

Measuring predictability: a cross-species approach

Contact: Hal Stern, Department of Statistics, University of California, Irvine (sternh@uci.edu)

Background

Two recent studies, by Molet et al. (2016) and Davis et al. (in press) with rats and humans, support the hypothesis that the predictability of maternal signals influences cognitive and emotional development. Both studies use a measure of entropy rate to characterize predictability of maternal sensory signals. This document describes the common cross-species approach in which maternal sensory signals are represented as a sequence of observed behaviors and then the predictability of the sequence is measured through its entropy rate.

Characterizing predictability of maternal sensory signals

Rodent Data

Rodent maternal sensory signals were recorded by directly observing the rodent dams (and their pups) during 50-minute windows twice per day for eight days. The set of actions considered as characterizing maternal behavior were, licking/grooming the pups (LG), carrying pups (C), eating (E), nursing (N), nest building (NB), off pups and no other activities (O), or self-grooming (SG). Each dam could only be performing a single behavior at any given time point. This provides a continuous time series of the dam's behavior. Figure 1A is a visualization of this time series for one rodent mother. We combine data from the different observation periods; thereby assuming the series of behaviors is a stationary process in time. The Conte Center research program posits that predictability of maternal sensory signals is a critical factor for the developing brain. As a result, we focus on transitions between behaviors. To do so we ignore the duration of a particular behavior and create a discrete-time sequence of distinct behaviors. This sequence is illustrated in Figure 1B.

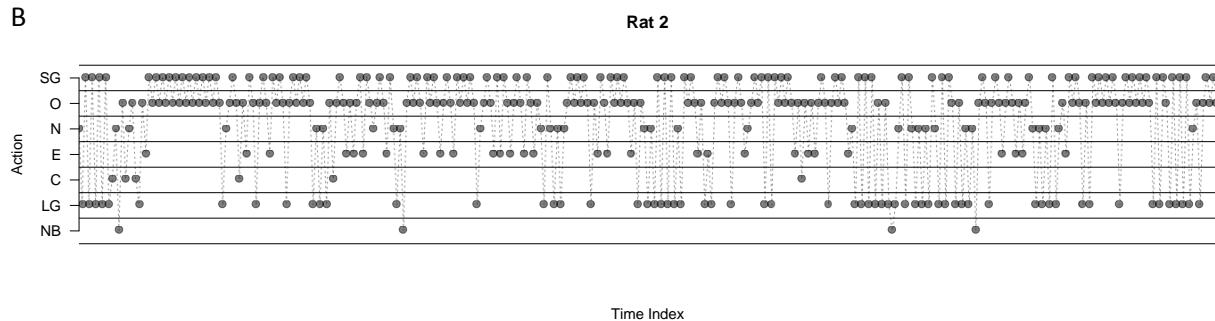
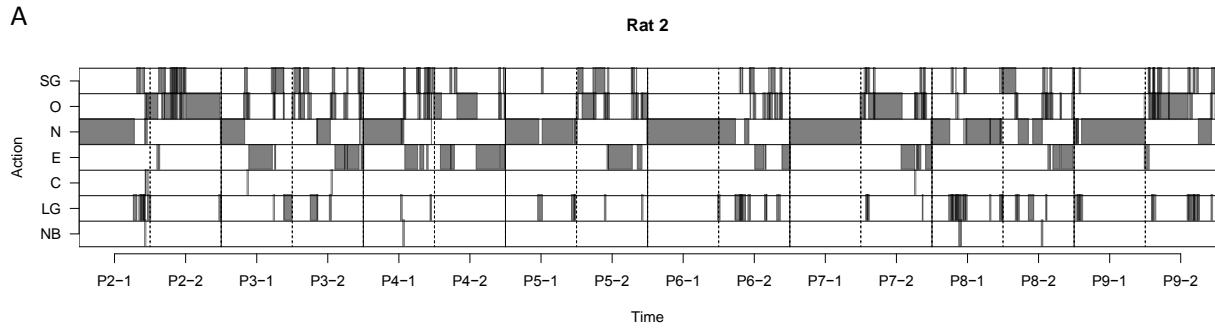


Figure 1. 1A displays a visualization of maternal behavior for two 50-minute periods per day over postnatal days P2-P9. Each row of the figure is a different action. The vertical lines crossing all rows in the figure correspond to beginning/ending of the 50-minute windows. Within each row, the solid blocks are durations during which each action was performed (these may appear as lines if the duration is short enough). 1B is a visualization of transitions in a dam's behavior. The data are the same as in 1A but each behavior has been reduced to a single dot so the duration is not pictured. Divisions between observation periods are not shown.

