a journal of the Society for Music Theory

Volume 1, Number 2, March 1995

On Considering a Computational Model of Similarity Analysis

Thomas R. Demske

KEYWORDS: similarity, REL, clustering, decision, post-tonal analysis

ABSTRACT: A clustering program is proposed for identifying salient similarity relationship trends within populations of pitch-class sets abstracted in post-tonal analysis. That program requires an evaluation component, but attempts to define one are frustrated. The similarity relation itself is too abstract to imply guidelines for its own application, and other potential criteria resist formal implementation. Even less rigorous applications of the more commonly used tools in post-tonal analysis are susceptible to the concerns raised here.

---

[1] A few months ago, I considered applying conceptual clustering techniques to identify context-specific similarity relations among pitch-class sets. That project quickly stalled for lack of suitable evaluation criteria, a problem which I take as the subject of this essay. The issues here are not necessarily new, but they are persistent. Recent SMT e-list interest in formal analysis suggests the timeliness of reviewing these issues now from an applicational standpoint.

[2] Many fields use clustering to establish meaningful distinctions within arbitrary data sets. The data are events described as n-tuples and interpreted as points in n-dimensional space. Points lying closer together according to some distance specification "cluster" into groups distinguished by boundaries. Clustering comes in many flavors. The conceptual clustering model in Michalski and Stepp 1983 is interesting because it determines cluster membership, not only through predefined metrics, but also by the resultant shape of the cluster—its "conceptual cohesiveness." Conceptual clusters are context-specific, in that events cohere on the basis of their relationships to other events in the data set.

[3] I was first attracted to conceptual clustering through ongoing research into melodic grouping in piano scores, an early stage of which is documented in Demske 1993. Clustering is fruitfully applied there to isolate and address causes of rule base misfire. Might the technique also prove useful in post-tonal analysis, despite the far more diffuse nature of the problem domain? Given a collection of pitch-class set segmentations abstracted from a composition, the clustering implementation would "learn" the most salient, structural trends particular to that population.

[4] Cluster membership indicates relatedness, an idea which is fundamental in pitch-class set analysis. Most analytical assertions eventually require boundaries in a conceptual space to distinguish one or more sets (or sets of sets, etc.) from others. Mainstream analytical set theory tends to recognize only a few types of musically motivated boundaries. These account for the familiar TnI set classes, transformational partitionings, "referential" collections and their subsets, and so

forth. In all cases, kinship and contrast between objects disposed respectively on one or both sides of a boundary provide the potential for a syntax. This simple duality can spawn sophisticated arguments about form, association, and process.

[5] "Similarity" in pitch-class set theory refers narrowly to special inclusion-based relationships. These are commonly understood as reflecting degrees of context-free, "aural similitude" (Morris 1980), although some writers avoid any perceptual interpretation Forte 1973. Isaacson 1990 surveys several similarity functions proposed over the years, while offering his own "IcVSIM" as an alternative.[(1)] Most similarity functions take two sets as arguments and return a single, usually scaled value determined through subset counts; Morris's ASIM, Rahn's ATMEMB, and Lewin's REL are perhaps the best known examples (Morris 1980; Rahn 1980; Lewin 1980).

[6] Similarity functions seem an ideal test bed for clustering because of the apparent simplicity with which they model distance in set-class space. However, conceptual clustering is a learning algorithm. Without a means for evaluating prospective clusters, the algorithm's output will likely make little sense. How might similarity-based clusters be evaluated, assuming the identification of "salient trends" as a loosely-defined goal?

[7] I use Lewin's REL here for illustration, although other similarity index functions would work as well. Consider a pivot, P = {0147}, and two target sets, T1 = {01369} and T2 = {047a}. REL(P,T1) = 0.783 and REL(P,T2) = 0.636. Drawing a boundary around all sets Y such that 0.783 >= REL(P,Y) > 0.636, P and T1 cluster together on one side, while T2 falls on the other. This boundary might be treated like any other in analysis. However, nothing in the interpretation of REL compels us to draw any particular boundary. Assume another set T3 = {0134}, where REL(P,T3) = 0.529. We could keep our original boundary, or now draw a new one separating P, T1, and T2 from T3. In fact, we could draw a boundary separating P from all of the target sets, or another which includes P and the targets all on one side. **Figure 1** shows the relationships just described, with vertical bars indicating potential boundary sites (the vee is explained later):

