

INTERPRETING RA MODELS OF NOTE-ONSET INTERACTIONS TO DISCERN AND EXPLAIN CLAVE DIRECTION

Mehmet Vurkaç

Seattle University, Electrical and Computer Engineering
901 12th Ave., BA 209, Seattle WA 98122, USA

ABSTRACT

Reconstructability Analysis (RA) was used to generate and evaluate models of sequences of musical note onsets. These sequences were classified into four classes in three musical contexts based on a musical grammar akin to “harmony” but concerning the timing of note events. (Having emerged only in certain societies and from the cultural interactions between Yorùbá, various Angolan peoples, Iberians, and the native peoples of the pre-Columbian Americas, this musical grammar is found only in some of the musics of South America and the Caribbean.)

A variety of search approaches and search criteria were used in the *OCCAM 3* modeling engine, including BIC, AIC, and information, to extract classification information from rhythmic sequences. (‘Rhythm’ here refers to the timing of note events, not necessarily to any steady pulse, repetition, or percussive instrumentation.) The models discovered reflect trade-offs between complexity (degrees of freedom) and simplicity in how they prioritize some note-event interactions over others.

These close to 10,800 randomly generated idealized patterns have a 16-dimensional input space and a four-dimensional output space (for a total of 20). By limiting the output classes to whether a certain clave direction was present or not, the search space dropped to 17 dimensions. Hence, it was paramount to develop search trade-offs. The approaches developed for traversing the search space efficiently are detailed in the paper. Interpretation of these models is compared with several criteria for clave-direction determination deduced from systematic observations of four master musicians (and from deep exposure to the underlying musical practice).

BIC was found to be the most beneficial modeling criterion, with information second, and AIC third. The models discovered through RA provide insight into how clave direction arises in all possible rhythm sequences in that they highlight certain rhythmic schemata known to be strong indicators of clave direction while leaving out others as well as highlighting other rhythmic relationships for discerning clave direction that have been discovered through musicological (qualitative) means.

In some cases, models considered in this study show trade-offs suggesting that sufficient musical insight may be gained by considering interactions of fewer note events. In other cases, the interpretation of clave through an analogue to algebraic elimination gains unexpected support through RA modeling.

We thus demonstrate the ability of RA to model an intricate and culturally specific (not broadly accessible) musical construct in terms of discrete note events and their

Interpreting RA Models of Note-onset Interactions

interactions in such a way as to mirror a human understanding of the corresponding musical practice.

Keywords: RA, culture, music, rhythm, clave

INTRODUCTION AND BACKGROUND

The musical concept, clave direction, addressed in this work is one that is not familiar to many people. Furthermore, among those familiar with it, it is rather controversial, and has been likened to politics and religion because of the strong feelings and opinions it engenders among musicians, music-theorists, and ethnomusicologists.¹ The overall research program that this article and corresponding presentation form a small part of aims to offer a logically consistent and parsimonious², data-driven, and culturally informed explanation of this musical practice.³

Taking a systems view of clave, we first present the notions of music theory and musical grammar which lead to the notions of consonance (Parncutt, 2008), accordance (Locke, 1996, 86), or harmony (Vurkaç, 2012a) as a frameworks of temporal organization—a rhythmic code. (We emphasize “temporal” over “rhythmic” to stress the point that clave direction concerns all instruments (including vocals), not just percussion (Cruz, 2004, 60; Vurkaç, 2012a, 48–49). While both clave itself and the research program this work is part of are broader in scope, the focus of the present article is the use of RA to interpret clave direction based on the interaction terms in RA models of note onsets (the beginnings of sounds, i.e., ‘rhythm’ in the music-theoretic sense).⁴

RA was used to generate and evaluate models of such sequences of musical note onsets. These sequences were classified into four classes in three musical contexts based on a musical grammar akin to “harmony” but concerning the timing of note events. (Having

¹ Cruz *et al.* explicitly argue that “[i]t’s a great understatement to say that ‘clave is something that should be felt, and not talked about!’” (Cruz, 2004, 62).

² This explanation, in addition to being free from exceptions (when applied within any genre), is based on a small set of defensible axioms derived from the literature inclusive-or from the analysis of data. In contrast, Cruz and co-authors state, “the more you learn, the more you’ll realize there *is* to learn about clave” (Cruz, 2004, 58). A consistent and parsimonious analytical explanation of clave direction, on the other hand, can remove the increasing complexity that one gets exposed to when clave explanations are instance-based and presented as exceptions to a vague but convenient starting point. This point is demonstrated in (Vurkaç, 2012a).

³ Although clave is far from being a household term, it is a central organizational element of numerous highly popular musical genres around the world. One sampling of these musical forms and genres follows: *son, salsa, songo, mozambique, changüi*, and the many forms of *rumba* from Cuba; *konpa* and *méringue/kadans* from Haiti; *calypso* from Trinidad & Tobago; classic *plena* from Puerto Rico; *scratch/fungi* from the Virgin Islands; *biguine* and *zouk* from Guadeloupe and Martinique; *mento* from Jamaica; *zydeco* and *second-line* from the US; *choro/chorinho, maracatu de baque virado, frevo, axé, samba-reggae, samba-Afro, samba-salsa*, and the various forms of *samba*-proper from Brazil; and *candombe* from Uruguay (not to be confused with the religion, *Candomblé*, from Brazil, and its music).

⁴ In justification of the use of onsets, we offer the following. “We distinguish a melody by its pattern of duration and accentuation; notes that fall on downbeats or other important rhythmic junctures usually become the tune’s most recognizable ones. ***In fact, the stress pattern is so important to melody that we can identify familiar tunes when their rhythm is tapped out on a single pitch – although we do not do as well when the tune’s rising/falling tones are used but they are made the same length*** (Jourdain, 2002, 81; emphasis added)

Interpreting RA Models of Note-onset Interactions

emerged only in certain societies and from the cultural interactions between Yorùbá, various Angolan peoples, Iberians, and the native peoples of the pre-Columbian Americas, this musical grammar is found only in some of the musics of South America and the Caribbean.)

