Learning to Represent a Multi-Context Environment: More than Detecting Changes

Ting Qian[*,1], T. Florian Jaeger[1,2], Richard N. Aslin[1]

[1]Department of Brain and Cognitive Sciences
[2]Department of Computer Science
University of Rochester, Rochester, NY, USA

Running title:
Learning in multi-context environment

Correspondence:
Ting Qian
Department of Brain and Cognitive Sciences
Meliora Hall, Box 270268
University of Rochester
Rochester, NY 14627-0268 USA
tqian@bcs.rochester.edu

Abstract

Learning an accurate representation of the environment is a difficult task for both animals and humans, because the causal structures of the environment are unobservable and must be inferred from the observable input. In this article, we argue that this difficulty is further increased by the multi-context nature of realistic learning environments. When the environment undergoes a change in context without explicit cueing, the learner must detect the change and employ a new causal model to predict upcoming observations correctly. We discuss the problems and strategies that a rational learner might adopt and existing findings that support such strategies. We advocate hierarchical models as an optimal structure for retaining causal models learned in past contexts, thereby avoiding relearning familiar contexts in the future.

Keywords: multi-context environment; contextual ambiguity; representation learning; contextual cue; change detection

1 Introduction

Learning requires a mechanism that infers from observable events in the environment a minimally sufficient hypothesis of the unobservable underlying structures. This hypothesis not only serves as an efficient representation of the causal relations in the environment, at least for a particular task, but also enables the learner to generalize to events that have not been observed. For example, if the task involves predicting the consumption of different food items in a school cafeteria, then a reasonable approximation is to tally the quantity of each food item that was consumed over some running average of the past (e.g., the prior month). However, there is considerable variation in these tallies across hours of the day, days of the week, and specific occasions such as holidays. Thus, in order to prevent more than the occasional dissatisfied customer, the manager of the cafeteria must develop a fairly flexible model that can modulate its predictions of the demand for food items dynamically given the values of these key variables. We will refer to these key variables as *contexts* and the cafeteria environment as an example of a *multi-context* environment. Each context in such an environment is typically associated with a distinctive causal structure. In the present article, we argue that most realistic environments are inherently multi-context, and that learning a flexible model that embeds information about contexts is the general task that confronts naive learners. To successfully accomplish this task, learners must be able to 1) infer (with uncertainty) whether a context change has occurred; 2) adapt to a changed context and learn new causal models if necessary; and 3) represent contexts along with corresponding causal models in an optimal manner. We discuss each of these three aspects next.

Context changes often signal that a different underlying causal model now applies. However, contexts are rarely explicitly labeled in the input available to the learner, and many contextual cues that are easily observable are not relevant to the underlying causal model. The canonical case, then, involves implicit contexts that must be discerned by the learner, often by noting that the current causal model does not provide an adequate fit with the most recent input. Thus, the first challenge of learning in a multi-context environment is to detect context changes from unexpected observations alone. This would be a trivial problem if the causal relations within each context were strictly deterministic. Consider the cafeteria example again. If the consumption rate of bottled milk during breakfast hours is exactly 10 bottles per minute, it is not difficult to conclude that breakfast is over when the rate drops to 1 bottle per minute. However, such deterministic relations are rare in reality. It is possible that the *average* consumption rate of bottled milk is 10 bottles per minute during the BREAKFAST context, but occasionally, it might be as low as 2 bottles or as high as 20. The uncertainty resulting from random and probabilistic variations creates a difficult situation for the manager: if a large lecture class, originally scheduled at 9 A.M., is cancelled because the professor's return flight from a conference is delayed by bad weather, then the demand for milk at the cafeteria may be altered idiosyncratically - the manager may observe a decrease as students are likely to get up later and skip breakfast. Unaware of the implicit context (i.e. CLASS CANCELLED), the manager is now faced with the problem of *contextual ambiguity*: should the manager interpret this decrease as acceptable random variations in the regular BREAKFAST context or as the representative characteristic of a changed context?

Resolving contextual ambiguity is only the first step of learning in a multi-context environment. Once a learner arrives at the conclusion that a different context has come into effect, they must also decide how to adapt to the changed context. Here, a learner at least two choices. They can either learn a new model and associate it with the context, or retrieve from memory a causal model learned for a past context, which closely resembles or even matches the current context. The need to learn a new causal model arises when the learner encounters a novel context. Consider a *new* manager of a school cafeteria. Although the new manager may draw upon her experience of working in a cafeteria at a different university, there remains the possibility of encountering novel contexts on the current campus. For example, students at the current university may prefer sleeping in over attending classes on Friday mornings, which would require reduced stocking of bottled milk on those days. Like a naïve learner in any task, the new manager not only has to learn the average quantity of milk to stock (i.e. the model), but also has to associate it with Friday mornings (i.e. the appropriate context). The difficulty lies in the fact that there are often no explicit cues for the manager to gain sudden insight into *what the appropriate context is*: Instead of using FRIDAY MORNING, the manager could just as easily consider the weather on that particular day. The benefits of identifying the appropriate contexts, on the other hand, also extend to the second choice of adapting to the change in context: reusing a learned model. If the learner has correctly associated the causal model (e.g., decreased demand for bottled milk) with the relevant context (e.g. FRIDAY MORNING), then, in theory, they will be able to retrieve and reinstate the model when the target context is effective again (e.g. next Friday).