The sequence of behaviors can be summarized by considering how often specific transitions occur; e.g., how many times self-grooming behavior is followed by nursing. A matrix can be used to organize the transition frequencies with rows indicating the initial behavior in each pair of actions and columns the behaviors being transitioned into. Normalizing each row by the row total provides the proportion of time each “column” event follows each “row” event. These are known as transition probabilities. We visualize this in Figure 2. Note that because we have concatenated results from different observation periods there is a small probability of transitioning from a behavior to the same behavior. The matrix of transition probabilities in Figure 2B is the summary of the data that is used to compute the behavioral entropy rate (see discussion below).

A

Transition Counts

From Action	C	E	LG	N	NB	O	SG	
C	0	0	1	2	0	3	0	6
E	0	0	1	2	0	13	13	29
LG	1	1	0	18	1	16	25	62
N	0	0	19	0	3	8	1	31
NB	0	0	1	0	0	3	0	4
O	5	26	21	3	0	0	57	112
SG	0	2	19	5	0	70	0	96
	C	E	LG	N	NB	O	SG	

B

Transition Probabilities

From Action	C	E	LG	N	NB	O	SG	
C	0	0	0.17	0.33	0	0.5	0	
E	0	0	0.03	0.07	0	0.45	0.45	
LG	0.02	0.02	0	0.29	0.02	0.26	0.4	
N	0	0	0.61	0	0.1	0.26	0.03	
NB	0	0	0.25	0	0	0.75	0	
O	0.04	0.23	0.19	0.03	0	0	0.51	
SG	0	0.02	0.2	0.05	0	0.73	0	
	C	E	LG	N	NB	O	SG	

Figure 2. 2A displays in matrix form the number of times a transition occurs from one behavior (the row labels) to another (the column labels). 2B shows the corresponding empirical transition probability matrix where each row is normalized to have sum one.

Human Data

In our human research, predictability of maternal sensory signals to the infant is characterized in the context of a semi-structured free play session with the mother and her infant. Here we overview the basic steps involved in coding maternal signals and using the resulting data to develop the transition matrix that is used to calculate entropy rate as an index of predictability of maternal signals.

Mother-child interaction during the ten-minute free play session is digitally recorded. Behaviors are then coded to characterize auditory, visual, or tactile sensory signals to the child. Coded behaviors that are used to define sensory inputs are listed in Table 1. Because we are interested in characterizing patterns of maternal signals, all behaviors are coded continuously in real time. For details on coding procedures see the separate coding manual available at the Conte Center website (available at <https://contecenter.uci.edu>). Figure 3 shows an example of data collected from one mother-child pair. The visualization depicts the time series of each of the three types of sensory signals (auditory, tactile, visual) during the 10-minute interaction.

Table 1 – Description of behaviors used to characterize maternal sensory input to her infant.

Sensory Signal	Actions/Behaviors included
Auditory	Verbal utterances including speech and laughing
Tactile	Touching, holding, supporting or carrying the child
Visual	Manipulating object while child is observing

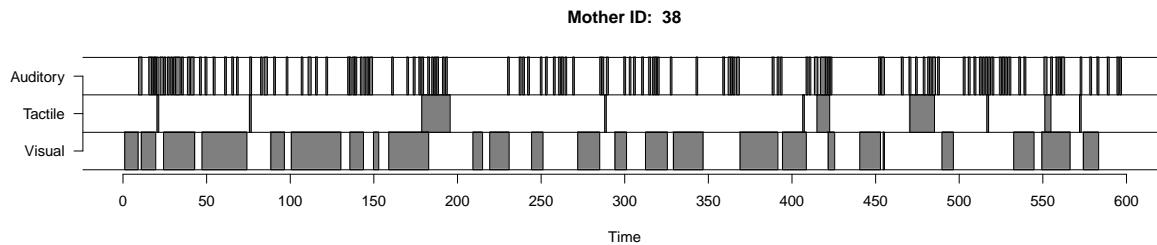


Figure 3. Visualization of the time series of sensory signals coded in the context of a mother interacting with her infant during the semi-structured play session.

