Learning with Limited Memory: Bayesianism vs Heuristics*

Kalyan Chatterjee
Penn State University

Tai-Wei Hu†
University of Bristol

May 18, 2021

Abstract

Bayesian analysis is considered the optimal way of processing information. However, it often leads to problems for decision-makers with constrained cognitive capacity. Modelling such constrained capacity by finite automata, we answer two questions in the context of Wald's (1947) sequential analysis, namely in what environments is optimal Bayesian analysis possible even with constraints; also, when it is not possible what simplifications in the analysis enable us to obtain a satisfactory outcome. We identify two features of the simplified analysis: information stickiness (ignoring information) and rule stickiness (ignoring small differences in the environment).

JEL classification: D81, D83.

Key words: Imperfect recall, bounded rationality, bounded memory, heuristics, behavioural biases.

*We benefited from discussions with Miguel Ballester, Faruk Gul, Ehud Kalai, Mamoru Kaneko, Yuval Salant, Eran Shmaya, and Olivier Tercieux. We thank seminar participants at Northwestern University, University of Oxford, University of York, Florida State University, One World Mathematical Game Theory Seminar, Penn State, and the 2021 NBER Decentralization Conference. We also thank Konstantin Guryev for a careful reading of the draft.

†Corresponding author; email address: taiwei.hu@bristol.ac.uk

1 Introduction

Bayesian inference is generally considered as the optimal way to process information in economic models. Even Tversky and Kahneman (1974), who examine systematic deviations from Bayesian conclusions, consider such deviations as biases or "mistakes." However, Bayesian analysis can be too complicated to pursue successfully due to cognitive limitations. For example, a court in England ruled out its use because it would introduce complications of a computational nature and prevent proper consideration of important factors.[1] Indeed, Simon (1956) suggests that "a great deal can be learned about rational decision making by taking into account, at the outset, the limitations upon the capacities and complexity of the organism, and by taking account of the fact that the environments to which it must adapt possess properties that permit further simplification of its choice mechanisms."

Following Simon's suggestion, we examine Bayesian inference by imposing limitations on the decision-maker's capacity for information processing. We ask two questions. First, in what environments can Bayesian learning be implemented by a boundedly-rational decision-maker with a finite capacity to process information? Second, what are the characteristics of the second-best procedures when Bayesian learning is infeasible? The aim of the first question is to identify "simple" environments where Bayesian inference is appropriate for both normative use and descriptive modelling. The goal of the second question is to distinguish biases arising from constrained optimality against genuine "mistakes." To match the finite ability we impose on the decision-maker, we consider a classical model of Bayesian inference that features an (essentially) finite stopping time.

Specifically, we study the Wald (1947) sequential sampling problem of hypothesis testing. The decision-maker aims at matching her action with the unknown state of nature, either High (H) or Low (L), based on a sequence of informative signals. The decision-maker has to decide either to acquire more information, which is costly due to discounting, or to take a terminal decision. For an unconstrained decision-maker, it is optimal to follow Bayes rule to update beliefs and to take the terminal action only when she is sufficiently convinced of the true state of nature. A positive cost of information acquisition implies that the learning stops within finite time almost surely under the optimal rule.

We model imperfect information processing by restricting the decision-maker to use strategies that can be implemented by a *finite automaton*.[2] This consists of finitely many

[1] The judge in R v Adams [1996], Royal Court of Justice, summarised that, "Quite apart from these general objections, as the present case graphically demonstrates, to introduce Bayes theorem, or any similar method, into a criminal trial plunges the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task."

[2] This modelling approach has been recently carefully evaluated by Oprea (2020), and is supported by Banovetz and Oprea (2020) with experimental evidence.

mental states, and a transition rule taking the process from a given state to another depending on the signal received in that period. The mental states are partitioned into *updating* states and *action* states. The process unfolds as the decision-maker starts with an initial state based on her prior belief, and, after observing a signal, decides which mental state to go to (a decision that might involve randomized transition to several states). An action state corresponds to a terminal action, and reaching one ends the process. This restriction induces two properties. First, it limits the decision-maker's capacity to store information. Second, it limits her computational ability to update complicated posterior beliefs, as the finiteness of the set of mental states implies a limited capacity to distinguish close numbers.

Main results

Most of our results can be illustrated with two signals, h and ℓ, that move the posterior toward states of nature H and L, respectively. We first identify two classes of "simple" information structures (environments in Simon's terminology) where Bayesian learning can be implemented by (deterministic) finite automata. The first class consists of "models of breakthroughs," where one of the signals, say ℓ, fully reveals state of nature L. The second class consists of information structures where the log likelihood-ratios of the signals are rational proportions of one another. Outside these two classes, however, we show that Bayesian learning is generically infeasible, despite the finite nature of the Wald problem.

Our main results characterize the constrained optimal rule for a given memory constraint, where the decision-maker is constrained to finite automata with no more than K updating states. Our results highlight two behavioural biases under constrained optimality. The first bias features lack of response to new information. In our framework this means that informative signals do not trigger transition in mental states, or *information stickiness*. The second implication is lack of flexibility to adopt new rules to changes in the information structures, or *rule stickiness*. This second implication reveals the heuristic nature of the constrained optimal rule induced by bounded rationality—a simple rule is optimal for a range of complex environments.

Our first main result shows that information stickiness is a prominent feature in the models of breakthroughs whenever the memory constraint binds. In this class of models, since an ℓ-signal fully reveals L, it leads to immediate action; in contrast, receiving an h would gradually increase the posterior on H until it reaches the threshold for the other action. A finite automaton can implement this rule if equipped with enough memory capacity, with updating states ranked according to the corresponding posterior on H and an h-signal leading to the next updating state until it is optimal to take the terminal action. However, for a given number of updating states, K, the constraint becomes binding if the discount factor is above a threshold. When it is binding, we fully characterize the constrained optimal rule, which involves randomization in every period in which an h-signal is observed.

This result implies that constrained optimal rules feature information stickiness. Indeed, randomization implies that all updating states are sticky: an informative signal (h) may not trigger a transition, which results in inaccurate inferences and observable deviations from standard Bayesian learning. Intuitively, the limited number of updating states prevents the decision-maker from optimal learning, but randomization substitutes additional mental states by delaying transition. In fact, randomization can be a very efficient substitute for memory when the discount factor is high, and we show that a small automaton (with one updating state only) using optimal randomization can guarantee a payoff arbitrarily close to the first-best for sufficiently high discount factors. Moreover, this small randomized automaton does better than a deterministic automaton with several more states. The key insight here is that randomization is an effective substitute for complexity when limited information storage capacity is pronounced.

Our second main result features optimal heuristics that process information by simple cancellation rules, implying *rule stickiness*. In a symmetric situation, an h-signal moves the posterior on H higher and an ℓ moves it lower but they are of similar strength. When the strengths are exactly the same, a deterministic finite automaton can implement the unconstrained optimal rule: with updating states ranked according their corresponding posterior on H, an h-signal pushes the updating state to the higher one while an ℓ-signal pushes to the lower one. Now, if the strengths of the two signals are not exactly the same but close, the updating process becomes very complicated and no finite automaton can implement the unconstrained optimal rule. Our main finding here is that when the relative strength of the two signals is similar but is complex, the optimal finite automaton is deterministic and treats the strengths of the two signals as exactly the same. In fact, we show that it is optimal to stick to this simple cancellation rule even when larger automata are available.

We extend this rule-stickiness result to other simple cancellation rules. If the log likelihood-ratio of one signal is a rational multiple of the other, then a similar cancellation idea still works. For example, one high signal may cancel two low signals. If the cancellation is not exact, we can use "approximate" probabilities to cancel out likelihood ratios that are close to each other in a specific sense, and we extend the above rule-stickiness result to environments near any signal structures that feature rational ratios. What constitutes a simple rule depends on the underlying constraint, but a common key feature is that given the constraint, the same heuristic rule is optimal for a range of complex environments. We also provide a numerical example to illustrate that the range can be fairly wide.

The identified heuristic rule is reminiscent of Benjamin Franklin's "decisional balance sheet,"where he suggests listing pros and cons for a given choice and cancelling a pro with a con if they are approximately of the same importance, or a pro with two cons if the weights match.[3] With multiple signals, one with a low enough likelihood ratio (a small signal) is ignored in the approximation, even though Bayesian analysis would take it into account, and hence information stickiness also occurs. These results then demonstrate the optimality of using imprecise or qualitative probabilities as heuristics to guide the behaviour, as our decision-maker behaves *as if* she faces a simple environment where signals can be cancelled with one another according to a simple rule, but in fact the precise inference in the true environment requires complex computations.

