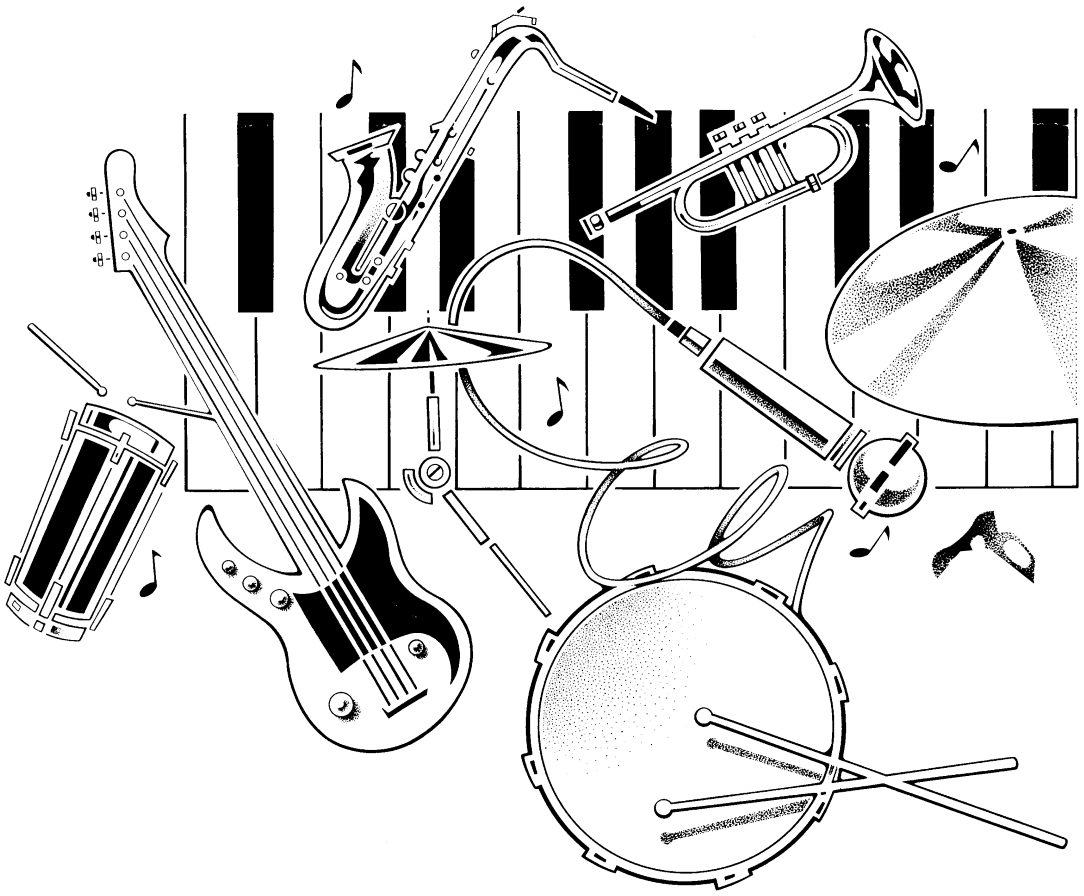
AI Magazine Volume 13 Number 3 (1992) (© AAAI)



A Conversation with Marvin Minsky

Marvin Minsky and Otto Laske

The following excerpts are from an interview with Marvin Minsky which took place at his home in Brookline, Massachusetts, on January 23rd, 1991. The interview, which is included in its entirety as a Foreword in the book Understanding Music with AI: Perspectives on Music Cognition *(edited by Mira Balaban, Kemal Ebcioglu, and Otto Laske), is a conversation about music, its peculiar features as a human activity, the special problems it poses for the scientist, and the suitability of AI methods for clarifying and/or solving some of these problems. The conversation is open-ended, and should be read accordingly, as a discourse to be continued at another time.*

…"is formalizing the right idea?"

The Notion of "Formalizing" Musical Knowledge

OL: I wish we could read people's brains when they are engaged in music-making. Since that is, alas, beyond the state of the art of AI, we are forced to deal with such epistemological issues as how to represent knowledge on the basis of what people tell us they do, which in most cases isn't very close to what they are actually doing. A further problem is that we can't use what they are telling us in the form it is reported; rather, we have to translate their verbalizations into even more formal code that is still farther removed from actual musical activities than the initial report. So, of course, two questions immediately come to mind, viz., how should one formalize musical knowledge, and, can it be done effectively, i.e., so as to generate some kind of musical action?

MM: And that in turn raises the question of "is formalizing the right idea?" There are many kinds of reasons for writing descriptions, and here we ought to have a softer concept than "formalize." When we write down things we know about other crafts, we don't usually feel any need to imitate mathematicians. Why should we always feel compelled to do that when describing our musical dispositions?

OL: Indeed, this convention is not much questioned. One thinks backwards from one's research goals, and since one knows how to compute, one concludes with little hesitation that one ought to formalize. The term "formalizing" is linked to the notion of computation …

MM: Quite so—ever since the earliest days of computing. Maybe because most of the first pioneers were concerned with mathematical matters. Consequently, what we call "computer science" quickly became quite technical. In some domains of computer development that was a good thing. But in other areas, particularly in the semantics of programming, it seems to me that the currently popular formalisms are premature and unnecessarily limited in conception, being inadequate for expressing some of the most important aspects of computation. For example, mathematical logic plays a dominant role in contemporary formalizations, and yet is quite inept at expressing the kinds of heuristic knowledge we need for developing Artificial Intelligence.

OL: "Formalization" typically amounts to making a post facto summary of some state of affairs that is based on specific interpretations. One looks at some description and attempts to "straighten out" things that don't fit the chosen formalism, in order to make them look more "logical."

Problems with Formalization

MM: Well, yes, I like how you put that. Indeed we can use logic to make things look more logical but we can also use it to produce illusions. I suppose that the term originally referred to matters involving understanding and knowledge. But in current computer science "logical" now means mainly to try to express everything in terms of two quantifiers, viz., "for all X," and "for some X" and two values, viz., "true" and "false." But those menus seem just too small to me! We need much richer ideas than that.

OL: In the past, we've had other logics, such as, for instance, Hegel's dialectical logic where one doesn't work with two quantifiers, but with the notions of thesis, antithesis, and synthesis. (Antithesis dates from Plato's idea of "Other" in the dialog entitled Sophistes.) So, one always knew there is some other way …

MM: Well, I'm not familiar with any applications of dialectic logic, but I certainly would favor going beyond binary logics. However, the extensions that I've seen in the so-called modal logics seem no improvement to me. Yes, they can express some ideas like "it is possible," or "it is necessary," and so forth, but extensions like those seem inconsequential. We much more urgently need ways to express ideas like "usually" or "it is often useful to," or "for all objects that resemble X in regard to aspect Y."

