# On Second Thoughts, Selective Memory, and Resulting Behavioral Biases


Philippe Jehiel[1] Jakub Steiner[2]

[1]PSE and UCL

[2]Cerge-ei and Edinburgh U


Zurich
January 2018

# Two goals

## robustness

- rich literature on inattention
- strong assumptions on information frictions
- we: a detail-free prediction

## hesitation

- focuss on decision process
- literature: information aggregation
- we: aggregation friction

**“inattention” models**:

$$\max_{\text{cognition process} \in C} \text{expected payoff}$$

**endogenous information**:
$\Rightarrow$ strong link between preferences and stochastic choice

**a complication**:
economists don’t observe $C$

**our contribution**:
information-processing model robust to details of $C$

**information processing models**:

- agent observes a sequence of signals
- perfect aggregation of information

**speed accuracy trade-off**:
$\Rightarrow$ joint predictions about decision time and choice

**a complication**:
speed-accuracy complementarity

**our contribution**:
information-aggregation friction

# Overview

reduced-form model of targeted search of information

procedural model of targeted search of information

behavioral applications

sophisticated agents

# Overview

reduced-form model of targeted search of information

procedural model of hesitation

behavioral applications

sophisticated agents

# Literature

**error-management** (interdisciplinary): Johnson et al. (2013)

**rational inattention**: Sims (2003), Matejka & McKay (2015)
entropy of beliefs cannot be reduced by more than $\kappa$

- identification of preferences from choice data
- source of identification: the assumed constraint

**microfoundations of cognitive constraints**:

- based on the sequential-sampling model of Ratcliff (1978): Hebert & Woodford (2016), Morris & Strack (2017)
- finite automata:
  Compte & Postlewaite (2012), Wilson (2014)

**identification of cognitive constraints**:
Caplin & Dean (2015), Oliveira, Denti, Mihm, & Ozbek (2017)

# Outline

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

**primitive signal distribution**: $p(x \mid \theta)$,

$x \in \{0, 1\}$, $p(\theta \mid \theta) > 1/2$, fix strategy $a(x) = x$

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

**primitive signal distribution**: $p(x \mid \theta)$,

$x \in \{0, 1\}$, $p(\theta \mid \theta) > 1/2$, fix strategy $a(x) = x$

**targeted information search**: the agent chooses search intensities $\beta_x > 0$ for each $x \in \{0, 1\}$

effective signal distribution: $r(x \mid \theta; \beta) = \frac{\beta_x p(x|\theta)}{\sum_{x'} \beta_{x'} p(x'|\theta)}$

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

**primitive signal distribution**: $p(x \mid \theta)$,

$x \in \{0, 1\}$, $p(\theta \mid \theta) > 1/2$, fix strategy $a(x) = x$

**targeted information search**: the agent chooses search intensities $\beta_x > 0$ for each $x \in \{0, 1\}$

effective signal distribution: $r(x \mid \theta; \beta) = \frac{\beta_x p(x|\theta)}{\sum_{x'} \beta_{x'} p(x'|\theta)}$

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

**primitive signal distribution**: $p(x \mid \theta)$,

$x \in \{0, 1\}$, $p(\theta \mid \theta) > 1/2$, fix strategy $a(x) = x$

**targeted information search**: the agent chooses search intensities $\beta_x > 0$ for each $x \in \{0, 1\}$

effective choice rule: $r(a \mid \theta; \beta) = \frac{\beta_a p(a|\theta)}{\sum_{a'} \beta_{a'} p(a'|\theta)}$

# Binary example

**task**: $A = \Theta = \{0, 1\}$, prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$, and $0$ otherwise

**primitive signal distribution**: $p(x \mid \theta)$,

$x \in \{0, 1\}$, $p(\theta \mid \theta) > 1/2$, fix strategy $a(x) = x$

**targeted information search**: the agent chooses search intensities $\beta_x > 0$ for each $x \in \{0, 1\}$

effective choice rule: $r(a \mid \theta; \beta) = \frac{\beta_a p(a|\theta)}{\sum_{a'} \beta_{a'} p(a'|\theta)}$

