A Transformer-based "Spellchecker" for Detecting Errors in OMR Output

Timothy de Reuse
timothy.dereuse@mcgill.ca

Ichiro Fujinaga
ichiro.fujinaga@mcgill.ca

McGill Schulich School of Music École de musique Schulich
CIRMMT Centre for Interdisciplinary Research in Music Media and Technology
DDMAL Distributed Digital Music Archives & Libraries Lab

To help humans correct inaccurate OMR output, we train a machine-learning model to highlight regions with errors.

INTRODUCTION

- All Optical Music Recognition (OMR) processes **require some human correction**
  - Manual correction is expensive and tedious!
- Errors by OMR processes often are **musically unlikely** relative to the genre. (See Fig. 1)
- Could we catch nearly all OMR errors just by highlighting things that "look wrong?"

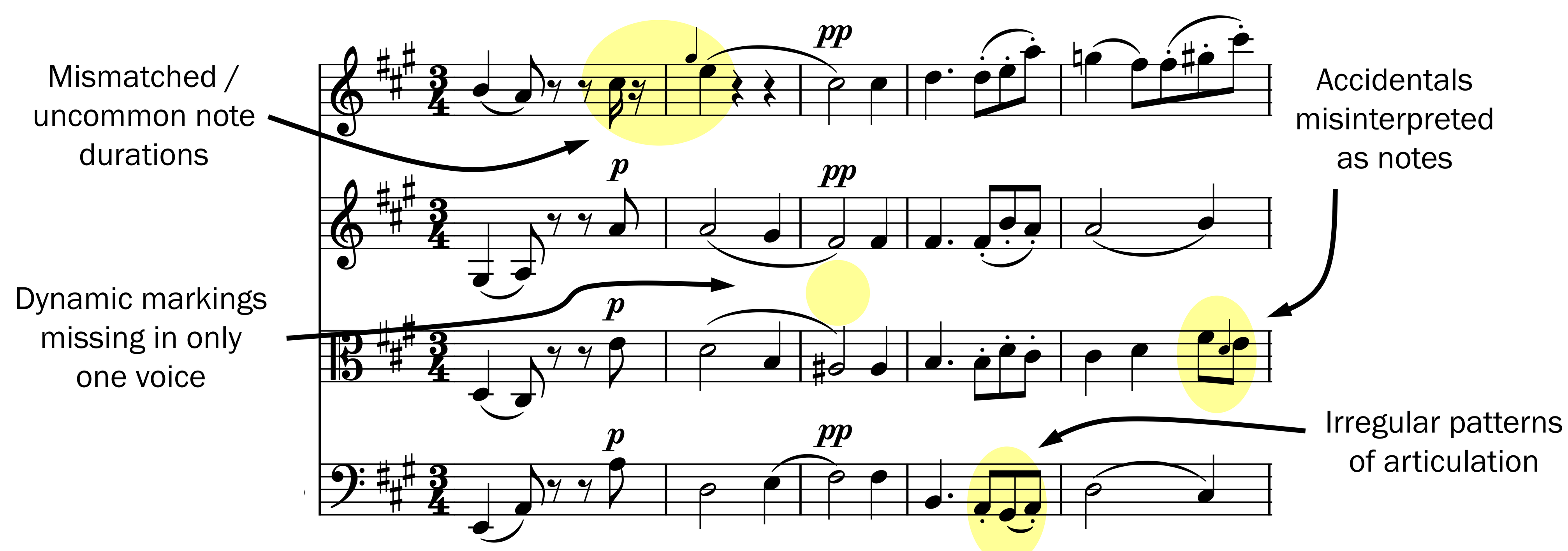

Figure 1: Mendelssohn's *String Quartet in A Major*, Op. 13, Mvt. 1, mm. 4–8, run through Photoscore's OMR.

WHAT IS AN ERROR?

