Style Analysis for Folk Melodies, with Classification using Inductive Learning

1. J.A. Carter 2. M.R. Brown 3. B. Eaglestone

1 & 2
University of Derby, UK
School of Maths & Computing
Kedleston Rd.
Derby DE22 1GB, UK
Email: j.a.carter@derby.ac.uk,
m.r.brown@derby.ac.uk

3.
University of Sheffield
School of Information Studies
S10 2TN, UK
Email: b.eaglestone@sheffield.ac.uk

Key Words: Music analysis, machine learning, inductive learning, classification.

Abstract

This paper addresses the area of style classification of folk melodies. The study undertaken consisted of the analysis of a series of folk melodies with different cultural backgrounds. The method of analysis used was that of Thomas O'Canainn (1993), which represents a statistical approach.

Thirty sixteen bar melodies were used for the study, fifteen of which were American and the other fifteen Irish. O'Canainn's analysis was carried out on all of the melodies enabling a set of ten attributes for each melody to be derived. This data was then passed through Quinlan's (1993, 1998) C5 algorithm for machine learning, to assess the suitability of the attributes as means for classification of the melodies.

The results were compared with an earlier study (Carter et. al. 1999), where the same melodies were analysed using elements of Lerdahl and Jackendoff's Generative Theory of Tonal Music (GTTM), (1983, 1996) and the results again evaluated using C5.

Both of these studies contribute to a larger project to develop a hybrid cognitive method for the analysis of western traditional folk music, which will be implemented as a computer based tool using an Artificial Intelligence approach. The tool will aid in the identification of the most important features that characterise a set of tunes with a particular style or cultural background.

1. Introduction

As a fiddle player I am aware that I am able to identify the cultural origin of many tunes fairly accurately, and that it is possible to improvise and develop further melodies of a particular style, or in some cases to play a tune of one culture in the style of another. This ability has developed through listening to and playing music over a long period of time, but with only little formal study of musical theory. As a computer scientist with a particular interest in Artificial Intelligence (AI), I have decided to attempt to replicate this ability as closely as possible using a computer.

AI techniques have been used successfully for a variety of musical applications (for example: Cope 1991, 1996; Robbie & Smaill 1995; Horowitz 1995; Wiggins & Smaill. 1998). It is the (long term) intention to use AI techniques in order to implement an automatic cognitive hybrid system for the analysis of folk melodies. AI as a study and simulation of intelligent behaviour, lends itself to problems that require a more human approach. Wiggins & Smaill, (1998) write,

> *'While AI itself will never be able to solve these problems [problems that require an intuitive approach] in general, its techniques are often able to go further than the standard 'unintelligent' approaches.'*

The employment of such techniques often goes alongside an interest in the *way* that a problem is solved, as well as achieving a desired result. It is not surprising then that the study of AI overlaps with many other disciplines, for example, psychology, linguistics, and music.

The value of taking an AI approach is multi-faceted; it can enable us to achieve a better understanding (by simulation) of human intelligence (Searle 1997, 1999); and conversely, attempting to automate human intelligence could help us develop better machines (Cawsey, 1998). From a musical perspective, Robert Rowe (1995) illustrates the two fold gains in AI/music research in the following two quotes:

> *'From the point of view of artificial intelligence research, such [AI/music] applications [he cites Cope 1991, Ebcioglu 1992, Barucha & Todd 1989] are attractive because of music's rich capacity to support a wide variety of models';*

> *'From the point of view of computer music composition and performance, artificial intelligence is of central importance because it directly addresses the modelling of human cognition'.*

The work carried out for this paper is one of a series of preliminary studies marking the beginning the identified larger project. The main aim of the larger project is to develop a hybrid cognitive method for the analysis (using both structural and performance information) of western traditional folk music, and then to verify appropriate AI techniques as means of automating the analysis process. The automated analysis tool will enable the identification of the most important features that characterise a set of tunes with a particular style or cultural background.

The study described in this paper uses a statistical approach as identified by O'Canainn (1993) to analyse two sets of 16 bar fiddle tunes (one set being Irish and the other American). The results of the analysis provide a means of deriving a set of attributes (initially ten) that describe certain characteristics of the melodies. The fiddle tunes selected were all in common time, in order to minimise the variables.

