A Study of Artificial Intelligence for Creative Uses in Music

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Cameron Lange
Spring, 2020

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

A Study of Artificial Intelligence for Creative Uses in Music

The Quest for Artificial Creativity in Music

The growth of artificial intelligence over the past two decades has allowed computers to achieve incredible feats that were once considered inconceivable, from identifying images and piloting cars in real time to winning a game of Jeopardy against human contestants (Marr, n.d.). Through deep learning and natural language processing, artificial intelligence (AI) has even made its way into spaces and industries defined by human creativity, including music and art. Some artists in the music industry already employ AI composition tools throughout the process of crafting and recording a song (*A.I. Songwriting Has Arrived. Don't Panic*, 2018).

This research paper examines current developments of artificial intelligence in the music industry and analyzes the motivations for and potential societal impacts of producing "creative" AI. To study this topic, the STS framework of technological momentum was utilized to determine how the relationships between musicians and technology have shifted over time, while actor-network theory was employed to investigate the potential impacts of creative musical AI on various groups, including computer scientists, musicians, and the general public.

Researching the Impact of Musical AI

What is the current state of artificial intelligence within the music industry, and how would the ability of AI to emulate human creativity within the music-making process impact society?

The main method of exploring this research question is organizing case studies on current usage of artificial intelligence to produce music. AI is currently being used in several capacities within different groups; computer scientists use AI to compose entirely new pieces of music from scratch, while singers and songwriters use commercially available computer programs to

help compose instrumental pieces of songs, known as “beats.” Each of these different use cases is examined through the lens of the overarching research question, meaning that the impact on society of AI used in both original music composition and as a tool for human artists are taken separately. This important distinction allows for a more robust understanding of exactly how “creative” AI can shape society, as there is a relatively wide gap between the worlds of classical music composition and popular music production. Within the realm of popular music, streaming and charting data is also analyzed to help determine the current cultural relevance of AI as a production tool.

For each case study, available primary resources for original music composition include technical papers and blog posts about artificial intelligence programs and algorithms, which additionally contain motivations for research and potential progress through comparisons to classically composed music. Resources for cases of AI in popular music include mainly streaming data and interviews that are analyzed and interpreted to determine the impact of potentially widespread replacement of human producers with computer programs.

Small scale Turing tests, in which computer programs try to ‘fool’ humans, are also conducted with both computer-composed music and computer-assisted popular music to determine the current state and convincingness of existing AI programs in music. A full list of survey questions is in the Appendix.

The Beginnings of Creative AI and the Importance of Music

The term “artificial intelligence” was coined in 1956 by John McCarthy, the “father of AI,” and later refined by Carnegie Mellon University and the Massachusetts Institute of Technology to be loosely defined as a system that could efficiently solve problems and learn by themselves (“World-Information.Org”, n.d.). Since that time, artificially intelligent computer

programs have grown and become almost ubiquitous in the culture of the 21st century. AI has mainly been used to simplify and automate human processes, but more recently, computer scientists have aimed to produce computer programs that can mimic human creativity, with some even calling creativity the "ultimate moonshot" for artificial intelligence (*The quest for AI creativity*, 2015). One area that AI has ventured into more recently is music, with sophisticated programs used to generate music appearing in both original music composition and as an assistive tool for popular musicians (*A.I. Songwriting Has Arrived. Don't Panic*, 2018).

Music has continuously been a defining piece of human culture throughout history, and it remains one constant between the many subcultures and societies across the globe. Despite there being no formal definition for music that spans all geographic locations, cultures, and time periods, "all peoples of the world sing" and "all peoples of the world have some form of instrumental music" (Trehub et al., 2015). This omnipresence of the desire to express oneself through rhythm and melody is a testament to the immense importance of music to human culture. Musician and neuroscientist Daniel Levitin notes that throughout history, music making has been as natural to humans as basic biological functions such as walking and breathing (Levitin, 2006). One of the main goals of music creation is the expression and incitement of emotion. During an interview with Thought Economics' Vikas Shah, award-winning composer Hans Zimmer stated that music acts as its own language, a language that allows him to hit "emotional targets" far more easily than any spoken or written words ("The Role of Music in Human Culture," 2017).

