

A Comparison of Human Creativity and Generative Music Production

Dr. Esther Fee Feichtner¹

¹Artificial Intelligence in Culture and Arts (AICA) - University of Music and Performing Arts Munich (HMTM), Munich University of Applied Sciences (HM)

March 23, 2024

ABOUT MUSICAL TASTE

People like what is familiar because it often already carries emotional connections, briefly reviving them and causing resonance. Moreover, we are energy minimizers, and familiar things do not require as much energy to analyze and process as new ones. If an acoustic experience is so new that we find little to no connection to something we've already experienced, we immediately discard the heard as "That's not my style!" or "Is this even music?" The proportion of new content can be greater the more different music styles we know and the better we have learned that it is worth the effort to consciously listen to the utterly unknown, and that the heard can be abstracted into metastructures to which we can find similarities within ourselves.

For light musical entertainment, the entirely new component should be rather small, and for this purpose, the algorithm behind generative artificial intelligence (AI) could not have been conceived more aptly. It draws its entire creative potential from the provided training data, exclusively from recurring patterns with the highest probability. What sounds so mundane is, however, very similar to human creativity.

CREATIVITY AND HUMAN SELF-PERCEPTION

Why have we been constantly reading articles about creativity since the beginning of 2023, although a technique in the field of machine learning, which was described in its basic form with the perceptron in the 1950s (Rosenblatt (1958)), is now just faster and has more training data available?¹

When the first known algorithmic composition - the Illiac Suite - was brought into the world by its human midwives Hiller and Isaacson in 1959, the outcry was incomparable (Hiller and Isaacson (1979)), nor when it was possible in the 1990s to generate new solos over chord progressions in different styles created by the user with the software Band-in-a-Box (Gannon (1991)). In these cases, human expert knowledge gained through experience and analyses was codified into rules and applied by a computer. This rule-based machine learning (also called "Symbolic Artificial Intelligence (AI)" or "Good old-fashioned AI") built upon expert knowledge, and the derived results were comprehensible.

In contrast, the capabilities of generative AI are unpredictable, and we cannot be shown the rules and conclusions a neural network has drawn from its training data like in the examples above. What we see are seemingly creative results that, it must be said, surprised us nonetheless.

This made it clear to us that what we perceive as creativity has an awful lot to do with data processing that can be digitalized, which almost does not require the active human, and this fact throws human self-perception into disarray.

Humanity has already necessarily gone through a specification phase at least twice. When steam engines took over the work of workers without needing sleep, food, and without demanding wages, who often defined their value in society through their labor capacity, it left for humans the qualities of

¹It should not go unmentioned that the solution to the vanishing gradient problem (Hochreiter (1991)) and other algorithmic improvements also contributed their part to the leap in capabilities of AI.

knowledge and logic, or broadly summed up, intelligence. When machines suddenly could perform calculations much faster and more precisely than we could, the species human had to call upon another virtue only humans would be capable of.

And here we stand, reading texts, looking at pictures, or listening to music that, until recently, could only have been created through the application of human creativity, now seemingly coming out of nowhere and with minimal effort before our eyes.

The definition of intelligence has benefited from machine competition and gained in profile with new facets, such as emotional intelligence. There, aspects can be better, well, or not at all fulfilled by computational processes. And creativity?

CREATIVITY AS A METHOD

Creativity begins with perception, followed by a filter deciding which impressions make it into our sphere of attention and how consciously it will be saved. For instance, we would not only save the scene of a tree on a hill in the sun but also individual aspects like the leaf shape, the colors, the smell, the warmth, and roughness of the bark, and so on. We might focus on the colors and only recall the smell when triggered by a particular catalyst. When we engage in creativity, we draw from this reservoir and combine stored impressions with varying granularity. Thus, attention and creativity are strongly linked.

Compositions with very coarse granularity of already heard elements result in new pieces of music that are very similar to other pieces in sound, whole chord sequences, rhythms, or other musical dimensions. We can speak of using large impression molecules. Whereas compositions that seem completely novel and unprecedented to us are composed almost from impression atoms, finely dissected impressions. A strong example is the use of individual sine tones instead of natural tones², or entire frequency ranges as in Stockhausen (Stockhausen (1956)).

The selection of our impression molecules and their combination corresponds quite exactly to what happens in transformers³ (Vaswani et al. (2017)). In AI, the decision processes for selection and combination use the probability of an element in connection with a *tag* or *keyword* from the learned data as a criterion. This part of creativity is - quod erat demonstrandum - technically feasible.

So, are machines creative? Yes and no. If we expand the definition of creativity to include the ability to select molecules and combine them in a way that does not yield the most probable outcome but something that is a particularly accurate reflection of a personal emotion or idea. This underlying personal idea, against which results are repeatedly checked, is also why machines leave humans far behind in one aspect. Algorithms are not inhibited by fear and can create not just one but often a multitude of results in seconds. Not being able to have emotions also means not fearing that future recipients will criticize the result, which has been compared so often with the inner idea, which in turn is strongly connected to one's identity. However, human creativity is not only inhibited by the fear of others' judgment but also by one's self-doubts about whether one can meet one's image, which can repeatedly stall the creative process.

The fearless creation of algorithms is actually only possible for humans in flow, in the highest mode of creativity.

Thus, a more complete definition of creativity emerges: The dissection of perceived, selection, and recombination (technically feasible) with the ability to compare an (intermediate) result with an inner theme - an emotion or idea. This ability, in turn, requires the capability to feel emotions and consciousness.

But, as we have already read about the nature of general musical taste, is this depth of creativity with the required emotionality actually necessary to create music? Or is this just a desperate wish to still assert oneself as the pinnacle of creation as a human?