Finally, our results speak to the less-is-more effect in psychology. Parpart, Jones and Love (2018) note, for example, that we often "ignore a great deal of information in the input data. One puzzle is why (such) heuristics can outperform full-information models, such as linear regression, which make full use of the available information. These 'less-is-more' effects, in which a relatively simpler model outperforms a more complex model, are prevalent throughout cognitive science."In the model of breakthroughs, a small stochastic finite automaton out-performs large deterministic ones; in the more symmetric case, simple cancellation rules outperform more complicated rules. Both results show the optimality of simple rules implementable with small finite automata. However, constrained optimal rules also feature biased beliefs in a formal sense. To describe such biases, we need to put our model in the context of some previous literature, to which we turn next.

Methodological contributions and related literature

Piccione and Rubinstein (1997) pioneered the concept of modified multi-self consistency to describe (biased) beliefs for games with imperfect recall, or imperfect memory. We extend that concept in our context, which states that optimal transitions in a finite automaton of a given size satisfies a form of "sequential rationality," and characterize the corresponding "beliefs" associated with each updating state. This approach is also shared by Wilson (2014),[4] who studies bounded memory in a model where the time horizon is exogenously determined (albeit stochastically). Hence, the learning problem in Wilson (2014) is infinite in nature and the unconstrained optimum is never implementable by finite automata. Wilson (2014) also obtains information stickiness as a prominent feature of the constrained optimal rule, and the beliefs are always biased.

In contrast, there are two classes of environments where unconstrained optimal rules can

[3]A copy of Franklin's letter to Joseph Priestley is at https://founders.archives.gov/documents/Franklin/01-19-02-0200

[4]See also Kocer (2010).

be implemented by (deterministic) finite automata in our framework. In these benchmark cases, the beliefs with respect to the modified multi-self consistency in fact coincide with the ones obtained by Bayes' rule. Once we deviate from these simple environments, however, constrained optimal rules feature biased beliefs. In the model of breakthroughs, the bias comes from randomization. In the class where the decisional balance sheet is optimal, the constrained optimal rule is deterministic, but beliefs are biased because the heuristic rule applies *wrong* cancellation rules, despite its constrained optimality.

However, these beliefs, together with modified multiself consistency, allow for a novel technique to obtain our results. For the model of breakthroughs, it allows us to fully characterize the constrained optimal rule for any given number of updating states. To our knowledge, this is the first result with a closed-form solution for any constraint in a class of information structures. For the rule-stickiness result, we need to demonstrate the optimality of a deterministic rule, and the difficulty is to rule out any deviation by randomization. The modified multiself consistency provides a necessary condition to rule them out: randomization is optimal only if the decision-maker is indifferent (according to the associated beliefs) between two updating states. Thus, to demonstrate that randomization plays no role when the relative strengths of the signals are close to a rational proportion, we show that the optimal finite automaton always features strict preference for transition in the benchmark simple environment and remains so in nearby environments. To our knowledge, ours is the first to show this heuristic feature of the constrained optimal rule, and this can only happen when the constrained optimal rule is deterministic.

This paper is also related to a growing literature that introduces frictions in information processing to explain behavioural biases. Salant (2011) rationalizes another well-known heuristic rule, the satisficing criterion, with state-complexity in a model with finite automata. Compte and Postlewaite (2012) explore heuristics and how these lead to numerical beliefs, focusing on an explanation of beliefs that takes into account only extreme evidence, much like the example in Wilson (2014)'s framework with high discount factors. In contrast, our heuristics are based on qualitative probabilities. Another friction in the literature is bounded recall, a special case of a finite state automaton.[5] In this strand of work, Jehiel and Steiner (2019) consider an inference problem with endogenous termination, but focus mainly on the case with memory of one where the decision-maker takes an observation and decides to act or to discard the observation and take another one. The main result is a necessary

[5] Kuhn (1953) already discusses memory in extensive form games with the notion of (im)perfect recall, and Piccione and Rubinstein (1997) point out the benefit of randomization under imperfect recall. More recently, people have studied bounded recall in repeated games, e.g., Barlo, Carmona and Sabourian (2009) on folk theorems with a memory of one and later more general bounded recall. There are also models that use finite automata in repeated games where randomization occurs in equilibrium, e.g., Ben Porath (1993).

condition for optimal randomization called "second-thought freeness,"which is similar to the indifference condition implied by multi-self consistency. In contrast, we tackle the optimality of randomization under a general complexity constraint, and the relationship of the optimal complexity of an automaton to the discount factor.

2 The model

The model is based on the classical sequential hypothesis testing problem of Wald (1947). Consider a decision-maker (DM) who faces a decision problem as follows. There are two states of nature, $\theta \in \Theta = \{H, L\}$, and the prior probability over $\theta = H$ is $\mathbf{P}_0(H) = \pi_0$ and over $\theta = L$ is $\mathbf{P}_0(L) = 1 - \pi_0$. The DM has to choose from two actions, $a \in A = \{a^H, a^L\}$, to match the state of nature, with the utility function

$$\begin{aligned} u(a^H, H) &= u^H(> 0),\ u(a^H, L) = 0, & (1) \\ u(a^L, L) &= u^L(> 0),\ u(a^L, H) = 0. & (2) \end{aligned}$$

The state of nature, however, is not observable to the DM. Instead, the DM can observe a sequence of signals, and the sequence is i.i.d. conditional on the state. The set of possible realizations at each date is denoted by signal x from a finite set X, which are i.i.d., conditional on θ, across time. Conditional on state of nature θ, $x \in X$ occurs with probability μ_x^θ. Learning, however, is costly, and the DM discounts the future with discount factor $\delta \in (0, 1)$. Thus, if the DM takes an action in period 0 without seeing any signals, then her payoff is not discounted. Otherwise, she can choose to see one signal at period 1 (continue), and if she makes a decision in the end at period 1, her payoff is discounted by δ. Of course, she can choose to continue in period 2 as well, and so on. For learning to have some value, we assume

$$\mu_x^H/\mu_x^L > 1 > \mu_{x'}^H/\mu_{x'}^L \text{ for some } x, x' \in X. \quad (3)$$

Let $X^* = \bigcup_{t=0}^{\infty} X^t$ denote the set of all partial histories of signal realizations, and let $\emptyset$ denote the empty history. A typical non-empty partial history is denoted by $\mathbf{x} = (x_1, ..., x_t)$, where $x_s \in X$, $s = 1, .., t$. We use $\mathbf{x} \subseteq \mathbf{y}$ to denote the fact that $\mathbf{x}$ is an initial segment of $\mathbf{y}$ (strict inequality means the two are not equal). A decision rule is then a function $f : X^* \to \{a^H, a^L, c\}$, where c denotes continue. With no loss of generality, we assume that if $f(\mathbf{x}) \neq c$, then $f(\mathbf{x}) = f(\mathbf{y})$ for all $\mathbf{y} \subset \mathbf{x}$ such that $f(\mathbf{y}) \neq c$ and $f(\mathbf{x}) = f(\mathbf{z})$ for all $\mathbf{z}$ such that $\mathbf{x} \subseteq \mathbf{z}$. Given a decision rule and a prior $\mathbf{P}_0$, the associated expected payoff is given by

$$U(f) = \sum_{\theta \in \Theta} \mathbf{P}_0(\theta) \left\{ \sum_{\mathbf{x} \in X^t, t=0,1,2,..} \left\{ \left[\prod_{s=1}^{t} \mu_{x_s}^{\theta} \right] \delta^t u^\theta : f(\mathbf{x}) = a^\theta; f(\mathbf{y}) = c \text{ for all } \mathbf{y} \subset \mathbf{x} \right\} \right\}. \quad (4)$$

Note that if $f(\emptyset) \neq c$, then the DM's payoff is not discounted. The following proposition characterizes the optimal decision rules when there is no memory constraint. We use π to denote a generic posterior after some observations of signals.

Proposition 2.1. *Assume (3). There are two thresholds, $\bar{\pi}$ and $\underline{\pi}$ such that the optimal decision rule is to take a^H for posteriors $\pi > \bar{\pi}$, take a^L for $\pi < \underline{\pi}$, and take c for $\pi \in (\underline{\pi}, \bar{\pi})$. Moreover, the threshold $\bar{\pi}$ strictly increases with δ while $\underline{\pi}$ strictly decreases, with* $\lim_{\delta \to 1} \bar{\pi}(\delta) = 1$ *and* $\lim_{\delta \to 1} \underline{\pi}(\delta) = 0$.

