentator. Our position paper discusses computational commentary much more broadly than review, but we aspire to an implementation and return to the possibility of such a reviewer (e.g., of ICCC submissions).

Desiderata for Computational Critics

Several interacting capabilities seem desirable in a critic, computational or human, though these capabilities will vary with the types of commentary. *First*, computational critics should understand aspects of the medium-specific traits and formal characteristics of the creation (Hayles 2004). Formal elements in literary expressions would include style (tone, diction, syntax, and structure), for instance; cinematography, mise-en-scene and montage in the case of film or other visual works; material texture in sculptures; compositional styles in music; clarity, sectioning, formatting in conference submissions; to name a few. Approaches in this vein include Russian formalism, New Criticism, and Barthes (1967) declaration of "The Death of the Author." Attention to formal aspects enables deeper understanding of the effects that a given creative product wields upon its viewers/readers. An understanding of medium-specific traits by an interpretative agent can also inform remediation of material between media.

Second, computational critics can use their personal and interpersonal understandings for finding connections to authorial intent in the creation, whether the critic reads the authorial intent "correctly" or not. Authorial intent can enrich the more immanent aspects of creative expression, as in biographical criticism (e.g. by Samuel Johnson, in *Lives of the Poets*) or the Romanticist vision of the creative genius (e.g. Wordsworth: a poem should be "the spontaneous overflow of powerful feelings" to be "recollect[ed] in tranquility" *From the Preface to the 2nd edition of Lyrical Ballads*, quoted Day, 2). A music video, for example, might enrich (or displace) a listener's internal remediation of a song; a conference review can suggest a related and productive direction for research.

Third, computational critics should be able to reason about the products (i.e., critiques) that they produce, and their relationship to the subject creation. One desirable characteristic of a critique is conceptual cohesiveness of the argument in the critique itself -- does the critique offer an interesting and informative interpretation of the subject artifact? Does a review adequately summarize a paper? If so, then at least in the "mind" of the agent writing the critique, the artifact is interpretable. Our concept of *interpretability* of artifacts is related to Bodily and Ventura's (2018) concept of *explainability*, which we elaborate later.

Fourth, a computational critic will, ideally, have the capacity to situate products within socio-historical contexts - namely engage in social criticism, perhaps with recommended citations in a review. Creative expressions not

only represent, but also critically reflect on and inspire reality; their contextual significance is therefore crucial to assessing the perspectival originality of a given work.

Fifth, a computational critic should be able to gauge how readers, viewers, and appreciators will react to the subject, thereby endowing paratextual value to the affective dynamics of a creative expression. Wolfgang Iser's "reader-response theory" (1978), Stanley Fish's (1970) "affective stylistics", David Bleich's emphasis on the subjective dimension of reader response (McCormick 1985), Norman Holland's (1989) focus on the psychological motivations that affect the reader's mode of engagement, and the attention Fish directed to the social and communal nature of reader response (Regis 1976), all fall under this category.

Critics Influence Creative Ecosystems

Several of these desirable capabilities highlight the social aspects of criticism. Humans are typically cognizant that other humans will critique their creative works, suggesting an ability to apply theory of mind when creating (Slater and Bremner, 2011). In the future, perhaps, computational creators will be cognizant that AIs (and humans) will comment as well. Computational critics might reasonably provide, at scale, what Ventura (2017) calls *external evaluation* -- critiques offer feedback that the subject's creator can use to learn or revise its own aesthetic(s). This possibility of computational critics acting, at scale, on human creators particularly, perhaps children, suggests Ethical Considerations too, which we discuss later.

Finally, criticism is a creative act in and of itself, for it requires an understanding of the myriad faculties of creativity, which must then be communicated in a persuasive manner, typically in natural language. Criticism has aspects of both interpolation and extrapolation, such as connecting the dots in a mystery and elaborating on the possible feelings of a minor character in a novel, respectively. The multivalence of criticism leads to benefits to the ICCC community of developing computational critics, perhaps the most obvious of which is that criticism represents a literary tradition that is under-represented in computational narrative generation, and in some human literary traditions as well (Smith, 1991).

Outline of the Paper

In the remainder of the paper we elaborate generative and interpretive perspectives; summarize prior work on evaluating creativity; forward cognitive architecture and ontology considerations for computational critic design, and address ethical considerations of computational critics. We conclude with a summary of main points, but a summary grounded in a prospective implementation of a computational reviewer of conference papers, notably ICCC, which we think is achievable as a proof-of-concept in one year.

