ABSTRACT
Is there a distinction between New Interfaces for Musical Expression and New Interfaces for Controlling Sound? This article begins with a brief overview of expression in musical performance, and examines some of the characteristics of effective “expressive” computer music instruments. It becomes apparent that sophisticated musical expression requires not only a good control interface but also virtuosic mastery of the instrument it controls. By studying effective acoustic instruments, choosing intuitive but complex gesture-sound mappings that take advantage of established instrumental skills, designing intelligent characterizations of performance gestures, and promoting long-term dedicated practice on a new interface, computer music instrument designers can enhance the expressive quality of computer music performance.

Keywords
Expression, instrument design, performance, virtuosity.

1. INTRODUCTION
The title and popularity of this conference, New Interfaces for Musical Expression, now in its sixth year, demonstrates the international interest in the design of new methods by which to enhance the expressive power of computer music. In this article we examine some generally accepted assumptions about computer music expression, and some practices associated with new interfaces, in order to draw attention to the question of whether musical expression in performance is being adequately addressed in much current research on realtime computer music interfaces.

2. WHAT IS EXPRESSION?
2.1 Common definition

expression: felicitous or vivid indication or depiction of mood or sentiment; the quality or fact of being expressive [15]

expressive: effectively conveying meaning or feeling [16]

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME 06, June 4-8, 2006, Paris, France.
Copyright remains with the author(s).

2.2 Composition vs. performance
Music can indicate mood or sentiment, and can convey meaning or feeling simply by its organization (composition) of sound elements [3], and the performer of music—often called the “interpreter” in the case of composed music—provides expression by evincing that organization, by adding shape and nuance to the given materials. So it is important to distinguish whether we are talking about expression in composition—expressive characteristics of musical materials and their organization by a composer—or expression in performance—expressive gestural nuance in real time. For the purpose of this article, we are referring to the latter—the nuance that a live performer adds to the available materials. The New Grove Dictionary of Music and Musicians notes that this form of expression encompasses “those elements of a musical performance that depend on personal response and that vary between different interpretations.” [4] In the case of “programmable instruments” and live control of compositional computer music algorithms, the distinction between compositional expression and performative expression may be blurred somewhat; the performer may be shaping primary characteristics of the composed/improvised musical materials themselves. What we are specifically concerned with in this discussion, however, are those characteristics of the live performance that enhance expressive communication beyond that which is contained in the materials on a notated page or a pre-programmed algorithm.

2.3 Performers bring expression to music
Poepel [22] described a mechanism by which “performers code expressive intentions using expressive-related cues” (including “tempo, sound level, timing, intonation, articulation, timbre, vibrato, tone attacks, tone decays and pauses”) and listeners “receive musical expression by decoding” these cues. This implies that performer expression, like language, depends on a set of conventional signifiers and an understanding of those signifiers shared by both performer and listener. These cues are generally at a different logical level than that of each individual parameter of a sound, and thus are not easily emulated with simple one-to-one mappings of gesture to sound parameter. “From the perspective of a musician, live performances allow him or her to make real-time choices that affect the interpretive variation in the music. These choices give the performance a trajectory and a range of variation that define the expressiveness and power of the performance. Techniques for creating this variation involve subtle control over

---

1 The New Grove article on “expression” treats primarily the organizational aspects of expression in composed music.
aspects such as timing, volume, timbre, accents, and articulation—sometimes implemented on many levels simultaneously.” [20] Thus, what we think of as musical expression in performance usually involves the performer’s contribution of culturally understood variations of specific sonic attributes at the note level—e.g. intonation, timbre, vibrato speed and depth, etc.—and attributes of the musical structure at the phrase level—rubato, crescendo, etc.

