A Memory Based Model of Bounded Rationality

Sendhil Mullainathan

published in Quarterly Journal of Economics, 2002

January 16, 2012
Introduction

- Can memory imperfections help explain observed economic behavior?
- Casual observation suggests that recollections shape beliefs.
- An individual forecasting her income uses not only aggregate unemployment data (hard information) but also specific information drawn from more personal experiences.
  - the anecdotes of a recently unemployed friend.
  - the news about a foreign firm’s plans to enter her industry.
- The physiological limitations of human memory about qualitative information shed light on biases in the inference process.
Introduction

- The sources of bounded rationality can be diverse, and there are potentially many different approaches to it.
- An advantage of focusing on memory is that scientific research on it is far more advanced than research on higher-order cognitive functions such as problem solving techniques.
- This permits a model more grounded in scientific evidence.
- Two stylized facts are employed to model the memory technology.
  - Rehearsal: remembering an event once makes it easier to remember that again.
  - Associativeness: similarity of the memory to current events facilitates recall.
A model of bounded rational agents faces the same problem as the one of hyperbolic discounting.

An agent may or may not know that her perception is biased in a particular way – sophisticated or naive.

Each of the decision rules has its appeal and a characterization of both is undoubtedly necessary.

As a first step, this paper investigates the case of naive agents to draw the implications of limited memory.
Introduction

- Associativeness in recall generates a central property of beliefs: an event affects beliefs not only through the information it conveys but also through the memories it evokes.
- This property implies that even completely uninformative signals can influence beliefs by altering the set of recalled memories.
- Even if one disregards these signals as noise, they have an indirect effect by altering the perceptions of the past.
- The added weight of these triggered memories leads to an over-reaction to news.
Introduction

- Another key component of the model is rehearsal which generates the persistence of evoked memories.
- Even if the information in the original event has been discredited, the memories it triggered continue to be more memorable and hence continue to influence beliefs.
- Individuals under-react to news that invalidates or revises old information.
Consider an individual who forms expectations about a state variable.

For the sake of the analysis, the state variable is taken to be equivalent with permanent income which moves for a variety of reasons.

Forming forecasts thus requires combining a diverse set of information.

Some of this information is “hard” or readily available in records: income in prior months, unemployment rate or GDP.

Other information is “soft” or harder to capture in records: a friend in a similar position being fired or a boss telling that you are one of the best employees.
Let $y_t$ be the income at time $t$ which follows

$$y_t = \sum_{k=1}^{t} \nu_k + \varepsilon_t,$$

where $\varepsilon_t$ is a transitory shock distributed $N(0, \sigma^2_x)$ and $\nu_k$ is a permanent shock.

$y_t$ is observed by the individual and represents the hard information.

Each period with probability $p$, an event $e_t$ occurs, which will be the soft information.

Each event has two components: an informative component $x_t$ and an uninformative one $n_t$. 
Setup

- When there is no event, \( e_t = \emptyset \) and will generally take \( x_t = 0 \) and \( n_t = 0 \).
- Conditional on an event occurring, they are distributed
  
  \[ e_t = (x_t, n_t) \sim F(x_t, n_t), \]

  where \( E(x_t) = 0 \) and \( E(n_t) = 0 \).
- The covariance \( \sigma_{xn} \) measures whether the neutral component typically appears with positive or negative information.
- The permanent shock is defined as
  
  \[ \nu_t = x_t + z_t, \]

  where \( z_t \sim N(0, \sigma_z^2) \).
Memory will be modeled as a stochastic map that transforms true history into perceived history.

Let history $h_t$ be a vector that includes $y_k$ and $e_k$ for $k < t$.

Memory maps $h_t$ into a random variable $h_t^R$.

- Past values of income $y_t$ are hard information and will be recalled perfectly.
- Events characterize soft information and are more prone to be forgotten.
Let recalled history be $h^R_t = (e^R_1, e^R_2, ..., e^R_{t-1}, y_1, y_2, ..., y_{t-1})$.

Notice that $e_t$ is transformed into a random variable $e^R_t$ whose value is governed by

$$e^R_t = \begin{cases} 
  e_k & \text{with probability } r_{kt}, \\
  (0, 0) & \text{with probability } 1 - r_{kt}.
\end{cases}$$

The probability that event $e_k$ is recalled at time $t$ is denoted by $r_{kt}$ where these probabilities are applied independently across events.

