Endogenous Learning with Bounded Memory

Yilmaz Kocer

USC

November 19, 2012

Introduction

- Learning through choices, i.e. by trial and error:

Introduction

- Learning through choices, i.e. by trial and error:
  - Consumer search or job search models with learning, Rothschild (1974), Bergemann & Valimaki (1996), Adam (2001)

Introduction

- Learning through choices, i.e. by trial and error:
  - Consumer search or job search models with learning, Rothschild (1974), Bergemann & Valimaki (1996), Adam (2001)
  - A monopolist learning about the unknown demand, Easley & Kiefer (1988)

Introduction

- Learning through choices, i.e. by trial and error:
  - Consumer search or job search models with learning, Rothschild (1974), Bergemann & Valimaki (1996), Adam (2001)
  - A monopolist learning about the unknown demand, Easley & Kiefer (1988)
  - A firm learning about different technologies, investing in alternative unknown projects, Roberts & Weitzman (1981), Cowan (1991), Jovanovic & Nyarko (1996)

Introduction

- Learning through choices, i.e. by trial and error:
  - Consumer search or job search models with learning, Rothschild (1974), Bergemann & Valimaki (1996), Adam (2001)
  - A monopolist learning about the unknown demand, Easley & Kiefer (1988)
  - A firm learning about different technologies, investing in alternative unknown projects, Roberts & Weitzman (1981), Cowan (1991), Jovanovic & Nyarko (1996)
  - Social learning and strategic experimentation, Bolton & Harris (1999), Keller, Rady & Cripps (2005)...etc.

Introduction

- Learning through choices, i.e. by trial and error:
  - Consumer search or job search models with learning, Rothschild (1974), Bergemann & Valimaki (1996), Adam (2001)
  - A monopolist learning about the unknown demand, Easley & Kiefer (1988)
  - A firm learning about different technologies, investing in alternative unknown projects, Roberts & Weitzman (1981), Cowan (1991), Jovanovic & Nyarko (1996)
  - Social learning and strategic experimentation, Bolton & Harris (1999), Keller, Rady & Cripps (2005)...etc.
- Most models of learning assume that **the agent can keep track of all past payoffs and actions**, and act accordingly.

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**
- **What should you remember/learn from past experiences? How should you act upon imperfect memories of the past?**

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**
- **What should you remember/learn from past experiences? How should you act upon imperfect memories of the past?**
- **Would memory limitations bias choices in a particular way?**

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**
- **What should you remember/learn from past experiences? How should you act upon imperfect memories of the past?**
- **Would memory limitations bias choices in a particular way?**
- Can bounded memory account for biases observed in the experimental literature ?.. Meyer & Shi (1995), Anderson (2001), Gans et.al (2007)

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**
- **What should you remember/learn from past experiences? How should you act upon imperfect memories of the past?**
- **Would memory limitations bias choices in a particular way?**
- Can bounded memory account for biases observed in the experimental literature ?.. Meyer & Shi (1995), Anderson (2001), Gans et.al (2007)
    - **under-experimentation..**the agent does not experiment as much as the received theory suggests.

Motivation and Question

- **How does a decision maker learn about uncertain alternatives when he has limited memory?**
- **What should you remember/learn from past experiences? How should you act upon imperfect memories of the past?**
- **Would memory limitations bias choices in a particular way?**
- Can bounded memory account for biases observed in the experimental literature ?.. Meyer & Shi (1995), Anderson (2001), Gans et.al (2007)
  - **under-experimentation..**the agent does not experiment as much as the received theory suggests.
  - **occasional experimentation** the agent never ceases experimentation completely, even when it is optimal to choose the safe option.

