This paper has been accepted for conference proceedings, meaning it has undergone a single round of peer review for inclusion in such proceedings. Subsequent minor revisions have been made based on these reviews; however, the version presented here has not undergone any additional review.
Learn more about our research at https://nivlab.princeton.edu

Last updated – May 16th 2023 v1
First posted online – May 16th 2023 v1
Accepted in Cognitive Science Conference Proceedings 2023 (v1)

Modelling Rumination as a State-Inference Process

Rachel L Bedder*, Sashank Pisupati, & Yael Niv (2023)

*Corresponding Author, rlbedder@princeton.edu

Modelling Rumination as a State-Inference Process

Rachel L Bedder[1] **(rlbedder@princeton.edu)**
Sashank Pisupati[1,2] **(pisupati@princeton.edu)**
Yael Niv[1] **(yael@princeton.edu)**
[1] Princeton Neuroscience Institute & Department of Psychology
Princeton University, Princeton, NJ, 08540 USA
[2] Limbic Ltd

Abstract

Rumination is a kind of repetitive negative thinking that involves prolonged sampling of negative episodes from one's past, typically prompted by a present negative experience. We model rumination as an attempt at hidden-state inference, formalized as a partially-observable Markov decision process (POMDP). Using this allegorical model, we demonstrate conditions under which continuous, prolonged collection of samples from memory is the optimal policy. Consistent with phenomenological observations from clinical and experimental work, we show that prolonged sampling (i.e., chronic rumination), formalized as needing to sample more evidence before selecting an action, is required when possible negative outcomes increase in magnitude, when states of the world with negative outcomes are *a priori* more likely, and when samples are more variable than expected. By demonstrating that prolonged sampling may allow for optimal action selection under certain environmental conditions, we show how rumination may be adaptive for solving particular problems.

Keywords: clinical; rumination; POMDP; latent state; inference; computational psychiatry

Introduction

Rumination has been defined as *"a class of conscious thoughts that revolve around a common instrumental theme and that recur in the absence of immediate environmental demands requiring the thoughts"* (Martin & Tesser, 1996). Trait rumination predicts the initial onset of depressive episodes (Abela, Brozina, & Haigh, 2002), and the likelihood of future episodes (Kuehner & Weber, 1999). Formalizing our understanding of rumination is especially challenging because rumination is an entirely internal and highly personal process. As such, datasets of the dynamics of ruminative episodes from initiation to termination do not exist, and there are no clear benchmarks for what computational models ought to explain.

The goal of this paper is to provide a computational model of why one might engage in a protracting thinking process that is inherently unpleasant (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). We hypothesize that rumination's normative origins lie in attempting to understand why some negative affective experience has occurred. This is achieved through trying to infer the (hidden) state of the world that the experience belongs to, by sampling from memory episodes that are thematically similar to the initial negative experience. For example, one might infer that their boss yelling at them is part of the state *'everyone is in a bad mood on Mondays,'* by recalling other instances where co-workers acted irrationally at the beginning of the week. By inferring the hidden state with sufficient certainty, the individual can select an action they believe will lead to positive outcomes (and resolve the negative affect) in this state of the world (e.g. opt to work from home on Mondays). We model this inference and action-selection process as a partially observable Markov decision process (POMDP; Kaelbling, Littman, & Cassandra, 1998) and simulate sampling of information from various hidden states of the world in order to explore conditions under which prolonged sampling may be the optimal policy.

We focus on three simulations reflecting different beliefs an individual may hold that may lead them to engage in chronic rumination. First, rumination may be exacerbated by a negativity bias: we show that when potential negative outcomes increase in magnitude, the optimal number of samples required before choosing an action increases. Second, we show that if inference begins with an *a priori* expectation that the hidden state of the world is negative, the optimal number of samples also increases. Finally, we show that when samples are more variable than expected, the optimal number of samples also increases.

