Chapter 19

Semi-rational Models of Conditioning: The Case of Trial Order

Nathaniel D. Daw, Aaron C. Courville, and Peter Dayan

June 15, 2007

1 Introduction

Bayesian treatments of animal conditioning start from a generative model that specifies precisely a set of assumptions about the structure of the learning task. Optimal rules for learning are direct mathematical consequences of these assumptions. In terms of Marr's (1982) levels of analyses, the main task at the computational level would therefore seem to be to understand and characterize the set of assumptions from the observed behavior. However, a major problem with such Marrian analyses is that most Bayesian models for learning are presumptively untenable due to their radically intractable computational demands.

This tension between what the observer *should* and what she *can* do relates to Marr's (1982) distinction between the computational and algorithmic levels, Chomsky's (1965) distinction between performance and competence, and Simon's (1957) notion of bounded rationality. As all these examples suggest, we need not simply abandon normative considerations in favor of unconstrained, cheap and cheerful heuristics. Indeed, the evolutionary argument often taken to justify a normative approach (that organisms behaving rationally will enjoy higher fitness) in fact suggests that, in light of computational costs, evolution should favor those who can best and most efficiently *approximate* rational computations.

In short, some irrational models are more rational than others, and it is these that we suggest should be found to model behavior and its neural substrates best. Here, in search of such models, we look to the burgeoning and theoretically sophisticated field studying approximations to exact Bayesian inference.

The major difficulty facing such analyses is distinguishing which characteristics of observed behavior relate to the underlying assumptions, and which to approximations employed in bringing them to bear. This is particularly challenging since we do not know very much about how expensive computation is in the brain, and therefore the potential tradeoff between costs and competence. As with all model selection questions, the short answer to the question of assumption versus approximation is, of course, that we cannot tell for sure.

In fact, that assumption and approximation are difficult to distinguish is a particularly acute problem for theorists attempting to study normative considerations in

isolation, which motivates our attempt to study both together. In practice, we may hope that credible candidates for both assumptions and approximations have different and rather specific qualitative fingerprints. In this chapter, we explore these ideas in the context of models of Pavlovian conditioning and prediction experiments in animals and humans. In particular, we address the recent trenchant discussion of Kruschke (2006), who argued against pure Bayesian learning models in favor of a particular heuristic treatment based on the effects of trial ordering in tasks such as backward blocking (Lovibond *et al.*, 2003; Shanks, 1985; Wasserman and Berglan, 1998) and highlighting (Kruschke, 2003, 2006; Medin and Bettger, 1991). His 'locally Bayesian' model, while drawing on Bayesian methods, neither corresponds to exact inference nor is motivated or justified as an approximation to the ideal.

While we agree with Kruschke that features of the data suggest something short of ideal reasoning in the statistical models we consider, we differ in the substance of the alternative modeling frameworks that we employ. On our analysis, at the computational level, effects of the ordering of trials bring up the issue of assumptions about how task contingencies change. The qualitative fingerprint here is *recency* – for most credible such models, recent trials will provide more information about the present state of affairs than distant trials. We show that some trial order effects emerge from optimal inference. For others, notably highlighting, which appears to be inconsistent with recency, we consider the effects of inferential approximation. We show how *primacy* effects, as seen in highlighting, qualitatively characterize a number of simplified inference schemes.

We start by describing the Kalman filter model of con ditioning (Dayan *et al.*, 2000), which arises as the exact inference process associated with an analytically tractable, but highly simplified, Bayesian model of change. We show that this model leads to certain trial order effects, including those associated with backward blocking; we then consider inferential approximations in this framework and their implications for trial ordering in highlighting. We conclude with a discussion of how approximations might be explicitly balanced by reasoning about their accuracy.

2 Learning as Filtering

2.1 The Generative Model

Consider a prediction problem in which, on trial t, the subject observes a possibly multidimensional stimulus $\mathbf{x}_t$ and must predict an outcome r_t. In a classical conditioning experiment, $\mathbf{x}$ might be a binary vector reporting which of a set of stimuli such as tones and lights were present and r some continuously distributed amount of food subsequently delivered. (We denote possible stimuli as A, B, C and write a unit amount of food as R.) In this case, the animal's prediction about r is of course measured implicitly, e.g. through salivation. In a human experiment using, for instance, a medical cover story, $\mathbf{x}$ might report a set of foods (e.g.$\mathbf{x}$=AC), with r being a binary variable reporting whether a patient developed an allergic reaction from eating them (e.g. $r = R$ or 0).

We briefly review a familiar statistical approach to such a problem (e.g. Dayan and Long, 1998; Griffiths and Yuille, this volume). This begins by assuming a space of

hypotheses about how the data ($\mathcal{D}$, a *sequence* of $\mathbf{x}\rightarrow r$ pairs) were generated. Such hypotheses often take the form of a parameterized stochastic data generation process, which assigns probability $P(\mathcal{D}|\theta)$ to each possible data set $\mathcal{D}$ as a function of the (initially unknown or uncertain) parameter settings θ. Then, conditional on having observed some data (and on any prior beliefs about θ), one can use Bayes' rule to draw *inferences* about the posterior likelihood of the hypotheses (here, parameter settings),$P(\theta|\mathcal{D})$. Finally, to choose how to act on a new trial T, with stimuli $\mathbf{x}_T$, the subject can calculate the posterior distribution over r_T using the likelihood-weighted average over the outcomes given the stimuli and each hypothesis:

$$P(r_T|\mathbf{x}_T,\mathcal{D}) = \int \mathrm{d}\theta P(r_T|\mathbf{x}_T,\theta)P(\theta|\mathcal{D}) \tag{1}$$

The interpretive power of the approach rests on its normative, statistical foundation. Indeed, the whole procedure is optimal inference based on just the generative model and priors, which collectively describe what is known (or at least assumed) about how the data are generated.

The generative models of conditioning based on these principles typically split into two pieces. First, they assume that included in θ is one or more sets of values $\mathbf{w}_t$ that govern a distribution $P(r_t|\mathbf{x}_t,\mathbf{w}_t)$ over the output on trial t given the stimuli. Second, they assume something about the probabilistic relationship between the parameters $\mathbf{w}_t$ and $\mathbf{w}_{t+1}$ associated with successive trials. The job of the subject, on this view, is to estimate the posterior distribution $P(\mathbf{w}|\mathcal{D})$ from past outcomes so as to predict future outcomes. We discuss the two pieces of the models in turn.

2.2 Parameterized Output Distribution

One standard assumption is that the outcome r_t on trial t is drawn from a Gaussian distribution:

$$P(r_T \mid \mathbf{x}_T,\mathbf{w}_T,\sigma_o) = N\left(\sum_j w_{tj} \cdot x_{tj},\sigma_o^{\,2}\right) \tag{2}$$

Here, the weights $\mathbf{w}_t$ specify the mean outcomes $\mathbf{w}_{tj}$ expected in the presence of each stimulus j individually, and σ_0 is the level of noise corrupting each trial. We typically assume that σ_0 is known, although it is formally (though not necessarily computationally) easy to infer it too from the data.

There are various points to make about this formulation. First, note that Equation 2 characterizes the outcome r_t conditional on the input $\mathbf{x}_t$, rather than modeling the probability of both variables jointly. In this sense it is not a full generative model for the data $\mathcal{D}$ (which consist of both stimuli and outcomes). However, it suffices for the present purposes of asking how subjects perform the task of predicting an r_t given an $\mathbf{x}_t$. We have elsewhere considered full joint models (Courville *et al.*, 2003, 2004, 2006).

