Cognitive complexity
and the structure of musical patterns

Jeffrey Pressing
Department of Psychology
University of Melbourne
Parkville, Victoria
Australia 3052
j.pressing@psych.unimelb.edu.au

Introduction

Patterns are fundamental to all cognition. Whether inferred from sensory input or constructed to guide motor actions, they betray order: some are simple, others are more complex. Apparently, we can erect a simplicity-complexity dimension along which such patterns may be placed or ranked. Furthermore, we may be able to place the *sources* of patterns—such things as mental models, neural modules, growth processes, and learning techniques—along a parallel simplicity-complexity dimension.

Are there psychological forces pushing us along this dimension? Seemingly so. The drive to simplify and regularize is seen in familiar *Gestalt* ideas of pattern goodness and *Prägnanz*, which hold in many situations. The drive towards complexity can be seen in the accretional processes of individual development, or the incremental sophistication of many cultural behaviours such as systems of musical design (e.g. musical harmonies in mid to late 19th century Europe), or new scientific methodologies (e.g. increasing sophistication in brain imaging techniques).

Quite beyond this, the handling of complexity is a central issue in human skill and learning, providing the drive to efficiency in resource-limited processes like memory, attention, or multi-tasking. It has parallel implications in systems engineering, organization management or adaptive neural network design problems. In these cases, complexity is something to be circumscribed by various information-management strategies while functional performance specifications are maintained. Yet another picture of complexity is as an emergent property, one embodying sophisticated implicit order springing up as a cumulative consequence of simple, typically iterative nonlinear processes operating in certain ranges of their control parameters. This is an approach derived from dynamical systems theory, and it has been used to model phenomena as diverse as social interactions, development, perception, motor skills, and brain organization, to say nothing of areas outside of psychology.

I aim here to examine the utility and reconcilability of these different aspects of the concept of complexity. This is to be done by examining some central questions:

What is complexity and how is it achieved?
How can complexity be measured?
What are the consequences of complexity for human behaviour?

In addressing these questions, it will be useful to primarily draw examples a single well-structured domain, music. However, the validity of the concerns is not limited to music.

What is complexity and how is it achieved?

Complexity is multi-faceted. One perspective on it might be called *hierarchical complexity*, which refers to the existence of structure on several or many scales or levels. These may be scales of space or time, or levels within a domain-specific functional space. The different structural levels may be apparently autonomous, or causally related (the first case providing greater complexity). Examples include intricate sets of interrelated rules (e.g., medical diagnosis), or conflicting and partially overlapping sets of case-based knowledge (e.g., legal precedents relevant to a particular case), hierarchically ordered sensory stimuli (e.g., a visual array corresponding to looking at a tree) or hierarchically ordered performance plans (e.g., a pianist's memory for a Chopin *Polonaise*).

Structural complexity is an important aspect of artistic design for three main reasons: a) hierarchical order increases the potential impact of the artistic statement by increasing the range of elements that can be coherently included within it; b) different hierarchical levels can vary in their complexity, so that access via at least one level is possible regardless of the sophistication level of the receiver (listener or performer); and c), multiple levels of order increase the potential for repeated listening, since attention can wander from level to level, defining a unique trajectory for each experience with the work.

Composers have exploited the potential for ambiguity or multiplicity of reference that hierarchical complexity makes available. In J. S. Bach's *Die Kunst der Fuge*, notes function as elements of linear processes operating at different time scales, and are at the same time compatible with a vertically-oriented chordal progression process. In many types of West African drum ensemble music, the organization of time line patterns used to coordinate the ensemble has been shown to follow an anti-Gestalt basis, intentionally promoting polymetric multiplicity of interpretation (Pressing, 1983).

Another slant on complexity focuses on time behaviour and change. Systems that show a rich range of behaviours over time, or adapt to unpredictable conditions, or monitor their own results in relation to a reference source, or can anticipate changes in self or environment, we may take to be complex. Examples are adaptive filter systems, human artistic creation, improvisatory performance, or fluid intelligence in general. This might be called *dynamic* or *adaptive complexity*.