Background

This paper integrates clinical observations and experimental findings on rumination and depression with a normative model of sequential sampling. Worry, another kind of repetitive negative thinking, has been suggested to be normative under some forward-looking planning models (Gagne & Dayan, 2022). It is less clear whether rumination's dwelling on the past has a similar normative origin; one such account uses an active-inference framework where sampling of behavioral policies is not able to reduce uncertainty (Berg, Feldmann, Kirchner, & Kube, 2022).

Clinical work often describes rumination as comprising thematically linked memories and their abstractions. For example, a ruminative episode may be prompted by thinking "*why can't I sleep?*", leading to a series of thoughts about all the things one didn't achieve or get done in the last week, leading to abstracted thoughts such as "*I am a failure*". Clinical explanations for the function of rumination include gathering evidence towards a conclusion by dwelling on previous episodes, and avoiding negative consequences by thinking about a problem (or solution) rather than acting and risking aversive consequences (Watkins, 2018).

To model ruminative episodes of this kind, we desire a normative model of the process of sequential sampling of thoughts or memories, and its termination, as a function of some desired objective (e.g., maximizing long-term outcomes) under some assumptions about the world (e.g., expectations of positive/negative outcomes and their generalization properties). The POMDP framework satisfies all these criteria, and has been used in other domains to model similar sequential sampling and termination processes (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012). Unlike in standard reinforcement-learning models that are formulated using Markov decision processes, states in a POMDP are not directly observable. As such, in many conditions, it is advantageous to use sampling actions to gather information and reduce uncertainty about the state before committing to a terminating action. In contrast to standard Bayesian models of sequential sampling, POMDPs make explicit the short- and long-term costs and benefits of different actions, adopting the reinforcement-learning objective of maximizing long-term reward, and allowing sampling to be biased by this objective (Dayan & Daw, 2008). Such models have been used previously in computational psychiatry, for example to explore inferential abnormalities and delusion severity in schizophrenia (Baker, Konova, Daw, & Horga, 2019). This paper represents a foundational attempt to apply such models to the entirely internal process of a stream of thought.

Approach

In our model, at each time step, an agent must choose one of three actions consisting of a sampling action, and two terminating actions that end the ruminative episode and lead to reward if the agent has correctly inferred the true state of the world (and to punishment otherwise). The agent makes this choice based on the long-run expected rewards of each action. Crucially, sampling is likely to increase the agent's certainty in its current state, such that the agent is more likely to choose the correct (i.e., rewarding) terminating action. However, sampling also comes at the cost of prolonging the ruminative episode. In the first two simulations, the agent's knowledge of the observation distributions (i.e., how likely is each state to produce different types of samples) matches the distributions we generate samples from. In the third simulation, we explore what happens when this is not the case.

We investigate how beliefs about possible hidden states of the world may cause protracted sequences of sampling under the optimal action-selection policy. We formalize these beliefs by modulating the agent's knowledge about the reward and loss outcomes of hidden states, the agent's initial belief about the hidden state, and the variance of the observation distributions. The optimal policy determines the certainty required before the agent makes a terminating action (i.e. the agent's decision threshold), and thus the amount of sampling in each condition. In this way, we explore under what conditions chronic rumination, formalized as prolonged sampling, may be normative.

General Methods

All simulations were implemented in Python. Code is available at https://github.com/RachelBedder/rumination_cogsci

Partially Observable Markov Decision Process

We model an environment with two possible hidden states, X_1 and X_2. The true state of the environment determines the probability of obtaining each of two outcomes (R_1, R_2) when selecting either of the two terminal actions (A_1, A_2). Using a simple POMDP, we model beliefs about the true state of the environment as a probability distribution B. At each timestep, the agent can either choose one of the two terminating actions, or a third action ($Sample$) that provides an observation o from the hidden state and potentially reduces uncertainty regarding the current state of the world. Choosing to sample accrues a cost c. Each timestep thus involves three stages: belief updating based on the recent observation, valuation and choice.

