Homepage

Reprinted from the Proceedings of the IEEE International Workshop on Robot and Human Communication, Tokyo, Japan, Sept. 1992.

Time-Forms,
Nature's Generators and Communicators of Emotion

Keynote paper, IEEE International Workshop on Robot and Human Communication, Tokyo, Japan, Sept. 1992.

Manfred Clynes
CNMAT
University of California, Berkeley
1750 Arch Street, Berkeley 94709, and
Microsound International Ltd. Box 143, Sonoma Ca 95476

Abstract - **Dynamic forms, called sentic forms, are described as language elements of a natural, biologically evolved language of communicating and generating emotions. These forms are genetically programmed into the central nervous system, and can be stored and recognized by computers and robots, and can serve as basis of real time emotional communication between them and humans. Music and art also utilize these forms to store and embody emotional meaning. A double-stream theory of music is outlined - including two principles of unconscious musicality which can realised on a computer - allowing one to create first-rate meaningful interpretations without manual performance or dexterity. New social opportunities and dangers are discussed arising from simulation and virtual reality which may exceed the average human in emotional eloquence. Real-real time is introduced as a concept to include human timeconsciousness.**

I. INTRODUCTION

This paper is a short review of new thinking and practical applications arising from the discovery and study of time-forms and their biologically evolved meanings in communicating and generating emotion.

In communicating between men and machines, the use of discursive symbols has been general practice. Symbols come in two kinds: discursive and nondiscursive [1]. Discursive symbols have one-to-one relationship with what they denote. In non-discursive symbols human imagination is brought into play; such a symbol can act to evoke meanings and indirectly feelings. It used to be thought that music and the arts made use largely of such non-discursive symbols [1,2,3].

A major step forward was achieved in realizing that the evolution of nature has designed specific dynamic forms which act not in a symbolic way but *directly* on the central nervous system. [4,5,6] These forms, like laughter and yawning, require no symbolic translation. They constitute the "words" of a natural language of emotional communication. Over two decades of studies have been devoted to isolate these forms and elucidate their function [7,8,9,10,11]

Human communication systems use zeros and ones, or dots and dashes, to transmit information. But nature's own emotion communication system has codesigned the sender, receiver, and message units with meaning evolved by nature: the message units themselves (the sentic forms) have analog form which *act like keys in locks of our nervous system.*

The specific dynamic forms - "words" - are produced through appropriate, prewired modulation by the sender, as analog forms. The receiver has demodulation filters that recognise these forms, like keys in locks. In our brain, the amygdala, a special structure of the midbrain, acts as a "gatekeeper" both to modulate and demodulate the specific dynamic forms, or "words"[12]. Several output-input modes may be chosen for this process, moreover: the auditory, visual, tactile or motor systems.

The biologically evolved communication system inherently encodes and decodes emotional meaning - with specific dynamic

forms (in the range of 1-10 sec duration). Without this remarkable symbiotic design, we would exist in emotional isolation. Through it we can touch one another emotionally in the present: we can share our emotions in our stream of life, and outside this stream through art and music.

This auto- and cross-communication system a has developed its own vocabulary. There is a class of qualities of experience which can be communicated inherently by specific dynamic forms and are *contagious.* These may be regarded as basic emotions. The emotional quality is transmitted from one individual to the other, in whom it is generated in turn. Such transmitting of emotional qualities may be observed in animals also (cf. innate release mechanisms), in the behavior of flocks or herds. In humans it may be seen for example, in crowd behavior, political oratory by demagogues, concerts of music, theatre, as well as in intimate behavior.

The basic dynamic forms, called sentic forms, may be evident in a gesture, in a tone of voice, in a musical phrase or dance step. It is the character of the form, not the particular output modality that determines its emotional meaning. We have isolated the dynamic forms for a number of emotions especially, anger, hate, grief, love, sex, joy, and reverence [4,5,6]. Like laughter and yawning these forms themselves appear to be largely universal and can not be arbitrarily learned - one can only discover what is already there inherently, "hard-wired".

Societies differ in the degree of suppression, the frequency of use, and the choice of output modality for communicating these emotional qualities using their biologically given dynamic forms .

Precision of Dynamic Form

A consequence of the biologic design of key-lock relationship is that the power of transmission becomes a function of the ***precision*** with which the form is realized. A deviation from the biologically designed form will tend to diminish its power to generate the emotion. If the deviation is sufficiently large, no recognition will take place at all - the key will not fit.

An interesting question is raised as a consequence: what kind of distortions are acceptable to the process, and what kind of distortions suppress the emotional meaning? This becomes one of the central questions to elucidate in a theory of emotional language transmission. Jamming of the emotional quality in the transmission process may be produced by one class of distortions, but not by another. Thus, the concepts of noise and of signal to noise ratio need to be looked at in a different light from that of man- made transmission systems. A new kind of mathematics needs to be developed that can distinguish between interfering and non-interfering types of distortion and noise.

The gain of the transmission of emotionally meaningful form is a form function. The degree of the perfection of the form governs the power of generation. It is therefore not sufficient to have a form that is conventionally 'similar' to a desired expressive quality. It is necessary to redefine the meaning of 'similar' in this context.

Such a form could be a caricature, and can also be perceived as mimicking. Mimicry may be useful when no real emotion communication is desired [7,9], i.e. no real-real time emotion communication (for the meaning of real-real time, see later in this paper). Then the form merely acts to remind a person of the emotional quality in a relatively "untouching" way. If one wants to have genuine emotional communication, the forms need to be precisely realized. These forms thus operationally define what we *sense* as a 'sincere' or "from the heart" expression of emotion. The conditions for this are biologically given.

Therefore by simply being faithful to this precision of dynamic form we can achieve a mode of communication between man and machine, and between machine and man, which will be felt as not machine-like.