To calculate the posterior, it is in fact more convenient to work with log likelihood-ratios. For any posterior, $\pi \in (0,1)$, as probability over H, define

$$r_\pi \equiv \ln[\pi/(1-\pi)].$$

Then, for any given interim posterior π, the log likelihood-ratio of the posterior, denoted by π', after the partial history of signal realizations, $\mathbf{x} = (x_1, ..., x_t)$ is then

$$r_{\pi'} = r_\pi + \sum_{s=1}^{t} r_{x_s} = r_\pi + r_{\mathbf{x}}, \quad (5)$$

where $r_{\mathbf{x}} = \sum_{s=1}^{t} r_{x_s}$, and $r_x \equiv \ln(\mu_x^H/\mu_x^L)$ for each $x \in X$. Thus, $r_{\mathbf{x}}$ is the sum of the log likelihood-ratios of signals from the partial history $\mathbf{x}$.

3 Finite automata and multi-self consistency

To model bounded rationality, we focus on decision rules that can be implemented with finite automata. We consider two classes of finite automata: deterministic ones and stochastic ones. Note that since the unconstrained optimal rule identified in Proposition 2.1 involves no randomization, we only need to consider the first class to determine whether the unconstrained optimal rule can be implemented by a finite automaton. The second class, however, is relevant when Bayesian learning is infeasible.

Given the set of signals, X, a *deterministic finite-state automaton*, abbreviated as DFSA, $M = (Q, \tau, q^o)$, consists of the following objects: a finite set of *updating* states Q, a transition function $\tau : Q \times X \to Q \cup \{q_H, q_L\}$, where $q_\theta \notin Q$ is an *action state*, corresponding to taking action a^θ, $\theta = H, L$ (and the game is terminated whenever an action state is reached), and an initial state $q^o \in Q \cup \{q_H, q_L\}$. Given M, we can extend the transition function τ to all partial histories $(x_1, ..., x_t) \in X^*$ by introducing another function λ_M that maps partial histories in X^* to $Q \cup \{q_H, q_L\}$, with the obvious restriction that once it maps to an action state the function stops there, defined by induction as follows: $\lambda_M(\emptyset) = q^o$; given $\lambda_M(x_1, ..., x_{t-1}) = q_{t-1} \in Q$,

$$\lambda_M(x_1, ..., x_t) = \tau(q_{t-1}, x_t). \tag{6}$$

We say that *a decision rule f can be implemented with the DFSA M* if for all $\mathbf{x} \in X^*$,

$$\begin{aligned} \lambda_M(\mathbf{x}) &\in Q \text{ whenever } f(\mathbf{x}) = c; \qquad (7) \\ \lambda_M(\mathbf{x}) &= q_\theta \text{ whenever } f(\mathbf{x}) = a^\theta \text{ and this is the first time along } \mathbf{x} \text{ for this to happen.} \end{aligned}$$

We focus on the DM with a constraint on the number of updating states in the finite automaton. In particular, we assume that only finite automata with fewer than K updating states are feasible. When the constraint, K, is binding, the insight from the literature is that randomization can be optimal (e.g., Hellman and Cover (1971) and Wilson (2014), in a different context). Accordingly, we introduce stochastic finite-state automata (SFSA), and allow for randomization in the finite automata used to implement the decision rules. We use $\tau(q, x; q')$ to denote the transition probability, from updating state $q \in Q$ to $q' \in Q \cup \{q_H, q_L\}$, after receiving signal x.[6] Note that DFSA is also a SFSA.

We first characterize the optimal SFSA for a given K by some necessary conditions. The key state variable in the unconstrained optimal rule in Proposition 2.1 is the posterior belief that is updated according to the signals received. Following the methodology in Wilson (2014), our necessary conditions also characterize a "belief" for each updating state in the optimal SFSA, based on modified multi-self consistency according to Piccione and Rubinstein (1997).

Consider an optimal SFSA, M, under the constraint $|Q| \leq K$. We begin with implications of optimality for action states. Given a state of nature θ and the action state q_θ, the expected payoff accumulated from q_θ conditional on θ (adding together discounted payoffs for all paths

[6] It is without loss of generality to assume that no randomization occurs when taking the final action, a result noted in Kalai and Solan (2003).

from the initial state to q_θ) is then given by $\mathbb{P}(q_\theta|\theta)u^\theta$, where for any $q \in Q \cup \{q_H, q_L\}$,

$$\begin{aligned}\mathbb{P}(q|\theta) = \mathbf{1}_{q=q^o} + \delta \sum_{x_1 \in X} \mu^\theta_{x_1} \tau(q^o, x_1; q) + \delta^2 \sum_{x_1,x_2 \in X, q_1 \in Q} \mu^\theta_{x_1} \tau(q^o, x_1; q_1) \mu^\theta_{x_2} \tau(q_1, x_2; q) \\ + \delta^3 \sum_{x_1,x_2,x_3 \in X, q_1,q_2 \in Q} \mu^\theta_{x_1} \tau(q^o, x_1; q_1) \mu^\theta_{x_2} \tau(q_1, x_2; q_2) \mu^\theta_{x_3} \tau(q_2, x_3; q) + \cdots\end{aligned} \tag{8}$$

and where for any $q \in Q$, $\mathbf{1}_{q^o=q} = 1$ if $q = q^o$ and $= 0$ otherwise. When $q = q_\theta$, the quantity $\mathbb{P}(q_\theta|\theta)$ has a natural interpretation: if δ were equal to one, then it would be the total probability that the SFSA M reaches q_θ; but since $\delta < 1$, the time when M reaches q_θ and takes action a^θ matters for the expected payoffs and in $\mathbb{P}(q_\theta|\theta)$ they are discounted accordingly. If M is optimal and if $\mathbb{P}(q_\theta|\theta) > 0$, then it follows that, for $\theta' \neq \theta$,

$$\mathbf{P}_0(\theta)\mathbb{P}(q_\theta|\theta)u^\theta \geq \mathbf{P}_0(\theta')\mathbb{P}(q_\theta|\theta')u^{\theta'}, \tag{9}$$

for otherwise one can use action $a^{\theta'}$ at q_θ to obtain higher expected payoffs. Now, (9) is a form of "sequential rationality," which states that conditional upon reaching the action state q_θ, the action chosen must be optimal according to some beliefs. In this case the log likelihood-ratio of such belief at q_θ would be given by

$$r(q_\theta) \equiv \ln\left[\frac{\mathbf{P}_0(H)\mathbb{P}(q_\theta|H)}{\mathbf{P}_0(L)\mathbb{P}(q_\theta|L)}\right] = r_{\pi_0} + \ln\left[\frac{\mathbb{P}(q_\theta|H)}{\mathbb{P}(q_\theta|L)}\right]. \tag{10}$$

Our concept of modified multi-self consistency then extends this form of "sequential rationality" to updating states. For those states, the sequential rationality is about transitions, which happen after receiving new signals and hence we need to extend the formulation of "beliefs" to include signals. In particular, for each $q \in Q$, define

$$r(q) = r_{\pi_0} + \ln\left[\frac{\mathbb{P}(q|H)}{\mathbb{P}(q|L)}\right] \text{ and } r(q,x) = r_{\pi_0} + \ln\left[\frac{\mathbb{P}(q|H)}{\mathbb{P}(q|L)}\right] + r_x. \tag{11}$$

Hence, in $r(q)$ we use the analogous formula to Bayes' rule to include the information contained by signal x. For later purposes, we also use π_q to denote the "belief" of H at updating state q, that is, $r(q) = \ln[\pi_q/(1-\pi_q)]$, and similarly $\pi_{q,x}$ is such that $r(q,x) = \ln[\pi_{q,x}/(1-\pi_{q,x})]$. To formulate the analogous notion of "sequential rationality" for transitions from updating states, we also need to calculate the continuation value for each updating state. These continuation values are determined by the simultaneous equations:

$$V_q(\theta) = \delta \sum_{q',x} \tau(q,x;q')\mu^\theta_x V_{q'}(\theta), \tag{12}$$

while for action state q_θ, it is $V_{q_\theta}(\theta) = u^\theta$ and $V_{q_\theta}(\theta') = 0$ for $\theta' \neq \theta$.

Now we are ready to define modified multi-self consistency.

