The Parley Classical Music Composition Generator

SIMON COLTON

1 OVERVIEW

The Parley system generates pieces of classical music in a way which celebrates the composition process. Parley communicates aspects of the music it makes through the production of a score and a video performance of the composition which contains talking points to stimulate engagement and intrigue. It is a neuro-symbolic AI system which combines the power of deep learning models with the control afforded by rule-based approaches. It designs a composition, generates music, then listens to what it has produced so that it can edit and analyse a final version. The long-term aim is for Parley to act as an independent composer which educates and entertains, producing beautiful music and adding to musical culture.

2 MOTIVATION AND ETHOS

In [CB22], Berker Banar and I put forward the idea that, while AI systems can be great tools for human composers, they also have the potential to add to musical culture *themselves*. To do this, we argue, they would need autonomy to invent musical forms and concepts, to innovate with novel instrumentation and accompaniment styles and to challenge assumptions by producing music which defies classification. Moreover, they would need to communicate their process, outputs and ideas via standard musical formats like sheet music, audio files and video performances.

We are building the Parley system according to the ethos that it could, one day, itself be considered a classical music composer. This throws up many technical challenges, but also societal ones. In particular we need to address what it would mean for an AI system to have some level of creative personhood. Alison Pease, Berker Banar and myself have approached this topic in [PCB23], where we discuss notions of agency, self-expression, individuality and responsible behaviours associated with human and AI creativity. Parley will be an experimental system within this context, and we plan to test out theories for how AI systems can be accepted as creative individuals using Parley in an outreach programme.

It is easy to have a knee-jerk fear of AI systems with creative autonomy, as we are constantly fed the myth that they could *replace* human artists, composers and designers. This is furthest from our thoughts with the Parley system. Classical music composition is a deeply human endeavour that we aim to celebrate with Parley, which composes music much more like people do than many other automated approaches. We aim for it to inspire, entertain and educate novice composers, to possibly provide musical ideas for established composers, and to produce beautiful and challenging music for anyone who is interested.

3 A NEURO-SYMBOLIC APPROACH

Parley produces episodic music in given forms such as ABA, whereby dynamics and tempo can change during an episode, and each episode can have a different flavour. It is designed to mimic to some extent some of the processes that human composers undertake during composition. In particular, Parley:

- Follows musical rules expressed symbolically to initially produce some music.
- Listens to the music it has produced and makes edits accordingly, to improve musicality and expression.
- Tries to balance the music over the episodes to address consistency while aiming for novelty.

Fig. 1. Composition before and after editing based on listening models (with edited notes highlighted in orange).

Parley is python software which is ported to a Colab notebook available here: https://tinyurl.com/parleyV2. The code repo is available at: https://github.com/simoncolton/Parley/. It is designed as a modular system in which various music generation, manipulation, analysis and communication routines can be run sequentially, building and communicating a musical composition along the way. The modules include:

- **Designers** which build the basics of a musical composition such as a bar structure and choice of instrumentation.
- **Generators** which add musical content, for instance melody lines, harmonisations and accompaniments.
- **Editors** which take an existing composition and make suitable edits, often after listening to it (see below).
- **Analysers** which describe a composition, for instance in terms of discordance, or what it sounds like.
- **Exporters** which produce the music in various formats including PDF scores, MIDI, MP3 and video files.
- **Communicators** which express details via margin notes on the video performance or on-screen messages.

There are currently 2 designer, 8 generator, 10 editor, 3 analyser, 6 exporter and 1 communicator module in the codebase. In [CBC23], we provide a technical description of some of the modules of the Parley system. Parley builds on top of existing music systems, including: MuseScore for musical score generation; ffmpeg for audio and video manipulation; and fluidsynth to export MIDI files in audio formats.

The most recent development in the project has been the integration of listening models trained by the Music Technology Group (MTG) at the Universitat Pompeu Fabra in Barcelona (www.upf.edu/web/mtg). With good accuracy, these models can classify music into genres and describe music in terms of emotions (e.g., happy), scenarios (e.g., christmas) and mood (e.g., deep). We have found that they can tell whether an edit in, say, a piece of Parley's piano music shifts the music in a particular direction, e.g., to sound more upbeat or melancholic. These small edits add a great deal to the musical quality of the outputs produced by Parley.

