# Orchestration within the Sonification of basic Data Sets

*Charlie Cullen*

Conservatory of Music
Dublin Institute of Technology
Rathmines Road, Dublin, Ireland
charlie.cullen@dit.ie

*Eugene Coyle*

School of Control Systems & Electrical Engineering
Dublin Institute of Technology
Kevin Street, Dublin, Ireland
eugene.coyle@dit.ie

**ABSTRACT**

The use of sonification as a means of representing and analysing data has become a growing field of research in recent years and as such has become a far more accepted means of working with data. Existing work carried out as part of this research has focused primarily on the sonification of DNA/RNA sequences and their subsequent protein structures for the purposes of analysis. This sonification work raised many questions as regards the need for sequences to be set to music in a standard manner so that different strands could be analysed by comparison, and hence the orchestration and instrumentation used became of great importance.

The basic principles of sonification can be rapidly extended to include many different data elements within a single rendering, and thus the importance of orchestration grows accordingly. Existing work on the use of rhythmic parsing within a sonification had suggested that far more information could be represented when orchestrated in a rhythmic manner than when simply reconstituted in single musical block. The principle was further extended to include the allocation specific instruments and pitches within rhythmic patterns so that each sonic event would convey the data it was intended to represent. To this end a fictional database of employees in a company was created as a means of developing the principles required for more effective sonification through orchestration.

The employee database was intended as a means of using a straightforward data set to analyse the effect of basic changes in instrumentation and orchestration rather than the data itself. The allocation of chord intervals or melodies to different data elements allowed the data to be represented in different ways at output in order that these differences would eventually highlight some form of framework for effective sonification of data sets with multiple elements.

## 1. ORCHESTRATION IN SONIFICATION

The use of orchestration as a tool for the better organisation and representation of music in a piece is of great importance in any work- particularly in pieces that contain many different instruments such as orchestral works or big band ensembles. The principles involved are often scaleable for different numbers and types of instrument and so lend themselves towards application within sonification as a means of combining different data within a single rendering.
Any orchestration of a piece must be combined with its instrumentation so that each part of the work is best represented individually and collectively within the final piece. In the examples given here basic instrumentation choices were made that made use of the range available within the General Midi standard. As such they can by no means be regarded as definitive or indeed final, but rather serve as markers towards future development in that area.

### 1.1. Basic Orchestration Definitions

A musical orchestration can be described as a piece of music that has been balanced for performance by a particular set of voices or instruments [1]. The basic premise of orchestration as it is applied to modern music is the organisation of different voices or instruments playing different musical parts so that each part can be heard individually and also within the overall context of the piece as a whole.
Even at a basic level the use of orchestration allows groups of instruments to interact in a piece and still retain their individual parts within the whole work.

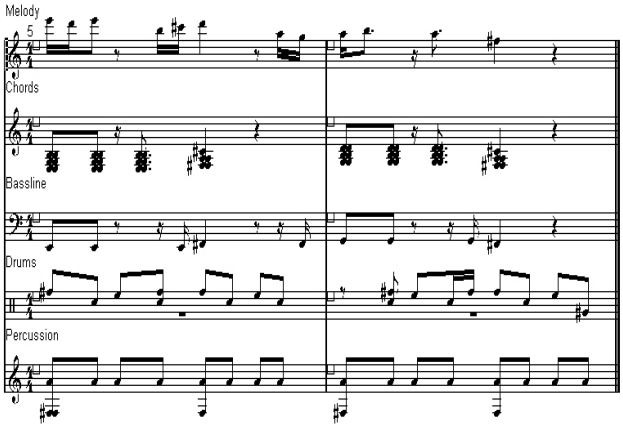


*Figure 1: An example of basic orchestration in 5 parts.*

From the above score it can be seen that at least five different elements of data are being conveyed within a single musical passage, elements that themselves could be used to represent many different data values within a defined set.
The use of musical orchestration for the conveyance of greater amounts of information within even the most basic of music frameworks (e.g. bass, chords and melody) is potentially of great use within sonification, especially when used in conjunction with rhythmic patterns as a means of segmenting data for representation.