Belief Updating. On each time step, the agent views an observation o and updates their belief about the hidden state using Bayes rule (Equation 1). In our simple model with only two hidden states, we define $B = P(X_2)$ and $P(X_1) = 1 - B$. We use B' to denote the subsequent belief state. We discretize the belief state in steps of 0.01, and the observation distribution, $o \in (0, 100]$, in steps of 1. We assume $P(o|X)$ is a known observation distribution.

$$B' = \frac{P(o|X_2)B}{P(o|X_1)(1-B) + P(o|X_2)B} \tag{1}$$

Valuation. The value of terminal actions A_1 and A_2 is determined by the outcome for that action in each possible state of the world, weighted by the agent's current belief of occupying that state:

$$Q(A_2, B) = R(A_2|X_1) \cdot (1-B) + R(A_2|X_2) \cdot B \tag{2}$$

and similarly for A_1. We assume that $R(A|X)$ is a known reward function.

The value of sampling given an agent's current belief, $Q(Sample, B)$, is determined recursively as a function of the cost of sampling, c, and the value of each subsequent belief state, $V(B')$, weighted by the probability of transitioning to it, $P(B'|B)$:

$$Q(Sample, B) = c + \sum_{B'} P(B'|B) \cdot V(B') \tag{3}$$

$$V(B') = max\{Q(A_1, B'), Q(A_2, B'), Q(Sample, B')\} \tag{4}$$

We calculated state values using value iteration (Sutton & Barto, 2018), concluding when the largest difference between iterations fell below 1 (1% of the baseline reward in all simulations).

We then used the Q values of each action in each belief state to determine the decision thresholds for choosing A_1 and A_2, defined as the most uncertain belief state B where the optimal action is no longer sampling.

Choice. We assumed a policy that selects the highest-valued action:

$$\pi(B) = argmax\{Q(A_1, B), Q(A_2, B), Q(Sample, B)\} \tag{5}$$

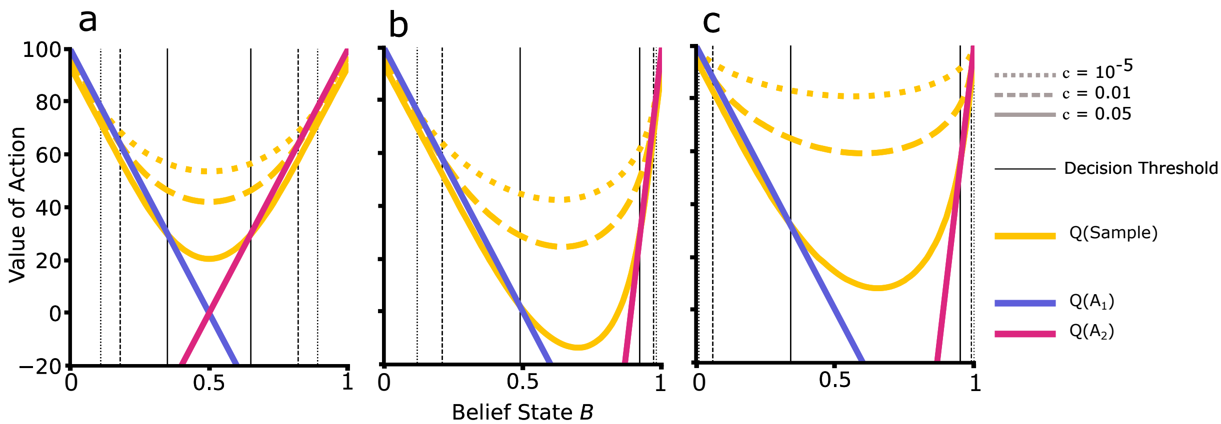

Figure 1: **Values of all actions in each belief state.** Using the optimal policy, the Q value of each action (A_1 (blue), A_2 (pink) and *sampling* (yellow)) is shown as a function of the belief state B, and for three sampling costs $c \in \{10^{-5}, 0.01, 0.05\}$. Decision thresholds (black vertical lines) indicate the minimum degree of certainty required for the optimal policy to select A_1 (black line closest to $B = 0$) or A_2 (line closest to $B = 1$). **a.** Baseline conditions: all rewards are +100 and all losses are -100 (observation distributions: $o(X_1) \sim \mathcal{N}(\mu = 45, \sigma^2 = 20^2)$, $o(X_2) \sim \mathcal{N}(55, 20^2)$). **b.** The loss for A_2 in X_1 is -800. **c.** The standard deviation for both observation distributions is 15 and the loss for A_2 is -800.

