MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 616 February, 1981

Music, Mind, and Meaning

Marvin Minsky

ABSTRACT: Speculating about cognitive aspects of listening to music, this essay discusses: how metric regularity and thematic repetition might involve representation frames and memory structures, how the result of listening might resemble space-models, how phrasing and expression might evoke innate responses and, finally, why we like music -- or rather, what is the nature of liking itself.

MUSIC, MIND, AND MEANING

Why do we like Music? Our culture immerses us in it for hours every day, and we all know how it touches our emotions, but no one thinks of how music touches other kinds of thought. It is astonishing for us to have so little curiosity concerning so pervasive an environ-mental influence. I'll speculate here about what we might discover, if we were to study musical thinking.

Have we the tools for such work? Years ago, when Science still feared Meaning, the new field of research on Artificial Intelligence supplied ideas about "representation of knowledge" helped out in several fields; I'll use them here. But aren't such tools too alien for anything so subjective and irrational, aesthetic and emotional as music? Not at all; I think the problems are much the same, and those lines wrongly drawn: only the surface of Reason is rational. [1].

Besides, much of what we now know of the mind emerged, this century, from other subjects once considered just so personal and inaccessible. Freud's work on dreams and jokes uncovered the Unconscious, and Piaget's work on children's talk and play initiated developmental psychology. Why did the work of Freud and Piaget have to wait for modern times? Before them, children seemed too childish, and humor too funny, for science to take them seriously.

Why do we like music? We all are reluctant, in music and art, to examine our sources of pleasure or strength. In part we fear success itself -- that Understanding might spoil Enjoyment. And rightly so; Art often loses power when its psychological roots are exposed. No matter; when this happens we'll go on, as always, to seek more robust illusions!

I feel that Music Theory has gotten stuck at trying too long to find Universals. Of course, we'd like to study Mozart's music the way those scientists analyze the spectrum of a distant star. Indeed, we find in every musical era some almost universal practices. But we must view these with suspicion. For, they might show only what those composers felt *should* be universal; if so, the search for truth in art becomes a travesty in which each era's practice only parodies the theory of its predecessor. (Imagine formulating laws of television screenplay, taking it as natural phenomenon, uninfluenced by custom or constraint of commerce.)

The trouble with the search for universal rules concerning thought is that our memories and thinking processes interact as we grow. We don't just learn about things, we learn ways to think about things; then we learn to think about *that*, and then about *that*. Before long, our ways of thought become so complicated that, I suspect, we cannot understand the details of any any mind, at any moment, without knowing the principles that guided its growth. The anatomy is too obscure without its embryology. Much of this essay speculates on how listening to music engages previously acquired personal knowledge of the listener.

As for the laws of liking music, it has become taboo for music theorists to ask why we like what we like: our seekers have forgotten what they're searching for. To be sure, "there's no accounting for tastes" -- in general. No matter: if different people have different preferences, we mustn't simply ignore the problem; instead we must try to account for how and why *that* happens! We must enlarge our aspirations to see that music theory isn't only about music, but about how people process it. To understand any Art, we must look below its surface, into the psychological detail of its creation and absorption.

If it sounds harder to explain minds than songs -- still, sometimes making problems larger makes them simpler! The theory of Equations' Roots seemed hard for centuries, within its little world of real numbers -- but suddenly seemed simple, once Gauss exposed the larger world of (so-called) complex numbers. Music, too, should make more sense once seen through listeners' minds.

SONATA AS TEACHING-MACHINE

Music makes things in our minds, but afterwards most fades away. What is it that remains? In that old Mozart story, the wonder child heard a concert, then wrote down the score. I don't believe it; history documents so few such tales that they would seem mere legend. (Though, by that argument, so would seem Mozart, too.) In any case, most people don't even remember the themes of an evening's concert -- yet, when the tunes are played again, they're recognized. Something must remain in the mind to cause this, and perhaps what we learn is not the music itself but *a way to hear it.* I'll explain.

Compare a sonata to a teacher! He calls attention, either dramatically or by the quiet trick of speaking softly. Next comes the careful presentation of the elements: useless to introduce too many ideas, or develop them too far; until the themes are learned, the listeners can't build on them. So, at first one repeats a lot. Sonatas, too, explain first one idea, then another, and recapitulate it all, just to be sure. **[2].**

Thus *Expositions* show the basic stuff, the atoms of impending chemistries, and how to make some simple compounds from those atoms. Then, in *Developments*, those now familiar compounds, made from bits and threads of beat and tone, can clash or merge, contrast or join together. We find it hard to remember details of things that don't fit easily into familiar frameworks -- things that seem meaningless. But I prefer to turn that around: a thing has meaning only to one who already knows some ways to represent and process what is meant -- who knows its parts and how they're put together. **[10].**

That's why sonatas start with simple things -- as do the best of talks and texts -- repeating basics several times before presenting larger things. No one remembers, word for word, all that was said in any lecture, or played in any piece. But if you understood it once, you now

own new networks of knowledge, about each theme and how it changes and relates to others. Thus, no one could remember Beethoven's Fifth Symphony entire, from single hearing. But neither could one ever hear again those first four notes as just four notes! Once but a tiny scrap of sound; it's now a Known Thing -- a locus in the web of all the other things we know, whose meanings and significances depend on one another. [3].

If sonatas are lessons, what are the subjects of those lessons? The answer's in the question! One thing the Fifth Symphony taught our culture is how to hear those first four notes. The surface form is just *descending major third, first tone repeated thrice.* At first, that pattern can be heard two different ways -- as *fifth and third in minor mode* -- or *third and first, in major.* But once we've heard the symphony, the latter is unthinkable -- a strange constraint to plant in all our heads! Let's see how it is taught.

