Artificial intelligence, education and music : the use of artificial intelligence to encourage and facilitate music composition by novices

Thesis

Artificial Intelligence, Education and Music

The use of Artificial Intelligence to encourage and facilitate music composition by novices

Simon Holland

MSc Cognition, Computing and Psychology, University of Warwick, 1985
BSc Mathematics, University of Reading, 1976

Submitted for the degree of Doctor of Philosophy in Artificial Intelligence
The Open University, Milton Keynes

July 1989

Author's number: M7023015
Date of submission: July 1989
Date of award: 5th September 1989

HIGHER DEGREES OFFICE Christine

LIBRARY AUTHORISATION FORM

STUDENT: S. HOLLAND

SERIAL NO:

DEGREE: PhD

TITLE OF THESIS: ARTIFICIAL INTELLIGENCE, EDUCATION & MUSIC

I confirm that I am willing that my thesis be made available to readers and maybe photocopied, subject to the discretion of the Librarian.

SIGNED: Simon Holland

DATE: 16/3/90

Abstract

The goal of the research described in this thesis is to find ways of using artificial intelligence to encourage and facilitate music composition by musical novices, particularly those without traditional musical skills. Two complementary approaches are presented.

We show how two recent cognitive theories of harmony can be used to design a new kind of direct manipulation tool for music, known as "Harmony Space", with the expressivity to allow novices to sketch, analyse, modify and compose harmonic sequences simply and clearly by moving two-dimensional patterns on a computer screen linker to a synthesizer. Harmony Space provides novices with a way of describing and controlling harmonic structures and relationships using a single, principled, uniform spatial metaphor at various musical levels; note level, interval level, chord level, harmonic succession level and key level. A prototype interface has been implemented to demonstrate the coherence and feasibility of the design. An investigation with a small number of subjects demonstrates that Harmony Space considerably reduces the prerequisites required for novices to learn about, sketch, analyse and experiment with harmony - activities that would normally be very difficult for them without considerable theoretical knowledge or instrumental skill.

The second part of the thesis presents work towards a knowledge-based tutoring system to help novices using the interface to compose chord sequences. It is argued that traditional, remedial intelligent tutoring systems approaches are inadequate for tutoring in domains that require open-ended thinking. The foundation of a new approach is developed based on the exploration and transformation of case studies described in terms of chunks, styles and plans. This approach draws on a characterisation of creativity due to Johnson-Laird (1988). Programs have been implemented to illustrate the feasibility of key parts of the new approach.

Dedication

To Caroline, Simon, Peter, and my Mother and Father
with all my love

Acknowledgments

Thanks to my supervisor, Mark who has born the brunt of the lunacy, the diversions, the modal harmonic ostinati, and kept me on planet earth throughout. Mark provided much needed support. He also provided an inspiring and energetic example of how to do outstanding research in Artificial Intelligence and Education. Without Mark's assured expertise in AI and his practical musicianship, I would not have had the confidence to start research in AI and Music.

This work would not have been started (or completed in a reasonable time scale) without Tim O'Shea, my co-supervisor. At the beginning, Tim found a simple way to get me to overcome my own doubts about tackling this challenging, interdisciplinary subject. At countless critical points, Tim solved practical difficulties and found just the right few words to motivate me. In the last few weeks before submission, Tim displayed exceptional generosity, horse sense, humour, intelligence and creativity. Tim invented ways of cutting the amount of work to do in half. He patiently and humorously picked garbage out of faxed drafts in the middle of the night. (The remaining garbage is due to my own cussedness). Tim is a first rate researcher and nurturer of research, but, more rarely, he is also a generous and open person.

I would like to thank Mark Steedman for his suggestion that Longuet-Higgins' theory might be a good area to explore for educational applications.

I would like to record my gratitude to George Kiss, a fine researcher and a conscientious teacher for being my mentor when I began Artificial Intelligence research.

