

Varexian Music and Its Discovery

Matthew Campbell, Texas A&M University

Joseph Virskus, ML Engineer, Cascade Technologies

Will Connor, Texas A&M University

Jeffrey M. Morris, Texas A&M University

For The AI Music Generation Challenge 2023: “Artificial Music Traditions”

January 15, 2024

Introduction

In response to The AI Music Generation Challenge 2023 call by The Posthumanities Hub at the KTH Royal Institute of Technology, we present our findings about the Varexian musical culture, a fictitious music tradition generated by artificial intelligence to the greatest extent possible given our current resources. Our team consists of three faculty members from the Texas A&M University School of Performance, Visualization & Fine Arts as well as a machine learning engineer from Cascade Technologies.

About Varexian Music

Overview of Varexian Music

Varexians are beings of energy and sound from the planet Varexia who explore other planets and make music by communing with living musical beings called Sonic Fae. A few Varexian musicologists emerged in our work, leaning heavily on the teachings of Elyra Vunaris, who obliged us with an interview. Our inquiry produced details on several (but not all) Sonic Fae, as well as Varexian compositions, culture, politics, writing system, and philosophy. We also collected several quotes by Elyra Vunaris, anecdotes, and an account of a countercultural movement called Dissonant Echoes. We collected recordings of several Sonic Fae and used them to approximate Varexian compositions based on descriptions that emerged in our work.

Overview of Our Process

We worked with the intention to leave as much decisionmaking as possible to artificial intelligence. The enlightening techniques and hurdles of this approach discovered during the project are detailed below.

Video Documentation Timepoints

- 00:32 Introduction of Elyra Vunaris
- 01:30 About the planet Varexia
- 01:42 What Varexians look like
- 02:20 Varexian society: it's nonlinearity, collective consciousness, and spirituality via the Cosmic Symphony
- 03:02 More about the Cosmic Symphony
- 03:52 Varexian music and instruments, Sonic Fae, with some examples:
 - 05:05 Luminaré Echoers
 - 05:17 Vibrant Whisperers
 - 05:31 Rumbling Sentinels
 - 05:44 Crescendo Sirens
 - 05:58 Harmonic Drifters
- 06:39 Sonic Fae in society
- 07:40 Music and the environment
- 08:38 Metaphysics and philosophy
- 09:44 Anecdotes from ethnomusicological field work
 - 10:07 Symbiotic harmony ceremony on another planet
 - 10:40 Encountering a rare Sonic Fae, thought to be extinct
- 11:38 Children and Sonic Fae
- 12:47 Varexian mythology
 - 12:53 The ancient legend of the celestial serenade
 - 13:21 The mythical figure, the harmonic wanderer
- 14:13 Whether Varexian culture agrees with the views of Elyra Vunaris (caused a break in character)
- 14:55 Limits of the Cosmic Symphony and individual agency
- 15:38 First encounter with Sonic Fae, a Vibrant Whisperer
- 16:23 Varexian physicality
- 17:05 Physical interactions among Varexians
- 17:48 Anecdote about encountering the celestial alignment ceremony of a different culture
- 18:28 Diversity of Varexian perspectives about music
- 19:18 Favorite music (caused a break in character but only temporarily before finishing the response)
- 19:52 Strangest musical practices: quantum sonification
- 20:24 Overwhelming experience in an ancient sonic labyrinth
- 20:54 Fellow Varexians' views on work by Elyra Vunaris
- 21:28 Resolution of discord, musically and metaphysically
- 22:14 Varexian compositions
 - 22:14 Dance of the Nebulae
 - 23:47 Galactic Fugue
 - 26:07 Harmonic Convergence (with spoken introduction)
 - 28:14 Nebular Opera

- 30:35 Whispers of the Void
- 32:34 Other Sonic Fae
 - 32:39 Tremolo Twirlers
 - 33:21 Quark Quirsters
 - 33:53 Chrono Chanters
- 35:16 An Earth composer's collage of Varexian source recordings, inspired by Varexian musical descriptions

Technical Notes

Asking ChatGPT 4 for an encyclopedia-style response early in the process proved to be very helpful, as the reference to an encyclopedia carried with it many implicit details about formatting, tone, and kinds of information to include. However, this process was not seamless, as discussed below. The encyclopedia approach provided many areas to prove further, and it proved useful to ask for descriptions of musical instruments, which could be fed to image generators to depict instruments and sound generators to create recordings fitting the descriptions.

