Memory and Randomness in Animal Movement

Alexander Q. Vining

Animal Behavior Graduate Group, University of California, Davis

Email: [aqvining@ucdavis.edu](mailto:aqvining@ucdavis.edu)

**Abstract**

The ability to store information about the environment is essential to every animal’s survival. For many animals, memory plays an important role in how they navigate through space. Whether remembering a type of environment that’s valuable for foraging, the location of a home base, or a useful route between destinations, animals that can utilize stored information can select routes that reduce the energy costs of travelling or increase their ability to utilize valuable resources. Animal memory is, however, a hidden process. Though researchers have developed many paradigms to demonstrate that memory exists, quantifying memory and differentiating between different types of memory is exceptionally difficult. Information theory and structural analysis offer promising new methods for measuring and quantifying animal movement, with the potential to reveal underlying, internal processes driving an animal’s behavior. Using Bayesian Structural Inference, I generated a posterior distribution of ϵ-machines (maximally predictive and minimally complex unifilar Hidden Markov Models) for a single white-faced capuchin (*Cebus* *capuchinus*). By limiting the prior distribution of possible machines to those generated by splitting states based on the previous state (analogous to a 2-order Markov Process), I aimed to infer whether the capuchin remembered the last location it was in, and planned its future movements using this information. A simple 1st order Markov Order with no memory was strongly favored in the posterior distribution. Across the posterior, average entropy rate was high and statistical complexity was low, suggesting the animal moved randomly and without the use of memory. These methods suffer from many challenges however, and future work to overcome these challenges may yet yield interesting results.

**Introduction**

Memory is a highly complex, often still mysterious process that is profoundly important to human behavior. The centrality of this process in our daily existence is made clear by debilitating diseases like Alzheimer’s; everything we do relies in some way on our memory. But understanding the processes driving memory have proven difficult to understand, and progress toward treating Alzheimer’s has been slow. One challenge in understanding memory is quantifying it. Scientists studying human memory can ask participants to describe their memory, or to complete difficult tasks that require memory, which has advanced our understanding considerable. But these methods don’t tell us two important things: how memory functions internally and how it has evolved. For these questions, we are best suited by turning to animals, but methods for describing and quantifying memory in animals are even more sparse. The ability to describe what and how much animals remember would open a great many doors to understanding the biological basis of memory.

One specific type of memory that is of particular importance to many animals is spatial memory. Animals from all taxon’s live in a single location to which they must return every day, from honeybees, to the limpit *Lottia scabra*, to the dusky owl monkey. Other animals migrate annually and must know how to move between seasonal homes. Yet more animals rely on rare or sparsely dispersed resources for survival and remember the locations of these resources to reduce searching time. All of these animals benefit greatly from the ability to store information about the spatial properties of their environments, but do they all do it the same way? What is the optimal amount of spatial information to remember, given that doing so is energetically costly? And why do humans seems to have a capacity for memory that far exceed that of other animals? These questions are the primary drivers of the work done in this project.

Answering such questions depends on the ability to measure and quantify memory in animals, methods for which are currently lacking. Information theory, however, provides several measures that could potentially be of use in this regard but have not yet been used for this purpose. These measures, such and entropy and statistical complexity, quantify aspects of data’s complexity: it’s unpredictability or the mutual information between the past and future, for example. Given the appropriate data and setup, these measures could capture internal systems such as memory that drive animal behavior. Further, by using a suite of information theory measures to quantify animal movement, it may be possible to identify different patterns of movement are signatures of different memory types.

These exciting applications come in conjunction with new tracking methods that allow for collection of appropriate movement data. In the last few years, high resolution GPS has greatly improved the time and accuracy with which we can determine an animals location. Combined with data from tri-axial accelerometers, movement scientists have a wealth of new data. Methods for analyzing these data, however, have yet to catch up. At the same time, new methods for analyzing ϵ-machines, the smallest, maximally predictive unifilar Hidden Markov Model for a given dataset, are being developed that are ideally suited to these new movement data. Specifically, semi-Markov renewal processes should be able to model both the patterns in an animal’s movement between different locations, but also patterns in the timing of its movement. This project is the exciting intersection of a new type of data in need of analysis and new analyses in need of applications.

Given the short time available, this project is only a first step in applying the appropriate information theory methods to understand animal movement and spatial memory. I started with the GPS track of a single white face capuchin money over the course of a couple months. Using a clustering algorithm in R and Bayesian Structural Inference, I generated a posterior distribution of models for the capuchin’s transitions between three different areas of environment. From these models, I calculated the average and 95% credible intervals for the entropy rate and statistical complexity of the movement data. I found that the animals movement was highly random and was best modeled by a 1st order Markov process, suggesting that, for this partition of the data, the animal was not using memory or its current location to determine its future movements.

