Outline

Summary: In this research project, we try to understand the scope and current state of the use of AI in generating jazz music. Although AI now plays a role in several aspects of the digital music experience, from how we find new music on streaming services and how that music is produced using AI production tools, we focus mainly on full-scale AI music generators. Principally, we consider three recent projects in this field, one that generates raw audio files based on a massive, 1.6 million song corporas, and two others that generate music improvisations in symbolic form. The music generated by this work is rich and complicated in its musical composition, and raises many questions about the progress and significance of computer generated music, and its importance in the broader context of musical production. After analyzing and comparing the musical elements produced by these models, we discuss a number of articles that attempt to make sense of this relationship between computers and music. In this section we consider questions such as whether computer art is as valuable as human art and whether computers are even theoretically capable of producing music that rivals human creativity. We consider these questions in context of the musical analysis of modern models that we discuss earlier, and conclude with any remaining open questions and a discussion of areas that deserve further research.

- **Introduction**
  - AI and music
    - In recent years, AI has been increasingly used across different realms of the music world.
    - In music production, tools like Magenta Studio bring AI tools to mixing and mastering in music production.
    - In music streaming services such as YouTube and Spotify, which are now the dominant platform for how we consume music in the US, AI is used extensively in reccomendation systems that intimately shape how we discover new artists.
  - Brief history and context of computer-generated music and jazz.
    - Computers have a long history in the production of music. Alan Turing, one of the fathers of computer science, used an early computer to generate music in 1951
    - Recent advances in AI, especially in language creation with models like GPT-3 has raised the question of whether similar advances can be made in music generation.

- **Musical Analysis of AI Jazz Projects**
  - Jukebox (2020) is "a neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist styles" developed by AI research group OpenAI.

  - Jukebox is a pioneering project in its rejection of symbolic music generation in favor of raw audio generation: "But symbolic generators have limitations—they cannot capture human voices or many of the more subtle timbres, dynamics, and expressivity that are essential to music."

  - The model was trained on 1.2 million songs, (half were in English) across a full range of genres and styles.

  - Limitations: "For example, while the generated songs show local musical coherence, follow traditional chord patterns, and can even feature impressive solos, we do not hear familiar larger musical structures such as choruses that repeat."

  - The team also notes that they "shared Jukebox with an initial set of 10 musicians from various genres to discuss their feedback on this work. While Jukebox is an interesting research result, these musicians did not find it immediately applicable to their creative process given some of its current limitations."

  - I would like to dive deeper into many of jazz generated samples. Some of the most interesting are ones that are generated "with unseen lyrics". For example, this sample produces "Ella Fitzgerald" singing lyrics about the Jukebox AI.

  - Another interesting sample is this one in the style Nat King Cole which includes extensive piano solos.

  - BebopNet is a neural net for generating jazz improvisations using a symbolic approach. Due to this, it has a much "cleaner" sound than Jukebox does, because it only has to generate the sequences of notes which can then be produced digitally.

  - Trained on a dataset of jazz solo transcriptions from a variety of artists including Charlie Parker, Sony Rollins, Cannonball Aderley, Stan Getz.

  - In my analysis, I'd like to look at the samples produced by BebopNet, and see specifically how they compare to human-produced solos from artists in the training set. In this, I could see what aspects of jazz are represented in the AI's generation.

  - Music Composition using LSTM Recurrent Neural Networks is a paper where the authors specifically trained the machine learning model on symbolic 12-bar blues samples. The samples were custom generated by the authors: "The melody line was built using the pentatonic scale (Figure 3)

commonly used in this style of music [blues]. Training melodies were constructed by concatenating bar-
long segments of music written by the first author to fit musically with each chord."

  – In my analysis, I can look at the samples generated by the model and see what elements of the blues are captured in the music, and how this compares with some of the elements emphasized in the training data or in blues in general.

- **Discuss/compare musical qualities of each AI jazz project.**
  – Specifically: How does OpenAI's "raw audio" approach compare to the symbolic generation used by the other models analyzed?