Thanks to Claire O'Malley, who was supportive and generous in the last, crucial, phase of the write up, at a time when she had much more pressing demands on her time. I would like to thank people who gave their time to read drafts of the thesis, producing vital improvements; Peter Desain, Henkjan Honing and Sara Hennessey. Thanks to Fiona Spensley, Alistair Edwards, Mike Baker, John Sloboda, Trevor Bray and Mark Steedman for feedback on particular chapters or earlier drafts. Special thanks are due to Simon Bento for doing a wonderful job on proof-reading (any finally escaped typos are my fault - I look forward to reading Simon's thesis). Special thanks to Barry Jones for very kindly taking time to rescue me from various musical errors.

Thanks to Chris Fry, Kevin Jones and Tom Green who generously took time to give me useful pointers when I was getting in started in AI and Music. Thanks for encouragement, useful discussions, support and shared joy from other researchers in AI and music and other (more or less) related fields; Dave Levitt, Henry Lieberman, folk at the Media Lab; Peter Desain, Henkjan Honing, Sterling Beckwith, David Wessel, Lelio Camilleri, Linda Sorisio, Christopher Longuet-Higgins, Jeanne Bamberger, Wally Feurzeig, Gerald Balzano, Mike Baker, Eric Clarke, Ian Cross, Stephen Pope, Mira Balaban, Kemal Ebcioglu, Rick Ashley, Derek Ludwig (and others at NorthWestern); Colin Wells, Nigel Morgan, Keith Waters, Anders Friberg, Lennart Fahlen, Steve McAdams, Xavier Rodet, Jean Baptiste Barrière and others at IRCAM; Marilyn Taft Thomas, Roger Dannenberg and folk at Carnegie Mellon; Stellan Ohlsson, Jeff Bonar and others at LRDC; Alan Borning, Bill Gaver, Fred Lerdahl, Alan Marsden, Stephen Page, Mike Greenhough, Christoph Lischka, Neil Todd, Tony Watkins, Simon Emmerson and Pete Howell.

Thanks for strong practical and moral support to Eileen Scanlon, Diana Laurillard, Pete Whalley, Sterling Beckwith, Nancy Marlett, Rachel Hewson, Richard Southall, and Jayalaksmi.

Special thanks to Olwyn Wilson, Dave Wakely (watch out, National Sound Archive), Tony Durham, Conrad Cork, Phil Butcher, Dianne Murray, Steve Cook, Mike Brayshaw, Di Mason, Pat Cross and Rick Evertsz.

Thanks for various kinds of help and encouragement to Pat Fung, Fiona Spensley, Anne Blandford, Norman White, Gordon Burt, Nicola Durbridge, Kathy Evans, Rae Sibbitt, Roshni Devi, Laurence Alpay, Rod Moyse, Betty Swift, Kate Stainton-Ellis, Beryl Crooks, Anne Wood, Margaret Harkin, Helen Boyce, Arthur Stutt, John Domingue, Tony Hasemer, Tim Rajan, Marc Eisenstadt, Hank Kahney, Don Clarke, Yibing Lee, Ann Jones, John Close, Richard Joiner, Bradley, Ed Lisle, Mike Fox Jitu Patel, Dave Kay, Simon Nuttall, Phil Swann, Benedict Heal, Hansa Solanki, and *all* in the Centre for Information Technology in Education, the Institute for Educational Technology, the Computer-Aided Learning Research Group and the Human Cognition Research Laboratory.

Thanks to various musicians who helped: Herbert Balsam, Martin Brems, Tony Jefferies, Jonathan Brown, Grant Shipcott, Sean Graham, John Winter, Guillame Ormond, John Walsh, John Close, Mike Baker, Rod Moyse, John Bickersteth, Fred de Borst, Gerda Koetje, Dave Gollancz, Richard Middleton, Donald Burrows, Rev. Cousins, Dean Lloyd, Frank Hawkins and Chris Clark.

I owe a considerable typographical debt to Rob Waller and Rachel Hewson.

Emergency squad: very big thanks to Simon Bento, Pete Whalley (not for the first time), Claire O'Malley, Katerine Bielaczyc, Tom Vincent and Aileen Treacy.

Utterly final emergency squad: Tim O'Shea, Claire O'Malley, Simon Bento, Sara Hennessy, Rae Sibbitt, Roshni Devi, Benedict Nugroho Budi Priyanto, Dana Sensuse, Bill Ball, Claudine and the security team. A special thank you to Dave Perry for help at this time.