In order to determine which attributes characterise the sets of melodies effectively, a machine learning technique known as inductive learning was used. The software used for this purpose is Ross Quinlan's C5 algorithm (1998), and given a training set it builds a decision tree that is then used to classify further examples. The software provides statistics that indicate the accuracy of the classifications made, as well as allowing the user to establish which of the given attributes were key to the classification process.

The results of the experiment are compared with an earlier work (Carter et. al. 1999) where the same melodies were analysed using the grouping and metrical components of Lerdahl and Jackendoff's GTTM (1983,1996). The attributes derived from this analysis process were similarly passed through Quinlan's C5 algorithm to enable identification of the key characteristics of the melodies, with respect to cultural background.

The results are discussed, and in the light of these, recommendations for further study are made.

2. The melodies

The tunes used for the experiment are all 16 bars long and are in common time. Half of the tunes are Irish, and are taken from ‘O’Neill’s Music of Ireland’, Miles Krassen (1976); the other half are American and are taken from ‘Fiddle Case Tune Book: Old Time Southern’, Stacey Philips (1989). Both sources are written for the fiddle and include typical performance information (slurring/phrasing, ornamentation), though both acknowledge that individual performers may vary the way that they play a particular tune.

3. O’Canainn’s method of analysis

This is an example of a fairly simple analysis process designed for work with folk melodies. O’Canainn applied the method to a series of melodies, showing that in Irish dance tunes there is a tendency to ‘concentrate on only a few notes of the available scale, and to return to these again and again throughout the tune’. His method is based on the idea of note frequency. Points are allocated to notes for the reasons indicated below:

- each occurrence of a note (from any octave)
- any note falling on a strong beat
- the highest note on its first appearance
- the lowest note on its first appearance
- any note proceeded to by an interval greater than a fifth
- the first stressed note
- a long note (with respect to the modal note length)

His own analyses showed that the most important notes resulting from this analysis were not necessarily the same as those that would result from a more formal analysis of key signature. The notes with the maximum and second maximum frequency of points were not necessarily the tonic and the dominant as might be expected, and hence he concluded that the tunes examined had complex tonality.

4. Machine Learning

Machine Learning is about the development of computer programs that improve automatically with experience. Various types of learning have been defined, the simplest being rote learning. Most of the work on machine learning has centred on inductive learning, or learning from example. Examples of such methods include Winston’s learning program (1975), version space learning, and decision tree induction. There are a number decision tree learning algorithms available; in particular those of Quinlan (ID3, 1986; C4.5, 1993; C5, 1998) have been used for a wide range of classification problems.

Other approaches to machine learning include Genetic Algorithms, which are influenced by evolutionary theories and can be used for a variety of search related problems; and Neural Networks, which are modelled on the structure of the brain. These provide alternative yet equally proven methods, but have not been used for this study.

4.1Quinlan's Algorithms for Inductive Learning

Quinlan's work on ID3, C4.5 and more recently C5, has made substantial contributions to machine learning and knowledge-based systems. The process of induction allows a set of examples to be used in order to create a decision tree, which can also be represented as a rule set. The algorithm achieves this by discovering and analysing patterns found in sets of data. This process is measuring the information needed to identify the class of cases provided in the training set, and such a measure is known in information theory as entropy.

The data presented to the algorithm for classification purposes must have the following characteristics:

- There must be a fixed collection of attributes or properties that describe the object or case, and these may have either discrete or numeric values.

- The classes must be pre-defined as this is supervised learning.

- Each case must belong to one class or another, and there should be substantially more cases than classes.

- There must be enough data for patterns to be established. The amount required varies depending on the complexity of the classification task.

(Quinlan, 1993)

The algorithm needs to have a training set in order to build the classifier, or decision tree. It is then able to make classification predictions on further data sets based on the tree it has already built. There are a number of ways of using the software to give more accurate predictions without increasing the available number of cases. One of these is known as cross-validation trials. This option uses the complete set of cases, and from this, takes a number of different samples as training sets. It builds a new tree each time, and amalgamates the output to give an overall result based on all of the trees. This means that every case is used at some time both as training data, and as sample data for classification.