When an event is forgotten, it is exactly as if no event occurred that period.
To specify $r_{kt}$, one needs to turn to the scientific evidence drawn from research by biologists and psychologists.

Rehearsal states that recalling a memory increases future recall probabilities: repetition strengthens memories.

Associativeness states that events more similar to current events are easier to recall: hearing a friend talk about his vacation will invoke memories of one’s own vacations, as events serve as cues that help find lost memories.

Both rehearsal and associativeness have a strong experimental basis as well as intuitive appeal.
Three parameters are employed to formalize these ideas: the baseline recall probability $m$, $\rho$ which quantifies rehearsal, $\chi$ which quantifies associativeness.

All are between zero and one and $m + \rho + \chi < 1$.

Let $R_{kt}$ denote the random variable which equals one if event $k$ is recalled at time $t$, with $R_{(t-1)t} = 1$ and $r_{(t-1)t} = 1$.

With this notation, we can write

$$r_{kt} = m + \rho R_{k(t-1)} + \chi a_{kt}.$$
The second terms indicates that an event recalled in the last period gets a boost of $\rho$.

The third term captures associativeness where $a_{kt}$ measures the similarity of event $e_k$ to $e_t$.

The events $e_k$ and $e_t$ are two points on a plane.

Letting $c$ be a closeness function (an inverse distance function), similarity is defined as

$$a_{kt} = \frac{1}{2}[c(x_t - x_k) + c(n_t - n_k)] = \frac{1}{2}[e^{-(x_t-x_k)^2} + e^{-(n_t-n_k)^2}].$$

where $a_{kt} = 0$ if either $e_k$ or $e_t$ is a nonevent.

Define $f_{kt} = 1 - r_{kt}$, $F_{kt} = 1 - R_{kt}$, and $f = 1 - m - \rho$. 
In this setup, the dynamics of recall is given by

$$E(f_{kt} \mid e_k) = (f - \chi E(a_{kt} \mid e_k)) \frac{1 - \rho^{t-k}}{1 - \rho}.$$ 

Recall probabilities decay exponentially over time: more distant memories have a higher chance of being forgotten.

Also, $E(a_{kt} \mid e_k)$ increases memorability.

- $E(a_{kt} \mid e_k)$ is defined as vividness which measures how strongly associativeness affects a memory.
- $E(x_k a_{kt} \mid e_t)$ is defined as evocativeness of event $e_t$ which captures the average information of the associated events.
Basic results: perfect memory forecasts

- The stochastic process generates a signal extraction problem: the individual must separate out the permanent shock to $y_t$ from the transitory ones.
- The posterior at time $t$ will be distributed normally with mean $\hat{y}_t$ and variance $\hat{\sigma}_t^2$.
- In steady state, these beliefs will equal

\[
\hat{y}(h_t, e_t) = x_t + \sum_{k=1}^{t-1}[\lambda^{t-k}x_k + (1 - \lambda^{t-k})\Delta y_k],
\]

\[
\hat{\sigma}_t^2(h_t, e_t) = \sigma_*^2 := \frac{1}{2}(\sigma_\nu^2 + \sqrt{\sigma_\nu^2(\sigma_\nu^2 + 4\sigma_\varepsilon^2)}),
\]

where $\Delta y_k = y_k - y_{k-1}$ is the change in income and $\lambda = \sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_*^2)$ is the associated long-run error-to-truth ratio.
Basic results: perfect memory forecasts

- $x_k$ influences forecasts one-for-one: its impact is the sum of two terms, a direct effect $\lambda^{t-k} x_k$ and an indirect effect from $\Delta y_t = x_k + z_k + \varepsilon_k - \varepsilon_{k-1}$.
- $\Delta y_k$ enters with weight $1 - \lambda^{t-k} < 1$.
- $y_k$ influences forecasts at $\lambda^{t-k-1}(1 - \lambda) < 1$ because it enter in $\Delta y_k$ and $\Delta y_{k+1}$.
- $n_t$ has zero impact as expected: neutral components convey no information.
Basic results: limited memory expectations

- It is assumed that the forgetful individual applies the forecasting rule to the recalled history: that is, she takes the recalled history as the true history.

