

NOMINAL CROSS RECURRENCE AS A GENERALIZED LAG SEQUENTIAL ANALYSIS FOR BEHAVIORAL STREAMS

RICK DALE

Department of Psychology
The University of Memphis, Memphis, TN 38152, USA
radale@memphis.edu

ANNE S. WARLAUMONT

School of Audiology and Speech-Language Pathology
The University of Memphis, Memphis, TN 38105, USA
anne.warlaumont@memphis.edu

DANIEL C. RICHARDSON

Cognitive, Perceptual, & Brain Sciences
University College London, London, WC1E 6BT, UK
daniel@eyethink.org

Received February 22, 2010; Revised May 4, 2010

We briefly present lag sequential analysis for behavioral streams, a commonly used method in psychology for quantifying the relationships between two nominal time series. Cross recurrence quantification analysis (CRQA) is shown as an extension of this technique, and we exemplify this nominal application of CRQA to eye-movement data in human interaction. In addition, we demonstrate nominal CRQA in a simple coupled logistic map simulation used in previous communication research, permitting the investigation of properties of nonlinear systems such as bifurcation and onset to chaos, even in the streams obtained by coarse-graining a coupled nonlinear model. We end with a summary of the importance of CRQA for exploring the relationship between two behavioral streams, and review a recent theoretical trend in the cognitive sciences that would be usefully informed by this and similar nonlinear methods. We hope this work encourages scientists interested in general properties of complex, nonlinear dynamical systems to apply emerging methods to coarse-grained, nominal units of measure, as there is an immediate need for their application in the psychological domain.

Keywords: recurrence; cross recurrence; psychology; coupling; language

1. Introduction

Psychologists often collect time series in the form of “behavioral streams”: measures on a nominal scale that reflect events in a person’s behavior or experiences during the day or a laboratory task. For example, a psychologist may simply register any point at which a person is smiling or frowning (or, neither). Doing this throughout the day will produce a behavioral stream of nominal “states” that a person was in: *smile*, *frown*, or *neither*. There are several ways to carry out this coding [Bakeman *et al.*, 2005]. Commonly, this coding activity produces a stream that is a time series, as commonly construed: a sequence of states,

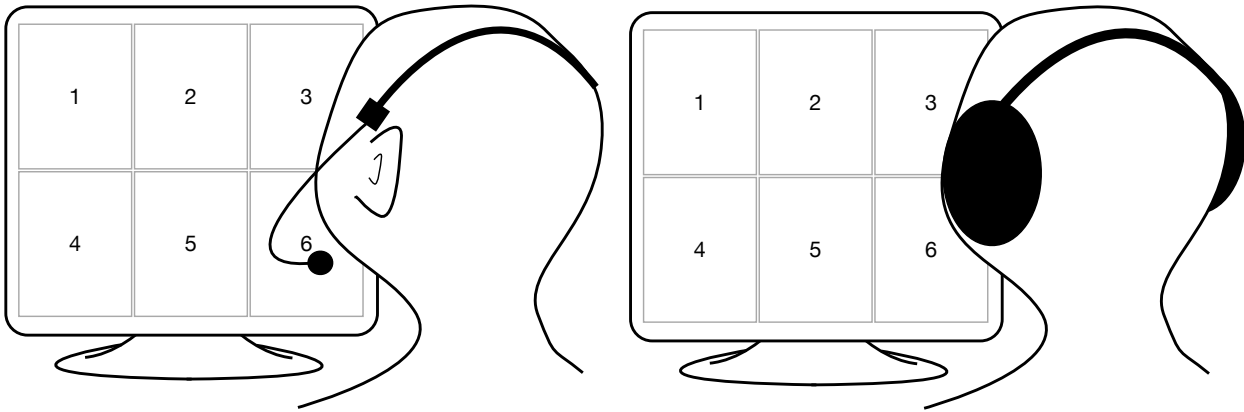


Fig. 1. The experimental context for collecting nominal time series. The left person is speaking about a 2×3 grid displaying television characters while the right person listens in. The eye movements generate coarse-grained nominal series of numbers 1-6 for approximately 60 seconds while the speaker narrates.

obtained by regular sampling, along some dimension representing what a person is up to. This data-collection context is pervasive in psychology, and there are many qualitative and discrete methods available for it [Agresti, 2002; Strauss, 1987]. Among the longest standing and most prominent frameworks for extracting dependencies and patterns in these behavioral streams is called lag sequential analysis [Sackett, 1979; Bakeman & Gottman, 1997; Bakeman & Quera, 1995a]. Lag sequential analysis is a wide-ranging application of the analysis of contingency tables, such as through log-linear modeling [Bakeman & Quera, 1995b].