Thanks to the experimental subjects for their hard work.

Thanks to Peter Mennim, Anne, Gerry, Rachel, Christine, Lindsay, Noleen, Sinead, Aine, Steven Brown, Richard Wrigley and Penelope for injecting a little sanity.

Bill Harpe, Wendy Harpe, Stevie Smith, Martin Brems, Judy Bates, Sally Morris, Niel, Stephen Knox, Eddie Tagoe, Rie Toft Hansen and Vivi Mitts for alternative reality training and companionship.

Typhun, Beyhan and Dr. Torsun for initial encouragement.

Thanks and apologies to those who accidentally aren't thanked properly by name.

Big thanks to Lydia, Tina, Howard, Ralph, Claire, Phillipa, Andrew, Thomas, Jo, Ursula, Christine, Barbara, Sharon, Phillip (first novice Harmony Space user), David, Francis, Julie and all of the family.

The most heartfelt thanks of all are due to my wife Caroline. This research would not have been possible at all without her sustained love, support, sympathy, humour and help. Lots of love.

Thanks to my sons Simon and Peter for love, help, fun, education and for reminding me that there are things infinitely more important than research. Big hugs.

Thanks to my mother and the memory of my father. Both are fine musicians and exceptional people whom I love, admire and respect deeply. Lots of love.

This research was supported by ESRC studentship C00428525025. The OU research committee gave an equipment grant. IET gave a study extension. These are all gratefully acknowledged.

Contents

PART II HARMONY SPACE

Chapter 5 Harmony Space

Chapter 6 Adapting Longuet-Higgins' theory for Harmony Space

Chapter 7 Harmony Space using Balzano's Theory

PART III TOWARDS AN INTELLIGENT TUTOR FOR MUSIC COMPOSITION

Chapter 8 An architecture for a knowledge-based tutor for music composition

PART IV CONCLUSIONS

Chapter 13 Novices experiences with Harmony space

Chapter 14 Conclusions and further work

TABLE OF FIGURES

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 1: Introduction

The goal of this research is to find ways of using artificial intelligence to encourage and facilitate the composition of music by novices, particularly those without traditional skills. It will be argued that the research also has implications for intelligent tutoring systems in general, because the fostering of open-ended approaches is relevant to any sufficiently complex domain (e.g. engineering design, architectural design, etcetera).

1 The problem

The goal of music composition can vary from composer to composer and occasion to occasion and cannot normally be stated precisely beforehand. Almost the only goal that applies to music composition in general is "compose something interesting" (Levitt, 1985). The original focus of the research was on novices who have little or no formal education in music, and are teaching themselves, perhaps with the help of their friends, outside a formal educational setting. As Camilleri (1987) points out in a critical review of an earlier report on part of this research (Holland,1987a), this aim fits well with the general educational orientation of the institution which has supported the research (the Open University). Following this original aim, our illustrations and vocabulary are drawn predominantly from popular music and jazz, since it seems good educational practice to focus on music that is highly motivating to potential students. In fact, it turns out that most of the work is equally applicable to western tonal music in general.

2 General Approach

2.1 Introduction to the approach

Much of the current work in open-ended computer-based learning environments still draws its inspiration ultimately from work on the Logo project (Papert, 1980). Logo is a programming language that attempts to link, amongst other things, the idea of a function to a child's intuitions about her own body movement. The validity of the claims made for the educational effectiveness of Logo are now hotly contested (see, for example, Kurland, Pea, Clement and Mawby (1986)). However, it can at least be claimed with some justice that Logo bears a special relationship to mathematics due to the more or less central role played by the mathematical function in both mathematics (Kasner and Newman, 1968) and Logo (Hayes and Sutherland, 1989). In many other domains, such as music, it is far less clear that functional application is a particularly appropriate metaphor. When designing an exploratory microworld for some new domain, we advocate that metaphors or formalisms should be sought that lie at the heart of *that particular domain* in order to be given appropriate computational embodiment. This advocated approach, which we will illustrate in Part II contrasts with the more 'traditional' microworld approach of addressing a non-mathematical domain by adding superficial, domain-specific facilities to a Lisp-like programming language.