From Generation to Interpretation

A common assumption is that human creative endeavors are produced from within a (self-)conscious liberal subject that designs one's creation. Intentionality, in this light, may be understood as the vector of consciousness emanating out of a creator. We call this the Romanticist generative model of human creativity.

Computational creativity is indebted to the Romanticist vision of a creator with complex internality (Wang 2000), but the current state of AI is not at a level where phenomenological intentionality is the core propellant of creating agents. Indeed, many would question whether a computational agent has intent at all. Without intent, computationally created artifacts might be conceptually indistinguishable from natural objects that we deem to be beautiful, sublime, or otherwise evocative as Kant observes in Critique of Judgment (1973). Heidegger explains that a work of art (intentionally-created expressions) must reveal manifold meanings that survive the tests of time, giving rise to new values and interpretations across contexts, cultures, and temporalities (2002, pp. 1-56). Calling this function "worlding," Heidegger subtly identifies imaginative intent as a driving force behind creativity of human art.

Computational Intentionality & Authenticity

Computational creativity researchers do not currently require the high bar of consciousness for a claim of intentionality (Ventura, 2017), but only that the computational creator guides search by goals and/or utilities, for example. Another suggestion of intention is that an agent can "explain" why decisions were made in creating artifacts, perhaps by AI credit and blame assignment.

Another aspect of the internal processing of computational creators is the experience that the agent draws upon. For instance, Colton et. al. (2018) closely examines expectations for authenticity in a computational creation, drawing on Walter Benjamin's notion of the "aura" (2008), among others. While Colton et. al's description of the aura highlights the spatial and psychological "distance" between the artist/artwork and the appreciators as the source of a given work of art's ability to inspire affective impact, we would also like to call attention to Benjamin's point about how the "aura" gravitates toward the material progeny, characteristic singularity, and contextual originality with the advent of mechanically reproduced (e.g. Andy Warhol) or readymade art (e.g. Marcel Duchamp's Fountain).

One of Colton et. al.'s (2018) cautions is that discovery that an AI has no human experience may result in its rejection by humans, and that appropriate actions can be taken. In something of a contrast, Gunkel (2017) suggests that "*... some brands of aesthetic theory, like the various versions of formalism, will be more open to and accommodating of machine-generated content than others, like Romanticism and its veneration of the figure of artistic genius.*"

Seen in this light, the value of a computationally created product becomes subject to external reception, as well as inherent qualities of the work in question or the software developers of the agent that created the work.

External Interpretation of a Creation

While approaches that underscore the networked effect of the creative expression's external impact such as reader-

response, social, or formal criticism may be readily employed to assess the "aura" of a computationally creative agent, said aura cannot be attributed to the creator as an actor with conscious intent. Google's Deep Dream is a case in point. The abstract visuals it produced were appreciated as a "technological novelty," resulting in commercial success, but the metrics of their assessment were geared more towards this novelty itself than the meaning the visuals convey in comparison to other works that share stylistic traits, as would be the case with human-made abstract paintings. The Magical Realism Bot is another example, where the artistic effect of cognitive dissonance it employs largely banks on the readers' marvel at the algorithm's ability to emulate the stylistics of magical realism, as well as the aphorisms themselves. The novelty of a computer-made object then, ultimately falls under the eye of the beholder (i.e., p-creativity as described by Boden, 2004), which reverts agency back to human interpreters.

This leads to another point on the Romanticist generative model, and again related to responses of Colton et., al. (2018) when faced with unmasking of inauthentic agents. Jean Baudrillard maintained that representation could copy the original and become a reality unto its own in his conceptualization of the "hyperreal" (1994). The output of computational creativity could be deemed "hyperreal" in that they do not intentionally attempt to represent any existent modes of reality, but instead process data/artifacts to generate new variations. The originality of any "hyperreal" object or phenomenon is assessed not based on its formal qualities but more on its contextual significance; for instance, the Epcot World Showcase section in Disney World is deemed hyperreal because its simulation of world cultures belies its aspiration to authenticity by reflecting the stereotypes pertaining to cultures rather than their genuine reality as lived experience.

In sum, the generative model is productive from our perspective in that it pushes humans to reflect on the parameters of creativity and why it has been considered as uniquely human. But creativity is not simply centrifugal -- creativity is social and multivalent too, dependent upon the multilateral dynamics of the social context across times and cultures, there is the need for an interpretive agency, which can be computationally modeled. This is why we propose the interpretive model as an addition to the (Romanticist) generative model, asserting that interpretation is an exercise of imagination. An act that bridges the subjective and the objective, critique demonstrates that the two modes of agency (generative and interpretive) are co-constitutive rather than binary oppositions.

