A Parsimonious Cognitive Architecture for Human-Computer Interactive Musical Free Improvisation

Adam Linson, Chris Dobbyn and Robin Laney

Faculty of Mathematics, Computing and Technology
Department of Computing
The Open University, Milton Keynes, UK

Abstract. This paper presents some of the historical and theoretical foundations for a new cognitive architecture for human-computer interactive musical free improvisation. The architecture is parsimonious in that it has no access to musical knowledge and no domain-general subsystems, such as memory or representational abilities. The paper first describes some of the features and limitations of the architecture. It then illustrates how this architecture draws on insights from cybernetics, artificial life, artificial intelligence and ecological theory by situating it within a historical context. The context presented consists of a few key developments in the history of biologically-inspired robotics, followed by an indication of how they connect to James Gibson's ecological theory. Finally, it describes how a recent approach to musicology informed by ecological theory bears on an implementation of this architecture.

1 Introduction

This paper presents some of the historical and theoretical foundations for a new cognitive architecture for human-computer interactive musical free improvisation. The paper first describes some of the features and limitations of this architecture. It then situates the architecture within a historical context by presenting a few key developments in the history of biologically-inspired robotics. This is followed by an indication of how these developments connect to James Gibson's ecological theory. Finally, it describes how a recent approach to musicology informed by ecological theory bears on an implementation of this architecture.

2 Features and Limitations of the Architecture

Although the architecture seeks neither to model human-internal cognitive structure, nor to closely emulate human musical behaviour, it does aim to provide musical behaviour that supports engaging improvisational interaction with an expert human performer. This is achieved with domain-specific low-level subsystems for dealing separately with pitch, loudness and timing, respectively, in both input and output. Subsystem output is combined into discrete monophonic note

streams, which are in turn combined into continuous polyphonic note streams. The resulting musical behaviour is open-ended and highly responsive to a collaborative human co-performer. The architecture allows for reciprocal influence between human and computer, and it ultimately facilitates engaging musical collaboration, according to preliminary findings from on-going empirical research. The architecture is parsimonious in that it has no access to musical knowledge and no domain-general subsystems, such as long-term memory or representational abilities (for more on the relation between cognitive architecture and knowledge, see Laird (2012); for more on the interaction between domain-general and domain-specific cognitive abilities, see Gerrans and Stone (2008)).

Two key limitations of the parsimonious architecture, as implemented in a current prototype agent, Adam Linson's *Odessa*, are counterbalanced as described below. The first key limitation, that the agent lacks any knowledge, is addressed by drawing on principles of the 'subsumption architecture' (Brooks 1999). In an agent with a subsumption architecture, intelligent behaviour can arise from the interaction dynamics between agent and environment using neither representation nor reasoning, but rather, via hardwired "competing" behaviours, organised into layers. In the case of *Odessa*, the agent's layers are organised in order to give it the appearance of an intentional agent to a human co-performer (for more on the role of intentional agency in this context, see Linson, et al. (2012a)). Although the agent does not merely mirror the human input, it does, at times, reorganise some of the input elements, providing a perceptible degree of resemblance. While stored knowledge could be expected to enhance the agent's musicality, the current implementation serves as an existence proof that a musical agent can function effectively without musical knowledge.

The second key limitation is that, without long-term memory, the agent lacks the internal capacity for developing the music in the course of a performance. However, this does not mean that no musical development takes place in a given human–computer collaboration. Although here, too, an internal mechanism for musical development could be expected to enhance the agent's performance, in the absence of one, the agent's continuous short-term responsiveness serves to offload long-term musical development onto the human co-performer. This offloading takes place while nonetheless preserving a constant (non-hierarchical) reciprocal influence between human and computer, made possible by the fact that, within the dynamic musical collaboration, the human and computer performers function as tightly coupled subsystems. Thus, a performance may exhibit musical development due to the reflective practice of the expert human improvisor, despite the computer agent's inability to handle long-term dependencies.

This parsimonious architecture, implemented in *Odessa*, draws on insights from cybernetics, artificial life, artificial intelligence and ecological theory. The remainder of this paper presents a few key developments in the history of biologically-inspired robotics. It then connects these developments to James Gibson's ecological theory. Finally, it describes how a recent approach to musicology informed by ecological theory bears on the design of *Odessa*.