Assuming that the learner has the ability to reinstate a previously learned causal model, does it mean that the learner must be capable of storing and representing multiple contexts simultaneously? Although intuitively, the answer to this question has to be a strong "yes" (since learning a new causal model should not lead to elimination of the an old one), it is not immediately transparent how these multiple contexts and their corresponding causal models are organized in the mind of the learner. Are contexts represented without order, as in "a bag of contexts / models", or are they structurally organized? For example, do learners represent the relations between different contexts so that the changes in one context may be generalized to another? A rational approach might predict that contexts with similar causal models are clustered to achieve an efficient representation as well as to highlight the relationships among contexts. How can these intuitions be captured in a formal model for learning in multi-context environments?

In the rest of this article, we integrate existing findings that are relevant to the issue of learning in a multi-context environment. Our primary goal is to offer a comprehensive overview that brings together insights from across various literatures of cognitive science, so that one may come to realize what is yet to be investigated and understood. Additionally, we outline the directions for future research. How the learner determines when a change in context is relevant and then learns a new causal theory must, we claim, involve building hierarchical models (or heuristic approximations of them). Such a hierarchical model must include the storage of multiple contexts so that the unexpected input serves as a trigger to shift from one causal model to another, rather than simply updating the current model to improve the fit. Finally, we hypothesize that contexts themselves are structurally rich components that may share *cues*, so that it is possible to infer whether the environment has returned to a previous context at the time of a context change.

2 Detecting a context change

In a realistic learning task, the learner has to rely on observations that unfold over time to form hypotheses about the environment. If the environment consists of a single context, the sequential nature of the input is less likely to be a problem since an optimal learning strategy, as prescribed by Bayesian belief updating, is available (for general discussions on Bayesian modeling of cognition, see Griffiths et al., 2008; Jones & Love, 2011). Similarly, if the learner is given explicit information regarding which context they are currently in, there is no contextual ambiguities to solve. However, in most cases (such as the cafeteria example), the environment might change from one context to another implicitly, leaving the learner with the difficult task of estimating where one context ends and another one begins. The difficulty is further compounded by the sequential availability of the input – recognizing the emergence of a different context must be achieved in an on-line manner rather than with post-hoc analysis. Detecting context changes is commonly referred to as a *change detection* problem in many studies (e.g. Behrens et al., 2007; Yu, 2007).

While monitoring for unexpected observations in the input is an intuitive strategy for detecting context changes, at the core is the problem of interpreting ambiguity in the unexpected data: they can be interpreted as outliers if we assume the environment is still in the same context as before, or, they can also be interpreted as representative samples of a new context that is already in effect. As mentioned in the Introduction, we refer to this type of ambiguity as *contextual ambiguity*. How do learners resolve contextual ambiguity? Can they do so optimally? A satisfying answer to these questions requires a definition of optimality in the context of resolving contextual ambiguity. We discuss the factors that have been shown to influence how the learner resolves contextual ambiguity before presenting our definition of optimal ambiguity resolution.

2.1 Prediction error

Prediction error is widely recognized as one factor that can be used to adjudicate between outliers versus a true context change. In typical experimental settings, prediction error is either explicitly signaled by the degree of reduction in reward on a trial-by-trial basis (Behrens et al., 2007; Pearson et al., 2009; Nassar et al., 2010) or assumed to be (subconsciously) computed by learners who seek to optimize overall gains (e.g. Fine et al., 2010). Large prediction errors, especially when they persist over time, imply a change in context, while small prediction errors are likely to be random deviations in the current context. Thus, on average, learners will resolve contextual ambiguity faster when the new context differs greatly from the previous context. In the animal conditioning literature, the partial reinforcement extinction effect describes exactly that situation – after the extinction of reward, animals stop displaying the conditioned behavior more quickly when the behavior was trained with a high reward rate than with a low reward rate (Tarpy, 1982; Pearce et al., 1997). Going from a high reward rate environment to the extinction stage results in more prediction errors than from a low reward rate environment. Similarly, during foraging, animals tend to stop visiting a depleted food source more quickly if the source location was previously associated with a high return of food (Kacelnik et al., 1987; Dall et al., 1999; Freidin & Kacelnik, 2011).

When human learners are tested in a similar experimental paradigm known as the "bandit game", which features sequential choices among several alternatives with various rewarding rates, they tend to show higher learning rates when experimenters change reward rates without announcing the changes (Behrens et al., 2007; Nassar et al., 2010). Intuitively, high learning rates can accelerate the process of learning a new causal model, which helps quickly minimize the ongoing prediction error. The more important finding is, however, that the learning rate positively correlates with the magnitude of prediction error (Courville et al., 2006; Nassar et al., 2010). This implies that human learners potentially react to context changes in an optimal (or at least near-optimal) fashion: with small prediction errors, the learner adjusts their current behavior conservatively since small errors are likely to be random variations; with large prediction errors, the learner adopts a high learning rate to catch up with what is probably a changed context.