**the problem**:

$$\max_\beta \sum_{\theta, a} \pi_\theta u(a, \theta) r(a \mid \theta; \beta)$$

# Optimality condition

**notation**:

$r^*(a \mid \theta)$—choice rule induced by optimal $\beta^*$

$a_1$, $a_2$—two conditionally independent draws from $r^*(a \mid \theta)$:
joint distribution of $\theta$, $a_1$, and $a_2$ is

$$\alpha(\theta, a_1, a_2) = \pi_\theta r^*(a_1 \mid \theta) r^*(a_2 \mid \theta)$$

# Optimality condition

**notation**:

$r^*(a \mid \theta)$—choice rule induced by optimal $\beta^*$

$a_1$, $a_2$—two conditionally independent draws from $r^*(a \mid \theta)$:
joint distribution of $\theta$, $a_1$, and $a_2$ is

$$\alpha(\theta, a_1, a_2) = \pi_\theta r^*(a_1 \mid \theta) r^*(a_2 \mid \theta)$$

## Proposition

*Optimal choice rule $r^*(a \mid \theta)$ satisfies*

$$\mathrm{E}_\alpha[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}_\alpha[u(a_2, \theta) \mid a_1 = a],$$

*for every $a$ chosen with a positive probability.*

# Optimality condition

**notation**:

$r^*(a \mid \theta)$—choice rule induced by optimal $\beta^*$

$a_1$, $a_2$—two conditionally independent draws from $r^*(a \mid \theta)$:
joint distribution of $\theta$, $a_1$, and $a_2$ is

$$\alpha(\theta, a_1, a_2) = \pi_\theta r^*(a_1 \mid \theta) r^*(a_2 \mid \theta)$$

### Proposition

*The agent at the end of the constrained-optimal process is indifferent between terminating and repeating it.*

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

set $\beta = (1, 1)$, so that $r = p$

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

set $\beta = (1, 1)$, so that $r = p$

$\mathrm{E}[u(a_1, \theta) \mid a_1 = 0] < \mathrm{E}[u(a_2, \theta) \mid a_1 = 0]$, since:

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

set $\beta = (1, 1)$, so that $r = p$

$\mathrm{E}[u(a_1, \theta) \mid a_1 = 0] < \mathrm{E}[u(a_2, \theta) \mid a_1 = 0]$, since:

- if $a_2 = a_1$, then second thought is inconsequential,

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

set $\beta = (1, 1)$, so that $r = p$

$\mathrm{E}[u(a_1, \theta) \mid a_1 = 0] < \mathrm{E}[u(a_2, \theta) \mid a_1 = 0]$, since:

- if $a_2 = a_1$, then second thought is inconsequential,
- if $1 = a_2 \neq a_1 = 0$, then $\theta = 1$ is more likely

# Discussion

the result follows from the marginal argument

**is "second thought" beneficial?** example:

- uniform reward and prior,
- $.9 = \Pr_p(\theta = 1 \mid x = 1) > \Pr_p(\theta = 0 \mid x = 0) = .8$

set $\beta = (1, 1)$, so that $r = p$

$\mathrm{E}[u(a_1, \theta) \mid a_1 = 0] < \mathrm{E}[u(a_2, \theta) \mid a_1 = 0]$, since:

- if $a_2 = a_1$, then second thought is inconsequential,
- if $1 = a_2 \neq a_1 = 0$, then $\theta = 1$ is more likely

$\Rightarrow \beta_0 \searrow$ is beneficial

# Second-thought-free rules

## Definition

A choice rule $r(a \mid \theta)$ is second-thought-free if for each action $a$ chosen with positive probability:

$$\mathrm{E}[u(a_1, \theta) \mid a_1 = a] \geq \mathrm{E}[u(a_2, \theta) \mid a_1 = a].$$

# Second-thought-free rules

### Lemma

*If a choice rule $r(a \mid \theta)$ is second-thought-free, then for each action $a$ chosen with positive probability:*

$$\mathrm{E}[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}[u(a_2, \theta) \mid a_1 = a].$$