A third slant on complexity is provided by algorithmic information theory, which derives from an idea originally suggested by Gregory Chaitin and Andrei Kolmogorov. This *generative complexity* is the length of the shortest program which can generate the object in question, when the program is written in a *universal programming language* (C, LISP or the coding of the universal Turing machine are three examples). It is known that the only (trivial) limitation is that coding lengths may differ between any two universal languages by a characteristic constant. This perspective has seen recent development as an area of computer science known as *information-based complexity*, to which we now turn.

Information-based complexity

In this approach the starting point is a target problem that requires a solution. Since solutions are seldom exact, we are actually seeking an approximate solution to a problem within a certain tolerance e. According to the theory of information-based complexity, a problem is framed by specifying three things (Traub, Wasilowski, & Wozniakowski, 1988):

- a *problem formulation*, which identifies relevant variables, what is to be approximated, and what error criterion is to be used;
- *relevant information*, which is not simply a list of facts but also a description of what operations are available to gain information from problem elements; and
- a *model of computation*, which says what operations are allowed and how much they cost (in terms of time to process or number of operations or even monetary units).

Given this, *complexity* is defined as the minimal cost of computing an approximate solution to the problem that has an associated error less than or equal to e. So instead of basing complexity on number of structural levels or capacity for adaptation, here complexity is based on a minimum set of resources (cost).

Notice that to complete this definition of complexity we have to specify how cost and error are defined. Cost in musical terms we can interpret in relation to cognitive and motor loading; error tolerance will depend on contextual standards of evaluation. In computer science, several significant cases of error and cost *settings* are distinguished: worst case, average case, or probabilistic. These refer to the operation of our system over a class of cases, and whether we respectively seek to minimize the worst case error, the average case error, or the worst case error when we ignore certain classes of rare events (Traub, Wasilowski, & Wozniakowski, 1988). Cost may be further based on relative or absolute error. Further distinctions are made on the basis of whether the process utilizes random information in decision-making, and whether it is necessary to accommodate noisy (possibly erroneous) information. Results are typically expressed in a function relating complexity to error tolerance e and other parameters.

One particular interpretation of this approach is MDL, the Minimum Description Length Principle, which can be seen as a principled version of Occam's razor (Zemel, 1995). This principle states that (Rissanen, 1989, cited in Zemel, 1995):

The best model to explain a set of data is the one minimizing the summed length (in bits) of:
1. The description of the model, and
2. The description of the data, when encoded with respect to the model.

The conception of complexity as computational cost also has the advantage that it can handle the effects of previous learning, which may predispose memory-guided (possibly biased) parsing of input and construction of output. Formally the change only requires that the theory-defining costs be framed in *conditional* terms: what is the shortest program *given* a certain knowledge base and *given* a certain production basis, with the cost of using elements from these bases set perhaps at much lower levels due to developed automaticity of processing.

How relevant is this to musical complexity? Although the dominant essence of music is not taken to be problem-solving in any culture that I am aware of, the idea that composing or performing music entails problem-solving activity that is subject to aptitude, memory, training and attentional constraints is widely acknowledged. Thus, in musical terms we face such archetypical problems as

- exposure to a pattern via audition or notational display, with the task of perceptually coding it, analyzing it, literally reproducing it, or producing a variant of it; or,
- creation of a new pattern that must be effective in a nominated context,

where, in each case, we are allowed a certain tolerance of timing, pitch, timbre, etc., in our solution. Solution details will be conditioned by previous learning, and the state of previous learning is assessable.

How can complexity be measured?

If the concept of complexity is to have computational power, we must be able to measure it. Complexity evaluation boils down to some central concerns, regardless of the type of complexity involved. And there is of course no need for the three types of complexity to be mutually exclusive.

