A Memory-Based Theory of Beliefs
Draft: Preliminary, please do not quote or cite.

Danell Benguigui   Maxime Cugnon de Sévricourt*

1st February, 2024

Abstract

We propose a model of conditional beliefs based on memory. Memory is construed as a count of past realizations of a random variable. All past observations are kept in the long-term memory, but, when asked to evaluate a conditional probability, the agent samples past observations from it. The probability of sampling a relevant experience is increasing in its frequency in the long-term memory and under some assumptions, sampling results in subjective Bayesian beliefs. To capture common belief biases, we let "irrelevant" memories interfere with the recall process, allowing for a broader spectrum of beliefs to be generated. Our main result is the characterization of the conditional beliefs that are consistent with such a recall process.

Keywords: Conditional beliefs, behavioral microeconomics, memory, belief biases, expectations, non-Bayesian updating, information.
JEL classification: D83, D84, D91.

1 Introduction

Since Lucas (1976), economic theory has been marked by a shift towards a rational expectations framework. This paradigm shift steered macroeconomics towards models that are robust to policy changes, but it fails to acknowledge the cognitive biases documented in the literature. The seminal work of Tversky and Kahneman (1974) on the representativeness heuristics in particular brought these biases to the forefront. In response, researchers are working towards integrating these psychological insights into economic theory, (see Benjamin (2018) for a recent survey on belief biases). We contribute to this expanding body of research by introducing a model that explores the role of memory in shaping beliefs. A recent strand of literature focuses on incorporating belief biases, as documented by Tversky and Kahneman (1974), into large-scale macroeconomic models. These studies are pivotal in explaining financial data, credit cycles, and reactions to news. For instance, Gennaioli and Shleifer (2010) develops a theory of inference called 'local thinking,' which captures some form of the representativeness heuristic through selective recall of information. A particular model that generates belief biases that are explained by the representativeness heuristic is known as 'diagnostic expectations' (Bordalo et al., 2018), and has been incorporated into macroeconomic models (Bordalo et al. (2019), Bordalo et al. (2020), Bianchi et al. (2021), L'Huillier et al. (2021)...). Diagnostic expectations is instrumental in explaining why we observe overreactions to news at the individual level and underreactions at the consensus level in the data. Other models, like Bordalo et al. (2022), investigate the role of similarity between hypotheses and data in explaining belief biases. Drawing inspiration from Bordalo et al. (2022), our model differs from theirs on at least 3 counts. First, we investigate the role of past observations on belief formation, whereas they focus on the content of the memories themselves, in particular how similar two different events can be. Second, our framework is more general as we allow for any finite sample space, unlike Bordalo et al. (2022), who restricts their analysis to binary variables. Finally, our aim is to fully characterize the beliefs consistent with our model, while they restrict their analysis to examples that fit their framework and illustrates the specific biases they seek to model.

*Department of Economics, Georgetown University, dcb97@georgetown.edu, mc2534@georgetown.edu

The human brain's cognitive limitations form a cornerstone of our model. Despite the vast storage capacity of its long-term memory, estimated at 2.5 million gigabytes, the brain's working memory can hold only about 7 ± 2 items (Miller, 1956). This gap between storage capacity and working memory is what motivates our focus on the recall process. Our model contends that an understanding of belief formation in economics is intrinsically linked to these memory recall mechanisms, consistent with the brain's cognitive limitations.

In our analysis, beliefs depend on recalled events, and are therefore stochastic. However, we focus on expected beliefs as they offer a more relevant perspective for understanding the aggregate patterns of belief formation under different memory recall conditions. This approach allows us to simplify our analysis while getting rid of variations that are only introduced by the sampling process and do not seem particularly relevant to our analysis.

Our model aims at capturing intuitive properties of the recall process and leads to an extended definition of rational beliefs. While beliefs consistent with the model have to obey some monotonicity conditions that we detail thereafter, they do not have to satisfy some conditions of rationality, such as summing to unity or satisfying Bayes'rule.

Our investigation is driven by a central research question: What beliefs are consistent with the limited recall of past observations? This question is pivotal in exploring the relationship between memory recall and belief formation.

In pursuing this, we delve into the specifics of our model, examining different recall processes and their impact on belief formation. This exploration allows us to understand how different limitations and capacities of memory recall affect the evolution and structure of beliefs in economic agents.

In our model, when an agent is asked about their conditional belief of an event A given B, they try to remember past observations. We model this process as past observations competing for slots in the working memory, which has a fixed limit. The probability that a type of event enters the working memory depends on the number of past observations that confirm or disconfirm the hypotheses. We also allow for irrelevant observations to interfere in the recall process, although such observations are discarded. This idea of competing memories allows us to model each recall attempt as a contest between players, where the effort exerted by the players is the number of observations of different types. We can therefore interpret the function linking past observations and probability of recall as a Contest Success Function (CSF), properties of which have been extensively studied in the literature (Skaperdas, 1996; Corchón and Dahm, 2010). Hence, we start our analysis by building a general model of recall, then we show that if the recall probability function displays the standard properties of contest success functions, then the agent is Bayesian and finally we investigate non-Bayesian beliefs by relaxing some of these assumptions in an intuitive manner, and characterize the beliefs that the model can accommodate in that case.