Assessing Creativity

Our expectation is that computational critics (and criticism) will be judged by their (its) conversancy about and synthesis of diverse works, creators, aesthetics, genre, and other critics. While understudied, we believe that computational criticism and critics will build on prior work on characterizing and assessing human and computational creativity. For example, computational creators internally assess their intermediate states in the act of creating, as well as assessing the outputs of creation. The measures and mechanisms used to make these assessments are undoubtedly related to, if not identical to some of the assessments that will be made by computational critics. Abilities to reason about criteria in this section have implications for Desiderata for Computational Critics discussed earlier.

This section focuses on prior work for assessing "artifacts as manifestations of creativity" -- many refer to these as "creative artifacts" and we will use that as shorthand too. In what follows we focus here on criteria as applied to the artifacts that will be critiqued, but in all cases, different instantiations of these broad classes of criteria also can be applied to the artifact that the critic creates – a critique.

After criteria applied to artifacts, we touch upon characterizations and assessments of creative processes. Most prior work in assessing creativity assumes implicitly or explicitly a generative perspective, and so we will also comment on how some of the criteria for creativity might be adapted to an interpretative perspective and a computational critic more specifically. For example, to repeat the opening sentence of this section, critiques, which are the creative artifacts of critics, are often valued for their recognition of the diversity of perspectival values, where diversity of perspective and aesthetic may not be as relevant when discussing generative creative agents.

Novelty

Novelty refers to an overall assessment, often as a metric, of the differences and similarities between an artifact and others of the same kind. The generalization of "same" to "comparable" kind, to include some that span media and form, stems uniquely from an interpretive perspective. For instance, a written work can certainly be compared to other writings of the same author and genre in terms of their techniques and narrative content, but additionally, comparisons can be made to audiovisual media (e.g., most obviously a film adaptation of a novel). Indeed, comparability can be as broad as remediation between media allows (e.g., a bible story of Daniel in the Lion's Den as a painting at the National Galleries). Colton (2010) also imagines very broad possibilities, based on shared ontologies and other resources, which we return to later.

Novelty can be judged relative to all comparable artifacts in the world, or it can be individualized to the artifacts experienced by an agent. The former is akin to Boden's (1992) definition of historic (h-)creativity (where 'creativity' is limited to the novelty dimension), and the latter is akin to Boden's (1992) notion of psychological (p-)creativity (novelty). An interpretive perspective suggests we generalize to a continuum that bridges small to large groups of persons. In sum, computational critics and their critiques will presumably be judged by the breadth of their knowledge base of artifacts, their comparable kinds, and their exposure to agent populations of varying scope.

Value

Artifacts can be novel but ineffective for utilitarian purposes or in eliciting an affective response. We identify four categories of value, which would be variously used by interpretive agents, including critics of art, film, book, engineering products, and humanistic works.

Exchange value corresponds to monetary valuations. Creative works (as in the domain of art) can be assigned exchange value when treated as commodity, in which case aesthetic appeal may be converted into monetary values (utility). The professionalization of art has boosted this tendency. For example Google's Deep Dream was sold for an exorbitant price at auction (Business Insider 2016).

Value can also refer to a **utilitarian value**. Maher and Fisher (2012) suggest that just as descriptive intrinsic attributes could be used to identify novelty, utility or functional attributes like processing speed or recyclability for laptops could be used to create a utilitarian space.

Value can be a way of measuring **technical sophistication**, reflecting the "craft" aspect of material manifestations of creativity. The value of a written prose piece, for instance, can be judged based on its skillful navigation of syntax, diction, structure, narrative feasibility, and other aspects that make the prose more appealing, logical, enjoyable, and rhetorically more effective. Technical sophistication would be used to partially assess the value of critiques themselves. In visual art, the creator's ability to effectively employ color combinations, apply proper strokes or perspectivization can determine the quality of the product.

Lastly, value can be seen in terms of the **affective force** that a creative work evokes. While the above forms of value are always subject to social contextualization and subjective judgment too, this type of value is more pronouncedly non-axiological; namely, it is less likely to be mapped on to differential hierarchies that determine whether a certain type of object/phenomena is superior or inferior to others. Value can simply inhere in instances where a creative output could emotionally "move" the recipient.