OL: So you think computer science—or at least AI—needs other, richer kinds of descriptive ideas and terms?

The Need for a Variety of Methods

MM: Yes indeed. I want AI researchers to appreciate that there is no one "best" way to represent knowledge. Each kind of problem requires appropriate types of thinking and reasoning—and appropriate kinds of representation. For example, logical methods are based on using rigid rules of inference to make deductions. This works well inside formal domains—those artificial worlds that we imagine for ourselves. But to cope with the unknowns and uncertainties of reality, we must base our actions on experience—and that requires us to reason by analogy, because no two situations are ever quite the same. In

0738-4602/92/$4.00 ©1992 AAAI

The modern tendency, to… say "anything is music," is bad for the mind…

turn, this means that we have to be able to recollect similar experiences and understand which differences are relevant. But in the logical world, ideas like "similar" and "relevant" are alien because logic can only answer questions like "what is this an instance of?" or "what is this a generalization of?" The trouble is that concepts like instance and generalization apply only to ideas—because no actual object or event can be an instance of anything else. However, real things can be seen as related—at least in an observer's mind—by apparent similarities of structures, effects, or useful applications. Certainly this is the case in music.

OL: Yes, I think reasoning in terms of similarity is very pertinent to music, especially since notions like "similarity" and "contrast" are often the only ones that give a musical actor a handle on what he is dealing with (especially in the sonic domain, i.e., in sound synthesis and orchestration). And since much of musical knowledge is really action knowledge (i.e., knowledge derived from, and destined for, pursuing action), that is of great relevance to music-making. Music is something we do, not just something we understand, and much of what we try to understand regarding music is meant to lead to the making of it.

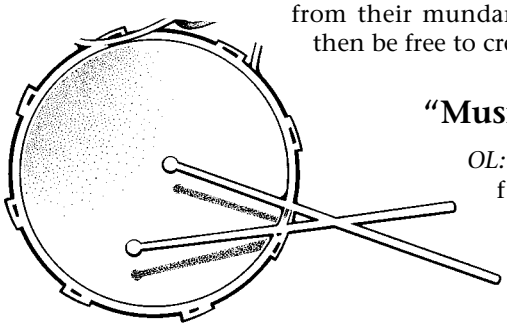
MM: And it is hard to understand all that making and understanding because they involve so many different mechanisms—I like to think of them as like many different animals inside each brain. When you hear a piece of music, different parts of your brain do different things with it, but we know too little about those different processes. One obstacle to understanding such matters is that psychologists still strive for a science that resembles physics, where it is usually better to treat different explanations as competitive, instead of looking for ways to combine them. That strategy indeed works well in physical science—presumably because there actually are only a very few fundamental laws of nature. But our ancestors had to deal with many different kinds of practical problems, and this led to the evolution of brains that have hundreds of distinct regions with significantly different micro-architectures and different principles of operation.

OL: You were saying that in listening to music, a lot of things are happening simultaneously, and our task seems to be to understand the interrelationship between the different structures and processes involved in musical reaction and understandings.

MM: Yes, and we can still only guess what some of them are. Certainly, our musical apprehensions involve quite a few partially separable processes involved with rhythm, melody, harmony, timbre, texture, and many other local phenomena—and each of these appears to involve multiprocessing aspects of their own, such as timbral and contrapuntal voice separation. Sometimes it seems that one can sense some of the distinctness of those processes, as when it seems that one part of the mind is annoyed at the monotonously repetitive rhythmic structure of a certain composition—while other parts of the mind don't mind this at all—perhaps because they treat those repeating structures as structures not deserving attention themselves but serving as skeletons or scaffoldings, like a branching Christmas tree on which you hang the decorations. In this view, the significant features are the higher level differences between musical portions or segments that are otherwise extremely similar or analogous. It is those higher-level recognitions that let us treat the repetitive aspects of the music not as an irritating monotony but merely as a textural background.

OL: In all media of communication one encounters a lot of redundancy, just to make it possible to get across those few gems …

MM: Precisely. But still, perhaps among all the arts, music is distinguished by this sublimely vulgar excess of redundancy, and we should try to understand its possible neurological consequences. I think Lukas Foss once remarked that anything repeated often enough can become interesting. Perhaps this phenomenon can be seen not as a paradox but as evidence that supports the idea of multiple processing levels. The function of the repetition is then to anesthetize the lower levels of cognitive machinery. (We know that this is the usual rule in neurology: decay of response to constant or repetitive signals.) But the result of this could be to suddenly and strangely free the higher levels of the brain



from their mundane bondage to reality—to then be free to create new things.

"Music" or Musics?

OL: Would you say that holds for all kinds of music? "Music" seems to be a notion like "God" or "love," something everybody can identify with, but which actually covers many different, even opposite, phenomena ...

MM: I think calling so many different things by the same name, "music," certainly makes it hard to think about those subjects. The modern tendency, to be tolerant and say "anything is music," is bad for the mind, both of the listener and of the critic also. (I don't know whether it's bad for the composer, if he can make a living.) A portmanteau word like "music" that is used for so many activities cannot be an element of a serious discussion.

OL: So, we would have to speak of "musics," and would have to define what we mean in any particular instance, which would be very cumbersome.

MM: Yes, like having to say we are going to talk about certain German music from the eighteenth century, or Indian music from such a place at a certain time. And then one can ask, to what extent do these engage similar mental activities?

OL: How should we proceed, music being really a universal?

MM: For serious analysis, I think we simply must avoid such universal. When I write about some mechanism of intelligence or learning, I try not to use words like "intelligence" and "learning"—except in the title or summary.

Music as a Label for Societal Acceptance

OL: That is indeed the usual practice in studies in AI and music, where you find, for instance, systems for doing harmonic analysis, or systems that generate compositional material—these are all specific kinds of things, and the claim is not that we know what music is. It seems to me that, viewed in the light of such studies, the term "music" expresses rather an acceptance on the side of society that something is o.k. The composer, as composer, doesn't care whether something is music. He is driven to generate something, and if an audience finds that what he produced is acceptable, or is "music," he is likely to be happy, and otherwise he is not—but it's essentially not his doing as much as it is society's.

MM: That raises an exciting and interesting question—that many people are reluctant to consider, which is the question: what is, or ought to be, o.k.? One thing I like to do is to consider the major human activities, and try to get people to ask: "is it ok?" All over the world many people listen to music for hours each day; in this country many spend substantial portions of their incomes on recordings, high-fis, personal earphone devices, rock concerts, and tolerate background music in their workplaces, restaurants, airplanes, and what not. Is that all right? Similarly, we ought to wonder whether it is reasonable to engage in sports. I ask people, "isn't there something funny about grown people gathering in a huge stadium to see other grown people kicking a ball from one end to the other?" Each of those persons is using a multitrillion synapse brain. It would be fun to ask the religious ones to consider whether it is not a sin to waste such wondrous hardware on watching adults kicking balls around? My own view is that this is less a sin than a symptom—of infection by a parasitic meme (namely, one that carries the idea that such an activity is o.k.) which has self-propagated through our culture like a software virus, a cancer of the intellect so insidious that virtually no one thinks/dares to question it. Now, in the same way we see grown people playing and working in the context of popular music that often repeats a single sentence or melodic phrase over and over and over again, and instills in the mind some harmonic trick that sets at least part of one's brain in a loop. Is it o.k. that we, with our hard earned brains, should welcome and accept this indignity—or should we resent it as an assault on an evident vulnerability?