2 Model

We propose a model (following Bordalo et al. (2022)) of belief formation based on past observations. When asked about their beliefs, the agent retrieve observations from their long-term memory to form conditional beliefs. We let Ω be a finite sample space, and $\Sigma \equiv 2^{\Omega}$ be its power set. Because Ω is finite, Σ is a σ-algebra. We also let $\Sigma^* \equiv 2^{\Omega} \backslash \{\emptyset\}$. We assume there exists a revelation mechanism so that the econometrician observes the conditional beliefs of an agent over Σ. Formally, we define conditional beliefs as the following:

Definition 1. *A function $b : \Sigma \times \Sigma^* \to [0, 1]$ is called a conditional belief function.*

We assume that the agent has observed N realizations of events, stored in their long-term memory. We can write their memory database as a function $m : \Sigma \to \mathbb{Z}_+$ where $m(A)$ represents the number of realizations equal to members of A recorded in the long-term memory of the agent. $m(\Omega) = N$. We also set $m(\{\emptyset\}) = 0$. For A and B disjoint, we also have $m(A \cup B) = m(A) + m(B)$. The memory database is the count of realizations that correspond to the events from the sample space that are evaluated. Consider an example: A is the event "price of wheat high" and B is the event "there is a drought", then $m(A)$ is the number of times the agent has observed a high price of wheat in the past, $m(B)$ is the number of realizations of droughts observed by the agent, and $m(A \cap B)$ is the count of simultaneous observations of a high price of wheat and a drought. In this model, as we look at conditional beliefs, the object of interest is the formed conditional belief of some event $A \in \Sigma$ conditional on $B \in \Sigma^*$. Note that if $B = \Omega$, then we are studying the unconditional belief over A.

The next part of the model is the working memory. When the agent is asked to report $b(A|B)$, we assume past observations that come to mind may belong to one of these events: $A \cap B$, $A^c \cap B$, $A \cap B^c$ and $A^c \cap B^c$. Together, they define the cross partition of the collections $\{A, A^c\}$ and $\{B, B^c\}$ that we denote:

$$\mathcal{A}_{A,B} = \{A \cap B, A^c \cap B, A \cap B^c, A^c \cap B\}$$

Note that $\mathcal{A}$ can contain at most 3 empty elements, for a given non-empty Ω. Consider our wheat price example, then when asked to report their belief about a high price of wheat given that there is a drought, the agent may recall events when the price was high and there was a drought, when the price was not high and there was a drought, when the price was high for other reasons than a drought, and when the price was not high and there was no drought. We can split this cross partition in two: relevant information and irrelevant information:

Relevant information	Irrelevant information
$A \cap B$	$A \cap B^c$
$A^c \cap B$	$A^c \cap B^c$

Traditional Bayesian updating requires that

$$p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{p(A \cap B)}{p(A \cap B) + p(A^c \cap B)}. \tag{1}$$

In other words, knowledge about the relative frequencies of the events $A \cap B$ and $A^c \cap B$ is sufficient to compute the conditional probability of A given B. This is why we call those events "relevant information". However, Kahana (2012) shows that interference of irrelevant data is an essential concept to understand the inner workings of human memory. Hence, we allow for $A \cap B^c$ and $A^c \cap B^c$ to have an effect on the probability of recall of A given B. When trying to recall an instance of an increase in the price of wheat when there was a drought, instances of price increases of wheat in the absence of a drought may come to mind.

When asked to report a belief, the agent samples psst observations from their memory database to their working memory. Our approach here differs from Bordalo et al. (2022) as we fix a limit to the number of sampling trials that can be done. According to Miller (1956), working memory only allows for a very limited number of items, which motivates our imposition of a fixed number of trials $n \in \mathbb{N}$, the capacity of the working memory.

Let $\mathcal{N} = \{0, 1, \ldots, N\}$ be a finite set in $\mathbb{Z}_+$ that represents all the possible counts of realizations in the long-term memory of an agent, with $N = m(\Omega)$. Let $\aleph_{A,B} = (m(A \cap B), m(A^c \cap B), m(A \cap B^c), m(A^c \cap B^c))$ be the element-wise application of $m(\cdot)$ to $\mathcal{A}$. $\aleph_{A,B}$ represents the vector of counts for each element of $\mathcal{A}$. Let $p^i(\aleph_{A,B})$ be an arbitrary contest success function $p^i : \mathcal{N}^4 \mapsto \mathbb{R}_+$ that yields the probability of the i-th element of $\aleph_{A,B}$ being successfully sampled in the working memory.

