Jieun Oh
Prof. Eleanor Selfridge-Field
Music 253: Musical Information
March 20, 2007

Final Essays

Essay 1: Encoding of Melodic Material

OVERVIEW

I chose a repertory of classical flute music from approximately Baroque Period to 20th Century (excluding modern, as I am not too familiar with very modern pieces that feature special techniques). As a wind instrument, flute is unique in that it involves controlling a stream of air using the embouchure beyond just fingering the notes. Specifically, I find tonguing, breathing, and vibrato to be important features of the instrument; representation of flute music (whether it be notation, sound, or analytical) should ideally capture at least some of these features. In the following sections, I discuss the hypothetical encoding of melodic material in the repertory of classical flute music for (a) notation, (b) sound, and (c) analytical applications. I will discuss potential problems encountered in each of these domains and further describe to what extent they are solvable.

A. NOTATION DOMAIN

Description of Encoding

Tonguing: In addition to traditional articulation marks such as legato or accent that suggest how one should tongue (i.e. soft “dah” for legato and harsh “ttah” for accent), notation for the flute repertoire should have the flexibility of stating the type of tonguing for a given passage of music: single, double, or triple tonguing. Often, these different types of tonguing are at best inferred from the tempo of the piece, but since the range of tempo and occasions that are feasible for each of the three types of tonguing overlap, sometimes it is difficult to tell which is optimal (or intended by the composer). Of course, this
would not be as relevant if the sound generated by the different types of tonguing were indistinguishable; however, the reality is that they are not so.

For instance, although they all produce six tongued notes in the following example, single tonguing six times (Tah Tah Tah Tah Tah Tah) sounds different from double tonguing three times (Tah-kah Tah-kah Tah-kah), which is also different from triple tonguing two times—of which there are three types: (Tah-kah-tah Tah-kah-tah), (Tah-kah-tah Kah-tah-kah), and (Tah-tah-kah Tah-tah-kah). Not only is the “kah” syllable generally weaker than the “tah” syllable, the way you group them results in a different accentuated beat as the tendency is to give a bit more emphasis on the first of each grouped notes.

For these reasons, it would be much clearer to the musicians if the notation indicated the type of tonguing (1, 2, or 3) with which syllable to use (‘t’ or ‘k’) using small font above the music notes for which the choice of articulation is ambiguous. [This would be analogous to how piano or violin music often has fingerings or bowing instructions indicated above the note-head; see below for example.]

**Breathing:** While a great portion of flute notation indicates where to breathe using a breath mark (as in choral music), I find two sources of problems with the traditional notation: (1) individual differences on how often one needs to breathe, and (2) how much time can (should) one take to breathe. The former is especially relevant for players like me with smaller lung capacity who must take a breath more often than what’s indicated on a page. The latter is relevant in that not all breath marks have the same purpose: some breath marks should be observed, even if one does not really “need” it, in order to give a sense of phrasing or to create a pause; others are there to suggest to the performer places where it would be the least interruptive to breathe—although no breath may be the optimal choice—in which case the breathing should be done hastily, in tempo.

Thus, in this hypothetical encoding of notation domain, one would be able to differentiate between mandatory breaths (the breath mark < ’ > would be circled) and optional breaths (the breath mark < ’ > would not be circled, just like in traditional notation), and one would also be able to tell how desirable it is to breathe in a given place by having undesirable breaths shown in lighter shade of gray. Of course, there are some problems with this method, which is discussed in the following section.
**Vibrato:** As far as the notation domain is concerned, I do not think special vibrato instruction is necessary, because the appropriate width and frequency of the vibrato is better determined by our ear than by a fixed number; in fact, vibratos sound artificial and machine-like when the pulses line up with the beats in the music.

Potential Problems & To What Extent They Are Solvable

**Tonguing:** Specifying the ‘t’ and ‘k’ syllables as well as tonguing types take away freedom from musicians who want to use the type of tonguing that suits them the best. Within triple tonguing, for instance, there usually are individual preferences to one of the three—I myself prefer “Tah-kah-tah Tah-kah-tah” over the other two; and encoding the specific syllables may become distracting and useless if they are not followed. This problem can be solved by noting these primarily in etudes for pedagogical
purposes, and in instances for which the composer would really like the performer to use one way of tonguing over another in an ambiguous situation.

