

Evaluating contingencies by a dual system of learning the structure and the parameters of the environment

Tamas J. Madarasz, Joseph E. LeDoux (tamas, joseph.ledoux)@nyu.edu

Center for Neural Science, NYU, 4 Washington Place, New York, NY 11201 USA

Joshua P. Johansen (johans@brain.riken.jp)

RIKEN Brain Science Institute, 2-1 Hirosawa Wako-shi Saitama, Japan 351-0198

Abstract

How does the brain identify stimuli that are relevant for predicting important events and how does it distinguish spurious relationships from truly predictive ones? We examined two contrasting theoretical frameworks: in the first, learning proceeds by considering a fixed hypothesis of the environment's statistical structure (the set of predictive and causal relationships) and adjusting strength parameters for these relationships to optimize predictions. In contrast, the second approach directly assesses ambiguity in predictive relationships by evaluating multiple hypothesis of the environment's statistical structure. We compared these frameworks in an animal model of aversive conditioning, allowing us to also manipulate the underlying brain systems. We show that when facing novel predictive stimuli, rats initially adopt a structure learning strategy, but switch to updating parameters during subsequent learning.

Keywords: Bayesian modeling; Animal cognition; Representation; Causal Reasoning

Introduction

To enhance their chance of survival human and nonhuman animals learn to make predictions based on sensory cues in their environment. However, it is not clear how they identify stimuli that are relevant for specific predictions, or how they distinguish predictive relationships from coincidence when evidence about the relationship is ambiguous. How subjects evaluate this type of ambiguity is a central question both in the field of classical conditioning (where predominantly animal subjects learn to predict impending threats or rewards from experience), and in studies of human causal learning, with previous research identifying many important parallel phenomena between two the fields.

Established accounts of classical conditioning (e.g. Rescorla & Wagner, 1972, for an overview see Pearce and Bouton, 2001), as well as some Bayesian accounts of causal judgments (e.g Holyoak & Cheng, 2011) have characterized the learning process in these tasks in terms of subjects fitting parameters in a fixed generative (or discriminative) model of the environment, with different cues competing with each other to predict important outcomes. Such parameter learning or cue competition approaches don't explicitly distinguish ambiguity in the environment's statistical structure (the set of all predictive and causal relationships), and uncertainty about the strength of established associations (e.g. the probability with which an outcome follows a cue). Instead they assume that subjects either have inherent knowledge about which variables matter for a specific task or prediction, or that they learn relationships between all the environmental

variables, potentially having to fit a very large number of parameters in complex real-world environments.

A drawback of such a learning strategy is that when the structure of the environment is not known in advance (e.g. when encountering novel stimuli), and inability to quickly distinguish spurious from predictive relationships can lead to incorrect predictions, especially when the sampling from the environment is limited, or when a large number of environmental variables are present. Falsely assuming predictive relationships can lead to overfitting (so-called data-fragmentation, Koller & Friedman 2009), by having to explain a combinatorially large space of stimulus interactions based on a limited number of observations. This in turn can lead to poor generalization for future predictions.

A contrasting approach is to directly evaluate competing models of the environment's statistical structure based on the evidence sampled from the environment. Since structure learning also considers sparser sets of statistical dependencies between variables, it will often lead to better predictions when some variables in the environment do not in fact correlate. Such a structure learning model for causal learning was introduced by Griffith & Tenenbaum (2005), and has proven successful in characterizing a broad range of causal judgments by humans. In subsequent work Lu et al. (2008) have also argued that strength and structure queries elicit causal judgments that are empirically and theoretically distinguishable, depending on the wording of the queries, with strength or structure learning models providing better fits respectively.