Music as a Device for Directing Human Activity

OL: You have suggested somewhere that music is often used for escaping the painfulness of thought.

MM: Certainly that is how music seems to be used at times. It seems very much as though a person can exploit music as an external intervention (in contrast to using an internal, perhaps chemical, regulatory system) to suppress one or another part of the brain—e.g., parts that might otherwise be occupied with sexual or social or other types of thinking which that person may presently not want to entertain. Clearly, people use

music for directing their mental activities. After all, that's what it means when we speak of music as stimulating, or as soothing, as like an opiate for relieving pain or anxiety. Or encouraging us to march and fight, or to sorrow at a funeral.

OL: The idea that music can be used to control inner states is, of course, a very old idea. There is after all such a thing as music therapy…

MM: There are also other ways of listening to music that people like you and I use a lot. Sometimes when hearing some music, I react in the ways we just mentioned, but at other times, when my attention is drawn to music, I find myself more concerned with how to make music that sounds that way. What's that sequence? How would you finger it? How did the composer get that idea? What would I have to learn to be able to do that myself? Heavens, how appropriate to double the horn and the oboe here. Did the composer figure that out anew or get the idea from some previous piece? Thus music can be the most stressful of all activities—from the perspective of the potential composer, because each intriguing new idea portends some unmastered aspect of ability, or newly recognized deficiency in one's own musical machinery. From this viewpoint, it can be very stressful to find oneself forced (from somewhere else within oneself) to like a piece of music without understanding the psycho-musical trick that makes it so effective.

OL: So, in that case music is not something that gives pleasure.

MM: Right. And even when it does give pleasure, there is no reason to take that pleasure at "face value." What pleases you can also control you, by causing you to do something other than you would have done otherwise. The way an adolescent can be enslaved by, and infatuated with, a pretty face borne by a person with no other evident merit. I think Paul Goodman once suggested that "Pleasure is not a goal; it is a feeling that accompanies an important seeming activity."

Music as a Tool for Emotional Exploitation

OL: Apparently, then, music can be used in part as an exploitative device?

MM: Yes. To be a popular composer, what is it you must do? Perhaps you must learn enough tricks that cause people to have pleasant sensations without too much stress. What are the tricks for making something that catches on in a listener's mind, and keeps repeating long after the performance? There is a superb novel, *The Demolished Man*, by Alfred Bester that depicts how the tunesmiths of the future develop jingle-composing to such a degree that it can be used, in effect, for mental assassination. A victim infested by such a tune is helpless to do any useful work, and must hire a specialized therapist to help remove the tune from memory.

…the enterprise of making an artificial intelligence involves not only some tools but also a larger-scale outlook.

OL: However, wouldn't such a use of music entail knowing the person one is imprisoning very well?

MM: That is a profound question: how universal are our musical techniques? And even if there are powerful universals, there must also be powerful non-universals and, so, perhaps in the future when we have better brain reading instruments, people may commission composers to produce works designed not for audiences, but for particular clients? Surely you can write a much better piece of music for a single listener ("better," of course, in that person's view). And then perhaps it will be considered most masterful to write a piece that only that single client will adore, and everyone else will abhor.

OL: Now, if such a work were possible, wouldn't we have almost no reason at all to understand music as something that has universality, and addresses itself to some characteristic people share?

MM: Good question. One could argue that even if we find compositional techniques that are widely universal, in the sense that they evoke strong and similar reactions in most listeners, those techniques would by that fact be in some sense rather shallow—because, in being so nearly universal, then it must simply be because of filling some niche, or exploiting some mechanism, or taking advantage of some bug that all human brains have. Consider how many people tend to tap the foot to the rhythm of a piece of music without knowing that they're doing it. An alien being might regard that as some sort of mechanical bug, even if humans regard it as natural.

The Present AI and Music Scene

OL: If we look at what people in AI and music do these days, we'll find that problems like those we have discussed are on nobody's

agenda. What one tries to do, instead, is to go a single step beyond verbalizing music's effect on people, or verbalizing how one composes music, by "formalizing" verbal information (mostly public, not personal information). For instance, in this book, you'll find a discussion of issues in knowledge representation, and of attempts to rationally reconstruct certain musical activities such as composing, performing, learning, and analyzing music. The presupposition that is made in such studies seems to be that we already know what music is, and that the goal of research is just a matter of spelling out that knowledge more clearly (whereby it is not always evident for whom). I think a very strong and, in my view, pernicious, idea in the field—if today it can be called a field at all—is that you can take AI as it is, and "apply" it to music, where "music" is a medium we already know well. As a consequence, one is harnessing the means of present-day AI to the task of "explicating" music, rather than to try to re-invent AI on the basis of musical knowledge other than just verbal knowledge.

MM: Yes, well, that's largely true—and perhaps it is also because AI isn't advanced enough yet, even to explain how we use language—for example, to understand stories. That surely would be an obstacle if it turned out that much of musical thinking involves those same mechanisms. This reminds me of how, in my childhood, there were musicians who talked about music theory. But when I asked what music theory was, it turned out that "music theory" was little more than nomenclature—the classification of chords and sequences. A syntax with no semantics whatever—with virtually no ideas how the music conveys its meanings, or whatever it is that makes us react. The situation is a good deal better, in AI-based language theories, because we have the beginnings of theories about the nature of "stories" and how they affect us. For example, Wendy Lehnert developed a nice theory about the most usual sorts of constraints on the structures of acceptable stories. A respectable story must introduce some characters—typically, a protagonist and an antagonist, and some sort of problem or anomaly that must eventually be resolved. Lehnert shows how such plots could be assembled, recursively, from elementary activities that she calls "plot units." These presumably are what help you construct, elaborate, comprehend, and (perhaps most important of all) remember, a story. For example, according to this model, if character A does something bad to character B then, eventually, A must compensate by doing something good for B—or else something bad must also befall A, so that B can "get even." Of course there can be exceptions to this—but not too many, or else the reader will be compelled to ask, "what is the point; this story rambles; it makes no sense; it is not a story at all." To "understand" a story one needs a well-controlled agenda of concerns about conflicts among its elements.

OL: So, you want to look at music as a kind of "pseudo-story?" (I did that myself in a 1986 paper entitled "Toward a Computational Theory of Musical Listening").