The Fifth declares at once its subject, then its near-identical twin. First comes the theme. Presented in a stark orchestral unison, its minor mode location in tonality is not yet made explicit, nor is its metric frame yet clear: the subject stands alone in time. Next comes its twin. The score itself leaves room to view this transposed counterpart as complement -- or as a new beginning. Till now, fermatas hide the basic metric frame, a pair of twinned four-measure halves; so far we've only learned to hear those halves as separate wholes.

The next four-measure metric half-frame shows three versions of the subject, one on each ascending pitch of tonic triad. (*Now* we're sure the key is minor.) This shows us how the subject can be made to overlap itself, the three short notes packed perfectly inside the long tone's time-space. The second frame-half does the same with copies of the complement ascending the dominant seventh chord. This fits the halves together in a most familiar frame of harmony; in rhythm, too, the halves are so nearly congruent that there's no room to wonder how to match them -- and attach them -- into one eight-measure unit.

The next eight-measure frame explains some more melodic points: how to smooth the figure's firmness with passing tones, and how to counterpoise the subject's own inversion, inside the long note. (I think that this evokes a sort of sinusoidal motion-frame idea that later helps to represent the second subject.) It also shows compression of harmonic time; seen earlier, this would obscure the larger metric unit, but now we know enough to place each metric frame precisely on the after-image of the one before.

(Cadence.) (Silence.) (Almost.) (Total.)

Now it's the second subject-twin's turn to stand alone in time. Conductor must select a symmetry: to answer prior cadence, start anew, or close the brackets opened at the start; can he do all at once and still maintain the metric frame? In any case the student hears a long, long unison F (subdominant?) in which he gets to do his homework! For, underneath that silent surface sound, one hears one's mind rehearsing what was heard.

The next frame shows the theme again, descending now by thirds. (We see it was the

dominant ninth, not subdominant at all. Fooled us that time -- but never again.) Then, *tour de force*, the subject sounds ascending every scale degree. This new perspective shows us how to see the four-note theme as an appogiatura -- and then, descending on each tonic chord-note, we're shown how to see it as a fragment of arpeggio. (That last descent completes a set of all four possibilities: harmonic and directional. Is this deliberate didactic thoroughness, or just an accidental outcome of the other symmetries?) Finally, the subject's interval is squeezed to nothing, still surviving -- even gaining strength --as single tone.

(It always seemed to me a mystery of art: the impact of those moments in quartets when texture turns to single line, and *fortepiano* shames *sforzando.* But think: that very act, by which the surface shows the least, must make the largest difference underneath. Shortly, I'll propose a scheme in which such sudden, searching changes wake a lot of *Difference-Finders.* That very change wakes yet more *Difference Finders*, and that still more in turn. And that's how sudden silence makes the whole mind come alive.)

We're told all this in just one minute of the lesson, and I've touched but one dimension of its rhetoric; besides explaining, teachers beg and threaten, calm and scare, use gesture, timbre, quaver, sometimes even silence. (Most vital, that, in music, too. In fact, in the Fifth, it's the start of the subject!) Such "lessons" must teach us as much about triads and triplets as mathematicians have learned about angles and sides! Think how much we were told about minor second intervals, in Beethoven's treatise [B].

Why on earth should anyone want to learn such things? Geometry is practical -- for building pyramids, for instance -- but of what use is music-knowledge? [4]. Here's one idea. Each child spends endless days in curious ways; we call it "play". He plays with blocks and boxes, stacking them and packing them; he lines them up and knocks them down. What's that all about? Clearly, he is learning Space! But how, on earth, does one learn Time? Can one Time fit inside another, can two of Them go side by side? In Music we find out!

Many adults retain that play-like fascination with making large structures out of smaller things -- and one way to understand music involves building large mind-structures out of smaller music-things. So that drive to build music-structure might be the same one that makes us try to understand the world. (Or is it just an accidental mutant variant; evolution often copies extra needless stuff, and minds so new as ours must still contain some.)

Sometimes, though, we use our music as a trick to misdirect our ways to understand the world. When thoughts are hurtful, we have no way to make them stop. We can attempt to turn our minds to other matters, but doing this (some claim) just leaves the bad thoughts gnawing underneath. Perhaps that music some call "background" can tranquilize by turning under-thoughts from bad to neutral -- to leave the surface thoughts deprived of affect by diverting the unconscious. The "meanings" we assemble, in that detached kind of listening, could be wholly self-contained in solipsistic networks -- webs of meaning-like cross-reference that nowhere touch "reality". In such a self-constructed world, one needs no truth or falsehood, good or evil, pain or joy. Music, in this unpleasant view, serves as a

fine escape from tiresome thought.

SYNTACTIC THEORIES OF MUSIC

Contrast two answers to: "why do we like certain tunes?"

"---because they have certain structural features."
"---because they resemble other tunes we like."

The first looks for the laws and rules that make tunes pleasant. In language, we know laws for sentences; that is, we know the forms they need to have to be syntactically acceptable -- if not the things they need to make them sensible, or even pleasant to the ear. But as to *melody*, it seems, we only know some features that can help -- we know of none we cannot do without. I don't expect much more to come from searching for the formal rules of musical phrase.

The second seeks significance outside the tune itself, just as to ask "which sentences are meaningful?" takes us outside of shared linguistic practice into each person's own tangled webs of thought. And, there, these preferences feed upon themselves, as in all spheres: we tend to like things that re-mind us of the other things we like. So some of us like music that resembles songs and carols, rhymes and hymns we liked in childhood. But this begs the question: if we like tunes like ones we like, where does this music-liking start? I'll come back to that later.