[8] Comparing REL values tells us only where boundaries may be drawn, not where they should be drawn. Barring *ad hoc* provisions in the interpretation, all sets are similar with respect to each other. Only the degree of similarity varies. This makes it difficult to evaluate competing clusters, such as those distinguished by alternately drawing boundaries through the third and fourth vertical bars in Figure 1 above. An *a priori* cutoff point X would help, whereby only targets T with REL(P,T) >= X are considered "similar" in an absolute sense. (TnI "equivalence" supports analogous appeals to the absolute.) But could such a point be meaningfully determined here? Statistical analysis of how the REL values at issue are disposed around an average or mean begs another question: are all of REL's returns with respect to a given pivot commensurable in a meaningful way (I am not sure that they are), or are they merely indications of a partial ordering?

[9] Other boundary ambiguities must be considered. We can compare two returns of REL only when an instance of the same set class appears in both calls. Rahn 1989 demonstrates this tangentially with his DATM function, derived from ATMEMB, but the same applies to REL and comparable similarity index functions. That argument common to both calls is the pivot, P. **Figures 2–4** show how the potential boundaries change when our old target sets are cycled through P. {01369} is the pivot at Figure 2. Unlike the situation earlier at Figure 1, {047a} and {0147} at Figure 2 will always appear on the same side of a boundary. Compared to Figure 1, taking {047a} as the pivot at Figure 3 means that {047a} and {0134} can no longer be grouped as a pair apart from {01369} and {0147}. In Figure 4, it suddenly becomes possible to place the pairs, {0134} and {01369}, and {0147} and {047a}, on opposite sides of a boundary.

[10] Figures 1–4 together represent a tiny bubble in that "(staggeringly complex) network of relations" revealed by functions like REL (Rahn 1980, 494). The downside to this attractive notion is that it means more potential conflicts among REL-derived clusters. Is {0147} more related to {01369} than it is to {047a}? Expressed through the totality of REL returns in Figures 1–4, the answer is yes and no. For example, {0147} and {01369} cluster together without {047a} under the boundary drawn through the vee at Figure 1, while {0147} and {047a} cluster together without {01369} under the two boundaries drawn through the vees at Figure 2.

[11] Critical REL returns are interpreted directly at Figure 1 and indirectly at Figure 2, suggesting that one criterion for evaluating conflicting clusters might be to prefer the more musically intuitive one. "Intuition" is difficult to pin down, however. For example, Rahn 1980 claims special intuitive status for ATMEMB, but others may find the notion of "mutual

embedding" less compelling. All else being equal, ATMEMB(A,B) remains constant so long as the sum of EMB(/X/,A) and EMB(/X/,B) remains constant, and so long as neither term equals zero.(2) This means A and B can differ considerably in the number of X types they separately include, and yet still return high ATMEMB values. REL seems more "intuitive" in this respect because it takes the product of EMB(/X/,A) and EMB(/X/,B). While REL(A,B) still increases along with either one of EMB(/X/,A) and EMB(/X/,B) (assuming nonzero values), it does so more quickly when *both* subterms increase. This makes sense, since we would expect—in the abstract—that two sets with convergent inclusion capabilities will have more "aural similitude" than two sets which diverge in their capabilities.

[12] The mechanics of ATMEMB, REL, and comparable functions present other potential barriers to intuition.(3) But those pale in comparison to the problems which arise when similarity function returns are viewed in their totality. Assume two sets WT1 = {0246} and WT2 = {0248}. Considered in isolation, REL(WT1,WT2) = 0.846 bespeaks urgency, emphasizing the difference in "aural similitude." That changes by adding CHR = {0123} to the mix. WT1 and WT2 now seem quite similar, by virtue of their relationship to CHR: an absence of odd interval classes binds the wholetone sets together. Relationships are thus colored by other relationships. In Rahn's vast network, where everything is connected, the result is a flat, homogeneous black (or white, depending on perspective). But when only a portion of the network is considered, glimmers of color surface, forming kaleidoscopic patterns of arbitrary complexity. Tracking those patterns can strain intuition mightily.

[13] So, who is to say whether the direct relation described in paragraph 10 above is more intuitive than, and thus necessarily takes precedence over, the indirect one? Rahn 1989 briefly explores the idea of following only "paths 'of least resistance'" through a network, while Block and Douthett 1994 offer a new way to focus on specific paths.(4) But neither indicates whether certain paths should generally be preferred over others, given an arbitrary collection of sets. Both Hindemith 1942 and Friedmann 1990 informally consider interval class 1 content as being especially determinative of chord quality. Perhaps those paths reflecting aspiration toward high semitone counts— REL hierarchies based on chromatic pivots—are therefore to be preferred *a priori*.

