Notes of the Modeling Repetition and Variation Project at the WiMIR 1st Annual Workshop

http://bit.ly/2OlhTKl

Morning: 10:00 - 11:00

Attending (please add your name if it's missing):
- Anja Volk (Utrecht University)
- Hendrik Vincent Koops (Utrecht University)
- Iris Yuping Ren (Utrecht University)
- Juan Pablo Bello (New York University)
- Eric Nichols (Microsoft)
- Jaehun Kim
- Marcelo Rodriguez (Yousician)
- Chanhong
- Jeng Cheng

Goals

- Juan: get to agreement to the definition of the task. What is the task definition?
- Tejaswinee: What do mean by "patterns"? What is the time scale?
- Eric: Agreement between annotators
- Difference between musical styles/time eras, pattern means something else in different styles
- Eric: composers are not the right source to get the definition of a "pattern"
- Juan: what is "ground truth"? disagreement is a fundamental property of the tasks we are interested in.
- Juan: Good idea for this workshop to have multiple annotation paradigms, to compare annotation methods and annotations

Afternoon 13:00 - 16:00

- Anja: I stopped when it became an "intellectual game", what is the right time window to make annotations?
- Vincent: similarity: when are patterns still patterns: what is the extent of variation?
- Eric: The problem over overlap: beginning of a pattern could be the end of another pattern
- Eric: Degrees of difference: what are minor variations or really a different pattern
- Eric: Filler material
- Eric: rhythmic patterns are important, sometimes in decisions the pitches do not matter in deciding that a pattern repeats
- Anja: Visually deciding what a pattern is can be different from listening
- Tejaswinee: Very small figures and arches that contribute to pattern discovery
- Tejaswinee: difference between ornamental and explicit trills
- Tejaswinee: importance of rhythm: if I can clap two excerpts in a similar way, I consider them to be similar
- Anja: what Tejaswinee annotated as important is just noise to me
- Patterns early in the piece establish patterns further on
- Tejaswinee: Indian music has more "filler" in the beginning, in contrast to western classical music

- Juan: my annotations have two main motives, start and line 5
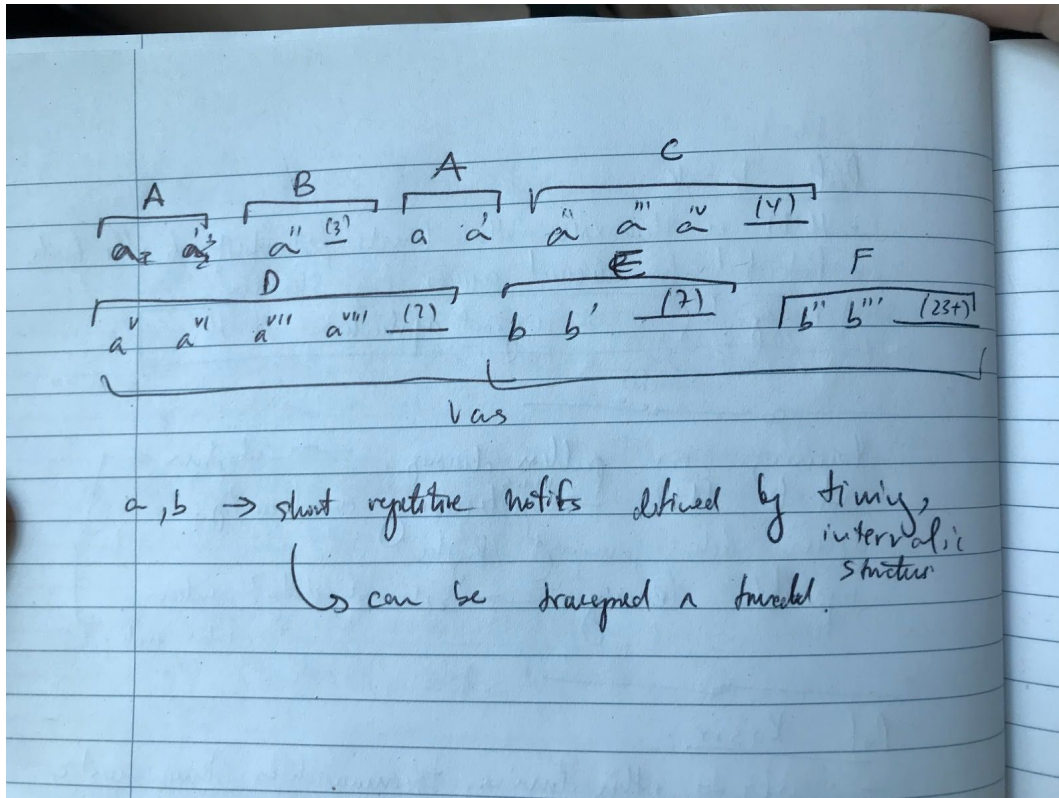
- Juan: what is the core material of the piece? What is filler?
- Juan: The performance emphasizes the rests, it influences decision making on what is a pattern or not