With hierarchical complexity, we can begin by counting hierarchical levels. Then we are left with the problem of evaluating complexity of the individual levels, which may somehow be weighted to give a global complexity. The evaluation problem at each individual level is then nearly the same as when our complexity is not differentiated into levels—as in some adaptive or generative cases.

To quantify the complexity on each scale / level requires some computational notion of order or the sources of order. Information theory, one well-developed attempt, bases information on entropy, giving the highest information content to random noise, since any redundancy reduces information content (Shannon, 1948). This is known to be implausible with regard to human affairs, since psychological engagement requires some amount of order—white noise in any sensory medium is boring. Furthermore, information theory defines information in relation to the probabilities of all other inputs that might have been encountered, and this is in general hard to specify, especially in ecologically realistic circumstances. A similar dilemma of context dogs the spin-off known as (psychological) coding theory, where special purpose coding languages are required for each context. This has been recently well discussed by Chater (1996), to whom I refer the reader for details.

Another theoretical position, as described above, is to use a measure of the complexity of the process generating the patterns. In algorithmic information theory, this is the length of the program generating the data and model. A familiar approach in this same spirit is to characterize the complexity by the number of free parameters needed to explain a substantial fraction of all the variance in the data. If we boil the data down to N underlying factors or clusters or defining parameters then N can be an index of complexity. We can weight the relative viability of various competing models by including penalties for too many free parameters, as in the AIC (Akaike Information Criterion).

A fundamental problem here though is that all parameters are not created equal when it comes to generating evident complexity. Although under some circumstances it can be proved that numbers of parameters alone provides a good estimate of complexity, as with neural networks (Amari, 1995), in the general modelling context, one particular parameter may explain a very small percentage of the variance in the data, another a very large percentage. An even more worrisome bugaboo is of course nonlinearity. Our apparently high dimensional white noise dominated data may in fact arise from a discrete iterative nonlinear one dimensional system operating in a chaotic domain.

This problem can be approached by the use of nonlinear dynamical indicators of complexity. There are many such indicators, but they may be broadly classed as either dimensional or entropic. The dimensional indicators include the correlation dimension D2 (Grassberger and Procaccia, 1983) and its various modifications $D2_i$ (Farmer, Ott, and Yorke, 1983) and $PD2_i$ (Skinner, Goldberger, Mayer-Kress, and Idekere, 1990). These suppose the system to be a nonlinear deterministic one, and attempt to find the number of dimensions (which may be noninteger) characterizing the attractor in phase space that describes the system's behaviour over time. The correlation dimension has been linked with musical timbre by Monro & Pressing (in press). For psychological time series, which are typically short and noisy, Gregson & Pressing (in press) have recommended $D2_i$ or $PD2_i$ as being more robust.

The entropic indicators look at the rate of loss of information over time of behaviour associated with the attractor. Approximate Entropy, Kolmogorov Entropy, and Lyapunov exponents are three such indicators (Pincus, 1991; Schuster, 1995). Discussion of the details of these indicators is outside the scope of this article. However, the presence of white or coloured noise sources in the data generating system can limit the validity of these approaches. Nevertheless, they are being refined to greater utility. For further discussion Gregson & Pressing (in press) and Abarbanel (1995) are recommended.

Two other approaches to complexity measurement are more pragmatic. In the first we equate pattern complexity with difficulty in learning. Whether the learner is human, animal, or a computational system like a neural network, we can examine time (or number of cases) required to learn a pattern (or solve a problem) to a certain level of accuracy, and asymptotic accuracy level or cumulative number of errors incurred during learning. These must be counted as relative measures, and they may differ depending on whether learning is shown by recognition or recall / production. The main problem here is trying to be certain whether we are predominantly measuring the complexity of the pattern to be learned or the complexity of the resources of the learning system.