Simulations. In all simulations, the agent completed 1000 trials under each combination of reward and cost conditions. Observations were generated from state X_2, which we refer to as the "true state" (and X_1 as the "alternative state"). A_2 was the action with the rewarding outcome in state X_2, and A_1 the rewarded action in X_1 (the other action in each state was punished; rewards and punishments varied by simulation and will be detailed below). Each trial began with the agent making a choice, which either terminated the trial or produced a sample observation and led to the next choice. We limited the maximum number of samples on each trial to 100.

Modulation of rumination by the magnitude of potential negative outcomes

People with depression often overestimate the impact of negative outcomes (Alloy & Ahrens, 1987). In our first simulation, we demonstrate that when prospective losses increase, the number of samples taken by the optimal policy increases.

Methods

In this simulation, the two hidden states generated observations from Gaussian distributions with means of 45 and 55 respectively, and a standard deviation of 20. Agents had full knowledge of these observation distributions. The true state of the environment was always X_2, for which the reward for choosing A_2 was +100 and the loss for choosing A_1 was -100. Figure 1a shows the values of all actions under the optimal policy for this condition. In the alternative state X_1, the reward for choosing A_1 was +100, whereas the potential loss for choosing A_2 was -12.5, -25, -50, -100, -200, -400, or -800 across conditions (Figure 1b shows the values when the potential loss is -800).

We tested how increasing the loss magnitude for A_2 in the alternative state affected the average number of samples taken

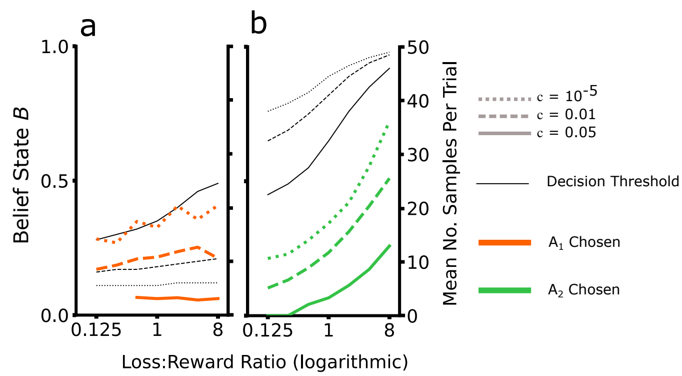

Figure 2: **Decision threshold evolution and number of observations sampled for increasing (potential) negative outcomes.** The sampling policy (length of rumination) is presented for sampling costs $c \in \{10^{-5}, 0.01, 0.05\}$, with the loss to reward ratio 0.125 to 8 (logarithmic scale on x axis). Black lines show decision thresholds (left axis) for an optimal agent to select **(a)** A_1, or **(b)** A_2 for different sampling costs. Colored lines denote the mean number of samples chosen on each of the trials (right axis). As the loss-reward ratio increased, agents chose A_2 (the correct action) less frequently: this action was chosen in 91.5% of trials when the ratio was 0.125 (c = 0.01), in 85.2% of trials when the ratio was 1, and in only 79.4% of trials when the ratio was 8 (a potential loss of -800 versus a gain of +100). When the cost of sampling was high (c = 0.05) and the loss-reward ratio was low, agents immediately chose A_2 after very few (or even zero) samples.

by the optimal policy on each trial. Also note that X_1 was never the true state, so the potential loss in that state is exactly that – a potential threat that never materializes, and therefore its magnitude is left to the imagination of the agent.

