Digital Creativity
2003, Vol. 14, No. 2, pp. 67–73

1462-6268/02/1402-0067$16.00

Embedded intelligent music—or iHiFi the intelligent HiFi

John Eacott and Mark d'Inverno

University of Westminister, UK

john@informal.org, dinverm@wmin.ac.uk

Abstract

One consequence of current technological advances is that we can imbue everyday objects such as our fridges, ovens, fires, doors, thermostats, tvs and stereos with processing power. Not only will this enable them to make decisions based on their current state and model of the world about their current behaviour (the oven may automatically turn itself off when there is an excess of smoke) it will also allow communication with others to make decisions (the fridge may choose to defrost some vegetables when the phone communicates that there is a recent message from a friend who is, unexpectedly, intending to show for some dinner). This leads to the notion of *Ambient Intelligence* where intelligent computational entities are interwoven into the very fabric of our lives. We envisage a scenario where music devices can not only make intelligent, sympathetic decisions about sound generation in order to satiate the particular sound requirements of a user, but would also interact with other devices both musical and otherwise in a massively dynamic and unpredictable environment. We also envisage that there will be a massive shift in the pattern of behaviour relating to music consumption; a move from the *passive consumer* to the *active creator*. The emerging field of computer science which is concerned with building systems which are inherently distributed, dynamic, open and social in this sense is know as *intelligent agents*. Whilst we are quite clear that we are in the very early tentative stages of this work, we nevertheless set out some realistic medium-term achievable functionalities for what we call *Intelligent Embedded Music*, where intelligent, interactive music generation could continually and dynamically surround and sustain our day to day existence. We call this the *Intelligent HiFi (iHiFi)*.

Keywords: algorithmic and generative music, intelligent agents, responsive environments

1 Introduction

Have you ever put a CD or record on that you had been looking forward to, and been disappointed because you found that the music is too slushy, too hectic, too fast, too slow, too out of date, or simply too *by someone else*! Have you ever thought it would be great if you could adjust some aspects of your music collection so it was just right for the mood you are in right now? Have you ever wondered what a 70's punk album would sound like with a Garage beat? Have you even considered what it would be like if the HiFi could serve you up just what you wanted at that particular time? If it knew you, your mood, your day, your environment, your taste and so on perhaps it would be able to play *the* single piece which was just right for you at that time.

The function of a HiFi is to reproduce

Digital Creativity, Vol. 14 No. 2

pre-made artifacts and the function of a studio is to create these. An intelligent HiFi is an extension of the domestic HiFi and used in a domestic environment to provide music as a source of entertainment according to taste, mood and activity. Unlike a standard HiFi it has an intelligence that gives it a flexibility to modify and adapt the content of the music to provide a personalised or enriched experience for the user. In other words owners *would create not consume*, which would lead to a huge change in emphasis for the whole of music creation and consumption both for purchasers, musicians and the music industry.

We could scale-up this notion and envisage some other scenarios where an intelligent music system could be of benefit:

- playing the most therapeutically beneficial music in hospitals;
- creating sounds which elicit the most productivity in offices and working environments;
- soothing effect in nurserys or waiting rooms; and
- most excitement in a dance club.

With the recent great strides made in technological advancement such a vision is not unrealistic. However, coupled with the technology has been dramatic growth in a field of computer science concerned with building a new generation of software called *Intelligent Agents* (d'Inverno 2001a).

2 Intelligent agents

What we require of our iHiFi is that it is intelligent, autonomous, flexible, dynamic and proactive software system that controls sound generation in our home. This kind of software is typically referred to as an intelligent agent (d'Inverno 2001). For the last ten years researchers in this field have attempted to provide definitions of what, exactly, constitutes an intelligent agent but there is no common consensus (Luck 2000). However, there is agreement on the kind of properties we wish to see embodied in software which is situated in a dynamic environment. These have been referred to as *dimensions of agency*, and we briefly survey some of them below and how they might impact on the functionalities of our iHiFi.

Autonomy is taken to mean that an agent can act without the intervention of others. Autonomy is often regarded as a relative notion; the more autonomous an agent is, the less supervision it needs. If it is not autonomous, then it may always be under the direct supervision of another system or user (Wooldridge 2002). Our iHiFi would have a range of behaviour between non-autonomous (in simply playing a CD, for example, the machine has no ability to make any decisions) to autonomous (where the agent would compose or arrange a piece of music for they user when they sit down with the paper for the first time that evening).

Reflection characterises an agent able to reason about its behaviour. This would be necessary if it was clear that the music performed had not pleased it's user, for example, by playing something inappropriate at a social function. It would need to ascertain why it made certain decisions about what to perform in order to learn how to do the right thing next time.

