Contemporary Music Studies

A series of books edited by Peter Nelson and Nigel Osborne, University of Edinburgh, UK

Volume 1
Charles Koechlin (1867–1950): His Life and Works
Robert Orledge

Volume 2
Pierre Boulez – A World of Harmony
Lev Koblyakov

Volume 3
Bruno Maderna
Raymond Fearn

Volume 4
What's the Matter with Today's Experimental Music? Organized Sound Too Rarely Heard
Leigh Landy

Volume 5
Linguistics and Semiotics in Music
Raymond Monelle

Volume 6
Music, Myth and Nature, or The Dolphins of Arion
François-Bernard Mâche

Volume 7
The Tone Clock
Peter Schat

Volume 8
Edison Denisov
Yuri Kholopov and Valeria Tsenova

Volume 9
Hanns Eisler: A Miscellany
Edited by David Blake

Volume 10
Brian Ferneyhough: Collected Writings
Edited by James Boros and Richard Toop

Please see the back of this book for other titles in the Contemporary Music Studies series

Readings in Music and Artificial Intelligence

edited by

Eduardo Reck Miranda
Sony Computer Science Laboratory
Paris, France

harwood academic publishers
Australia • Canada • France • Germany • India
Japan • Luxembourg • Malaysia • The Netherlands
Russia • Singapore • Switzerland

Contemporary Music Studies

A series of books edited by Peter Nelson and Nigel Osborne, University of Edinburgh, UK

Volume 1
Charles Koechlin (1867–1950): His Life and Works
Robert Orledge

Volume 2
Pierre Boulez – A World of Harmony
Lev Koblyakov

Volume 3
Bruno Maderna
Raymond Fearn

Volume 4
What's the Matter with Today's Experimental Music? Organized Sound Too Rarely Heard
Leigh Landy

Volume 5
Linguistics and Semiotics in Music
Raymond Monelle

Volume 6
Music, Myth and Nature, or The Dolphins of Arion
François-Bernard Mâche

Volume 7
The Tone Clock
Peter Schat

Volume 8
Edison Denisov
Yuri Kholopov and Valeria Tsenova

Volume 9
Hanns Eisler: A Miscellany
Edited by David Blake

Volume 10
Brian Ferneyhough: Collected Writings
Edited by James Boros and Richard Toop

Please see the back of this book for other titles in the Contemporary Music Studies series

Readings in Music and Artificial Intelligence

edited by

Eduardo Reck Miranda
Sony Computer Science Laboratory
Paris, France

harwood academic publishers
Australia • Canada • France • Germany • India
Japan • Luxembourg • Malaysia • The Netherlands
Russia • Singapore • Switzerland

Copyright © 2000 by OPA (Overseas Publishers Association) N.V. Published by license under the Harwood Academic Publishers imprint, part of The Gordon and Breach Publishing Group.

All rights reserved.

No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, without permission in writing from the publisher. Printed in Singapore.

Amsteldijk 166
1st Floor
1079 LH Amsterdam
The Netherlands

British Library Cataloguing in Publication Data

Readings in music and artificial intelligence. –
(Contemporary music studies; v. 20)
1. Music – Data processing 2. Artificial intelligence – Musical applications
I. Miranda, Eduardo Reck, 1963 –
780.2'85'63

ISBN: 90-5755-094-6

CONTENTS

Introduction to the Series vii

Preface ix

1 Regarding Music, Machines, Intelligence and the Brain: An Introduction to Music and AI
Eduardo Reck Miranda 1

2 Music, Intelligence and Artificiality
Alan Marsden 15

3 Musical Knowledge: What can Artificial Intelligence bring to the Musician?
Geraint Wiggins, Alan Smaill 29

4 Symbolic AI versus Connectionism in Music Research
Petri Toiviainen 47

5 On the Potential of Machine Learning for Music Research
Gerhard Widmer 69

6 Computer Analysis of Jazz Chord Sequences: Is *Solar* a Blues?
François Pachet 85

7 Musical Pattern Extraction and Similarity Assessment
Pierre-Yves Rolland, Jean-Gabriel Ganascia 115

8 Interactive Music Systems in Ensemble Performance
Robert Rowe 145

9 Artificial Intelligence Architectures for Composition and Performance Environments
Antonio Camurri 163

10 Dynamic Programming for Interactive Music Systems
Roger Dannenberg 189

11 The Mechanization of Intelligence and the Human Aspects of Music
Carlos Gustavo Guerra 207

12 Artificial Intelligence and Music Education
Brian Smith 221

13 Artificial Intelligence in Music Education: A Critical Review
Simon Holland 239

List of Contributions 275

Index 279

INTRODUCTION TO THE SERIES

The rapid expansion and diversification of contemporary music is explored in this international series of books for contemporary musicians. Leading experts and practitioners present composition today in all aspects – its techniques, aesthetics and technology, and its relationships with other disciplines and currents of thought – as well as using the series to communicate actual musical materials.

The series also features monographs on significant twentieth-century composers not extensively documented in the existing literature.

Nigel Osborne

PREFACE

Music and Artificial Intelligence has recently been established as an important discipline in many music-related curricula throughout the world. There have been one or two advanced books and conference proceedings available, but there has never been a publication which introduces the fundamentals of the discipline to beginners and points students and researchers to other references. *Readings in Music and Artificial Intelligence* is an attempt to bridge this gap. The editor commissioned a number of top Music and AI researchers to write an original chapter introducing their field of expertise. The reader should therefore regard this book as a collection of key introductory texts from a variety of standpoints, at times contradictory.

The book begins with a brief introduction to Music and AI, followed by a discussion regarding the computer as an intelligent aid for composition. The chapter introduces the concept of musical grammars and touches on one of the most fundamental dilemmas in AI: knowledge representation. In chapter two Alan Marsden drives a coach and horses through crucial philosophical issues that have always been taken for granted in AI research. Marsden questions whether we should not definitively acknowledge that human intelligence is incompatible with computers. AI systems, suggests Marsden, would be better off if they were designed to explore the intrinsic capabilities of the computer itself (e.g. processing speed and massive manipulation of data) rather than be designed to emulate the human mind.