This brief paper describes this method and compares it to cross recurrence quantification analysis (CRQA). Specifically, we show that when quantifying the relationship between behavioral streams of two people interacting in some way (e.g., in conversation), lag sequential analysis can be seen as proportional to measures obtained from some version of CRQA. Extending this lag sequential method by integrating it with CRQA has the benefit of further connecting psychological explorations to a growing understanding of complex systems and their behavior. Already it is being pursued in the analysis of behavioral time series on continuous scales, such as postural coordination between two people talking [Shockley *et al.*, 2003]. CRQA, a method that may be described as a form of generalized cross-correlation [Marwan *et al.*, 2007], naturally extends to nominal behavioral streams by using numeric codes and a radius of zero to reflect state matches [Orsucci *et al.*, 1999]. Indeed, related methods have been developed in natural language processing with dot plots [Ducasse *et al.*, 1999] and sequencing methods in molecular biology [Von Heijne, 1987]. Below, we first summarize our sample human data used here to demonstrate lag sequential analysis and CRQA. We then compare these methods, and show that CRQA loosely encompasses sequential analysis, but can also reveal nonlinear patterns in coupled systems.

2. Example Data Streams

In the past two decades, eye movements have become a common source of behavioral data in psychology and are often used to produce nominal time series [Spivey *et al.*, 2009]. Consider Fig. 1. The numbered panels compose a shared visual space that participants can discuss. During a *separated* interaction (e.g., by talking on the phone), participants discuss the panels (presented on two computer monitors), and their eyes are tracked while they do this. At 33ms intervals, a nominal measure can be extracted that simply represents the panel number that is being fixated at that moment (see Fig. 2). We have used this context to carry out several studies employing simple recurrence measures from CRQA [Richardson *et al.*, 2007; Richardson & Dale, 2005; Richardson *et al.*, 2009]. Here we use a set of subject pairs ($N = 6$) from an earlier experiment for demonstration [Richardson & Dale, 2005]. In this work, one participant listens to the language of another participant while they both look upon a panel of characters from two prominent television shows in the US (*Friends* and *The Simpsons*). A 2×3 visual array was used containing pictures

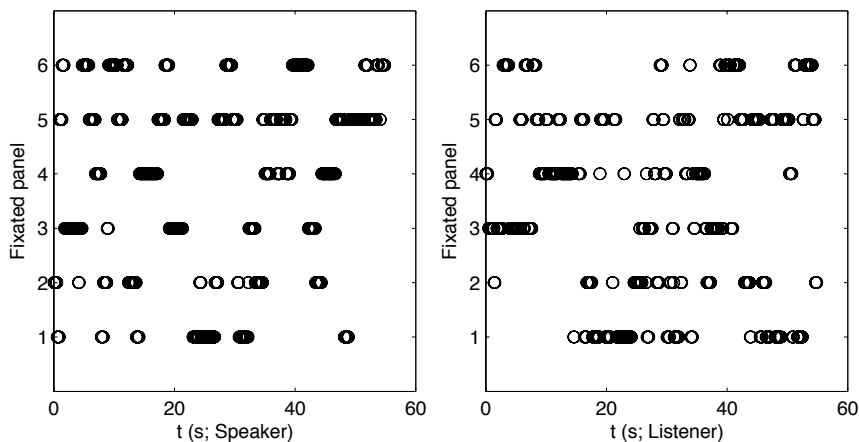


Fig. 2. The two nominal time series from an example dyad among the 49 from the original study. The language of the speaker proceeds for approximately one minute while the other participant listens in. Each 33ms interval produces a numeric value representing the panel that was fixated. The occasional absence of a numeric value at some time points means that participants were not looking at any of the panels.

of the characters. Because only one member of a pair is doing the talking, the second participant (the listener) is coupled to the speaker and his or her behavioral stream (i.e., nominal eye-movement time series) exhibits a distinct lag relationship. We show this below.

3. Sequential Analysis of the Stream

Lag sequential analysis is a suite of methods related to contingency table analysis, including a variety of extensions that have become part of a standardized set of analytic strategies and techniques [Bakeman & Quera, 1995a]. To begin in the case of our example data, one would produce a contingency table (CT) of relative eye positions at a chosen lag. A CT represents the relationship between states in one time series (e.g., the speaker) and states in another time series (e.g., the listener) at a given lag, τ . Rows of the CT may reflect speaker states and are ordered consistently with the columns of the CT , representing the listener's states. Numbers in the cells represent times each state in a row's relevant time series (the speaker) was followed with lag τ by the corresponding column's state in the listener. This can be expressed simply as:

$$CT_{i,j}(\tau) = \sum_{t=1}^{t=T-\tau} q(t) \quad (1)$$

$$q(t) = \begin{cases} 1 & \text{if } x(t) = i \text{ and } y(t + \tau) = j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Where $CT(\tau)$ is the contingency table at lag τ , $CT_{i,j}(\tau)$ is the cell entry for the i^{th} and j^{th} states between nominal time series x and y , and T is total time. The function $q(t)$ is simply a membership-sum function for the CT which specifies if at time t for its relative lag whether $x(t)$ and $y(t + \tau)$ have the same specified states i and j ($= 1$) or not ($= 0$). (Note, t and τ are in sample units, where $T = N33ms$, and N is the number of samples.)