Breathing: The notation method that I proposed for breathing unfortunately has many potential problems. First, the purpose of a breathing mark is usually a mix of the two reasons mentioned above (that the music needs it there for phrasing and that the player needs it to take in more air), as opposed to only one of the two. Therefore, the decision whether or not to “circle” the breath marks to indicate “the music needs it there, so you have to take a breath” would turn out to be a subjective one. Also, determining the shades of gray to be used for the mark is problematic in that the desirability of breathing in a given place cannot be accurately determined without taking into consideration the previous places in which the performer chose to breathe (in order to maintain the consistent phrasing pattern or to not have the performer breathe more often than is necessary). This problem would be difficult to solve unless there is a complex algorithm that personalizes breathing marks by taking into consideration both the phrasing structures of the piece and the performer’s breathing capacity.

B. SOUND DOMAIN

Description of Encoding

I assumed that the sound encoding would be similar to MIDI in many ways: tracking key number and duration, along with other information needed to reproduce the sound digitally. While MIDI has many inherent shortcomings (such as “indistinguishability” of enharmonics), my proposed encoding of flute repertory in the sound domain assumes the basic set-up of MIDI and attempts to add new features to it to accommodate some of the more unique features of the instrument.

Tonguing: We could distinguish a “tah” from a “kah” by creating a unique hexadecimal representation for the latter (“Tah” would be the default tonguing). Thus, for instance, each note that should be tongued using a “kah” should be indicated of this fact in the way similar to how its key-num and attack velocity are noted.
Breathing: The question of whether to encode breathing or not is directly related to the bigger issue of whether we want our MIDI event timings to reflect the physical notational duration or the actual performed duration. In the former case, the encoding of breath marks should not result in any audible differences in the MIDI sound output; it would only be there so that the breath marks show up if we happen to export a file to a software in the notation domain (i.e. Sibelius or Finale). In the latter case, however, breath marks would certainly result in a difference in its sound output: at the least, there should be a (slight) gap of time needed to model the time required for someone to take a breath, and we may even want to artificially generate a kind of noise heard when we take a breath (assuming that this “noise” would sound realistic in a way that is not distracting).

Vibrato: Encoding vibrato in the sound domain is tricky in that it would require a very large number of sampling to mimic the exact width and pulse of vibratos at any given point in music. On the other hand, if we simply set all vibratos to be of equal width and frequency throughout the entire piece of music, the generated sound would seem very artificial. Thus, if our primary reason for encoding vibrato in the sound domain is to characterize it qualitatively for analytic purposes (rather than for recreation of the exact sound, for which a recording better serves the purpose), we may encode this information by discretizing a given piece of music based on vibrato-quality, then qualifying the vibratos used in each of these sections based on the amplitude and frequency of the vibrato on a scale from 0 to 5.

Potential Problems & To What Extent They Are Solvable

Tonguing: We should be careful to add this feature only if we want to encode and preserve the slightly different sound quality of “kah”, and if such encoding would be programmed in such a way that indeed mimics the sound of a “kah” tonguing.

Breathing: Obviously, similar to how MIDI files that were produced from an actual performance do poorly when one attempts to export it to the notation domain, MIDI files that regard breathing as an event (or even a series of event: inhale followed by a quick pause to “reset” embouchure) have the potential to experience similar problems. For instance, exporting a sound-domain file that incorporated breathing into the notation domain may generate unwanted rests and pauses in music, causing
misalignment between beats and notes. To minimize this problem, one should clearly label this section in the sound encoding as “a breathing mark” so that the events that occur within this section do not get treated as sound and rests with durational values, but rather get translated as a breath mark symbol in the notation domain.

**Vibrato:** As described above, the notation that I devised would not result in a sound output that exactly captures a flutist’s vibrato. Rather, it qualifies the amplitude and frequency of the vibrato using a scale so that an extremely wide and somewhat quick vibrato would be labeled as having amplitude 5 and frequency 3. From this, we can attempt to encode a more realistic-sounding vibrato by discretizing the piece in smaller chunks, each of which would be characterized by more precise scales of amplitude and frequency that allows decimal places (i.e. amplitude 4.96 and frequency 3.20).