John Smith, who works with IBM's Watson, one of today's more well-known AI programs, notes that there is no set definition for creativity, and that for this reason it may be difficult to train computers to be truly creative. There are also concerns, however, on "how far AI … should go in the creative process" (*The quest for AI creativity*, 2015). These concerns can

easily carry over to the emergence of computer-generated music, as the implications of a computer imitating something so emotionally-based and rooted in human culture could be massive.

Technological Momentum and Actor-Network Theory in AI

The first of two main STS frameworks used to answer this overall research question is technological momentum, first developed by technological historian Thomas Hughes to describe time-dependent sociotechnical relationships. This theory asserts that technology and society are very tightly interconnected and that each group heavily influences the other, but that these relationships are not set in stone. As time moves on and technology and society each change, so does their shared relationship (Hughes, n.d.). Hughes draws from two different previously developed STS frameworks: technological determinism, which claims that technology irrefutably influences society (Smith, n.d.), and the social construction of technology (SCOT), which argues instead that the role and purpose of technology is shaped solely by social influences (Klein & Kleinman, 2002). Technological momentum is somewhat an amalgamation of these two contrasting theories by positing that society takes more of a deterministic role over technology at the onset of that technology's discovery and implementation, but that once technology becomes ingrained in society, it gains the ability to reverse roles and influence society itself.

The second major framework used throughout this research paper is actor-network theory (ANT). This theory, developed by STS scholars Michel Callon, Bruno Latour, and John Law, is far more abstract than technological momentum, and separates all aspects of a sociotechnical system into actors interconnected by a series of networks that signify relationships and influences. One important aspect of ANT is that it attempts to trace exactly how a current network came to be, and looks primarily to current "primary actors" to examine exactly how the

specific network was built (*Cressman—Overview of ANT.pdf*, n.d.). Actor-network theory also has the ability to regard non-human elements as "actors", and human elements as "networks". This framework effectively allows for a duality of interpretation: all human and non-human factors can be looked at as "actor-networks", but cannot be reduced to one simple concept. This theory is especially useful when examining music, which is such a large cultural and social entity that it can behave both as an actor as well as a series of relationships between many interconnected human and machine groups.

The Growth and Influence of Musical AI

Artificial intelligence is on the cusp of emulating musical creativity through classical single instrument arrangements and recordings, though it is not yet able to rival human compositions at this point in time. AI in popular music, on the other hand, is currently used by musicians more for effective and efficient production assistance rather than true composition. Artificially intelligent production tools have become sophisticated enough to be used viably in popular genres such as pop, hip-hop, and rock. Though not yet mainstream, the use of such tools is already gathering attention as an issue that will become polarizing to musicians, producers, and consumers in coming years as a balance between creativity, authenticity, and opportunism is sought after.

The purest form of artificial intelligence and machine learning used in music comes in the emulation of classically composed pieces of music, as these arrangements are often performed by a single instrument and involve many complex rhythms and patterns that AI programs can model. One promising use of AI to generate music comes from neural networks, computer programs "modeled loosely after the human brain" that are particularly useful for recognizing patterns in complex data (Nicholson, n.d.). Machine learning developer Alex Yu used neural

networks in a 2019 attempt to produce an extended piece of piano music arranged entirely by a neural network. Yu's motivation for the project was simply curiosity based on a history of playing the piano as a child, and the recognition that the complexities of music theory may fit well into an artificially intelligent program. After iterating through several variations of neural networks, Yu's final product utilized an algorithm called "Wave2Midi2Wave", which transcribes, generates, and synthesizes a coherent piece of audio that "sounds like it was a recording of an actual person playing the piano," though there are parts that "may not sound too pleasant," especially to an experienced listener (Yu, 2019). The algorithm used by Yu incorporates three cutting edge models trained on one of the largest and most robust existing musical data sets with over 172 hours of recordings, and comes very close to producing a final product that could pass for a piece of classically composed music.