Definition 3.1. A SFSA M satisfies *modified multi-self consistency* under $\mathbf{P}_0$ if

1. for each updating state $q \in Q$ with $\sum_\theta \mathbf{P}_0(\theta)\mathbb{P}(q|\theta) > 0$, each signal x, and any $q' \in Q \cup \{q_H, q_L\}$ such that $\tau(q, x; q') > 0$,

$$\pi_{q,x} V_{q'}(H) + (1 - \pi_{q,x}) V_{q'}(L) \geq \pi_{q,x} V_{q''}(H) + (1 - \pi_{q,x}) V_{q''}(L) \text{ for all } q'' \in Q; \quad (13)$$

2. for action states,

$$\pi_{q_H} u^H \geq (1 - \pi_{q_H}) u^L \text{ and } (1 - \pi_{q_L}) u^L \geq \pi_{q_L} u^H. \quad (14)$$

Clearly, (14) follows immediately from (9) and (10), and (13) extends (14) to updating states but takes signals into account according to (11). Now we are ready to give a full characterization of optimal SFSA. We say that two updating states q and q' are *equivalent* if $V_q(\theta) = V_{q'}(\theta)$ for both $\theta = H, L$. Then, we can keep the same ex ante payoff by eliminating one of them.

Theorem 3.1. *Any optimal SFSA under the constraint $|Q| \leq K$ satisfies modified multi-self consistency. Moreover, suppose that M is an optimal SFSA without equivalent states among those with $|Q| \leq K$. We rank the updating states in M according to*

$$r(q_1) \leq r(q_2) \leq \cdots \leq r(q_K),$$

with the convention that if $r(q_i) = r(q_{i+1})$, then $V_{q_i}(H) \leq V_{q_{i+1}}(H)$. Let $\Delta V^\theta_{i,j} \equiv V_{q_i}(\theta) - V_{q_j}(\theta)$, with the convention that $V_{q_0}(\theta) = V_{q_L}(\theta)$ and $V_{q_{K+1}}(\theta) = V_{q_H}(\theta)$, and let $\bar{r}_i = \ln\left(\Delta V^L_{i,i+1} / \Delta V^H_{i+1,i}\right)$. Then, for any $q \in Q$,

$$\tau(q, x; q_i) > 0 \text{ only if } r(q, x) \in [\bar{r}_{i-1}, \bar{r}_i], \quad (15)$$

where $\bar{r}_{-1} = -\infty$ and $\bar{r}_{K+1} = \infty$, and $\tau(q, x; q_i) = 1$ if $r(q, x) \in (\bar{r}_{i-1}, \bar{r}_i)$.

Theorem 3.1 gives a structural characterization of the optimal SFSA similar to Proposition 2.1 in the following sense. If we let $\bar{r}_H \equiv \bar{r}_{K+1}$ and let $\underline{r}_L \equiv \bar{r}_0$, then Theorem 3.1 states that $\tau(q, x; q_H) > 0$ only if $r(q, x) \geq \bar{r}_H$ and $\tau(q, x; q_H) = 1$ if the inequality is strict, that is, it is optimal to take action a^H whenever the "posterior" log likelihood-ratio $r(q, x)$ is above a threshold. Similarly, taking action a^L is optimal whenever such posterior is below $\underline{r}_L$. However, different from Proposition 2.1, the posteriors are not necessarily updated according to Bayes' rule. In fact, from the proof we know that $r(q_i) \in [\bar{r}_{i-1}, \bar{r}_i]$ for all updating states q_i in an optimal SFSA M, and the thresholds $\bar{r}_i$'s are constructed from the optimality condition (13) and (14). An important implication is that, if $\tau(q, x; q_i) > 0$ and $\tau(q, x; q_j) > 0$ for $i < j$ in the optimal M, then it must be the case that $r(q, x) = \bar{r}_i = \bar{r}_{j-1}$, that is, randomization in transition can occur only if there is indifference according to the beliefs. Wilson (2014) also has a similar characterization result. However, a crucial difference is that in our framework there exist "simple" environments where the unconstrained optimal rule is implementable and the beliefs given by (11) in fact coincide with Bayesian learning, but in Wilson (2014)'s framework the unconstrained optimal rule is never implementable by SFSA and beliefs are always biased.[7] The following proposition gives sufficient conditions for this to happen. Recall the function λ_M given by (6) for a given DFSA M.

Proposition 3.1. *Suppose that the DFSA M satisfies the following conditions:*

(O1) If $\lambda_M(\mathbf{x}) = \lambda_M(\mathbf{y}) = q \in Q$, then

$$r_{\pi_0} + r_{\mathbf{x}} = r_{\pi_0} + r_{\mathbf{y}}. \tag{16}$$

(O2) If (a) $r_{\pi_0} + r_{\mathbf{x}} \geq r_{\bar{\pi}}$ or (b) $r_{\pi_0} + r_{\mathbf{x}} \leq r_{\underline{\pi}}$ and this does not hold for any $\mathbf{y} \subset \mathbf{x}$, then $\lambda_M(\mathbf{x}) = q_H$ in case (a) holds or q_L in case (b) holds.

Then, M implements the unconstrained optimum. Moreover, for each $q \in Q$, $r(q)$ is computed according to (16) using any partial history $\mathbf{x}$ such that $\lambda_M(\mathbf{x}) = q$, that is,

$$r(q) = r_{\pi_0} + r_{\mathbf{x}}. \tag{17}$$

Moreover, if the optimal DFSA is unique, $\bar{r}_{i-1} < r(q_i) < \bar{r}_i$ for all updating states q_i.

The sufficient conditions (O1) and (O2) in Proposition 3.1 ensure that M implements the unconstrained optimal rule; indeed, it follows immediately that under M, an action is taken if and only if the posterior has crossed the upper bound $\bar{\pi}$ or the lower bound $\underline{\pi}$ for the first time. Equation (17), however, is less obvious. The key observation leading to this result is the following. The likelihood ratio, $\pi_q/(1-\pi_q)$, is equal to the ratio $[\mathbf{P}_0(H)\mathbb{P}(q|H)]/[\mathbf{P}_0(L)\mathbb{P}(q|L)]$

[7] Indeed, beliefs in Wilson (2014) are also given by expressions to similar to $\mathbb{P}(q|\theta)$. However, a crucial difference is that the sum $\sum_{q \in Q} \mathbb{P}(q|\theta)$ is endogenously determined by the transition to the action states in our model, while the analogous sum in Wilson (2014) is equal to a constant for any SFSA—reflecting the fact that the stopping time is exogenously determined there. The endogenous stopping time is almost surely finite in our model, which allows for such simple environments and unbiased beliefs to exist.

by (11). But by (8), the term $\mathbb{P}(q|\theta)$ is the discounted sum over the probabilities that q is reached. More precisely, for each partial history $\mathbf{x}$ such that $\lambda_M(\mathbf{x}) = q$, let $\mathbb{P}(\mathbf{x}|\theta)$ denote its probability under state of nature θ. If $\mathbf{x}$ has length n, then it enters the sum $\mathbb{P}(q|\theta)$ by $\delta^n \mathbb{P}(\mathbf{x}|\theta)$. Now, (O1) implies that the ratio $[\mathbf{P}_0(H)\mathbb{P}(\mathbf{x}|H)]/[\mathbf{P}_0(L)\mathbb{P}(\mathbf{x}|L)]$ is constant for all $\mathbf{x}$ such that $\lambda_M(\mathbf{x}) = q$, which coincides with the ratio according to Bayes' rule. As a result, we can cancel δ^n's and obtain the desired result.

When $X = \{h, \ell\}$ with $r_h > 0 > r_\ell$, there are two classes of information structures under which we have a DFSA M that satisfies the sufficient conditions, (O1) and (O2), in Proposition 3.1. The two classes are

(I1) $r_\ell = -\infty$;

(I2) $|r_h|/|r_\ell|$ is a rational number.

We give detailed analyses of how to achieve the unconstrained optimal rule by DFSA in the next two sections, but here we give two impossibility results stating that that these two classes of information structures are essentially the only ones where finite automata can implement Bayesian learning.

Theorem 3.2. *Suppose that $X = \{h, \ell\}$ and $r_h > 0 > r_\ell$ with $r_h < +\infty$.*

1. *Given a prior $\pi_0 \in (0, 1)$ and given K, for δ sufficiently high, the unconstrained optimal decision rule cannot be implemented by a DFSA with $|Q| \leq K$.*

2. *Suppose that*

$$r_{\underline{\pi}} < r_{\pi_0} + r_\ell, \;\; r_{\pi_0} + 2r_h < r_{\bar{\pi}}. \tag{18}$$

Then, generically in μ_h^H the optimal decision rule cannot be implemented by a DFSA.