The term *resemble* begs a question, too: what are the rules of musical resemblance? I'm sure that this depends a lot on how melodies are "represented" in each individual mind. And in each single mind, I'm sure, the different mind-parts do this different ways; the same tune seems (at different times) to change its rhythm, mode, or harmony. Beyond that, individuals differ more. Some listeners squirm to symmetries and shapes that others scarcely hear at all; some fine fugue subjects seem banal to those who sense but single line. My guess is that our contrapuntal sensors harmonize each fading after-memory with others yet to play; perhaps Bach's mind could do this several ways at once. (But even one such process might suffice, for choosing what one next should try to play. "Try" is enough for improvisers, since they -- like stage magicians -- know enough "ways out" to keep the music going, when bold experiments fail.)

"Feature-based" explanations can't begin to describe such processes as that. Much better are the "generative" and "transformational" (e.g., neo-Schenkerian) methods of syntactic analysis -- but only for the simplest analytic uses. For, at best, the very aim of syntax oriented theories is misdirected; they aspire to describe the things that minds produce -- without attempting to describe *how* they're produced. This only works on very simple

things; for complex ones, taxonomy must yield to causal explanation. Therefore, to really understand how memory and process merge in "listening", we'll simply have to use much more "procedural" descriptions -- that is, the kinds that can describe *how processes proceed.* [5]

I don't see why so many theorists find this disturbing. It's true that this new power has a price: one can *say* more, with computational description, but *prove* less. Yet not so much is lost as many think; for Mathematics never *could* prove very much about such complicated things. Theorems often tell us complex truths about the simple things, but only rarely tell us simple truths about the complex ones. Believing otherwise is wishful thinking, "mathematics envy".

And for that price, a gain. Many musical problems that resist formal solution will become tractable anyway, in future simulations that grow artificial musical semantic networks -- perhaps by "raising" simulated infants in traditional music cultures. [6]. It will be exciting when one of these first shows a hint of real "talent".

SPACE and TUNE

When one enters a room, one seems to see it all at once; not so with any symphony. "Naturally," one might declare, for "hearing extends in time while vision extends in space." But it really takes time to see new scenes, though we're usually unaware of this. That Consciousness sees seeing as so instant and immediate, is certainly the strangest of our "optical" illusions.

Still, music can immerse us in some stable-seeming worlds. I'll try to explain this by arguing that hearing music is like seeing scenery. Only, instead of treating this analogy light-heartedly, I'll hammer it to death -- asserting that good music *really* makes the mind act very much the way it does when seeing things. [7]. And no mistake: I meant to say "good" music! This little theory doesn't apply to every bag of musical tricks, but only to those certain kinds.

Our eyes are always flashing sudden flicks of different pictures to our brains -- yet none of that saccadic action leads to any sense of change or motion in the world; each thing reposes calmly in its "place"! What makes those objects seem so stable, when their images jump and jerk so? What makes us such innate Copernicans? I'll first propose how this might work in Vision, and then in Music. But first I have to say some things about the way the mind regards itself!

When we speak about illusion, we always talk in terms of someone being fooled -- be it another or oneself. "*I* know those lines are straight", one says, "but they look bent to *Me*".

Well, gosh, -- who are the different I's and Me's in that? We're all convinced that somewhere in each person struts a central Self, atomic, indivisible. (And secretly we hope that it is also indestructible.)

Instead, I say, inside each mind work *many* different "agents". I can't describe this theory here in much detail (its sources are describe in note [11]) but all we really need is this rough sketch: each agent knows what happens to some others, but little of what happens to the rest. In this view, we can see how little it can mean to say a thing like "Eloise was unaware of X" -- unless one says more about which of her mind-agents were uninvolved with X. Thinking itself is mainly making mind-agents work together; the very core of fruitful thought is breaking problems into parts, and then assigning every part to agents good at just that kind of job. Among our most important agents are those that make these management assignments, for they are the agents that embody what each person knows about what he knows. But "self-awareness" is a luxury, and usually an impractical one, for managers who really manage cannot afford to know everything that their subordinates do.

In that division of labor we call "seeing", I'll suppose that one agent of the mind -- called "*Feature-Finder*" -- sends messages (about features it finds on the retina) to another agent -- "*Scene-Analyzer*". The latter draws conclusions from the many messages it receives and sends its own, in turn, to other mind-parts. For instance, *Finder* might inform about some scraps of edge and texture; then *Analyzer* finds that these might fit some bit of shape.

Perhaps those features emanate from some part of a table-leg, but knowing such a thing is not for agents at this level. What Analyzer can do, though, is broadcast a shape-gram message to that host of other agents specialized for finding Vision-, a matter more involved with memory and learning. (There's one such agent, at the least, for every kind of thing this mind has learned to recognize.) Thus, we can hope, this message reaches *Table-Maker*, an agent specialized to gather evidence for tables in the field of view. After many such stages, descendants of those messages finally reach *Space-Builder* -- the agency that purports to inform yet other agencies about Real Things in real World-Places.

Now we can see one reason why perception seems so effortlessness: while the messages from *Analyzer* to *Table-Maker* are *based* on evidence that *Finder* supplied, the messages themselves need say little about *Finder* itself, or what it did. Partly this is because it would take *Analyzer* too long to explain all that; but in any case the recipients could make no use of all that information -- not being engineers or psychologists, but just little specialized nerve-nets. [8], also [11].