Decision Making under Memory Constraints

- Imperfect recall and games with absentmindedness, Piccione & Rubinstein (1997)

Decision Making under Memory Constraints

- Imperfect recall and games with absentmindedness, Piccione & Rubinstein (1997)
- Memory imperfections, Mullainathan (2002)

Decision Making under Memory Constraints

- Imperfect recall and games with absentmindedness, Piccione & Rubinstein (1997)
- Memory imperfections, Mullainathan (2002)
- Bounded recall, Lehrer & Solan (2009)

Decision Making under Memory Constraints

- Imperfect recall and games with absentmindedness, Piccione & Rubinstein (1997)
- Memory imperfections, Mullainathan (2002)
- Bounded recall, Lehrer & Solan (2009)
- Bounded memory as a "finite automaton": as a neutral/ flexible storage model of memory, it **does not dictate the agent what to remember and what information to discard**: Cover & Hellman (1970,1976), Abreu & Rubinstein (1988), Kalai & Solan (2003), Wilson (2004), Borgers & Morales (2004), Salant (2007), Monte (2008)..

Bounded Memory as Finite Number of Memory States

- **A helpful analogy**: using a ring on your finger as a way of remembering things.

Bounded Memory as Finite Number of Memory States

- **A helpful analogy**: using a ring on your finger as a way of remembering things.
- Each finger corresponds to a "memory state".

Bounded Memory as Finite Number of Memory States

- **A helpful analogy**: using a ring on your finger as a way of remembering things.
- Each finger corresponds to a "memory state".
- A strategy describes for each memory state:
  1) What to do (what action to take, or what guess to make...)
  2) How to update the memory state (move the ring to a different finger) in response to incoming information.

Bounded Memory as Finite Number of Memory States

- **A helpful analogy**: using a ring on your finger as a way of remembering things.
- Each finger corresponds to a "memory state".
- A strategy describes for each memory state:
  1) What to do (what action to take, or what guess to make...)
  2) How to update the memory state (move the ring to a different finger) in response to incoming information.
- It imposes no constraints on how an agent might use his memory: there is just a capacity constraint, and the agent is assumed to make the best use of it.

Bounded Memory as Finite Number of Memory States

- **A helpful analogy**: using a ring on your finger as a way of remembering things.
- Each finger corresponds to a "memory state".
- A strategy describes for each memory state:
  1) What to do (what action to take, or what guess to make...)
  2) How to update the memory state (move the ring to a different finger) in response to incoming information.
- It imposes no constraints on how an agent might use his memory: there is just a capacity constraint, and the agent is assumed to make the best use of it.
- The histories are lumped into N categories where N is the number of memory states (fingers). It measures how "finely" you can remember history.

Wilson (2004), Hypothesis Testing

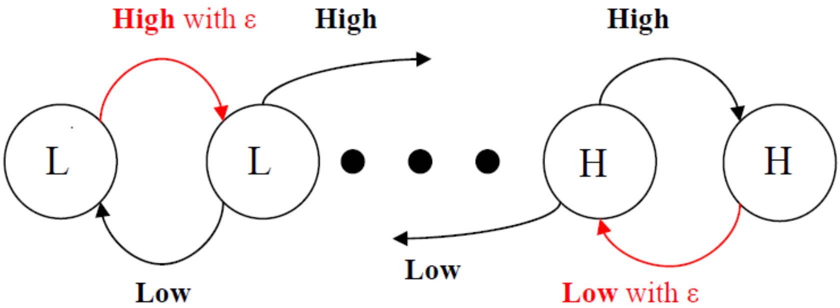
- $s = H$ or L, $\pi_H = 1/2$. The agent receives a signal *High*, *Med*1, *Med*2 or *Low* each period and guesses the state of the world.

Wilson (2004), Hypothesis Testing

- $s = H$ or L, $\pi_H = 1/2$. The agent receives a signal *High*, *Med*1, *Med*2 or *Low* each period and guesses the state of the world.
- He tries to maximize the discounted sum of correct guesses.

Wilson (2004), Hypothesis Testing

- $s = H$ or L, $\pi_H = 1/2$. The agent receives a signal *High*, *Med*1, *Med*2 or *Low* each period and guesses the state of the world.
- He tries to maximize the discounted sum of correct guesses.
- He has a finite number of memory states, and chooses 1) what to guess at each memory state and 2) how to update the memory state after receiving a signal. What is the optimal memory rule?