In many cases, one may simply be interested in whether states are being matched between the two series. In other words, this analysis would focus on whether the speaker and listener are “doing the same thing” at lag τ . In this case, one simply attends to the entries along the diagonal ($i = j$) of the CT , since rows and columns are ordered consistently for both time series. When a CT is of size 2×2 the phi coefficient gives the correlation between the two binary variables:

$$\phi = \frac{CT_{1,1}CT_{2,2} - CT_{1,2}CT_{2,1}}{\sqrt{\prod CT_{\Sigma}}} \quad (3)$$

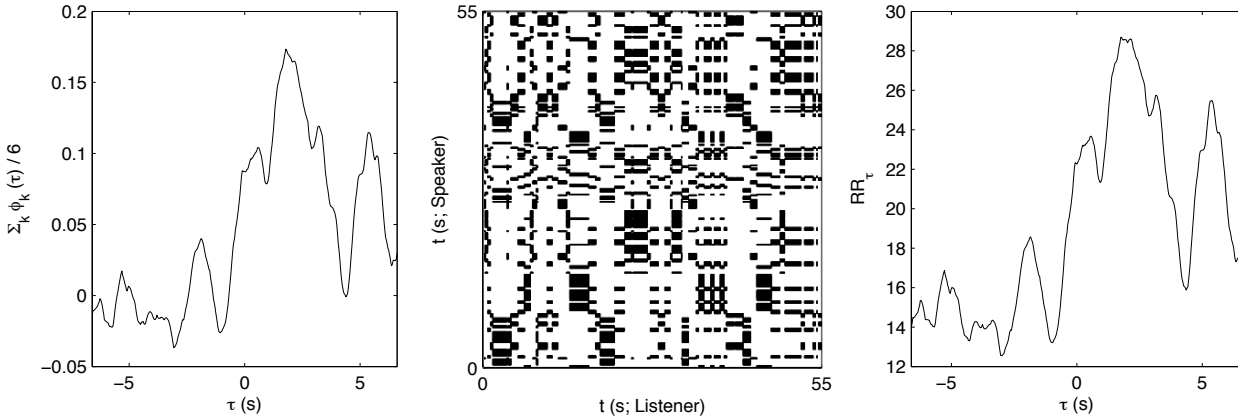


Fig. 3. The left-most plot is a mean phi coefficient between matching states, at given lags, across the 6 categories, for the time series shown in Fig. 2; The middle plot shows a cross recurrence plot (CRP) between the nominal two time series shown in Fig. 2 with $m = 1$, $\tau = 1$ sample (33ms), and $\epsilon = 0$; the right-most plot shows the diagonal-wise recurrence rate, RR_τ , of that CRP. Note the similarity between the mean phi coefficient on the left and the diagonal-wise recurrence rate to the right.

CT_Σ represents the set of four marginal sums (row and column) of the CT . When the CT is larger than 2×2 , as it is in this case, one can reduce it to the test of interest in the following simple way. First, let k be one of the states in the series. Create a 2×2 CT where states are recoded into k and not- k . Here, let $\phi_k(\tau)$ represent the phi coefficient for this given category k at lag τ .

We can repeat this procedure for each possible category, $k = 1, \dots, 6$. The mean phi coefficient will now give the average binary correlation between the time series across the 6 categories across all lags. This is shown in Fig. 3, left. In this sense, lag sequential analysis is carrying out an average binary cross-correlation between the two time series. Fig. 3 shows that there is a highest correlation at a lag of approximately $\tau = 2s$, showing that the optimal covariation is occurring at approximately a two-second delay for this particular listener to couple his or her eye movements to the speaker.

Often it is desirable to measure states across the two time series at lag τ that are not necessarily the same, matching states. In terms of the CT , this is equivalent to inspecting the off diagonals, where $i \neq j$. For example, though we do not carry out this analysis here, it may be the case that fixating one panel (i.e., television character) may accompany looks to a different panel (i.e., another relevant character). These contexts arise regularly in lag sequential analysis. The procedure just described for using mean $\phi_k(\tau)$ to characterize exact matching is adapted easily. Instead of recoding into k and not- k , we recode columns and rows differently depending on the test of interest. This test may consist of finding lags at which one time series is in state $k = i$ while the other time series is in state $k = j$. This reduces to constructing a CT with rows reflecting i and not- i states, and columns j and not- j . The cell entry $CT_{i,j}(\tau)$ in this 2×2 table will now represent the off-diagonal sum in the original, full table. The phi coefficient can now easily be calculated in these cases and averaged with any other off-diagonal matches of interest.