Unexpectedness

Another common characteristic for judging creativity is unexpectedness or surprise. Grace and Maher (2014) have a good taxonomy and survey of unexpectedness. We call out two works here: novelty in a future space of anticipated (projected) artifacts (Maher and Fisher, 2012), and changes to the posterior (post-artifact) distributions (e.g., Baldi & Itti, 2010) and/or to conceptual structures (Grace, et. al., 2015), over a space of artifacts. Our reason for calling out these interpretations of unexpectedness is that they make historical context explicit, through a "weighting" of the past, even "curve fitting", for an assessment of the current artifact. The historical perspective suggests another characteristic which we would want critics to assess -- authority.

Authority and authority

By authority we mean a dimension that is akin to its use in social/citation networks (Kleinberg, 1999), with inward and outward pointing links to each node. Nodes with a higher fan-in, for example, suggest greater authority. In the context of computational creativity, an artifact's history within a social network is reflected in the links between artifacts. Every time a variation on (aka derivation of) a creative artifact is made, its authority is increased. This perspective suggests that it cannot be evaluated until sometime after an artifact's introduction, because variations and interpretations of it must be made, and this may take time.

However, not only will the valued critic be able to trace and evaluate the descendants (in links) of an artifact, but a critic may be able to judge the capacity of an artifact to be reinterpreted and varied. Potential authority refers to the capacity for imaginative latitude in adapting an artifact. This category aligns with the principles of reader-response theory, but also highlights a creative artifact's inherent qualities that encode a greater degree of interpretive freedom. For instance, works that are deemed "canonical" retain their appeal across temporal and cultural transitions because they address issues that resonate with appreciators in a universal manner. In the case of Shakespeare, most notably, his ability to capture the most pressing concerns of the human society and heart has enabled his works to be reinterpreted, generating new value and appreciation centuries after their original debut. Impressionist paintings still evoke marvel in the viewers' eyes by accentuating the central role of light and their interplay with human perception. In sum, potential authority can be measured by an artifact's capacity, as determined through intrinsic and social factors, to command universality and generate new values. The specifics of such a measure remain a challenge.

Sosa and Gero (2005) use the term "authority" to mean hierarchical authority among agents/persons within a field of endeavor. More generally, they say *"... in agent societies with strong social ties uneven hierarchies generate powerful opinion leaders that exert the role of gatekeepers to the domain. In contrast, in social networks with weak ties, influence is distributed among adopters and the expert judgments tend to vary over time. Consistent with Gardner's (1994) observation, the former social arrangement generates higher variance in the distribution of prominence whilst the latter yields more egalitarian distributions." (Sosa and Gero, 2005, p. 26)*

This is a different, but a related kind of authority, -- authorities of persons, not artifacts. It is yet another kind of authority, an awareness of which a critic should be assessed on, even if the critic adjusts for or ignores it in the critique. Colton (2010) made a similar point.

Like many conferences ICCC 2019 has moved to blind review because it *"encourages submissions from authors in adjacent research fields."* Though the call stops short of stating the rationale of this brand new policy as one of mitigating the influence, through review, of power hierarchies by opening gates to those outside the community, that desire to mitigate and expand is a natural inference.

Authenticity

We have already talked at length about Colton et. al.s (2018) concept of authenticity, and this too can be used by

critics, particularly when they know the creator and its capacity, if an AI, for empathy and other feelings. We return to authenticity under Ethical Considerations.

Explainability and Interpretability

Our conceptualization of interpretability is related to what Bodily and Ventura (2018) conceptualize as explainability, but rather than being an internal capacity of a creative agent, explainability can also happen external to the creator, in another agent, as part of an interpretive capacity that we call interpretability.

Bodily and Ventura give broad examples of explainability by a creative agent, to include stating its "feelings" and goals in creating an object, as well as the processes ("logical rules") used in the creation. Specifically, explainability is the capacity of an agent to explain why its creation is consistent with the agent's aesthetic values. Our second-to-last value type of technical sophistication is particularly related to the concept of explainability, but other types can probably be cast here too (e.g., "why did that computational painting bring so much at auction?")

By interpretability we mean explainability (by the critic, not the creator) using any (or a selected) set of aesthetic values that are at the disposal of the critic. Because the critic is not generating the artifact, "proving" interpretability by demonstration -- that is, finding an interpretation -- has non-deterministic aspects, that can be addressed by searching a knowledge base of past products and aesthetic rules for those that fall within an aesthetic class, as well as using other information like authorship, location of creation, and other metadata to hone in on relevant concepts using established ontologies known by the critic. AI argumentation (Bench-Capon and Dunne, 2007) will be important in the writing of critiques that stand up to scrutiny, including lawsuits, as well as demonstrating to humans what a well-supported critique looks like, which has Ethical Considerations as well.

