**Program Schedule**
**September 5, 2015**
Psychology Research and Training Centre, 2nd Floor, Room # SBB 238
105 Bond Street, Ryerson University

**Schedule:**

| | |
|---|---|
| 08:30 am-09:00 am | Registration |
| 09:00 am-09.10 am | Welcome (Naresh/Frank) |
| 09:10 am-10:05 am | Keynote - **Morwaread Farbood:** *Can music cognition help us build better computer-music software?* |
| 10:10 am-11:10 am | **Sebastian Stober, Avital Sternin, Adrian M. Owen, Jessica A. Grahn:** *Similarity and feature learning for EEG recordings of music perception and imagination*<br>**Naresh Vempala & Frank Russo:** *Understanding the role of features in emotion recognition using the MAP dataset*<br>**Dakota Killpack:** *A correlated topic model of structural schemata in popular music* |
| 11:10 am-11:20 am | Break |
| 11:20 am-12:00 pm | **Speed poster talks** (see page 2) |
| 12:00 pm-01:30 pm | Lunch/Poster session |
| 01:30 pm-02:25 pm | Keynote - **Erik Schmidt:** *Music discovery in the long tail at Pandora* |
| 02:25 pm-03:05 pm | Invited Talk - **Wang Ye:** *Developing MIR systems for learning and wellbeing* |
| 03:05 pm-03:25 pm | **Michael Barone, Robert Rawlins, Tasmine VandenBerg, Matthew Woolhouse:** *Digital music consumption with respect to socio-cultural development* |
| 03:25 pm-03:40 pm | Break |
| 03:40 pm-04:40 pm | **Jotthi Bansal, James Renwick, Matthew Woolhouse:** *Expression of acoustic features across multiple genres within individual download collections*<br>**Frank Russo & Naresh Vempala:** *A new cognitive measure of melodic similarity*<br>**Michael Barone, Kurt DaCosta, James Renwick, Matthew Woolhouse:** *Grail: A scalable music metadata collection interface* |
| 04:40 pm-05:00 pm | Announcement of Student Prizes/Closing Remarks |

**Speed poster talks:**

1. **Ian Macchiusi**
   "Working in Audio is More Fun": The Digital Audio Workstation, Rendering and the Language of Manipulation
2. **Catherine Renwick, James Renwick, Matthew Woolhouse**
   Genre Transition Probabilities Using Markov Chain Matrices
3. **Kurt DaCosta, Michael Barone, Matthew Woolhouse**
   Music Recommendation Using Multi-Attribute Categorical Production Metadata
4. **Paolo Ammirante, Frank Russo**
   Low-skip bias: The distribution of skips across the pitch ranges of vocal and instrumental melodies is vocally constrained
5. **Yihao Fang, Michael Barone, Matthew Woolhouse**
   Sentiment Analysis of Mandarin Pop Lyrics Using Multi-Emotion Profiles
6. **James Renwick, Catherine Renwick, Matthew Woolhouse**
   What's My Age Again? Date-of-Birth Estimation Using Musical Memories

**Keynote Addresses:**

**Morwaread Farbood, New York University**

Can music cognition help us build better computer-music software?

In this talk, I discuss several research projects that combine aspects of music cognition and computer music. These projects are discussed in the context of my personal experiences with the methodological divide between engineering/computer science and cognitive psychology research. While seamlessly merging these approaches is not necessarily possible, I offer some perspective on how methods from the two disciplines can be employed in a complementary way.

**Erik Schmidt, Pandora**

Music discovery in the long tail at Pandora

Finding the music of the moment can often be a challenging problem even for humans with well-versed musical tastes. These challenges further explode into myriad complexities when attempting to construct algorithmic approaches for automatic playlist generation. A variety of factors play a role in influencing a listener's perception of what music is appropriate on a given seed (e.g., musicological, social, geographical, generational), and these factors vary across different contexts and listeners. Furthermore, as opposed to more traditional recommender systems which need only to recommend a single item or set of items, Pandora's recommenders must provide an evolving set of sequential items which constantly keep the experience new and exciting. The talk will present an overview of recommendation at Pandora, followed by a deep dive into the challenges of recommending content in the long tail. A mixture of experts based system will be discussed as well as the many data resources required to bring small stations to life for both new and indie artists. Both content and context will be discussed including the development of powerful machine listening systems which leverage the vast data resources of the music genome project.