Finally, we may engage domain-specific experts to assess pattern complexity. This allows the assessment of patterns in either their intended (in music, notated) form or their realized (in music, performed) form. Yet high inter-rater reliability in domains like artistic expression can be difficult to achieve. My own experience in a pilot survey of expert musicians' ratings of complexity of musical motives, done some years ago, confirmed that consensual judgment is difficult to achieve, apparently due to an evaluative pluralism that reflects the pluralism of late 20th century musical languages and values. In other areas like language or sport performance this may be less problematic.

Sidebar 1: Complexity and maximum likelihood

Before examining how humans handle complexity, it will be useful to illuminate two established yet contrasting views on the nature of perceptual organization. One theory, initially raised by Helmholtz (1910/1962), is based on the *likelihood principle*: sensory input is organized into the most probable event, object or process consistent with the input. The second theory, developed by Wertheimer and subsequent Gestalt psychologists, is founded on the *simplicity principle*: sensory input is organized into the simplest possible coding consistent with the input (Chater, 1996). The simplicity principle is evidently very close to Minimum Description Length principle.

Chater (1996) argues that the simplicity and maximum likelihood criteria are identical. He submits two proofs: one using Shannon's information theory, and one using Kolmogorov's theory of complexity. The first proof is limited in psychological application for the reasons described above, but the equivalence proof based on Kolmogorov theory makes a stronger claim for our attention.

The equivalence between the two views is shown by straightforward mathematics, which I will not review here, for the underlying idea is simple. The essential point (Solomonoff, 1964) is that if events or perceptual objects are generated or naturally describable by processes (e.g., programs), the optimal description of such events is not best based on their own *a priori* equal likelihood, but on the *a priori* relative likelihood of encountering programs (that is, coherent algorithmic processes) that generate such events. Clearly, short coherent programs are more likely to be randomly generated than longer ones. Hence, likelihood of occurrence and simplicity of description form natural dual representations of the same "selection" process.

This is a nice result, though its relevance for human experience is tempered by several considerations, as raised by Chater (1996) himself. First, the intrinsic computational approach takes no account of the goals or interests or the class of potential actions which the perceiving agent may execute. These may drive productive description into a form that is not the simplest representation. Second, organisms do not normally have time to find the simplest or most likely description, but must be content with an approximate solution that is contextually sufficient given limited time and resources. This clearly accords with the perspective of algorithmic information theory given above.

These points are particularly relevant when we consider production tasks, and especially those subject to stringent real-time constraints. Music performance or sight-reading provides the classic normative case. At the time of performance, the performer has certain information structures (perceptual schemata) and production structures (generalized motor programs, dynamical systems) available, and not others, and these will critically shape organization of both input and output. These will be used to generate an effective solution, which may be far from the best. The push towards simplicity is conditional on the existing data and production structures.

For example, a pianist may produce an awkward fingering passage during sight-reading that later study will allow to be replaced by a better, simpler, but rarer fingering pattern that is much harder to find in the space of possibilities, given normative fingering practices. An example which readily comes to mind is the first theme of the last movement of Beethoven's Appassionata, where an initially counterintuitive but functionally more natural fingering of crossing the 4th right hand finger over the 3rd in ascent in moving from A^b to C is recommended by some editors. See Figure 1, measure 20. Variants of this motive permeate the movement, and this motor pattern can be