Messages, in this scheme, go various ways. Each motion of eye or head or body makes *Finder* start anew, and such motions are responses (by muscle-moving agents) to messages that *Analyzer* sends when he needs to resolve ambiguities, or fetch more details. *Analyzer* himself responds to messages from "higher up"; for instance, *Space-Builder* may have asked "is that a table?" of *Table-Maker*, who replies (to himself) "Perhaps, but it should have another leg -- there", so he asks *Analyzer* to verify this for him, and *Analyzer* gets the job

done -- by making *Eye-Mover* look down and to the left. [9].

When you look up, you're never frightened that the ground has disappeared -- although it certainly has dis-appeared. This is because *Space-Builder* remembers all the answers to its questions -- and never changes any of those answers without reason; moving eyes or raising head provides no cause to exorcise that floor inside your current spatial model of the room. My paper on "frame-systems" [F] says more about these concepts, but here we only need these few details.

Now, back to our illusions. While *Finder* is not instantaneous, it's very, very fast: a highly parallel pattern matcher. Whatever *Analyzer* asks, *Finder* answers in an eye-flick, a mere tenth-second or so (or less if we have image-buffers). And still more speed comes from the way we mentioned just above, in which *Space-Builder* can often tell himself, from his own high-speed model memory, about what's been seen before. I argue that all this speed is another root of our illusion: *if answers seem to come as soon as questions asked, they'll seem to have been there all along.*

The illusion is enhanced another way -- by "expectation" or "default". Those agents know good ways to lie and bluff! Aroused by only partial evidence for "table", *Table-Maker* supplies *Builder* with fictitious details about some "typical table" while its servants find out about the real one! Once so informed, *Builder* can quickly move and plan ahead, with little risk, ready to make corrections later. This only works, of course, if the prototypes are good -- but that's what intelligence is all about.

As for "awareness" of how all such things are done, there simply isn't room for that. *Space-Builder* is too remote and different to understand how *Finder* does its work of eye-fixation. Each part of mind is unaware of almost all that happens in the others. (That's why we need psychologists; we *think* we know what happens in our minds -- because those agents are so facile with "defaults" -- but, actually, we're almost always wrong.) True, each agent needs to know *which* of its servants can do *what* -- but as to *how*, that has no place or use inside those tiny minds inside our minds.

Finally, we return to how both music and vision build things in our minds. Eye-motions show us real objects; phrases show us music-objects. We learn a room with body motions; large music-sections show us "music-places". Walks and climbs move us from room to room; so do transitions between sections. Looking back in vision is like recapitulation in music; both give us time, from time to time, to revise or review conceptions of the whole.

So, hearing a theme is like seeing a thing in a room. An allegro is like the room itself, and the whole sonata is like an entire building. I do not mean to say that Music builds just the sorts of thing that *Space-Builder* does. (I know that's too naive -- comparing sound and place like in a little child's poem. Still, truth may lie in simple thoughts.) I do mean to say that composers stimulate coherency by engaging *the same sorts of inter-agent coordinations* that Vision uses to produce its illusion of a stable world -- of course, using different agents.

I think the same is true of talk or writing, the way these very paragraphs make sense -- or sense of sense -- if any.

COMPOSING AND CONDUCTING

In Seeing, we can move our eyes; Lookers can choose where they shall look, and when. In music one must listen *here* -- that is, the part played *now*; it's simply no use asking *Music-Finder* to look *there* -- because it isn't *then*, now.

But then, if composer and conductor choose what part you hear, does not this ruin our analogy? When *Music-Analyzer* asks its questions, how can *Music-Finder* answer them -- unless miraculously the music happens to be playing what it wants at just that instant? If so, then how can music paint its scenes -- unless composers know exactly what the listeners will ask at every moment? How to ensure, when *Analyzer* wants something -- *now*, that just that "something" will be playing -- *now*?

Well, that's just the secret of music, one end to the other, of writing, playing, and conducting! Music need not, of course, confirm the listener's every expectation; each plot demands some novelty. Still, whatever the intent, control is required, or novelty will turn to nonsense. And if allowed to think too much himself, the listener will find unanswered questions in any score, about accidents of form and figure, voice and line, temperament and difference-tone.

So, every music artist must anticipate and pre-direct the listener's fixations; he draws attention *here*, distracting it from *there*: forcing hearer (again, like a magician) to ask just questions that the composition is about to answer. Only by establishing such pre-established harmony, can music make it seem that something's there.

RHYTHM AND REDUNDANCY

A popular song has a hundred measures, a thousand beats. What must the Martians imagine we mean by those measures and beats, measures and beats! The words themselves reveal an awesome repetitiousness. Why isn't music boring?

Is hearing so like seeing, that we need a hundred glances to build each music-image? Some repetitive musical textures might serve to remind us of time-persistent things like wind and stream. But much of sound is one-time: you must hear a pin drop *now* or seek and search for it; that's why we have no ear-lids. Poetry drops pins -- says each thing once, or less. So

does some music.

Why, then, do we tolerate music's relentless rhythmic pulse? There's no one answer, for we hear in different ways, on different scales. [10]. Some of those ways portray the spans of time directly, but others speak of musical *things*, in worlds where time folds over on itself. And there, I think, is where we use those beats and measures.

What I mean is this: Music's metric frames are transient templates -- used for momentary matching. Its rhythms are "synchronization pulses" used to match new phrases against old -- better to notice Differences and Change. As these are sensed, the rhythmic frames fade from our awareness, their work done; the messages of higher-level agents never speak of them, and that's why music isn't boring!

Differences and Change! Good music grows from tiny roots, by careful steps. And see how cautiously we handle novelty, enclose the new -- like sandwiches -- between repeated sections of familiar stuff! The clearest kind of change is near-identity -- in Thought just as in Vision. Slight shifts in view may best reveal an object's form -- or even that it's there at all.