Wilson (2004), Hypothesis Testing

- Wilson (2004) finds that for sufficiently high discount factors,

Wilson (2004), Hypothesis Testing

- Wilson (2004) finds that for sufficiently high discount factors,
- The agent optimally moves *only after* the most informative signals and only to an adjacent memory state,

Wilson (2004), Hypothesis Testing

- Wilson (2004) finds that for sufficiently high discount factors,
- The agent optimally moves *only after* the most informative signals and only to an adjacent memory state,
- *ignores information* at extreme memory states 1 and N with probability close to 1.

Wilson (2004), Hypothesis Testing

- Wilson (2004) finds that for sufficiently high discount factors,
- The agent optimally moves *only after* the most informative signals and only to an adjacent memory state,
- *ignores information* at extreme memory states 1 and N with probability close to 1.
- The implied behavior is consistent with many biases in the experimental literature: under/overconfidence, confirmatory bias, belief polarization, first impressions matter.

Wilson (2004), Hypothesis Testing

- Wilson (2004) finds that for sufficiently high discount factors,
- The agent optimally moves *only after* the most informative signals and only to an adjacent memory state,
- *ignores information* at extreme memory states 1 and N with probability close to 1.
- The implied behavior is consistent with many biases in the experimental literature: under/overconfidence, confirmatory bias, belief polarization, first impressions matter.
- In this model information is free and one dimensional.

Endogenous Learning with Bounded Memory

- My goal is to develop a model of learning with bounded memory where learning is by trial and error.

Endogenous Learning with Bounded Memory

- My goal is to develop a model of learning with bounded memory where learning is by trial and error.
- Attitudes towards experimentation: Exploitation / Exploration trade-off:

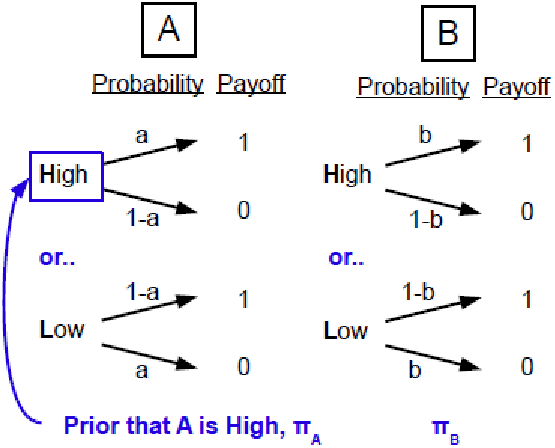
Endogenous Learning with Bounded Memory

- My goal is to develop a model of learning with bounded memory where learning is by trial and error.
- Attitudes towards experimentation: Exploitation / Exploration trade-off:
  - **short run (myopic) motive** to choose the better alternative with the current information.

- My goal is to develop a model of learning with bounded memory where learning is by trial and error.
- Attitudes towards experimentation: Exploitation / Exploration trade-off:
  - **short run (myopic) motive** to choose the better alternative with the current information.
  - **long run (learning) motive** to choose the "less explored" alternative for better future decisions.

Endogenous Learning with Bounded Memory

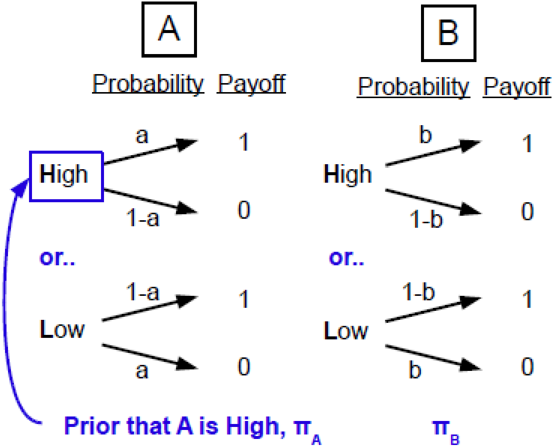
- My goal is to develop a model of learning with bounded memory where learning is by trial and error.
- Attitudes towards experimentation: Exploitation / Exploration trade-off:
  - **short run (myopic) motive** to choose the better alternative with the current information.
  - **long run (learning) motive** to choose the "less explored" alternative for better future decisions.
- How do you organize your memory when you're learning about two things? (multidimensional uncertainty)