Even the physical gestures made by the performer affect the listener’s perception of the music. In the visual arts, a viewer can respond to specific traits of a pencil line, such as its smoothness, which can be “a tactile quality, a pleasurable one. [One’s] reaction to the line is not just to its formal qualities (continuity, for instance, or darkness or lightness), but also a kinesthetic sympathy (the Italians call it syntony) with the hand that drew it—the pressure, the weight, the gestural control, etc. Is the line, then, expressive or does it seem expressive because it is a trace of what I perceive/read as an expressive human action?” [Murata, M., personal correspondence]. Similarly, when one hears a sound, one can imagine and empathize with the physical gesture that might have created the sound; in this way, sounds imply gesture, even choreography. Conversely, viewing the physical gesture that a performer makes can influence the listener/viewer’s perception of the sonic expression. Computer interfaces can dissociate gesture from result to varying degrees by the way that software intermediates the relationship between gesture and resulting sound. (A one-to-one correspondence such as a mallet striking a marimba is an example of a simple gesture-result relationship, while a finger pushing the play button on a CD player exemplifies the opposite extreme in which a simple neutral gesture produces a complex musical result.) Jordà [19] evaluates this relationship as the “efficiency” of the interface, defined as the ratio of “musical output complexity” to “control input complexity”, but acknowledges that these are “quite fuzzy terms”, and that while computer-mediated controllers can provide more “efficiency” than most acoustic instruments, they often lack the “expressiveness” (flexibility, diversity of micro-control, etc.) of traditional instruments.

2.4 Can a machine be expressive?

Just as philosophers and computer scientists have debated the question of machine intelligence for decades [20, 25], there continues to be debate as to whether true musical expression (conveying meaning or feeling) can be produced by a computer. The fundamental question in both cases is whether the appearance of intelligence or expression is sufficient to believe that intelligence or expression exists. Emulations of expressive performance have been attempted by means of rule-based programs (e.g., [7, 10, 14]) and by machine learning, notably case-based reasoning (e.g., [1, 2, 3]). However, it is questionable whether systematic emulation of performer expression derived from other musical contexts is the same as what human performers do each time they interpret a composition. Computer musicians working with realtime performance systems have expressed the cautionary view that “just as the –ivity suffix in the word ‘interactivity’ connotes ‘a quality of’ interaction that can only be artificial in a machine, ‘expressivity’ for a computer can only be a demonstration of an artificial or simulated quality of being expressive in the sense that we apply it to human music making: the conveyance of meaning or feeling.” [11] “Be aware though: music instruments, being machines...cannot be expressive...since machines do not have anything to express. However, they can be used to transmit human expressiveness, and they can do that in many ways and with different degrees of success.

Our instruments will achieve this for better or worse, in the measure they permit the performer to transform a mental musical representation into musical gesture and sound.” [19] These musicians feel that the expression comes from the performer, and the instrument enables—and ideally facilitates and amplifies—that human expression. Although we may speak of an “expressive instrument” for the sake of brevity, it is important to recognize that we usually mean “an instrument that affords expression”, that is, “an instrument that enables the player to be expressive”.

3. CONTROL AND EXPRESSION

3.1 Control ≠ Expression

The mere presence of a finely calibrated instrument does not guarantee that it will be put to an expressive use. It might be said that the ability to control a sound generator, and the means of that control, are the tool or the medium by which expression is made possible. But it is important to note that one should not therefore equate control with expression. The performer’s expression is the significant content that is made possible to convey by the ability to control the sound generator in real time.2

The most basic need for a controller is that it accurately capture the data provided to it by the human interface. Another basic need is that the software provide correspondences between input data and output sound that are sufficiently intuitive for both performer and audience. It has been suggested that “the expressivity of an instrument is dependent on the transparency of the mapping for both the player and the audience.” [13], be it through direct mapping schemes, or more sophisticated gesture analysis (e.g. [8, 23]). Transparent or not, the correspondences must be learnable, repeatable, and sufficiently refined to enable control of the sound that is both intimate (finely detailed) and complex (diverse, and not overly simplistic).