Converging evidence for the role of prediction error is also provided by imaging and multi-electrode recording studies. It has been suggested that the brain region known as the anterior cingulate cortex (ACC) represents predictions errors at the time of outcome (see Yu, 2007; Rushworth & Behrens, 2008; Pearson et al., 2011, for reviews and opinions on the role of ACC) or related quantities (e.g. the "volatility" of an environment; Behrens et al., 2007). More recent studies also suggested that prediction error belongs to the set of variables that are encoded by the neurons in the ACC to guide choice behavior in general (Hayden et al., 2011).

In the above scenarios, the information about prediction error is assumed to be immediately available once the learner has made a decision. However, there are other cases where such an assumption does not hold. For example, when prediction errors are derived from rewards, the learner will experience delayed prediction errors if rewards are given out in batches rather than on a trial-by-trial basis. How should the learner detect a context change in these situations? If learners adopt the same strategy as in an environment with immediate feedback, the overall loss will likely be widened because the incorrect causal model will be applied for a much longer period of time. So far, little to no empirical research has been conducted to investigate what kinds of strategies learners actually use to detect context changes in an environment coupled with delayed prediction errors.

2.2 Estimation uncertainty

Although large and small prediction errors are correlated with different presumed explanations for outliers, there are two types of prediction errors that are worth distinguishing. In the first case, the learner makes a substantial number of prediction errors because a good model of the environment has not yet been formed. Those prediction errors are the result of random guessing and are thus unhelpful for the purpose of resolving contextual ambiguity. The other type of prediction error arises when the learner is confident that the current causal model has been sufficiently refined to be a good theory for the current context, and then becomes genuinely surprised by the inadequate fit with the most recent input. From the rational decision-making perspective, only this second type of prediction error is meaningful to the learner (the solution to the former is simply to collect more data). Thus, one expects that when facing a particularly difficult task (due to either complexity or limited sampling), learners will be less likely to reach a low-uncertainty estimate of the current causal model, and they will consequently fail to recognize new contexts as easily as they have done in the studies reviewed above.

Unfortunately, none of the studies that we are aware of have addressed this issue directly within a single experimental paradigm. However, an artificial language learning experiment has provided some interesting insights. In Gebhart et al. (2009), learners listen to two artificial languages presented successively in a single session (with equal amount of exposure and without an overtly signaled change point). Under these conditions, only the first language is learned. The crucial difference between artificial grammar learning paradigms and simple decision-making tasks (such as the bandit games in Behrens et al., 2007) is that learners in the latter environment are able to reach asymptotic performance relatively effortlessly. On the contrary, learners cannot easily reach asymptotic performance in an artificial grammar learning experiment due to the high-dimensional nature of the linguistic input (Gerken, 2010). Therefore, the high uncertainty associated with the model of the first language prevents the learners from resolving the contextual ambiguity and learning a second grammar. Another experiment, in which subjects were tested with a variant of the famous Wisconsin Card Sorting task, showed that learners failed to detect when the sorting game entered a new context (characterized by changes in the reward rules) as optimally as a Bayesian learner (Wilson & Niv, 2012). Presumably, this is also because it is difficult to reach low estimation uncertainty when context changes result in structural differences in the causal relations, which is a more demanding learning task. Future studies, however, must test the hypothesis of estimation uncertainty directly within a single experimental paradigm to further our understanding of this issue.

2.3 Prior expectation for context change

What happens if learners approach the problem of resolving contextual ambiguity with a bias towards looking for changes in context? Put differently, will believing that there are multiple contexts prior to learning improve the recognition of changes? A variant of the foregoing artificial language learning experiment was conducted, where not only the subjects knew that there would be two languages (i.e., contexts), but also they experienced a 30-second silent pause between these two languages (Gebhart et al., 2009). With this change, subjects readily learned both languages. The bias towards changes can also be introduced by the use of more subtle explicit cues (e.g. subjects learn separate models when each context is coupled with a speaker-voice cue: Weiss et al., 2009), or by familiarizing learners with the pattern of a multi-context environment prior to conducting the target trials (Gallistel et al., 2001). These findings suggest that the prior expectation for a change in context enhances the ability of recognizing context changes in subsequent sequential input.

Is having a prior expectation for changes in context beneficial for learning in realistic and ecologically valid environments? This is largely an empirical question that awaits much more experimental investigation (see Green et al., 2010 for relevant discussions). Theoretically, it is not difficult to see that such a prior expectation is only advantageous when it matches the frequency of context changes in the environment. If the prior expectation for context change is comparatively weak, learners would simply ignore contextual ambiguity and miss the new context. However, if it is too strong, learners may effectively treat each minor deviation as a signal for a new context in the environment – thus overfitting the data. In that case, no stable learning can be achieved.