MM: Exactly, that's just what I was reaching for. At least in some kinds of "music," as in most kinds of stories, there mustn't be too many loose ends, unresolved problems, irrelevant material, and pointless excursions. However, acceptable compositions differ from acceptable stories in permitting far more repetition. In works of music, there is a lot of redundancy, as shown by the fact that, in much of classical music, measures are of the same length, and there is a binary tree structure: two phrases are repeated to make what music theorists call a period, and two periods repeat to make a musical sentence. In my 1980 paper I suggested the obvious thing: that certain parts of the brain might apply to this a hierarchical analysis that, in effect, parses this input into a binary (or, rarely, ternary) tree-like structure; then the meaning can be extracted from the now easily recognizable differences between the corresponding parts. Why else would you repeat something so many times except either to learn it by repetition, or to point out to the listener certain small differences between sequences. For instance, we notice in a typical 4-line tune that one of the phrases goes up the first time, near the end, and the second one goes down, that's a sort of pair of parentheses; and we get used to that.

OL: Yes.

MM: In a computation-based treatment of musical expression you would expect to see attempts to describe and explain such sorts of structure. Yet the most "respectable" present-day analyses—e.g., the well-known Lerdahl & Jackendoff work on generative grammars for tonal music—seem to me insufficiently concerned with such relationships. The so-called "generative" approach purports to describe all choices open to a speaker or composer—but it also tries to abstract away the actual procedure, the temporal evolution of the compositional process. Consequently, it cannot even begin to describe the choices composers must actually face—and we can understand that only by making models of the cognitive con-

straints that motivate an author or composer. I suspect that when we learn how to do that, many regularities that today are regarded as grammatical will be seen as results of how the composer's motivations interact with the knowledge-representation mechanisms shared by the composer and the listener. In any case it seems to me that, both in music and language, one must understand the semantics of tension-producing elements—at least in the forms that resemble narrative. Each initial discord, be it melodic, rhythmic, harmonic, or whatever, can be seen as a problem to be later resolved. A lot of what a composer does is setting up expectations, and then figuring out how to frustrate them. That gives the composer some problems to solve. The problems and their solutions are then like elements of a plot, and composition becomes a kind of story telling.

OL: If we could understand composing as a variety of story-telling, and performed music as a pseudo-story, wouldn't that help us to arrive at a theory of musical discourse?

The Heart of AI

MM: Yes, indeed. And if we hope to apply AI ideas to that, we'll want to exploit the best tools that AI can offer—different sorts of processing, different ways to represent knowledge, and so forth. But the enterprise of making an artificial intelligence involves not only some tools but also a larger-scale outlook. And AI will fail to illuminate music, just as linguistics did (because of trying to isolate meaning apart from syntax) if we use only the parts but reject the heart. For just as linguistics has a heart—to find how we communicate, so AI too does have a heart—to find out how machines can be made to solve significant problems. Now, in past years we have seen a certain amount of applying AI tools to traditional music analysis. But I would like to see more from the heart of AI, the study of problem solving, applied to issues of how you solve musical problems.

Making a Composer

OL: So, then, for you to apply AI to music, if one can say apply…

MM: … would be making composers, or at least listeners…

OL: By "making," do you mean to produce a robot-like creature that does certain things like, observably composing?

MM: Yes, indeed. And in the case of listening it would have to know when to say "oh, this is exciting," or "how very tender," and the like. I haven't seen much of that.

OL: In Japan, one has built a robot that is capable of reading music and then playing it on the piano.

MM: Yes, that fellow at Mazda.

OL: Is that something you have in mind here?

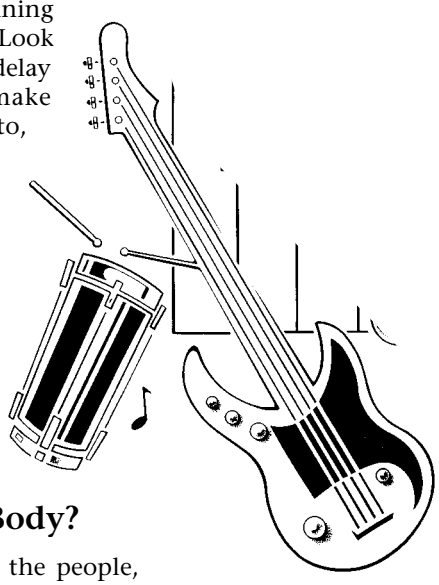
MM: Not at all. Because I'm more concerned about what happens at larger scales of phrase and plot. Our listening machine would have to understand the music well enough to recognize from each moment to the next which problems have been solved, and which remain open.

OL: How would such understanding have to become manifest?

MM: Well, for example, an understanding listener can hear a piano concerto and appropriately say things like "that was a good idea here in the cadenza, but he didn't carry it through." I'd want the robot to make similar analyses.

OL: To do that, the robot must be able to recognize solutions, good or bad. How would it communicate such solutions to others?

MM: One way might be to have it write the sorts of sentences that critics write. Or to have them work more in the musical realm by performing as a teacher does, explaining differences by demonstration—"Look how much better it would be to delay a little these notes here, and make those near the end more staccato, like this, and this." And of course if our machine turned out to able to produce interesting enough interpretations, then we might be satisfied by that alone—if many listeners were to agree that "really, that performer has a lot of good ideas about this music, and brings out stuff that I didn't realize was there."



Does Music Need a Body?

OL: For me, this brings to mind the people, especially at MIT, who have begun to build robotic insects. This "nouvelle AI" seems to be an approach that is in some contrast to the top-down symbolic approach AI has used in the past. I don't know whether you see the matter quite my way. Regarding a musical robot, for instance, I would want to know how it is programmed. Is it programmed on the basis of symbolic representations (i.e., high level constructs), or rather on the basis of a multitude of simple circuits, or of neural networks—or do all of these have a role to play?

MM: Well, surely understanding music

requires many levels, and a good deal of cultural knowledge.

OL: Don't you also need a body level to build an artificial composer or a listener, to render actions?

MM: I don't think that this will be important.

OL: So you don't consider this to be a critical problem—that emotion, as distinguished from cognition, in all known cases manifests itself through an organic body?

