Proceedings of the European Society for Aesthetics

Volume 11, 2019

Edited by Connell Vaughan and Iris Vidmar Jovanović

Published by the European Society for Aesthetics

esa

Proceedings of the European Society for Aesthetics

Founded in 2009 by Fabian Dorsch

Internet: http://proceedings.eurosa.org
Email: proceedings@eurosa.org
ISSN: 1664 – 5278

Publisher
The European Society for Aesthetics

Department of Philosophy
University of Fribourg
Avenue de l'Europe 20
1700 Fribourg
Switzerland

Internet: http://www.eurosa.org
Email: secretary@eurosa.org

Proceedings of the European Society for Aesthetics

Volume 11, 2019

Edited by Connell Vaughan and Iris Vidmar Jovanović

Table of Contents

Artificial Creativity and Generative Adversarial Networks

Jens Dam Ziska[1]

Kunstakademie Düsseldorf

ABSTRACT. This paper argues that although past and current attempts at programming creative machines have yielded noteworthy results, these attempts ultimately fall short of genuine creativity. Most importantly, we are yet to see machines and programs which not only traverse a creative domain to produce novel products, but which do so in a manner which itself is creative and where this creativity is not better ascribed to the creator of the program. Until we know whether it is possible in principle for a machine to achieve these feats, it remains unclear whether genuine artificial creativity is at all possible.

1. Introduction

Computer programs already compose music, write poetry, paint paintings, and assist in scientific discoveries. Does this mean that such programs are creative, or that they have the potential to be creative? This question is sometimes used to raise fears of a future in which artists and scientists have been replaced by creative machines, but philosophically it also raises further questions about the nature of human creativity. Even if machines will never

[1] Email: jens.dam.ziska@gmail.com

make human creativity obsolete, seemingly creative programs may still offer us a glimpse into the workings of human creativity.

In this paper, I argue that it is still an open question whether artificial creativity is at all possible. Although classical and contemporary attempts at producing artificial creativity yield many noteworthy results, they ultimately fall short of genuine creativity. Most importantly, we are yet to see machines and programs which not only traverse a creative domain to produce novel products, but which do so in a manner which itself is creative. As long as machines cannot achieve this feat, it remains unclear whether they can teach us anything about human creativity.

2. Classical AI

According to Margaret Boden (2005), creative machines can dispel some of the romantic myths that stand in the way of a scientific understanding of creativity. In particular, artificial intelligence can help dispel the claim that human creativity must remain outside the scope of scientific explanation because creative acts are the unpredictable acts of geniuses who bring something new into existence - a new theory, a new work of art, etc. - which could not have been foreseen. If we can use AI to program machines that simulate such creativity closely enough, perhaps we may eventually also uncover the rules and heuristics that govern human creative thought.

Boden's sentiments are shared by many computer scientist, but what is perhaps more surprising is that they are also shared by many artists. One example is Harold Cohen who in addition to being an accomplished painter became rightly famous for programming AARON, a drawing program intended by Cohen to demonstrate that artistic creativity is governed by rules which artificial intelligence can help elucidate. AARON comes in a range of versions. Early versions could draw abstract forms resembling rocks, sticks, or even birds whereas later versions are able to draw human figures such as acrobats. Abstract-AARON produces its drawings by first selecting a starting-point at random and then completing the drawing by following a set of IF-THEN rules that specify what should be done next depending on what has already been drawn. Abstract-AARON, however, cannot consider its drawing as a whole or learn from what it has drawn, as it has no memory (*ibid.* p. 157).

Later versions of the program are more complex. Contrary to abstract-AARON, acrobat-AARON can both plan its drawings before it begins to draw and check if it is following the plan as it is drawing. The later versions of AARON are able to draw human figures because a computational model of the human body is incorporated into them. This model provides an outline of the human body that can be varied and coupled with a range of details so as to produce an indefinite number of drawings that are all different from one another. AARON, however, cannot draw any kind of figure. A number of constraints limit what it can do. For example, the

figures must have two arms and the composition must obey certain strictures (*ibid.* pp. 162-3).

Can Cohen's program tell us anything about how creative artist work? According to Boden, it can so long as we do not try to claim too much on the program's behalf. AARON is not creative in the substantive historical sense that it creates something which has never been seen before. Rather, Boden argues that AARON is creative in the more modest sense that later versions of the program can do something which earlier versions could not. According to Boden, "[c]reativity is the ability to come up with ideas or artefacts that are *new, surprising and valuable*" (*ibid.* p. 1). If we see earlier and later versions of AARON as one continuous program, AARON seems to satisfy all three of these features. The program is able to come up with new artefacts, since none of its drawings are the same. Moreover, those artefacts are surprising, since one cannot predict what the next drawing will look like based on previous drawings, even if the program makes its drawings all by itself using strict procedures. Last but not least, AARON can produce artefacts that are valuable given that many of its drawings are aesthetically pleasing.

