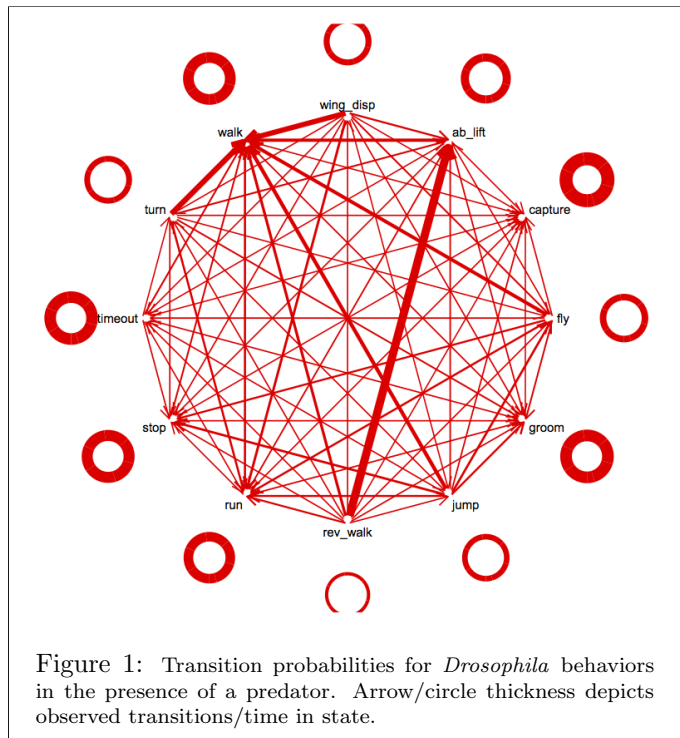


Background: Understanding the targets of natural selection on behavior is challenging in large part because behavior consists not just of isolated actions, but as extended sequence of actions driven by shifts in underlying motivational states [1, 2]. Discerning the motivational states corresponding to behavioral transitions is difficult because behavior unfolds in a continuous stream, and yet identification of these states is critical for understanding the mapping between genetic (e.g. in neuromodulatory systems) and behavioral variation [3]. One way to address this is the use of Hidden Markov Models (HMM). HMMs use any kind of sequence data (or transition matrices derived from sequence data) to estimate latent (hidden) variables, with transitions between hidden states influencing observed behavioral states (emission probabilities). HMMs are used in image processing, text processing, bioinformatics (gene finding) as well as for the study of behavior [4–11]. With respect to the study of behavior, the hidden variables are often used to model underlying motivational states [12–15]. Importantly, HMMs allow for the incorporation of additional predictors and hierarchical structure to assess biological factors (sex, species) and reduce assumptions (stationarity and homogeneity) that have plagued other forms of sequential analysis. Exciting recent development of random effects HMMs [12, 16–18] may allow for the rigorous evaluations of questions pertaining to behavioral syndromes (personality), yet such extensions have not yet been realized nor implemented in a general framework.

Despite the potential of HMMs, the meaning of the hidden (latent) states is rarely obvious. Currently it is unclear how the hidden factors can be interpreted as motivational or neurophysiological states. Thus in addition to issues of estimation and inference inherent in all data modeling approaches, whether HMMs lead to biologically relevant inferences about underlying mechanisms is unclear. Evaluation of hidden factors in contexts like gene finding is relatively easily (detecting transcripts or gene function), but is less straightforward for behavior. Thus there is a critical need to both extend and evaluate HMMs in systems where motivational states can be directly evaluated, and thereby to correlate latent variables in HMMs to validated control mechanisms.

Objectives: The long term goal of this research is to understand the evolution of underlying motivational states, and observed sequence of behaviors. This proposal seeks to extend and evaluate the suitability of HMMs to address such questions, using a two-stage interdisciplinary strategy. We will use behavioral sequence data from several biological systems (where motivational states can be experimentally manipulated through the presentation of biologically meaningful stimuli) and from *Avida*, where we can directly examine the underlying execution of operations leading to the observed behaviors. This research will have implications for the study of behavioral sequence analysis specifically as well as Hidden Markov Models more generally. In particular, we expect to explore methods for validating the biological traits corresponding to latent variables detected by HMMs as well as extending HMMs to incorporate random effects (i.e personality) and hierarchical structures.



