Title	Algorithmic categorisation in formal music analysis
Author	Anagnostopoulou, Christina.
Qualification	PhD
Year	2001

Algorithmic Categorisation in Formal Music Analysis

Christina E. Anagnostopoulou

Ph.D.

Faculty of Music

University of Edinburgh

2001

Abstract

A new method of formal music analysis is presented, *Categorisation Analysis of Music*, which is based on Paradigmatic Analysis of music, as devised by Nattiez. The new method, which is independent of musical style, is applied to three pieces of diverse character: a Scottish folk tune, a solo flute piece by Debussy and a piece from a piano sonata by Boulez.

As in Paradigmatic Analysis, the analyst produces a classification of the segments of a piece of music. However, the new method is set in a more formal framework in which the individual steps of the analysis are clearly delineated and the criteria for the classification are explicitly defined. Furthermore, in extension to Paradigmatic Analysis, the resulting classification is hierarchical, and new findings from categorisation theory are brought in to enhance the methodology and to act as a bridge to cognitive modelling of categorisation in music analysis. The new method is computationally modelled with an unsupervised neural network algorithm, thereby further formalising the classification process.

The rationale behind this new approach is to allow for a formal analysis without restricting the analyst's freedom of choice: the method acts as a framework for the analysis, making explicit previously intuitive decisions of the analyst, while the analyst remains free to choose his/her own analytical criteria.

The thesis is divided into two parts: the first part provides necessary background and describes the new method in detail in musical terms. An example analysis is demonstrated for a Scottish folk tune. The second part describes the computational model of the method. The purpose of the computational analyses is two-fold: first, to produce entirely new analyses, and second, to reproduce human analyses in order to find the criteria on which the analyst based his/her decisions. Two more pieces of different character are analysed using the computational model: *Syrinx* by Debussy, which has previously been analysed by Nattiez in the framework of Paradigmatic Analysis, and *Parenthèse* by Boulez. The thesis concludes with an evaluation of the method and a discussion of the extent to which it is possible and useful to be formal in music analysis.

Acknowledgements

It has been my good fortune to have had Raymond Monelle, Alan Smaill and Nigel Osborne as my supervisors. They have been helpful, inspiring and patient during these years; without their comments, advice and support, this thesis would have not been produced.

Raymond has been supportive and understanding during difficult times, not only with discussions and advice of the academic type, but also general philosophical observations. His book *Linguistics and Semiotics in Music* has been an inspirational source for the ideas developed in this thesis. Alan helped with the computational as much as with the musical part, and Nigel brought his enthusiasm and general knowledge to the difficult final stage, helping me to find the motivation to complete the writing-up.

This thesis would also have been impossible without the support and help of Gert Westermann. In particular, the paper I have written collaboratively with him has helped determine the shape of this thesis.

I am also indebted to Karin Höthker, Dominik Hörnel and Darrell Conklin for the many discussions and work together that have influenced various ideas articulated in the computational part. Thank you Darrell, Dominik and Karin for the invaluable help, exciting collaboration, good german chocolate, home-made chocolate cake parcels, Abba and Steps song analyses, fun times writing papers during sleepless nights, and for your friendship.

Thanks also to all the people at the Music Informatics and MMM Group, and especially Peter Nelson; Hugh Trappes-Lomax for useful discussions on linguistic discourse, cohesion and semantics; Mike Ramscar for discussions on similarity and categorisation, and for the best red wine; Alex Lamont, Irène Deliège and Catherine Monelle; Mick Power, Xing-Hui Liang and Bill Donovan.

Thank you Craig, Neil, John, Alexios and Robert for solving various computing problems and allowing me to work in AI; Peter for digging out a workstation; Andrew and everyone at the faculty office; British Foundation of Women Graduates for the scholarship; Helen Pain, Mike Ramscar, Graham Hair and Raymond Monelle for the various teaching jobs.

Thank you Al, Bridget, Cathy, Dimitri, Eleni, Juliana, Maarten, Miguel, Thomas, Victor, Zeta; thank you Volker; and thank you Fred for putting up with me towards the end, for all the help, and for all the muffins.

Thank you Dad and Mum for everything; Sofia, Aliki and Poppy.

Declaration

I have composed this thesis myself and it reports original research that has been conducted by myself unless otherwise indicated.

Edinburgh, 21st August 2001

Christina Anagnostopoulou

Christina E. Anagnostopoulou

Table of Contents

Chapter 1

Introduction

1.1 The Unanswered Question

In cultural studies, questions have traditionally been answered from the perspective of intuition or straightforward historical description. This approach can seem unsatisfactory because answers that have been obtained in such a way cannot be well integrated with each other, and, what is more, such answers are so dependent on the circumstances in which they were given that they are neither reproducible nor do they possess a value that stretches beyond their time. A more formal approach to cultural studies seemed thus desirable.

The first major attempt on such a formalisation emerged in the form of *Structuralism*. Instead of approaching human questions from the perspective of intuition or straightforward historical description, it was possible to turn to a more formal methodology. Such a methodology could then be developed in the study of fields like mythology, kinship structures and ritual. Structuralism was first developed in anthropology by Claude Levi-Strauss and in linguistics, and from there it spread to literary work, film, drama, legal study, history and many other fields.

In music, an obvious structuralist influence has been the publication of the

Fondements d'une Sémiologie de la Musique, by Jean-Jacques Nattiez, in 1975. The idea was that every analytical criterion could be declared before analysis began, and anyone could check the analysis by simply "operating the system". In consequence, it was assumed, music analysis would cease to be a critical study, and would become exact, objective – and closer to science.

However, from the very beginning, Nattiez's "description of the neutral level" (Nattiez, 1975) left much unexplained. In the conclusion to the writer's famous analysis of Varèse's *Density 21.5*, he compared his result to that achieved by James Tenney with a computer (Nattiez, 1982b). The two analyses were significantly different.

Nattiez's approach seemed to indicate that a formal approach to music analysis is not a trivial matter. Initially, a few researchers, concluding that Nattiez was simply not systematic enough, pressed on to formalise further and more consistently and to push subjectivism further and further out of the picture. An example of such an approach, based on logic, has been carried out by Borthwick (1995). But Borthwick's system progressively collapses as his book progresses, and he closes with some brilliant, but entirely intuitive analyses of works by Tippett.

There had been very few useful analyses according to the systematic principles of Nattiez, mainly on monophonic music or song melody, though it was striking that they covered a vast range of music: 16th-century English song, Brahms, Debussy, Varèse, Messiaen, Xenakis. The central unanswered question – how to formalise the material sufficiently for every detail of the analysis to be accounted for – was never really answered. Musicologists, on one hand, moved on to different ways of thinking and abandoned these ideas. On the other hand, there have been various attempts to achieve analyses using computers, but these were of limited musicological scope.

The central topic of this thesis is to develop such a formalised approach that goes beyond Nattiez' attempts in this direction, making explicit and formalising each step of the analytical process. Perhaps the limitation of various previous computational approaches to music analysis has been that they put too much emphasis on an automated process, and too little on the intentions of the analyst while doing the analysis. In this thesis, we are taking a different approach: our formal model serves as a framework for an analysis that is however still done by a human analyst. The analyst is free to make his/her own decisions, but in the formal framework, each of these decisions has to be formulated explicitly, making them consistent and reproducible. This framework allows for a transparent and reproducible analysis of musical pieces; at the same time, it can also be used to "verify" existing analyses that have been carried out in a more traditional, intuitive way. In doing so, it becomes possible to elucidate the intuitions of the analyst, and to examine the consistency of the analysis.

In the rest of this introductory chapter, we first look at music analysis as a discipline (section 1.2), then concentrate on formal music analysis (section 1.3) and especially its methods and techniques. Section 1.4 explains the motivation for the work presented in this thesis, and a brief description of the formal analysis method that is developed here is given in section 1.5.

1.2 Music Analysis: Past and Present

Music Analysis, as part of the broader discipline of musicology, and at the same time a musicological discipline in its own right, is engaged with the investigation, explication and at times evaluation of musical pieces from distinct points of view. There have been many different approaches throughout the history of western music theory, some discursive and interpretative, some more formal, looking at specific aspects of music such as harmony or motivic relations. In the last few decades, two new methodologies have been added to the analysis by the individual analyst: the use of computer and the introduction of psychological experimental methods for the investigation of perception issues and cognitive processes, mainly those related to the listener.

The purposes of music analysis can be diverse. Analysis may be regarded as part of the training of the composer; all early treatises of analysis were announced as composition primers – for example Riepel (1754); Koch (1802); Marx (1841). By the end of the 19th century, there was a demand for "Music Appreciation" guides - books which would help the intelligent listener to get more out of their concerts. This tendency also continued in the program notes of Tovey (1936), written for concerts of Edinburgh's Reid Orchestra between the wars. In the area of education, teachers of music history, harmony, counterpoint and all other music disciplines find analysis an indispensable part of a teaching curriculum. Music analysis is considered to play a major role towards the understanding of structural, textural and other characteristics of a piece or corpus of pieces.

Yet analysis, together with theory, and quite apart from such practical uses, constitutes the identification of a scholarly field. In fact the two are so much joined together that they can be thought of as the two sides of the same coin. Music is music because it has been identified as such by appropriate theory, and analysis is a vital part for this process. This in itself is a motivation for carrying out work in the field. However, no such reason of existence or motivation for music theory and analysis is needed, not even a musical one. As the field exists on a meta-level of music, motivation can be totally intrinsic: music analysis *per se*.

1.2.1 Traditional Music Analysis

We can think of traditional music analysis following two broad trends: the *morphological* and the *discursive*.

Traditionally, music analysis, especially in educational settings, is morphological: it concentrates on the overall form of a piece and defines structural sections and subsections. Because Classical and Romantic music was sectional and tonal, oriented towards structural cadences, it was natural to separate a movement into structural items which acquired traditional titles: first subject, transition, second group, episode, and so on. This is seen in a developed form in the *Einleitung zur Komposition* (Marx, 1841). Observing in particular the music of Mozart and Beethoven, Marx wrote an exemplary description of Sonata Form, which became the basis of most analytical writing until now. The piece of music is split into sections, and these sections are named, usually after the first letters of the alphabet. A section that is considered to be similar to a previous one might get the same letter name, or a variation of a name (such as A and A′).

Harmonic and motivic elements are discovered to support the claims on form. Consequently, a *harmonic analysis* is a specialised way of analysis that deals with the harmonic progressions and harmonic structures in a piece, whereas a *motivic analysis* is concerned with the appearance and variation of motives through a piece of music. These more specialised analyses were usually, but not solely, encountered within the framework of morphological analysis, as being complementary and contributing towards the identification of forms and structures.

Morphological analysis encouraged composers of the time to keep to the same, well-studied, sectional forms, and the evolution of musical forms was very slow until the end of the romantic era.

The second category of traditional music analysis is discursive, or interpretive analysis. This usually takes the form of a long essay that discusses the music in question in a less methodical and rigid way, where the analyst gives his/her own literary interpretation of the piece, often by the use of metaphors to the world outside music, and often in a poetic and romantic manner. Such an example is the eminent work of D.F. Tovey, also mentioned above. The following quote is taken from his *Essays in Musical Analysis*, and is part of the analysis of Beethoven's Violin Concert in D Major, Op. 61:

> *Nothing can be really final in a movement so etheral and so static as this larghetto has been from the outset: there is only one way to prove that the vision is true, and that is to awaken in the light of common day and enjoy that light with the utmost vigour and zest. (...) The violin extemporises a cadenza and plunges into a finale, beginning with one of those drastic rondo themes with which Beethoven loves to shock the superior person (or would if he had time to think of him).* (Tovey, 1936), p.94

The aims of this type of writing, apart from elucidating various morphological aspects of the piece, were to create an atmosphere and mood similar to the one that the piece itself evoked, often by the use of metaphors; to create a fine, eloquent text which, as Schumann once said, should be as elegant as the object it criticises.

Contemporary Music Analysis

The changes brought by the 20th Century in all aspects of social, artistic and intellectual life could not have left music analysis untouched. In recent years, from the seventies onwards, there has been a tendency in music analysis to become more formal and less interpretative. This tendency for formality has not been unique to music analysis: as it has been mentioned at the beginning, many other disciplines started to become more formal and, following the structuralist movement in one respect, attempted to use a more scientific procedure.

Music analysis has been one of the last to catch up on advances that happened in other theoretical fields, such as philosophy, psychology, literature analysis, and anthropology.

Although this tendency in music analysis started late in the 20th century, there has been a striking forerunner as early as the 1920s. The work of the Austrian theorist Heinrich Schenker exhibited an entirely new approach to formal music analysis. In his view, music can be thought of as being multi-layered, from the surface to a deep-structure that most pieces of tonal music share. Schenker's ideas, very progressive for the time, find their equivalent in language only in the fifties and sixties with the work of Chomsky (1965) on generative and universal grammar.

Formal music analysis has therefore built on traditional morphological analysis rather than discursive analysis and there are many principles of morphological analysis that persist in current formal analysis. The next section discusses formal music analysis in more detail.

1.3 Formal Music Analysis

If one would like to name a starting point for formal music analysis, then this could be the publication of the well-known article by Ruwet in 1966, *Méthodes d'analyse en musicologie* (Ruwet, 1966), translated in Ruwet (1987). In this article, Ruwet sets the beginning of a methodology that has been dominant in the field to the present date. However, it was not until Nattiez's *Fondements d'une sémiologie de la musique* in 1975 that music analysis was put onto a more stable philosophical ground, by bringing in the ideas of structuralism, and of the *neutral level* of description as a sequence of structuralism.

The neutral level is one of the three levels that was associated with a work of art, together with the *aesthesic* and the *poietic* levels. Although the definition of these levels came before him, it was Jean Molino (1975) that combined the three together for the analysis of art; Nattiez later named Molino's approach the *tripartition of the levels*.

The neutral level denotes aspects of the work of art itself with no human relations attached to it: the painting, the book, the sound object. The poietic level describes the relation of the artist to his/her work; how he/she perceives and explains it, the influences that society and fellow artists have on him/her in the production of the work. The aesthesic level describes the relation of any perceiver of the work of art; it is concerned with how we perceive and comprehend a piece of art.

Nattiez argues that one should first study the neutral level alone in order to gain objectivity and then, should one wish to do so, relate this to the aesthesic and poietic level. However, his and Molino's views have been attacked by various scholars, who argue that there is no neutral level in any piece of art, since when we look at the neutral level we apply our own preconceptions, ideas and perceptive mechanisms to it, and therefore move onto the aesthesic level. Others claim that the neutral level, even if it exists, carries no interest in itself if it is not connected to those who are related to it, that is, to the artist and the perceivers.

Structuralism, which affected so much of the thinking of the last century, is thought to have had a significant impact on music analysis. Structuralist thinking teaches the breaking of the object of study into units and the study of the relations between these units, how the "system" operates. As such it is a way of comprehending the whole. Traditional morphological music analysis uses these principles too, by breaking a piece into sections and naming the sections according to similarity. However, after the structuralist influence, musical pieces start being studied at a much lower level, that is, much smaller segments and relations between these small segments rather than whole sections. If pieces were not split into segments, various relations between single notes were found, which is again a very low level of analysis. An example of this was Schenkerian Analysis.

Did music analysis need structuralism to reach the state it has reached? In one respect yes, in another no. Music can be seen as a science and as an art at the same time. In analogy to linguistic communication, it can be thought of as both close to language and to literature. As a science, music analysis did not need structuralism to progress; the methods that music analysis uses have mostly been borrowed from the scientific disciplines. As an artistic and cultural medium, it follows the history of other similar disciplines and in that case structuralism has a big part to play. Both ways of seeing music are potent.

1.3.1 Methods and Techniques

Bent (1987), pp.80-81, defines six "methods of operating" for music analysis:

1. reduction technique,
2. comparison and recognition of identity, similarity or common property,
3. segmentation into structural units,
4. search for rules of syntax,
5. counting of features,
6. reading and interpretation of expressive elements, imagery, symbolism.

For our purposes we prefer the term *technique* of operating rather than *method*. Bent uses the term *technique* above only in relation to reduction. We reserve the term *method* for the various types of musical analysis such as Schenkerian, Paradigmatic, and so forth.

Apart from the last technique, the other five are specific to formal analysis. We attempt here a more general classification of the techniques used in music analysis. Various music analysis methods of more or less formal character have been developed over the years. All of them use either or both of two analytic techniques: *grouping* and *categorising*.

Grouping involves the gathering together of several usually adjacent notes, adjacent chords or musical segments into one unit. These units "respect the score": a unit is a note or a small part of the score, and consists of one or more notes, or combinations of notes. The grouping technique comprises Bent's reduction technique and segmentation into structural units and other groupings of notes, such as pitch class sets.

The most obvious kind of grouping is the segmentation of a score. There can be many ways and criteria for segmentation, and this is discussed more in chapter 3. Other, more subtle and refined ways of grouping include Schenkerian Analysis (Forte and Gilbert, 1982) and Lerdahl and Jackendoff style Analysis (Lerdahl and Jackendoff, 1983). Motivic Analysis (Réti, 1962) also groups several notes into motives. Pitch class set Analysis (Forte, 1973) divides notes into groups which can, at a following stage, be assigned pitch class sets.

Categorising involves the classification of the artificial units that are constructed by grouping. The categorisation puts these units into categories according to various criteria, usually based on similarity. Categorising does not respect the score: units from various parts of the score can be thought of as belonging to one category. The most obvious example of categorising is Paradigmatic Analysis.

Categorisation is not a simple process. Bent's "comparison and recognition of identity, similarity or common property" and "counting of features" are all part of the same complex process. In fact, these are not even separable, as will be argued in subsection 2.2.1, and throughout the thesis.

Most methods of analysis make use of both of these techniques. Paradigmatic Analysis is a method that uses segmentation and categorisation. The two are separate processes which can be carried out either in turn or at the same time.

Pitch class set analysis and motivic analysis also use categorising. In Paradigmatic Analysis, the output of the analysis is the categorisation itself. In pitch class set analysis, the categories are known *a priori*, so it is a matter of assigning groups of notes to prime forms of groups that are listed in catalogues. In motivic analysis, one tries to trace motives across bigger pieces of work and see how they develop through the piece. Categorisation in pitch class set analysis is discussed in more detail in the next chapter.

Bent also mentions "search for rules of syntax" as one of the techniques. This is different from the rest of the techniques in that the analyst stops interpreting the score as segments or groups of notes, but as a whole, placing the analytical units and findings into the bigger framework.

1.4 A Call for a New Methodology

Most formal methods of music analysis share a significant complication. Although their aim is to be scientific and to get objective results, the criteria for various analytical decisions are not always explicitly set. This is also the case for Paradigmatic Analysis, where the criteria for segmentation and categorisation are often unclear. Ruwet's attempt at this has been the most successful to date, although his method is not directly applicable to all types of music, but only to a limited set. In Schenkerian Analysis, one is often in doubt about which of the notes might be more important than the others, even though there are rules on the procedure that has to be followed and the criteria to be applied. In pitch class set analysis, the criteria for set membership are very strict, and truly formal. However, the way to look for sets, and how much these sets should expand has not being made fully explicit.

More specifically, methods that use categorising as an analytical technique, like paradigmatic, motivic and pitch class set analysis, all face the problem of *similarity*: how can similarity be defined in music? What are the criteria for deciding if two musical passages are similar, and how can we make this process more clear and explicit?

With these kinds of questions, one has to be cautious in the case of music analysis, unlike sentence-level formal linguistics, and very much like linguistic discourse and text analysis. This is because in music, context is crucial to the comprehension of a piece, especially if one looks at similarity judgements of segments within a piece. Context can be anything from the close context for a segment, ie., the piece, to a higher-level context, which can be the style of the composer, the era, and so on. Two passages of music that can appear very similar in one context, can strike as very different in another. It should be noted here that context can also have the meaning of extra-musical associations and / or cultural settings. We do not deal with this way of interpreting context here.

Therefore, if we want to embrace all music and provide a methodology that can be used independent of musical style, and that will, at the same time, be objective, we have to allow for style differences. In effect, we cannot define any specific criteria for similarity. We cannot, for example, say that any semi-quaver stepwise upward movement is similar to any other semi-quaver stepwise upward movement because whether it is similar or not depends on the context.

However, what we can do, instead of pre-defining such criteria for similarity, is to allow the analyst to define his/her own criteria. The role of the formalisation is then to establish rules that govern the treatment of these criteria and the handling of the similarity judgements according to these criteria.

At the same time we extend Paradigmatic Analysis to obtain more information about the categories of a piece that can be useful especially for comparative analysis, which has been one of the purposes of Paradigmatic Analysis.

The motivation for a new methodology can be summarised to the following points:

- There is a need for a truly formal method of analysis, at least to the degree that a formal method in music analysis can exist. This can be achieved by the existence of common principles for the communication between analysts, and the evaluation of others' work. This is particularly salient in there comparative analysis of various pieces, where, without a uniform methodological setting, any comparison would not be meaningful.
- At the same time, the methodology of the analysis should not restrict the analyst's musical choices - it should rather provide a framework for analysis.
- Similarity and categorisation are two concepts crucial to music analysis, and especially to Paradigmatic Analysis. There is a vast literature on similarity and categorisation stemming from philosophy and psychology. Although this literature has affected the research on music psychology, it has not yet truly reached music analysis.
- A formal model of the methodology proposed here will ensure that the process is transparent and reproducible.
- The model can, in future work, be used towards cognitive modelling of the analytical process in music, and especially of categorisation. Although cognitive modelling has been studied in the area of listening, not much work has been done in the case of the analyst who has the score in

front of him/her. Analysing music, as separated from listening or other musical tasks, is a significant musical task that needs to be further investigated. Although in this thesis the model is not analysed in terms of a cognitive pertinence, some ideas are discussed on how this might be achieved in the final chapter.

It should be noted that, although segmentation is an important part of most analysis methods, it is not dealt with in this thesis. Segmentation is taken as given, and we concentrate in the process of categorisation.

1.5 A new method: the Categorisation Analysis of Music

In this thesis we develop a formal framework for music analysis, *Categorisation Analysis of Music*. This method is related to Paradigmatic Analysis in that the piece of music is first segmented and in that the resulting segments are then categorised. However, there are also significant differences to Paradigmatic Analysis, since the aim of the Categorisation Analysis is to redress the above-mentioned points in traditional Paradigmatic Analysis. The differences are that:

- The criteria for the classification have to be explicitly set before the actual classification takes place.
- The classification is carried out by a computational algorithm.
- The classification is hierarchic, and the relations between classes are made explicit.
- There are probabilistic prototypical values for each class.

The architecture of the methodology is modular (fig. 1.1). Each step of the analysis is realized as an individual component, and for each component the form of the inputs that it receives and of the outputs it produces are clearly specified. The modular character of the system allows for a transparent, step-wise execution of the analysis. Further, the individual components of the system can be substituted as long as they produce the required output forms given the specified input forms. The form of the various results that can be obtained are shown in figure 1.2.

In this thesis, the Categorisation Analysis method is applied successfully to three pieces of music of different character: a Scottish folk tune, a solo flute piece by Debussy and a piece from a piano sonata by Boulez. Out of the various kinds of results, emphasis is given on the hierarchic classification and the relations between classes.

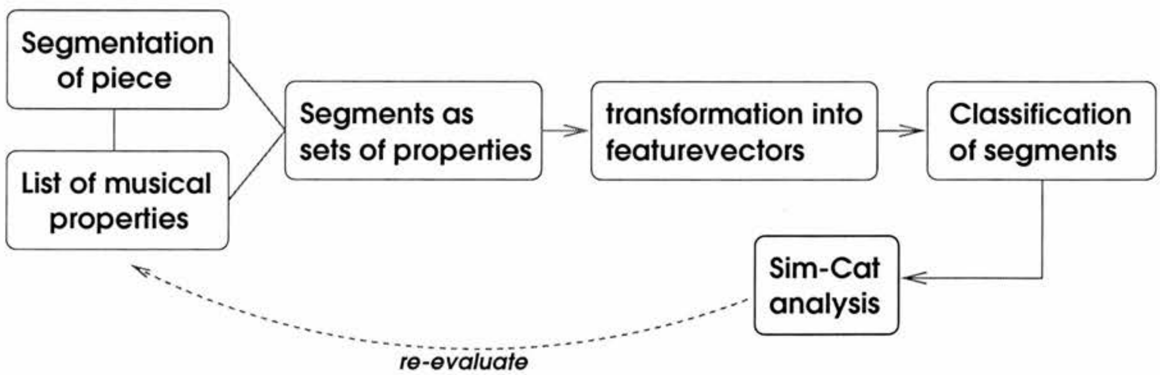

Figure 1.1: A general overview of the system for the Categorisation Analysis of Music. The modular character allows for the substitution of components with equivalent ones. The final analysis, named "Sim-Cat" here, is short for the various results; these are based on similarity and categorisation, and are shown in more detail in figure 1.2.

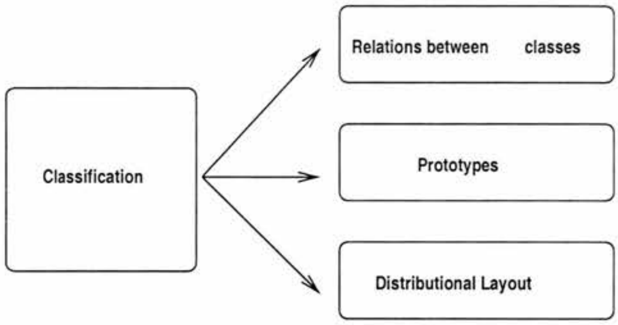

Figure 1.2: The various results in a Categorisation Analysis of Music.

1.6 The Rest of the Thesis

The next chapter explores the background information that is essential for the remainder of the thesis. It contains:

- a more detailed picture of traditional Paradigmatic Analysis, concentrating on Ruwet and Nattiez's ideas,
- tracing the concept of categorisation in pitch class set analysis,
- an introduction to formal and computational modelling, and why this might be beneficial for any theory,
- an introduction to some of the main theories of similarity and categorisation in cognitive psychology and philosophy, and
- a selection of current work on computational music analysis that deals with categorisation.

Chapter 3 describes the Categorisation Analysis method, without making reference to computational modelling.

Chapter 4 gives a sample analysis using the method. The piece is a Scottish folk tune, *All the Blue Bonnets are over the Border*. Although the approach is formal, no use of the computational categorisation is made. The discussion section (4.10) points out the need for a computational algorithm in order to make the method more reliable.

Chapter 5 describes the computational model of the Categorisation Analysis of Music. It explains the collaboration between the analyst and the system, and the active role that the analyst has to play in the analytical process.

In chapter 6, two pieces are analysed: *Parenthèse* by Pierre Boulez, a movement from his 3rd piano sonata, and *Syrinx* by Debussy. The famous second analysis of *Syrinx* by Nattiez (1975) is used as a point of reference. This shows how this system can be used not only to produce new analyses, but also to reproduce existing ones. When an analysis is reproduced, one could argue that the criteria used for achieving the computational analysis can be the same that the human analyst used for his/her analysis. This is one of the claims that follow the formal modelling approach.

Chapter 7 evaluates the method and the analyses that have been presented during the thesis, the advantages and shortcomings, and compares the present system to others.

Chapter 8 gives some conclusive remarks on categorisation in general, and in music analysis in more detail, and suggests diverse directions for further research.

Chapter 2

The background

2.1 Introduction

This chapter discusses the background of the thesis. Since our area is interdisciplinary, drawing from music analysis, computer modelling and to a lesser extent cognitive science, the chapter gives a brief account of all these areas. It also explores some related contemporary work that has been carried out by other researchers. Owing to space limitations, we have been selective on the choice of theories, methods and relevant work presented.

Paradigmatic Analysis has been the main starting point of Categorisation Analysis, and a large section of the chapter (2.2) is devoted to it and its main founders, Ruwet and Nattiez. First, a description of the method is given, followed by the motivation for such kind of analysis. The problems and limitations of Paradigmatic Analysis are discussed at the end of the section. The linguistic parallel is also explored, not only in relation to phonology, where Nattiez reasoned the analogy lies, but also to discourse structure, where we argue for closer parallels.

Other well-established methods of analysis also use categorisation, as was explained in the introductory chapter; the most interesting use of categorisation takes place in pitch class set analysis. The way that this method is based on the categorisation concept is described in section 2.3. In the same section, two other, less used methods of analysis are revised, category and feature analysis. They bear resemblances to our method, especially in their use of attributes in order to describe the music.

The next section (2.4) gives a brief account of various theories of similarity and categorisation in cognitive science, which have their foundations in both philosophy and psychology and that have influenced our work in defining the Categorisation Analysis method.

An introduction to formal and computational modelling follows in section 2.5. Emphasis is given to why modelling is important for a theory, and therefore this section offers a partial explanation of why this thesis takes the modelling approach to music analysis. Knowledge representation and choice of algorithm, two key issues in the area of artificial intelligence, are discussed, including a short introduction to neural networks, the specific type of algorithm used in this work.

Finally, section 2.6 presents a selection of current work on computational music analysis that deals with representation and categorisation.

2.2 Paradigmatic Analysis

This section first provides a description of the method of Paradigmatic Analysis in its traditional form, as introduced principally by Jean-Jacques Nattiez, and then proceeds to explain Nicholas Ruwet's earlier methodology and how this relates to Nattiez' paradigm. The motivation of Paradigmatic Analysis, as claimed by its main founders, is discussed in subsection 2.2.2. Subsection 2.2.4 considers briefly the work of another successful analyst, Marcelle Guertin (1981, 1990). Subsection 2.2.5 explores the linguistic parallel with phonology and then also explores a different parallelism between Paradigmatic Analysis and the discourse level of language.

2.2.1 Description of the method

Paradigmatic Analysis consists of two main tasks: firstly the segmentation of the score of a piece and secondly the categorisation of the resulting segments according to their similarity. There are various constraints on how this process is carried out, and on how the two main tasks are related.

- The *segmentation* of the score depends, to a large extent, on the principle of repetition. Ruwet (1966) has been more explicit than Nattiez on this: there is a segment boundary where a repetition of some previously encountered material starts and where it ends. For material that has been encountered before, the segment boundaries remain the same.

  Ruwet's first attempt at Paradigmatic Analysis required equal length segments. Nattiez extended this and accepted that segments of different length should also be able to be compared and judged as similar. This is since, according to Schönberg, inequality of length might be a defining characteristic (Monelle, 1992, Schönberg, 1967).

- The resulting segments are grouped into categories. The first occurrence of a segment in each class is called the *paradigm*, and subsequent segments are compared to paradigms in order to determine their category membership. The paradigms therefore play the role of class prototypes.

- The *categorisation* of the segments according to similarity is carried out intuitively, and in some cases the reasons why segments have been classified together are made explicit. Sometimes an inventory of some properties of the segments is drawn after the classification has taken place, such as the third analysis of *Syrinx* by Nattiez (1975).

- One important aspect of Paradigmatic Analysis is that the analysis is written in such a way that it is possible to follow the original score: it is written as a table in which each column forms a class as defined by a paradigm. This table is then filled with the segments from left to right, and each new segment is put into the column corresponding to its class paradigm. No segment can be written down in a row before its predecessors, and a new row is started when the column for the class of the present segment has been passed in the current row. Therefore, after the analysis is complete, the score can be read from the table, from left to right, top to bottom.

Paradigmatic Analysis makes the strong interdependence between the individual steps obvious: the segmentation of the piece relies on the occurrence of previously encountered material, but this material need not occur in exactly the same way as before. Therefore, a notion of similarity between segments must exist already at the earliest stages of the analysis, that is, at the time of the creation of the segments themselves. Similarly, the establishment of paradigms requires a global view of similarity between segments: for each new segment, a decision has to be made whether to make it a new paradigm or to group it with an already existing paradigm.

However, the problem remains. What counts as similarity, and what are the criteria for such judgements? How is the segmentation related to the categorisation procedure? This question and more on the scope and limitations of Paradigmatic Analysis will be discussed in the next chapter.

Semiotic analysis and distributional analysis are two terms often used to describe Paradigmatic Analysis. However, these two terms have broader meanings and include other kinds of analysis and concepts that go along with them.

In this thesis, we restrict ourselves to Paradigmatic Analysis and do not discuss semiotic and distributional analysis, or the whole area of semiotics.

2.2.2 The initial motivation and what followed

Ruwet's starting point

The beginning of Paradigmatic Analysis in music came with Nicolas Ruwet's famous article *Méthodes d'Analyse en Musicologie* (Ruwet, 1966). In this article, he describes a method for segmenting a melody which is based on repetition. The purpose is not to come up with a categorisation of segments, but rather to come up with a good and fully explicit segmentation that is based on similarity. It is a "a machine for the discovery of elementary elements".

Ruwet's concern was to provide a scientific procedure that gives an objective analysis, and this is why he names the procedure a "machine". The result of a scientific procedure would be a process which, when followed again in exactly the same way, would give exactly the same results. However, he admits that it would be better used as a process of verification, in order to check existing analyses which were based on intuition, rather than to produce new ones. This is a point that was also made by Zellig Harris (1951), in linguistics, from whom Ruwet was greatly inspired when constructing his method (see section 2.2.5):

> *The procedure is much more a procedure of verification, meant to keep a check that the analysis is coherent, a discovery procedure in the strict sense of the term. Doubtless it will always be possible to apply it rigorously in the given order, and the same results would be obtained, but it is much more economical and quick to use it to verify the results of an analysis obtained sometimes quite rapidly in a purely intuitive manner.* Ruwet (1972) p.117.

Nattiez's motivation

It is often said that Nattiez's main contribution was to link Ruwet's ideas to the concept of the tripartition of levels, and to the neutral level more specifically (for example Monelle, 1992, p.94). This might be the case in terms of the aesthetic issues related to Paradigmatic Analysis; however, Nattiez did substantially more than that in practical terms: he introduced explicitly the concept of categorisation in music analysis, and created the Paradigmatic technique. In the past, the idea of categorisation was never explicit, and was present only insofar as it helped to show something else – in the case of Ruwet, it showed the segmentation; in the case of Seeger (1960), it showed similarity relations between melodic contours.

Nattiez also realised, like Ruwet, that there can be multiple analyses of the same piece and by using the same method, according to different criteria: "what is important in an analysis is to make explicit one's criteria" (Nattiez 1975, p.240 and 340). He aims to describe musical events exhaustively; give an account that can later be interpreted in some way, either by the same analyst, in order to produce some second step of the analysis, or by others, to read and understand. Paradigmatic Analysis can be an interesting and sometimes essential pre-analysis to other kinds of analysis, for example comparative, stylistic and syntagmatic analysis.

It was not surprising that Paradigmatic Analysis originated from an ethnomusicologist since it can be argued that is a particularly suitable method for all kinds of music. However, there have been various oppositions to this, for example Blacking (1981).

Nattiez wants to define Music Theory in its own terms, without the interference of psychological and sociological factors, and not as a unified phenomenon, that is to be the same in the three levels. His aim is an analysis of the neutral level of music. His and Ruwet's intention was to create an "objective", scientific analysis – to give remarks on the thing itself, before interpreting it poietically or aesthesically. Nattiez's idea of separation is to carry out the analysis first, and then to check it against the other two levels, for example, perform psychological experiments, look for evidence from the composer, and so on. There have been many objections to the way Nattiez saw the neutral level of music. However, one cannot deny that, even if his conception of it is debatable, the idea is a very useful one in that it distinguishes the pure musical object from cognitive and perceptive processes. This distinction is noteworthy not only for music, but also for the development of music psychology and cognitive modelling.

2.2.3 Problems and limitations

The role of intuition

Apart from the neutral level, what Paradigmatic Analysis has been mainly criticised for is the difficulty in keeping the procedure formal while performing an analysis, because of the large role of intuition. One of the main opposers to Paradigmatic Analysis has been Nicolas Cook. Below, he cites Nattiez:

> *People decide to associate several units in a single paradigm because of semantic and psychological criteria that they do not express consciously.* Cook (1987), p.180.

Although this is a criticism often heard, one could argue that Cook has made a special case of criticising Paradigmatic Analysis because he also criticised music theory and analysis of this type for not making a psychological account of music. His criticism is again discussed at the conclusions chapter of this thesis.

It is true that intuition plays a large part in this kind of analysis. What is more, it is obvious that no two analyses are going to be the same. The grouping of objects into categories can be performed in many different ways, and it can be argued that there are no right and wrong answers, only answers which are consistent or non-consistent to one's criteria.

The criteria for the categorisation of the segments are usually not made explicit. That means that the reason segments are considered to be similar to each other is not always clearly stated. This is a problem when one analyst tries to understand and evaluate the analysis made by another analyst. If these criteria were made more explicit, then the whole approach would become more acceptable.

Paradigmatic headings

Another problem of Paradigmatic Analysis is the emphasis it places on the first instance of a category segment, the paradigm. Paradigms are supposed to be the most significant occurrence of the category, and this concept can be compared to a prototype. However, the character of music is such that the first occurrence is not always prototypical.

Paradigms as first occurrence might make more sense in the case of listening, where the listener might use this occurrence as a reference point for the other occurrences of a similar motif or segment. However, the analytical process is not incremental like listening, and its purpose is not the link to the aesthesic level of listening, at least with this kind of analysis.

Hierarchic classification

Nattiez's categorisations of segments are problematic in that the relations *between* the categories, or indeed between the segments within a category, are not specified. Out of the resulting classification, two classes might be very similar to each other, that is share a lot of common features, and others might be very different. The same holds for segments within a category: there might be identical segments, slightly varied ones and very different ones in certain respects. Categorising objects into classes without having further sub-categories and hyper-categories, is limited, for this reason. In a hierarchic categorisation, relations between categories, or members within a category, would be shown.

The repertoire

It has been a common criticism to Paradigmatic Analysis that, although the existing analyses cover a large range of epochs and styles, the method is mainly applied to monophonic pieces. Also, as with all methods which require a vertical segmentation, strictly polyphonic music can be problematic because of the segmentation.

2.2.4 Other work on Paradigmatic Analysis

Apart from Nattiez's famous examples of Paradigmatic Analysis, there have been various other attempts, such as Morin (1979), Naud (1975) and Morris (1989). Elisabeth Morin compares two sixteenth-century variation sets on the song *John come kiss me now*. One set of variations is composed by Byrd and the other by Thomkins. These are two very simple songs and Morin makes a separate analysis of the rhythmic and the melodic elements. She also separates the right from the left hand. Her analysis brings out clearly the differences between the two pieces.

However, the most interesting other example of Paradigmatic Analysis, yet again to come from French-speaking Canada, is by Marcelle Guertin (Guertin, 1981, 1990). In principle Guertin is following Nattiez's paradigm, but in practice she is taking the analysis much further, producing valuable results.

Guertin performs a comparative analysis of Debussy's Préludes in order to get some information about the style of the composer. In it, she concentrates on the melodic line of the preludes. By using the Paradigmatic Analysis technique, she notices certain kinds of regularities on the syntactic level of the pieces, and she is thus able to form a theoretical model on the composition of these melodies. Her work is one of the few examples where Paradigmatic Analysis indeed serves one of its original purposes: to aid towards syntagmatic principles and comparative and stylistic analysis.

2.2.5 The linguistic parallel

For certain types of automated musical tasks, it might be convenient to build large musical dictionaries of musical segments, or to rewrite the segments into a specific order that would facilitate further analysis. This kind of analysis emerged first in linguistics, and more specifically in phonological studies with Zellig Harris (1951) and it is called taxonomic or distributional analysis. The tradition was later adopted by musicians, who found this method particularly well-suited to the nature of music, and now it has become a well-known technique in music analysis.

In Harris' distributional analysis, small differences in the sound cause differences in the meaning. Repetition of a sound, or phoneme, gives it a status that makes it a unit. Zellig Harris looked at recurrences of these units. This was his Paradigmatic Analysis on the phonological level of language, which was based on repetition. Monelle observes that

> *The kind of recurrence that structures language happens in music, too; the unification of a long passage by the constant interworking of small motives is a familiar feature. It is unusual in language, however, for items to recur* syntagmatically, *in immediate succession. Phonemes are hardly ever repeated successively in the syntagmatic chain, and successive repetition of morphemes occurs only in rhetorical or poetic utterances. But music, unlike language, often repeats phrases syntagmatically in a very simple and regular way.* Monelle (1992), p.65-66

Harris' work did not have any apparent relation to the purely semantic level of language. However, there has been a considerable amount of work in linguistic discourse and semantics which might be thought closer to Paradigmatic Analysis than distributional analysis. Anagnostopoulou (1997) (see also appendix) draws a parallel between the paradigmatic and syntagmatic level of music and lexical collocation and reference of the linguistic discourse. This point is also discussed in the conclusions chapter of the thesis.

2.3 Categorisation in Pitch Class Set Analysis, Category and Feature Analysis

Of the other standard methods of music analysis, pitch class set analysis makes an interesting use of categorisation, quite different from Paradigmatic Analysis. In this section we look at how the categorisation works in this method, and then briefly move on to two other methods, category and feature analysis.

In pitch class set analysis:

- the objects to be categorised are sets of notes.
- there exist pre-defined categories, each category has a name, and these categories are listed by Forte (1973). These categories exist quite apart from any musical score. There can be various sets of notes that belong to each of the categories.
- There are relations between these categories; some relations are explicit and some are not. For example, the hierarchy of classes (sets, subsets and hyper-sets) is explicit.
- The objects for categorisation are not notes that belong to a set, but complete sets. Complete sets have to fulfill necessary and sufficient conditions for belonging to one of the pre-specified categories.

The analyst has two tasks: The first task is to define the objects that are to be categorised, and the second is to see in which one of the pre-defined categories the objects belong to.

The **first task** is very challenging: the analyst has to decide which notes to group together to form a set. He/she has to decide not just the boundaries of a *string* of notes, but sometimes the task is much more difficult: the notes do not have to be successive, and they do not all have to be part of the same chord. There are no explicit rules on how to choose which notes form a unit, a set. The analysts have to use their experience with the kind of music as well as their musical intuition. There are two heuristic rules that help in this process, both of which have fundamental definition problems:

1. to look for phrase boundaries, chords, motives. Sets can be found horizontally, vertically, diagonally, any selection of notes that are "justifiably" near each other. Notes from a sequence can be omitted. Musical sense has to play a role, especially when analysing atonal music.

2. to look for categories that have already appeared, and try to find sets that would belong to these categories. This is because some kind of repetition is desired, and it could be an indication that a category is the right one if it repeats itself a few times during a piece.

The problem with the second rule is that it is result-oriented. That means that the results already achieved act as a guide to the rest of the analysis.

The **second task** is, having found a set, to define which of the categories it belongs to. This is a very formal and clear procedure that the analyst has to follow. With some easy calculations every set of notes can be reduced to a prime form which gets its name from the list by Forte (1973). This name denotes the category the set belongs to. In a categorisation task of pre-specified categories, this task is equivalent to finding the attributes that describe the specific object and make it a member of some *a priori* category.

Another problem with this method, not an insoluble one though, is that this kind of method restricts itself to the limitations of its description language. The description language of the music is made up of pre-defined categories that Forte has devised. Although the method looks at relations between these categories, it does not look in depth at how various sets of notes are related to each other musically. The common feature here is that both sets can be reduced to the same pitch class set.

Compared to pitch class set analysis, Paradigmatic Analysis has some similar problems. In both cases, the objects that are to be categorised have to be specified. In pitch class set analysis this consists of finding the pitches that form a set, and in Paradigmatic Analysis this consists of finding the pitches that form a segment. This is the grouping task described in chapter 1. In the case of Paradigmatic Analysis, the task is more acceptable because segments are also segments of time during the music, that is successive notes on a score.

In Category and Feature Analysis:

Category analysis and feature analysis are two very similar methods of analysis. Category analysis was mainly developed by LaRue (1992). He analyses music according to 5 categories, namely Sound, Harmony, Melody, Rhythm, Growth, and each of these is divided to sub-categories. For example, melody is divided into range, motion, patterns, and so on. He uses three values for these categories: Large, Middle and Small, which he calls dimensions. His analyses are tables of categories and dimensions.

Feature analysis uses a set of features, such as a particular chord, or interval, or texture, or even a whole pattern, and counts the instances of these features through the piece or set of pieces. This is used in categorising works, usually by creating dendrograms in terms of "affinity" (Lincoln, 1970).

2.4 The Concepts of Similarity and Categorisation

> *Funes remembered not only every leaf of every tree of every wood, but also every one of the times he had perceived it ... Not only was it impossible for him to comprehend that the generic symbol dog embraces so unlike individuals of diverse size and form; it bothered him that the dog at three fourteen (seen from the side) should have the same name as the dog at three fifteen (seen from the front).* Borges (1964), quoted in Eysenck and Keane (1998).

Human beings are not like Funes because they are capable of organising their knowledge and experiences more economically into more general abstract categories of concepts. Once these concepts have been formed, they can be structurally and often hierarchically related (Collins and Quillian, 1969; Luger and Stubblefield, 1998). This is one of the most important characteristics of human knowledge.

In this section we discuss a selection of the main theories of categorisation in philosophy and cognitive science which are relevant and have influenced the methodological decisions taken for the development of the categorisation method. Finally, we briefly mention some work on music similarity that has been carried out, not in the area of music analysis and computational approaches, as these are covered by other sections, but in music psychology, the area where most work on musical similarity has taken place.

Before getting into the subject of categorisation, it is worth clarifying three terms in the literature: *property*, *feature* and *attribute*. Following Eysenck and Keane (1998), the terms are going to be used interchangeably. However, there are slight differences between them and it seems appropriate when talking about the knowledge representation issue to make the distinction. The generic term *property* refers to any predicate that can be asserted of some or all of the members of a category (Mechelen *et al.* (1993), p.15). The term *attribute* refers to a specific type of property: a property that has a number of mutually exclusive alternative possibilities, termed its values. For example, colour can be thought of as an attribute, where possible values can be red, green, and so forth. The term *feature* is found earlier in literature (Mechelen *et al.* (1993) p.15) in structural linguistics, and also in the artificial intelligence literature, and in particular in neural networks. A *feature* is a special form of *attribute*: it can be present or not present. For example, in describing a musical segment, a possible feature is the presence or absence of a musical ornament. Katz and Fodor (1963) use the term extensively in *Structure of a Semantic Theory*.

The term *attribute* is generally preferred in psychology, as a more general one, whereas *feature* is preferred in computational disciplines. This is because in computer science, attributes are often turned into features with binary values using a procedure that is described in chapter 5.

The classical view: Defining-attribute theories

The traditional theory of what makes a category, stretching back as far as Aristotle, is the *Defining-Attribute Theory*. The idea here is that a category can be defined by a set of necessary and sufficient features. All objects having these features belong to the category, and all other objects do not. For example, the definition of a square is an object with four sides of equal length, and right angles. Every object that has both of these properties is a square, and no other object is.

The defining-attribute view has been very influential both in philosophy and in psychology. In philosophy, it was elaborated by Frege at the end of the nineteenth century as one of the basics of formal logic. Frege (1952) distinguished between the *intension* and the *extension* of a concept: the intension gives the list of the defining attributes that an object needs in order to be a member of a category, and the extension is the set of entities that are members of this concept. In psychology, and science in general, this view has formed the basis for many theories of human processing – for example Smith and Medin (1981), Medin and Smith (1984) –, and it has been the foundation of the whole of logic-based artificial intelligence. Here, categories are represented in terms of their features, and they can be related to each other hierarchically. For example, the category "square" is a sub-category of "quadrilateral" from which it inherits all attributes (i.e., "has four sides of equal length"). The attribute "has right angles" is specific to the square and distinguishes it from the other sub-category of "quadrilateral", "rhombus". In summary, the theory says that each entity has to have a singly necessary and jointly sufficient set of attributes in order to be an instance of a specific category. There are clear boundaries between what belongs and what does not belong to a category, and between one category and another category. All instances are equally representative of a category, and there are no prototypes.

At the same time, concepts (or categories) can be hierarchically organised. That means that categories can be combined to create super-categories, and categories can be split into sub-categories. Collins and Quillian (1969) describe this notion by arranging concepts into *semantic networks*, and the way this hierarchical process takes place is explained in chapter 5, section 5.4.1.