The second aspect of creating a piece of music involves the actual performance, which artificial intelligence has also made great strides in emulating. In 2016, a research team from Google DeepMind produced a new algorithm known as WaveNet, which has the ability to generate raw audio waveforms that sound highly natural and realistic. This program creates waveforms manifested as speech, musical instruments, or any particular sound used to train the artificially intelligent algorithm (Oord et al., 2016). One particular use case of the WaveNet program involves training the algorithm on samples of classical piano pieces and then producing new audio files from scratch in attempts to emulate live instrumentation (*Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset: Online Supplement*, n.d.). A Turing test survey conducted for this research project compares samples of human-performed piano compositions to samples of piano compositions generated using WaveNet. In the first set of songs, only 27% of participants correctly identified that the WaveNet-synthesized song was

computer generated while the training sample was not. In the second set of songs, however, 73% of participants were able to correctly distinguish the WaveNet sample from the human-performed sample. These results indicate that while there is some validity to the assertion that an algorithm such as WaveNet is able to recreate a realistic representation of a live piano, there are still improvements that must be made to AI before computer programs can entirely replace human performance.

Artificially intelligent programs are still in the early stages of generating music composition, as evidenced by the fact that even with state-of-the-art technology, the final output of these programs is still marginally distinguishable from real human arrangement and performance. By the framework of technological momentum, at this point society influences computer scientists to continue on the path of musical creativity largely due to curiosity and the prospect of new discovery. As AI becomes more sophisticated and widely used within music, however, the deterministic roles of society and technology will reverse, with the capabilities of AI influencing musicians and computer scientists and changing the landscape of popular music as we know it.

Classical music, however, tends to reach smaller audiences and has less social significance in the 21st century, with genres such as pop, hip-hop, electronic, and country dominating the musical world in the ears of consumers. The more culturally relevant uses of musical artificial intelligence, therefore, are those that involve the creation and production of 'popular' music that falls into these more listener-friendly genres. Though there are currently no mainstream applications of popular music that are entirely produced and performed by computers, artists are beginning to discover that AI can be a valuable tool in the song-making process.

The use of computers to aid in various aspects of the music-making process is not a recent development. Platinum-selling artist David Bowie used a computer program called the "Verbasizer" throughout the creation of his 1995 album *1. Outside*. This program was a digital version of Bowie's "cut-up techniques for generating ideas and lyrics" (Smith, 2013) that took a string of text as input and cut it into different phrases to be used for lyrics. Another popular use of technology in music emerged in 1998 with the release of Cher's "Believe," which was the first obvious instance of the use of Auto-Tune, a pitch-correction technology that is estimated to be in up to 99% of pop music today (Reynolds, 2018). Each of these cases, especially Auto-Tune, was driven by a need for change by the artist in the ever-changing landscape of popular music. Musicians are constantly influenced by society and look for ways to separate from the competition, with Auto-Tune being a way to "forge new sonic signifiers for age-old emotions" (Reynolds, 2018). Despite its widespread usage, however, Auto-Tune also has its critics within the music industry who say that it has created a landscape that is too uniform and devoid of the 'human' elements of singing. One such critic is guitar maker Paul Reed Smith, who went as far as to assert that Auto-Tune "completely destroyed Western music" (Reynolds, 2018). The rise and subsequent criticism of Auto-Tune in music may foreshadow the future path of artificially intelligent programs used in music production, with technological advances moving quickly in recent years.

One recent example of artificial intelligence used in popular music comes from singer-songwriter Taryn Southern, who garnered fame from her appearance on the hit show *American Idol*. In 2018, Southern released her debut album *I am AI*, which was composed with the assistance of AI programs Amper and IBM Watson Beat (Reilly, 2018). Programs such as Amper have the capability of producing an instrumental beat focused around a specified genre or

mood. Southern utilized these features to create "basic random structures" of instrumental beats for each song on her album before reworking the small snippets to fit her desired product (Kharkovyna, 2019).

The different uses of computers by musicians over the last three decades have varied drastically, mirroring an industry that is rapidly changing and evolving. From Bowie's Verbasizer in the 1990s to the prominence of Auto-Tune in the early 2000s and the emerging artificial intelligence programs of today, computers have very clearly had a deterministic influence on the way society views and consumes music. The undertaking of such projects, however, is simultaneously a product of how quickly the music industry changes, and how most popular artists need to constantly reinvent, rediscover, and refresh their own sounds and styles while releasing a constant supply of new music to satisfy fans. The majority of new music today is consumed through streaming services such as Spotify and Apple Music, with over 77% of new music being listened to through these and similar services in 2018, compared to 17.3% through traditional album sales and 5.7% through single sales (Wang & Wang, 2019). The owner of a song's copyright, however, earns only a fraction of a cent per each song stream, with industry averages falling between $0.006 and $0.0084 (Wang & Wang, 2018). As a result of such low returns for the relative amount of work put into creating and performing a song, many artists feel pressured to speed up their creative process. Taryn Southern asserts that the use of AI programs like Amper can be used as tools to expedite rather than replace creativity. As a solo artist, Southern does not have extensive practice or training in all of the skills necessary to arrange and produce a song, and she notes that "it could take weeks" to successfully compose and perform a guitar arrangement without that very specific skill set, for example (Reilly, 2018). As a result, Southern is able to save not only valuable time but also money to surround herself with