How can we improve?

- Eric: One way to deal with multiple annotations is majority voting. But maybe having multiple reference annotations is a better idea than having a single ground truth.
- Get extra information from annotators
- Juan: Talking about each other annotations make me understand different annotations
- Juan: Annotations from algorithms could come with explanations, that clarify why it is a pattern
- Anja: Instead of just an output, also give a reason for the algorithm output
- Iris: could features serve as an explanation of a choice of an algorithm?
- Juan: you would need a lot of user data to create good explanations
- Tejaswinee: what is the spectrum of annotations? What are qualitative criteria to describe annotations? What descriptions could there be?
- Juan: For example: a unit could be explained by a melodic contour. Show which elements are robust in different variations of a unit. Which dimensions are changed and which are kept the same.
- Anja: what is the limit of transformation?

- Eric: Fundamentals of musical composition - Schoenberg.
  - https://monoskop.org/images/d/da/Schoenberg_Arnold_Fundamentals_of_Musical_Composition_no_OCR.pdf

- Anja: how can we make evaluation less strict, perhaps more fuzzy? Different degrees of probability of patterns/similarity.
- Juan: annotation process, both listening and reading helped making annotations.
- Eric: I want to listen to it a couple of times and take my time. It takes work. The music is too simple if it one takes one take.
- Vincent: Too much analyzing makes me find patterns where there maybe are none: over analyzing a piece
- Eric: It takes practice to annotate a piece
- Tejaswinee: I preferred to work with only the score,
- Anja: in what way do we annotate: in depth, like a composer or more "superficial"?
- Vincent: important to keep in mind what the goal of an annotation is, for what purpose is it gathered?
- Anja: what is the point of pattern finding algorithms? What is the purpose?
- Iris: Algorithms to better understanding structure of a piece
- Eric: How to process music we are hearing? Finding patterns is important for understanding music in general, it is basic MIR
- Eric: Multiple annotations could aid in better understanding how humans process music.
- Juan: Educational context. Pattern finding as an annotation tool.
- Iris: Etudes are used in educational context, it has well defined repetitions.
- Tejaswinee: can we use generational systems? Generate variations of a pattern to investigate variation and see the similarity of a piece.
- Juan: use interactive learning in a practical context to get good data on a large scale.
- Vincent: it is an intellectual academic exercice, how to make it "cool" for users
- Anja: application: learning a new piece
- Vincent: Metronaut

Conclusions

- Vincent:
  - Variations and patterns are very fuzzy concepts, meaning that their units are not always clear, which leads to lower inter-annotator agreement. One way of obtaining annotations with more agreement could be obtaining more task-specific annotations.

- - - Finding patterns and variations is a high level task, compared to for example chord label annotations, this is problematic for obtaining large amounts of annotation data, and creating crowd-sourced annotation tasks at scale.
    - I really like the suggestion of including explanation and rationale in annotation tasks, which could extent training and evaluation automatic pattern/variation finding systems. Sometimes disagreement
    -
- Eric:
  - We need to define the space of possible annotations for our specific task.
    - E.g.,: We will label motifs, repetitions, degree of repetition, melodic sequence, cadential material, "filler", etc.
    - I'd like a nicer term than "filler" for "stuff that's not core motivic content"
  - To scale up annotations a software interface tailored to the task would be helpful.
  - Annotations should be done at different levels of hierarchy.
  - It's important to allow temporal overlap between multiple motifs
  - Multiple annotators can be used to improve confidence in annotations and reduce ambiguity. In the case of clear ambiguity, perhaps multiple co-existing annotations should be recorded (and metrics modified accordingly)
  - It would be nice to annotate the specific ways in which motifs are transforms of one another (rather than just saying A, A', A'', etc.) Schoenberg's composition text may provide a list of possible motivic variation types.