Many aspects of the defining-attribute view have been challenged by philosophical and psychological arguments. The most severe problem with the defining-attribute view is that experimental evidence has shown that there simply are no necessary and sufficient features that define category membership. In fact, even before experimental evidence proved this, Wittgenstein had raised this point in his theory of family resemblance:

> *Consider for example the proceedings we call games. I mean board-games, card-games, ball-games, Olympic-games, and so on. What is common to them all? [...] For, if you look at them you will not see something that is common to them all, but similarities, relationships, and a whole series of them at that. [...] similarities crop up and disappear. And the result of this examination is: we see a complicated network of similarities over-lapping*

and criss-crossing: sometimes overall similarities, sometimes similarities of detail. Wittgenstein (1953) p31-32.

Other research has addressed the implications of the defining-attribute theory and has shown them to be not consistent with the psychological evidence. Conrad (1972) suggested that some attributes are more salient than others. For example, the fact that a salmon is pink might be more relevant than the fact that it has fins. Other research has suggested that categories are less clear-cut than predicted by the theory, and, most importantly, that categories possess an internal structure: some members are judged as more representative than others (Rosch 1973). For example, a British person is less representative of a European than a French person, or a robin is a more typical bird than an ostrich.

Various refinements of this theory, taking into account the above problems, resulted into new theories and models, like the feature-comparison theory, which claims that there are two kinds of features, the defining ones and the characteristic ones. The characteristic attributes can be used to evaluate how typical a member is of its category. However, none of these refinements have been able to address all criticisms, and a very different approach emerged in the form of prototype theories.

2.4.0.1 Prototype theories

Prototype theories approach the problem of categorisation from a different angle. The intension of a category is not defined by a set of defining attributes, but with a prototype which lies in the "central space" of this category's concept.

Different variants of prototype theories define the prototypes in different ways: sometimes, a prototype is seen as a set of characteristic attributes for a concept, for example Rosch (1978). These attributes are not shared by all members of the category, there might not even be a single object that possesses them all, but the more of the attributes an object has, the more typical it is for the concept. By contrast, another variant of prototype theories, which are also called *exemplar theories* argue for the prototype to be the best existing example of the concept, and membership in the category is defined by similarity to this exemplar (Medin and Shaffer, 1978).

Prototype theories can overcome many of the problems faced by defining-attribute theories: there is no delimiting set of necessary and sufficient conditions for membership in a category; instead, members can show family resemblances by just sharing some of the features. Further, a concept displays an internal grading: the most typical member is the prototype, and typicality of the other members is defined via the similarity to that prototype. As a consequence, category boundaries can be fuzzy: an object might be located just between two prototypes and might thus be considered a member of both categories.

A sub-branch of prototype theories are conceptual hierarchies in prototype theory (Rosch *et al.*, 1976), which combine prototype theory and Collins and Quillian's work on hierarchies (Collins and Quillian, 1969). Categories are assumed to be organised on three levels: the basic level, the superordinate level, and the subordinate level. Basic level categories (e.g., chairs, beer) have a minimal overlap in their attributes, that is, they have maximal inter-category difference and they have a maximal intra-category similarity. The basic level is the most "natural" to humans, where they would instantly categorise concepts. This level is not fixed but its position can change according to individual differences, expertise and cultural exposure. Objects at the subordinate level (e.g., one of my kitchen chairs, a pint of lager) are characterised by more attributes that overlap between categories. At the superordinate level (furniture, drinks), few attributes apply to the members of a concept and the description on that level is therefore not very informative.

Despite their explanatory power, prototype theories also do not come without problems. The most important one is that relations between attributes are not discussed, although this information can be useful in the categorisation process. Experimental evidence shows that people rely more on relations between attributes than on attributes themselves, especially in cases with large data sets (Medin *et al.*, 1990; Goldstone *et al.*, 1991).

Another problem of prototype theories is their poor definition of similarity relations and inadequacy to account for category membership in certain cases. Barsalou (1989) shows some examples of *ad hoc* categories, such as, for example, things to take out of a burning house. This is also a coherent category, although there is no prototype and no defining attributes.

Prototype theories do not explain the categorisation process, the what and why of concepts. It is hard to represent causation using attributes alone. Rosch (1978), when discussing whether these theories form a good basis for modelling categorisation, says that:

> *with specific concern to modelling: "pure" prototype theories appear to describe results rather than any effective proceduralisation of how these results are brought about.*

2.4.1 Similarity in music psychology

Apart from Nattiez and the whole trend of Paradigmatic Analysis, the work which focuses on musical similarity is to be found mainly in the area of cognitive musicology rather than music analysis. In particular, experiments have been carried out in listening, and researchers have concentrated on what listeners perceive to be similar or different. The reader is referred to Lamont and Dibben (2001), Chapin (1982), Deliège (1992), Deliège (2001), Deliège and Melen (1997), Edworthy (1985), Pollard-Gott (1983), Welker (1982), and Zbikowski

(1999).

2.4.2 Paradigmatic Analysis and Categorisation Theories

How does Paradigmatic Analysis relate to Defining-attribute and Prototype theories of categorisation?

Paradigmatic Analysis employs an exemplar view of categories in which the first occurrence of a segment, the paradigm, takes the role of exemplar. Subsequent segments are then assigned to the categories based on their similarity to the exemplars. The categories are defined intuitively: no explication of characteristic or even defining attributes is made, and the choice of the first occurrence of a pattern as the exemplar is arbitrary. The assignment of the segments to paradigms is based on the intuition and skill of the analyst; no formal similarity function is defined to determine the closest exemplar. As a consequence, the analyst makes a clear decision with which paradigm to group each segment, but this decision neglects the psychological fact that the boundaries between the categories might be fuzzy and that some segments might be legitimately assigned to two or more different paradigms. Furthermore, in Paradigmatic Analysis only one level of categories is considered and the relationship between the classes is not evaluated.

Our approach, described in the following chapters in more detail has various differences to Paradigmatic Analysis. By taking a formal approach, the method described in this thesis yields results that go beyond those of Paradigmatic Analysis: our computational model does not take an exemplar view but a characteristic attribute approach to categorisation: the prototype lies in the *centre* of the category as defined by its attribute space, reflecting the statistical properties of the attribute values of the category members. The similarity of each segment to the prototype can be precisely measured as the Euclidean distance in attribute space. This method makes explicit those segments that are ambiguous in their classification, having equal distances to several prototypes. Further, the formal categorisation process uncovers hierarchical relations between classes so that the analysis of a piece becomes possible on multiple levels. We will return to these points in chapters 3 and 5.

2.5 Formal and Computational Modelling

One of the core arguments of this thesis is that formal models provide several advantages over the more intuitive application of theories. In this section we look at some principles and motivation for formal modelling and then look at the issues of knowledge representation and neural networks in classification.

2.5.1 The General Methodology of Formal Modelling

The observation of behaviour, or introspection, can lead to the development of a theory about various processes or relationships. In whatever way a theory might be formed, it is in its advantage if it can be tested: A theory is useful if we can prove that it has no magical steps in it, no processes that rely on intuition, and that all the steps are clearly defined.

In order to achieve this, a model is required. In transforming a theory into a formal model, often inconsistencies that were not even anticipated before are uncovered. These can concern vague formulations of inputs to the system, or "magical steps" within the theory where e.g. data is transformed in unspecified ways between steps, or where implicit additional assumptions have been made that were not believed to be part of the theory.

A theory is important for the understanding of any procedure. However, if intuition is necessary to understand the working of a theory, then the understanding might stem from the intuition and not from the theory. If a theory can be carried out step by step by some exterior processor, then we can eliminate this problem. This processor does not necessarily have to be a computer. Any implementation of a theory, either computational or merely a formal procedure, is a model of the theory.

The explicitness of the model allows for its testability: once every step has been precisely defined, it is possible to verify for every input how it produces a certain output. Apart from testing a theory, another reason that we might need a model is to be able to make predictions about the outcome of theories, in order to check the effectiveness of the theory by checking its predictions. The results we get can give us more information about the theory, and this can allow us to go back and re-evaluate the theory. This procedure can be repeated until satisfactory results are obtained.

2.5.2 Knowledge Representation

A key issue in Artificial Intelligence is that of knowledge representation. The way knowledge is represented can alter the results of any model dramatically. Appropriate and efficient knowledge representation is therefore needed in order to solve any problem. This subsection discusses some issues of knowledge representation that are related to the rest of the thesis. Mylopoulos and Lavesque (1984) propose a general taxonomy of representational schemes:

1. Representation in logic, using expressions in formal logic
2. Procedural representation, using "if ... then ..." rules for problem solving
3. Network representation using graphs and nodes, and
4. Structural representation, which is like the above, only nodes can be whole nets (structures) themselves

In *network representations* there exist nodes and arcs that connect the nodes. Nodes can be any concepts or objects, and the arcs show the relations between these concepts or objects. Examples of network representations are semantic networks and conceptual graphs. Inheritance is often a property of these representations. An example of a semantic network which preserves inheritance is developed by Collins and Quillian (1969): A canary inherits all the properties of a bird, and a bird inherits all the general properties of an animal, a fish also inherits all the properties of an animal, but has nothing to do with those of birds, and so forth. This is an elegant and economic way of representing properties.

Structural representations are richer network representations in two ways: firstly they allow a node to be a network representation itself. Secondly each node can be a more complex data structure with slots and values. Frames are structural representations.

Further developments in the domain of knowledge representation seek to connect properties between them, describing relations between various properties. This results in complex associative structures as described in the two last options of knowledge representation above.

The issue of knowledge representation as applied to music is discussed in the next chapter and continued in chapter 5, where more information on the method of analysis and its computational model respectively appears.

2.5.3 Neural Networks and Classification

Classifying objects is a fundamental task which has been studied in depth in disciplines such as formal learning theory, computer science, and artificial intelligence.

Neural Networks have, in the past fifteen years, become very powerful systems that can be used in all sorts of tasks and applications. Their value for the modelling of psychological processes has been firmly established (McClelland *et al.*, 1986; Ellis and Humphreys, 1999). Most such networks were originally designed as models of the brain at some level of abstraction and their ability to generalise to new data and learn underlying rules in sets of inputs has made them an attractive alternative to symbolic, rule-based approaches in many areas of cognitive modelling. The abilities of these networks, especially in subtle discrimination tasks, mean that they are often used in an engineering sense as the best solution to a particular problem, independently of their biological origin.

Although there exist a great variety of neural network models, all consist of the same simple building elements, called *units* or *nodes*. Units are connected between them with weighted connections that are often simply called *weights*. A unit in a neural network can receive input from one or more other units, or external input. The unit adds up the inputs it receives, and then, according to this collective input and an activation function, it either gets activated or not. Different neural network models differ in their connection patterns, in the form of the unit activation function, and in the algorithms to update the weights in order to learn a certain task or achieve a certain goal. A newer class of models, like the one used in this thesis, also change their architecture during the learning process by adding units and connections.

In *supervised* models, the network learns to associate a set of inputs with a set of corresponding outputs: the inputs are presented to the units of the *input layer* which propagate their resulting activation, perhaps through one or several *hidden layers*, to an *output layer*. The output that is produced by the network is then compared to the desired output, and the connection weights are adjusted so that when the same pattern is presented again, the actual and the desired outputs will correspond better. The best-known of such supervised algorithms is the error back-propagation algorithm (Rumelhart *et al.*, 1986).

While supervised algorithms rely on an external teacher that tells the network the desired output for a certain input, in *unsupervised* learning there is no teaching signal. Instead, the network discovers the underlying structure and regularities in an input data set. Unsupervised models are therefore used in clustering and classification tasks. Well-known such algorithms are competitive learning (Rumelhart and Zipser, 1986) and Kohonen's Self-Organizing Feature Map (Kohonen, 1982).

Pattern recognition and classification are two tasks which neural networks are particularly good at. Many need a substantial supervised training period, but some can create their own categories unsupervised. In this thesis we choose the unsupervised paradigm: the analyst does not have access to a "teacher" that tells him/her to which class each segment has to be assigned. Instead, he/she discovers the classification based on his/her (intuitive) choice of properties that describe the segments. In using a formal approach, we can apply unsupervised learning in a novel way: the input data set consists of the segments of a piece that are explicitly described in terms of their features. The unsupervised algorithm then discovers the underling structure of this data set and performs a classification. We can then compare this formal classification with that done by the analyst. If they are different, we do not adapt any weights in the model as in the supervised approach, but instead adapt the *input representations*, that is, the description of the segments. If we achieve a classification that corresponds to that of the analyst, we can take the formal description of the segments as a model of the descriptions employed by the analyst in his/her classification. Since the descriptions used by the analyst are usually informal and intuitive, we can by this method also uncover inconsistencies: some segments might be categorised in the same way as by the analysts by employing a certain representation of the segments, while other segments might require a different representation. In chapter 5 we examine in more detail the algorithm used for our model, namely the *Growing Neural Gas* algorithm (Fritzke, 1995).

2.6 Related Work on Representation and Categorisation in computational music analysis

Computer systems need a special music representation in order to understand and work with music. The choice of a representation language is the crucial first step for the further working of any automated system, and the choices made will influence the result. We have to decide what to represent, that is which of the information we have at our disposal is relevant in the solution to the problem, and how to represent it so that not only the computer understands it, but that it is also clear and efficient in a way that will not obscure the solution process.

The next important step in producing a system is which parts to automate and how. Clarity and efficiency are two important issues in algorithms, as is architecture of the whole system. For example, in a task such as ours, a system should be modular, which means that for every single separate task there is a separate algorithmic unit, and these are totally distinct from each other.

One of the earliest and most influential works in computational music analysis has been the *Proposal for a Grammar of Melody* by Baroni and Jacoboni (1978). The approach is rule-based, and the application is on Bach Chorale melodies. Throughout the work, the authors critically evaluate the role of the computer in music analysis.

For some general issues, such as categorisation, there are plenty of algorithms available, each with its own advantages and disadvantages in the implementation and in the results they give, and one can choose between them. For example, some algorithms can give a hierarchic categorisation of objects, while others can process the objects incrementally and others cannot.

In the last few years, there have been various representation formalisms and classification algorithms that have been used specifically for the classification of musical segments or whole pieces. Anagnostopoulou *et al.* (1999); Höthker *et al.* (2000) have investigated the impact the choice of representation and algorithm has on the classification analysis of a piece or pieces of music. They have found that the choice of representation makes a significant difference on the results, while the choice of algorithm has little effect on the results. Therefore, what is important when choosing an algorithm are the *benefits* that the algorithm has to offer. For example, one algorithm might give a hierarchic classification. Another algorithm can allow for overlapping classes, and so on. This is very important for music analysis, when adjusting one of these algorithms to fit a musical problem. It depends on the musical analyst what kind of such benefits are important for his/her analysis. Chapter 7 gives a more in depth comparison of the various algorithms with our approach and what each one has to offer to the analyst. In the following subsections we describe some interesting and relevant systems of representations and algorithms.

2.6.1 The CHARM system for music representation and an algorithm for Paradigmatic Analysis

The CHARM system, as presented in Smaill *et al.* (1993) was designed as a general purpose musical representation system. That means that is should be possible to encode any kind of music, of any style, and any kind of relation between various units (whether these are notes, phrases, and so forth) into the system. The important feature of this system is that it uses the concept of a *constituent* and that it is hierarchical.

The work first appeared in Wiggins *et al.* (1989) and developed further in Smaill *et al.* (1993). The system makes use of the *event* structure, which is a combination of a unique identifier number plus pitch, onset time, duration and timbre information. From these events, by using several functions, it is possible to deduce higher-level information such as whether some notes have the same pitch or duration, and so on.

Constituents are larger data structures of various types which include events and capture higher-level structural information. Constituents can also form part of higher level constituents. In the same paper, they give an interesting application of Paradigmatic Analysis of *Syrinx* by Debussy, thus focusing on concepts of similarity between events and constituents. Their algorithm is based on the Ruwet approach to Paradigmatic Analysis (Ruwet, 1972) and runs as follows:

- as a first step, the program looks for repetitions of identical phrases. The first occurrence is named motif, and the repeats are called derivations. All of these phrases are then removed from the rest of the program.
- as a second step, the program looks at the remainder of the piece, in order to identify similar phrases to the motifs found above, under some criteria for similarity.
- the second step is repeated until no more music is left and all the music has been successfully attributed to categories.

The similarities that they consider, apart from identity, are:

- if two phrases are identical apart from a longer first note
- if a phrase is an exact transposition of another
- if a phrase is an exact transposition of the pitches, but durations are allowed to be different.

The authors note:

> *It might be argued that these similarities are rather* ad hoc; *indeed, it is generally necessary to use similarities in this kind of analysis which are suited to the style of writing used by the composer. However, note that the program is modular over the set of similarities, which can be regarded as further data which can be easily updated.*

The idea that the similarity has to be defined according to the context (i.e., the piece or the style under analysis), is a recurring one in this thesis.

Westhead and Smaill (1994), describe a system for automatically characterising musical style using motifs. Motifs are patterns of rhythms or pitches common to more than one piece of music in a style. The Style Analysis with Motifs (SAM) system uses an original classification technique and is able to learn to distinguish between different styles with a success rate of over 95%. It is suggested that since motifs can be used to automatically discriminate between different styles effectively, they may be very important in the way humans achieve the task.

2.6.2 The representation formalism of Multiple Viewpoints

The work carried out by Conklin and Witten (1995) introduces a novel approach to music representation which is general and flexible, called *Multiple Viewpoints of Music*.

In this approach, the music score is translated into streams of *viewpoints*. Each *viewpoint* models some specific type of musical phenomenon derived from the musical surface; for example, melodic contour, duration, interval from a tonic referent pitch, or melodic intervals. A melody is represented as a sequence of basic events; tuples of pitch, duration, and start time.

A *linked viewpoint* is a combination of two or more viewpoints that models several derived types occurring simultaneously for the same event.

A *threaded* viewpoint models phenomena that occur at defined places within a melody; for example, at the beginning of a bar or phrase or at every quarter note pulse. The `gis221` viewpoint, named after Lewin's Lewin (1987) generalized interval system 2.2.1, assumes a value that is the difference in start–time between two events.

The notion of *viewpoint patterns* appears in Conklin and Anagnostopoulou (2001) (see also appendix), where the authors look at the representation and discovery of melodic patterns within multiple viewpoints. This is a novel approach in the treatment of musical similarity, and on the selection of the most musical significant patterns. The authors note:

> *There can be repetitions within different musical parameters, such as intervals or successions of intervals, melodic motion (contour), relative rhythmic values, middle or fundamental structure, harmonic progressions (implicit or explicit), register, dynamics series, pitch class sets, and so on. Approaches to pattern discovery in music analysis have so far concentrated on the similarity relationships between note patterns rather than on recurrent patterns within these musical parameters. However, in a music analysis task of any kind, it would make more sense to be able to capture these recurrent patterns: they are more general, look at a deeper level of similarity within the musical corpus and make explicit* where exactly *the similarities between the patterns lie. [...] A (...) filtering looks at all discovered patterns and selects the most musically important ones by introducing the notion of the* longest significant patterns *in a musical corpus.*

2.6.3 Categorisation in the GCTMS theory

Cambouropoulos (1998) has devised a *General Computational Theory of Musical Structure* (GCTMS) which "may be employed to obtain a structural description (or set of descriptions) of a musical surface". As an initial music representation scheme, he use the CHARM notation described above. The theory is divided into several discrete steps, which cover a *General Pitch Interval Representation* and *General Chord Representation*, various stages of segmentation, pattern matching, the categorisation of segments using the *UNSCRAMBLE* algorithm (Cambouropoulos and Smaill, 1997) and a temporal organisation of the categories which result to the musical structure information.

Of this theory, particular relevance to the present thesis is the categorisation process with the *UNSCRAMBLE* algorithm. This is an unsupervised categorisation algorithm, which takes as input segments of a musical piece and some properties that describe these segments and it outputs a set of plausible classifications, in a dynamically evolving process. The knowledge about the final categories is explicit in terms of properties and weights, and therefore it is possible to predict potential membership of new segments into the categories.

2.6.4 Classification in MELONET and MELOGENET

Two systems, MELONET and MELOGENET have been developed for the learning of musical style and for melody completion (Hörnel and Höthker, 1998; Hörnel, 1998; Hörnel and Höthker, 1999). Both systems use a classification procedure of segments of melodies in order to define the musical structure. The algorithm employed for this procedure is the Ward algorithm (Ward, 1963), which is unsupervised and results in a hierarchic classification.

2.6.5 Dynamic programming techniques for pattern discovery used in classification: the work of P.-Y.Rolland

In Rolland (1998a), Pierre-Yves Rolland has devised the *Star-Center* algorithm, which is used mainly for pattern extraction within sequences. Sequences can be anything from DNA-base sequences to melodic sequences. The goal of pattern extraction often involves some similarity judgement and classification, that is, how can a discovered pattern be considered similar to one encountered previously.

There are two basic differences between this algorithm and the Paradigmatic analysis way of categorising: this algorithm allows for classification with overlapping classes, that is an object can be part of more than one class. Connected to that, there is the second difference: it does not have the problem of Paradigmatic Analysis of "similarity going further away" within a class.

2.6.6 Other approaches

The existing classification theories and techniques from computer science, and neural networks in particular, have only generally influenced music analysis until very recently; the only example of an earlier attempt using a classification algorithm has been by Gjerdingen (1991), where he uses an unsupervised neural network, ART by Carpender and Grossberg (1998).

2.7 Conclusions

This chapter has provided a diverse background in the relevant areas the thesis deals with. The common theme in the entire chapter has been categorisation based on similarity. First we looked at Paradigmatic Analysis, as well as two other kinds of music analysis that make a stated use of categorisation and saw the treatment of categorisation in them, the motivation as well as the limitations. Then we proceeded to theories of similarity and categorisation from a philosophical and psychological point of view, since a lot of work has been done in this area, and categorisation is considered to be one of the main human cognitive processes. Then we moved to the computational approach, describing why modelling is suitable and how neural networks can treat categorisation. Categorisation is a significant issue in artificial intelligence since a lot of problem solving tasks involve this process. Finally, we looked at some contemporary computational music analysis systems that use categorisation as part of their processing. In this chapter we did not discuss in detail the relevance of this material to the present work. This will make sense only after introducing our own method and will therefore appear during the rest of the thesis. This is especially the case with other computational approaches.

We have introduced two approaches of categorisation in music analysis so far. The first one is the purely music analytical one, in the style of Nattiez, and the second one is the computational one. The computational one uses categorisation more as an application in order to achieve some result, for example describe a structure or generate some melodies, and does not look in-depth at the process of categorisation (with some exceptions – see chapter 5). These two approaches have not yet properly converged, and what is more, they have not made appropriate use of all the material that exists in the theoretical and experimental world of philosophy and psychology, as far as music is concerned.

Computer applications, apart from modelling a certain process, are designed to be useful and perform specific tasks. In the case of Paradigmatic Analysis this is not straight-forward because of all the issues it raises and the problems one encounters.

The rest of the thesis will first address the limitations and problematic issues of Paradigmatic Analysis and then investigate the intersection of the two approaches, the music analytical and the computational one, and show how computers can help us solve some of the problems posed in music analysis.

Chapter 3

The Categorisation Analysis of Music

3.1 Introduction

The main idea behind traditional Paradigmatic Analysis, namely the comparison of melodic segments in order to study similarity relations within a piece of music, has been a valuable one. However, several problems arise when the method is put into practice; which is why most criticism refers to its limitations rather than this fundamental concept. The previous chapter explained Paradigmatic Analysis and discussed its problematic issues. In this chapter a new method of analysis is proposed, *Categorisation Analysis of Music*, which stems from Paradigmatic Analysis. The new method tries to overcome the limitations of paradigmatic analysis and extend it in appropriate ways. In brief, it differs from traditional Paradigmatic Analysis in that:

- it formalises further the whole procedure - to the degree that this is possible,
- it makes the criteria for the categorisation choices explicit,
- it introduces multiple levels of categorisation and different categorisation rules,

- it discovers relations within and across categories,
- it introduces prototypes rather than paradigms.

The above points are explained in this chapter. First the motivation for the categorisation analysis as a method for music analysis is discussed. Then a detailed description of the methodology follows, with a special emphasis on the introduced differences. In the next chapter, a sample piece of music is analysed and the method is evaluated. This and the next chapter are purely devoted to music analysis; the computational model, which is only mentioned here, is allotted to chapters 5 and 6.

3.2 Motivation

This section aims at answering the following three questions:

- *What* are the changes to traditional Paradigmatic Analysis?
- *Why* are these changes important?
- *How* are they achieved?

There are three improvements to traditional Paradigmatic Analysis in categorisation analysis, which provide the answer to the *what* question:

1. The categorisation analysis method is more formal.
2. There is more information available at the end-result of the analysis.
3. Findings from research on similarity and categorisation are brought in.

Each of these points is elaborated in the following subsections, and for each of them the *why* and *how* questions are answered.

3.2.1 A more formal analysis

Chapter 1 discussed what constitutes a formal method of analysis, why it might be desirable and what are its drawbacks.

It is common practice in music analysis for each analyst to use his/her own method and own way of analysing a piece; this is a very interesting way of looking at music because it provides diversity in analytical thinking. However, one could argue that it would be more useful and clear if some of these analyses used a similar general framework, a common setting where a priori criteria would be specified and results could therefore be juxtaposed. In that way, a meta-analyst could evaluate the analyses and draw conclusions on the similarities and differences of the various views. This does not mean that each analyst would have to do the same kind of analysis; the common framework should not restrict the analyst's freedom, it should only make the choices clear to everyone. Therefore, the requirement would be that the analyst cannot anymore decide on things *ad hoc*, but has to base his/her decisions on criteria that he/she has chosen. We explain how this can be done in this chapter, and in chapter 5 we explain why the introduction of a computational tool might help in this procedure.

The analyst's intuitions do not vanish if they have to be made more explicit. In music analysis we are often faced with analytical choices that have a poor explanation or none, and this makes the evaluation of an analysis hard. It would be interesting not just to see an analytical decision, but to also know where this decision came from; then it is easier to agree or disagree.

When various pieces are analysed using exactly the same method and criteria, we can also compare the results in a more objective way. In this common framework it will be more straightforward to carry out stylistic and comparative analyses. So far, although this has been a significant goal of Paradigmatic

Analysis, this has not been the case apart from few exceptions, for example Morin (1979). Also, a comparison of analyses of the same piece by different analysts should be, in the same way, more straightforward.

The main motivation behind categorisation analysis is to provide a formal framework for music analysis which is based on categorisation. This is achieved by first splitting the methodology into *steps*, each step producing a specific result, and the reasons or criteria for each choice being made explicit to the degree that this is possible. This constrains the analyst into being explicit and formal. At the same time, the analyst's freedom is not restricted; he/she is free to make analytical choices based on musical sense. In this way not only do we know why each step takes place, but the analyst is also persuaded to adhere by his/her choices and therefore to either avoid or simply notice inconsistencies.

If the criteria for the analysis, or the categorisation in this case, are stated, then it is legitimate to have various categorisations; it is expected that different people would choose different criteria for their analysis.

One of the aims of categorisation analysis is that it can be modelled computationally. In that respect, the method of analysis should be as formal as possible. The reason to make it computational is to test and ensure that the framework is indeed formal throughout, with no intuitive "gaps".

This thesis is chiefly concerned with music analysis rather than music. A possible extension of this work is to create a cognitive model of the analyst. Just as performing, listening, composing, improvising are cognitive musical tasks, so is analysing from the score itself. The analyst has to make choices that show his musical understanding of the piece, especially when categorisation is involved. Categorisation shows the analyst's perception of what is "similar" and what is "different" in the specific context, as well as which musical features are significant for these choices. Comparing analyses of various people can reveal whether there are common underlying mechanisms that have to do with cognition. In order to create such a cognitive model, the theory has to be formal. This point is elaborated more on the conclusions chapter, as a possible further extension.

3.2.2 More available information

Paradigmatic Analysis produces a categorisation of segments, which is usually at one level only . This means that there are no further subdivisions of each category into smaller categories, or that these categories are not linked to create hyper-categories. Ruwet, in his famous article *Methods of Analysis in Musicology* (Ruwet, 1966, 1972) defines a two levelled *segmentation*, one represented with capital letters and one with small letters, but not any levels of categorisation. Nattiez's first two analyses of Syrinx (Nattiez, 1975) also use different levels of segmentation, although they are presented as unrelated.

Furthermore, the criteria for putting each segment into a specific category are not usually made explicit. If there is some list of criteria, like in Nattiez's 3rd Syrinx analysis (Nattiez, 1975), this list is firstly not complete and secondly does not say which criteria apply to which choices. This information on its own does not say very much, and more information could be made available as the result of an analysis, both in terms of a multi-levelled categorisation and of explicit criteria used for the categorisation.

This extra information does not necessarily have to be used for the further comparative analysis, but the fact that it is available to be used might be significant in itself. We obtain this by having information on levels of categorisation, on criteria for categorisation, on relations between categories and on prototypes of categories; this will become apparent during the description section of this chapter.

Knowing the relations between segments which belong to the same category is important but not sufficient. By categorising segments in the *same* category we have implied some closeness, that is similarity, between them. Categorising segments into a *different* category implies some distance, that is difference between them. It is of equal importance to look at this distance between categories, which can vary from category to category, from piece to piece, from analysis to analysis. One should not only study why segments are grouped together into a category, but also why they are separated, and an analysis would be incomplete without this. One way of looking at the relation between different categories is to look if there are common features between them, contrasting features (for example slow - fast), or just simply unrelated features. This gives an account of the difference. The way for doing this is explained below, in the relevant section of the methodology.

This has a direct effect on the concepts of repetition, variation, transformation, but also of transcription and "arrangement". These terms all use the concept of similarity to a varying degree. Although they are acceptable for general purpose speaking about music, they seem inadequate when used in music analysis, as they are not very informative: what exactly is a repetition? Is a segment and its transposition considered to be repetition? How far can variation or transformation go before they are considered unrelated to the original segment? This issue of similarity is usually dependent on context. Two segments that are considered a repetition or variation or transformation in one piece might not be considered so in another.

A more informative way of looking at repetition, variation and transformation in music is by considering it in terms of various degrees of similarity. This is a more scientific way, given that none of these traditional three terms is well-defined, even the much used "repetition". Degrees of similarity can be a more objective way of calculating the distance between two segments.

The idea of similarity substituting repetition, variation and transformation is not new. It is encountered in pitch class set theory, where the similarity depends on the member of each pitch class set. In visual art, there is also the concept of similarity rather than repetition, variation and transformation.

Paradigmatic Analysis does not deal with *difference*, only with similarity. It could be argued that this is not necessary, since difference is a complementary concept to similarity and that a relatively long distance in a similarity measure implies a difference. But even if so, a study of similarity within a piece of music would be incomplete without the furthest possible similarity relation (ie. the notion of objects with no features in common), and paradigmatic analysis disregards this respect.

Difference, or furthest similarity relation, can be observed by comparing categories to each other – how much difference is there between categories? This can be very important when studying a composer's style, when one needs to observe the uniformity or diversity in the use of musical material.

The importance of difference also becomes apparent when one considers the fact that different people would produce different analyses of the same piece. That might suggest that some categories are closer together than others, and to just divide segments into categories, without other levels of categorisation, might not be enough. Deliège (1993) has noted the importance of difference in music, and suggests that there are two principles in music – the principle of similarity and the principle of difference.

Finally, although Paradigmatic Analysis is supposed to be followed by a *Distributional Layout*, that is the distribution of the segments in time, or what is called a syntagmatic analysis, this is not always the case in practice.

3.2.3 New research from similarity and categorisation

The previous chapter discussed the current research on similarity and categorisation in the domain of cognitive science. Recent experimental findings, as described in chapter 2, suggest that the human categorisation process is based on the following principles:

1. there can be overlapping categories, which means that an object of one category can belong to another category, too. This is equivalent to overlapping extensions of concepts.

2. There are unifying properties across different categories. That means that in Frege's terminology, *semantic intensions* can overlap.

3. There are no necessary and sufficient conditions for category membership.

4. There are prototypes in terms of probabilistic values.

There is a dilemma, on which of these principles to incorporate into a method for music analysis, and how to achieve this. The dilemma stems from the double motivation for this method of analysis: on one side we would like a powerful method which can analyse anything from a single piece to large datasets of music in a clear and efficient way. From the other side, we desire a cognitive inspiration to the degree that this is possible, in order to use, at a later stage, this method as the basis to a cognitive model of the analytic procedure in music, and to create a method that feels natural to use.

This is a common tradeoff encountered in artificial intelligence work which aims to be close to cognitive science, between the efficacy of a model and the cognitive principles behind it.

Since the aim of the thesis is music analysis and not a cognitive model, we adopt the three out of the four principles that will help our design of a method to be used more efficient in music analysis, especially for very large data sets of comparative analysis. We omit the first principle, the overlapping categorisation. However, its incorporation to a future extension of a cognitive model would be possible and necessary.

3.3 Description of the Methodology

3.3.1 General description

We divide the method of Categorisation Analysis of Music into six steps: *Segmentation, Feature Description of Segments, Categorisation, Relations within a category and across categories, Prototype discovery*, and *Distributional layout*. These are shown in figure 3.1. The main reason that this method is divided into seven steps is, apart from clarity, that each part yields a *concrete result*. For example, the result of segmentation is a segmented piece of music; the result of the feature description of segments is a set of attributes describing each segment; the result of categorisation is a set of categories of segments; the result of prototypes is a collection of the most characteristic properties of a category.

The division is not related to the order in which the steps are carried out; that means that the steps do not necessarily have to be carried out sequentially. However, there is one restriction: The first three steps, that is segmentation, feature description of segments and categorisation, have to be carried out before the other four steps. This is simply because of practical reasons. We cannot have, for example, prototypes of categories if we do not have the segments themselves. For this particular example it is important to note that here we do not follow the strict prototype theories (see chapter 2), where the formation of a category depends on the distance from a prototype.

In figure 3.1 the three first steps, segmentation, feature description of segments and categorisation, are linked together. This indicates the way that the results are achieved: these three steps cannot be totally separated from one another regarding the way they are carried out, although their *results* are totally separable. It has been argued, both in the analysis (Nattiez, 1975) and in the psychology of music (Deliège, 2001), that segmentation and categorisation are inseparable, and one depends on the other. It has also been argued, both in philosophy and cognitive psychology (Ramscar and Hahn, 1998) and in computational approaches (Cambouropoulos and Smaill, 1997) that similarity (that is the common sets of attributes) and categorisation are bound together. We are not interested in either psychological processes, or the analysts' preference on ordering, or how the results are achieved; only in the results themselves. This is why we are not going further into this issue, which has occupied psychologists recently.

The other three steps, namely relations within and across categories, prototype discovery and distributional layout are products of the categorisation step. They are totally separable between them, and the order does not matter. The only restriction is that they have to follow the first three steps, since they take the initial categorisation step as the starting point.

In the next subsections we describe the six steps and their results in detail, and we discuss how the claims of the previous section, regarding the motivation are evaluated.

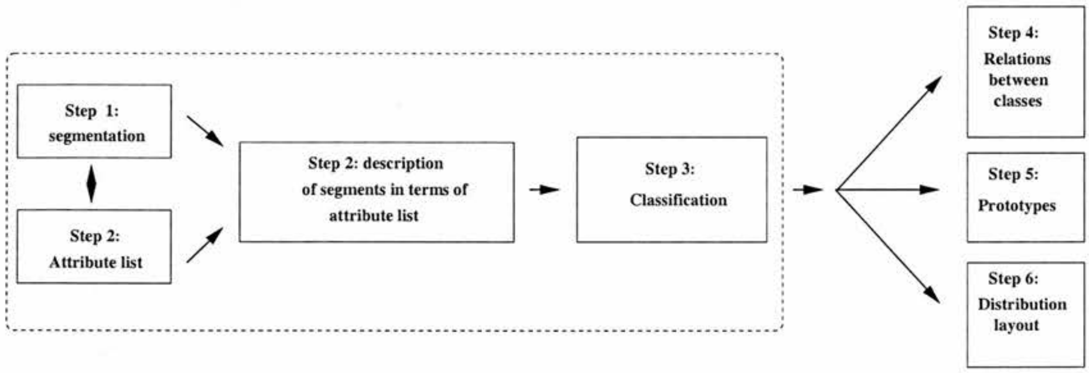

Figure 3.1: An overview of the Categorisation Method

3.3.2 Segmentation

As has been stressed in the introductory chapter, this thesis does not explicitly deal with segmentation. It is rather taken as fixed or given, and the thesis concentrates on the other steps of the analysis, which are related to similarity and categorisation. Therefore, it is only briefly discussed what kind of segmentation is needed.

Chapter 1 considered segmentation as a case of grouping strategy. The process assigns consecutive notes to units that we call segments. As with all kinds of structural analysis, the relations between these units can be made apparent. The unit in this method is the *segment*, and not, for example, the note, the bar, the motif or the phrase. In fact, a segment can be any of these, or any length of an excerpt of the musical piece. The segments can be of variable length in the same piece, and it is up to the analyst's judgement to decide what will constitute a musical segment. The reasons why we choose the segment and not any of the more conventional melodic distinctions are the following:

- The categorisation analysis is a general method of analysis that should be appropriate for any kind of music. *Motives* only exist in certain kinds of music. Furthermore, even in a music which can be characterised as motivic, for example Beethoven's piano sonatas, there are clearly other types of melodic constructions that appear, such as transition passages. Therefore the use of motives is very limited.

- The use of *phrases* is limited for the same reason: not all music can be divided into phrases. Furthermore, the term *phrase* is ambiguous in its use.

- *Bars* are just metric distinctions and segmenting in bar lines most of the time does not make any musical sense.

- The *note* does not carry any significant properties on its own (Seeger, 1960). All intervallic and contour information would be lost. However, when at a later stage we describe and give attributes to segments, then what we really describe is relations between notes.

It is not trivial to define a musical segment, and the degree of unity that a segment must entail. In linguistic syntactic analysis, there are more clear constituents, but in music this is not the case. In semantic analysis, semantic constituents are not as clear as syntactic ones, but ambiguity exists to a much lesser extent than in music. A musical segment, as we define it here, can be any portion of music. The only restriction here is that there have to be enough segments for comparison and categorisation. The whole piece can be considered as a segment if there are enough pieces. Therefore the minimum needed for comparison is two segments, which can be of arbitrary length.

In order to keep the segmentation process as general as possible, we allow segments to be *of variable length* and *overlapping* within the same piece. By accepting a variable-length segmentation we take into account the fact that there can be longer or shorter units and passages of music that according to the analyst could not be further split. We prefer this kind of flexible segmentation because it is very rarely that an equal length segmentation can be appropriate in music, perhaps only in very uniform pieces. Methods of analysis that use a fixed-length segmentation can usually be applied to a very limited set of music. For example, Höthker *et al.* (2000) use 9 two-part Bach Inventions, and the segmentation is at every crotchet beat. Also, by being able to cope with flexible segmentation, the fixed-length one is not excluded. Chapter 2 shows an example where fixed-length segmentation works for the purposes of the specific analysis.

Overlapping segmentation occurs when a segment starts before the previous one has ended, so that there is a passage of music, sometimes of only one note, which belongs to both segments. Overlapping segmentation also allows for two parallel segmentations of the same passage of music, whether this passage is short, long, or even the whole piece, when both segmentations would make sense. For example, in Brahms's *Intermezzo*, Op.119 No.3 – which Nattiez analysed using two different segmentations – we see that we don't want to lose any information; keeping just one of the segmentations would mean missing out some interesting repeated motifs. This point is explained in more detail in Monelle (1992). One could think of this as *musical ambiguity* where both interpretations are valid.

Hierarchic segmentation, that is dividing each segment into subsegments, recursively if desired, is attempted in chapter 6, on *Parenthèse* by Boulez. This kind of segmentation is not particularly well-suited to our method, and the drawbacks as well as potential solutions are discussed in that chapter.

There is no necessity for the analyst to consider all the segments that stem out of the segmentation step. There might be small fragments of music that can be left out at the consequent steps. This omission can vary between one note, or one rest, up to big fragments of music.

How much can be left out? It depends on the purpose of the analysis. An alternative to segmentation is to just consider a selection of musical fragments, that are in some way significant to the analyst, to continue the further analysis. These might be, for example, a motif and its traces through a big piece, in order to study its transformations and variations, or it could be a selection of interesting musical phrases. In computer music, it is very common to consider only certain significant patterns, and this procedure is called *pattern discovery* (for example, Conklin and Anagnostopoulou 2001, also see appendix).

In general, segmentation and identification of similarity are two processes difficult to separate. An obvious point to segment a string of music is where a repetition starts. This can be one of the criteria for segmentation. However, as it has been stressed in the subsection above, the two processes have clearly distinguished results, and here they are also clearly separated.

As a result of this closeness of the two processes, the categorisation method can evaluate different segmentations and, given a number of different segmentations, it can distinguish which one is the most appropriate. This depends on the results one can get from the classification steps. If the results are thought to be unsatisfactory, a new segmentation can improve these results.

3.3.3 Feature description of segments

The three terms *feature, property, attribute,* although they have their own special connotations, as described in chapter 2, in general can be treated as being interchangeable (Eysenck and Keane, 1998). This is the approach taken from here onwards, unless otherwise specified.

The purpose of this step of the analysis is to create a list of musical attributes that are relevant to the piece and each segment can be described in terms of these attributes. The reason for this description is that the desired result is a categorisation that depends on specific criteria, and these attributes will form the criteria according to which the categorisation (which is the next step) will be carried out. The choice of attributes depends solely on the analyst. Here we discuss what kinds of attributes are possible, and in the next subsections we explain the methodology of this step.

The question that first needs to be addressed is what constitutes an attribute or property of a musical segment, and then how can attributes be criteria for categorisation. To explain what an attribute is, one can look at feature theories, for example, Neisser (1964), which form one class of pattern recognition theories. Eysenck and Keane (1998), p.48, explain:

> *A pattern consists of a set of specific features or attributes. For example, a face could be said to possess various features: a nose, two eyes, a*

> *mouth, a chin, and so on. The process of pattern recognition is assumed to begin with the extraction of features from the presented [...] stimulus.*

However, with experimental evidence, these kind of theories proved not sufficient for explaining and adequately describing patterns. What was missing was the structural description, that is, how these attributes were related to each other (Bruce and Green, 1990). For example, the description of a face would contain all the above attributes plus relations between the attributes.

Marr and Nishihara (1978) discuss the hierarchical organisation of attributes in object recognition, and especially in visual form. Their most well-known example is of the human body. They say that the human body can first be represented as a cylinder. Then, at a closer look, this cylinder can be broken to cylinder for the main body, two cylinders for the arms and two cylinders for the legs. Each of these cylinders in turn can be broken down to further cylinders, and so forth.

In our approach we use attribute descriptions which are simple, and in certain cases are hierarchical. A taxonomy of attributes is introduced below, and explained further in chapter 5, since it is mainly a computational issue. For our purposes, *any* description of the segment of music, whether this description is structural or not, in fact anything that the analyst wants to note as potentially interesting, will be considered here as an attribute or a set of attributes. Some attributes might be:

- a specific descriptive feature, not associated with any other features; for example, the existence of a tritone, a grace note, an ascending melodic line, an arpeggio, a pause. These could either be present or not.

- a structural relation between features; for example a cadence, a long note of a specific tonality, or a specific *pattern*, for example a minim followed by a pause, or the first four notes in Beethoven's 5th Symphony.

- some hierarchic features, or properties, preserving inheritance, as described in chapter 2; Figure 3.2, shows an example of such a structure, which is discussed in more detail there.

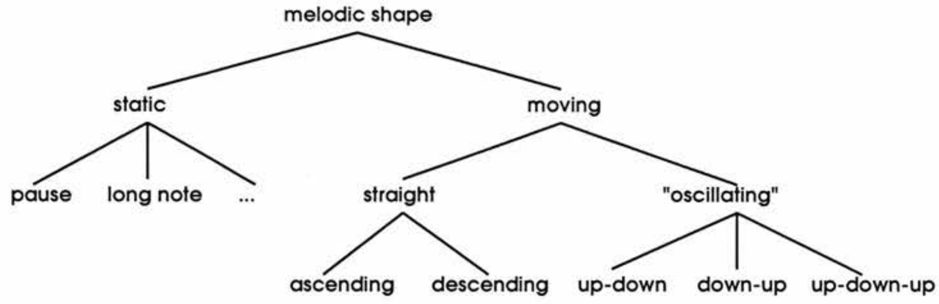

Figure 3.2: Semantic structure with built-in inheritance

The background of inheritance, and relations between properties in that way, is diverse. As mentioned above, in vision it was examined by Marr and Nishihara (1978); in structural linguistics, by Greimas (1966). In music, the work of Greimas has been advanced by Raymond Monelle (Monelle, 1992, 1991a). However, a big part of the literature comes from knowledge representation in artificial intelligence. Chapter 5 discusses in more length the issue of knowledge representation.

The *feature description of segments* step can be divided into the following sub-steps:

- **a.** The feature extraction of each segment,
- **b.** the inventory of all the features to form the *Attribute List* of the piece, and
- **c.** the description of each segment according to this attribute list.

Each one of these sub-steps is explained below.

a. The feature extraction from each segment

For the first segment, the analyst decides which features of this segment are potentially interesting for his/her analysis. There are no rules on how many these should be, how general or specific, or concerning which musical aspects. For example, a rhythmic analysis would concentrate only on rhythmical features; a comparative analysis of many pieces might pick several general, higher-level features; the tracing of a specific interval across many pieces might include only the presence of this interval; and so forth.

Moving on sequentially to the next segment, the analyst has to do the above, and in addition to examine any major differences to the previous segment that should be noted down.

It is expected that there cannot be an exhaustive description for each segment; not all of the properties can be recorded. The purpose here is *not* to be able to reproduce the segment from its description. That would only be possible in the case of the lowest-level possible description, which in a formal grammar would be equivalent to terminal nodes; in here that would be a mere description of the notes – for example quaver G4 followed by crotchet C4, and so forth. This is clearly a reproduction of the score, does not make any abstraction of musical properties and is not useful for further categorisation. The analyst can be selective. What is important for one analyst might not be for another, and it should not be forgotten that the purpose of this method is to preserve this analytic freedom. However, there are cases that such features might want to be recorded. For example, if the pattern [G4, G4, G4, Eflat4] is very important for the piece of music, then it can be noted as it is and be considered as one feature.

b. The Attribute List of the Piece

All the features from all the segments are concatenated into a list, creating the *Attribute List* of the Piece. Features that are recorded more than once for different segments in the previous sub-step, are now recorded only once in this list. The purpose is to create a collection or inventory of all the possible features that were found in the description of the piece's segments.

If the purpose is comparative analysis of various pieces, then the procedure is the same: all the features of the various pieces under analysis are concatenated into a single list that is used for the further analysis.

c. Description of segments according to the Attribute List

The purpose is to describe each segment in terms of all the attributes of the attribute list of the piece. Since the list is a collection of all the attributes of all the segments, some of these attributes of the list are present on each segment and some are not. This can be notated on a table, where on one side are the segments and on the other side are the attributes. Then each segment can take a yes/no value against each of the attributes, depending on whether this attribute exists or not in the segment. This way of transcribing musical segments into sets of features has been used before both in semiotics and in computer applications of music analysis. Table 6.1 in chapter 6 gives an example of such a table. In that table, whenever attribute exists, it is notated with a "y", otherwise it is left blank.

3.3.4 Categorisation

Having described each segment as a set of features, the categorisation of the segments according to similarity becomes clearer, since it does not depend on the music any more, but on the feature description of the segments. The description language of the objects to be classified has changed.

There exist many theories and methods of categorisation, and the main ones are discussed in chapter 2. In here, the following principles hold:

- The number of categories is not restricted, so the analyst has to decide on how many categories he/she wants.

- There can be shared features among different categories.

- There are no necessary and sufficient conditions for category membership. That means that not all the segments of a category need to share a specific feature, as it has been explained in the previous chapter.

There is no single correct way of categorising objects. Each analyst will come up with his/her own categorisation. This point is demonstrated in chapter 6. This is acceptable and indeed desirable, as long as the criteria for the categorisation are made explicit and other analysts can follow the way the categorisation was achieved. It follows that with a different feature description (and weights) a different categorisation will be produced.

3.3.5 Relations within and across categories

In most cases there will be no unifying common features for each category, and there will also be features in one category that are shared with other categories. That shows that some categories from the previous step might be closer, in terms of common features, to some of the other categories and further to others. Figure 3.3 illustrates an imaginary illustration of segments belonging to categories. We can imagine a potential distribution of the segments in the n-dimensional space, where each dimension represented by a different attribute. In this figure, for illustration purposes, we reduce the n-dimensional space into two dimensions. We can observe that some categories are closer and some are further apart, a relation that Nattiez does not examine.