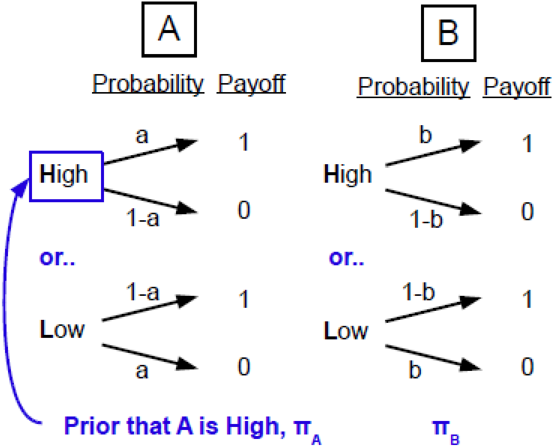
- Types are independently chosen at the start, fixed throughout.

The Model - Problem

Standard Two Armed Bandit Problem



- Types are independently chosen at the start, fixed throughout.
- Agent does not know the types and has a prior for each.



- Types are independently chosen at the start, fixed throughout.
- Agent does not know the types and has a prior for each.
- He chooses either A or B each period to maximize the discounted sum of payoffs.

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,

The Model - Bounded Memory Strategy

Finite Automaton

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,
2. **Action Rule:** at each memory state, which alternative to choose,

The Model - Bounded Memory Strategy
Finite Automaton

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,
2. **Action Rule:** at each memory state, which alternative to choose,
3. **Transition Rule:** at each memory state, depending on the alternative chosen and the payoff received, how to update the memory state, i.e. which memory state to go to.

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,
2. **Action Rule:** at each memory state, which alternative to choose,
3. **Transition Rule:** at each memory state, depending on the alternative chosen and the payoff received, how to update the memory state, i.e. which memory state to go to.

- The agent chooses a strategy at the start in order to maximize **expected β-discounted sum of payoffs**.

The Model - Bounded Memory Strategy

Finite Automaton

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,
2. **Action Rule:** at each memory state, which alternative to choose,
3. **Transition Rule:** at each memory state, depending on the alternative chosen and the payoff received, how to update the memory state, i.e. which memory state to go to.

- The agent chooses a strategy at the start in order to maximize **expected β-discounted sum of payoffs**.
- I assume that the agent can **randomize** in each component of the strategy.

The Model - Bounded Memory Strategy

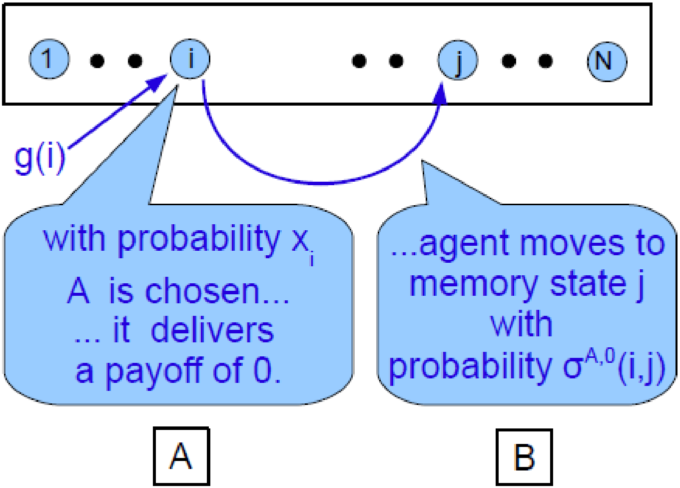
Finite Automaton

A bounded memory strategy for the agent consists of:

1. **Initial memory state:** which memory state to start at,
2. **Action Rule:** at each memory state, which alternative to choose,
3. **Transition Rule:** at each memory state, depending on the alternative chosen and the payoff received, how to update the memory state, i.e. which memory state to go to.