The ideal solution for the learner would be to estimate the frequency of context changes in the environment *before learning begins. However, such a strategy is only possible when the learner is familiar with the task environment and can anticipate the start of the learning process. Estimating the frequency of context changes in a novel environment, whose cues and features are entirely different from what the learner has encountered before, is indeterminate because there is no certainty about the type of changes and when they occur. The question of interest is then: how strong a prior the learner has for context changes in these novel environments?* While experimental evidence on this issue is thin, we do know that prior expectations for context change, in the absence of explicit instruction from the experimenter or explicit cues from the environment, must be relatively moderate. Such insights come from experiments where the context of the environment alternates frequently, resulting in an unrealistically volatile causal structure. In those conditions, learning is either virtually non-existent (Clapper & Bower, 2002) or substituted by a heuristic strategy that heavily depends on recent exemplars (Summerfield et al., 2011). The tendency of preferring locally stable and coherent observations is also seen in young infants: in the absence of suggestive information, infants are more likely to assume that a sequence of observations consists of correlated samples with common properties rather than independent samples randomly drawn from the whole population (Gweon et al., 2010).

3 Adapting to the changed context

Once a context change is hypothesized to have occurred, the learner must decide how to adapt to the changed context. If the context is novel, the learner has no choices other than to infer a set of new causal relations from observations. If the context is familiar, however, the learner may retrieve from memory the causal model of a past context and use it to predict future observations. Instead of discussing both scenarios directly (which we will cover slightly later), here we focus on two theoretical assumptions that must be in place to make these scenarios possible: the capacity of storing multiple contexts and the organization of these contexts in memory.

3.1 In with the new, while retaining the old?

When the environment presents a novel context, a new causal model should be generated to represent the dependencies between the variables of interest. To achieve this goal, the learner can either update the current causal model, parametrically or structurally, or learn a second model that will co-exist in parallel with the previous one. Existing accounts, such as associative strength theories (e.g. the Rescorla-Wagner model; Rescorla & Wagner, 1972) or reinforcement learning models (see Payzan-LeNestour & Bossaerts, 2011 for an example), have typically assumed the former theoretical position. Such a theoretical position is also shared by the more recently proposed change detection models (see Box 1) and sequential sampling models (see Box 2), both of which are intended to explain how ideal learners should behave in multi-context tasks.

— Insert Box 1 approximately here (box content on page 14) –

— Insert Box 2 approximately here (box content on page 16) —

However, disrupting or erasing the causal model learned under a past context (also known as catastrophic interference in connectionist terms; French, 1999) might not be a rational choice, especially when the environment may revert back to a past context. Experimental findings suggest that animals and humans do not simply abandon knowledge of past contexts. For example, in conditioning experiments, animals that have gone through extinction still possess a trace of the learned dependencies between the conditioned stimulus and response, which can spontaneously recover (e.g. Sissons & Miller, 2009), be renewed (e.g. Bouton & King, 1983) or be reinstated (e.g. Thanellou & Green, 2011) under the right conditions. Adult barn owls can rapidly re-adapt to an abnormal association between auditory cues and locations in visual space if they have previously learned such abnormal audio-visual dependencies when they were young (Knudsen, 1998; Linkenhoker et al., 2005). Humans also routinely switch back and forth between a certain set of contexts, without relearning a causal model each time a previously encountered context is active (for example, becoming familiar with a foreign accent does not lead to a complete relearning of your native accent). It is impossible for learners to display such behaviors without, implicitly or explicitly, representing multiple contexts concurrently. A theory for learning in multi-context environments must include a hypothesis about how these contexts are stored.

3.2 A bag of contexts?

Nevertheless, more behavioral and theoretical studies are needed to understand whether learners *optimally* represent learned models of past contexts, as would be predicted by a theory of a rational learner. When a past context has little to no chance of reappearing in the future, it seems unnecessary to store its information in memory (c.f. Anderson & Schooler, 1991). When a past context is quite common overall, or when a repetitive pattern of environmental changes has appeared, learners will benefit greatly if its information remains readily available through the learning process. In addition, in order to efficiently retrieve a causal model of a past context from memory, the learner must implement mechanisms that support the identification of familiar contexts. In the case where there are observable cues co-occurring with the advent of contexts, it is possible to index contexts with these cues for later retrieval. This is especially helpful as most contexts do not come with explicit labels – the use of co-occurring cues may serve as the functional labels for these contexts, which will then become easily retrievable (García-Gutiérrez & Rosas, 2003; Rosas & Callejas-Aguilera, 2006; Abad et al., 2009). In the case where there are no cues whatsoever, we expect learners to have a more difficult time identifying familiar contexts, potentially because such identification would have to rely on the assessment of multiple existing models.

These types of optimal learning decisions call for a sophisticated theory that, in our opinion, must extend beyond a process of parameter or structural revision of a *single* causal model. This is because at the end of the day, the outcome of the learning process should be more than a snapshot of the latest context of the environment, but rather an organized body of knowledge summarizing various forms of causal relations in the environment, past and present. We outline a picture of such a model – in the form of a Bayesian hierarchical model – in the next section. Finding the answers to these questions can greatly supplement our understanding of how animals and humans learn multiple causal models for multiple contexts to solve a particular task through sequential observations.