It should be noted immediately that AARON is creative only on the assumption that all versions of it are part of the same program. Boden admits that if we do not grant this assumption, AARON's creativity is much less radical, since each version only does what Cohen has programmed it to do. There is no autonomous development from one version to another.

Cohen has to step in and do the necessary programming before the next version can do anything new. On its own, AARON cannot go from drawing abstract landscapes to drawing acrobats. Yet, notwithstanding this limitation, Boden argues that AARON's performance is comparable to that of an artist painting in a personal style.

This comparison is apt up to a point. There are, however, significant differences between AARON and an artist painting in a personal style. Whereas each version of AARON gets its style instilled into it once and for all when Cohen programs it, artists develop their personal style over time. An artist who merely adopts his or her style from another artist is not recognised as creative and may not even be recognised as an artist at all. When we read about the great painters of the Renaissance in Vasari's account of the period, they are invariably compared and esteemed by their ability to change the course of painting (Vasari 1998). The same goes for modern painters. Mondrian's creative achievement did not consist in his ability to paint abstract pictures using lines and rectangles of different colours – most people can do that with some practice – but in his developing this style.[2]

[2] The above is in line with Kant when he says that rather than being a matter of following rules "[g]enius is the talent (natural gift) that gives the rule to art" (Kant 2005, 5: 307). Kant seems here to agree that there are rules of art, but argues that the people we recognise as being the most creative, geniuses, do not produce art by following these rules. They give us new rules. Yet, 'rule' here does not mean the kind of rule used by AARON. Instead, it is an exemplar that stands out as an excellent example setting a new standard for

It therefore seems that AARON lacks something fundamental to be called creative. To be truly creative, the program should also be able to develop its own style. The criticism put forward here does not contradict Cohen's claim that painting is rule-governed. Judging from their appearance, it seems likely that Mondrian's paintings are constructed according to certain rules. What makes Mondrian a truly creative painter, however, is that he developed those rules himself. Perhaps we could say that the development of a new set of rules was the primary act of creation that made a new kind of painting possible whereas Mondrian's following of those rules was merely a secondary act of creation that realised his artistic vision. This of course leaves open whether the production of rules is itself governed by further rules. Yet, Cohen's failure to come up with a program that can change its style of drawing suggests that the answer is no, or, if yes, that the rules must be significantly different from the ones used by AARON.

With the distinction between the production and the following of rules, we are also nearing the distinction between creativity and skill.[3] It is rules of skill that most resemble the rules used by AARON. These are rules that trained artists must often master, but how creative the artists are depends on how inventive they are in applying the rules. Like a trained artist, AARON can be said to master a range of skills. However, unlike a creative professional artist, AARON is not free to employ its skills in

other artists to emulate.

[3] Gaut (2009) explores this distinction in more detail.

inventive ways, since it has been programmed to use the rules a certain way. In this respect, AARON strikes us as anything but creative. This problem cannot be circumvented through the addition of more rules. Even if we could somehow add second-order rules which specify how to select and apply a given set of first-order rules so they increase inventiveness, a new space would only open up for unruly creativity, since the second-order rules could themselves potentially be applied in more or less inventive ways by a human artist.

3. Generative Adversarial Networks

AARON's failure to emulate a genuinely creative artist is symptomatic of a general flaw afflicting all attempts at reproducing creativity using classical AI. They all rely on pre-given rules and heuristics which a computer then applies in a plodding manner which would not be described as creative had it been performed by a human.[4] Even if such computer simulations can deliver seemingly creative results, this does not mean that the procedure issuing in those results was itself creative. As Hubert Dreyfus notes, when we aim for psychological explanation, it is not enough that a computer simulation manages to imitate the input/output functions performed by a human being when these functions may be satisfied in many ways. The

[4] Cf. Novitz (1999), p. 74.

program must also simulate the cognitive processes that people actually go through when they generate the output from the input (Dreyfus 1972, p. 80).

This is an issue that current research on computational creativity seems increasingly to recognise. Much of this research does not attempt to identify the supposed rules and heuristics which define a given style of thought. Instead, this research aims to generate creative outcomes by training neural networks on databases of manmade artworks until these networks are able to produce similar works themselves. So-called Generative Adversarial Networks are made up of two sub-networks: a discriminator which has access to a training set of manmade images and a generator which produces new images. When these sub-nets are set up so that the generator tries to produce images which the discriminator will mistake for a real manmade image, the two sub-nets will eventually reach an equilibrium at which they begin to produce outputs which look like already existing art.

