A Cognitive Model in Design of Musical Interfaces

Anna Sofie Christiansen
University of Copenhagen
anna@cmuc.carnegie-mellon.edu

Abstract

Computer music interfaces present the performer/user with new options for controlling computer-generated sounds. Design of real-time music: interfaces or electronic instruments has previously often been carried out as a secondary feature subordinated to the capabilities of the sound-processing software or the sound producing system. In my paper I will suggest a model for instrumental interfaces in accordance with research in virtual reality, cognition and dynamic theory to be applied in design of musical interfaces, in order to ensure that the performer has a sufficiently differentiated sound control to ensure that he can rely on the expressive means, enhancing the performer's options for giving an individual and intuitive interpretation.

1 Background

In interactive computer music the field of sound synthesis has often been given preference to that of sound control. The consequences for the performance of interactive computer music might be more severe than we think. Standard interactive systems based on MIDI are, till now, far from allowing us take full advantage of sophisticated synthesis techniques for real-time purposes.

The crucial point in design of interfaces is the mapping between human actions or to the domain of computer generated response. In this paper I will take my point of departure in the direct gestural modeling of sound. Human interaction with traditional instruments shows that musical expression is performed across several parameters of sound [see, e.g. Rowe 1993]. Traditionally, parameters such as dynamics, duration and pitch have been computer-controllable, but expressiveness requires also taking advantage of, e.g., differentiated timbral variations within the single tone.

Newman and Laurel described the importance of relying on inter-human concepts creating a computational metaphorical approach in design of the link between action and response in the computer/human interaction. Thus, the mapping could be conceptualized as a prepositional communication between human to the user and the designer [Boden 1991]. The task to be accomplished in the communication can be considered as how to make explicit "that which has been left implicit" [Shank & Abelson 1977]. The prevailing model of interfaces has relied on the assumption that coordination between the human gesture and the computer's response can be represented as a solely conceptual model. The nature of interaction ras thus focused on mental tasks, ignoring totally the differentiated sensing of the human body. [Johnson, 1987] emphasized the significance of a non-propositional component in human cognition, a component closely connected to the human body.

2 Representing Direct Physical Action

In the following I will outline the representative level of interfaces. The act of playing music on an instrument involves a combination of cognitive as well as physical actions. The act of sound production is thus represented to the performer in several domains, involving notation, style and physical motion. The interface involves the domain of physical action in that it constitutes:

- A physical representation to a performer of agencies that enables him to act upon a sound-generating computer application.

The performer's actions are cleared by a sensing mechanism, and the software maps the performer's physical and/or sonic gestures to cause events generated by the computer. The correspondence between the performer's physical actions and the resulting sound thus requires a conceptualization between cause and effect.
• The performer's conceptualization of the link between his actions and the computer-generated events.

The performer's conceptualization of the link between cause and effect should relate to the physical causality, because the understanding of cause and effect in the physical world is generally understood by all people, and thus subject to intuitive understanding.

Research in cognition and dynamic theory have lately focused on the role of the human body in cognition, which opens up the possibility of integration with the domain of physical action. In my opinion, this integration is crucial to the performance of music. The human action generates, through the mapping, a response from the computer. The response also functions as a feedback mechanism that facilitates the performer's orientation on the interface.

3 Physical Feed-Back of Musical Instruments

The mechanical production of sound on acoustic instruments provides a highly differentiated sonic image, which is of crucial importance for the performance of music. The differentiation in sound control requires an intense training of the performer, and simplification of the control mechanisms reduces the variation in sound, or limits the freedom of the performer. Design of interfaces must therefore aim at providing a sufficiently differentiated sound, with control mechanisms capable of providing a sufficient representation of this differentiation to the performer. That is: modeling an appropriately differentiated sensory feed-back with a direct correspondence in the sonic result.

The modeling of the feed-back should take into account cognitive aspects of human interaction with acoustical instruments. I consider the sensory domains of primary significance to musical performance as being:

• Kinesthetic
• Tactile
• Aural

The importance of aural feed-back is widely acknowledged throughout the field of computer music, but the interdependency between aural and physical feed-back is frequently ignored. The point is whether it is reasonable to imagine that a highly differentiated control of sound should be possible solely as an aural, visual and mental representation. Research in tactile sensation has shown that the human bandwidth of perception depends on physical receptors [Johnson and Cott 1991]. In my opinion physical interaction with acoustic instruments benefits from the feedback of physical action (kinesthetic, tactile) that cannot be accomplished in a solely mental representation. Sufficient differentiation in the sonic image can not be represented to humans without relying on the increased bandwidth in multiple sensory feed-back.

4 The Embodied Schemata

Traditional instruments have a direct mechanical action that corresponds to our experiences of physical action/reaction in our surrounding world. Our motoric schemata are thus adjusted to respond as a reflex to actions that represent structures similar to those of the physical world. Schank and Abelson emphasize that creating a causal invariance will facilitate understanding. The differentiated motoric pattern in the performance of music should thus have a sonic equivalent in the mechanical oscillating body.