Social Interactions

We have already addressed social interactions at length, with information like authorship; authority as citation and influence within a community; and a vision for an ecosystem of creators and commentators.

Assessing Creative Processes

Up to now we have addressed creativity manifest in artifacts, as well as social processes that surround them. The processes that create artifacts internal to the creator can also be part of the critiques, if they are known. Stokes (2005) suggests, using an AI search framework, that an ideally creative process is not too under-constrained (e.g., without an aesthetic guide) and not too over-constrained (e.g., a near straight line to a goal, as in a mediocre "new" car design). This dichotomy is related to the exploration and exploitation tradeoff, which is ubiquitous in AI, and with continuous space for characterizing creativity between unconstrained exploration and rigid exploitation.

Stokes asserts that any creativity starts with an ill-defined problem space, which the creative agent can better define. It is this defining process where much of the agent's creativity is found, and the ability to better define a problem space is yet another possible definition of intentionality. For example, "*Monet precluded dark-light contrasts*" (Stokes, 2005, p. xiiv) and Jimmy Breslin sought out interviews that no one else thought to do (e.g., Breslin, 1963). These definitional choices help to ensure the novelty, unexpectedness, perhaps value of the final artifact, long before it is finished. Novelty and value can be applied to the earliest internal states of the creation process. Harkening back to Baudrillard and Heidegger, the idea of "originality" is always contextual. Originality can be, for instance, referring to whether the creating agent was able to produce the output mainly based on one's own imagination without starting the work using a similar object as a seed, though this too illustrates a binary that can be sublated.

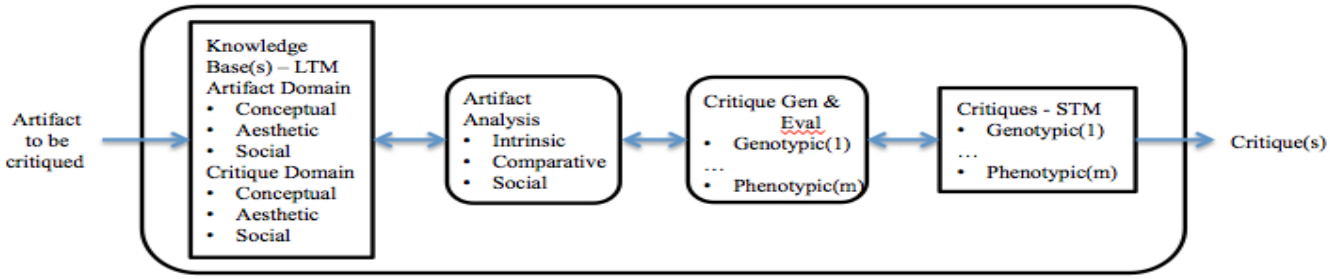
While a computational critic will need generative processes in producing its critiques, and will use much of the technology developed for other creators of narrative (e.g., Riedl and Bulitko, 2013), its ability to peer into "thought" processes of creators will be limited, so this sophisticated theory of mind (Prabowo & Thelwall, 2013) would not be a near-term focus, but we pose it as a challenge.

Architectures and Ontologies

Critics should have a capacity to summarize an artifact; to compare it to other artifacts for purposes of assessing its apparent creativity; to reason about its creator, where possible; to place the artifact in social and cultural context; and to organize and write an interesting and "fair" critique.

Figure 1 brings together our various discussions by illustrating a general architecture for a computational critic. It profitably compares and contrasts with Stiny and Gips (1978, Ch 2), inviting a deeper analysis and synthesis later.

An artifact to be critiqued is presented to the system of Figure 1, and undergoes analysis of its intrinsic properties and metadata, as well as comparisons to other works and abstractions in a knowledge base to assess novelty, value, unexpectedness, authority, interpretability relative to aesthetics at its disposal, and other characterizations.



The critic will perform these assessments by consulting conceptual, aesthetic, and social knowledge bases. For example, assessments will typically compare an artifact to a set of artifacts based on a set of intrinsic relationships (e.g., the ratio of the length of a car's hood to the total length of the car, in an engineering example), which are often unary, also called attributes (e.g., the color of a car). If artifacts are represented by attributes only, then this representation is typically referred to as a feature vector. If object representations include one or more greater-than-unary relationships, then these are relational representations, which are naturally visualized as graphs. Triples representations represent all knowledge in binary relations between objects; such uniformity is desirable in many settings. The relationships used to do comparisons can be the union of relationships ever used for instances of a kind of artifact, or they can be individualized, both with regard to what relationships are included in comparisons and in how these relationships are weighted for producing an overall assessment. Extreme points in knowledge representation create another continuum on which critics and critiques will fall.