In the language of contest success functions, $m(X_i)$ is the effort exerted by player i and p^i is the probability that player i wins the contest. We consider the sub-contest between relevant information sets, $p^i_K(\aleph_{A,B})$ the probability of the i-th element of $\aleph_{A,B}$ to be successfully sampled, when only the elements $k \in K$ are competing. The irrelevant information sets may influence the result of this sub-contest but may not be sampled. Let $W_{A|B}$ be the number of successful recalls of A conditional on B. Then we assume:

$$W_{A|B} \sim \text{Binomial}\left(n, p^1_K(\aleph_{A,B})\right)$$

We sample events such that, if A conditional on B fails to be recalled, then A^c conditional on B is recalled and takes a spot in the working memory. We thus have that $W_{A^c|B} = n - W_{A|B}$ where

$W_{A^c|B}$ also follows a binomial distribution. After resolving successes and failures of recall, the agent forms a conditional belief Π such that

$$\Pi_{A|B} = \frac{W_{A|B}}{W_{A|B} + W_{A^c|B}} = \frac{W_{A|B}}{n}$$

As $W_{A|B}$ follows a binomial distribution, we have:

$$\mathbb{E}[W_{A|B}] = np_K^1(\aleph_{A,B})$$

. Thus, we can obtain the expected conditional belief function:

$$\mathbb{E}[\Pi_{A|B}] = p_K^1(\aleph_{A,B}) \equiv b(A|B)$$

Our focus is on expected beliefs, we leave the variations introduced by the sampling process to future work. Our next step is to study different models of recall and look at their implications for the beliefs.

3 Results

3.1 Subjective Bayesianism

In this section, we study a simplified recall model, where recall only depends on one elementary recall function, and we examine its implications for the beliefs. We introduce 3 axioms on the expected beliefs of agents.

Axiom 1 (Unitarity of unconditional beliefs)**.** *We say that b satisfies unitarity of unconditional beliefs if $b(\Omega|\Omega) = 1$.*

This axiom is equivalent to the second Kolmogorov axiom of probability. Note that non-negativity is already implied by our definition of a conditional belief function.

Axiom 2 (Finite additivity of unconditional beliefs)**.** *Let $A_1, A_2, \ldots, A_N$ be any disjoint sets of Σ, then we say that b satisfies finite additivity of unconditional beliefs if*

$$b\left(\bigcup_{i=1}^{N} A_i \middle| \Omega\right) = \sum_{i=1}^{N} b(A_i|\Omega)$$

This is the third Kolmogorov axiom, stating that the belief of a union of disjoint sets is the sum of beliefs of each disjoint set. The reported probability of observing a high price of wheat and a medium price of wheat is equal to the probability of observing a high price plus the probability of observing a medium price, under this axiom.

Axiom 3 (Bayesian conditional beliefs)**.** *We say that b is Bayesian with respect to unconditional beliefs if for all sets $A \in \Sigma, B \in \Sigma^*$*

$$b(A|B) = \frac{b(A \cap B|\Omega)}{b(B|\Omega)}$$

This last axiom says that beliefs must satisfy Bayes rule with respect to unconditional beliefs. Combining those 3 axioms imposes standard assumptions on the belief function: it must be a probability distribution, and must be Bayesian. However, those conditions do not imply that the unconditional probabilities, or Bayesian priors, are correct with respect to some objective probabilities. Hence, a distortion could be applied to unconditional beliefs and the conditional beliefs would all satisfy those axioms while yielding wrong results compared to an objective probability. We can now define a specific contest success function

$$\tilde{p}_K^i(\aleph_{A,B}) = \frac{\sum_{x \in A \cap B} r(x)}{\sum_{y \in B} r(y)}$$

where $r : \Omega \mapsto \mathbb{R}_+$ is called an elementary recall function. We choose this form for the recall process as it only depends on one single univariate function of the sample space, which means that each elementary event will have its own "strength of recall". It does not allow for recalls specific to certain subsets.

Definition 2 (r generates b)**.** *An elementary recall function $r : \Omega \mapsto \mathbb{R}_+$ generates belief b if for all $A \in \Sigma$, and $B \in \Sigma^*$, we have:*

$$b(A|B) = \frac{\sum\limits_{x \in A \cap B} r(x)}{\sum\limits_{y \in B} r(y)}$$

Theorem 1. *The following are equivalent:*

1. *There exists $r : \Omega \to \mathbb{R}_+$ that generates b.*
2. *b satisfies unitarity of unconditional beliefs, finite additivity of unconditional beliefs and Bayesian conditional beliefs.*

Proof. In Appendix A. □

This result tells us that beliefs which follow this recall process must be subjective Bayesian. In particular, if conditional beliefs do not reflect the true conditional probabilities, then the unconditional beliefs must also disagree with unconditional probabilities. Note that this recall process excludes the possibility of interference of irrelevant information. Hence, belief biases such as overreaction to characteristic events are not fully characterized by subjective Bayesianism. See Appendix B for more details on this matter.

Corollary 1. *r is unique up to a positive linear transformation.*

Proof. In Appendix A. □

This result tells us that any r that generates b belongs to a unique family of linear transformations that all generate b, or in other words, the beliefs are insensitive to the absolute counts of each events, only the relative counts matter.