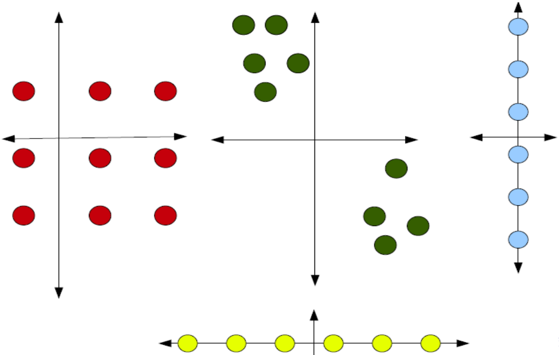
- The agent chooses a strategy at the start in order to maximize **expected β-discounted sum of payoffs**.
- I assume that the agent can **randomize** in each component of the strategy.
- **What is the optimal bounded memory strategy,** for a sufficiently high discount factor β**?**

The Model - Bounded Memory Strategy



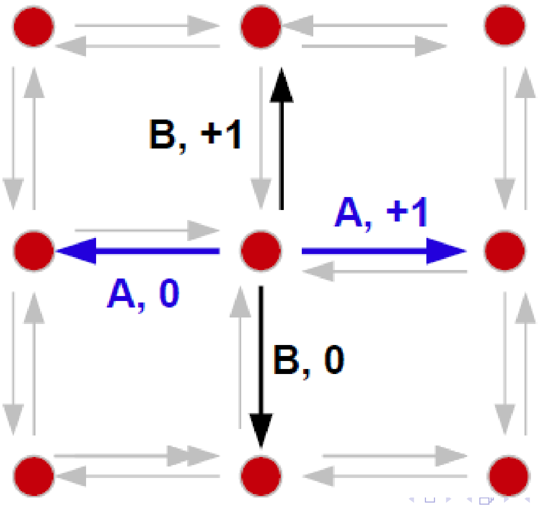
Beliefs at Memory States - Examples

- **Given a strategy (g, x, σ), one can compute the agent's beliefs about the alternatives' types at each memory state using Bayes' rule.**
- How would the **optimal** beliefs look? The axes are beliefs on A's type and beliefs on B's type, and each point corresponds to the induced belief at a memory state:



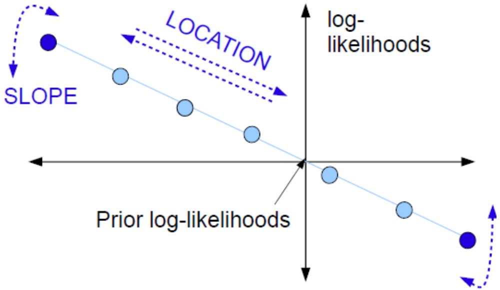
Beliefs at Memory States - Examples

For example, this strategy keeps a rough summary of each alternative's type, independently:



Optimal Limit Beliefs

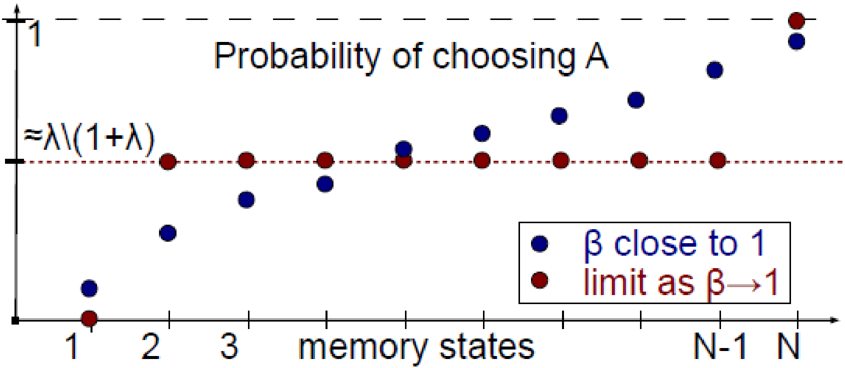
As $\beta \rightarrow 1$, **optimal** (conditional) limit beliefs are **evenly and linearly spaced on the log-likelihood** space with a **negative** slope :