# Second-thought-free rules

## Lemma

*If a choice rule $r(a \mid \theta)$ is second-thought-free, then for each action $a$ chosen with positive probability:*

$$\mathrm{E}[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}[u(a_2, \theta) \mid a_1 = a].$$

**proof**: Suppose not:

$$\mathrm{E}\left[u(a_1, \theta) \mid a_1\right] \geq \mathrm{E}\left[u(a_2, \theta) \mid a_1\right]$$

# Second-thought-free rules

## Lemma

*If a choice rule $r(a \mid \theta)$ is second-thought-free, then for each action $a$ chosen with positive probability:*

$$\mathrm{E}[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}[u(a_2, \theta) \mid a_1 = a].$$

**proof**: Suppose not:

$$\mathrm{E}\left[\mathrm{E}\left[u(a_1, \theta) \mid a_1\right]\right] > \mathrm{E}\left[\mathrm{E}\left[u(a_2, \theta) \mid a_1\right]\right]$$

# Second-thought-free rules

## Lemma

*If a choice rule $r(a \mid \theta)$ is second-thought-free, then for each action $a$ chosen with positive probability:*

$$\mathrm{E}[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}[u(a_2, \theta) \mid a_1 = a].$$

**proof**: Suppose not:

$$\mathrm{E}\left[u(a_1, \theta)\right] > \mathrm{E}\left[u(a_2, \theta)\right]$$

# Solution

## reformulation

### Proposition

*Optimal choice rule is second-thought-free,*
*regardless of the primitive distributions $p(x \mid \theta)$.*

primitive information structure $p$ unobservable to outsiders
$\Rightarrow C = \{r(a \mid \theta; \beta, p) : \beta \in B\}$ is unobservable

yet, the outsider can identify preferences:

e.g. suppose that an outsider observes

- uniform prior, and
- equal shares of both error types, $r(1|0) = r(0|1)$

$\Rightarrow$ losses from both types of error are equal

# Outline

**task**: $a \in A$, $\theta \in \Theta$, $|A|, |\Theta| < \infty$, $u(a, \theta)$, prior $\pi_\theta$

**primitive decision processes**: $p(a \mid \theta)$
$a$—action recommendation

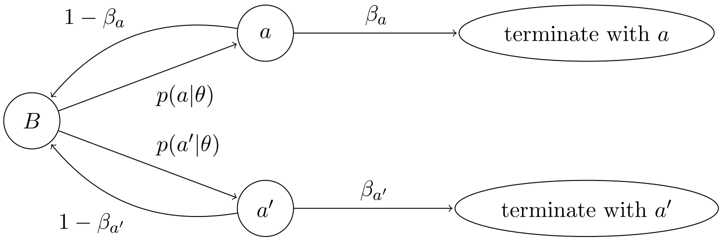
set of the feasible primitive processes $P$

**repetitions**:
termination strategy: $(\beta_a)_a$

- agent draws action recommendation $a$ from $p(\cdot \mid \theta)$
- terminates with probability $\beta_a$
- restarts with probability $1 - \beta_a$ and draws $a'$, ...

**effective choice rule**: $r(a \mid \theta; p, \beta)$—probability that the agent terminates with action $a$

**information-aggregation friction**: will be relaxed


$1-\beta_a$
$a$
$\beta_a$
terminate with $a$
$p(a|\theta)$
$B$
$p(a'|\theta)$
$a'$
$\beta_{a'}$
terminate with $a'$
$1-\beta_{a'}$

# Repeated-cognition problem

the agent chooses

- the primitive process $p$ and
- the termination strategy $\beta$

to maximize her expected payoff:

$$\max_{p \in P, \beta \in B} \sum_{\theta \in \Theta, a \in A} \pi_\theta u(a, \theta) r(a \mid \theta; \boldsymbol{p}, \beta)$$