It is important to note that there are many variants to these kinds of analyses, and we show only the most rudimentary. In addition, many issues of coding, and the testing and reporting of statistical significance are addressed in lag sequential analysis that we do not consider here for lack of space [Bakeman & Gottman, 1997]. However, the general observations we have made reflect the kernel analyses characteristic of lag sequential analysis.

4. CRQA of the Stream

CRQA, as described in detail in various places [Marwan *et al.*, 2007; Marwan & Kurths, 2002; Zbilut *et al.*, 1998], is easily applied to these example eye-movement time series by choosing any dimension (m) and time delay (τ) parameter values, and simply setting the radius, ϵ , equal to zero so that only exact matches will be considered instances of recurrence (Orsucci *et al.*, 1999; unless otherwise noted, all notation is taken from Marwan *et al.*, 2007). We have found in past research that in our nominal time series an embedding

window size of $m = 1$ (i.e., no embedding) is sufficient to obtain basic experimental effects [Dale & Spivey, 2006]. Unless otherwise noted, we use this parameter setting below.

A cross recurrence plot (CRP) between the example dyad’s time series presented in Fig. 2 is shown in Fig. 3. Present in the plot are different “textures” [Eckmann *et al.*, 1987] that occur with nominal states that may be sustained in the persons’ behavior. These produce large boxes (stacked rows of vertical/horizontal lines) reflecting periods of time during which the participants are fixating the same region. In previous work [Richardson & Dale, 2005] on eye-movement time series using CRQA, we extracted a very simple measure: the diagonal-wise recurrence rate, RR_τ , given by calculating recurrence rate across a range of diagonals around the line of synchronization (LOS) [Marwan *et al.*, 2007]. Doing this on the example time series from Fig. 2 gives the RR_τ plot shown in Fig. 3. This fundamental measure provides the bridge between basic lag sequential analysis and CRQA, as we briefly discuss next.

5. Sequential Analysis \subset CRQA

In the case of testing state matches ($i = j$ in the CT), the diagonal-wise recurrence, RR_τ , is the percentage of matching along the diagonal of the $CT_{k,k}(\tau)$, $k = 1, \dots, 6$.

$$RR_\tau = \frac{\sum_{k=1}^6 CT_{k,k}(\tau)}{N - \tau} \quad (4)$$

Here $N - \tau$ is the number of elements along that diagonal (each subsequent diagonal, specified by τ , has a corresponding smaller size relative to the total size of the time series, N). Given the form of the phi coefficient, which makes use of the product of $i = j$ entries (see Eq. 3), then RR_τ is proportional to the mean phi across k states, $\phi_k(\tau)$. This is evident in comparing the left and right portions of Fig. 3.

In the case of testing for coincidence of two different states at some lag, as noted above, a 2×2 CT can easily be constructed to produce a phi coefficient. In CRQA, one may produce a new recurrence plot to reflect similar questions of interest. To do this, we can define a function that can be applied to time series x and y such that they produce a new pair of time series permitting the construction of a CRP for which RR_τ will reflect coincidence of these non-matching states, $i \neq j$:

$$F(x(t), i) = \begin{cases} 1 & \text{if } x(t) = i \\ \emptyset & \text{otherwise} \end{cases} \quad (5)$$

This function applied to y as $F(y(t), j)$ provides two new time series for which $x(t) = i$ and $y(t) = j$ will produce a point on the CRP, but any \emptyset values can be ignored. Now any recurrence point on the CRP will reflect diagonals in the corresponding converted 2×2 CT . Therefore, in general, any test for contingency in lag sequential analysis will have a proportional test for association along the diagonals of some CRP. (Such a CRP may be better termed a “cross coincidence” plot rather than a cross recurrence plot, as it reflects coincidences of two different nominal values.)

6. CRQA $\not\subset$ Sequential Analysis

Any diagonal or off-diagonal analysis of any $CT_{i,j}(\tau)$ is producible in the basic diagonal-wise RR_τ measures that can be taken from a CRP. Yet, there are numerous measures that CRQA naturally provides that the analytic framework of lag sequential analysis is not specifically designed to obtain. These are primarily analyses over the various lines and structures in the plot and their distribution. Such novel measures, even in what psychologists sometimes term “high-level” abstracted categorical measures, may provide signatures that have direct interpretations in the language of dynamical systems: coupling or attractor strength, bifurcation, onset to chaos, and so on.

An example of bifurcation in our experimental context would be the following scenario: Listener-speaker dyads begin to move to a topic that produces quasi-periodic fixations from one panel to another. For example, a narrative may shift into a topic that is relevant to two particular panels, before which it was relevant to only one. In this case, participants would suddenly exhibit quasi-periodic shifting from one

panel to the other together. In this context, diagonal lines in the plot would remain high (high *DET*), but overall recurrence would be lowered (lowered *RR*). *DET* would remain high because recurrence points can still largely fall on diagonal structures (though these now reflect coordinated sequences of eye movements to multiple panels), but *RR* drops because the eyes do not focus on the exact single state as much as before the period-doubling shift (when participants were perhaps both mostly looking at one panel). The reverse may also hold (e.g., in period-diminishing bifurcations).