The particular relations in the knowledge bases would be based on ontologies that could be hand crafted and/or built through processes of conceptual blending and analogical reasoning (Colton, 2010). There are other ontologies that cover diverse topics that would be good starting points (e.g., BBC's Ontologies).

Endowing computational critics with knowledge, to include of other agents, their works, and relationships between them, will likely be one distinguishing characteristic in the design and behavior of computational critics from computational generators of artifacts. So far, this latter group is not concerned with other creators, nor diverse aesthetics, compared to their human creator counterparts.

The architecture of Figure 1 is informed by Ventura's (2017) architecture for a (generative) computational creativity system. In particular, Ventura proposes an architecture for CC systems that differentiates "genotype" and "phenotype" representations and evaluations. A genotype of a creative artifact is a private precursor to a public phenotypic representation. Instead of one, we imagine that there will be a series of genotypic representations, starting with a concept map (Nowak & Cañas, 2006), with subsequent genotypes adding annotations from the various knowledge bases outlined, then following revision steps similar to narrative generation found in Callaway and Lester (2002), for example, including a narrative planner (i.e., narrative organizer, sentence planner, revision component, and surface realizer). This planning and writing apparatus, together with creativity pointing evaluation functions at each step of genotype and phenotype, would fill the Gen&Eval module of Figure 1.

Other architectures include Romero et. al. (2003), which used computational critics for rating candidates during an evolutionary creative algorithm. They also have a vision of a society of creators and commentators.

Ethical Considerations

There are many ethical concerns that computational critique implicates, to include so called fake news and memes, both of which are in easy reach of generative computational creativity; modeling civility and productivity in commentary, even generosity (Coman, et. al., 2018); the obligations and conventions of citing AI commentators; the legal responsibility and liability of AI commentaries; the implications of discovered inauthenticity to certain persons, like children; to name but a few. We save most of these for another day, and concentrate on two concerns that follow from the (Romanticist) generative perspective.

Implications of the Generative Perspective

While persuasive in its appeal to highlight the subjectivity of the creator, the generative perspective carries the risk of reinforcing anthropocentrism when applied to computational creativity. Under its rubric, the AI either acquires a metaphorical equivalence to the genius figure who possesses a singular talent to be appreciated, or is dismissed as an instrumental actuator of human will, denigrated to the level of a mere tool. Either way, the exaltation of the creator becomes a celebration of what we deem to be essentially human. This process, in turn, re-fetters humanity to a hierarchical ontology that has been plaguing us for eons, for the gist of the undergirding logic here is that those who possess special talent or skills (and subsequently their creations) are "better" than others. Attributes that are commonly attached to creativity, such as novelty or authenticity, demonstrate the danger of this discursive framework. These characteristics also apply to computational critics, because they will generate artifacts as well.

Commentary at Scale of Human Creations

What will an ability to computationally critique works imply – particularly critiquing human creative works and particularly when done at scale? Even if computational commentators are intended to critique computational generative agents, there seems no reason that these commentators could not be applied at scale to human works, and not

just to professional human creators, but all groups, including children. Computational commenters, particularly mediocre ones that gain from anthropomorphism, could squelch creativity rather than beneficially add to creative ecosystems. There is already debate on the pros and cons of automated essay graders on the mindset of students. (Smith, 2018). Granted, robo-graders are rudimentary commentators at best, but their effects can be significant. Our intent here is that computational creativity research raises the level of commentary, and indeed, the best computational commentators may model what many agree is noble behavior by removing or managing cultural bias. We take up this latter theme in the next section.

Implications of Adherence to Authenticity

Authenticity presupposes the existence of a grounded truth that deserves to be respected. High-fidelity to the "real" can render alternative modes of representation and perception marginal or even inferior, as we have seen in Plato's Allegory of the Cave (2013). Platonic idealism, which suggests that there exists a transcendental "truth" while their material instantiations are ephemeral and therefore inferior "shadows," is in essence a precursor of Rene Descartes' substance-dualist affirmation of humanity's superiority over other forms of being, as he had asserted in "Animals are Machines" (Harrison, 1992).