Corollary 2. *Let b be conditional beliefs generated by some recall function r. $b(x|\Omega) = m(x)/m(\Omega)$ if and only if $r(x) = \alpha m(x)$ for some $\alpha > 0$.*

Proof. In Appendix A. □

Unconditional beliefs must reflect memory frequencies from the long-term database if and only if the recall of x is proportional to the number of times x was observed. This result is important as in the long-run, we can assume that agents observe enough realization so that the frequencies of the long-term memory converge to the true probabilities, but we still observe biases in the beliefs.

3.2 Interference of Irrelevant Data

We have seen that subjective Bayesian beliefs can be represented by an elementary recall function $r : \Omega \mapsto \mathbb{R}_+$. However, as discussed in the introduction, belief biases can also come from interference of irrelevant data. Bordalo et al. (2022) find strong supporting evidence of this claim when tested with experimental data. To allow for this property of the recall process, we must think about the properties of $p^i(\aleph_{A,B})$. We make 4 natural assumptions on this recall probability function and then look at their implications for the beliefs.

Assumption 1. *$\sum_{i=1}^{4} p^i(\aleph_{A,B}) = 1$ and $p^i(\aleph_{A,B}) \geq 0$ for all $i \in \{1, \ldots, 4\}$ and all $\aleph_{A,B}$. For some event $X_i \in \mathcal{A}$, if $m(X_i) > 0$, then $p^i(\aleph_{A,B}) > 0$.*

This condition says that the CSF must satisfy the properties of a probability distribution and when a player exerts a strictly positive effort, then their probability of winning should also be strictly positive.

Assumption 2. *For all $i \in \{1, \ldots, 4\}$, $p^i(\aleph_{A,B})$ is non decreasing in $m(X_i)$ and nonincreasing in $m(X_j)$, for all $j \neq i$.*

A player's probability of success is nondecreasing in their own effort but nonincreasing in every other player's effort. In our model, the probability of recalling A given B should be nondecreasing in the number of observations of $A \cap B$ but nonincreasing in the number of observations of $A^c \cap B$ for example.

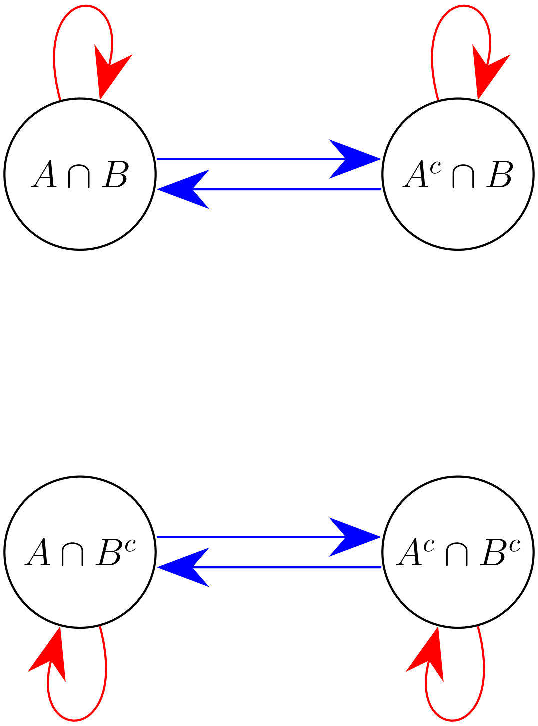

Figure 1: Graphical representation of Assumption 3

Assumption 3. *For any permutation π that is identity ($\pi(x_i) = x_i \forall i$) or $\pi(x_i) = x_j \Rightarrow \pi(x_j) = x_i$ such that $\pi(x_1) = x_2 \Rightarrow \pi(x_3) = x_4$; we have: $p^{\pi(i)}(\aleph_{\pi(X)}) = p^i(m(X_{\pi(1)}), \ldots, m(X_{\pi(4)})), \forall i \in \{1, 2, 3, 4\}$*

This relaxes the usual anonymity assumption of contest success functions: only specific permutations are anonymous. Specifically, only permutations of $m(A \cap B)$ with $m(A^c \cap B)$. In our model, full anonymity would imply that the probability of recalling events corresponding to $A \cap B$ needs to be nonincreasing in the same way in $A^c \cap B^c$ as in $A^c \cap B$, but this is not very intuitive in our model: different irrelevant data can impact the probability differently. Arrows indicate anonymous permutations in the contest success function. In red, we have the condition that identity permutations are anonymous, and in blue, if the first element of $\aleph_{A,B}$ is permutated with the second element, then the third has to be permutated with the fourth to preserve anonymity.

Assumption 4. $p^i_K(\aleph_{A,B}) = \dfrac{p^i(\aleph_{A,B})}{\sum_{k \in K} p^k(\aleph_{A,B})}, \forall i \in K$ *and* $\forall K \subseteq \{1, 2, 3, 4\}$ *with* $|K| \geq 2$.

This allows a sub-contest to be run with properties consistent with the larger contest. When we evaluate the probability of recalling A given B in our model, we are effectively running a sub-contest where $K = \{1, 2\}$, the effort of player 1 is $m(A \cap B)$ and the effort of player 2 is $m(A^c \cap B)$.

Now that we have discussed the assumption on the recall process, we study their empirical content for the beliefs, by introducing a set of axioms.