The third chapter is by Geraint Wiggins and Alan Smaill. Here the authors present a thorough introduction to the benefits of AI for music research and extend the notion of knowledge representation introduced in the first chapter. Wiggins and Smaill focus on an interesting representation technique of their own called Charm. Then, Petri Toiviainen introduces the reader to a completely different concept: connectionism (or neural networks). Connectionism is regarded by many as a rival of (traditional) AI but each approach has its own merits and pitfalls. Toiviainen presents a general survey of the relationship between these two fields of research and illustrates the benefits of neural networks for music. Although neural networks are highly prized for their ability to learn, Gerard Widmer follows with a discussion of traditional (or symbolic) AI methods for machine learning and puts his assumptions in black and white by means of an example application.

试读结束：需要全本请在线购买： www.ertongbook.com

The next two chapters are dedicated to the benefits of AI in systems for music analysis. François Pachet introduces the basics of computer-aided music analysis and describes his own spectacular system for the analysis of jazz chord sequences. A chapter by Pierre-Yves Rolland and Jean-Gabriel Ganascia follows with an in-depth discussion on the role of pattern extraction techniques in computer tools for music analysis and composition.

AI techniques have also found wide popularity in interactive systems for music composition and performance. Interactive music systems change their behaviour in response to the actions of a performer. Robert Rowe introduces Cypher, a system that has the ability to "listen" and react during a musical performance, as if the computer was a fellow musician in a jam session. The following chapter by Antonio Camurri complements this by presenting the notion of multimodal environments for integrating interactive music systems with visual media and multiple actuators (e.g. dancers and actors). Then Roger Dannenberg proposes Dynamic Programming as a programming strategy for the implementation of interactive music systems.

The last three chapters are dedicated to music education. Music educators see great potential in using AI for building intelligent teaching systems. To begin with, Carlos Gustavo Guerra introduces some fundamental issues in musical education from a holistic approach and criticises the undiscerning use of computers in current educational trends. Next, Brian Smith discusses a number of issues that are central to the design of music tutoring systems, giving various examples to illustrate his points of view. Finally, an extensive survey of current approaches to using AI in music education is given by Simon Holland. A number of existing AI systems for music education are invoked to illustrate the key issues, techniques and methods associated with these approaches.

I would like to thank all the contributing authors for their support, enthusiasm and expertise. I hope that this book will stimulate researchers and musicians to explore this new and exciting field of investigation.

Eduardo Reck Miranda

1

REGARDING MUSIC, MACHINES, INTELLIGENCE AND THE BRAIN: AN INTRODUCTION TO MUSIC AND AI

Eduardo Reck Miranda

The Musical Brain

From a number of plausible definitions for music, the one that frequently stands out in musicological research is the notion that music is an intellectual activity; that is, the ability to recognize patterns and imagine them modified by actions. We understand that this ability is the essence of the human mind: it requires sophisticated memory mechanisms, involving both conscious manipulations of concepts and subconscious access to millions of networked neurological bonds. In this case, it is assumed that emotional reactions to music arise from some sort of intellectual activity.

Different parts of our brain do different things in response to the stimuli we hear. Moreover, music is not detected by our ears alone; for example, music is also "heard" through the skin of our entire body (Storn, 1993). The brain's response to external stimuli, including sound, can be measured by the activity of the neurons. Two measuring methods are commonly used: PET (Positron Emission Tomography) and ERP (Event Related Potential). PET measures the brain's activity by scanning the flow of radioactive material previously injected into the subject's bloodstream. Despite its efficiency, this method is rather controversial because the long term side effects of the radioactive substances to the health of the subject are not entirely known. ERP uses tiny electrodes placed in contact with the skull of a person to measure the electrical activity of the brain. As far as the health of the subject is concerned, ERP is safer than PET, but the measurement is different. Whilst PET scans give a clear cross-sectional indication of the area of the brain where the bloodflow is more intense during the hearing process, ERP gives only a voltage level vs. time graph of the electrical activity of the areas of the brain where the electrodes have been placed.

Our understanding of the behaviour of the brain when we engage in any type of musical activity (e.g., playing an instrument or simply imagining a melody) is merely the tip of an iceberg. Both measuring methods have brought to light important issues that have helped

researchers uncover the tip of this iceberg. PET scans have shown that listening to music and imagining listening to music activate different parts of the brain and ERP graphs have been particularly useful to demonstrate that the brain expects sequences of stimuli that conform to established circumstances. For instance, if you hear the sentence "A musician composes the music", the electrical activity of your brain will tend to run fairly steadily. But if you hear the sentence "A musician composes the dog", the activity of your brain will display significant negative electrical response immediately after the word "dog".

The human brain seems to respond similarly to musical incongruities. Such behaviour obviously depends upon one's understanding of the overall meaning of the language in hand. A number of enthusiasts believe that we are born "programmed" to be musical, in the sense that almost no-one has difficulties in finding coherence in simple tonal melodies (Robertson, 1996).

Understanding Intelligence with AI

The understanding of the behaviour of the human brain is not, however, identical to the understanding of intelligence. Physical measurements of brain activity may certainly endorse specific theories of intelligence, but not all theories seek endorsement of this sort.

One of the goals of AI is to gain a better understanding of intelligence, but not necessarily by studying the inner functioning of the brain. The methodology of AI research is largely based upon logics, mathematical models and computer simulations of intelligent behaviour (Luger and Stubblefield, 1989).

Of the many disciplines engaged in gaining a better understanding of intelligence, AI is one of the few that has special interest in testing its hypotheses in practical day-to-day situations. The obvious practical benefit of this aspect of AI is the development of technology to make machines more intelligent; for example, thanks to AI computers can play chess and diagnose certain types of diseases extremely well.

It is generally stated that AI as such was "born" in the late 1940's, when mathematicians began to investigate whether it would be possible to solve complex logical problems by automatically performing sequences of simple logical operations. In fact AI may be traced back far before computers were available, when mechanical devices began to perform tasks previously performed only by the human mind, including some musical tasks (Levenson, 1994).

Is intelligence synonymous to the ability to perform logical

operations automatically, to play chess or to diagnose diseases? Answers to such types of questions tend to be either or ambiguous biased to particular viewpoints. The problem is that once a machine is capable of performing such types of activities, we tend to cease to consider these activities as intelligent. Intelligence will always be that unknown aspect of the human mind that has not yet been understood or simulated.