Theorem 3.2 (1) states that in any non-trivial information structure, the constraint K will eventually be binding as δ converges to one. This is because the threshold $\bar{\pi}$ converges to one as δ converges to one. For any prior $\pi_0 < 1$, it requires arbitrarily many h's to reach $\bar{\pi}$, and this would then require many updating states. This demonstrates that the unconstrained optimal rule is complex as it requires large *information-storage capacity.* As we will see in Section 4 below, this result will be particularly pronounced in class (I1) and constrained optimal rule will feature information stickiness.

Theorem 3.2 (2), in contrast, shows that any generic information structure is complex in that *no* finite automaton can implement Bayesian learning, as long as δ is above a minimum threshold according to (18) to avoid trivial cases. Genericity is needed to exclude information structures in class (I2). Specifically, we show that whenever the ratio $|r_h|/|r_\ell|$ is a *normal* number, i.e., for any finite patterns of the form $t_1 t_2 \cdots t_n$, where $n \in \mathbb{N}$ and $t_1, ..., t_n \in \{0, 1, .., 9\}$, we can find a segment that coincides with that pattern in its decimal expansion, no DFSA can implement the unconstrained optimum. Since normality is a generic property,[8] almost all information structures are too complex to implement Bayesian learning by finite automata. While the proof is technical in nature, the intuition for Theorem 3.2 (2) is the following. When the relative likelihood ratio of the two signals is a complicated number, the computation of posteriors exactly according to Bayes rule requires many updating states to place many different sequences of signals into different categories. When the ratio is a normal number, these computations that are relevant for final actions turn out to require an unbounded number of categories.

As we will see in Section 5 below, we use information structures in class (I2) as benchmark cases to study constrained optimal rules when this *computational* aspect of complexity is pronounced as in Theorem 3.2 (2). We show there when the underlying environment is too complex in this sense, the constrained DM optimally simplifies the environment to act as if it is one that can be implemented by a DFSA, a rule-stickiness result.

4 Model of breakthroughs and information stickiness

Here we consider information structures in class (I1), which also have been used in the literature to model "breakthroughs."[9] In this model, one signal is much more informative than the other. In particular, one signal, say the signal ℓ, is fully revealing, and hence represents the "breakthrough." Accordingly, here we assume that for some $\mu \in (0, 1)$,

$$\mu_h^H = 1, \ \mu_\ell^L = \mu. \tag{19}$$

This implies that $r_h > 0$ and $r_\ell = -\infty$, and the h-signal increases the posterior on H.

We first consider the unconstrained optimal rule. One convenient feature of this model is that it is optimal to take a^L whenever the signal ℓ is received, as the posterior on L immediately jumps to one. Hence, the threshold $\bar{\pi}$ can be determined in closed-form:

$$\bar{\pi} u^H = \delta[\bar{\pi} u^H + (1 - \bar{\pi}) \mu u^L]. \tag{20}$$

To understand (20), note that when π is slightly below the threshold $\bar{\pi}$, the optimal rule dictates the DM to take a^H if signal h is received next period and to take a^L if ℓ is, and these

[8] The notion of genericity here is that the set of real numbers that are not normal has Lebesgue measure zero, a fact proved by Borel (1909).

[9] See, for example, the one used by Che and Mierendorff (2019).

are the payoffs on the right-side of (20). At the threshold $\bar{\pi}$ the DM is indifferent between this and taking a^H immediately, whose payoffs are on the left-side of (20). The lower bound, $\underline{\pi}$, is determined by comparing the optional value by accumulating h-signals against that by taking a^L immediately.

Of course, the thresholds are meaningful only if it is optimal to wait for at least one more period, which requires δ to be sufficiently high. We use δ_0^b to denote such a threshold for the discount factor. For $\delta < \delta_0^b$, the optimal choice is to take the action immediately: a^H if $\pi_0 > \pi^o$ and a^L if $\pi_0 < \pi^o$, where

$$\pi^o = u^L/(u^H + u^L). \tag{21}$$

We focus on the case where $\pi_0 \geq \pi^o$, and the unconstrained optimum is characterized by the number of h-signals before taking a^H (of course, whenever an ℓ is received a^L is taken). The other case is similar but the optimum also requires comparison against taking a^L immediately. That case adds no significant insight but makes the analysis more complicated, and we discuss it at the end. The following lemma characterizes the optimal decision rule for the unconstrained optimum under $\pi_0 \geq \pi^o$.

Lemma 4.1. *Let $\pi_0 \in [\pi^o, 1)$ be given. There exists a sequence $\delta_0^b < \delta_1^b < .... < \delta_N^b < ...$ (which depends on π_0) such that for any $N \geq 1$ and any $\delta \in (\delta_{N-1}^b, \delta_N^b)$, the unconstrained optimum is achievable by a DFSA with N updating states but not by one with fewer. For $\delta < \delta_0^b$, it is optimal to take a^H immediately.*

The sequence is constructed as follows, using the result from Proposition 2.1 that $\bar{\pi}$ strictly increases with δ and converges to one with δ. Since $\pi_0 \geq \pi^o$, δ_0^b is the lowest $\delta \geq \delta^o$ such that $\pi_0 \geq \bar{\pi}(\delta)$. Similarly, δ_N^b is the lowest δ such that $r_{\pi_0} + N r_h = r_{\bar{\pi}(\delta)}$ for $N \geq 1$. For $\delta < \delta_0^b$, it is optimal to take a^H immediately. For $\delta \in \Delta_N^b \equiv (\delta_{N-1}^b, \delta_N^b)$, the unconstrained optimum is implementable with the following DFSA, denoted by $M^{b,N}$, where b stands for breakthrough and $N \geq 1$ stands for the number of updating states:

$$Q = \{q_1, q_2, ..., q_N\} \text{ and } q^o = q_1; \tag{22}$$

$$\tau(q_j, h) = q_{j+1} \text{ for } j = 1, ..., N-1, \text{ and } \tau(q_N, h) = q_H; \tag{23}$$

$$\tau(q_j, \ell) = q_L \text{ for all } j = 1, ..., N. \tag{24}$$

According to (22), $M^{b,N}$ has N updating states that start from q_1. According to (24), the transition is such that it goes to the terminal state q_L that takes action a^L upon receiving the signal ℓ. Finally, according to (23), upon receiving h, the state goes up by one until the highest one, q_N, which goes to q_H if another h comes. A graphical representation of $M^{b,N}$

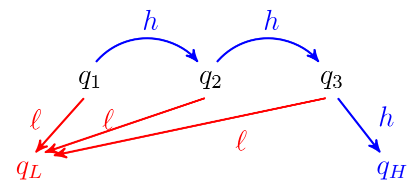

Figure 1: $M^{b,N}$ with $N = 3$

with $N = 3$ can be found in Figure 1.

Constrained optimal rule

Now we turn to constrained optimality. Let π_0 be a given prior and let K be a given constraint on the number of updating states. Lemma 4.1 implies that if $\delta \leq \delta_K^b$, then the unconstrained optimum is implementable. For $\delta > \delta_K^b$, however, the constraint K is binding. As in the unconstrained case, it is optimal to take a^L whenever ℓ is received. Indeed, $r_\ell = -\infty$ implies that $r(q, \ell) = -\infty$ as well for any updating state q by (11). Since $V_q(L) < u^L = V_{q_L}(L)$ for any $q \neq q_L$ due to discounting, Theorem 3.1 implies that it is optimal to transit to q_L whenever ℓ is received. One can show that in this case the optimal deterministic finite automaton with $|Q| \leq K$ is $M^{b,K}$. But can randomization deliver a better payoff? Since it is optimal to transit to q_L whenever ℓ is received, randomization can only occur when h is received. For this to happen at some updating state q, Theorem 3.1 implies there must be two updating states, q' and q'', which have the same expected continuation values according to the belief $r(q, h)$ given by (11).