Note that some maternal behaviors are roughly instantaneous (e.g., a mother's utterance) while others last for an interval of time (e.g., holding a child in the lap). Because our goal is to identify transitions between behaviors, as discussed above in the rodent section, we focused on the initiation of behaviors in analyzing the human behavioral data (e.g., initiation of a touch or a verbal utterance). It is necessary to standardize identification of discrete behaviors (e.g., to identify whether "Look." "It is a truck." is one event or two). We use verbal utterances to demonstrate our procedure for doing so. The start of each utterance was initially coded from video as a separate event. Thus, the example, "Look." "It is a truck." would initially be coded as 2 events. The instantaneous event time for each utterance was then extended to have a duration of 1-second. Thus, each event was right-padded with a one-second duration to standardize the lengths of the utterances and then instances that overlapped were merged into one event. This means that two maternal utterances separated by 1 second or more were characterized as two events and two maternal utterances for which the initiation was separated by less than 1 second were characterized as a single event. The result of this can be seen in auditory behavior portion of Figure 3. The time interval of one second was chosen based upon an empirical investigation of 10 randomly selected videos indicating that the average maternal utterance to her infant in this context was one second. This can be modified in the software

described below. Although the issue of combining events most frequently affected auditory behaviors, a similar approach was applied to other brief events (e.g., affectionate touch).

The human maternal sensory behaviors are not mutually exclusive (i.e., more than one event can happen at the same time). For example, a mother can be speaking while also touching the child. The eight possibilities listed in Table 2 illustrate all of the possible combinations. To analyze the human data we identify any change between one of these eight sensory combinations and another combination as a transition.

We visualize the resulting sequence of maternal behavior states for one mother in Figure 4A (parallel to the rat data in Figure 1A). As with the rodent, the duration of each state can be collapsed, as the “order” of the behaviors is the only thing that is important for assessing predictability of transitions between sensory signals. Transitions, therefore, are characterized as changes among any of the eight possibilities in Table 2. We visualize this in Figure 4B (parallel to the rat data in Figure 1B). Finally, based upon this sequence of events, the table of transition counts and the empirical transition matrix are constructed. They are provided in Figure 5 (parallel to the rat data in Figure 2).

Table 2 – Descriptions of behavioral categories used to characterize transitions

Grouping of Categories	New Category Descriptions
No Behavior	No Auditory, Tactile, or Visual Stimulation
Single Behavior	Only Auditory Stimulation Only Tactile Stimulation Only Visual Stimulation
Combinations of Two Behaviors	Both Auditory & Tactile Stimulation Both Auditory & Visual Stimulation Both Visual & Tactile Stimulation
Combination of All Behaviors	All States: Auditory, Tactile, & Visual Stimulation

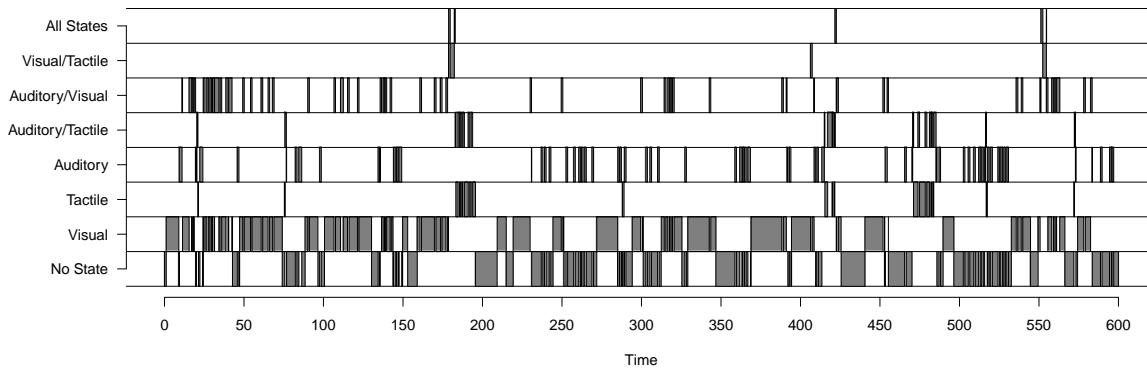
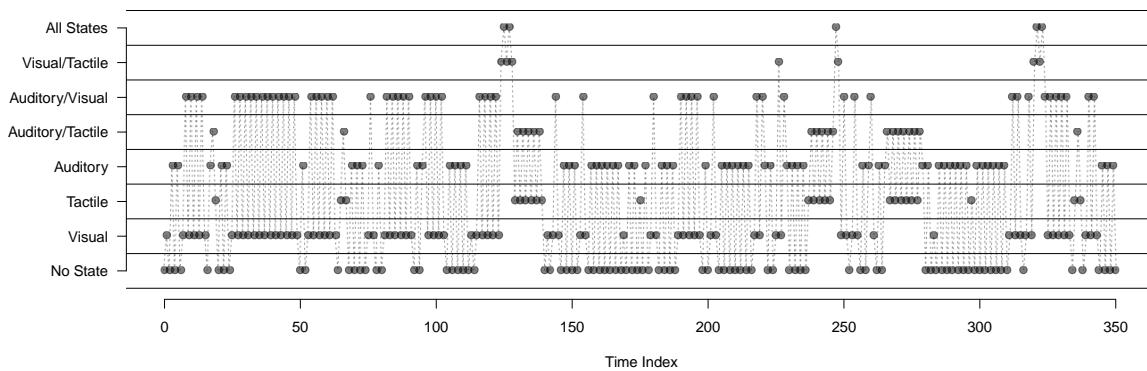
A**Mother ID: 38****B****Mother ID: 38**