Deliberation is the ability to manipulate symbolic representations of an environment. This is necessary because the iHiFi needs to have an internal model of the current situation and the intended effect (however, multi-layered) of the current performance (d'Inverno et al 1998, d'Inverno and Luck 1998). The agent can then see whether its model of the intended effect matches (in some way) its perception of the actual effect.

Reactivity is the ability of an agent to respond to changes in its environment within an appropriately small amount of time. For example, our iHiFi would need to respond to people entering or leaving a room, the posture

or position of someone in the room.

Pro-Activeness refers to agents having longer term agendas. As well as reacting to environmental changes such agents have goals that may not be satisfied immediately, and which will typically require planning. Our iHiFi certainly needs this quality - it should be able to pro-actively adapt or initiate music to elicit some predicted mood response (say) and subsequently judge the success of those decisions.

Goal-based Perception. At any one time there will be a massive number of percepts available to an agent. (Either through other software interfaces (agents, databases) or from sensors in the physical world.) In general, there will not be enough processing power to consider all the percepts, and so an agent should be able to determine which are the most relevant percepts to it at any one time *based on its current objectives*. For example, relevant percepts to the agent might be the posture and movement of its user in determining their current mood.

Social Awareness refers to the fact that agents should be able to model other agents (both human and computational), their goals, motivations, actions, perceptions, and the social structures that exist in any social environment (d'Inverno 2001). This is a key ingredient of the functionality of our iHiFi. First, it needs to model the social relationship of its users - a parent might have rank over the child about performance, a candlelit dinner will be very different from a business meeting at home and so on. Second, and less obvious is the need to model other iHiFi's and their users. It can then learn about the success or failure of attempts to satiate the sound needs of its users and others. one could even envisage scenarios in which iHiFi's worked collaboratively on a piece. Boy meets girl and invites her back to dinner - their respective iHiFis construct the perfect symphony for them. If things work out that night then perhaps further downs the line our iHiFi might be instructed to choose or even compose their wedding music!

Communication. There have been many attempts to define a language for communication between agents. The two most important initiatives are KIF (Knowledge Interchange Format www.logic.stanford.edu/kif/kif.html) and FIPA (Foundation for Intelligent Physical Agents - www.fipa.org). Ideally we would require a common language for the representation of different music formats (audio, midi, algorithmic). What is required, and what is needed in the algorithmic music community, is a movement towards a language to describe the form, structure and process of algorithmic music.

Mobility can refer to properties of both software and physical agents. Software mobile agents are programs that can be dispatched from a computer and transported to a remote machine for execution. How might this be beneficial to our system? It might be amusing to have someone else's intelligent agent come and control your iHiFi for the day just to see how bad their tastes are!

Any reader interested in more information on agent technology the reader is invited to consider a few of the following references: (d'Inverno 2001, Jennings 2000, - Wooldridge 2002, Luck 2003) or look at the excellent AgentLink home page (www.agentlink.org). The major conference is called The Joint International Conference on Autonomous Agents and Multi-Agent Systems and is held this year in Melbourne, Australia (www.aamas-conference.org).

3 Algorithmic music

Although our iHiFi might conceivably reads music in any form, the most flexible and determinable representation would be an algorithmic one. An audio signal is fixed structure in time and space; an algorithm has an intrinsic logical, mathematical structure ideal for intelligent manipulation.

If music is provided in algorithmic form

the iHiFi is able to access not only different layers of the music but the thoughts and ideas, which make it. By adjusting these *thoughts* that exist inside the iHiFi in the form of intelligent algorithms, variations, subtle or radical, on the original music can be achieved.

The practice of creating generative algorithms for the creation of music has developed as has the level of participation, awareness and use of generative music has increased. Software products such as

Max/MSP (Zicarelli 1990), SuperCollider (Mccartney 2002)and PD (Puckette 2002) have emerged which provide powerful and usable environments for musicians with varying degrees of computer coding experience to use.

Although all of those packages are marketed as general purpose sound tools a large proportion of users exploit their potential for non-deterministic or generative processes. This use is now finding exposure by a wide range of musicians from architectural sound designers, installation designers, electro-acoustic composers to producers of dance music. Generative sound is becoming almost a de-facto standard in the creation of music by a significant number of musicians and DJs. Subsequently, a new phenomenon has emerged - generative artists with record deals! Every individual creating a generative system will have their own set of approaches and values, which are expressed through its algorithms, the construction of which is probably the key concern in creating algorithmic music (Eacott 1999). Recently there have been significant advances in the development of algorithmic approaches that create and manipulate music in satisfying ways across a whole range of forms (Collins 2002).