When we discussed sonatas, we saw how different metric frames are matched together, making it easy to discern the musical ingredients. Once frames were matched then we could see how what was *there* a major third was changed to seconds *here* by adding passing tones; that what was once a seventh-chord is now a dominant ninth. Thus matching lets our minds see different things, from different times, together. Fusing all those matching lines of tone from different measures -- like Television's separate lines and frames -- it's that lets us make those magic music-pictures on our mind-screens.

How do we make our Music-Agents do this kind of work for us? We must have them organized in some special structure specialized for finding differences between frames. Here's a four-level scheme that might work:

Feature-Finders listen for simple time-events,
like notes, or peaks, or pulses
Measure-Takers notice certain patterns of time-events
like 3/4, 4/4, 6/8
Difference-Finders observe that the figure here is
same as that one there, except a perfect fifth above.
Structure-Builders perceive that three
phrases form an almost regular "sequence".

I won't give much detail; that's for the future. But many such ideas are seen already in research on Vision. [11]. First, the *Feature-Finders* search the sound-stream for the simplest sorts of musical significance: entrances and envelopes, the tones themselves, the other little, local things. Then *Measure-Takers* look for metric patterns in those small

events, and put them into groups, thus finding beats and postulating rhythmic regularities. Then the *Difference-Finders* can begin to sense events of musical importance -- imitations and inversions, syncopations and suspensions. Once these are found, the *Structure-Builders* can start work on a larger scale.

The entire four-level Agency is just one layer of a larger system in which analogous structures are repeated on larger scales. At each scale, another level of order (with its own sorts of Things and Differences) makes larger-scale descriptions, and thus consumes another order of structural form. As a result, notes become figure and those turn to phrase, and those into sequence -- and notes become chord, and those make progression, and -- so, on and on. Relations at each level, turn to Thing at next above; more easily remembered and compared. This time-warps things together, changing tone into tonality, and line into polyphony.

The more regular the rhythm, the easier the matching goes, and fewer Difference agents are excited further on. Thus, once used for "lining up," the metric structure fades from your attention, because it's represented as a fixed and constant object-frame (like the floor of the room you're in) -- until some metric alteration makes the measure-takers change their minds. [12].

And so the regularities are hidden from attention while expressive nuances are sensed and emphasized and passed along. Rubato or crescendo, ornament or passing tone -- the alterations at each level become the objects for the next. The mystery is solved; the brain's so good at sensing difference at each stage that it forgets the things themselves, whenever they're the same. And as for liking music, that depends on what remains.

"SENTIC" SIGNIFICANCE

Why do we like any tunes in the first place? Do we simply "associate" some tunes with pleasant experiences? Should we look back to the tones and patterns of mother's voice or heartbeat? Or, could it be that some themes are *innately* likable? All these theories could hold truth, and others too -- for nothing need be single-cause inside the mind. [13].

In [C], Manfred Clynes, a physiologist and pianist, describes certain specific temporal sensory patterns, and claims that each is associated with a certain common emotional state. For example, in his experiments, two particular patterns (that gently rise and fall) are said to suggest states of love and reverence; two others (more abrupt) signify anger and hate. He claims that these and other patterns -- he calls them "sentic" -- arouse the same effects through different senses -- that is, embodied as acoustical intensity, or pitch, or tactile pressure, or even visual motion -- and that this is cross-cultural. The time-lengths of these sentic shapes, the order of one second, could correspond to parts of musical phrases.

Clynes studied the muscular details of instrumental performances with this in view, and concluded that music can engage emotions through these sentic signals. Of course, more experiments are needed to verify that such signals really have the reported effects. Nevertheless, I would quite expect to find something of the sort for a quite different reason; namely, to serve in the early social development of children. For, sentic signals, if they exist, would be very useful in helping infants to learn about themselves and others.

All learning theories require brains to somehow impose "values" on events -- implicit or explicit in the choice of what to learn to do. Most such theories say that certain special signals are involved in this, called *reinforcers*. For many goals, it should suffice to use some simple, "primary" physiological stimuli like eating, drinking, relief of physical discomfort. But human infants must learn social signals, too. The early learning theorists in this century assumed that social sounds (e.g., of approval) could *become* reinforcers -- by association with innate reinforcers -- but real evidence for this was never found. If parents could exploit some innate sentic cues, this might explain that mystery.

This might also touch another, deeper problem: of how an infant forms an image of its own mind. Self-images are important for at least two reasons. First, external reinforcement can be but one part of human learning; the growing infant must eventually learn to learn from inside -- to free itself from parent. With Freud, I think that children replace and augment the real one with self-constructed inner parent-image. Second, one needs a self-model simply in order to make realistic plans for solving ordinary problems. For example, one must know enough about one's own disposition to be able to assess which plans are feasible. Pure self-commitment doesn't work; one simply can't carry out a plan that one will find too boring to complete, or that is too vulnerable to other, competing interests. Both reasons point to needs for models of one's own behavior. But how could a baby be smart enough to build such a model?

Innate sentic detectors could help, by teaching children about their own affective states. For, if distinct signals arouse specific states, the child can associate those signals with those states. Just knowing that the states exist -- *that is, having symbols for them* -- is half the battle. If those signals are the same in others as in oneself then, from social discourse, one can learn some rules about those states' behavior. Thus a child might learn: *Conciliatory* signal changes *Angry* to *Affectionate*. Given that sort of information, a simple learning machine should be able to construct a "finite-state person-model". This model would be crude at first, to be sure -- but half of that job, too, is getting started. And once the baby has a crude model of some Other, he can copy and adapt it to begin work on making his model of himself.