Towards Intelligent Music Machines

Music is without doubt one of the most intriguing activities of human intelligence. By studying models of this activity, researchers attempt to decipher the inner mysteries of both music and intelligence. From a pragmatic point of view, however, the ultimate goal of Music and AI research is to make computers behave like skilled musicians. Skilled musicians should be able to perform highly specialized tasks such as composition, analysis, improvisation, playing instruments, etc., but also less specialized ones such as reading a concert review in the newspaper and talking to fellow musicians. In this case the music machine would need to have some basic understanding of human social issues, such as sorrow and joy. Will computers ever display such highly sophisticated and integrated behaviour? More optimistic enthusiasts believe so (Minsky, 1985).

As happens in other areas of AI research, however, computers have so far been programmed to simulate most specialized tasks, but fairly independently from each other. Current research work is now looking for ways to integrate the ability to perform a variety of such tasks; e.g., mechanisms from systems for music analysis are aggregated to systems for composition in order to allow for the computer to autonomously compose music in the style of analysed pieces.

It is debatable whether musicians want to believe in the possibility of an almighty musical machine. Musicians will keep pushing the definition of *musicality* away from automatism for the same reasons that scientists keep redefining intelligence. Nevertheless, AI is helping musicians to better operate the technology available for music making and to formulate new theories of music (Balaban *et al.*, 1992).

Formal Grammars

The notion of *formal grammars* is one of the most popular, but also controversial, notions that has sprung from AI research to fertilize the grounds of these flourishing new theories of music (Cope, 1991). Formal

grammars appeared in the late 1950s when linguist Noam Chomsky published his revolutionary book *Syntactic Structures* (Chomsky, 1957). In general, Chomsky suggested that people are able to speak and understand a language mostly because they have mastered its grammar. According to Chomsky, the specification of a grammar must be based upon mathematical formalism in order to thoroughly describe its functioning; e.g, formal rules for description, generation and transformation of sentences. A grammar should then manage to characterise sentences objectively and without guesswork. Chomsky also believed that it should be possible to define a universal grammar, applicable to all languages.

The study of the relationship between spoken language and music is as old as the music of the Western culture. It is therefore not by accident that Music and AI research has been strongly influenced by linguistics and particularly by formal grammars. Many musicologists believe that Chomsky's assumptions can be similarly applied to music (Lerdhal and Jackendoff, 1983; Cope, 1987; Holtzman, 1994). A substantial amount of work inspired by the general principles of structural description of sentences has been produced, including a variety of useful formal approaches to musical analysis; e.g., Schenkerian-like techniques (Cook, 1987; Smoliar, 1979).

A Brief Introduction to Formal Grammars

Figure 1 below illustrates an example of a grammatical rule for a simple affirmative sentence: "A musician composes the music".

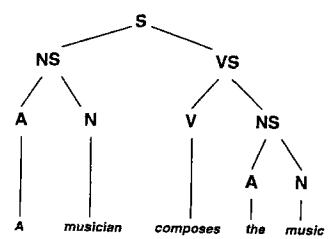

Figure 1 Example of a grammatical rule.

The rule in Figure 1 is saying that:

(a) S = NS + VS (a sentence S if formed by a noun-sentence NS and a verb-sentence VS)
(b) NS = A + N (a noun-sentence NS is formed by an article A and a noun N)
(c) VS = V + NS (a verb-sentence VS is formed by a verb V and a noun-sentence NS)

Such a rule can be programmed into a computer in order to generate sentences automatically. In this case, the computer must also be furnished with some sort of lexicon of words to choose from. For example:

A = {the, a, an}
N = {dog, computer, music, musician, coffee}
V = {composes, makes, hears}

Note that the lexicon defines the three classes of words required by the rule: articles (A), nouns (N) and verbs (V). Given the above rule and lexicon, the computer can be activated to generate meaningful sentences, such as "A computer composes the music" or "A dog hears the musician", but also nonsense ones, such as "A musician composes the dog" and "A coffee hears the computer". The production of nonsense could be alleviated by defining tighter rules, but a notion of semantics would also be necessary for better results.

The rule above is called generative because it is used to generated sentences from scratch. *Transformational rules* work similarly to the generative ones. In this case, the computer is usually programmed to verify if the sentence to be transformed is syntactically correct. This verification is often done by simply matching the *generative rule* that would have generated the sentence. For example, a transformational rule to change the order of simple sentences could be defined as follows:

IF:

$S_{(o)} = NS_{(n)} + VS_{(m)}$ *and*
$NS_{(n)} = A_{(n)} + N_{(n)}$ *and*
$VS_{(m)} = V + NS_{(m)}$

THEN:

$S_{(t)} = NS_{(m)} + VS_{(n)}$ *and*
$NS_{(m)} = A_{(m)} + N_{(m)}$ *and*
$VS_{(n)} = NS_{(n)} + V$

In plain English, the rule above reads as follows: *If the sentence to be transformed is composed of a first noun-sentence followed by a verb-sentence, and the first noun-sentence is formed by an article and a name, and the verb-sentence is formed by a verb and a second noun-sentence of the same format of the first noun-sentence, then the transformed sentence will be formed by the second noun-sentence followed by a new verb-sentence composed of the first noun-sentence followed by the verb.*

Applying this transformation rule to the sentence "A musician composes the music" will give the result "The music a musician composes". Punctuation could also be included in the rule in order to produce "The music, a musician composes."

Generative and transformational rules to generate and transform musical "sentences" can be similarly defined.

An Example of a Music Formal Grammar

In order to define a grammar for music one should carefully consider what the constituents of the rules will be; e.g., notes, phrases, melodies, chords, etc. For the purposes of the example below, we defined the constituents of our grammar in terms of 5 fundamental notions:

(a) the notion R_n of a reference note (e.g., R_1 = C4)
(b) the notion of interval I_n between two notes (e.g., I_1 = perfect 5th)
(c) the notion of direction D_n of the interval (e.g., D_1 = upwards)
(d) the notion of sequence SEQ_n
(e) the notion of simultaneity SIM_n

An example of a generation rule from our grammar is defined as follows (Figure 2):

(a) $SIM_1 = SEQ_1 + SEQ_2$
(b) $SEQ_1 = [I_5,D_1] + [I_8,D_1] + [I_{11},D_1]$
(c) $SEQ_2 = [I_5,D_2] + [I_8, D_2]$

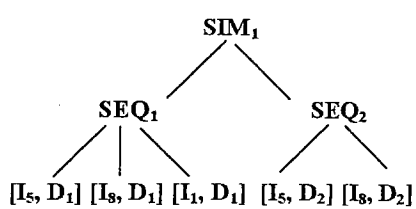

Figure 2 An example of a music grammar.