Results

Sampling increased when prospective losses were greater. As the loss-reward ratio increased, the decision threshold on belief state B that the agent required before choosing the correct action A_2 also increased, as did the number of samples taken (Figure 2b). For example, when the loss was one quarter of the reward, the threshold was $B \geq 0.69$ and the mean number of samples required was 6.64, whereas when the loss was four times larger than the reward, A_2 was chosen only when $B \geq 0.94$, and the mean number of samples required for this level of certainty was 20.77 (for sampling cost $c = 0.01$). Intuitively, as the magnitude of the potential loss for choosing A_2 in state X_1 increases, the optimal policy requires increased certainty that the state is indeed X_2 before selecting A_2.

Modulation of rumination by pessimistic initial beliefs

One possible reason for ruminating is believing that you are more likely to be in a disadvantageous state of the world. In the next simulation, we demonstrate that when the belief state starts *a priori* closer to the hidden state with a greater prospective loss (a pessimistic initial belief), the optimal policy requires more samples to choose the rewarding action. However, in extreme conditions, specifically, a pessimistic initial belief, a high cost of sampling and large prospective loss, the optimal policy is to not sample at all but rather to immediately choose the "safer" action A_1 that would avoid catastrophic losses. This suggests that rumination may only be initiated when it is perceived as not entirely futile.

Methods

In the previous simulation, agents began with $B = 0.5$, that is, they assumed states X_1 and X_2 were equally likely to be the true hidden state. Here, we varied the initial belief state between 0.125 and 0.875 in increments of 0.125, and the loss for A_2 in X_1 was -12.5, -100 or -800 across conditions.

Results

Increased sampling is required to make the correct choice in pessimistic initial belief states. As the initial belief in state X_2 decreased, the number of samples taken before correctly choosing A_2 increased. In contrast, the incorrect action A_1 was "rashly" chosen after few samples when the initial belief in X_2 was very low (Figure 3). Intuitively, when the initial belief state was close to a decision threshold, only a small number of samples were needed for the belief state to exceed the decision threshold. Thus, if the initial belief state was close to the decision threshold for selecting A_1, and the observation distributions had high variance, A_1 was sometimes selected if a few noisy samples generated from the true hidden state X_2 erroneously provided evidence for the alternative hidden state X_1. When samples supported X_2, more samples were needed as the belief had to change considerably before an action was chosen.