Another method that improves the prediction accuracy is that of boosting. This is a feature of C5, the most recent version. Boosting also requires a number of classifiers to be built. Each time a new one is constructed, the algorithm concentrates on the cases that it predicted wrongly in the previous attempt. When all the trials are complete, each classifier votes for its predicted class. To determine the final class the votes are counted and the result is often significantly better than any of those from the individual trials. (Quinlan, 1998).

5. Method

Each of the thirty melodies was analysed according to the rules outlined in **section 3**. The number of points for each note in every melody was then used to derive a series of attributes that describe features of the melody. The attributes were derived as follows:

- The note with the highest number of points
- The note with second highest number of points
- The note with the lowest number of points
- The note on which the first stressed note falls
- Percentage of points allocated to the key signature tonic
- Percentage of points allocated to the key signature dominant
- Mean number of points per note
- Standard deviation of points for each note
- Percentage of points allocated for a long note
- Percentage of points allocated to a note proceeded to by a large interval (fifth or more)

('Note' as termed above refers to the pitch of a note within a scale, but a note can be from any octave. E.g. an occurrence of middle C and of the C above, would be counted as 2 occurrences of the note C.)

Values for the first four attributes were described using the standard naming for notes of a diatonic scale (i.e. tonic, supertonic, medient, subdominant, dominant, submedient, leading note). The remaining attributes all had numeric (and continuous) values.

The results of this analysis were then used as input for the C5 decision tree induction software. The software provides ways of improving the reliability of the results for small sample sizes such as this (fifteen for each class). The boosting option (see section 4) is an example of this and was used for each test. Reliability of the results is also improved by making the training set as large as possible which causes the test data to be small, but the data can be passed through the algorithm a number of times and summarised as compensation for this.Tests were also carried out using cross validation trials. Rulesets were generated for all trials.

6. Results

The results were positive and showed that the melodies could be classified according to cultural background using the attributes from the analysis. Use of the boost option gave the best results in terms of correct predictions. The most prominent features identified as contributors to successful classification are listed below, with references to the associated trees and rules generated by C5. A number of trials were carried out and the trees shown below are given as typical examples.

Characteristics of Irish melodies:

- Fewer points allocated for long notes. (see Tree 1)
- Higher mean number of points per note. (see Tree 3)
- Lower percentage of points on key signature tonic & key signature dominant. (see Trees 3 & 4)
- Second maximum number of points were allocated mainly to either the supertonic, the subdominant, or the leading note. (see Tree 2)
- Lower standard deviation of points per note.

Characteristics of American melodies

- More points allocated for long notes. (see Tree 1)
- Lower mean number of points per note. (see Tree 3)
- Higher percentage of points on key signature tonic & key signature dominant. (see Trees 3 & 4)

- Second maximum number of points were allocated mainly to either the tonic, the medient, or the submedient. (see Tree 2)
- Higher standard deviation of points per note.

Tree 1

Decision tree:

```
Percentage of points for long note <= 5.9: irish (17.0/3.0)
Percentage of points for long note > 5.9: american (13.0/1.0)
```

The numbers at the end of each leaf node (e.g. 17.0/3.0) indicate the number of cases mapped to the leaf (17 on the first leaf), and the number classified incorrectly by it (3 on the first leaf). Some cases may only be partially mapped to a particular leaf node, this results in some of these values being non-whole numbers.

Extracted rules:

```
Rule 0/1: (cover 17)
      Percentage of points for long note <= 5.9
      ->  class irish  [0.789]

Rule 0/2: (cover 13)
      Percentage of points for long note > 5.9
      ->  class american  [0.867]
```

The number in square brackets at the end of a rule indicates the confidence with which the prediction is made.