Aim 1: *Extend and evaluate the interpretability of HMMs for underlying motivational states in behavioral analysis using Avida.*

Aim 2: *Evaluate motivational states in biological systems (*Drosophila* and bees) where these can be manipulated via evolutionary or experimental manipulations.*

PROPOSED RESEARCH

Aim 1: We will set up simplified experiments with focal Avidians. Utilizing a cross-section of Avida organisms evolved during previous experiments on foraging strategies [19], and predation and predator-avoidance behaviors [20], we will record both their realized behavioral sequences and the actual execution order (trace) of their genetic programs. We will analyze these genetic programs with their execution traces to identify the sensory inputs and internal states that were used in formulating decisions. We will also feed the behavioral sequences produced into HMM generators to measure the number of internal states predicted by the HMM, and how each influence behavior. This phase of research will enable us to determine the correspondence of estimated HMMs with the actual control system used by the digital organisms, allowing calibration of our HMM estimation techniques to maximize correspondence for use in the subsequent aims. In addition this will allow us to evaluate how estimation is influenced by important empirical parameters such as time allotted to observations, number of observed states, and number of observed transitions.

Aim 2: While behavioral sequence data is widely available, rarely are the inferred motivational states explicitly validated by evidence for the proximate mechanisms of state transitions. We will use the primary systems of Dworkin and Dyer, namely searching behavior in bees and anti-predator behaviors in *Drosophila*. As seen in Figure 1, *Drosophila* exhibit a wide array of behaviors in the presence of predators [21], with non-random transitions between them. We will manipulate predation risk to alter underlying motivational states in populations evolved with and without a predator (zebra jumping spider) for ~60 generations. Dyer will use both previously collected and new data to address the underlying motivational states in two species of bees. Prior work has shown that

experiencing changes in the quality and amount of food alters behavioral sequences related to the learning of new resources [22,23] or searching for resources not yet found. This occurs in a species specific manner [24,25]. Figure 2 shows data from two bees of different species as they search without reward among three locations over the course of an hour. We will investigate whether an HMM approach can explain behavioral sequence variation involving changes in some hidden states independent of others or if states are modulated in a correlated way. This may guide future work exploring gene expression dynamics for genes known to modulate foraging behavior in insects.

We will use several libraries in R (DEPMIXS4, HIDDENMARKOV, HMM, HMM.DISCNP, RHMM) and Python (GHMM WRAPPER, SCIKIT-LEARN, HMMPYTK, PYHMMTB) to assess performance of the HMMs. As there is no general purpose software for hierarchical or random effect HMMs [12,16–18], PI Dworkin will prototype them in the STAN probabilistic programming language for hierarchical modeling using MCMC [26,27]. A general purpose parser will be written to generate files in appropriate input formats.

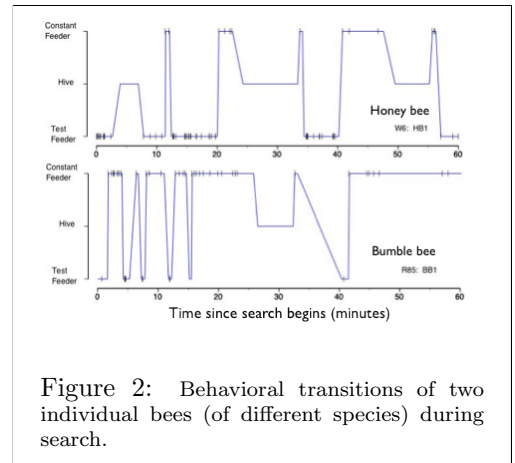


Figure 2: Behavioral transitions of two individual bees (of different species) during search.

