Lyndon Chen
5th November 2024

Resilience in the face of AdversiCathy: an unforeseen continuation of *Stripsody* (1966)

My first exposure to "new" music, though I didn't know it at the time, was in Mr. Poyser's third-grade music class, during which he showed us different forms of musical notation. There are, he explained, a vast number of ways one might express instructions to produce a noise, from Steve Reich's *Clapping Music* to John Cage's *4' 3"*; scores were not confined to vast reams of black and white blobs. The one that I remember the most vividly, however, was *Stripsody*, written by American vocalist and composer Cathy Berberian (who had the somewhat peculiar habit of inserting her own name into the titles of her compositions). Listening to a recording was a distinctly new, or rather, novel experience, as even our eight-year-old sensibilities were somewhat affronted that this comic book-inspired squawking might be held in the same vanguard of Important Music To Be Studied as, say, Mozart or Beethoven. Walter Benjamin, therefore, should have attributed me in positing that "The conventional is uncritically enjoyed, and the truly new is criticized with aversion", in Part XII of his essay on cultural criticism, *The Work of Art in the Age of Mechanical Reproduction*. It is difficult to claim that any art, and therefore thought, is "truly" new—Berberian was certainly not the first to use graphic scores, and it is impossible to know for sure whether her use of onomatopoeia and cartoon elements was the first the world has ever seen. Nevertheless, I find it important to pay homage to Berberian and her work as firstly, to continue to challenge the pre-existing notions and structures of art is empowering in its inherent "tension between moral and aesthetic passion", as Susan Sontag identifies in her essay *Notes on Camp*; and secondly, because I feel that Berberian would approve of more art in the world that, at its core, is just a little bit silly.

The model was trained on a dataset that is a) not mine, b) relatively recent, and c) quite small, which arguably introduces a smorgasbord of problems into the work, but also raises a number of intriguing questions. The original graphics were conceptualised and structured by

Berberian, but were designed by Italian artist and satirist Roberto Zamarin—to whom would licensing fees be paid to? To nobody at all, one might argue; in Chapter 10 of *Art in the Age of Machine Learning*, Sofian Audry draws parallels between "machine learning remixes" and the similar imitations of styles and genres as carried out by the Dadaists "using pictures to make collages and photomontages", or sampling existing music to create one's own. Furthermore, I would argue that my transgression is minor within the total potential of AI misuse—it fails to satisfy any of the criteria mathematician Cathy O'Neil describes as indicative of a "weapon of math destruction". It is incredibly transparent, being open-source code; it is regulated by the YCRC (and my own computer science capabilities); and it has been used as a learning tool, not a position where it might have dominion over a human existence, and so lacks anything to even challenge in the first place. The model is, however, scalable—Berberian's original score can be performed in roughly six minutes, so the results aren't as impressive as they might be for other larger collections of graphic scores, such as those of Morton Feldman or Cornelius Cardew. Of these, however, I chose to use Berberian's score as the training images due to both time constraints and the greater entertainment value.

It is due to this reduced data set, therefore, that the predictions are not very visually striking—generated visual art with today's computing potential seems synonymous with Refik Anadol's large-scale architectural presentations and rooms with floor-to-floor screens. This may partially be due to the training time; twenty-four hours may not have been enough time to capture the nuances in Zamarin's illustrations, in which the positioning, typeface, and shading conveys a lot of the character Berberian brings out onstage. I was particularly disappointed that it didn't manage to reproduce any of the shaky handwriting on the second page of music ("YOU STUPID KITE, COME DOWN OUT OF THAT TREE!"). My favourite pages of the predicted score, however, capture Berberian and Zamarin's use of juxtaposition and movement—the original score contains such a diversity of sounds (which Berberian translates to numerous, highly-entertaining facial expressions and gestures) that, packed together with little

context or segue, give *Stripsody* its distinctive character. I was also surprised at the diversity of how precise the musical staves were in the final predictions, as I had observed them being mostly accurate in earlier iteration; a number of the 100 seeds I used produced lines that curved randomly about the page, which I haven't used in order to be able to stay true to Berberian's interpretation and performance as possible.

I suspect that there aren't many works on the syllabus that bear an immediate resemblance to this project as a black-and-white cartoon strip of mostly white space is quite a mundane thing to harness the power of generative AI for, and so I suspect that whoever responds to this might struggle initially to grasp something tangible or meaningful about this work. I was, however, quite inspired by an article I came across, written by Google DeepMind research engineer Yennie Jun, titled "Can AI read music?". After testing and evaluating the "visual reasoning capabilities of Large Language Models", Jun concludes that (as of April 2024) various models struggle to identify, let alone evaluate what humans might consider the basic elements of music theory: key signature, time signature, rhythm, for example. My initial response to this article was to wonder if the same scores entirely notated out in solfège (and whatever the equivalent for rhythm is) might be recognisable, since the models would find this much easier to parse, though the prompt might be a lot longer. Intriguingly, however, Jun's remark on the irony of Anthropic's model being named "Opus", yet being unable to correctly name a single score from a picture of an excerpt, seems to echo one of the misconceptions about AI computer scientist Melanie Mitchell lays out in her essay *Why AI is Harder Than We Think*: "the lure of wishful mnemonics".

One might critique this work for not fitting the requirement of being a visuals-forward work, and that the accompanying audio is far more interesting than the predicted images themselves. My response to this differs depending on the perspective of such a critique. If it is merely an issue of the lack of visual stimulation, then I would turn to monochrome artworks of Robert Rauschenberg or Ad Reinhardt as examples of things that might be "boring" to look at it

that still convey meaning in themselves. As far as my own work is concerned, I argue it to be a testament to the importance of intentionality in creating, interpreting, and critiquing art; the contemporary artist's refrain of "you could have done it, but you didn't" comes to mind. More convincing, in my opinion, is the view that this intentionality is derived from the entertainment factor of the silly noises on the audio file, that is, not primarily from the visual component of the work. I would instead argue that the audio associated with both this work and *Stripsody* are auxiliary to the visual stimuli; by a rather first principles argument, the audio was created in response to the visuals. Additionally, I believe that the visual aspect of this work both creates the framework for the audio to be meaningful; presenting the audio by itself would not be a sufficient continuation of *Stripsody*, nor would it be distinguishable from outtakes from a sound effects booth. Finally, it is the score itself that is responsible for challenging our preconceptions of art. While the audio itself may be more engaging, by itself, it is merely a means of enabling the viewer to access the surprising visual juxtaposition of comic book fonts and the highly stylised, pristine space that art occupies in our collective worldview, somewhat similar to Marshall McLuhan's argument of the medium conveying the message more so than the result itself.

Perhaps the fact that this work needs defending is indicative of its quality as a project, though I take some comfort in the knowledge that it would be impossible for a submission to stand entirely on its own as a piece of art. I live in the hope that Audry's proposal that analysis of generated art is "adjacent", as opposed to directly equatable to, the mechanisms by which we conceptualise and understand human art; that is, by responding poorly to this work, you might just not know any better.