A Systems View of Clave: Music-Theoretic Background on Clave Direction

Temporal Harmony

Clave direction (henceforth *CD*) is a type of temporal harmony (Vurkaç, 2012), or rhythmic consonance (Parncutt, 2008). Just as the notion of musical harmony deals with the vertical (simultaneous) and horizontal (sequential) relationships (*harmoniousness*, according to a combination of the physics of music, innate aspects of human pitch perception, and European cultural constructs and values) in terms of the pitches of notes (the frequencies of each fundamental), *CD* is a similar system in that it deals with *harmoniousness* in a subjective cultural manner, but with respect to the time aspect of music. It says that the timing and spacing of note onsets in any given pattern—especially though not limited to repeated patterns—can encode a preference for a broad yet restricted class of rhythmic options (timings and spacings).

Such preferences are culturally specific (Parncutt, 2008; Cazden, 1945)—they do not carry the same meaning or weight outside of a given musical idiom—and are far from obvious. (For example, even copying a large portion of a pattern does not guarantee that the derivative pattern is in the same direction as that pattern.)⁵ To communicate this, Agawu refers to “commonplaces rich in associative meaning for cultural insiders” (Agawu, 2003, 73).

Music Theory

In music, concepts of harmony and consonance reflect human experience (perception) and the emergence and cognition of social practices. Such cognition constitutes a *music theory*, a description of cultural practices. The goal of a music theory is “to be a formal description of the musical intuitions of a listener who is experienced in a musical idiom” (Clarke, 2004).

What is typically considered *the* music theory is not the only one. It primarily describes music of western- and northern-European origins and musics of the era of globalization (due to the worldwide influence of Euro-American culture). The court musics that developed elsewhere around the world—such as in Indonesia, Thailand, China, Persia, Turkey, and north and south India—have their own systems of music, which may or may not have been worked into formal theories. These systems are related to but do not exactly correspond or conform to the music-theoretic explanations that exist for the

⁵ For our purposes, we define “pattern” as that property—of a sequence which has some aspects in common with other sequences and some aspects not in common with those other sequences—which engenders categories and enables classification of different sequences into categories.

Interpreting RA Models of Note-onset Interactions

baroque, classical, romantic, and other periods of European court music, and later, for global popular music.⁶

When we examine any of those musical practices, it is as if there is a code of what to do and what not to do (or do a lot less of) in each one. Poignantly, one author of a music reference book directly asks “What is the secret code behind so many musical compositions?” (Martineau, 2008, back cover). (It turns out to be highly mathematical.) He continues, “... music theory, in essence, is primarily descriptive and not prescriptive. The tendencies and practices in music are only observed and catalogued upon analysis, after the fact.” (Martineau, 2008, 1). Music theories, then, are descriptions of musical practices. They arise out of the effort to explain how music (in a certain culture at a certain time) is typically made. They are derivative of the practices themselves. Hence there can be separate music theories reflecting the emergent preferences of different cultures in their practices of music-making.

Is clave such an emergent property of a network of musics? Considering the musicological, music-theoretic, and mathematical literature on the various aspects of clave *direction*, we find that clave is a central organizational principle in a certain family of musics—it is an emergent grammar.

Clave Direction as a Music-Theoretic Grammar

Much of the evidence for *CD* as a music-organizational principle in the literature—Herder (1972), Stone (1985, 142), Soebbing (1988), Tobias (1988), Malabe *et al.* (1990, 9–33), Mauleón-Santana (1993, 48; 1999, 8–17), Washburne (1997, 66), Machado *et al.* (2002, 10), Lehmann (2002), Rodriguez (2003, 41–44), Cruz (2004), Sublette (2004), and Spiro *et al.* (2006, 12–17)—is summarized in the introduction to (Vurkaç, 2012a).

Further evidence for the rhythmic-regulative role and temporal-harmony function of “wide-sense clave”—also Chor’s non-surface “framework for rhythmic organization, informing and constraining rhythms” (Chor, 2010, 59)—is found in Wilson (2010), Garibaldi *et al.* (1999, 66), Kwabena Nketia (1963, 78), Stover (2009, 8), and Peñalosa (2009, 84).

In addition to expressing the importance of clave and *CD*, many of the aforementioned authors—Cruz *et al.*, Malabe *et al.*, Peñalosa, Rodriguez, and especially Mauleón-Santana and Lehmann—present analyses of vertically in-clave sets of musical patterns in which they point to individual notes that determine the clave direction of the phrase in question. While we agree—and in fact find it very important to recognize—that a single note (or a single rest) can and often do determine clave direction, the purpose of the wide-

⁶ We find at least two modes in which the musical framework discussed in this article also qualifies as a distinct music theory, one being that its notion of harmony is temporal, not pitch-oriented (although “regular” European pitch harmony is also mostly retained in these musics) and the other being that the primary temporal element of European music theory, meter, does not carry the same role or meaning in these musics (Arom, 1991; Chor, 2010, 18).

Interpreting RA Models of Note-onset Interactions

sense, or *balanço*⁷, framework is the unambiguous determination of *CD* in every situation (consistently within a given context), with ease (parsimoniously), unambiguously, and without having to resort to special cases.

Background Work on Clave Direction as a Music-Theoretic Grammar

Consequently, the research program of the *balanço* framework started out with about five years of participant observation and literature review. After this period, a double-blinded randomized presentation of a selection of onset vectors (generic rhythm patterns) culled from a corpus of more than 65,000 randomly generated rhythms was carried out with four professional musicians who are (or were in one case, as one is deceased) experts in the idiom. The *balanço* framework was developed based on the combined observation of all these sources, and was used to classify 10,800 onset vectors in three contexts.

These 10,800 randomly generated idealized patterns constitute some of the vertices⁸ of a 16-dimensional input space and a four-dimensional output space (for a total of 20). By limiting the output classes to whether a certain clave direction was present or not, the search space dropped to 17 dimensions. Hence, it was paramount to develop search trade-offs. Interpretation of these models is compared with several criteria for clave-direction determination deduced from systematic observations of four master musicians (and from deep exposure to the underlying musical practice).