Fig. 4 shows six dyads whose measures appear to indicate bifurcation by these considerations. We calculated windowed-recurrence measures by a sliding window of size approximately 6s across the eye-movement time series. In some cases, *RR* remains high then drops, while *DET* remains stable. Often, one can see potential periods of apparent instability, indicated by the brief drop in *DET*, that may indicate a transitioning between these two stable “topics” induced by the driving system (i.e., the speaker).

This makes CRQA, in a central respect, a generalization of lag sequential analysis. This generalization may be termed a solution to the “problem of multiple *CT* realizability.” Any individual *CT* may in fact be produced by a wide variety of differing sequential patterns. This makes any table “multiply realizable” in the actual observed dynamics of the relevant systems. Yet CRQA maintains these temporal patterns and quantifies them in the plot, thus “unfolding” a given *CT* into its dynamical realization for any observed pair of time series.

It is important to note that this does not mean that sequential analysis could not be modified or extended in its own investigative context to accommodate these kinds of questions. Another value of CRQA relative to sequential analysis is that the latter usually chooses a small subset of lags of interest, while CRQA simultaneously treats all possible lags in the CRP. What we argue for here instead is that CRQA provides a natural framework in which to explore the dynamic, nonlinear properties of systems even when coarse-graining them with nominal measures. We revisit the importance of this kind of nonlinear analysis framework for the cognitive sciences in the conclusion.

7. Example: A Coupled Logistic Map

A question may remain whether nominal variables are capable of revealing such dynamical signatures in coupled systems. Even in the eye-movement nominal time series we have collected, participants occasionally look away from the display, thus making the measurements partly intermittent. Here we very briefly show that in a coupled logistic map that has been used in communication research [Buder, 1991], nominal cross recurrence reveals these events: bifurcation and onset to chaos. Previous research has shown RQA’s potential for detecting such transitions in the logistic map and other systems [Marwan *et al.*, 2007; Trulla *et al.*, 1996]. In this very simple demonstration, we will coarse-grain such a coupled system and only intermittently measure from it (see below). CRQA is still capable of revealing these dynamical signatures as the plots in Fig. 4 suggest for human systems.

Buder [1991] coupled two logistic maps, and interpreted their output as levels of “involvement” of each of the communicating systems. The two systems A and B operate according to Eqs. (6-7):

$$A_{t+1} = \lambda_A(1 - A_t)(1 - [A_t - B_t])A_t \quad (6)$$

$$B_{t+1} = \lambda_B(1 - B_t)(1 - [B_t - A_{t+1}])B_t \quad (7)$$

For the purpose of this simulation, we will imagine these two systems to be coupled with $\lambda_A = \lambda_B$, and refer to the parameter simply as λ . We conducted a simulation of these two logistic maps in a manner similar to the human experiments. 200 iterations of the maps were run to produce a time series of “interactions” at each value of λ between 2.3 and 3 at increments of .001. For each such cycle, we simulate two things: coarse-graining (as in the panels in the display in human experiments), and intermittent sampling of the state (as in the participants occasionally looking away). The first of these was done by partitioning the state space of the systems into 6 equally-spaced intervals. The second strategy was conducted by randomly removing 10% of the data from any given λ value in its respective 200 iterations (i.e., 20 entries in the