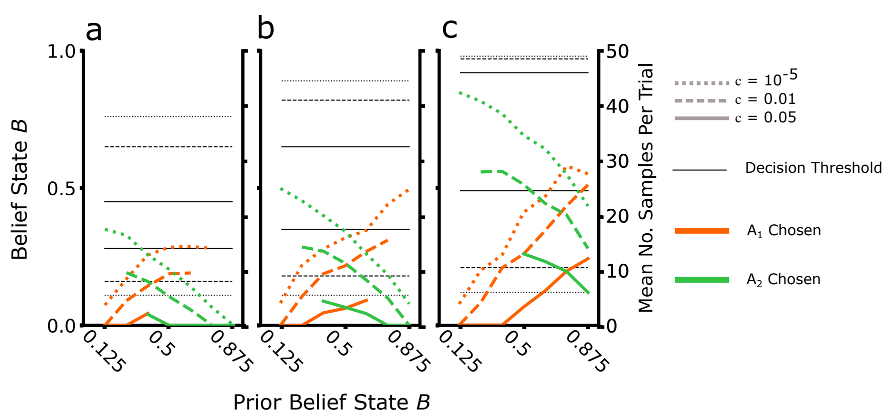

Figure 3: **Varying initial belief states.** Simulation results are presented for sampling costs $c \in \{10^{-5}, 0.01, 0.05\}$ with potential losses of (**a**) -12.5, (**b**) -100, and (**c**) -800. The agent's initial belief about the hidden state increases on the x axis from $B = 0.125$ (more likely to be X_1) to $B = 0.875$ (more likely to be X_2). Black horizontal lines denote the most uncertain belief state B in which the optimal policy selects A_2 (left y-axis), with the higher line of each line style indicating the decision threshold for A_2, and the lower for A_1. The colored lines denote the mean number of samples chosen on each trial (orange for when A_1 was chosen, green for when A_2 was chosen, right y-axis). When the potential loss was -12.5 (**a**), the correct choice was chosen more often as the initial belief state increased, both when the sampling cost was low ($c = 10^{-5}; B = 0.125 : 36\%, B = 0.5 : 92\%, B = 0.875 : 100\%$) and high ($c = 0.05; B = 0.125 : 0\%, B = 0.5 : 100\%, B = 0.875 : 100\%$). Similarly, when the potential loss was -800 (**c**), the same relationship was observed when costs were low ($c = 10^{-5}; B = 0.125 : 34\%, B = 0.5 : 92\%, B = 0.875 : 99\%$) and when costs were high ($c = 0.05; B = 0.125 : 0\%, B = 0.5 : 34\%, B = 0.875 : 94\%$).

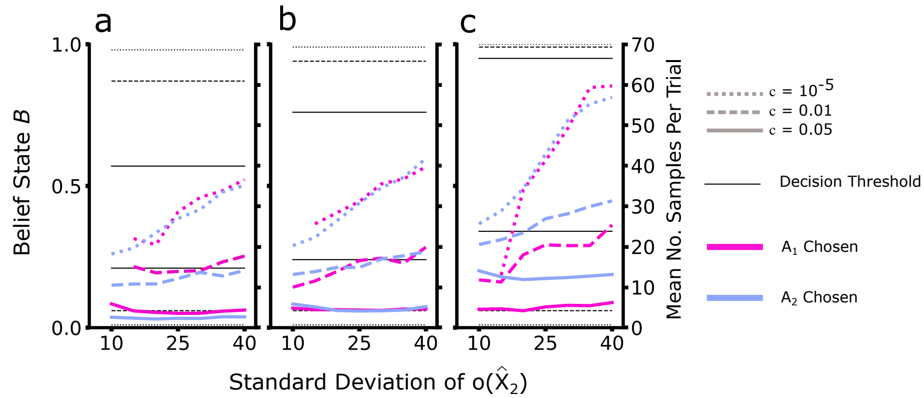

Figure 4: **Varying standard deviation of observation distributions.** Simulation results for sampling costs $c \in \{10^{-5}, 0.01, 0.05\}$ with potential losses of (**a**) -12.5, (**b**) -100, and (**c**) -800. Black horizontal lines denote the most uncertain belief state B an optimal agent must have to select A_2 (left y-axis) with the higher line of each line style indicating the decision threshold for A_2, and the lower for A_1. The colored lines denote the mean number of samples chosen on each trial (pink for when A_1 was chosen, light blue for when A_2 was chosen, right y-axis).

As potential losses increase, so does sampling when initial beliefs are pessimistic. When the initial belief state is pessimistic ($B < 0.5$), rash (incorrect) decisions were more common as the loss outcome for A_2 in X_1 increased, because the punishment for "incorrectly" inferring X_2 in a world where X_1 is true (and thus choosing the terminal action that would lead to a catastrophic punishment) outweighed the risk of making the incorrect action in a true X_2 (Figure 3c).

When it is costly to sample, pessimistic or optimistic initial belief states reduce the likelihood of beginning a sampling sequence. When the loss-reward ratio was equal (Figure 3b), as the initial belief state B approached 0 or 1, the optimal policy terminated with minimal to no sampling. For example, for high sampling cost (c = 0.05), the trial was terminated immediately without sampling when the initial belief state was lower than 0.25 or higher than 0.75. For greater loss-reward ratios (Figure 3c), the punishing action A_1 was selected immediately even in more uncertain pessimistic belief states ($B \leq 0.375, c = 0.05$). Intuitively, when sampling was expensive and there were significant potential losses at stake for selecting A_2, expected returns could be maximised by selecting the action A_1. Selecting A_1 avoids the potential of the catastrophic (-800) loss if the true state is X_1, and the benefits of more certainty given sampling are outweighed by the cost of that sample.