4 A hierarchical framework for learning in multi-context environments

The hierarchical Bayesian modeling framework has been successfully applied to a wide range of cognitive phenomena (e.g. Kemp et al., 2007; Kemp & Tenenbaum, 2008; also see Lee, 2011, for a review). In fact, most existing Bayesian models of change detection fall into the category of hierarchical models, where the volatility parameter is treated as a hyper-parameter (Behrens et al., 2007; and most notably the nested volatility model in Wilson et al., 2010). While we also advocate a hierarchical Bayesian approach for modeling learning behaviors in a multi-context environment, our primary goal is to understand whether the learner forms a hierarchical *representation* of the environment. Previous modeling efforts, on the other hand, have typically emphasized the issue of whether and how learners can dynamically adapt their strategies when contexts change. We argue that only when a generative model simultaneously represents multiple contexts and their corresponding causal models, will the ideal learner be able to attribute unexpected observations to the right sources, and retain and reuse causal models from past contexts (see Kording et al., 2007, for similar ideas).

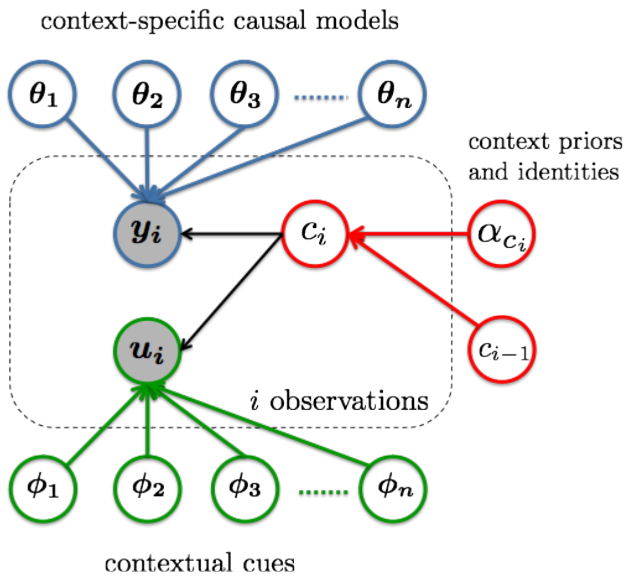

Figure 1: One potential hierarchical model for representing information learned in a multi-context environment.

Figure 1 shows one possible realization of such a hierarchical representation. For simplicity, consider an example where the causal models differ across contexts only in their parameter values, shown as $\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \boldsymbol{\theta_3} \ldots \boldsymbol{\theta_n}$ in the figure (bold symbols denote vectors of variables). There are three components in this hierarchical representation. The first component (highlighted in blue) consists of the contexts and causal models, each of which describes a theory of how the observations of interest $\boldsymbol{y_i}$ are generated from the parameters $\boldsymbol{\theta}$. Importantly, the parameters of the causal model of each context are individually represented, thus allowing for the storage of multiple contexts and avoiding catastrophic interference between these contexts. The second component is the mechanism that infers the identity of the currently active context c_i (highlighted in red). This decision process in turn depends on two variables: the hyperparameter α_{c_i}, which reflects the likelihood of context c_i coming into effect without explicit cues, and the inferred identity of the previously encountered context c_{i-1}. The identity of the currently active context corresponds to only one of the causal models (i.e. one of $\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \boldsymbol{\theta_3} \ldots \boldsymbol{\theta_n}$). Thus, once the identity of the current context has been *correctly* inferred (which might not be true due to probabilistic nature of the model), it can prevent the irrelevant contexts from being used to explain the observed data $\boldsymbol{y_i}$ or being revised to fit unrelated data. In other words, the dependence between $\boldsymbol{y_i}$ and c_i, as shown in the figure, serves as a regulator that chooses the appropriate context as needed.

The third component in the hierarchical representation is the *optional* cuing mechanism (highlighted in green). When covarying cues $\boldsymbol{u_i}$ are available, the values of these cues will depend on the identity of the contexts and the causal relations between contexts and these cues (the effect of $\boldsymbol{\phi}$ on $\boldsymbol{u_i}$). Therefore, these cues, in theory, serve the same functional purpose as the observations of interest $\boldsymbol{y}$ – evidence for inferring the identity of the current context. There is a vast literature on how humans may be able to optimally combine two sources of information to perform inferences (Ernst & Banks, 2002; Knill, 2007; Toscano & McMurray, 2010, to name a few). By building this cueing mechanism into the hierarchical representation, we are also making the assumption that learners should take advantage of the covarying cues as an extra source of information when available.

To be clear, Figure 1 is only meant to illustrate one of the many possible ways of constructing a hierarchical model to capture context-sensitive learning. Many details, such as the prior for the appropriate number of $\boldsymbol{\theta}$ variables and any hyperparameter reflecting the relationships between them, are not shown in the figure. Our goal here is to provide a concrete sense of what a hierarchical framework may look like for future modeling efforts. Experimental studies, especially those designed to test the effect of recognizing past contexts, are needed to further tease apart the factors that affect learning in a multi-context environment.