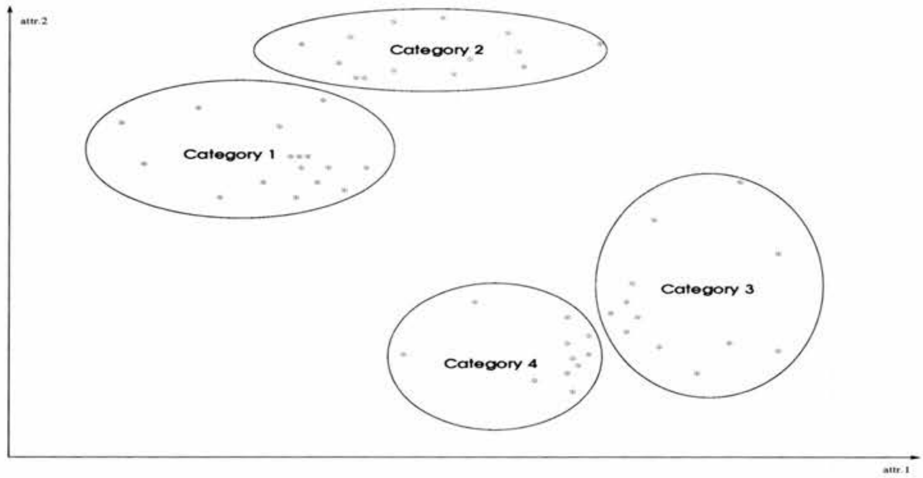

Figure 3.3: Nattiez' categories can be closer or further apart between them, depending on how many attribute values they share. Here is a possible distribution of *segments* in a 2-dimensional input space, where dimensions represent segment attributes.

In the same figure we can also see that some segments could be closer together than other segments within the same category. In order to study relations of segments within each of the categories, we apply the same categorisation procedure again, in each one of the categories. The result is a subcategorisation in each category, which results in new subcategories. For relations across categories we create a table where on one side are the categories and on the other side the features. The reason for creating this table is to see which features are shared and among which categories. From this we can create hypercategories, that is, put together some of the existing categories to create bigger categories. This last step is particularly significant for the study of difference, or distance between categories; an analysis that looks only on similarity and not on difference would be incomplete. Figure 3.4 shows how the categories of various categorisations might be related to each other. The categorisation that has already been carried out in the previous step is called here the original categorisation and in the diagram it is represented in bold squares.

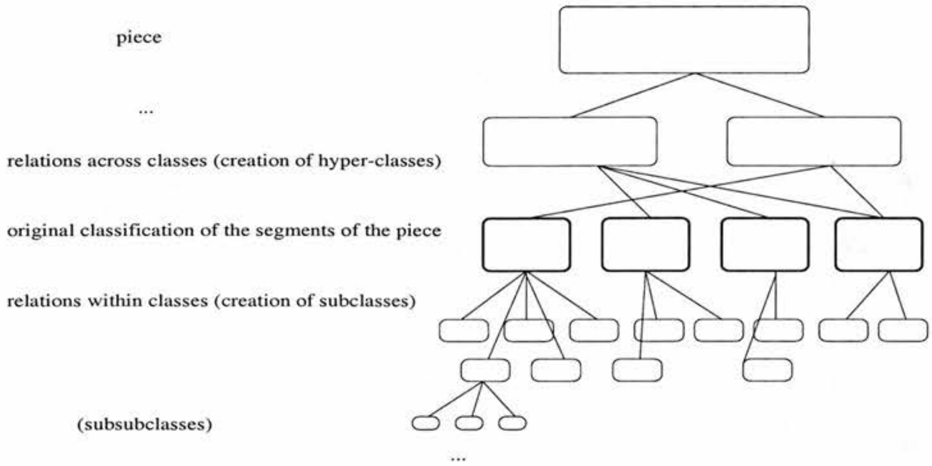

Figure 3.4: Different levels of categorisation can result from the original categorisation

3.3.6 Prototype discovery

The concept of prototype and its use in categorisation theories is discussed in chapter 2. In here we follow the view that a prototype is a collection of characteristic features and not the best example out of the segments of the category. There is no delimiting set of necessary and sufficient properties for a prototype; it is simply a collection of the most frequent ones. Here the prototype is artificially constructed after the categorisation. This definition is refined when introducing the computational model in chapter 5. In the next chapter there are examples of prototypes of classes.

3.3.7 Distributional layout

The focus of this step is on the succession of the various categories and features in time. The first category encountered is thus named *a*, the second one *b*, and so forth. The piece then can be transcribed into a sequence of letters, for example *ababcdeffefg* ..., each letter denoting a category. This is rather like traditional analysis, where form is studied as binary, ternary, and so forth. The difference here is that we are not looking at structure, and that this is at a much lower lever, the segment level. The string of letters is named *Abstract Motif Sequence* (Hörnel and Höthker, 1998). Hörnel and Höthker first used this concept of syntagmatic analysis when they looked at children's' songs.

Another way to illustrate the distributional layout is a graph, where the x-axis represents time in segments and the y-axis has markers, each marker representing a category or a feature. In this way, it is immediately obvious what is the distributional layout of the material used by the composer. Anagnostopoulou (1997) has used this technique to display cohesive and referential chains in linguistic discourse. The full paper can be found in the appendix. Figure 4.11 in the next chapter shows a distributional layout of the classes of the piece under analysis, a Scottish Folk melody.

This step can be useful in stylistic and comparative analysis, when the composer might favour some distribution in a set of pieces. It is particularly interesting in comparative analysis, when one can observe a specific distribution that a certain composer can favour, as it is shown in the appendix , where various 2-part Bach inventions are analysed.

Guertin (1981) has also looked at the distributional layout in her analyses of Debussy's preludes. However, she arrives at a generative rule at the end, which is very interesting but which can be found in so few other pieces. Therefore we are not concerned with generative rules here.

3.4 Discussion and Conclusions

This chapter has presented the motivation and description of categorisation analysis method in musical terms. Already, without any additions from artificial intelligence, the method works and can be used as a more refined form of Paradigmatic Analysis, in a more formal way, and making more information available. The discussion so far has been abstract, without the help of any examples. The next chapter illustrates the method by analysing a Scottish folk tune, which is relatively simple, in order to demonstrate the methodology.

However, although some further formalising can be obtained by using computational techniques, as it is, the categorisation procedure is not really formal, although the criteria are made explicit. The problem is how the analyst is using these criteria when categorising segments. Furthermore, the discovery of prototypes, without the help of an algorithm, might not be always correct. Chapter 5 examines these problems and proposes a computational tool.

Chapter 4

Sample analysis using the Categorisation Analysis method

4.1 General Introduction

This chapter demonstrates how Categorisation Analysis works in practice. The piece analysed here is a simple monophonic Scottish folk tune, which does not make full use of the techniques of Categorisation Analysis but which is appropriate for demonstration purposes.

The Scottish folk tune is *All the Blue Bonnets are over the Border* (Logan, 1947). The score is edited for bagpipe, where a lot of acciacaturas have been added by the editor for the use of bagpipe playing. In here we keep the tune only and do not consider the acciacaturas.

The piece has not got a key signature. However, F and C sharp are implied, and G sounds somewhere between a G and a G sharp. This is the natural tuning of the bagpipes. Figure 4.1 shows the piece.

The analysis is split into the six steps discussed in the previous chapter, *segmentation, feature description of segments, categorisation, relations within and across categories, prototypes and distributional layout*. After the analysis, a section follows where alternative analyses of various steps are shown.

Figure 4.1: All The Blue Bonnets Are Over The Border. Segmentation is shown by the number of segment on top, at the beginning of each segment.

4.2 Segmentation

The segmentation of the piece is shown in figure 4.1. We use a non-overlapping and fixed-length segmentation, where each segment is half a bar long. The first note, the anacrusis, is left out, and the segmentation coincides with the bar lines and the middle of the bar. This is one of the assumptions made in order to keep the analysis simple, since we are not interested in the results at this stage. If the first anacrusis was taken into account, one could find other anacruses in the piece; then the last note of some segments could be considered an anacrusis for the next segment, and we would have to use overlapping segmentation. The features would then be more complicated. However, this case of overlapping segmentation is considered at the end of this analysis, at the section where alternative analyses are encountered.

Here there are 32 segments all together. Each segment takes a number from 1 to 32, according to their temporal order in the piece. Since segmentation is not discussed at length in this thesis, we are not going into more depth on why this segmentation was chosen, and instead we will take it as a given first step.

4.3 Feature Description of Segments

4.3.1 Feature extraction from the segments

It is expected that each analyst will discover his/her own attributes for the analysis. Here, by considering the first segment, we observe the following:

- The melodic contour is steady, unchanged.
- It is just As.
- It is 2 notes long.
- It is a crotchet followed by a quaver.

Moving to the second segment, we can extract the following features. In doing so we keep in mind what differentiates this segment from the previous ones. Therefore we have:

- The melodic contour is moving, in fact it is rising.
- It is A arpeggio notes.
- It is two notes long.
- It is a crotchet followed by a quaver.

So far we see that the first two segments share some features and are differentiated by some other ones. Moving on to the third segment we see more differences:

- The melodic contour is not straight, but "oscillating". Oscillating is the word we are going to use for any non straight melodic movement, that is melodic movement which moves in at least two directions.
- The melodic contour changes direction at least once.
- It starts and ends on the same note.
- It is three notes long.
- It has a dotted rhythmic pattern.

It is interesting to see what happens when one considers segment 6: It is also dotted, but the dot this time is on the second note of the segment, as opposed to the first note of the segment, as it was at segment 3. Therefore here is a new feature that is added to the list and which will make a difference later on for the description of segments. Segment 6's description can be:

- Moving melodic contour, oscillating
- up-down
- starts on the same note that it ends
- three notes long
- dotted rhythmic pattern
- *dotted on the second note*

The procedure continues in the same way. At the end, we have collected the following features for our analysis, which are appended together in the Attribute List of segments.

4.3.2 Attribute List of segments

Table 4.3.2 shows the attribute list that the segments are going to be described at the next step. It is obvious from this table that some features that were found during the description of the segments are left out, while new ones are added. This is explained below.

By having the categorisation in mind, we judge that some of the features are already enough for the future categorisation. If this proves to be wrong, then there will be a re-evaluation of the features and more features will be added. This will be seen in the alternative analysis section. It should also be noted that for each analyst different features and different number of features will be enough. This will be more discussed at the alternative analysis section. Such a feature is whether the notes are part of an A arpeggio, or whether the segment ends at the same note it started with.

Some features are redundant. For example, whether there are two or three notes in the segment; this coincides with whether the segment is dotted or not. In a dotted segment there are always three notes and in a non dotted segment there are always two notes. Also, if there were a feature shared by *all* the segments, then this would be redundant too, since it would not play any role in the categorisation. An example of such a feature would be *register between C4 and C6*, since all the piece is within that register.

The feature that is added is *dotted on the first note*. This is to differentiate it from *dotted on the second note*. It could have been named *non-dotted on the second note*; that would have been equivalent. The dot on the second note of the segment produces a very interesting rhythmic pattern which is very characteristic for this kind of music and it is called *scotch snap*. Another feature that is also added is *non-dotted*. This is again to distinguish the dotted patterns from the non-dotted ones. An alternative way, more economical, of notating features

Melodic contour information
steady
moving
straight moving
up
down
oscillating moving
up-down
down-up
Rhythmic information
dotted rhythmic pattern
non dotted rhythmic pattern
dotted on the first note
dotted on the second note

Table 4.1: The Attribute List for *All the Blue Bonnets are Over the Border* that is used for the further analysis in order to describe the segments.

is discussed in the alternative analyses section at the end of this chapter. This issue is also discussed in the knowledge representation section of chapter 5.

We will find that the above features are adequate for a further meaningful categorisation. However, if the categorisation is not good enough, we can always go back and change the features according to which the segments are going to be classified later on.

Figure 4.2 shows how these features are related to each other. The way that features are connected to each other is described in the previous chapter, and in more detail in the knowledge representation section of chapter 5.

We have two sets of features: the ones concerning **melodic contour** and the ones concerning **rhythm**. Melodic contour can be divided into two opposing cases; either it is *steady* or *moving*. If it is *moving*, it can either move in a *straight*

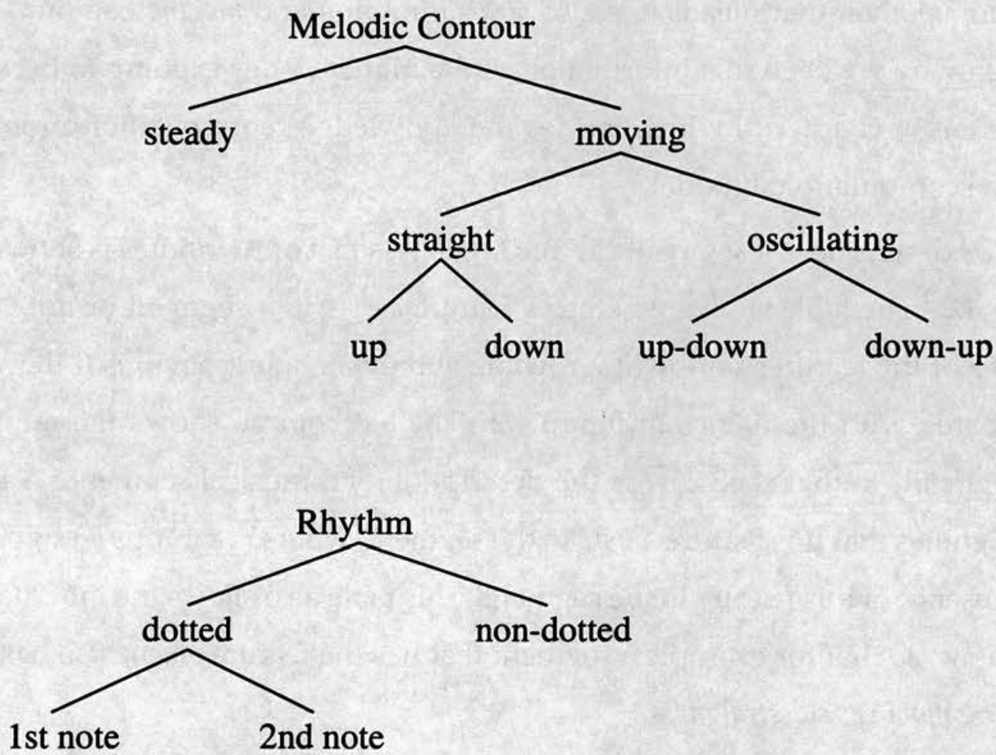

Figure 4.2: The network for the attributes that describe the segments

motion or it can be *oscillating* (as defined above). If it is moving in a *straight* motion it can either go *up* or *down*, while if it is *oscillating* it can either go *up-down*, or *down-up*, at least for the specific piece.

Inheritance is the property of this tree, which means that any property or feature on the lower nodes of the tree carries with it all the properties or features of the higher nodes that it is connected to. Thus, if a segment has the feature *up*, then that means it is also *straight* motion and *moving* contour. The reason why we keep this information of the higher nodes is going to become apparent in chapter 5, which tackles the knowledge representation problem for the computational model.

The description of segments according to this List of Attributes is shown in table 4.2. The table shows whether a feature exists in a segment or not. The names of the features on the first row are abbreviated, but obvious if they are compared with the names in figure 4.2. The left column shows the number of segment, so that each row is the description of a musical segment. A plus (+) signifies that this feature exists in the segment, whereas a minus (-) signifies the absence of this feature in the segment. This table also shows the inheritance relations, so that for example a segment that has the feature *down*, also has the feature *moving* and *straight*.

One might think that not all these columns are needed. An example of redundancy might be the two columns marked *down* and *up*: This information could have been contained into one column only, since whatever has a minus on the *up* feature has also got a plus on the *down* feature and vice versa. Another example is the *oscillating* column: one could argue that it is not necessary since the two *oscillating* cases, *down-up* and *up-down* are covered by two other columns. This alternative way of representation is shown in the next section where alternatives to this analysis are shown. In general however, these are decisions left to the analyst. Here their only role is clarity, but there is a reason

segm	stead	mov	str	osc	U	D	UD	DU	dot	nodot	1std	2ndd
1	+	-	-	-	-	-	-	-	-	+	-	-
2	-	+	+	-	+	-	-	-	-	+	-	-
3	-	+	-	+	-	-	-	+	+	-	+	-
4	-	+	+	-	-	+	-	-	-	+	-	-
5	-	+	-	+	-	-	+	-	+	-	+	-
6	-	+	-	+	-	-	+	-	+	-	-	+
7	-	+	+	-	-	+	-	-	+	-	+	-
8	-	+	+	-	+	-	-	-	-	+	-	-
9	+	-	-	-	-	-	-	-	-	+	-	-
10	-	+	+	-	+	-	-	-	-	+	-	-
11	-	+	-	+	-	-	-	+	+	-	+	-
12	-	+	+	-	-	+	-	-	-	+	-	-
13	-	+	+	-	-	+	-	-	+	-	+	-
14	-	+	+	-	+	-	-	-	+	-	+	-
15	-	+	-	+	-	-	+	-	+	-	-	+
16	-	+	+	-	-	+	-	-	-	+	-	-
17	-	+	+	-	+	-	-	-	+	-	+	-
18	-	+	-	+	-	-	-	+	+	-	+	-
19	-	+	+	-	-	+	-	-	+	-	+	-
20	-	+	+	-	-	+	-	-	+	-	+	-
21	-	+	+	-	+	-	-	-	+	-	+	-
22	-	+	+	-	+	-	-	-	+	-	+	-
23	-	+	+	-	+	-	-	-	+	-	+	-
24	-	+	+	-	-	+	-	-	-	+	-	-
25	-	+	+	-	+	-	-	-	+	-	+	-
26	-	+	-	+	-	-	-	+	+	-	+	-
27	-	+	+	-	-	+	-	-	+	-	+	-
28	-	+	+	-	-	+	-	-	+	-	+	-
29	-	+	+	-	-	+	-	-	+	-	+	-
30	-	+	-	+	-	-	-	+	+	-	+	-
31	-	+	-	+	-	-	+	-	+	-	-	+
32	-	+	+	-	-	+	-	-	-	+	-	-

Table 4.2: Feature description of segments of All the Blue Bonnets are over the Border. The left column shows the number of segment, and the top row shows a syntomography of the name of feature.

when there is a computational model. More on this knowledge representation issue is discussed in chapter 5.

4.4 Categorisation

There are seven categories, named here **a, b, c, d, e, f, g**. Table 4.3 shows the categorisation of the segments into these categories. The top row shows the names of the categories, which are arbitrary, **a** being the first category encountered, **b** the second, and so forth. The numbers, as before, indicate the segments. The way the table is written is in the style of a Paradigmatic Analysis; that means that if read from left to right, the table follows the score. This is also obvious form the numbers of the segments, since they follow the order they are encountered in the piece.

The categorisation carried here is non-overlapping, which means that no segment belongs to two categories. Because the categories are so clear in this piece, there is no need for an overlapping categorisation.

Which features and in which order played a role for this categorisation? Category **a** consists of segments that are steady melodic contour, that is they are stationary. Category **b** and **g** are ascending, the difference being that in category **b** the segments are non dotted whereas in category **g** they are dotted. Category **c** has the down-up segments. Category **e** has the up-down segments. Note that these do not need to be separated further into dotted/non-dotted because they are all dotted. Categories **d** and **f** contain the down segments, category **d** has the ones which are non dotted while category **f** has the ones that are dotted.

Note that we have not taken into account the scotch snap, that is whether the dot is on the first or the second note of the segment. This was an analytical choice made at this point: We decided to disregard this feature. However, it is going to play a role later, during the sub-categorisation, that is the relations within categories.

If the analyst is not happy with the resulting categorisation, he/she can go back to the previous step of the analysis and revise the attributes that were used. For example, if two segments that, according to his/her judgement should have belonged to different categories but here belong to a single category, need to be separated, then a feature that introduces the *difference* of the two segments can be added to the list. For example, here we might have wanted to separate further category **d** into segments that finish in the tonic note (A) and segments that finish in the dominant note (E). We can go back and re-evaluate the my features, in this case that is add this feature. If we want it to affect only category **d**, and not the other categories that might also have segments ending on A or E, then we could specify the feature further into "descending with two notes that finish on A" and "descending with two notes that finish on E". The "descending with two notes" part of the feature specifies that only those segments in category **d** will be affected.

4.5 Relations within and across categories

First relations within categories are described, and then relations across categories.

4.5.1 Relations within categories

For each category we are going to carry out an extra categorisation, internal to the category. Some of the features of the feature list are going to reappear, while new ones might appear for the sub-categorisation. Below each category is analysed separately.

Category a

The two segments that belong to this category, 1 and 9, are identical so no further categorisation is possible.

Category b

Category **b** consists of segments 2, 8 and 10. We can split this category into two subcategories, **b1** and **b2**, according on whether there is an interval of a second or third between the two ascending notes of the segment. Another feature one could consider, and which would give exactly the same sub-categorisation, would be whether the first note is B or C. Category **b1** starts with C whereas category **b2** starts with B. To consider this would be redundant for the categorisation. The analyst might wish to consider it if he is looking for this specific feature across a wide range of pieces, but here this is not the case, therefore one of the two features is enough. Figure 4.3 shows the features and the sub-categorisation of category b.

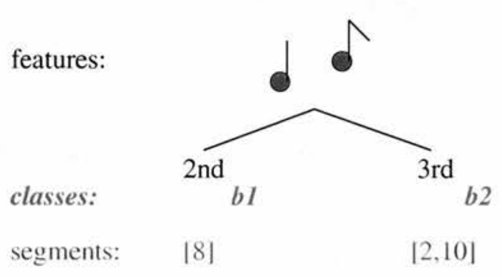

Figure 4.3: Features of the sub-categorisation of category b

Category c

Category **c** consists of segments 3, 11, 18, 26 and 30. It can be divided into two subcategories, **c1** and **c2**, depending on whether the two intervals between the three notes are a fifth or a third. There is also another feature that could have been considered for this distinction, and that is whether the first note is an A or an E. This feature would have given exactly the same sub-categorisation. Figure 4.4 shows the relations between the features and the sub-categorisation.

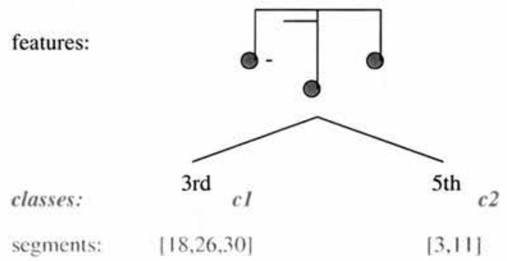

Figure 4.4: Features of the sub-categorisation of category c

Category d

Category **d** consists of the segments 4, 12, 16, 24 and 32. Again here there can be two different features that can be considered for the sub-categorisation, but they both give the same sub-categorisation. The falling interval between the two notes of this category can be either a 4th or a 2nd. The first would create the subclass **d1** whereas the second would create the subclass **d2**. The other two features that would give the same sub-categorisation is whether the first note is and A or a B. Figure 4.5 shows the sub-categorisation and the features for this category.

Category e

Category **e** consists of the segments 5, 6, 15 and 31, and can be divided into two subcategories, **e1** which consists of the segment 5, and **e2** which consists of the segments 6, 15 and 31. All the segments of this category are up-down and dotted. The difference between **e1** and **e2** is that in **e1** the dot is on the first note, while in **e2** is on the second; this last pattern is an example of the *scotch*

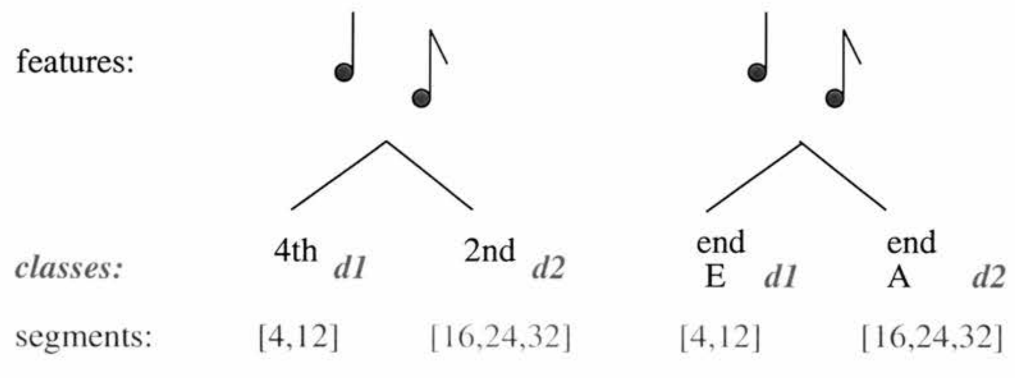

Figure 4.5: Features of the sub-categorisation of category d

snap.

Another categorisation would have been possible if one looked at another feature, whether the segment starts and ends on the same note or not. Both these features are shown in figure 4.6. However, our choice here is that the feature regarding the dot is more significant than the starting and ending note, therefore the categorisation depends on this important feature.

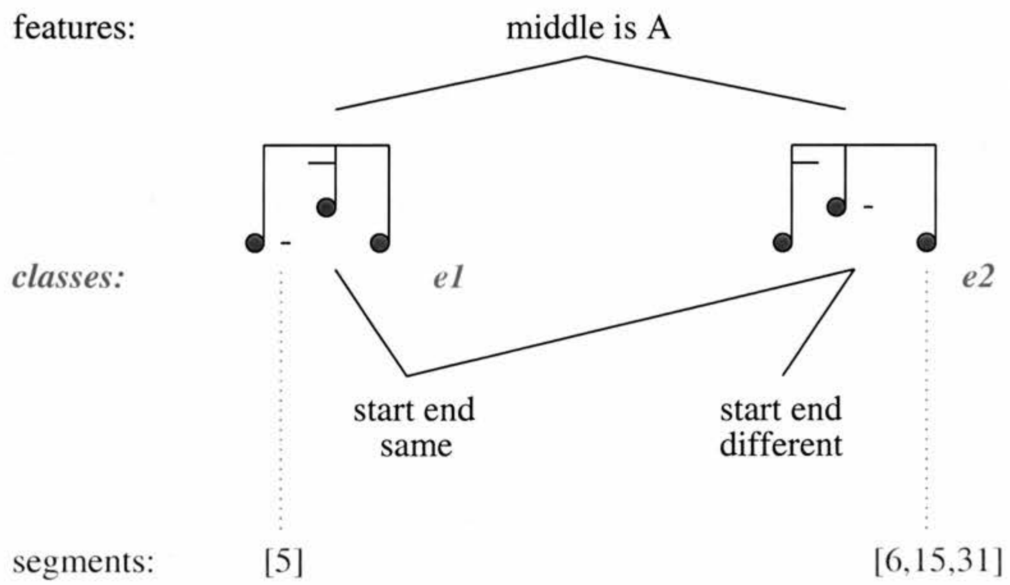

Figure 4.6: Features of the sub-categorisation of category e

Category f

Category **f**, the largest category, consists of segments 7, 13, 19, 20, 27, 28 and 29. They are all dotted segments with three notes that go downwards. There are two major subcategories - those who go down in thirds and those who go down stepwise. However, the stepwise subclass can be further split into starting from F, starting from C, and starting from A. If we carry this "starting from" into the first subclass, then we see that there is only one case, starting from E. So in this case we have sub-subcategories, 4 in total, depending first on the interval and then on the starting note. It is interesting to note that if we had left out the first feature, the categorisation would have still been the same; so we only really need the starting note features to distinguish the 4 subcategories. Figure 4.7 displays this categorisation. The first subcategory, **f1**, has an additional feature: it is an A arpeggio.

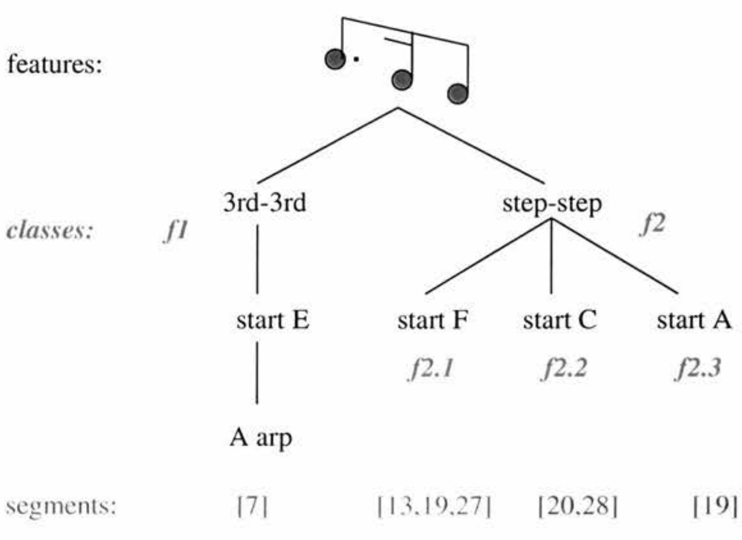

Figure 4.7: Features of the sub-categorisation of category f

Category g

Category **g** is the second biggest category after category **f**, with 6 segments, 14, 17, 21, 22, 23 and 25. It is interesting to note that the features found in this category in order to carry out the sub-categorisation are the same as the ones found with category **f**, the only difference being that intervals that in the first category, instead of starting on E we have starting on A. These two categories which obviously share a lot of features are discussed further in the next section, where relations across categories are observed. Figure 4.8 shows the relations and the sub-categorisation in category **g**.

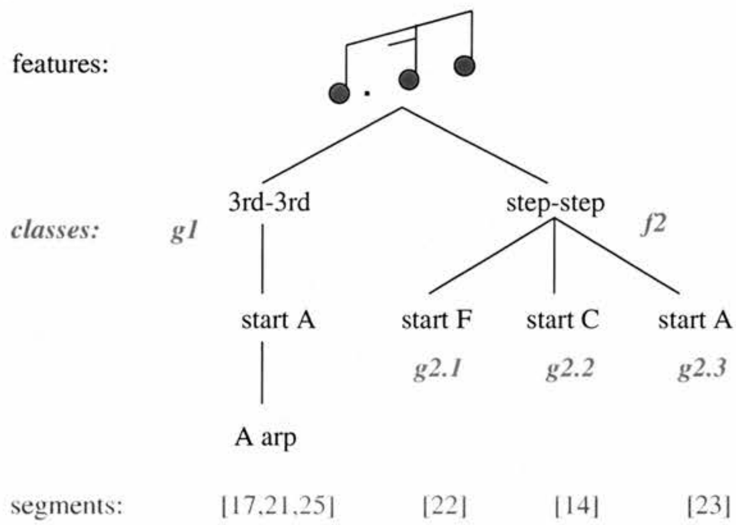

Figure 4.8: Features of the sub-categorisation of category g

4.5.2 Relations across Categories

We look at relations across categories in three ways: the first has to do with the categorisation so far; the second has to do with shared features found above, in relations within categories; the third has to do with producing different categorisations. Each one of these is discussed below.

a. Hyper-categorisation of the existing categorisation

The first way of looking at relations across categories is to take the categories produced in the categorisation step (table 4.3) and see how they can be grouped together to create bigger categories, or hyper-categories, and then group these together to create even bigger categories and so forth until we reach one class, all the segments of the piece together. This is helped by the second step of the analysis, that is feature description of segments, as shown in table 4.2. The result of the hyper-categorisation is shown in figure 4.9; this figure also shows according to which features the hyper-categorisation at each point takes place.

The figure makes use of all the features that were used in the feature description of segments, and no new features. It can be read from top down:

- if we had only two categories, then these would be the segments that have a steady melodic line and the segments that have a moving melodic line.

- If we had three categories, then these would be segments with steady melodic line, segments with straight melodic line and segments with oscillating melodic line. The three categories were achieved by breaking down the moving melodic line of the previous step.

- If we had five categories, these would be segments with *steady* melodic line, segments with *ascending*, *descending*, *up-down* and *down-up* melodic lines. The five categories were achieved by breaking down the *straight* category of before into *up* and *down*, and the *oscillating* category into *up-down* and *down-up*.

- So far no rhythm information has been used. This is introduced here: We get seven categories by keeping the *steady* category and dividing the other four categories of the previous step into two: segments that are

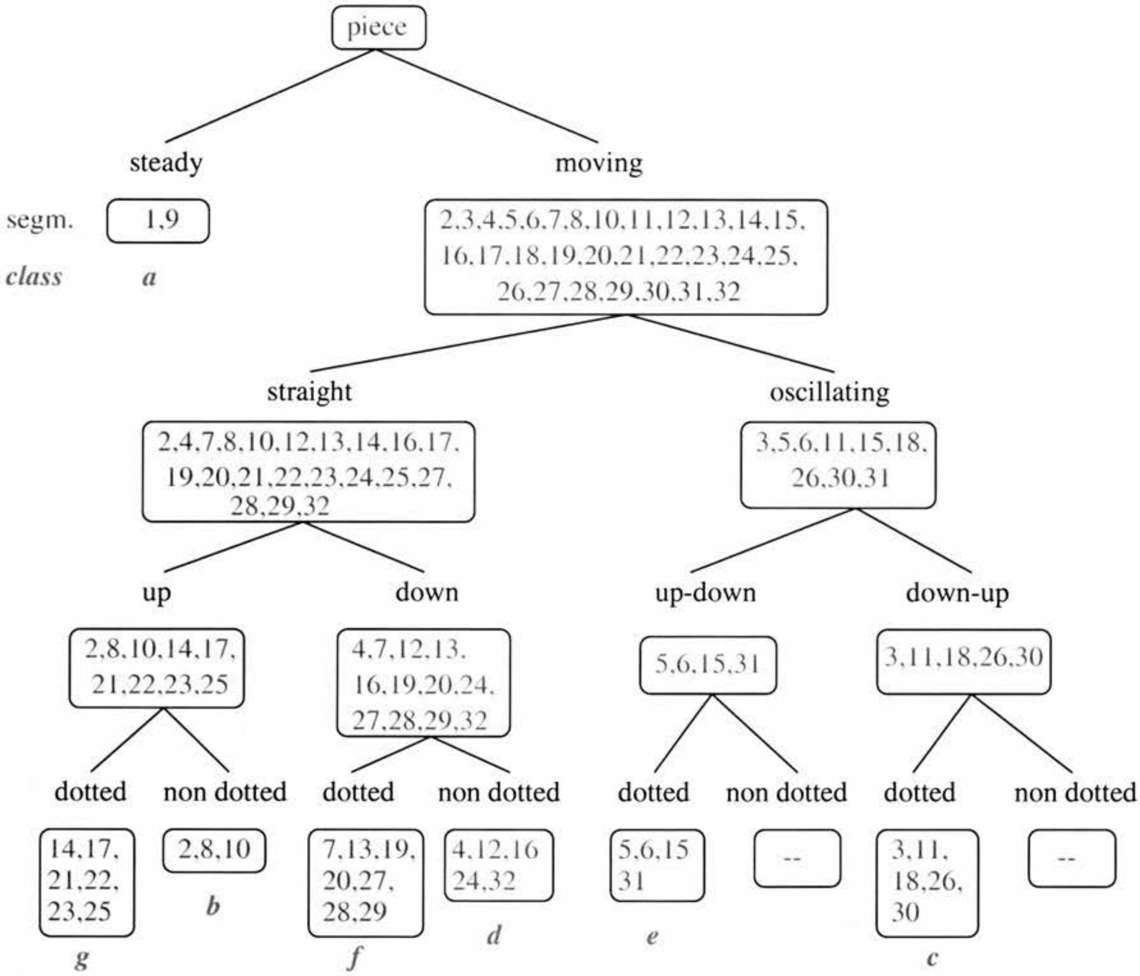

Figure 4.9: Relations across categories: Hyper-categorisation.

dotted and segments that are *non-dotted*. This is not applied to the *steady* category; it could have been applied. The reason was that there are only two identical segments and we would have one resulting category anyway. The result of this step is not nine categories but only seven, since there are no instances for all these categories. The seven resulting categories are the categories that were presented in the categorisation step, table 4.3.

In here there has been an analytical choice: we decided that all the contour features were more important than the rhythmic features and gave us the first categorisations. It could have been the other way round, that is first to categorise with the rhythmic features and then with the contour features. That is shown below, on the third way of carrying out this step.

This step can also be used as an explanation on how we got the final categorisation of the categorisation step, in table 4.3.

b. Sharing of other features

A lot of new features were considered at the previous analysis step, the relations across categories. It is interesting to see which of these features are shared across categories, too. Or, if there are any other features that are shared between the categories. The following remarks concerning shared features can be made:

- Categories **a**, **b** and **d** have all got segments of two notes, a crotchet followed by a quaver. Categories **c**, **e**, **f** and **g** consist of segments of three notes, and all are dotted.

- Categories **c** and **e** have both got segments that either start and end with the same note or start and end with different note.

- Categories **f** and **g**, the two largest categories, share a lot of features: apart from the three notes. dotted, straight moving, they also share the step-step or 3rd-3rd progression, the step-step is also an A arpeggio, the step-step can start from either F, C or A.

- There are possibly other shared features, one can never notate everything. (and this is not the purpose anyway)

c. Other ways of looking at relations across categories

The categorisation of the third step could have depended on other features, or on the same features but with different order, or on the same features that were linked differently:

- If the categorisation of the third step depended on other features, then we would have had a different categorisation. Examples of such features could be the features found in the previous step: starting and ending note are the same, starting and ending note are different, intervals, A arpeggio, and others. A different categorisation is seen in the alternative analyses section.

- If the order of the features were different, for example if the rhythmic features were taken into account before the contour features then the categorisation would have been the same in this case. This is not generally the case though, and with more complicated pieces a different categorisation would have resulted.

- If the same features were linked differently, as for example in figure 4.10, then the categorisation at this case would have been again the same, but in other more complicated pieces it would have been different. In relating the features differently, that means that the features that were inside the tree change, because these are the links between the lowest level features.

4.6 Prototype Discovery

The previous chapter explained how prototypes do not need to be realised musical segments but a collection of the most common features of the set of segments of each class. Our approach to prototypes is explained in more detail

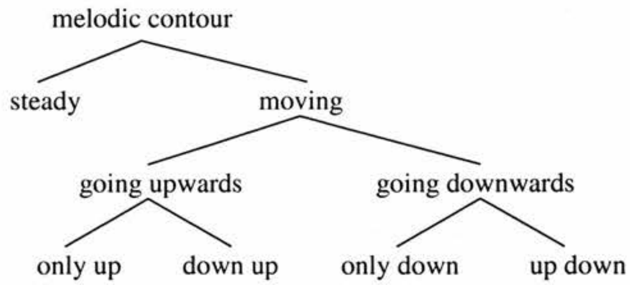

Figure 4.10: A different way of relating the same low level features that were used in the feature description of segments.

in 2.4.2 and 3.3.7. Although in our approach prototypes cannot be realised, in this analysis, because the piece is so simple and it is usually the case that at least half the members of each category are identical, the prototypes can be realised.

- For category **a**, the case is very simple: we have two identical segments, and the prototype is identical to them, too.
- For category **b**, we have three segments, two of which are identical. The prototype is one of these two segments.
- For category **c**, we have five segments, three of which are identical. The prototype is one of these segments.
- For category **d**, we have 5 segments, three of which are identical. The prototype is one of these three segments.
- For category **e**, we have 4 segments, two of which are identical. The prototype is one of these.
- For category **f**, we have 7 segments, three of which are identical, and all the other are different; the prototype is one of these three segments.

- For category **g**, we have 6 segments and three are identical; the prototype is one of these three segments.

4.7 Distributional Layout

Figure 4.11 shows the distributional layout of the seven categories over time in segments. The X axis shows time in segments, and each segment is attributed a category, as shown with the circle on the graph. This kind of information is interesting to compare with other pieces.

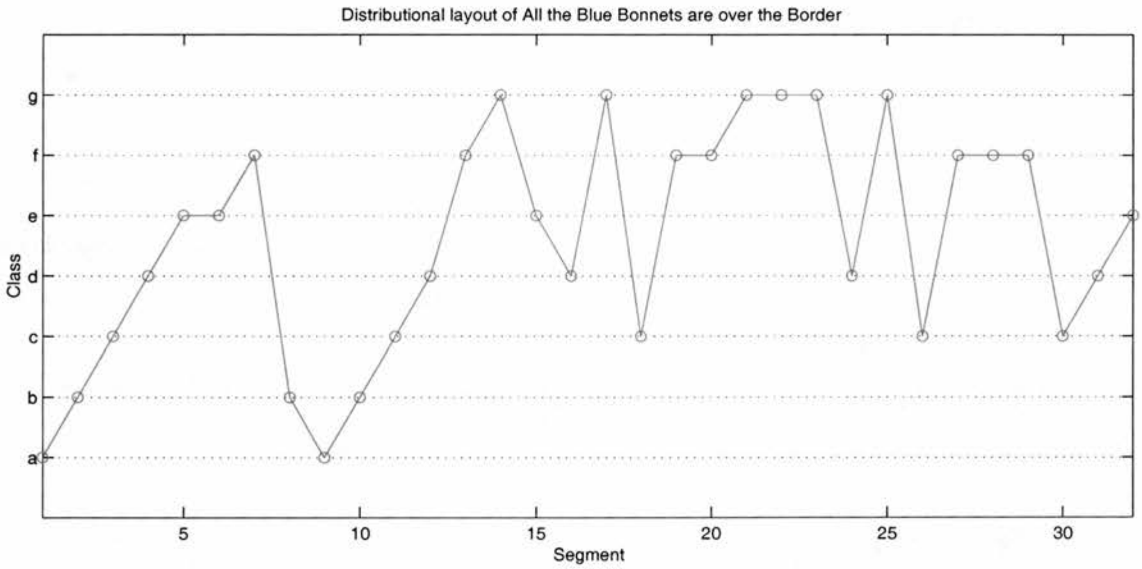

Figure 4.11: Distributional layout of the Scottish tune: x axis shows time in segments and y axis shows categories. The connections between the points in the graph has no meaning. It is there only for visual purposes.

The above observations are sufficient to differentiate between the segments. There is no need to point out features that are common to all segments because that would not make any difference to the categorisation, and our purpose for describing the segments is not the description itself, but the categorisation. Thus, there is no need for the observation that all segments finish with a quaver, simply because *all* of them do. This feature, if included in our list of features, would be redundant.

Similarly, we do not need to say that all segments have got the same overall duration if one overlooks the anacrusis. Also, we don't need to say that some segments are longer when there is an anacrusis because this is captured in the "anacrusis" attribute (ie. whether there is an anacrusis or not).

There is no reason to distinguish between 2 note segments and 3 note segments because this is captured on whether they are dotted or not: all dotted ones are 3 notes, all non dotted are 2 notes long, since the time length of each segment is the same (not taking into account the anacruses, which is a separate attribute of course).

4.8 Alternative Analyses

It has been stressed throughout this thesis that the analyst's freedom is a very important issue, and that each analyst can have his own analytical opinion; indeed, there is not one single correct way of analysing a piece of music (this has been discussed in the introductory chapter). In this section we look at some alternatives that we could have been chosen for the above analysis; these are either different choices or different ways of notating the same choices as above.

There are three alternatives in this section: first, there is an alternative segmentation, and we see the difference it makes in the choice of features. Second, we take the same features of the analysis of the Scottish tune and notate them differently; this makes no difference to the following categorisation. Third, we take two different feature descriptions and we see the role they play for the categorisation.

4.9 Overlapping segmentation

The first difference we are looking at is a different segmentation and how this affects the rhythmic features of the piece. Only a part of the full analysis is carried out.

Figure 4.12 shows the segmentation for the same piece; this time the anacruses that were left out above are taken into account. That means that the last note of a segment is also the first note of the next segment. This kind of segmentation is overlapping. Since segmentation is not the purpose of the thesis we are not going to go into detail on how the segmentation was achieved. It is the analyst's choice to come up with a segmentation, and here it is taken as fixed in order to concentrate on the results a different segmentation has on the other steps.

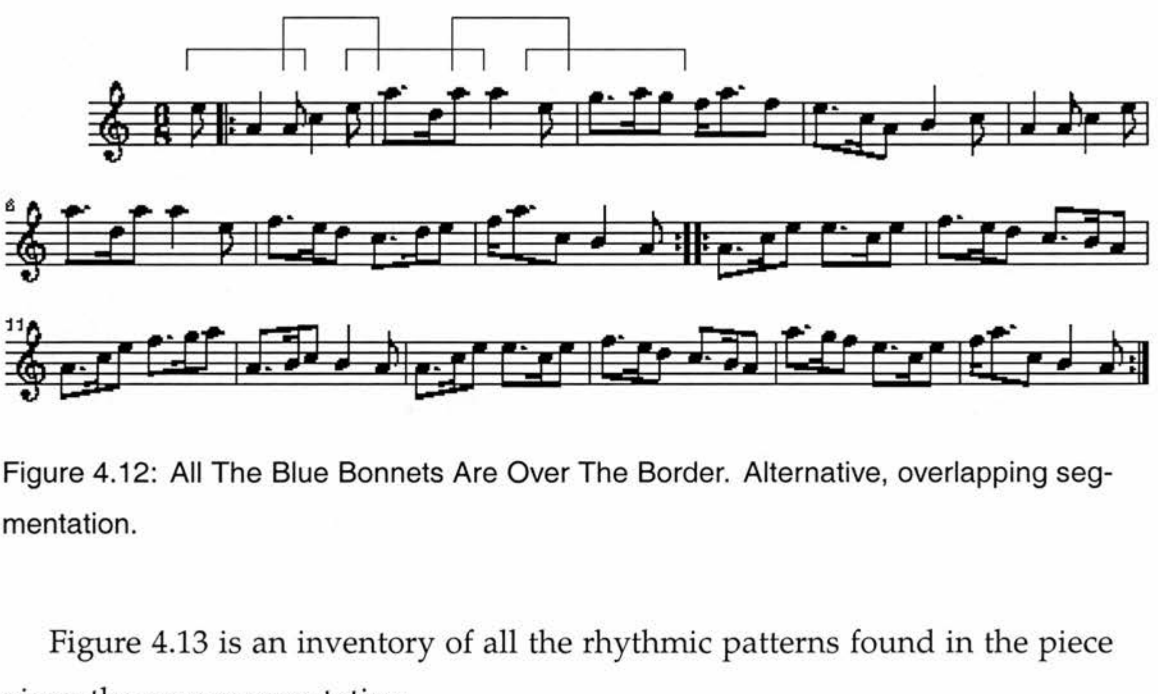

Figure 4.12: All The Blue Bonnets Are Over The Border. Alternative, overlapping segmentation.

Figure 4.13 is an inventory of all the rhythmic patterns found in the piece given the new segmentation.

These rhythmic attributes are connected with each other as shown on the network of figure 4.14.

It is interesting to note that only the "dotted on the first note" segments can

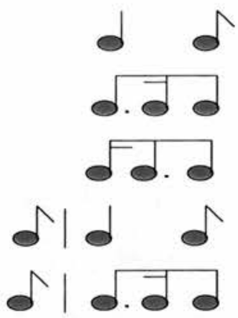

Figure 4.13: The rhythmic patterns

have anacrusis, therefore only these are linked to the anacrusis attribute, and not the 2nd note segments. That is indeed an interesting feature to be traced across other Scottish folk tunes.

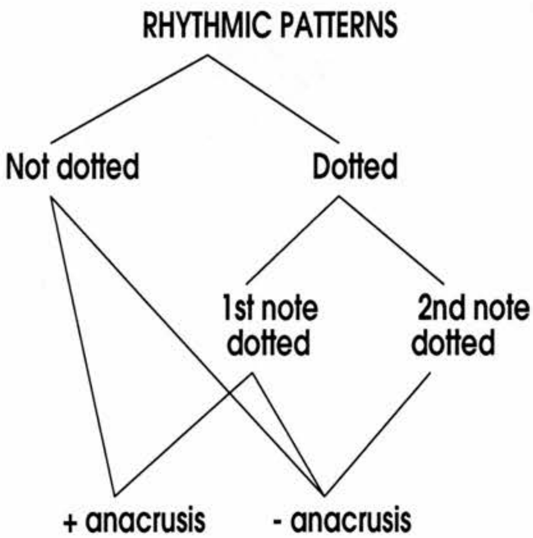

Figure 4.14: The network for the rhythmic attributes that describe the rhythmic patterns

From figure 4.13 we can go to figure 4.14. The other way round is not possible. This is because our description of the segments is not exhaustive, and it is not our purpose to reproduce them from the description. For this we would need a different kind of segment description, for example `note1= E5, quaver` and so forth. This is discussed in detail in chapter 6, after the computational model of Categorisation Analysis has been presented in chapter 5 because for the formal model this issue of features is very critical and needs to be described in that context.