These networks have two advantages. The first is that they do not follow heuristics like the programs of classical AI, but do instead seek to emulate how the brain processes information. The second advantage is that they have access to and respond to a database including canonical works much in the way that artists respond to exemplary works of art. Yet, these networks do not generate anything creative in their current form. After all, they are set up to produce works which look like already existing art. They

can therefore at most be likened to an artist who has learnt to imitate a certain style.

Perhaps it is possible to augment these network, however, so that they generate genuinely creative outputs. Elgammal et al. (2017) describe a program which represents an attempt to do just this. This program, which Elgammal et al. call a Creative Adversarial Network, is set up so that instead of trying to generate images in an already existing style, it attempts to generate novel images which are ambiguous between different pictorial styles. This type of network has produced some remarkable results. When asked to compare a series of abstract pictures generated by the network with a series of abstract pictures from the Art Basel 2016 art fair, a group of respondents rated the images by the network as more aesthetically pleasing than the works from the art fair (the program is less successful when it is asked to produce figurative work).

It is, however, still too early to say whether this type of network holds any promise of producing genuine creative works of art. As the authors themselves admit, an evaluation of aesthetic pleasantness does not entail that the work is also creative. One might indeed expect the opposite. Radically creative work is almost by default less likely to be found aesthetically pleasing than derivative work which is made to please the aesthetic sensibility of its time. The history of art is full of artists whose work was initially derided only to be recognised as groundbreaking by later generations. Even after Manet had made a name for himself in Paris, his

paintings continued to be ridiculed not only for their unfinished look, but also for their subject matter. "I cannot imagine what can have made an intelligent and distinguished artist choose so absurd a composition", one critic wrote of *Le Dejeuner sur l'herbe.*[5] Conversely, the history of art also contains many examples of artists who found instant success, but who are now thought of as little more than opportunists who knew how to pander to the aesthetic taste of their time.

The main reason for attributing creativity to Elgammal's network can therefore not be that people find its outputs more pleasing than some comparison class of paintings. Rather, the main reason for attributing creativity to the network lies in its ability to deviate from stylistic convention while still producing images that are recognised as having aesthetic value. It is this feature that puts the network in contention for satisfying all three of Boden's marks of creativity: *novelty*, *surprisingness*, and *value*. This also exempts the network from the main objection that was levelled at Cohen's drawing program. In contrast to AARON, the network does not adhere to a single style that it has been programmed to explore. Elgammal et al. do not employ exactly the same criteria as Boden, however. Instead, they follow Colton (2009) in taking novelty, skill, and an ability to

[5] William Bürger (Théophile Thoré), *Salon de 1863*, in *Salons de W. Bürger*: "Je ne devine pas ce qui a pu faire choisir a un artiste intelligent et distingue une composition si absurde […]". Quoted in Fried (1996), p. 297 and p. 570, n. 83.

assess its own creation as the relevant criteria for judging whether a system is creative:

> Our proposed system possesses the ability to produce novel artifacts because the interaction between the two signals that derive the generation process is designed to force the system to explore creative space to find solutions that deviate from established styles but stay close enough to the boundary of art to be recognized as art. This interaction also provides a way for the system to self-assess its products. The quality of the artifacts is verified by the human subject experiments, which showed that subjects not only thought these artifacts were created by artists, but also rated them higher on some scales than human art. (2017, pp. 20-21)

Is this enough to render the network creative? I wish to raise two worries that it is not. First, even if the network is able to depart from existing styles, this does not mean that the network is able to generate its own distinct style as opposed to producing work which is merely ambiguous between different already existing styles. This is akin to an artist experimenting with different styles at the same time. Such a process can lead to creative outcomes, but stylistic ambiguity by itself does not suffice for creativity. The way in which an artwork is ambiguous between different styles must itself be inventive for the work to be creative. Yet, it is not clear that the network achieves stylistic ambiguity in an inventive manner.

The second worry grows out of this first worry. It is worth noting that the way in which the network deviates from stylistic norms is modelled after Colin Martindale's evolutionary theory of artistic creativity. According to Martindale (1990), all artists work within a set of conventions which define a particular style. Yet, creative artists are also under pressure to maintain the "arousal potential" of their art in response to spectators becoming habituated to their art. Creative artists therefore have to produce ever more novel and surprising works while still remaining within the parameters of what spectators deem aesthetically pleasing. Eventually, however, this process will lead to a breakdown of stylistic conventions, thus making a shift in style inevitable only for the whole cycle to be repeated.