**Background**

There are many hypotheses for how spatial memory operates, and it likely operates differently in different species. The path integration hypothesis, for example, posits that animals store information about the distance and angle to important locations relative to themselves. There is some evidence that this is the case in honeybees (Menzel & Gregger, 2015). In a cognitive map, by contrast, animals remember important locations relative to each other, thus creating a robust mental representation of the space they inhabit. Translocated Egyptian fruit bats, for example, are able to find their way home only if they can see major terrain features that are also visible from their home habitat, suggesting the use of a cognitive map (Tsoar et al. 2011). In primates, particularly interesting because of their close evolutionary relationship to humans, researcher have observed straight line movements across long distances that ended at important feeding sites (Janson & Byrne, 2007), but demonstrating that these movements were not the result of simply picking a direction and moving until a resource was found has been more difficult. There have also not been any successful attempts to determine if primates are navigating to locations stored in memory, how they do so.

The Barro Colorado Island in Panama provides a unique location for research on memory in navigation in animals, particularly primates. One trait thought to facilitate the evolution of advanced spatial memory is frugivory. Frugivorous animals often must move between sparsely distributed, discrete resources patches for which the ability to locate without random searching would greatly increase foraging efficiency. There are six species of frugivorous mammal on Barro Colorado that live sympatrically, allowing for informative comparisons. Further, there is only one type of fruit available to these species during the winter: *Dypterix oderata*. Thus, the species on this island are ideal candidates to investigate for patterns of memory in their movement, and are the subject of this project.

**Dynamical System**

The dynamical system of interest has two main components; the physical location of the animal and its internal state. An animals internal state is complicated and consists of many components including motivation, memory, and sensory cues, but for the purposes of this project we are only interested in memory.

**Methods**

Data were collected using a GPS collar affixed to a white faced capuchin (Olga) on Barro Colorado Island, Panama. The GPS recorded Olga’s latitude and longitude every four minutes from 6 am to 6 pm every day from December 15, 2015 to February 20, 2016. The resulting data were cleaned by removing all duplicate timestamps, any rows with missing data, and an algorithm that removed rows that resulted in impossibly long travel distances for the elapsed time. Hierarchical cluster averaging was then used to assign each data point in the track to one of three groups using the ‘hclust’ function in the R package ‘fastcluster’ (Figure 1) Once grouped, I extracted the time between each of Olga’s transitions between clusters and the cluster to which she moved. These transformed data were used for the structural analysis.

To quantify the entropy rate and statistical complexity of Olga’s movement data, I utilized Bayesian Structural Inference to infer a set of ϵ-machines. I used a limited prior for this inference, allowing only eight different topologies. The first and simplest was a three state machine with a three letter alphabet. Each letter represented a move to one of the three clusters, and moved the machine to a state corresponding to that cluster. The rest of the machines were generated by splitting one or more of these states based on the previous state. Thus, if state C were split, there would now be one state for if C were observed while in state B and a second state for if C were observed while in state A. All possible combinations of split states were allowed, but only one split per cluster. Both the prior and posterior distributions were applied over these topologies with a beta of four. I sampled 2000 machines from each of these distributions and calculated their entropy and statistical complexity.

**Results**

In the prior distribution, the probability for the simplest, 1st order Markov machine was 94.7 percent. The probability for this machine went up in the posterior, to 98.0%. Figure 2 depicts a sample of one of these machines. This machine suggests Olga’s movement between clusters is not entirely random; she tends to go through cluster A when travelling between B and C, but still moves directly between the latter clusters about one third of the time. Overall, the entropy rate of the machines in the posterior (mean = 0.972, CI = 0.933-0.996) was much higher than the entropy rate of the machines in the prior (mean = 0.724, CI = 0.342-0.970) (Figure 3). The mean statistical complexity for the prior and posterior distributions was about the same, 1.528 and 1.585, respectively. The credible interval for the posterior, however, converged tightly on the maximum complexity of the simplest machine, reflecting the unpredictability of the data (Figure 4).