Table 4.4 shows the description of the segments taking into account the anacrusis. Instead of plus and minus this time we use "yes" and "no" depending whether a feature exists in a segment or not.

The pitch attributes are more complicated than the rhythmic ones. In figure 4.15 we see melodic contour, with anacrusis taken into account. Note that the attribute "straight" does not exist any more (ie., all the segments that were "straight" before, they all have an anacrusis, which is not the same note, so they are not "straight" any more). The same things can be shown for the intervals, taking anacrusis into account. Again, there would be many more attributes to take into account. It is important to note that now we have different segments.

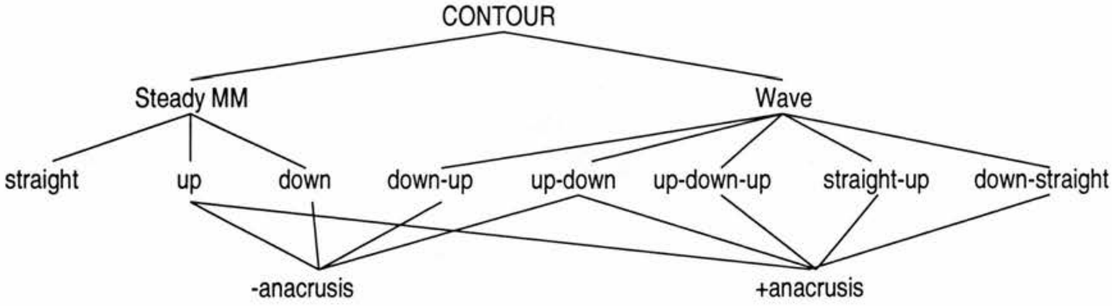

Figure 4.15: Pitch attributes networks

4.9.1 Alternatives in the representation of the feature description

Back to the original segmentation and choice of features, here is a different way for representing the same features in the same segment description. Tables 4.5 and 4.6 show the same information in a more economical way. Here the columns are condensed and opposed features are put together into one column. In that way, the values are not yes or no, plus or minus anymore, but actual features themselves. Again this is an issue of knowledge representation which is going to be explored further in the knowledge representation section of the computational model.

a	b	c	d	e	f	g
1	2	3	4	5		
				6	7	
	8					
9	10	11	12		13	14
				15		
			16			17
		18			19	
					20	21
						22
						23
			24			25
		26			27	
					28	
					29	
		30		31		
			32			

Table 4.3: Categorisation of segments in *All the Bonnets are Over the Border*. The letters on the top row denote names of classes, and the numbers denote numbers of segments, as they were numbered in figure 4.1. The layout is similar to Paradigmatic Analysis.

segm	dotted	anacrusis	dot 1st note	dot 2nd note
1	no	yes	no	no
2	no	yes	no	no
3	yes	yes	yes	no
4	no	yes	no	no
5	yes	yes	yes	no
6	yes	no	no	yes
7	yes	no	yes	no
8	no	no	no	no
9	no	yes	no	no
10	no	yes	no	no
11	yes	yes	yes	no
12	no	yes	no	no
13	yes	yes	yes	no
14	yes	no	yes	no
15	yes	no	no	yes
16	no	no	no	no
17	yes	yes	yes	no
18	yes	no	yes	no
19	yes	no	yes	no
20	yes	no	yes	no
21	yes	yes	yes	no
22	yes	no	yes	no
23	yes	no	yes	no
24	no	no	no	no
25	yes	yes	yes	no
26	yes	no	yes	no
27	yes	no	yes	no
28	yes	no	yes	no
29	yes	no	yes	no
30	yes	no	yes	no
31	yes	no	no	yes
32	no	no	no	no

Table 4.4: Description of the segments of segmentation 1 in terms of rhythmic features.

segm	steady?	straight?	up?
1	yes	-	-
2	no	straight	up
3	no	oscil	-
4	no	straight	down
5	no	oscil	-
6	no	oscil	-
7	no	straight	down
8	no	straight	up
9	yes	-	-
10	no	straight	up
11	no	oscil	-
12	no	straight	down
13	no	straight	down
14	no	straight	up
15	no	oscil	-
16	no	straight	down
17	no	straight	up
18	no	oscil	-
19	no	straight	down
20	no	straight	down
21	no	straight	up
22	no	straight	up
23	no	straight	up
24	no	straight	down
25	no	straight	up
26	no	oscil	-
27	no	straight	down
28	no	straight	down
29	no	straight	down
30	no	oscil	-
31	no	oscil	-
32	no	straight	down

Table 4.5: Description of the segments in terms of melodic contour features.

segm	dotted?	where?
1	no	-
2	no	-
3	yes	1st note
4	no	-
5	yes	1st note
6	yes	2nd note
7	yes	1st note
8	no	-
9	no	-
10	no	-
11	yes	1st note
12	no	-
13	yes	1st note
14	yes	1st note
15	yes	2nd note
16	no	-
17	yes	1st note
18	yes	1st note
19	yes	1st note
20	yes	1st note
21	yes	1st note
22	yes	1st note
23	yes	1st note
24	no	-
25	yes	1st note
26	yes	1st note
27	yes	1st note
28	yes	1st note
29	yes	1st note
30	yes	1st note
31	yes	2nd note
32	no	-

Table 4.6: Description of segments in terms of their rhythmic features.

Figure 4.16 shows FS5. "FS5" is a different contour tree to the one used in the previous examples. The node "oscillating" is missing, and instead we have other nodes, and other priorities.

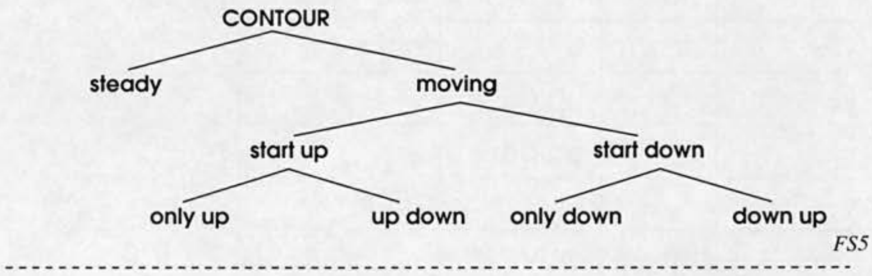

Figure 4.16: FS5

4.10 Discussion

In this chapter we presented a sample analysis of a simple monophonic Scottish tune in order to introduce the similarity and categorisation methodology in practice.

This method proves successful in analysing such a piece. The results we obtain are not interesting because this is such a simple piece. However, if the same method were applied to various Scottish tunes, there could be a comparative analysis between them which would yield interesting results in discovering more facts on the style of Scottish folk tunes. Comparative analysis, especially in folk music, is a good application of this type of method. However, it can present problems if it is carried out manually, like here, without the help of a computer.

One problem can be the massive amount of data that can be produced by undertaking such a task. Another problem is a problem of methodology, if this method is left without being computationally modelled:

segm	moving?	start?	only up?	only down?
1	no	-	-	-
2	yes	up	only	-
3	yes	down	-	down up
4	yes	down	-	only
5	yes	up	up down	-
6	yes	up	up down	-
7	yes	down	-	only
8	yes	up	only	-
9	no	-	-	-
10	yes	up	only	-
11	yes	down	-	down up
12	yes	down	-	only
13	yes	down	-	only
14	yes	up	only	-
15	yes	up	up down	-
16	yes	down	-	only
17	yes	up	only	-
18	yes	down	-	down up
19	yes	down	-	only
20	yes	down	-	only
21	yes	up	only	-
22	yes	up	only	-
23	yes	up	only	-
24	yes	down	-	only
25	yes	up	only	-
26	yes	down	-	down up
27	yes	down	-	only
28	yes	down	-	only
29	yes	down	-	only
30	yes	down	-	down up
31	yes	up	up down	-
32	yes	down	-	only

Table 4.7: Description of segments according to the new feature description.

There has been one point of the methodology which was discussed in the previous chapter, but which has not been illustrated with this analysis: that there is no need to have necessary and sufficient conditions, in terms of attributes, for segments to belong to a category. In this analysis, because it is a simple case, it happened that, although there were attributes shared across categories, there were also defining conditions (attributes) for category membership. This was because the piece was so simple and of homogeneous character.

However, in more complicated cases, where indeed there are no such attributes that define the categories, the categorisation process is much more complicated. If one thinks of pieces of more complicated character, and longer length, then this is not possible, unless one wants to produce a very large number of categories.

Therefore, keeping this process manual by the analyst leads to losing some of the formal disposition of this method. There is a need for a way of formalising this process without restricting the criteria for the classification. The next chapter discusses how to overcome this problem by introducing a computational tool for the categorisation step, and at the same time keeping this analytical freedom.

Chapter 5

The Computational Model

5.1 Introduction

This chapter presents a computational model, or tool, for the Categorisation Analysis of Music (CAM). The model is not fully automatic; the analyst has to play an active role, making musical and analytical choices which constrain the system in order to produce the results. In this way, analytical freedom is preserved, while the method is still formal.

The system architecture is modular, allowing for the substitution of individual modules with alternative ones. Should the final analysis be unsatisfactory, the analyst can re-evaluate his/her original knowledge representation choices.

In our computational approach we focus on two main issues that are central to Artificial Intelligence: knowledge representation and choice of algorithm. First the motivation for a computational approach is discussed. This is followed by a full description of the system: an overview of the architecture, the knowledge representation components, the algorithmic component and the results. The chapter concludes with a critical evaluation and a number of potential further extensions.

5.2 Motivation

The computational model of the categorisation method serves as a tool for analysts, providing a well-defined algorithm for the clustering of segments, without restricting the choice of the classification criteria, which in this case are the musical properties.

The classification process becomes formal, since it is carried out by an algorithm. In cases of music pieces with numerous segments, it would be very difficult to manually categorise segments without inconsistencies, unless the segments were very different from each other and the classes were completely distinct. This is rarely the case in music, where so many properties have to be taken into account.

This work could contribute towards a musical analytical *workbench*, where the analyst could decide on tools for various types of analysis. Various representations, algorithms and system architectures would be at hand. Such a project would not only involve the CAM tool, but also motivic, Schenkerian, and other models for other types of analysis. This idea is described in more detail in the conclusions chapter, where future work is discussed.

By having to be formal in a computational sense, the issue of music knowledge representation has to be investigated in depth for two reasons; an algorithm requires such a representation, and also the results have to be musically acceptable. It will be argued that an inappropriate knowledge representation will give unsatisfactory results. By investigating the issue of knowledge representation, we observe that it is virtually impossible to create a database that contains *all* musical knowledge, so that any piece could be analysed with the same database. Rather, a selection of musical knowledge has to be made.

5.3 A general overview of the CAM system

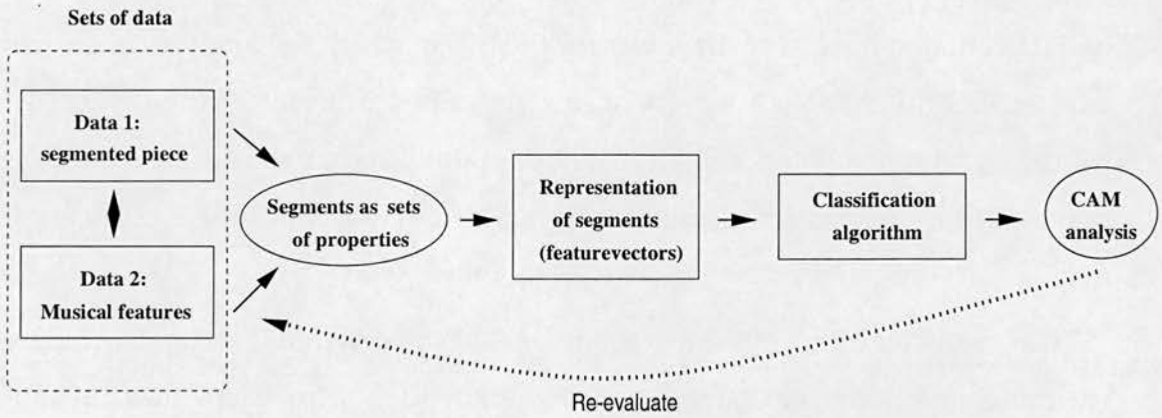

Figure 5.1: A general overview of the CAM system. The squares represent the components of the system, and the ellipses represent results from the components.The modular character allows for the substitution of components with equivalent ones.

The purpose of the CAM system is to take a set of musical segments and classify them according to similarity. The end classification is multi-levelled, that is with classes, hyperclasses and subclasses, and with a property description for each of these. Figure 5.1 shows a general overview of the system.

The central component in a classification system, such as CAM, is the classification algorithm. A classification algorithm takes as input certain objects, and outputs a classification of these objects according to how similar they are. In this case, the objects that need to be classified are segments of music. These segments have to be represented in a suitable way so that the algorithm can understand and accept them as a valid input. The way they are represented constitutes the issue of knowledge representation. The problem of music knowledge representation is particularly challenging.

Two sets of data are needed for the system to work: these are shown in figure 5.1, the overview of the system. The first set consists of the *numbered segments* of the piece. The segmentation has been carried out in advance, either by an analyst or imported from some other algorithm – as mentioned in previous chapters, segmentation is not dealt with in this thesis. The second set of data contains the *List of Attributes* for the piece, which the analyst constructs as described in chapter 3. Segments are described in terms of the List of Attributes, and represented in feature-vectors with binary values. The classification algorithm takes these feature vectors as input and gives the classification results.

The character of the system is modular. This means that it is constructed by various components, or *modules*, that are independent from each other but connected. Each component is responsible for a different subtask of the analytic procedure which can be solved independently. This gives the system clarity, since all the subtasks are clearly describable, but also flexibility and variety in use, since any of these components could be substituted by an alternative one, without affecting the other components. For example, should the analyst wish to use a different classification algorithm, they are free to choose one.

The output of the classification module constitutes a Categorisation Analysis. If the analyst is not satisfied with the results obtained, he/she can *reevaluate* the feature representations. For example, if two segments, which the analyst considers to be different, are grouped together by the model, then the analyst can introduce a feature into data set 2 that distinguishes these segments from one another. Based on the resulting new representation of the segments, a different classification will be obtained. This process is repeated until a musically satisfactory classification is achieved.

The algorithm produces various classifications, from two classes to as many as possible, depending on the input. These classifications are hierarchically connected. Each class in each of the classifications has its own prototype, a prototype being here a list of all the properties that have been met in the class, together with their probabilities.

In short, we envisage a potential use of the system as follows: the analyst takes a segmented piece of music, and from each segment derives various property descriptions. All the properties are appended to a single list, the Attribute List, and then all segments are described against this List, as explained in chapter 3. The descriptions are turned into a feature-vector notation. The algorithm takes this notation as input and outputs a hierarchic classification of the segments, together with prototypes for each class. If the resulting analysis is not satisfactory, the analyst can revise the initial choice of features. This process can repeat until a satisfactory analysis is produced. In the sections below, this procedure and each component of the the system are described in more detail.

5.4 The Music Knowledge Representation

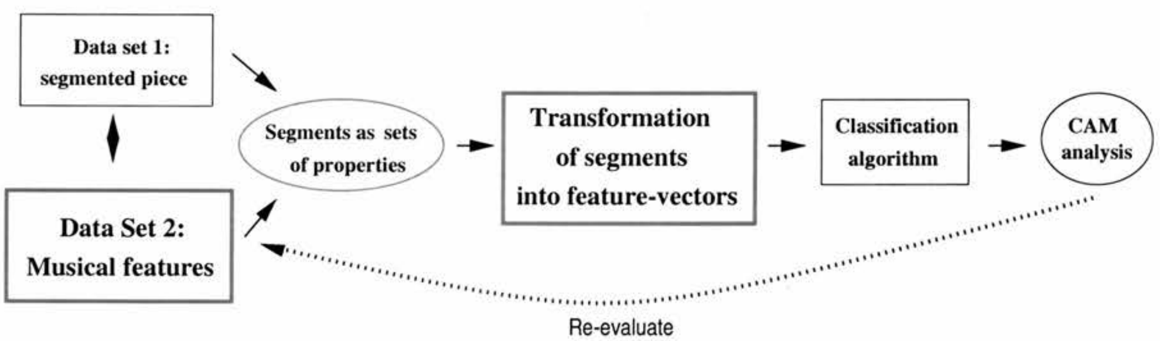

Figure 5.2: The two modules described in this section are in red boxes.

This section deals with the music representation, that is *what* we need to represent and *how* we represent it. Figure 5.2 shows which of the system's components are responsible for the music representation: the second set of data, "musical features", and the transformation of the segments into feature-vectors. Each module is described separately. However, first a theoretical account of various representation choices that have been made is given: we build on the discussion that started in chapter 2 and continued in chapter 3.

5.4.1 Knowledge Representation, tailored to our approach

Section 2.5.2 provided a background on the vast issue of knowledge representation, and section 3.3.3 took the discussion further to what is considered to be a suitable representation in the music domain, as well as explaining how we can get to an Attribute List of a piece. In this section we explain how we formalise the segment description and turn it to an input suitable for the computer. We go into more detail on the kinds of music representation and stress hierarchic representations.

The task of the system is to classify segments of music into categories based on similarity. The questions that naturally arise are what is similarity and how can we measure it. From the general literature, especially on cognitive mechanisms, it appears that almost all kinds of similarity and categorisation accounts are based on some form of descriptive property information, even the prototype theories. Therefore, there is a need to extract this property information from the object of study, in this case the musical segments. Various features, or properties, need to be extracted in order to describe the segments so that these segments can be compared afterwards. One plausible way would be to *describe each segment as a set of properties*. What makes two musical segments identical, similar or different will depend crucially on the property selection and on the way of representing this property selection.

In section 2.4 , following Eysenck and Keane (1998), the terms property, feature and attribute are going to be used interchangeably. In this chapter the term property is going to be used for all attributes and features, while the terms attribute and feature will be reserved for their specific meaning.

Although much work on knowledge representation actually forms part of cognitive studies that look at *mental* representations, the discussion here bears no connection to psychology. However, in the conclusions section of the chapter the cognitive plausibility of the ideas presented here will be discussed.

5.4.2 A taxonomy of musical descriptions and the Musical Features Component

What can constitute a *musical* property has been briefly discussed in chapter 3. Here we attempt a more thorough taxonomy of various kinds of descriptions and properties that can be encountered in music, based on the distinctions discussed above. We divide musical descriptions into three broad groups:

1. single property descriptions,
2. conceptual hierarchies,
3. functional descriptions.

In order to explain these, it might be beneficial to start with a relatively well-founded concept, such as *head*, and look at its potential descriptions. A very crude single property description could be: eyes, nose, lips, ears, hair, skin, and so forth. The description consists of a set of features, without relations between these features. A structural description, or conceptual hierarchy, would further include connections between these: two eyes that contain irises, two ears, and so on. Another type of representations, relational representations, would explain the relation between these features: a nose under and in the middle of the two eyes, ears on the side, hair on top, and so forth. We do not deal with relational descriptions here. A functional description would consist of the function of each feature to the whole: an organ to breathe, an organ to see, and so forth.

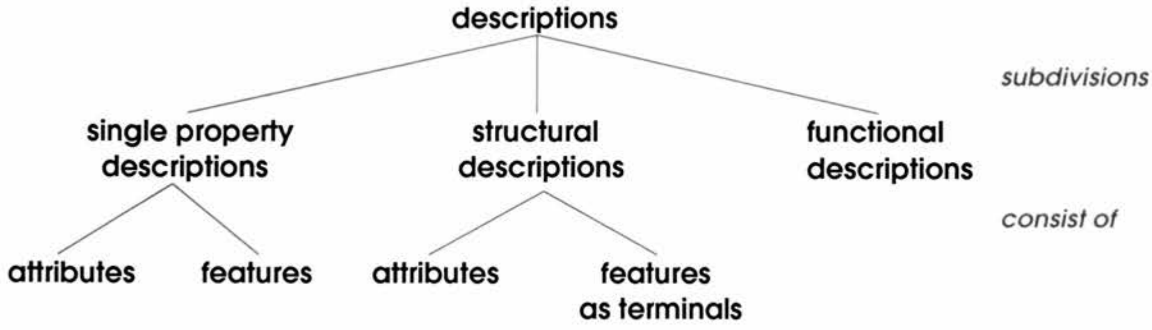

Figure 5.3: The various descriptions of information for similarity judgements

Single property descriptions can either be formed by attributes with many values or by features with yes/no values.

An example of a feature in music would be *grace note*, which could take the value *yes* and *no* depending on whether there is a grace note or not in the specific segment; this is shown in figure 5.4.

Figure 5.4: A feature with binary values

An example of an attribute with many values is *register*, which could, for example, take the values *low, middle, high*, as shown in figure 5.5. In a different case, it could take other values, for example octave numbers if the instrument were piano and if we were interested in such detail. Even more, it could also take conjunctions of values or properties, depending on the instances we find in the musical segments under analysis. An example of this is shown in figure 5.6.

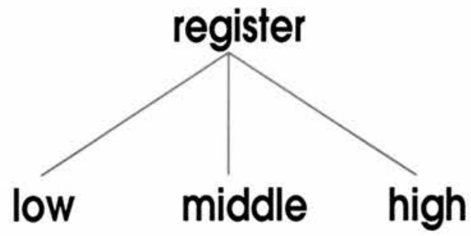

Figure 5.5: An attribute with values, values being disjoint

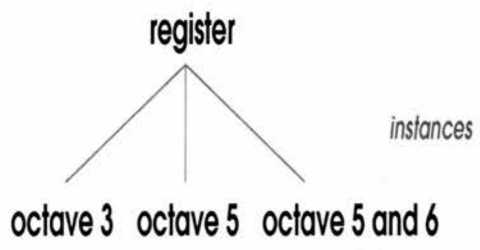

Figure 5.6: An attribute with some values, values being several instances

In **conceptual hierarchies** only attributes are encountered; the only exception is terminal nodes, that is nodes which are not connected any further to more than one node, where features can also be found. Figure 5.7 is a schematic diagram of a potential conceptual hierarchy.

As an imaginary example, we want to describe a musical segment as a melodic movement which is straight and goes upwards. These are in fact two properties: one is straight, and the other upwards. Therefore we have created a two-level hierarchy, with the upper level being straight and the two opposing properties at the lower level being upwards and downwards. This is shown in figure 5.8.

In a single piece it is possible to find segments where the melody does not move, that it is stationary, either because of a long note, or a pause, or repeated same note, and so forth. These properties can also be added to our structure. A stationary melodic movement is opposed to a moving

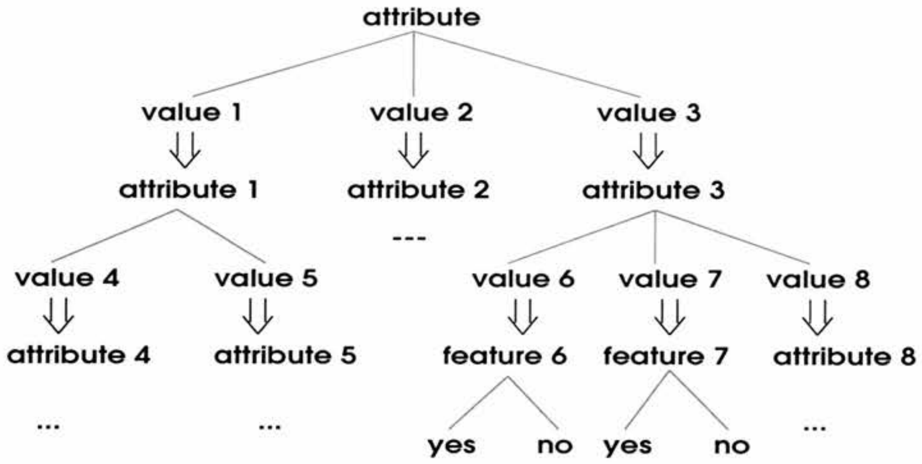

Figure 5.7: A possible structural description - Michalski's "attribute tree".

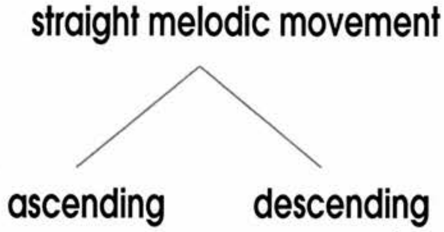

Figure 5.8: Properties of melodic shape linked hierarchically.

melodic movement. A moving melodic movement can be straight, or it can be other things, for example oscillating (that means it includes at least two directions, for example up down, and so forth). Therefore we need to add a higher level to the existing structure of figure 5.8 that will include moving and stationary melodic movement. We can also add the oscillating property, if we have any segments for which this is true in the piece). The structure thus becomes as in figure 5.9.

In the figures above, all properties are linked with inheritance that has to do with the terms themselves and not with the piece. By definition a melody can either move or remain static (Monelle, 1992); a moving

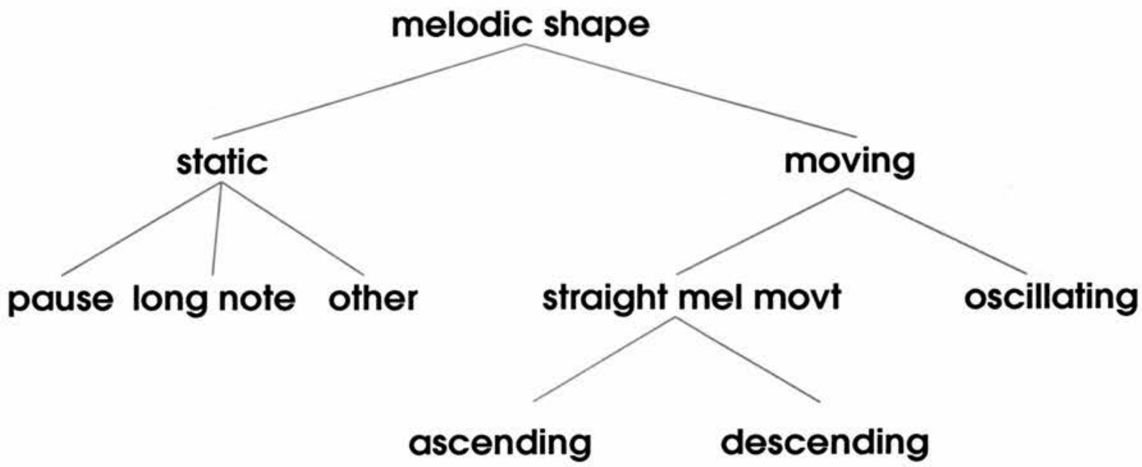

Figure 5.9: feature with binary values

melody can either be straight (1 direction) or oscillating (at least 2 directions); and so forth. However, let us say that out of ascending melodies, half of them use quaver notes and the other half semi-quaver notes, and that we do not have any other instances of ascending melodies in the piece. We can add this to the structure, since inheritance this time is preserved by the instances of the piece. The structure then becomes as in figure 5.10.

Monelle (1991b) uses similar versions of conceptual hierarchies in music, by applying one of the earliest structural linguistic theories, Greimas' *Sémantique Structurale* (Greimas, 1966) to music.

A "property" here can be anything that describes a musical segment: a property (like slow, fast), or a specific pattern (for example a tritone). As mentioned above, anything can be considered as a property, even a specific pattern of notes: for example, concerning the tritone, it is the *use* of the tritone, which can be found in many different occasions and pitches. This kind of argument has led people to call such properties semantic rather than syntactic: the use of the tritone would thus be a

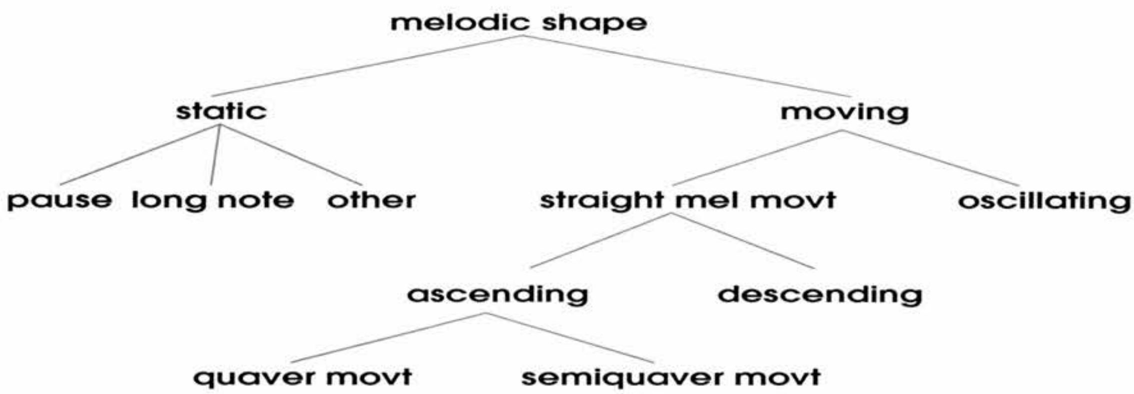

Figure 5.10: feature with binary values

semantic and not a syntactic property.

Functional descriptions describe a musical segment in terms of its function in the piece, for example "main theme" or "coda". These are disregarded in the representation approach because we are interested only in the locally defined musical similarity and not in the functional. For example, if two segments both serve as coda in different places of the same piece, but in terms of musical properties (pitch, rhythm, timbre, texture) are totally different, then it might not be desirable to attribute them any similarity which will enforce them to perhaps share the same class later on. This could be thought of as an artificial rather than actual similarity. It is better to not make any assumptions about functional descriptions that have to do with the placing in the whole piece, and let these perhaps be discovered at the end. An example of such a discovery is found in the next chapter, where the symmetrical structure of a piece by Boulez is revealed.

5.4.3 An attempt at a general module

Our method of representation involves a choice of musical properties or features in order to describe the music. These have to be decided by the analyst.

No one has attempted to use a general module that would be a list of all possible descriptions and aspects of music, appropriate for all kinds of musics and styles. In the computational analysis literature, there are two approaches: in the first case the analyst chooses specific features for his/her analysis, as we do above – for example Cambouropoulos (1998), Conklin and Witten (1995); in the second case there is a fixed representation used for each note, or *event*, where some specific characteristics of the note are described, such as pitch, duration, and so on – for example Hörnel (1998), Rolland and Ganascia (2000). The CHARM system (Smaill *et al.*, 1993), belongs to the second category, but allows for hierarchically built events.

In an imaginary general model which could account for all kinds of music, there should be an inventory of "all musical properties" which would account for all three levels of description:

- the sound level,
- the note level,
- the pattern/segment level.

Although it is possible to imagine a number of possibilities for properties at the sound and note level, in the case of the pattern/segment level, it is impossible to capture every possibility. There can be so many instances and various lengths of patterns and segments, and descriptions that go with them, even at the most basic level (for example, "quaver followed by a minim").

We can thus presume that a general module that can accommodate for all kinds of music is almost impossible. In language, the same problem has been encountered with the creation of large knowledge-based systems. The famous CYC project (Lenat and Guha, 1990) represents a valuable experiment on building common-sense knowledge into a database, which proved too hard a task.

How can one database account for all kinds of music, across all styles, including ethnic musics? Perhaps building such databases is even not necessary. It seems that *context* is the most important factor when deciding on the content of such a knowledge base, context being what creates cohesion in a musical piece, the piece itself, that is the style of the composer and the era.

5.4.4 The "Transformation into feature-vectors" Module

At this stage each segment is described as a list of properties, some of these properties being related between them. This list has to be turned into a *feature vector* so that it can serve as input to the algorithm.

Each of the properties takes a slot (dimension) in the feature-vector. This way of encoding is explained below. We use high-dimensionality, one-out-of-n encoding. The reason for that is that clearer representation and maximum similarity are more important than efficiency in this case. We take various cases of properties, attributes and features.

- for a *feature with binary values*, whether this feature exists or not, only one slot from the vector is needed. For example, for the existence of a grace note in the musical segment:
  - 1 (for the existence of a grace note)
  - 0 (otherwise)
- for an *attribute with several values*, more slots are needed. For example, if looking at instrument register in a flute piece, then a description of segments according to register could be:
  1. low octave (for a segment that makes use of the notes in the lowest octave only)

2. middle octave (for middle octave only)
3. high octave (for high octave only)
4. low and middle (for a segment that makes use of the notes of both these registers)
5. ... and so forth

Not all of these instances of the attribute register might be needed. In the module chosen by the analyst, only those cases that appear in the piece are needed, and out of those only those that the analyst is interested in, or considers to be significant for the analysis.

The slots needed for this feature are as many as its instances: the first slot represents the low octave, the second represents the middle octave, and the third the high octave. The previous four instances are thus turned into:

1. 1 0 0 (low octave)
2. 0 1 0 (middle octave)
3. 0 0 1 (high octave)
4. 1 1 0 (low and middle octave)

- Each feature, or node in the tree of hierarchical organisation is represented as above. The reason that all nodes are represented, and not just the bottom level ones is that the goal is to achieve maximum similarity in the vector forms of segments that share bottom level features.

5.5 The Classification Algorithm

There are numerous classification or *clustering* algorithms in the literature that can be used here. We chose the one gave the best musical results and the format of the results that was appropriate for our work (such as hierarchic classification and prototypes). This was a neural network, unsupervised and constructive algorithm, *Growing Neural Gas* (Fritzke, 1995). Below we describe Growing Neural Gas and its properties that are significant for a CAM analysis.

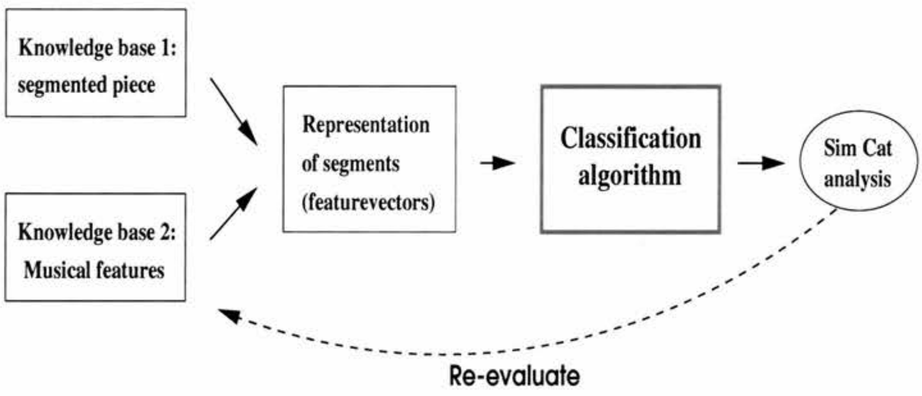

Figure 5.11: The module described in this section

Growing Neural Gas is an unsupervised neural network algorithm that grows units while it learns. Each unit corresponds to the prototype of one cluster. An input signal, i.e., a feature-vector representing a musical segment through n binary features, can be viewed as having a position in the n-dimensional input space, and the units of the network are positioned in the same space. When an input signal is presented, the unit of the network which is closest to it (measured by Euclidean distance) is moved towards this signal by a fraction of the distance to this signal, together with its topological neighbours. The distance between the signal and the winning unit is added to a local error variable of this unit. The winning unit and the second closest unit are then connected by an edge, or the age of the edge is reset to zero, if it already exists. The edges reflect neighbourhood relations between the network units. At each step, all edges in the network are aged, and edges which have reached a pre-defined maximum age are deleted. This process ensures a continuous updating of the neighbourhood relations.

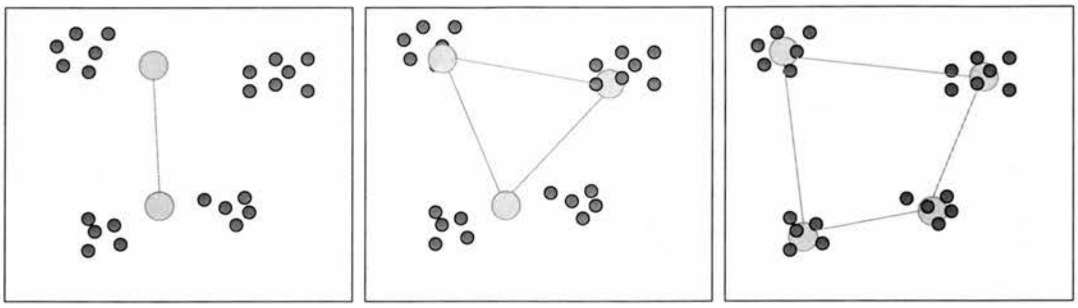

Figure 5.12: How the algorithm works: the small circles represent segments in the input space; the large circles represent the units that the algorithm adds at every set of interations; the green lines represent edges between units.

A new unit is inserted into the network at regular intervals, between the unit with the highest accumulated error and its neighbour with the highest error. The built-up structure of the network reflects the distribution and density of the input signals: the units move towards the input signals, and a high density of inputs in an area will lead to more units being allocated in this area. Figure 5.12 shows the development of a network in a two-dimensional input space with four distinct clusters. The network starts with two units and can therefore distinguish only between the two main clusters. In effect, the network answers the question: if there were two clusters, what would their prototypes be? The fact that the algorithm starts by two initial units is really an artifact of the algorithm: it always needs a winner and a second winner which are both moved towards the input signal and connected with an edge. After a certain number of epochs (presentation of the input signals), a new unit is inserted and the units move to the positions indicated in the second picture. When the fourth unit is inserted, the units distribute over the four clusters.

In principal, insertion of units proceeds forever. The GNG algorithm thus lets the analyst define the level of grainedness of his/her analysis and does not impose *a priori* constraints on the number of clusters. That means that the algorithm produces all the classifications from 2 to as many classes as possible depending on the input, and the analyst can then decide on how many classes he/she wants. Each unit forms a prototype of a cluster, expressed in the probability distribution of the feature values of their cluster members. These prototypes are frequency-dependent, since the position of each unit is updated with each presentation of an input signal. Neighbourhood relations between clusters are expressed in the connections between the network units.

GNG is an especially appropriate algorithm for this type of music analysis since if offers:

- no restriction in the number of output classes,
- a hierarchic classification, when observing the results while the algorithm runs,
- probabilistic values of the properties that appear in the resulting classes,
- it is non incremental; that means that it looks at *all* the segments several times (depending on the iterations) before deciding on the next level of classification,
- it is unsupervised, so there is no need for a priori training of the algorithm.

5.6 The Results

The GNG algorithm is trained on the musical feature-vectors for a large number of iterations until the input cannot be any further subdivided. A new unit is inserted at every specific set of iterations, for example every 200 iterations. This procedure is usually very fast.

During the experiments the various training parameters of the algorithm are changed heuristically, until a good result is achieved. Being satisfied with this, one can experiment with the input representation until an even better result is achieved. Since the algorithm's parameters are constant, the initial feature selection can be revised in order to obtain a better final analysis.

Re-evaluation, as will be shown in the next chapter, plays a very important role in this method. It is possible to tune the results to what the analyst finds acceptable. For example, if two segments that the analyst considers to be different are grouped together by the model, he will introduce a feature that distinguishes the features by one another. Based on the resulting new representation of the segments, a different classification will occur.

5.7 Conclusions and Further Work

This chapter presented the computational model for the Categorisation analysis. The system is not fully automatic, requiring the analyst to play an active role in the analysis by making various choices that the system takes into consideration. The system has a modular architecture, enabling the user to substitute modules for other equivalent ones. Thus it is possible to change segmentation, choice of property descriptions, representation into feature-vectors and clustering algorithm.

The issue of algorithms needs consideration when one has to decide which kind of algorithmic properties are best suited to a specific analysis. Growing Neural Gas algorithm is a very efficient and appropriate algorithm for this kind of analysis, giving a hierarchic classification, no set number of classes and probabilistic properties for each class.

The current work builds towards a larger "workbench" for musical analysis, where the analyst is free to choose between various representations, algorithms and architectures for their analysis. Such a project should involve not only CAM analysis to be available, but also other analytic techniques.

Other future extensions involve the development of a cognitive model based on the current work. A cognitive model of the analyst will be developed where the categorisation and representation procedures are going to be modelled. This will require the use of psychological experiments and the building of a cognitive model based on the CAM architecture which will reproduce the existing results and predict further results. IT is the predictive power of the model that will give it its cognitive plausibility. At a second step, the inner architecture and procedure will be broken into subtasks and these will be further experimentally tested and computationally modelled. A lot of insight can be obtained by similar work carried out in language processing – for example see Hahn and Ramscar (2001).

However, the first further extension would be an automatic extraction-of-properties module that would offer suggestions to the analyst (and not replace them). Such a module could make use of well-known feature extraction algorithms, which would be adapted to the specific problem.

Finally, one could keep investigating further the impact that varying the initial representation has on music categorisation. This might be the future of music categorisation: to predict and distinguish between good and bad representations by looking at resulting categorisation analyses. This is the role that music analysis can play for artificial intelligence advances.

Using the CAM system can also be a way of evaluating various segmentations: a bad segmentation will give equally bad categorisation results. A segmentation that makes sense will give sensible results. One reason for this is that a good segmentation, amongst other criteria, will be based on repetition, and repetition is captured by the categorisation results.

The next chapter presents some analyses obtained by using the CAM system. The various further extension of the model, briefly mentioned above, are discussed in more detail in the Conclusions chapter.

Chapter 6

Boulez and Debussy: Two Analyses

6.1 Introduction

The approach described in the previous chapters is here applied to two pieces, with different tasks in mind:

- *Parenthèse* from Pierre Boulez' Third piano sonata. This is a challenging serial and non-monophonic piece, and the task is to produce an analysis that suggests new findings about the piece and that demonstrates the structure, as this is described in writing by the composer (Boulez, 1975).

- *Syrinx* by Jean Claude Debussy for solo flute. The tasks here are to investigate Nattiez's famous 2nd analysis (Nattiez, 1975), to examine in depth the criteria to produce an analysis of this piece and to demonstrate how the re-evaluation procedure can refine the results of the analysis.

The rest of the chapter is organised as follows: there are two main sections, each devoted to a different piece. Section 6.2 briefly describes Boulez' Third piano sonata and *Parenthèse*, and discusses the challenges that this piece poses to the analyst. Then, we explain in detail how our formal approach to categorisation works in this case, including segmentation of the piece, representing the segments in terms of musical features, and clustering these representations with a computational algorithm. Section 6.2.3 describes the categorisation experiments that were carried out, and the results of these experiments are presented in section 6.2.3. In section 6.2.4 we discuss these results and suggest directions of future research. The work presented in this section is based heavily on Anagnostopoulou and Smaill (2000).

In section 6.3 we discuss the Syrinx analysis. This section originated from Anagnostopoulou and Westermann (1997), and has here been developed to study the various features used for the analysis and how these influence the final result.

6.2 Boulez' *Parenthèse*: A New Analysis

6.2.1 Boulez' Third Piano Sonata

According to Stoianova (1978), Pierre Boulez' Third piano sonata is based on difference as much as on similarity, in that there are various strong relationships between the movements. *'Repetition is vital, although it is "a repetition-difference" within the circumstances of the serial writing [...] In reality, it is a different kind of repetition, which is the principal generator of dodecaphony and serialism.'* Stoianova here talks about the compositional technique of a serial composition, where the composition is based on the series and its various subsets when constructing the piece. Segments of music that are based upon the same set have some resemblance which might not be obvious, ie. not *iconic*, but which is an underlying one.

Parenthèse consists of 6 fragments of music that are obligatory to play, and in between them are 5 fragments of music in parentheses, which are optional to play. According to Stoianova [Stoianova, 1978, p.140], *Parenthèse* is the

> *'microcosm of the symmetrical structure of the whole sonata. The presence of the obligatory and optional (in parentheses) fragments implies the co-existence of two symmetric structures: a circular symmetry of the obligatory groups and another similar one of the groups in parentheses.'*

In order to capture all the aspects of its structure, the study of the entire piece can be split into three parts: first, the analysis of the six obligatory fragments, second, the analysis of the optional fragments in parentheses, and third, the relation between obligatory and optional fragments. Here we demonstrate a full analysis of the first part, that is, the obligatory fragments of the piece, at least to the degree that a "full" analysis might be possible in music.

Within *Parenthèse*, one can observe various similarity relations between its segments, that can be used at a later stage for the analysis: first, the dodecaphonic "repetition-difference" that Stoianova talks about, which is based on the use of pitch class sets, and second, the more obvious similarity relations in musical properties such as rhythm and tempo, tonal centres, intervals, contour, and way of playing.

The method of analysis that we present in this thesis aims to show how these relations can be achieved. The aim of this analysis is, on the one hand, piece-specific: to demonstrate the structure of the obligatory part of the piece, the "circular symmetry" that Stoianova describes, and to see which attributes play an important role. On the other hand, a more general aim is to demonstrate how the Categorisation Analysis, shown to work for monophonic pieces in the previous chapters and also in Anagnostopoulou and Westermann (1997), can be applied to a non-monophonic, atonal piece of music with very rich internal relations.

6.2.2 The Analysis

First the piece is broken down hierarchically into smaller units, and then each of the segments is described as a set of properties. By using the term *piece*, we henceforth mean the obligatory fragments that are analysed here. The description of the segments is then turned into an appropriate computational input in the form of feature vectors, and the classification algorithm takes this input and produces a hierarchic classification of the segments. The result of this process is a categorisation analysis that makes similarity relations explicit. In the following sections, we describe each step in detail.

Segmentation

In most formal methods of analysis, the music piece is first split into segments. The precise way in which the piece is segmented has a profound influence on the outcome of the analysis. In *Parenthèse* , segmentation is an easier task than for most pieces, since in most places the segmentation points are clearly indicated by the composer. We define segment boundaries:

- at the beginning and end of the fragments in parentheses,
- where the piano stave is marked by the composer with *V* sign to denote a break point,
- where there is a more or less obvious change of texture, that is, between segments 2a and 2b, and 4c and 4d. This also coincides with the change of a pitch-class set, and this segmentation therefore corresponds to the so-called imbrication method (Forte, 1973).

The resulting segmentation of the piece is shown in figure 6.1. We denote the obligatory fragments with numbers 1,. . . , 6. These fragments are then further divided into segments 1a, 1b, 1c, 2a, 2b, 2c and so on.

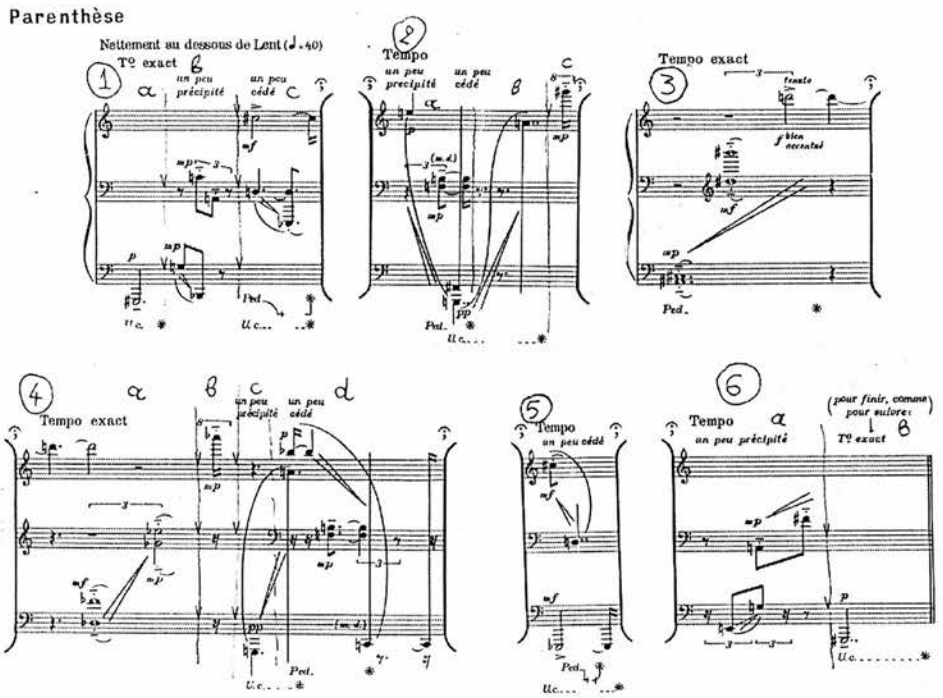

Figure 6.1: The obligatory fragments of *Parenthèse* and their sub-segments.