- The optimal strategy fixes the **slope** and the **location** of this line segment (it carries the point of prior log-likelihoods)
- **The optimal solution to a seemingly complex problem boils down to the choice of two parameters !! (**λ for the slope, q for the location of the belief line**)**

Optimal Strategy for Sufficiently High Discount Factors

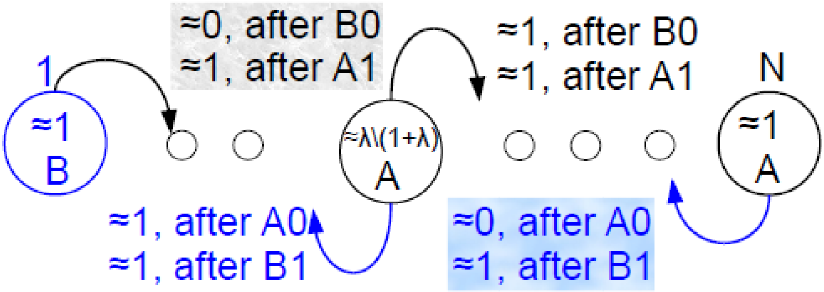
Action Rule



Optimal Strategy for Sufficiently High Discount Factors

Transition (Updating) Rule

- q measures the relative exit probabilities out of extreme memory states, 1 and N.
- λ measures the relative likelihood that transitions are due to A observations rather than B observations.



Optimal Strategy for Sufficiently High Discount Factors

1. Choose (almost always) B at memory state 1, and A at memory state N.

Optimal Strategy for Sufficiently High Discount Factors

1. Choose (almost always) B at memory state 1, and A at memory state N.
2. Make sure that you are (almost always) at the extreme memory states 1 or N. (q determines the **location** of the belief line).

Optimal Strategy for Sufficiently High Discount Factors

1. Choose (almost always) B at memory state 1, and A at memory state N.
2. Make sure that you are (almost always) at the extreme memory states 1 or N. (q determines the **location** of the belief line).
3. Move to the next memory state on the right **either** after a payoff of 1 from A **or** after a payoff of 0 from B, and to the left after opposing payoffs, with the following restriction:

Optimal Strategy for Sufficiently High Discount Factors

1. Choose (almost always) B at memory state 1, and A at memory state N.
2. Make sure that you are (almost always) at the extreme memory states 1 or N. (q determines the **location** of the belief line).
3. Move to the next memory state on the right **either** after a payoff of 1 from A **or** after a payoff of 0 from B, and to the left after opposing payoffs, with the following restriction:
4. λ is the likelihood that an adjacent transition is due to an observation on A rather than B. This likelihood determines the **slope** of the belief line.

Results

- **No exploitation/exploration trade-off** asymptotically as $\beta \to 1$; the extra consideration of using what is learned does not constrain the amount of learning.

Results

- **No exploitation/exploration trade-off** asymptotically as $\beta \to 1$; the extra consideration of using what is learned does not constrain the amount of learning.
- It is optimal to **store joint information** about the uncertainties; optimally a rough summary of the **relative ranking of alternatives is encoded**, i.e. the beliefs about A and B are **negatively correlated, even though alternatives' types are independent**.

Results

- **No exploitation/exploration trade-off** asymptotically as $\beta \longrightarrow 1$; the extra consideration of using what is learned does not constrain the amount of learning.
- It is optimal to **store joint information** about the uncertainties; optimally a rough summary of the **relative ranking of alternatives is encoded**, i.e. the beliefs about A and B are **negatively correlated, even though alternatives' types are independent**.
- Biases in updating beliefs in the short run: After receiving good news about A (payoff of 1), the agent moves to a memory state with worse opinions on B, **even though no information on B is received** !!