# Outline

# Effective choice rule characterization

selective repetitions allow for targeted information search

## Lemma

*The agent terminates with action $a$ in state $\theta$ with probability*

$$r(a \mid \theta; \beta, p) = \frac{\beta_a p(a \mid \theta)}{\sum_{a' \in A} \beta_{a'} p(a' \mid \theta)}.$$

**proof**:

$$r(a \mid \theta) = \beta_a p(a \mid \theta) + \left( \sum_{a' \in A} p(a' \mid \theta) \left(1 - \beta_{a'}\right) \right) r(a \mid \theta),$$

and rearrange

# The main result

**Proposition**

*If a choice rule solves the repeated-cognition problem, then it is second-thought-free.*

regardless of the set of the feasible primitive processes $P$

**proof**: the first-order condition

# Imperfect recall

## our agent cannot condition on past recommendations

team-equilibrium interpretation (Piccione & Rubinstein 1997)

$a$-self:

- draws inferences about $\theta$ from:
  - $a$ drawn from $p(a \mid \theta)$
  - from the fact that previous selves have not terminated
- maximizes the “team payoff”
- is indifferent btwn terminating and passing to the next self

# Partial preference identification

an outsider observes:

- choice rule $r^*(a \mid \theta)$
- prior $\pi_\theta$

identifying conditions linear in $u$:

$$\mathrm{E}[u(a_1, \theta) \mid a_1] = \mathrm{E}[u(a_2, \theta) \mid a_1] \text{ for all observed actions}$$

$|A| - 1$ independent conditions

**binary setting**: full utility identification up to affine transformations

**otherwise**: partial identification

# Partial preference identification

an outsider observes:

- choice rule $r^*(a \mid \theta)$
- prior $\pi_\theta$

identifying conditions linear in $u$:

$$\sum_\theta \pi_\theta u(a_1, \theta) r(a_1 \mid \theta) = \sum_{\theta, a_2} \pi_\theta u(a_2, \theta) r(a_1 \mid \theta) r(a_2 \mid \theta)$$

$|A| - 1$ independent conditions

**binary setting**: full utility identification up to affine transformations

**otherwise**: partial identification

# Outline

# Binary setting

$A = \Theta = X = \{0, 1\}$, arbitrary prior $\pi_\theta$

$u(a, \theta) = u_\theta > 0$ if $a = \theta$ and $0$ otherwise

$\tilde{p}(x \mid \theta)$, increasing likelihood ratio: $\frac{\tilde{p}(1|0)}{\tilde{p}(0|0)} < \frac{\tilde{p}(1|1)}{\tilde{p}(0|1)}$

when $|X| > 2$, agent ignores all but the two most precise signals

**two summary parameters**:

- relative a priori attractiveness of action $1$: $R = \frac{\pi_1 u_1}{\pi_0 u_0}$
- distinguishability of the two states: $d = \frac{\tilde{p}(1|1)\tilde{p}(0|0)}{\tilde{p}(0|1)\tilde{p}(1|0)}$

# Primitive decision processes

four pure strategies $\sigma : X \longrightarrow A$

set $P$: four primitive decision processes $p(a \mid \theta; \sigma)$

**Lemma**

*There exists a solution in which $\sigma$ is the identity function.*

# Feasibility condition

## Lemma

*The effective choice rule $r(a \mid \theta; \beta)$ satisfies, for each positive $\beta$:*

$$\frac{r(0|0;\beta)r(1|1;\beta)}{r(0|1;\beta)r(1|0;\beta)} = d.$$

**proof**: $\frac{r(1|1)r(0|0)}{r(0|1)r(1|0)} = \frac{\frac{\beta_1 p(1|1)}{\sum_a \beta_a p(a|1)} \frac{\beta_0 p(0|0)}{\sum_a \beta_a p(a|0)}}{\frac{\beta_0 p(0|1)}{\sum_a \beta_a p(a|1)} \frac{\beta_1 p(1|0)}{\sum_a \beta_a p(a|0)}}$