Now, Proposition 3.1 allows us to verify this necessary condition when δ is close to the boundary, δ_K^b. Indeed, since $M^{b,K}$ satisfies the sufficient conditions (O1) and (O2), Proposition 3.1 implies that we can use Bayes' rule to compute the beliefs when δ is close to δ_K^b, at least as an approximation. At the boundary $\delta = \delta_K^b$, indifference occurs at q_K, where $\pi_{q_K,h} = \bar{\pi}$, and the expected continuation value according to this belief from q_K coincides with that from q_H. This follows from the fact that at $\bar{\pi}$, the DM is indifferent between waiting for exactly one more period, a strategy corresponding to q_K, and taking a^H, corresponding to q_H. For $\delta < \delta_K^b$, $r(q_K, h) > \bar{r}_K$ and hence taking a^H is strictly preferred; for higher δ's we have the opposite inequality and hence waiting is strictly better but constrained by K. An induction argument then shows that we have the same indifference at q_i for all $= 1, ..., K-1$ as well—$r(q_i, h) = \bar{r}_i$ at δ_K^b. This suggests that randomizing between q_i and q_{i+1} at q_i when

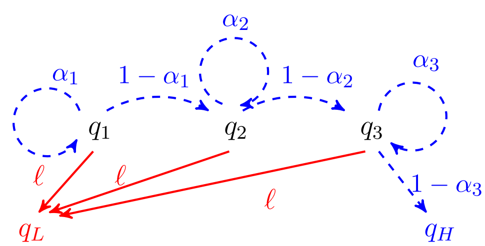

Figure 2: $M^{b,3}(\alpha_1, \alpha_2, \alpha_3)$ with $K = 3$

receiving h can be optimal, as in the following SFSA, denoted by $M^{b,K}(\alpha_1, ..., \alpha_K)$:

$$Q = \{q_1, q_2, ..., q_K\} \text{ and } q^o = q_1;$$
$$\tau(q_j, h; q_j) = \alpha_j = 1 - \tau(q_j, h; q_{j+1}) \text{ for} j = 1, ..., K-1;$$
$$\tau(q_K, h; q_K) = \alpha_K = 1 - \tau(q_K, h; q_H);$$
$$\tau(q_j, \ell; q_L) = 1 \text{ for all } j = 1, ..., K.$$

Compared to (22)-(24), there is randomization when $\alpha_i \in (0, 1)$ for some $i \in \{1, ..., K\}$, and $M^{b,K} = M^{b,K}(0, ..., 0)$. See Figure 2 for a graphical representation.

The following theorem shows that it is in fact optimal to have $\alpha_i > 0$ for all i in $M^{b,K}(\alpha_1, ..., \alpha_K)$ whenever the constraint is binding.

Theorem 4.1. *Let $\pi_0 \in [\pi^o, 1)$ and let memory constraint $K \geq 1$ be given.*

1. *If $\delta \leq \delta^b_K$, the unconstrained optimum is implementable.*

2. *If $\delta > \delta^b_K$, then the optimal SFSA takes the form $M^{b,K}(\alpha_1, ..., \alpha_K)$ with optimal $\alpha_1 = \alpha_2 = ... = \alpha_K = \alpha \in (0, 1)$.*

Theorem 4.1 (1) follows immediately from Lemma 4.1. To prove Theorem 4.1 (2), we need two steps: first we show that any optimal SFSA takes the form $M^{b,K}(\alpha_1, ..., \alpha_K)$, and then we prove that optimal $\alpha_1 = ... = \alpha_K = \alpha$ for some $\alpha > 0$. The main insight to achieve the first step is to observe that we can use backward induction to solve for the optimal SFSA in this model: once q_i is reached, the values of q_j with $j < i$ do not matter for optimality of transitions among higher updating states. Then we use Theorem 3.1 to show that randomization can only occur between adjacent updating states and show by induction that optimal transition always takes the form as in $M^{b,K}(\alpha_1, ..., \alpha_K)$. For the second step, we show that optimal α is strictly positive whenever $\delta > \delta^b_K$ by considering the first-order condition at $\alpha = 0$. We also show that the ex ante payoff is symmetric in $(\alpha_1, ..., \alpha_K)$ and the optimal $(\alpha_1, ..., \alpha_K)$ is symmetric as well. When $\alpha_1 = ... = \alpha_K = \alpha$, we also use $M^{b,K}(\alpha)$ to denote $M^{b,K}(\alpha_1, ..., \alpha_K)$. Theorem 4.1 then shows that randomization is necessary for optimality whenever the memory constraint is binding in the model of breakthroughs. To our knowledge this is the first result that fully characterizes constrained optimal rule for any K for a class of information structures, as well as when randomization is optimal.

This result illustrates information stickiness. Indeed, in $M^{b,K}(\alpha)$, each updating state is *sticky*, as with probability α there will be no transition when receiving h-signal. This also shows that randomization is an efficient substitute for memory capacity. Indeed, whenever $\delta > \delta^b_K$, the unconstrained DM prefers to keep learning after K high signals, but the constrained DM cannot do that without being stuck with perpetual waiting under the high state of nature. Randomization allows for a stochastic horizon for indecision, and we have demonstrated the value of randomization to relax the memory constraint.[10] Finally, whenever the DM is constrained, the posterior according to (11) coincides with the one obtained by Bayes' rule only when breakthrough occurs. When the DM takes action a^H, then randomization is used in all updating states up to the current one, and beliefs given by (11) are biased.

We conclude this section with some comments about the case where $\pi_0 < \pi^o$. When δ is low and hence π_0 is either below or only slightly above $\underline{\pi}$, the expected payoff from the unconstrained optimum is not too different from the expected payoff by taking a^L immediately. If the memory constraint is binding (i.e., $K < (r_{\bar{\pi}} - r_{\pi_0})/r_h$), the payoff from the best SFSA with no more than K updating states is bounded away from the payoff from the unconstrained optimum. As a result, for such cases no SFSA with no more than K states can do better than taking a^L immediately. However, as δ gets larger, it becomes more efficient to use randomization, and we can show that the optimal SFSA requires strict randomization. This last result, in fact, is proved with a very small SFSA, which demonstrates the efficiency of using randomization in this context, a result we turn to next.

Efficiency of randomization

Theorem 4.1 shows that randomization is always useful whenever the DM is constrained, and this happens especially when the cost of information acquisition vanishes or when δ goes to one and when $\mu^H_h = 1$. The following result shows a more radical result for large δ's and high μ^H_h's: a small SFSA can be very efficient.

Theorem 4.2. *Let $\pi_0 \in (0,1)$ be given. For any given ε, there exists $\bar{\delta} < 1$ such that for all $\delta > \bar{\delta}$ and all $\mu^H_h \geq \delta$, the expected payoff from $M^{b,1}(\alpha^*)$ with optimal α^* is within ε compared against the unconstrained optimum.*

[10] See also the example in Hellman and Cover (1971), which is in the context of their (different) problem.

The payoff under the constrained optimum as δ goes to one approaches $\pi_0 u^H + (1 - \pi_0)u^L$, the payoff under full information. As a corollary, Theorem 4.2 implies that, for any given K, for sufficiently high δ, $M^{b,1}(\alpha^*)$ delivers a strictly better payoff than any DFSA with no greater than K updating states, including $M^{b,K}$. Among SFSA, however, since $M^{b,1}(\alpha^*)$ only uses one updating state, one can do better with more states and with randomization. Nevertheless, Theorem 4.2 shows that with randomization the value of the additional updating states diminishes as δ goes to one. Note that the result generalizes to more than two signals in a straightforward manner.

5 Rule stickiness and qualitative probabilities

Now we turn to information structures where no single signal fully reveals the state of nature. In this case, the optimal rule would depend on the relative strength of the signals. We begin with the conceptually simplest case with two signals that are of similar strengths, and then move to the general case. Our main results demonstrate the heuristic nature of the constrained optimal rule in these environments, where the decision-maker employs the same simple rule for a range of log likelihood ratios of the underlying signals.

5.1 Nearly symmetric case

Theorem 3.2 implies that generically the unconstrained optimum cannot be implemented by any DFSA. To approach the constrained optimal decision rule for a given upper bound of updating states, K, we begin with $X = \{h, \ell\}$ and the extreme case where the signals are completely symmetric: for some $\mu \in (0, 1)$,

$$\mu_h^H = \mu = \mu_\ell^L. \tag{25}$$

As we will see below, this knife-edge case will help us study the constrained optimal rule for the more general case where μ_h^H and μ_ℓ^L can be different but close.