The classified vectors were used in a variety of RA model searches. Information, BIC, and AIC for candidate models were used to determine the best models resulting from different searches. This article interprets these models as rhythmic components and compares the insights they supply with the explicit statement of the wide-sense *balanço* framework.

Why this approach and not Markov models or recurrent neural networks? Just as we perceive and interpret words as a whole—we see the word ‘rhythm’ not as the sequential letters *r-h-y-t-h-m*—we do not build it up after seeing *r-h* and postulating ‘rhinoceros’, ‘rhetoric’, or ‘rhapsody’. Likewise, while musical rhythm does unfold over time, the cognition of clave direction is incomplete (and unreliable) until the complete pattern is known.⁹ See the example “*Caixas ‘em cima’*” in (Gonçalves *et al.*, 2000, 51).¹⁰

⁷ This concept (Vurkaç, 2012a and 2012b) is closely related to Stover’s “clave pendularity” (Stover, 2009, 7 and 12).

⁸ The curse of dimensionality (Bellman, 1961) is a concern here. The present author’s current work includes the generation of real-valued vectors by processing raw audio so as to address one of the many problems brought about by the curse of dimensionality.

⁹ Hence, the way *CD* perception may build as the notes are heard in sequence may be of interest in terms of the role of Bayesian model-building in the psychology of rhythm perception, it is assumed that the influence of that process on human *CD* recognition is minimal.

¹⁰ Note, further, that patterns used by “samba schools” change over time, and that the change documented in the variation from the first (cited) to the second edition of this book (Gonçalves *et al.* 2012, 61) for the same instrument as played by the same group illustrates the accuracy of the *balanço* framework.

Interpreting RA Models of Note-onset Interactions

Background Work on RA Modeling

A variety of search approaches and criteria were used in the *OCCAM 3* modeling engine (<http://dmm.sysc.pdx.edu>) to extract information from rhythmic sequences. ('Rhythm' refers to the timing of notes, not necessarily to any steady pulse, repetition, or percussive instrumentation.) The models discovered reflect trade-offs between capturing complex interactions and the ability to generalize.

The Selection of RA Models through Search Strategies

This section is adapted from Vurkaç (2011). For search, evaluation, and fitting, *OCCAM 3* uses criteria like information, α (chi-squared significance for the full change from reference to model), incremental α (from last model to current model), dAIC¹¹ (the difference in "an information criterion"—AIC), dBIC (the difference in the "Bayesian information criterion"—BIC), and percent correct on test.

The task at this stage was to identify the right set of conditions and then to choose among the search criteria. No model is admissible unless its α value suggests that the likelihood of being wrong in claiming this model is different from the reference is less than 5%. As a result, α is used only to check the admissibility of selected models.

The main difference between AIC and BIC is in the choice of penalty. In AIC, the penalty is a function of the number of parameters in the model (Claeskens *et al.*, 2003, 2). In BIC, the penalty is a function of sample size as well as degrees of freedom, with the main emphasis on sample size (Hastie *et al.*, 2001, 233; Wolpert *et al.* 1995, 215).

The model-selection literature indicates, further, that the main problems with AIC are that it under-penalizes as sample size approaches infinity (Hastie *et al.*, 2001, 235; Cavanaugh, 1999, 222) and that it under-penalizes when the sample size down to about triple the number of parameters (Hurvich, 2003). The latter was not a concern because it implied a sample size on the order of 50. None of the data sets used were that small.

While BIC is proven to be asymptotically consistent (chooses the best model if sample size equals infinity), it is known to underfit if the underlying reality generating the data is not of a finite-parameter nature (Claeskens *et al.*, 2003, 2). Due to the premise that greater complexity reduces the probability of achieving a target mapping, and because BIC penalizes more heavily for model complexity, it appeared to be the best choice for the search criterion even though it was known to over-penalize (Cavanaugh, 1999, 224).

A decision between AIC and BIC depends on two factors: dataset size and number of model parameters. Since the latter varies in the course of each RA search, and since the former varies during the course of the overall research, even known asymptotic behaviors could only serve as imprecise guidelines. The decision, then, was to carry out every RA search once with AIC, once with BIC, and once with information, along with an empirical study of RA search criteria and their overfitting behavior.

¹¹ "dAIC" is the change in AIC from one model to the next (in the direction of search selected when setting up *OCCAM 3*). The best model has the minimum xIC, which is the maximum dxIC (Vurkaç, 2011, 34).

Interpreting RA Models of Note-onset Interactions

Three sets of pilot RA runs were performed: loopless, depth-wise (width-one) all-model, and wider all-model searches informed by the depth-wise searches. Plots of percent correct for training and test data were created with respect to various criteria to identify the onset of overfitting, resulting in reducing the set of criteria under consideration to dDF (the change in the number of degrees of freedom from one model to the next), dAIC, dBIC, α , and information. It was then determined that dDF and α are to be used in a confirmatory mode, not for searching. The models discussed here come from the best model in each of the following search types:

- Information (confirmed or tempered by α), carried out according to two search heuristics: *mixed-model bore-rotate* with feature selection (“MBR-i”) and *all-model bore-expand* without feature selection (“ABE-i”), and
- AIC-followed-by-BIC (confirmed or tempered by α), carried out according to the same search heuristics, with AIC for the *mixed-model bore-rotate* searches (“MBR-a”) and BIC for the *all-model bore-expand* searches (“ABE-b”)

What is meant by *bore-rotate* and *bore-expand* are, respectively, moving quickly into the depths of the model tree with initially a very small search width (considering a small number of models, one or two per level) and a large search depth (“boring”) to determine what size models are justifiable and following up such a search with a “rotate” step of moderate width and extreme depth informed by the maximum model size from the “bore” step, and for the latter strategy, following up the same “boring” approach with a very wide search (20 models per level) and moderate depth (for computational-load concerns).

Selected RA Models

The models were found as follows. (The ‘Z’ variable, output, is not shown in these listings, but every model component includes the output class.)