MUSIC AND EMOTIONS

How much emotion needs to be infused into music depends on how much emotion the function requires that a piece of music must have in a given context at a minimum. These functions - or tasks - range from elevator music to improvised live concerts. Let's consider a few edge cases.

Elevator music serves to avoid silence so that elevator passengers do not feel pressured to start a conversation and to reduce the likelihood of potential claustrophobia. Thus, the music could be a great jazz standard performed by top-notch interpreters, but it does not have to be. For this purpose, a generated

²Natural tones are composed of individual sine oscillations.

³A transformer is a network architecture that many generative AI systems are based on.

piece, in which melodic solos over typical jazz chord progressions are automated and randomly created, is more than sufficient. The more typical, the better, as it does not demand any attention and thus no energy from the listener, who experiences no acoustic surprises.

On the "required emotions" axis, the following scenario is a small step further: A jingle for an arbitrary (in our example, non-artistic) podcast. Function: A short motif that serves as recognition for the podcast. Thus, without any musical knowledge, one could generate 50 possible jingles and would only need one's taste to make the emotional match.

Far on the other side of the scale are musical experiences that embody the function of deep empathic connection between people. For instance, anyone who has seen Joe Cocker perform on stage (see, for example, *lomey* (2007)) knows that just watching an artist deeply immersed in his work and conveying it through strong expression in sound and body language acts as an empathic catalyst on the viewer, in other words: synchronizes the emotions of the audience and the artist.

In general, it's technically possible for any music that needs to fulfill a function, which does not require a human artist with passion, empathy, and intention, to be replaced by an AI system. This applies to many scenarios in today's music production where the listener's main focus is not on the depth of creation of a piece of music. For example: music for advertising, where its purpose can be articulated, chord accompaniment for practicing at home, background music in restaurants, music to encourage dancing or celebrating, and many others. As mentioned earlier, in many cases, the purpose is even better fulfilled the less novelty is found in the music piece.

What remains is art.

THE NEW ROLES OF AI

With the technical capabilities of generative AI, it is possible to generate music that is sufficiently suitable for many areas of application and can no longer be distinguished from purely human works. However, the more meaning a piece of music is supposed to have through its function, the more important the capability of human emotion and empathy becomes in the background.

What does this mean for today's musician? To completely turn away from music because AI will soon take over everything, or should one, as a "serious artist," distance oneself from AI to remain authentic? To approach this question, we must consider at least three different roles that AI can take on in creative work:

- *AI as a Simple Tool* with arbitrary or trivial human input.
Example: "3-second jingle for a podcast about South America travel"
- *AI as Inspirator and Tool*: The confrontation with a potential end result leads to an internal stance by the artist, who must match it with their inherent idea. However, this idea is continually developed through the perception of the generated output, which itself is based on the works of all artists that have flowed into the training and the prompt of the active artist. Thus, one can speak of inspiration and re-inspiration.
 - A particular advantage of human-AI "co-creativity": Interaction with AI instead of a human takes away creative fears. And since genuine creativity cannot arise in an environment of "premature judgment," even the impression of a creativity accelerator can emerge, although there simply are no impediments.
 - Disadvantage: Creative synergy, as in human-human co-creativity, is made possible by a shared enthusiasm for a shared vision. This emotional synchronicity is missing in human-machine co-creativity.
- *AI for Modality Transfer*: As mentioned earlier, a fully trained artificial neural network that delivers great results is still a black box to the human mind and not formulable in a rule-based manner. However, this disadvantage also has tremendous potential. Thus, it is possible for AI to learn difficult or non-formulable correlations that are present in a dataset, possibly even across the boundaries of individual modalities. Here are a few examples:
 - Image to Sound
 - Gesture and facial expression to intensity (for example, in the analysis of a conductor)

- Emotion in word to emotion in facial expression
- ...

This third role is the not-so-frequently used superpower of AI that can also open new doors in human creativity.

CAN AI CREATE ART?

If AI is also capable of suggesting emotions, are there still limits? Can AI create art?

In the role of an inspiration supporter and as a translator from one form of expression to another, AI is used as a tool to implement a human vision. It's interesting to consider how a work is to be judged that was created without artistic intention.

According to Tasos Zembylas (Zembylas (1997)), the definition of art is subject to constant change, naturally fueled by such disruptive technological leaps as those brought about by generative AI. Thus, we cannot avoid dealing with the extent to which the personality of an artist is indeed required for the concept of art.

Humans are, first and foremost, social beings. Therefore, our connection to others and shared exchange are of utmost importance to us. This forms the root of our empathy and our interest in discovering what moves others internally. This is also why we harbor a different interest in art created by humans as opposed to a random product, even if we could no longer distinguish it, either acoustically or visually, from a work made by humans.

As mentioned at the beginning, humans are also inclined to minimize effort, and because attempting to fully engage with something and gain an understanding of a work of art requires emotional openness and energy, we need to be convinced that there is an intention and thus an emotion behind the facade, which we can experience in successful synchronicity.

Since what is the reason to devote oneself to something that was created without devotion?

REFERENCES

- Gannon, P. (1991). *Band-in-a-Box*. PG Music. *Inc., Hamilton, Ontario*, 1999.
- Hiller, L. A. and Isaacson, L. M. (1979). *Experimental Music; Composition with an electronic computer*. Greenwood Publishing Group Inc.
- Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen. *Diploma, Technische Universität München*, 91(1):31.
- lomey (2007). Joe Cocker - you are so beautiful (nearly unplugged).
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.
- Stockhausen, K. (1956). *Studie II*. Number 3. Universal Edition.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Zembylas, T. (1997). Kunst oder Nichtkunst: Über Bedingungen und Instanzen ästhetischer Beurteilung.