(That function defines as well *exaggeration* which also acts as impediment to 'sincere' sharing.)

A human tends to be very quickly seduced to project an *entity* into a machine that communicates in this way. One needs in fact to develop a resistance to avoid being "sucked in". *Continuing and varied interaction* between man and machine with *truly expressive* gestures and tones of voice on the part of the machine, or of genuine musical expressiveness, make it mountingly difficult to maintain one's awareness that it is "only a machine".

As it is increasingly becoming possible, by incorporating the dynamic forms now known to be truly and 'sincerely' expressive of specific emotional qualities, to program a machine to interact with the human so that the expressions are experienced as not being machine-like but "living", we need therefore to look as far as we can now, with some urgency, to consider the differences between a man and such a machine - especially emotionally.

Anyone who has seen the newest little toy dogs walk, jump and beg may have had the almost irresistible response to feel: "isn't he cute!" Applying that paradigm to the far more sophisticated dialogues and simulations that are possible with this new knowledge, it is not hard to realize that we will need to 're-arm' ourselves emotionally to maintain our human identity. This will not be easy to do because we are *biologic prisoners of those forms,* and *cannot escape* the feelings of livingness associated with the spontaneous production of such forms, and especially the more so in dialogue.

It is not so much the ability of such a machine to "think" that would cause this problem, but the implied and varied feelings which it would express so well. The emphasis is on *so well* . It becomes an artistic achievement, so to speak, to program a machine so that it communicates in emotionally meaningful dynamic forms, indistinguishible from those produced by humans - and by humans moreover who are *most effective in their emotional communication.*

Until now, it has been the province of great art to provide a "living" storage of emotionally meaningful form. The great performer of music Pablo Casals could express more emotional meaning in shaping one single note than another performer perhaps in an entire concert. Inwardly perceived first, the high resolution of time-form required in all the variables concerned in producing the musical meaning and sound, was the province of his art.

That power of pure biologic form, transferred into a computer and machine can potentially result in more powerful communication of emotional qualities than the *average* human can produce most of the time.

This is a prospect that we need to face as the Pandora's box of sentic forms is now open. Society needs to learn to use them in new creative ways (see footnote 1) and to live with their now conscious knowledge - as well as with the results produced by ever-better "mechanization" of incorporating these living forms in dead machines. Looking at a mannequin may intrigue us to a degree. But when that mannequin can variously and expressively converse with us and even sing with more eloquent and "living" dynamic shapes than most humans can produce, our feelings towards the mannequin will change almost irresistably. To force ourselves to regard such a "living" mannequin as a machine may after a while tend to make us wary of *real* humans to whom we might tend to apply the same filter, as if they were such mannequins. This, if successful, would recreate an isolation and alienation from which nature had rescued us millions of years ago when it developed these very forms. If we cannot *trust these forms any more,* what *can* we trust?

A new kind of emotional education will be required. We know too little yet to be able to say how such new emotional communication will turn out.

The most difficult programming eventually will be to incorporate true empathy, as distinct from sympathy, in the dynamic expressions of the machine [see 7, 9 for distinctions between empathy and sympathy]. Such expressions are not commonly found in theatre. They need to be based on a knowledge that the machine has of the human that far exceeds the current question the machine is being asked; they need to be based on insight which comprises his potentiality. Once that is achieved however, it will be very difficult to avoid *feeling* an entity, a benevolent, perhaps even loving entity within the machine.

Conversely, a hostile, malevolent persona can be programmed, for purposes that may be devilish, or perhaps merely protective against intrusion.

Personality structure could in fact eventually be introduced which would govern the types of emotional responses produced in specific dialogue and situations.

Furthermore, even the personality structure could be altered (or switched) in response to different *questioners* , in a preprogrammed way, so that "strangers", for example may be treated entirely differently from "friends", within a whole field of discourse. This could be experienced as two (or more than two) different people residing in the machine. And these different personalities could even function simultaneously, from the same machine, independently of one another - being friendly, and hostile, and indifferent simultaneously to different users or customers!

It thus becomes increasingly important to attempt to elucidate the real distinctions between man and machine, especially in view of

1) the largely prevelant concept that increasing complexity of machines might be the required prelude to their becoming conscious and

2) the ability of computers and machines to simulate not just human thinking but emotional communication in the biological language devised by nature with a precision and power that will not merely equal that of humans but may exceed it.

TIME FORMS

The central nervous system has evolved the capacity to perceive transient time forms with great precision - an ability that is used in the communication of emotional qualities. Because of the transient nature of these forms they rarely receive recognition in human language by specific words (some exceptions: sigh, caress). Nevertheless they represent real entities to the nervous system, as do visible objects which are frequently named. The time forms concerned are in the general range of about 1 to 10 seconds, and the precision may be as high as 1 part in 500. Small changes in the transient form affect the meaning. The dynamic forms behave in parametric space like functions with a number of troughs or minima (solutions for the parametric values for which the forms are most powerful (singularities)) separated by saddle regions. As the shape changes from its optimal form, by varying the parameters, it becomes gradually less meaningful and powerful. If we transform the form gradually from one basic emotion to another, it first becomes meaningless before it starts to express the second emotion. The dynamic forms for specific emotions may thus be described as *islands of meaning in a sea of meaninglessness.*

The parametric description of such forms associated with its meaning and the effect of distortions requires a new mathematical treatment, or a new kind of mathematics for its appropriate representation. We have described [6,7] the averaged expressive forms for each of the emotions studied in terms of Laplace transforms as responses to unit impulses as follows:

$$O / I = b \left(\frac{\tau_1 \tau_2 \tau_3 \delta}{(1+\tau_1\delta)(1+\tau_2\delta)(1+\tau_3\delta)} \right) \left(\frac{\tau_4\delta}{(1+\tau_4\delta)} + \Omega k_p \frac{\tau_5\delta}{(1+\tau_5\delta)} \right)$$

where:

δ	is the Laplace Transform
O	is the Laplace Transform of u(t), the sentic form.
I	is the unit impulse input
τ_1, τ_2	are time constants, parameters having characteristic values (eigenvalues) for each specific emotion
k_p	is a passion coefficient (positive for sex, hate; zero for joy, love and reverence)
Ω	is an operator which has the value 1 for $du/dt \geq 0$ 0 for $du/dt \leq 0$
b	is a scale factor

While this description gives a good representation of the transient form it does not portray the changes in meaning and in the power of communication as the form changes. We are presently working on a mathematical formulation that can reflect this important part of its natural behavior. It can be said that techniques of least square differences and other currently used curve-fitting approaches are not suitable for this.

PROPERTIES OF SENTIC TIME FORMS

We may list the remarkable properties of these dynamic forms, that act as multiple body-mind windows [25], and their action in the central nervous system as follows:

1. Its power of communication depends on its faithfulness to the characteristic form.
2. There is a one-to-one correspondence between the quality to be expressed and the particular dynamic form. There is an inherent coherence between meaning and dynamic form. This coherence is biologically determined and genetically conserved.
3. The expressive time forms are preprogrammed by the brain as single entities (trajectories) for each act of expression.
4. Iteration of the form tends to augment the intensity of the emotion state it generates over a certain number of repetitions to a peak from which intensity will recede with continued repetition. Repetition is most effective when it is not strictly predictable, i.e.not 'mechanical', but rather when the intervals between repetitions vary somewhat, in an unpredictable, quasi-random manner. This variation is of the order of 5-10% of the repetition rate.
5. Mixed, or compound emotions are expressed not through an algebraic addition of the expressions of the emotions of which it is composed, but through a telescoping of two separate expressions in time: the two component emotion expressions are seamlessly *spliced* somewhere in the middle (at a variable point). The expression of melancholy, for example, a compound emotion, begins with love, then is spliced to end in sadness. The love expression cannot

complete itself and turns into a sadness expression, somewhere about half-way through the expression; however this is a single preprogrammed entity of 5-6 seconds duration, not two separate expressions.

6. Each emotion expression needs to be completed, otherwise a frustration is felt. Interupted expressions produce frustration *regardless of whether they express positive or negative emotions*.
7. Humans can express only one sentic form (single or compound, see 5 above) at any one time "sincerely" - regardless of which and how many output modes they may employ - ie. sentic expression is a single channel sytem. If several output forms are used simultaneously, they will tend to embody the same sentic expressive form. A human cannot, for example, express anger with the voice and love with a gesture simultaneously, and *both* sincerely.

A computer however is under no such restriction since for it "sincerity" is not bound to a single channel.

Different classes of distortions will be differently effective to degrade meaning. Some types of distortions are quite tolerable and affect meaning little. Other types are highly detrimental. It is important to be able to distinguish the different types of distortion. In the visual arts, for example, small missing gaps in lines can be readily supplied by the eye, or more precisely, by the central nervous system, while even small distortions in the shape of the lines will be felt as serious detriments. Similar distinctions apply to the dynamic forms. Thus, meaning suffers less from limitations of high frequency response between 10,000 and 20,000 Hz say, than through the effects of wow and flutter, and the relative shaping of individual musical phrases, i.e. sub-audiofrequencies. These sub-audiofrequencies (ie. transient shapes) are demodulated by the amygdala, a different brain structure than is involved in auditory perception per se.

While the effects of hi-fi have been greatly promoted, relatively little attention has been paid to the concept of fidelity of meaning. Within a relatively restricted frequency band it is yet possible to have a highly faithful rendition in regard to meaning, as well as a highly degraded one. Thus the tone of voice, for example, the minute timing and loudness inflections of the voice can carry vast ranges of meaning, even in a limited frequency domain. The extensions of the frequency band may change the range of meanings only comparatively little. (Arthur Schnabel's interpretations of Beethoven gained very little when the frequency response was improved from the 78 rpm reproductions). Even whispered speech can carry a large range of emotional meaning in its inflections.

There is thus a need to develop a more extensive theory of communication of qualities through dynamic forms [25]. The work which we have conducted over the last two decades attempts only to make a beginning in this.

APPLICATIONS

We have identified and isolated specific dynamic transient forms, called *sentic forms, specific for each emotion,* that naturally serve to auto- and cross- communicate basic human emotions, such as anger, love, grief, or sexual desire. They also serve to *generate* emotions through repeated expression.

These time domain forms are distinct from facial expression, which has been predominantly investigated by researchers in a static setting (though animation by film artists has been an active contribution). The forms work expressively in the various modes of spatio-temporal expression [26]. For example, they can communicate and generate emotion over the telephone as sound shapes, and through touch expression, gesture, dance, and music.

Measured through touch, sentographically (see footnote 2), they can be transformed into sounds expressing the same emotion conserving the dynamic form. Such expressive sounds derived from touch, have been tested on hundreds of subjects inter-culturally, and will be demonstrated.

Sentic forms can be stored in computers and used to modulate a variety of outputs which thereby become emotionally expressive. They can also be used to modulate the tone of voice.

The relation of dynamic form to meaning has also been most particularly studied in great detail with music. The microstructure of music has been shaped in accordance with meaningful forms so as to convert the dead notes of a score into living performances [13,14,15,16,17,21].

The generating power of repeated expression of these forms have been used to create a sequence of emotions for an individual, each expressed repeatedly called *sentic cycles,* that has therapeutic benefits and benefits to well-being[7, 9, 18]. This programmed sequence of about half hour duration can also be stored on the computer, to guide the timing of expressions. The measured expressive shapes can also be stored, and displayed on the screen [19].