- *ABE-b* model

AB:AC:AJK:AS:BD:EFG:FJ:HJ:JKM:JN:JR:P:QS:RS

- *ABE-i* model

*AB:AC:AG:AK:AM:AR:AS:BD:BN:BQ:BR:CD:EFG:FJ:FQ:FR:GH:GP:GQ:GS:HJ:HP
:HQ:HS:JKR:JN:JQ:KM:MQ:NQ:PQS:RS*

- *MBR-a* model

AC:AJK:CG:FG:KM:Q:R

- *MBR-i* model after RA fitting

ABCQ:ABR:BFJQ:BMQR:CJK:CKR:FG:GR:JKM:JKQ:JKR:JQR:KMQ:KQR

The original context for the model search was to use a number of promising RA models to set the representational bias of artificial neural networks of the MLP type. If the models were useful in this or any other clave-oriented regard, they would also be able to

Interpreting RA Models of Note-onset Interactions

provide insights into the *CD*-discernment problem. The goal of the present article is to look at high-performing RA models for the clue insights that they can provide.

RA Representation and Rhythm Representation

In the *OCCAM 3* representation used, the unambiguous letters *A, B, C, D, E, F, G, H, J, K, M, N, P, Q, R, and S* were used for the onset locations, and *Z* for the binary output class. In other words, the variables corresponds to the idealized (quantized, non-expressive) TUBS subdivisions¹² as follows.

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
A	B	C	D	E	F	G	H	J	K	M	N	P	Q	R	S

Figure 1: Rhythmic subdivisions with shortened (single-character) musician utterances and the selected RA variable labels (with the letters 'I', 'L', and 'O' left out to avoid confusion with the numerals 0 and 1)

Among the various other ways onset sequences are represented in ethnomusicology and music information retrieval is the binary representation. A 16-bit vector can represent any choice of event versus no event in each of the timing spots in Fig. 1 above. For example, simply keeping time with quarter notes can be written as 1000100010001000. In addition to the TUBS and regular music notations, we will have occasion to use the binary representation. The variables *A, E, J, and P*, then, represent the quarter-note onsets. It is critical to point out why the quarter-note onsets are highlighted here.

A Brief Discussion of Meter, Time Signature, and Timeline

In the European court-music tradition, its scholarship, and its notation, the downbeat, the beginning of each bar of music, where a ‘bar’ is determined by the meter (and communicated by the time signature), has a more specific accentual meaning than it might when musics originating in other cultures get notated or interpreted through European conventions of music theory and notation.

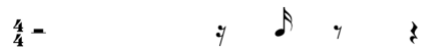
Only *A* among the RA variables is considered a downbeat in strict (European) music theory. In West African and Afro-Latin folkloric musics, meter and time signature have only a convenience function—perhaps mainly for outsiders wishing to notate or otherwise analyze and interpret such music. Because of this, the appropriate choice of time signature in notating Cuban, Brazilian, various West African, and other musics is an ongoing debate (Cruz *et al.*, 2004, 58–66; Vurkaç, 2011, 102 and 122–124). That West African and West African-derived Latin American timelines (many of which act as *CD*-defining patterns) do not have the exact role as meter has been argued even by the leading experts in music-theory and ethnomusicology. Synthesizing the work of such experts as Kubik, London, and Arom, Chor has argued that “In much of the music of sub-Saharan Africa, temporal organization is not necessarily governed by meter. Instead, it uses time-line patterns, which differ from meter in several important ways. ... [T]he notes are not organized into strong and weak [by timelines as they are by a meter]. [On the other hand],

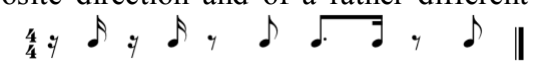

¹² The “time-unit box system” (TUBS) was developed in 1962 by Philip Harland, Robert Bonsu, and Robert Ayitee, as reported in (Koetting, 1970).

Interpreting RA Models of Note-onset Interactions

the time signature of a piece of metric music says relatively little about the music's character.” (Chor, 2010, 17–18, incorporating Arom, 1991)

We see that timelines (such as clave patterns and other clave-defining patterns) operate differently to inform musicians than meters do. Although, as pointed out by Chor, timelines, and especially *CD*-defining patterns do not organize notes into strong and weak bar by bar, they do provide a large scale guideline to every musician in an ensemble about the probabilities (expected frequencies of occurrence) of note events (onsets) at a variety of points in every phrase.

To see why this is challenging enough to necessitate analysis through modeling and machine learning, consider the following two examples. The single-note pattern,  || (0000000001000000) is sufficient to send the message that we are in the 2:3 direction (assuming that tempo and phrase references were already being provided by clave-neutral¹³ parts). The placement of that single note (preferably in a prominent way) tells every clave-aware musician in any ensemble that the direction is 2:3, and as a result, that dozens of familiar rhythms (and thousands of improvisational options) are discouraged.

Likewise, but selecting an example of the opposite direction and of a rather different character, the 3:2 *sambadouro* pattern,  || (0101001010010010), has a great deal (its entire second half) in common with the 2:3 (yes, opposite) *clave son*:  || (0010100010010010).

What we are finding is that the clave concept is multivalent and not straightforward. Template-matching is out of the question. The rhythmic (accentual) organizational function of *A*, *E*, *J*, and *P* is unavailable in this context. *B*, *F*, *K*, and *Q*, the strong offbeats immediately following each tactus (Vurkaç, 2012a, 42, Fig. 3) are expected to be the most informative onsets—an insight that the RA analysis below debunks to muddy the waters further. Nonetheless, before running any model searches, it was important to record what we expected the RA models to find. Based on musical experience, we believed we knew about the roles and usefulness of different onset groupings, and wondered if RA could—and if so, whether RA would—find these. In addition to the meter-and-downbeat-motivated preceding discussion, we expected *A*, *B*, *C*, *D*, *J*, *K*, *M*, and *N* (the first and third four-note schemata) to be more predictive than *E*, *F*, *G*, *H*, *P*, *Q*, *R*, and *S* (the second and fourth four-note schemata), and in addition, for *B*, *F*, *K*, and *Q* to be more predictive than *D*, *H*, *N*, and *S*.