Returning to music, it seems just barely possible that we conceal, in the innocent songs and settings of our children's musical cultures, some lessons about successions of our own affective states. Sentically encrypted, those ballads could encode instructions about conciliation and affection, aggression and retreat -- just the sorts of knowledge of signals and states that we need to get along with others. In later life, more complex music might

illustrate more intricate kinds of conflict and compromise, ways to fit goals together to achieve more than one thing at a time. Finally, for grown-ups, our Burgesses and Kubricks fit Beethoven's Ninths to Clockwork Oranges.

If reader finds all this far-fetched, so do I. But before rejecting it entirely, recall the problem: *why do we have Music, and let it occupy our lives with no apparent reason*? When no idea seems right, the right one must seem wrong.

THEME AND THING

Beethoven's Fifth: what is its subject; is it just those first four notes? Does it include the twin, transposed companion, too? What of the other variations, augmentations and inversions? Do they all stem from a single prototype? In this case, yes.

Or do they? For later in the symphony the theme appears in triplet form, to serve as counter-subject of the scherzo. Three notes and one, three notes and one, three notes and one; still they make four. Melody turns into monotone rhythm; meter's converted to two equal beats. Downbeat now falls on an actual note, instead of a rest. With all those changes, the themes are very different, yet still the same. Neither allegro or scherzo subject alone can be the single prototype; separate and equal, they span across musical time.

Well, then, is there some more abstract idea they both embody? This is like the problem, raised by Wittgenstein in [W], of what words like "game" mean. In [F] I argue that, for Vision, "chair" can be described by no single prototype; better to use several prototypes, connected in relational networks of samenesses and differences. But I doubt that these serve well to represent musical ideas; there are better tools in contemporary Artificial Intelligence research, like constraint systems, conceptual dependency, frame systems and semantic networks. That's where the action is, today, in dealing with such problems. [A].

Still, what we really want to know is: "What is a *good* theme?" Without that bad word, "good", I don't think the question is well-formed -- because anything's a theme, if everything is music!

Now, let's split that question into (1) "what mental conditions or processes do pleasant tunes evoke?" and (2) "what do we mean by "pleasant"? Both questions are hard, but the first is *only* hard; to answer it will take much thought and much experiment. Good.

The second question is very different. Philosophers and scientists have struggled mightily to understand what pain and pleasure are. I especially like Dennett's [D] explanation of why that's been so difficult. He argues that there simply isn't any one such thing as "pain" at all; instead "it" works in different ways at different times, and all those ways have too little

in common for the usual sorts of definition. He's right, I think, but then -- if pain's no single thing -- why do we talk and think as though it were, and represent "it" with such spurious clarity?

I claim this is no accident: illusions of this sort have special uses. They play a role connected with a problem facing any Society (in or outside the mind) that learns from its experience. The problem is how to assign the credit and blame, for each accomplishment or failure of the Society as a whole, among the myriads of agents involved in everything that happens. To the extent that the agents' actions are decided locally, so must also *these* credit decisions be made locally.

How, for example, can a mother tell that her child has a need (or that one has been satisfied), before she has learned specific signs for each such need? That could be arranged, by evolution gathering together signals from many different internal processes concerned with needs, and providing them with a single, common, output -- an infant's sentic signal of discomfort (or contentment). By a genetically pre-established harmony, this would evoke a corresponding central state in the parent. We'd feel this as something like distress, as we do when babies cry.

The satisfaction side needs signals, too. Suppose, among the many things a child does, there's one that mother "likes" -- and so she makes approving sounds. The child's just been walking *there*, and holding *this* just so; and thinking *that*, and speaking in some certain way. How can its mind find out *which what* was "good"? The trouble is, each thing it did just *then*, must owe, in turn, to little plans it made before. We can't "reward" a *going to some certain place*: you can't reward an act. One only can reward the agency that *selected* that strategy -- and the agency who wisely activated *that* agency -- and so on.

To do this, one must propagate, to all those agencies and processes, some message that they all can use to evaluate what they did -- the plans they made, their strategies and the things that they constructed. But all these various recipients have so little in common that such a message, to work at all, must express the essence of oversimplification. So "good" is a centralization that enables a tutor, inside or outside a society, to tell its members that one or more of them has done something good -- that satisfies some need -- without having to understand which ones, or how, or even why.

Now words like "satisfies" and "need" have many senses, yet we seem to understand that phrase -- the same illusion of substantiality that fools us into thinking it tautologous, unworthy of study, to ask *"why do we like pleasure"*? The social discourse levels where we use such clumsy words as "like," or "good," or "that was fun" must (by that very poverty of word and sign) most coarsely crush together many different meanings. "Good" is no symbol that *means* as "table" does. Instead, it names an injunction: "activate all those (unknown) processes that correlate and sift and sort, in learning, to see what changes (in myself) should now be made". "Like", then, is a name we use for when we send such structure-building signals *to ourselves.* **[14].**

And that's another reason we like music. Liking's just how certain mind-parts make the others learn the things they need to understand that music. Hence *liking* (and its relatives) are at the very heart of understanding what we hear. So Affect and Aesthetic do not lie in other academic worlds, that music theories safely can ignore. Those separate worlds are Academic-Self deceptions, used to make our problems seem like someone else's. **[15]**

NOTES

[1]. I don't mean that understanding Emotion is easy, only that understanding Reason is harder. Our culture has a universal myth in which we see emotion as more complex and obscure than intellect. Indeed, emotion might be "deeper" in some sense of prior evolution, but this need not make it harder to understand; in fact, I think today we actually know much more about emotion than reason.