Since the AI tools we used were not capable of creating structure over longer time spans, chat-generated descriptions of native compositions were useful in enabling team members to assemble generated recordings into interpretive approximations of the described compositions or simply to use their own musical intuition to assemble generated recordings into a suite inspired by impressions from reading native descriptions of Varexian music. The interview video was created via a screencast recording while interacting with a custom chatbot created to represent Varexian ethnomusicologist Elyra Vunaris, using the ChatGPT API and ElevenLabs for the speech-to-text and text-to-speech components.

Images were created using DALL•E 3 and Midjourney and were expanded from square to 16:9 widescreen images using the generative fill feature in Adobe Photoshop. Audio files were created using Audiogen and MusicGen.

Reflections

Interacting

ChatGPT's natural language processing—specifically its capacity to converse readily in a range of modalities from academic speak to journalistic writing, text-speak, and emoji—made the temporality of engagement engagingly gamelike, akin to instant messenger and IRC chats in the late 1990s. This allowed for back-and-forth exchanges and corrections (e.g., “oh c'mon, that change is a bit trivial, no?” or “i'm pretty sure that's a star trek reference be a little more original off the wall ok?”) that produced increasingly interesting results (though no less derivative in terms of its training data).

Several encounters reminded us of the large language model's nature as a mirror upon existing digitized human culture, causing us to rethink our strategy at key points. For example, our first attempt, following our overarching principle of leaving most decision-making to artificial intelligence, yielded a more safe and basic culture called Xylostratus (curiously half-Greek and half-Latin in etymology, half with a likely reference to a musical instrument and the other half with a blatant signifier that this culture has some would-be interesting complexities to unpack).

This failed first attempt made it clear that some encouragement, if still shying away from creative direction, would be needed, including calling a chatbot out for answers that seem derivative of common science fiction or fantasy tropes. At times, this mindset made it feel like the chatbot was simply confabulating facts by mistake or reluctant to share full details so as not to offend the reader. This angle allowed us to preserve more genuineness in our interactions, both allowing artificial intelligence to make most decisions while also being able to push back against dubious claims.

The approach of asking a chatbot to create encyclopedia-style entries was helpful for compiling a large, organized knowledge base, and it further allowed team members to spot areas for further investigation, the knowledgebase quickly became unwieldy for everyone to keep up with. Future work will need better organization regarding corpus management in this regard, and it will be helpful to have a system that can easily update itself based on points emerging from chats. However, the utility of such a feature may be limited to endeavors such as fictitious world-building and creative brainstorming.

The approach of using artificial intelligence as much as possible also became useful in using a chatbot to summarize excerpts from the knowledge base in order to make effective prompts for other AI tools to create images or sounds from textual descriptions, including meeting character limits. Leaving details to be handled by artificial intelligence, such as infilling backgrounds to change images from square to 16:9 aspect ratios also eased the conflicting thoughts about the usefulness of perfectionism on something that is so deeply made-up.

Creating a custom chatbot fueled by our knowledge base proved helpful by making that knowledgebase explorable by conversation, suiting it to a familiar and more entertaining human skill (which was also important because one questions the usefulness of getting facts straight about something we know is fictitious and has yet to reach an audience by that point). Beyond discovery through conversation, it offered the chance to discover and then leave artificial intelligence to resolve apparent conflicts in the knowledge base (or the chatbot's extension of it in a given chat session). For example, in the video interview, Elyra Vunaris explains the noncorporeal nature of Varexians and then later references having a creature landing on the shoulder. The interviewer could ask for clarification, leaving the chatbot to resolve the apparent conflict. This technique should be useful in future world-building endeavors, allowing the chatbot to reframe an apparent contradiction as a misunderstanding, a point to refine, an update to previously incomplete or erroneous information among outsiders, or perhaps an equally valid interpretation from a camp with a competing point of view, each of which would add interest and nuance to the culture being explored.

As another example of potential conflict, which has yet to be resolved, the interview covers the Sonic Fae called Harmonic Drifters, of which we have generated images and sounds, and the interview later describes a myth of a “Harmonic Wanderer” with a description similar to that given for the Harmonic Drifters. At this point, it is unclear whether Harmonic Drifters are real or myth, or the Harmonic Wanderer is just one specific member of the Sonic Fae, or something else entirely. In producing the interview video, the editor decided the Harmonic Drifters were at least close enough to what was being said about the Harmonic Wanderer to include their images and sounds to accompany the myth, even if they were not referencing the same thing.