In the following experiments, we use three levels of segments: the undivided high-level segments 1,..., 6, the low level segments 1a, 1b, and so on, and an intermediate level where we combine certain adjacent low-level segments: for example, the low level segments 1a and 1b form the intermediate level segment 1ab. By this we hope to capture similarities that exist between the different segmentation levels.

High Level: Obligatory fragments	1,2,3, etc
Middle Level: Segment combinations	1ab, 1bc, 2ab, etc
Low Level: Segments	1a, 1b, 1c, etc

Description of Segments as Sets of Properties

Like the choice of segmentation, the choice of properties to describe the segments has a profound influence on the results of the computational classification: the algorithm groups the segments according to their similarity, and this similarity is determined by the property values for each segment. What makes two music segments identical, similar or different will be defined by the property selection, and on the way of representing the properties. That means that two segments, although they might be different in terms of musical notation and sound, in here might be identical if they have the same properties. Whereas the choice of properties is made by the analyst, the categorisation method of analysis shows precisely how this choice influences the resulting analysis.

In developing a set of properties, a segment is analysed in terms of various musical properties that seem important for its description and for its differentiation to other segments. Then, all properties that have been chosen for the segments are combined into a list, and each segment is described in terms of this list of properties.

The description of a segment by a list of properties is not complete: it is impossible, based on the properties, to re-create the particular segment they describe. This is because not all aspects of the music are formalised into properties. Instead, the properties contain all information about a segment that are considered important for the further analysis of the piece. Different analyses warrant different properties: in the case of a rhythmic analysis, one would describe the rhythmic properties of each segment in detail, and in an analysis aiming to compare certain features across a wider music repertoire would only emphasise those specific features.

Two kinds of properties can be used for describing a musical segment in this piece:

- properties that are true for a part of the segment, for example, the existence of a specific interval in the segment, and
- properties that are true for the whole of the segment, for example a rising melodic movement.

In our approach we mainly make use of the second kind of properties, with the exception of specific rhythmic and intervallic patterns that describe merely part of a segment.

Table 6.1 shows the properties that we use in the analysis, and the segments in which they are found. The properties considered here are:

- The existence of various pitch-class sets and certain common subsets that they share. The composer has chosen sets that are very similar to each other in terms of common subsets, and we make this similarity relation explicit by introducing these subsets as features. In order for a pitch-class set to be true for a segment, *all* the notes of the segment have to belong to the pitch class set. Table 6.1 uses a separate notation for when this is not the case.

property	1a	1b	1c	2a	2b	2c	3	4a	4b	4c	4d	5	6a	6b
3-1(12)			y			y						y		
4-1(12)	y-	-y-	-y											
7-2				y-	-y-	-y			y-	-y-	-y			
6-9				y-	-y			y-	-y	y-	-y			
5-2	y-	-y					y	y					y-	-y
5-5				y							y			
all								y-	-y-	-y-	-y			
inv 012	y-	-y	y	y	-y-	y	y	y	-y	y-	y	y	y-	-y
inv +3	y-	-y	-y	y	-y-	-y	y	y	-y	y-	y		y-	-y
inv +5	y-	-y		y-	-y-	-y	y	y	-y	y-	-y		y-	-y
inv +7				y	-y-	-y		y-	-y	y-	y			
longn	y		y				y	y				y		y
Q,Q		y											y	
4note		y											y	
SQdot						y			y					
triplet		y		y							y		y	
exact	y						y	y	y					y
précip		y		y						y			y	
cédé			y		y	y					y	y		
mf+			y				y	y				y		
cresc	y-	-y-	-y				y					y	y	
dimin			y					y					y-	-y
steady	y	y		y-	-y-	-y			y	y	y			
G#/Aflat	y					y			y					y
G,G#,A	y-	-y			y-	-y			y-	-y-	-y		y-	-y
D							y	y				y		
C#,D,D#			y				y	y				y		
semit		y	y	y	y-	-y	y	y			y	y	y	
tritone		y		y			y				y			
third		y		y			y	y					y	
wob		y											y	
down1			y					y				y		
down2				y							y			
up2							y							

Table 6.1: The lowest-level obligatory segments (1a, ..., 6b) and the properties that are true for each segment. When a property exists in a segment, then this is marked by a "y". When there is a property that is true for a bigger segment but not for the low-est level, then this is marked in the lowest-level segments that the bigger segment is made from, by using "y-", "-y", "-y-", according to which adjacent the property is shared with. The first part of the table contains the pitch-class sets and their common subsets, the second part contains the rhythmic patterns, the third part contains the directions by the composer on tempo and way of playing, the fourth part contains tonal centres and specific intervals and the last part contains contour information.

- The existence of various rhythmic patterns. These, in contrast, do not require for all the notes of the segment to belong to the specific rhythmic pattern. However, all the notes of the pattern have to be found sequentially in the right order in the segment. These patterns are shown on the second part of table 6.1: *longn* stands for long note, that is a white (non-filled) note, *Q,Q* stands for quaver followed by quaver, *4note* stands for two sets of quaver notes, perhaps as triplets, that are intermingled, and *triplet* has the obvious meaning.

- Tempo and dynamic descriptions. The composer is very specific about which tempo and dynamic descriptions he uses, and these are important for the distinction of the segments and the overall structure of the piece, so in a classification task they should be part of the segment description attributes. The third part of table 6.1 shows these descriptions and where they appear.

- Tonal centres, which in this case are single tones rather than keys, and relations between tones, significant intervals that the composer seems to favour. These are on the fourth part of the table.

- Contour information is displayed at the final section of table 6.1. *wob* stands for "wobble", that is up and down movement, in no matter which direction and how many times, *down1* stands for downwards movement one time, *down2* stands for downwards movement twice and *up2* stands for upward movement twice. For these to be true, all notes of the segment have to comply to the specific contour feature.

Table 6.1 also shows how each segment is "translated" from musical notation to a set of properties. The reason for this transformation is to achieve, at a next step, a consistent classification. Describing the segments in terms of properties (*cf.* table 6.1) results in a 34-bit feature vector for each segment, making it thus appropriate computational input for the classification module.

As it has been explained above, there are two kinds of properties in this piece - the properties that are true for the whole of the segment and the properties that are true for part of the segment. Some of the properties that are true for a whole segment sometimes actually continue at the next segment or segments. In order to notate this phenomenon, the sharing of properties across adjacent segments, we use a specific notation: *y* where the property is true for the segment alone, and *-y*, *y-*, *-y-* if the property continues on the left, right or both sides of the segment respectively. This is important because there is a difference of meaning in three segments having all y, y, y as properties and the three segments having y-,-y-,-y as properties. For example, in the case of a PC set, the first case denotes a repetition of the whole set in each of the three segments, whereas in the second case all the notes of the three segments are needed to form the set.

This is important in our study because we use more levels of segmentation than the lowest level that table 6.1 presents.

Classification

The classification of the segments, that are now represented as feature vectors, is carried out with the GNG algorithm and its parameter values as described in the previous chapter. The differences we get to the traditional Ruwet/Nattiez Paradigmatic approach are that the classification process depends totally on the choice of properties, the representation and the algorithm, thus avoiding unfounded results. The classification proceeds in an approximately hierarchic way, from the whole piece being considered as one class to each segment being considered as a separate class. An example of hierarchic classification is shown below at the results. Instead of paradigms, the algorithm develops probabilistic prototypical values of the class properties, directly showing similarities and differences between the classes.

6.2.3 Experiments

We performed four experiments:

In the first experiment, the classification algorithm was trained on the feature vectors that represent the segments on the smallest level only: 1a, 1b, 1c, 2a, 2b, 2c, 3, 4a, 4b, 4c, 4d, 5, 6a, 6b. The properties that stretch over adjacent smallest-level segments were not taken into account.

In the second experiment, the algorithm was again trained on feature vectors representing the smallest-level segments, but this time they were enhanced with those features that stretch over segment boundaries. For example, if segment 1 has a property *a* that is not reflected in its sub-segments 1a, 1b, and 1c, then here these sub-segments inherited this global feature.

In the third experiment, all segmentation levels were represented in parallel and the algorithm was trained on the full set of lowest-level segments 1a,...,6b, the highest level segments 1,...,6, and middle-level segments such as 1ab, 4bcd, and so on. In contrast to experiment 2, the lowest-level segments were only represented by their own properties and not the shared ones.

In the fourth experiment, we considered only a selection of eight segments drawn from all the levels: 1ab, 1c, 2, 3, 4a, 4bcd, 5 and 6.

By comparing the developing network architecture over a period of insertion of units, we were able to observe the hierarchy of classes.

Results

Table 6.2 shows the results of experiments 1 and 2, when the number of classes is 5. In computational terms this is when the algorithm has inserted five units, each after 200 iterations. Table 6.3 shows the results in the same two experiments, when there are 7 and 8 classes, and respectively 7 and 8 units inserted by the algorithm.

Class	Exp 1	Exp 2
Class I	2a, 4d	2a, 2b, 2c, 4b, 4c, 4d
Class II	3, 4a	1c, 5a
Class III	1a, 2b, 2c, 4b, 4c, 6b	1b, 6a
Class IV	1b, 6a	3, 4a
Class V	1c, 5a	1a, 6b

Table 6.2: The experimental results in the two first experiments when the number of classes is 5.

Class	Exp 1	Exp 1	Exp 2
Class I	2a, 4d	2a, 4d	2c, 4b
Class II	3, 4a	3, 4a	1c, 5a
Class III	1a, 6b, 4b	1a, 6b	1b, 6a
Class IV	1b, 6a	1b, 6a	1a, 6b
Class V	1c, 5a	1c, 5a	3, 4a
Class VI	2b, 4c	2b, 4c	2a, 4d
Class VII	2c	2c	2b, 4c
Class VIII	–	4b	–

Table 6.3: The experimental results in the 2 experiments when the number of classes is 7 or 8.

The results of experiments 1 and 2 are all intuitively acceptable, although those from experiment 2 seem slightly better. In table 6.2, at the results of experiment 1 with 5 classes, 1a, 2b, 2c, 4b, 4c, 6b belong to the same class. This classification would have been better if segments 1a and 6b were in a different class from the others, since they are characterised by the use of long notes whereas the other segments contain shorter notes. This difference could be enhanced by introducing an extra feature *note-length* in the list of properties describing the segments. This is an example of re-evaluation of the segment descriptions.

Table 6.3 shows the classification for experiment 1 with 7 and 8 classes. Here the same segments are separated into three classes when the overall number of classes is 7. Therefore a bigger number of classes produces more satisfactory results in this case.

Whereas experiment 1, which does not incorporate properties that stretch over segment boundaries, emphasised the iconic similarity between segments, in experiment 2 the structural similarity between segments is enhanced due to the added more "global" features relating to higher-level segments. Here, all the subsegments of segments 2 and 4 are in the same class. Even though the iconic similarity of these segments is low (for example 4b and 4d), they both share the global properties of segment 4.

Figure 6.2 shows the progression of the classification in experiment 2 from 2 to 10 classes. This is an interesting example of hierarchic classification, which shows the symmetrical structure of the piece.

In experiment 3 all levels of segments are taken into account. The results for 5 and 8 classes are shown in table 6.4. In this case we often get segments and their subsegments classified in the same category, since they share many of their properties (for example, segments and subsegments of 2 and 4). This problem cannot be avoided in such a setting and the results need further interpretation in order to be valid, for example some sorting mechanism. For this reason, 5 classes seems to be too few classes for an acceptable classification. When the number of classes increases to 8, the results improve: 3 and 4a are correctly classified into a category of their own, and the same holds for 1b and 6a. It is interesting to see segment 4 on a category of its own, since it is the

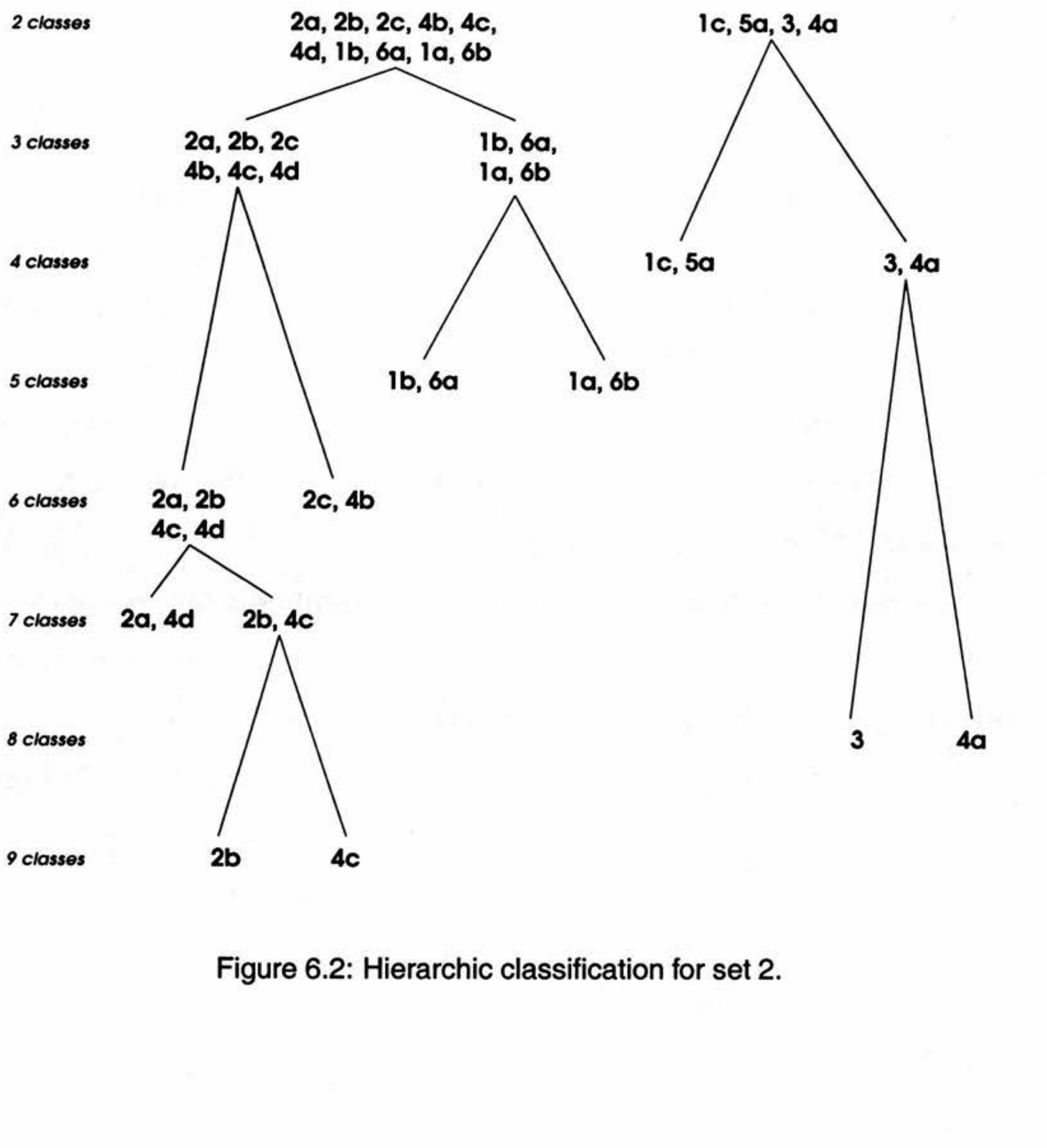

Figure 6.2: Hierarchic classification for set 2.

Class	Exp 3, classes 5	Exp 3, classes 8
Class I	1c, 3, 4a, 4, 4ab	1c, 5a, 4ab
Class II	1a, 2b, 2c, 4b, 4c, 6b, 2bc	1a, 2b, 2c, 4b, 4c, 6b
Class III	2a, 4d, 2, 2ab, 4cd, 4bcd	1, 6, 1ab, 1bc
Class IV	1b, 6a, 1, 6, 1ab	2a, 4d, 2, 2ab, 4cd, 4bcd
Class V	1bc	4
Class VI		1b, 6a
Class VII		2bc
Class VIII		3, 4a

Table 6.4: The experimental results in the third experiment when the number of classes is 5 and 8.

longest segment of all. Segments 2 and 4bcd are placed in the same category and are an example of similarity across the different segmentation levels.

Experiment 4 is an outcome of the interpretation of the above results. It is the simplest experiment because we consider only a selection of 8 segments across levels. These are chosen in order to show the structure of the piece that was almost revealed with the previous experiments; here the aim is to show this more explicitly. Table 6.5 shows the resulting classification when having 4 classes: the first and last segment, 1ab and 6, are classified together, and the same is true for 1c and 5, 2 and 4bcd and 3 and 4a. These segments are almost mirror images of each other, and define the symmetrical structure of the piece. It is important to note that no information about the symmetrical structure was input to the system.

6.2.4 Discussion

In this section we applied the Categorisation Analysis method to the analysis of Boulez' *Parenthèse* from the Third Piano Sonata, taking into account the

Class	Experiment 4
Class I	1c, 5
Class II	2, 4bcd
Class III	1ab, 6
Class IV	3, 4a

Table 6.5: The experimental results in experiment 4, with 4 classes.

obligatory fragments of the piece. The resulting hierarchic classification defines the similarity and difference relations between classes and between segments. We demonstrated how a classification analysis is appropriate for this piece and how it brings out the symmetrical structure that the composer intended. This method of analysis, shown previously to work on more traditional kinds of music, is shown here to be appropriate for an atonal and non-monophonic piece of music.

The results give many interesting insights on the obligatory fragments. In terms of internal relations, it is a very rich piece, each note situated in its position for a variety of reasons, forming part of an overall plan. More specifically, we see that the piece also has an interesting tonal structure, evolving mainly around G sharp at the beginning and end, and around D in the middle of the piece. The pitch class sets used are very similar to each other, segments 2 and 4 sharing sets, and the same for segments 1, 3 and 6. Dynamics and tempo seem to be very important for the segmentation and difference between subsegments, whereas contour information seems to be reflecting the symmetrical structure of the piece.

The issue of hierarchic segmentation in a classification task poses interesting challenges to the analyst. When classifying all the levels at the same time, on the one hand we get interesting similarities across levels, but on the other hand we get similarities between segments and their subsegments which are redundant. A sorting mechanism is needed for this problem.

6.3 *Syrinx* by Debussy

6.3.1 Nattiez' Analysis

Not much information exists in the historical literature about this piece by Debussy. However, it has been very popular with analysts because of its short, monophonic and, as its name suggests, flowing character. Nattiez produced three different outstanding analyses of Syrinx Nattiez (1975). These have been such good examples of Paradigmatic Analysis practice that other writers have commented on them: Monelle (1992), Bent (1987), Cook (1987). Other well known writings on paradigmatic analyses are Dunsby (1983), Nattiez (1982b), Dunsby and Whittall (1988).

In this section we study the second analysis of Nattiez. We produce various analyses using our model in order to investigate Nattiez's analysis and also show how the re-evaluation mechanism on the initial criteria works in order to refine an analysis.

Nattiez's segmentation

We take as given the segmentation from the second analysis that Nattiez produced on Syrinx, in Nattiez (1975), page 334-337, shown here in figure 6.3. The segmentation can be seen on his paradigmatic chart, where it is mostly clear where segment boundaries are. Segments are also marked by a number, which is given progressively, as each new segment appears; however, in case of *identity* with a previously encountered segment, Nattiez uses an old number, the number of the first occurrence of this particular segment. There are also 10 cases where his segmentation is overlapping.

We use the above segmentation, but rename all the segments, giving a new number to each segment, so two identical segments get different numbers since they do not occur at the same time. This is shown in figure 6.3. We

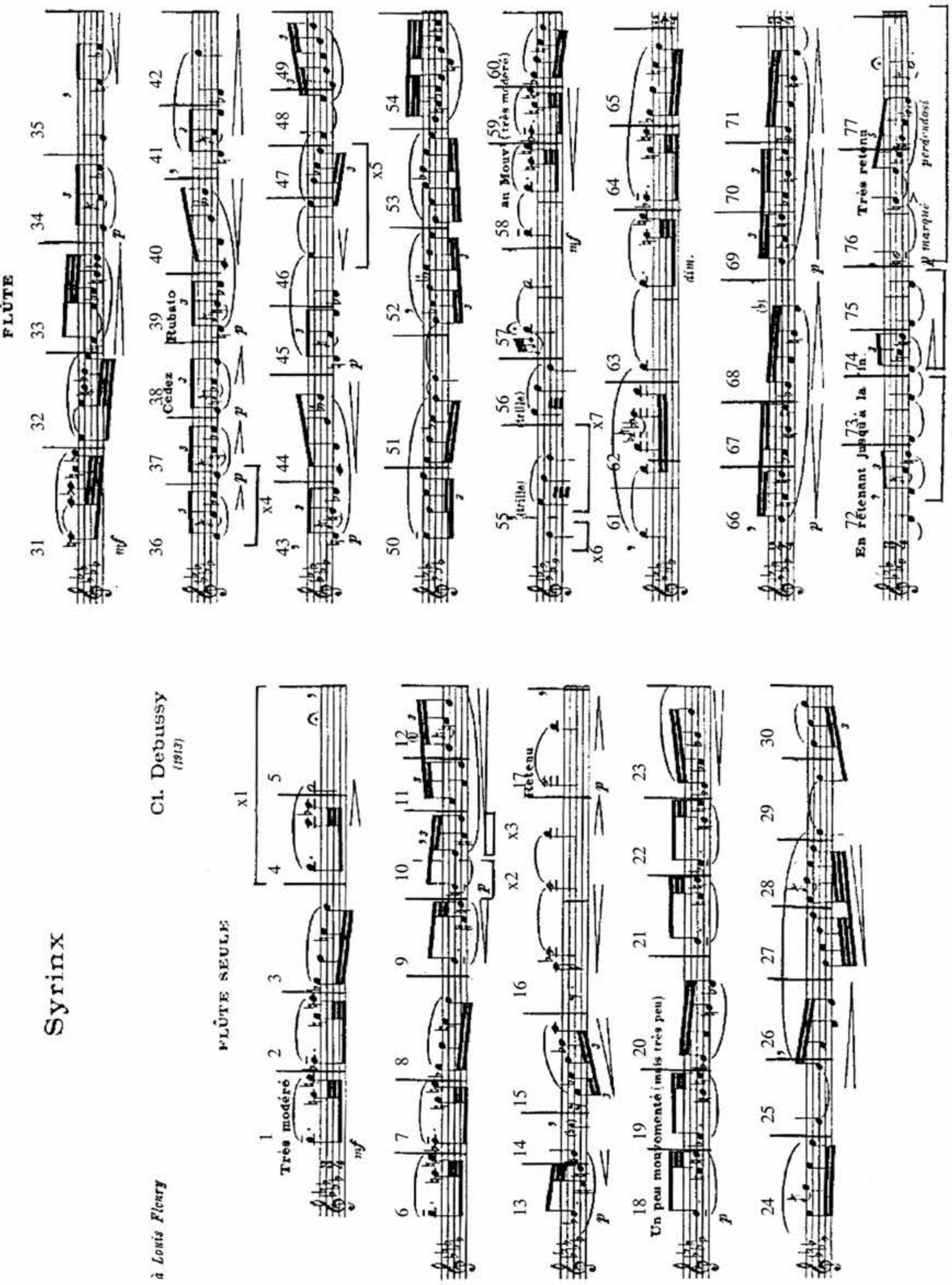

Figure 6.3: *Syrinx* by Debussy, with the second segmentation by Nattiez.

also include the overlapping segmentation cases. There is a total of 87 segments, out of which 77 are sequential, covering the whole piece, and 10 are overlapping. The 77 segments get numbers from 1-77, while the overlapping segments get names x1,x2,...,x10.

6.3.2 The First Experiment: three attributes

The first observations one can make about this piece are that: the contour seems to be important, since the piece always seems to flow upwards and downwards; that the first rhythmic pattern of the first segment seems to be quite catchy and characteristic, since it is immediately repeated and keeps occurring throughout the piece; finally, that there seem to be some long notes every now and then that break the flow.

Starting with these simple observations, we can formalise them and test them on the categorisation of the piece. Looking at the categorisation that will be produced, we can make alterations and refinements on the properties we have chosen. This procedure can go on until we are satisfied with the final analysis.

Formalising the observations

The general mechanism of formalisation of musical observation is explained in more detail in chapters 3 and 5. Here we show how each one of the three observations is turned into a multi-valued attribute or feature, and consequently a part of a feature-vector that will be an input to the system. As with the previous piece, for each of the attributes, it is stated whether they are true for the whole or for part of the segment.

Melodic contour: Based on figure 5.8 from chapter 5, melodic contour could be subdivided and formalised as shown in figure 6.4. Each of the seven features can take one position in the feature-vector, and take the value 1 is true and 0 otherwise. These features have to be true for the whole of the segment. The only exception is when a note is repeated (and therefore there is a bit of a "steady" contour in the middle of movement). Nattiez has chosen the segments in such a way that there are only these 7 instances: either one direction of the melody (up or down) or two (up-down, down-up). We use the term *steady* for when the melody does not move either upwards or downwards.

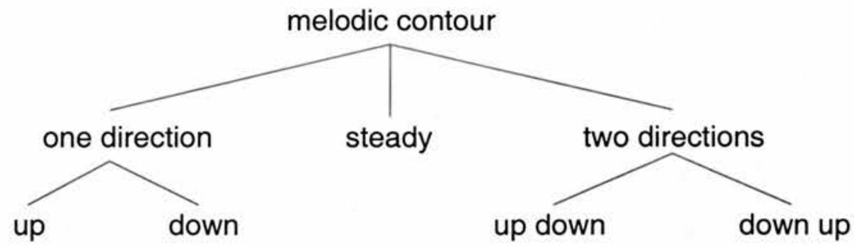

Figure 6.4: *Syrinx* by Debussy, with the second segmentation by Nattiez.

Long note: This is a rhythmic pattern. We decide to call a long note anything longer than a dotted crotchet. Segments are not restricted to having only a long note, they can just include one.

Catchy pattern of first segment This rhythmic pattern, *dotted-quaver, 32nd, 32nd,* does not have to be alone in a segment either. The segment can contain more notes.

The 9 features above, when represented each by a position in the vector, will take value 1 if they exist and value 0 if they don't exist in the specific segment. The 7 first positions of the vector will be on the contour, the 8th position on the long note, and the 9th on the dotted rhythmic pattern of the first segment. For example, the vectors for the first 3 segments are:

Segment 1	1 0 0 0 0 0 1 0 1
Segment 2	1 0 0 0 0 0 1 0 1
Segment 3	0 0 0 1 0 1 0 0 0

The experiment

We ran the algorithm, a new unit inserted every 200 iterations, and the final number of classes we got was 13 classes. It is not possible to get any further classification because the input features are not adequate to differentiate between segments that belonged to the resulting classes.

Results

We obtained all possible classifications, from 2 to 13 classes. Here we show the resulting classification with 13 classes, and then the classification with 8 classes. We discuss the results and then point out the differences.

Figure 6.5 shows the classification on the actual score, when having 13 classes. Each class is represented by a different colour. For example, segments 1,2,6, and all the others that are marked with red all belong to the same class. The choice of colours for each class is random – two colours that are close in the spectrum do not denote two classes that are more similar.

Table 6.6 shows the same results as figure 6.5, that is the classification with 13 classes, in a table form. This makes it easier to compare it to the classification of 8 classes, also in the same table.

Interpretation and Discussion

Even with these very simple first observations, the results of the 13-class classification seem to be acceptable. We notice that distinctions are based mainly on contour, but rhythm seems to be playing a role too. Objects within a class seem to be united by contour information; however, the same contour can be

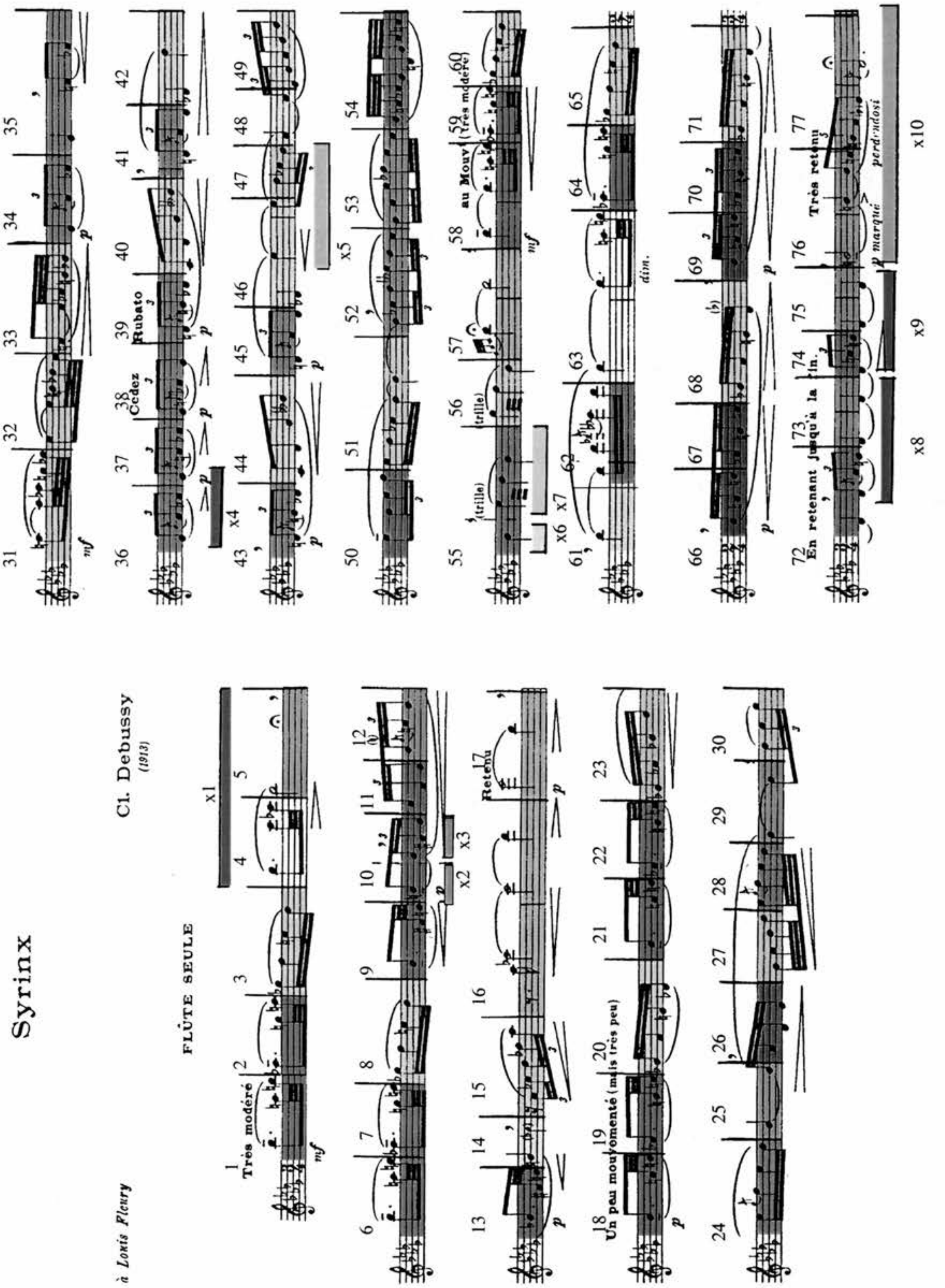

Figure 6.5: The classification results of Experiment 1, with 9 features and 13 classes.

Class	First Experiment, 9 features, 13 classes
Class I	17, 76, 77, x5, x10
Class II	x1, x8, x9
Class III	63
Class IV	5, 25, 29, 73, 75
Class V	15, 23, 27, 30, 40, 42, 44, 46, 49
Class VI	10, 11, 26, 50, 54, 66, 67, 69, 70, x3
Class VII	x2, x6
Class VIII	3, 8, 20, 31, 32, 33, 47, 51, 56, 60, 65, 68, 71, x7
Class IX	4
Class X	12, 24, 28, 34, 35, 36, 37, 38, 39, 41, 43, 45, 52, 53, 55, 62, 72, 74, x4
Class XI	14, 48, 61
Class XII	16, 57
Class XIII	1, 2, 6, 7, 9, 13, 18, 19, 21, 22, 58, 59, 64
	First Experiment, 9 features, 8 classes
Class I	17, 76, 77, x5, x10
Class II	**4**, 12, 24, 28, 34, 35, 36, 37, 38, 39, 41, 43, 45, 52, 53, 55, 62, 72, 74, x1, x4, x8, x9
Class III	1, 2, 6, 7, 9, 13, 18, 19, 21, 22, 58, 59, **63**, 64
Class IV	5, 14, 25, 29, 48, 61, 73, 75
Class V	15, 16, 23, 27, 30, 40, 42, 44, 46, 49, 57
Class VI	10, 11, 26, 50, 54, 66, 67, 69, 70, x3
Class VII	x2, x6
Class VIII	3, 8, 20, 31, 32, 33, 47, 51, 56, 60, 65, 68, 71, x7

Table 6.6: Results of *Syrinx*, First experiment, with 9 features, in 13 (first part) and 8 classes (second part). Notice the differences of the 8-class to the 13-class classification: in Class II: segments x1,x8 and x9 of the 8-class results form a class of their own in the 13-class results. Also, the same for Class IV: segments 5, 25, 29, 73 and 75. In Class V: segments 16 and 57. The other classes remain the same, apart from segments 4 and 63 (in bold) that form classes of their own.

shared by other classes. For example, segments 1, 2 and 9 share the contour down-up and belong to the same (red) class. Segment 10, on the other side, although shares the same contour and one might think of it as belonging the the same class, belongs to a different class because it contains a long note (as long note was defined). Furthermore, segments 66 and 67, although perhaps intuitively they could be in the same (red) class, they are not. This is because their rhythm is different, although the pitches are the same as in 18 and 19, and the exact pitch classes and intervals as segments 1 and 2. This does not affect the classification; it was not described in the input vectors, therefore the system does not "know" of this similarity.

In the same way, segments 31, 32 and 33, that in Nattiez' analysis form a separate class, here are just classified according to contour, and therefore belong to the same class as segment 3 (yellow class). This is because in our original representation the difference between these three segments and the rest of the class was not formalised, and the system therefore did not "know" about the difference.

In this classification, there are two classes that potentially carry the most interest, each with a single segment: segment 4 and segment 63. Segment 4 shares the rhythmic features with the class of segments 1 and 2 (red class) and the contour features of the the class of segment 12 (turquoise). Similarly, in the case of segment 63, the segment contains the contour and dotted pattern of the segment 1, but also contains a long note, and this differentiates it from the class where segment 1 belongs (red class), as well as from the class that contains only one long note, such as segment 5 (violet).

The 13-class classification is the final, that is the lowest possible level that our algorithm gives us, and therefore gives all the possible distinctions between the segments, as realised depending on our initial knowledge representation. In order to see how the algorithm sees these solitary segments, leaning towards which class, as well as how classes are closer and further apart to each other, one has to see the classifications the algorithm produced when having fewer number of classes.

When having only 8 classes, we see that classes that were distinct in the case of the 13-class result, are now put together. This is the relation of classes in hierarchic classification.

- The class that contains segments 1 and 2 (red class) now contains segment 63, which in the previous case formed a class of its own. This is because, in terms of our feature representation, there is only one difference between segment 63 and 64, or 1, which in the case of eight classes is not sufficient to allow segment 63 to be in a separate class.
- Segment 4 is now classified according to contour, together with the class that contains segment 12 (turquoise class). This can be explained: in our initial representation, contour has been more important than the dotted rhythmic pattern, since it occupies five positions in the feature vector, whereas the dotted rhythmic pattern occupies only one. Nattiez also classified this segment according to contour. From this we can assume that if we wanted to reproduce Nattiez's analysis, contour has to be more important in our representation than the specific rhythmic pattern. This is discussed further below.
- The class of segment 5 (violet) and the class of segment 14 (light grey) are merged here into the same class. This at first seems to make sense when having a fewer number of classes, since the segments could intuitively be classified together; however, this happens here because of a not precise enough representation: the violet class of segment 5 is characterised of two features: steady melodic line and long note. The light grey class of segment 14 is characterised by a steady melodic line and nothing else.

The two vectors are thus:

Segment 5 0 0 0 0 1 1 0
Segment 14 0 0 0 0 1 0 0

One can see the similarity of the vectors, although they only share the presence of one feature position.

- The class of segment 15 (pink) and the class of segment 16 (light blue-purple) are also merged, since the difference between them is only the addition of the long note.

6.3.3 Second Experiment: Adding interval and register information

The purpose of this experiment is to enhance the initial representation with further information on intervals and register, observe some of the differences this makes on the results, and try to see why these happen. Interval information is encoded as a feature: 1 if the interval exists in the segment, 0 otherwise. Register is made out of three features that are different instances (see below). Contrary to the interval features, all the notes of a segment have to have the register features in order for these to be true. 6 more positions in the vectors of Experiment 1 were taken by intervals, and 3 by register:

- the existence of a semitone in the segment,
- the existence of a leap in the segment,
- the existence of a 3-semitone in the segment,
- the existence of a fourth (including augmented and diminished),
- the existence of a fifth (including augmented and diminished),

- the existence of an octave or bigger interval,
- flute's lower register,
- flute's higher register,
- crossing both registers.

The results are shown in table 6.3.3 for 8 classes.

Class	Second Experiment
Class I	34, 36, 37, 38, 39, 41, 43, 45, 72, 74, x4
Class II	3, 8, 20, 31, 32, 33, 47, 51, 60, 65, 68, 71, x5
Class III	5, 14, 25, 29, 48, 61, 73, 75, 76, 77, x2, x6, x10
Class IV	12, 24, 28, 35, 52, 53, 55, 62, x1, x8, x9
Class V	1, 2, **4**, 6, 7, 9, 10, 13, 18, 19, 21, 22, 50, 58, 59, **63**, 64
Class VI	15, 23, 27, 30, 40, 44, 49
Class VII	11, 26, 54, 66, 67, 69, 70, x3
Class VIII	16, 17, 42, 46, 56, 57, x7

Table 6.7: Results in *Syrinx*, second experiment, with 20 features, in 8 classes.

Discussion

Table 6.7 shows that the classification we get when adding interval and register information is not significantly different to the one on the first experiment. This means that the extra information might not fundamentally crucial. However, the results also show that the 3-semitone interval is quite significant in the piece, and the semitone interval was not present in several of the segments, therefore not as redundant as one might have thought. There are some changes in the second experiment:

- Segment 4 is now classified in the same class as segment 1 (red class). That means that rhythmic shape is more important than contour in this classification. However, this is not the only common feature that segments 1 and 4 share: there is a higher-level similarity of contour, that of the feature *two directions*. The new similarities of register and of intervals (both segments include a semitone) reinforce the similarity between the segments and 1 and 4. By having more positions on the vector, the contour difference between the segments becomes less significant.

- Segment 63 now belongs to the same class as segment 1 (red), whereas in the first experiment it did not. This is because the long note at the beginning of the segment has become less important by the use of extra features, which are actually enhancing the similarity between this segment and the rest of its class.

Nattiez classified segment 4 according to melodic shape and not to rhythm. The problem with this is that since segment 4 is the first occurrence for such a melodic shape, it becomes a paradigm. However, further segments are more and more varied, and the result is that segment 4 is in no way prototypical of the whole category. In our analysis, this problem was avoided because prototypes represented the weighted average of all class members and not the first occurrence of a class member.

6.4 Conclusions

In this chapter we have demonstrated how the model presented in the previous chapter works in practice. We applied the method to two pieces of different character and different analytical requirements: *Parenthèse* by Boulez, an atonal piano piece, and *Syrinx* by Debussy, a short solo flute piece with interesting melodic and rhythmic features. The method is particularly suitable to find appropriate feature selections and representations by looking at the resulting classifications.

However, there are various significant relations between the segments that this method of analysis, as it is at the moment, cannot discover. This is the case especially across sequential segments and their deep structure. For example, although Syrinx is a particularly good piece for this method of analysis, there are still very significant relations of the segments that are left out:

Long segment made of:	*threaded pitch relations:*
Segments 1-2-3	B flat (B natural) - A flat A (natural) - G flat
Segments 9-10-11-12	B flat - B natural - C - D flat
Segments 31-32-33	A natural - E flat - A natural

These kind of relations need a further syntagmatic analysis in order to become apparent. The start for this is the abstract segment sequence mentioned in chapter 3. Sequences of segments and of specific properties are needed in order to bring out such relations. Conklin and Anagnostopoulou (2001) solve this by finding the longest significant patterns within various multiple viewpoints (the same as the definition of properties in this thesis). In this approach, initial segmentation is abandoned in favour of pattern discovery. Although the rationales of the two approaches are different, they can be combined to solve some of the issues that this method creates. This is discussed again in the final chapter.

With the analysis of Syrinx we demonstrated that the results depend on the initial representation, that is the choice of properties according on which each segment is described. A different choice of properties would yield different results. However, a bad resulting classification would show that the initial properties were not chosen carefully, and a re-evaluation of these properties is needed. In that way, the analyst can revise the initial properties. This procedure can theoretically go on until an acceptable classification is produced.

Chapter 7

Evaluation of the method and model

This chapter is divided into two parts: first the evaluation of the method presented in this thesis is discussed, and then its representation, algorithm and general framework are compared with other relevant research.

7.1 An evaluation of the Categorisation Analysis method

In this thesis we described a new method of music analysis, *Categorisation Analysis of Music*, which is based on Paradigmatic analysis. As in Paradigmatic Analysis, the analyst produces a classification of the segments of a piece of music. The method is independent of style, based on general principles which act as a framework, to ensure a consistent analysis, making sure not to restrict the analyst's freedom at the same time. In particular, Categorisation Analysis advances Paradigmatic Analysis in the following ways:

- The new method is set in a more formal framework in which the individual steps of the analysis are clearly delineated. Each step has its own defined results.
- Nicholas Cook comments on Paradigmatic Analysis (Cook, 1987, p.181):

> *the first difficulty with semiotic analysis – as with formal analysis in general – is that while it allows us to make precise statements about music, it doesn't always seem to be clear precisely what we mean by these statements.*

In our approach, the use of criteria, which are the musical properties that the analyst decides to base his/her analysis upon, are explicitly defined. The analyst is free to choose his/her own criteria, but has to follow certain formal requirements in expressing these criteria. In this way the analysis is intelligible and it becomes possible to evaluate it by other analysts.

- These formal requirements produce a novel way of representing the musical score, in such a way that it can then become the input to a categorisation algorithm.
- The new method is computationally modelled with an unsupervised neural network algorithm. This means that the categorisation procedure is formal and there are no inconsistencies, according to the initial criteria.
- If the end result of the analysis is not satisfactory, one can re-evaluate the choice of the initial criteria, i.e. the musical properties that the analyst chose for the score description.
- The above points ensure the more formal approach, without restricting the analyst's freedom of choice and creativity. The method acts as a framework for the analysis, making explicit previously intuitive decisions of the analyst, while the content rests with the analyst.
- As an extension to Paradigmatic Analysis, the resulting classification is hierarchical. The hierarchical classification points to relations and distances between classes.

- New findings from categorisation theory are brought in to enhance the methodology and to act as a bridge to cognitive modelling of categorisation in music analysis. For example, the concept of paradigm is dropped in favour of prototypes, and prototypes are defined as the set of probabilistic values of properties. There are no necessary and sufficient conditions for class membership.

- As with Paradigmatic Analysis, the method is able to deal with any kind of music, since it is based on general principles such as context dependent similarity. Here we have analysed 3 pieces of diverse character: a Scottish folk tune, *Syrinx* by Debussy and *Parenthèse* from Boulez' 3rd Piano Sonata and obtained interesting results for all three pieces (see chapters 4 and 6).

In our approach we encountered the following problematic issues:

- As in Paradigmatic Analysis, segmentation, similarity and categorisation are strongly interdependent. The segmentation relies on previously encountered material, whether identical or similar. The concept of similarity in this approach is expressed by description of properties of the segments. Categorisation depends on these properties, as well as on segmentation. In theoretical terms, it is virtually impossible to separate the three. However, in practical terms – which is what we are interested in here, as the purpose is music analysis – it becomes possible to separate them: the three steps, segmentation, description of segments as properties and categorisation, each have their own specific output result. The result of the segmentation is a segmented piece, no matter how this has been achieved. The result of the "description of segments as sets of properties" step is the transformation of each segment into a list of properties, and the result of categorisation is the division of segments into classes.

- It is possible to repeatedly revise the initial representation, that is the selection of properties, until a good categorisation is achieved. However, here we do not deal with segmentation, or possible re-evaluations of segmentation according to the resulting categorisation. What we can do, however, is if we have a good result, to assume that the segmentation has been a good one too. A bad segmentation would result in a bad categorisation. However, segmentation is an ongoing problem for music analysis, since all the subsequent steps depend on it. In section 8.2 we discuss an alternative approach to segmentation.

In our approach we explored the limits of formalism in music analysis while respecting the analyst's freedom. There can be many arguments as to why a formal approach to music analysis might be desirable. However, there is a tradeoff between formalisation and freedom in music analysis. For a fully formalised, or automated method, the style of the music has to be not only restricted, but also chosen in such a way that it is strict itself, uniform and easily reproducible. This is why many computational music analyses researchers choose to work with Bach chorales or other early music uniform repertories, or folk music of a specific place.

Eric Clarke agrees by saying that *"each piece can make use of a substantial number of principles that are specific only to that work, and which are consequently inexplicable (or at least explicable only at a very general level) in terms of a broad and general structural theory."*

7.2 Comparison with other representations, algorithms and systems

In this section we compare various parts of the system to others. Subsection 7.2.1 presents an investigation on how various representations and algorithms influence the classification results of an analysis. Subsection 7.2.2 compares the CAM system to two other existing systems.

7.2.1 Investigating the Influence of Representations and Algorithms in a Categorisation Analysis

This subsection describes work that has been carried out in collaboration with Dominik Hörnel and Karin Höthker, University of Karlsruhe (Anagnostopoulou *et al.*, 1999; Höthker *et al.*, 2000) – see appendix.

There have been various formal and computational models for music classification, using different architectures, representations and algorithms. The authors investigate the impact of varying the following two aspects:

- the knowledge representation of the musical segments, specifically the choice of musical features that describe the segments and the way of representing these features,
- the algorithms used for the classification of these segments.

A distance function allows the comparison of the various classification results in an objective way. The pieces we chose to analyse were nine two-part Bach inventions.

Methodology

For the above study the authors use a computational model which is an enriched version of the CAM model described above (Figure 7.1). As with CAM, the architecture is modular and therefore allows the experimentation by substituting different modules with other equivalent ones, in this case knowledge representations and algorithms.

The segmentation is not varied throughout the experiments. For the nine Bach inventions, a steady and fixed-length segmentation is used, by inserting a break on every crotchet beat. The segments are transformed into feature-vectors, and then the clustering algorithm performs the classification of these segments. Different classification results can then be compared to each other, for example by using human evaluation as a reference point and an objective distance function calculating the distance between the two. Although different approaches in the literature have used different model architectures, representation and classification components are common to all of these models and can be tested independently of the general framework.

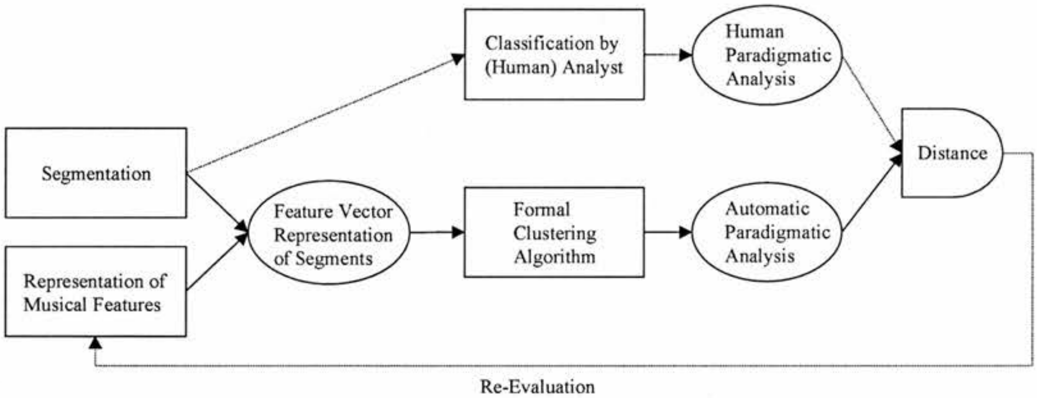

Figure 7.1: A formal model of classification

Knowledge Representations

In order to become an appropriate input for the classification module, each segment needs to be transformed into a feature-vector. The questions we address here are *which* musical features to represent and *how*. Figure 7.2 shows the 7 different representations we chose for our investigation, and how these are related to each other. In Figure 7.3, the representation of a single motive is shown using all seven kinds of representation. In more detail,

C1 describes whether the next note goes up (1), down (-1) or is stationary (0). We use the semi-quaver beat as a unit, that means a quaver can be thought

C1-I-R	**C2-I-R**	contour,intervals, rhythm
C1-I	**C2-I**	contour, intervals
C1	**C2**	contour
pitch		low level

Figure 7.2: The various representations used: C indicates melodic contour information, I interval information and R rhythm information. C1 and C2 are two principally different ways of representing melodic contour information.