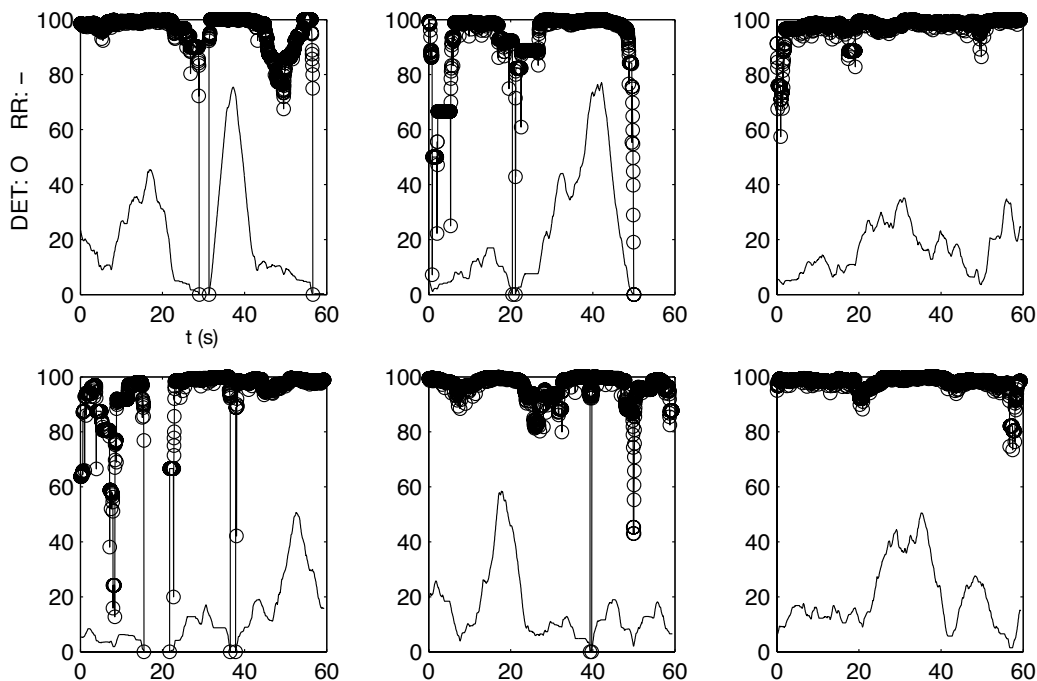


Fig. 4. Six example dyads with DET and RR calculated over sliding windows of size 6s. DET and RR are occasionally partly independent measures that may indicate transitions between stable, coupled modes between listener-speaker dyads. Consider the bottom-right panel. Here, DET is always high, indicating eye-movement sequences taking place overall, but the RR changes. It rises in the middle of the narrative, indicating that the overall eye-movement sequences may focus on fewer panels than around it (thus increasing overall RR). Something akin to “period-diminishing” occurs as the narrative transitions into to 30s, then back to a “period-doubling” transition into lower RR but still high DET in the latter portion of the narrative.

200-element time series for each λ). CRQA measures remain capable of revealing subtle fluctuating in the bifurcations and onsets to chaos as shown in Fig. 5, right. In short, CRQA is capable under coarse-graining and intermittent data sampling to capture interesting transitions in the coupling dynamics of two systems.

This capacity of CRQA may be expected from certain perspectives, such as the equivalence between symbolic itineraries and a system’s phase space through generating partitions in symbolic dynamics. In fact, when the dimensionality of a system (or its description) is $m = 1$, then the coarse-graining of a system may be equivalent to the ϵ -tube [Marwan *et al.*, 2007] used to capture the original system’s orbits. And it is also important to note that misplaced partitions can change apparent underlying topology [Davidchack *et al.*, 2000]. Nevertheless, we do not intend to argue that the values shown in Fig. 5 estimate any dynamical invariants, as are commonly discussed in this literature. Instead, for the purpose of psychological investigation, the relative values of RR and DET provide new windows onto potential transitions (qualitatively identified) in coupled human systems in the kind of research contexts described here.

8. Conclusion

In this paper we have briefly described the essence of the lag sequential approach, and demonstrated its proportionality to nominal CRQA. For nominal scales of measure, CRQA can be seen as an extension of lag sequential analysis, as CRQA retains the various coincident events and their temporal patterns that make up a CT , while permitting complementary measures, such as RR and DET , to be calculated. It is also natural for all lags to be computed simultaneously (and visualized either in recurrence plots or in diagonal lag profiles) in CRQA, whereas this is rarely done in lag sequential analysis. The promise of the measures extracted from CRQA to show the transitions of stable modes in a coupled dynamical system was demonstrated in a simple simulation, which appeared to match the same kind of relative CRQA measures seen in 6 example dyads from previous experimental data. These patterns are promising, and demand

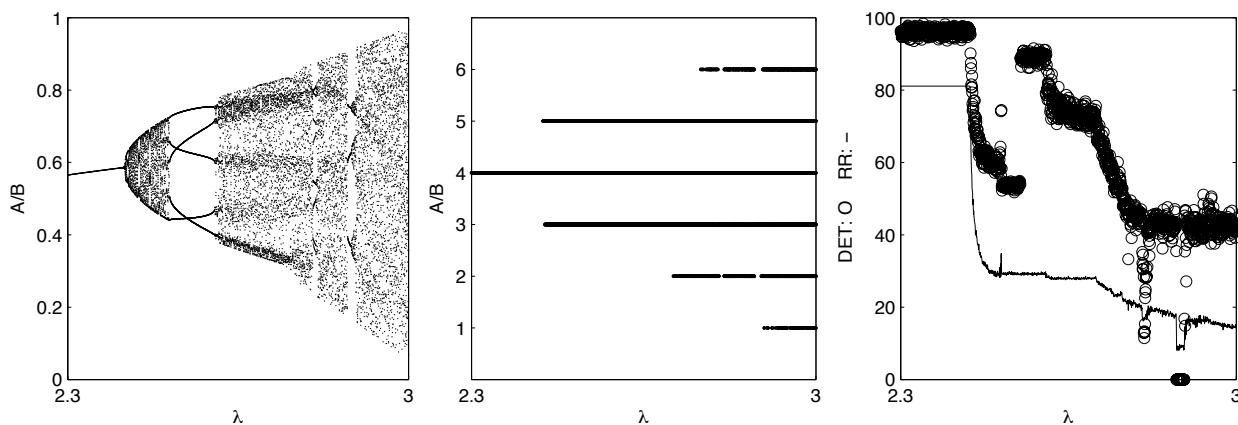


Fig. 5. Left panel shows the A/B states in the final 10 values of the 200 iterations at each λ ; middle panel shows the coarse-graining of the system into 6 bins; right panel shows recurrence measures DET and RR changing relatively across bifurcation and onset to chaos. These analyses based on the 10% data loss and coarse-graining, as described in the text.