The *lexicon* for our grammar includes the following:

I = {minor 2nd, major 2nd, minor 3rd, major 3rd, perfect 4th, augmented 4th, perfect 5th, minor 6th, major 6th, minor 7th, major 7th, octave, none}
D= {upwards, downwards, none)

The pair $[I_5,D_1]$, for example, indicates that the interval is a perfect 4th (the fifth element of the set I) and the direction of the interval is upwards (the first element of the set D). Thus, the rule above reads as follows: *A certain musical passage is composed of two sequences played simultaneously. One sequence is formed by 3 notes and the other is formed by 2 notes. The notes of the former sequence are calculated from a given reference note in this order: a perfect 4th upwards, a minor 6th upwards and a major 7th upwards. The notes of the latter sequence are calculated, using the same reference note, in this order: a perfect 4th downwards and a minor 6th downwards.* By establishing that the reference point R = C4, the musical passage shown in Figure 3 can be generated:

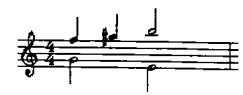

Figure 3 Musical passage generated by the grammar.

Note that rhythm has not been included in our grammar, in order to keep our examples as simple as possible.

A transformational rule for the above type of passage could, for example, create a new sequence SEQ_3 by joining the two sequences into a single simultaneous event ($SIM_2 = SEQ_1 + SEQ_2$), followed by the original first sequence SEQ_1 (Figure 4):

IF:

$SIM_1 = SEQ_1+SEQ_2$ *and*
$SEQ_1 = [I_5,D_1] + [I_8,D_1] + [I_{11},D_1]$ *and*
$SEQ_2 = [I_5,D_2] + [I_8, D_2]$

THEN:

$SEQ_3 = SIM_2 + SEQ_1$ *and*
$SIM_2 = SEQ_1 + SEQ_2$

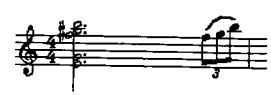

Figure 4 An example of a musical transformation

The computer could be programmed to produce an entire musical composition by successive activation of a variety of generative and transformational rules (Figure 5):

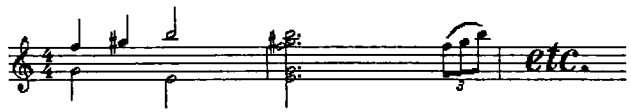

Figure 5 Musical material produced by successive activation of rules.

The music produced by computer programs of this kind is usually labelled as *algorithmic composition*. Algorithmic composition, however, does not necessarily mean composition produced by grammars, but all music produced automatically by any kind of computer program. Most composers working with computers nowadays tend to be cautious with such labelling.

Who Composed *Entre l'Absurde et le Mystère*?

Entre l'Absurde et le Mystère is a piece for chamber orchestra produced by *CAMUS*, a computer system designed by the author (Miranda, 1993; 1994). CAMUS uses cellular automata-based simulations of biological behaviour to produce sequences of music structures (e.g., melodies, chords, clusters, etc.)

The public warmly applauded its performance by The Chamber Group of Scotland in 1995 in Edinburgh. Martyn Brabbins, the conductor, was reluctant to believe that a computer had generated the piece and generally members of the audience found the piece to be pleasant. The general wonder of that evening was: "Was the piece really composed by a computer?"

This question is debatable and has serious ideological implications. To our point of view, a distinction between *author* and *meta-author* should be made in such cases. The ultimate authorship of the composition here should be to the person who designed and/or operated the system. Even in the case of a program that has the ability to program itself, someone is ultimately behind the design and/or the operation of the system. Similarly, one would hardly consider that Pierre Boulez's *Polyphonie X* (Nattiez, 1993), for example, was composed by the *serialism system*.

Semantics

Formal grammars are suitable for the description of the *syntactical rules* of a language but they do not guarantee *meaningful formations*. If one wishes to program a computer to produce meaningful sentences automatically,

some notion of *semantics* must be included in the system. Although semantics is a concept primarily related to spoken languages, it also applies to music to a certain extent.

Semantics, often referred in AI jargon *as declarative knowledge*, is a crucial aspect of AI research. A substantial amount of research work is dedicated to the design of methods to represent knowledge effectively (Brachman and Levesque, 1985).

Knowledge Representation

Designers of AI systems require *knowledge representation* techniques that provide representational power and modularity. They must capture the knowledge needed for the system and provide a framework to assist the systems designer to easily organize this knowledge (Bench-Capon, 1990; Luger and Stubblefield, 1989).

The Internal Representation Hypothesis

The primary assumption in AI is that mental activity is mediated by *internal representations*. Although there is no consensus as to what these representations actually are (some regard them as neurophysiological states, whilst others may define them as symbols or even images), the traditional approach to AI assumes that *intelligent activity* is achieved through:

(a) the use of symbols to represent a problem domain
(b) the use of these symbols to generate potential solutions to problems
(c) the selection of a suitable solution to a problem.

The use of an adequate knowledge representation technique is therefore one of the most important keys for the design of successful AI systems.

A brief survey of knowledge representation paradigms

Logic representation: first-order predicate calculus

A number of logics have been developed in philosophy and mathematics to represent knowledge; for example *propositional calculus* and *first-order predicate calculus*. The first-order predicate calculus is largely used in AI systems.

The first-order predicate calculus provides a well-defined language for describing and reasoning about qualitative aspects of a system. It can denote objects of a domain by using simple symbols and

can express relationships between objects, assertions and denials of these relations, and logical relations between these statements (Luger and Stubblefield, 1989).

The first-order predicate calculus is sufficiently general to provide a foundation for other models of knowledge representation. AI problem domains however often require large amounts of highly structured interrelated knowledge. Some high-level notion of structure is needed to help the systems designer represent complex concepts in a coherent way. The first-order predicate calculus alone does not provide this help.