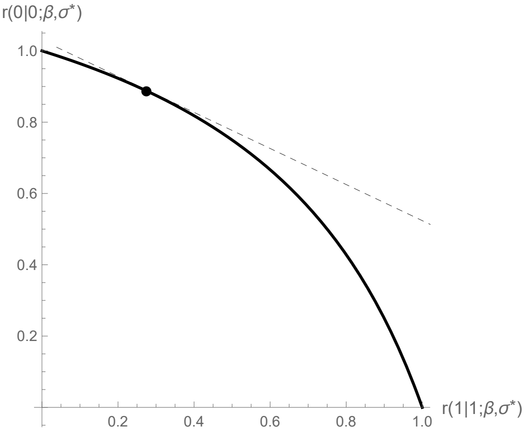
# Feasibility condition

## Lemma

*The effective choice rule $r(a \mid \theta; \beta)$ satisfies, for each positive $\beta$:*

$$\frac{r(0|0;\beta)r(1|1;\beta)}{r(0|1;\beta)r(1|0;\beta)} = d.$$

**proof**: $\frac{r(1|1)r(0|0)}{r(0|1)r(1|0)} = \frac{p(1|1)p(0|0)}{p(0|1)p(1|0)} = d$



r(0|0;β,σ*)
1.0
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1.0
r(1|1;β,σ*)

# Solution

feasibility $+$ second-thought-free $+$ normalizations $\Rightarrow$

## Proposition

1. *when* $R \geq d$*, the agent always chooses* $1$*,*
2. *when* $R \leq 1/d$*, the agent always chooses* $0$*,*
3. *when* $R \in (1/d, d)$*, then*

$$r^*(1 \mid 1) = \frac{dR - \sqrt{dR}}{(d-1)R}, r^*(0 \mid 0) = \frac{d - \sqrt{dR}}{d-1},$$

$$\frac{\beta_1^*}{\beta_0^*} = \frac{dR - \sqrt{dR}}{\sqrt{dR} - R} \frac{\tilde{p}(0 \mid 1)}{\tilde{p}(1 \mid 1)}.$$

# Confirmation bias

symmetric primitive information: $p(1 \mid 1) = p(0 \mid 0)$

symmetric incentives: $u_1 = u_0$

asymmetric prior: $R = \frac{\pi_1}{\pi_0} \in (1, d)$

## Corollary

*targeting of the a priori likely state:* $\beta_1^* / \beta_0^* > 1$

**intuition**:
set $\beta_1 = \beta_0$ so that $r = p$

consider the surprising recommendation. repetition leads to

- the same recommendation (inconsequential), or
- action switch (beneficial)

# Speed-accuracy complementarity

psychologists: delayed choices are less accurate

the setting from the previous slide

### Corollary

*$Pr_{\pi,r^*}(a=\theta|t)$ decreases with the response time $t$.*

**intuition**:
probability of the correct choice is lower in $\theta=0$

delay indicates that the agent has repeatedly received the surprising recommendation

delay indicates $\theta=0 \Rightarrow$ high error probability

# Overweighting of rare events

**state of the aviation**: uniform prior

- safe: flight accident probability $10^{-6}$
- dangerous: flight accident probability $10^{-5}$

**agent**:

- draws one flight from the realized flight distribution
- observes whether the flight was eventful
- announces safe/dangerous state
- receives payoff $1$ if correct

**optimal search intensities**:

- $\beta_{\text{accident}}/\beta_{\text{no accident}} \approx 316,000$
- probability of the correct guess $\approx 0.76$ in both states

# Outline

# Salience

Kahneman:

*"Our mind has a useful capability to focus on whatever is odd, different or unusual."*

**color-recognition task**: $\{azure, indigo, red\}$

agent is likely to hesitate after a blue impression since it is noisy

$\Rightarrow$ optimal bias in favor of the distinct red color

# Formalization

**task**: $\Theta = A$, $|\Theta| > 2$, $u(a, \theta) = 1_{a=\theta}$, uniform prior

**primitive physiological process**:
$p(\theta' \mid \theta)$—probability of perception $\theta'$ in state $\theta$

**assumptions**:
sufficient precision: $p(\theta' \mid \theta) \leq p(\theta \mid \theta)$ for all $\theta' \neq \theta$
symmetry: $p(\theta' \mid \theta) = p(\theta \mid \theta')$

