SALAMI: Structural Analysis of Large Amounts of Music Information

David De Roure
Oxford e-Research Centre
University of Oxford
and School of ECS
University of Southampton
dder@oerc.ox.ac.uk

J. Stephen Downie
Graduate School of Library and
Information Science
University of Illinois at
Urbana-Champaign
jdownie@illinois.edu

Ichiro Fujinaga
Schulich School of Music
McGill University
Montreal
ich@music.mcgill.ca

SALAMI (Structural Analysis of Large Amounts of Music Information) applies computational approaches to the huge and growing volume of digital recorded music that is now available in large-scale resources such as the Internet Archive. It is set to produce a new and very substantive web-accessible corpus of musical analyses in a common framework for use by music scholars, students and beyond, and to establish a methodology and tooling which will enable others to add to the resource in the future. The SALAMI infrastructure brings together workflow and Semantic Web technologies with a set of algorithms and tools for extracting features from recorded music which have been developed by the music information retrieval and computational musicology communities over the last decade, and the project uses "controlled crowd sourcing" to provide ground truth annotations of musical works.

1. The Challenge

Structural analysis of music (formal analysis) is one of the most fundamental analyses performed by music scholars including music theorists, musicologists and ethnomusicologists. The formal analysis usually precedes any other types of analysis because it provides the overall view of the piece. The main goal of formal analysis is to find similar sections within a piece of music and label these sections, such as ABA and ABCB'A. With further analysis, these sections can be marked with predefined labels such as Intro, Verse, Bridge, Chorus, Verse, and Outro (popular music) or Introduction, Exposition, Development, Recapitulation, and Coda (sonata form). This is illustrated in Figure 1.

This formal analysis is potentially useful in classifying different genres of music and it can be used to compare different styles of composition within a composer's works or between composers. It can also be used to understand historical influences over time and location. By analysing large sets of music, new discoveries can be made about these questions. Another important aspect of formal analysis is that it can be applied to almost any music, anything from Russian folk songs to Byzantine music to Miles Davis or to electronic music. Furthermore, there are many musics in the world where other types of analysis, such as harmonic, motivic, or rhythmic, have little meaning.

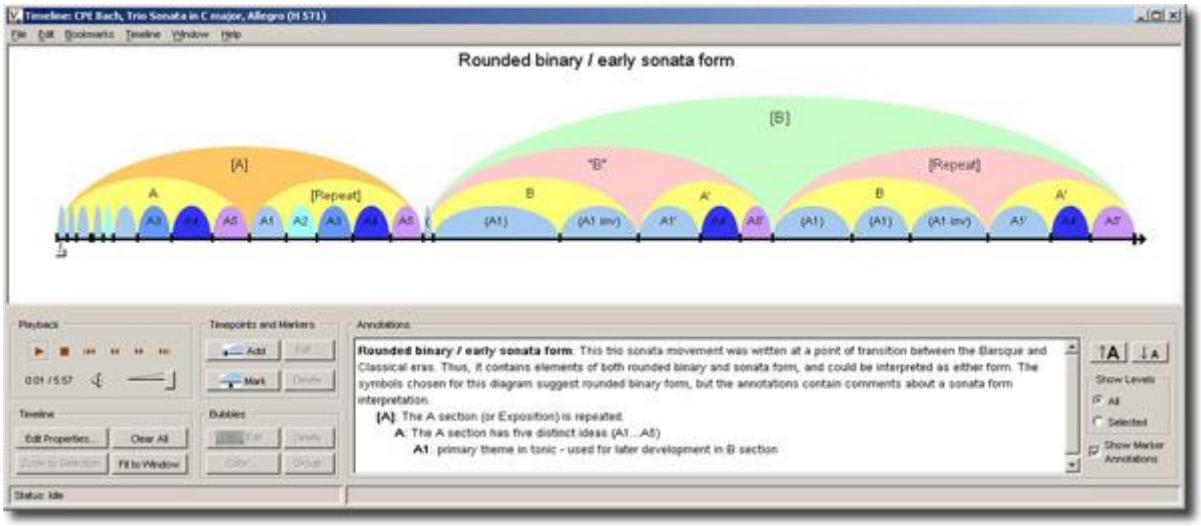

Figure 1. Sonata form, illustrated in Audio Timeliner, an audio annotation and analysis tool from Indiana University for creating and labelling bubble diagrams. These diagrams can be used to navigate music or other audio for detailed study.

Traditionally, structural analysis of music has been done manually, with very few exceptions. This is a time-consuming task and only a small sample of music has been analysed. Moreover, since there are no standard formats for describing the structure, let alone a standard machine-readable file format, there is no simple way to compare large amount of music based upon its various internal structures. Even within the field of Music Information Retrieval (MIR) research, with its evolving computational approach to music analysis, the majority of structural analysis work has been performed on a few hundred pieces, at best.

