

DDSP études for Tenor Saxophone & Violin

Guilherme Coelho

The New Centre for Research & Practice
guilherme.coelho@thenewcentre.org

Abstract. This work presents a collection of études and considerations on the use of the DDSP model (Engel, Hantrakul, Gu, & Roberts, 2020) as a quasi-object and epistemic tool that introduces sound objects and performances in the practice of AI interfaces for musical expression, machine musicianship and human-machine interactions. This research follows the work on AI and improvisation by George E. Lewis, and on Robert Rowe’s work on music and machine musicianship, on digital musical instruments as epistemic tools and cultural probes by Thor Magnusson, on machine learning and creativity by Rebecca Fiebrink, and Google’s Magenta research work and development of machine learning models for music.

Keywords: Neural Audio Synthesis, Machine Musicianship, Musicking, Idiomatic Performances, Transhumanist Musical Literacies

1. Introduction

Currently, music is experiencing another technological evolution of musical tools, sparked by peculiar, increasingly “conative” machine learning systems that can adapt to content and present it for peculiar forms of cognition – encouraging new modes of perception. New tools, such as the DDSP model, express new forms of ontology and epistemology in music, inspiring new forms of *musicking*¹ (Small, C. 1998) and machine musicianship (Rowe, 2001). New tools bring forth new practices and ways of perceiving art, and this work explores this territory.

2. On the études presented

The études stem from translations of its author, using audio inputs of synthesizer recordings. They explore this model's ability towards musical literacies that explore new modes of music interaction and modes of being outside of the author's reach and the respective instrument idiom. The études are mainly addressed as idiomatic performances, and they should be listened to as outputs from a machine musicianship filled with blemishes and moments of alluring idiosyncrasy. The model exhibits expressive, disjointed urges towards improvisational explorations that encourage modes beyond established practices of instrumental techniques of this instrument - showcasing ‘extended DDSP techniques’, expanding the limits available in the organic

¹ “To music is to take part, in any capacity, in a musical performance, whether by performing, by listening, by rehearsing or practicing, by providing material for performance (what is called composing), or by dancing.” (Small, C. 1998)

forms of these instruments and exhibiting singular forms of musicality. The DDSP model is an interoperable, moldable tool with conative behavior exhibiting agency and identity forms - presenting novel interpretations of input material and an ability to explore new forms of expressivity. These études specifically explore the limits of the interstitial spaces of synthesis of the respective representations of tenor saxophone and violin and the ability of this model to generate trans-idiomatic performances. I am not a trained instrumentalist and would not have reached these performances if not for this tool's interaction.

This work focusses on the DDSP Timbre Transfer Collaboratory. The Collaboratory enables the translation of inputs into the audio styles of four different instruments: violin, flute, trumpet and tenor saxophone. Each instrument has had a small training dataset of around 13 minutes of expressive instrumental performances from the MusOpen Royalty-Free Music Library. The model outputs audio in 16kHz.

3. Thoughts on the phenomenological experience of listening to these études

Here, I present some considerations on the listening environment caused by these DDSP translations and what this phenomenological experience might entail. Before I begin to address these thoughts, I would like to clarify that the following argument isn't an overarching assumption on the matters that I am about to discuss, but rather personal observations from a practitioner's point of view.

Carlos E. Perez wrote an effective analogy that is presented here to compare how machine learning models generate and process content, and the way listeners approach these études.

'DE3p Larennlg mhica3ns wrok smiliair to hOw biarns wrok. Tehse Machnies wrok by s33ing f22Uy pa773erns and cnonc3t1ng t3Hm t0 fU22y cnonc3tps. T2hy wRok l4y3r by ly43r, j5ut lK1e A f1173r, t4k1NG cmopl3x sc3n3s aNd br3k41ng tH3m dwon itno s1pmLe iD34s' (Perez, 2018).

Listeners approach these études similar to how these machine-learning models learn to generate content. These models attempt to make sense of the data they are 'hearing' according to what they have been taught, and in this listening experience the listener ends up experiencing a similar process: looking for moments of coherence, patterns of comprehensibility and moments of clarity according to the institutionalised notion of music as a form of language that humans, as socio-cultural beings, have engrained in their perceptions. Carlos E. Perez also states that, 'With Deep Learning, another kind of cognition, specifically intuition, is being automated. It is different from logical reasoning in that it is model free. Specifically, it begins without a knowledge representation of a problem and then creates one using the process of inductive reasoning. This is the same abstract process that human intuition performs: learning from induction' (Perez, 2018). Human intuition comes from an interaction with the designed world under structural confinements made within such definitions; machine-learning intuition comes from the data that it interacts with outside of structuralist boundaries and biological constraints. These pieces showcase similar dynamics in human approaches to listening to music and demonstrate that machine learning can be

used to approach music from different kinds of cognition and interpretation. That machine learning can lead to specific epistemic forms of listening, in this case, suggests that humans listen to these études in ways similar to machine-learning models: attempting to find sense and patterns according to the structuralist notions of what music is.