Second, in Equation 2, the mean of the net prediction is assumed to be a *sum* of the predictions w_{tj} associated with all those stimuli x_{tj} present on trial t. This is ecologically natural in some contexts, and is deeply linked to the Rescorla–Wagner (1972) model's celebrated account of cue combination phenomena such as blocking and conditioned inhibition. However, there are other possibilities in which the stimuli are

treated as *competing* predictors (rather than cooperating ones: Jacobs *et al.*, 1991a,b). For instance, one alternative formulation is that of an additive mixture of Gaussians, which uses an extra vector of parameters $\boldsymbol{\pi}_t \in \theta$ to capture the competition:

$$P(r_t \mid \mathbf{x}_t, \mathbf{w}_t, \sigma_o, \boldsymbol{\pi}_t) \propto \sum_j \pi_{tj} x_{tj} \mathcal{N}(w_{tj}, \sigma_o) + (1 - \boldsymbol{\pi}_t \cdot \mathbf{x}_t) \mathcal{N}(w_{t0}, \sigma_o) \tag{3}$$

Here, on each trial, a single stimulus j is chosen from those present with probability π_{tj} (or a background stimulus with the remaining probability) and its weight alone provides the mean for the whole reward. This is known to relate to a family of models of animal conditioning due to Mackintosh (1975; see Dayan & Long, 1998; Kruschke, 2001), and formalizes in a normative manner the notion in those models of cue-specific attentional weightings, with different stimuli having different degrees of influence over the predictions.

Finally, the Gaussian form of the output model in Equation 2 is only appropriate in rather special circumstances (such as Daw *et al.*, 2006). For instance, if r_t is binary rather than continuous, as in many human experiments, it cannot be true. The obvious alternative (a stochastic relationship controlled by a sigmoid, as in logistic regression) poses rather harder inferential problems. The common general use of the Gaussian illustrates the fact that one main route to well-found approximation is via exact inference in a model that is known to be only partially correct.

2.3 Trial Ordering and Change

Equation 2 and its variants capture the characteristics of a single trial, t. However, the data $\mathcal{D}$ for which the subject must account are an ordered series of such trials. The simplest way to extend the model to the series is to assume that the observations are all independent and identically distributed (IID), with ,$\mathbf{w}_t = \mathbf{w}, \forall t,$ and:

$$P(\mathcal{D} \mid \mathbf{w}, \sigma_o) = P(r_1 \ldots r_T \mid \mathbf{x}_1 \ldots \mathbf{x}_T, \mathbf{w}, \sigma_o) = \prod_{t=1}^{T} P(r_t \mid \mathbf{x}_t, \mathbf{w}, \sigma_o) \tag{4}$$

Since the product in Equation 4 is invariant to changes in the order of the trials t, *exact* inference in this model precludes any effect of trial ordering. Kruschke (2006) focuses his critique on exactly this issue.

However, the assumption that trials are IID is a poor match to a typically nonstationary world (Kakade and Dayan, 2002). Instead, most conditioning tasks (and also the real-world foraging or inference scenarios they stylize), involve some sort of change in the contingencies of interest (in this case, the coupling between stimuli $\mathbf{x}$ and outcomes r, parameterized by $\mathbf{w}$). If the world changes, an ideal observer *would not* treat the trials as either unconditionally independent or having identical distributions, but instead, different outcome parameters $\mathbf{w}_t$ may obtain on each trial, making:

$$P(\mathcal{D} \mid \mathbf{w}, \sigma_o) = P(r_1 \ldots r_T \mid \mathbf{x}_1 \ldots \mathbf{x}_T, \mathbf{w}, \sigma_o) = \prod_{t=1}^{T} P(r_t \mid \mathbf{x}_t, \mathbf{w}_t, \sigma_o) \tag{5}$$

Nonstationarity turns the problem facing the subject from one of inferring a single $\mathbf{w}$ to one of *tracking* a changing $\mathbf{w}_t$ in order to predict subsequent outcomes.

To complete the generative model, we need to describe the change: how the $\mathbf{w}_t$ that applies on trial t relates to those from previous trials. A convenient assumption is first-order, independent Gaussian diffusion:

$$P(\mathbf{w}_{t+1} \mid \mathbf{w}_t, \sigma_d) = \mathcal{N}(\mathbf{w}_t, \sigma_d^2 \mathbf{I}) \tag{6}$$

As for the observation variance σ_o^2, we will assume the diffusion variance σ_d^2 is known.

Together, Equations 2, 5, and 6 define a generative model for which exact Bayesian inference can tractably be accomplished using the Kalman (1960) filter algorithm, described below. It should be immediately apparent that the assumption that $\mathbf{w}_t$ is changing gives rise to trial ordering effects in inference. The intuition is that the parameter was more likely to have been similar on recent trials than on those further in the past, so recent experience should weigh more heavily in inferring its present value. That is, the model exhibits a recency bias. But note that this bias arises automatically from normative inference given a particular (presumptively accurate) description of how the world works.

Here again, the Gaussian assumption on weight change may accurately reflect the experimental circumstances, or may be an approximation of convenience. In particular, contingencies often change more abruptly (as between experimental blocks). One way to formalize this possibility (Yu and Dayan, 2003, 2005) is to assume that in addition to smooth Gaussian diffusion, the weights are occasionally subject to a larger shock (e.g. another Gaussian with width $\sigma_j \gg \sigma_d$). However, the resulting model presents substantial inferential challenges.

2.4 The Kalman Filter

Consider the generative model of Equations 2, 5, and 6. As is well known, if prior beliefs about the weights $P(\mathbf{w}_0)$ take a Gaussian form, $\mathcal{N}(\hat{\mathbf{w}}_0, \Sigma_0)$, then the posterior distribution having observed data for trials up to $t-1$, $P(\mathbf{w}_t|\mathbf{x}_1 \ldots \mathbf{x}_{t-1}, r_1 \ldots r_{t-1})$ will also be Gaussian, $\mathcal{N}(\hat{\mathbf{w}}_t, \Sigma_t)$. That is, it consists of a belief $\hat{\mathbf{w}}_t$ about the mean of the weights and a covariance matrix Σ_t encoding the uncertainty around that mean. Because of the Gaussian assumptions, these quantities can tractably be updated trial by trial according to Bayes theorem, which here takes the form (Kalman, 1960):

$$\hat{\mathbf{w}}_{t+1} = \hat{\mathbf{w}}_t + \Sigma_t (r_t - \hat{\mathbf{w}}_t \cdot \mathrm{x}_t) \tag{7}$$

$$\Sigma_{t+1} = \Sigma_t - \boldsymbol{\kappa}_t \mathbf{x}_t \Sigma_t + \sigma_d^2 \mathbf{I} \tag{8}$$

with *Kalman gain* vector $\mathbf{k}_t = \Sigma_t \mathbf{x}_t^T / (\mathbf{x}_t \Sigma_t \mathbf{x}_t^T + \sigma_o^2)$. Note that the update rule for the mean takes the form of the Rescorla–Wagner (1972) (delta) rule, except with each stimulus having its own individual learning rate given by the appropriate entry in the Kalman gain vector.

The Kalman filter of Equations 7 and 8 is straightforward to implement, since the sufficient statistics for all the observations up to trial t are contained in the

fixed-dimensional quantities $\hat{\mathbf{w}}_t$ and $\boldsymbol{\Sigma}_t$. Indeed, this is why the assumptions underlying the Kalman filter have been made as approximations in cases in which they are known not to hold.

This completes the normative treatment of learning in a non-stationary environment. In the next section, we apply it to critical examples in which trial order has a significant effect on behavior; in Section 4, we consider approximation schemes employing inexact inference methods, and consider the sorts of normative trial order effects they capture, and the non-normative ones they introduce.

3 Backward Blocking and Highlighting

We illustrate the effects of trial ordering in the Kalman filter model using two key order sensitive paradigms identified by Kruschke (2006), backward blocking (e.g. Lovibond *et al.*, 2003; Shanks, 1985; Wasserman and Berglan, 1998) and highlighting (Kruschke, 2003, 2006; Medin and Bettger, 1991). In particular, we consider the effect of a priori nonstationarity in the Kalman filter ($\sigma_d^2 > 0$) compared, as a baseline, against the same model with $\sigma_d^2 = 0$, which is equivalent to the IID assumption of Equation 4 and is therefore trial ordering invariant.