Sloboda characterizes the performance of the highly specialized performer — the expert — as somebody who is able to deploy several skills simultaneously, and who employs schemata that can coordinate action and feed-back. Schank and Abelson suggest that this specific knowledge is used to interpret and participate in events we have been through many times. The performance of music involves specific knowledge about sensation on several levels. Sloboda suggests that a specialized performer is constrained to a pool of real-time resources, but, by

\[ A \text{ special case is the organ, where an indirect mechanical response has been modeled to accommodate to a notion of direct physical action.} \]

Christiansen

ICMC Proceedings 1996
automating execution of some skills, he can manage to coordinate multiple layers of specialized skills simultaneously.

The model of the schemata can successfully be applied to music, in that it assigns data structures to represent generic concepts stored in memory. Schemata exist for generalized concepts underlying objects, situations, events, sequences of events, actions, and sequences of actions [Rumelhart & Ortony, 1977]. Performance of complex tasks relies on knowledge derived from several sub-schemata (conceptual or motoric), and the schemata are simultaneously adjusted to fit new situations if inconsistencies between assumption and feedback occur. Kinesthetic and tactile feed-back are thus important in that they convey sensing information to help the performer orient himself on the instrument and adjust his behavior in accordance with the feedback. His physical adjustments are encapsulated in a chain of cognitive and embodied schemata [Johnson, 1987] that knows, e.g., how to infer an action of changing air pressure by adjusting the embouchure from a particular fingering.

Dynamic theory suggests that integration of the embodied schemata in the cognitive model of musical instruments, requires a consideration of the dynamic aspects in multi-sensory perception. The cognitive system is here no longer considered to interact by passing messages, but co-evolves with them [Goldner and Port]. The relation between action and feedback is thus no longer solely depending on knowledge, but on performance. This approach leaves room for context-sensitive aspects of human behavior, because the understanding of coevolving transactions between body and mind, where the boundaries between perception, action and cognition are reduced to a single process, that blurs the distinction between knowledge and performance [Thelen]. The performer is no longer a cognitive system relying on mental representation, but a multi-sensory system who's mental and behavioral act is emerging in context.

The system consisting of performer and instrument is complex, but any configuration of the interface will serve to enhance certain patterns of action. These enhanced patterns will thus emphasize certain behaviors, and ensue a consistent performance. Sloboda's statement of live performance of music is relying on automated processes, seem to suggest a parallel to the trajectory.

5 The Musical Interface

Dynamic theory seems to imply that cross-fertilization between mental and motoric domains, is unavoidable. Therefore we are forced to rely on both domains in musical interfaces whether we intend such an integration or not. We can further increase the performer's sensual bandwidth, in relying on multi-sensory feedback, and thus enhance a cognitive representation of a medium sufficiently differentiated for sound control.

The degrees of freedom in musical interfaces are reduced by the limitations of the performer's mental and physical abilities (but also by the resolution of the sound control provided by the interface). But a good musical interface enhances the performance of music by giving preference to specific trajectories that are of musical significance. By this means, considerations of the coevolution of physical gestures and mental understanding offer a significantly increased potential for a good cognitive understanding of the control mechanisms available.

The interface must thus seek to enhance favorable trajectories by offering an unambiguous representation to the performer with a sufficient amount of feedback. The musical interface should thus permit the performer to take advantage of these trajectories, in that the execution of these trajectories will depend on the orientation the performer can extract from the multi-sensory feedback. If the interface thus includes multi-sensory feedback, the sensory bandwidth of the performer will enable him to orient himself in controlling a much more differentiated sound generating system. The problem is now how we model the interface taking these constraints into account.

It might be useful here to look into design theories employed in virtual reality. The all-accommodating virtual environment has to provide a sufficient amount of reality-like features in order to simulate a physical reality that we all know quite well. The designer of musical interfaces who considers musical performance as a solely mental behavior, must realize that by incorporating significant features from the performer's physical reality, a reasonable mapping between sound and gesture will enhance the potential of the performance.
The performer's artistic means of expression depend on a highly differentiated sound control, which enables him to create an individual and intuitive interpretation, that, in itself, is an essential constituent of musical performance. Designing the interface — keeping in mind the significance of tactile and kinesthetic feedback — and thus ensure that a differentiated control mechanism is optimally represented to the performer. This will enable the performer to take advantage of the means of expression available on acoustic instruments with a direct mechanical action, where the performer has control over subtle inflections in sound. This corresponds to a detailed sonic image and thus enhancing the expressive performance of music.

In general, there is little standardization of electronic instruments. Furthermore, rehearsal time is constrained and education of performers seldom leaves options to experiment with electronic instruments. I find it of crucial importance for the field of computer music to consider the constraints that interfaces designed with little regard to the nature of musical performance impose on the integration of computer music in the standard concert repertoire. I am fully aware of constraints due to the expenses of equipment etc., but ensuring that musicians can rely on their previous training in performing on electronic instruments, will contribute to make better performances, and enhance portability which might lead to an increased interest for electronic music.

References