Now, to be sure, we know a bit about the surface ways of Reason -- the ways we organize and represent ideas we get. But whence come those ideas themselves -- that so conveniently fill these envelopes of order? A poverty of language shows how little this concerns us: we "get" ideas; they "come" to us; we are "reminded of". I think this shows that ideas come from processes obscured from us, with which our surface thoughts are almost uninvolved. Instead, we are entranced with our emotions -- so easily observed, in others and ourselves. Perhaps the myth persists because emotions (by their nature) draw attention, while the processes of Reason (much more intricate and delicate) must hide themselves in privacy, to work best left alone.

In any case, those old distinctions -- Feeling, Reason and Aesthetic -- are like the Earth and Air and Fire of an ancient Alchemy. We'll need much better concepts for a working psychic chemistry.

[2]. Music has many forms, and there are many ways to teach. I do not say that Beethoven consciously intended to teach at all, but still, he was a master of inventing forms for exposition -- including forms that swarm with more ideas, and work our minds much harder.

[3]. Learning to recognize is not the same as memorizing. A mind might build an agent that can sense a certain stimulus, yet build no agent that can reproduce it. How could such a mind learn that the first half-subject of Beethoven's Fifth -- call it "A" -- pre-figures the second half-- call it "B"? Simple: an agent *A* that recognizes A sends a message to another agent *B*, built to recognize B. That message serves to "lower *B*'s threshold" so that after *A* hears "A", *B* will react to smaller hints of B than it would otherwise. Result: that mind "expects" to hear B after A, e.g., it will discern B, given fewer or more subtle cues, and might complain if it cannot. And yet that mind cannot reproduce either theme, in any generative sense. The point is that inter-agent messages need not be in surface music-languages, but in codes that bias certain other agents to behave in different ways.

Andor Kovach pointed out to me that composers don't dare use this simple, four-note motive any more. So memorable was Beethoven's treatment, that now an accidental hint of it can wreck another piece by moving listener's mind into that other, unintended place.

[4]. True or not, it's often said that Mathematicians are unusually involved with music, but not the other way around. Perhaps both share a liking to make simple things more

complicated -- but Mathematics may be too constrained to satisfy that want entirely, while Music can be rigorous *or* free. The way the Mathematics game is played, most variations lie outside the rules, while Music can insist on perfect canon, or tolerate a casual accompaniment. So Mathematics might need Music, too, but not the other way around. A simpler theory: since music engages at earlier ages, some mathematicians *are* those missing mathematical musicians.

[5]. In science one always first explains in terms of what can be observed. (Earth, Water, Fire, Air.) But things that come from complicated processes need not show their natures on the surface. (The steady pressure of a gas conceals those countless, abrupt micro-impacts.) To talk of what such things might mean or represent, one has to speak of how they're made.

But one can't do that, for the mind, without good ways to describe complicated *processes.* Before computers came, *no* languages were good for that. Piaget tried algebra and Freud tried diagrams, other psychologists used Markov chains and matrices -- none came to much. Behaviorists, quite properly, had ceased to speak at all. Linguists flocked to formal syntax, and made progress for a time, but reached a limit: transformation grammar shows the contents of the registers (so to speak), but has no way to talk of what controls them. This makes it hard to say how surface speech relates to underlying designation and intent -- a baby-bath condition. The reason I like ideas from Artificial Intelligence research is that there we tend *first* to seek procedural description -- which seems more right for mental matters.

[6]. We can already write computer programs that write better music than most people can, but still that music is quite "bad." I don't know quite what to do about this; our lowest standards are so high that we find it hard to distinguish early progress from worthless noise.

[7]. Edward Fredkin suggested to me the theory that music-listening might exercise some innate "map-making" mechanism. When I mentioned the puzzle of music's repetitiousness, he compared it to how rodents explore new places: first one way a little, then back to home. Do it again a few times. Go a little further. Try small digressions, but frequently return to base. Both men and mice explore new territories that way, making mental maps lest they get lost. Music might portray this building process, or even exercise those very organs of the mind.

[8]. Only in the past few centuries have painters learned enough technique and trickery to simulate reality. (Once so informed, they often now choose different goals.) Thus, *Space-Builder,* like an ordinary person, knows nothing of how vision works, nor of perspective, foveae, or blind-spots. We only learn such things in school: millenia of introspection never led to their suspicion -- nor meditation, transcendental or mundane. The mind holds tightly to its secrets -- not from stinginess, nor shame, but simply that it does not know them.

[9]. Nor is *Scene-Understander* autonomous; his questions to *Analyzer* are responses to

requests from others. There need be no First Cause, in such a network. See [13].

[10]. What is the difference between merely *knowing* (or remembering, or memorizing) and *understanding*? We all agree that to understand something you must know what it means -- and that's about as far as we ever get. I think I know why that happens. A thing or idea seems meaningful, only when we have *several* different ways to represent it -- different perspectives and different associations. Then you can turn it around in your mind, so to speak: however it seems at the moment, you can see it another way; you never come to a full stop. In other words, one can *think* about it. If there's only one way -- if it just sits in your mind doing nothing -- you wouldn't call it thinking.

So a thing has any real "meaning" only when it has several; if you understood it just one way, you didn't understand at all. That's why the seekers of the "real" meanings never find them -- and this holds true especially of words like "understand" itself!

[11]. The idea of interconnecting *Feature-Finders*, *Difference-Finders*, and *Structure-Builders* are well exemplified in Winston's work, [L]. Measure-Takers would be kinds of Frames, as described in [F]. The idea of "societies of agents", developed in [P], [K], and [J] comes from my work with Seymour Papert.