C1-I-R	0 0.1 1 1 0.3 0 -1 0.2 0
C1-I	0 0.1 1 0.3 -1 0.2
C1	0 1 -1
pitch	1.4 0 1.4 1 1.6 0 1.5 0
C2-I-R	0 0 0 0 0.17 0 0 0 0 0 0 0 0 0 0.17 0.17 0.17 0 0 0 0 0 0 0 0 0.1 0.33 0.3 0 0.2 0
C2-I	0 0 0 0 0.17 0 0 0 0 0 0 0 0 0 0.17 0.17 0.17 0 0 0 0 0 0 0 0 0.1 0.3 0.2
C2	0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0

Figure 7.3: Representation example of a segment.

of as two identical semi-quavers. This assumption does not affect the contour information.

C1-I To C1, interval information is added: 0.2 for an interval of a second, 0.4 for a fourth and so forth. 0.1 shows the interval of a "prime", that is the same note.

C1-I-R To the above, rhythmic information is added by notating whether there is a new note on each semi-quaver position (0) or not (1).

C2 is the kind of conceptual hierarchy described in chapter 5. Each node of the tree is represented by a separate position in the feature-vector (1 if the feature exists, 0 if not). The tree is a structure that preserves inheritance, so the entire path that leads to the terminal node that is 1 is also 1. In this case, the network is:

- straight melodic motion
  - up
  - down
- oscillating
  - up down
  - down up
  - etc
- stationary

C2-I The same interval information as above is added.

C2-I-R To the above the same rhythm information is added.

Pitch, the last representation, is a very crude low level representation that contains only information on absolute pitch and rhythm.

The Algorithms

The algorithms that the authors use for their experiments are: Kohonen's self-organising feature map (Kohonen, 1982), Ward (Ward, 1963), a Kohonen-Ward combination as defined in the above mentioned papers, Growing Neural Gas (Fritzke, 1995), and Star Center (Rolland, 1998b). A detailed description of these is found in the above relevant papers in the appendix.

The Distance Function

The various resulting analyses were evaluated by comparing them to each other and to an analysis by a musicologist. In order to get general and reproducible results, the authors use a distance function reflecting the degree of similarity between two analyses, which was originally devised by K. Höthker.

Since Categorisation Analysis analysis abstracts from motive instances by assigning them to motive classes, the distance function is computed based on a sequence of class labels, the *abstract motive sequence* (e.g. *ababccdc*). The distance function is independent of the actual class labels. This is a reasonable property for a distance function on abstract motive sequences since the outcome of Paradigmatic Analysis is not affected by relabelling classes.

Experiments

For each of the 9 Bach pieces, seven vector-file versions were created, using each one of the different knowledge representations. These 9 pieces (numbers 1, 2, 5, 7, 8, 9, 11, 13, 15) were split into a training set (1, 7, 8, 11, 15) and a test set (2, 5, 9, 13). The algorithms were trained on the training set by using the human analysis of each piece for fitting the parameters (e.g. train the Kohonen networks or determine the threshold value for the Star Center algorithm). The trained algorithms were then applied to the test set. Results were obtained for all combinations of algorithm and representation. The results were translated into abstract motive sequences and compared to each other and to the human analyses, using the distance function described above.

Results

The abstract motive distances were computed on the training and test set using the representations and algorithms presented above. The authors report the results on the test set, but got very similar ones on the training set thereby confirming the homogeneity of the Bach two-part inventions.

The results show that the omission of interval and rhythm information does not notably change the distance. The contour information clearly emerges as the most prominent feature for classifying musical structure in our specific experimental setting. The more sophisticated C2 representation proved to be general enough for the representation of all the inventions of the test and training set.

One can see that the algorithms have influence on the outcome as well, although to a much lesser extent: for our task, Kohonen-Ward and GNG are better than the other algorithms. The Kohonen network which transforms the data into a two-dimensional space presents a suitable pre-processing method for the Ward algorithm.

A detailed report of the results can be found in the appendix.

Discussion

Both ways of representing contour, C1 and C2, have advantages and disadvantages. C1 is the most efficient and intuitive. However, unwanted a priori similarity judgements are not avoided: imagine a segment with stationary melodic contour, C-C-C, one ascending- descending, C-E-C, and one descending- ascending, C-A-C. This kind of representation will assume that the second segment is closer to the first than it is to the third, which might not be intuitively obvious in music analysis. C2 overcomes this problem by introducing the upper node "oscillating" in the tree, which is common to the second and third segment, therefore increasing their similarity. However, this procedure, apart from being computationally very expensive due to its dimensionality increase, has the further problem that the similarity is encoded a priori. One could even argue that segments are already pre-classified and one does not need the classification algorithm anymore. There is a trade-off.

The real-valued vector representation which created the problem above with C1, has the same effect in interval information, making similarity assumptions about the various intervals depending on their distance. For example, an interval of a second is more similar to an interval of a third than to an interval of a fourth. This is something that might not be intuitively sensible.

It is interesting to note that rhythm is implicitly encoded in both C1 and I information, by mentioning the position of each note on the semi-quaver beat sequence of the segment.

7.2.2 Comparison with other systems

Two other pieces of work are not mentioned in the above investigation: the UNSCRAMBLE algorithm (Cambouropoulos and Smaill, 1997; Cambouropoulos, 1998) and the Multiple Viewpoints Representation (Conklin and Witten, 1995). In chapter 2 we gave a brief description of these works. Here we compare them with our approach.

The Multiple Viewpoints Approach

The music representation formalism of multiple viewpoints (Conklin and Witten, 1995) can be used to describe music at varying levels of abstraction. Viewpoints are functions that operate on the basic score representation, producing an attribute of an event or of a set of events. A score is described as a sequence of elements of a viewpoint, for example, a sequence of contours or of melodic intervals. Complex viewpoints are constructed by linking or threading pairs of viewpoints.

It is possible, using this formalism, to simulate the conceptual hierarchy description of our approach. For example, the C2 contour representation mentioned above (which is the same as the conceptual hierarchy representation we describe in chapter 5) can be simulated using a viewpoint that threads melodic contour on every event that changes direction in contour. By this simple construction, a segment can be viewed as a sequence of terminal node C2 identifiers. In order to include the non-terminal nodes of the hierarchy, extra viewpoints would have to be deduced from the existing ones.

Viewpoints do not allow for similarity of viewpoint elements. Rather, similarity is achieved by equality of more abstract viewpoints. For example, in our representation the contour up-down is somewhat similar to the contour down-up (both oscillating), but if we used a viewpoint, they would not be similar. However, one could define an oscillating viewpoint (a higher level of abstraction) for which the up-down and down-up would have identical viewpoint elements.

Another significant difference between the two approaches is that in our representation the properties refer to whole segments, whereas viewpoints are applied to individual notes rather than patterns. However, viewpoints are powerful enough in that they can deduce pattern properties (or viewpoints). This is described in Conklin and Anagnostopoulou (2001), see also appendix.

The UNSCRAMBLE categorisation algorithm

UNSCRAMBLE (Cambouropoulos, 1998; Cambouropoulos and Smaill, 1997) is a symbolic, unsupervised categorisation algorithm which can be usually applied to a single piece of music. Like in CAM, the system needs two data sets as inputs: the segmented piece of music and the properties according to which the segments are described. These properties acquire weights that change during the process of the classification, promoting those properties that differentiate the classes between each other the most. The final classification is is non hierarchical and the number of classes produced is fixed. The algorithm finds the optimum number of classes. This is different to our approach in that the system produces all possible numbers of classes, where hierarchic relations can be observed and and the final number of classes rests with the analyst.

7.3 Conclusions

This chapter presented the computational model for the Categorisation Analysis. The system is not fully automatic, requiring the analyst to play an active role in the analysis by making various choices that the system takes into consideration. The system has a modular architecture, enabling the user to substitute modules for other equivalent ones. Thus it is possible to change segmentation, choice of property descriptions, representation into feature-vectors and clustering algorithm.

In this chapter we looked at the initial motivation for the Categorisation Analysis method and how this was met by the work described in this thesis. We then compared our method to other approaches in computational music analysis that make use of the categorisation process. We described the work by Anagnostopoulou *et al.* (1999); Höthker *et al.* (2000), which investigates the influence of representations and algorithms in music classification tasks. The next chapter concludes the thesis with some general discussion and observations in categorisation theory and in formal music analysis, and describes potential interesting paths for further research.

Chapter 8

Conclusive Remarks and Further Work

> [Jorge Luis Borges mentions] *a "certain Chinese encyclopaedia" in which it is written that "animals are divided into: (a) belonging to the Emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l)* et cetera, *(m) having just broken the water pitcher, (n) that from a long way off look like flies".*
> Foucault (1992, p.XV)

Foucault, with this passage, wants to show that one's categories depend on one's point of view, and that other points of view – especially those of other places and other ages – are literally unthinkable. In the animal kingdom which Borges mentions, well-known and rigid taxonomies have existed at least since Aristotle (350BC), who divided animals into genera, species within the genera and various subcategories within these. This classification was close to the system we use today, first developed by the Swedish naturalist Carolus Linnaeus (1707-1778). Linnaeus separated animals and plants according to certain physical similarities and gave identifying names to each kind.

However, such classifications, based on similar and common properties, are not the only kind possible. Foucault shows that such positivistic and rational methods of classification were not always current, and that other methods may yield surprising results. No matter how unthinkable other systems might appear to be, they can still enter into a fruitful dialogue with the scientific system. Of course, Borges' strange Chinese classification is as unexpected to us as anything can get, but to different contexts, such as ancient Chinese culture and mythology, this kind of classification might have been more meaningful.

One cannot argue that one approach is right and the other wrong. Categorisation, which has been the main thread in this work in music analysis, is a resilient concept, and there are many methods which produce very diverse results. In the first place, the goal may not so much be to get the "right system", but to state one's criteria and to see where the system takes you. The Chinese classification arouses our interest exactly because it states its criteria for classification, and we can therefore evaluate it and find it interesting, surprising, unthinkable, or satirical.

This is especially the case in music, where the objects to be classified, whether they are segments or whole pieces, are not so easily definable and their description and variety depends mostly on context. In this thesis we have shown that the detailed description of musical entities is vital for any musical categorisation task which is based, in Nattiez's terms, on the Neutral Level. This is precisely because of the character of music and the importance of immediate and general context (the piece and the composer's style). A categorisation without its criteria stated would be meaningless for the purposes of music analysis. A categorisation with explicit criteria is coherent and therefore acceptable to that extent. The analyst can thus evaluate various categorisations and decide on their interestingness and usefulness. Others can also study his/her analysis and be able to understand and evaluate it.

Categorisation is a vital process in music analysis. Most formal methods of analysis use categorisation at some point in their methodologies. Paradigmatic Analysis is the one which is based solely on categorisation. In this thesis, two approaches to categorisation were discussed: the first comes from music analysis, and in particular from Nattiez's approach, and the other from psychological categorisation theories. We have critically merged a psychological view, namely that of categorisation based on prototypical probabilistic properties, with the musicological approach. Our objective has been to develop music analysis, and we have created the Categorisation Analysis of Music.

As in Paradigmatic Analysis, the analyst produces a classification of the segments of a piece of music. However, the new method is set in a more formal framework in which the individual steps of the analysis are clearly delineated and the criteria for the classification are explicitly defined. Furthermore, in extension to Paradigmatic Analysis, the resulting classification is hierarchical, and the new findings from categorisation theory act as a bridge to cognitive modelling of categorisation in music analysis. The new method is computationally modelled with an unsupervised neural network algorithm, thereby further formalising the classification process.

The rationale behind this new approach is to allow for a formal analysis without restricting the analyst's freedom of choice: the method acts as a framework for the analysis, making explicit previously intuitive decisions of the analyst, while the analyst remains free to choose his/her own analytical criteria.

We have analysed three pieces of different character: a Scottish Folk tune (chapter 4), *Parenthèse* from the Third Piano Sonata by Boulez, and *Syrinx* by Debussy (chapter 6). With these analyses we demonstrated how the method can yield interesting results in different kinds of musical repertoire.

8.1 On Music Analysis

Cook (1987) has several objections to Paradigmatic Analysis, which would also apply to our approach:

> *The problem is this: how much of what matters about music is retained in the translation from sound-experiences to abstract categories such as*

> *"ascending conjunct line"? Can we say anything important about the experience of a given line simply by classifying it as the opposite of lines which are descending or disjunct? Aren't we in danger of making precise statements about musical scores which have only the vaguest connection with the music we experience?* Cook (1987), p.181

Cook here seems to misinterpret Paradigmatic Analysis. This becomes clear in the last sentence, where he makes the connection between "statements on the musical score" and the "music we experience". There is the assumption here that music analysis aims to reflect the musical experience. This is, however, precisely what Nattiez tried to avoid with his method of analysis. He wanted a methodology that steered clear of any such experiences.

On the other hand, Cook is again not correct because "statements on the musical score" are *indeed* descriptions of the way we experience music. Since an analysis is an interpretation by a single analyst of a piece or pieces of music, the analysis describes his/her own experience and interpretation of the music.

This is what Nattiez tried to avoid in his analysis, and indeed Cook (1987, p.183) goes on to criticise Nattiez's methodology for this reason:

> *Let us just think what it would mean for an analysis to be genuinely scientific and objective. It would mean that you could get the right results simply by following given procedures correctly: intuitive judgements about the music (I feel that ...) would not be involved.*

However, Nattiez's related statement had been:

> *People decide to associate several units in a single paradigm because of semantic or psychological criteria that they do not express consciously. We do not seek to downgrade the role of intuition at the outset of the analysis.*
> Nattiez (1982a)

Cook, dismissing altogether Paradigmatic Analysis when applied to a single piece, suggests that this method should move towards comparative analysis. He claims that since there cannot be any objective and scientific criteria for the analysis, and since in comparative analysis there is at least the measurement against the other pieces, there is no reason of existence for paradigmatic analysis, unless it is used as a first step in comparative analysis.

Cook did not foresee that it is possible to accept intuitive judgements about music, and still have an objective and scientific analysis. The aim of this thesis has been to show how it is possible to have objective and scientific analyses of a single piece of music, without restricting the intuition of the analyst.

As has been discussed in earlier chapters, there is no one single, correct way of categorisation. Categorisation is a cognitive process, and there are various ways and criteria to achieve it. It is true that the human factor plays an important role in this kind of analysis, as Nattiez admits. What is more, it is unlikely that any two analyses are going to be the same. The grouping of objects into categories can be carried out in many different ways, and there are no right and wrong answers, only answers that are consistent with one's criteria or not. Categorisation should not be examined as a strict task, unless we have explicit criteria.

However, this does not mean that a categorisation cannot be formal. Instead of speaking about right and wrong analyses, perhaps it would make more sense to redefine these as consistent and non-consistent, or bad analyses. A satisfactory analysis would be any analysis that is consistent to one's criteria, whatever these might be, and a non-satisfactory one if it is inconsistent.

What Nattiez argues for above and what he has demonstrated on several occasions – for example, see Nattiez (1997)– is the acceptance of several analyses of the same piece as valid. He recommends that a true semiotic analysis be based on the superimposition of a number of separate interpretations rather than merely on one.

Nattiez has been very progressive in this statement, recognising the true nature of his analytical method and aiming for the freedom of analytical expression. No analysis should claim that it is the one and only way of looking at a piece or pieces of music. Complementary analyses of the same piece can say a lot more than a single interpretation.

In a larger number of superimpositions of similar analyses, one can observe statistically significant general trends that can form the base of the cognitive modelling of the music analysis task.

As mentioned in the introductory chapter of the thesis, psychology and cognitive science are rapidly gaining territory in music, albeit mainly in musical listening, offering explanations of the psychological processing during the analysis that would otherwise remain mere speculations.

8.2 Future Developments

Further computational advances

The model can be extended in various ways. Firstly, it is possible that the analyst wishes to attribute different importance to different features. Such a weighting of features should be incorporated into an extended model. In the present model, the weighting is carried out manually, by inserting a feature as many times as the desired weight.

After the analyst defines the List of Attributes, the description of each segment according to this List of Attributes could perhaps be computationally modelled. However, this would require the analyst to not only come up with the List of Attributes in the first place, but also define functions for each attribute. Certain functions, for the most common attributes, can be predefined, but the analyst would still need to define some new ones, depending on the piece under analysis.

In order to be usable, the model requires an interface, so that it is easier and friendlier to use by musical analysts. In this way the music analyst will not be required to know the inner workings of the system, but instead rely on a graphical interface for choosing options and defining new ones.

Towards a musical workbench for music analysis

The work presented here can form part of a large scale analytical tool, a Musical Workbench, that will offer many methods of analysis, machine learning and generation, an idea discussed in detail in Anagnostopoulou *et al.* (2000). The various methods and techniques would be gathered into this workbench, which will be able to analyse, learn and generate melodies and thereby test melodic features for their relevance in the context of a given style. The resulting system would be a tool constructed for musicologists – music analysts, ethnomusicologists, music students – and for cognitive scientists who investigate cognitive modelling of music analysis. MELOLAB, the name given to the future system, would also provide a platform for computer scientists which allows them to test existing and develop new algorithms on challenging tasks such as categorisation, multi-scale learning and the learning of hierarchical relationships. This will create an experimental basis for developing a generative theory of musical style and thereby open a horizon towards the integration of musicological, mathematical and cognitive viewpoints and explanations of music.

Cognitive Modelling of the analytical process

The psychological validity of the various analytical methods and results has become an important issue in the human sciences of today. Therefore, any new method of analysis would benefit from experimental evidence that demonstrates its cognitive plausibility.

Much work has already been carried out in studying the cognitive processes involved in listening, performing and improvising music. However, less has been done in analysing music, apart from very low level experiments on series of notes, rather than whole pieces of music. The time is ripe for cognitive modelling of tasks involved directly in music analysis of whole pieces. Analysing music is a significant musical task, just as much as listening, performing, learning or improvising, which reveals the general musical understanding of a piece or pieces of music.

The Categorisation Analysis of Music has been built in such a way that it permits a potential extension on the cognitive modelling of the analyst. The procedure is clearly defined, with no magical steps, as is ensured by its formal character. Moreover, it is separated into modules which can be modelled separately. Finally, by having a non-fixed number of classes, it is possible to model any number of classes that are produced by the human subjects, and the process of re-evaluation can propose plausible initial choices of musical properties for the categorisation. The method could be used as a starting point to a cognitive model of musical categorisation.

Segmentation and Pattern Matching

Segmentation is a problematic issue in any kind of analysis. The results of the analysis will depend crucially on the segmentation chosen. While the present version relies on a given segmentation of a piece, in principle a revision of this initial segmentation could be incorporated into the classification process.

However, there is a different way of approaching the segmentation issue which, for this kind of analysis, would be equally (if not more) valuable. Instead of dividing the piece of into segments, one can look for pattern discovery and pattern matching first. At a subsequent step, the rest of the piece, in between patterns, could also form segments. Conklin and Anagnostopoulou (2001, also see appendix) present a new way of pattern representation and discovery. One of their future goals is to apply the method to categorisation analysis, thereby overcoming the segmentation issue altogether.

On the relation between music and language

Linguistic discourse analysis and music analysis have developed in parallel and have studied similar phenomena in their own domains. However, no systematic comparison of both techniques and their potential results exists to date. In Anagnostopoulou (1997, see also appendix), linguistic discourse and music are compared with respect to *cohesion*. This is a new perspective of analogy between language and music in three respects: Firstly, the focus is on discourse rather than sentence level. This is appropriate as music and linguistic discourse are both instances of human communication: they carry some kind of "meaning", they are intelligible, and above all they are both inherently temporal. Secondly, a semantic rather than a syntactic relation is examined, namely cohesion (Halliday and Hasan, 1976), which makes it possible to address a semantic level in music. And finally, the investigation focuses on the surface level by looking at associative features, without making any claims concerning underlying structure. From this comparison, the term *Musical Cohesion* is established. However, as with any parallels between language and music, and although there are self-evident similarities, one should not neglect their differences.

This leads us towards new thinking in the area of music semantics. If the mechanism of cohesion is the same in musical and in linguistic discourse, and if the linguistic cohesion is almost solely based on the semantics of concepts, this hints towards the existence of semantics in music that are similar to the semantics in language. Since this goes beyond the scope of the present discussion, the reader is referred to Anagnostopoulou (1997) and Anagnostopoulou and Ramscar (fc); for the notion of musical coherence from a slightly different perspective, this of syntax rather than semantics, see Baroni (1998).

Syntagmatic analysis

The principles of similarity and difference are principles common to the vast majority of musical repertoire. It can be argued that they are responsible to a large extent for cohesion and coherence within the musical piece.

The musical cohesion of a piece can be studied by looking at the syntagmatic axis of the analysis. The syntagmatic axis of an analysis consists of the relations between categories, as these are distributed over time. The 6th step of Categorisation Analysis (Chapter 3) deals with the syntagmatic axis of an analysis. The start for this is the abstract segment sequence mentioned in chapter 3.

However, it would be more interesting to investigate the arrangement of musical properties, rather than segmental categories, over time. Sequences of segments and of specific properties are needed in order to bring out such relations. Conklin and Anagnostopoulou (2001, also see appendix) solve this by finding the longest significant patterns within various multiple viewpoints (multiple viewpoints are the same as the definition of properties in this thesis). In this approach, initial segmentation is abandoned in favour of pattern discovery. One next step of this approach will be to create a syntagmatic analysis of the various viewpoints, or properties.

Musical Repertoire

Categorisation Analysis can in principle be applied to any kind of musical repertoire. Here we have applied it to a Scottish folk tune, a 20th century flute piece and an atonal piano piece. However, more applications need to be made in order to check the results in more demanding pieces of music, such as a larger-scale multi-vocal score, harmonic and polyphonic texture, as well as more contemporary music.

Polyphonic music poses an interesting challenge: in a strictly polyphonic score, such as for example a four-part Palestrina motet, the breaking points for the segmentation do not occur simultaneously in the four voices. A false conception exists, that segmentation has to cut "through" the score, so that the break points have to be at the same time in all the voices. However, it is not clear why this should be the case. It is interesting to find patterns in the different voices, and how these are repeated by other voices, since this is one of the characteristics of the music. Also, it would be interesting to see the vertical relations of the voices at any given moment in time, without this meaning necessarily that the segmentation has to be made totally "through" the piece.

Another interesting extension of this work would be to apply the Categorisation Analysis method to electronic and electroacoustic music. The principles of music that this method concentrates on, namely context similarity and difference within a piece, are also very apparent in electronic and electroacoustic music, where often some initial sound or sounds are amenable to various kinds of transformations, with varying degrees of similarity.

There can be so many exciting further paths to computational music analysis, and the area is still developing. Using the computer for the various music analysis purposes can be highly beneficial and inspiring in many ways. However, the relation of human analyst and computer should never be ignored and it is crucial that future computational analysis and composition systems do not underestimate this perspective. Else we run the risk of building very complicated systems, but with limited musical interest or use.

Bibliography

Anagnostopoulou, C. (1997). Cohesion in linguistic and musical discourse. In *Proceedings of the Third Triennial ESCOM Conference*, Uppsala, Sweden.

Anagnostopoulou, C. and Ramscar, M. (fc). What are concepts? a cognitive approach to "meaning" in music and language. Liège, Belgium. to be submitted to ESCOM 10th Aniversary Conference, topic "Musical Creativity".

Anagnostopoulou, C. and Smaill, A. (2000). Similarity and categorisation in Boulez' *parenthèse* from the third piano sonata. In *Proceedings of the 3rd ESCOM and 6th International Conference of the Music Perception and Cogntion*, University of Keele, UK.

Anagnostopoulou, C. and Westermann, G. (1997). Classification in music: A computational model for paradigmatic analysis. In *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece.

Anagnostopoulou, C., Hörnel, D., and Höthker, K. (1999). Investigating the influence of representations and algorithms in music classificaiton. In *Proceedings of the AISB'99 Symposium on Musical Creativity*, Sussex, UK. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Anagnostopoulou, C., Hörnel, D., and Höthker, K. (2000). MeloLab: A Musicological Workbench for learning-based investigation of Melodic Cohesion. Research Proposal submitted to the VW Foundation.

Aristotle (350 B.C.). The history of animals.

Baroni, M. (1998). Musical rules and musical signification: hermeneutic problems. *Contemporary Music Review*, **17**(2).

Baroni, M. and Jacoboni, C. (1978). *Proposal for a Grammar of Melody*. Les Presses de l'Université de Montréal, Montreal, Canada.

Bent, I. (1987). *Analysis*. The New Grove Handbooks in Music. Macmillan, London, UK. with W. Drabkin.

Blacking, J. (1981). The problem of "ethnic" perceptions in the semiotics of music. In W. Steiner, editor, *The Sign in Music and Literature*. University of Texas Press, Austin, Texas, USA.

Borges, J. (1964). *Funes the Memorious*. Penguin, London.

Borthwick, A. (1995). *Music Theory and Analysis - the limitations of logic*. Garland Publishing Inc., New York and London.

Boulez, P. (1971, 1975). *Boulez on Music Today*. Faber and Faber, London, UK.

Bruce, V. and Green, P. (1990). *Visual Perception: Physiology, Psychology and Ecology*. Lawrence Erlbaum Assocs Ltd., Hove, UK, 2nd edition.

Cambouropoulos, E. (1998). *Towards a General Computational Theory of Musical Structure*. Ph.D. thesis, University of Edinburgh.

Cambouropoulos, E. and Smaill, A. (1997). Similarity and categorisation inextricably bound together: the unscramble machine learning algorithm. In *Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation*, Edinburgh, Scotland. University of Edinburgh.

Carpender, G. and Grossberg, S. (1998). The ART of adaptive pattern recognition. *IEEE Computer*, **21**(3).

Chapin, S. (1982). Extracting an unfamiliar theme from its variations. *Psychomusicology*, **2**.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Collins, A. and Quillian, M. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, **8**, 240–247.

Conklin, D. and Anagnostopoulou, C. (2001). Representation and discovery of multiple viewpoints patterns. In *Proceedings of the International Computer Music Conference*, La Havana, Cuba.

Conklin, D. and Witten, I. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, **24**.

Conrad, C. (1972). Cognitive economy in semantic memory. *Journal of Experimental Psychology*, **92**.

Cook, N. (1987). *A Guide to Musical Analysis*. J.M. Dent and Sons Ltd, London,UK.

Deliège, I. (1992). Recognition of the Wagnerian leitmotiv: Experimental study based on an excerpt from Das Rheingold. *Musikpsychologie*, **9**.

Deliège, I. (1993). Mechanisms of cue extraction in memory for musical time. *Contemporary Music Review*, **9**(1 and 2).

Deliège, I. (2001). Prototype effects in music listening: an empirical approach to the notion of imprint. *Music Perception*, **18**(3).

Deliège, I. and Melen, M. (1997). Cue abstraction in the representation of musical form. In I. Deliège and J. Sloboda, editors, *Perception and Cognition of Music*. Hove: Psychology Press.

Dunsby, J. (1983). Music and semiotics: the Nattiez phase. *The Musical Quarterly*, **LXIX**.

Dunsby, J. and Whittall, A. (1988). *Music Analysis in theory and practice*. Faber Music, London and Boston.

Edworthy, J. (1985). Melodic contour and musical structure. In I. C. &. R. W. P. Howell, editor, *Musical Structure and Cognition*. Academic Press, London, UK.

Ellis, R. and Humphreys, G. (1999). *Connectionist Psychology*. Psychology Press.

Eysenck, M. and Keane, M. (1998). *Cognitive Psychology - A Student's Handbook*. Psychology Press, East Sussex, UK.

Forte, A. (1973). *The Structure of Atonal Music*. Yale University Press, New Haven.

Forte, A. and Gilbert, S. (1982). *Introduction to Schenkerian Analysis*. W.W. Norton, N.Y.,London.

Frege, G. (1952). On sense and reference. In P. Geach and M. Black, editors, *Translations from the philosophical writings of Gottlob Frege*. Blackwell, Oxford.

Fritzke, B. (1995). A growing neural gas algorithm learns topologies. In G. Tesauro, D. S. Touretzky, and T. K. Lean, editors, *Advances in Neural Information Processing Systems 7*, pages 625–632, Cambridge, MA. MIT-Press.

Gjerdingen, R. (1991). Using connectionist models to explore complex musical patterns. In *Music and Connectionism*, pages 138–149. MIT Press.

Goldstone, R., Medin, D., and Gentner, D. (1991). Relational similarity and the non-independence of features in similarity judgements. *Cognitive Psychology*, **23**.

Greimas, A. (1966). *Sémantique Structurale.* Larousse, Paris, France.

Guertin, M. (1981). Différence et similitude dans les préludes pour piano de Debussy. *Revue de Musique des Universités Canadiennes, Montreal*, **2**.

Guertin, M. (1990). *De la Lecture à l'Audition d'un texte musical.* Les Presses de l'Université de Montréal, Montréal, Québec, Canada.

Hahn, U. and Ramscar, M. (2001). *Similarity and Categorisation.* Oxford University Press, Oxford.

Halliday, M. and Hasan, R. (1976). *Cohesion in English.* English Language Series. Longman, London and N.Y.

Harris, Z. (1951). *Methods in Structural Linguistics.* University of Chicago Press, Chicago, USA.

Hörnel, D. (1998). A multi-scale neural-network model for learning and reproducing chorale variations. In W. Hewlett and E. Selfridge-Field, editors, *Melodic Similarity: Concepts, Procedures, and Applications*, Computing in Musicology 11. MIT Press.

Hörnel, D. and Höthker, K. (1998). A learn-based environment for melody completion. In *Proceedings of the XII Colloquium on Musical Informatics*, Gorizia, Italy.

Hörnel, D. and Höthker, K. (1999). Künstlerische Neuronale Netze lernen Musikstile. *Symposium des DVSM, Frankfurt/Main*, **13**.

Höthker, K., Hörnel, D., and Anagnostopoulou, C. (2000). Investigating the influence of representations and algorithms in music classification. *Computers and the Humanities*, **35**(1).

Katz, J. and Fodor, J. (1963). The structure of a semantic theory. *Language*, **39**.

Koch, H. (1802). *Musicalisches Lexicon.* Hildesheim Olms 1964, Frankfurt, facsimile edition.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59–69.

Lamont, A. and Dibben, N. (2001). Motivic relations and the perception of similarity. *Music Perception: Special issue on Similarity and Categorisation in Music, guest editor I. Deliège*, **18(3)**.

LaRue, J. (1992). *Guidelines for Style Analysis.* Harmonie Park Press, Michigan. USA.

Lenat, D. and Guha, R. (1990). *Building Large Knowledge Based Systems*. Addison-Wesley, Reading, MA, USA.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.

Lewin, D. (1987). *Generalized Musical Intervals and Transformations*. Yale University Press, USA.

Lincoln, H., editor (1970). *The Computer and Music*. Cornell University Press, NY and London.

Logan (1947). *Logan's Collection of Highland Bagpipe Music*. Paterson, London. Arrangements by Pipe-Major Ross.

Luger, G. and Stubblefield, W. (1998). *Artificial Intelligence - Structures and Strategies for Complex Problem Solving*. Addison Wesley Longman, Reading, MA, USA.

Marr, D. and Nishihara, K. (1978). Representation and recognition of the spatial organisation of three-dimensional shapes. *Philosophical Transactions of the Royal Society B London)*, **B200**, 269–294.

Marx, A. (1841). *Die Lehre von der musikalischen Komposition*. Breitkopf and Härtel, Leipzig.

McClelland, J. L., Rumelhart, D. E., and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychological and Biological Models*. MIT Press, Cambridge, MA.

Mechelen, I. V., Hampton, J., Michalski, R., and Theuns, P. (1993). *Categories and Concepts - Theoretical Views and Inductive Data Analysis*. Academic Press, San Diego, CA, USA.

Medin, D. and Shaffer, M. (1978). Context theory of classification learning. *Psychological Review*, **85**.

Medin, D. and Smith, E. (1984). Concepts and concept formation. *Annual Review of Psychology*, **35**, 113–138.

Medin, D., Goldstone, R., and Gentner, D. (1990). Similarity involving attributes and relations: Judgements of similarity and differences are not inverses. *Psychological Science*, **1**.

Monelle, R. (1991a). Music and semantics. In *Musical Signification - Proceedings of the 2nd International Congress on Musical Signification*, Helsinki, Finland.

Monelle, R. (1991b). Structural semantics and instrumental music. *Music Analysis*, **10**(1-2).

Monelle, R. (1992). *Linguistics and Semiotics in Music*. Number 5 in Contemporary Music Studies. Harwood Academic Publishers.

Morin, E. (1979). Essai de stylistique comparée. Technical report, University of Montreal, Montreal.

Morris, D. (1989). A semiotic investigation of Messiaen's abîme des oiseaux. *Music Analysis*, **8**(1-2).

Mylopoulos, J. and Lavesque, H. (1984). An overview of knowledge representation. In M. Brodie, J. Mylopoulos, and J. Schmidt, editors, *On Conceptual Modeling*. Springer-Verlag, New York, USA.

Nattiez, J. (1975). *Fondements d'une sémiologie de la musique*. Union générale d'éditions, Paris, France.

Nattiez, J. (1982a). *Three Musical Analyses*. Toronto Semiotic Circle. Toronto University, Toronto, Canada.

Nattiez, J. (1982b). Varèse's Density 21.5. *Music Analysis*, **I**.

Nattiez, J. (1997). The Tristan's cor anglais solo: A semiological analysis. In *Proceedings of the 3rd Triennial ESCOM Conference*, Uppsala, Sweden.

Naud, G. (1975). Apercus d'une analyse sémiologique de nomos alpha. *Musique en Jeu*, **17**.

Neisser, U. (1964). Visual search. *Scientific American*.

Pollard-Gott, L. (1983). The emergence of thematic concepts in repeated listening to music. *Cognitive Psychology*, **15**.

Ramscar, M. and Hahn, U. (1998). What family resemblances are not: The continuing relevance of Wittgenstein to the study of concepts and categories. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, University of Wisconsin, Madison, USA.

Réti, R. (1962). *The Thematic Process in Music*. Macmillan, New York.

Riepel, J. (1754). *Anfangsgründe zur musicalischen Setzkunst*. Montag, Regensburg, Germany, 2nd edition.

Rolland, P.-Y. (1998a). *Découverte automatique de régularités dans les séquences et l'application à l'analyse musicale*. Ph.D. thesis, Université Paris VI.

Rolland, P.-Y. (1998b). Flexpat: A novel algorithm for musical pattern discovery. In *Proceedings of the XII Colloquium on Musical Informatics*. AIMI Association.

Rolland, P.-Y. and Ganascia, J.-G. (2000). Musical pattern extraction and similarity assessment. In E. Miranda, editor, *Readings in Music and Artificial Intelligence*. Harwood Academic Publishers.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. Lloyd, editors, *Cognition and Categorization*. Laurence Earlbaum Assocs.

Rosch, E., Simpson, C., and Miller, R. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, **2**.

Rumelhart, D., Hinton, G., and Williams, R. (1986). Learning internal representations by error propagation. In D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*, pages 318–362. MIT Press, Cambridge, MA.

Rumelhart, D. E. and Zipser, D. (1986). Feature discovery by competitive learning. In D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*, pages 151–193. MIT Press, Cambridge, MA.

Ruwet, N. (1966). Méthodes d'analyse en musicologie. *Revue Belge de Musicologie*, **20**.

Ruwet, N. (1972). *Langage, Musique, Poésie*. Editions du seuil, Paris, France.

Ruwet, N. (1987). Methods of analysis in musicology. *Music Analysis*, **6**(1-2). translated and introduced by M. Everist.

Schoenberg, A. (1967). *Fundamentals of Musical Composition*. Faber and Faber Limited, London, UK.

Seeger, C. (1960). On the moods of a music logic. *Journal of the American Musicological Society*, **13**.

Smaill, A., Wiggins, G., and Harris, M. (1993). Hierarchical music representation for composition and analysis. *Computers and the Humanities*, **27**.

Smith, E. E. and Medin, D. L. (1981). *Categories and Concepts*. Harvard University Press, Harvard.

Stoianova, I. (1978). *Geste-Texte-Musique*. Union Générale d'Editions, Paris, France.

Tovey, D. (1936). *Essays in Musical Analysis: Concertos*, volume III. Oxford University Press.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, **58**.

Welker, R. (1982). Abstraction of themes from melodic variations. *Journal of Experimental Psychology: Human Perception & Performance*, **8**.

Westhead, M. and Smaill, A. (1994). Automatic characterisation of musical style. In A. Smiths, A. Smaill, and G. Wiggins, editors, *Music Education - An Artificial Intelligence Perspective*. Springer-Verlag, London, UK.

Wiggins, G., Harris, M., and Smaill, A. (1989). Representing music for analysis and composition. In C. Lischka, editor, *Proceedings of the 2nd IJCAI AI/Music Workshop*.

Wittgenstein, L. (1953). *Philosophical Investigations*. Macmillan, New York, USA.

Zbikowski, L. (1999). Musical coherence, motif, and categorization. *Music Perception*, **17**(1).

Appendix: Selected Publications

- Anagnostopoulou, C. (1997). Cohesion in Linguistic and Musical Discourse. *Proceedings of the Third Triennial ESCOM Conference*, Uppsala, Sweden.
- Anagnostopoulou, C. and Westermann, G. (1997). Classification in Music: A Computational Model for Paradigmatic Analysis. *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece.
- Anagnostopoulou C., Hörnel, D. and Höthker, K. (1999). Investigating the influence of representations and algorithms in Music Classificaiton. *Proceedings of the AISB'99 Symposium on Musical Creativity*, Sussex, UK.
- Anagnostopoulou, C. and Smaill, A. (2000). Similarity and Categorisation in Boulez' *Parenthèse* from the Third Piano Sonata. *Proceedings of the 3rd ESCOM and 6th International Conference of the Music Perception and Cogntion*, University of Keele, UK,
- Conklin, D. and Anagnostopoulou, C. (2001). Representation and Discovery of Multiple Viewpoints Patterns. *Proceedings of the International Computer Music Conference*, La Havana, Cuba.

Lexical Cohesion in Linguistic and Musical Discourse

Christina Anagnostopoulou*
University of Edinburgh
Faculty of Music and Dept of Artificial Intelligence
chrisa@music.ed.ac.uk

Abstract: An analogy is presented between music and linguistic discourse in terms of lexical cohesion. Linguistic discourse and music are two different manifestations of human communication. They are both meaningful and coherent, and both are processes inherently temporal. Lexical cohesion is a semantic relation manifested in the lexical level of language. A short sample of discourse is analysed and compared with a sample music analysis, regarding lexical cohesion. Their similarities are discussed, and the term Musical Cohesion is established.

1 Introduction

Linguistic discourse analysis, namely the analysis of text, and music analysis have developed in parallel and have studied similar phenomena. However, no systematic comparison of both techniques and their potential results exists to date. In this paper, linguistic discourse and music are compared with respect to cohesion. First, a linguistic analysis of a short sample of text is performed to demonstrate the concept of lexical cohesion. The musical counter-part follows, with an extended and modified paradigmatic analysis. The two are juxtaposed, making the similarities explicit, and the term Musical Cohesion is established. Finally, the property of linguistic cohesion contributing significantly to coherence and intelligibility is discussed in terms of music.

The analogy between language and music is here investigated from a perspective which is different to various influential existing approaches in three respects. Firstly, the focus is on discourse rather than sentence level. This seems appropriate since both music and linguistic discourse are instances of human communication, carry some kind of meaning, are intelligible, and above all both are inherently temporal. Secondly, a semantic rather than a syntactic relation is examined, namely cohesion,

*I would like to thank Hugh Trappes - Lomax and Raymond Monelle for helpful discussions

which makes it possible to address a semantic level in music, and finally, the investigation focuses on the surface level by looking at associative features, without making any claims concerning underlying structure.

2 The Discourse Analysis Background

Discourse[1] is any linguistic passage that forms a semantic unity, spoken or written, of whatever length or form. The major and necessary factor that causes a linguistic passage to be a text rather than an arbitrary string of sentences is *cohesion*. Cohesion occurs when some element in the discourse either presupposes the existence of another for it to be interpreted, or is semantically linked to one. For example, in the following beginning of a discourse: "A thermodynamics professor had written a take home exam for his graduate students (...)", the interpretation of his presupposes the existence of thermodynamics professor. This type of cohesion is called *reference*.

The type of cohesion discussed here is *lexical cohesion*, a semantic property manifested on the lexical level by the use of specific words that are either identical or semantically close. It can be divided into reiteration and collocation.

Reiteration includes the exact repetition of a word, a synonym, super-ordinate, or general word, for example, in "I turned to the ascent of the peak. The ascent is perfectly easy", the word ascent is repeated. Instead of its second occurrence, one could have also used the climb, task, thing as examples of reiteration ([4], p.279).

Collocation is manifested by the use of words that are semantically related in some more distant way, but can still be thought of as belonging to the same semantic network. Examples are the pairs exothermic-endothermic and answer-proof in "Is hell exothermic or endothermic? Support your answer with a proof."

The above examples demonstrate only pairs of related words, where in fact there can be whole strings, named *cohesive chains*. In "As for souls entering hell, lets look at the different religions that exist today (...)", the chain is souls-hell-religions.

Lexical cohesive chains are sets of words classified together according to semantic closeness or similarity. There can be several chains in a text, running in parallel, and interweaving. They can be global (during the whole discourse) or local (for a part of it).

2.1 An example analysis

In the following text, five cohesive chains are displayed by using different font styles. Reference is also noted (by an asterisk), but only when the presupposed item belongs to one of the lexical chains. Reiteration and collocation are not distinguished.

[1] The approach described here follows [4], which is considered to be the standard and most well-accepted account on cohesion. Note that the terms Discourse and Text are used interchangeably.

Soon |her* eye| fell on a little glass box that was lying under the table: she* opened it, and found in it a very small |**cake**|, on which* the words "|**eat**| me*" were beautifully marked in |**currants**|. 'Well, I*'ll |**eat**| it*," said |Alice|, "and if it* makes me* |LARGER|, I* can reach the key; and if it* makes me* |SMALLER|, I* can creep under the door; so either |*way*| I*'ll get into the garden, and I* don't care which happens!"

She* |**ate**| a little bit, and said anxiously to herself*, "Which |*way*|? Which |*way*|?" holding |her* hand| on the top of |her* head| to feel which |*way*| it* was |GROWING|, and she* was quite surprised to find that she* remained the same |SIZE|: to be sure, this generally happens when one |**eats**| |**cake**|, but |Alice| had got so much |*into the way*| of expecting nothing but |*out-of-the-way*| |THINGS| to happen, that it seemed quite |*dull*| and |*stupid*| for life to go on |*in the common way*|.

So she* set to work, and very soon finished off the |**cake**|*.

(text quoted in [4], p.319).

Figure 1 shows how the above chains (together with their references) unfold through time: the x-axis represents the word number of the discourse, and the level on the y-axis represents the different chains: The first chain (her eye, Alice, her hand, her head, Alice) is shown at level 5, the second (cake, eat, currants, eat, ate, eats, cake, cake) at level 4, the third (larger, smaller, growing, size, things) at level 3, the fourth (way, way, way, into-the-way, out-of-the-way, in-the-common-way) at level 2, and the last (dull, stupid) at level 1. The chains alternate, and some of them are global like "Alice" and "cake", and some local, like "dull".

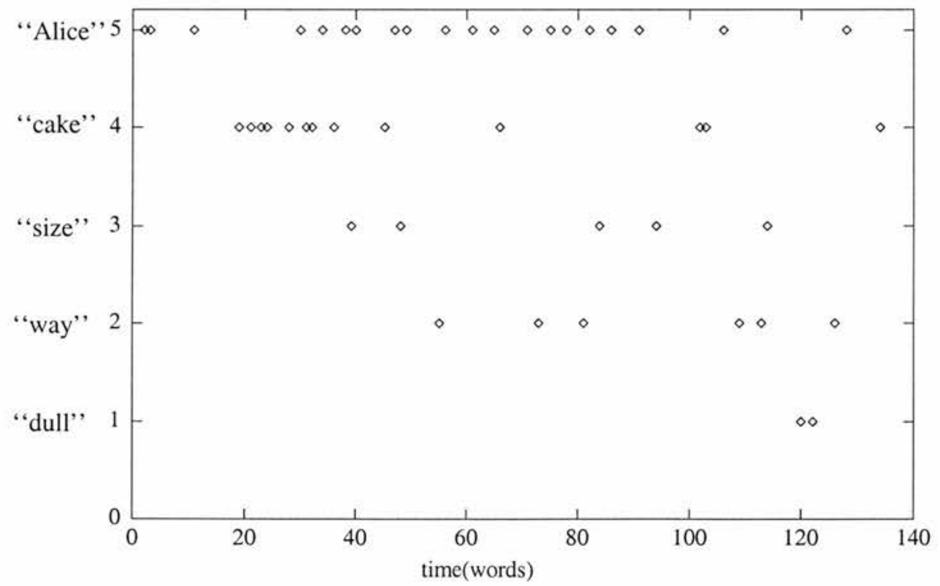

Figure 1: Cohesive Chains from linguistic discourse

3 The Music Analysis

Repetition, variation and transformation have been studied extensively in music. Prime examples are paradigmatic [6] and motivic analysis, and various other classifications of musical material have been carried out, for example [3, 2, 1].

The method of analysis chosen here is a type of paradigmatic analysis, carried out using a computational model of paradigmatic analysis [1]. The analysed piece is Debussy's *Syrinx* for solo flute. The results, some of which are shown here, were very close to Nattiez' second paradigmatic analysis of the same piece [6]. In order to obtain a classification, each musical segment is described as a list of features. These features are chosen by the analyst, and they can be any musical property, for example concerning melodic shape, rhythm, and whatever the analyst chooses to be his criteria for classification. The segments (described as lists of features) are classified by the paradigmatic analysis algorithm. The categorisation of the segments is hierarchical: there can be many levels, and categories can be divided into subcategories. Also, segments that are not repeated are left out.

Figure 2 shows three sample classes: classes A and B are global, whereas C is local. Most of these segments (apart from the ones in class C) are repeated throughout the piece, but here they are included only once in each class[2].

Figure 2: Some of the classes from an analysis of Debussy's *Syrinx*.

[2]Due to space limitations, the whole analysis is not included here. The method of paradigmatic analysis is considered familiar.

4 Comparison

From the above analyses, two points can be made regarding the similarities between discourse and music.

- Classes of objects can be observed where the objects share common properties and are classified together because of their similarity. There exist various classes with contrasting material which nonetheless can also share certain properties, and form a hyper-class at a higher level.
- These classes are distributed over time. Some classes can be local (like the "dull" chain and the C class above), and some can be global, like "Alice" or A. Classes alternate, following no specific rule). However, one could observe some patterns of sequences, for example occurrence of class 1 might always be followed by occurrence of class 2.

4.1 Musical Cohesion

The formation of classes in both linguistic discourse and music depends on similarity: repetition and variation. In discourse analysis the effect created by these principles is called cohesion. In music, since there is no term for such an effect, the linguistic term can be adopted as *musical cohesion*.

It is important to note that cohesion is a relation that appears in the text itself, it is visible or audible, and can be brought out and studied. Therefore, it is a relation that appears at the *Neutral Level* of Discourse[3]. Similarly, musical cohesion appears on the Neutral Level of music. It can be studied objectively and formally.