Expectations for RA Models based on a priori Conjectures

In the *ABE-b* model, there are indeed six components exclusively from the first and third four-note schemata (*AB:AC:AJK:BD:JKM:JN*), only four components exclusively from the second and fourth four-note schemata (*EFG:P:QS:RS*), and four that contain variables from both (*AS:FJ:HJ:JR*).

¹³ For the concept of clave-neutral patterns, see (Mauleón-Santana, 1999, 10) and (Cruz, 2004, 80), and (Vurkaç, 2012a, 51–52).

Interpreting RA Models of Note-onset Interactions

In the *ABE-i* model, there is an almost even distribution of components when examined from the same point of view: nine exclusively from the first and third four-note schemata (*AB:AC:AK:AM:BD:BN:CD:JN:KM*), an unexpected twelve exclusively from the second and fourth four-note schemata (*EFG:FQ:FR:GH:GP:GQ:GS:HP:HQ:HS:PQS:RS*), and eleven that contain variables from both (*AG:AR:AS:BQ:BR:FJ:HJ:JKR:JQ:MQ:NQ*).

In the *MBR-a* model, three, three, and one also fails to meet our *a priori* expectation. In the *MBR-i* model, there are two components exclusively from the first and third schemata (*CJK:JKM*), two more that are exclusively from the second and fourth schemata (*FG:GR*), and ten across: *ABCQ:ABR:BFJQ:BMQR:CKR:JKQ:JKR:JQR:KMQ:KQR*.

The Interpretation of RA Models through the Balanço Framework

Having eliminated some insight-based possibilities, we start the model analysis in earnest with the *ABE-b* model, listing its components as rhythm cells. Rather than come up with an *ad hoc* justification for each component, we match them up (as well as can be done) to the *CD*-discernment criteria in the *balanço* framework: 1) *partido-alto*; 2) isolated missing downbeats on 1 or 3 (IMD); 3) gross (large-scale) relativeness; 4) *terceira Mocidade* (a partial indication of *partido-alto*); 5) fine relativeness (relative strengths of schemata and algebraic cancellation of onsets across schema bridges, first to third, and second to fourth); 6) template-matching to *teleco-teco* (assisted by the heuristics “zipper” for 3:2 and “bookends” for 2:3); 7) the “African” clave sense (antecedent/consequent relationship); 8) the partial (“hanging”) IMD; 9) *how it feels*; and 10) direct template-matching to the so-called “bossa clave” (Vurkaç, 2012a, 63).

AB: criteria 3 and 6

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
A	B														



AC: criteria 1 and 3

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
A		C													



AJK: criteria 2 and 5

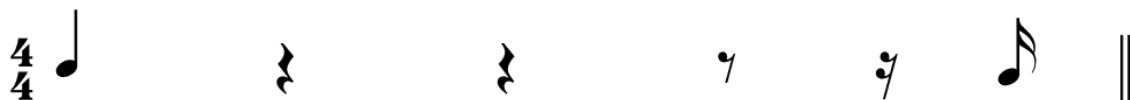
1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
A								J	K						

Interpreting RA Models of Note-onset Interactions



AS: criterion 6

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
A															S



BD: criteria 1 and 8

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
	B		D												



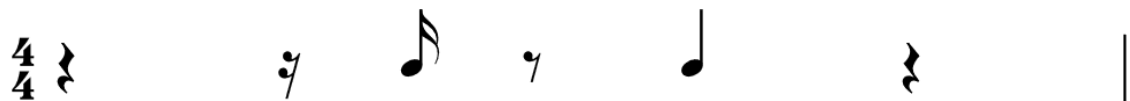
EFG: criterion 3

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
				E	F	G									



FJ: criterion 3

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
				F				J							



HJ: criterion 6

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
							H	J							

Interpreting RA Models of Note-onset Interactions



JKM: criterion 3

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
								J	K	M					



JN: criteria 3 and 10

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
								J			N				



JR: possibly criterion 5, or none

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
								J						R	



P: possibly criterion 1, although too little information in this case

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
												P			



QS: criteria 1 and 8

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
												Q		S	

Interpreting RA Models of Note-onset Interactions



RS: criterion 8

1	e	&	a	2	e	&	a	3	e	&	a	4	e	&	a
														R	S



This model (*AB:AC:AJK:AS:BD:EFG:FJ:HJ:JKM:JN:JR:P:QS:RS*) suggests that the first and second sixteenth-note onsets (*AB*), the first and third sixteenth-note onsets (*AC*), the first and third downbeats (with the 16th off the third downbeat: *AJK*), the first and last onsets of the sequence (*AS*), the first two offbeats of the first tactus (*BD*, a highly typical schema in clave-based musics), the first three onsets of the second tactus (a very unexpected combination), the third downbeat and the preceding offbeat (a reference to criterion 2), and so on, are relevant to determining the clave direction of a pattern, and are more relevant than the interactions not addressed by this model.

We expect that such a model would find many relevant relationships in onset sequences for determining *CD* but would also miss several. The interactions not captured by this model but captured in another may turn out to be crucial to determining the direction of many sequences. Consider that the *ABE-i* model has the component *AK* instead of the *ABE-b* model's *AJK*, a more economical term that may capture sufficient musical intuition. The prominence of two-onset interactions in the top model (*ABE-b*), further, supports the insight that clave direction can often be inferred through an analogue to algebraic elimination: criterion 5 (Vurkaç, 2011, 125–127; Vurkaç, 2012a, 63).