In this paper we contrasted these two learning strategies in an implicit learning task, using animal subjects in an aversive conditioning paradigm. We show that neither strategy alone gives a good account of animal behavior, and instead find that when presented with novel stimuli, animals initially pursue a structure learning strategy, followed by a switch to parameter learning on subsequent conditioning episodes. We focus on a simple example of an ambiguous cue-outcome relationship that arises when an outcome occurs both in the presence and absence of a sensory cue, the so-called degraded contingency effect (Rescorla, 1968). In the first part of the paper we show that cue competition, a characteristic of associative and Bayesian parameter learning models, is not required to learn a reduced cue-outcome contingency between a novel predictive stimulus and salient outcome. Instead, in accordance with a structure learning account, we find that changes in the strengths of different associations are dissociable. In

our third experiment we explore a wider range of conditioning phenomena and perform quantitative model comparison to show that a structure learning approach better explains the data than a variety of (Bayesian and associative) parameter learning models. Finally, in Experiment 4 we show that learning in a subsequent conditioning session is explained by switching to parameter learning, using the distribution over structures learned during initial exposure to the stimuli.

Experiment 1

Intermixing unsignaled outcomes with outcomes signaled by a discrete sensory cue in classical conditioning experiments is known to lead to reduced behavioral responding to the cue during subsequent memory tests. This reduction has been interpreted as a sign of competition between learning an association to the conditioning Context (the static conditioning chamber that is continually present during the conditioning phase of the experiment), and the discrete predictive cue (also referred to as conditioned stimulus, or CS). In particular, in an aversive (or fear) conditioning setting, a strong association formed between the conditioning context and the shock outcome (also called the unconditioned stimulus, or US) is claimed to reduce subsequent learning of the tone-shock association. Alternatively, a strong contextual association has been proposed to be competing with the tone-CS at the time of memory expression (Stout & Miller, 2007). We first wanted to determine whether predictions of the cue competition models were supported when ambiguity in the ability of a given cue to predict the outcome was high. To test this we examined the relationship between Context and Tone memory strengths, while also varying the order of tone shock pairings and unsignaled shock USs.

Method

Subjects 79 male Sprague-Dawley rats (Hilltop) weighing 275-300g on arrival, individually housed on a 12h light/dark cycle, and given food and water ad libitum.

Materials and Stimuli Animals were fear conditioned in a sound-isolating conditioning chamber (Context A). The two predictive stimuli were a 30s, 5kHz auditory conditioned stimulus (the Tone), consisting of thirty consecutive auditory pips with pips at 1HZ, and the conditioning chamber (Context). The predicted outcome (or US) was a 1s, 1mA footshock. Tests for contextual fear memory strength were conducted in the original conditioning chamber, Context A. Tests for the Tone fear memory were conducted in a different chamber (Context B, different shape, size, lighting and odor from Context A). Memory strength was evaluated by scoring rats' freezing behavior during the tests, with freezing defined as the cessation of all bodily movement with the exception of respiration-related movement.

Procedure Each animal was taken to the conditioning room, and placed into the conditioning chamber, where it was given a series of signaled and unsignaled footshocks with random intertrial intervals (ITIs) of around 2min. Tone-shock

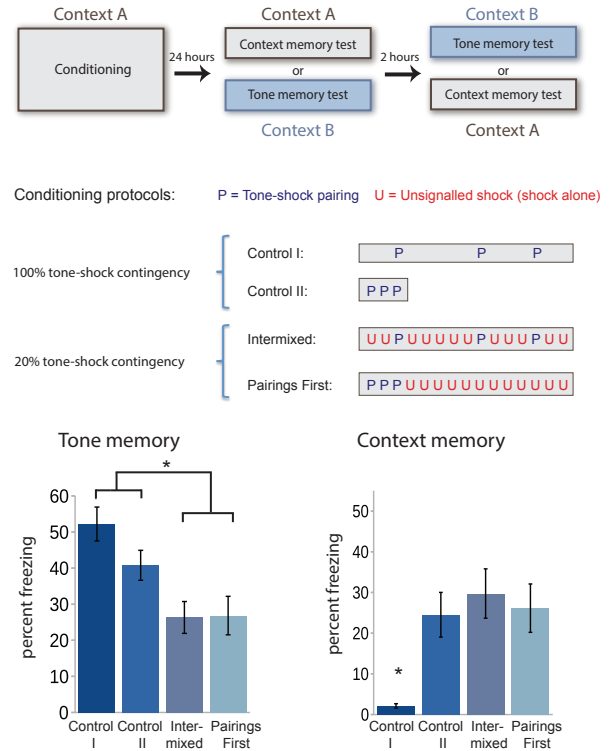


Figure 1: Reduced Tone-shock contingency results in reduced Tone memory, irrespective of trial order and with or without contextual learning or changes in Context memory strength.

pairings consisted of a presentation of the Tone stimulus, with the last pip coterminating with the footshock. Unsignaled footshocks meant the arrival of a footshock without prior warning.