Tree 2

Decision tree:

```
second max points on = tonic: american (3.2)
second max points on = supert: irish (3.9)
second max points on = med: american (8.7/1.6)
second max points on = subdom: irish (1.6/0.8)
second max points on = submed: american (4.0/0.8)
second max points on = LN: irish (4.7)
second max points on = dom:
:...Percentage of points for long note <= 5.9: irish (1.6)
    Percentage of points for long note > 5.9: american (2.4)
```

Tree 3

Decision tree:

```
Mean <= 21.57: american (10.5/2.6)
```

```
Mean > 21.57:
:...Percentage of points for long note <= 5.9: irish (11.7)
    Percentage of points for long note > 5.9:
    :...percentage on Key Sig Dominant <= 14.29: irish (1.5)
        percentage on Key Sig Dominant > 14.29: american (6.3)
```

Tree 4

Decision tree:

```
Percentage of points for long note > 5.9: american (11.9/0.8)
Percentage of points for long note <= 5.9:
:...percentage on Key Sig tonic <= 28.19: irish (11.9)
    percentage on Key Sig tonic > 28.19:
    :...Mean <= 21.57: american (3.9)
        Mean > 21.57: irish (2.2)
```

Extracted rules:

```
Rule 4/1: (cover 11.9)
     percentage on Key Sig tonic <= 28.19
     Percentage of points for long note <= 5.9
     ->  class irish  [0.928]

Rule 4/2: (cover 7.3)
     Mean > 21.57
     Percentage of points for long note <= 5.9
     ->  class irish  [0.892]

Rule 4/3: (cover 11.9)
     Percentage of points for long note > 5.9
     ->  class american  [0.868]

Rule 4/4: (cover 3.9)
     percentage on Key Sig tonic > 28.19
     Mean <= 21.57
     ->  class american  [0.830]
```

7. Discussion

The most important characteristic was the number of points allocated due to the note being longer than the modal length. This is not surprising, since an examination of the scores shows that the Irish melodies are often a stream of notes of equal length, with longer notes appearing mainly at the end of a section. Many of the American melodies however have more variation in the rhythm.

This could also account for the variation in the mean number of points per note (higher for Irish), since the note frequency will be greater if there are less long notes. Using the mean as an attribute is probably not so helpful in retrospect since it could be hiding more significant features.

The American tunes tended to have more points on the key signature tonic and dominant than the Irish tunes, suggesting a less complex tonality (as defined by O'Canainn, 1992).

The standard deviation of points per note appeared in some trees as a means of classification, but these were relatively few and resulted in a high proportion of errors when used to classify. Nevertheless this suggests further study of the range and distribution of notes within a melody. The occurrences of notes within its specific octave could also add more value to this.

A number of trees attached some importance to the second maximum number of points awarded, (supertonic, subdominant, leading note for Irish melodies; tonic, medient, submedient for American melodies). For a tune without complex tonality the expected result for this would be the dominant, so this shows an interesting feature but is not specific enough as it stands to be very useful. There was no pattern to the note for which the maximum points were awarded, although for all of the tunes there was a tendency for this not to be the tonic (which would be the expected result). This tendency was greater for the Irish melodies.

7.1 Comparison with earlier work

An earlier study of the same melodies using the grouping and metrical components of Lerdahl and Jackendoff's GTTM (Carter et. al. 1999) showed that it was possible to classify the tunes successfully according to the positioning of group boundaries and consideration of metrical information. The analysis was carried out using the well-formedness and preference rules.

The analysis using O'Canainn's method seems to give equally good results in terms of correct classifications, however it is more difficult to interpret the reasons.

Some of the rules in GTTM overlap with the reasons for awarding points in O'Canainn's method (e.g. a group boundary is created in GTTM for large intervals, and for long notes; in O'Canainn's method points are allocated for intervals greater than a fifth, and also for long notes).

The key features identified by GTTM were as follows: deviations from metrical structure (caused by frequent occurrence of triplets in Irish tunes); and group boundaries due to the position of phrase endings and beginnings.

O'Canainn's method does pick out different features to those identified by GTTM. For example, GTTM didn't pick out the more frequent occurrence of long notes in American tunes even though there is a rule saying a group boundary occurs where there is a longer gap between consecutive notes. This is probably because the long notes in the American tunes appear in batches close together and another preference rule in GTTM says that small groups must be avoided, thus long notes did not necessarily cause a new group boundary. Occurrence of significantly sized intervals appeared to happen to a similar degree in both types of analysis.

The different levels of complexity in tonality identified by O'Canainn's method were not apparent in the results of the analysis using GTTM.

