Harmonic Reductions as a Strategy for Creative Data Augmentation

Néstor Nápoles López
McGill University, CIRMMT
nestor.napoleslopez@mail.mcgill.ca

Ichiro Fujinaga
McGill University, CIRMMT
ichiro.fujinaga@mcgill.ca

ABSTRACT

In this paper, we introduce a technique for generating large collections of artificial training examples, which can be used to train chord labeling, key detection, and roman numeral analysis models. The technique consists of using roman numeral analysis annotations of existing datasets to generate *harmonic reductions* of the chords implied by the original annotations. The artificially generated examples ignore the original notes of the annotated example (i.e., the specific "voicings" of the chords), replacing them with voicings suggested by a rule-based voice leading algorithm. A relatively large number of artificial examples can be generated from a single annotated progression using this technique. For example, 10 different voicings in 12 different keys would result in 120 artificial examples generated out of one annotated chord progression.

The voicings suggested for different keys do not necessarily overlap, given that the range of the voices and other variables are taken into account by the rule-based algorithm. This results in data augmentation with potentially unique voicings in each key, contrary to what would be obtained by simply transposing the artificial examples to a different key.

We show the process of applying this technique to a dataset of annotated Bach chorales from the KernScores website. Similar datasets with roman numeral analysis annotations could be used with this approach to generate a large number of artificial training examples for training machine learning models.

1. INTRODUCTION

Roman numeral analysis annotations provide detailed information about the underlying harmonic context of a musical piece. For example, an annotated degree such as `"C:V/V"` can imply a tonal function (dominant), chord label (D major), and key context (a tonicized G major key) at a specific onset of the piece. Thus, these annotations are useful for harmonic analysis and key detection models. However, these annotations are also very expensive to produce, because they involve not only the identification of chords but the identification of ambiguities, non-chord tones [1], ornamentations, and other compositional techniques that obscure the harmonic activity of the music.

Given these complexities, only a handful of expert-curated datasets currently exist [2–6]. Combining all the examples in the existing datasets, a few hundred files are currently available with annotated roman numeral analysis.

When a machine learning algorithm is trained with annotated examples from the literature, like those from existing datasets, the model has to deal with all of the complexities of harmonic analysis at once (ambiguous chords, non-chord tones, ornamentations, etc.), and doing so with a relatively small number of examples. Thousands (or even tens of thousands) of examples would presumably increase the performance of machine learning models dealing with harmonic analysis. Particularly, the ones that involve deep neural networks.

We consider that harmonic reductions could be a useful way of producing many artificial training examples through data augmentation. Furthermore, the generated examples have a reduced harmonic complexity compared to the musical examples from the literature (i.e., the ones provided in existing datasets), which could be an additional advantage when training a machine learning model.

2. HARMONIC REDUCTIONS

A harmonic reduction is a process of removing ambiguities, non-chord tones, ornamentations, and other artifacts out of the original music, maintaining only a musical rendition of the harmonic movement of the piece. These reductions are often voiced or "harmonized" in the form of block chords and four-part arrangements that can be performed on the piano. Nevertheless, when voicing a harmonic reduction, it is common to pay a great deal of attention to maintaining voice-leading considerations in the musical reduction. For example, maintaining the notes within the ranges of a singing voice, avoiding parallel fifths and octaves, and resolving dissonances. Describing the correct methodology for harmonizing a sequence of roman numerals (or figured bass) annotations has been a popular subject and assignment in textbooks and classes concerned with tonal harmony of the common practice period.

We consider that this compromise of (1) maintaining a sensible and stylistic voice-leading practice in the harmonic reduction and (2) removing complexities (e.g., non-chord tones, ornaments, and other artifacts) provide an opportunity for using harmonic reductions as a creative data-

Attribution: Néstor Nápoles López, Ichiro Fujinaga, "Harmonic Reductions as a Strategy for Creative Data Augmentation", *Extended Abstracts for the Late-Breaking Demo Session of the 21st Int. Society for Music Information Retrieval Conf.,* Montréal, Canada, 2020.

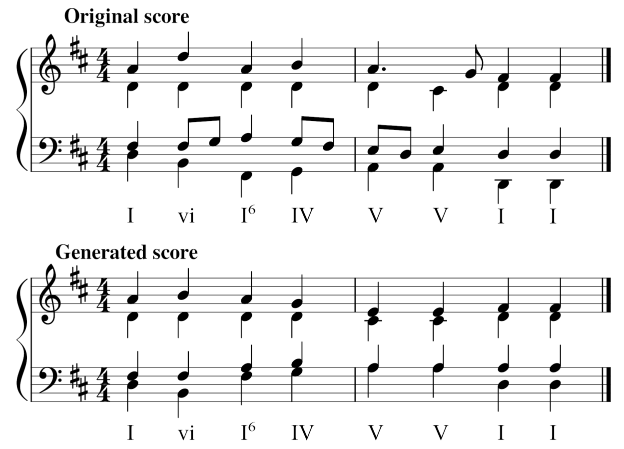

Figure 1. First phrase of J. S. Bach's BWV 262 chorale (above) and an artificially generated harmonic reduction (below). The generated score was derived from the roman numeral annotations in the original score and a rule-based algorithm for voice leading. No information about the original notes was utilized for the generation.

augmentation technique.

Theoretically, a model could learn the vocabulary of chords, and the basic voice-leading "style" of the common practice period from a dataset of harmonic reductions. Techniques such as transfer learning or others could later *specialize* the model in the harmonic style of a specific composer or period to perform accurate chord labeling, key analysis, or roman numeral analysis annotations of unseen data.