- Let $\hat{y}_t^R(h_t^R, e_t)$ denote the mean and $\hat{\sigma}^2_t^R(h_t^R, e_t)$ denote the variance of a (nave) forgetful posteriors.

- The assumption of naive forecasts means that $\hat{y}_t^R(h_t^R, e_t) = \hat{y}_t(h_t^R, e_t)$ and $\hat{\sigma}^2_t^R(h_t^R, e_t) = \hat{\sigma}^2_t(h_t^R, e_t)$.

- Forgetful forecasts look just like perfect recall forecasts except that forgotten events ($R_{kt} = 0$) are excluded.
Basic result: limited memory forecasts

Let $err_t = y_t - \hat{y}_t$ and $err_t^R = y_t - \hat{y}_t^R$ be the forecast errors.

Also, define $err_t^m = \hat{y}_t - \hat{y}_t^R$ as the memory error.

Note that $err_t^R = err_t + err_t^m$.

Proposition

The impact of event $e_t$ on time $t$ beliefs does not depend on its vividness but does depend on its evocativeness. On the other hand, its impact on time $t + j$ beliefs depends on both vividness and evocativeness.
Basic results: limited memory forecasts

▶ Vividness plays no role in how an event influences beliefs at the time it occurs.
▶ It only matters as time passes by increasing memorability.
▶ Evocativeness influences beliefs contemporaneously.
▶ An event with positive evocativeness disproportionately draws forth positive memories leading to a more positive forecast.
▶ Moreover, since these triggered memories persists by rehearsal, evocativeness also influences future beliefs.
Basic results: limited memory forecasts

Proposition

Let \( e_t = (0, n_t) \) be an uninformative event but with nonzero neutral component. This event influences beliefs if and only if \( \sigma_{xn} \neq 0 \). The sign of this influence equals \( \text{sign}(\sigma_{xn}n_t) \).

- Even though the individual disregards a signal with \( x_t = 0 \) as completely uninformative, her belief is still shaped by the memories this event triggers.
- A positive neutral cue \((n_t > 0)\) selectively evokes other positive neutral cue memories.
- If \( \sigma_{xn} > 0 \), these memories will on average have \( x_k > 0 \) and hence the event will selectively evoke positive information memories.
Overreaction and underreaction

Proposition

Forecast errors are negatively correlated with the information in the latest event:

\[ \text{cov}(y_t - \hat{y}_t^R, x_t) = \text{cov}(err_t^R, x_t) < 0. \]

The extent of this overreaction increases with \( \chi \) and \( \lambda \):

\[ \frac{\partial \text{cov}(err_t^R, x_t)}{\partial \chi} < 0, \quad \frac{\partial \text{cov}(err_t^R, x_t)}{\partial \lambda} < 0. \]

- Good information may lead to a rosy view of the past, which leads to forecasts that are too large.
- The effect of \( \lambda \) arises because it measures the importance of history and hence the importance of selective recall.
Overreaction and underreaction

Proposition

Let $T > t$. When events are very memorable ($f$ low, $\chi$ and $\rho$ large), then

$$\text{cov}(\text{err}_t^R, \text{err}_{t+1}^R) > 0.$$
Overreaction and underreaction

Proposition

Suppose that forgetting probabilities are small, so that $f$ is low and $\rho$ and $\chi$ are high. Then,

$$\text{cov}(\Delta \hat{y}^R_{t+1}, \Delta y_{t-1}) < 0.$$ 

When these probabilities are large, however,

$$\text{cov}(\Delta \hat{y}^R_{t+1}, \Delta y_{t-1}) > 0.$$
There are two effects that govern belief dynamics: forgetting and overreaction.

Forgetting induces a positive correlation (underreaction) between beliefs and lagged information; overreaction induces a negative correlation.

When events are memorable, the overreaction effect dominates.
Conclusion

- There are several questions left open by this model.
- This paper focuses on the naive case, but the sophisticated case also seems to be of interest.
- Associativeness as formulated in this paper has a failing: while current events can trigger related memories, the memories that one recalls cannot themselves trigger other memories, an extension referred to as association chain.
- To sum up, the paper provides a model of memory limitations which builds on two basic facts from scientific research.