## 2. APPLICATIONS OF ORCHESTRATION TECHNIQUES TO BASIC SONIFICATIONS

Many databases relate to the basics of employee wage structures and their corresponding data such as age, position, gender, attendance and so forth. It was felt that such a basic structure would be a good place to start when seeking to

develop a more comprehensive framework for orchestration within data sonification, and hopefully any analysis could be performed in as straightforward a manner as possible. A basic database was set-up containing a number of fictional employees in a company. Within this framework various simple elements were defined for possible analysis [see Appendix].

### 2.1. Initial Database Analysis for Orchestration

Previous work on the use of rhythmic intervals [2] as a delimiter for sonification of data suggested that any definition of the data in basic rhythmic terms should reflect elements of the data both common and important to all of the data in the set- effectively requiring a primary key [3] within the database structure. Rhythmic structure is potentially more important than melodic orchestration (particularly in large data sets or data which falls within a very narrow range of values) and so even at this stage it is important to define not only why each data value occurs but when.

In the database example, the definition of employee number was chosen as the primary key within the database, and its numerical and sequential nature make it ideal for defining each entry in the database. It must be noted that the allocation of primary and secondary keys in database schema design is not always based upon numerical values (e.g. alphabetical listings) but even within these structures it should be a relatively straightforward process to assign some form of numerical increment to each instance within the database and thus apply the principle of basic rhythmic delimiting.

### 2.2. Instrumentation and Orchestration of Database Elements

Instrumentation can be defined as 'the art of combining the sounds of a complex of instruments to form a satisfactory blend and balance' [4] and as such is a technique that has much potential when applied in data sonification. The simple arrangement of different elements of the sonified data in terms of their sound and timbre allows a far more cohesive and robust template for sonification to be produced. By defining different data elements as specific instruments relative to the data they convey a far more transparent representation of the data can be achieved.

Within the above structure, the definition of bass notes and chord intervals was allocated relative to different aspects of employee information such as wage bracket or seniority. Herein, an employee with a specific job title is allocated a particular bass note within an octave, and this is then related to his or her salary (major chord intervals) and supervisor (basic melody). The definitions were kept wilfully mundane and straightforward (major intervals within a single key and rhythmic interval) in order that such basic information could be sonified and analysed as quickly and clearly as possible without additional distraction.

| *Assume Key Signature of C Major* |
|---|

| **PrimaryKey:** Empl. No | 1 crotchet in 4/4 |
|---|---|
| **SecondaryKey:** Salary | Bassline |
| 10,000- 14,999 | C |
| 15,000- 19,999 | D |
| 20,000- 24,999 | E |
| 25,000- 29,999 | F |
| 30,000- 34,999 | G |
| **Position** | Chord Interval |
| Shop Floor | (C&E) |
| Stock Control | (D&G) |
| Purchasing | (C&Coct) |
| Development | (E&A) |
| Management | (F&B) |
| **Supervisor** | Melody |
| Supervisor 1 | (C, E, G) |
| Supervisor 2 | (G, E, C) |
| N/A (ie no manager) | no melody |

*Table 1: Orchestration choices for test database.*

The intervals defined were also intentionally straightforward in order that the effects of each different aspect of orchestration (and hence orchestration) could be assessed in isolation as well as collectively. The third element of information considered in these sonifications was the supervisor responsible for each member of staff, and indeed in the case of the management themselves the absence of a melodic signature proved just as effective a means of defining them as separate as it was to define an additional melody for them. This highlights one of the potential considerations when designing a GUI driven application of sonification tools in that the less the user has to learn the quicker they will adapt and understand the principles of any new avenue of analysis.

At this stage, the output sonification would contain information about employees concerning their wages, job title and immediate superior all indexed by their employee number. The use of rhythmic intervals allowing all of the data to be orchestrated in such a manner as to be transparent to the listener.