Elgammal's network therefore presupposes a specific theory of creativity. This in itself is a significant methodological concession. Whereas Boden's hope was that artificial intelligence could be a means to understanding creativity, the explanatory order has now been reversed. Now we first have to possess an adequate theory of creativity before we can construct machines which emulate this creativity. This is to relinquish one of the original missions of artificial intelligence, namely to elucidate the nature of thought via the construction of thinking machines. Instead, the human capacity for creative thought reappears as the crucial explanandum in need of an explanation.

Yet, it is not clear that we will ever possess an adequate theory of creativity that can be applied in the building of creative machines.

Creativity may well be what Morris Weitz (1956) terms an "open concept" which resists definition, since it is by nature emendable and corrigible. If that is the case, we cannot give a general theory of creativity, since it will always be a matter of contention whether we should extend or restrict the concept to include more or less things in its domain. What we can do, of course, is to close the concept by restricting ourselves to a certain historical context much as Martindale does when he describes past creative cycles. In that case, we are using the concept in a purely descriptive sense to describe how creativity manifests itself at a specific time and place.

What we cannot do, however, is to generalise from such a contextualised description of creativity to a general theory of creativity. Even if Martindale's theory is adequate as a theory of past creative cycles, it may not generalise to future instances of creativity. If the concept of creativity is an open concept, it also has an evaluative component which invites us to consider whether an idea, object, or procedure deserves to be classified as creative regardless of whether it has been classified this way in the past. For the same reason, a machine may not be creative even if the procedure that it is instructed to follow has proved to be creative in the past. After all, it will be open to contention whether a procedure which is designed to mimic past creative achievements itself deserves to be called creative or merely repetitive.[6]

[6] Indeed, some might argue that it is impossible to program a machine to follow a procedure which mimics past creative achievements without thereby making the machine

4. Conclusion

If what I have argued is correct, we are still some way from seeing genuine artificial creativity. Although past and current attempts at producing artificial creativity yield many noteworthy results, they ultimately fall short of genuine creativity. Most importantly, we are yet to see machines and programs which not only traverse a creative domain to produce novel products, but which do so in a manner which itself is creative. What is more, we are also yet to see programs which are able to develop their own distinct style where this style amounts to more than mere ambiguity between already existing styles. Until we know whether it is possible in principle for a machine to achieve these feats, it remains unclear whether genuine artificial creativity is at all possible.[7]

References

Boden, Margaret A. (2005), *The Creative Mind: Myths and Mechanisms: second edition*, Abingdon: Routledge.

uncreative. This, one might argue, would remove the element of surprise and originality which is characteristic of creativity. Far from being free to come up with surprising and original solutions to creative problems, the machine would merely be predetermined to re-enact the trajectory which some past creative achievement has originally formed. I develop this point in a longer version of this paper.

[7] This research was made possible by a postdoctoral fellowship from the Fritz Thyssen Foundation.

Colton, Simon (2008), 'Creativity Versus the Perception of Creativity in Computational systems', *AAAI Spring Symposium: Creative Intelligent Systems*, Vol 8.

Dreyfus, Hubert (1972), *What Computers Can't Do*, London: Harper & Row.

Elgammal, A., Liu, B., Elhoseiny,, M. & Mazzone, M. (2017), 'CAN: Creative Adversarial Networks Generating "Art" by Learning About Styles and Deviating from Style Norms', *Proceedings of the Eighth International Conference on Computational Creativity*, pp. 96- 103.

Fried, Michael (1996) *Manet's Modernism: or, The Face of Painting in the 1860s*, Chicago: Chicago University Press.

Gaut, Berys (2009), 'Creativity and Skill', in: Bardsley, K. Krausz, M. and Dutton, D. (eds.) *The Concept of Creativity in Science and Art*, second edition, E. J. Brill.

Kant, Immanuel (2005), *Critique of the Power of Judgment*, (ed.) Guyer, P. Cambridge: Cambridge University Press.

Martindale, Colin (1990), *The Clockwork Muse: The Predictability of Artistic Change*, New York: Basic Books.

Novitz, David (1999), 'Creativity and Constraint', *Australasian Journal of Philosophy*, Vol. 77, No. 1, pp. 67-82.

Vasari, Giorgio (1998), *The Lives of the Artists*, Oxford: Oxford University Press.

Weitz, Morris (1956), 'The Role of Theory in Aesthetics', *The Journal of Aesthetics and Art Criticism*, Vol. 15, (1), pp. 27-35.