management, publicists, and other necessary pieces of a successful team in hopes of making a living as an artist.

Southern is at the forefront of the AI emergence within the music industry, at a point where progress and relevance are largely dictated by technological advancements. As has occurred with Auto-Tune, however, it seems likely that as computer-generated music becomes more mainstream, society will then dictate the direction in which AI will move as it relates to the music industry. As is demonstrated by Hughes' framework of technological momentum, these relationships between computer scientists, musicians, and consumers will shift and grow over time along with the technology. Eventually, as musical AI becomes more sophisticated, it will either be accepted culturally as a part of popular music or rejected and pushed away. Already, some musicians and members of the music industry describe Southern's use of AI to produce an entire album as "unnerving", and some even claim that she is "cheating" the music-making process (Deahl, 2018). In a small Turing Test conducted to test the current viability of the Amper software used by Southern, only 27% of participants were able to correctly identify that Southern's "Break Free" utilized a computer-generated instrumental background while a second song, similar in genre, tempo, and mood, was entirely human-generated. These results would suggest that such programs are able to generate snippets of music within popular genres that are nearly indistinguishable from those produced by humans.

It remains to be seen whether new advancements in the coming years lead to AI being the next big technological advancement in music after Auto-Tune. Based on the responses of musicians such as Black Eyed Peas member Will.i.am, who refers to AI as a nothing more than a new "tool" to aid musicians rather than replace them, computer-assisted compositions could soon become the new norm within popular music. Such a development would suggest that artificially

intelligent programs used in popular music are already at the turning point within the theory of technological momentum where technology shifts from being shaped by society to driving its own fundamental changes in users and consumers.

Currently, AI acts largely as a 'network' as opposed to an 'actor' in the relationships between musicians and listeners. Musicians are still the figures associated with the output of songs, while the technology is only a tool that musicians can utilize. As AI grows, however, computer scientists and their creative programs could turn into actors themselves if a point is reached where each and every aspect of a song is generated entirely through a computer program.

One main limitation of this project involves the rapidly changing landscape of computer science and artificial intelligence. Despite research having been conducted on some of the most recent and technologically advanced algorithms and products, new research and products will emerge in years, months, or even weeks that could drastically impact the interactions between computer scientists, musicians, and consumers described in this paper. Additionally, the Turing test survey conducted was limited by the relative lack of available music samples to test with as well as a relatively small pool of test subjects. A more robust and comprehensive test conducted with more music samples on a larger number of subjects would work to validate the results discussed in this project.

One exciting possibility for future research on this topic involves the discussion of artificially generated pieces of classical music with figures who are very knowledgeable about music theory – a professor within the Music department, for example – and would be able to more easily distinguish programmed music from human-composed music. Conducting a Turing

test on a pool of such educated subjects would also provide a new benchmark for the success of such computer programs.

Conclusion

Despite recent efforts in computer science and artificial intelligence, even the most effective algorithms are not entirely capable of passing a Turing test for total music generation and production, and programs used by popular artists still require some human input and manipulation to achieve usefulness. The rise of AI programs within popular music in particular, however, is already becoming a divisive issue within the music community, and the divide between those who want to maintain the purity and creativity of songwriting and those who look to use AI as a tool to advance themselves in the industry will only deepen as AI becomes more prevalent.

Appendix

Turing Test Survey Questions

Song Set X:

Song 1: [url]

Song 2: [url]

Which of the two songs listed utilizes AI? i.e. which song(s) appear to be computer generated, as opposed to being composed by a real human?