2.2 Developing a learning environment appealing to intuitions in several modalities

Another source of power in microworlds can arise from linking a domain metaphor to simple intuitions in many modalities (sound, gesture, animation etc.). The learning environment for harmony discussed in part II of the thesis links cognitive theories of harmony at a number of levels to intuitions in three modalities; visual, auditory and body movement. This learning environment uses the cognitive theories to make abstract harmonic concepts concrete and manipulable. This allows abstract concepts to be treated in a computer -based environment using all the advantages of direct manipulation methods (Shneiderman, 1982), (O'Malley, in press).

2.3 The need for new intelligent tutoring systems approaches in domains that demand creative behaviour

One of the most interesting consequence of the choice of domain from an intelligent tutoring systems (often abbreviated to 'ITS') point of view is that it is inappropriate for the system to be directly prescriptive or critical. There already exist a number of rule-based intelligent tutors for music composition which have a prescriptive and critical style. For example, Stephen Newcomb's (1985) LASSO is a tutor for a paradigm or style of music known as 16th Century Counterpoint, and Marilyn Taft Thomas's (1985) MacVoice is an intelligent tutor for a musical task known as voice leading - which is part of the related task of four-part harmonisation. Both of these pioneering musical Intelligent Tutoring Systems are prescriptive and critical and it is appropriate for their particular domains that they should use this approach. Why then could not a tutor for popular music composition be prescriptive in the same way? The answer is that 16th Century Counterpoint and four part harmonisation are genres that (as taught in schools) can be considered to have stable and known rules, goals, constraints and bugs.[1] In the case of popular music it might be possible to write a series of distinct expert components for a range of genres or paradigms of composition but this would not satisfactorily address the problem. This is particularly so in the case of self educators driven by personal motivation rather than a need to satisfy an examination or course requirement. The reasons for this are twofold. Firstly, genres in popular culture are still rapidly evolving from a rich set of background genres. A new song can easily combine features, rules and goals from genres not previously mixed. Designing a series of expert components that could perspicaciously characterise each genre or paradigm, be freely interchangeable and clearly exhibit inter-relationships between the styles, remains an unresolved research problem. Secondly, any new piece of music worth its salt invents new rules, goals and constraints specifically for that piece, and may consequently violate some rules normally expected of the genre. In this constantly evolving domain it is hard to see how any intelligent tutoring system could know enough to be prescriptive or critical without the risk of misunderstanding all the most innovative and interesting ideas.

If we are claiming that a prescriptive or critical style is not appropriate for the domain, does this mean that we are limited to a non-directed microworld or environment approach? To the contrary, we will argue that microworlds are useful for such tasks, but that they need complementing with other ITS approaches.

[1] This is an over-simplification because as we will note in chapter 4, work on formal models of musical processes has shown that text book rules are grossly inadequate - but the assumption of a stable, understood style is often made by educators.

2.4 A framework for individually relevant guidance

An important issue in the design of microworlds is the issue of providing individually relevant guidance. For example, the music microworld Music Logo (Bamberger, 1986), comes with a well-written manual containing graded work sheets. Anyone, novice or expert who worked through the book would learn a lot about music or see it anew from a fresh angle - but the issue of educationally exploiting individual differences in areas of motivation and ability is not considered. In a school context, this is easily addressed - the teacher can consider each individual's special interests and point them to different worksheets in different orders. She can suggest modified versions of the exercises to take into account likes and dislikes - so maximising motivation. In the case of a self-educator, opportunities are missed to greatly accelerate the learning process and fuel motivation.

