The Art of Teaching Computers: The SIMSSA Optical Music Recognition Workflow

ICHIRO FUJINAGA[1]

Abstract: In many machine learning systems it would be effective to create a pedagogical environment where both the machines and the humans can incrementally learn to solve problems through interaction and adaptation. We are designing an optical music recognition workflow system where human operators can intervene to correct and teach the system at certain stages so that both parties can learn from the errors and, consequently, the overall performance is increased progressively as more music scores are processed.

Keywords: Optical music recognition, machine learning, interactive learning

1. Introduction

As part of the SIMSSA (Single Interface for Music Score Searching and Analysis) project [1], we are developing a series of browser-based interfaces for the different stages of our OMR (optical music recognition) workflow: image preprocessing, music symbol recognition, musical notation recognition, and final representation construction (see Figure 1). In the first two stages we integrate human input with the aim of teaching the computers to improve the performance.

We are proposing the idea of a pedagogy for learning machines as the study of the methods and activities of teaching machines, and the creation of an environment where humans can learn the art of how to teach machines running learning algorithms. In order to achieve this, we first need to understand how humans interact with a machine-learning component and then build a clever workflow to take advantage of the intelligence of the human and the ability to perform fast calculations of the computer.

Bieger et al. [2] proposed a conceptual framework for teaching intelligent systems. In this framework, the interaction between a human teacher and a learner (i.e., a computer system) has the goal of teaching the learners to gain knowledge about a specific task. We hypothesize that by knowing the learner, and how the learner reacts to correction and new input, teachers can adapt their pedagogical strategies to improve the learner.

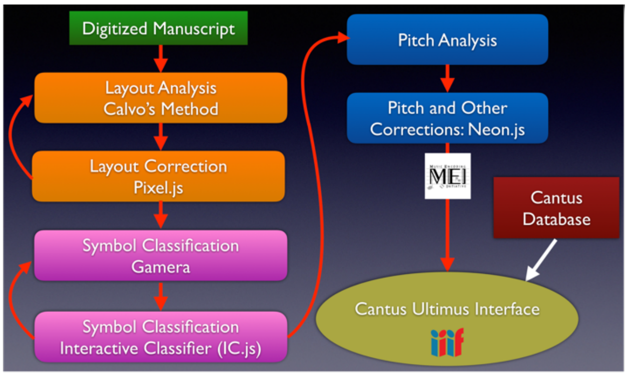

Figure 1: OMR workflow for neume notation.

2. Teaching machines to read scores

Our ultimate goal is to read and extract the content from digitized images of music documents. For scalability and because of the large varieties of Western music notation systems found, especially in the period before the 18th century, we deploy supervised machine learning strategies. Rather than creating a large set of ground-truth data for training, we create interactive environments where humans will incrementally teach the computers to learn to read music scores.

As the computer learn we are interested in learning about the type of errors that the computer makes, in order to choose better training data to improve the performance of the computer and allow teachers (humans) of the system to understand and gain knowledge about where learners (machine learning models) make mistakes in order to modify the behavior of the learners.

References

[1] Fujinaga, I., Hankinson, A. and Cumming, J. 2014. Introduction to SIMSSA (Single Interface for Music Score Searching and Analysis). In *Proceedings of the International Workshop on Digital Libraries for Musicology*, 100–102. London.

[2] Bieger, I, Thórisson, K. R., and Steunebrink, B. R. 2017. The pedagogical pentagon: A conceptual framework for artificial pedagogy. In *International Conference on Artificial General Intelligence (Lecture Notes in Computer Science, vol. 10414)*, T. Everitt, B. Goertzel, and A. Potapov (Eds.). Springer, Cham, 212–222.

Acknowledgments

This research has been supported by the Social Sciences and Humanities Research Council of Canada, ComputeCanada and the Fonds de Recherche du Québec - Société et Culture.

1 Schulich School of Music, McGill University