Another detrimental form of adherence to "authenticity" can be seen in the domain of ethnic literature. Racial and ethnic minority writers are often defined mainly by their hyphenated identity (e.g. Asian-American, African-American), and are therefore expected to serve as metonyms of their identificatory peers. A case in point is the unfavorable response that Korean-American writer Chang-rae Lee encountered upon publishing his third novel Aloft (2004) in which the protagonist is a white Italian American man. Having received critical acclaim by exploring the trope of being torn between two cultures or the stereotype of the "forever foreigner" in his debut novel Native Speaker (1995), and the traumatic legacies of coloniality across Japan and Korea from the perspective of an immigrant in his second novel A Gesture Life (1999), Lee directly countered the general readership and critics' expectation that narratives pertaining to his own identity as an Asian immigrant are the only ones that carry authentic value. An ironic development, given how the predominantly white male writer-base in the cultural industry within the U.S. has been granted the liberty to assume a wide variety of identities in their fictional imaginaries, which attests to the "invisibility of whiteness" (Reddy, 1998).

But computational criticism can be an "equalizer" along dimensions of bias. This approach opens up new understandings of creativity, and art, in tow, granting agency to the medium itself in new ways. The concept of medium agency is already acknowledged by media studies (that a thing or phenomena can have agency or in other words do things and make things happen) without necessarily being a subject. N. Katherine Hayles's (2017) exploration of nonhuman cognition, in this regard, is relevant.

Concluding Remarks and Future Work

This position paper has argued broadly for the importance of considering computational interpreters as part of the computational creativity landscape. We have summarized prior work on evaluating creativity in a generative system, with attention to how the interpretive perspective can add to this literature; forwarded cognitive architecture and ontology considerations for computational critic design; and addressed ethical considerations of computational critics. In the future near term, a deeper comparison with Stiny and Gips (1978) is needed. We also aspire to implement a computational critic in the near term.

On opening we suggested that a peer review of a conference paper submission would be a good first implementation of a computational critic. A review is grounded by a specific paper, in a specific context, which constrains the analysis immensely. Of the five desirable capabilities of the Introduction an ideal review instantiates all of these, with comments on scientific clarity and organization; suggestions for related research; self checking that the review is on point and helpful; suggestions of related work and otherwise appropriate citations; and conveying other helpful suggestions in the review that improves the paper. An ideal conference reviewer would also be aware of the many specific criteria of Assessing Creativity, like novelty and value, and would comment on and/or score these.

Conference submission reviews, which are text, also critique text, which means that a number of computational tools are available, like source-text summarization, related-work search, and topic modeling are available for a proof of concept. All in all, a proof of concept of a computational reviewer for a particular conference like ICCC seems achievable in one year.

Acknowledgements

This work was supported in part by NSF #1521672 "Collaborative Research: CompSustNet: Expanding the Horizons of Computational Sustainability". Special thanks to the students in the Vanderbilt University Course on the Ethics of Artificial Intelligence, and to three anonymous reviewers for their very helpful comments, including pointers to Stiny and Gips (1978) and Stiny (2015); "owning" the newness of the interpretive perspective; ambitions for implementation and agenda; and importance of connecting concepts across the paper.

References

Baldi, P. & Itti, L. (2010) "*Of Bits and Wows: A Bayesian Theory of Surprise with Applications to Attention*" Neural Networks. V.23, N.5: 649–666.

Barthes, R. (1967) "*The Death of the Author*" in Image, Music, Text: 143-148. Trans. Stephen Heath. Fontana.

Baudrillard, J. (1994) Simulacra and Simulation. The University of Michigan Press.

Bench-Capon, T.J.M & Dunne, P. (2007) "*Argumentation in artificial intelligence*"Artificial Intelligence,V.171

Benjamin, W. (2008) The Work of Art in the Age of Its Technological Reproducibility, and Other Writings on Media. The Belknap Press of Harvard Univ. Press.

Boden, M. (1992) *The Creative Mind*. London: Abacus.

Bodily, P. M., & Ventura, D. (2018) "*Explainability: An Aesthetic for Aesthetics in Computational Creative Systems*" International Conference on Computational Creativity: 153-160.

Breslin, J. (1963). Digging JFK grave was his honor. The New York Herald Tribune, Opinion, November 1963.

Callaway, C. B., & Lester, J.C. (2002) "*Narrative Prose Generation*" Artificial Intelligence,V.139,N.2: 213–52.

Colton, S. (2010) "*Towards ontology use, re-use and abuse in a computational creativity collective*" In Kutz, O., Hois, J., Bao, J., Grau, B.C. (eds.) Modular Ontologies – Proceedings of the Fourth International Workshop (WoMO'2010): 1–4, IOS Press.

Colton, Pease, A., Saunders, R. (2018) "*Issues of Authenticity in Autonomously Creative Systems*" International Conference on Computational Creativity: 272-279

Coman, A., Mueller, E.T., Mayer, M. (2018). International Conference on Computational Creativity.