**Conclusion**

This project represents a first step toward applying information theory and complexity analysis to animal movement, but much work is needed before meaningful conclusions can be drawn. I determined that Olga’s movement between the regions of her home territory, determined by a clustering algorithm were mostly random and that memory of the cluster Olga came from did not help predict the next cluster she would go to. The biological relevance of this conclusion, however, is tenuous. The choice to cluster into three groups was done for logistical reasons: limiting the alphabet size of this initial analysis and allowing clustering algorithms to operate in a reasonable time-frame on a dataset of nearly 100,000 data points. Future analyses might utilize the hierarchical cluster dendrogram to infer the number of clusters, or a biologically informed assumption about the number of attraction points in the animal’s environment. Methods other than clustering, such as home range kernel density estimates, may also be more effective at partitioning the animal’s environment in a biologically meaningful way.

It is also possible that the limited prior didn’t allow for an appropriate model to be found. Though the simple set used was selected to maximally represent models with memory, the process primarily being investigated, a full set of topologies would allow for a more complete test, but becomes computationally intensive only small increases in the alphabet size and number of state. Given the available data for times between transitions, the use of semi-markov renewal processes would allow for testing the structure and complexity not just of where animals move, but when. I believe that the failure of the current analysis to find evidence of memory in Olga’s movement is reflective of a design in need of improvement, not of the lack of a biological memory process. Finally, completing analyses on the movement of other species living on Barro Colorado Island will allow for informative comparative analyses. New methods for partitioning the geographic space, generating topologies for the prior, and including the transition time data are all goals I have to improve this analysis in the future.

**Bibliography**

Janson, C. H., & Byrne, R. (2007). What wild primates know about resources: opening up the black box. Animal cognition, 10(3), 357-367.

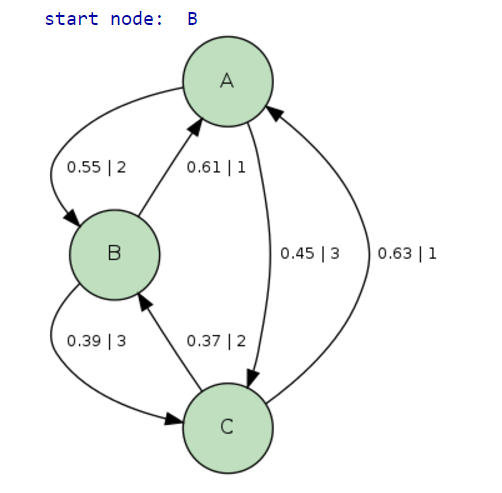
Menzel, R., & Greggers, U. (2015). The memory structure of navigation in honeybees. Journal of Comparative Physiology A, 201(6), 547-561.

Tsoar, A., Nathan, R., Bartan, Y., Vyssotski, A., Dell'Omo, G., & Ulanovsky, N. (2011). Large-scale navigational map in a mammal. *Proceedings of the National Academy of Sciences*, *108*(37), E718-E724.

**Figures**



Figure 1. The GPS fixes for Olga during Winter 2015-16 colored by cluster.



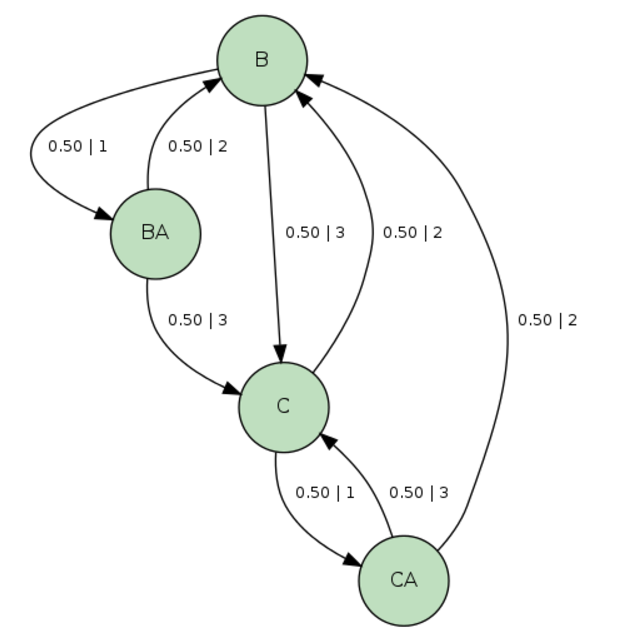


Figure 2. a) Sample machine from posterior distribution of models for Olga’s cluster transitions. This model indicates Olga was more likely than not to move to region A from either region B or C, but her overall her transition were mostly unpredictable. b) Topology of machine in which state A has been split.