Modulation of rumination by an overly deterministic model of the world

In the previous simulations, the agent's model of the observation distribution in each hidden state was consistent with the ground truth. However, a feature of ruminators is that they tend to be overly optimistic about the probability with which ruminating will lead them to discover the true world state (Watkins, 2018). Thus, in our final simulation, we modelled the effect of such erroneous meta-beliefs by generating samples with higher variance than the agent expected. We demonstrate that when beliefs about the breadth of a state's possible observations are mismatched with reality, specifically, when the agent expects more informative observations (narrower observation distributions) than the world provides, this can also lead to increased sampling.

Methods

In this simulation, we generated samples from a new observation distribution, $o(\hat{X}_2) \sim \mathcal{N}(55, \hat{\sigma}^2)$ with $\hat{\sigma}$ adjusted in steps of 5 from 10 to 30. The agent assumed a standard deviation of 15 for observations from both states X_1 and X_2 (Figure 1c shows the values of all actions for this condition). That is, the agent assumed observation distributions that overlapped less than in the previous simulations. Again, in different conditions, the loss for A_2 in X_1 was $-12.5, -100$ or -800.

Results

Unexpectedly varied observations increased sampling. As the standard deviation of $o(\hat{X}_2)$ increased, so did the number of samples required before selecting a terminating action (Figure 4). For example, when the loss for A_2 in X_1 was equal to the loss for A_1 in X_2, the average number of samples before choosing A_2 was 26.45 when the true standard deviation of $o(\hat{X}_2)$ was 20, increasing to 41.81 when the true standard deviation of $o(\hat{X}_2)$ was 40 (for $c = 10^{-5}$) (Figure 4b). When the loss for A_2 increased to -800, extra sampling was more pronounced, with the number of samples taken increasing from 34.37 to 57.63 as the standard deviation of $o(\hat{X}_2)$ increased. This is because the agent had strict decision thresholds given the assumption of a less variable observation distribution, and required many (noisy) samples to reach that desired high level of certainty. And as in previous simulations, as the potential loss increased, more certainty was required.

Discussion

Rumination is a chronic and painful symptom that increases vulnerability to depression, anxiety, and other psychopathologies (Watkins & Roberts, 2020). We suggest a new inference-based framework for understanding ruminative episodes as a normative information-sampling strategy and explore conditions that promote more protracted such sampling. Using a computational model, we show how episodes of rumination can become difficult to terminate due to exaggerated beliefs about potential negative outcomes for one's actions, pessimistic initial beliefs about the state of the world, and when the world is estimated to be more deterministic than it actually is. Thus, normative "rumination" exists, and can be exaggerated by additional sampling biases.

In our first two simulations, the agent had an accurate model of the true hidden state of the environment. These simulations demonstrate that even under these conditions, beliefs about the environment (e.g., the potential for catastrophic losses in X_1) normatively prolong sampling. However, in rumination, we envision sampling as memory recall – drawing on past experiences to try to "think through" a situation to determine its latent state and the resultant correct course of action. Therefore, although in these simulations samples always reduced uncertainty about the true hidden state, in rumination this may not be the case. In particular, during a ruminative episode, memory recall is biased towards negative events (Watkins, 2018). Thus, in our third simulation we modelled a degree of mismatch between the true state of the world and its model within the agent. We demonstrated that when the variance of the observations from the true state is greater than what agents expect, this can lead to protracted sampling. Notably, an extreme degree of mismatch between the samples observed and the samples expected (by the optimal policy) can lead to failure modes, where the agent is never able to reach a decision threshold.

In this paper, the agent's decision to sample is determined by an optimal policy that assumes that all samples are independent, identically distributed, and neutral in valence. Although this is a simplification, it is unclear whether people indeed consider their own memory biases (e.g., sampling predominantly negative events) when evaluating the utility of rumination as a problem-solving strategy. A second simplification we make is that the decision thresholds are calculated before an episode of sampling begins, and do not change during the episode. One alternative is to recalculate the thresholds iteratively once a ruminative episode begins to take into account observed sampling biases. Further extensions of this model should consider the types of biases that may influence sampling from memory, how they can be incorporated into determining the decision threshold, and whether it can be reached during sampling. First, rumination typically involves selectively retrieving negative memories, which can be formalized as memories where the punishing action was chosen, biasing expectations about the reward function of the hidden state. Second, samples from memory are likely to be activated in sequences based on the similarity of various features (Polyn, Norman, & Kahana, 2009; Tomita, Barense, & Honey, 2021), which can be modeled using a dependent sampling process (e.g., a Markov Chain Monte Carlo process, (Gershman, Vul, & Tenenbaum, 2009)).