MM: Not really, because I don't expect that emotional behavior will turn out to be singularly hard to simulate. I'm inclined to agree with Niko Tinbergen that the basic emotions come from the activities of various almost separate processes, brain-systems that each have different, and rather clear-cut goals. For example, in the brain stem at least a dozen such systems have evolved and, to a large extent, they behave like different animals. In *The Society of Mind* I called them "Protospecialists." When you're low on sugar, the protospecialist for "hunger" gets activated—and causes you to apply your available knowledge to find a way to get food. If you're too cold, then another specialist gets activated, suppressing the others; it uses the available knowledge to get out of the chill and into the sun, or to cover your body with insulators, or to huddle together with your friends. I think most people assume that emotions are very deep and complicated, because they seem so powerful and hard to understand. But my view is just the opposite: it may be largely because they are basically simple—but wired up to be powerful—that they are hard to understand. That is, they seem mysterious simply because they're separate and opaque to your other cognitive processes. However, there is also another thing that makes these emotions hard to understand, at least in their older, more adult forms. It is that although an emotion may be simple by itself—for example, a brain center that is genetically wired to learn actions that keep one warm—the end result of what it may learn can become arbitrarily complicated. The problem is that each infantile emotion, such as hunger, defense, nutrition, or sex, can eventually acquire a huge cognitive system for achieving its own goal, perhaps by exploiting the other emotions. And presumably virtually every adult cognitive activity develops in some complex way from infantile activities. So I see as false the commonsense distinction between "intellect" and "emotion." It would be better to emphasize the development over time from simple to complex. Nevertheless, music appears to depend for some of its effects on being able to arouse emotions; therefore, in order to build a competent composing or listening machine, we might need to understand such activities well enough to simulate them. Some of this could surely be done with the sorts of decentralized agents and agencies being used in those insect-like AIs—but I don't suppose they would work well for this unless we managed to program them with roughly the required emotions and cognitions. I have never heard of any non-artificial insect that could learn to prefer Schubert to Mendelsohn.

OL: Right. These artificial insects would have to climb the tree of knowledge first.

The Need for an Emotional Culture

MM: Yes, indeed. And to use them inside our music machine, they might have to evolve to be quite similar to the ones inside our human brains. Our reactions to the sound of a violin may exploit our reactions to the pitch centers of female voices; the sound of a cello may seem more male. A machine that was really competent to listen to nineteenth century classical chamber music might well need some knowledge-understanding of human social affairs—about aggression and conciliation, sorrow and joy, and family, friendship and strangership. And of course there are other constraints on how we use those instruments. Generally the cello will have to play longer and slower notes that emphasize foundational notes of the chord and there's nothing sexual about that—but nevertheless, our human perceptual experience will tend to associate the higher notes with children's voices, the middle octave with female, and the lower one with male voices. It might even be natural to link yet lower tones with those of large fierce animals.

OL: There is sufficient historical evidence that suggests what you are saying is entirely to the point. The violin was developed in Renaissance Italy, by people who were very eager to simulate the human voice. For them, the human voice was the telos which they strove to re-embody in the instruments of the violin "family."

MM: I have to admit that the first time I encountered a modern synthesizer, the most exciting stop for me was the choral stop, the high female voices. Having been a pianist all my life, it was so astounding to be able to touch a key, and have those soprano voices come out, and basses, tenors, contraltos. Completely entrancing. For an hour or two, like a fairyland. But eventually it seemed

…until our machines have access to commonsense knowledge, they'll never… understand stories…

wrong that I couldn't make these voices talk, they were always making the same bleating sound. But what a marvelous experience….

Requirements of Being a Music Scientist

OL: We were wondering why none of these ideas about cognition and emotion seem to be of relevance in present-day AI and music research…

MM: I think that's because the AI people have suffered from the same misconception that most cognitive psychologists have suffered from, viz., the idea that "well, we'll do the easy things first, like understanding memory and simple reasoning and so forth; but emotions are surely too difficult, so let's put off researching them for now." I once came across a statement by Freud in which he complains along lines like this: "people think I work on emotions because those are the profound, important things. Not so, What I'd really like to understand is common sense thinking. And it is only because that's so difficult, so incredibly complicated, that I work on emotions instead—because emotions are so much more simple."

This is why I like Douglas Lenat's project to build a commonsense data base. I'm sure that Freud would have liked it too. This is a big project; the CYC data base will involve millions of different items, fragments of our cognitive machinery. But until our machines have access to commonsense knowledge, they'll never be able to understand stories—or story-like music. One reason that AI has not gone very far in such domains is because researchers have been afraid to say: "I think emotions are simple enough that we can make useful models of them." I'm not saying that emotions are trivial; they surely involve some complicated machinery. But you can't make progress unless you're willing to begin with simple theories, to serve as first approximations to the science we'll eventually need to understand musical activities.

OL: Could you explain that a little further?

MM: Sure. I mean that one should not be daunted by the apparent complexity of emotions, because it may be feasible to get a good start by making what may seem to be oversimplifications. You could begin by saying "maybe there are only three basic emotions"—happy, sad, and whatever you like. The "sentics" model of Manfred Clynes stipulates seven. It doesn't matter, so long as you start somewhere. I want to avoid the disaster that has befallen syntactic linguistics which, for all its apparent accomplishments, is dead in the water because of failing to include even the simplest caricature of a theory of what words mean. In my view, syntax has led to some useful discoveries, but to fewer than would have come from building more comprehensive models of language production and understanding. So I would like to see some music-theorists start with models based on simulating a few postulated emotions, each linked to a few procedural rules about how various sorts of rhythmic, harmonic, and melodic elements might work to arouse, suppress and otherwise engage a listener's various feelings and attitudes. Unless you start somewhere you'll go nowhere.

OL: Are you saying: the best way to approach music is to go at it from some hypothesized basic emotions, rather than from ideas about problem solving, or the idea that music is a sort of intellectual game like chess, or is something explainable by psychoacoustics?

Music as a Pseudo-Story

MM: Well, emotion might not be the right word, but we must somehow formulate and engage what we believe are the important musical problems. And clearly, an important aspect of "understanding" music experience is the listener's experience of apprehensions, gratifications, suspense, tensions, anxieties, and reliefs—feelings very suggestive of pains and joys, insecurities and reassurances, dreads and reverences, and so forth. So, just as a good story confronts us with conflicts and resolutions, so do good musics, which take over the listener's disposition with feelings

like: "I'm worried that something bad will happen unless this conflict is resolved." Problems and solutions. Once I was trying to finish a two-part invention but there was something wrong with it. I asked a musical friend for help and was surprised to hear a simple answer. "Look here, you went up into this octave but then you never made anything of it. You can't do that. When you let a voice get into a new range, like breaking new ground, you must have a good reason for it, and you must also think of a way to get that voice out of there afterwards. But here you just left those notes hanging up there, and I ended up wondering what will happen to them." Now I wonder if this doesn't reflect some sense of territoriality instinct, in which the listener is disturbed by not knowing who controls which area....

OL: Well, it also relates to rules of musical discourse, doesn't it? If you put forth an argument, you have to follow it up ...

MM: Right—and that reminds me again of the constraint-like operations of Lehnert's Plot Units. Those dangling high notes evidently require follow-up, some resolution, and unless the composer brings them down, the listener is left with some unresolved conflict, concern, or problem. As though one of your emotional proto-specialists were left in an active state. After all, what are musical problems if not problems having to do with the resolution of conflict?