  – Does raw audio present a step forward in the realm of music generation, or is the symbolic approach ultimately more useful?

    • What musical elements and jazz continuities (if any) are found in the raw audio approach? Could look specifically at timbre elements, since the authors mention that representing timbre was one of their main motivations for pursuing a raw audio approach.

  – The two other projects both focus on jazz and blues improvisation from symbolic data, though one was trained on transcriptions of jazz soloists while the other on hand-made data. How do these two results compare in terms of their musical qualities?

  – Is there any novel qualities found in the AI jazz?

    • Lyrically, a fruitful place to look would be the No computer art could match the emotional heft of a human original No computer art could match the emotional heft of a human original No computer art could match the emotional heft of a human original Jukebox project, where the AI generated music of singers singing lyrics that the research group had written. How does these lyrics compare in their musical quality to the real lyrics sung by, for example, Ella Fitzgerald or Billy Holiday?

- **Cultural and wider impacts of AI-generated music and art**
  – Discuss whether computer-generated art is less valuable than human art, specifically with reference to the projects we studied here.

    • In Jonathan Jones' opinion piece in the Guardian he discusses an AI model that is able to reproduce Rembrandt portraits, a task not entirely dissimilar to the one we are studying here in the realm of music. Jones says the computer productions are "a horrible, tasteless, insensitive and soulless travesty of all that is creative in human nature," and that "no computer art could match the emotional heft of a human original."

- Is Jones validated in this opinion in light of the musical elements we studied earlier? Or does his criticism stand to lose ground as the quality of the AI generation increases?

- Oliver Roeder takes a different stance in his article, arguing that because humans create computers and AI models, "all computer art is human art," and thus just as valuable. In this sense, a computer is just another tool for an artist as is paint or a saxophone.

- Roeder contextualizes this argument with the deepjazz project, which was created in a weekend by a Princeton CS student, and can generate jazz music in from a small input of just a few songs.

– How does AI inform our understanding of human music and sound? Is music a cognitive human process that can ultimately be modeled by sufficiently intelligent machines? Or does it consist of something more fundamental to human nature?

- The book Understanding Music with AI (available at Burling library) compiles a collection of articles and essays on the role of AI in music. This review by Jason Vantomme gives an overview of some of the important ideas discussed in the essays.

- One essay discusses general problems in modeling musical activities, and argues that in order to recreate a human musical production, the model must "consider the broader issues of what actually constitutes musical behavior." How does this argument fair in light of the AI projects we have studied, and compared to Christopher Small's "musicking"

- Another article argues that "computation, alone, might not be powerful enough to describe all of musical thinking," though the review suggest Kugel's essays is rich, and worthy of closer study. Still, it would be interesting to consider this argument in light of our previous research.

- The book Machine Models of Music (also available at Burling) is another anthology of seminal essays and papers in the realm of computer generated music from the 50s-90s. In this review by Ann McNamee she suggests that some of the article's musical analysis is flawed, but other articles which discuss foundational work in computer generated music and has worthy discussions of the value of these models.

- The article, 'Live Algorithms for Music: Can Computers be improvisers?' discusses the challenge of creating computer systems that can interact with human musicians: "The fundamental question is the capacity for computational processes to exhibit "creativity." The problems inherent in computer music performance are considered, in which computers are quasi-instruments or act in proxy for another musician."

- In light of the work in this article, we may consider how the AI projects we considered may be used by musicians, or if there are other areas where AI could contribute to musical creators.

- One area where I'd like to bring in more sources is in more recent essays, academic articles or books studying AI music since the recent advances in technology in the last 10 years.

- Another area worth more research is in the ethics of AI music generation. There is much work on studying the ethics of language models trained on large corpora that often reproduce racist or bigoted language, but I wonder if the same problem exists in music generation. If so, what would the implications be?

- **Conclusion**
  - Restate most interesting findings and thesis from research.

  - Discuss aspects of research that deserve more attention or further study.

  - Conclude with any open questions or interesting points about musical cognition or jazz uncovered during the research.