future attention in research contexts where lengthy nominal time series are collected from participants (e.g., in text-based analyses: Doxas *et al.* [2007]; Dale & Spivey [2005]).

8.1. Sequential Analysis and CRQA

Despite the relationship we have identified above, lag sequential analysis remains crucial for “event local” analyses in which one wishes to know exactly what occurs from one lag to another, among a small set of previously specified lags. Because the CT provides a window onto the specific patterns that are being matched, their relative frequencies (by inspecting the marginals of the CT), including off-diagonal cells, it offers a highly transparent analysis if one can so choose a small number of lags for inspection. This makes lag sequential analysis a powerful and indispensable analytic framework for nominal time series.

However, if one does not have a specific lag in mind, and wishes to carry out extensive exploratory analyses on temporal relations, then even the RR_τ measure in CRQA provides a computationally cheap means of performing data mining over a nominal time series. In fact, the goodness-of-fit statistics in a multinomial cross-correlation of category values, which may also be shown to be proportional to RR_τ , is computationally expensive, and special methods are developed for it [Komarek & Moore, 2003]. Even in its simplest application, which the authors have used in previous work [Richardson *et al.*, 2007; Richardson & Dale, 2005; Richardson *et al.*, 2009], CRQA is a valuable extension of lag sequential analysis across any range of lags desired (for another recent extension, see also Quera [2008]).

8.2. Cognitive Science and CRQA

For decades it has been recognized that two humans, when interacting, seem to synchronize or coordinate in a variety of ways [Chapple, 1970; Condon & Ogston, 1967]. These ways include low-level synchrony, such as of postural patterns [Shockley *et al.*, 2003], and coordination of higher-level states, such that the same state is likely to occur at approximately the same time between two interacting people (e.g., use of the same grammatical sequences: Dale & Spivey [2006]). The advent of automated measurements—such as eye tracking, as we demonstrated here—permits the application of more sophisticated quantitative techniques to this general research question about coordinating cognitive systems. Cognitive systems exhibit not just periodic behavioral patterns such as postural sway, but also stochastic behavioral events such as the use of a particular word or sentence. Ideal analytic frameworks would be ones that (1) fit both continuous and/or periodic measurement schemes along with nominal schemes (such as words, etc.), (2) are resistant to noise given the inherent stochasticity of cognitive systems in context, and finally (3) are naturally suited to nonlinear dynamical explanations that have recently been urged for cognitive systems but have been in relatively short supply in cognitive science [Van Orden *et al.*, 2003].

CRQA is ideal for this research context. In the current paper, we quantified the coupling of eye-movement patterns in participants who engaged in a one-way interaction while they looked upon a scene. A relatively simple analysis over the diagonals of a CRP provides the quantification of state coordination (i.e., eye movements), and is proportional to mean phi coefficients across category values. The nominal cross recurrence method we employ has an evident relationship to lag sequential analysis, but provides a framework of analysis that is designed for capturing nonlinear, dynamical properties of one system or two coupled systems. The further development of these methods, either by coarse-graining continuous human measurements (where useful) [beim Graben *et al.*, 2000] or extracting coded categories, will bring this sophisticated analysis of nonlinear dynamical systems into higher levels of cognitive science, where they are needed.

