How to make a machine listen?*

Henkjan Honing

Music Cognition Group, ILLC, Universiteit van Amsterdam

honing@uva.nl ; http://www.hum.uva.nl/mmm/

Imagine what it would look like: A machine that is able to listen and react in a human and musical way... What should such a machine know, what should it listen for, how could it respond, and how can we compare and evaluate them?

The design of a "listening machine" that embodies the musical and listening skills common to most humans turns out to be a full-fledged research program and is part of the scientific enterprise generally known as music cognition. Like language, music is a uniquely human capacity that arguably played a central role in the origins of human cognition, and as such an important topic in the cognitive sciences. In this field of research computational modeling (formalizing a theory in the form of a computer program and relating it to human behaviour) is an influential methodology that has contributed to a further understanding of music as a process in which the performer and the listener play a central role.

This focus on music performance and perception not only enriched music research itself - musicology typically restricts its focus to musical scores in their historical and cultural context -, it also started to influence the main issues addressed in other areas of scientific research; It slowly diminished the "trade deficit" that music research built up over its existence in the sense that it borrowed more from other disciplines, such as psychology and computer science, than it contributed to them. Until recently music was only a minor topic in psychology reference books, and only occasionally is it a topic in high-impact journals like *Nature* or *Science*. However, in recent years several disciplines, ranging from the humanities to the social and natural sciences, show a growing interest in the scientific study of music. Music was even advocated by Robert Zatorre as "the food of neuroscience" (*Nature*, 17 March 2005), as it plays a central role in several recent research projects in Europe and the

* Published as Honing, H. (2006). How to make a machine listen? *Intensive Science*, 21. Sony Computer Science Laboratory: Paris.

US ranging in topics from motor-skill to memory and emergent cognition to emotion.

SCSL should also be credited for acknowledging music as a more than relevant and important domain, giving it a prominent place in their research program next to language and robotics. Still, music research is still a relatively lagging behind as compared to fields like linguistics, simply due to the amount research effort spent on it. A (slightly dubious) advantage of this situation is, however, that new insights or discoveries are relatively easy to obtain in this domain. This might well make music cognition an attractive domain for young scientists (and especially PhD students).

Furthermore, research at SCSL reflects a general move in computer science from artificial to cognitive intelligence. The Continuator, developed by François Pachet, is a clear example of an artificial listening machine that is entering the cognitive domain. While the technology that is used (so-called Markov chains) is relatively well known and has formed the bases of many successful automated learning systems, including speech and music, the additional representational and biasing mechanisms make it a system musicians find engaging and challenging. Furthermore, a recent (however informal) Limited Turing Test revealed that is was virtually impossible to distinguish between a real musician and the Continuator when both took alternating turns.

Still, the musical mind is of course more than just an advanced Markov system. Human cognition consists next to perception and different kinds of memory, also of attention, expectation and mechanisms that allow for cognition to emerge. The future of listening machines, like the Continuator, is to try and incorporate these mechanisms, for as far as we currently understand them, and to make musical machines that can relate to musicians in an even more engaging way, As such being considered a serious musical partner on stage, at home or virtually via the internet. Besides a scientific challenge this is also a technological challenge that will have a major impact on a large variety of music-related technologies of which the possibilities and implications are yet unforeseen. Still, the Continuator demonstrates that a large variety of users (ranging from young children to professional musicians) enjoy to be involved in music facilitated by these new technologies.

IntensiveScience 21

DUETS WITH THE CONTINUATOR

Children interacting with the Continuator. The joy experienced during play is enormously seductive.

How to Make a Machine Listen?

[Henkjan *Honing*]

Imagine what it would look like: a machine that is able to listen and react in a human and musical way... What should such a machine know, what should it listen for and how could it respond? How can we compare and evaluate these kinds of machines?

The design of a "listening machine" that embodies the musical and listening skills common to most humans turns out to be a full-fledged research program and is part of the scientific enterprise generally known as "music cognition". Like language, music is a uniquely human capacity that arguably played a central role in the origins of human cognition and is as such an important topic in the cognitive sciences. In this field of research, computational modelling (formalising a theory in the form of a computer program and relating it to human behaviour) is an influencial methodology that has contributed to a further understanding of music as a process in which the performer and the listener play a central role.

This focus on music performance and perception not only enriched music research itself – musicology typically restricts its focus to musical scores in their historical and cultural context –, it also started to influence the main issues addressed in other areas of scientific research. It slowly diminished the "tacit deficit" that music research built up over its existence in the sense that it borrowed more from other disciplines, such as psychology and computer science, than it contributed to them. Until recently music was only a minor topic in psychology reference books, and only occasionally is it a topic in high-impact journals like *Nature* or *Science*. However, in recent years several disciplines, ranging from the humanities to social and natural sciences, have shown a growing interest in the scientific study of music. Music was even advocated by Robert Zatorre as "the food of neuroscience" (*Nature*, 17 March 2005), as it plays a central role in several recent research projects in Europe and the US ranging in topics from motor-skill to memory and emergent cognition to emotion.

Sony CSL should be credited for acknowledging music as a more than relevant and important domain, giving it a prominent place in their research program next to language and robotics. Still, music research is still relatively lagging behind when compared to a field like linguistics, simply due to the amount of research effort spent on it. A (slightly dubious) advantage of this situation is, however, that new insights or discoveries are relatively easy to obtain in this domain. This might well make music cognition an attractive domain for young scientists (and especially PhD students).

Furthermore, research at Sony CSL reflects a general move in computer science from artificial to cognitive intelligence. The Continuator, developed by François Pachet and his team, is a clear example of an artificial listening machine that is entering the cognitive domain. While the technology that is used (so-called Markov chains) is relatively well-known and formed the basis for many successful automated learning systems, including ones for speech and music, the additional representational and biasing mechanisms make it a system that musicians find engaging and challenging. Furthermore, a recent (however informal and limited) Turing Test revealed that it was virtually impossible to distinguish between a real musician and the Continuator when both took alternating turns.

Still, the musical mind is of course more than just an advanced Markov system. Human cognition consists next to perception and different kinds of memory, also of attention, expectation and mechanisms that allow for cognition to emerge. The future of listening machines, like the Continuator, is to try and incorporate these mechanisms, for so far as we currently understand them, and to make musical machines that can relate to musicians in an even more engaging way. As such they could be considered as a serious musical partner on stage, at home or on the internet. Besides a scientific issue this is also a technological challenge that will have a major impact on a large variety of music-related technologies of which the possibilities and implications are yet unforeseen. Still, the Continuator demonstrates that a large variety of users (ranging from young children to professional musicians) enjoy to be involved in music facilitated by these new technologies. ▪

Henkjan Honing *heads the Music Cognition Group (MCG), part of the Department of Musicology, the Institute for Logic, Language and Computation (ILLC), at the University of Amsterdam (UvA), and conducts research in music cognition using theoretical, empirical and computational methods.*

Fragment of exhibition cataloque published for the occasion of the exhibition **Intensive Science**, *held at La Maison Rouge in Paris on 6-7 October 2006 to celebrate the tenth anniversary of the Sony Computer Science Laboratory in Paris.*