By demonstrating how a POMDP framework can be adapted to explore rumination, we hope to show the value of adopting existing formalisations to understand clinical phenomena such as repetitive negative thinking. While we focused on a subset of dynamics, our reason for adopting a state-inference based framework was that it can be easily extended to explore other important ruminative dynamics, such as the relationship between concrete (e.g. "*I performed badly on yesterday's exam*") and abstract thought (e.g. "*I am a failure*"). These could be formalised as individual samples and generalised hidden states, respectively. Furthermore, each trial need not be treated as independent (as we have done here). Just as with ruminative episodes, an individual can accrue over trials a long-run estimate of the value of engaging in sampling at all, which may be reflected in the cost of sampling.

Finally, we acknowledge that the conditions we use for these simulations (loss magnitudes, observation distributions) were purposely selected to demonstrate conditions under where extended sampling sequences would be observed. While this represents a highly selective space of the simulations we could have encompassed (e.g., we did not include conditions with increasing rewards), this does not detract from the value this approach may have for researchers to test their hypotheses about how beliefs about states of the world may interact with repetitive thinking.

Acknowledgments

We thank Sam Zorowitz for helpful comments on this manuscript. We thank Peter F Hitchcock, Paul B Sharp, and the Niv Lab for general feedback on the project. This work was supported by NIH Grant R01MH119511.

References

Abela, J. R., Brozina, K., & Haigh, E. P. (2002). An examination of the response styles theory of depression in third-and seventh-grade children: A short-term longitudinal study. *Journal of abnormal child psychology*, *30*, 515–527.

Alloy, L. B., & Ahrens, A. H. (1987). Depression and pessimism for the future: biased use of statistically relevant information in predictions for self versus others. *Journal of personality and social psychology*, *52*(2), 366.

Baker, S. C., Konova, A. B., Daw, N. D., & Horga, G. (2019). A distinct inferential mechanism for delusions in schizophrenia. *Brain*, *142*(6), 1797–1812.

Berg, M., Feldmann, M., Kirchner, L., & Kube, T. (2022). Oversampled and undersolved: Depressive rumination from an active inference perspective. *Neuroscience & Biobehavioral Reviews*, 104873.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(4), 429–453.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*(11), 3612–3628.

Gagne, C., & Dayan, P. (2022). Peril, prudence and planning as risk, avoidance and worry. *Journal of Mathematical Psychology*, *106*, 102617.

Gershman, S., Vul, E., & Tenenbaum, J. (2009). Perceptual multistability as markov chain monte carlo inference. *Advances in neural information processing systems*, *22*.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence*, *101*(1-2), 99–134.

Kuehner, C., & Weber, I. (1999). Responses to depression in unipolar depressed patients: An investigation of nolen-hoeksema's response styles theory. *Psychological medicine*, *29*(6), 1323–1333.

Martin, L. L., & Tesser, A. (1996). *Ruminative thoughts* (Vol. 9). Psychology Press.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on psychological science*, *3*(5), 400–424.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological review*, *116*(1), 129.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tomita, T. M., Barense, M. D., & Honey, C. J. (2021). The similarity structure of real-world memories. *bioRxiv*. doi: 10.1101/2021.01.28.428278

Watkins, E. R. (2018). *Rumination-focused cognitive-behavioral therapy for depression*. Guilford Publications.

Watkins, E. R., & Roberts, H. (2020). Reflecting on rumination: Consequences, causes, mechanisms and treatment of rumination. *Behaviour Research and Therapy*, *127*, 103573.