5 Considerations for single-context laboratory experiments

If animal and human subjects can readily detect new contexts without being explicitly instructed to do so, then we have reason to suspect that subjects will involuntarily look for context changes even in laboratory experiments where subjects are expected to learn a causal model for a fixed but unknown context. In a variety of such behavioral tasks, subjects exhibit an automatic and seemingly suboptimal behavior: they put an undue emphasis on the sequence of past observations, even when these observed stimuli are independent samples from the same causal model. Two notable instances of such suboptimal behavior in the literature are the hot hand illusion (Gilovich et al., 1985) and the tendency of reinforcing local patterns (e.g. Cho et al., 2002; Maloney et al., 2005; Gökaydin et al., 2011). While the conventional interpretation is that learners are irrational in that they perceive spurious correlations between past and upcoming outcomes, these seemingly suboptimal behaviors may well be the result of learners automatically inferring multiple contexts (e.g., hot hand context vs. cold hand context) from the sequential input (for similar opinions, see Jones & Sieck, 2003; Yu & Cohen, 2008; Wilder et al., 2010; Green et al., 2010). More generally, the bias for perceiving multiple contexts may also hold the key to explaining order effects in learning (e.g. Sakamoto et al., 2008; Rottman & Keil, 2012). At the same time, it raises the concern that such a bias may lead to misinterpreted experimental findings because participants readily adapt to what they perceive to be changes in contexts (perhaps subconsciously). The above cited studies are in fact the best examples to show that the use of balanced designs in experiments do not effectively prevent participants from "inappropriately" adopting this bias (see Jaeger, 2010 for similar discussions).

6 Conclusions

Recognizing context changes in the environment helps learners build or choose the appropriate causal model and make accurate predictions about the consequences of their actions. In this article, we have addressed several questions about what we believe is the canonical case of context learning: when the changes in context are implicit rather then being explicitly noted by a 'teacher'. Current research findings suggest that learners are able to resolve contextual ambiguity and thereby recognize a new context by only observing sequential input, albeit with some limitations. Recognizing a new context is, however, only a part of the bigger picture. How do learners store the causal models of past contexts? Can learners reuse previously learned causal models? Crucially, the definition of rationality should rely on one important issue: given a change in context, should the learner build a new causal model or try to reuse, and potentially update, an old one? How should the learner decide? These intriguing questions are open for future research.

Box 1: Bayesian change detection models

Detecting a change in context is an important step in learning a rich representation of a multi-context environment. The traditional approach to change detection comes from studies of controlled stochastic processes (e.g. Shiryaev, 1978), where the goal is to find an optimal policy for mapping observations to stopping decisions (i.e., whether or not to consider that a context has ended). While the solutions are useful for many engineering applications, it is often difficult to attach a cognitive interpretation to the algorithms used in those solutions.

Here we focus on the Bayesian change detection approach that has recently become popular in the cognitive science community. As a computational-level theory, these models describe how a rational observer *should* learn a causal model given a particular formulation of the problem (Marr, 1982). Consider a simple scenario where the goal is to predict the number of automobiles that pass through a given intersection in each 24-hour period. The parameter of interest is θ, which refers to the number of automobiles being driven from point A to point B. The causal model to be discovered by the learner specifies the relation between the parameter θ and the observation y, the number of automobiles passing through the intersection. However, at any given time step, a change in context might happen (e.g., road construction), which will alter the previous relation in effect and yield unexpected observations. Detecting the change then depends on how likely the learner is to attribute the unexpected observations to a change in the value of θ. The change detection approach assumes the determining factor here is the learner's expectation of the volatility of θ. If θ is assumed to be changing smoothly and with little variance (i.e. non-volatile), then learners will tend to view unexpected observations as outliers and keep the value of θ unchanged. If θ is assumed to be capable of abrupt changes of substantial magnitude, learners will more likely update the value of θ when observing unexpected data.

Formally, the volatility of an environment, represented by a hyper-parameter α, can range from 0 to 1: With probability α, θ_t will be the same value as θ_{t-1}; with probability $1-\alpha$, θ_t will be randomly drawn from a predefined reset distribution p_0. Thus, if α is 1, then learners are essentially assuming a single-context environment, where the value of θ is the same at each time step. If its value is 0, then learners are essentially assuming a completely chaotic multi-context environment, where the value of θ at the preceding time step has no predictive value over the current time step at all. Any intermediate value reflects the degree to which learners are biased against single-context environments. Additionally, the value of α, i.e. the degree of volatility, can change over time as well.

This model gained its popularity due to its conceptual simplicity and the range of phenomena it can explain (Cho et al., 2002; Yu & Cohen, 2008; Wilder et al., 2010; Wilson et al., 2010; see also Nassar et al., 2010; Mathys et al., 2011 for variants that are claimed to be cognitively more plausible; and Summerfield et al., 2011; Wilson & Niv, 2012 for cases where the Bayesian change detection model is not the best descriptor of human behavior). A significant drawback of this class of models, however, lies in its memory-less learning mechanism. Once the ideal learner detects a change in context, it learns the new parameter settings by overriding those of the old context. This is undesirable since animal and human learners have clearly demonstrated the ability of holding onto knowledge learned from past contexts.