Figure 4. 4A displays the continuous time series of maternal sensory signals in the context of interacting with her child and using the eight categories shown in Table 2. Using the eight possible combinations, a time series of mutually exclusive behaviors is created, similar to the rodent data in Figure 1A. 4B displays the discrete time series of human maternal sensory signals where each sensory signal has been reduced to a single dot so duration is not pictured. This is similar to the rat data shown in Figure 1B.

A

Transition Counts

	NS	V	T	A	A/T	A/V	V/T		
From Action	NS	22	5	62	0	0	0	0	89
V	22	0	0	0	0	56	3	0	81
T	6	0	0	0	18	0	0	0	24
A	61	0	0	0	2	2	0	0	65
A/T	0	0	18	1	0	0	0	1	20
A/V	0	57	0	2	0	0	0	0	59
V/T	0	2	1	0	0	0	0	4	7
A/T/V	0	0	0	0	0	1	4	0	5
	NS	V	T	A	A/T	A/V	V/T		

B

Transition Probabilities

	NS	V	T	A	A/T	A/V	V/T		
From Action	NS	0	0.25	0.06	0.7	0	0	0	0
V	0.27	0	0	0	0	0.69	0.04	0	0
T	0.25	0	0	0	0.75	0	0	0	0
A	0.94	0	0	0	0.03	0.03	0	0	0
A/T	0	0	0.9	0.05	0	0	0	0.05	0
A/V	0	0.97	0	0.03	0	0	0	0	0
V/T	0	0.29	0.14	0	0	0	0	0.57	0
A/T/V	0	0	0	0	0	0.2	0.8	0	0
	NS	V	T	A	A/T	A/V	V/T		

Figure 5. 5A displays in matrix form the number of times a transition occurs from one behavior (or combination of behaviors) as indicated by the row label to another as indicated by the column label. 5B shows the corresponding empirical transition probability matrix where each row is normalized to have sum one. This is similar to the rat data illustrated in Figure 2.

Measuring Degree of Predictability – Entropy & Entropy Rate

For both the human and rodent data, the empirical transition matrix (Figure 2B, Figure 5B) is a summary of the way in which a mother transitions among sensory signals and this matrix has significance when modeling behavior. In our approach the observed sequence of behaviors is modeled as a time-homogeneous, first-order stationary Markov chain and the entropy rate of this Markov chain serves as our measure of the predictability of the process. A Markov chain is a specific type of stochastic process (i.e., a sequence of random variables) with a finite state space (i.e., a limited set of behaviors). In a first-order Markov chain the probability distribution of the next observation is related only to the most recently observed behavior or state (i.e., earlier behaviors don't effect the current transition). This is precisely the information contained in the matrix of transition probabilities. The assumption of time homogeneity implies that this probability function does not change with time, while the stationary assumption implies that in the long term the distribution of behaviors will approach a stationary distribution π (i.e., a regular pattern describing the relative frequency of the different behaviors). The entropy rate of a Markov chain is a quantitative measure of the degree to which a future behavior can be

predicted from the current behavior. We provide a brief introduction to the concept and then describe how it is calculated for a Markov chain.