OL: Any book on counterpoint will instruct you that "having made a particular step, there are only so many things you can do," as well as "not following up this step is a mistake in terms of the rule system we are adhering to." Whereas you were saying, following up some compositional decision is required for being emotionally or story-wise consistent in one's music.

Where All Those Rules Go ...

MM: Yes, and that raises questions about the numbers of and characters of those "rules"—and those two dimensions are far from independent. Indeed, this is what I was thinking about in that analogy with syntax. For if you tried to analyze contrapuntal music, you could write rule after rule to constrain its surface structure, but you'd never get quite enough rules to tell composers what to do. However, I suspect that if we managed to build the right sort of model for musical plots and conflicts, we might end up with a more compact theory that solves many more problems. A theory that takes account of both transmission and reception, of composing and listening.

Chapters 22 and 26 of my book *The Society of Mind* propose such a theory of how language works. The idea is that when a speaker explains something to a listener, the goal is to produce in the listener a structure that resembles a certain semantic network in the speaker's brain. Many linguistic tricks are used for controlling the growth of that network. My conjecture is that mechanisms of this sort could lead to good descriptions of what utterances are understandable—and could do this, I suspect, with simpler and fewer rules than the usual kinds for describing sentential grammar—because many grammatical rules that seem separate and independent might result from these deeper mechanisms. It is OK to begin by looking at surface regularities.

Musicological Taboos

OL: It's certainly a step forward to have begun to describe musical products in terms of the rules that may underlie their generation (although in most cases, such as Lehrdahl and Jackendoff, such rules are rarely if ever tested empirically, and therefore have little cognitive reality). Rules codify actions, or elements making up an action. For the longest time, a taboo has existed in music theory and musicology regarding any attempt of understanding music as an activity. Researchers have typically concentrated on music as an object, an artifact, a product, rather than studying the process by which music is generated, both in the composer and the listener. How do you explain this taboo? Are you aware of analogous taboos in other sciences?

MM: I think the taboo made sense in the past simply because before the advent of modern theories of complex information processing, there really was no useful way to think about how minds might do such things. In my view, the theories of philosophers before the era of Freud, Piaget, and Tinbergen were too primitive to provide much insight. Freud recognized that higher forms of human thought involved the pursuit of goals by acquiring and exploiting knowledge. But it was not until the 1960s that workers like Allen Newell and Herbert Simon formulated adequate theories of systems with goals. To be sure, there were earlier attempts to base psy-

chology on simple principles of association—as in the models of Pavlov and Skinner—but it was never clear how these could lead to higher levels of cognition. Nevertheless, crude surface-behavioral descriptions became dominant, at least in American Psychology—and in my view this included both Skinner and Chomsky, despite their famous debate; neither of them seemed comfortable with the idea of making models of the internal processes that underlie behavior. In any case, outside of psychoanalytic circles, making complex theories about mental processes did become taboo.

OL: This taboo might be called "paradigmatic," then; it is built into the paradigm of what good research was supposed to be.

MM: I think so, considering the lack of progress on most other mental activities. But beginning with the 1950s, AI and computer science started to produce an enormous quantity of new concepts. Just look at the lingo—a "stack," a "push down list," "default inheritance." In these fields, there emerged literally a thousand new ideas and terms for describing mental processes, whereas before, maybe there were a few dozen, in language. So, humanity was just not prepared to understand anything as complicated as the process of a thinking machine. Until it had some practice with it, due to the computer.

OL: Do you have the impression, as I do, that the taboo we've talked about is still in force, even in AI and music today?

MM: Yes, especially in music. There, the taboo is often justified by arguing that too much inquiry will spoil it. There's the apprehension that if you understood music, you would lose your interest and destroy the beauty of it.

OL: There is also this Western notion of a "work of art," as something that is just too good to be drawn into the question of how it was made, and the further notion that it would be best to forget that it was made at all (something that my teacher Th. W. Adorno was the first to notice in his sociology of music) …

MM: Right—and we know what Freud would say about any such notion, viz., that, if there is such a fear, it's because you're worried that there's not so much there, and you're repressing your worries. The more angry you get when it's questioned, the more you are giving yourself away.

I once had an extended argument with a chemist friend because I said: "why don't we synthesize good wine? Why be at the mercy of the weather and the vintages?" And he replied: "That simply can't be done; it is impossible; the sense of taste is infinitely complicated." And I replied: "Well, why not analyze some wine; we might find that there are only a few important chemicals in each good wine, and then we just figure out the amounts of them." And he said: "No, there must be a million chemicals." I said: "And you can taste them all, and they all matter?" "Yes," he said, "they all stick together to form one indescribable whole." The argument seemed interminable and I wondered, did this outstanding chemist unconsciously fear that chemistry itself was inadequate? Did he fear the thought that he himself might be "merely" chemical? Or did he fear that all his likes and dislikes could be explained—and then there'd be no mystery to life? Is there a danger if we understand music, or art, or liking food, there will be nothing important left to do? I'm not worried, myself, about running out of mysteries.

How Do Works of Music Relate Their Process?

OL: There is another question worth investigating, viz., the relationship between a work and the process that produces it. For me at least, that question is the crux of any inquiry deserving the name "musicological." It is not enough to make explicit the structure of the work (as is done in music analysis and traditional musicology), nor is it sufficient to make explicit the structure of the process that produced the work (as is done in AI and Cognitive Science). The crucial musicological question regards the link between product and process, the issue of whether one can describe the link, and in what way. And this problem, of relating product and process, for me has always been one of the crucial issues in all of the human sciences, not only in musicology.

MM: Very interesting idea. We're used to seeing completed works—but it would be interesting to see the steps in between. Suppose you were to watch an artist making a painting. Of course, that doesn't show you the artist's thoughts, but it shows some of the planning and structure. How would that affect your relation to the painting? Similarly when you hear music, the steps are concealed—but some composers do work at the keyboard, and it would be interesting to have recordings of those activities. Beethoven made many revisions in some of his manuscripts of Opus 111, some bits written over many times, and pages of alternative sketches; but only the final results are performed. We should ask the next Beethoven to provide us with recordings of his improvisato-

The crucial musicological question regards the link between product and process…

ry sketches. But perhaps that won't be necessary if, over another few decades, we find ways to more directly record a composer's actual brain activities.

This touches on the relation between composing and improvising. In composing you can be out of real time, and make revisions, and cover your tracks. Improvisors cannot be quite so devious—and therefore, in some cases, they may communicate more successfully. And this reminds me of a different experience when I was a student; reading books on mathematics always seemed peculiarly difficult and always took a long time. But one day, I ran across a book by John von Neumann on *Mathematical Foundations of Quantum Theory* and it was the clearest, most pleasant mathematics book I'd ever read. I remember understanding his explanation of the Metric Density theorem as like a real time experience, like being inside another person's head. (This is a theorem about probability or, more precisely, about measurable sets. It says, for example, that if a certain subset U of a square S has probability larger than zero—that is, if when you throw a dart at the square there is a non-zero probability of hitting a point in U—then there must exist smaller squares in S in which the U-points have probabilities arbitrarily close to 1.) I mentioned this to my friend Stanislaw Ulam, who had worked with von Neumann, and Ulam thought he recalled that von Neumann had been in a hurry when he wrote the book, it was a first draft and he never revised it. And so, he was writing down how he thought about the problems.