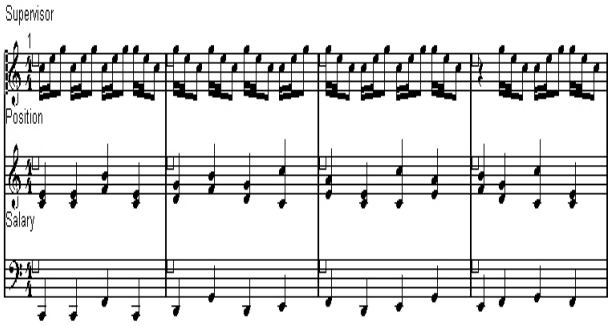


*Figure 5: Output sonification comprising bass, chords and melody.*

The use of orchestration in a sonification of such a data set is obviously driven by the information that is required to be analysed- e.g. location is not an obvious issue in a small firm when analysing wage structures. A basic orchestration such as this may not appear to be of immediate benefit in that the data it works with is fairly straightforward to analyse- even from the initial spreadsheet itself. However, the potential to further define additional elements of the database within this same sonification could suggest more useful application than the initial chord/melody structure.

## 3. RHYTHMIC ORCHESTRATION OF BASIC SONIFICATIONS

As already stated, previous work had centred on sonification and its subsequent rhythmic parsing in order that elements of

the sonified data would be more transparent to the user. In the above data set, additional elements of data can be represented by percussive sounds as a means of conveying more information than with the simple chord/melody structure.

| **Drums**: *Allocation split between Gender and Location* | |
|---|---|
| **Gender** Female | Snare Drum |
| **Gender** Male | No Sound |
| **Location** (floor) | HiHat |
| **Percussion** Age | |
| 20-30 | Long Guiro |
| 30-40 | Chimes |
| 40-50 | Sleigh Bells |

*Table 2: Drum and Percussion orchestration for test database.*

The use of specific drum and percussion sounds is typical of basic orchestration in that percussion is largely used for rhythmic definition and augmentation within a piece rather than as a specific motif that would be considered the basis of the work. To this end it is prudent that the percussion sounds used stand out from the other sounds used in the orchestration so that they can be heard in isolation or conjunction with other elements as required.

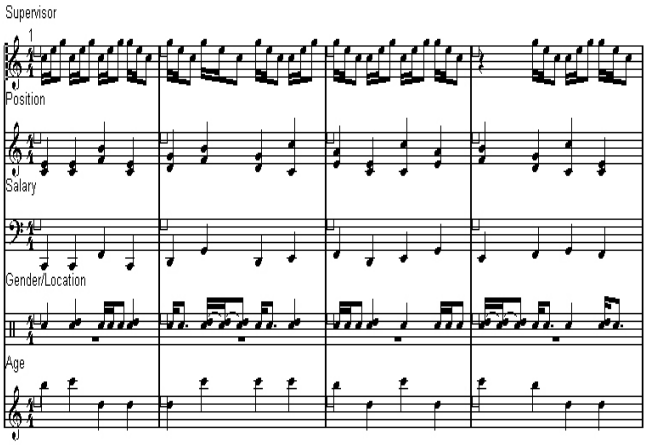


*Figure 6: Output sonification including drums and percussion.*

At this stage a 4 bar sonification has been produced that represents 6 different elements of a basic employee database. The important factor is not the data set being analysed or even the choice of musical pitches and chord intervals used in the output sonification (although this is easily just as important) but rather the notion of layering multiple data elements from a common set within a fairly simple musical structure.

Once a basic sonification template such as this has been set up for the relevant data set (obviously subject to further change) then it can be considered for rhythmical analysis on the basis of representing the data in as obvious a manner as possible.
The first and most important aspect of rhythmic parsing in sonification is the time signature [5] to be used as it will dictate how many elements can be parsed together effectively before clarity begins to be lost. A 3/4 time signature has proven effective for other types of data sonified as part of this research [2] and other signatures may also prove equally useful depending on the data being analysed, but in keeping with the straightforward nature of the examples a 4/4 signature allows for a great degree of flexibility and familiarity to most listeners.