**definition**:
$\theta_1$ is more distinct than $\theta_2$ if for all $\theta_3 \neq \theta_1, \theta_2$, $p(\theta_3|\theta_1) < p(\theta_3|\theta_2)$

**result**:
If $\theta$ is more distinct than $\theta'$, then the optimal perception discriminates in favor of $\theta$: $r(\theta \mid \theta') > r(\theta' \mid \theta)$.

# Outline

# Sophisticated decision procedures

**limitations assumed so far**:

- coarse termination strategy
- very small memory

**the model accommodates agents who**:

- condition termination on “rich” information
- remember some information from previous runs
- choose how much to forget

# Rich information

$\mu(x \mid \theta)$—the available statistical experiment

agent conditions terminations and choice on $x$

termination probabilities $(\gamma_x)_x$ and action choice $a = \sigma(x)$

effective choice rules $p(a \mid \theta; \gamma, \sigma)$

$P$—set of all feasible choice rules in the new model

Optimal choice rule is second-thought-free.

**Proof**: embed the new model into the main model:

let the agent choose any $p \in P$ and $(\beta_a)_a$

optimal $r(a \mid \theta; p, \beta)$ is second-thought-free

the set of all feasible $r(a \mid \theta; p, \beta)$ coincides with $P$:

$r(p(\gamma, \sigma), \beta) = p(\gamma', \sigma)$ where $\gamma'_x = \gamma_x \beta_{\sigma(x)}$

# Information from previous runs

## Wilson (2014) with endogenous termination

primitive statistical experiment $\mu(x \mid \theta)$

set $M$ of mental states $m$, the pocess starts at $m_0$

at the end of each run of $\mu$, the agent can terminate or transition into a mental state $m$ and continue

mixing conditional on $(x, m)$

action choice $a = \sigma(x, m)$

Optimal choice rule is second-thought-free.

**proof**: embed the setting the main model

optimal rule in the embedded model is second-thought-free

the embedding doesn't expand the set of the feasible rules

## Partial forgetting

agent comprehends up to $N$ signal draws from $\mu(x \mid \theta)$

at each signal history $h$, $|h| \leq N$, she can mix over:

- terminate and choose $\sigma(h)$, or
- return to any truncation $h'$ of $h$, or
- if $|h| < N$, then acquire a new signal

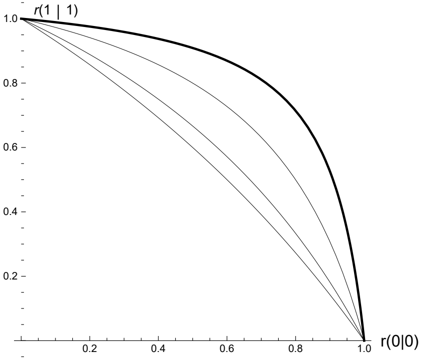
Optimal choice rule is second-thought-free.

the same proof

# Binary setting

The solution for the simple agent extends to all three examples

**argument**: each $p \in P$ induces a set of feasible rules.
the rule $p$ with the maximal perceptual distance dominates all



# Opportunity cost

linear opportunity cost of time

the memory management problem becomes

$$\max_{\beta,p} \mathrm{E}_{r(\beta,p)}[u(a,\theta) - t]$$

the second-thought-free condition:

$$\mathrm{E}[u(a_1,\theta) \mid a_1 = a] = \mathrm{E}[u(a_2,\theta) - t \mid a_1 = a] \text{ if } \beta_a \in (0,1)$$

# Conclusion

a model of the second (and further) thoughts

productive hesitation:

selective repetitions of a decision process affect correlation between the choice and state

the optimal effective decision process is second-thought-free: the condition doesn't refer to the unobservable $P$

confirmation bias, speed-accuracy complementarity, overweighting of rare events, salience