3.1 Backward Blocking

Table 19.1 details an experiment in which stimuli *A* and *B* are paired with reinforcement over a number of trials (we write this as $AB \rightarrow R$), and responding to *B* alone is then tested. Famously, predictions of *R* given *B* probes are attenuated (*blocked*) when the $AB \rightarrow R$ training is preceded by a set of $A \rightarrow R$ trials, in which *A* alone is paired with reinforcement (Kamin, 1969). One intuition for this *forward blocking* effect is that if reinforcement is explicable on the basis of *A* alone, then the $AB \rightarrow R$ trials do not provide evidence that *B* is also associated with reinforcement. This particular intuition is agnostic between accounts in which stimuli cooperate or compete to predict the outcome.

The next line of the table shows the experimental course of *backward* blocking (Shanks, 1985), in which the order of these sets of trials is reversed: $AB \rightarrow R$ trials are followed by $A \rightarrow R$ trials and then by a test on *B* alone. Backward blocking is said to occur if responding to *B* is attenuated by the $A \rightarrow R$ post-training. The intuition remains the same – the $A \rightarrow R$ trials indicate that *B* was not responsible for reinforcement on the $AB \rightarrow R$ trials. However, what makes backward blocking interesting is that this lack of responsibility is only evident *retrospectively*, at a point that *B* is no longer provided.

Table 19.1. Experimental paradigms

	Phase 1	Phase 2	Test	Result
Forward blocking	$A \rightarrow R$	$AB \rightarrow R$	*B*?	*R* attenuated
Backward blocking	$AB \rightarrow R$	$A \rightarrow R$	*B*?	*R* less attenuated
Highlighting	$AB \rightarrow R \times 3n$ &	$AB \rightarrow R \times 1n$ &	*A*?	*R*
	$AC \rightarrow S \times 1n$	$AC \rightarrow S \times 3n$	*BC*?	*S*

Backward blocking is an example of 'retrospective revaluation,' in which subsequent experience (with *A*) changes the interpretation of prior experience (with *B*). As we will discuss, the mere existence of retrospective revaluation strongly constrains what sort of inferential approximations are viable, because particular information must be stored about the initial experience to allow it to be reevaluated later. Critically, backward blocking tends to be weaker (that is, responding to *B* less attenuated) than forward blocking (e.g. Lovibond *et al.*, 2003). Since forward and backward blocking just involve a rearrangement of the same trials, this asymmetry is a noteworthy demonstration of sensitivity to trial order (Kruschke, 2006), and thus refutation of the IID model.

The simulation results in figure 19.1a confirm that forward and backward blocking are equally strong under the IID Kalman filter model. It may not, however, be

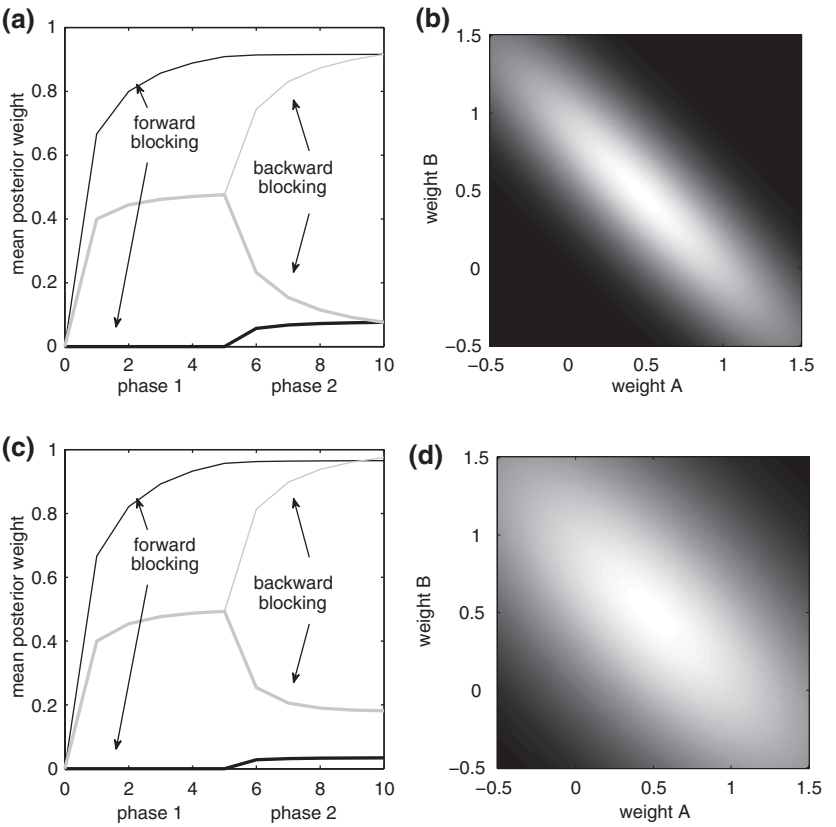

Fig. 19.1. Simulations of forward and backward blocking using Kalman filter model. a: Estimated mean weights $\hat{w}_A$ (upper thin lines) and $\hat{w}_B$ (lower thick lines) as a function of training in IID Kalman filter ($\sigma_d^2 = 0; \sigma_o^2 = 0.5$); the endpoints for forward and backward blocking are the same. b: Joint posterior distribution over w_A and w_B at start of phase 2 of backward blocking; the two weights are anticorrelated. c & d: Same as a & b, but using non-IID Kalman filter ($\sigma_d^2 = 0.1$); backward blocking is attenuated.

obvious *how* the rule accomplishes retrospective revaluation (Kakade and Dayan, 2001). Figure 19.1b illustrates the posterior distribution over w_A and w_B following $AB \rightarrow R$ training in backward blocking. The key point is that they are anticorrelated, since together they should add up to about R (1, in the simulations). Thus, if w_A is greater than $R/2$, then w_B must be less than $R/2$, and vice-versa. Subsequent $A \rightarrow R$ training indicates that w_A is indeed high and w_B must therefore be low, producing the effect. In terms of the Kalman filter learning rule, then, the key to backward blocking is the off-diagonal term in the covariance matrix Σ, which encodes the anticorrelation between w_A and w_B, and creates a *negative* Kalman gain κ_{tB} for stimulus B during the $A \rightarrow R$ trials in which B is not present (Kakade and Dayan, 2001).

Figure 19.1c shows the same experiments on the non-IID Kalman filter. Here, consistent with experiments, backward blocking is weaker than forward blocking. This happens because of the recency effect induced by the weight diffusion of Equation 6 – in particular (as illustrated in Figure 19.1d), the presumption that w_A and w_B are independently jittered between trials implies that they become less strongly anticorrelated with time. This suppresses retrospective revaluation. Forward blocking is not similarly impaired because the presumptive jitter on w_A is mean-preserving and does not therefore attenuate the belief (from $A \rightarrow R$ trials) that A is responsible for subsequent reinforcement on $AB \rightarrow R$ trials, in which $r_t - \hat{\mathbf{w}}_t \cdot \mathbf{x}_t = 0$.

3.2 Highlighting

The phenomenon of highlighting, which involves the base rates of outcomes, is rather more challenging for the Kalman filter (Kruschke, 2006). In this paradigm (see Table 19.1), three stimuli A,B,C are associated with two outcomes R and S according to $AB \rightarrow R$ and $AC \rightarrow S$. However, although equal numbers of both trial types are delivered, they are presented unevenly across the course of training, with $AB \rightarrow R$ predominating early (e.g. by a factor of three), and $AC \rightarrow S$ late (by the same factor).