Considering this model component by component, we find *AB*, *AC*, *AS*, *BD*, *P*, *QS*, and *RS* as the components with only events on the outside, and *EFG*, *FJ*, *HJ*, *JKM*, and *JN* as the components with only events on the inside. *AJK* and *JR* are components with events on both sides. Out of these components, the one that stands out the most is *BD*. This is the quintessential Brazilian schema, the rhythmic cell that is the signature of almost all Brazilian music from *maracatu de baque virado* to *axé*, from *MPB* to *bossa nova*. Its presence is a very strong indicator of *CD*, although neither a necessary nor a sufficient condition. This is because it's not solely the presence of the 0101 schema but where in a phrase it shows up that matters. Repeating the same schema two beats (180°) away could cancel out the effect of this occurrence (*would* do so in the absence of other information). We point to the band *monobLoco*'s use of a similar trick to appear to indicate *CD* with a schema almost as strong (0100), repeat that schema 180° later, and leave the *CD*-determining role to a weaker pair of schemata (1001 and 1000) which establish *CD* through relative offbeatness and hanging IMD (Vurkaç, 2012a, 39, 44–45, 50–52, 61–62): 0100100101001000. Hence we see that for the *BD* cell to be fully informative, some other model components must show up: either the same schema 180° later (*KN*, which

Interpreting RA Models of Note-onset Interactions

does not show up in this model) or components addressing the interaction of the variables *A*, *B*, *C*, or *D* to the variables *J*, *K*, *M*, or *N*. We see this in the components *AJK*, *JKM*, and *JN*. Nonetheless, we can ask why RA selected these components in a different manner than its sharp focus on the characteristic schema *BD*. This is a recurring question that comes out of the present analysis. One possible answer is that given the number of variables, RA searches simply have not had the opportunity to examine the later sections of the patterns (the letters further down the alphabet) due to already reaching risking overfitting. This could be investigated by forcing searches to start at a variety of predetermined candidate models from different regions of the search space.

A very different set of model components are *AB*, *AC*, *AS*, *QS*, and *RS* in one half of the phrase and *EFG*, *FJ*, *HJ*, and *JN* in the other half. These components appear to focus on interactions within a schema (as in component *BD*) and together address the potential for each of the two clave directions. For example, *AC*, if both variables were 1, would suggest a strong likelihood of 2:3 direction for the entire pattern, *AB* a moderate likelihood of 3:2 for the same pattern, and *AS* a small likelihood of 3:2. On the other hand, and here is the strength of the modeling process and the usefulness of a large randomly generated data set, a mixture of 1 and 0 for in each of these components would provide a great deal of additional information. For example if *AB* were 01 and *AC* were 00, together they would suggest a similar strong indication toward 3:2 just as *BD* being 11 would have. In the other half of the phrase, the component *EFG* can achieve the same effect, indicating a strong likelihood for 2:3 when it takes on the values 101 or 001, a strong likelihood for 3:2 when it takes on 010, a weak indication of 3:2 with 011 and with 110, and very little information about *CD* with the values 000, 100, and 111.

On the other hand, note the component *JKM*, which is *EFG* 90° later and in the same half of the phrase. Taking these components together we can get clues about whether no *CD* exists in a given pattern, which happens either through clave-neutrality¹⁴ or the complete absence of any clave sense, or clave-incoherence (Vurkaç, 2012b, 332).

If *EFG* concatenated with *JKM* were to be 001001, we would get a rather strong indication of 2:3 direction; with 010010, a slightly stronger indication of 3:2 direction; with 011011, a moderate indication of 2:3 direction; with 100100 a moderate indication of 3:2 direction; with 101101, a very strong indication of 3:2 direction; and with 110110, a somewhat strong indication of 2:3 direction. In all the foregoing examples, *EFG* = *JKM*. If *EFG* were not to take on the same values as *JKM*, we may find evidence of clave-incoherence: 101010 is a strong indicator that the complete pattern could turn out to be clave-incoherent. We can see this more clearly when we add the missing *H* and *N* variables. If *H* and *N* were both 0, the two consecutive schemata result in 10100100 on the inside of the pattern, setting the stage for clave incoherence unless equally strong or

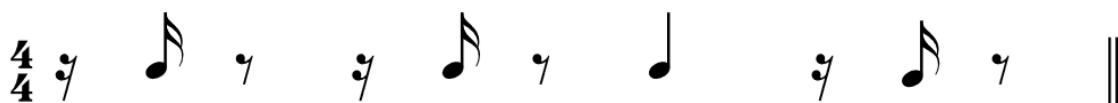
¹⁴ This is very important for making music danceable by avoiding the “loss of total effect” problem that shifts the perceived starting points of phrases; i.e., reinterpreting timelines as if they were metric, as discussed by Merriam who put forth an insightful conjecture for the origin of the clave sense: “The continual use of off-beating without respite would cause a readjustment on the part of the listener, resulting in a loss of the total effect.” Clave, then, “is a device whereby the listener’s orientation to a basic rhythmic pulse is threatened but never quite destroyed” (Merriam, 1959, 16).

Interpreting RA Models of Note-onset Interactions

even stronger schemata were to reinforce each other on the outside of the patter (such as 1010101001001010 or 0101101001000101 or 0101101001000111). And finally, the components *EFG* and *JKM* can also serve to contribute to establishing the clave-neutrality of the overall pattern if they both took on values such as 100 and 111.

We have discussed the potential contributions of the components of one RA model. Let us next consider a model with very different components. The *MBR-i* model has the components *ABCQ*, *ABR*, *BFJQ*, *BMQR*, *CJK*, *CKR*, *FG*, *GR*, *JKM*, *JKQ*, *JKR*, *JQR*, *KMQ*, and *KQR*, a more equitable distribution of schemata from around the phrase. (Recall that the previous model had only two components from both sides of the phrase.) Let's view a subset of these, the ones that reach across the two halves, visually.

BFJQ: criteria 5 and 8



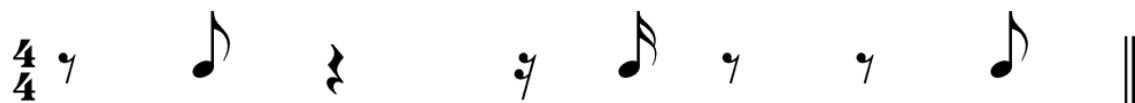
BMQR: criteria 1 and 2



CJK: criteria 1 and 6



CKR: criteria 1 and 8



GR: none without more information (neutral)



JKQ: criterion 5

Interpreting RA Models of Note-onset Interactions



KMQ: criteria 5 and 8



KQR: criteria 5 and 8



For these model components, let's first look at the ones that isolate the onsets with the same function (those that are 90° or 180° apart): *GR* does this, and nothing else; all but one of the remaining components include this feature along with some other variables: *BFJQ*, *BMQR*, *CJK*, *CKR*, *KMQ*, *KQR*.