The unconstrained case

For any given π_0, we can derive the DFSA that implements the unconstrained optimal rule from Proposition 2.1, which states that the threshold $\bar{\pi}$ strictly increases with δ. By symmetry, we may only consider the case where $\pi_0 \geq \pi^o$, where π^o is given by (21). To derive the memory size needed to implement the unconstrained optimum, we derive thresholds on the discount factors. First, consider the threshold below which taking immediate action is

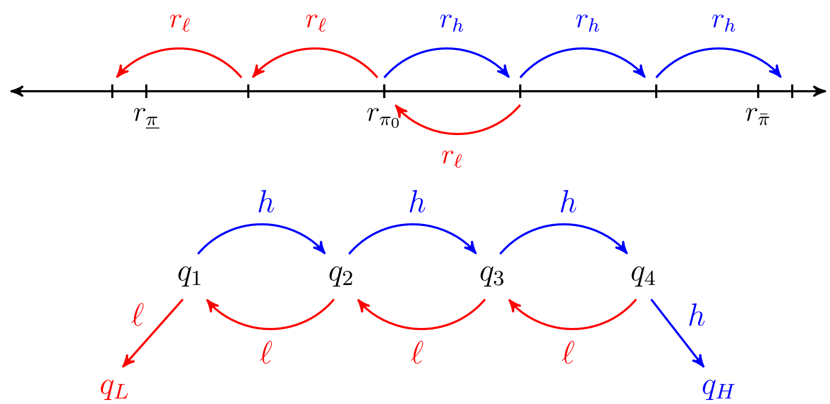

Figure 3: $M^{s,(N-1)}$ with $N = 5$

optimal: the threshold δ_0 is the unique solution to $\pi_0 = \bar{\pi}(\delta)$. If $\delta < \delta_0$, $\pi_0 > \bar{\pi}$ and taking a^H immediately is optimal, and the optimal DFSA only needs one action state, q_H.

For δ higher, the optimal DFSA is determined by the number, $N(\delta)$, given by

$$N = \left\lceil \frac{r_{\bar{\pi}(\delta)} - r_{\pi_0}}{r_h} \right\rceil + \left\lceil \frac{r_{\underline{\pi}(\delta)} - r_{\pi_0}}{r_\ell} \right\rceil, \tag{26}$$

that is, $N(\delta)$ is the sum of the number of h-signals needed to cross $\bar{\pi}$ from π_0 (the first term) and the number of ℓ -signals needed to cross $\underline{\pi}$ from π_0 (the second term). To implement the unconstrained optimum, from the initial state q^o corresponding to π_0, we need $\left\lceil \frac{r_{\bar{\pi}(\delta)} - r_{\pi_0}}{r_h} \right\rceil - 1$ other updating states before reaching q_H, and $\left\lceil \frac{r_{\underline{\pi}(\delta)} - r_{\pi_0}}{r_\ell} \right\rceil - 1$ before reaching q_L. This leads to $N-1$ updating states in total. The corresponding DFSA, denoted by $M^{s,(N-1)}$, with $q^o = q_k$ and $k \equiv \left\lceil \frac{r_{\underline{\pi}(\delta)} - r_{\pi_0}}{r_\ell} \right\rceil$, is given by:

$$Q = \{q_1, ..., q_{N-1}\};\ \tau(q_1, \ell) = q_L,\ \tau(q_{N-1}, h) = q_H;$$
$$\tau(q_i, h) = q_{i+1} \text{ for all } i = 1, .., N-2;\ \tau(q_{i+1}, \ell) = q_i \text{ for all } i = 2, .., N-1.$$

The superscript s in $M^{s,(N-1)}$ refers to "symmetric," or "similar", whose meaning will become clear later. The $N-1$ updating states are ranked: when receiving a high signal, the states transit up until a^H is taken; when receiving a low signal, the states transit down until a^L is taken. A graphical description of the situation in terms of prior and the thresholds for $N = 5$ and $k = 2$ and the corresponding optimal DFSA, $M^{s,(N-1)}$ with $N = 5$ and $q^o = q_2$, can be found in Figure 3. The following lemma fully characterizes the optimal DFSA that implements the unconstrained optimal rule.

Lemma 5.1. *Let π_0 be given. There exists a threshold $\delta_0(\pi_0)$ below which the optimal rule* *is to take one of the actions immediately. For any $\delta > \delta_0$, there exists a number $N \geq 2$ such that the unconstrained optimal rule is implementable by $M^{s,N-1}$ with $q^o = q_k$ and k solving*

$$r_{\underline{\pi}} - kr_\ell \leq r_{\pi_0} \leq r_{\bar{\pi}} - (N-k)r_h, \tag{27}$$

but it is not implementable by any DFSA with fewer than $N-1$ updating states. Moreover, $N(\delta)$ weakly increases with δ and goes to infinity as δ goes to one.

The number N in Lemma 5.1 is given by (26), and since $\bar{\pi}(\delta)$ strictly increases and $\underline{\pi}(\delta)$ strictly decreases with δ, $N(\delta)$ weakly increases with δ. For each $N \geq 2$, define

$$\Delta_N^s = \operatorname{int}\left(\{\delta \in (\delta_0, 1) : N(\delta) = N\}\right), \tag{28}$$

the interior of the set of δ's such that $N(\delta) = N$. Note that Δ_N^s takes the form $(\underline{\delta}_N^s, \bar{\delta}_N^s)$ if non-empty, and Δ_N^s may be empty for some N's. Since symmetry implies that

$$r_{\bar{\pi}(\delta)} + r_{\underline{\pi}(\delta)} = 2r_{\pi^o} \text{ for all } \delta < 1,$$

with π^o given by (21), when δ increases slightly from δ_0, it may require more than one ℓ-signals to cross $\underline{\pi}$, depending on how far π_0 is from π^o. This implies that $\Delta_N^s = \emptyset$ for all $N < N_0 \equiv \lim_{\delta \downarrow \delta_0} N(\delta)$. As δ increases, both $\frac{r_{\bar{\pi}(\delta)} - r_{\pi_0}}{r_h}$ and $\frac{r_{\underline{\pi}(\delta)} - r_{\pi_0}}{r_\ell}$ increase in (26), but the increase leads to an increase in N if one of them, or both, has an increment of exactly one. If $r_{\pi_0} \neq r_{\pi^o} + Mr_h/2$ for any $M \in \mathbb{N}$, then N increases by one once at a time after N_0 and hence Δ_N^s is nonempty and $\bar{\delta}_N^s = \underline{\delta}_{N+1}^s$ for any $N \geq N_0$. Otherwise, N increases by two each time δ hits the boundary, and Δ_N^s is nonempty if and only if $N - N_0$ is even, and $\bar{\delta}_N^s = \underline{\delta}_{N+2}^s$ when $\Delta_N^s \neq \emptyset$. Finally, since $\mu_h^H = \mu_\ell^L$ and hence $r_h + r_\ell = 0$, the DFSA $M^{s,N-1}$ with $q^o = q_k$ according to (27) satisfies the conditions (O1) and (O2) for $\delta \in \Delta_N^s$, and hence it implements the unconstrained optimum by Proposition 3.1. Lemma 5.1 shows that one cannot use a smaller DFSA to implement the unconstrained optimum. Its proof requires a characterization of decision rules implementable by a DFSA of a given size, a well known result called *Myhill-Nerode Theorem* (see Theorem 6.1 in the Appendix).

Constrained optimal rule

Now we study the constrained case. We analyse this by perturbing two parameters, μ_h^H and δ, for a given $\pi_0 > \pi^o$ and $\delta \in \Delta_N^s$. First consider perturbation in μ_h^H under the memory constraint $K \geq N-1$. Since $\delta \in \Delta_N^s$, when μ_h^H is exactly equal to μ, the constraint is not binding. However, Theorem 3.2 implies that, if $N > 4$ and $2 < k < N-1$, then generically the unconstrained optimum cannot be implemented by any DFSA if we disturb μ_h^H even slightly. As we shall see later, in fact, for any $N > 2$, the number of updating states required for the unconstrained optimum is of the order of $1/|\mu_h^H - \mu|$. Thus, when μ_h^H deviates from μ, the constraint K becomes binding in the sense that the unconstrained optimum is no longer implementable with K updating states.

What would the optimal SFSA constrained by K updating states be like? To answer this question, we appeal to Theorem 3.1 and look for potential indifference. When μ_h^H is close to μ, we only need to consider local deviations from $M^{s,N-1}$ for possible optimal randomization. However, different from the model of breakthroughs, there is no indifference at the limit case where $\mu_h^H = \mu$, which suggests that no randomization is needed for optimality for small perturbation over μ_h^H. To see this, Proposition 3.1 implies that $r(q_i)$ coincides with the posterior at q_i according to Bayes' rule when $\mu_h^H = \mu$. Hence, $\pi_{q_k} = \pi_0$ as $q^o = q_k$, and

$$\pi_{q_i} V_{q_i}(H) + (1 - \pi_{q_i}) V_{q_i}(L) > \pi_{q_i} V_{q_j}(H) + (1 - \pi_{q_i}) V_{q_j}(L) \text{ for all } j \neq i. \tag{29}$$

This follows immediately from the fact that $\pi_{q_i} V_{q_i}(H) + (1 - \pi_{q_i}) V_{q_i}(L)$ is the expected payoff of the unconstrained optimum when the prior is π_{q_i}, and $M^{s,N-1}$ with $q^o = q_i$ implements that optimal value but not $M^{s,N-1}$ with $q^o = q_j$ for any $j \neq i$. Since for each $i < N - 1$, $\pi_{q_i,h} = \pi_{q_{i+1}}$, (29) implies that it is strictly optimal to transit to q_{i+1} from q_i when receiving signal h according to (13). A symmetric argument holds for receiving signal ℓ.