- Take **agnostic representation**[1] of all our input (See Fig. 2c)
- The **Needleman-Wunsch algorithm** performs **sequence alignment**:
  - What operations we would need to do change one sequence into another?
- Our task: *Given a musical score with errors,* ***where must we perform operations to correct it****, according to a Needleman-Wunsch alignment?*
- Binary classification: For each symbol, **would it require an operation, or not?**

DATASETS

- The Mendelssohn String Quartets OMR dataset[2]:
  - 28 total movements run through PhotoScore OMR – not nearly enough
- Use large dataset of string quartets, **augmented with OMR-like** errors
  - All quartets of: Beethoven, Haydn, Mozart, Schubert, plus assorted others
- Using Long Short-Term Universal Transformer[3] (LSTUT) with ~8M parameters

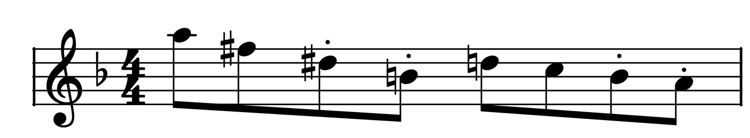

Figure 2a: Mendelssohn's Op. 13, Mvt. 4, mm. 19, 1st violin part.

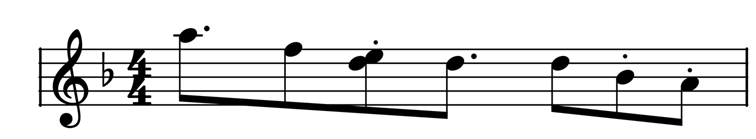

Figure 2b: Fig. 2a, run through PhotoScore's OMR.

Aligned Original	Aligned OMR Output	Operation	Error?
8th.pos11.startBeam	8th.pos11.startBeam	-	-
accid.pos9.sharp	dot.pos12	REPLACE	X
8th.pos9.continueBeam	8th.pos9.continueBeam	-	-
accid.pos7.sharp	_	INSERT	X
8th.pos7.continueBeam	8th.pos7.continueBeam	-	-
_	^	DELETE	X
_	8th.pos8.continueBeam	DELETE	X
^	^	-	-
dot.pos10	dot.pos10	-	-
accid.pos5.natural	8th.pos7.endBeam	REPLACE	X
8th.pos5.endBeam	_	INSERT	X
^	_	INSERT	X
dot.pos8	dot.pos8	-	-
accid.pos7.natural	_	INSERT	X
8th.pos7.startBeam	8th.pos7.startBeam	-	-
8th.pos6.continueBeam	_	INSERT	X
8th.pos5.continueBeam	8th.pos5.continueBeam	-	-
^	^	-	-
dot.pos8	dot.pos8	-	-
8th.pos4.endBeam	8th.pos4.endBeam	-	-
^	^	-	-
dot.pos6	dot.pos6	-	-

Figure 2c: The Needleman-Wunsch alignment between (2a) and (2b). We consider an error to be present in the OMR output anywhere that the alignment prescribes some operation.

RESULTS

- Best performance: Recall of 99%, Precision of 51%
  - We can exclude **half the score** from correction and only **miss 1 in 100 errors**
- Model often identifies **too large a region** as erroneous
  - Identifies **general location** of error – is that good enough for humans?
- Failure (little better than chance) on partially-corrected OMR
  - Cannot find the types of errors that humans miss
  - Difficult to find deleted elements

TWO EXAMPLES (FIGURES 3 & 4)

OMR errors are marked in red, while regions predicted to contain errors by our model are marked in gray boxes. Note: A correct symbol may be erroneous according to our alignment method if, to correct the score, one needs to insert a symbol after it.

Figure 3a: Mendelssohn's *String Quartet in E♭ Major*, Op. 12, Mvt 1, mm. 212–217, 1st violin part.

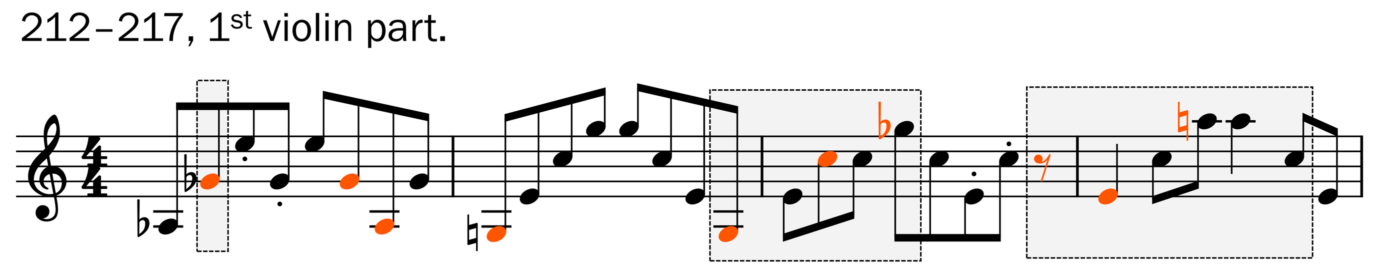

Figure 3b: Fig. 3a, run through PhotoScore OMR, with errors marked.