During conditioning (Fig. 1), animals were given either three massed tone-shock pairings before, or three spaced pairings intermixed with 12 unsignaled shocks (Pairings first and Intermixed groups respectively, all with 20% contingency). Control I and II animals were given three CS-US pairings only (100% contingency). The Control I training protocol was identical to the Intermixed group, with the three CS-US pairings spaced, but with all UUS omitted. The control II group received massed CS-US pairings identical to the Pairings First group, with the subsequent UUSs omitted, and conditioning terminated after the third CS-US pairing (see Figure 1, middle). Contextual and tone-evoked freezing was measured 24h later. Context memory tests consisted of 5min spent in Context A, and the test for Tone memory consisted of 5 presentations of the 30s Tone stimulus in Context B, with presentations separated by random ITIs of around 2min. The two tests were separated by 2 hours, and the order varied. Total time spent freezing during the 5 tone presentations, and during the context test were calculated and are represented as percentage of total time spent freezing (divided by 2.5min and 5min respectively). Error bars represent standard errors.

Results and Discussion

We found that animals showed similar low levels of tone-evoked freezing in both reduced contingency conditions (Fig. 1, bottom left), with a two-way ANOVA showing significant main effect for contingency ($p < 0.0001$). Animals were therefore sensitive to the ambiguity of the CS-US relationship, and demonstrated the ability to integrate contingency information irrespective of the temporal order of training trials, contradicting a traditional cue competition based contextual blocking account of contingency degradation. Further, we observed a reduction in Tone memory strength between the Pairings First and Control II groups without a corresponding change in Context memory strength (Fig. 1, bottom right), suggesting that a simple give-and-take between the different associations at the time of memory expression also fails to account for contingency learning.

Experiment 2

To better understand the influence of contextual associations on learning the tone-shock contingency, and to directly test for cue competition during learning and/or retrieval, we tested if animals were sensitive to the reduced tone-shock contingency even in the absence of learning any contextual associations. We therefore infused the NMDA-receptor antagonist APV into the dorsal hippocampus (DH) prior to conditioning, a manipulation known to block the formation of contextual memories (Kim, DeCola, Landeira-Fernandez & Fanselow, 2011).

Methods

Subjects 36 adult male rats similar to those in Exp. 1.

Materials and Stimuli We dissolved the NMDA-receptor antagonist APV in saline, and infused this mixture into the dorsal hippocampus. Stimuli and materials during conditioning and testing were identical to Exp. 1.

Procedure Prior to the conditioning session, animals received either APV and saline, or saline only injections. Around half of the animals in both drug conditions were then trained using the Pairings First training protocol, and the rest of the animals with the CTL II protocol. Conditioning and testing proceeded as in Exp. 1.

Results and Discussion

Figure 2 shows that APV infusions significantly impaired contextual learning as expected, but had no effect on learning the Tone-shock contingency, providing further evidence against cue competition as a sufficient or necessary mechanism for contingency learning.

Models

In this section we give details of the models formalizing the different (structure vs. parameter) learning strategies, using the representational formalism of Bayesian networks. To