Network representation: graphs

A network representation also provides the means to denote objects of a domain and relations between them by using simple symbols. The advantage of network representations over logic representations is that the former can provide some high-level notion of structure that helps the systems designer to represent taxonomically structured information. The philosophy behind a network representation is that one reasons about a concept or object by relating it to other concepts or objects of the domain.

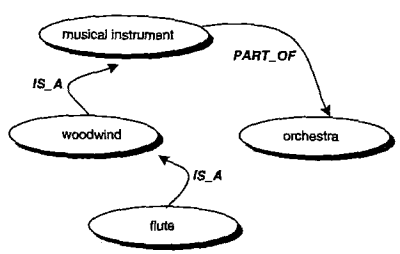

Figure 6 An example of a simple semantic network.

Graphs technique is an example of network representation; it provides a means to explicitly represent objects and relations by using nodes and arcs. A number of graphs techniques have been developed and used in AI systems; for example, conceptual graphs and semantic networks.

A *semantic network*, for instance, consists of a network of *nodes* linked by *arcs*, so that nodes represent *concepts*, or *object*, and arcs represent *relations* between them; both nodes and arcs are usually labelled (Figure 6).

In Figure 6, there are 4 objects ("musical instrument", "orchestra", "woodwind" and "flute") and 2 types of relations ("PART OF" and "IS A"). This semantic network represents the following facts:

A musical instrument is part of the orchestra.
A woodwind is a musical instrument.
A flute is a woodwind.

Other facts can be inferred from the network, in addition to the facts which are explicitly represented. For example:

A flute is part of the orchestra.

One of the advantages of a graphs-based representation is that facts come from the definition of links and associated inference rules that define specific mechanisms, such as the *inheritance mechanism*. In the case of the above example, "flute" inherited the fact that it is part of "orchestra".

In itself a graphs-based notation of relationships is not so different from the first-order predicate calculus. The power of a network representation is that it provides an explicit method to represent objects and relations, and promotes the organization of knowledge into *class hierarchies* and the inheritance mechanism.

Structured representation: Frames

Network representations allow for the representation of knowledge using explicit links between single objects in a knowledge base. Structured representations however extend network representations by providing a means to organize large networks of knowledge into a collection of separate networks, each of which represents some stereotyped situation or *class of objects*.

Frames technique is a type of structured representation. Frames technique allows for the representation of complex structures by encapsulating multiple *attributes* of situations, or objects, into single units, or *classes of objects*, in the domain.

A frame is a data structure whose components are called slots. Slots have names and accommodate various types of information: a value, a link to other frames or procedures to calculate its value. A slot may also be left incomplete.

As in semantic networks, the most useful feature of frames is the inheritance mechanism; when a frame represents a class of objects and another frame represents a superclass of this class, then the class frame inherits from the superclass the values for its incomplete slots. Examples of frames:

试读结束：需要全本请在线购买： www.ertongbook.com

FRAME: *musical instrument*
part_of: *orchestra*
FRAME: *woodwind*
is_a: *musical instrument*
excitation_method: *air stream*
resonance_method: *pipe*

Note that slots are similar to the arcs of a network representation. Slots however have the advantage that they can hold procedures to perform some function, in addition to links to other concepts.

Structured representations thus extend network representations by representing complex objects as interconnected structured single entities, rather than as one single large network.

Conclusion

Mathematics and logics undoubtedly play a dominant role in the formalisation of intelligence for AI research. But is formalization the right approach to express intelligent behaviour? Is it right to distinguish between mind and body, semantics and syntactics, knowledge and abstract representation schemes?

The great majority of AI work to date has assumed that intelligence can be simulated by encapsulating chunks of data into static "packets" of information. Intelligent activity is then performed by an "engine" that picks and combines appropriate packets of information stored in memory in order to achieve specific goals. In this case, knowledge is often classified into two main groups: *declarative* (e.g., the semantics of the grammar or the "meaning" of the packets of information) and *procedural* (e.g., the grammar itself or the "how" the engine should function).

In fact humans store knowledge in a more complex way. Our brains are stubborn systems which cannot be deconstructed so neatly. Human intelligence is formed by both *conscious* and *unconscious* elements distributed at different levels of layers in our mind.

Only the conscious ones can be objectively accessed and manipulated. We seem not to have access to the levels which do most of our thinking. When we think, we certainly change our own rules and rules that changes the rules, and so on, but we cannot change the lower layers; i.e., neurons always function in the same way.

Modern approaches to AI seek inspiration from this neurophysiological model. Traditional AI research methods are not necessarily inspired by neurophysiology but have, nevertheless, produced fruitful results. Perhaps the most fruitful of them all is the conclusion that there are many types of intelligence and each have their own characteristics. Music certainly involves a very distinct type of intelligence and it is up to Music and AI researchers to find the right approaches to it.

9. References

Balaban, M., Ebicioglu, K. and Laske, O. (Editors), *Understanding Music with AI: Perspectives on Music Cognition*, Menlo Park: AAAI Press/MIT Press, (1992).

Bench-Capon, T.J.M., *Knowledge Representation: An Approach to Artificial Intelligence*, London: Academic Press, (1990).

Brachman, R.J., and Levesque, H.J. (Editors), *Readings in Knowledge Representation*, San Mateo: Morgan Kaufmann Publishers, (1985).

Chomsky, N., *Syntactic Structures*, The Hague: Mouton, (1976).

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

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

Cope, D., An Expert System for Computer-assisted Composition, *Computer Music Journal*, (11)4, 30–46. MIT Press, (1987).

Holtzman, S.R., *Digital Mantras*, Cambridge/MA: MIT Press, (1994).

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

Leveson, T., *Measure for Measure: A Musical History of Science*, New York: Simon and Schulster, (1994).

Luger, G.F. and Stubblefield, W.A., *Artificial Intelligence and the Design of Expert Systems*, Redwood City: Benjamin/Cummings, (1989).

Miranda, E.R., Cellular Automata Music: An Interdisciplinary Project, *Interface*, 22, 3–21, (1993).

Miranda, E.R., Music Composition Using Cellular Automata, *Languages of Design*, 2, 105–117, (1994)

Minsky, M. *The Society of Mind*, New York: Simon and Schuster, (1985).

Nattiez, J-J., *The Boulez-Cage Correspondence*, Cambridge: Cambridge University Press, (1993).