- Juan:
  - This is clearly an area where the ground truth paradigm is insufficient, and any computational task cannot be simply defined as a classification problem
  - Multiple interpretations require an understanding of the rationale behind the decisions made by each annotator. Discussing our thought process was very illuminating, so perhaps we should require our algorithms to be transparent with their decisions as well.
  - Human-in-the-loop computing seems like an appropriate paradigm for the analysis of repetition and variation. Perhaps this could be conducive to ecologically-valid setups for data collection and validation, e.g. in score annotation for performers, or music theory exercises for students.
- Iris:
  - Degrees types of patterns: gap fillers (high note density related), explanation in what are they. Prototypes and variations. Categorisation.
  - Different degrees of listeners, composers, musical activities are performed differently
  - Annotation process can be incorporated in the loop of improving the algorithms. Harder to incorporated
- Anja:

- ○ our evaluation methods in MIR force us to ask for yes/no answers that humans in many contexts would not have; specifically in the case of musical patterns; hence a very precise number can nevertheless be pretty meaningless
  - ○ Musical patterns are definitely very important to how humans perceive and understand music, so it is a very important topic of research, but we need to deal with multiple possible plausible interpretations/annotations.
  - ○ How do we get to ecologically valid annotations of patterns?
  - ○ Once we have multiple annotations: how do we evaluate the pattern discovery algorithms? Is it better to let algorithms search for "candidate patterns" and let humans decide how useful they are? In what context useful? E.g. in learning to play a new piece?
  - ○ How do we distinguish the parts of "meaningful repetitions" and "filling parts" where there are lots of repetitive elements but as a human we would not annotate them as "significant parts" of the composition? How does an algorithm distinguish that?
- ● Tejaswinee:
  - ○ Score annotation has its limitations, but annotation tasks could also be of different kinds, for example annotation through movement contours, tracings and other ways.
  - ○ Annotation tasks could be subclassified into a large family of variances. Annotation for education, for ease of playing a new score.
  - ○ It is important to remember that doing this task of pattern discovery is itself important for musical education.
  - ○ Score pattern finding, and pattern finding through listening appear to be separate tasks to me. As a listener, i am sensitive to a lot more energy and 'non score' material.
  - ○ Annotation and generation can be considered as two separate components to workaround the same problem of pattern <finding> or <fitting>, or variation finding. For example, if meaningful variations can be generated, and they can be generated across several different types of meaningful dimensions, then some of the repetition occurrences in compositions themselves could be accounted for in the 'distance metric' of such a generator.
  - ○
- ● Changhong:
  - ○ It would be helpful if the annotation process be easier and less ambiguous, eg. conducting an investigation of the candidate annotations to create a label set with explanations and agreement levels for users to select directly. This process itself can also be a data collection process by letting annotators add their new descriptions and reasons. Each round of data collection can be used to train design features and optimize algorithms.
  - ○ Transfer the annotation process to a well-designed task/game is a good solution to this problem. Also, user background can be considered in the task, that is, different user different tasks.

https://github.com/hvkoops/wimir2018

Annotation Process:

(A thought)

Listen to the following pieces and annotate the salient melodic patterns with

1) Rate how relevant this pattern is to this piece:

1 = Not as relevant

2 = Relevant

3 = Highly relevant

2) One word to label the type of this pattern: rhythmic, intervallic, create your own

3) A short description on why you find it to be a pattern: no word limit but as concise as possible.

4) How difficult it was for you to decide whether it's a pattern and why:

1 = Not difficult

2 = Inbetween

3 = Very difficult

Discussion:

Put into application for education

Without sheet music

Contour

Annotation from everyone:

https://photos.app.goo.gl/1a7Tg8Jq36RsZdhH8