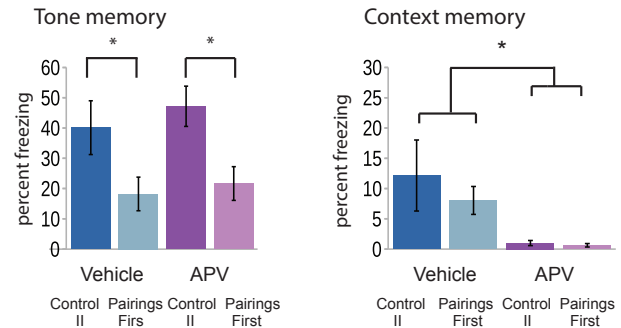


Figure 2: Hippocampal APV injections impair the acquisition of contextual aversive memories, but have no effect on learning the reduced tone-shock contingency. A two-way ANOVA on the Context memory data showed main effect for drug ($p < 0.01$). A two-way ANOVA on the Tone memory data showed a main effect for contingency, ($p < 0.01$). No interactions were significant.

generate predictions for animal behavior, the models have to specify the functional form by which concurrently present cues are combined to predict an outcome. Here we use the so-called noisy-OR generating function, corresponding to the assumption that different cues predict an outcome independently of each other. We also explored simulations with a rectified linear function (typical of traditional associative models), but this choice did not affect the conclusions in the paper, though it slightly worsened the fit of each type of model.

Using these generating functions, a parameter learning model needs to introduce extra variables to be able to explain the findings in Experiment 2. Similarly to the original causal learning models, we therefore introduced an additional 'Background' variable. This Background variable represents the sum of all unobservable or unspecified influences in the environment (and in particular on the US occurrence), and might in principle allow parameter learning to account for the results of Exp. 2. We can formulate our Structure Learning Model (SLM) both with, or without a Background variable, achieving similar model fits in both cases. Having the Background variable that is always assumed to have a predictive connection to the US obviates the need to specify a prior distribution for the probability of US occurrence when all the predictive stimuli are absent, leading to a simpler and perhaps more principled model.

Structure Learning Model (SLM)

SLM learns a posterior probability distribution over the different possible constellations of predictive relationships in the environment (represented by the different graph structures in Fig. 3), given observations during conditioning. During retrieval, the strength of an association is evaluated by calculating the posterior probability of a connection (a direct edge, or path in the graph) between the corresponding cue and outcome, using a model-averaging procedure.

We calculated the posterior distribution over different

Bayesian network structures, without assuming or learning specific parameter values ω_i for the edges. We considered the six possible graph structures $G_i \in \mathbf{G}$ that can lead to different predictions about the US. In Graphs 1 and 2, leaving out, or adding the edge *Context* \rightarrow *Tone* is irrelevant when making predictions about the US, we therefore considered only one of each of these pairs of functionally equivalent graphs (the one with no $C \rightarrow T$ edge)

By Bayes' rule, the posterior probability of each graph is given by

$$P(G_i|D) \propto \int_{\omega_{G_i}} P(G_i) \cdot P(\omega_{G_i}|G_i) \cdot P(D|G_i, \omega_{G_i}) d\omega_{G_i}$$

where priors over the edge parameters $\omega_1, \dots, \omega_4$ were uniform and independent. $P(G_1)$, the prior probability of graph 1, was a free parameter, with the other graphs having equal priors $P(G_{i,i>1}) = \frac{1-P(G_1)}{5}$. Since the conditioned stimuli are largely neutral at the beginning of experiments, this corresponds to $P(G_1)$ being close to 1. The posterior probability of an edge from the Tone to the US,

$$\sum_{G_i} P(G_i|D) \cdot I_{(T \rightarrow US \in G_i)}$$

determined the strength of the Tone-shock association (where I is the indicator function). Context-US association was determined by the posterior probability of an edge from Context to the US or, when no direct edge exists, the probability of an indirect path, weighed by an estimated strength of the Context-Tone connection.

$$\sum_{G_i} P(G_i|D) \cdot I_{(C \rightarrow US \in G_i)} + \gamma \cdot P(G_4|D)$$

where $\gamma = \alpha \cdot \frac{\text{number of trials with Tone}}{\text{total number of trials}}$ for a constant α that we fit.