3   Cybernetics, Nouvelle AI and Ecological Psychology

According to Alan Turing's view, as presented in his well-known account of the so-called "Turing test", intelligence can be assessed behaviourally in an unstructured, open-ended interactive context (Turing 1950). This view has close parallels to the work of cybernetics pioneer W. Grey Walter. Walter is perhaps best known for his autonomous robotic tortoises built in 1948-49 (Holland 2003), and first published about in the same year as Turing's 1950 paper (Walter 1950). Walter's ideas were partly based on his own empirical research on the human brain, and partly on the work of his colleague and fellow cybernetics pioneer, W. Ross Ashby. Ashby proposed a dynamic systems model of a brain, based on the premise that a small collection of simple subsystems, tightly coupled but capable of temporary independence, could produce a wide range of complex behaviours (Ashby 1960). He maintained that despite the simplicity of his model, nothing prevented it from being theoretically scaled up to human-level behaviour. Although Walter's robotic tortoises were far from human-level intelligence, they did exhibit basic lifelike behaviour through their interactions with each other, with humans and with their physical environment.

Several decades later, Rodney Brooks developed the 'subsumption architecture', a concept he applied to building insect-like robots and other biologically analogous models (Brooks 1999). For Brooks, like Walter before him, the basic premise of his robot design is that a small number of simple interconnected components can lead to complex emergent behaviour. The mechanically embodied robots built with the subsumption architecture were rapidly prototyped with "real world" empirical testing. This meant that at an early stage of development, they were deployed, for example, in crowded offices, and could be iteratively adjusted depending on how well they negotiated such everyday environments. This approach stood in contrast to the elaborate pre-deployment development of earlier GOFAI robots, such as the Stanford Research Institute's *Shakey*, and their evaluation in carefully constructed specialised environments.

Another hallmark of the approach to robotics evident in both Walter and Brooks is that, rather than focusing solely on an agent's internal construction, they emphasise its adaptive behaviour in terms of feedback loops between perception and action, or, more broadly, between the agent and its environment. This theoretical orientation bears a strong resemblance to that of ecological psychology, a field which largely emerged from the work of James Gibson. Ecological psychology fundamentally concerns itself with the tight coupling of perception and action, and of agent and environment, in the context of adaptive behaviour; Gibson is also the source of the well-established notion of *affordances* (Gibson 1977, 1979). Simply put, an affordance is a feature of an environment that presents a possible action to an agent, relative to the agent's capabilities.

4   Traditional, Critical and Ecological Musicology

With respect to the study of music, traditional musicological theory generally regards classical western notation as sufficient to grasp the salient aspects of

music. More recently, however, critical approaches to musicology have served to identify, for example, aspects of musical improvisation that previously escaped traditional musicological theory. Significantly, it has been shown that there are psychosocial dynamics at work in improvisation that are not reducible to musical structure defined solely in terms of traditional western musical notation (Sansom 1997). This definition of structure is not only the primary object of traditional musicological analysis, but also the predominant focus of computer music research. In the critical mode of analysis, musical structure is not denied, but rather, it is viewed as a vehicle for socioculturally situated and context-dependent meaning production and exchange.

Drawing on the work of Gibson, Eric Clarke has advanced a new strain of musicology and music psychology research which he describes as "an ecological approach to the perception of musical meaning" (Clarke 2005). In this context, he addresses improvised music, pointing out how performers of collective improvisation forcefully demonstrate the ecological link between perception, action and meaning (see Clarke (2005), p. 152-154). Additional writings on improvisation within critical musicology, though not explicitly ecological, can be viewed in terms of this perception-action-meaning linkage. For example, in the context of jazz improvisation, Ingrid Monson examines performative musical irony, looking at what can be understood as high-level affordances for creative meaning production (Monson 1994). In another account of jazz improvisation, Benjamin Givan examines ways in which the low-level physicality of performers' bodies and instruments can become central affordance-rich vehicles that facilitate creative exploration (Givan 2009). And, in a direct consideration of free improvisation, Matthew Sansom focuses on general interconnections between musical structure and social dynamics, highlighting the ways in which nominally extramusical structure can afford new directions in underlying musical structure, as it unfolds in real time (Sansom 1997). These qualitative approaches to analysis identify salient features of music that remain invisible (or, at best, marginal) to traditional musicology, computational musicology and most quantitative approaches to musical perception (see Linson, et al. (2012b)).