Axiom 4 (Unitarity of Complements). *For all A and $B \subseteq \Omega$:*

$$b(A|B) + b(A^c|B) = 1 \tag{2}$$

We do not require unconditional beliefs to sum to one but we require their complements to sum to one. This axiom is a consequence of the anonymity condition of the recall process. $p^1_K(\aleph_{A,B})$ must be equal to $p^2_K(\aleph_{A^c,B})$.

Axiom 5 (Subset Certainty)**.** *$\forall A, B \subseteq \Omega$, if $B \subseteq A$, then $b(A|B) = 1$*

Definition 3 (Covering Order)**.** *We say A' is greater than A in the B-covering order, and we write $A' \succeq_B A$, if $A \subseteq A'$ and $A' \setminus A \subseteq B$.*

In other words, A' is greater than A in the B-covering order if A' contains "more of" B than A, and elements not in B are held constant.

Axiom 6 (Monotonicity in Covering Orders)**.** *Let A, A', B, and B' be events in Σ. If $A' \succeq_B A$, and $B' \succeq_A B$ then:*

$$b(A|B) \leq b(A'|B') \tag{3}$$

Axiom 7 (Union-Preservation of Ordering)**.** *For all events A, A', B and C such that $A, A' \subseteq B$, $C \subseteq B \setminus (A \cup A')$, we have that if $b(A|B) \leq b(A'|B)$, then $b(A \cup C|B) \leq b(A' \cup C|B)$.*

This axiom imposes that if, given the same conditional B, a belief is greater than another, then this order should be preserved when adding more events (the same ones) from B into the evaluated events.

Proposition 1. *If p and m generate b, then b satisfies unitarity of complements, subset certainty, monotonicity in covering orders and union preservation of ordering.*

Proof. In Appendix C □

Conjecture 1. *If b satisfies unitarity of complements, subset certainty, monotonicity in covering orders and union-preservation of ordering, then there exist p and m that generate b.*

Proof. In progress. □

Proposition 1 together with Conjecture 1 constitute our representation theorem.

4 Conclusion

In this paper, we develop a model of belief formation through memory, where the recall process plays a central role. We characterize what beliefs are consistent with the process we posited. Then, we discussed the importance of departures from Bayesianism through interference of irrelevant data and developed assumptions on the recall process and attempt to characterize their empirical content for the induced beliefs. Our next step is to prove the converse of Proposition 1, which is in progress. Then, we aim to verify which common belief biases could be captured by this model. We believe that our model could be useful to implement in financial applications, as well as generally in macroeconomic models that are sensitive to individual beliefs.

Appendix A: Proofs from Section 3.1

Theorem 1. The following are equivalent:

1. There exists $r : \Omega \to \mathbb{R}_+$ that generates b.
2. b satisfies unitarity of conditional beliefs, finite additivity of unconditional beliefs and Bayesian conditional beliefs.

Proof. Consider $A \in \Sigma$ and $B \in \Sigma^*$, we want to show that r generates b if and only if b satisfies axioms 1-3.

1. We can start by assuming r generates conditional belief b:

$$b(\Omega|\Omega) = \frac{\sum_{x \in \Omega} r(x)}{\sum_{y \in \Omega} r(y)} = 1$$

Thus satisfying axiom 1.

Now consider $A_1, A_2, \ldots A_N$ sequence of N disjoint sets in Σ, then:

$$\begin{aligned}
b\left(\bigcup_{i=1}^{N} A_i \middle| \Omega\right) &= \frac{\sum_{i=1}^{N}\sum_{x\in A_i} r(x)}{\sum_{y\in\Omega} r(y)} \\
&= \sum_{i=1}^{N} \frac{\sum_{x\in A_i} r(x)}{\sum_{y\in\Omega} r(y)} \\
&= \sum_{i=1}^{N} b(A_i|\Omega)
\end{aligned}$$

which proves that axiom 2 holds. Finally, using the following:

$$b(A|B) = \frac{\sum_{x\in A\cap B} r(x)}{\sum_{y\in B} r(y)}, \qquad b(B|A) = \frac{\sum_{x\in A\cap B} r(x)}{\sum_{y\in A} r(y)}$$

$$b(A|\Omega) = \frac{\sum_{x\in A} r(x)}{\sum_{y\in\Omega} r(y)}, \qquad b(B|\Omega) = \frac{\sum_{x\in B} r(x)}{\sum_{y\in\Omega} r(y)}$$

We can compute:

$$\begin{aligned}
\frac{b(B|A)b(A|\Omega)}{b(B|\Omega)} &= \frac{\left(\sum_{x\in A\cap B} r(x)\right)\left(\sum_{x\in A} r(x)\right)\left(\sum_{y\in\Omega} r(y)\right)}{\left(\sum_{y\in A} r(y)\right)\left(\sum_{y\in\Omega} r(y)\right)\left(\sum_{x\in B} r(x)\right)} \\
&= \frac{\sum_{x\in A\cap B} r(x)}{\sum_{y\in B} r(y)} = b(A|B)
\end{aligned}$$

and thus axiom 3 holds.