Parameter learning Model (PLM)

PLM predicts behavioral responses based on learning the posterior mean of the edge parameters ω in the maximally connected graph (Graph 6), starting from some prior distributions over the edge parameters. For maximum flexibility of this model, these priors were allowed to be independent and to be different for each edge, but were assumed to be fixed before conditioning begins, so that they are shared by animals across all conditioning protocols. For parameter $\omega_{j,k}$ (for the edge $X_i \rightarrow X_j$) using the joint prior over we have

$$\hat{\omega}_{j,k} = E(\omega_{j,k}|D, G_6) = \int_{\omega|G_6} \omega_{j,k} \cdot P(\omega) \cdot P(D|G_6, \omega) d\omega$$

The model predicts that freezing responses are explained by the probability of a shock calculated given the stimulus present during testing, using standard probabilistic inference in Graph 6 with parameters $\hat{\omega}$. We restrict the priors for the edge parameters to come from a Beta distribution, fitting the model thus means finding a pair of parameters for each of the four prior beta distributions (8 parameters in total), such that they best explain the behavioral data across all training protocols.

Learning both structure and parameters (PSLM)

Learning a full posterior over the Bayesian network representations included first learning a distribution over the graph structures as in SLM, and then learning a posterior distribution for the parameters present for each structure, similarly to PLM. For each graph, predictions are calculated using the posterior mean edge parameters, and these predictions are then averaged, weighed by the posterior probability of each graph.

Associative models

We also included in the model comparison two advanced associative cue competition models that extend the classical Rescorla-Wagner model to include more complex interactions between stimuli. Van Hamme and Wasserman's (1994) extension (HW-RW) implements cue competition during learning, while in the sometimes competing retrieval model (SOCR) cues compete during performance (Stout & Miller, 2007), with the two models using different strategies to capture covariance information in the cue-outcome relationships. These models can't account for the results from Experiment 2, but it was important to see if they could account for the purely behavioral results that didn't involve neural manipulations.

Experiment 3

To enable model fitting and comparison we collected further behavioral data in a manner similar to Exp. 1, but using varied numbers of tone-shock pairings and unsignaled shocks, allowing us to test which models can simultaneously explain behavioral phenomena under different conditions of ambiguity. In particular, when USs arrive only in the presence of the Tone (i.e. only tone-shock pairings are given), the association between Context and US is itself highly ambiguous: it is not clear whether predictive power should be attributed to just the Context, or just the Tone, or both. We fitted our models using the combined behavioral results from Experiments 1 and 3, and used the best-fit parameters for each model to predict the results of the brain manipulation in Exp. 2.

Methods

Subjects 117 adult male rats similar to those in Exp. 1.

Materials and Stimuli Materials and Stimuli were identical to those in Exp. 1.

Procedure Animals were conditioned using different sequences of tone-shock pairings and unsignaled shocks, as detailed in Fig. 4.

Results and Discussion SLM provided a better quantitative fit than PLM or the associative cue competition models, while also using fewer free parameters (Table 1 and successfully accounted for standard learning curves as well as contingency evaluations under ambiguity, including the effects of contingency degradation and the u-shaped learning curve for the context memory strength during overshadowing

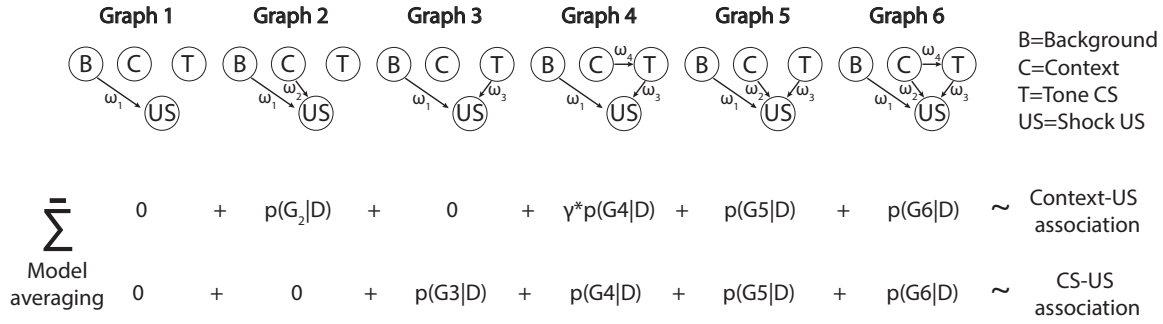


Figure 3: The graph structures used in the Structure Learning Model (with the Background term). PLM uses graph 6 to learn the four edge parameters.

