The Chinese Music Box

“Music has no subject beyond the combinations of notes we hear, for music speaks not only by means of sounds, it speaks nothing but sound.” — Eduard Hanslick

Formalists, such as the philosopher Roger Scruton (1976) and the composer Igor Stravinsky follow Hanslick. They agree that (purely instrumental) music strips all the representational properties from sound. In contrast to portraits and landscape painting, one can understand music without understanding what it is about. Listening to music is like looking through a kaleidoscope.

These austere opponents of representationalism focus on the internal structure of the sounds. Their aesthetic detachment is unsurpassed. Music is closed off from the world – offering a perfect refuge.

This structural approach to musical understanding is analogous to computationalism about speech comprehension. According to the computationalist, a suitably programmed computer can understand natural language. This principle is assumed by the Turing test: a computer can think if an interrogator cannot tell, just by conversing with it by teletype, whether it is a human being or a computer.

John Searle argues that Alan Turing’s test is too broad:

Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in
the room to pass the Turing Test for understanding Chinese but he does not understand a word of Chinese. (Searle 1999, 115)

Since the man in the room does not understand Chinese by virtue of his capacity to appropriately manipulate symbols, neither can any computer. A computer is just a physical instantiation of a formal system – a device for changing input strings into output strings. The strings are meaningless to the computer. It is all syntax and no semantics.

Now consider a musical analogue of the Turing test: If the computer’s performance in an interrogative duet is indistinguishable from that of a human musician, then the computer understands music.

The best-known example of an interrogative duet is the “Dueling Banjos” scene in the movie “Deliverance”. A guitarist visiting the Appalachian backwoods notices a blank-faced retarded boy holding a banjo. He tests whether the boy can play the banjo by strumming some rudimentary chords on his guitar. The boy feebly replicates what was played. The guitarist tries more advanced melodies. The boy keeps up. As the tempo accelerates, the boy’s play overtakes the guitarist’s. “I’m lost!” confesses the delighted guitarist. But when he extends a hand of congratulation, the boy shirks away. The idiot savant does not understand the meaning of a handshake.

A computer might also prevail over an inquisitive musician. Would that superior performance show that computer understood the music?

Searle might be tempted to answer no by adapting his linguistic thought experiment to music. Would the man in the Chinese Music Box be merely simulating understanding of music? The man’s instruction manual resembles musical notation: When you hear sound sequence A, B, C, respond with X, Y, Z. As Searle emphasizes in debates about the Chinese Room, the man may have memorized the manual. According to the formalist, syntactic facility is sufficient for musical understanding. The man might be missing extra-musical associations, such as homages to past composers (as when Charlie Parker quotes Stravinsky's Rite of Spring in his jazz solo on ‘Repetition’). But this type of “musical” appreciation is an accretion, not part of the music itself.

Or so says the formalist. My own opinion is that formalism is a deep insight that goes slightly too far, neglecting marginal semantic phenomena such as leitmotifs and
onomatopoeia (Kivy 1984). Yet, I am almost as opposed to crediting the computer with musical understanding as language understanding. One explanation of this disproportionate reaction is that musical understanding is perceptual. A computer cannot hear a sound as a musical sound rather than a speech sound – despite the computer’s sensitivity to sound waves. Possibly some of our aversion to crediting speech comprehension to computers is also due to their lack of speech perception (a delicate point that gets swamped by Searle’s semantic insight).

References