**Invited Talk:**

**Wang Ye, National University of Singapore**

Developing MIR systems for learning and wellbeing

The use of music as an aid in healing body and mind has received enormous attention over the last 20 years from a wide range of disciplines, including neuroscience, physical therapy, exercise science, and psychological medicine. We have attempted to transform insights gained from the scientific study of music and medicine into real-life applications that can be delivered widely, effectively, and accurately. We have been trying to use music in evidence-based and/or preventative medicine. In this talk, I will describe a few clinically-focused tools to facilitate the delivery of established music-enhanced therapies, harnessing the synergy of sound and music computing (SMC), mobile computing, and cloud computing technologies to promote healthy lifestyles and to facilitate disease prevention, diagnosis, and treatment in both developed countries and resource-poor developing countries. I will present some of our past and ongoing research projects that combine wearable sensors, smartphone apps, and a cloud-based therapy delivery system to facilitate music-enhanced physical and speech intervention, as well as the joys and pains working in such a multidisciplinary environment.

**Abstracts:**

***Low-skip bias: The distribution of skips across the pitch ranges of vocal and instrumental melodies is vocally constrained***

**Paolo Ammirante, Frank Russo***

***SMART lab, Ryerson University**

Skips are relatively infrequent in diatonic melodies and are compositionally treated in systematic ways. This treatment has been attributed to deliberate compositional strategies that are also subject to certain constraints. Study 1 showed that ease of vocal production may be accommodated compositionally. Number of skips and their distribution within a melody's pitch range were compared between diverse statistical samples of vocal and instrumental melodies. Skips occurred less frequently in vocal melodies. Skips occurred more frequently in melodies' lower and upper ranges, but there were more low skips than high ("low-skip bias"), especially in vocal melodies. Study 2 replicated these findings in the vocal and instrumental melodies of a single composition (Bach's Mass in B minor). Study 3 showed that among the instrumental melodies of classical composers, low-skip bias was correlated with the proportion of vocal music within composers' total output. We propose that, to varying degrees, composers apply a vocal template to instrumental melodies. Implications of this research for receptive aspects of music are discussed.

***Expression of acoustic features across multiple genres within individual download collections***

**Jotthi Bansal, James Renwick, Matthew Woolhouse***

***McMaster University**

This research examines the extent to which acoustic features pertaining to a user's preferred genre are expressed in songs they consume in other genres. For example, do people who prefer fast-tempo dance music prefer fast-paced pop or rock music? The data used in the study are from a five-year data-sharing agreement between McMaster University and MixRadio, and consist of over 1.3 billion music downloads from countries across the globe. Previous study of user-download histories revealed that some genre based subgroups are more heterogeneous than others with respect to genre diversity. Methodologically, the current research calculates the median and mean

values for each acoustic feature and compares these with the corresponding values within genre-based subgroups. Results are pending, however initial exploration suggests that certain acoustic features of a user's preferred genre are indeed expressed in songs they consume in other genres. The cognitive implications of this for music processing are explored.

***Grail: A scalable music metadata collection interface***

**Michael Barone, Kurt DaCosta, James Renwick, Matthew Woolhouse***
***McMaster University**

We introduce Grail, an interface in development that links the identity spaces of open access metadata from diverse online music services. Companies such as Last.fm, MusicBrainz, and Discogs provide valuable access to their comprehensive corpuses of music metadata. As the music industry continues to transition to online models of delivery, such as Pandora and MixRadio, development of sophisticated decision-making applications increasingly relies upon accurately linking identity-space data. Currently, these massive libraries are somewhat disjoint, and linking these databases presents unique challenges. Grail eases data collection and enhances discovery for multidisciplinary music informatics. We describe the architecture of Grail, which is scalable to petabytes of information. Grail is being used to develop research applications that are being tested on a database of 1.36 billion downloads provided by MixRadio. We explore the difficulties in collating digital-music information from multiple sources as well as the scalability of future music information retrieval.

***Digital music consumption with respect to socio-cultural development***

**Michael Barone, Robert Rawlins, Tasmine VandenBerg, Matthew Woolhouse***
***McMaster University**

Previous research has linked composite measures of human development, such as those contained within the United Nations' Human Development Report (HDR), and global-music consumption patterns. With an appreciation for the explanatory limitations of correlation, we conduct a fine-grained analysis that attempts to discover which components of human development are closely related to music consumption and those which are unrelated. Our methodology is (1) to identify salient patterns within MixRadio's music-consumption database consisting of over 1.36 billion music downloads, such as genre dispersion and audio summary information derived from The Echo Nest, and (2) to calculate the degree of association between these features and the components of the indices in question. The indices used were from the United Nations Human Development Report and World Values Survey. The former consists of factors such as life expectancy, access to education, suicide rates and economic measures; the latter consists of religious affiliation, social and political attitudes, and information relating to post-materialism. Our findings provide insights into the degree to which music is socially embedded and a vehicle of cultural expression.