The results show a mixture of what appear to be recency and primacy effects. In particular, tested on A after training, subjects predict R (the more common outcome in the first block); but tested on the novel combination BC, subjects predict S (the more common outcome in the second block of trials). Note that in the balanced form of the task presented here (Medin and Bettger, 1991; Kruschke, 2006) overall, B (and indeed A) is paired with R exactly as many times as C with S, so any asymmetry in the predictions must result from trial ordering.

These equalities imply that the IID model does not exhibit highlighting, a fact confirmed by the simulations in Figure 19.2a. To make the point in an extreme way, we assumed that $R = 1$ and $S = -R = -1$, so the two outcomes are in direct competition.[i] What may be less immediately obvious is that the non-IID model also fails to show

[i] This assumption may be more or less appropriate to particular empirical settings, depending for instance on whether the cover story and response requirements frame the outcomes as mutually exclusive. In any case, our models and arguments extend to the case with two nonexclusive outcomes.

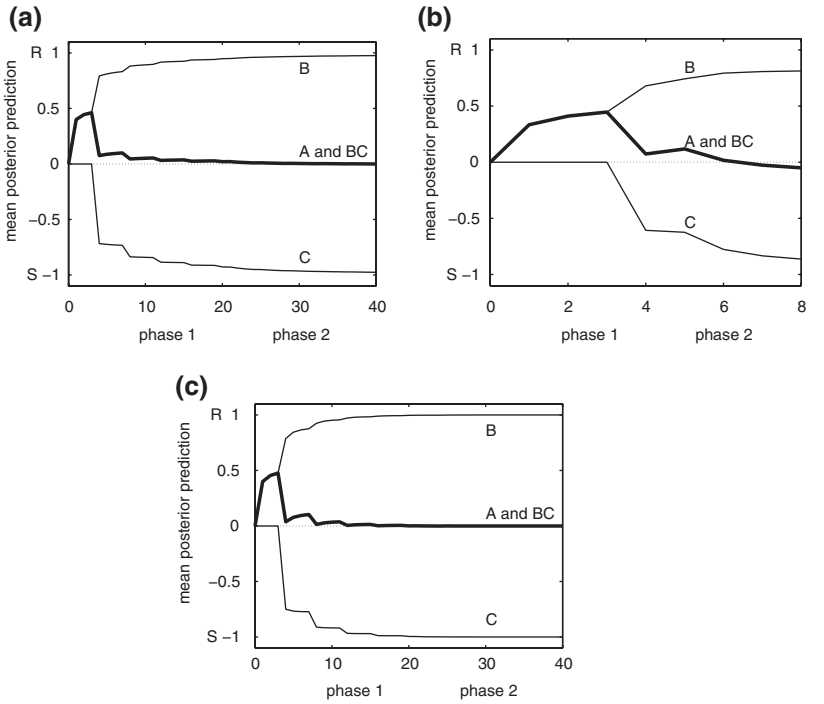

Fig. 19.2. Simulations of highlighting using the Kalman filter model. Development of mean estimates for test stimuli ($\hat{w}_A$ and $\hat{w}_B + \hat{w}_C$) illustrated as thick line (both are the same, see main text); $\hat{w}_B$ and $\hat{w}_C$ illustrated individually with thin lines. a: IID Kalman filter ($\sigma_d^2 = 0; \sigma_o^2 = 0.5$); b: non-IIDKalman filter ($\sigma_d^2 = 0.1$) shown with few trials and $\sigma_o^2 = 1.0$ to demonstrate pre-asymptotic behavior. c: Aymptotic behavior of non-IID Kalman filter with many trials ($\sigma_d^2 = 0.1; \sigma_o^2 = 0.5$) .

highlighting (Figure 19.2b). This can be seen in two ways; first the recency bias implies that both *A* and *BC* slightly favor *S*, the outcome predominantly received in the second block. Second, it is straightforward to verify that it will always be true that $\hat{w}_A = \hat{w}_B + \hat{w}_C$, for all parameters and after any sequence of the two trial types. Thus the model can never capture the pattern of results from highlighting, in which *A* and *BC* finish with opposite associations.

Further, given sufficient trials, even the non-IID Kalman filter will actually conclude that $\hat{w}_A = 0$, and use only $\hat{w}_B$ and $\hat{w}_C$ to predict *R* and *S* respectively – with all predictions balanced even though the base rates are locally skewed (Figure 19.2c). It is intuitive that the Kalman filter should be asymptotically insensitive to the base rates of stimuli, since it is attempting only to estimate the probability of *R conditional* on the stimuli having occurred, i.e. regardless of their base rate. The mechanism by which this occurs is again dependent on retrospective revaluation: initially, the Kalman filter attributes the predominance of *S* trials in the second block to both *A* and *C* (Figure 19.2b); given more experience, and through the medium of the

anticorrelation in the posterior between w_A and both w_B and w_C, it revalues A as wholly unpredictive and attributes all S to C (Figure 19.2b).

3.3 Summary

We have so far examined how trial ordering effects arise naturally in a simple Bayesian model. Because they follow from assumptions about change, these generally involve some sort of recency effect, though this can be manifest in a fairly task-dependent manner.

Backward blocking is a straightforward consequence of the generative model underlying the non-IID Kalman filter. The pattern of results from highlighting is not: Quite uncharacteristic for inference in a changing environment, the latter seem to involve in part a primacy effect for $A \rightarrow R$.

One noteworthy aspect of these investigations is the importance of retrospective revaluation to both experiments. Backward blocking, of course, is itself a retrospective revaluation phenomenon; however, that it is weaker than forward blocking indicates that the revaluation is less than perfect. Similarly, one feature of highlighting is the failure retrospectively to determine that stimulus A is unpredictive, after it had been initially preferentially paired with R. (This is particularly clear in a version of highlighting discussed by Kruschke, 2003, 2006, which starts just like backward blocking with a block of only $AB \rightarrow R$ trials.) Retrospective revaluation is closely tied to Bayesian reasoning, in that it typically seems to involve reasoning about the whole distribution of possible explanations (as in Figure 19.1d), rather than just a particular estimate (as in the Rescorla–Wagner model, which fails to produce such effects).

As we discussed in the introduction, there are at least three strategies to follow in the face of the failure of this simple model to account for highlighting. The first is to downplay the emphasis on principled reasoning and seek a heuristic explanation (Kruschke, 2006). The second is to consider it as a failure of the generative model and to seek a more sophisticated generative model that perhaps better captures subjects' beliefs about the task contingencies. While there are doubtless exotic beliefs about change processes and cue-combination rules that would together give rise to highlighting, we have not so far discovered a completely convincing candidate. Instead, we would suggest that the Kalman filter's behavior is characteristic of inference in a changing environment more generally. As we have seen, trial order sensitivities in Bayesian reasoning ultimately arise from the a priori belief that trials are not identically distributed. A reasonable general assumption is that trials nearer in time are more similar to one another than to those farther away – predicting, all else equal, a recency bias. Together with the fact that failures of retrospective revaluation are characteristic of a number of well-founded inferential approximation strategies, as we discuss below, this observation motivates the third approach: to consider that the phenomenon actually arises from a failure of the brain to implement correct inference.

4 Approximate Inference

In the face of generative models that are much more complicated and less tractable than that in Equations 2, 5, and 6, statisticians and computer scientists have

developed a menagerie of approximate methods. Such approximations are attractive as psychological models because they offer plausible mechanistic accounts while maintaining the chief advantage of Bayesian approaches: *viz* a clear grounding in normative principles of reasoning.