2. Now, we want to show that if b satisfies axioms 1-3, then there exists $r : \Omega \to \mathbb{R}_+$ that generates b. Let b be conditional beliefs satisfying axioms 1-3, then let us define r so that $r(x) \equiv b(x|\Omega)$ for all $x \in \Omega$. We will first show that r generates unconditional beliefs $b(A|\Omega)$ by induction on the cardinality of A, and then show that it also generates all the conditional beliefs as well. Let $A \subseteq \Omega$. For $|A| = 1$, letting $x \in A$, r generates b if

$$b(A|\Omega) = \frac{r(x)}{\sum_{y\in\Omega} r(y)}$$

by definition, we have:

$$b(A|\Omega) = \frac{r(x)}{\sum_{y\in\Omega} b(y|\Omega)}$$

applying finite additivity, we get:

$$\begin{aligned}
b(A|\Omega) &= \frac{r(x)}{b\left(\bigcup_{y\in\Omega}\{y\}\middle|\Omega\right)} \\
&= \frac{r(x)}{b(\Omega|\Omega)}
\end{aligned}$$

Using unitarity of unconditional beliefs, we have:

$$b(A|\Omega) = r(x)$$

as $b(\Omega|\Omega) = 1$

and thus r generates unconditional beliefs $b(A|\Omega)$. Now suppose that for all $A \subseteq \Omega$ with $|A| = k$, r generates $b(A|\Omega)$, that is :

$$b(A|\Omega) = \sum_{x\in A} r(x)$$

Now, consider a set $A' \subseteq \Omega$ of cardinality $|A'| = k + 1$, we need to show:

$$b(A'|\Omega) = \sum_{x\in A'} r(x)$$

because A' is of cardinality k, there exists $A \subseteq \Omega$ of cardinality k and $x \in \Omega$ such that $A' = A \cup \{x\}$ and by axiom 2, we have:

$$\begin{aligned}
b(A'|\Omega) &= b(x|\Omega) + b(A|\Omega) \\
&= r(x) + \sum_{y\in A} r(y) && \text{(by our induction hypothesis)} \\
&= \sum_{y\in A'} r(x)
\end{aligned}$$

to complete the proof, we consider any conditional belief $b(A|B)$ with $|B|$ an arbitrary number. Then, by using our definition for $r(x)$ we can write:

$$\begin{aligned}
\frac{\sum\limits_{x\in A\cap B} r(x)}{\sum\limits_{y\in B} r(y)} &= \frac{\sum\limits_{x\in A\cap B} b(x|\Omega)}{\sum\limits_{y\in B} b(y|\Omega)} \\
&= \frac{b(A\cap B|\Omega)}{b(B|\Omega)}
\end{aligned}$$

As b satisfies axiom 3, we have:

$$\frac{\sum\limits_{x\in A\cap B} r(x)}{\sum\limits_{y\in B} r(y)} = b(A|B)$$

which completes the proof.

□

Corollary 1. r is unique up to a positive linear transformation.

Proof. Let r and $\tilde{r}$ both generate conditional beliefs b. We want to show that there exists $\alpha > 0$ such that $r(x) = \alpha\tilde{r}(x)$ for all $x \in \Omega$. To that effect, let $x \in \Omega$. Because both r and $\tilde{r}$ represent b, we have:

$$\frac{r(x)}{\sum\limits_{y\in\Omega} r(y)} = \frac{\tilde{r}(x)}{\sum\limits_{y\in\Omega} \tilde{r}(y)} = b(x|\Omega)$$

Rearranging, we get:

$$r(x) = \frac{\sum\limits_{y\in\Omega} r(y)}{\sum\limits_{y\in\Omega} \tilde{r}(y)}\tilde{r}(x)$$

Letting $\alpha = \frac{\sum_{y\in\Omega} r(y)}{\sum_{y\in\Omega} \tilde{r}(y)}$ proves the claim. □

Corollary 2. Let b be conditional beliefs generated by some recall function r. $b(x|\Omega) = m(x)/m(\Omega)$ if and only if $r(x) = \alpha m(x)$ for some $\alpha > 0$.

Proof.

1. Assume $b(x|\Omega) = m(x)/m(\Omega)$ for all $x \in \Omega$. Then, because r generates b, we have, for all $x \in \Omega$:

$$\frac{r(x)}{\sum\limits_{y\in\Omega} r(y)} = \frac{m(x)}{m(\Omega)}$$

   Equivalently, we can write:

$$r(x) = \frac{m(x)}{m(\Omega)} \sum_{y\in\Omega} r(y)$$

   which proves the first part of the claim.

2. Let us assume there exists $\alpha > 0$ such that $r(x) = \alpha m(x)$. We have:

$$\begin{aligned} b(x|\Omega) &= \frac{r(x)}{\sum\limits_{y\in\Omega} r(y)} \\ &= \frac{\alpha m(x)}{\sum\limits_{y\in\Omega} \alpha m(y)} \\ &= \frac{m(x)}{m(\Omega)} \end{aligned}$$

□

Appendix B: Characteristic Overreaction

Following the definition of diagnostic expectations from Bordalo et al. (2018), we write a more general formulation called "characteristic overreaction".