Robertson, P., *Music and The Mind*, television production for Channel 4 first shown May 1996, UK.

Smoliar, S., Computer Aid for Schenkerian Analysis, *Communications of the ACM*, 110–115. (1979)

Storr, A., *Music and The Mind*, London: Harper Collins Publishers, (1993).

2

MUSIC, INTELLIGENCE AND ARTIFICIALITY

Alan Marsden

Introduction

Computers are machines. Intelligence is a human characteristic, and though it is often taken to be the characteristic which distinguishes us from animals, computers rarely approach the intelligence even of animals. One of the characteristics of machines is that they are man-made, and in that sense artificial (the sense of artificial as "unreal" will be discussed briefly below). The characteristic which distinguishes them from other artificial things is *behaviour*, and this characteristic is one which they share with humans and animals. In fact everything has behaviour, in that everything responds in a particular way in interaction with an environment: if one drops a ball it bounces; if one drops a glass it smashes. The real distinction between machines and other artificial objects cannot be made without reference to human values and intentions: we value machines because of their behaviour and not because of other characteristics (e.g. their shape, dimensions and solidity, as in the case of chairs). We use machines to extend our own behaviour. A class of machines which has become particularly important during this century is machines whose behaviour concerns information. This class contains such ancient machines as the printing press and such common ones as the telephone – it is a mistake to regard *information technology*, at least in this sense, as something new.

The characteristic which computers have which is genuinely new, and which sets them apart from other information-processing machines, is that their behaviour is not only controllable by the user (this is an important characteristic of all useful machines) but that their behaviour is *definable* by the user. Other machines can have this characteristic, both information-processing machines and others, but only within tight constraints. In the case of a computer, on the contrary, its behaviour is highly unconstrained, at least in the domain of the processing of information (the possibilities for physical behaviour are usually very limited). In fact, the ideal computer is a *universal processing machine* which is capable of performing any kind of behaviour in the domain of abstract information processing. At the level of programming, the "input" which a computer reads is a definition of a kind of behaviour, or in other words, a definition of an abstract machine. If computers are thus intended to be

able to mimic any kind of behaviour, it is not surprising that there should be interest in programming computers to behave in ways that are human-like and which could be called intelligent. There has also been interest in having computers perform musical tasks, whether it be playing music, processing music, or creating music. Whether behaving in a musical manner implies behaving in a human manner is discussed below. For now, it is sufficient to note that the combination of the two – the intention to behave in a human-like fashion and to perform a musical task – is the topic of this chapter.

History

An argument is presented below that no attempt to have a computer perform a musical task can be totally unconcerned with the issues of Artificial Intelligence, but customarily Music-AI has included only those musical computer systems which have involved a degree of complexity which is not the complexity of mathematical formulae, nor the complexity of large quantities of data, but rather a kind of complexity of ideas. As in other domains, certain tasks have been considered to involve intelligence while others have not. (This is a problematic issue, which will be returned to below.) Sound synthesis, for example, is an area which has attracted a great deal of very successful work, but little of it is regarded as being in the domain of Music-AI because it has concerned acoustic and psycho-acoustic phenomena and the mathematics of signal processing rather than being concerned with thinking. Similarly, the vast area of systems for capturing, processing and using performance data via sequencers and the like is excluded from the domain of Music-AI, as are systems for music notation. A brief historical survey is presented here, organised around different architectures of systems.

Early attempts at programming computers to perform musical tasks took an *algorithmic approach*. The objective was to describe the procedures which must be performed in order to produce a musical result. An example of high-quality work of this kind can be found in the research of Longuet-Higgins and his co-workers (Longuet-Higgins, 1978; Longuet-Higgins and Steedman, 1971). The objective of this work was a system which could transcribe music played on a keyboard (the work began in the days before MIDI) to music notation. This involves resolving issues about the representation of pitch (should a note be written as C sharp or D flat, for example), which involves determining key, and issues about the representation of rhythm, which involves both determining metre and coping with the variations from metronomic playing of a real

performance (which can be quite severe). Algorithms with a moderate to high degree of success for these tasks were designed and implemented in the language Pop-11. Projects which have also taken an algorithmic approach have been directed at tasks as diverse as composition (Ames and Domino, 1992; Cope, 1991) and transcription of lute tablatures (Charnassé and Stepien, 1992). While this approach can produce good results, if those results are to be applicable in other programs to perform other tasks, then it is up to the researcher to make certain that the algorithms are suitably designed and explained. Some authors, (Longuet-Higgins among them) are excellent at explaining what their algorithm does; others are not so. The algorithms themselves, without explanation, cannot be expected to be transferable to a program to perform another task, however similar. At issue here is really the nature of the principal objective of research in Music-AI. Is it to design and implement computer systems which perform musical tasks (an engineering objective), or is it to discover and explain the knowledge which underlies these tasks (a cognitive-science objective)? Most researchers would claim the latter, but this can only be tested by achieving the first objective to some degree also.

While every computer program ultimately comes down to algorithms, there has been considerable interest in devolving the translation from knowledge to algorithm to the computer so that the representation in which a system is expressed can be more directly a representation of the knowledge underlying a particular task. A number of formalisms intended to achieve this have been designed. The one which has most often been used in music, usually because of a perceived similarity with language, has been formal grammars. Another early example of Music-AI is the *harmonic analysis* system of Winograd (1968). The core of this was a *systemic grammar* which described the configurations of chords, harmonies and tonalities possible in homophonic tonal music such as the chorale harmonisations of J.S. Bach. This gave an extremely clear exposition of the "knowledge" of tonal theory. The grammar could be applied in the analysis of a piece of music to discover how the grammar accounts for the piece, and thereby, by reporting the steps of the derivation, producing a harmonic analysis of the piece. However, many different analyses were possible for any one piece (musicians will be familiar with the idea of different possible analyses, but they might be surprised at quite how many were allowed by Winograd's grammar, which was quite a faithful reproduction of classical tonal theory.) The part of the system which derived analyses, therefore, called the "parser", had to be quite complex and make use of other, procedural knowledge in order to arrive at harmonisations which were acceptable. In principle a grammar should be applicable in either direction, i.e. to either

analyse music or produce music. It might be possible to use Winograd's grammar to produce harmonisations, but Winograd did not attempt this. A well-known grammar which did produce music was that of Baroni *et al.* (1992), who produced a number of grammars to generate chorale melodies, eighteenth-century French chansons, and the text repetition patterns of Legrenzi arias. In both Baroni et al. and Winograd's work, the business of translating the grammar to an algorithm was not devolved to the computer, as suggested above, but coded by hand. In the case of Kippen and Bel's *Bol Processor* (1989; 1992), however, the computer system operated directly on the grammars. The Bol Processor was a system intended to assist in the understanding of a style of *tabla drumming* found in North India. It was capable both of producing new pieces of music, and of analysing existing pieces. Their publications also include excellent discussions of the principles of using grammars in this kind of work and of some of the issues involved.