Thus, control is a precondition for enabling expression, but is not in and of itself sufficient. Expressive control requires at a minimum that the interface provide accurate capture of gesture, and that the mapping of input to sonic result be situated at the appropriate level of structural detail (microscopic, mid-level, or macroscopic). Expressivity can be enhanced by intelligent recognition of gesture in order to characterize the gesture and make the appropriate mapping.

3.2 Simple and complex mapping schemes

In trying to design an instrument that will enable expression, it is necessary to consider how the performer will provide musical expression, notably how the performer’s gesture will affect the sound. Simple one-to-one mapping of input control data to a particular sound parameter is essential in many cases in order for the performer to have precise control, but such control is not equal to expression. Expressive control relies on more sophisticated use of the control input information, such as through one-to-many mapping of control data to a combination of parameters, recognition of complex characteristic gestures, or other methods that enable the

2 The English philosopher R.G. Collingwood formulated “his celebrated distinction between art and craft, according to which craft is a means to an end and must therefore be conducted according to the rules laid down by that end, whereas art is not a means but and end in itself, governed by no external purpose.” [4] The building of instruments/controllers is the craft of enabling expressive control, whereas the expressive use of the instrument is an art.
simultaneous and multi-dimensional shaping of combinations of parameters. Good performers use this complex multi-parametric shaping to encode a meaningful variation from a norm, be it by adding nuance not specified in the score of a composition or by varying established standards of consistency (steady tempo, discrete scale steps of pitch). For example, vibrato in a flute or a violin sound, which is rarely notated but which is a generally accepted method of playing espressivo in Romantic and contemporary music, is a simultaneous modulation of pitch, loudness, and timbre (i.e., frequency, amplitude, and spectrum), with these multiple modulations themselves being shaped (modulated) by the performer.

It is one thing to create a controller with simple mappings that even a novice can use with satisfying results without training (e.g., [8]), but it is quite another to develop an instrument that provides maximal control, diversity, and efficiency, in order to best enable expression by a skilled user. For an instrument to be considered potentially expressive by a trained musician, it must necessarily have a certain degree of complexity in the relationship between input control data and sonic result. So it is reasonable to expect that such an instrument will have a certain learning curve (c.f. [19], p. 176 ff.); a performer will require a certain amount of training and practice to achieve good control of it. For high-quality musical expression, an instrument should be mastered; the performer should achieve a level of virtuosity.

4. WHITHER VIRTUOSITY?

4.1 Common Definition

virtuosity: great technical skill [17]

First we wish to clarify that our use of the term “virtuosity” refers to a person having complete mastery of an instrument, such that s/he can call upon all of the capabilities of that instrument at will with relative ease; we are not referring simply to extravagant displays of extreme speed or dexterity. (Indeed, a computer is capable of playing at speeds much greater than humans, so playing fast notes on a computer instrument is no longer necessarily a display of virtuosity.)

4.2 Virtuosity Facilitates Expression

A primary value of virtuosity is that it transfers much of the knowledge of how to control the music to the subconscious level for the performer; basic functionality of the interface is no longer necessarily foremost in the performer’s thoughts. “The human operator, once familiar with the system, is free to perform other cognitive activities whilst operating the system.” [18] When control of the instrument has been mastered to the point where it is mostly subconscious, the mind has more freedom to concentrate consciously on listening and expression.

4.3 Lack of Virtuosity Inhibits Expression

How much time is needed to develop mastery or virtuosity on a musical instrument? Of course there is no definitive answer to this question, but it is probably safe to say that virtuosi have almost invariably spent years of highly focused practice and experience on their instrument. It is also safe to say that almost all sophisticated musical instruments have evolved over many years or even many centuries of technological refinement, focused development of technique over a long period by numerous different musicians (often competing, but also usually learning from each other), and a production of a large body of repertoire that contributes to both the technological and the technical advances of the instrument. Is it then naïve to think that masterful performance will occur on an instrument that was designed, developed, built, composed for, and rehearsed only within the last year or even the last few months or weeks?

