Examiners' commentary 2017–2018

CO3346 Sound and music – Zone A

General remarks

Overall performance in this examination was good. There were a few weak papers and some strong ones, with only a very small handful of fails. What follows is a brief discussion of the individual questions on this paper, with hints towards the answers expected by the examiners.

Comments on specific questions

Question 1

Computational models of music cognition

Many candidates chose this question, usually with good responses.

Part (a) was material directly from the subject guide and most candidates were able to correctly say that grouping is about combining musical events together, for example chords or bass-line and melody, and that expectation involves modelling or predicting what a person might expect next when listening to a piece of music.

Similarly for part (b), simply knowing the material in the subject guide allowed for a good answer. Any of the Gestalt principles could have been used, for example for proximity, the visual aspect may be that things that are physically close together would be perceived as a group, while for musical aspects, sounds that are close together in time may be perceived as musical change having not occurred. Whichever Gestalt principle was chosen, it was essential to demonstrate that it is the perception that is under consideration here.

Most candidates were able to explain, for part (c), that Schellenberg developed a model that incorporated pitch information, as well as the concepts of registral direction (revised) and registral return. Given a pair of consecutive pitches in a melody, *x1* and *x2*, we can compute the conditional probability *p(x1 — x2)* as being the number of times that *x1* is followed by *x2* , divided by the number of times that *x1* occurs regardless of what follows it.

The final part, part (d), required some ability to demonstrate understanding of concepts learned, rather than only being able to recall information from the subject guide. For example, for sub-part (i), a good answer would be that it is useful to experimentally validate the predictions of a cognitive model of musical expectation, so that you can know if your model can make predictions about real human expectation.

A probe tone experiment, for (ii), could take the approach of playing a melody to several people and asking them which note they anticipate next. For sub-part (iii), examiners required an annotated diagram of the user interface of whatever was proposed, which not all candidates provided. What was expected was some kind of melody player, with a 'next note' selector, and finally, for sub-part (iv), data would most likely be pairs of melody sequence with next expected note, and the analysis would involve computing the probabilities of predicted notes for different melody sequences.

Question 2

Interactive Sound using Pure Data

This was a reasonably popular question, again answered well by most candidates.

For part (a), the four categories expected would have been object, number box, message and data table, with corresponding explanations of what each of these are. A few candidates included a bang but this is not a category.

Part (b) was usually answered well, and is material that comes from the subject guide. For frequency modulation, oscillators modulate the frequency of other oscillators, with appropriate gain and frequency to generate audible side-band frequencies, while with subtractive sound synthesis, a spectrally rich waveform is sent through filters to remove (subtract) partials. Modulation can be applied to the parameters to make the sound dynamic. In terms of efficiency, examiners expected the response to include consideration of the CPU use required to generate sines vs. filtering, and so on.

The main requirement of part (c) was the application of knowledge and techniques, to demonstrate understanding and obtain actual answers. The technique being used in the patch is frequency modulation, and it's being used so that the oscillator modulates the frequency of another oscillator.

The appropriate ranges are for A, —1 : 1, for B, —100 : 100, for C, 100 : 300 and finally for D, —1 : 1. To play a note with a different pitch, you would edit the number object with 200 in it.

The rest of the question required demonstration of understanding through explanation and discussion. For part (d), candidates did not always include comparisons with other software environments in their discussions, and nor did their answers always address the technical skills required for using Pure Data.

For part (e) examiners expected an explanation that sampled sound is pre-recorded sound, for example an instrument recording, while synthesising sound generates sound on the fly, usually from a model or using mathematical calculations.

Finally, for part (f), a good answer might be that line is an object that can be used to iterate through the samples in a table to play them back, while tabread is an object that can be used to look up samples in an array using an index. Many candidates gave some very general answers, which indicates a more limited understanding of the material.

Question 3

Algorithmic composition and musical interaction

This was a very popular question, chosen by almost all candidates and answered reasonably by most.

Part (a) was a simple question from the subject guide. Three characteristics are rhythm, pitch and timbre.

Part (b) indicated very specific requirements in the answers expected. Candidates had to identify, for (i), one each of a characteristic that the system does have and does not have. The system is clearly reflective, as it reflects the notes that the musician plays; and it is not innovative as it does not generate its own notes. Not all candidates provided justification of this sort for their claims. For (ii), some examples could be: maximise reflection by extracting note sequences and searching for fragments with the same note sequence; you could maximise transparency by visualising the data the system is collecting from the player and how it matches to, say, YouTube media. One candidate provided examples for three of these, but examiners considered only the first two. It is important to answer the question being asked, rather than to simply give information.

Part (c) expected candidates to explain that the concept of a metrical hierarchy is that there are several metres in a piece of music, for example, the interval between the fastest notes, the interval between start of each bar, and so on. It makes beat tracking difficult as it is not clear which metrical hierarchy the beat needs to fall on.

Part (d) related to material in the subject guide and required a diagram indicating the two different situations. To show good beat tracking, the lines would be in a similar place, running from the centre to the circumference, within a circle, while in the other they would be in a range of different locations, though also running from centre to circumference.

Finally, for part (e), any coherent and sensible approaches were accepted. One example could be that the process would follow something along these lines:

```
Set interval to first offset value.
Computer predicted offset values from 0 with that interval.
Count how many matches there are to the actual sequences and how
many non-matches.
Increase the interval, repeat.
Select the interval with the highest hit to miss ratio
```

Question 4

Music Information Retrieval

This question was chosen by some candidates, and again answered reasonably.

For part (a), examiners expected a sensible discussion of the fact that genres are not formally defined such that we cannot absolutely assign a genre. Therefore, it is a problem with low specificity as many answers are correct.

Part (b) relates directly to material in the subject guide. The bag-of-frames approach combines all sequential frames into a single feature, ignoring time, while with a sequence approach, frames are compared in sequence.

Following on from this, for part (c) it is most likely that you would prefer a sequence approach, as the bag-of-frames approach does not consider changes in the song, as it ignores time. However, you would probably want to do a moving average bag of frames vs. 'bag-of-frames seen so far' feature.

Part (d), the final part, needed a work through of a recommender system design, to fulfil certain requirements. This was given in a step by step way, and each stage was awarded part of the marks. Sub-part (i) required a high level description, such as 'the system collects data about users' listening habits, then finds users with similar listening habits. Then it finds the songs that user 1 has not bought, but that user 2 has bought, and recommends them". Sub-part (ii) was about the data such a system would require, such as lists of songs that users are listening to. Sub-part (iii) required insights that one approach could be to extract musical notes from music, and reject items with similar note sequences. Finally, for (iv), a diagram was essential, plus a database including users and some algorithm that converts the user data into a search query on the database.