Another formal systems for representing knowledge applied in Music-AI is KL-ONE, a well-developed system of knowledge representation, derived from frames and semantic networks, which expresses knowledge in terms of concepts and roles, and defines inheritance and other relations between them. Here again, the intention is to allow a clear expression of knowledge which is susceptible to direct implementation by computer. Furthermore, this knowledge is, in principle at least, expressed abstractly without any reference to its application in any particular task. KL-ONE is used to provide the symbolic layer of HARP, a hybrid system applied to a number of musical tasks, often involving real-time interaction between a performer and a music-production system (Camurri *et al.*, 1994); see chapter by Antonio Camurri in this volume.

One of the problems Kippen and Bel identified in developing their Bol Processor grammar was the difficulty of knowing what should go into a grammar: how is the researcher to determine what the rules of the grammar should be? The common paradigm has been to make a first attempt, to examine its results, then, on a rather *ad hoc* basis, to attempt some revisions to the grammar which will correct the errors of the previous results. The cycle of testing and revision then begins again. Such a strategy will probably never produce a perfect system, though it might approach perfection, but the *ad hoc* nature of the rule revision is disconcerting: how can the researcher have any confidence that the revisions are the best to propose in the circumstances? It is a characteristic of an intelligent animal that it learns from its experience and performs better next time in similar circumstances. In fact, this behaviour is more characteristic of intelligence than is behaving well in every circumstance.

One of the goals of Artificial Intelligence, then, is systems that learn, and these can be found in Music-AI also. Kippen and Bel attempted to build learning into their system so that rule strengths could be adapted automatically and so that at least some of the new-rule generation process could be automated (1989). Musical learning systems, however, are best exemplified in the work of Widmer, who has completed projects which learn counterpoint rules (Widmer, 1992) and which learn expressive performance (Widmer, 1996). *EMI* system (Cope, 1991), which learns to compose music in the style of the music given to it, does not properly belong in this category of intelligent learning systems because the learning requires a considerable degree of input from the user of the system also. While it is true that intelligent animals often learn best with teachers, these teachers do not interfere with the functioning of the animal in any way other than the normal channels of interaction. (Teachers do not resort to brain surgery, in other words.) Furthermore, it is a, characteristic of intelligent animals that they learn spontaneously, and it is this characteristic that is most sought in AI research in learning.

A number of characteristics of intelligent behaviour, including the one of *spontaneous learning* just mentioned, gave rise at the end of the 1980s to a new paradigm in computing variously called *connectionism, parallel distributed processing*, or *neural networks*. Two of the most important motivations were the observation that intelligent behaviour could not possibly arise from the mechanisms proposed by traditional symbolic AI approaches at the speed at which it does in animals. Furthermore, it is a characteristic of intelligent animals that, in surroundings which they have never before encountered, and therefore surroundings for which they have no perfectly applicable knowledge, they are able to perform tolerably well. Traditional AI systems, however, when presented with something somewhat different from their intended task, generally perform spectacularly badly. This is sometimes referred to as brittleness. In the new paradigm, which is clearly explained in Leman (1992) and other sources, the behaviour of a system results from the net effect of the behaviour of a number, possibly a very large number, of simple but interacting processing units. When appropriately configured, such systems are capable of learning, in the sense that their behaviour approaches the desired behaviour. Furthermore they typically perform moderately well with unfamiliar input rather than exhibiting the brittleness of classical systems. Such systems have been used with remarkable success in such diverse domains as *tonal theory* (Leman, 1994, 1995), the *classification of timbre* (Cosi *et al.*, 1994), and the *quantisation of rhythm* (Desain and Honing, 1992). Desain and Honing (1992) include a direct comparison of a classical and a network system performing the

same task. From the engineering perspective, such systems often perform well. From the cognitive science perspective, however, they involve a total shift of philosophy. It is inappropriate to use a network system in the hope of discovering the rules of tonal harmony, for example, at least in the form that they are traditionally expressed. The knowledge which a network system acquires during its learning is distributed through the connections of the network; one cannot necessarily examine the state of the network after training and directly extract from it a rule in the form "if X then Y", as one often can from a learning system based on classical computing.

The philosophical shift has justifications other than the utility of such network systems, expressed in Leman (1993), Lischka (1991) and Kaipainen (1996), but it is important to realise quite how different it is from the cognitive science which gave rise to grammars, KL-ONE, and the like. Nor should it be thought that the new paradigm has supplanted or should supplant the former one. Much recent work involves both kinds of computing (e.g. Camurri and Leman, 1992; Goldman *et al.*, 1995), often assigning *subsymbolic* processing to a network while *symbolic* processing is carried out using a more traditional kind of architecture. However, care must be taken in ensuring that the mixture of the two philosophies is sound in the goal of improving. The understanding of musical behaviour – the cognitive-science goal which was argued above to the fundamental to Music-AI – is not to be compromised.

Philosophy

In a precise discussion of Music-AI, there are three terms to be defined: "music" "artificial" and "intelligence". Some adumbrated definitions were given above. "Artificial", for example, was taken to mean man-made and not occurring naturally in the universe. By this definition music is also artificial, as is any other human product. A tightening of the definition is warranted, restricting the word "artificial" to refer to human products which are intended to emulate something else (which probably, but perhaps not necessarily, occurs naturally), hence artificial pearls, etc.