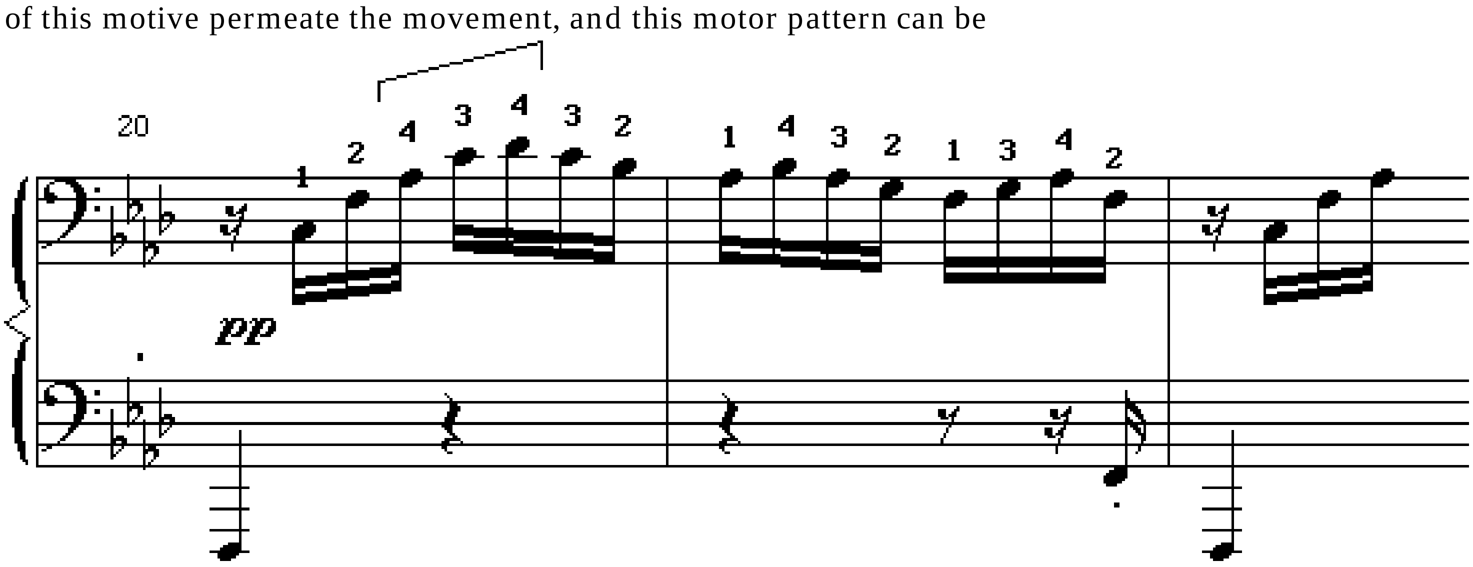

Figure 1: first theme, Mve 3, Sonata No. 23, Op. 57 By L. van Beethoven.

transferred widely because it does not entail thumb crossing and so is not limited by white/black topography. It is functionally more natural, hence expressively simpler, because it automatically supports the phrasing of the melody, providing an accent on the C and tending to shorten the length of the preceding A^b. Initially more natural fingerings, those which previous normative training will tend to bring up, all tend to contradict the phrasing of the melody. My own experience with this is that as a student performer this "counterintuitive" recommendation felt unnatural and I persevered with different normative fingerings, never reaching public performance standard. When I returned to it many years later, this fingering felt immediately appropriate and viable.

The point of this example is that due to performance traditions and habits of the hand, the "counterintuitive" fingering is simpler and more effective, yet rare. This suggests that this limitation on the equation of maximum likelihood and simplicity cannot be patched up by the equating of *conditional* likelihood with *conditional* simplicity. The essential problem is that of encountering new situations, where contextually effective simplicity can be very hard to find (rare). This is not a minor concern, since this must be a major component of fluid intelligence or creative action in performance and problem-solving.

Sidebar 2: Complexity of musical patterns

To make the above considerations more concrete, let us take a test set of musical patterns and assess their complexity. Consider a set of musical patterns made up of conventionally distinct note events defined as all the repeating patterns satisfying the following criterion:

> Members of the set of patterns consist of a repeating cycle of 4 units duration, with 2 tones (one A, one B) at any two positions of the 4-position cycle.