OL: He improvised the book ...

MM: Yes, and I found that once I had read it, I never had to go back. Whereas most other math books have been made terribly tight; all extra or "unnecessary" words removed, and all traces of how the author actually thinks about the problem. My point is that it can take much longer to read a shorter book with the same content. So too, in music, it sometimes might be nice to have the composers' first drafts—that is, from composers we like, and who improvise at the keyboard ...

Knowledge Acquisition in Music

OL: Of course, making music with computers gives you the possibility of doing knowledge acquisition. In the OBSERVER and the PRECOMP projects (see Laske, this volume) I have presented children as well as adults with a compositional problem to be solved by using a particular program for computer-aided composition. During these experiments, I have generated what H. Simon would call a "protocol," i.e., a document capturing the actions the problem solver had used to find a solution to the given problem.

MM: So, these composers have to shape a piece by applying some operators to the material given them ...

OL: Yes, one might call that task-based knowledge acquisition, since knowledge is here being elicited in the framework of a specific task for which an appropriate environment has been built using a computer. The resulting protocol is not a verbalization of expert knowledge, as in today's knowledge acquisition for expert systems; rather, it is a document that simply captures the usage of problem solving operators over the time of the experiment. It is then the task of the knowledge elicitor/analyst ("musicologist") to understand that operator sequence in terms of its control structure, and to relate his procedural insights to what he/she declaratively knows about the structure of the product ("work") composed during the experimental session....

MM: Then you could also ask the composers why they did what they did, I suppose...

OL: You could, perhaps not children, but professional composers, or people who are learning to be composers ...

MM: They might not be able to tell you anything useful, though. One just doesn't know those processes. It's just like asking somebody why they used a certain clause in their sentence, ... what could they tell you?

OL: Not much indeed. I don't in any way presume such processes would be conscious knowledge... Exactly for this reason, I decided to avoid asking musicians to verbalize what they do (except in a study on music analysis called KEITH published in 1984). I wanted to catch them in the act of music-making.

MM: But you could look over the (action) protocol, and maybe find things that you never otherwise find ...

OL: Yes, that's my point. For the first time in music history, we are able to produce empirical traces of a musical process; we can then study such a process in terms of the actions it is composed of. Common sense tells me that a musical form derives from the process that produced it, and I would think, therefore, that the control structure of that process is intimately linked to the musical form emerging from it. (See also Marsella/Schmidt in this volume.)

Music as a Bug between Brain Areas

MM: Yes—and perhaps in some years from now we'll be able to see those processes even more directly by using high-resolution brain activity imaging instruments. I would certainly like to be able to see what's happening in my brain when I improvise, because I have virtually no direct insight into that; I'm thinking about other things when I do it—sometimes even about some completely different piece of music. I suspect that my music production machinery is mostly in my motor brain regions—that is, in the front half of the brain—more than in the back or sensory half of the brain. (I really hate to hear so many people repeating superstitions about the left and the right sides of the brain; the brain has hundreds of different parts, so you can divide it in two in many ways.) The result of this is really annoying to me; I have great difficulty writing down music on paper until my hands play it on the keyboard so that I can hear it. Now, I don't actually mean hearing, literally, because I can play it on an imaginary keyboard in the air. But I can't hear the music nearly so well without moving the fingers. This probably means that in order to close the loop, I have to use some bundles of nerve fibers that tunnel under the central sulcus from the motor regions back to the sensory regions of the cortex. It would be nice to understand this—and there's no reason why not, with new instruments.

OL: If you are right with your hunch that to understand how humans make music we need to know how they use their brains, then, of course, we are right at the question of how music developed in the history of the human brain.

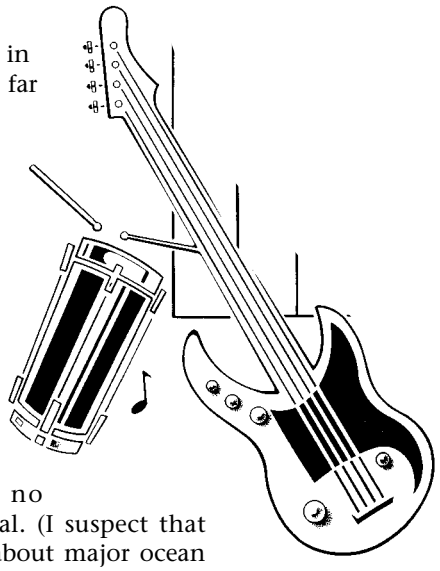
MM: Yes, and that should be extremely interesting because music seems to have no evolutionary origin. So I suspect that in each musical person there has been some early incident in which some musical process-knowledge comes to occupy some piece of brain that isn't dedicated to something else, and it probably happens somewhat differently in each person. I don't think anybody has mapped this very much.

OL: You just said something for me quite astonishing, viz., "Music seems to have no evolutionary origin."

MM: Well yes, in the sense that one can't see anything much related to it in animals descended from our ancestors.

OL: There are no traces …

MM: … of musical concerns or abilities in our ancestors or relatives—so far as I know, no sign of musical interest in any other primates. Nor, so far as I know, in any other animals. I know of no animals that even tap their feet to rhythms or anything like that. I don't consider bird-songs to be songs; yes, they have communicative functions but no reason to relate that to music. Or whale songs. We call them songs but we still have no idea about their functions— and again, I see no reason to call them musical. (I suspect that they contain information about major ocean currents or coastline features or lists of where various individuals have been seen, or other functional things like that. But no one knows.) Of course we all meet romantic people who maintain that their plants enjoy and thrive in musical environments. But the only careful experiment I've seen only demonstrates that vibrations tend to retard plant growth by injuring rootlets. My best guess is that music became possible because of some anatomical innovation that just happens to facilitate interactions between other, older functions—for example between some of the brain that does planning for paths in space and some of the parts involved with language, or story-like memory systems. If that were the case, it might explain why hearing certain kinds of sounds might come to give you the feeling that you understood something, or give you the experience of being in some other place. If so, we could even regard music as being related to some sort of "hardware bug" perhaps involving brain areas concerned with visualization, kinetic imagery, language, or whatever—some combination of structures that could lead to almost autonomous kinds of activity that we label "musical."