Figure 4a: Mendelssohn's *Fugue*, Op. 81/4, mm 12–14.

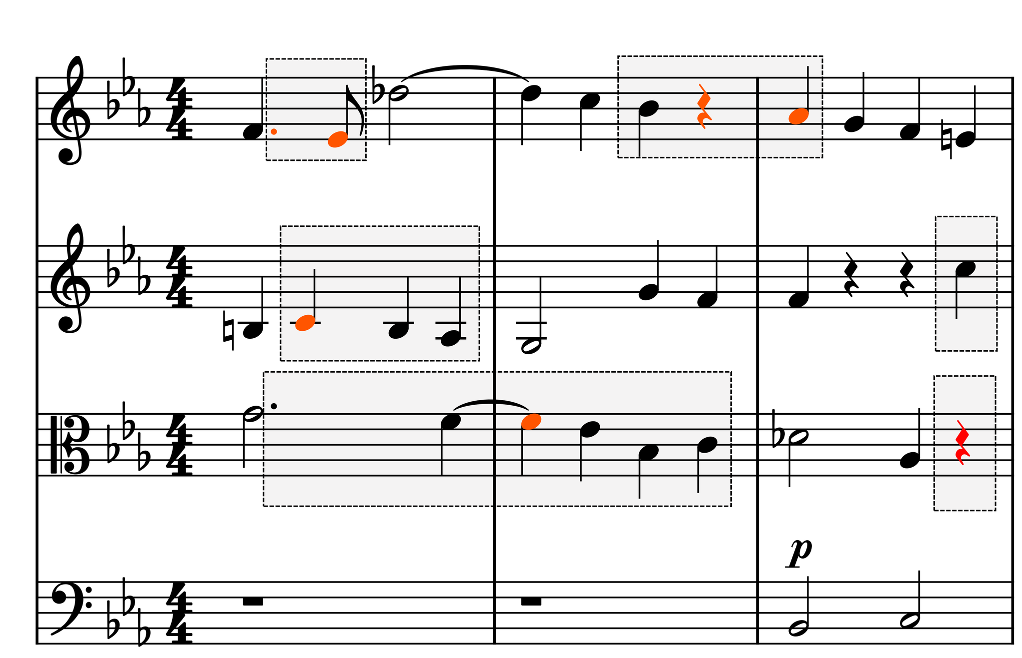

Figure 4b: Fig. 4a, run through PhotoScore OMR, with errors marked.

[1] Calvo-Zaragoza, Jorge, and David Rizo. 2018. "End-to-End Neural Optical Music Recognition of Monophonic Scores." *Applied Sciences* 8 (4): 606.

[2] Jacob deGroot-Maggetti, Timothy de Reuse, Laurent Feisthauer, Samuel Howes, Yaolong Ju, Suzuka Kokubu, Sylvain Margot, Néstor Nápoles López, Finn Upham. 2020. "Data Quality Matters: Iterative Corrections on a Corpus of Mendelssohn String Quartets and Implications for MIR Analysis", in *Proc. of the 21st Int. Society for Music Information Retrieval Conf.,* Montréal, Canada.

[3] de Berardinis, Jacopo, Samuel Barrett, Angelo Cangelosi, and Eduardo Coutinho. 2020. "Modelling Long- and Short-Term Structure in Symbolic Music with Attention and Recurrence." In *Proc. of the 1st Joint Conf. on AI Music Creativity*. Stockholm, Sweden.

This research is supported in part by funding from the Social Sciences and Humanities Research Council of Canada, the Fonds de Recherche du Québec – Société et Culture, and compute resources from Calcul Québec and the Digital Research Alliance of Canada.

23rd International Society for Music Information Retrieval Conference
Bengaluru, India, December 4–8, 2022

Paper and code available at:
github.com/timothydereuse/transformer-omr-spellchecker