2.5 Broad architecture of a tutor for music composition

Given that we cannot be overtly prescriptive in our domain but want to open up the possibility of tackling the issue of providing individually relevant guidance, how can we proceed, since neither the open-ended microworld approach nor the traditional interventionist tutoring approach is adequate for both purposes? In broad terms, our answer is to use a non-prescriptive knowledge-based tutor linked to one or more musical microworlds. The point of such an architecture is that the knowledge base could provide abstract knowledge and individually relevant guidance for exploring the microworlds, and the microworlds give principled visual and gestural alternative representations to auditory experience and predominantly textual, conceptual knowledge. This kind of approach coincides with the broad framework of Guided Discovery Learning, proposed by Elsom-Cook (1984) as a framework for combining the benefits of learning environments with those of interventionist tutors. Broadly speaking, Guided Discovery Learning methodology proposes the framework of providing a microworld in which discovery learning can take place, integrated with a tutor that is capable of executing teaching strategies that range from an almost completely non-interventionist approach to highly regimented strategies that give the tutor a much larger degree of control than the student. The place of the

ITS architecture of this project within the Guided Discovery Learning paradigm is discussed in Holland and Elsom-Cook (in press). Of course, Guided Discovery Learning does not in itself solve the problem of tutoring in open-ended domains as outlined earlier, but it does give a framework for combining discovery learning in a computer environment with the provision of individually relevant guidance.

Foundations for a new tutoring approach for open-ended domains are developed based on the exploration and transformation of case studies described in terms of chunks, styles and plans. This approach draws on a characterisation of creativity due to Johnson-Laird (1988). Interventionist teaching strategies for providing individually relevant guidance are not presented in the thesis. The new tutoring framework looks very well suited to supporting a wide range of active teaching strategies but the work is concentrated on finding solutions for the core problems that had to be solved to make a tutor possible at all: the devising of ways of characterising strategic considerations about chord sequences in ways that are both suitable to the modelling of creative processes and at the same time potentially communicable to novices. Although a fully implemented tutoring system is not described, programs to illustrate the feasibility of key parts of the new approach have been implemented. One of the main techniques developed is to describe chord sequences in terms of plans and manipulate the plans using a musical planner, which we will now discuss.

2.6 Complementing microworld representations with abstract knowledge

A major issue related to the issue of individually relevant guidance (and one that microworld approaches in general have only recently started to tackle) is the issue of providing *abstract knowledge* to complement direct experience and to guide exploration of the domain. The need for this can be illustrated in the case of music composition as follows. Music composition is about more than just familiarity with raw musical materials and the way they are perceived. When we listen to any piece of music, we are influenced by expectations based on other pieces of music, and by styles that we already know. To explore effectively ways of using the raw materials, we need some way of relating them to musical knowledge about existing works, styles and sensible musical behaviour in general. We introduce a new way of characterising abstract knowledge about harmonic sequences called "musical plans", to try to make explicit plausible reasons why existing pieces use raw materials in the way they do, and to characterise a range of common intelligent musical behaviors. Musical plans suggest various ways of organising raw musical materials to satisfy various common musical goals.

3 Relevance for ITS in domains other than music

Almost all work in intelligent tutoring systems up until now has been restricted to domains such as arithmetic, electronics, physics and programming in which there are clear-cut goals and in which it is relatively easy to test formally if a solution is right or not. In domains demanding creative behaviour, on the other hand, goals may not be precisely definable before the event, and "good" solutions may be very hard to identify formally. We will argue that existing ITS approaches are inadequate for such domains, and that finding ITS methodologies for dealing with open-ended domains of this nature is an important task for ITS as a whole. Tutoring creative behaviour cannot be dismissed as being relevant to only a handful of "unusual" domains; to the contrary, this issue is of central concern if intelligent tutoring systems are ever to progress beyond highly restricted domains into real world areas with less well-defined goals and methods. We can postulate that in any knowledge-rich domain with a sufficiently large search space, searching the space can be treated as an open-ended creative task. Any domain at all, even arithmetic, electronics, physics or programming may need to be considered creatively when a requirement arises to extend the domain or reconceptualise it. Indeed, some teachers and psychologists might argue that extension and reconceptualisation lie at the heart of effective learning in general. For these reasons, progress towards architectures for tutoring creative tasks has implications for the discipline as a whole. This project does not solve the problem of tutoring in a creative domain in general, but makes some first steps.

A second reason why intelligent tutoring systems for music composition may have important implications for the discipline as a whole, is that due to the demanding nature of the domain, research in ITS for music can act as a valuable *forcing function* for ITS methodology and architectures, introducing unusual perspectives and fresh techniques.