Tools for inferential approximation may crudely be split into two categories, though these are often employed together. Monte Carlo techniques such as particle filtering (e.g. Doucet *et al.*, 2000) approximate statistical computations by averaging over random samples. While these methods may be relevant to psychological modeling, the hallmarks of their usage would mainly be evident in patterns of variability over trials or subjects, which is not the focus of the present work. We will focus instead on deterministic simplifications of difficult mathematical forms (e.g. Jordan *et al.*, 1999), such as the usage of lower bounds or maximum likelihood approximations. One critical feature of these approximations is that they often involve steps that have the consequence of discarding relevant information about past trials. This can introduce trial order dependencies, and particularly effects similar to primacy. In this section, we will demonstrate some simple examples of this.

4.1 Assumed Density Filtering

The Kalman filter (Equations 7, 8) updates its beliefs recursively: the new belief distribution is a function only of the previous distribution and the new observation. In many cases, we may wish to maintain this convenient, recursive form, but simplify the posterior distribution after each update to enable efficient approximate computation of subsequent updates. Such methods are broadly known as *assumed density* filters (see Minka, 2001, who also discusses issues of trial ordering). Typically, the posterior distribution is chosen to have a simple functional form (e.g. Gaussian, with a diagonal covariance matrix), and to have its parameters chosen to minimize a measure (usually the so-called Kullback–Liebler divergence) of the discrepancy between it and the best guess at the true posterior. Because of this minimization step, this approximation is sometimes called *variational* (*Jordan et al.*, 1999).

Clearly such an approximation introduces error. Most critical for us is that these errors can be manifest as trial ordering effects. In the Kalman filter update, the previous belief distribution can stand in for all previous observations because the posterior distribution is a *sufficient statistic* for the previous observations. The recursively computed posterior equals the posterior conditioned on all the data, and so for instance the IID filter (the Kalman filter with $\sigma_d = 0$) can correctly arrive at the same answer no matter in what order trials are presented. In backward blocking, for instance, $\hat{w}_B$ is retrospectively revalued on $A \rightarrow R$ trials without explicitly backtracking or reconsidering the previous $AB \rightarrow R$ observations: the posterior covariance Σ summarizes the relevant relationship between the variables. A simplified form of the posterior will not, in general, be a sufficient statistic; how past trials impact the posterior may then depend on the order they arrived in, even in cases (e.g. the IID filter) for which the exact solution is order-invariant. This can disrupt retrospective revaluation, since the ability to reinterpret past experience depends on its being adequately represented in the posterior.

4.1.1 Simulations

Perhaps the most common assumed density is one in which the full posterior factorizes. Here, this implies assuming the joint distribution over the weights is separable into the product of a distribution over each weight individually. For the Kalman filter, this amounts to approximating the full Kalman filter covariance matrix Σ by just its diagonal entries,[ii] thus maintaining uncertainty about each weight but neglecting information about their covariance relationships with one another.

Since we have already identified the covariance terms as responsible for backward blocking we may conclude immediately (and simulations, not illustrated, verify) that this simplification eliminates backward blocking while retaining forward blocking.

A subtler trial order dependence also arises in the form of a robust highlighting effect (Figure 19.3a). This traces to three interlinked features of the model.

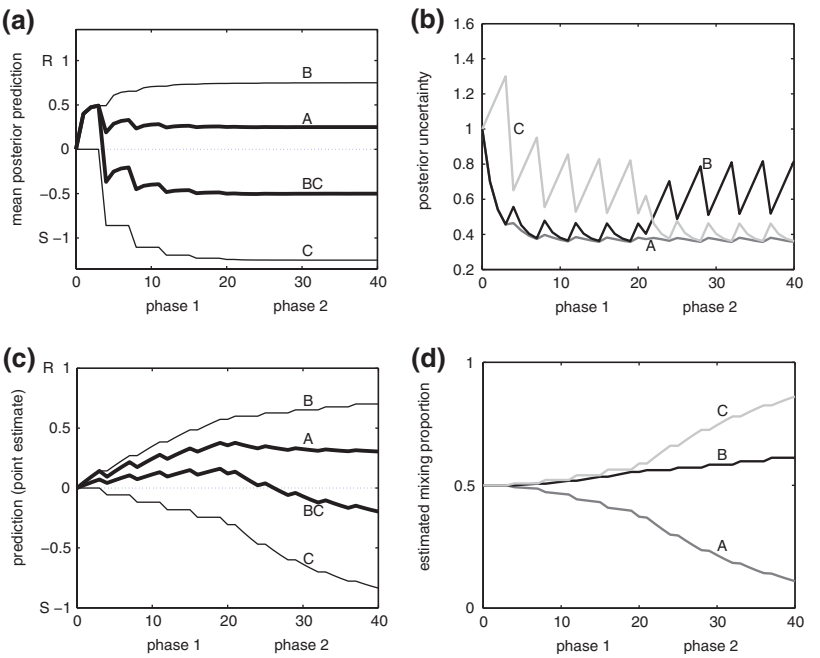

Fig. 19.3. Simulations of highlighting in approximate Bayesian reasoning. (a) Highlighting in Kalman filter with diagonalized assumed covariance ($\sigma_d^2 = 0.1; \sigma_o^2 = 0.5$). (b) Posterior uncertainty about each weight as a function of training; $\hat{w}_C$ is more uncertain in first phase, highlighting it. (c) Highlighting in additive Gaussian mixture model using EM (η= 0.1; illustrated are the point estimates of the three individual weights and, for compound *BC*, the net outcome expectation weighted by mixing proportions). (d) Development of estimates of mixing proportions with training; *C* is highlighted.

[ii] This minimizes the KL-divergence from the full covariance Gaussian among the class of all diagonal distributions.

First, much like attentional variables in associative accounts of highlighting (Kruschke, 2003), in the Kalman filter, the uncertainties about the weights (the diagonal elements of the covariance matrix) control the rate of learning about those weights (Dayan and Long; Kakade and Dayan, 2002). More uncertain weights get a larger gain κ and a bigger update from Equation 7; when stimuli are observed, the uncertainties about them (and subsequent learning rates) decline, whereas when stimuli are unobserved, their uncertainties increase. This means (Figure 19.3b) that on $AC \rightarrow S$ trials in the first block, C (which is rarely observed) learns more rapidly than A (which is commonly observed). Conversely, A is likely to be paired with R the first few times it is seen, when its learning rate is highest. Second, the weights of presented stimuli must interact additively to produce the outcome (Equation 2).A's association with R will therefore reduce B's association with it (since the two weights must additively share the prediction on $AB \rightarrow R$ trials), whereas C's association with S will be correspondingly enhanced by additionally having to cancel out A's opposing prediction of R. Finally, since the covariance is not represented, A is never retrospectively revalued as a nonpredictor – its initial association with R instead persists indefinitely as a primacy effect. There is a continuum of possible values $\hat{w}_A$, $\hat{w}_B$ and $\hat{w}_C$ that together add up to explain exactly the results of both trial types (specifically $\hat{w}_B = 1 - \hat{w}_A$ and $\hat{w}_C = -1 - \hat{w}_A$ for any $\hat{w}_A$); lacking revaluation, this model sticks with the first one it finds.

Note also that the venerable Rescorla–Wagner model results from one further simplification over this one: the assumption that the learning rates are simply constant (if perhaps stimulus-dependent), i.e. that the uncertainties are never updated. This is motivated by the fact that, under special circumstances, for instance, if each stimulus is presented on every trial, the Kalman filter ultimately evolves to a fixed asymptotic learning rate (the value at which information from each observation exactly balances out diffusion in the prior). However, from the perspective of highlighting, this is a simplification too far, since without dynamic learning rates C's learning about S is not boosted relative to A. Similar to the full Kalman filter, no highlighting effect is seen: the local base rates predominate. The lack of covariance information also prevents it from exhibiting backward blocking.