5   Deployment context

By design, *Odessa* does not impose limitations on the behavioural and sonic complexity of the human performance. While its output may not conform to the expectations of traditional musicological analysis, its open-ended complex interactive musical behaviour provides affordance-rich material for an expert human improvisor to engage with in real time. In these respects, the design draws, in part, on techniques pioneered by Walter and Brooks, as well as on insights from Clarke's ecological musicology.

On one hand, considerations of agent–environment interaction apply here, with the idea that an interactive agent for musical improvisation should be able to cope with the real-time dynamics of its unstructured musical environment, like the physical environments of Walter's and Brooks' robots. On the other hand,

for human–computer musical free improvisation, the interaction model relates more to Turing's test, because the musical environment – like a conversation – consists in an open-ended agent–agent interaction (see Lewis 2007). In designing an agent for freely improvised music, both modes of interaction are relevant.

In such music, even in the limited case of duets, there are simultaneous sonic-musical streams between the players in both directions. Everything in these streams is heard by each co-performer either as a potential response to what was played, or as a potential affordance for a new response – or as both of these at once. Musical streams from each agent – whether human or computer – may sometimes support one another, sometimes provide contrast or resistance, and sometimes steer the interaction in a new direction.

6 Conclusion

Contrary to a frequently encountered assumption, it is not necessary for a computer co-performer to be mistaken for a human in order for the human–computer interaction to be engaging. An expert human improvisor can collaborate with an agent that has a unique way of making music, thus producing a shared musical outcome. The computer's musical output and, indeed, even the collaborative musical result, are an entirely separate matter from the unfolding dynamics of the collaborative mutual engagement. The on-going empirical studies of *Odessa*'s collaborative abilities seek to demonstrate the effectiveness of a parsimonious cognitive architecture in this context.

7 Acknowledgments

References

Ashby, W.R.: Design for a brain: The origin of adaptive behaviour. Chapman & Hall (1960)

Brooks, R.A.: Cambrian intelligence: The early history of the new AI. MIT Press (1999)

Clarke, E.F.: Ways Of listening: An ecological approach to the perception of musical meaning. Oxford University Press (2005)

Gerrans, P., Stone, V.: Generous or parsimonious cognitive architecture? Cognitive neuroscience and theory of mind. The British Journal for the Philosophy of Science **59**(2) (2008) 121–141

Gibson, J.J.: The theory of affordances. In Shaw, R., Bransford, J., eds.: Perceiving, acting, and knowing: Toward an ecological psychology. Lawrence Erlbaum (1977) 67–82

Gibson, J.J.: The ecological approach to visual perception. Houghton, Mifflin and Company (1979)

Givan, B.: Thelonious Monk's pianism. Journal of Musicology **26**(3) (2009) 404–442

Holland, O.: Exploration and high adventure: The legacy of Grey Walter. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **361**(1811) (2003) 2085–2121

Laird, J.: The Soar cognitive architecture. MIT Press (2012)

Lewis, G.: Improvising tomorrow's bodies: The politics of transduction. E-misférica **4.2** (2007)

Linson, A., Dobbyn, C., Laney, R.: Interactive intelligence: Behaviour-based AI, musical HCI and the Turing test. In Müller, V., Ayesh, A., eds.: Revisiting Turing and his test: Comprehensiveness, qualia, and the real world (AISB/IACAP symposium, Alan Turing Year 2012). (2012a) 16–19

Linson, A., Dobbyn, C., Laney, R.: Critical issues in evaluating freely improvising interactive music systems. In Maher, M., Hammond, K., Pease, A., Pérez y Pérez, R., Ventura, D., Wiggins, G., eds.: Proceedings of the Third International Conference on Computational Creativity. (2012b) 145–149

Monson, I.: Doubleness and jazz improvisation: Irony, parody, and ethnomusicology. Critical Inquiry **20**(2) (1994) 283–313

Sansom, M.J.: Musical meaning: A qualitative investigation of free improvisation. PhD thesis, University of Sheffield (1997)

Turing, A.M.: Computing machinery and intelligence. Mind **59**(236) (1950) 433–460

Walter, W.G.: An imitation of life. Scientific American **182**(5) (1950) 42–45