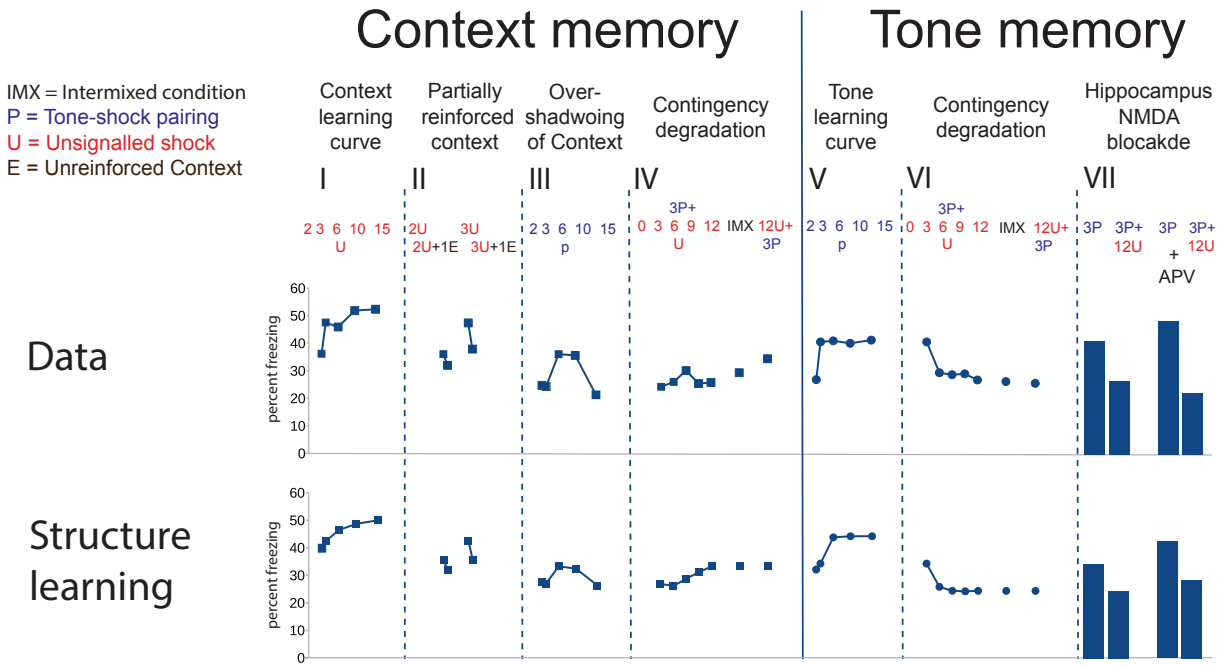


Figure 4: Behavioral data and model fit for SLM.

(Fig.4 I,II,III). In addition, SLM but not PLM or the associative models, successfully predicted the effects of hippocampal manipulations. The fixed structure models all failed to explain the behavioral data, both because the memory strengths didn't covary according to a cue competition principle, and because these interactions were different under the different ambiguity conditions. While it might be possible to extend the associative models to better fit our dataset by adding further model parameters and variables, it is unlikely that this would lead to a principled and general framework for how animals evaluate ambiguity. In contrast SLM was robust to changes in specific components of the model, and provided low error both using an alternative formulation with no 'Background' variable ($MSE = 15.39$), or using a rectified additive generating function ($MSE = 16.64$), suggesting that it is the principle of evaluating different models of the environment that enables it to match observed behavior. SPLM

provided a similar fit to SLM, but performed worse according to measures controlling for extra model parameters, such as adjusted R-squared. SLM thus provided the best fit for the behavioral data and predicted the effects of neural manipulations on learning, suggesting a close correspondence between structure learning and the strategies animals use to resolve ambiguity.

Experiment 4

In the final experiment we examined how animals adopt their learning strategy if conditioning sessions are distributed over time, in particular if trials carrying information about the reduced tone-shock contingency are separated in time from the full contingency trials.