### C. ANALYTICAL APPLICATIONS DOMAIN

**Description of Encoding**

Incorporating information about tonguing, breathing, and vibrato in an analytical applications domain (such as Humdrum) should not be too difficult to imagine, as it would be quite analogous to specifying dynamics in an additional column that runs parallel to the “main” column in Humdrum. Again, in proposing an improved encoding scheme for flute music in the domain of analytical applications, I assume that we already have a Humdrum-like system that can already adequately handle traditional western music.

**Tonguing:** We can encode tonguing types using an additional column that runs parallel to the main one that denotes the consonant of the proper tonguing type that should be used. For instance, ‘t’ would be for regular tonguing, ‘k’ would be for the second syllable of double tonguing, ‘d’ would be for legato tonguing, and ‘g’ would be for the second syllable of legato double tonguing, and so on. We could even convey extremes using double letters: ‘tt’ would be accented tonguing, while ‘ddd’ would be a super-legato tonguing. A “.” would suggest no tonguing; that is, part of a slur or a tie. This way, we can
even get rid of the articulation marks, traditionally conveyed using parentheses and brackets, in the main column that encodes pitch and duration.

**Breathing:** Similarly, we can add a column devoted to breathing running in parallel. A ‘B’ may be for “required” breaths, a ‘b’ for optional breaths, and ‘nb’ for no breaths (for places where the phrasing should not be interrupted by a breath). And these letters can be placed on the same row as the note (from the main column) after which the breathing should occur. As in other domains, incorporating breathing can be problematic because deciding on how “desirable” it is to breathe is often a very subjective task.

**Vibrato:** Another column can be used to convey the quality of the vibrato suitable for given section(s) of music. Similar to how vibrato quality was encoded in the sound domain, a numerical scale used to characterize the relative amplitude and the frequency of vibrato used. However, I feel that there is no need to have two separate parameters of amplitude and frequency of the vibrato, since the two are usually directly related, such that increased amplitude is usually accompanied by increased frequency as a way of achieving heightened intensity. A series of “.” in between two rows of vibrato markings would suggest that the previous value is used until the new value is encountered; a series of “<” would suggest a smooth intensification of vibrato, while a series of “>” would indicate a smooth weakening of intensity.

*Potential Problems & To What Extent They Are Solvable*

**Tonguing:** I actually quite like my hypothetical encoding tonguing (and articulation in general) described above. A potential problem would arise when a piece can be performed in multiple instruments (one that utilizes tonguing and one that does not--i.e. string instruments). However, I feel that it is fine to convey articulation types using these letters for other instruments (including non woodwind/brass instruments) despite the fact that these notations were originally derived specifically from tonguing.

**Breathing:** Breathing causes more problems in that they depend hugely on the performer of the music, and those who approach the music from a perspective other than a performer may not even notice it. The column for breathing can hopefully used for conveying phrasing in addition to breathing (so that its use can be more universal) –by suggesting places where small (‘b’) and large (‘B’) phrases end.
**Vibrato**: Though the choice of vibrato intensity is also quite subjective, I feel that it is generally more predictable than breathing. Again, these markings can be interpreted in a more broad manner as “levels of intensity” at a given moment in the piece, and in doing so we can hope to minimize the inevitable problem of subjectivity that arises when notating music based solely on a specific performance of a chosen musician.

![Taken partly from Ex 3 of SCORE exercise: Hypothetical encoding for Analytical Domain]

```
**kern**  **tonguing**  **breathing**  **vibrato**
*staff1*  *  *  *
*clefG2*  *  *  *
*k[b-]*  *  *  *

16e  t  .  .  \// intensity = 3
16f  k  // double  .  \// gradually intensify
16g  t  .  .  
16f  k  .  .  
16g  t  .  .  
16e  k  .  .  
16f  t  .  .  
16g  k  .  .  
16a  t  b  // if needed  .
16b- k  .  .  
16cc t  .  .  
16b- k  .  .  
16a  t  .  .  
16b- k  .  .  
16a  t  .  .  \// ... until here
16g  k  .  .  4  // intensity = 4

8a  tt  // accented  B  // do breathe  .  \// maintain intensity
8a  t  .  .  
8g  t  .  .  
8g  tt  .  .  
8a  tt  .  .  
8a  t  .  .  
8g  t  .  .  
8g  =  .  .  

16a  d  // legato  b  // if needed  \>  // decrease intensity
16b- g  // legato double  .  \>  
16a  d  .  .  \>  
16b- g  .  .  \>  
16cc g  .  .  \>  
16dd g  .  .  \>  
16cc d  .  .  \>  
16b- g  .  .  \>  // ... until here
2a  dd  // super legato  .  2  // intensity = 2
*  .  .  .  *
```
CONCLUSION