References

- [1] Sibly R, McFarland D (1976) On the fitness of behavior sequences. *American Naturalist* 110: 601–617.
- [2] Scholes III E, Sodhi N (2006) Courtship ethology of carola’s parotia (*Parotia carolae*). *The Auk* 123: 967–990.
- [3] Houston A, McFarland D (1976) On the measurement of motivational variables. *Animal Behaviour* 24: 459–475.
- [4] Feldman A, Balch T (2004) Representing honey bee behavior for recognition using human trainable models. *Adaptive Behavior* 12: 241–250.
- [5] Krogh A (1998) An Introduction to Hidden Markov Models for Biological Sequences. In: *Computational Methods in Molecular Biology*.
- [6] Rimey RD, Brown CM (1991) Controlling eye movements with hidden Markov models. *International Journal of Computer Vision* 7: 47–65.
- [7] Morgan B (1976) Markov properties of sequences of behaviours. *Applied Statistics* 25: 31–36.
- [8] Stanke M, Waack S (2003) Gene prediction with a hidden markov model and a new intron submodel. *Bioinformatics* 19: ii215–ii225.
- [9] Eddy SR (1998) Profile hidden markov models. *Bioinformatics* 14: 755–763.
- [10] Rabiner L (1989) A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* 77: 257–286.
- [11] Brand M, Oliver N, Pentland A (1997) Coupled hidden markov models for complex action recognition. In: *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. IEEE*, pp. 994–999.
- [12] Zucchini W, Raubenheimer D, MacDonald IL (2008) Modeling Time Series of Animal Behavior by Means of a Latent-State Model with Feedback. *Biometrics* 64: 807–815.
- [13] MacDonald IL, Raubenheimer D (1995) Hidden Markov Models and Animal Behaviour. *Biometrical journal* 37: 701–712.
- [14] Carola V, Mirabeau O, Gross CT (2011) Hidden Markov model analysis of maternal behavior patterns in inbred and reciprocal hybrid mice. *PLoS ONE* 6: e14753.
- [15] Patterson TA, Basson M, Bravington MV, Gunn JS (2009) Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology* 78: 1113–1123.
- [16] Altman RM (2007) Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting. *Journal of the American Statistical Association* 102: 201–210.
- [17] Maruotti A (2011) Mixed Hidden Markov Models for Longitudinal Data: An Overview. *International Statistical Review* 79: 427–454.

- [18] Schliehe-Diecks S, Kappeler PM, Langrock R (2012) On the application of mixed hidden Markov models to multiple behavioural time series. *Interface focus* 2: 180–189.
- [19] Bartlett FA, Dyer FC, Pennock RT (2012) Evolution of migratory-like behavior in avidians. *Artificial Life* 13.
- [20] Wagner AP, Zaman L, Dworkin I, Ofria C (2013) Arms races and strategy chases promote the evolution of prey intelligence. arXiv preprint arXiv:13101369 .
- [21] Parigi AA, Porter C, Pitchers W, Dworkin I (2013). The anti-predatory behavioral repertoire of *Drosophila melanogaster*. doi:10.6084/m9.figshare.774612. URL <http://dx.doi.org/10.6084/m9.figshare.774612>.
- [22] Wei CA, Dyer FC (2009) Investing in learning: why do honeybees, *Apis mellifera*, vary the durations of learning flights? *Animal Behaviour* 77: 1165–1177.
- [23] Wei C, Rafalko S, Dyer F (2002) Deciding to learn: modulation of learning flights in honeybees, *Apis mellifera*. *Journal of Comparative Physiology A* 188: 725–737.
- [24] Townsend-Mehler JM, Dyer FC, Maida K (2011) Deciding when to explore and when to persist: a comparison of honeybees and bumblebees in their response to downshifts in reward. *Behavioral ecology and sociobiology* 65: 305–312.
- [25] Townsend-Mehler JM, Dyer FC (2012) An integrated look at decision-making in bees as they abandon a depleted food source. *Behavioral Ecology and Sociobiology* 66: 275–286.
- [26] Stan Development Team (2013). Stan: A c++ library for probability and sampling, version 2.1. URL <http://mc-stan.org/>.
- [27] Hoffman MD, Gelman A (2013) The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* in press.