Box 2: Sequential sampling methods

Sequential sampling models are another approach to learning in multi-context environments. These models are inspired by sequential Monte Carlo sampling techniques, which are commonly used to approximate Bayesian inference in analytically non-tractable problems. In the cognitive science community, the particle filter, one of the most common sequential sampling algorithms (e.g. Sanborn et al., 2010), has been successfully applied to learning tasks where there are changes in context (Brown & Steyvers, 2009). In a particle filter model, the learner is assumed to simultaneously entertain a limited number of hypotheses (called particles) about the values of parameters in the environment. This contrasts with the Bayesian change-detection approach, where learners are assumed to maintain full uncertainty about the estimates of the volatility (i.e. α) and state (i.e. θ) parameters. At the beginning of the learning process, random values of θ are assigned to the particles since the learner has not made any observation of the environment. Each particle is then repeatedly updated according to subsequent observations. If a particle reflects a theory of the environment that is consistent with a new observation, then it is likely to be retained. Otherwise, the particle is likely to be reset and its value resampled from the hypothesis space. Since this sampling process is stochastic, there is always some chance that a few particles are inconsistent with the current state of the environment. These inconsistent particles are useful for detecting context changes in the environment. When the learner encounters an unexpected observation, particles that used to be consistent with the previous context now need to be reset, while those that were previously inconsistent are retained and duplicated, thus achieving the goal of detecting changes.

While we are not aware of any study directly testing the different predictions made by the change detection and the particle filter models, one crucial difference exists between them. The particle filter model, due to its stochastic nature and the limited number of observations in sequential sampling, is suited for predicting individual-level results (Brown & Steyvers, 2009; Yi & Steyvers, 2009; Frankenhuis & Panchanathan, 2011). The change-detection model, because its goal is to characterize rational behaviors, is suited for predicting average behavior. Patterns of individual learning outcomes tend to be different from group-averaged learning outcomes (Newell et al., 2001; Gallistel et al., 2004). Particle filter models can readily accommodate such differences – a single run of a sequential sampler tends to yield unpredictable patterns, but the average of many runs, by definition, reflects the expected properties of the probability distribution that is being sampled from (see Daw & Courville, 2008, for a similar argument).

References

Abad, M. J. F., Ramos-Alvarez, M. M., & Rosas, J. M. (2009). Partial reinforcement and context switch effects in human predictive learning. *Q. J. Exp. Psychol.*, *62*(1), 174–88.

Anderson, J., & Schooler, L. (1991). Reflections of the environment in memory. *Psychol. Sci.*, *2*(6), 396–408.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.*, *10*(9), 1214–21.

Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *J. Exp. Psychol. Anim. Behav. Process*, *9*(3), 248–65.

Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cogn. Psychol.*, *58*(1), 49–67.

Cho, R. Y., Nystrom, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P. J., et al. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cogn. Affect. Behav. Neurosci.*, *2*(4), 283–99.

Clapper, J. P., & Bower, G. H. (2002). Adaptive categorization in unsupervised learning. *J. Exp. Psychol. Learn. Mem. Cogn.*, *28*(5), 908–923.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends. Cogn. Sci.*, *10*(7), 294–300.

Dall, S., Mcnamara, J., & Cuthill, I. (1999). Interruptions to foraging and learning in a changing environment. *Anim. Behav.*, *57*(1), 233–241.

Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Adv. Neural. Inf. Process. Syst.*, *20*, 369–376.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–33.

Fine, A., Qian, T., Jaeger, T., & Jacobs, R. (2010). Is there syntactic adaptation in language comprehension? In *Proceedings of CMCL Workshop* (pp. 18–26). ACL.

Frankenhuis, W. E., & Panchanathan, K. (2011). Individual Differences in Developmental Plasticity May Result From Stochastic Sampling. *Perspect. Psychol. Sci.*, *6*(4), 336–347.

Freidin, E., & Kacelnik, A. (2011). Rational Choice, Context Dependence, and the Value of Information in European Starlings (Sturnus vulgaris). *Science*, *334*(6058), 1000–1002.

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends. Cogn. Sci.*, *3*(4), 128–135.

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.*, *101*(36), 13124–31.

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *J. Exp. Psychol. Anim. Behav. Process*, *27*(4), 354–372.

García-Gutiérrez, A., & Rosas, J. M. (2003). Context change as the mechanism of reinstatement in causal learning. *J. Exp. Psychol. Anim. Behav. Process*, *29*(4), 292–310.

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cogn. Sci.*, *33*(6), 1087–1116.

Gerken, L. (2010). Infants use rational decision criteria for choosing among models of their input. *Cognition*, *115*(2), 362–6.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cogn. Psychol.*, *17*(3), 295–314.

Gökaydin, D., Ma-Wyatt, A., Navarro, D., & Perfors, A. (2011). Humans use different statistics for sequence analysis depending on the task. In *Proceedings of CogSci 2011* (pp. 543–548).

Green, C. S., Benson, C., Kersten, D., & Schrater, P. (2010). Alterations in choice behavior by manipulations of world model. *Proc. Natl. Acad. Sci. U.S.A.*, *107*(37), 16401–6.

Griffiths, T., Kemp, C., & Tenenbaum, J. (2008). Bayesian models of cognition. In *Cambridge Handbook of Computational Psychology* (pp. 59–100).