"Music" is notoriously difficult to define (for a straightforward discussion of some of the issues, see Davies, 1978), but all agree that while it involves sound, it is impossible to define solely in terms of sound. The classic test case is John Cage's piece for piano *4'33"*, during the performance of which the only sounds heard are those which happen to occur in the environment – the performer is not instructed to make any sounds at all. If this piece, in which any sound can occur, is to be taken as music, then any sound is music and so all sounds are music. This is clearly

unsatisfactory as a definition of the word as normally understood. Even if this extreme case is not admitted as a piece of music, it is not difficult to name pieces in which all kinds of normally non-musical sounds have been included, and it is extremely difficult to find physical differences between the sounds which characterise music and those that do not. Thus definitions of music generally make reference in some way or other to human activities, whether composition, performing or listening. If, then, the very definition of music requires reference to human activities, any computing system which is supposed to perform a musical task must also take account of those human activities. As an example, consider a sound synthesis system, a common kind of musical computing system which is not normally considered an example of artificial intelligence. In designing any such system, choices must be made about the frequency ranges to be accommodated (and hence the sampling rates to be used). For a musical system, the appropriate choices are to set the frequency range to the maximum humanly audible range, since the results are intended to be listened to by people and not bats or any other animal with a different audible range. Pursuing the example further, suppose that the designer wishes the user of the system to be able to specify the sound output in terms of individual sound events, which we might call "notes", and to specify the time of occurrence for each note. This will require some reference to the phenomena by which we segment a stream of sound into separate events, and also an understanding of where the perceived "start-time" of a note is in relation to the physical beginning of the sound, its amplitude envelope, etc. Going yet further, the user might want to be able to specify the grouping of notes into phrases and have this phrasing reflected in the synthesised sound. This would require an understanding of the relation of variations in timing and other factors to perceptions of phrase beginnings and endings (see Todd, 1985; Sundberg, Friberg, and Frydén, 1991). The point of the argument is that if any system is to be musical it must make reference to human behaviour, and to that extent *any* musical system must involve artificial intelligence. There is no obvious place at which to draw a boundary between where one must take into account human behaviour which is not intelligent, and where one must take into account behaviour which is intelligent. By this argument, furthermore, the discipline of Artificial Intelligence becomes not a peripheral specialisation but a core element of successful computer science.

"Intelligence" is the most difficult of the three terms to define, and the one whose definition is most contentious. It was suggested in the introduction above that artificial intelligence meant programming computers to behave like people. Later, spontaneous learning and

performing with moderate success in unfamiliar surroundings were suggested as characteristics of intelligent behaviour. A third definition is suggested by a common usage of the word "intelligent" with respect to software. An "intelligent help system", for example, is one which determines the information to be provided to the user on the basis of the user's recent activities. In other words the behaviour of the system is sensitive to its environment. This is true of every piece of software – its output it determined by its input – but here there is a significant difference in the domain of the input. Normally software uses a very restricted input; so-called intelligent software instead attempts to receive input from as much as possible of its environment. Clearly this is related to the definition of intelligent behaviour as performing moderately well in unfamiliar surroundings, because attention is payed to the totality of the surroundings. Furthermore, if the environment is taken to include the past, then this definition of intelligence as behaving appropriately in the environment will include learning also. However, computers generally have extremely limited channels for receiving input from the environment, and considerable work is needed in this area if we are to see behaviour which is really intelligent under this definition. In fact, if we really want an intelligent computer to behave in the same way in which a human would in a given environment, including that environment's past, then the computer would have to have the same channels of input, the same memories, the same means of acting upon the environment, and indeed the same objectives. In short, the computer would *be* that person. Artificial intelligence under this definition, then, is an impossible goal.

Some of these difficulties are overcome by limiting the channels of communication, as in the definition of intelligence encapsulated in the *Turing test*, proposed by Alan Turing at the very beginning of the discipline of Artificial Intelligence. The test is as follows. Two rooms have teletypes (the technicalities are not significant – any restricted means of communication usable by both computers and humans would do) as the only means of communication with the outside world. In one room is a computer connected to the teletype, in the other a person. Those on the outside may ask questions via the teletype, in a restricted domain. If they cannot tell from the responses to the questions which room contains the computer and which the person, the computer has passed the test and may be described as intelligent. A musical version of this test could be proposed also. (For a similar argument making a point related to the one above about the essentially human nature of musical activity, see Cross, 1993.) Two rooms are set up with a channel by which music is communicated to the outside world. We might also allow a channel by

which some sort of feedback (applause, perhaps, or other pieces of music) goes into the room. In one room is a composer; in the other is a computer. The test is passed when those outside the rooms cannot tell which contains the computer. While it might be possible for a computer to pass this test in practice (i.e. in an empirical sense), there is an argument that a computer could never pass the test in principle (i.e. in a rationalist sense). (While the test might appear inherently empirical, because it fundamentally involves observations, it is generally not conducted in practice but as a "thought experiment", and so is not empirical at all.) It is often argued that originality is an essential characteristic of music. (From the perspective of composing, this is commonplace; for a perspective from listening, see Kunst, 1978). Computers are digital automata, and so their behaviour is always, in principle at least, predictable and therefore *cannot* be original. Thus a computer cannot, in principle, pass this test. There is a persuasive counter-argument that dynamic systems, and so-called chaotic systems in particular, can be deterministic, in the sense that their future state is entirely determined by their current state, but yet unpredictable. In fact such systems have been used for creating both music and visual art (the visual examples are quite well known; see Little, 1993 for a musical example). However, this depends, in principle, on the dynamic system operating in an infinite domain (e.g. using rational numbers), and computers can only simulate this by a finite domain of very many elements. The argument in principle, therefore, remains. The argument in practice will not be defended because clearly it is a hopeless task for a person to know all the details of the state of a computing system, finite as the number of possible states might be. Indeed, it is now to matters of practice that we will turn.

Pragmatics

If the goals of Music-AI suggested above – behaving in a completely human-like way and composing music indistinguishable from humanly-composed music – are impossible to achieve, what should Music-AI realistically aim at? In fact, we frequently want superhuman, and therefore non-human, behaviour from computers. We often want computers to process data in larger quantities, at greater speed and with greater accuracy than is humanly possible. In these cases, putting aside questions about whether the computer's behaviour is really intelligent, it is precisely because it is artificial (other-than-human) that it is useful. Thus the real goal in developing a computer system is often for it to behave in a human-like manner in some respects but in a non-human-like

试读结束：需要全本请在线购买： www.ertongbook.com