Now that we know sentic forms, we can communicate with emotional meaning between man and computer, or robot, in the time domain. While the robot will not feel the emotion, it will be able to discriminate among emotions and in turn, send dynamic signals to humans in the biologically based language devised and used by nature in communicating and generating emotion - in sharing emotion among individuals.

COMMUNICATION THROUGH TIME FORMS IN MUSIC

In music the precision of time forms in communicating qualities is exemplified par excellence. It provides an excellent laboratory for the study of small changes in time form and their effects.

Time forms in music appear to function through two streams, simultaneously produced or perceived by the central nervous system.

One stream is the pulse, a repetitive phenomenon which is automatized after the first instance utilizing the property of the Central Nervous System we have called Time Form Printing. A dynamic form or movement of the order of one second duration may be repeated automatically by the nervous system without subsequent specific attention being paid to the form - we can put it into 'automatic repeat mode' by a simple act of will! For example, one may make a movement pattern with the arm say, in a triangular shape, or in an ellipse, and once begun, one may reiterate this pattern an indefinite number of times, without paying further attention to it. The shape of this movement will tend to be conserved throughout the repetitions. Stopping such a movement however, or changing it to a new form requires a momentary attention, to direct it to move with the new form; and it will then continue to repeat with the changed form, again without requiring further attention. While we do such a repetitive movement with the arm, we can talk and tell a story, a second stream, without one stream interfering with the other.

Such a two-stream process goes on in music and appears to be of the essence. In Western music of the late eighteenth and nineteenth centuries the pulse seems to represent the intimate identity of who is telling the story, that is, the composer - as was noted by Becking [20]. Along with this repetitive stream there is a second stream, the unfolding emotional story of the music. The two streams are perceived in parallel.

Different styles of music emphasize the two streams differently. Rock and roll has mostly the pulse, Gregorian Chant mostly the story, classical music of the 19th century a balance between the two. Notably, much of present day avant-garde music has lost the pulse as a separate stream. (This tendency seems to have originated to some degree with the impressionistic (vs. expressionistic) music of Debussy.) In folk music and most of ethnic music there also is a fine balance between the two streams. The pulse here expresses the ethnic identity, rather than the composer's, and the microstructure of the ethnic pulse appears to be related to the rhythmic fine structure of the spoken language.

In one of the great drawbacks of an otherwise toweringly invaluable music tradition, *Western notated music is dead if performed as written.* The musical thought of the composer cannot be notated in detail, only as a skeleton. The performer has to flesh out the skeleton with meaningful musical microstructure in order to re-create the living musical thought.

Two principles which convert the dead musical notes to living music have been discovered [12].

1. The Pulse,
2. Predictive Amplitude Shaping.

Each applies to one of the streams described.

1. The Hierarchic Pulse.

The pulse is a repetitive *combined time and amplitude warp* applying throughout the music to groups of notes forming the pulse matrix, typically a group of four nominally equal notes.

We have determined the microstructure for the pulse for a number of composers as a pulse matrix combining specific time and amplitude warps [13,14,15,16,21,24]. This pulse matrix functions hierarchically on several levels, in a relatively attenuated mode for the higher levels. By applying the hierarchal pulse structure to a musical composition we can obtain performances of considerable quality and distinction in terms of musical meaning. Such a performance of Mozart's Sonata K330 will be demonstrated at this meeting. The attenuation levels at the various hierarchical levels and the hierarchical form

of the pulse structure need to be chosen with musical judgment, for each composition. The pulse *is therefore applied with musical judgment,* not in a purely mechanized way. But the application of the pulse and the choice of parameters occurs generically to the entire piece. Of the several levels of the pulse, typically three, the lowest two are composer specific, the highest piece specific. For piano music the pulse alone can provide much of the musicality required to bring music to life. Other interpretive variables which need to be entered include the balancing of voices, specific dynamic indications prescribed by the composer in the score, micropauses at the boundaries of sections and occasionally elsewhere, choices of tempo, and of changes in general tempo.

Like the gait of a person, or handwriting, the pulse represents the unique personality of a composer and provides his 'point of view'. or 'presence'. It can be readily realized with MIDI. Pulse matrix values are now known for Beethoven, Mozart, Haydn, Schubert, Schumann, Mendelssohn and some other composers. The pulse is also highly applicable to popular music, enhancing its livingness.

2. Predictive Amplitude Shaping

The second, parallel, stream of the music process is the unfolding of the musical story as a chain of melody and is continually changing and developing as the story unfolds. In this stream the amplitude contour shape of each separate note contributes to the musical meaning. A singer, string player or wind player can and does shape each note distinctively and differently according to the musical meaning. How the amplitude envelope shapes of each note are related to the melodic structure was discovered and is described by the principle of Predictive Amplitude Shaping, a *second principle* underlying the microstructure of music. It too, describes an unconscious element of musical thought, involved in musicality, not notated.

Where musical instruments allow it, every tone of a piece of music is shaped differently according to expressive needs. In a good performance the shapes of the tones are by no means uniform, but what guides their shape has not been systematically described either by music interpreters or theorists.

By using computers to generate melodies using sinusoidal sound only, and varying the amplitude envelope shapes and the durations of the tones, a principle was discovered which we call Predictive Amplitude Shaping, which appears to mirror functions of musicality which govern the shapes employed in meaningful interpretations [13].

This principle makes the amplitude shape of the present note deviate from a basic shape in a specific manner depending on what the next note is going to be. The envelope shape is skewed forward in proportion to the *tangent of the pitch-time curve* at that note. This means that the shape of the *present* note is skewed forward if the next note is going to be higher in pitch, and it is skewed backwards if the next note will be at a lower pitch, depending also on the time when the next note occurs.