- Song 1 only
- Song 2 only
- Both Song 1 and Song 2
- Neither Song 1 nor Song 2

This question style was repeated three times, once for each set of two songs. Each song set contained one song that utilized artificial intelligence and one that did not.

References

Cressman—Overview of ANT.pdf. (n.d.). Retrieved October 24, 2019, from https://collab.its.virginia.edu/access/content/group/1f0a8b93-9859-49c7-a64e-6091d596faf7/Readings/Cressman%20-%20Overview%20of%20ANT.pdf

Deahl, D. (2018, August 31). How AI-generated music is changing the way hits are made. The Verge. https://www.theverge.com/2018/8/31/17777008/artificial-intelligence-taryn-southern-amper-music

Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset: Online Supplement. (n.d.). Retrieved February 21, 2020, from https://storage.googleapis.com/magentadata/papers/maestro/index.html

Hughes, T. Technological Momentum.pdf. (n.d.). Retrieved from https://collab.its.virginia.edu/access/content/group/1f0a8b93-9859-49c7-a64e-6091d596faf7/Readings/Hughes%20-%20Technological%20Momentum.pdf

Kharkovyna, O. (2019, July 30). Artificial Intelligence and Music: What to Expect? Medium. https://towardsdatascience.com/artificial-intelligence-and-music-what-to-expect-f5125cfc934f

Klein, H. K., & Kleinman, D. L. (2002). The Social Construction of Technology: Structural Considerations. Science, Technology, & Human Values, 27(1), 28–52. https://doi.org/10.1177/016224390202700102

Levitin, Daniel. THIS IS YOUR BRAIN ON MUSIC: Understanding a Human Obsession. Penguin Books, 2006.

Marr, B. (n.d.). The Most Amazing Artificial Intelligence Milestones So Far. Forbes. Retrieved December 10, 2019, from https://www.forbes.com/sites/bernardmarr/2018/12/31/the-most-amazing-artificial-intelligence-milestones-so-far/

Nicholson. (n.d.). A Beginner's Guide to Neural Networks and Deep Learning. Pathmind. Retrieved February 21, 2020, from http://pathmind.com/wiki/neural-network

Oord, A. van den, Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., & Kavukcuoglu, K. (2016). WaveNet: A Generative Model for Raw Audio. ArXiv:1609.03499 [Cs]. http://arxiv.org/abs/1609.03499

Reilly, D. (2018, October 25). A.I. Songwriting Has Arrived. Don't Panic. Fortune. https://fortune.com/2018/10/25/artificial-intelligence-music/

Reynolds, S. (2018, September 17). How Auto-Tune Revolutionized the Sound of Popular Music. Pitchfork. https://pitchfork.com/features/article/how-auto-tune-revolutionized-the-sound-of-popular-music/

Smith, C. (2013, March 8). Ty Roberts: From Working With David Bowie To Co-Founding Gracenote. Hypebot. https://www.hypebot.com/hypebot/2013/03/ty-roberts-on-the-trail-from-working-with-david-bowie-to-co-founding-gracenote.html

Smith, M. R. (n.d.). Technological Determinism in American Culture. . . The MIT Press., 35.

The quest for AI creativity. (2015, September 11). IBM Cognitive - What's next for AI. http://www.ibm.com/watson/advantage-reports/future-of-artificial-intelligence/ai-creativity.html

The Role of Music in Human Culture. (2017, August 26). Thought Economics. https://thoughteconomics.com/the-role-of-music-in-human-culture/

Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1664). https://doi.org/10.1098/rstb.2014.0096

Wang, A. X., & Wang, A. X. (2018, August 8). How Musicians Make Money—Or Don't at All—In 2018. Rolling Stone. https://www.rollingstone.com/music/music-features/how-musicians-make-money-or-dont-at-all-in-2018-706745/

Wang, A. X., & Wang, A. X. (2019, January 3). Album Sales Are Dying as Fast as Streaming Services Are Rising. Rolling Stone. https://www.rollingstone.com/music/music-news/album-sales-dying-as-fast-as-streaming-services-rising-774563/

World-Information.Org. (n.d.). Retrieved October 24, 2019, from http://world-information.org/wio/infostructure/100437611663/100438659360

Yu, A. (2019, March 8). Generating Music With Artificial Intelligence. Medium. https://towardsdatascience.com/generating-music-with-artificial-intelligence-9ce3c9eef806