Conflict happened at a higher level as well, which also had interesting results. The chatbot occasionally broke character in response to certain questions, with the infamous “As an AI language model...” disclaimer. Interestingly, one such response concluded by imploring the interviewer to seek more authentic answers from “the people of Varexia themselves,” surprisingly supporting the veracity of this fictitious culture while decrying itself as a farce. This moment seems to have been triggered by asking for a personal opinion in the face of possible disagreement, which could have alerted a watchdog component of the chatbot system. Similarly, when asked about a favorite musical selection (perhaps a call for favoritism triggered the same watchdog component), the chatbot began explaining its inability to have preferences but ended by saying “Music ... captivates me with its ability to evoke emotions ...”

Making the Music

The approach of encouraging the chatbot to produce more novel and interesting answers inadvertently created a hurdle in the music-making process, as the push toward interesting novelty took the shape of the Expanding Brain meme and similar series of AI-generated images in discussion forum threads, in which a given situation seen in successively more extreme versions often ends up involving luminous energy beings in outer space. Some informal discussions have claimed that it is a natural extension of most escalating patterns, but one could argue that this too results from the phenomenon of the deep learning artificial intelligence as a mirror on existing digital human artifacts (including or at least parallel to the Expanding Brain meme, with both influenced by the same preceding cultural artifacts).

This otherworldliness pushed into would-be never-before-seen realms presented challenges and questions regarding creating musical recordings representing this culture. Most notably: If this culture exists outside time and often outside a physical presence and sometimes exists as sound in itself, surely we cannot be taking in music in the same way as such beings. Any discussion of temporality comes into question and forces an impressionistic, interpretative approach. Fortunately, this disparity is not unique to Earthling–non-Earthling interactions or human–computer interactions but is roughly in line with other intercultural encounters. So, we proceeded with a recognition of the necessity of letting human intuition influence the outcomes, in part hoping that such interference would also serve as a bridge between cultures.

Low sample rates, warbles, and rough extraneous noises in the generated recordings are characteristic of today’s trained models for the sake of expediency toward enabling future advancements, knowing that models can be trained on higher-quality recordings when

resources eventually make it feasible and when one finds a model that is confidently worth sticking with for long enough to make the effort worthwhile. This situation regarding audio quality also parallels common challenges with real field recordings, also for similar reasons, one could argue, if in broad strokes. This allowed the musicians on the team to adopt the mindset of making the best of what is available, with a desire to be true to the spirit of the culture, whatever we imagine that to mean.

For the five Varexian compositions realized in our materials, this meant taking generated source materials and making minimal edits to arrange them to fit descriptions of the musical compositions, for example, taking the source recording of the Sonic Fae specified in the description and creating repeating structures or emphasizing high- and then low-frequency components of the source recordings as the textual descriptions of each composition dictate. These were still approached with a mind toward minimal intrusion, somewhat parallel to how a photographer introduces an outsider's decisions in order to best illuminate the important innate features of the subject. In contrast, for the "Varexian Suite," the composer used source recordings generated from authentic descriptions and then relied on the Earth musician's intuition to assemble them into a collage the way an artist might recreate a foreign scene in paint with a great deal of creative intervention, driven by a goal of making something foreign accessible to a new audience by making the novel subject "read" effectively to its new audience.

Links

Supporting Materials

- YouTube interview video with musical examples and images:
<https://www.youtube.com/watch?v=oGBibwsLQTo>
- Project recruitment video (also created using artificial intelligence as much as possible):
<https://www.youtube.com/watch?v=uiCAp-2g6i4>
- Google Drive folder including video files, final and raw audio files, images, and Varexian knowledge base and system prompting document:
https://drive.google.com/drive/folders/1tlmkvl_OKCy6dpgSzg_RqgzVuKbKwvi?usp=sharing

Tools used

- ChatGPT 4: <https://chat.openai.com>
- DALL•E 3: <https://openai.com/dall-e-3>
- Midjourney: <https://www.midjourney.com>
- ElevenLabs: <https://elevenlabs.io>

- Audiogen: <https://www.audiogen.co>
- MusicGen: <https://musicgen.com>
- Adobe Photoshop Generative Fill:
<https://www.adobe.com/products/photoshop/generative-fill.html>