The basic voice-leading patterns for four-voice harmonic reductions can be approximated using rules, in a similar way that they have been taught in harmony textbooks for many years. In an initial experiment with this technique, a dynamic programming algorithm generated many four-voice chorale-like renditions of the music's harmony, given an annotated chord progression with roman numeral analysis. The algorithm that we chose was implemented by Eric Zhang [1] and it considered four voice-leading aspects when suggesting an artificial voicing for a chord progression: (1) penalize voice overlaps, (2) penalize large melodic leaps within a single voice, (3) avoid parallel fifths and octaves, (4) resolve the dissonances of dominant seventh chords.

3. PRELIMINARY EXPERIMENT LABELING THE CHORDS OF 63 BACH CHORALES

For the preliminary experiment, we predicted the chord labels of 63 annotated Bach chorales [5].

69 Bach chorales annotated with roman numeral analysis are displayed in the Verovio Humdrum Viewer website. [2] From this set, we utilized 63 of them that have a time signature of `4/4`.

The chorales in the dataset were split in three sets: training (49), validation (7), and test (7).

[1] `https://github.com/ekzhang/harmony`
[2] `https://verovio.humdrum.org/`

Training set	Test accuracy
1. Bach	91.9%
2. Generated	88.3%
3. Bach + Generated	**92.5%**

Table 1. Results of a preliminary experiment training a GRU model to predict the pitch class set (chord labels) of annotated Bach chorales. The same model is trained with three different training sets.

A MLP, LSTM, and GRU models were trained using the original training set (49 annotated chorales) and evaluated in the validation set to find the best model.

The models were trained using 4-measure chunks of music sampled at sixteenth notes. In total, each training example consisted of 64 time steps of piano-roll (input) and pitch-class-set (output) representations. Each time step in the input vector consisted of a many-hot encoding of the sounding notes, similar to the MiniBach model described in [7]. [3] As in the MiniBach model, a special "hold" symbol was utilized to represent held notes. Each time step in the output consisted of a one-hot encoded vector with the pitch-class set implied by the roman numeral annotations. Similarly, a "hold" symbol was used for continuing harmonies. All pitch-class sets considered could be resolved to a common chord label (e.g., $\{0, 4, 7\}$ = Cmaj).

In the initial search, the best results were achieved using the GRU model, which obtained 95% of accuracy in the validation set. The models were evaluated with a categorical accuracy metric. Using the selected model, we tested the effect of our generated harmonic reductions when used as part of the training set. The validation portion was now included as part of the training set.

The harmonic reductions were generated by retrieving all the roman numeral annotations in a chorale of the training set. The original notes were discarded, and 10 voicings for the chord progression were generated using the dynamic programming algorithm.

Three short experiments were conducted on the test set by training the selected GRU model with: (1) the original 56-chorale training-validation set, (2) the 560 generated harmonic reductions (10 voicings of each chorale in the training set), and (3) the two training sets combined ($56 + 560 = 616$ chorales).

The results are shown in Table 1. The training set of harmonic reductions (second experiment) does not achieve a better performance than the original training set, even though there are 10 times more training examples. However, the training set of harmonic reductions contains no information about the original notes and ornaments utilized by Bach. Considering that, the results are reasonable. When combining the original training set with the generated harmonic reductions, the accuracy of the model in the test set increased. This suggests that harmonic reductions could work as a creative data augmentation technique.

[3] Except that voices were not encoded independently but as a single piano-roll representation.

4. REFERENCES

[1] Y. Ju, N. Condit-Schultz, C. Arthur, and I. Fujinaga, "Non-Chord Tone Identification Using Deep Neural Networks," in *Proceedings of the 4th International Workshop on Digital Libraries for Musicology*, ser. DLfM '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 13–16, event-place: Shanghai, China. [Online]. Available: https://doi.org/10.1145/3144749.3144753

[2] M. Neuwirth, D. Harasim, F. C. Moss, and M. Rohrmeier, "The Annotated Beethoven Corpus (ABC): A Dataset of Harmonic Analyses of All Beethoven String Quartets," *Frontiers in Digital Humanities*, vol. 5, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fdigh.2018.00016/full

[3] J. Devaney, C. Arthur, N. Condit-Schultz, and K. Nisula, "Theme And Variation Encodings with Roman Numerals (TAVERN): A New Data Set for Symbolic Music Analysis." in *Proceedings of the 16th International Society for Music Information Retrieval Conference*. Málaga, Spain: ISMIR, Oct. 2015, pp. 728–734. [Online]. Available: https://doi.org/10.5281/zenodo.1417497

[4] M. Gotham, D. Tymoczko, and M. Cuthbert, "The RomanText Format: A Flexible and Standard Method for Representing Roman Numerial Analyses," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*. Delft, The Netherlands: ISMIR, Nov. 2019, pp. 123–129. [Online]. Available: https://doi.org/10.5281/zenodo.3527756

[5] C. S. Sapp, "Computational Chord-Root Identification in Symbolic Musical Data: Rationale, Methods, and Applications." *Computing in Musicology*, vol. 15, 2007.

[6] N. Nápoles López, "Automatic Harmonic Analysis of Classical String Quartets from Symbolic Score," Master's thesis, Universitat Pompeu Fabra, Dec. 2017. [Online]. Available: https://doi.org/10.5281/zenodo.1095617

[7] J.-P. Briot, G. Hadjeres, and F. Pachet, *Deep learning techniques for music generation*. Springer, 2019, vol. 10.