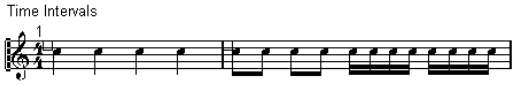


**Figure 6: Basic rhythmic intervals**

The use of 4/4 gives subdivisions of 4 crotchets, 8 quavers and 16 semi-quavers within a single bar of time, and these subdivisions can be used as a means of parsing the different data elements requiring to be sonified. In the above example a basic trio of instruments was created that conveyed employee information regarding salary, age and position. It should be already noticeable that the basic rhythmic delimiter of a bar has been used as a means of incrementing each record in the data set- i.e. each employee has been allocated a bar of music within which all information about them required to be sonified must be conveyed.

It was found that users kept up with initial sequences for a while but quickly fell behind or got confused to the point of being unable to differentiate between intervals they has already recognized earlier in the sequence. This problem was found to be largely associated with the lack of any perceivable rhythm or gap in the music that would allow the listener time to digest and process the information they had just heard, not unakin to reading a book with no spaces or punctuation. The addition of basic rhythmic structures proved highly effective in allowing users to focus on smaller chunks of information in sequence and thus make a more informed judgment about their structure.

Subsequently, in the sonification of the employee database the addition of a basic crotchet rest delimiter in every bar (as in the DNASon software) served to break the information into far more manageable segments, particularly in the bass line which would often be the first element to become indiscernible within even a basic test sequence (perhaps also due in part to the definition of pitches within a single octave major scale).

The orchestration was then segmented further by placing the chords and melody on different beats from the bass, drums and percussion in order that different musical information within the sonification could be represented in as transparent a manner as possible. The use of the crotchet rest as a basic delimiter for all data in each bar gives way to smaller delimiters applied to specific elements of the data set as required. In this case, the melody line is offset by a semi-quaver from the beat of the bassline and drums in order that it does not become lost among the differing percussive timbres employed in the sonification (although assumption has been made here as to the timbre of the melody instrument itself).

The chord intervals were also offset by a quaver to give the user time to assess each event in its individual and collective entirety and also provided the fringe benefit of giving more movement to the rhythm of the sonification as a whole. With these basic delimiters in place, the final sonification of the employee database was rendered thus:

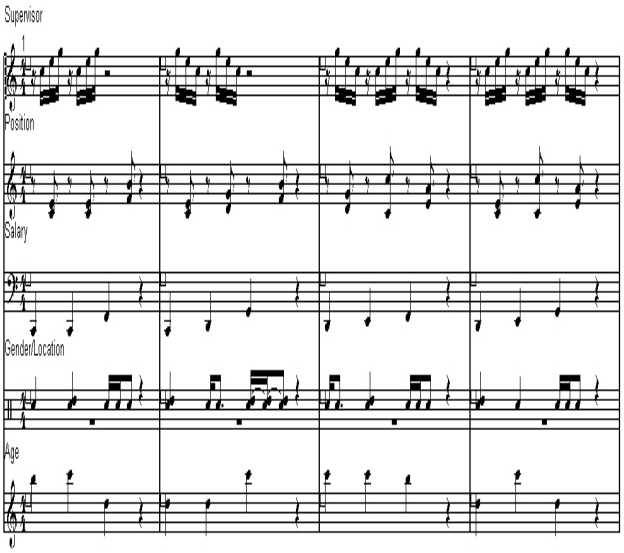


*Figure 7: Final output sonification including rhythmic parsing.*

The existing orchestration details seven potential data elements- employee number, position, salary, supervisor, gender, age and location. As the amount of data in the sonification increases even experienced listeners may develop difficulties in discerning which element is which and how they relate to each other and so the orchestration of the different elements involved begins to take on a greater degree of importance.

## 4. CONCLUSIONS AND FUTURE WORK

Using the orchestration template for basic sonification, a simple database of employee related information has been sonified in a fairly straightforward and obvious manner. The emphasis is largely on the obvious, as the templates suggested by such basic orchestration techniques will hopefully lend themselves to sonifications that can be easily analysed and understood by any and all regardless of musical knowledge or experience. The principles defined here are intentionally simple in order that they may eventually be utilized within a non-musical analysis framework (i.e. by users who potentially have no musical education or training whatsoever).