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proc. Natl. Acad. Sci. U.S.A.*, *107*(20), 9066–71.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.*, *14*(7), 933–9.

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cogn. Psychol.*, *61*(1), 23–62.

Jones, M., & Love, B. C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav. Brain Sci.*, *34*(04), 169–188.

Jones, M., & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *J. Exp. Psychol. Learn. Mem. Cogn.*, *29*(4), 626–640.

Kacelnik, A., Krebs, J. R., & Ens, B. (1987). Foraging in a changing environment: An experiment with starlings (sturnus vulgaris). In M. L. Commons, A. Kacelnik, & S. J. Shettleworth (Eds.), *Quantitative Analyses of Behavior: Foraging* (Vol. VI, p. 63-87). Psychology Press.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Dev. Sci.*, *10*(3), 307–21.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proc. Natl. Acad. Sci. U.S.A.*, *105*(31), 10687–92.

Knill, D. C. (2007). Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of vision*, *7*(7), 5.1–24.

Knudsen, E. (1998). Capacity for Plasticity in the Adult Owl Auditory System Expanded by Juvenile Experience. *Science*, *279*(5356), 1531–1533.

Kording, K. P., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat. Neurosci.*, *10*(6), 779–86.

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *J. Math. Psychol.*, *55*(1), 1–7.

Linkenhoker, B. A., Ohe, C. G. von der, & Knudsen, E. I. (2005). Anatomical traces of juvenile learning in the auditory system of adult barn owls. *Nat. Neurosci.*, *8*(1), 93–8.

Maloney, L. T., Dal Martello, M. F., Sahm, C., & Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proc. Natl. Acad. Sci. U.S.A.*, *102*(8), 3164–9.

Marr, D. (1982). *Vision*. W.H. Freeman.

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.*, *5*(May), 39.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.*, *30*(37), 12366–78.

Newell, K. M., Liu, Y. T., & Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychol. Rev.*, *108*(1), 57–82.

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS computational biology*, *7*(1), e1001048.

Pearce, J. M., Redhead, E. S., & Aydin, A. (1997). Partial reinforcement in appetitive Pavlovian conditioning with rats. *Q. J. Exp. Psychol.*, *50B*(4), 273–94.

Pearson, J. M., Hayden, B. Y., Raghavachari, S., & Platt, M. L. (2009). Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr. Biol.*, *19*(18), 1532–7.

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends. Cogn. Sci.*, *15*(4), 143–51.

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

Rosas, J. M., & Callejas-Aguilera, J. E. (2006). Context switch effects on acquisition and extinction in human predictive learning. *J. Exp. Psychol. Learn. Mem. Cogn.*, *32*(3), 461–74.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: observations and interventions. *Cogn. Psychol.*, *64*(1-2), 93–125.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.*, *11*(4), 389–97.

Sakamoto, Y., Jones, M., & Love, B. C. (2008). Putting the psychology back into psychological models : Mechanistic versus rational approaches. *Mem. Cognit.*, *36*(6), 1057–1065.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychol. Rev.*, *117*(4), 1144–67.

Shiryaev, A. N. (1978). *Optimal stopping rules*. New York: Springer-Verlag.

Sissons, H. T., & Miller, R. R. (2009). Spontaneous recovery of excitation and inhibition. *J. Exp. Psychol. Anim. Behav. Process*, *35*(3), 419–26.

Summerfield, C., Behrens, T. E., & Koechlin, E. (2011). Perceptual classification in a rapidly changing environment. *Neuron*, *71*(4), 725–36.

Tarpy, R. (1982). *Principles of animal learning and motivation*. London: Scot, Foresman.

Thanellou, A., & Green, J. T. (2011). Spontaneous recovery but not reinstatement of the extinguished conditioned eyeblink response in the rat. *Behav. Neurosci.*, *125*(4), 613–25.

Toscano, J. C., & McMurray, B. (2010). Cue Integration With Categories: Weighting Acoustic Cues in Speech Using Unsupervised Learning and Distributional Statistics. *Cogn. Sci.*, *34*(3), 434–464.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech Segmentation in a Simulated Bilingual Environment: A Challenge for Statistical Learning? *Lang. Learn. Dev.*, *5*(1), 30–49.

Wilder, M., Jones, M., & Mozer, M. (2010). Sequential effects reflect parallel learning of multiple environmental regularities. *Adv. Neural. Inf. Process. Syst.*, *22*, 2053–2061.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2010). Bayesian online learning of the hazard rate in change-point problems. *Neural Comput.*, *22*(9), 2452–76.

Wilson, R. C., & Niv, Y. (2012). Inferring relevance in a changing world. *Front. Hum. Neurosci.*, *5*(January), 189.

Yi, M., & Steyvers, M. (2009). Modeling human performance in restless bandits with particle filters. *J. Prob. Solv.*, *2*(2), 81–102.

Yu, A. J. (2007). Adaptive behavior: humans act as Bayesian learners. *Curr. Biol.*, *17*(22), R977–80.

Yu, A. J., & Cohen, J. D. (2008). Sequential effects: Superstition or rational behavior? *Adv. Neural. Inf. Process. Syst.*, *21*, 1873–1880.