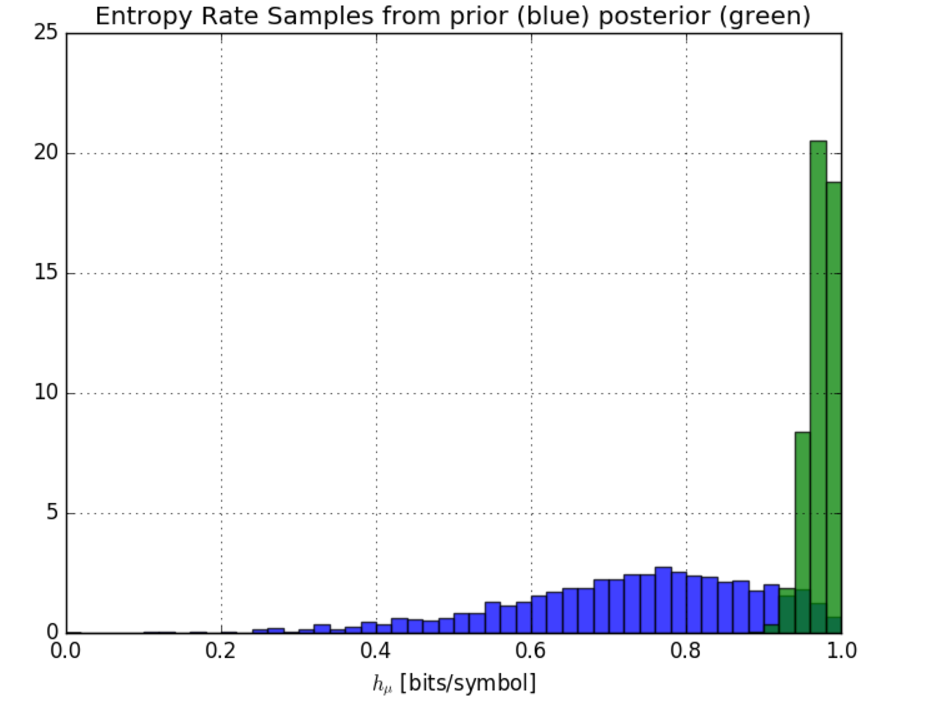


Figure 3: Entropy rates of models in the prior and posterior distributions for Olga’s transitions between clusters. The models in the posterior strongly tended toward the maximum entropy of the models in the prior.

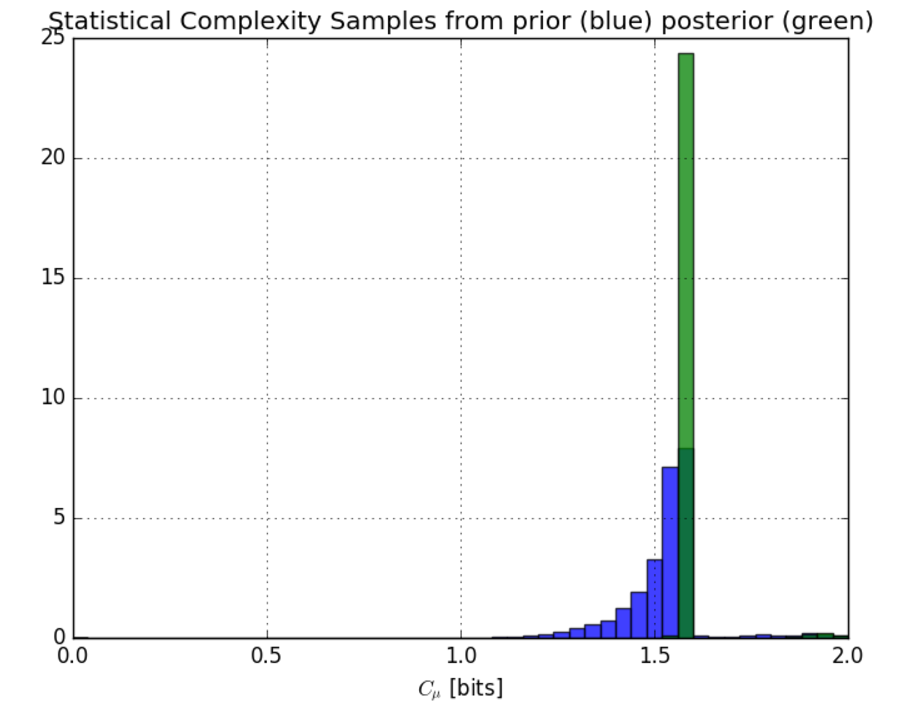


Figure 4. Statistical complexity of models in the prior and posterior distributions of Olga’s transitions between clusters. The prior distribution shows the most models around the maximum statistical complexity of the simplest topology in the prior, with a decreasing number of models with less statistical complexity and very few models with more statistical complexity. The posterior distribution shows a strong convergence on the maximum statistical complexity for the simplest model topology.