The principle relates the shapes organically to melodic structure, so that the shape of the present tone *implicitly* presages what tone *will follow.* This gives a feeling of continuity and a continuity of feeling to the performance. It markedly enduces musicality into the phrasing of melodies, and appears to apply to melodies in general, i.e., it is not composer specific. The mean form however from which the skewing takes place seems to have composer-specific aspects.

In creating the shapes of envelopes for the amplitude of the musical tones we have departed from the industry standard of specifying attack, decay, sustain and release. These parameters were derived from the properties of a piano or keyboard tone, they reflect the musical machinery of a piano. We note however, that musical thought does not think in terms of such discontinuities. Rather, the musical thought of a tone is shaped as *a continuous curve.* Such continuous curves of amplitude contour are naturally produced by the human voice, string instruments, wind instruments, wherever the tone can be continously shaped. This ability is in fact a source of *superiority* of such instruments over the comparatively fixed amplitude contours produced by pianos (and even more so harpsichords and organs which cannot even readily accomodate the amplitude warps of the composers' pulse matrix - but not so the clavichord!).

The shapes of individual tones, and their range of variation in music can be largely described through only two parameters of a beta function [13,21]. We have used this method to calculate customized shapes for each tone based on the melodic structure, and the above principle. The basic shape from which skewing takes place applies to all notes, short or long (except for very long notes eg. pedalpoints).

These organically varied shapes seem to represent musical thought better the conventional methods. They make even scales sound more musical, and often result in phrasing as specially indicated by composers.

They relate to the inner gestures that govern the emotional expression of the music and allows the form of the music to

correspond to the inner desired shape.

Vibrato structure and changes in timbre within each note can also be guided by related principles in an integral way, customizing every note.

This second principle shall be of great use especially in the next technologic generation.

It is not readily realised through MIDI, in contrast to the Hierarchic Pulse, but will be of great value in the next technologic generation for the general user to achieve living musicality.

APPLICATIONS

Both the pulse and predictive amplitude shaping can readily be incorporated into a computer program which can perform music in a living manner as a consequence, with emotional meaning as desired by the user of the program. This allows any user who has no manual dexterity from ages 8 to 80 to interpret great music using only their own sense of musicality.

Composers can now specifiy in a micorscore the way they wish their music to be performed.

Using these principles teaches one how to improve one's musicality in a profound and subtle way. It is a powerful pedagogic means.

The knowledge that we have obtained through the realization and application of these principles of unconscious musical thought, through computers, is significant to AI [27], to musicians and to the general public. In a sense it is the achievement of the virtual reality of the music performer - the simulation of music interpretation and performance, so that every smallest detail is known (through the global adjustment of the principles, using one's musicality. Less than 2 % of the notes need 'manual, individual' adjustment). It has implications also to the subtleties of rhythms and inflections of speech in conveying shades of emotional meaning.

Appendix A gives details of the demonstrations of its function, given as part of this presentation.

EMOTIONAL COMMUNICATION WITH A COMPUTER OR ROBOT

TIME CONSCIOUSNESS: REAL-REAL TIME

More than any sense perception, such as visual, auditory or tactile experience, time consciousness is intimately linked to consciousness. Without seeing, hearing, touching, or other sensory inputs, consciousness is still engaged with time consciousness. Mental events are ordered in time, apart from sensory inputs. One may think, for an obvious example, of a musical rhythm. The tempo of such a rhythm for a given meaning depends on our time consciousness.

Time consciousness has different aspects for short term, intermediate term and long term periods. In communicating emotions and qualities with dynamic forms we are concerned with shorter periods, times under 10 seconds, as entities or units communicating of emotional meaning, and periods of the order of one hour for the experience of communicative structures ('stories') built from multiple units (e.g. a symphony). In these regions we have observed stabilities of timing of the order of one part in 500 [22, 23]. Such stable timing has been observed regardless of time of day, temperature of the environment, body temperature, including fever of 2-3 degrees above normal, and to a degree, variable acoustics. Studies of timings of performances of musical pieces well known to the performers on different occasions over several years have documented this. Experiments with tapping and mentally rehearsing portions of the same piece, over many years by the same person has illustrated similar stability [22].

These findings suggest that if the intended meaning attributed to the musical piece does not change (i.e. the concept of the music is not altered), a high degree of stability is observed in the performance timing and consequently, we may say, in the individual rate of time consciousness. Moreover, that a particular piece of music, or dynamic expressive form, will tend to have similar meaning to different persons *at a given tempo* suggests that their relative time consciousnesses cannot be too widely different. One could conjecture at most differences of the order of 5% between individuals, but it could be considerably less than that. *This* rate of time consciousness does not appear to change with age over an adult lifetime, and may be even similar at an earlier age. Clearly, there are neurobiologic clocks operating in the brain which are involved in this stability, and most likely they are of molecular rather than neuronal character [23].

A quite different aspect of our experience with time is found in relation to our sense of boredom and excitement. Here periods of the order of hours may seem relatively long, or short, depending on our engagement in activity and its degree of fascination. On this scale the experience of passage of time appears to be also highly age dependent: For a child an hour may seem an interminable period, especially when it is filled with an unwanted activity or condition. As one gets older an hour seems to shorten progressively, and days and weeks appear to go by considerably quicker, with a ratio for a sixty year old of perhaps as much as 3:1 compared with childhood. That this should happen although the time consciousness involved in the expression of dynamic expressive forms (eg. music) remains unchanged is very remarkable and appears to attest to the existence of different, less precise processes in this range.

A third aspect of time consciousness is found in the estimate of longer periods, of the order of a few hours to one day, and involve functions which make it possible, for example, to wake up at a specific time without the benefit of alarm clocks, simply by an act of will, i.e. a pre-determined period entered by the mind into itself. This ability does not seem to degrade with increasing age, and can also be remarkably precise.