4.2 Maximum Likelihood and Expectation Maximization

Similar phenomena can be seen, for similar reasons, in other approximation schemes. We exemplify this generality using a representative member of another class of inexact but statistically grounded approaches, namely those that attempt to determine just a maximum likelihood point-estimate of the relevant variables (here the means $\hat{\mathbf{w}}$) rather than a full posterior distribution over them. Often, these methods learn using some sort of hill climbing or gradient approach, and, of course, it is not possible to employ Bayes' theorem directly without representing some form of a distribution. As for assumed density filters (and for the Rescorla–Wagner model, which can also be interpreted as a maximum-likelihood gradient climber), the failure to maintain an adequate posterior distribution curtails or abolishes retrospective revaluation.

Another example is a particular learning algorithm for the competitive mixture of Gaussians model of Equation 3. We develop this in some detail as it is the canonical

example of a particularly relevant learning algorithm and is related to a number of important behavioral models (Kruschke, 2001; Mackintosh, 1975), though it does have some empirical shortcomings related to its assumptions about cue combination (Dayan and Long, 1998). Recall that, according to this generative model, one stimulus out of those presented will be chosen and the outcome will then be determined solely based on the chosen stimulus' weight. Learning about the weight from the outcome then depends on unobserved information: which stimulus was chosen. *Expectation-maximization* methods (Dempster *et al.*, 1977; Griffiths and Yuille, this volume) address this problem by repeatedly alternating two steps: estimating the hidden information based on the current beliefs about the weights ('E step'), then updating the weights assuming this estimate to be true ('M step'). This process can be understood to perform coordinate ascent on a particular error function, and is guaranteed to reduce (or at least not increase) the error at each step (Neal and Hinton, 1998).

An online form of EM is appropriate for learning in the generative model of Equation 3. (Assume for simplicity that the weights do not change, i.e., that Equation 4 obtains.) At each trial, the E step determines the probability that the outcome was produced by each stimulus (or the background stimulus 0), which involves a Bayesian inversion of the generative model:

$$q_{tj} \propto x_{tj}\pi_{tj} \exp(-(r_t - w_{tj})^2/\sigma_0)$$

Here, the background responsibility $\hat{\pi}_{t0} = \max(1 - \hat{\boldsymbol{\pi}}_t \cdot \mathbf{x}_t, 0)$ and the constant of proportionality in q_t arrange for appropriate normalization. The model then learns a new point estimate of the weights $\hat{\mathbf{w}}$ and the mixing proportions $\hat{\pi}$ using what is known as a partial M step, with the predictions associated with each stimulus changing according to their own prediction error, but by an amount that depends on the responsibilities accorded to each during the E step:

$$\hat{w}_{t+1,j} = \hat{w}_{tj} + \eta q_{tj}(r_t - \hat{w}_{tj}) \tag{9}$$

$$\hat{\pi}_{t+1,j} = \hat{\pi}_{tj} + \eta x_{tj}(q_{tj} - \hat{\pi}_{tj}) \tag{10}$$

Here, η is a learning rate parameter, which, as for the Rescorla–Wagner rule, by being fixed, can be seen as a severe form of approximation to the case of continual change in the world.

4.2.1 Simulations

Like most other more or less local hill climbing methods, the fact that the M-step in this algorithm is based on the previous, particular, parameter settings (through the medium of the E-step) implies that there are trial order effects akin to primacy. As Figure 19.3c illustrates, these include highlighting, which arises here because the responsibilities (and the estimated mixing variables $\hat{\boldsymbol{\pi}}$ that determine and are determined by them) take on an attentional role similar to the uncertainties in the diagonalized Kalman filter account. In particular, *A* and *B* share responsibility *q* (Figure 19.3d) for the preponderance of *R* in the first block. This reduces the extent to which $\hat{w}_B$ learns about *R* (since the effective learning rate is ηq_B), and the extent to which *B*

contributes to the aggregate prediction during the *BC* probe (since *B*'s contribution to the expectation is proportional to $\hat{\pi}_B$). Meanwhile, by the second block of trials, the model has learned that *A* has little responsibility ($\hat{\pi}_A$ is low), giving a comparative boost to learning about $\hat{w}_C$ during the now-frequent *S* trials. Less learning about $\hat{w}_A$ due to its lower responsibility also means its association with *R* persists as a primacy effect. These biases follow from the way the evolving beliefs about the stimuli participate in the approximate learning rule through the determination of responsibilities – recall that we derived the model using the IID assumption and that *optimal* inference is, therefore, trial order independent.

Note that unlike the diagonal Kalman filter example of Figure 19.3a, the highlighting effect here doesn't arise until the second block of trials. This means that this EM model doesn't explain the 'inverse base rate effect,' (Medin and Edelson, 1988) which is the highlighting effect shown even using only the first block when *R* predominates. One reason for this, in turn, is the key competitive feature of this rule, that the predictions made by each stimulus do not interact additively in the generative rule (Equation 3). Because of this, while stimuli may share *responsibility* for the outcome, the net prediction doesn't otherwise enter into the learning rule of Equation 9, which still seeks to make each stimulus account for the whole outcome on its own. In highlighting, this means *A*'s association with *R* cannot directly boost *C*'s association with *S* during the first phase. The same feature also causes problems for this model (and its associative cousins) explaining other phenomena such as overshadowing and inhibitory conditioning, and ultimately favors alternatives to Equation 3 in which cues cooperate to produce the net observation (Dayan and Long, 1998; Hinton, 1999; Jacobs *et al.*, 1991a).

Despite this failure, the competitive model does exhibit forward blocking, albeit through a responsibility-sharing mechanism (Mackintosh, 1975) rather than a weight-sharing mechanism like Rescorla–Wagner (simulations not illustrated). More concretely, stimulus *A* claims responsibility for *R* on $AB \rightarrow R$ trials, due to already predicting it. This retards learning about *B*, as in blocking. However, given that it is based only on a point-estimate, it retains no covariance information, and so, like Rescorla–Wagner, cannot account for backward blocking.

4.3 **Summary**

We have shown how two rather different sorts of inferential approximation schemes, in the context of two different generative models for conditioning, both disrupt retrospective revaluation – abolishing backward blocking and producing a highlighting effect. Exact Bayesian reasoning is characterized by simultaneously maintaining the correct likelihood for every possible hypothesis. This is what enables retrospectively revisiting previously disfavored hypotheses when new data arrive, but it is also the main source of computational complexity in Bayesian reasoning and the target for simplification schemes. In short, primacy effects – the failure retrospectively to discount an initially favored hypothesis – are closely linked to inferential approximation.

We have used extreme approximations to expose this point as clearly as possible. While the experiments discussed here both suggest that retrospective revaluation is

attenuated, that effects like backward blocking exist at all rules out such extreme approaches. In the following section, we consider gentler approximations.

5 Blended and Mixed Approximations

So far, neither the Bayesian nor the approximately Bayesian models actually exhibit the combination of recency and primacy evident in backward blocking and highlighting. In fact, this is not hard to achieve, as there are a number of models that naturally lie between the extremes discussed in Sections 3 and 4. The cost is one additional parameter. In this section, we provide two examples, based on different ideas associated with approximating the Kalman filter.

5.1 Reduced-rank Approximations

In Section 4.1, we considered the simplest possible assumed density version of the Kalman filter, in which the posterior fully factorized, having a diagonal covariance matrix (Figure 19.3a). This method fails to exhibit backwards blocking, since it cannot represent the necessary anticorrelation between the predictions of the two CSs that arises during the first set of learning trials.

A less severe approximation to the posterior is to use a reduced-rank covariance matrix. We use one that attempts to stay close to the *inverse* covariance matrix, which (roughly speaking) characterizes certainty. An approximation of this sort allows the subject to carry less information between trials (because of the reduction in rank), and can also enable simplification of the matrix calculations for the subsequent update to the Kalman filter (Treebushny and Madsen, 2005).