***Music recommendation using multi-attribute categorical production metadata***

**Kurt DaCosta, Michael Barone, Matthew Woolhouse***

***McMaster University**

Recommender systems have been examined quantitatively using a variety of methods including opinion mining, audio feature extraction, and social network analysis. We consider a novel addition to recommender systems using collaborative filtering. A feature space is created based on user preference for production attributes. Production metadata includes roles such as sound engineers, producers, lyricists, and session musicians that are vital to a recording, but are frequently overlooked in its creation. We combine this potentially neglected data with play-history information, including track recency, to create a production-preference profile (PPP) per person. These PPP's are used to generate similarity weightings with respect to a set of recommended seed tracks. Evaluation of this approach is currently being undertaken.

***Sentiment analysis of Mandarin pop lyrics using multi-emotion profiles***

**Yihao Fang, Michael Barone, Matthew Woolhouse***

***McMaster University**

Sentiment analysis is a subfield of natural language processing (NLP) which attempts to identify the emotional categories of text. While these techniques have produced impressive findings using machine learning, many issues remain. In particular, automated-sentiment analysis classifies emotional content dichotomously, restricting classification to only one emotion. This categorical approach fails to capture the affective nuances of subjective artistic expression in text. We examine whether general linguistic patterns exist in Mandarin Pop lyrics using multi-emotion profiles coupled with machine learning and NLP. Using crowd sourced survey data, we gather multiple emotion profiles for Mandarin Pop song lyrics. A multi-class support vector machine, constructed from these profiles and n-gram NLP models, is applied to the song lyrics. Kfold cross validation is then applied to analyze the efficacy of classification in the training data. We report the accuracy of this model for sentiment analysis and discuss implications of future directions based on our findings.

***A correlated topic model of structural schemata in popular music***

**Dakota Killpack***

***Princeton University**

Since their popularization in 2003 by David Blei, topic modeling techniques have enjoyed widespread use on account of their applicability to finding human-intelligible "topics" in unstructured text and image data. The fashionableness of topic models extends to music: recent research has demonstrated the successful extraction of harmonic "topics" from raw audio features, managing to capture perceptually salient aspects of chord vocabulary and scalar modal quality in songs. Although significant scholarship in music theory and cognition has been directed towards finding structurally significant recurrent patterns in rhythmic units greater than the measure, paying particular attention to differences in cognitive constraints on the perception of rhythm at varying timescales, little work has taken place towards the application of topic models to musical features beyond single notes and chords. This study seeks to bridge the disparate worlds of music theory and music informatics by using topic modeling techniques to advance a cognitively based approach to the unsupervised extraction of human-intelligible structural features from symbolic music data (in particular, from the McGill SALAMI and Billboard

projects). To this end, I examine correlations between topical content and sectional segmentation in songs, identifying possible harmonic and rhythmic mental schemata for each segment of a song.

***"Working in audio is more fun": The digital audio workstation, rendering and the language of manipulation***

**Ian Macchiusi***
***York University**

The computer's visual representation of sound has revolutionized the creation of sample-based music through the interface of music software, or—as it is commonly called—the Digital Audio Workstation (DAW). With the rise of DAW-based composition in popular music styles, many artists' sole experience of musical creation is through the computer screen. I assert, the particular sonic visualizations of the DAW, propagate certain assumptions about music, influencing aesthetics and adding new visually-based parameters to musical composition. One example is the operation of rendering—a process where one representation of sound is transformed or converted to another. Many artists utilize rendering to create a more complex visual representation of their sounds. This increased complexity and its corresponding feedback when manipulated, is often linked to feelings of control, mastery and pleasure in the artist, which in turn, encourages more manipulation. An analysis of the influence of rendering will be furnished through the depiction of the user interface concept of direct manipulation, as well as an examination of online communities such as Reddit and futureproducers.com. Specifically, I believe this vast collection of internet materials provides a fascinating glimpse into the computer musician's processes, habits and biases that are engendered by the visualizations of the DAW.

***Genre transition probabilities using Markov chain matrices***

**Catherine Renwick, James Renwick, Matthew Woolhouse***
***McMaster University**

We use statistical analysis of users' digital-music download patterns to determine song, artist and genre transition probabilities. The investigation uses MixRadio's global music-consumption database consisting of 1.36 billion downloads, which contains detailed temporal information. We classify the transitions between pairs of songs by genre, and derive probabilities representing average-user download trends by country. Probability values are then deployed in a Markov-chain transition matrix to demonstrate the temporal relationships between genres. Perhaps unsurprisingly, we find that users have a strong likelihood of downloading songs of the same genre in succession, ranging from 60-90%. Results also show that in some instances, the number of times a user transitions from one genre to the next can be 10-20 times greater than in the reverse direction. The asymmetry between transition probabilities could be explained by a number of factors, including disparity of genre representation within the database (although normalization procedures discount this to some degree), language barriers, or, more interestingly, that certain genres point users in particular musical directions.