One of Schaeffer's motivations was to breach the fortress of musical tradition and open up 'fissures of randomness for the poetic adventure to slip through' (Rob Young, 2010), and AI models can act as tools that explore such notions. Musique concrète repositioned the act of listening and composing to new ways of making and experiencing music outside of structuralist external systems of reference. In *Search of a Musique Concrète*, Schaeffer writes:

'The object forces us to listen to it, not by reference, but just as it is, in all the reality of its substance. As it doesn't say much, and certainly not what we would like it to say, once we have heard it, it makes us fall silent. In this silence we perceive new disturbances' (Schaeffer, 1952).

The behaviours and material epistemologies presented by this model make it an epistemic tool: a tool leading to the conception of distinctive models that presents the delegation of creative interpretation in object-oriented terms according to the datasets that it is trained on. Interactions with this tool can exploit the underlying structure of input material and provoke new approaches and new ways of musical interaction. Examples of AI in music are often redundant and used as an end to achieve statutory contexts instead of a means to explore them, and this model can lead artists, researchers and listeners to extend the scope of their approaches and insights by engaging with this technology. By attempting to adapt to the DDSP object, we can listen to what the object tells us and learn from its singular object-oriented ontology.

Music mostly exists within confined structural confinements of language, genre and generic tropes institutionalised by culture among people from an early age (Williams, 2001). AI models can be used as a lens towards post-structural and nonstructural paradigms that exhibit what is valuable according to human interpretation. AI enables the possibility to build foundations and concepts free from pre-established notions, and where each model can even have its own environment. Practice-based engagements with AI tools can allow musicians to expand structurally determined understandings of music and expand concomitant limitations of one's view of one's practice. These models can expand creative practices and invoke critical reflection towards the acknowledgment of structuralism, limitations of music as a form of language and the intertextuality at play. My engagement with the DDSP model has brought forth acknowledgements of the polysemic nature of artistic artifacts outside of the confinements of structuralist human perception and the dictation of music as a militarised language. Any expertise or form of practice is reliant and built upon by the preceding cultural and knowledge structures at play – AI can be used as a prominent tool of post-structuralist and epistemic investigation, allowing to break through wider cultural indoctrinations, dominant forms of meaning and the structural systems of which any system of knowledge is a part. In *The Liberation of Sound*, Edgard Varèse stated the following:

'Our musical alphabet is poor and illogical. Music, which should pulsate with life, needs new means of expression, and science alone can infuse it with youthful vigor' (Varèse, 1966).

Interactions with the DDSP model invites new considerations of performance behaviors and socio-cultural spaces for new explorations of music and research to emerge. This model presents the potential for new musical expression where previously defined modes are open to new definitions and usage, resulting in new musical interactions. AI systems can be used to challenge the commonalities of musical language and its auditory practices - expanding cognitive and psychological models of creativity, and their relation with existing psychological architectures and epistemological accounts. New musical possibilities, phenomenological approaches, and forms of practice manifest themselves wherever new types of technology are introduced to practitioners, and this model brings forth new ways of musicking.

4. Conclusion

The DDSP model is a quasi-object that provides explorations of otherness and other possible and legitimate musical explorations that fleet human systems of value. This model can be perceived as a tool that can produce interactions that lead to unique heretical ideas, distinctive behaviours and creative trajectories towards modes that aren't necessarily reduced to pre-existing systems of value, perception and socio-cultural and philosophical formulations regarding music - proposing a reframing of phenomenology and approaches to music-making.

This model enables practitioners to explore particular languages, vernaculars and models of syntax - translating input material of a variety of musical and "non-musical" sources into new forms of expression in an environment of trans-idiomatic creativity. Interactions with the DDSP model can open new reflections of artistic practices, performance contexts and music practices for new explorations of music to emerge. My experience with the DDSP model, in particular, has enabled me to form new methods of critical and reflexive listening that are object-oriented and not bound to considerations of what an instrument should sound like and how it should be played. Further, this tool has enabled new forms of expression and *musicking* in my practice.