In our memories experiences are ordered in time, both in short term memory and long term memory, although the method of tagging may well be different for each. (If one could not tell the order in time of two consecutive events, all thought, logical or otherwise would be impossible.) Thinking, like a computer program, is not reversible in time. Time consciousness is the only solid scientific evidence we have for the direction of time. Physical laws do not really define it (they work equally for both directions) and the laws of thermodynamics allow us only statistical inferences. (see footnote 3)

Furthermore, time consciousness occurs in the present. The contents of consciousness changes, or, we can say, the content of the present changes, but the present (itself) remains the same, unchanging.

Physicists have brainwashed us into a habit of regarding time as a straight line, going from left to right in which the present is considered as a mathematical point which moves along this line at an unspecified rate. It is, in fact, meaningless to physics to ask what the rate of movement of this point is, along the line. What events take place are described by a change from T1 to T2 ("as T goes from T1 to T2"!). The present does not enter at all. Physics cannot deal with the present.

But the past is gone and the future is not yet here. All that is here is the present. Physics gives us an insufficient view of what exists, leaving out the present - which is all that really exists - and in which the eternal (ie.at some level unchanging) laws of physics operate.

But not so, time consciousness. Through it we are aware of the present moment and may distinguish it from both the past and the future. How it does that is unknown. We can say, however, that time consciousness is entirely *relative to the species* in which it is found. This relativity of time consciousness has nothing to do with the theory of relativity. It concerns rather the concept that individuals from a different galaxy might for example experience time in such a way that for them, night and day might appear as a flicker. Or, a fly on earth may well have a different time consciousness from a human. There is nothing absolute about time consciousness.

For a stone which may be assumed to have no time consciousness, it has no meaning to say that a given time period is either long or short. Thus, a billion years *for a stone* are not long and a microsecond not short, as it has no time consciousness.

For simplicity we propose that we denote human real time as real-real time *when considering communicating with computers and robots.*

HOW DO COMPUTERS AND ROBOTS EXIST IN TIME?

Clearly, computers and robots have no time consciousness. If and when they do acquire time consciousness, this would have to be especially engineered to be scaled according to human time consciousness or real-real time, for compatibility with humans. Materially, a computer is not essentially different from a stone. Unlike a stone, however, it operates with a succession of logical operations. Each operation is actually carried out between the ticks of the computer clock. Changing the rate of the computer clock, or even making it uneven has no effect on the outcome of its calculations. The numbers by which successive operations are given time tags are no different from any other series of numbers which the computer may store and handle.

At each tick of a computer clock the computer is in fact totally stationary as far as its calculating functions are concerned. It is, in fact, "dead". Only *between* ticks does it actually "work". And here the time taken is not known, and does not enter into the picture as long as its fast enough so it can be completed before the next tick occurs.

How does the computer therefore relate to the present?

Can we consider the computer to operate according to the image of time which we have learned from physicists? Its activities may be considered to span a period of from T1 to T2. The result of the calculation depends only on the sequence of logical operations or instructions carried out and is itself totally independent of how long it might have taken to complete them.

When a computer operates in real time it harmonizes the sampling rate of input and output so as to reproduce or produce output behavior corresponding to the physical changes which it attempts to model. Its calculations merely have to be fast emough to keep up with the sampling rate.

It is not concerned with human time consciousness nor that of any other species, nor with the human experience of the present.

But when a computer performs a time-form such as emotionally expressive music or speech, its real time operations need to reflect the real-real time experience of humans. In that domain a 1% difference in the time scale will already *noticeably alter meaning.* A 10% change is a considerable change for the meaning of music, for example. In speech a 10% change will not change the meaning of the words but will alter the effectiveness of the emotional "overtones" which the message carries. Accordingly therefore when communicating emotional qualities to humans the computer cannot choose its own convenient scaling but must match the output to human time consciousness, or real-real time. This requirement is unnecessary when a computer prints out verbal messages or presents visual information such as emotionally expressive faces, for example, as long as those faces are not in expressive motion.

Departure from real-real time results in the emotional qualities acquiring a Mickey Mouse character. They are still recognizable as representing particular emotional qualities, but lose most of their emotional impact. To understand this better we need to consider that the emotions as transmitted by these dynamic forms, which we have called sentic forms, also contain and imply *cognitive substrates* . Thus, we have shown experimentally that with the feeling of love, there is an openess and guilelessness as part of the inherent biological program [11]. Even a small lie effectively blocks the experience of love at that time. Similarly, grief affects memory function: the ability to learn, and short term memory are diminished by the feeling of grief, and the consequent loss of interest in interaction with the environment. *The cognitive substrates are the first to disappear when the dynamic forms are distorted* [24]. Thus the Mickey Mouse characteristics imply that the emotions denoted are robbed of their cognitive substrates. But it is precisely the cognitive substrates that provide the elements of profoundity which may be experienced through really true emotional expressions and of its sequences. Thus a piece of music may seem profound not merely because it contains a sequence of emotions but because the sequence of emotions implies a series of cognitive substrates which give it a widely and deeply probing story.

In the use of computers and robots to communicate emotions to humans, one therefore has a choice to communicate merely signs a la Mickey Mouse or to represent genuine, "sincere" human expressions and emotions. The particular applications will decide which approach may be more appropriate. If it is decided to use "sincere" expressions we have then the power in real-real time to make these expressions more powerful and more convincing than the average human would tend to produce under most average conditions. We have at our fingertips the knowhow to make these expressions powerfully communicative, contagious, and seductive in the manner of the very best that any human can do. (see footnote 4) We can optimize it *beyond* the abilities of the *average* human: we can optimize it to the degree of which our most powerful art is capable. Whether we chose to do so and for what reasons and needs, surely will comprise a new branch of social ethics which badly needs to be developed.