This yields the following 6 patterns,

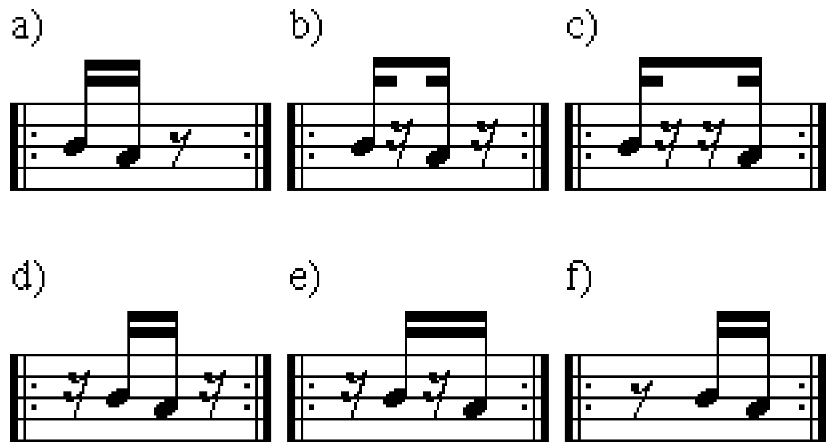

Figure 2: 6 related repeating rhythmic patterns

plus 6 more with identical rhythmic structure but with reversed note order (AB rather than BA). There is no reason why note order will affect complexity of this pattern and hence it is sufficient to treat the 6 cases of Figure 2.

Note that, given perfectly literal and metronomic performance, as from a computer notation package realization, or an idealized literal performer, and apart from an identified starting point, the aural results from cases a, c, d, & f are identical (forming category 1), as are cases b & e (category 2). Thus, if listeners were asked to discriminate between the patterns within each category played by a literal performer or such a software package, based on fade-in from zero volume of a looped recording (as done for example in the classic pattern perception work of Garner (1974) or Handel & Todd (1981) —see Pressing, 1983 for a music-related discussion) they would fail to distinguish members of the same category, as there would simply be no information in the recording allowing them to do so.

If, in contrast, the recordings of the patterns were made by musicians, then we should find differential evidence of distinct mental models (metrical frameworks) at work. The evidence will be of two kinds: dynamics and timing. Not only should mean and variance of dynamics and interonset intervals of the two notes betray the distinct mental model, but so should the patterns of correlation between these variables. Although this experiment has not yet been done, this supposition is to be expected on musical grounds, and on the basis of concordance with recent experimental work showing that structural equation modelling (which analyzes the performance covariance matrix) can distinguish between mental models used by musicians in patterns which have the same literal performance target (for timing of 4:3 polyrhythms, Pressing, Summers & Magill, 1996; for timing of certain African musical patterns, Magill & Pressing, in press).

How do such distinctions reflect underlying complexity? If we adopt the complexity-as-computational-cost idea, then there are cognitive costs incurred in maintaining the metrical framework, motor and cognitive costs in the physical production of the pattern, and cognitive costs in reconciling the pattern's placement within the metrical framework. Although there may well be interactive effects between these three costs, if we assume separability of costs as a starting approximation, then all 6 patterns will have the same framework cost CF, and all patterns within each category will have the same production cost, CP1 or CP2, for categories 1 and 2 respectively. The residual differences are due to the pattern-framework reconciliation cost CR, and if we consider that for slow to moderate tempos CP1 and CP2 are not likely to be very different (CP1 somewhat greater, to be sure), then the dominant effect on complexity in production will be via reconciliation cost CR.

Pattern-framework reconciliation corresponds in musical terms to the handling of syncopation. For the purposes here I will use the ideas of syncopation-generated cognitive cost found in the autonotation software package *Transcribe,* developed by Jeff Pressing, Xudong Cao, and Peter Lawrence (Pressing & Lawrence, 1993) for the Macintosh computer, and available from the author. *Transcribe* takes MIDI or audio input and produces a cognitively optimal musical notation. The package distinguishes 5 types of syncopation, which are ordered to provide an implicit measure of cognitive cost. Cost is specified at each meaningful level of pulse; in this case that comprises the 8th-note level, divided in 2 parts, and the quarter-note level, divided into 2 or 4 parts. Five syncopation types can occur, and they are, in order of increasing cognitive cost: *filled* (note at each of the four 16th-note positions or two 8th-note positions, e.g., pattern *b,* quarter-note level), *run* (note in first position followed by a run of others, e.g. pattern *a* at quarter-note level), *upbeat* (note in first position plus one or more consecutive pickup(s) to first position note, e.g. pattern *c* at quarter-note level), *subbeat* (this type cannot occur in a cycle of length 4), and (fully) *syncopated* (starting and ending on off-beats, e.g. pattern *d* at quarter-note level, pattern *e* at both quarter-note and 8th-note levels). In addition, a *null* case occurs when there is no subdivision of the given level—that is, if the time window has no notes or only a single on-beat note.