In many of the early examples it was assumed that algorithms should be applied primarily to pitch and rhythm (at phrase level), which are, to most listeners, the primary features of music. They are also aspects, which Snyder would describe as *basic level phenomena* being of a suitable magnitude, neither too large and unwieldy or too minute and detailed to contemplate (Snyder 2000). Recently there has been a gradual shift towards using generative processes on other aspects of sound including timbre, form, spatialization and even less quantifiable aspects like energy or intensity. (Without going too deeply into an ethno-musicological issues, for many people, when we think of musical 'composition' we think primarily about these three areas. Interestingly perhaps, these parameters are less significant in contemporary dance music culture. Although melody, and harmony are important, they take an equal status with timbre, style and energy. Since the emergence of breakbeat culture in the 90s there has been a significant increase in the number of recordings which have little or no tonal elements other than those provided from borrowed samples.)

... by adjusting these thoughts that exist inside the iHiFi (...) variations, subtle or radical, on the original music can be achieved

There are good reasons for this shift. From a cognitive or memory perspective, the single element that most listeners retain about a piece is melody, whereas the reason we get to *feel* the power and emotion of a piece is largely because of its arrangement and sonic qualities. Most, if not all of us are considerably less able to retain these aspects in our memory than we

are of melody. We are, in our ability to sing, whistle, clap and hit things, provided with excellent means to reproduce melodic and rhythmic components ourselves and have significantly less ability to recreate harmonic and timbral elements.

If we take for example a generative piece which is designed to be heard repeatedly, like a track on MORPHEUS for example, we should decide whether or not it is desirable that the listener learns and recognises its main elements (Eacott 2001a, Eacott 2001b). This issue will determine how, if at all, we use a generative process to determine primary identifying features.

However, it should be clear that the user would (in most cases) not wish to interact at the level of the algorithm and we would not consider that this would not be an appropriate functionality. However, what it does need to do, is interpret all the information about the user and their environment and somehow translate this into mathematical operations. This is the tricky bit. In the next two sections we consider the kinds of functionality that the iHiFi can provide at the level of a *specific* request and then at the more *intelligent* level where the software acts autonomously and proactively based on its model of the environment and its users' desires.

4 iHiFi in practice

Clearly there isn't any content in the most desirable form (i.e. algorithmic) available for the iHiFi yet. Composers and producers would have to start producing special versions of their music for iHiFi format. This format would use algorithmic constructions of the music. Once formed, aspects of the algorithmic content could be varied. We can easily list some examples of how this might work in practice where the user can give very *specific* instructions.

- Change the tempo.
- Change the tonality.
- Add or remove instruments and vocals. (Similar to the interactive CDroms from the 1990s)
- Choose from a library of preset styles and download style updates from the web. With algorithmic representations this would not imit the user to an either/or decision, it would be possible (say) to specify Trance with just a hint of Drum'n'Bass.
- Change the arrangement with the help of physical modelling algorithms which are able to simulate the acoustic properties and behaviours of acoustic instruments to symphony orchestra or solo classical guitar (Roads 1999).
- Re-harmonise a favourite piece using a range of emotional parameters and measures.
- If you have a favourite record that you have played to death and love every aspect of it but are now fed up with the melody then a e-compose melody feature would generate another track maintaining many of the qualities of your original.
- Your tracks would automatically incorporate style updates to keep up to date.
- By becoming aware of other elements of your music collection iHiFi would be able to create a collection of *optimised music*, which embraces the range of your taste.

The intelligent agent would carefully monitor these operations in an attempt to learn what it is that the user likes: which operations does the user perform to which kinds of music at what times on which day in which season and with what people? Once the agents have built up a model of the user it can then begin to generate sound and music *intelligently*.

5 iHiFi in intelligent ambient mode

All the above features are appropriate when a user is able and willing to specify quite precisely what it is they want. However, many if not most of us probably don't want to think too hard about it, we simply want the machine to make

educated selections about what is right for the occasion. Of course the operations described above could be useful if the iHiFi gets it wrong. The user could then massage the generated sound using the kinds of specific commands we have mentioned above (and over time we would expect our intelligent agent to get it wrong less and need less intervention). Eventually, as is not untypical in human nature, we want the things we want without even noticing it was the thing we wanted.

The model the agent has of the environment and user is key. Specific issues such time of day, day of week, season would be required as well as more general, context dependent information that could be gathered through various interfaces such as user mood, other people in environment, their social relationships, the type and form of their interaction and so on. It would not be difficult for sensors to recognise certain guests, understand their preferences, realise your social aspirations and the cocktails you are mixing, and play the most perfect, sublime dinner jazz in your favourite key.