Definition 4 (Characteristic Event)**.** *An event $A \in \Sigma$ is characteristic of $B \in \Sigma^*$ if*

$$\frac{m(A \cap B)}{m(B)} > \frac{m(A \cap (B^C))}{m(B^C)}$$

Definition 5 (Overreaction)**.** *A conditional belief $b(A|B)$ is an overreaction with respect to m if*

$$b(A|B) > \frac{m(A \cap B)}{m(B)}$$

Definition 6 (Characteristic Overreaction)**.** *A conditional belief function $b : \Sigma \times \Sigma^* \to [0, 1]$ exhibits characteristic overreaction with respect to some memory database m, if, for all $A \in \Sigma$ and $B \in \Sigma^*$, the following is true:*

$$\frac{m(A \cap B)}{m(B)} > \frac{m(A \cap (B^C))}{m(B^C)} \Rightarrow b(A|B) > \frac{m(A \cap B)}{m(B)}$$

This idea captures the diagnostic expectations dynamic of overreacting to news that are characteristic of a sample. We show that this idea is not compatible with our recall process from Theorem 1.

Proposition 2. *If an agent is subjective Bayesian, then they cannot display overreaction to characteristic events.*

Proof.

1. Let A and B be two events such that A is characteristic of B, and $B \neq \emptyset$ and $A \neq B$.
2. Because the DM is Bayesian, we have that: $b(A|B) = b(A \cap B|B)$.
3. For the DM to display characteristic over-reaction, it must be that $b(A|B) > p(A|B)$.
4. Consider event $C = A \cup B^C$. C cannot be characteristic of B, because $p(C|B) = p(A \cup B|B) < 1$ and $p(C|B^C) = p(B^C|B^C) = 1$.
5. But then because the DM is Bayesian, $b(C|B) = b(A \cap B|B) > p(A|B) = p(C|B)$, which implies that the DM is overreacting to C!
6. This implies that the DM is not overreacting to C^C given B, which is a characteristic event.
7. Therefore, the DM cannot display overreaction to characteristic events.

□

Appendix C: Proofs from Section 3.2

Proposition 1. If p and m generate b, then b satisfies unitarity of complements, subset certainty, monotonicity in covering orders and union-preservation of ordering.

Proof.

1. Let b be some conditional beliefs, and p and m be a CSF and a memory database that generate b.
2. Let us show that b satisfies unitarity of complements. Let A and B be two events. We have:

$$\begin{aligned} b(A|B) + b(A^C|B) &= \frac{p_1(\aleph_{A,B})}{p_1(\aleph_{A,B}) + p_2(\aleph_{A,B})} + \frac{p_1(\aleph_{A^C,B})}{p_1(\aleph_{A^C,B}) + p_2(\aleph_{A^C,B})} \\ &= \frac{p_1(\aleph_{A,B}) + p_2(\aleph_{A^C,B})}{p_1(\aleph_{A,B}) + p_2(\aleph_{A^C,B})} && \text{(by the properties of } p\text{)} \\ &= 1. \end{aligned}$$

3. Let us show that b satisfies Subset Certainty. Let A and B be two events such that $A \subseteq B$. We have:

$$\begin{aligned} b(A|B) &= 1 - b(A^C|B) \\ &= 1 - \frac{p_1(\aleph_{A^C,B})}{p_1(\aleph_{A^C,B}) + p_2(\aleph_{A^C,B})} \\ &= 1 - 0 = 1. && \text{(as } A^C \cap B = \emptyset\text{)} \end{aligned}$$

4. Let us show that b satisfies monotonicity in covering orders.

   (a) Let A, A', B, and B' be events such that $A' \succeq_B A$ and $B' \succeq_A B$.

   (b) First, we will show that $A \cap B \subseteq A' \cap B'$. Let $\omega \in A \cap B$. Because $A \subseteq A'$, $\omega \in A'$. Moreover, because $A' \setminus A \subseteq B$, we must have $\omega \in B$ and so, $\omega \in B'$.

   (c) Second, we want to show that $A'^C \cap B' \subseteq A^C \cap B$. Let $\omega \in A'^C \cap B'$. This implies that $\omega \notin A'$ and thus $\omega \notin A$. On the other hand, $\omega \in B'$, and so, by $B' \succeq_A B$, $\omega \notin B' \setminus B$, otherwise $B' \succeq_A B$ would imply $\omega \in A$, a contradiction. But then we must have $\omega \in B$.

   (d) Third, we need to show that $A' \cap B'^C \subseteq A \cap B^C$. Let $\omega \in A' \cap B'^C$. By $B \subseteq B'$, $\omega \in B'^C$ implies $\omega \in B^C$. But then, we have that $\omega \in A'$ and $\omega \notin B$, and so $\omega \in A$, otherwise a contradiction with $A' \setminus A \subseteq B$ would arise.