References

- Agresti, A. [2002] *Categorical data analysis* (Wiley-Interscience).
- Bakeman, R., Deckner, D. F. & Quera, V. [2005] “Analysis of behavioral streams,” *Handbook of research methods in developmental science*, ed. Teti, D. M. (Blackwell, Oxford).
- Bakeman, R. & Gottman, J. M. [1997] *Observing interaction: An introduction to sequential analysis*, 2nd ed. (Cambridge University Press).
- Bakeman, R. & Quera, V. [1995a] *Analyzing interaction: Sequential analysis with SDIS & GSEQ* (Cambridge University Press).
- Bakeman, R. & Quera, V. [1995b] “Log-linear approaches to lag-sequential analysis when consecutive codes may and cannot repeat,” *Psychological Bulletin* **118**, 272–272.
- beim Graben, P., Saddy, J. D., Schlesewsky, M. & Kurths, J. [2000] “Symbolic dynamics of event-related brain potentials,” *Physical Review E* **62**, 5518–5541.
- Buder, E. H. [1991] “A nonlinear dynamic model of social interaction,” *Communication Research* **18**, 174–198, doi:10.1177/009365091018002003, URL <http://crx.sagepub.com/cgi/content/abstract/18/2/174>.
- Chapple, E. D. [1970] “Experimental production of transients in human interaction,” *Nature* **228**, 630–633, URL <http://adsabs.harvard.edu/abs/1970Natur.228..630C>.
- Condon, W. S. & Ogston, W. D. [1967] “A segmentation of behavior,” *Journal of Psychiatric Research* **5**, 221–235.
- Dale, R. & Spivey, M. J. [2005] “Categorical recurrence analysis of child language,” *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, pp. 530–535.
- Dale, R. & Spivey, M. J. [2006] “Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation,” *Language Learning* **56**, 391–430.
- Davidchack, R. L., Lai, Y. C., Bollt, E. M. & Dhamala, M. [2000] “Estimating generating partitions of chaotic systems by unstable periodic orbits,” *Physical Review E* **61**, 1353–1356.
- Doxas, I., Dennis, S. & Oliver, W. [2007] “The dimensionality of language,” *Proceedings of the 29th Annual Cognitive Science Society*, pp. 227–232.
- Ducasse, S., Rieger, M. & Demeyer, S. [1999] “A language independent approach for detecting duplicated code,” *Proceedings ICSM*, p. 109118.
- Eckmann, J. P., Kamphorst, S. O. & Ruelle, D. [1987] “Recurrence plots of dynamical systems,” *Europhysics Letters* **4**, 973–977.
- Komarek, P. & Moore, A. [2003] “Fast robust logistic regression for large sparse datasets with binary outputs,” *Artificial Intelligence and Statistics (AISTAT)*.
- Marwan, N. & Kurths, J. [2002] “Nonlinear analysis of bivariate data with cross recurrence plots,” *Physics Letters A* **302**, 299–307.
- Marwan, N., Romano, M. C., Thiel, M. & Kurths, J. [2007] “Recurrence plots for the analysis of complex systems,” *Physics Reports* **438**, 237–329, doi:10.1016/j.physrep.2006.11.001.
- Orsucci, F., Walter, K., Giuliani, A., Webber, C. L. & Zbilut, J. P. [1999] “Orthographic structuring of human speech and texts: linguistic application of recurrence quantification analysis,” *International journal of chaos theory and applications* **4**, 21–28.

- Quera, V. [2008] "RAP: a computer program for exploring similarities in behavior sequences using random projections," *Behavior Research Methods* **40**, 21–32.
- Richardson, D. C. & Dale, R. [2005] "Looking to understand: The coupling between speakers and listeners eye movements and its relationship to discourse comprehension," *Cognitive Science* **29**, 1045–1060.
- Richardson, D. C., Dale, R. & Kirkham, N. Z. [2007] "The art of conversation is coordination: common ground and the coupling of eye movements during dialogue," *Psychological Science* **18**, 407–413.
- Richardson, D. C., Dale, R. & Tomlinson, J. M. [2009] "Conversation, gaze coordination, and beliefs about visual context," *Cognitive Science* **33**, 1468–1482, doi:10.1111/j.1551-6709.2009.01057.x, URL <http://dx.doi.org/10.1111/j.1551-6709.2009.01057.x>.
- Sackett, G. P. [1979] "The lag sequential analysis of contingency and cyclicity in behavioral interaction research," *Handbook of Infant Development*, ed. Osofsky, J. D. (New York: Wiley), p. 623649.
- Shockley, K., Santana, M. V. & Fowler, C. A. [2003] "Mutual interpersonal postural constraints are involved in cooperative conversation," *Journal of Experimental Psychology-Human Perception and Performance* **29**, 326–332.
- Spivey, M., Richardson, D. & Dale, R. [2009] "The movement of eye and hand as a window into language and cognition," *The Psychology of Action*, eds. Morsella, E., Bargh, J. & Gollwitzer, P. M. (Oxford University Press, New York), pp. 225–249.
- Strauss, A. L. [1987] *Qualitative analysis for social scientists* (Cambridge University Press).
- Trulla, L. L., Giuliani, A., Zbilut, J. P. & Webber, C. L. [1996] "Recurrence quantification analysis of the logistic equation with transients," *Physics Letters A* **223**, 255–260.
- Van Orden, G. C., Holden, J. G. & Turvey, M. T. [2003] "Self-organization of cognitive performance," *Journal of Experimental Psychology. General* **132**, 331–350, doi:10.1037/0096-3445.132.3.331, URL <http://www.ncbi.nlm.nih.gov/pubmed/13678372>, PMID: 13678372.
- Von Heijne, G. [1987] *Sequence analysis in molecular biology: treasure trove or trivial pursuit* (Academic Press, Inc. Orlando, FL, USA).
- Zbilut, J. P., Giuliani, A. & Webber, C. L. [1998] "Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification," *Physics Letters A* **246**, 122–128.