The unique characteristics of flute that I hoped to account for in my hypothetical encoding system had to do mostly with the way flute is a woodwind instrument. Among tonguing, breathing, and vibrato, I found breathing to be the most difficult feature to account for in my system because of its highly variable and subjective nature (depending largely on the performer and his/her interpretation of the music). While I do not feel too satisfied with my encoding of breathing (especially in the domain of notation and analytical application), I felt that my way of incorporating tonguing—or, articulation type, to be more general—to the various domains of musical information is quite reasonable and feasible.

Essay 2: Attributes of Musical Information in the Domain of Graphics (Question 2)

Musical Information in the domain of graphics is unique in that it must visually communicate to the music readers in a manner that allows them to obtain a very good sense of both the logical information and the intended sound that the music should produce. Because of this added visual component, the physical appearance of the musical information as printed on a page is of crucial importance in this domain. Thus, it is not enough to simply have the music note in the correct vertical position on a stave (for pitch) with a correct flag type (for duration); we must further make sure that its relative position with other objects are appropriate so that their logical relationships can be conveyed clearly. Also, the overall appearance of the music should also be considered, as we do not want the information cluttered on a page for instance, as that would make it tough on our eyes. Conversely, however, the meanings of the objects that get drawn on a page are much less important, as far as the graphical notation software is concerned. (That is, with an unrealistic assumption that the graphical domain does not have functionalities such as playing back the music or importing/exporting to softwares of different domains…). In this manner, of a primary concern in the graphics domain is optimizing visual communication to human readers (musicians) of the musical information.
Specifically, the following attributes of musical information are usually required to manipulate data in this domain: the *placement* and *shape* of (1) background information that by themselves do not produce any sound but gives important ideas of the overall piece (i.e. clef, key signature, time signature, tempo markings, repeat signs, double bars), (2) pitch and duration of notes (including information on how they are grouped when chords or multiple voices are involved), which are “sound units” that actually generate sound, and (3) extra markings that provides more detailed information and instructions as to how these individual sound units should be generated (i.e. dynamics, articulation marks, slurs and ties).

Knowing each of these requirements allows the graphics-producing software, such as SCORE, to pick out an appropriate musical symbol for it and to determine a *rough* placement of these symbols. For instance, knowing that the treble type is “treble,” it should know to take a symbol for a treble clef, and place it in the left-most feasible place on the stave so that it is in an appropriate alignment with the five lines of the stave. However, what makes excellent graphical information different from other domains of information is that, from this point on, there must exist an easy-to-use mechanism to manipulate and fine-tune the symbols so that they “look” good to our eyes. Some of these modifiable parameters include its horizontal and vertical positions of key features (such as beginning and end coordinates of a slur), size or length of the symbol, curvature and thickness, and types (i.e. dotted versus solid lines versus squiggly lines). And, of course, the whole point of going through the trouble of fine-tuning the objects is to achieve optimal visual communication so that the intended logical and sound information can be accurately inferred from the graphics. Another reason is to heighten the aesthetic level of the resulting output—in order to make them more desirable for consumers.

Therefore, in contrast to the sound domain (manipulations in which directly contributes to audible differences), manipulations in graphical information tend to influence the sound of the encoded music only indirectly, by changing the logical information that is conveyed by the graphical notation and/or by influencing the way musicians interpret the visual cues. Similarly, in contrast to the “logical” domain (changes in which directly influence how we understand the relationship of notes and measures and
voices, etc), manipulations of the graphical information may not guarantee that such changes in the conceptual understanding of the music would take place in the readers’ mind (although ideally it should).

To sum up, the graphics domain is like an easy-to-use interface for qualitatively conveying the intended sound to the human audience: its layout is quite clear and meanings are easy to understand by humans, as it has been developed and shaped by history to suit human visual cognition capabilities. Yet, because its information is rather qualitative and is meant to take in an additional layer of interpretation by the human musician, it produces different sound outcomes depending on who this “user” is; it does not provide the precise, quantitative instructions needed for machines to generate the intended sound out of it.