To calculate pattern complexity due to reconciliation, the following procedure may be used. We assign to each syncopation type the following weights (ranks): *null*, 0; *filled*, 1; *run*, 2; *upbeat*, 3; *subbeat*, 4; and (fully) *syncopated*, 5. Then we calculate the pattern complexity as the sum of the complexity ranks at the quarter note level and the average of the complexity ranks of the two eight-note windows making up the pattern. This yields

Pattern	Quarter note syncopation (weight)	Eighth note syncopations (weights)	Pattern complexity
a	run (2)	filled (1), null (0)	2.5
b	filled (1)	null (0), null(0)	1.0
c	upbeat (3)	null(0), upbeat (3)	4.5
d	sync (5)	upbeat (3), null (0)	6.5
e	sync (5)	sync (5), sync (5)	10.0
f	sync (5)	null (0), filled (1)	5.5

Table 1. Computed complexities of the six patterns of Figure 2.

Hence the ranking of complexity of the patterns would be, in increasing order, *b, a, c, f, d, e,* which is in accord with standard musical judgment. This complexity ranking allows us to make predictions about stability of these patterns and the likelihood of transitions occurring between them under conditions of speeded or otherwise stressed performance, using the simple reasoning advanced by Pressing (1995): more complex patterns will be less stable than simpler patterns, and spontaneous transitions between patterns due to skill limitations will always move in the direction of reducing complexity. Thus, we predict stability will increase with pattern type in the order *e, d, f, c, a, b,* and that spontaneous transitions will always be from left to right in this list. If we further assume that skilled performers can readily distinguish category 1 patterns from category 2 patterns cue to their very different motor demands, then spontaneous transitions should be dominantly intracategory, of the form $e \rightarrow b$ and $d \rightarrow (f,a,c), f \rightarrow (a,c), a \rightarrow c$. These are exactly the kinds of common errors found in performing these patterns, a fact especially notable in West African drum ensemble music, where repeating patterns of these kind are widespread. (This also appears to be the case in other repertoires where these kinds of figures occur, as in the inner (e.g., viola) parts in some of Johannes Brahms' symphonic works.)

In performance, experienced musicians soon become aware of such pattern errors, and will correct, if possible, back to the target rhythm. This shows that *intentional* transitions can move in the opposite direction in our pattern ranking, acting to increase complexity. Over the course of an extended learning trial, the performer may repeatedly get trapped in a spontaneous phase transition to a less complex version of the pattern, recognize the error, and then correct it either by direct phase adjustment (speeding up or slowing down until alignment is reached), or by stopping and then restarting after mentally realigning to the beat.

What are the consequences of complexity for human behaviour?

Space constraints prevent saying very much about this topic here. Evidently though, complexity is dealt with by humans by the development of automatic routines, information bases and heuristics (for interpretation, decision-making, action, etc.) that both circumvent the impact of normal limitations in memory, attention, and control, as described by classical formulations of skill learning (e.g. Fitts, 1964; Rasmussen, 1986), and extend the hierarchy governing the task to deeper and richer levels, building expertise. When complexity overwhelms the system's capacity, performance breaks down by faltering, stopping, or by substitution of a simpler behaviour that typically still retains some contextual fit, as we saw above in the musical pattern example. Adaptive practice reduces the likelihood of breakdown, and promotes rapid recovery to contextually apt behaviour.