[14] Such concerns raise the inevitable issue of perception. Evaluating similarity clusters with perceptual criteria is complicated because we can attend to the same music in many different ways. Still, there are limits. **Figure 5** graphically displays similarity relations between successive piano chords in the first movement of Messiaen's *Quatour pour le fin du temps*. The piano part is a recurring sequence of 29 chords set to a rhythmic ostinato of 17 durations; chord 30 is thus the same as chord 1. Higher ASIM, ATMEMB, and REL values plotted on the graph indicate greater similarity. (ASIM is adjusted to facilitate comparison with the other functions.) Figure 5 shows the three functions diverging in their values, but mostly agreeing about peaks, valleys, and plateaus. The relation expressed in moving from chord 1 to 2 appears above point 2 on the x axis; that from chord 2 to 3 appears above point 3; etc. The middle chord in each three-chord succession is the pivot, while the two flanking chords are the targets.

[15] The graph resembles a "harmonic fluctuation" analysis (Hindemith 1942), and should capture a comparable aspect under the most obvious interpretation of the similarity relationship. However, and despite great effort in attending, I can discern no correspondence at all between musical transitions in the *Quatour* and those represented symbolically on the graph. For example, the strong contrast that Figure 5 reports between chords 15 and 16 simply does not conform to Messiaen's smooth realization. Taking this perception into account suggests preferring a boundary through the vee at **Figure 6**. This is not entirely satisfactory, however, even allowing for the wide disparity in REL values. To my ear, the move between Messiaen's chords 16 and 17 is in fact more striking than that between his chords 15 and 16. Perhaps a boundary should therefore separate chords 15 and 16 from chord 17. The REL scheme of Figure 6 cannot accomodate that boundary, so another pivot is required. (see **Figure 7**).

[16] The pivot in Figure 7 no longer coincides with the original point of transition, a fact which obscures the harmonic fluctuation interpretation of similarities.(5) Perhaps this shows harmonic fluctuation to be an insufficient standard for perceptual cluster evaluation. That in turn suggests another complication when appealing to perceptual criteria. Exactly how similarity relationships might inform perception is speculative. Gradations of "smooth" chord succession, strength of motivic association, and degree of harmonic contrast in form delineation are only the first potential manifestations which spring to mind. The idea that any formal evaluation procedure could embrace all of the possibilities seems untenable. On

what bases would a partial set of possibilities be selected for implementation? Since the different criteria may address different—and possibly conflicting—aspects of perception, how would the application of one criterion be coordinated with that of another?

[17] Apart from the manner in which similarities shape perception, there is the further question of the extent to which they do so. For example, strong contrast is perceived in moving between Messiaen's chords 6 and 7, despite the initial plateau of Figure 5. This implies drawing a boundary through the vee at **Figure 8**, even though that is not an allowed boundary site. Nothing in the micro-network of REL relationships among Messiaen's 29 chords will support this boundary, since chords 6 and 8 are members of the same set class. Clearly, competing factors inaccessible to similarity functions and the pitch-class set model in general influence perception enormously.(6) In order to realistically evaluate similarity clusters according to perception, we must first have a means of isolating and gauging the influence of these factors.

[18] Such a means does not appear to be forthcoming at present, given current knowledge about music perception. Implementing practical event structures and distance functions whose resultant clusterings satisfy general perceptual criteria is thus but a dim hope. We cannot expect a simple automaton processing high-level abstractions to have much to do with perception. For now at least, the place for perceptual criteria remains in cautiously selecting input and very carefully weighing output.

[19] Analytical music theory usually holds a less naive view of context than that described in the chord-by-chord scenario above. Relationships not immediately perceptible can nevertheless inform musical structure. This suggests another possible evaluation criterion: goodness of fit between a boundary and an analytical goal. The clustering program could take as input segmentations together with an analytical paradigm. For example, a short movement strongly projects the large-scale succession of pitch-class sets shown at **Figure 9**. The analyst wants to emphasize that the movement is through-composed. A REL clustering solution appears at **Figure 10**, with two boundaries indicating progressive motion away from the opening set in "similarity space."