As an example, consider the pre- and post-edit passages of music in figure 1. This piece is in C major, and Parley was instructed to edit the bars in this passage to fit with the episode narrative of expressing melancholy. As can be seen in the orange (edited) version, Parley edited the pitches of two notes in each bar, and when listening to the music, it is clear that the objective has been achieved: the music does sound more melancholic than the pre-edited version. To do this, Parley tried out 20 random pitch edits per bar and passed the resulting music through the MTG neural listening models. The edits shown in figure 1 are the ones where the listening model said that the level of melancholy was increased the most. In each case, the edit introduces accidentals to the music which increased the intrigue and musicality of the bar.

4 A COLAB NOTEBOOK INTERFACE

Colab notebooks from Google enable not only open-source sharing of code, but also the running of AI programs enhanced with deep learning models, using GPU compute if required. We have ported Parley to a Colab notebook to demonstrate it for the IRCAM Generative Music Prize. This installation enables users to have a high level of control over the nature of the music produced, but equally, there are options for generating music which don't require any understanding of how it operates. Each parameter can be randomised so that the range of musical styles produced is quite high, and some of the music produced can be quite surprising.

Parley works in stages of generating, editing and analysing music, and produces draft versions of the composition which it shares with the user along the way. In particular, as portrayed in figure 2, for each piece it makes, Parley shares a lead sheet for the composition, a first draft which includes the accompaniment and a voice-leading melody and a second draft which has edits in the bass line and a harmonisation added. These can be listened to while waiting for the the final version to be presented in a video performance, as also portrayed in figure 2. In the video, there are margin notes which provide talking points. In figure 2, we see that the talking points include identifying which bars have been improved via listening model edits. The margin notes also include a description of the bars in terms of how they sound to the listening models. In offline experiments, we have calculated average activations for each of the listening model tags over hundreds of classical piano excerpts, so that Parley can indicate which bars in its compositions have greater (or sometimes much greater) activations than the average. For instance, bar 10 in the score at the bottom of figure 2 has been indicated – with an asterix in the margin notes – as sounding greater than average in terms of the tag "adventure".

5 FUTURE DIRECTIONS

We will soon be releasing Parley as an open source project to which we hope people will contribute modules that increase and diversify Parley's abilities as a classical music composer. As mentioned above, we aim for Parley to add to musical culture by inventing and communicating musical concepts. With the introduction of the listening models, we have moved closer to achieving this objective. That is, Parley now has a sophisticated way of not only telling what passages of music sound like, but also to edit them to emphasise the effect. We plan to introduce the listening models in many different ways, e.g., to choose chord sequences, melody invention, harmonisation, instrumentation and accompaniment schemes. Parley produces interesting music which is non-standard, and we hope to use the listening model approach to identify unusual ways to achieve music which sounds deep or jazzy, christmassy or melancholy.

ACKNOWLEDGEMENTS

REFERENCES

[CB22] S Colton and B Banar. Automatically adding to artistic cultures. In *Proceedings of the 12th International Conference on Artificial Intelligence in Music, Sound, Art and Design.* Springer, 2022. Available at: https://imaginative.ai/papers/colton_evomusart22.pdf.

[CBC23] S Colton, B Banar, and S Cardinale. Neuro-symbolic generation of music with talking points. In *Proceedings of the 14th International Conference on Computational Creativity*, 2023. Available at: https://computationalcreativity.net/iccc23/papers/ICCC-2023_paper_45.pdf.

[PCB23] A Pease, S Colton, and B Banar. On the notion of creative personhood. In *Proceedings of the 14th International Conference on Computational Creativity*, 2023. Available at: https://computationalcreativity.net/iccc23/papers/ICCC-2023_paper_149.pdf.

Fig. 2. Top: screenshot from the Colab notebook. Bottom: still image from a performance video.