4 Timeliness

The research is timely in a practical sense due to the coincidence of a number of theoretical, educational and technological developments. AI researchers have taken an active interest in AI and music research for some time; examples include Simon (Simon and Sumner, 1968), Winograd (1968), Minsky (1981a), Longuet-Higgins (1962), Steedman (1984) and Johnson-Laird (1988). This work has been very broad, ranging from surprising strides in clarifying music theory (Longuet-Higgins, 1962) and ways of relating music to other areas of human cognition (Lerdahl and Jackendoff, 1983), to finding flexible ways of representing general time-embedded highly interacting parallel processes (Rodet and Cointe, 1984). More recently, there has been a strong upsurge of research in the psychology of music, particularly research influenced by the cognitive paradigm (Sloboda, 1985). This has been associated with a shift of theoretical attention from low level phenomena to higher level musical processes in rich contexts. This has provided a stream of powerful explanations and theories for musical phenomena that had previously defied *formal* analysis for centuries, e.g. aspects of ;

- harmony (Longuet-Higgins, 1962), (Balzano, 1980),
- rhythm (Longuet-Higgins and Lee, 1984),
- voice-leading (Wright, 1986), (McAdams and Bregman, 1985), (Cross, Howell and West, 1988), (Huron, 1988)
- hierarchical reduction (Lerdahl and Jackendoff, 1983),
- rubato (Todd, 1985).

Up until recently, much of this research had very little scope for application due to the rarity and expense of computer controllable musical instruments. The recent emergence of inexpensive high quality electronic musical instruments in the marketplace, coinciding with the recent establishment (Loy, 1985) of an industry standard (MIDI) for connecting computers to electronic musical instruments means that all of this work is now open for the first time to widespread practical application. These converging considerations mean that work done now in ITS and music is timely and capable of exploitation.

5 Roots of the research

Part II of the thesis was inspired directly by a single piece of work - Longuet-Higgins (1962). Some points of the interface design were influenced by Smith, Irby, Kimball, Verplank and Harslem (1982), and a later version of the interface was designed to take into account Balzano (1980). Part III of the thesis (the architecture for a tutoring system for music composition) was loosely inspired by work in machine learning due to Lenat (1982), but arose directly from consideration of the 'traditional' structure of intelligent tutoring systems, as described in, for example, Sleeman and Brown (1982), and simple observations about the nature of open-ended tasks. Inspiration for the constraint-based planner for music composition came from Levitt's work (1981, 1985).

6 Structure of thesis

The thesis is divided into 4 sections, together with some appendices.

Part I provides an introduction to the thesis as a whole. This part includes a selected review of previous work related to Artificial Intelligence, Education and Music.

Part II reviews two cognitive theories of harmony and shows how they can be used to design learning environments for tonal harmony. Some programs developed independently that relate to the work are discussed.

Part III outlines an architecture for tutoring music composition that combines the learning environment with sources of more abstract knowledge. A musical planner is the key abstract knowledge component of the architecture.

Part IV outlines a practical investigation carried out with Harmony Space, summarises the results of the thesis as a whole and looks at possibilities for further work.

Chapter 2: Early uses of computers in teaching music

In Part I of the thesis, we review the uses of computers in music education. Chapter 2 looks at early uses of computers in music education. Chapter 3 reviews recent trends in the musician machine interface. Chapter 4 is concerned with intelligent tutors for music. The final chapter includes pointers to related work in other fields.

In this chapter, we review early uses of computers in music education, with special attention to the needs of novices learning to compose. We review four representative and contrasted early approaches. The four approaches considered are: computer-assisted instruction (CAI) in music; music logos; tools for music education; and interactive graphical games. (Straightforward music editors and computer musical instruments are omitted, except for brief references.) The review is selective throughout. Applications may display characteristics of more than one division, so the categories cannot be taken as hard and fast. The structure of the chapter can be summed up as follows;

1 Computer-aided instruction in music education
2 Music logos
3 Tools for the student
4 Games
5 Conclusions on early approaches

We will begin by considering the first of the four early approaches, namely computer-aided instruction in music education.

1 Computer-aided instruction in music education

Historically, the first use of computers in teaching music or teaching any other