[20] The boundaries of Figure 10 are a good fit to the analytical description, but they are not a completely honest fit. **Figure 11** shows another REL clustering of the same sets. The single boundary here at least suggests a different form, the symmetrical, ternary structure distinguished by outer sections equally distant from B in similarity space (**Figure 12**). The goodness of fit criterion now appears suspicious, in that the analyst receives only supportive solutions, and is shielded from possibly damaging ones.

[21] REL and comparable similarity functions are not used much in analysis today. Still, my basic concerns extend beyond a particular function and its potential for rigorous application in a computer model. Post-tonal analyses typically involve countless decisions about boundaries, ranging from relatively low-level segmentation selections to high-level identifications of referential collections in refractory textures. The decisions interact and motivate one another freely across all levels of description, and often depend on abstract criteria whose perceptual relevance and provability is speculative. We all slip, sometimes conspicuously: dubious examples of "composed out" sets in undergraduate textbooks; untempered enthusiasm for a single octad class underlying a composer's entire output. Exuberant flexibility is essential to the post-tonal analytical dialectic. But I suggest checking it occasionally by asking what constraints we would impose on a machine approaching the same decisions.

Thomas R. Demske
Connecticut College
Department of Music
270 Mohegan Avenue
New London, CT 06320
USA
trdem@conncoll.edu

Works Cited

Block, Steven and Jack Douthett. 1994. "Vector Products and Intervallic Weighting." *Journal of Music Theory* 38/1: 21–41.

Demske, Thomas R. 1993. "Recognizing Melodic Motion in Piano Scores: Rules and Contexts." Ph.D. diss., Yale University.

Forte, Allen. 1973. *The Structure of Atonal Music*. New Haven: Yale University Press.

Friedmann, Michael L. 1990. *Ear Training for Twentieth-Century Music*. New Haven: Yale University Press.

Hindemith, Paul. 1942. *The Craft of Musical Composition*, v. 1. Trans. Arthur Mendel. New York: Associated Music.

Isaacson, Eric J. 1990. "Similarity of Interval-Class Content Between Pitch-Class Sets: The ICVSIM Relation." *Journal of Music Theory* 34: 1–28.

Lewin, David. 1980. "A Response to a Response: On Pcset Relatedness." *Perspectives of New Music* 18/2: 498–502.

Lewin, David. 1977. "Forte's Interval Vector, My Interval Function, and Regener's Common-Note Function." *Journal of Music Theory* 21/2: 194–237.

Michalski, Ryszard S. and Robert. E. Stepp. 1983. "Learning from Observation: Conceptual Clustering." In *Machine Learning: An Artificial Intelligence Approach*, R. Michalski, J. Carbonell, and T. Mitchell, eds. Palo Alto, CA: Tioga Publishing Company.

Morris, Robert. 1980. "A Similarity Index for Pitch-Class Sets." Perspectives of New Music 18/2: 445–460.

Rahn, John. 1989. "Toward a Theory for Chord Progression." *In Theory Only* 11/1,2: 1–10.

———. 1980. "Relating Sets." *Perspectives of New Music* 18/2: 483–498.

Footnotes

1. Isaacson has also written a handy, inexpensive, and well-documented program which calculates the values of various similarity functions for arbitrary sets. DOS and Windows versions are available.
Return to text

2. The embedding function EMB(/q/,r) countsthe multiplicity of q-type sets included in set r; see Lewin 1977.
Return to text

3. Blind subset polling is a basic source of such barriers. Two REL calls with the same pivot may yield identical results, and yet differ with respect to the types of subsets counted. Ignoring the degree of this difference when comparing REL value spreads strikes me as questionable.
Return to text

4. Block and Douthett's (1994) "vector product" parts with other similarity measures in that the pivot argument is not a set, but rather an arbitrarily specified standard of subset content. What was once merely a "staggeringly complex" network of relationships is now potentially infinite. Otherwise, vector product similarities continue to admit the same critique voiced here with respect to REL.
Return to text

5. Hindemith (1942) types chords according to intrinsic intervallic properties, which is very roughly analogous to holding one particular set of ideals as a constant pivot argument to REL. Consequently, variance between pivot and point of transition would not conflict with Hindemith's absolutist conception of harmonic fluctuation. The relativistic conception pursued here in the interpretation of Figure 5 conforms more closely to that implied in Rahn's (1989) "theory of chord progression."
Return to text

6. Rahn (1989) presupposes a "theory of instances" to account for discrepancies between the heard surface and DATM's predictions. That theory remains incomplete.
Return to text

Prepared by Cara Stroud and Rebecca Flore, Editorial Assistants