One positive way of using the new powers opened up by this Pandora's Box opened by the findings of our research is to provide great musical performances. How this can be done we will attempt to demonstrate with a performance of the Mozart Sonata K330, which will be brought into conjunction with the best available performances on CD by great artists for the first time. That this can be done now is only the first step in an endeavor that has many further possibilities, not easy to fathom.

Appendix A

Demonstrations of the Meaning and Precision of Time Forms

A Computer Interpretation of Mozart's Piano Sonata K330, which utilizes the composer-specific hierarchic pulse principle will be heard, along with 6 other performances of the greatest recordings available on CD of this Sonata. ***The music panel***

and the audience will be asked to rate the performances, and to pick* which one *is the computer performance. This test will compare the real-time performane of a computer in terms of musicality with the best real-real time human efforts.

Also played will be a computer interpretation of Bach's Air on the G string, for four independent voices, employing both Predictive Amplitude Shaping, and the Hierarchic Pulse.

The purpose is to demonstrate the degree of understanding of the principles of unconscious musical thought, of musicality.

In a second presentation, emotionally expressive *sounds* will be presented, which were transformed from *touch* expressions of the same emotion, by a transform that conserves the dynamic form - to illustrate that the nature of a particular emotion expression depends on the dynamic form, and not on the output mode, ie. is largely independent of the output mode. White urban touch expressions of specific emotions transformed to sound expressions were tested on Australian Aboriginees, for additional cross-cultural validation [13] .

Appendix B

A Short Note on the Development of Consciousness

In the context of this paper some remarks on the natural development of consciousness may be permitted. Experience of emotion is not possible without consciousness. Also, sensory experience, such as red for example, has continuity in time, and stability over long time (in addition to its unique quality), which are so far difficult to ground on the discontinuous events in time and in space of the multiple neuronal events accompanying the experience. Accordingly some thoughts on how consciousness may have arisen in nature may be assayed, in view of our developing knowledge of molecular biology. As a newly evolved phenomenon, it is of such importance that it seems not unlikely that its *potentiality* is *provided for in the laws of nature,* ie. that like water, say, it appeared in the evolution of the universe not as a total surprise. Contrary to a prevelant view according to which consciousness is deemed to become possible only when complexity increases to a sufficiently high degree, i.e. that consciousness is essentially *complexity*- related, the author puts forward a different view of what may be required for consciousness to arise. This approach is outlined briefly here, and will be described elsewhere more fully.

In exploring such a view, one should also consider the pervasive unconscious mental functioning, dreaming, and especially the unexplored questions of the boundaries between the unconscious and the autonomic, and whether unconscious mental function predates consciousness.

1. Content of Consciousness
   Humans are able to see, hear, smell and touch *simultaneously* without notable interference of one sensory experience with the other. This defines in size, and in also complexity, a minimum capacity of consciousness even without considering other mental functions.
2. Animals, even relatively primitive animals, appear to share this capacity of consciousness with humans in that they too can simultaneously see, hear, smell, touch and so on. Moreover, many primitive animals can sense some of these variables with higher resolution than can humans. Consequently, with regard to these functions one cannot take the view that animals have a smaller capacity of consciousness. The argument that such animals have no consciousness at all and act as reflex automatons is rejected (animals clearly appear to make decisions based on their sensory experience involving many of these variables, aneasthetics are used to eliminate consciousness in these animals, etc.).
3. One of the earliest qualities of experience developed in evolution is hunger. Experience of hunger - a remarkable "invention" of nature, replaced chemotaxis and served as *a fount of knowledge* for the animal concerning *what* to eat, *when* to eat, and *how much* to eat - all of which is encapsulated in the experiential entity, "hunger". But hunger, as well as experience of sexual attraction which also developed at an early stage, can exist and function only if there is consciousness.
4. It is suggested therefore as plausible that consciousness itself may have developed through *newly evolved genes* at a relatively low stage of evolution. It would be supposed that these genes for consciousness produce certain proteins or other gene products which cause consciousness in the brain. According to this view, it is thus not complexity per se but *these specific genes* that would account for the emergence of consciouness in evolution. (If this is so, the capacity for being a little bit conscious would appear somewhat like a little bit pregnant.)

   As progress is being made in the human genome project it is possible that such genes for consciousness could be identified in decades to follow. As we share so may of our genes with animals we may also share such consciousness creating genes with them.

5. In achieving consciousness through the interaction of proteins and other gene products in a totally unknown way it could be that a physical law would be invoked which is as yet unknown - a law which would in effect provide for the establishment of a Leibnitzian monad, *as a result of specific molecular interaction.* This kind of creation of *one* from *many,* similar to the creation of the oneness of the molecule from so many atoms (the potentiality of which preexists, predictable through the laws), and in some ways in effect reminescent of a field which at any one point automatically and necessarily summates the effects of many contributing sources in space, but obeying relationships as yet quite unknown would seem, according to this view, to be necessarily involved in the phenomenon of consciousness.

   Accordingly, machines that do not possess the gene functions for consciousness would not become conscious no matter how complex they might become.

   This does not preclude that those effective gene functions might be reproduced by an alternative molecular realisation, so that consciousness could then also be produced by a different configuration of matter than in natural evolution on this planet. But that would depend on the specific nature of the interaction and functions; it could also be that the solution realised by nature which we observe on earth is the only one possible.

References

[1] Langer, S. (1946) Philosophy in a new key.

[2] Langer, S. (1953) Feeling and Form, New York, Scribner

[3] Piechowski, M. (1981). "The logical and the empirical form of feeling", J. of Aesthetic Ed., 15,1,31-53

[4] Clynes, M. (1969). "Precision of essentic form in living communication," in *Information processing in the nervous system,* edited by K.N. Leibovic and J.C. Eccles, (Springer, New York), pp. 177-206.