Results

- The behavior will induce long sequences of choices of the currently favored alternative, interrupted randomly by short experimentation runs through intermediate memory states.

Results

- The behavior will induce long sequences of choices of the currently favored alternative, interrupted randomly by short experimentation runs through intermediate memory states.
- Optimally, an alternative is **probabilistically reinforced** after a good payoff from it, and inhibited after a bad payoff: a **stochastic reinforcement learning type strategy** is optimal under neutral memory constraints.

Results

- The behavior will induce long sequences of choices of the currently favored alternative, interrupted randomly by short experimentation runs through intermediate memory states.
- Optimally, an alternative is **probabilistically reinforced** after a good payoff from it, and inhibited after a bad payoff: a **stochastic reinforcement learning type strategy** is optimal under neutral memory constraints.
- **Specialization in learning:** If A is **substantially more informative** than B, all learning will be done on A and alternative B is treated as if it is safe (known).

Results

- The behavior will induce long sequences of choices of the currently favored alternative, interrupted randomly by short experimentation runs through intermediate memory states.
- Optimally, an alternative is **probabilistically reinforced** after a good payoff from it, and inhibited after a bad payoff: a **stochastic reinforcement learning type strategy** is optimal under neutral memory constraints.
- **Specialization in learning:** If A is **substantially more informative** than B, all learning will be done on A and alternative B is treated as if it is safe (known).
- Generically, the optimal bounded memory strategy generates under-experimentation and occasional experimentation.

Results - Biases in Learning Experiments

Under-experimentation and occasional experimentation

- Suppose one of the alternatives' type is known, i.e. a one-armed bandit problem.

Results - Biases in Learning Experiments

Under-experimentation and occasional experimentation

- Suppose one of the alternatives' type is known, i.e. a one-armed bandit problem.
- Meyer & Shi (1995), Anderson (2001), Steyvers & Lee (2008), Gans, Knox & Croson (2007) conduct this experiment and observe two departures from the full memory optimal:

Results - Biases in Learning Experiments

Under-experimentation and occasional experimentation

- Suppose one of the alternatives' type is known, i.e. a one-armed bandit problem.
- Meyer & Shi (1995), Anderson (2001), Steyvers & Lee (2008), Gans, Knox & Croson (2007) conduct this experiment and observe two departures from the full memory optimal:

1. Agents give up experimentation with the uncertain alternative rather early relative to the full memory optimal (**under-experimentation**).

Results - Biases in Learning Experiments

Under-experimentation and occasional experimentation

- Suppose one of the alternatives' type is known, i.e. a one-armed bandit problem.
- Meyer & Shi (1995), Anderson (2001), Steyvers & Lee (2008), Gans, Knox & Croson (2007) conduct this experiment and observe two departures from the full memory optimal:

1. Agents give up experimentation with the uncertain alternative rather early relative to the full memory optimal (**under-experimentation**).
2. Agents choose the uncertain alternative occasionally, even after choosing the known alternative for many periods (**occasional experimentation**).

Results - Biases in Learning Experiments

Under-experimentation and occasional experimentation

Both biases are consistent with optimal behavior under memory constraints for sufficiently patient agents:

- (**Under-experimentation**) Because he does not have the ability to store future information efficiently, the bounded memory agent does not appreciate the option value of information as much as the full memory agent. As a result, **he optimally explores less** with unfamiliar options.

Both biases are consistent with optimal behavior under memory constraints for sufficiently patient agents:

- (**Under-experimentation**) Because he does not have the ability to store future information efficiently, the bounded memory agent does not appreciate the option value of information as much as the full memory agent. As a result, **he optimally explores less** with unfamiliar options.
- (**Occasional experimentation**) The bounded memory agent will **never give up experimentation completely;** he experiments with the unknown alternative with positive probability after any history. Even at a memory state with the most pessimistic beliefs, it pays to experiment with a small probability.

Results - Biases in Learning Experiments

Under-experimentation and Occasional Experimentation