Although O'Canainn's method is statistical and has no claims as a cognitive model (unlike GTTM), there is evidence to suggest (Jarvinen et. al. 1999, Toiviainen et. al. 1999) that statistical analyses can provide adequate information to enable the classification of musical styles according to their perceptual similarities. As such, any useful features of such methods will be incorporated into the final hybrid cognitive analysis method.

Many of the findings of this study and the earlier work (Carter et. al. 1999) are not really surprising, and are probably less so to the accomplished music theorist. However the purpose of these studies is to find a way of capturing the knowledge and/or 'experience' of a musical listener in such a way that it can be used within a computer system to 'listen' and draw conclusions about the style of the music as effectively as a human listener might.

8. Conclusion

The analysis method described has proved to be useful in a limited way. It does add some information to the previous study using GTTM, though the results are not as specific. The choice of attributes could be improved as described in earlier sections in order to reveal more information. The study has identified some key characteristics of the groups of melodies that were not identified using GTTM in the earlier study.

The results were successful in that the analysis enabled the classification of the melodies, however only a few features were drawn out as being important, and the aim is to achieve an analysis method that provides a full description of the melodies. Further work is required to derive a hybrid method that fully describes the styles in terms of their perceptual similarities. The next stage will be to complete a study using GTTM fully, and then to look at the possible contribution of other existing cognitive approaches to analysis.

9. References

Bharucha, J. & Todd, P.,1989, Modelling the Perception of Tonal Structure with Neural Nets, *Computer Music Journal* 13(4):44-53.

Carter, J., Brown, M. Eaglestone, B., Hodges, R., 1999, Inductive Learning for Musical Style Classification, *Society for Music Perception and Cognition*, conference, 1999.

Cawsey, A., 1998, *The Essence of Artificial Intelligence*, Prentice-Hall.

Cope, D., 1991, *Computers and Musical Style*, Oxford University Press.

Cope, D., 1996, *Experiments in Musical Intelligence*, A-R Editions Inc.

Ebcioglu, K., 1992, An Expert System for Harmonizing Chorales in the Style of J.S.Bach, In M. Balaban, E. Ebcioglu, O. Laske, eds. *Understanding Music with AI*, MIT Press.

Horowitz, D., 1995, Representing Musical Knowledge: Processing melodic Lines in a Jazz Improvisation System, In *proceedings of International Congress in Music & AI*, pp 103-118.

Krassen, M, 1976, *O'Neill's Music of Ireland*, Oak Publications.

Lerdahl, F. & Jackendoff, R., 1983, *A Generative Theory of Tonal Music*, MIT Press, Cambridge, Massachusetts.

Lerdahl, F. & Jackendoff, R., 1996, *A Generative Theory of Tonal Music*, MIT Press, Cambridge, Massachusetts.

Jarvinen, T. Toiviainen, P. Louhiviori, J., 1999, Classification & Categorisation of Musical Styles with Statistical Analysis and Self-Organising Maps, In *proceedings of Artificial Intelligence & Simulation of Behaviour*, 1999.

O'Canainn, T., 1993, *Traditional Music in Ireland*, Ossian.

Phillips, S., 1989, *Fiddle Case Tunebook: Old-Time Southern*, Amsco Publications.

Quinlan, J.R., Kaufmann, 1993, *C4.5: Programs for Machine Learning*, Morgan.

Quinlan, J.R., 1998, Is See5/C5 better than C4.5?, *http://www.rulequest.com/see5-comparison.html, Rulequest Research.*

Robbie, C. & Smaill, A., 1995, Implementing Lerdahl & Jackendoff's Grouping Rules in an Interactive System, In *proceedings of 1995 International Congress in Music & AI*, Edinburgh, pp91-102.

Rowe, R., 1995, Artificial Intelligence and Musical Interaction, In *proceedings of 1995 International Congress in Music & AI*, Edinburgh, pp3-11.

Searle, J., 1997, *The Mystery of Consciousness*, Granta Books.

Toiviainen, P., Jarvinen, T., 1999, Exploring Improvisation Strategies of Jazz Pianists, *Society for Music Perception and Cognition, conference, 1999.*

Wiggins, G., Smaill, A., 1998, Musical Knowledge: What can Artificial Intelligence bring to the Musician?, *Department of AI Research Paper, No. 903*, University of Edinburgh.