It has been pointed out that cohesion is a semantic relation, manifested, in the case of lexical cohesion, at the lexical level. In the same way, musical cohesion is not a relation of the equivalent on the phonological level in language, but a semantic relation, manifested in sounds. The similarity criteria for classification were musical properties extracted from the musical segments (and not the segments themselves). These properties can be linked formally into a semantic web.

4.2 Discussion

Coherence is an attribute a text possesses on the aesthesic level, that is in its perception: it is a reaction that we have to a text that it "hangs together". Cohesion supports coherence, although it is not the only factor contributing to it. It could be argued that musical cohesion is a major factor contributing to coherence, although the degree to which this holds might vary in comparison to language.

[3] According to Molino's distinction of the three levels, Neutral, poietic, aesthesic, [5].

This issue, and the investigation of other significant factors contributing to coherence (like context) are beyond the scope of the current paper, but are a major direction for future work. Moreover, it would be interesting to investigate how other types of cohesion can be related to music, and in general to explore what discourse analysis has to offer for music analysis. However, as with any parallelism between language and music, although there are self-evident similarities, one should not neglect their differences.

References

[1] C. Anagnostopoulou and G. Westermann. Classification in muisc: A computational model of paradigmatic analysis. In *Proceedings of The International Computer Music Conference*, Thessaloniki, 1997.

[2] E. Cambouropoulos. The role of similarity in categorisation: Music as a case study. In *Proceedings of the Third ESCOM Trennial Conference*, Uppsala, 1997.

[3] R.O. Gjerdingen. Using connectionist models to explore complex musical patterns. In *Music and Connectionism*, pages 138–149. MIT Press, 1991.

[4] M.A.K. Halliday and R. Hasan. *Cohesion in English.* Longman, 1976, 1995.

[5] J. Molino. Fait musical et sémiologie de la musique. *Musique en Jeu*, 17:37–61, 1975.

[6] J.J. Nattiez. *Fondements d'une Sémiologie de la Musique.* Union Générale d'Editions, 1975.

Classification in Music: A Computational Model for Paradigmatic Analysis

Christina Anagnostopoulou (1)
Gert Westermann (2)
(1) Faculty of Music and Dept of Artificial Intelligence, University of Edinburgh
chrisa@music.ed.ac.uk
(2) Centre for Cognitive Science, University of Edinburgh
gert@cogsci.ed.ac.uk

Abstract

We present a computational model for the paradigmatic analysis of musical pieces, which is the classification of musical segments into similarity-based categories. The model requires the analyst to make explicit choices for the characteristics by which the musical segments are described. The classification of the segments into categories is determined by these characteristics and is performed by a self-organizing neural network algorithm. In this way, traditional problems associated with paradigmatic analysis, namely lack of consistency and objectivity, can be avoided. Moreover, the model extends the analytical technique by providing different levels of classification, prototypes for each class, and by showing relations between classes.

1 Introduction

The paradigmatic analysis (henceforth PA) of musical pieces has long been criticized for its reliance on intuition and the resulting inconsistencies [1]. In this paper, we describe a formal model for this task as a way to address such criticisms. In the next section, traditional PA is described and its shortcomings are discussed. The formal model is then presented, and its functioning is demonstrated by analyzing Debussy's *Syrinx* and comparing the results with J.J. Nattiez' (the leading figure in PA) second analysis of this piece [4]. We conclude with a discussion of the model and suggestions for further work.

2 Paradigmatic Analysis

PA consists in the segmentation of a piece of music and the classification of these segments into categories according to their similarity. The first occurrence of a segment in each class is called the *paradigm*, and subsequent segments are compared to these paradigms to determine their class membership. The paradigms therefore play the role of class prototypes.

The motivation for this kind of analysis is that repetition, variation, transformation, and contrast within a musical piece are made explicit. Further analysis is thus facilitated—this can be distributional, comparative or stylistic.

The main original goal of PA was to give a formal, objective account of the material used in a piece, not taking into account the composer's intentions or the listener's perceptions. In practice, however, the assignment of the musical segments to different classes usually relies on intuition: According to Nattiez, *"People decide to associate several units in a single paradigm because of semantic or psychological criteria that they do not express consciously."* (quoted in [1], p. 180). This lack of explicit criteria underlying the classification will naturally lead to inconsistencies in the analysis, and this has in fact been the main criticism of PA. Further criticisms address its limited character: there is only one level of classification, when subcategories could easily be identified and could prove to be useful. Relationships between different classes are not considered, although some classes will be more similar than others. Moreover, as in most other analytical techniques, the segmentation of a piece has been criticized as being usually informal, which is a problem for the subsequent classification. Finally, the paradigms against which other segments in a category are compared are merely a first occurrence and not necessarily prototypical of their

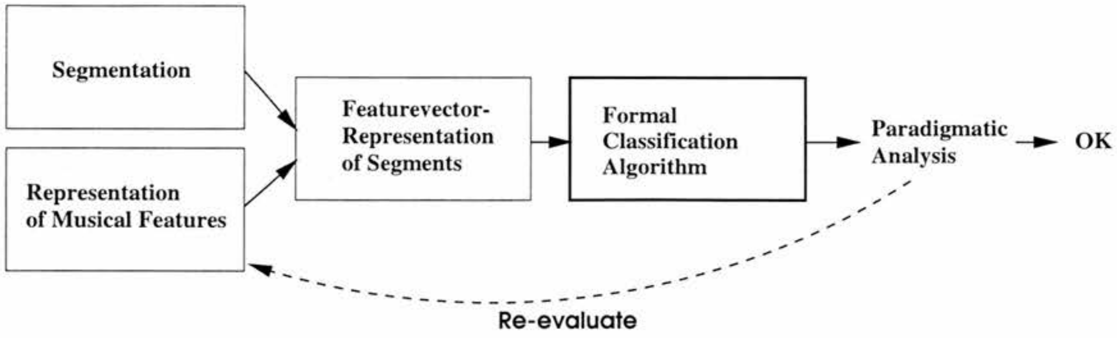

Figure 1: A formal model for Paradigmatic Analysis

category (see section 4 for an example from Nattiez' analysis), and potential inconsistencies may arise by comparing each segment to a paradigm which is not prototypical.

3 A Formal Model

Classifying objects is a fundamental task which has been studied in depth in other disciplines, such as formal learning theory, computer science, and psychology. However, the existing classification theories and techniques from those disciplines have not generally influenced music analysis (but see [3]).

In a formal model, the algorithm by which the segments of the piece are classified has to be explicitly defined. Also, the feature representations on which the classification is based have to be made explicit, so they can serve as input to the classification algorithm. A formal system further leads to the modularization of PA into different subtasks, each of which can be solved independently. Figure 1 shows such a formal model. In the first step, the musical piece is segmented, and the analyst has to decide on the way in which musical features are to be represented. In the second step, each segment is expressed as a list of features (a *feature vector*), which is then used as input to a classification algorithm. The output of this module constitutes a PA. If the analyst is not satisfied with the clustering obtained, he will re-evaluate the feature representations. For example, if two segments which the analyst considers to be different are grouped together by the model, he will introduce a feature that distinguishes these segments from one another. Based on the resulting new representation of the segments, a different classification will occur. This process is repeated until a satisfactory classification is obtained. This final classification will be based on explicit segments and features and will be free from inconsistencies.

In summary, a formal model of paradigmatic analysis serves as a tool for the analyst, forcing her to make her choices of representation explicit and providing a well-defined algorithm for the clustering of segments, without restricting the freedom to choose the classification criteria.

3.1 Segmentation

It is generally accepted that there is no single "correct" way of segmenting a piece of music. Segmentation is a problematic issue for any kind of musical analysis, and therefore ideally a system should accept any analyst's choice on segmentation. The modular character of the present system allows this approach. In that way, different segmentations can be compared—it is obvious that the most sensible ones will result in the most sensible classifications.

In our example we used J.J. Nattiez' segmentation from his second PA of Debussy's *Syrinx* for solo flute [4].

3.2 Representation of Musical Features

Each segment is described in the formal model as a list of features. The term "feature" is used here not only in the traditional sense, that is with binary values (yes/no) , but also more generally, to include multi-valued attributes, and in fact any hierarchic relation in a semantic network. An example of a feature with a binary value would be the existence of a grace-note in a segment. An example of a multi-valued feature would be instrument register: it could be the first octave of the flute, the second or the third, or any combination of these. Examples of hierarchic relations are shown below for melodic shape and rhythmic movement.

It is obvious that the results of the PA will depend crucially on the feature selection. The analyst

Segm.	Music Notation	Feat. vector	Class
1, 6		1 0 0 0 0 ···	I
4		0 1 0 0 0 ···	I or II
2, 7		1 0 0 0 0 ···	I
24, 28		0 1 0 0 0 ···	II
52		0 1 0 0 0 ···	II

Table 1: Classification examples for several segments from *Syrinx*.

will choose the features according to the desired outcome: for example, he might choose to focus on a rhythmic analysis, or compare several pieces of music according to a set of common features. Since similarity in music could be argued to be context-dependent (context being the piece or pieces under analysis), features can be low-level, piece-specific (e.g., the use of a specific interval), as well as very general musical properties (like *upward melodic motion*).

For our experiments we chose a combination of general and of piece-specific features, describing

- melodic shape (moving—up, down, or different combinations— or stationary).
- rhythmic movement (continuous—which can be quaver, semiquaver or demisemiquaver movement— or interrupted—by a dotted rhythm, syncopation, long note or a pause—).
- interval patterns (with instances being low-level successions of intervals) and
- instrument register (in order to describe transposition).

These features proved to be sufficient for the final classification.

The features describing a segment were concatenated to form a feature vector, in our case with 40 binary values, which was then used as input to the classification algorithm. Table 1 shows several segments out of the 79 from *Syrinx* with part of their feature vector representations.

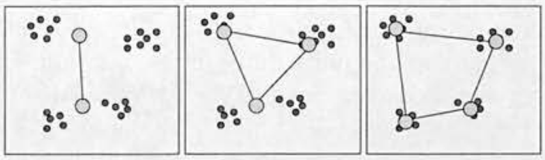

Figure 2: Construction of a GNG network. Small circles represent input data, and large circles connected with edges are the units of the network.

3.3 The Classification Algorithm

The classification algorithm which we chose to employ for our experiments was *Growing Neural Gas* [2]. This is an unsupervised neural network algorithm that grows units while it learns. Each unit corresponds to the prototype of one cluster. An input signal, i.e., a feature vector representing a musical segment through 40 binary features, can be viewed as having a position in the 40-dimensional input space, and the units of the network are positioned in the same space. When an input signal is presented, the unit of the network which is closest to it (measured by Euclidean distance) is moved towards this signal by a fraction of the distance to this signal, together with its topological neighbours. The distance between the signal and the winning unit is added to a local error variable of this unit. The winning unit and the second closest unit are then connected by an edge, or the age of the edge is reset to zero, if it already exists. The edges reflect neighbourhood relations between the network units. At each step, all edges in the network are aged, and edges which have reached a pre-defined maximum age are deleted. This process ensures a continuous updating of the neighbourhood relations.

A new unit is inserted into the network at regular intervals, between the unit with the highest accumulated error and its neighbour with the highest error. The built-up structure of the network reflects the distribution and density of the input signals: The units move towards the input signals, and a high density of inputs in an area will lead to more units being allocated in this area. Figure 2 shows the development of a network in a two-dimensional input space with four distinct clusters. The network starts with two units and can therefore distinguish only between the two main clusters. In effect, the network answers the question: If there were two clusters, what would their prototypes be? After a certain amount of epochs (presentation of the input signals), a new unit is inserted and the units move to the positions indicated in the second picture. When the fourth unit is inserted, the units distribute over the four clusters.

In principal, insertion of units proceeds forever. The GNG algorithm thus lets the analyst define the level of grainedness of her analysis and does not impose *a priori* constraints on the number of clusters. Each unit forms a prototype of a cluster, expressed in the probability distribution of the feature values of their cluster members. These prototypes are frequency-dependent, since the position of each unit is updated with each presentation of an input signal. Neighbourhood relations between clusters are expressed in the connections between the network units.

4 Experimental Results

The GNG algorithm was trained on the musical feature vectors for 2000 iterations, inserting a new unit every 100 iterations (2 minutes CPU time on a Sun Ultra workstation). Thus, the final classification consisted of 20 categories, and by comparing this to previous stages, the hierarchy of clusters could be observed.

We ran various experiments with different input representations. With our final representation (which is mentioned above) we obtained an intuitively sensible analysis which was surprisingly close to Nattiez' second paradigmatic analysis. Due to the lack of space it is impossible to give the various results obtained in full length. In table 1, segments 1, 2, 6 and 7 belong to the same class; segments 24, 28 and 52 belong to another. Segment 4 is a problematic segment in that it can be classified to either of these two categories, according to different input representations. It would be classified with the first category when rhythmic features are taken into account, and with the second if melodic shape is emphasized. In our experiments, after 2000 iterations, this segment formed a class by itself, but was linked by edges to both class I and class II.

Nattiez classified segment 4 according to melodic shape and not to rhythm. The problem with this is that since segment 4 is the first occurrence for such a melodic shape, it becomes a paradigm. However, further segments are more and more varied, and the result is that segment 4 is in no way prototypical of the whole category. In our analysis, this problem was avoided because prototypes represented the weighted average of all class members and not the first occurrence of a class member.

5 Conclusions

We have demonstrated a formal model of paradigmatic analysis. The model requires the analyst to make the categorization criteria explicit without restricting his particular choices. The classification produced by the model is then entirely based on these choices and depending on the obtained results, they can be revised by the analyst. The model yields different levels of classification, prototypes for each class and relations of classes of the same level.

The model can be extended in various ways. Firstly, it is possible that the analyst wishes to attribute different importance to different features. Such a weighting of features should be incorporated into an extended model. Secondly, while the present version relies on a given segmentation of a piece, in principle a revision of this initial segmentation could be incorporated into the classification process, combining the stages of PA in a single unified model. This will be our main direction of future research.

6 Acknowledgements

We are grateful to Peter Nelson for comments on a draft of this paper. The second author was funded by the ESRC (award no. R00429624342) and by the Gottlieb Daimler-und Karl Benz-Stiftung (grant no. 02.95.29).

References

[1] N. Cook. *A Guide to Musical Analysis*. Oxford University Press, 1987.

[2] Bernd Fritzke. A growing neural gas algorithm learns topologies. In G. Tesauro, D. S. Touretzky, and T. K. Lean, editors, *Advances in Neural Information Processing Systems 7*, pages 625–632, Cambridge, MA, 1995. MIT-Press.

[3] R.O. Gjerdingen. Using connectionist models to explore complex musical patterns. In *Music and Connectionism*, pages 138–149. MIT Press, 1991.

[4] J.J. Nattiez. *Fondements d'une Semiologie de la Musique*. Union Generale d'Editions, 1975.

Investigating the Influence of Representations and Algorithms in Music Classification

Christina Anagnostopoulou *; Dominik Hörnel † ‡; Karin Höthker †
* Faculty of Music, University of Edinburgh
† Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe
‡ Hauptabteilung Prozeßdatenverarbeitung und Elektronik, Forschungszentrum Karlsruhe
chrisa@dai.ed.ac.uk; dominik@ira.uka.de; hoethker@ira.uka.de

Abstract

Classification in music involves the segmentation of a piece and the categorisation of the segments depending on similarity based criteria. We present an investigation on how varying the knowledge representation and clustering algorithms influences the result. More specifically, we vary the selection of the features that describe each segment, the way these features are represented and the clustering algorithms. While doing this, we keep the other parameters, that is the overall model architecture, the music pieces and the segmentation fixed. We introduce a distance function to compare the results of algorithmic and human classification. We can show that the algorithmic results are very close to human analysis, if an appropriate representation has been found. The results allow an objective evaluation of various approaches to music classification in a uniform setting.

1 Introduction

Classification within a piece of music involves the breaking up of the piece into segments and the categorisation of these segments according to similarity based criteria. In music analysis, this method is also known as *Paradigmatic Analysis* (PA), originally proposed by Ruwet (1966) and developed further by Nattiez (1975). It is a widely used method because it provides a useful and "objective" first step for most further formal musicological study, like stylistic, comparative and motivic analysis.

There have been various formal and computational models for music classification, using different architectures, representations and algorithms. In this paper we investigate the impact of varying the following two aspects:

- the knowledge representation of the musical segments, specifically the choice of musical features that describe the segments and the way of representing these features
- the algorithms used for the classification of these segments.

We introduce a distance function which allows us to compare the various classification results in an objective way. The pieces we choose to analyse are 9 two-part Bach inventions.

2 Motivation

Several attempts have been made sofar to formalise the musical classification process, using various segmentation techniques, knowledge representations, clustering algorithms and general frameworks (e.g. Gjerdingen, 1990; Cambouropoulos and Smaill, 1997; Anagnostopoulou and Westermann, 1997; Hörnel, 1998; Rolland, 1998). A comparison of these existing formal models can tell us more about their generalisation properties with respect to the various musical styles. For example, testing classification methods on the *same* musical data set allows direct conclusions about their classification capabilities.

This comparison can serve as a first step towards creating a "toolkit" for music analysts, where various algorithms and representations are available for different musicological purposes. Apart from the analysis, this can also be used as a basis for the generation of a stylistically coherent musical (melodic) structure, for example for multi-scale neural network composition (Hörnel, 1998).

3 Methodology

In this section we first draw an outline for a classification model and then explain our experimental methodology and evaluation procedure. We propose a formal model of music classification, or PA, (cf. Figure 1) where the architecture is modular and therefore allows us to experiment by substituting different modules with equivalent ones. The segmentation module breaks the musical piece into small segments and the music representation module provides the knowledge for the description of these segments. The segments are transformed into feature vectors, and then the clustering algorithm performs the classification of these segments. Different classification re-

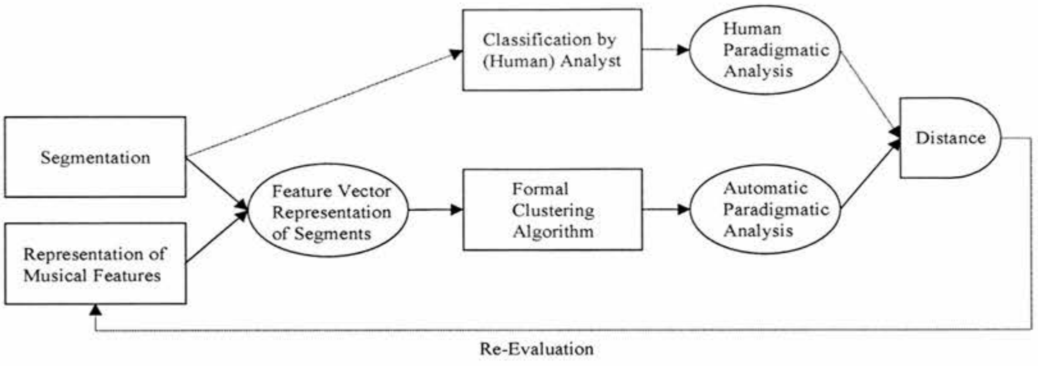

Figure 1: A formal model of classification

sults can be compared by using a human classification as a reference point and calculating a distance that represents the similarity between them. Although different approaches have used different model architectures, representation and classification components are common to all of these models and can be tested independently of the general framework.

3.1 Segmentation

Segmentation is an important issue for music analysis in general, since there is no single "correct" way of segmenting a piece of music, and different segmentations give very different results - especially in a classification task. Although we acknowledge the problem, we have chosen not to look at this issue here, and instead to keep a steady, fixed segmentation for all our experiments. This allows us to concentrate on the issues of knowledge representation, feature selection and clustering algorithms. The segmentation we used for our experiments is a break on every crotchet beat.

3.2 The Knowledge Representations

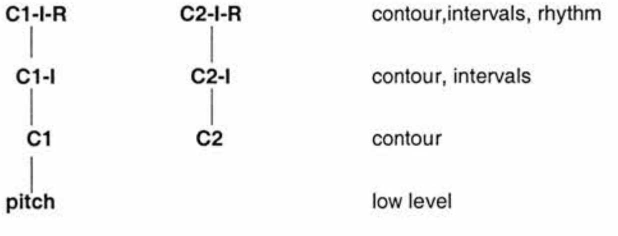

Figure 2: The various representations used: C indicates melodic contour information, I interval information and R rhythm information. C1 and C2 are two principally different ways of representing melodic contour.

In order to become an appropriate input for the classification module, each segment needs to be transformed into a feature vector. The questions we address here are which musical features to represent and how. Figure 2 shows the 7 different representations we chose for our investigation, and how these are related to each other. In Figure 3, a single motive is represented using these seven kinds of representation.

C1 describes whether the next note goes up (1), down (-1) or is stationary (0). We use the semiquaver beat as a unit, that means a quaver can be thought of as two identical semiquavers. This assumption does not affect the contour information.

C1-I To C1, interval information is added: 0.2 for an interval of a second, 0.4 for a fourth and so forth. 0.1 shows the interval of a "prime", that is the same note.

C1-I-R To the above, rhythmic information is added by notating whether there is a new note on each semiquaver position (0) or not (1).

C2 The features describing melodic contour are structurally related into a *tree* where inheritance is preserved. Each node of the tree is represented by a separate position in the feature vector (1 if the feature exists, 0 if not). Because of the inheritance, all the path that leads to the terminal node that is 1, is also 1. Figure 4 shows the features used for C2.

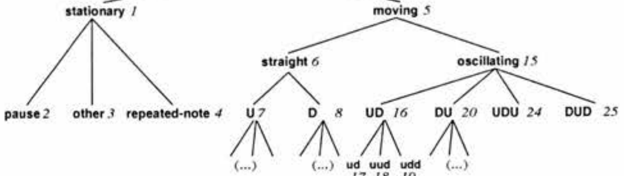

Figure 4: The tree-structure representation C2. The number next to each feature shows its vector position in figure 3. U=up, D=down.

C2-I The same interval information as above is added.

C2-I-R To the above the same rhythm information is added.

C1-I-R	0 0.1 1 1 0.3 0 -1 0.2 0
C1-I	0 0.1 1 0.3 -1 0.2
C1	0 1 -1
pitch	1.4 0 1.4 1 1.6 0 1.5 0
C2-I-R	0 0 0 0 0.17 0 0 0 0 0 0 0 0 0 0.17 0.17 0.17 0 0 0 0 0 0 0 0 0.1 0.33 0.3 0 0.2 0
C2-I	0 0 0 0 0.17 0 0 0 0 0 0 0 0 0 0.17 0.17 0.17 0 0 0 0 0 0 0 0 0.1 0.3 0.2
C2	0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0

motive 

Figure 3: A representation example

Pitch, the last representation, is a very crude low level representation that contains only information on absolute pitch and rhythm.

Out of the two ways of representing melodic contour, C1 is the most efficient and intuitive. However, unwanted a priori similarity judgements are not avoided. Imagine a segment with stationary melodic contour, *C-C-C*, one ascending - descending, *C-E-C*, and one descending - ascending, *C-A-C*. This kind of representation assumes that the second segment is closer to the first than to the third, which might not be intuitively obvious. C2 overcomes this problem by introducing the upper node "oscillating" in the tree, which is common to the second and third segment, therefore increasing their similarity. However, this procedure, apart from being computationally more expensive due to its dimensionality increase, has the further problem that certain similarities are encoded a priori.

3.3 The Algorithms

In this section we present an outline of the clustering algorithms used in our experiments. To illustrate the algorithms, we use a simple musical example which consists of five different 3-note segments (see Figure 3), and show how the algorithms classify these. For the purpose of the example, we keep a steady representation, i.e. C1-I in a two-dimensional vector form. This representation considers the interval similarity between motives one, two and three, but does not capture the pitch similarity between motive one and five.

Ward (Ward, 1963) is an agglomerative hierarchical data clustering algorithm. The idea is to gradually merge sets of data elements which are closest to each other. A distance measure D between two sets A and B is defined as the doubled difference between the homogeneity H of the merged set and the sum of the homogeneity of the original sets.

$$D(A, B) = 2(H(A \cup B) - (H(A) + H(B)) \qquad (1)$$

Using the euclidean distance, the homogeneity H of set S is computed as the variance Var without normalization.

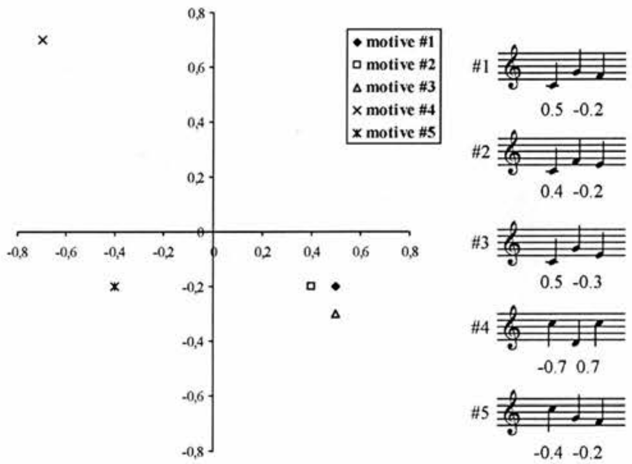

Figure 5: Musical Data

$$H(S) = \sum_{i=1}^{n} \| s_i - \overline{s} \|^2 = (n-1)Var(S) \qquad (2)$$

Starting with individual elements, the algorithm recursively merges pairs of element sets and recomputes the distance between them until all elements belong to one set. The result can be represented as a binary tree called *dendrogram* (see Figure 6).

Growing Neural Gas (Fritzke, 1995) is an unsupervised neural network algorithm which is able to learn the topological relations in a given set of input vectors. This is done by means of a simple Hebb-like learning rule which adapts the reference vector w_s of unit s by fraction ε according to an input signal ξ.

$$\Delta w_s = \varepsilon(\xi - w_s) \qquad (3)$$

Starting with two units, new units are inserted successively. To determine where to insert new units, local error measures E_s are calculated during the adaptation process for the unit s which is nearest to ξ.

$$\Delta E_s = \| \xi - w_s \|^2 \qquad (4)$$

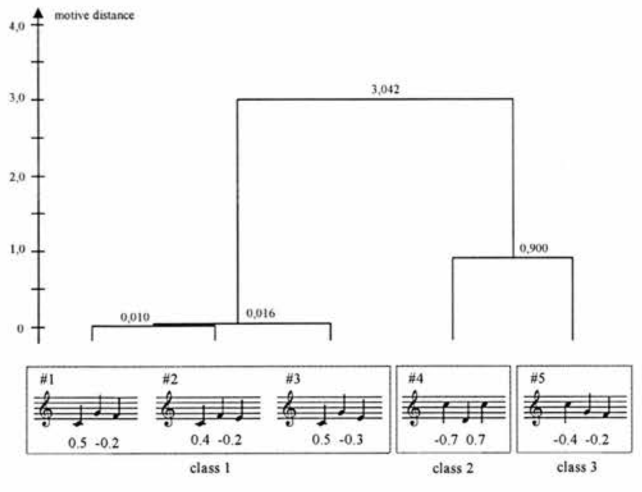

Figure 6: Dendrogram produced by the Ward algorithm

Each new unit is inserted near the unit which has accumulated the largest error. The algorithm has no parameters which change over time and is able to continue learning, adding units until a performance criterion has been met (e.g. a specified number of classes has been reached). Figure 7 shows the resulting network. Observing the structure of the net during the growth process gives insight into the hierarchical organisation and relation of the classes. The classification produced for the example is the same as for the Ward algorithm.

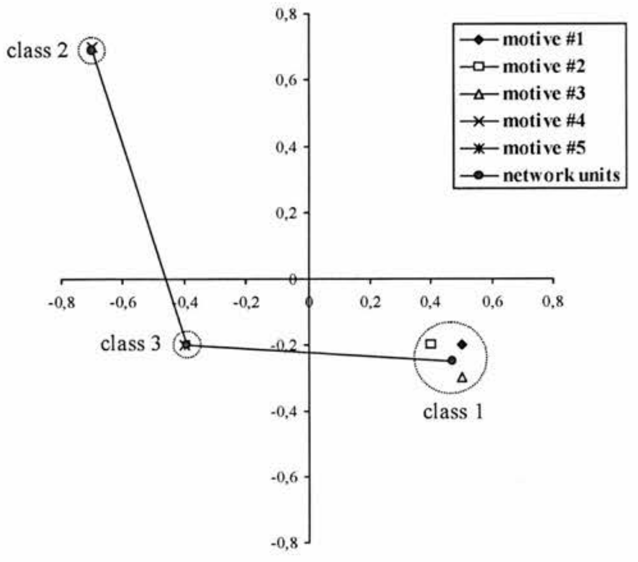

Figure 7: Growing Neural Gas Network with three units after 100 learning cycles

Star Center (Rolland, 1998) is an algorithm which is influenced by an application proposed in molecular biology (Gusfield, 1997). Given a similarity graph, it extracts a list of vertices (stars) in decreasing prominence. A star consists of a vertex - the prototype - and its neighbours and can be interpreted as a motive class. In the similarity graph, the vertices v_1 and v_2 are only connected if their similarity value reaches a specified threshold θ.

$$Similarity(v_1, v_2) \geq \theta \tag{5}$$

The prominence or *totalValuation* of a star v is computed as the sum of the similarity values to all neighbours.

$$totalValuation(v) = \sum_{v' \in adj(v)} value(v, v') \tag{6}$$

The set of stars is then sorted by decreasing *totalValuation*. Extracting the first k elements of the resulting list, one gets k motive class representatives (Figure 8).

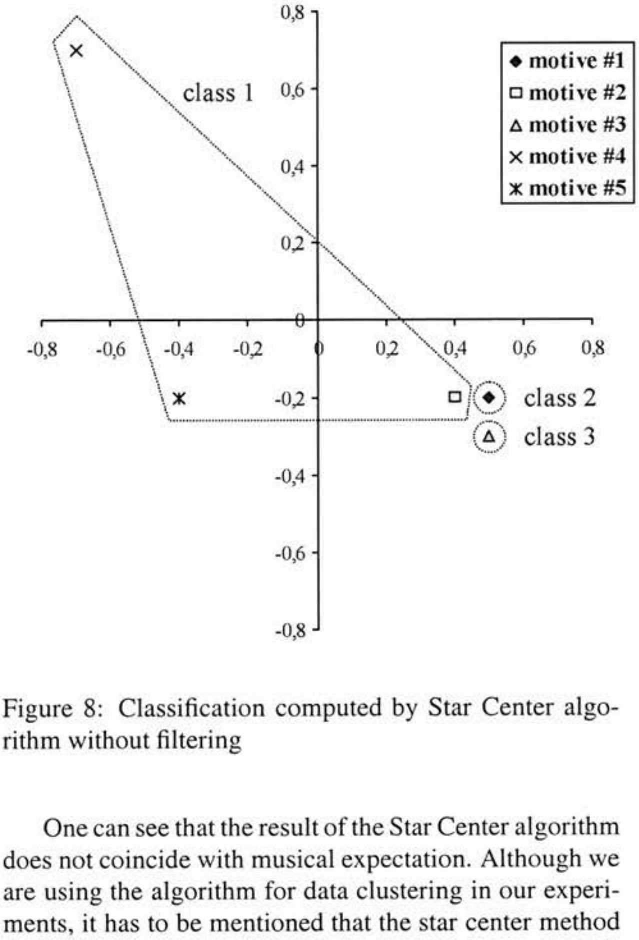

Figure 8: Classification computed by Star Center algorithm without filtering

One can see that the result of the Star Center algorithm does not coincide with musical expectation. Although we are using the algorithm for data clustering in our experiments, it has to be mentioned that the star center method has not been developed for this purpose, but more specifically for pattern extraction. It prefers star centers which are close to many other motives within an environment given by θ, in our case motive one, two and three. If we do classification, all other motives have small *totalValuation* and are therefore assigned to one of these prototypes instead of forming their own classes. For this reason we use a version of the algorithm that is modified in the following way: The resulting list of stars is filtered sequentially by eliminating all stars which are within the environment of another star with higher *totalValuation*. As a result, the filtered list S of stars fulfils the following condition:

$$value(v_i, v_j) < \theta : \forall v_i, v_j \in S, i \neq j \quad (7)$$

In using the *Kohonen* network (Kohonen, 1990), we transform each of the knowledge representations into a two-dimensional grid of a Kohonen feature map and use the Ward algorithm to cluster the data pairs afterwards.

In the training phase, the winner neuron with highest activation is determined for each input signal x_i. Then the network weights w_{ij} of the winning unit and of units within a certain neighbourhood radius are adapted according to

$$\Delta w_{ij} = e_j(t)(x_i(t) - w_{ij}(t)) \quad (8)$$

where $e_j(t)$ is a Gaussian function whose expansion depends on the neighbourhood radius. The radius is decreased over time to allow convergence of the network weights. In our example the Kohonen-Ward algorithm produces the same result as the Ward algorithm since the data is already two-dimensional.

3.4 The Distance Function

We evaluate PAs generated by different clustering methods by comparing them to an analysis done by a human analyst. In order to get general and reproducable results, we use a distance function reflecting the degree of similarity between two PAs.

Since paradigmatic analysis abstracts from motive instances by assigning them to motive classes, we compute our distance function based on a sequence of class labels, the *abstract motive sequence* (e.g. $ababccdc$).

When describing the structure of an abstract motive sequence $M = (m_1, ..., m_n)$, the key question is whether two motives m_i and m_j at different positions of the sequence belong to the same class or not. The *relationship function* $\mathrm{rel}_{ij}(M)$ captures this aspect:

$$\mathrm{rel}_{ij}(M) = \begin{cases} 1 & : \quad m_i = m_j \\ 0 & : \quad m_i \neq m_j \end{cases} \quad (9)$$

The matrix $(rel_{ij}(M))_{i,j=1}^{n}$ defines a graph on the sequence positions 1 through n which shows the motivic relationships within the sequence (Figure 9). Note that the graph associated to an abstract motive sequence is always a union of complete, disjoint subgraphs.

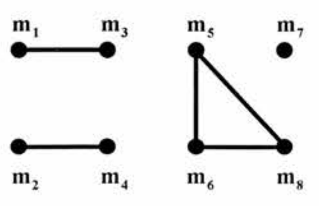

Figure 9: Graph for the abstract motive sequence $ababccdc$

Comparing the values of the relationship function allows to detect structural differences **between** motive sequences (still considering fixed positions i and j). The *exclusive-or function* $\mathrm{xor}_{ij}(L, M)$ identifies opposite relation values in abstract motive sequences M and L:

$$\mathrm{xor}_{ij}(L, M) = \begin{cases} 1 & : \quad \mathrm{rel}_{ij}(L) \neq \mathrm{rel}_{ij}(M) \\ 0 & : \quad \mathrm{rel}_{ij}(L) = \mathrm{rel}_{ij}(M) \end{cases} \quad (10)$$

The *distance function* for abstract motive sequences is then obtained by summing over all pairs of positions and normalizing with the maximally possible difference:

$$\mathrm{dist}(L, M) = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathrm{xor}_{i,j}(L, M)}{\frac{1}{2}n(n-1)} \quad (11)$$

The distance function is independent of the actual class labels. This is a reasonable property for a distance function on abstract motive sequences since the outcome of paradigmatic analysis is not affected by relabelling classes.

The distance function is a metric. By definition, a metric yields only values greater or equal to zero, it is symmetrical (i.e. $\mathrm{dist}(L, M) = \mathrm{dist}(M, L)$), it is zero if and only if the elements to be compared are identical. In the context of paradigmatic analysis we consider abstract motive sequences to be equal if their graphs are equal which means that relabelling classes does not lead to a different abstract motive sequence. While the previous properties follow immediately from the definition of the distance function, the last property of a metric, the triangle inequality $\mathrm{dist}(K, M) \leq \mathrm{dist}(K, L) + \mathrm{dist}(L, M)$, can easily be shown by using that $\mathrm{xor}_{ij}(L, M)$ is a metric for fixed i and j.

3.5 The Human Analysis

The results of the experiments with the various representations and algorithms are compared to a human analysis (see Figure 1). It is important to note that the human analysis serves merely as a point of reference for the comparison of the results and is not an "optimal" one, though it is considered to be a musically sensible one.

4 Experiments

For each of the 9 Bach pieces we created 7 vector file versions, using each one of the different knowledge representations. These 9 pieces (numbers 1, 2, 5, 7, 8, 9, 11, 13, 15) were split into a training set (1, 7, 8, 11, 15) and a test set (2, 5, 9, 13). The algorithms were trained on the training set by using the human analysis of each piece for fitting the parameters (e.g. train the Kohonen networks or determine the threshold value for the Star Center algorithm). The trained algorithms were then applied to the test set. We got results for all combinations of algorithm and representation. The results obtained were translated into abstract motive sequences and compared to each other and to the human analyses, using the distance function described above.

5 Results

We computed the abstract motive distances on the training and test set using the representations and algorithms presented above. Here we report the results on the test set, but we got very similar ones on the training set, thereby confirming the homogeneity of the Bach two-part inventions. We first varied the algorithms / representations along each dimension by

- fixing the algorithm and computing the mean of all representations
- fixing the representation and computing the mean of all algorithms
- fixing algorithm / representation and computing the distance to the human analysis

Table 1: Mean distances (of all representations) between algorithms and human analyses on the test set

	Star Center	GNG	Ward	Kohonen-Ward
Human	0.200	0.111	0.164	**0.099**
Star Center		0.158	0.194	0.180
GNG			0.104	**0.086**
Ward				0.144

Table 1 compares the algorithms and human analyses by averaging the distances for all representations. The best algorithm is Kohonen-Ward (dist = 0.099), but the GNG results are also close to human analysis. These two algorithms produced the most similar results (dist = 0.086).

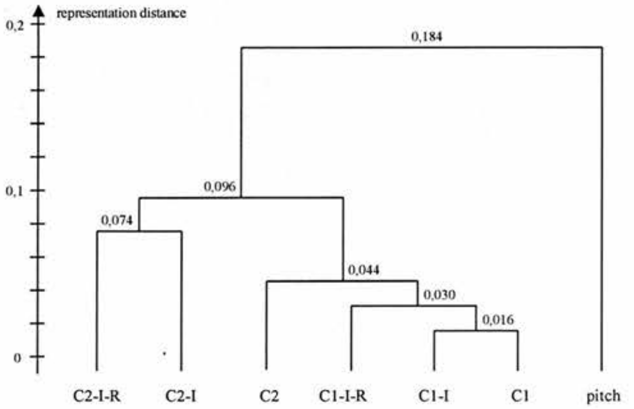

Figure 10: Dendrogram illustrating the mean distances (of all algorithms) between the representations on the test set

We then fixed the representations and clustered them (see Figure 10). The dendrogram reveals that the omission of interval and rhythm information does not notably change the distance. The pitch representation is clearly different from the other ones.

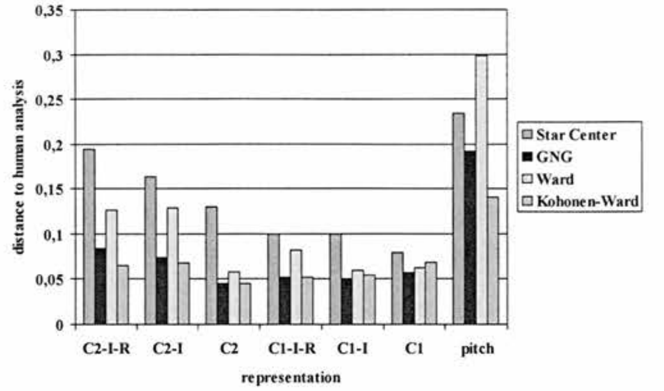

Figure 11: Mean distances to human analysis for the representations and algorithms

Finally we compared the distances to human analysis for all representations and algorithms (Figure 11) to find out the best combination. The C2 representation gives the best results except for the Star Center algorithm. C1 gives good mean results as well. Altogether the results are generally worse with additional interval and rhythm information. In particular the results are poor for the more complex C2 representation when using the Ward and the Center Star algorithm. There seems to be a scaling problem although the components of the representations had been normalized. The low-level pitch representation is inadequate as we expected. In summary the contour information clearly emerges as the most prominent feature for classifying musical structure in our specific experimental setting. The more sophisticated C2 representation proved to be general enough for the representation of all the inventions of the test and training set.

However, one can see that the algorithms have influence on the outcome as well: for our task, Kohonen-Ward and GNG are better than the other algorithms. The Kohonen network which transforms the data into two-dimensional vectors proves to be a suitable preprocessing method for the Ward algorithm. The Star Center algorithm in its current form seems to be not that appropriate for the classification task.

When using Kohonen-Ward algorithm and C2 representation on invention number 13 we obtain an example result of a good performance. The human analysis gave us the following result (abstract motive sequence):

```
abccbbdefgccfgccfcfcfcheabccbbdd
fgccfgccfgcfcfcfighejgfhjgfhjgfh
jgfhabddjgfijgfijgfiddabbbbkkgffjjfh
```

And the Kohonen-Ward algorithm produced:

```
abccbbdefgccfgccfcfcfcheabccbbdd
fgccfgccfgcfcfcfighejgfhjgfhjgfh
jgfhabddkgfikgfikgfiddabbbbcahffkkfh
```

The distance between them is 0.013.

6 Conclusions and Further Work

We have presented a formal model which allows an objective evaluation of various approaches to music classification in a uniform setting. We have shown that the algorithmic results are very close to human analysis, if an appropriate representation has been found. In particular, a tree-based contour representation proves to be appropriate in representing motivic properties typical of two-part Bach inventions. The best results were obtained by the Kohonen-Ward algorithm which combines neural preprocessing and hierarchical agglomerative clustering.

The results of our study raise a number of interesting issues concerning the relation between representations, algorithms and classification results. Further experiments need to be carried out, involving various knowledge representations and other algorithms that have been used in classification tasks; more specifically, our future main directions include

- exploring the properties of various representations and clustering algorithms in more detail, and investigating the impact of varying segmentations.
- working towards a *toolbox for music analysts*, including various representations and algorithms that could be suited to different kinds of music analysis and different pieces. For example, for a style recognition problem a more general representation like C1-I-R might be more appropriate. For a specific piece analysis, a more informed representation like C2-I-R might give better results.
- to work towards a *cognitive model of classification* in music. This will involve investigating essential features for human classification. This will be done as follows: instead of getting one human analysis, to get an "average" human PA, build a model to reproduce the results, and then see what features are important for such a task. At a second step, use these features to make predictions for classifications of new pieces.

Acknowledgements

Many thanks to Alan Smaill, Volker Steuber and Gert Westermann.

References

C. Anagnostopoulou and G. Westermann. Classification in music: A formal model for paradigmatic anlaysis. In *Proceedings of the ICMC*, 1997.

E. Cambouropoulos and A. Smaill. Similarity and categorisation inextricably bound together: The unscramble machine learning algorithm. In *Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation*, 1997.

B. Fritzke. A growing neural gas algorithm learns topologies. In G. Tesauro, D. S. Touretzky, and T. K. Lean, editors, *Advances in Neural Information Processing Systems 7*. MIT Press, 1995.

R.O. Gjerdingen. Categorization of musical patterns by self-organising neuronlike networks. *Music Perception*, 7(4), 1990.

D. Gusfield. Algorithms on strings, trees, and sequences. In *Computer Science and Computational Biology*. Cambridge University Press, 1997.

D. Hörnel. A multi-scale neural-network model for learning and reproducing chorale variations. *Melodic Similarity - Concepts, Procedures, and Applications. Computing in Musicology*, 11, 1998.

T. Kohonen. The self-organising map. *Proceedings of the IEEE*, 78(9), 1990.

J.J. Nattiez. *Fondements d'une Sémiologie de la Musique*. Union Générale d'Editions, 1975.

P.Y. Rolland. Flexpat: A novel algorithm for musical pattern discovery. In *Proceedings of the XII Colloquium on Musical Informatics, AIMI Association*, 1998.

N. Ruwet. Méthodes d'analyse en musicologie. *Revue Belge de Musicologie*, 20, 1966.

J.H. Ward. Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, 58, 1963.

Similarity and Categorisation in Boulez' *Parenthèse* from the Third Piano Sonata: A Formal Analysis.

Christina Anagnostopoulou [⋆] and Alan Smaill [†] [*]

[⋆] Faculty of Music, University of Edinburgh

[†] Division of Informatics, University of Edinburgh

chrisa@music.ed.ac.uk, smaill@dai.ed.ac.uk

1 Introduction

Categorisation is the process of detecting structures and similarities between the objects in the world, and grouping similar objects together into classes. This process lies at the basis of most human cognitive activities. Equally, similarity and difference relations play a fundamental role in the internal structure of a musical piece, and in our musical understanding (e.g., [Deliège, 96]). Many theories and analytical methods in music, such as traditional morphological analysis, paradigmatic analysis, pitch class set theory, motivic analysis and so forth, are based on similarity relations.

A problem with a categorisation-based approach to music analysis is that often the categories in musical pieces are chosen intuitively, making it difficult to justify the choice of a specific class for a musical segment, and introducing inconsistencies into the analysis. In this paper we address this problem by presenting a formal approach to categorisation which is based on a clustering algorithm that operates on well-defined descriptions of musical segments, and we apply this approach to the analysis of a musical piece, namely, Boulez' *Parenthèse*, a movement from his 3rd piano sonata.

The rest of this paper is organized as follows: in the next section, we briefly describe Boulez' 3rd piano sonata and *Parenthèse*, and we discuss the challenges that this piece poses to the analyst. Then, we explain in detail our formal approach

*Many thanks to Gert Westermann and Fred Howell.

to categorisation, including segmentation of the piece, representing the segments in terms of musical features, and clustering these representations with a computational algorithm. Section 4 describes the categorisation experiments that were carried out, and the results of these experiments are presented in section 5. In section 6 we discuss these results and suggest directions of future research.

2 Boulez' 3^{rd} Piano Sonata

Pierre Boulez' 3^{rd} piano sonata is based on difference as much as on similarity. There are various strong relationships between the movements that need not concern us here, where we aim to study the low-level relationships of a single movement. According to Ivanka Stoianova [Stoianova, 1978], 'repetition is vital, although it is "a repetition-difference" within the circumstances of the serial writing [...] In reality, it is a different kind of repetition, which is the principal generator of dodecaphony and serialism.'

Parenthèse consists of 6 fragments of music that are obligatory to play, and in between them are 5 fragments of music in parentheses, which are optional to play. According to Stoianova [Stoianova, 1978, p.140], *Parenthèse* is the

> "microcosm of the symmetrical structure of the whole sonata. The presence of the obligatory and optional (in parentheses) fragments implies the co-existence of two symmetric structures: a circular symmetry of the obligatory groups and another similar one of the groups in parentheses."

In order to capture this structure, the analysis of the entire piece can be split into three parts: first, the analysis of the six obligatory fragments, second, the analysis of the optional fragments in parentheses, and third, the relation between obligatory and optional fragments. In this paper, we demonstrate a full analysis of the first part, that is, the obligatory fragments of the piece.

Within *Parenthèse* we observe different similarity relations between its segments: first, the dodecaphonic "repetition-difference", which is based on the use of pitch class sets, and second, similarity relations in musical properties such as rhythm and tempo, tonal centres, intervals, contour, and way of playing.

The method of analysis that we present in this paper aims to bring out these relations. The aim is, on the one hand, piece-specific: to demonstrate the structure of the obligatory part of the piece. On the other hand, a more general aim is to demonstrate how the formal method of analysis, that has previously been shown to work for monophonic pieces ([Anagnostopoulou and Westermann, 97]) can be applied to a non-monophonic, atonal piece of music with very rich internal relations, and where a hierarchical segmentation is needed.

3 The Analysis

The analysis method is a formalised and extended version of Paradigmatic Analysis [Ruwet, 1996; Nattiez, 1975]. The formalisation consists in dividing the analysis process into discrete steps, fully specifiying the representations at each step, and performing the clustering of the musical segments with a well-defined algorithm.

The analysis process is illustrated in figure 1. First the piece[1] is broken down hierarchically into smaller segments, and then each of the segments is described as a set of properties. The description of the segments is then turned into an appropriate computational input in the form of feature vectors, and the classification algorithm takes this input and produces a hierarchic classification of the segments. The result of this process is a categorisation analysis that makes similarity relations explicit. In the following sections, we describe each step in detail.