***What's my age again? Date-of-birth estimation using musical memories***

**James Renwick, Catherine Renwick, Matthew Woolhouse***
***McMaster University**

We compare and contrast survey data on musical memories with patterns in digital streaming in an attempt to determine the age of users based on their consumption habits. User data was gathered from two different sources:

a crowdsourcing survey and Last.fm's API. The survey included data from ca 5,000 participants who recalled songs and the age they first heard them. The listening histories of ca 4,000 Last.fm users were recorded. Users from both datasets showed a strong tendency to listen to or recall songs released when they were between 15 and 20 years of age. We link these findings to psychological research showing a tendency for adults to have an increased recollection for events that happened during adolescence. Using a statistical regression model, an algorithm is created to estimate user date of birth based on musical preferences and this psychological effect. Lastly, the algorithm is tested to ascertain the veracity of this method for predicting user age in cases where this information is absent. The difference in accuracy is due to the rule-based system's tendency to over-assign a dominant function.

***A new cognitive measure of melodic similarity***

**Frank Russo, Naresh Vempala***

***SMART lab, Ryerson University**

We present an empirically derived measure, based on psychological and cognitive predictors, for determining similarity between two melodies with multiple-note changes. The derivation of our model involved three stages. In Stage 1, eight standard melodies were systematically varied with respect to *pitch distance, pitch direction, tonal stability, metric salience,* and *melodic contour*. Comparison melodies with a one-note change were presented in transposed and nontransposed conditions. For the nontransposed condition, predictors of explained variance in similarity ratings were pitch distance, pitch direction and melodic contour. For the transposed condition, predictors were tonal stability and melodic contour. In Stage 2, we added the effects of primacy and recency. In Stage 3, we introduced comparison melodies with two-note changes, which allowed us to derive a more generalizable model accommodating multiple-note changes. In a follow-up experiment, we show that our measure of melodic similarity yielded superior performance to the Mongeau and Sankoff similarity measure. A cognitive measure, such as the one described here, has the potential to extend the domain of similarity-finding methods in music information retrieval.

***Similarity and feature learning for EEG recordings of music perception and imagination***

**Sebastian Stober, Avital Sternin, Adrian M. Owen, Jessica A. Grahn***

*** Brain and Mind Institute,**

**University of Western Ontario**

Aiming to learn spatio-temporal features from the OpenMIIR dataset of EEG recordings taken during music perception and imagination, we propose two deep learning techniques based on convolutional autoencoding. Conventionally, autoencoders consist of an encoder that transforms the input into an internal feature representation and a decoder that reconstructs a signal from these features. For learning representative features, the reconstruction error of the inputs is minimized. To learn features that are stable across trials of the same stimulus, we propose cross-trial autoencoding where instead paired trials belonging to the same stimulus have to be reconstructed. For our second technique, similarity constraint encoding, these paired trials are extended into triplets by adding trials from different stimuli that need to be recognized as less similar than the pair based on the feature representation. This facilitates learning features that allow to distinguish trials from different stimuli and at the same time creates an embedding of the trials into a similarity space. Combining these two techniques and using only three trials per stimulus for training, we can correctly recognize the paired trials as most similar within the triplets for unseen trials with a significant accuracy of 70-75% (above 50% chance baseline) depending on the subject.

***Understanding the role of features in emotion recognition using the MAP dataset***

**Naresh Vempala, Frank Russo***

***SMART lab, Ryerson University**

Several previous studies have predicted listeners' global judgments of felt emotion on the basis of aggregate assessment of musical features, while others have predicted continuous judgments on the basis of moment-to-moment assessment of features. Prior attempts at modeling have been relatively successful, however, they have not provided sufficient insight regarding the relative contribution of audio features on emotion judgments. In this study, we collected continuous valence and arousal ratings of felt emotion from 60 participants for 60 music excerpts, as a part of the Music, Affect, and Physiology (MAP) dataset. We chose excerpts from four different musical genres, and designed our study such that 12 unique listeners heard each excerpt. We extracted audio features ranging from low-level features reflecting energetic and timbral aspects of music to mid-level features reflecting rhythmic and tonal aspects. We used artificial neural networks to predict the mean valence and arousal ratings of participants, at different levels of temporal analysis. Our analyses will address the relative contributions of audio features across levels of temporal analysis, and how they may differ across genres.