The vast majority of performances of computer music that involve new interfaces, new instruments, alternative controllers, etc. are more experimental than they are refined and virtuosic. They are generally performed by someone who has only recently encountered the instrument, has had relatively little time to explore and understand the subtleties of expressive capabilities it affords, may be dealing with an interface the mappings of which have only recently been programmed (let’s be honest, in some cases as recently as the dress rehearsal), and who does not truly have a performance-level mastery of the instrument as it is configured. In many cases, the performer is someone who is a composer or technician more than a professional instrumentalist or stage performer.

There is nothing wrong with this experimentation. Indeed, it is vital to the progress of this field. And in fact there is nothing so very wrong with putting this experimentation onstage in a less-than-refined form at demonstrations, workshops, and conferences. But it would be a mistake to pretend that such an onstage experiment is a good representation of the expressive capability of that instrument, or that it can—except in a few fortunate instances—be legitimately compared to a high-caliber professional virtuosic musical performance.

Schloss [24] has remarked that “some pieces performed nowadays claim to be interactive, but in reality they are simply not finished yet. So the performance involves the ‘baby-sitting’ and ‘knob-twiddling’ we might see on stage that is so unsatisfying to watch.” The lack of virtuosity on new musical interfaces is apparently another case of the “elephant in the corner”—a big bothersome issue that everyone knows is there but is hesitant to discuss.

4.4 New Instruments Modeled on Old Ones

One approach to improving virtuosity and expressivity in live computer music has been to design instruments modeled on existing acoustic instruments. Indeed, this is still an attractive approach to many in the field, as demonstrated by the NIME 2006 “special paper session” on Digital Interfaces for the Violin Family. Early designers of synthesizers, and designers of the MIDI protocol, recognized the value of taking advantage of the years of skill developed by large numbers of keyboardists. Designers of other commercial computer-enabled acoustic instruments (computer-captor-enhanced violins, saxophones, etc.) and instrumental controllers (Zeta Strados violin pickup and Synthony II MIDI processor, Yamaha G1D guitar pickup and G50 guitar MIDI converter, Yamaha WX5 wind controller, etc.) have also attempted to make controllers that will allow capable performers of those instruments’ acoustic counterparts to bring their expressive skills to computer music.

Computer interfaces that are closely modeled on existing acoustic instruments can reduce the learning curve for those performers who are experienced on the acoustic counterpart.

A particularly musically satisfying fusion of ‘technical’ and ‘artistic’ issues occurs in works such as Chopin’s Etudes for piano, which address both in a meaningful way.
5.2 Still Better Mapping Ideas

Defining correspondences between gesture and sound—i.e., mapping control data to sonic parameter(s)—has been the focus of an enormous amount of research (e.g., [28]). Some basic problems have been recognized, yet many still have not been satisfactorily solved. “Strategies to design and perform these new instruments need to be devised in order to provide the same level of control subtlety available in acoustic instruments.” [27] One problem is the need to have intuitive yet detailed control of a computer music instrument that might itself be vastly multi-dimensional. However, with “controllers that output more than 3 continuous streams from the same gesture, it can be exceedingly difficult to reliably reproduce a gesture in performance.” [21] Indeed, with a motion capture system, “a single performer wearing a standard set of thirty markers, with three coordinates per marker, produces a stream of 90 simultaneous continuous parameters available for musical control….This profusion of control data presents…a challenge of the limitations of awareness for the performer.” [12] A “fly-by-wire” strategy (i.e., a divergent one-to-many, mapping), whereby a small amount of control data provides the necessary guidance to a complex system, is implied for intuitive-yet-complex control of sound. Such a system requires some time to master.

