Writers and composers do not create prose and music in a purely sequential manner. Composition is an iterative process that includes applying edit operations and shifting attention to and from different parts of the piece. We use a Neural Editor for completion and synthesis of musical phrases by editing prototypes. An advantage of this approach is that it is explainable by design: it maintains the provenance of all edits performed until reaching the synthesized output, which allows us to fully visualize the editing process.

Prototype and edit approach to music completion.

Dataset:
- Collection of 3,428 classical music compositions, 8 composers.
- Tokenized main instrument, monophonic musical phrases.
- Vector representations for discrete musical note tokens.
- Split dataset into 95% training and 5% test sets.

Neural editor (Guu et al, TACL 2018).
- Bi-directional LSTM model
- Edit vectors apply various operations to musical notes:
  - {insert, delete, replace, move left, move right}
- Model trained by maximizing the marginal likelihood.

Human evaluation of Neural Editor synthesized music
- Scatter plot of Likert scale of original music vs. the synthesized
- Each composer is colored and positioned in one of 9 positions.
- Circle size represents number of votes for each combination.
- Marks above diagonal: synthesized as good or better than original.

Human evaluation of musical samples on a 1-5 Likert scale (mean ± standard deviation)

<table>
<thead>
<tr>
<th>Original</th>
<th>Neural Editor</th>
<th>MaskGAN (ICLR 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bach</td>
<td>Beethoven</td>
<td>Brahms</td>
</tr>
<tr>
<td>4.3±0.9</td>
<td>3.5±1.0</td>
<td>3.4±1.0</td>
</tr>
<tr>
<td>Handel</td>
<td>Haydn</td>
<td>Mozart</td>
</tr>
<tr>
<td>4.0±0.9</td>
<td>4.4±0.7</td>
<td>3.9±1.0</td>
</tr>
<tr>
<td>Schubert</td>
<td>Vivaldi</td>
<td></td>
</tr>
<tr>
<td>3.2±1.3</td>
<td>4.2±0.9</td>
<td></td>
</tr>
</tbody>
</table>

Source code: github.com/gjohnsen/musical-phrase-completion