[12]. Rhythm has other roles, of course. Societies (in minds or lands) with different functions see things different ways. Agents used for dancing do attend to rhythm, while other forms of music demand less steady pulses.

We all experience a phenomenon we might call "persistence of rhythm" -- in which our minds maintain the "beat" through episodes of ambiguity. I presume that this emerges from a basic feature of how agents are usually assembled: at every level, many agents of each kind compete. [K]. Thus, agents for 3/4, 4/4, and 6/8 compete to find best fits. However, once in power, each agent "cross-inhibits" its competitors. Thus, once 3/4 takes charge of things, 6/8 will find it hard to "get a hearing" -- even if the evidence on his side becomes slightly better.

When none of them has any solid evidence for long enough, then agents change at random, or take turns. Thus, anything gets interesting -- in a way --if monotonous enough! We all know how, when word or phrase is oft enough repeated, it -- or we -- begin to change; because the restless Searchers start to amplify minutiae, interpret noise as structure. This happens at all levels: for when things are regular at one, the Difference agents at the next will fail -- to be replaced by other, fresh ones that then re-present the sameness different ways. (Thus Meditation, undirected from the higher mental realms, fares well with the most banal of inputs from below.)

[13]. Theories about children need not apply to adults, because (I suspect) human minds do so much self-revising that things can get detached from their origins. One might end up liking both *Art of Fugue* and *Musical Offering*, mainly because each one's subject

illuminates the other, giving each a richer network of "significance". Dependent circularity need be no paradox here, for, in thinking (unlike logic) two things *can* support each other in mid-air. To be sure, such autonomy is precarious; once detached from origins, might one not drift strangely awry? Indeed so, and many people seem quite mad, to one another.

[14]. Most of the "uses" of music mentioned here -- learning about time, fitting things together, getting along with others, and suppressing one's troubles -- all seem very "functional" and overlook much larger scales of "use". Curt Roads pointed out to me: "every world above bare survival is self-constructed; whole cultures are built around common things people come to appreciate". These appreciations, represented by aesthetic agents, play roles in more and more of our decisions -- as what we think is beautiful gets linked to what we think is important. Perhaps, Roads suggests, when groups of mind-agents can't agree, they tend to cede decisions to those others concerned with what we call aesthetic form and fitness. With small effects at many little points, those cumulative preferences for taste and form can shape a world.

[15]. Many readers of a draft of this complained about its narrow view of music -- what about Jazz, or "modern" forms, or songs with real words; or monophonic chant and raga, gong and block, and all those other kinds of sound? And several claimed not to be so intellectual, to simply hear and feel and not build buildings in their minds. There simply isn't room to talk of all those things; besides, no composition can please everyone. I'll say just two more things.

First thing -- for those who argue music doesn't make them do so much construction: what makes you sure you know your mind so surely? It is ingenuous to think that you can "just react" to anything a culture works a thousand years to grow. A mind that thinks in terms of direct apprehension has more in its unconscious than it has in its philosophy.

Second: it makes little sense to vilify a view of music as insufficiently comprehensive. For what is "music" anyway -- all things played on all instruments? Fiddlesticks. All structures made of sound? That has a hollow ring. The things I said of the word "tune" hold true for "music", too: it does not follow that because a word be public, so must also be the ways it works on minds. Before one seeks the grail that holds the essence of all music, first see the folly of a simpler quest: to grasp the essence of one single noise, that "music" word itself.

ACKNOWLEDGMENTS

Conversations and/or improvisations with Maryann Amacher, John Amuedo, Betty Dexter, Harlan Ellison, Edward Fredkin, Bernard Greenberg, Danny Hillis, Douglas Hofstadter, William Kornfeld, Andor Kovach, David Levitt, Tod Machover, Charlotte Minsky, Curt Roads, Gloria Rudisch, Frederic Rzewski, Stephen Smoliar. In memory of Irving Fine.

REFERENCES

I apologize for lack of scholarship; I'm sure these ideas have many more antecedents. There are good reviews of recent work on musical cognition in [A].

[A]: *Computer Music Journal*, vols. 4(2), Summer 1980, 4(3), Fall, 1980.

[B]: Beethoven, Ludwig v. *Grosse Fugue*, in E-flat, Op. 133.

[C]: Clynes, Manfred. *Sentics*, Doubleday, New York, 1978.

[D]: Dennett, Daniel."Why a Machine Can't Feel Pain" in *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, Vermont: Bradford Books, 1978.

[F]: Minsky, Marvin. *A Framework for Representing Knowledge*, M.I.T., Artificial Intelligence Laboratory, AI Memo 306. Cambridge, Ma., June 1974. Condensed version in *The Psychology of Computer Vision*, Edited by P. H. Winston. New York: McGraw-Hill, 1975.

[J]: Minsky Marvin. *Jokes and the Logic of the Cognitive Unconscious*, M.I.T., Artificial Intelligence Laboratory, AI Memo 603. Cambridge, Ma., Nov. 1980.

[K]: Minsky, Marvin. "K-lines: A Theory of Memory." *Cognitive Science*, Vol. 4, No. 2 (April 1980), 117-133.

[L]: Winston, P.H. "Learning Structural Descriptions by Examples." *Psychology of Computer Vision*, Edited by P. H. Winston. New York: McGraw-Hill, 1975.

[P]: Minsky, Marvin. "Plain Talk About Neurodevelopmental Epistemology." *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, Ma., August 1977. Condensed version in *Artificial Intelligence*, edited by Winston and Brown, Vol. 1, MIT Press, 1979.

[W]: Wittgenstein, Ludwig. *Philosophical Investigations*, Oxford, 1953.