It is perceived as vital to translate many of the accomplished methods used in sonification into a framework that can be manipulated and hence understood at a general level so that it becomes a more integrated part of data analysis in general. To this end, a new application is being developed in the spirit of the existing DNASon software that will ideally deal with business data of various formats. It is hoped that eventually an application can be developed that takes fairly basic data sets and sonifys them in a similarly obvious manner for the purposes of analysis and hence development.

The DNASon software defined drag and drop GUI techniques that allowed users to allocate the various musical parameters required for the output sonification in as straightforward a manner as possible. It is intended that the development of an application for database sonification should follow similar lines as regards it user interface. It was found that the more manipulable and visual the user interface was the more transparent the underlying audio definitions became. Many test users who had no previous musical knowledge showed a greater understanding of chord interval assignment when this function was implemented as a drag and drop interface rather than a pull down menu (as indeed was the bass line pattern).

The development of future software will be within the C++ programming environment in an attempt to create more robust and useable applications than are possible with Visual Basic. After much initial testing with Visual Basic routines it was found that the software proved unstable in live user situations- even after the provision of the Flash front-end GUI designed to take some of the weight off the VB program itself. To this end new routines are being developed in C++ in an attempt to minimize these factors where possible.

It is eventually hoped to produce a basic application that will sonify data provided within a standard delimited database format. This application will ideally provide the basis for more investigation into the orchestration and rhythmic parsing of sonification data, as the initial work suggests that much could be obtained from such an endeavour.

## 5. REFERENCES

[1] Orchestration, Walter Piston, W.W Norton & Company, ISBN 0-393-09740-4

[2] Rhythmic Parsing of Sonified DNA and RNA Sequences, Charlie Cullen and Eugene Coyle, ISSC2003 Limerick.

[3] Database Systems 3rd Edition (pg79), Thomas Connolly and Carolyn Begg, Addison Wesley, ISBN 0-201-70857-4

[4] The New Grove Dictionary of Music and Musicians 2nd edition volume 12, Macmillan Publishers ltd 2001, ISBN 0-333-60800-3

[5] The AB Guide to Music Theory Parts 1&2, Eric Taylor, ABRSM Publishing, ISBN 1-85472-446-0

## 6. APPENDIX- FICTIONAL EMPLOYEE DATABASE AS USED IN TEST SONIFICATIONS

| E.No | 1st Name | 2nd Name | Position | Salary | Svisor | Gender | Age | Location |
|---|---|---|---|---|---|---|---|---|
| 1 | Joe | Soap | Shop Floor | 12,000 | J Doe | Male | 43 | 1st Floor |
| 2 | Mary | Soap | Shop Floor | 14,000 | J Doe | Female | 37 | 1st Floor |
| 3 | John | Doe | Management | 28,000 | | Male | 29 | 3rd Floor |
| 4 | Jane | Doe | Shop Floor | 11,000 | S Watt | Female | 26 | 1st Floor |
| 5 | James | Watt | Stock Control | 18,000 | S Watt | Male | 28 | 2nd Floor |
| 6 | Susan | Watt | Management | 33,000 | | Female | 31 | 3rd Floor |
| 7 | Tony | Curtis | Stock Control | 17,000 | S Watt | Male | 38 | 2nd Floor |
| 8 | Julia | Curtis | Purchasing | 22,000 | J Doe | Female | 37 | 1st Floor |
| 9 | Robert | Sands | Development | 26,000 | S Watt | Male | 41 | 3rd Floor |
| 10 | Donna | Rodgers | Shop Floor | 15,000 | J Doe | Female | 22 | 1st Floor |
| 11 | Tony | Blair | Purchasing | 24,000 | J Doe | Male | 36 | 1st Floor |
| 12 | Bertie | Ahern | Development | 31,000 | S Watt | Male | 27 | 3rd Floor |
| 13 | Virginia | Wolf | Management | 24,000 | | Female | 32 | 3rd Floor |
| 14 | Gore | Vidal | Stock Control | 27,000 | S Watt | Male | 46 | 2nd Floor |
| 15 | John | Lennon | Purchasing | 34,000 | J Doe | Male | 29 | 1st Floor |
| 16 | Aretha | Franklin | Stock Control | 29,000 | S Watt | Female | 21 | 2nd Floor |