5.3 Feedback

Instrumentalists rely on tactile and visual information as well as sonic information. A pianist can see and locate a specific key before playing it, can use the resistance of the key-action mechanism to help know how hard to press the key, and can use the feeling of adjacent keys to keep track of hand position. Similar examples can be found for almost any acoustic instrument. Thus, visual information (telling the player what is possible, and where controls are) and visual feedback (telling the player what happened) are very helpful in a new interface. Likewise, haptic (tactile) feedback is useful for gauging one’s progress on a continuum. Some new interfaces lack sufficient visual and haptic feedback: for example, video motion tracking software allows the performer unfettered movement, but provides only sonic feedback. Sonic feedback, while necessary and valuable for musical performers, is always retrospective; the sound has already occurred by the time the feedback has been received. One can learn to play such a “virtual” instrument virtuosically, but the learning curve is decidedly high.

5.4 Gesture Recognition

While accurate tracking of gestural information is crucial for good control and expression, the software that interprets that data can be made even more “intelligent” by analyzing characteristics of that gesture. This includes second- and third-order analyses of the input data—recognizing not only the input value, but also the speed and direction of change, acceleration of change, etc. Some implementation of pattern analysis, recognition, and categorization can also lead to more intelligent software. For example, in addition to tracking gesture, it is useful to know what kind of gesture it is, thus making it possible to associate meaning with that gesture. Techniques of pattern matching and gesture recognition have a long and well-developed history in the field of artificial intelligence, and much of that knowledge can be fruitfully applied to musical gesture.

5.5 Critical Discourse

Critique is a vital aspect of intellectual and artistic life, which obliges analysis, evaluation, and discussion, which in
turn leads to improvement. Conferences such as NIME, ICMC, SEAMUS, and SMC focus predominantly on technical presentations and music concerts for an audience of like colleagues; however, critical discourse regarding the quality of the music, aesthetic values, and effectiveness of new interfaces takes place mostly in private conversations over an after-concert beer. In addition to being the sites for exchange of technical information, these conferences can and should serve as ongoing public forums for evaluative aesthetic discourse, encouraging increased public critique and debate.

5.6 Repeat Performances
In order for listeners to appreciate and evaluate the expressive qualities of a performance, access to multiple interpretations—whether by the same artist or by different artists—is essential. However, since most of the leading computer music conferences focus their selection criteria for performances on either the composition or the technology, placing a premium on originality, repeat performances are rarely available. This is the opposite of the situation in classical music, where the most visible performances are usually pieces from a standard repertoire rather than newly composed works. Performance interpretation and expression are highly valued in the classical genre, while in live computer music circles the quality of the individual’s performance is usually a secondary consideration—or may even be impossible to evaluate because the piece is only heard once and the interface is so novel. If expression is truly a valued component of this new art form combining humans and machines, then time could be allocated at major public events for the display, critique, and contemplation of the unique qualities brought to life in a particular realtime performance. One result of such concert programming might be the emergence of a number of “classical” pieces in the genre, in which different performers’ interpretations could be critically compared, thus focusing attention on the expressive use of the interface, rather than its design characteristics.

6. CONCLUSION
If musical expression with new computer interfaces is to reach the level of sophistication achieved by major artists in other specialties (jazz, classical, etc), it will be necessary to encourage further development in the following areas: continued focused research on strategies for better mapping, gesture recognition, and feedback; dedicated participation by virtuosos (utilizing existing virtuosity and developing new virtuosity); repertoire development for—and multiple performances with—a given instrument as a way to further its development; and more opportunities for critical discourse, both within the community of practitioners and among non-practitioners. The future is rich with possibilities for involvement by a wide array of interested and talented artists and artisans.

7. ACKNOWLEDGMENTS
Thanks to music historian Dr. Margaret Murata for contributing her observations on historical and philosophical views of expression in music.

8. REFERENCES

At a recent electronic music conference, all the composers were called to the stage for a group photo, leaving the few performers in attendance to ponder the implications of their exclusion.