In *GR*, these two onsets form a powerful indicator because they could serve to determine membership in any of the four *CD* classes. If both *G* and *R* were 1 (and their schema neighbors were 0), strong opposing schemata at *A–D* and *J–M* would lead to a very strong membership in either 2:3 (for instance, 1010001001010010) or 3:2 (for instance, 0101001010100010). If *A–D* were an exact copy of *J–M*, (say, 0111001001110010), clave-incoherence would be easily established in the strict context and clave-neutrality would be easily established in the loose and firm contexts (Vurkaç, 2011, 20–21).

In a similar way, but with more information, *BFJQ* and *BMQR* combine onsets with the same function in the same and in opposite sides of the phrase. For example, if *BFJQ* were 1111, with all 0s in the rest of a vector, the 3:2 *CD* would be apparent. However, if the remaining variables took on a mix of values, although all four output classes would be possible, the determination of any of the four should be made easier by knowing *BFJQ*. (For example, using the schemata set of *xx11*, we would get *x111x1111x11x111*, the sixteen remaining options of which can be 2:3 if we had 1011 for the *xs* (1111011111111111), 3:2 if we had 0100 for the *xs* (0111111110110111), and neutral if we had 1111 for the *xs* (1111111111111111—and likewise if our schema had been *xx00*: 1100110011001100). The incoherent option would be possible with a different selection of the unspecified onsets: 0100010110100110. Similar examples can be anticipated for the other components, suggesting that the *MBR-i* model should be very successful at discerning *CD*. What we have found, instead, is a discrepancy between this *musician's intuition* and conclusions from model performances by RA criteria, the latter of which suggested strongly that *ABE-b* was the best model and that both the *MBR* models should be discarded. Perhaps, the *MBR-i* model with the bigger components reaching across the two halves attempts to encapsulate too many output classes in each of its selections of variables. Perhaps it overfits thus.

Interpreting RA Models of Note-onset Interactions

Summary of Interesting Findings

Not only are the RA models able to provide unexpected insights into the recognition of CD (in the selected context) in a variety of patterns (and to some extent, debunk expected insights), but the review of the pilot study also revealed that we should be able to obtain much better models than the models found so far.

The four RA models were able to learn the classification task with percent correct on test ranging 60–69%. These levels suggest a great deal of overfitting, as the same for models were able to learn the training set to 85–94% accuracy. However, the percent correct on test for MLPs prestructured by each of these RA models ranged 73–92% (with training accuracies at 84–92%, respectively, suggesting less overfitting, although still present. Further MLP experiments suggest BIC-model-prestructured MLPs’ test accuracies were capped at 91% (with analysis incomplete as of this writing), which underperform on test in comparison to fully connected MLPs whose performances appear to be in the tighter range of 92–93%. We interpret both these findings to mean that the selection of RA models—the traversal of the model space—has to be conducted better to achieve better insights into the musical problem and to continue developing an RA-based technique for biasing MLPs.

Conclusions and Future Work

The models considered in this analysis, then, suggest the unexpected trade-off that interesting musical insights may be gained by considering interactions of fewer onsets at one time than would be expected from domain knowledge. It *is*, however, absolutely essential for a human classifier of CD to consider the entire pattern, all 16 onsets. How can these findings be reconciled? We must notice that although the RA models consider the interactions of a few variables at a time, they also take into account numerous such interactions. For humans, keeping track of many numerical relationships is challenging, but perceiving the *Gestalt* becomes fast and reliable with experience. Such experience, however, if not acquired from birth through the early years, is elusive or at least very expensive in terms of one’s investment of time. Computational means like RA can come to the aid of struggling musicians in unexpected ways (as suggested by the present interpretation of some of the RA models), allowing the technologist to build educational tools to help aspiring musicians develop CD discernment.

We have demonstrated the potential of RA to model an intricate and culturally specific (not broadly accessible) musical construct in terms of discrete note events and their interactions in such a way as to sometimes mirror and sometimes challenge—thus, aid—human understanding of the corresponding cultural practice.

Directions of future work identified include forcing searches to start at a variety of predetermined candidate models from different regions of the search space so as to attempt to avoid the bias for more thorough model searches focused on certain convenient regions of the search space, the investigation of more (and more orthodox?) search strategies, potentially the development of hybrid search strategies, and the use of more powerful computing resources.

Interpreting RA Models of Note-onset Interactions

In addition, the use of RA for biasing neural networks, the use of evolutionary and immune-system algorithms and cellular automata to generate clave-concordant families of musical patterns (as a component of algorithmic composition), and the investigation of game theory as a means to incorporate simulations of the social evolution of clave-based musics into the framework of understanding these cultural practices are also identified as directions for future work.