Lastly, this author is of the opinion that there is a lot of untapped potential for future directions in the field towards approaches that are not based on statutory contexts or structural human notions of value, but on object-oriented forms and a posthumanist and poststructuralist attitude instead.² Research according to these theoretical frameworks can lead to distinct musical artefacts and foster novel and productive interdisciplinary considerations of the use of AI as a quasi-object towards new considerations and scholarly discussions.

² A posthuman and poststructuralist attitude in the sense of engaging and seeking beyond the limits of human perception, cognition, and knowledge imposed by statutory contexts and structuralist frameworks.

5. Future Work

Machine Learning has been explored to create new musical interfaces and instruments since at least the early 1990s. Despite the frequency of literature, development and research in this field, the literature has few examples of artistic research exploring specific models and of research interrogating and exploring matters of encapsulated knowledge, epistemology, object-oriented ontology and non-anthropocentric explorations of human interactions with these tools. However, work exploring Machine Learning through these principles, such as AlphaGo Zero³, reveal distinctive insights.

Most of the work that has explored these matters has come from the models' designers and programmers, whose goal is to achieve "coherent" results, rather than from researchers and artists who are captivated by this tool. Future work will showcase that AI tools such as the one used here as a case study can lead to novel spaces and possible modes of being, leading to novel insights and considerations of this art form., and will investigate how AI tools can be used as harbinger towards new ways of approaching and researching the field outside of human statutory contexts and structuralist paradigms.

This work focused on traditional instruments. Future work will explore training DDSP datasets that are not necessarily representative of a materialist musical object such as the violin. I am currently exploring modes of trans-idiomatic terrains and interstitial spaces of synthesis using this model with particular user datasets and input material to explore this model's sonic behaviours and its potential towards peculiar audio models and distinctive performance behaviours. This work will interpret input material to new findings and new artefacts – exploring non-linear interconnections of distinct forms of strata, investigating matters of contingency and materialism of the audio generation and interstitial spaces of translation of this model. Future work will explore performance contexts where instrumentalists interact with this model in order to explore extended techniques and other possible spaces of possible modes of being. E.g., DDSP models will be trained on a musician's performances, and that musician will interact with the DDSP model as a tool to explore his practice. Improvisatory sessions will be conducted and recorded.

This research will investigate how distinct forms of thought and practices exist and emerge through interaction with this tool and how it leads to the production of new knowledge and new ways of *music*ing. It will investigate the polysemic possibilities of this model and explore ways in which AI can open up new perspectives and lead to the production of particular languages, forms of syntax, synthesis models, idiosyncratic musical artefacts, ontological and epistemological accounts, and the development of the arts.

³ Hassabis, D., Silver, D. Deepmind. (2018). AlphaGo Zero: Starting from scratch. [online] Available at: <https://deepmind.com/blog/article/alphago-zero-starting-scratch>

References

Balaban, M., Kemal Ebcioğlu and Laske, O.E. (1992). Understanding music with AI : perspectives on music cognition. Cambridge, Mass.: Aaai Press.

Bostrom, N. (2019). Transhumanist Values. [online] Available at: <https://www.nickbostrom.com/ethics/values.html>.

de Laubier, Serge. (1988). "The Meta-Instrument" in Computer Music Journal. vol. 22 (1). pp. 25-29.

Engel, Hantrakul, Gu, & Roberts, (2020). DDSP: Differentiable digital signal processing. In International Conference on Learning Representations, 2020.

Fiebrink, R. (2017). Machine learning as meta-instrument: Human machine partnerships shaping expressive instrumental creation. In Musical Instruments in the 21st Century, pages 137–151. Springer.

Hassabis, D., Silver, D. Deepmind. (2018). AlphaGo Zero: Starting from scratch. [online] Available at: <https://deepmind.com/blog/article/alphago-zero-starting-scratch>

Lewis, G. (2018). “Why Do We Want Our Computers to Improvise?” The Oxford Handbook of Algorithmic Music. Oxford University Press: 6.

Magnusson, T. (2009). Epistemic Tools: The Phenomenology of Digital Musical Instruments. PhD thesis, University of Sussex

Perez, Carlos E. (2018). Artificial Intuition : The Improbable Deep Learning Revolution. Editorial: United States Of America, Createspace Independent Publishing Platform, 2018.

Rowe, R. (2004). Machine Musicianship. Cambridge, Mass: London, MIT, 2004.

Schaeffer, P., North, C., & Dack, J. (2013). *In search of a concrete music*. Berkely: University Of California Press.

Small, C. (2011). *Musicking*. Middletown: Wesleyan University Press.

Varese, E. and Wen-chung, C. (1966). The Liberation of Sound. *Perspectives of New Music*, 5(1), p.11.