More precisely, we approximate the inverse posterior covariance after one trial, $(\mathbf{\Sigma}_t - \kappa_t \mathbf{x}_t \mathbf{\Sigma}_t)^{-1}$, by retaining only those n basis vectors from its singular value decomposition that have the highest singular values, thereby minimizing the Frobenius norm of the difference. On the next trial we reconstruct the covariance as the pseudo-inverse of the rank-n matrix plus the uncertainty contributed by the intervening drift, $\sigma_d^2 \mathbf{I}$.

Figure 19.4a,b shows the consequence of using a rank-2 approximation ($n = 2$) to the covariance matrix. This results in highlighting without further disrupting backward blocking, which, in any case, only requires a two-dimensional posterior. A general prediction of this sort of resource bottleneck approach is that the effects of approximation should become more pronounced – e.g. retrospective revaluation more attenuated – for problems involving higher dimensional and more intricately structured posteriors.

5.2 Mixing Filters

A different possibility is to mix the exact and diagonally approximated Kalman filters more directly. Here the idea is that there may be mixing at the behavioral level of distinct underlying psychological and/or neural processes, one corresponding to each model. In some circumstances – for instance, when the diagonal elements of the covariance matrix are anyway small – the additional accuracy to be gained by maintaining the full covariance matrix may not justify the additional energetic costs relative to the particularly simple diagonal version. Such considerations suggest that the

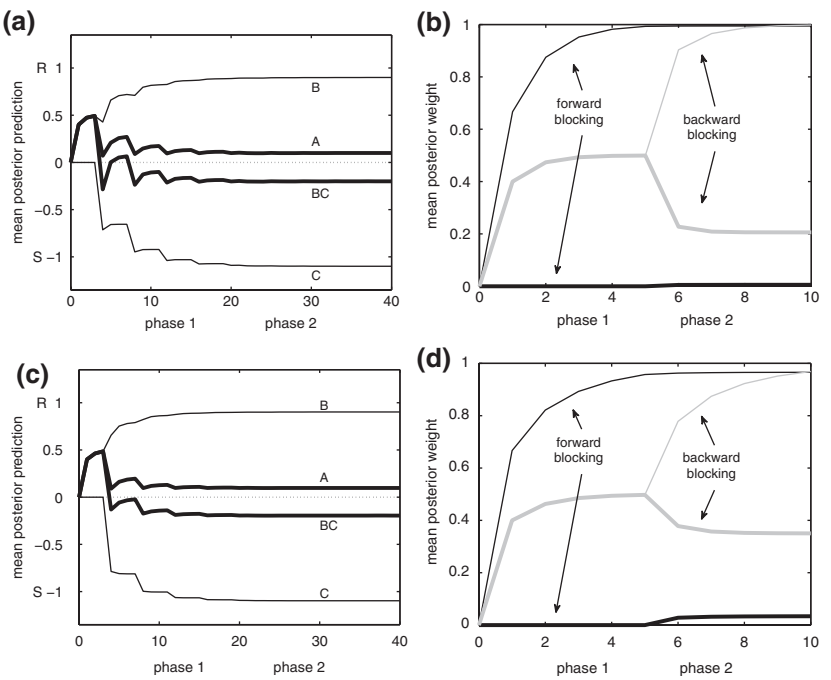

Fig. 19.4. Simulations of two approximate Bayesian models exhibiting highlighting and backwards blocking. (a,b) Reduced-rank covariance Kalman filter $(\sigma_d^2 = 0.1; \sigma_o^2 = 0.5; n = 2)$. (c,d) Blended full/diagonal covariance Kalman filter $(\sigma_d^2 = 0.1; \sigma_o^2 = 0.1)$.

brain could adaptively trade off whether to employ approximation based on a sort of meta-rational cost-benefit analysis. (In this case, blending would appear in results via the average over trials or subjects.) A slightly different version of this idea would suggest that subjects actually compute both forms simultaneously, but then reconcile the answers, making an adaptive decision how much to trust each, much as in other cases of Bayesian evidence reconcilation (Daw *et al.*, 2005). The 'exact' computation might not always be the most accurate, if in biological tissue the extra computations incur additional computational noise; it might therefore be worthwhile to expend *extra* resources also computing a less noisy approximation.

Figure 19.2c,d shows simulations of a model which performs mixing by multiplying the off-diagonal elements of the covariance $\mathbf{\Sigma}$ by 0.7 at each step. This restricts the efficacy of retrospective revaluation without totally preventing it, allowing both backward blocking, which is curtailed relative to forward blocking, and highlighting.

6 Discussion

6.1 Summary

In this chapter, we have focused on the intricacies of inference in Bayesian models of conditioning. We used theory and simulations to show how particular classes of effects

in learning (e.g. backward blocking) can arise from optimal inference in the light of a simple generative model of a task, and others (e.g. highlighting) from more or less extreme, but still recognizable approximations to optimal inference. This work on sensitivity to trial order is clearly only in its infancy, and the data to decide between and refine the various different models are rather sparse. However, just as we try to differentiate subjects' assumptions in exact Bayesian modeling, we hope in the future to adjudicate more definitively between different approximation methods by identifying tasks that better expose their fingerprints. Here, we have focused on trial order, but many similar issues arise in other areas of conditioning, such as stimulus competition.

6.2 Locally Bayesian Learning

One spur to study trial order was the recent article by Kruschke (2006). He pointed out the apparent contradiction between the recency in backward blocking and the primacy in highlighting, and noted the implications of the IID assumption for both phenomena. Kruschke framed these findings by contrasting classic associative learning models (which explain highlighting via ideas about stimulus attention) with a particular IID Bayesian model (which explains retrospective revaluation). Rather than addressing the IID assumption (which, of course, Bayesian models need not make), he proposed a 'locally Bayesian' model blending features of both of these approaches. This model consists of interconnected modules that are Bayesian-inspired in that each updates a local belief distribution using Bayes' rule, but heuristic in that the 'observations' to which Bayes' rule is applied are not observed data but instead synthetic quantities constructed using an ad-hoc message-passing scheme. Although the individual modules treat their synthetic data as IID, trial ordering effects emerge from their interactions. The theoretical status of the heuristic, for instance as a particular form of approximation to a well-found statistical procedure, is left unclear.

We have attempted to address the issues central to highlighting and backwards blocking in unambiguously Bayesian terms. We develop a similar contrast between exact and approximate approaches, but rather than seeing statistical and associative learning as contradictory and requiring reconciliation, we have stressed their connection under a broader Bayesian umbrella. The approximate Kalman filters discussed in Section 5 retain a precise flavor of the optimal solutions, while offering parameterized routes to account for the qualitative characteristics of both backwards blocking and highlighting.

It is also possible to extend this broader Bayesian analysis to the mixture model of Section 4.2, and hence nearer to Kruschke's (2006) locally Bayesian scheme. The mixture model fails to exhibit retrospective revaluation since it propagates only a point, maximum likelihood, estimate of the posterior distribution over the weights. This could be rectified by adopting a so-called *ensemble learning* approach (Hinton and van Camp, 1993; Waterhouse *et al.*, 1996), in which a full (approximate) distribution over the learned parameters is maintained and propagated, rather than just a point estimate. In ensemble learning, this distribution is improved by iterative ascent (analogous to E and M steps) rather than direct application of Bayes' rule.

One online version of such a rule could take the form of inferring the unobserved responsibilities, and then conditioning on them as though they were observed data

(see also the mixture update of Dearden *et al.* 1998). Since it conducts inference using synthetic in place of observed quantities, this rule would have the flavor of Kruschke's locally Bayesian scheme, and indeed would be a route to find statistically justifiable principles for his model. However, this line of reasoning suggests one key modification to his model, that the unobserved quantities should be estimated optimally from the statistical model using an E step, obviating the need for a target propagation scheme.