Fish, S. (1970) "Literature in the Reader: Affective Stylistics"

Grace, K. & Maher, M.L. (2014) "*What to expect when you're expecting: The role of unexpectedness in computationally evaluating creativity*" International Conference on Computational Creativity.

Grace, K. and Maher, M. L., Fisher, D. & Brady, K. (2015) "*Data-intensive evaluation of design creativity using novelty, value, and surprise*" International Journal of Design Creativity and Innovation, V.3, N.4: 125-147

Gunkel, D. (2017) "*Rethinking Art and Aesthetics in the Age of Creative Machines*" Philosophy & Technology, V. 30, N. 3: 263-265.

Harrison, P. (1992) "*Descartes on Animals*" The Philosophical Quarterly, V.42, N.2: 219-227.

Hayles, N. K. (2004) "*Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis*" Poetics Today, V.25,N.1: 67-90.

Hayles, N. K. (2017) Unthought : The Power of the Cognitive Nonconscious. The University of Chicago Press.

Heidegger, M. (2002) "*The Origin of the Work of Art*" in Off the Beaten Track, Cambridge University Press.

Holland, N. N. (1989) Poems in Persons: an Introduction to the Psychoanalysis of Literature. Columbia Univ Press.

Iser, W. (1978). The Implied Reader. Johns Hopkins Univ Press.

Kant, I. (1973) The Critique of Judgement. Oxford, Clarendon Press.

Kleinberg, J. (1999).Authoritative sources in a hyperlinked environment. Journal of the ACM, V.46, N.5. 604-632

Lyu, S., Rockmore, D. & Farid, H. (2004) A digital technique for art authentication. *Proceedings of the National Academy of Sciences,* V.101, N.49:17006–17010.

Maher, M.L. & Fisher, D. (2012) Using AI to Evaluate Creative Designs. ICDC 2012 - 2nd International Conference on Design Creativity.

McCormick, K. (1985) "*Theory in the Reader: Bleich, Holland, and Beyond*" College English,V.47,N.8: 836-850.

Novak, J. D. & Cañas, A. J., (2006) "*The Theory Underlying Concept Maps and How to Construct and Use Them*" Technical Report IHMC CmapTools 2006-01 Rev 2008-01 Institute for Human and Machine Cognition, Pensacola Fl.

Plato (2013) Republic. Harvard University Press.

Prabowo, Rudy & Thelwall, Mike. (2013). Sentiment analysis: A combined approach. Journal of Informetrics. 143-157. 10.1016/j.joi.2009.01.003.

Reddy, T. M. (1997) "*Invisibility/Hypervisibility: The Paradox of Normative Whiteness Transformations*" The Journal of Inclusive Scholarship and Pedagogy, V.9,N.2:55-64.

Regis, Edward (1976) "Literature by the Reader: The Affective Theory of Stanley Fish" College English 38

Riedl, M. & Bulitko. V. (2013) "*Interactive Narrative: An Intelligent Systems Approach*" AI Magazine. V. 34, N. 1.

Romero, P., Machado, A., Santos, A., & Cardoso, J. (2003) "*On the Development of Critics in Evolutionary Computation Artists*" Evoworkshops.

Slater, A. & Bremner, J. G. (2011) "*An Introduction to Developmental Psychology*" British Psychological Society, Blackwell Publishers.

Smith, B. (1991) "*Textual Deference*" American Philosophical Quarterly, V.28, N.1: 1-13.

Smith, T. (2018). More States Opting to 'Robo-Grade' Student Essays by Computer, NPR, aired June 30, 2018.

Sosa, R. and Gero, J. (2005) "*A Computational Study of Creativity in Design: The Role of Society*" Artificial Intelligence for Engineering Design Analysis and Manufacturing, V.19, N.4: 229-244

Stiny, G., & Gips, J. (1978). Algorithmic aesthetics : computer models for criticism and design in the arts. Berkeley: University of California Press.

Stiny, G. (2015). The critic as artist: Oscar Wilde's prolegomena to shape grammars. Nexus Network Journal, V. 17,N.3: 723-758

Stokes, P. (2006) Creativity from Constraints. Springer Publishing Company, New York

Ventura, D. (2017). "*How to Build a CC System*" International Conference on Computational Creativity.

Wang, O. N. C. (2000) "*Kant's Strange Light: Romanticism, Periodicity, and the Catachresis of Genius*" Diacritics, V. 30, N.4: 15-37.