OL: Looking back at ancient Greek culture, you'll find that the term *mousike* denotes something very specific, viz., the linkage between poetry and sound. This culture really has no term for instrumental music as we know it. Sound-making that was not linked to the human voice was considered as being a kind of "techne," and thus much more lowly than mousike. It seems, then, that the split between poetry and music, and the coming into being of music as we know it today (i.e., instrumental and electroacoustic musics) is a rather recent event, even in terms of the short history of humanity.

…perhaps most important of all, we need to know more about common sense reasoning…

MM: Right.

OL: Of course, one would have to look at other cultures.

MM: Surely. And in any case, however the brain is involved, music clearly interacts with many memory mechanisms. Not only in music but in other realms, it seems much easier to remember things that are grouped in regular temporal structures that resemble rhythm and meter. The other day, I complained to Carl Sagan that we have no good theory of why people are so attracted to regular rhythms. He pointed out that the mother's heartbeat is a prominent context of every baby's development. I'm sure that there is something to that—except that dogs and cats hear heartbeats too, but do not seem to tap their feet or otherwise show much sign of being affected by music.

A Musical Common Sense Data Base

OL: Indeed, music is a human privilege, a reflection of the human condition, and it probably holds the key for much that psychology is attempting to unearth about human nature. Alas, our time is almost up. Therefore, to conclude our conversation, could I ask you what you think has been achieved in the field of AI and music so far, and beyond that, what you think should be worked on, or might be worked on, in the next decade or so? Is that a fair question to put to you?

MM: A vital question. And to guess at an answer, we need to understand our position in history. AI is only 30 or 40 years old, but people are always asking: "What has it done lately?" or "What are its important achievements?" The trouble is, we can't really say, because what seems important at one moment may not be what turns out to have been important 10 or 20 years later. We can't be certain which are the good ideas yet. Perhaps we need more research on case-based reasoning. Or on structures like Lehnert's plot units, to understand more about the structures of stories—and of musical compositions. (In Roger Schank's recent theories of learning and reasoning, story-structures play critical roles.) Or, perhaps, most important of all, we need to know more about common sense reasoning, and the data structures that underlie it. Because no computer in the world today yet knows the meaning of enough words to understand a story.

OL: And this will remain true for some time to come …

MM: That is the gap that Doug Lenat is trying to fill, viz., by working on a common sense data base. And similarly we need some sort of musical common sense data base.

OL: So it would seem.

MM: In college, I attended one of Walter Piston's courses. He used to complain that no one knew enough about what makes good tunes. He said that there were only rules of thumb, like the rules for good manners or good behavior; a good melody must show some sense of direction and return; it shouldn't be too jerky—and it shouldn't be too smooth—but there were no useful formal theories about such things. And I remember thinking then—before I ever thought about AI—maybe there actually can't be such rules. Because perhaps it works a different way, by analogy. Perhaps I match each tune I hear with a hundred tunes that I learned when I was a baby. A machine to do this might have to know all those nursery rhymes and folk songs and lullabies and Christmas carols. Perhaps the reason why we like certain tunes is largely because of already liking other similar ones?

OL: Most likely.

MM: And if that were so, then the important thing about a tune would be how well it resembles some of those repertoried tunes—and what are their interesting differences. In any case, those music courses didn't seem to know how to tell me how to compose. They were infuriating …

OL: They still are …

MM: Now, today, things ought to be better. We have so many new good ideas. Case-based ideas, grammatical ideas, problem-solving

ideas. Expert-system ideas. So much has been accomplished that we may be ready for the renaissance of many an old inquiry. Perhaps the research community as a whole already has a critical mass of the needed insights—and is only waiting for someone to see a way to pull them together. Anyway, in my view the most critical thing, in both music research and general AI research, is to learn how to build a common music data base. But nobody in music research works on that yet, do they?

OL: Not that I know of … There exist, to be sure, small data bases of scores of the so-called Western tradition, but that is not what you have in mind here …

MM: It's mostly the same in AI. Most hope to understand language by using compact theory-tricks, like formal grammars. But that simply can't do enough by itself. You need some sort of data base, of experience. A few little stories about each word. And the same for music. Surely you can't react "properly" unless you possess some "stories" about each chord sequence, or melodic turn, or rhythmic variation.

OL: So, that's what you call a common sense musical data base: a collection of stories about important compositional or auditory constructs?

MM: Precisely. And consequently, despite all those popular, fancy, formal theories, we're missing all the substance beneath. We need a musical Lenat to start a musical CYC project.

Bibliography

Alker, H. R.; Lehnert, W. G.; and Schneider, D. K. 1985. Two Reinterpretations of Toynbee's Jesus: Explorations in Computational Hermeneutics. In *Intelligence and Text-Understanding: Plot Units and Summarization of Procedures,* ed. G. Tonfoni, 51–94. Parma, Italy: Edizioni Zara.

Laske, O. 1986. Toward a Computational Theory of Music Listening. In *Reason, Emotion, and Music,* eds. L. Apostel et al., 363–392. Ghent, Belgium: *Communication and Cognition.*

Laske, O. 1984. KEITH: A Rule System for Making Music-Analytical Discoveries. In *Musical Grammars and Computer Analysis,* eds. M. Baroni et al., 165–199. Florence, Italy: Leo S. Olschki.

Laske, O., and Drummond, S. 1980. Toward an Explicit Theory of Musical Listening. *Computer Music Journal* 4(2): 73–83.

Lehnert, W. 1982. Plot Units: A Narrative Summarization Strategy. In *Strategies for Natural Language Processing,* eds. W. Lehnert and M. Ringle. Hillsdale, N.J.: Lawrence Erlbaum.

Lehnert, W., and Loiselle, C. 1985. Plot Unit Recognition for Narratives. In *Intelligence and Text-Understanding: Plot Units and Summarization of Procedures,* ed. G. Tonfoni, 9–47. Parma, Italy: Edizioni Zara.

Minsky, M. 1989. Music, Mind, and Meaning. In *The Music Machine,* 639–656. Cambridge, Mass.: The MIT Press.

Minsky, M. 1986. *The Society of Mind*. New York: Simon and Schuster.

Minsky, M. 1980. K-Lines: A Theory of Memory. *Cognitive Science* 4:117–133.

Tonfoni, G., ed. 1985. *Intelligence and Text-Understanding: Plot Units and Summarization of Procedures*. Quaderni di Ricerca Linguistica 6. Parma, Italy: Edizioni Zara.

Marvin Minsky has contributed to the domains of symbolic description, knowledge representation, symbolic applied mathematics, computational semantics, and machine perception. His main concern over the past decade has been to discover how to make machines do commonsense reasoning. The foundations of his conception of human psychology are described in his book The Society of Mind. Minsky is a 1990 recipient of the Japan Prize that recognizes original and outstanding achievements in science and technology.

Otto Laske is a composer, and the President of NEWCOMP, Inc., The New England Computer Arts Association, Boston, Massachusetts.

…the most critical thing, in both music reserch and general AI research, is to learn how to build a common music data base.