2. An e-Research approach

There is now a tremendous volume of digital audio recordings available commercially and in private collections and online, covering many types of music. Most prior analytic research work has focused primarily on Western popular and "classical" music. In the Salami project we have secured permission to use important sources of music audio so that the range of different kinds of music that will be analysed will be of far larger variety than anything previously done – our dataset includes a wide variety of music from all over the world, from many time periods, and includes folk music, "classical" music, contemporary music, improvised music, and live music. The Internet Archive collection is by far our largest resource, with ~18,000 hours of audio including a substantial collection of live concert recordings (~66,000 pieces) which represent a novel source for structural analysis. The resources are shown in Table 1.

SOURCE COLLECTION	SOURCE ACCESS TYPE	TRACKS	HOURS	SIZE (TB) UNCOMPRESSED
Internet Archive	Open via Internet	276,000	18,333	11.00
DRAM	Subscription via Internet	8,000	2,300	1.38
IMIRSEL	Closed (Stored locally)	34,000	2,267	1.36
McGill	Closed (Stored locally)	32,000	2,133	1.28
Totals		350,000	23,267	15.02

Table 1. The SALAMI Audio Data Collections

To analyse this we have a set of algorithms and tools for extracting features from recorded music which have now achieved significant effectiveness, developed by the music information retrieval and computational musicology (CM) communities over the last decade and evaluated through a series of annual international events called the Music Information Retrieval Evaluation eXchange (MIREX) [1]. The ability to analyse music directly in the audio format is an important development: in the past, most music structural analyses have been conducted using only the musical scores that were readily available, especially with European "classical" music. The new audio-based structural information offers novel perspectives to music research especially for ethnomusicologists where no scores exist for many of the music cultures. The technical expertise needed to analyse music in audio format has prevented most music researchers from dealing with the actual performance of the music. With the recent revolution in MIR and CM research, many new tools and algorithms to analyse and to visualise music audio have been developed. The algorithms for structural analysis, such as autocorrelation over hybrid feature extraction, are particularly computationally intensive.

This scale of analysis (350,000 pieces or 23,000 hours) brings both opportunities and challenges. The algorithms chosen, modified and/or developed for use by SALAMI are being trained and evaluated using a set of ground-truth data based upon several thousand exemplars created by trained musicologists. To achieve this, the project is currently conducting an exercise in "controlled crowd-sourcing" whereby suitably qualified people can sign up to participate in the annotation task. This exercise is itself creating a unique resource for the community. Salami is developing a standardised ontology for music structure to facilitate use of analyses, and will be publishing the analysis data using a Linked Open Data approach. This is based on the foundational work of Raimond [2].

To provide the computational infrastructure for this scale of analysis, Salami is using the Meandre dataflow engine together with a very significant grant of 250,000 hours of supercomputing time at the National Center for Supercomputing Applications (NCSA). Meandre is an open-source dataflow execution framework designed to simplify the running of large-scale data mining/analysis applications on high-performance computing clusters, and it stores the operational data of each session run in RDF making it easier to acquire and integrate the provenance data. This framework is known as NEMA (Networked Environment for Music Analysis – see http://nema.lis.uiuc.edu/) and is an open and extensible webservice-based resource framework that facilitates the integration of music data and analytic/evaluative tools, including the myExperiment social web site [3] which has been extended to support the sharing of Meandre workflows.

3. Conclusion

Salami brings contemporary e-Research thinking to a significant musicological research challenge, bringing together computationally intensive music analysis algorithms with techniques for creating and sharing annotations and an infrastructure that is both technical (web services and workflow) and social (the ground truth collection). The initial project runs till June 2011 and will deliver new resources as well as putting in place the tools and methodology for the analysis activity to be sustained and enhanced by the community. These resources will add to a growing body of Linked Data in the music domain. In terms of the e-Research agenda, Salami will lead to new tools and method development, and is also enabling us to evaluate Linked Data as a platform for research enquiry.

Acknowledgements

SALAMI is a Digging into Data project (see http://www.diggingintodata.org/) and is funded by JISC, NSF and SSHRC. NEMA is funded by the Andrew W. Mellon Foundation. We are also grateful to NCSA for the donation of supercomputer resource, and to our partners including the Internet Archive for provision of content.

References

[1] Downie, J. S. (2008). The Music Information Retrieval Evaluation Exchange (2005-2007): A window into music information retrieval research. *Acoustical Science and Technology, 29*(4): 247–255.

[2] Raimond, Y., Abdallah, S., Sandler, M. and Giasson, F. (2007). The Music Ontology. Proceedings of the 8th International Conference on Music Information Retrieval. Wien, Österreich, September 2007, pp. 417-422.

[3] De Roure, D., Goble, C. and Stevens, R. (2009) The Design and Realisation of the myExperiment Virtual Research Environment for Social Sharing of Workflows. Future Generation Computer Systems 25, pp. 561-567.