Methods

Subjects 51 adult male rats similar to those in Exp. 1.

Table 1: Comparison of Model fits

Model	MSE	s.d. of the MSE (% freezing squared)	MSE Experiment 2	Model Parameters
SLM	13.44	0.02	34.03	2
SPLM	12.61	0.24	18.29	9
PLM	20.80	0.29	245.64	8
SOCR	28.26	-	245.58	5
HW-RW	67.06	-	848.88	8

Materials and Stimuli Identical to Exp.1.

Procedure Conditioning and Testing was distributed over 3 days. On Day1 CTL II+Exposure and Delayed Degradation groups received the CTL II conditioning protocol, Pairings First group was conditioned as before. On Day 2, the Delayed Degradation group received 12 unsignaled shocks, and the CTL II+Exposure group was exposed to Context A for an identical length of time. Animals in the Pairings First group weren't conditioned on Day 2. On Day 3, all animals were tested as in Exp. 1.

Results and Discussion

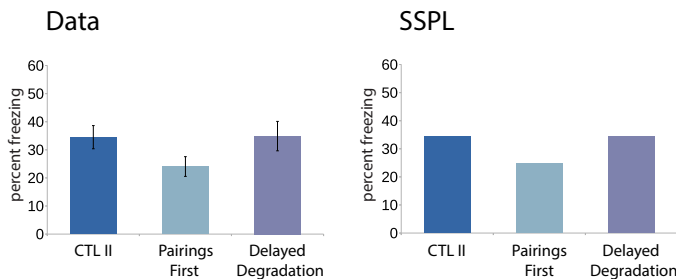


Figure 5: Data and Simulation for delayed contingency degradation

Whereas unsignaled shocks delivered within the same session decreased Tone memory strength as before (one-way ANOVA with Bonferroni correction, $p < 0.5$), the same number of unsignaled shocks delivered 24h after tone-shock pairings were given had now such effect. SLM and SPLM cannot account for this difference as their representation of these two conditioning protocols is essentially identical. However, the difference in the resulting memory strengths is well-predicted by a model in which animals switch from a structure learning strategy on initial encounter with the stimuli, to a parameter updating strategy on later exposures. The Switching from Structure to Parameter Learning model (SSPL) assumes that in the first session animals learn a distribution over possible graphical models as in SLM (therefore producing identical predictions to SLM for our previous data), and in the second session they only update the edge parameters for each

structure, using a uniform prior over the edge parameters in both cases. SSPL is thus similar to SPLM, but uses separate datasets (conditioning trials) for evaluating the different structures, and for learning parameters.

Conclusion

We showed using a combination of behavioral, neural and modeling techniques that animals use a structure learning strategy and evaluate different statistical models of their environment when encountering novel stimuli. Further, we demonstrated that once a distribution over these models is learned, they refine this representation by updating parameters in these statistical models. Further important questions include exploring the exact circumstances under which one or the other approach (or a mixture of the two) is preferentially recruited, and in particular understanding what changes in the environment can lead to reevaluating the statistical structure once it's already learnt. This could lead to a better understanding of the complex mechanisms involved in memory updating, and in particular inform new ways in which aversive associations might be permanently extinguished.

References

- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334-384.
- Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: the new synthesis. *Annual Review of Psychology*, *62*, 135-163.
- Kim, J. J., DeCola, J. P., Landeira-Fernandez, J., & Fanselow, M. S. (1991). N-methyl-D-aspartate receptor antagonist APV blocks acquisition but not expression of fear conditioning. *Behavioral Neuroscience*, *105*, 126-133.
- Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*.
- Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*, 955-984
- Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, *52*, 111-139.
- Rescorla, R. a. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*(1), 1-5.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II Current Research and Theory* (Vol. 21)
- Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): a formalization of the comparator hypothesis. *Psychological Review*, *114*(3), 759-783.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue Competition in Causality Judgments: The Role of Nonpresentation of Compound Stimulus Elements. *Learning and Motivation*, *25*, 127-151.