[5] Clynes, M. (1970). "Towards a view of Man", in *Biomedical Engineering Systems,* edited by M. Clynes and J. Milsum , (McGraw-Hill, New York), pp. 272-358.

[6] Clynes, M. (1973). "Sentics: biocybernetics of emotion communication", Annals of the New York Academy of Sciences, 220, 3, 55-131.

[7] Clynes, M. (1977). *Sentics, the touch of emotion.* (Doubleday Anchor, New York. New edition, Prism Press, Avery Publishing, New York, London, 1989.

[8] Clynes, M. (1980). "The communication of emotion: theory of sentics", in *Theories of Emotion Vol. 1,* edited by R. Plutchik and H. Kellerman, (Academic Press. New York) pp. 171-216.

[9] Clynes, M. (1988). "Generalised emotion, how it is produced and sentic cycle therapy" in *Emotions and Psychopathology* edited by M.Clynes and J. Panksepp, (Plenum Press, New York) pp. 107-170.

[10] Clynes M., and Nettheim, N., (1982). "The living quality of music, neurobiologic patterns of communicating feeling", in *Music, Mind and Brain: the Neuropsychology of Music* edited by M.Clynes, (Plenum, New York) pp. 47-82.

[11] Clynes, M., S. Jurisevic, and M.Rynn (1990). "Inherent cognitive substrates of specific emotions: Love is blocked by lying but not anger", Perceptual and Motor Skills, 70, 195-206.

[12] Aggleton, J.P. and Mishkin, M. (1986). "The amygdala: sensory gateway to the emotions", in *Emotion: Theory, Research and Experience,* edited by R. Plutchik and H. Kellerman, Vol 3, (Academic Press, New York).

[13] Clynes, M. (1983). "Expressive microstructure linked to living qualities" in *Publications of the Royal Swedish Academy of Music, No. 39* edited by J. Sundberg pp.76-181.

[14] Clynes,M. (1985a). "Secrets of life in music" in *Analytica, Studies in the description and analysis of music in honour of Ingmar Bengtsson. Publication of the Royal Swedish Academy of Music, No 47* pp. 3-15.

[15] Clynes, M. (1985b). "Music beyond the score", Communication and Cognit., 19, 2, 169-194.

[16] Clynes, M., (1987). "What a musician can learn about music performance from newly discovered microstructure principles, P.M. and P.A.M", in *Action and Perception of Music,* edited by A. Gabrielsson (Publications of the Royal Swedish Academy of Music, No. 55, Stockholm) pp. 201-233.

[17] U.S Patents 4,704,682, 4,763,257, 4,999,773, Jap. Pat. Pend., EEC.Patent.

[18] Patent pending.

[19] Patent pending.

[20] Becking, G. (1928). *Der musikalische Rhythmus als Erkenntnisquelle* (Filser, Augsburg, Germany).

[21] Clynes, M. (1986). "Generative principles of musical thought: Integration of microstructure with structure", Comm. and Cognition, CCAI, Vol. 3, 185-223.

[22] Clynes, M., and Walker, J., (1982). "Neurobiologic functions of rhythm, time and pulse in music", in *Music, Mind And Brain: the Neuropsychology of Music* edited by M. Clynes, (Plenum New York) pp. 171-216.

[23] Clynes,M. and J. Walker (1986) "Music as Time's Measure", Music Perception 4,1,85-120.

[24] Clynes, M. "Composer's pulses are liked best by the best musicians", in press.

[25] Clynes, M. (1990). Mind-body windows and music", Musikpaedagogische Forschung, Vol 11, 19-42 Verlag Die Blaue Eule, Essen, Germany

[26] Hama, H. and Tsuda, K., (1990). "Finger-pressure waveforms measured on Clynes' sentograph distinguish among emotions". Perceptual and Motor Skills, 70, 371-376

[27] Minsky, M., (1987), The Society of Mind, Simon and Schuster, NewYork

---

Footnotes

Footnote 1

One such way is Sentic Cycles, a simple art form of touch leading to emotional balance and increased joy of living, which has been developed by the author[9, 7]. Touch ExPress(tm) is a self-contained hand-held version of the Sentic Cycle kit .

Footnote 2

The sentograph is an instrument measuring transient pressure (or force, more precisely) in two dimensions independently; pressure of the middle finger is used for expressing the sentic forms, as voluntary actions ('voluntary' here is a technical term meaning deliberate, conciously initiated by an act of will, transmitted by the voluntary muscle system, just like the movement of the arm in throwing a ball to hit a target.)

Footnote 3

When we see a car moving say at about 20 miles an hour, or a person walking, ie. at moderate speeds we do not see a series of stroboscopic pictures, nor a blurred image, and we can estimate the speed of the car, yet a camera will either show it still if

the shutter speed is fast enough, or with a blur. From a single photo we can judge the speed only from the degree of blur. We cannot tell directly from a sharp photo whether the object is moving or not. (Nor can physics tell the momentum of an object from an instantaneous view-apart from the relativity contraction, it will look just like a stationary object.) The human (and an animal, probably) does it without the need for blurring (blurring occurs only at higher speeds, exceeding the system's capacity), using his or her time consciousness-through a *perceptual quality*, aided by ratesensitive receptors that are submerged from the successive images that would appear on a photograph. Thus he can tell the direction of movement, and so of time, without needing a blur, nor multiple images. How much can we still learn about dataprocessing from nature!

Footnote 4

People with the clinical condition of hypomania often have an abnormally effective and powerful production of sentic forms in speech and gesture; this helps them to convince and seduce others - specific biochemical causes of this heightened expressity are not known.