In order to make iHiFi there would have to be considerable research and development in creating a versatile, comprehensive yet easily usable format. The experience of creating MORPHEUS has given us some useful pointers however.

It's now well known how to specify music in terms of arrays of sequenced information. This can exist as midi sequences but a format like SuperCollider provides a more powerful and general framework (Mccartney yyyy). Since algorithmic music is programmed by humans it makes sense that the main notions around which we organise algorithms relate to *middle level* abstractions of music. These are abstractions that are neither to general nor too specific. They are, on a basic cognitive level, the main way musicians organise music (Snyder 2000). Historically, they are also tasks that, in larger productions, would be delegated to differ basic algorithmic song content could take the form of:

- melodic and rhythmic arrays for compositional content (primary melodic),
- rhythmic / lyrical / form features,
- sequence / arrangement / structural information stylistic *signatures* (Cope 1996),
- timbral generators in the form of electronic synthesis or physical modelling,
- mixing and spatialization,
- mood and texture alterations, and
- organic development of themes through loose interaction between agent and user (one might imagine a conversation where a theme is suggested and user counters with either specific or behaviour responses which alter this theme a little and perhaps suggest to the agent ways in which it might develop the theme even more (and so on)).

6 Conclusions

In this paper we have proposed some ideas for a music system based on advances in algorithmic and generative music and the field of intelligent agents and multi-agent systems. Such a project would also require input from a whole spectrum of different disciplines including behavioural psychologists, musicologists, computer scientists, cognitive scientists, sound architecture designers, interactive specialists and almost certainly some good lawyers!

Acknowledgements

Thanks to Michael Luck from Southampton University who has developed many of the ideas about agents presented here with the second author. We would also like to thank Philippe Naughton and Jason Knight of Thus Music Ltd., who have been invaluable in discussing possible directions of the project described in this paper. Thanks also to Nick Collins, Fredrik Olofsson, and all on the IDM Summer School

and the SC-WMIN mailing list. With thanks for the support by the University of Westminster who have enabled us to develop some of these ideas working on a project called the Intelligent Street. Please see our web pages for details.

References

Collins, N. (2002) Infinite length pieces: A user's guide in S. Hawkins (Ed.) *Proceedings of the Maxis Festival*, Sheffield Hallam University.

Cope, D. (1996) *Experiments in musical intelligence*, A-R Editions, Incorporated.

d'Inverno, M., D. Kinny, M. Luck, and M. Wooldridge (1998) A formal specification of dMARS. In *Intelligent agents IV: Proceedings of the Fourth International Workshop on Agent Theories, Architectures and Languages*, vol. 1365, pp 155-176. Springer.

d'Inverno, M. and M. Luck (1998) Engineering AgentSpeak(L), *Logic and Computation*, 8(3), 1998.

d'Inverno and M. Luck (2001a) Multi-agent systems research into the 21st century, *The Knowledge Engineering Review*, Cambridge University Press, 16(3), pp. 271-275.

d'Inverno, M. and M. Luck (2001b) *Understanding Agent Systems*, Springer, 2001.

Eacott, J. (1999) Generative music composition in practice, *Generative Art*, 1999.

Eacott, J. (2001a) Morpheus - emergent music. In S. Soddu (ed.) *Proceedings of the Generative Music Symposium*, Milan Polytechnic.

Eacott, J. (2001b) *Morpheus CDRom* (ed.) Mushi Mushi, London, 2001.

Jennings, N. (2000) On agent-based software engineering. *Artificial Intelligence*.

Luck, M. and M. d'Inverno (2000) A conceptual framework for agent definition and development, *Computer Journal*, 44(1) 1-20.

Luck, M., P. McBurney and C. Preist (2003) *Agent technology: Enabling next generation computing*. AgentLink.

McCartney, J. (2002) *SuperCollider audio synthesis environment*, http://www.audiosynth.com (accessed 06/06/2002)

Puckette, M. (2002) *PD*, http://crca.ucsd.edu/~msp/software.html (accessed 06/06/2002)

Roads, C. (1999) *The computer music tutorial*. MIT Press.

Snyder, B. (2000) *Music and memory*. MIT Press.

Wooldridge, M. (2002) *An introduction to multi-agent systems*. Wiley.

Zicarelli, D. (1990) *Max/MSP, Cycling 74/Ircam*, 1990-2001.

Mark d'Inverno is a professor of computer science at University of Westminister in London. His research focuses on the application of formal methods in providing theoretical and foundational formal frameworks in emergent technological fields of Computer Science. In particular, formal specification techniques are applied to describing a library of agent systems, languages and theories which can be used in the principled development of software. There is a new book from Springer on this subject co-authored with Michael Luck at Southampton University.