REFERENCES

- Agawu, V. K. (2003). *Representing African Music – Postcolonial Notes, Queries, Positions.*, Routledge, New York.
- Arom, S. (1991). *African polyphony and polyrhythm: Musical structure and methodology.*, Cambridge University Press, New York.
- Bellman, R. (1961). *Adaptive Control Processes: A Guided Tour.*, Princeton University Press, Princeton, NJ.
- Cavanaugh, J. E. (1999). A Large-Sample Model Selection Criterion Based on Kullback's Symmetric Divergence, *Statistics & Probability Letters.*, 44: 333–344.
- Cazden, N. (1945). Musical Consonance and Dissonance: A Cultural Criterion, *The Journal of Aesthetics and Art Criticism.*, 4(1):3–11.
- Chor, I. (2010). *Cognitive Frameworks for the Production of Musical Rhythm.*, UMI/Northwestern University, Evanston, IL.
- Claeskens, G. and Hjort, N. L. (2003). The Focussed Information Criterion, *Journal of the American Statistical Association*, 98: 900–945.
- Clarke, E. and Cook, N. (2004). *Empirical Musicology: Aims, Methods, Prospects*, Oxford University Press, Oxford.
- Cruz, T., Moore, K., Gerald, M., and Fiol, O. (2004). *The Tomás Cruz Conga Method, Volume I—The Beginning: Conga Technique as Taught in Cuba.*, Mel Bay Publications, Pacific, MO.
- Garibaldi, D., Diaz, J., and Spiro, M. (1999). *TimbaFunk by Talking Drums.*, Warner Bros. Publications, Miami.
- Gonçalves, G. and Costa, M. O. (2000). *Aprendendo A Tocar O Batuque Carioca: As Baterias Das Escolas De Samba Do Rio De Janeiro.*, Groove, Rio De Janeiro.
- Gonçalves, G. and Costa, M. O. (2012). *Aprendendo A Tocar O Batuque Carioca: As Baterias Das Escolas De Samba Do Rio De Janeiro, 2ª Edição, Revisada E Ampliada.*, Groove, Rio De Janeiro.
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.*, Springer-Verlag, New York.
- Herder, R. (1972). *Spanish Montunos – Ensemble Studies in Rhythmic Precision and Pitch Security*, Continuo Music Press, Inc., New York.
- Hurvich, C. (2003). *The Corrected AIC (AIC_C).*, NYU Stern, New York.
- Jourdain, R. (2002) *Music, the Brain, and Ecstasy: How Music Captures Our Imagination.*, Quill, Harper Perennial, New York.
- Koetting, J. (1970). Analysis and Notation of West African Drum Ensemble Music. *Selected Reports*, 3: 116–146.
- Kwabena Nketia, J. H. (1963). *African Music in Ghana.*, Northwestern University Press, Evanston, IL.

Interpreting RA Models of Note-onset Interactions

- Lehmann, B. (2002). *The Syntax of 'Clave' – Perception and Analysis of Meter in Cuban and African Music.*, Tufts University, Medford, MA.
- Locke, D. (1996). Africa: Ewe, Mande, Dagbamba, Shona, BaAka” in *Worlds of Music: An Introduction to the Music of the World's Peoples*, (J. T. Tilton, ed.), Prentice Hall, London.
- Machado, H. F., Muñoz, W., and Sadi, J. (2002). *El Toque de Candombe: Ritmo Tradicional Afro-Uruguayo / Traditional Afro-Uruguayan Rhythm.*, Mel Bay Publications, Pacific, MO.
- Malabe, F. and Wiener, B. (1990). *Afro-Cuban Rhythms for Drumset.*, Alfred publishing, Van Nuys, CA.
- Martineau, J. (2008). *(the world's most compact guide to) The Elements of Music: Melody, Rhythm & Harmony (an introduction to the universal language).*, Bloomsbury, New York.
- Merriam, A. P. (1959). Characteristics of African Music, *Journal of the International Folk Music Council*, 11: 13–19.
- Parncutt, R. (2008). Template-matching Models of Musical Pitch and Rhythm Perception, *Journal of New Music Research.*, 23(2): 145–167.
- Peñalosa, D. (2009). *The Clave Matrix: Afro-Cuban Rhythm: Its Principles and Origins.*, Bembe Books, Redway, CA.
- Rodriguez, A. (2003). *Traditional Afro-Cuban Concepts in Contemporary Music.*, Mel Bay Publications, Pacific, MO.
- Soebbing, H. W. (1988). Latin American Instruments and Rhythms, in *Percussion Anthology – A Compendium of Percussion Articles from The Instrumentalist Fourth Ed., including articles from 1946 through 1987*, The Instrumentalist Co., Northfield, IL.
- Stone, R. M. (1985). In Search of Time in African Music, *Music Theory Spectrum.*, 7: 139–148.
- Spiro, M. and Ryan, J. (2006) *The Conga Drummer's Guidebook.*, Sher Music Co., Petaluma, CA.
- Stover, C. (2009). *A Theory of Flexible Rhythmic Spaces for Diasporic African Music.*, University of Washington, Seattle, WA.
- Sublette, N. (2004). *Cuba and Its Music: From the First Drums to the Mambo.*, Chicago Review Press, Chicago, IL.
- Tobias, D. A. (1988). Latin-American Percussion, in *Percussion Anthology – A Compendium of Percussion Articles from The Instrumentalist Fourth Ed., including articles from 1946 through 1987*, The Instrumentalist Co., Northfield, IL.
- Vurkaç, M. (2011). *Prestructuring Multilayer Perceptrons based on Information-Theoretic Modeling of a Partido-Alto-based Grammar for Afro-Brazilian Music: Enhanced Generalization and Principles of Parsimony, including an Investigation of Statistical Paradigms.*, UMI/ProQuest, Ann Arbor, MI.
- Vurkaç, M. (2012a). A Cross-cultural Grammar for Temporal Harmony in Afro-Latin Musics: Clave, Partido-Alto, and Other Timelines, *Current Musicology.*, (94): 37–65.
- Vurkaç, M. (2012b). Balanço: The Contour of Relative Offbeatness, in *Bridges Towson: Mathematics, Music, Art, Architecture, Culture, Proceedings 2012* (R. Bosch, D. McKenna, and R. Sarhangi, eds.), Tesselations Publishing, Phoenix, AZ.

Interpreting RA Models of Note-onset Interactions

Washburne, C. (1997), The Clave of Jazz: A Caribbean Contribution to the Rhythmic Foundation of an African-American Music, *Black Music Research Journal.*, 17(1): 59–80.

Wilson, D. E. (2010). *Partido-Alto: Rhythmic Foundation Analysis of Aquarela Do Brasil.*, Youngstown State University, Youngstown, OH.

Wolpert, D. H. and Macready, W. G. (1995). *No Free Lunch Theorems for Search.*, Santa Fe Institute, Santa Fe, NM.