Conclusion

I have tried here to sketch the potentials and pitfalls of evaluating pattern complexity and using this to predict behavioural consequences. The domain of music has provided the foundation for the examples, even though the approach is by no means limited to the acoustic domain. The results are in encouraging accord with musical knowledge of performance practice, and they suggest that more systematic experimental investigations be performed.

References

Abarbanel, H. (1996). *Analysis of observed chaotic data.* New York: Springer Verlag.

Amari, S. (1995). Learning and statistical inference. In M. Arbib (Ed.) *The handbook of brain theory and neural networks.* (pp. 522–526) Cambridge: MIT Press.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review* 103, 566–581.

Farmer, J.D., Ott, E. and Yorke, J.A. (1983). Dimension of chaotic attractors. *Physica 7D*: 153–180.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.) *Categories of human learning* (pp. 243–285). New York: Academic Press.

Garner, W. (1974). *The processing of information and structure.* John Wiley and Sons: New York.

Grassberger, P., and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica* 9D: 189–208.

Gregson, R.G., & Pressing, J. (1997). Dynamic Modeling. to appear in L.G. Tassinary, J.T. Cacioppo, G. Berntson, (Eds.) *Principles of Psychophysiology: Physical, Social, and Inferential Elements.* Cambridge University Press: New York.

Handel, S., & Todd, P. (1981). Segmentation of sequential patterns. *Journal of Experimental Psychology: Human Performance and Perception*, 7(1), 41–55.

Helmholtz, H. L. F. von (1910 / 1962). *Treatise on physiological optics* (Vol 3; J.P Southall, Ed. & Tran.) New York: Dover.

Magill, J., & Pressing, J. (in press) Asymmetric cognitive clock structures in West African rhythms. *Music Perception.*

Monro, G., & Pressing, J. (in press). Sound visualization using embedding: The art and science of auditory autocorrelation. *Computer Music Journal.*

Pincus, S.M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences* 88: 2297–2301.

Pressing, J. (1983). Cognitive isomorphisms in world music: West Africa, the Balkans, Thailand and Western tonality, *Studies in Music*, 17, 38-61.

Pressing, J. (1995). Testing dynamical and cognitive models of rhythmic pattern production. in Glencross, D.J., & Piek, J.P. (Eds). *Motor control and sensory-motor integration: Issues and directions.* pp. 141–170, Amsterdam: Elsevier Science.

Pressing, J., & Lawrence, P. (1993). *Transcribe*: a comprehensive autotranscription program. *Proceedings of the 1993 International Computer Music Conference*, pp. 343-345.

Pressing, J., Summers, J., & Magill, J. (1996). Cognitive multiplicity in polyrhythmic performance. *Journal of Experimental Psychology: Human Perception and Performance* 22, 1127-1148.

Rasmussen, J. (1986). *Information-processing and human-machine interaction: An approach to cognitive engineering.* Amsterdam: North Holland.

Rissanen, J. (1989). *Stochastic Complexity in Statistical Inquiry.* Singapore: World Scientific.

Schuster, H.G. (1995). *Deterministic chaos: an introduction.* New York: VCH.

Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423, 623–656.

Skinner, J.E., Goldberger, A.L., Mayer-Kress, G., and Idekere, R.E. (1990). Chaos in the heart: implications for clinical cardiology. *Biotechnology* 8: 1018–1024.

Solomonoff, R. J. (1964). A formal theory of inductive inference, Parts 1 and 2. *Information and Control*, 7, 1–22, 224–254.

Traub, J. F., Wozniakowski, H., Wasilowski, G. W. (1988). *Information-based complexity.* NY: Academic.

Zemel, R. S. (1995). Minimum description length analysis. In M. Arbib (Ed.) *The handbook of brain theory and neural networks.* (pp. 572–575) Cambridge: MIT Press.