6.3 Bayes, Damn Bayes, and Approximations

At its core, the Bayesian program in psychology is about understanding subjects' behavior in terms of principles of rational inference. This approach extends directly beyond ideal computation in the relatively small set of tractably computable models into approximate reasoning in richer models. Of course, we cannot interrogate evolution to find out whether some observable facet of conditioning arises as exact inference in a model that is a sophisticated adaptation to a characteristic of the learning environment that we have not been clever enough to figure out, or as an inevitable approximation to inference in what is likely to be a simpler model. Nevertheless, admitting well found approximations does not infinitely enlarge the family of candidate models, and Occam's razor may continue to guide.

Waldmann *et al.* (2007; this volume) pose another version of our dilemma. They agree with Churchland (1986) that the top-down spirit of Marrian modeling is always violated in practice, with practitioners taking peeks at algorithmic (psychological) and even implementational (neural) results before building their abstract, computational accounts. However, unlike the approach that we have tried to follow, their solution is to posit the notion of a minimal rational model that more explicitly elevates algorithmic issues into the computational level.

We see two critical dangers in the Waldmann 'minimal rationality' programme, one associated with each of the two words. One danger Marr himself might have worried about, namely the fact that minimality is in the eye of the beholder (or at least the instruction set), and that our lack of a justifiable account of the costs of neural processing makes any notion of minimality risk vacuity. The second danger is that by blending normative considerations with incommensurate pragmatic ones, minimal rationality risks being a contradiction in terms. We agree with Waldmann and colleagues' criticism that rational theorists have sometimes been a bit glib relating theories of competence to performance, but we see the solution in taking this distinction more seriously rather than making it murky. Since computational and algorithmic levels involve fundamentally different questions (e.g. why versus how), we suggest preserving the innocence of the computational account, and focusing on approximations at the algorithmic level.

Finally, as we saw in Section 5.2, one important facet of approximate methods is that it is frequently appropriate to maintain multiple different approximations, each of which is appropriate in particular circumstances, and to switch between or blend their outputs. To the extent that different approximations lead to different behavior, it will be possible to diagnose and understand them and the tradeoffs that they (locally) optimize. Our understanding of the blending and switching process is less advanced.

In the present setting, the idea goes back at least to Konorski (1967) that Pavlovian learning can employ both a stimulus-stimulus pathway (which is more cognitive in this respect and echoes our full Kalman filter's representation of interstimulus covariance) and a simpler stimulus-reward one (perhaps related to our diagonalized Kalman filter); such processes also appear to be neurally distinguishable (Balleine and Killcross, 2006). In fact, there is evidence for similar behavioral dissociations coming from attempts to demonstrate retrospective revaluation in rats (Miller and Matute, 1996). When training is conducted directly in terms of stimulus-reinforcer pairings, no retrospective revaluation is generally seen (as with our diagonalized covariance Kalman filter), but revaluation does succeed in the more obviously cognitive case in which the paradigms are conducted entirely in terms of pairings between affectively neutral stimuli, one of which (standing in for the reinforcer) is finally associated with reinforcement before the test phase.

Parallel to this in the context of instrumental conditioning is an analogous division between an elaborate, cognitive, (and likely computationally noisy) 'goal-directed' pathway, and a simpler (but statistically inefficient) 'habitual' one (Dickinson and Balleine, 2002). In this setting, the idea of normatively trading off approximate value-inference approaches characteristic of the systems has been formalized in terms of their respective uncertainties, and explains a wealth of data about what circumstances favor the dominance of goal-directed or habitual processes (Daw *et al.*, 2005). It would be interesting to explore similar estimates of uncertainty in the mixed Kalman filters and thereby gain normative traction on the mixing.

References

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: An integrated view of amygdala function. *Trends in Neurosciences*, **29**, 272–279

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. MIT Press.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind-brain*. Cambridge, MA: MIT Press

Courville, A. C., Daw, N. D., Gordon, G. J., & D. S. Touretzky (2003). Model uncertainty in classical conditioning. In *Advances in Neural Information Processing Systems* (Vol. 16). Cambridge, MA: MIT Press.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2004). Similarity and discrimination in classical conditioning: A latent variable account. In *Advances in Neural Information Processing Systems* (Vol. 17). Cambridge, MA: MIT Press.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, **10**, 294–300.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, **8**, 1704–1711.

Daw, N. D., O'Doherty, J. P., Seymour, B., Dayan, P., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, **441**, 876–879.

Dayan, P., & Long, T. (1998). Statistical models of conditioning. *Advances in Neural Information Processing Systems*, **10**, 117–123.

Dayan, P., Kakade, S., & Montague P. R. (2000). Learning and selective attention. *Nature Neuroscience*, **3**, 1218–1223.

Dearden, R., Friedman N., & Russell S. J. (1998). Bayesian Q-learning. In *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI)*, PP. 761–768.

Dempster A. P., Laird N. M., & Rubin D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, **39**, 1–38.

Dickinson, A., & Balleine, B. (2002). The role of learning in motivation. In C. R. Gallistel (Ed.), *Stevens' handbook of experimental psychology Vol. 3: Learning, Motivation and Emotion* (3rd ed., pp. 497–533). New York: Wiley.

Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, **10**, 197–208.

Griffiths, T. L., & Yuille, A. (2007). Technical introduction: A primer on probabilistic inference. 2007. (this volume).

Hinton, G. (1999). Products of experts. In *Proceedings of the Ninth International Conference on Artificial Neural Networks (ICANN99)*, pp. 1–6.

Hinton, G., & van Camp, D. (1993). Keeping neural networks simple by minimizing the description length of the weights. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pp. 5–13.

Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991a). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science*, **15**, 219–250.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991b). Adaptive mixtures of local experts. *Neural Computation*, **3**, 79–87.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1999). An introduction to variational methods for graphical models. *Machine Learning*, **37**, 183–233.

Kakade, S., & Dayan, P. (2001). Explaining away in weight space. In *Advances in Neural Information Processing Systems* (Vol. 13). Cambridge, MA: MIT Press.

Kakade, S., & Dayan, P. (2002). Acquisition and extinction in autoshaping. *Psychological Review*, **109**, 533–544.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, **82**, 35–45.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 242–259). New York: Appleton-Century-Crofts.

Konorski, J. (1967) *Integrative activity of the brain*. University of Chicago Press: Chicago.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, **45**, 812–863.

Kruschke, J. K. (2006). Locally Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review*, **113**, 677–699.

Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science*, **5**, 171–175.

Lovibond, P. F., Been, S.-L., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory and Cognition*, **31**, 133–142.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 532–552.

Marr, D. (1982). *Vision: A computational approach*. San Francisco, CA: Freeman and Co.

Medin, D. L., & Bettger, J. G. (1991). Sensitivity to changes in base-rate information. *American Journal of Psychology*, **40**, 175–188.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General*, **117**, 68–85.

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General*, **125**, 370–386.

Minka, T. (2001). *A family of algorithms for approximate Bayesian inference*. PhD thesis, Massachusetts Institute of Technology.

Neal, R. M., & Hinton, G. E. (1998). A view of the EM algorithm that justifies incremental, sparse and other variants. In M. I. Jordan (Ed.), *Learning in graphical models* (pp. 355–368). Kluwer Academic Publishers.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning, 2: Current research and theory* (pp. 64–69). New York: Appleton Century-Crofts.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, **37B**, 1–21.

Simon, H. (1957). A behavioral model of rational choice. In *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*. Wiley, New York.

Treebushny, D., & Madsen, H. (2005). On the construction of a reduced rank square-root Kalman filter for efficient uncertainty propagation. *Future Generation Computer Systems*, **21**, 1047–1055.

Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2007). Causal learning in rats and humans: A minimal rational model (this volume).

Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, **51B**, 121–138.

Waterhouse, S., MacKay, D., & Robinson, T. (1996). Bayesian methods for mixtures of experts. *Advances in Neural Information Processing Systems*, **8**, 351–357.

Yu, A. J., & Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. In *Advances in Neural Information Processing Systems* (Vol. 15). Cambridge, MA: MIT Press.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, **46**, 681– 692.