   (e) Fourth, let us show what $A'^C \cap B'^C \subseteq A^C \cap B^C$. Let $\omega \in A'^C \cap B'^C$. For the sake of contradiction, assume $\omega \notin A^C \cap B^C$. If $\omega \notin A^C$, then $\omega \in A$. But then, $\omega \in A$ and $\omega \notin A'$, a contradiction. Alternatively, if $\omega \notin B^C$, this implies that $\omega \notin B$, a contradiction with $\omega \notin B$. Therefore, $\omega \in A^C \cap B^C$

   (f) From the four previous points, we can conclude:

$$\begin{aligned} m_{A \cap B} &\leq m_{A' \cap B'} \\ m_{A^C \cap B} &\geq m_{A'^C \cap B'} \\ m_{A \cap B^C} &\geq m_{A' \cap B'^C} \\ m_{A^C \cap B^C} &\geq m_{A'^C \cap B'^C} \end{aligned}$$

   (g) Therefore, by the properties of p, it must be that:

$$p_1(\aleph_{A,B}) \leq p_1(\aleph_{A',B'}).$$

   Or, in other words, $b(A|B) \leq b(A'|B')$.

5. Let us show that b satisfies union-preservation of ordering.

   (a) Let A, A', B and C be events such that $A, A' \subseteq B$, $C \subseteq B \setminus (A \cup A')$.

   (b) Because $A, A' \subseteq B$, we have that $A \cap B^C = A' \cap B^C = \emptyset$, and $A^C \cap B^C = A'^C \cap B^C = B^C$.

   (c) Therefore, $b(A|B) \leq b(A'|B)$ implies:

$$p(m_{A \cap B}, m_{A^C \cap B}, 0, m_{B^C}) \leq p(m_{A' \cap B}, m_{A'^C \cap B}, 0, m_{B^C})$$

   (d) But then, by the properties of p, this implies:

$$m(A \cap B) \leq m(A' \cap B) \tag{4}$$
$$m(A^C \cap B) \geq m(A'^C \cap B) \tag{5}$$

   (e) In turn, because $C \subseteq B \setminus (A \cup A')$, this implies:

$$m((A \cup C) \cap B) \leq m((A' \cup C) \cap B) \tag{6}$$
$$m((A \cup C)^C \cap B) \geq m((A' \cup C)^C \cap B) \tag{7}$$
$$m((A \cup C) \cap B^C) = m((A \cup C) \cap B^C) = 0 \tag{8}$$
$$m((A \cup C)^C \cap B^C) = m((A \cup C)^C \cap B^C) = m(B^C) \tag{9}$$

   (f) And so, again by property of p, we have that:

$$b(A \cup C|B) \leq b(A' \cup C|B)$$

   as desired.

□

References

Benjamin, D. J. (2018, October). Errors in probabilistic reasoning and judgment biases. Working Paper 25200, National Bureau of Economic Research.

Bianchi, F., C. L. Ilut, and H. Saijo (2021, March). Diagnostic business cycles. Working Paper 28604, National Bureau of Economic Research.

Bordalo, P., K. Coffman, N. Gennaioli, F. Schwerter, and A. Shleifer (2019, March). Memory and representativeness. Working Paper 25692, National Bureau of Economic Research.

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer (2022, 08). Memory and Probability*. *The Quarterly Journal of Economics 138*(1), 265–311.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020, September). Overreaction in macroeconomic expectations. *American Economic Review 110*(9), 2748–82.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *Journal of Finance 73*(1), 199–227.

Brandenburger, A. and E. Dekel (1993, February). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory 59*(1), 189–198. Copyright: Copyright 2017 Elsevier B.V., All rights reserved.

Caplin, A. and J. V. Leahy (2019, March). Wishful thinking. Working Paper 25707, National Bureau of Economic Research.

Chambers, C. P., Y. Masatlioglu, and C. Raymond (2023). Coherent distorted beliefs.

Coibion, O. and Y. Gorodnichenko (2015, August). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Corchón, L. and M. Dahm (2010). Foundations for contest success functions. *Economic Theory 43*(1), 81–98.

Gennaioli, N. and A. Shleifer (2010). What comes to mind. *Quarterly Journal of Economics 125*(4), 1399–1433.

Grether, D. M. (1980, 11). Bayes Rule as a Descriptive Model: The Representativeness Heuristic*. *The Quarterly Journal of Economics 95*(3), 537–557.

Kahana, M. (2012). *Foundations of Human Memory.* Oxford University Press, USA.

L'Huillier, J.-P., S. R. Singh, and D. Yoo (2021, February). Incorporating Diagnostic Expectations into the New Keynesian Framework. Working Papers 339, University of California, Davis, Department of Economics.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy 1*, 19–46.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review 101*(2), 343–352.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica 29*(3), 315–335.

Skaperdas, S. (1996). Contest success functions. *Economic Theory 7*(2), 283–290.

Tversky, A. (1977). Features of similarity. *Psychological Review 84*(4), 327–352.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124–1131.

Yegane, E. (2022). Stochastic choice with limited memory. *Journal of Economic Theory 205*, 105548.

Zhao, C. (2018). Representativeness and similarity. Working paper, Princeton University.