Entropy is a continuous measure that quantifies the predictability of a single random variable (see, e.g., Shannon (1948) and Cover and Thomas (2006)). As an example, consider rolling a single die and recording the number of spots observed. If the die were perfectly fair, each side of the die would have equal probability of one-sixth of occurring. This corresponds to maximum entropy and maximum unpredictability. If the die were such that the same side always landed on top, then the die would be perfectly predictable and have zero entropy. Mathematically, if X is a discrete random variable capable of taking on k values and p_i is the probability of the i^{th} value, then the entropy $H(X)$ is computed as

$$H(X) = - \sum_{i=1}^k p_i \log_2 p_i.$$

Entropy rate extends the concept of predictability from a single random variable to sequences of random variables. The entropy rate of a first-order Markov chain is calculated from the matrix of transition probabilities P where P_{ij} gives the probability of transitioning from state i to state j (as shown in Figures 2B and 5B, where the row index represents the “from” state and the column index represents the “to” state). The entropy rate of a first-order stationary Markov chain is,

$$H(\mathcal{X}) = - \sum_{ij} \pi_i P_{ij} \log_2 P_{ij}$$

where $\pi_i, i = 1, \dots, K$ is known as the stationary distribution and summarizes the long-run characteristics of the process. There are a number of approaches to finding the stationary distribution (see, e.g., Levin et al., 2009). Entropy rate is a continuous measure bounded between zero and $\log_2 K$, where K is the number of possible states. Therefore in the rat experiment the $H_{max, rat} = \log_2 7 = 2.807$ and for the human observational study $H_{max, human} = \log_2 8 = 3$.

Software for calculating the entropy rate

A software package for calculating the entropy rate of an observed Markov chain, Conte Center Behavioral Entropy Rate (ccber), has been created using the R programming language. The package and relevant documentation are available at github.com/bvegetable/ccber. The package incorporates a number of different functions that assist the user in summarizing an observed sequence of behaviors and calculating the entropy rate for the underlying process.

Application Specific Functionality – Behavioral Entropy Rate of Human Data

The package provides the capability to estimate the behavioral entropy rate (and optionally examine diagnostic plots for an individual) from observations recorded in a single Microsoft Excel file output from the Observer system (Noldus) (template available by emailing NRP@du.edu). In addition, the package provides the functionality to estimate entropy rate for all files in a directory (although in this case without the ability to generate diagnostic plots). The README at github.com/bvegetable/ccber provides an overview of the functions and their inputs.

General Functionality – Entropy Rate of Sequences (i.e., relevant to both Rodent Data and Human Data)

If the data are in another format, the package can still be used. By loading the ccber package into R, the user imports the functions that are used for estimating the entropy rate of an observed sequence. To apply these functions, the user is required to create an R object that contains the observed sequence of behaviors (e.g., a vector identifying which of the eight combinations in Table 2 has been observed such as: 1, 3, 5, 8, 6, 8, 6, 4,). Then a series of ccber functions creates the matrix of transition counts as in Figure 5A, creates the matrix of transition probabilities as in Figure 5B, calculates the stationary distribution of the process, and then finally calculates the entropy rate (or a single function can be used to compute the entropy rate directly from the sequence of behaviors.) The entropy rates for the rodent data of Davis et al. (In press) was calculated in this manner. The README at github.com/bvegetable/ccber provides an example of estimating the entropy rate of a first-order Markov chain using these functions.

References

Cover, T.M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edition. **Wiley-Interscience**: Hoboken, NJ.

Davis, E.P., Stout, S.A., Molet, J., Vegetable, B., Glynn L.M., Sandman, C.A., Heins, K., Stern, H., Baram, T.Z. (in press). Exposure to unpredictable maternal sensory signals influences cognitive development across-species. **Proceedings of the National Academy of Sciences**.

Levin, D.A., Peres, Y., Wilmer, E.L. (2009). *Markov Chains and Mixing Times*. **American Mathematical Society**: Providence, RI.

Molet, J., Heins, K., Zhuo, X., Mei, Y.T., Regev, L., Baram, T.Z., Stern, H. (2016). *Fragmentation and high entropy of neonatal experience predict adolescent emotional outcome*. **Translational Psychiatry** 6, e702; doi:10.1038/tp.2015.200

Shannon, C.E. (1948). *A mathematical theory of communication*. **Bell Systems Technical Journal**, 27:379-423.