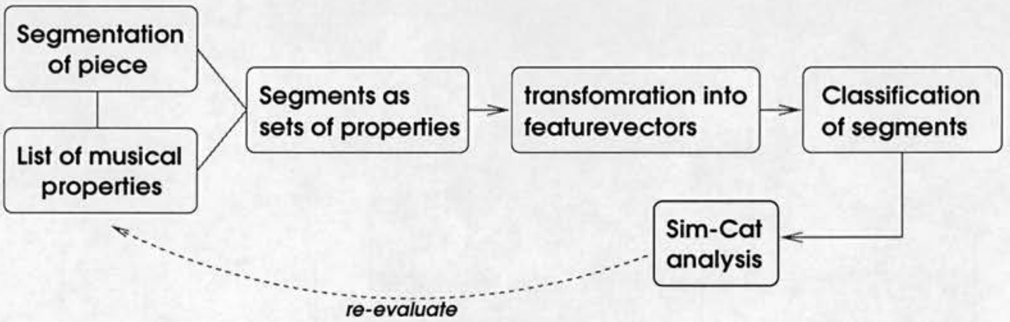

Figure 1: A general overview of the Similarity and Categorisation Method of Analysis.

3.1 Segmentation

In most formal methods of analysis, the music piece is first split into segments. It is important to consider that the precise way in which the piece is segmented has a profound influence on the outcome of the analysis.

[1]By using the term *piece* we mean the obligatory fragments that are analysed here.

In *Parenthèse*, segmentation is an easier task than for most pieces, since in most places the segmentation points are clearly indicated by the composer. We define segment boundaries

- at the beginning and end of the fragments in parentheses
- where the piano stave is marked with *V* as a break point
- where there is a more or less obvious change of texture, that is, between segments 2a and 2b, and 4c and 4d. This also coincides with the change of a pitch-class set, and this segmentation therefore corresponds to the so-called imbrication method of segmentation [Forte, 73].

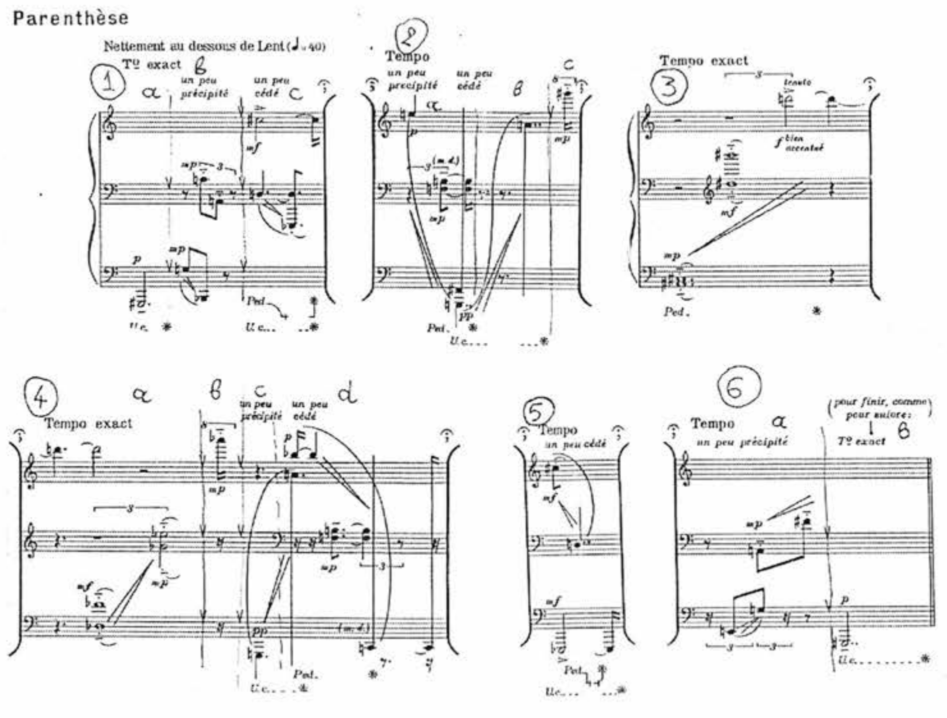

Figure 2: The obligatory fragments of *Parenthèse* and their sub-segments.

The resulting segmentation of the piece is shown in figure 2. We denote the obligatory fragments with numbers 1,..., 6. These fragments are then further divided into segments 1a, 1b, 1c, 2a, 2b, 2c and so on.

In the following experiments, we use three levels of segments: the undivided high-level segments 1,..., 6, the low level segments 1a, 1b, and so on, and an intermediate level where we combine certain adjacent low-level segments: for example, the low level segments 1a and 1b form the intermediate level segment 1ab. By this we hope to capture similarities that exist between the different segmentation levels.

3.2 Description of Segments as Sets of Properties

The term *property* is often used interchangeably with *feature* and *attribute*, and here we use it in the same way. A property value can denote the presence or absence of a feature in a segment (e.g., crescendo), or it denotes an attribute that can take one of several (mutually exclusive) values, e.g., key. In order to translate such a multi-valued property into a binary form which is required for the classification algorithm, we use a 1-out-ot-n encoding: out of the n possible values, the one that is present is set to 1, and all others to 0.

Like the choice of segmentation, the choice of properties to describe the segments has a profound influence on the results of the computational classification: the algorithm groups the segments according to their similarity, and this similarity is determined by the property values for each segment. What makes two music segments identical, similar or different will obviously depend crucially on the property selection, and on the way of representing the properties. Whereas the choice of properties is made by the analyst, the formal method of analysis shows precisely how this choice influences the resulting analysis.

In developing a set of properties, a segment is analysed in terms of various musical properties that seem important for its description and for its differentiation to other segments. Then, all properties that have been chosen for the segments are combined into one set, and each segment is described in terms of this set of properties.

The description of a segment by a list of properties is not complete: it is impossible, based on the properties, to re-create the particular segment they describe. Instead, the proeprties contain all information about a segment that are considered important for the further analysis of the piece. Different analyses warrant different properties: for example, a rhythmic analysis would describe the rhythmic properties of each segment in detail, and an analysis aiming to compare certain features across a wider music repertoire would emphasize those features.

Two kinds of properties can be used for describing a musical segment:

- properties that are true for a part of the segment, for example, the existence of a specific interval in the segment

- properties that are true for the whole of the segment, for example a rising melodic movement

In our approach we mainly make use of the second kind of properties, with the exception of specific rhythmic and intervallic patterns that describe merely part of a segment.

Table 1 shows the properties that we use in the analysis, and the segments in which they are found. The properties considered here are:

- the existence of various pitch-class sets and certain common subsets that they share. The reason to consider the common subsets is to reinforce similarity between the sets that the composer uses, which are indeed very similar to each other. In order for a pitch-class set to be true for a segment, *all* the notes of the segment have to belong to the pitch class set.

- the existence of various rhythmic patterns. These in contrast do not require for all the notes of the segment to belong to the specific rhythmic pattern.

- tempo and dynamic descriptions. The composer is very specific about which tempo and dynamic descriptions he uses, and these are important for the distinction of the segments and the overall structure of the piece, so in a classification task of this piece, they should be taken into account.

- tonal centres, which in this case are single tones rather than keys, and relations between tones, significant intervals that the composer seems to favour, and contour information.

Table 1 shows how each segment is "translated" from musical notation to a set of properties. The reason for this transformation is to achieve, at a next step, a consistent classification. For this reason we need to have the criteria set forth before the classification takes place.

Describing the segments in terms of properties (*cf.* table 1) results in a 34-bit feature vector for each segment, making it thus appropriate computational input for the classification module.

3.3 Classification

The classification of the segments that are represented as feature vectors, is carried out with a computational algorithm. This approach differs from the traditional Ruwet/Nattiez Paradigmatic Analysis in that

property	1a	1b	1c	2a	2b	2c	3	4a	4b	4c	4d	5	6a	6b
3-1(12)			y			y						y		
4-1(12)	y-	-y-	-y											
7-2				y-	-y-	-y			y-	-y-	-y			
6-9				y-	-y			y-	-y	y-	-y			
5-2	y-	-y					y	y					y-	-y
5-5				y							y			
all								y-	-y-	-y-	-y			
inv 012	y-	-y	y	y	-y-	y	y	y	-y	y-	y	y	y-	-y
inv +3	y-	-y	-y	y	-y-	-y	y	y	-y	y-	y		y-	-y
inv +5	y-	-y		y-	-y-	-y	y	y	-y	y-	-y		y-	-y
inv +7				y	-y-	-y		y-	-y	y-	y			
longn	y		y				y	y				y		y
Q,Q		y											y	
4note		y											y	
SQdot						y			y					
triplex		y		y							y		y	
exact	y						y	y	y					y
précip		y		y						y			y	
cédé			y		y	y					y	y		
mf+			y				y	y				y		
cresc	y-	-y-	-y				y					y	y	
dimin			y					y					y-	-y
steady	y	y		y?-	-y-	-y			y	y	y			
Gis/Aes	y					y			y					y
G,Gis,A	y-	-y			y-	-y			y-	-y-	-y		y-	-y
D							y	y				y		
Cis,D,Dis			y				y	y				y		
semit		y	y	y	y-	-y	y	y			y	y	y	
tritone		y		y			y				y			
third		y		y			y	y					y	
wob		y											y	
down1			y					y				y		
down2				y							y			
up2							y							

Table 1: The lowest-level obligatory segments (1a, ..., 6b) and the properties that are true for each segment. When a property exists in a segment, then this is marked by a "y". When there is a property that is true for a bigger segment but not for the low-est level, then this is marked in the lowest-level segments that the bigger segment is made from, by using "y-", "-y", "-y-", according to which adjecent the property is shared with. The first part of the table contains the pitch-class sets and their common subsets, the second part contains the rhythmic patterns, the third part contains the directions by the composer on tempo and way of playing, the fourth part contains tonal centres and specific intervals and the last part contains contour information.

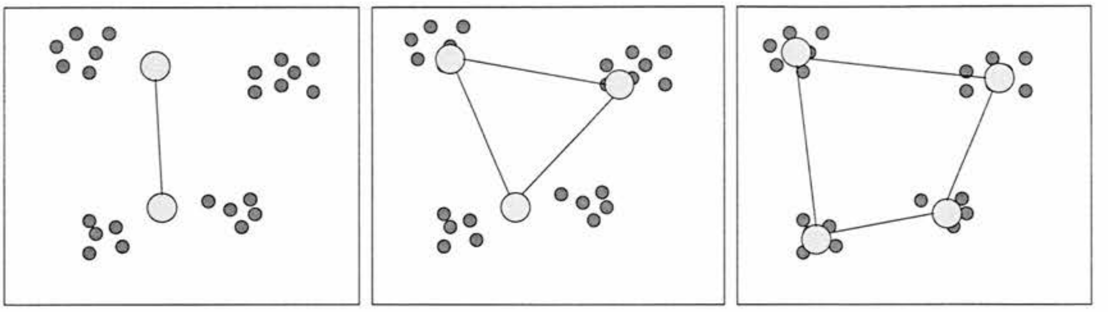

Figure 3: Construction of a GNG network. Small circles represent input data, and large circles connected with edges are the units of the network.

- the classification process is formalised and depends on explicit criteria, that is, the set of properties, thus avoiding intuitional and unfounded decisions.
- the classification proceeds in an approximately hierarchic way, from the whole piece being considered as one class to each segment being considered as a separate class.
- the algorithm develops probabilistic prototypical values of the class properties, directly showing similarities and differences between the classes.

The algorithm used here is an unsupervised neural network clustering algorithm, *Growing Neural Gas* (GNG) [Fritzke, 1995], which has been used previously for musical analysis of different musical styles and has been shown to produce valid results ([Höthker, Hörnel and Anagnostopoulou, 2000; Anagnostopoulou and Westermann, 97]).

The GNG consists of units that move towards the center of the classes, and during the learning process it adds units at a constant interval, effectively increasing the number of classes in the analysis. Each segment belongs to the class of its closest unit.

Figure 3 shows the development of a network in a two-dimensional input space with four distinct clusters. The network starts with two units and can therefore distinguish only between the two main clusters. After a certain number of presentations of the feature vectors (here 500 presentations of each vector), a new unit is inserted and the units move to the positions indicated in the second picture. When the fourth unit is inserted, the input units distribute over the four clusters.

In principal, insertion of units proceeds forever. The GNG algorithm thus lets the analyst define the level of grainedness of her analysis and does not impose *a priori* constraints on the number of clusters. Each unit forms a prototype of a cluster, expressed in the probability distribution of the feature values of their cluster members, and the distances between the units can be measured to gain information about the similarity between the classes.

4 Experiments

We performed four experiments:

In the first experiment, the classification algorithm was trained on the feature vectors that represent the segments on the smallest level only: 1a, 1b, 1c, 2a, 2b, 2c, 3, 4a, 4b, 4c, 4d, 5, 6a, 6b. The properties that stretch over adjacent smallest-level segments were not taken into account.

In the second experiment, the algorithm was again trained on feature vectors representing the smallest-level segments, but this time they were enhanced with those features that stretch over segment boundaries. For example, if segment 1 has a property *a* that is not reflected in its sub-segments 1a, 1b, and 1c, then here these sub-segments inherited this global feature.

In the third experiment, all segmentation levels were represented in parallel and the algorithm was trained on the full set of lowest-level segments 1a,. . . ,6b, the highest level segments 1,. . . ,6, and certain middle-level segments such as 1ab, 4bcd, and so on. In contrast to experiment 2, the lowest-level segments were only represented by their own properties and not the shared ones.

In the fourth experiment, we considered only a selection of eight segments drawn from all the levels: 1ab, 1c, 2, 3, 4a, 4bcd, 5 and 6.

By comparing the developing network architecture over a period of insertion of units, we were able to observe the hierarchy of classes.

5 Results

Table 2 shows the results of experiments 1 and 2, when the number of classes is 5 (that is, when the network has inserted 5 units). Table 3 shows the results in the same two experiments, when there are 7 and 8 classes.

The results of experiments 1 and 2 are all intuitively acceptable, although those from experiment 2 seem slightly better. Table 2 shows the results of experiment 1 with 5 classes: here, 1a, 2b, 2c, 4b, 4c, 6b belong to the same class. This classification would have been better if segments 1a and 6b were in a different class from the others, since they contain long notes whereas the other segments contain shorter notes. This difference could be enhanced by introducing an extra

Class	Exp 1	Exp 2
Class I	2a, 4d	2a, 2b, 2c, 4b, 4c, 4d
Class II	3, 4a	1c, 5a
Class III	1a, 2b, 2c, 4b, 4c, 6b	1b, 6a
Class IV	1b, 6a	3, 4a
Class V	1c, 5a	1a, 6b

Table 2: The experimental results in the two first experiemtns when the number of classes is 5.

Class	Exp 1	Exp 1	Exp 2
Class I	2a, 4d	2a, 4d	2c, 4b
Class II	3, 4a	3, 4a	1c, 5a
Class III	1a, 6b, 4b	1a, 6b	1b, 6a
Class IV	1b, 6a	1b, 6a	1a, 6b
Class V	1c, 5a	1c, 5a	3, 4a
Class VI	2b, 4c	2b, 4c	2a, 4d
Class VII	2c	2c	2b, 4c
Class VIII	–	4b	–

Table 3: The experimental results in the 2 experiments when the number of classes is 7 or 8.

feature *note-length* in the list of properties describing the segments. This is an example of re-evaluation of the segment descriptions.

Table 3 shows the classification for experiment 1 with 7 and 8 classes. Here we see that the same segments are separated into three classes when the overall number of classes is 7. Therefore a bigger number of classes produces more satisfactory results in this case.

Whereas experiment 1, which does not incorporate properties that stretch over segment boundaries, emphasised the iconic similarity between segments, in experiment 2 the structural similarity between segments is enhanced due to the added "global" features relating to higher-level segments. Here, all the subsegments of segments 2 and 4 are in the same class. Even though the iconic similarity of these segments is low (e.g., 4b and 4d), they both share the global properties of segment 4.

Figure 4 shows the progression of the classification in experiment 2 from 2 to 10 classes. This is an intuitively successful example of hierarchic classification, which shows the symmetrical structure of the piece.

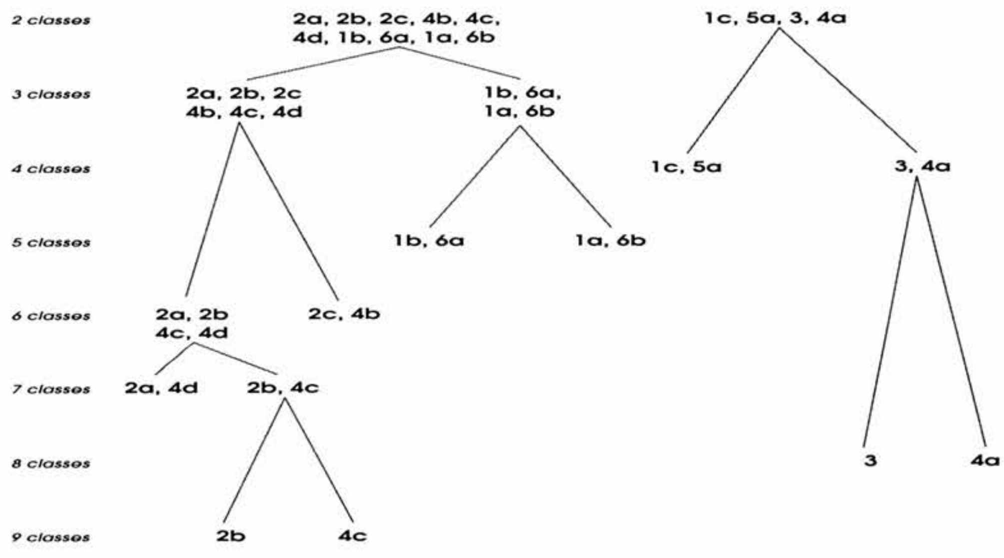

Figure 4: Hierarchic classification for set 2.

In experiment 3 all levels of segments are taken into account. The results for 5 and 8 classes are shown in table 4. In this case we often get segments and their subsegments classified in the same category, since they share many of their properties (for example, segments and subsegments of 2 and 4). This problem cannot be avoided in such a setting and the results need further interpretation in order to be valid. For this reason, 5 classes seems to be too few classes for an acceptable classification. When the number of classes increases to 8, the results improve: 3 and 4a are correctly classified into a category of their own, and the same holds for 1b and 6a. It is interesting to see segment 4 on a category of its own, since it is the longest segment of all. Segments 2 and 4bcd are placed in the same category and are an example of similarity across levels. In general, 8 classes seem to be sufficient for demonstrating the symmetry of the segments, although one needs to consider carefully which segments denote this and which are merely related subsegments of the same bigger segment.

Experiment 4 is the simplest experiment because we consider only a selection of 8 segments across levels. These are chosen in order to show the structure of the piece. Table 5 shows the resulting classification when having 4 classes: the first and last segment, 1ab and 6, are classified together, and the same is true for 1c and 5, 2 and 4bcd and 3 and 4a. These segments are almost mirror images of each other, and define the symmetrical structure of the piece.

Class	Exp 3, classes 5	Exp 3, classes 8
Class I	1c, 3, 4a, 4, 4ab	1c, 5a, 4ab
Class II	1a, 2b, 2c, 4b, 4c, 6b, 2bc	1a, 2b, 2c, 4b, 4c, 6b
Class III	2a, 4d, 2, 2ab, 4cd, 4bcd	1, 6, 1ab, 1bc
Class IV	1b, 6a, 1, 6, 1ab	2a, 4d, 2, 2ab, 4cd, 4bcd
Class V	1bc	4
Class VI		1b, 6a
Class VII		2bc
Class VIII		3, 4a

Table 4: The experimental results in the third experiment when the number of classes is 5 and 8.

Class	Experiment 4
Class I	1c, 5
Class II	2, 4bcd
Class III	1ab, 6
Class IV	3, 4a

Table 5: The experimental results in experiment 4, with 4 classes.

6 Conclusions

We presented a formal method of analysis based on categorisation of music segments according to similarity. We applied this method to the analysis of Boulez' *Parenthèse* from the 3$^{\text{rd}}$ piano sonata, taking into account the obligatory fragments of the piece. The resulting hierarchic classification defines the similarity and difference relations between classes and between segments. We demonstrated how a classification analysis is appropriate for this piece and how it brings out the symmetrical structure that the composer intended. This method of analysis, shown previously to work on more traditional kinds of music, is shown here to be appropriate for an atonal and non-monophonic piece of music.

The results give many interesting insights on the obligatory fragments. In terms of internal relations, it is a very rich piece, each note situated in its position for a variety of reasons, forming part of an overall plan. More specifically, we see that the piece also has an interesting tonal structure, evolving mainly around G sharp at the beginning and end, and around D in the middle of the piece. The pitch class sets used are very similar to each other, segments 2 and 4 sharing sets, and the same for segments 1, 3 and 6. Dynamics and tempo seem to be very important for the segmentation and difference between subsegments, whereas contour information seems to be reflecting the symmetrical structure of the piece.

The issue of hierarchic segmentation in a classification task poses interesting challenges to the analyst. When classifying all the levels at the same time, on the one hand we get interesting similarities across levels, but on the other hand we get similarities between segments and their subsegments which are redundant.

The results depend on the initial representation, that is the choice of properties according on which each segment is described. A different choice of properties would yield different results. However, a bad resulting classification would show that the initial properties were not chosen carefully, and a re-evaluation of these properties is needed. In that way, the analyst can revise the initial properties. This procedure can go on until an acceptable classification is produced.

The principles of similarity and difference are common principles to the vast majority of musical repertoire. It can be argued that they are responsible to a large extent for cohesion and coherence within the musical piece.

7 Bibliography

Anagnostopoulou, C. and Westermann, G, 1997, *Classification in Music: A Formal Model for Paradigmatic Analysis*, Proceedings of the International Computer Music Conference, Thessaloniki, Greece.

Deliège, I., 1996, Cue Abstraction as a Component of Categorisation Processes in Music Listening, in Psychology of Music, 20(2).

Forte, A., 1973, *The Structure of Atonal Music*, Yale University Press, New Haven and London.

Fritzke, B., 1995, *A Growing Neural Gas Algorithm Learns Topologies*, in G. Tesauro and D. S. Touretzky and T. K. Lean (eds), Advances in Neural Information Processing Systems 7, MIT Press.

Höthker, K., Hörnel, D. and Anagnostopoulou, C., 2000, *Investigating the Infuence of Representations and Algorithms in Music Classification*, in Computers and the Humanities, 34(4), in press.

Nattiez, J.J., 1975, *Fondements d'une Sémiologie de la Musique*, Union Générale d'Editions, Paris.

Stoianova, I., 1978, *Geste-Texte-Musique*, Union Générale d'Editions, Paris.

Ruwet, N., 1966, *Méthodes D'Analyse en Musicologie* in Revue Belge de Musicologie 20.

Representation and Discovery of Multiple Viewpoint Patterns*

Darrell Conklin
Computational Biology
ZymoGenetics, Inc.
Seattle, USA
conklin@zgi.com

Christina Anagnostopoulou
Faculty of Music
University of Edinburgh
Edinburgh, Scotland
chrisa@dai.ed.ac.uk

Abstract

An important problem in computational music analysis is the representation and automated discovery of recurrent patterns. In this paper we present a new method for pattern representation and discovery in a large corpus of music. Using the formalism of multiple viewpoints, music is viewed as multiple streams of description derived from the basic surface representation. Patterns are discovered within viewpoint sequences derived from the corpus for selected viewpoints. A statistical method is used to restrict attention to only those patterns which occur much more frequently than expected, where expectation is based on a Markov model of viewpoint elements. The concept of the longest significant patterns in a corpus is introduced. The method presented in this paper is designed to rapidly enumerate all longest significant patterns within a large corpus. An application of the method to the Bach chorales is presented.

1 Introduction

The low entropy of music is due to the inherent structural constraints in a musical style, and repetition of both *intra-* and *inter-* opus musical material. An important problem in computational music analysis is the representation and automated discovery of recurrent musical patterns. Patterns can be used for abstraction and compact representation of a work (Smaill et al. 1993); as musical building blocks for the paradigmatic analysis of a work (Nattiez 1975); as fragments for motive-based algorithmic composition (Rolland and Ganascia 2000); as keys for content-based music retrieval (Hsu et al. 1998); and for the recognition and distinction of musical style or authorship (Westhead and Smaill 1993).

Repetition in music occurs not only as repetition of exact pitches and durations, or as mere transposition into a different key, but is often much more subtle. There can be repetitions within different musical parameters, such as specific intervals, melodic motion (contour), relative rhythmic values, middle or fundamental structure, harmonic progressions (implicit or explicit), register, dynamics series, pitch class sets, and so on. Approaches to pattern discovery in music analysis have so far concentrated on the similarity relationships within pitch or transposed passages rather than on recurrent patterns within deeper musical parameters. However, in a music analysis task of any kind, it would make more sense to be able to capture these recurrent patterns: they are more general, look at a deeper level of similarity within the musical corpus and make explicit exactly where the similarities between the patterns lie. Pattern discovery algorithms should be able to handle the comparison of a small set of pieces as well as the processing of a large corpus with hundreds or thousands of pieces. Finally, such methods should have some selection procedures for the results, removing uninteresting patterns from consideration.

The topic of this paper is the discovery of general patterns that span a substantial number of diverse works. Our pattern discovery technique is based on the music representation formalism of multiple viewpoints (Conklin and Witten 1995), presented in this paper, where each viewpoint models some musical parameter. In this way we are able to search for patterns within these parameters, or viewpoints, rather than patterns of notes taken directly from the music, and we can capture the exact level where similarity occurs in the music. Pattern discovery is performed by building a suffix tree data structure with all multiple viewpoint sequences derived from the corpus for chosen viewpoints. A subsequent step finds those patterns that occur in a specified minimum number of pieces and that satisfy a statistical significance criterion. A further filtering looks at all significant discovered patterns and selects the *longest significant patterns* within the set. This paper presents an application of the method to the Bach chorales.

*To appear in *Proc International Computer Music Conference, Havana, 2001.*

2 Methods

2.1 Multiple viewpoints of music

The representation language we use for music is based on the formalism of multiple viewpoints (Conklin and Witten 1995). Viewpoints are functions, defined or selected by the music analyst, that operate on the basic representation. A viewpoint models some specific type of musical feature derived from the musical surface, for example, melodic contour, intervals, duration, or interval from a tonic referent pitch. A piece of music is therefore transformed into a higher level description derived from the basic surface representation. At the surface, a piece is represented as a set of sequences (tracks, voices) of events derived from a MIDI encoding: an event has a pitch, duration, and a start time. In addition to these event attributes, we assume some global attributes such as fermatas (used as phrase markers), key signature, and time signature.

In more detail, a viewpoint is a partial function which associates a *viewpoint element* with sequences. The notation $[\tau]$ denotes the range of this function; the set of valid viewpoint elements for a viewpoint τ. For example, for the melodic interval viewpoint, the viewpoint elements are integers. A viewpoint is a partial function, meaning that it may be undefined at certain locations. For example, the melodic interval viewpoint is undefined for a sequence containing only one event (see Figure 1).

For every viewpoint a *viewpoint sequence* function transforms a sequence of basic events into a sequence of defined viewpoint elements. The viewpoint sequence function simply applies the viewpoint to every element in the sequence, retaining those elements where the viewpoint is defined. For example, for the melodic contour viewpoint, this function transforms a sequence of pitches into a sequence of contour indicators (see Figure 1).

A *linked viewpoint* is a combination of two or more viewpoints that models other viewpoints simultaneously. A link between viewpoints can be defined using the constructor $\otimes$. For a linked viewpoint $\tau = \tau_1 \otimes \ldots \otimes \tau_n$, and any sequence m, $\tau(m)$ is undefined if $\tau_i(m)$ is undefined for any component viewpoint, else it is the tuple $\langle \tau_1(m), \ldots, \tau_n(m) \rangle$. The set of viewpoint elements is therefore the cross product of all sets of component viewpoint elements:

$$[\tau_1 \otimes \ldots \otimes \tau_n] = [\tau_1] \times \ldots \times [\tau_n] \tag{1}$$

For example, Figure 1 illustrates the linked viewpoint between melodic contour and duration. Note how the elements of this viewpoint are pairs of values, one from each component viewpoint.

A *threaded* viewpoint models the value of a *base viewpoint* at defined temporal or metric locations within a piece; for example, at the beginning of a bar or phrase or at every quarter note pulse. These defined locations are captured by a *test viewpoint*. A threaded viewpoint is defined only at locations where the test viewpoint is true. In this way, a viewpoint "threads" through a sequence, potentially ignoring non-adjacent surface events. Any viewpoint (even a linked viewpoint) can be used as a base viewpoint. The test viewpoint must have a Boolean (0 or 1) value. Given a base viewpoint τ and a test viewpoint θ, a threaded viewpoint can be defined using the constructor $\oslash$. The set of viewpoint elements for a threaded viewpoint is the cross product of the base viewpoint elements and the set of inter onset intervals:

$$[\tau \oslash \theta] = [\tau] \times [\texttt{ioi}] \tag{2}$$

The `ioi` viewpoint is the inter onset interval between two events; the difference between their start times.

For example, we can construct a viewpoint that measures the melodic interval between events that occur as the first event in a bar, or a viewpoint that threads through events that start on crotchet beats (see Figure 1).

2.2 Viewpoint patterns

A *viewpoint pattern* P_τ is a sequence of viewpoint elements for some viewpoint τ. A pattern *occurs* in a piece if it is contained in the viewpoint sequence for that piece (see Figure 2). The *length* of a pattern P is denoted $l(P)$. The *empty pattern* $\emptyset_\tau$ for a viewpoint τ has zero length and is defined to occur anywhere that the viewpoint is defined. Henceforth we omit the subscript from viewpoint patterns as the viewpoint should always be evident from context.

The *piece count* of a pattern is the number of pieces that a pattern occurs in. The *total count* of a pattern is its total number of occurrences, including repetitions within an individual piece.

Pattern scoring. The potential musical significance of a viewpoint pattern P with respect to a data set is evaluated by comparing the piece count of P with how many times we expect it to occur if pieces in the data set were generated from a Markov model of viewpoint elements.

Large differences between observed and expected counts indicate a potentially interesting pattern. A pattern is given a *pattern score*, which represents the magnitude of this difference:

$$\frac{(\#(P) - E(P))^2}{E(P)} \tag{3}$$

where $\#(P)$ is the observed total count and $E(P)$ is the expected total count (defined below) for the pattern P. The score for a pattern will increase with the difference between its observed and expected total count.

Expected count for a pattern. The expected total count for a pattern is the number of sites where the pattern could possibly occur multiplied by the probability of finding the pattern

Viewpoint	Sequence	Result
`st`	$\overline{e_7}$	36
`pitch`	$\overline{e_{12}}$	74
`int`	$\overline{e_1}$	undefined
`contour` $\otimes$ `dur`	$\overline{e_{12}}$	$\langle 1, 8\rangle$
`int` $\otimes$ `ioi`	$\overline{e_1}$	undefined
`int` $\otimes$ `ioi`	$\overline{e_{13}}$	$\langle -2, 8\rangle$
`st`	$\overline{e_7}$	$(8, 12, 20, 24, 30, 32, 36)$
`int` $\otimes$ `ioi`	$\overline{e_1}$	$()$
`int` $\otimes$ `ioi`	$\overline{e_7}$	$(\langle 0,4\rangle, \langle 7,8\rangle, \langle -3,4\rangle, \langle -2,6\rangle, \langle -2,2\rangle, \langle 0,4\rangle)$
`contour`	$\overline{e_{16}}$	$(0, 1, -1, -1, -1, 0, 1, 1, -1, 1, 1, -1, -1, -1, -1)$
`int` $\oslash$ `fb`	$\overline{e_{16}}$	$(\langle 4,12\rangle, \langle -4,12\rangle, \langle 2,12\rangle, \langle 5,12\rangle, \langle -3,12\rangle, \langle -4,12\rangle)$
`int` $\oslash$ `fph`	$\overline{e_{11}}$	$(\langle 4,48\rangle)$
`int` $\oslash$ `isq`	$\overline{e_7}$	$(\langle 0,4\rangle, \langle 7,8\rangle, \langle -3,4\rangle, \langle -4,8\rangle, \langle 0,4\rangle)$

Figure 1: A fragment from the chorale *Aus meines Herzens Grunde*, with some example applications of the viewpoint element function (top) and the viewpoint sequence function (bottom) for various viewpoints. Start time (`st`) is represented as semiquaver ticks from time 0, and pitch as a MIDI number. The first event in this fragment starts at tick 8. The notation $\overline{e_n}$ is an abbreviation for a sequence of events $(e_1, \ldots, e_n)$. The viewpoint `contour` refers to melodic contour, and `int` to melodic interval. The `ioi` viewpoint is the inter onset interval between two events. The test viewpoints `fb`, `fph`, and `isq`, used to construct threaded viewpoints, are true if an event is the first in a bar, first in a phrase, or on a crotchet beat, respectively.

Viewpoint	Pattern	Occurrences
`int`	$\emptyset$	$2, 3, \ldots, 16$
`int` $\otimes$ `ioi`	$\langle 0, 4\rangle$	2, 7
`dur`	$(4, 3, 2, 4)$	3, 6
`int` $\oslash$ `fb`	$\emptyset$	4, 7, 10, 12, 14, 16
`contour` $\oslash$ `fb`	$(\langle 1, 12\rangle, \langle -1, 12\rangle)$	2, 10

Figure 2: Examples of viewpoint patterns for the chorale fragment in Figure 1. Occurrences refer to the event number of the first event in the fragment where the pattern occurs. Note the use of the empty pattern $\emptyset$ which is defined anywhere that the viewpoint is defined.

in a random viewpoint sequence. These two quantities are defined here.

Consider a pattern P of length $l(P)$. In a single piece, there are $l(P)-1$ positions where the pattern cannot possibly occur, because it would extend past the end of the piece. Therefore, in a data set of n pieces, there are $n(l(P)-1)$ positions where the pattern cannot possibly occur. It follows that there are $\#(\emptyset)-n(l(P)-1)$ positions where the pattern P might occur in the data set. The expected number of occurrences $E(P)$ of a pattern P in the data set is therefore

$$E(P) = p(P) \times (\#(\emptyset) - n(l(P)-1)) \qquad (4)$$

That is, the number of times we expect to see a pattern P is the probability $p(P)$ of the pattern multiplied by the number of positions where the pattern could occur. Probabilities of viewpoint patterns are computed using a blended zero- and first-order Markov model of viewpoint elements seen in the corpus.

Statistical significance. It is useful to report a p-value for a pattern; the probability that an equal or greater pattern score could arise within random viewpoint sequences. Patterns with high p-values will frequently occur in random viewpoint sequences and therefore are not interesting.

An exponential probability distribution is used to model pattern scores. This p-value of a pattern must be adjusted to reflect that fact that we are evaluating its significance not in isolation but within all patterns found in the corpus. This is called a Bonferroni adjustment, and reflects the probability of finding a pattern of equal or greater score within all patterns tested. Given a particular pattern score, an adjusted p-value is computed by multiplying it by an adjustment factor which is simply the total number of patterns evaluated for significance.

Pattern discovery algorithm. The pattern discovery algorithm (Figure 3) employs a suffix tree data structure, which compactly stores all suffixes and substrings occurring within a data string. The algorithm proceeds as follows. First, for a viewpoint selected by the analyst, every piece is transformed into a viewpoint sequence. Then, every suffix of this viewpoint sequence is incorporated into the suffix tree. This suffix tree is scanned to produce the set of all patterns that occur within at least k pieces (we use $k = 10$ for most of our results). The size of this set is the adjustment factor used to compute pattern p-values. The statistical significance of each pattern in this set is evaluated, and insignificant patterns are discarded.

Longest significant patterns. The output from the algorithm above can include many patterns that are significant yet contained within longer significant patterns. To handle this effect, we place all significant patterns into a *subsumption taxonomy* (Woods 1991). This is a directed graph where nodes represent patterns and links represent pattern containment (subsumption). In a sense, the subsumption taxonomy can be viewed as an expansion of the suffix tree; nodes become explicit patterns rather than viewpoint elements.

Following the construction of the subsumption taxonomy, the longest significant patterns are found at the leafs (nodes subsuming no other nodes) of the data structure.

(a) For a selected viewpoint, transform all pieces in the corpus into viewpoint sequences.

(b) Incorporate the viewpoint sequence for every piece into a suffix tree.

(c) Search the suffix tree, building the set of all patterns occurring in at least k pieces.

(d) Compute a p-value for each pattern, discarding insignificant patterns.

(e) Build a subsumption taxonomy from all remaining significant patterns.

(f) Report the leafs of this structure as the longest significant patterns in the corpus.

Figure 3: The multiple viewpoint pattern discovery algorithm.

	events	
voice	total	average
soprano	9226	50
alto	11361	61
tenor	11570	63
bass	11809	64

Table 1: The event composition of 185 Bach chorales; the total number of events in a voice, and the average number in each voice in a chorale.

The chorale data set. This study uses 185 Bach chorales, comprising a total of about 40000 events (Table 1). Sections annotated by a repeat are not expanded. Even so, this data set has some redundancy, in the form of some transposed chorale melodies and transposed reuse of phrase material. For this study we do not attempt to remove this redundancy.

3 Results

About 20 viewpoints were encoded; most of them pertain to melodic and rhythmic aspects of the chorales. Several viewpoints are test viewpoints that are used mainly for linking and threading with other viewpoints. We have also encoded some viewpoints which model harmonic or vertical structures. An extended set of viewpoints and results will appear in a longer

	number of patterns			
k	total	significant	longest	
2	11305	7263	275	(16.9)
5	1285	237	81	(7.1)
10	535	96	33	(6.1)
25	174	40	15	(5.1)
50	75	23	9	(4.6)
100	28	9	6	(3.7)

Table 2: Numbers of patterns found within the soprano line of 185 Bach chorales, using the melodic interval viewpoint. The second column refers to the number of raw, unfiltered patterns occurring in at least k pieces. The third column refers to the the number of statistically significant patterns at a p-value of 0.01. The last column refers to the number of patterns remaining at the leafs of the subsumption taxonomy. The average length of the longest patterns is indicated in brackets.

paper. Here, the adjusted p-value cutoff for patterns was set to 0.01. In all experiments, unless specified otherwise, the parameter k (the minimum number of pieces a pattern must occur in) was set to 10.

In this section we present some results obtained with the pattern discovery algorithm. Our most interesting results came from the linked and the threaded viewpoints, where we identified patterns that captured deeper structure of the music. The patterns presented in Figure 4 are among the highest scoring patterns that were discovered for the particular viewpoints.

General behavior of the algorithm. Table 2 illustrates the behavior of the algorithm as a function of the parameter k (the minimum number of pieces a pattern must occur in). At lower values, the method discovers many patterns. The filtering effectiveness (from total patterns to longest significant patterns) can be as high as 98%. At $k = 2$ many long patterns are found; most of these long patterns arise from redundancy with the corpus. As k increases, the patterns found are shorter, as they are required to occur in more pieces. Even with $k = 100$, some significant patterns are found.

Melodic intervals in soprano and alto voices. For the soprano line, 33 longest significant melodic interval patterns were found. For the alto line, 29 were found. Referring to example 4(a), it is quite common in the soprano to have stepwise movement. The initial rising fourth suggests a harmonic progression of V-I, and it is likely that the I is on the strong beat.

Example 4(b) demonstrates a familiar feature of alto lines: flat melodic lines that serve mainly to fill in the harmony. The semitone movement suggests a leading note to tonic succession. A leading note is restricted: it can usually rise up to the

Figure 4: Some multiple viewpoint patterns discovered in the 185 Bach chorales. Each block illustrates an instance of a pattern, with its viewpoint, total count, piece count, voice, and pattern. An event is presented as a semibreve if its duration is not determined by the viewpoint pattern. The viewpoint `pcint` is the pitch class interval viewpoint. The dots between the crotchets in the threaded viewpoint signify that quavers or even semiquavers may occur between the indicated pitches. All viewpoint patterns are invariant under transposition.

tonic, or occasionally drop a third (in the middle voices). Here we have an example of both.

Melodic contour. A single significant melodic contour pattern, example 4(c), was discovered by our algorithm. This is a long line, spanning 12 events, that occurs within 10 pieces. It is of interest that the interval pattern of 4(a) is a specific instance of a portion of this contour pattern.

Pitch class intervals for bass. For the bass line, we used a pitch class interval viewpoint. A total of 57 longest significant pitch class interval patterns was found. Pattern 4(d) is an example of harmonic movement — the end of the segment shows a potential perfect cadence.

A well known common pattern in the bass line of the chorales is a chromatic stepwise movement, which is mentioned in most Bach chorale composition books. In example 4(e) we have found a pattern for this phenomenon. This pattern occurs 19 times, within 12 different pieces.

Linked interval and duration. For linked viewpoints, we were able to combine different parameters to see how these are related in the music. For example, for a linked viewpoint between melodic intervals and duration, we found 23 longest significant patterns. Figure 4(f) is an example of a leap in the soprano followed by rapid stepwise movement of the opposite direction, presumably to counterbalance potential singing mistakes in a congregation.

Threaded viewpoints. An example of a threaded viewpoint is one which describes the melodic interval at crotchet beats. A total of 335 patterns was discovered, and after the subsumption filter 17 remained. Figure 4(g) shows one threaded pattern. Occurrences of this pattern can include quavers or semiquavers in between the crotchets. These can be passing notes (at the last interval), consonant skips, échapées, cambiatas, and suspensions.

Threaded patterns are a step closer to the metric reduction or deeper structure of a work in the Schenkerian sense (Forte and Gilbert 1982). However, a metric reduction is more complicated than the process of the threaded viewpoint: for example, in case of a suspension, the harmonic note might not be on a crotchet beat. In that case, applying pattern discovery to a metric reduction of the score would yield better results.

4 Discussion

Pattern discovery in music is a difficult problem. Making truly new discoveries is rare, but computational techniques can contribute. This paper has presented a new formalism for describing musical patterns and a new algorithm for discovering them. The computational approach employed is to look for patterns which occur much more frequently than expected. The use of p-values for patterns can separate the truly significant patterns from statistical background. These significant patterns can be explored further by the music analyst.

There have been various approaches to automatic pattern discovery in music. Most approaches focus their analysis on a single piece for patterns (Cambouropoulos 1998; Hsu et al. 1998; Meredith et al. 2001) and are not directly applicable to the analysis of a large number of pieces. Though an artificial single piece might be constructed for these methods by joining several pieces together, since they have running times of a polynomial order in the length of the piece they may not be practical for the analysis of hundreds or thousands of pieces.

An approach that can naturally find patterns in two pieces is known as dynamic programming (Mongeau and Sankoff 1990). In this technique musical similarity is encoded into a distance function, pairs of transposed melodies are compared, and the common pattern is the trace of aligned elements. Iterations of this pairwise comparison are necessary in order to find patterns occurring within more than two sequences. By contrast, in our approach, knowledge is encoded into discrete modules, the viewpoints. Patterns are found not in a surface representation but rather in a deeper transformed representation. Rather than looking at similarity or partial similarity in the score, we shift the problem into the representation level, and look for identity. An identity at one or more viewpoints results from a similarity (of varying degree) in the music. Furthermore, since we seek identities within a transformed representation, our algorithm is computationally efficient and will find all of the patterns in a corpus.

Pattern discovery algorithms can produce voluminous output which must be filtered for both statistical and musical significance. This is usually done by preferring the longest, most frequent patterns. However, the properties of pattern length and frequency are inversely related, because longer patterns tend to occur less frequently. Balancing these two properties in a single measure is the essence of evaluating a pattern. Cambouropoulos (1998) uses a function of the three variables of pattern length, pattern frequency, and pattern overlap to rank patterns. The parameterization of the equation involving these three variables is performed manually by the investigator. Hsu et al. (1998) calibrate a minimum acceptable pattern length by running the method on synthetic random melodies.

It can be demonstrated that the pattern score of Equation 3 balances the two properties of frequency and length in a single measure, and avoids the need for pattern length threshold specification. It follows from Equation 3 that for a pattern P, if its frequency $\#(P)$ remains constant while its length increases, its expected frequency $E(P)$ will decrease while its score will increase. On the other hand, if its length remains constant while its frequency $\#(P)$ increases, its score will increase because $E(P)$ will remain constant.

In Nattiez' (1975) two paradigmatic analyses of Debussy's

Syrinx, we observe the need for longest significant patterns (first analysis) and most general patterns (second analysis). Both types of patterns are useful for further music analysis. Nattiez permits pattern overlapping in exceptional circumstances, when it is felt that both patterns are equally important and belong to different classes. In our method, pattern overlapping is allowed only when the overlapping patterns are not covered by a longer significant pattern.

An approach to musical style recognition (Westhead and Smaill 1993) and generation (Cope 1987) is to use a catalog of *signatures* that cover instances of the style. An interesting application of our method is to produce the *most general consistent patterns* occurring within a musical corpus. Consistency can be defined with respect to positive and negative examples of the style. The most general consistent patterns will be more useful than the longest significant patterns for the task of style recognition, as they are more frequent in the musical corpus and much more likely to occur in new, unseen examples of the style. For the task of style generation, general patterns are less likely to be recognizable as fragments from the pieces used to define the patterns.

Patterns are statistically significant if they are surprising with respect to a background model. Therefore, the closer the background model is to the style under consideration, the more subtle and interesting the discovered patterns will be. In this study we have used a fairly primitive Markov model as a background model. Alternatives to this are to parameterize the Markov model on another style, or on another voice within the corpus.

Although Bach chorales have traditionally been treated as exemplary harmonic sequences of a homophonic texture, our results show that voice-leading techniques are just as important as in the other works of J.S.Bach. Our model is especially suitable for teaching purposes in that it can contribute information on Bach chorale composition by the production of significant sequential patterns of the various viewpoints.

In summary, this paper has presented a new approach to pattern representation and discovery which is particularly well suited to various music analysis purposes. Based on the multiple viewpoint formalism, it produces explicit viewpoint patterns rather than similarity judgments of note patterns. Music is transformed into viewpoint sequences. An efficient suffix tree data structure is used to rapidly discover all patterns. A statistical method is used to restrict attention to only those patterns which occur much more frequently than expected. The significant patterns are organized into a subsumption taxonomy, and the longest significant patterns in a corpus are found at the leafs of the structure. The method presented here can be used to rapidly enumerate all patterns within a large corpus.

Future work will include application of the pattern discovery method to harmonic aspects of music, and a more extensive analysis of patterns discovered for multiple viewpoints. The interactions between melodic and vertical viewpoints will be used to provide further interesting insights to the corpus of the Bach chorales.

References

Cambouropoulos, E. (1998). *Towards a General Computational Theory of Musical Structure*. Ph. D. thesis, University of Edinburgh.

Conklin, D. and I. Witten (1995). Multiple viewpoint systems for music prediction. *International Journal of New Music Research 24*(1), 51–73.

Cope, D. (1987). An expert system for computer–assisted composition. *Computer Music Journal 11*(4), 30–46.

Forte, A. and S. Gilbert (1982). *Introduction to Schenkerian Analysis*. W.W. Norton.

Hsu, J.-L., C.-C. Liu, and A. Chen (1998). Efficient repeating pattern finding in music databases. In *Proc Int Conf on Information and Knowledge Management*, pp. 281–288. Association of Computing Machinery.

Meredith, D., G. Wiggins, and K. Lemström (2001). Pattern induction and matching in polyphonic music and other multidimensional datasets. In *Proc SCI 2001: Conference on Systemics,Cybernetics and Informatics*.

Mongeau, M. and D. Sankoff (1990). Comparison of musical sequences. *Computers and the Humanities 24*, 161–175.

Nattiez, J.-J. (1975). *Fondements d'une Sémiologie de la Musique*. Union Générale d'Editions.

Rolland, P.-Y. and J.-G. Ganascia (2000). Musical pattern extraction and similarity assessment. In E. Miranda (Ed.), *Readings in Music and Artificial Intelligence*, Chapter 7, pp. 115–144. Harwood Academic Publishers.

Smaill, A., G. Wiggins, and M. Harris (1993). Hierarchical music representation for composition and analysis. *Computers and the Humanties 93*, 7–17.

Westhead, M. and A. Smaill (1993). Automatic characterisation of musical style. In *Music Education: an AI approach*, pp. 157–170. Springer-Verlag.

Woods, W. (1991). Understanding subsumption and taxonomy: A framework for progress. In J. F. Sowa (Ed.), *Principles of Semantic Networks*, pp. 45–94. Morgan–Kaufmann.