Now we consider the extreme state, q_{N-1}. Upon receiving signal h, $M^{s,N-1}$ dictates a transition to q_H. This transition is also strictly optimal, since for all $j = 1, ..., N - 1$,

$$\pi_{q_{N-1},h} u^H \quad > \pi_{q_{N-1},h} V_{q_j}(H) + (1 - \pi_{q_{N-1},h}) V_{q_j}(L). \tag{30}$$

We note that (30) holds even at the boundary case when $\delta = \bar{\delta}_N^s$. Indeed, by (O2), $r(q_{N-1}, h) \geq r_{\bar{\pi}}$, and hence the left-side of (30) gives the optimal value under the unconstrained optimum for prior equal to $\pi_{q_{N-1},h}$, while the right-side is the expected payoff from $M^{s,N-1}$ with $q^o = q_j$ under that prior, which is suboptimal. In particular, at the boundary $\delta = \bar{\delta}_N^s$ and hence $r(q_{N-1}, h) = r_{\bar{\pi}}$, a DM with belief equal to $\pi_{q_{N-1},h} = \bar{\pi}$ is indifferent between taking u^H immediately and waiting. However, the crucial observation is that waiting here can be implemented only by $M^{s,N}$ with $q^o = q_N$ but not by $M^{s,N-1}$, as it takes N ℓ-signals to cross $\underline{\pi}$ instead of $N - 1$, as in $M^{s,N-1}$ with $q^o = q_{N-1}$. A symmetric argument also holds if at the boundary we have $r(q_1, \ell) = r_{\underline{\pi}}$.

Theorem 5.1. *Let π_0 be given and let $N > N_0$. Suppose that $|Q|$ is constrained by $K \geq N - 1$. Let k satisfy (27).*

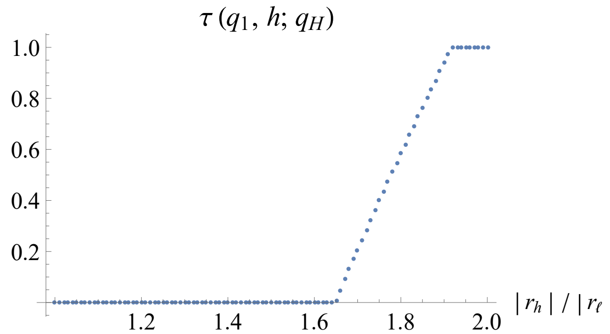

Figure 4: Optimal SFSA as $|r_h|/|r_\ell|$ increases

(1) Suppose that $\delta \in \Delta^s_N$. There exist $\underline{\mu} < \mu < \overline{\mu}$ such that for all $\mu^H_h \in [\underline{\mu}, \overline{\mu}]$, $M^{s,N-1}$ is the optimal SFSA with $q^o = q_k$.

(2) Suppose that $K = N - 1$ and $\mu^H_h = \mu$. Then, there exists $\epsilon > 0$ such that for all $\delta \in (\underline{\delta}^s_N, \bar{\delta}^s_N + \epsilon)$, $M^{s,N-1}$ is the optimal SFSA with $q^o = q_k$.

According to Theorem 5.1 (1), even when K is large compared to $N-1$, $M^{s,N-1}$ is still optimal whenever μ^H_h lies within a range around μ for a given $\delta \in \Delta^s_N$. Theorem 5.1 (2) extends this result to higher δ's, but in this case $K = N - 1$. Clearly, for higher δ's a bigger automaton will be useful. Thus, when the two signals have similar strengths, neither randomization nor bigger automata (up to a given size) are useful. Instead, it is optimal for the DM to simply cancel one h-signal by one ℓ and consider only the difference. However, how different μ^H_h can be from μ for the same rule to remain optimal depends on K and δ.

To illustrate the range in Theorem 5.1 (1), consider the following numerical example: $\pi_0 = 0.56$, $\mu^\ell = 0.7$, $u^H = u^L = 1$, and $\delta = 0.914 \in \Delta^s_3 = (0.890, 0.937)$. Theorem 5.1 (1) implies that $M^{s,2}$ is optimal with $q^o = q_2$, when holding r_ℓ constant, for a range of r_h's that are close to $|r_\ell|$. Now, holding r_ℓ fixed, we change $|r_h|/|r_\ell|$ from one to two. It turns out that there is only one transition where randomization can occur in the optimal SFSA, $\tau(q_1, h; q_H)$, whose optimal value is depicted in Figure 4, and other optimal transitions remain deterministic and are the same as those in $M^{s,2}$. The range for $M^{s,2}$ to be optimal is fairly large, including all information structures with $|r_h|/|r_\ell| \in [1, 1.65]$. Randomization only occurs for a small range, from $|r_h|/|r_\ell| = 1.65$ to 1.85, and the DM is treating one h-signal as cancelling out one ℓ-signal for lower ratios and treating one h-signal as cancelling out two ℓ-signals for higher ratios. Even for the range where randomization is optimal, it is randomizing between the two approximations.

Theorem 5.1 then shows that the same simple rule, $M^{s,N-1}$, is optimal for a range of parameters, and hence proves rule stickiness. It also demonstrates the optimality of *approx-*

imate probabilities: when $\mu_h^H \neq \mu$ but not too far from it, the optimal SFSA behaves *as if* μ_h^H and μ_ℓ^L are equal. This result provides a foundation for the use of *qualitative* probabilities, and the optimal rule corresponds to the *decisional balance sheet* attributed to Benjamin Franklin, according to which the decision maker should list "pros" and "cons" of different options and balance them with certain weights. This method is simple and optimal given the constraint, but it is biased: the corresponding beliefs π_{q_i}'s are biased in $M^{s,N-1}$ whenever $\mu_h^H \neq \mu$. However, unlike the model of breakthroughs, here the bias comes from the fact that one h-signal does not cancel one ℓ under Bayes rule, that is, $M^{s,N-1}$ employs a wrong cancellation rule whenever $\mu_h^H \neq \mu$. Moreover, Theorem 5.1 also illustrates the "less-is-more" effect, as $M^{s,N-1}$ performs better than *any* other SFSA with states no more than a given size for a range of information structures.

The main difficulty to prove Theorem 5.1 is to demonstrate that *any* deviations, especially those involving randomizations, from $M^{s,N-1}$ cannot improve ex ante payoff. Even for modest N, this potentially requires going through a long list by checking 2^N possible configurations (the set of updating states for which randomization can be optimal), each involving multi-dimensional optimization. We overcome this difficulty by appealing to Theorem 3.1, using the inequalities (29) and (30). Indeed, condition (13) requires the DM to be indifferent between the updating states that have positive probabilities to transit to. The strict inequalities in (29) and (30) ensure that the indifference condition (13) cannot be satisfied by local deviations. Finally, we also need to show that there are no other global deviations, which we do in the proof with an argument appealing to continuity and the uniqueness of the unconstrained optimal rule when $\mu_h^H = \mu$. This technique is also applicable to more general settings where the relative strengths other than one-for-one, a topic we turn to next.

5.2 General relative strengths

Here we generalize Theorem 5.1 to the case where $|r_h|/|r_\ell|$ is a general rational number, for example, two. The idea is that when the DM has a sufficiently large number of updating states, she can implement decision rules that are more complicated than cancelling one high signal by one low, but by using cancellations that are closer to the true relative strengths of the two signals. However, how close it can be will depend on the constraint, and will be approximated by the closest rational number that the DM can afford. First we characterize the unconstrained optimal rule when $|r_h|/|r_\ell|$ is a general rational number.

Consider an information structure where $|r_h|/|r_\ell| = M_h/M_\ell$, where M_h and M_ℓ are two mutually prime natural numbers, and, with no loss of generality, we assume that $M_h < M_\ell$.

For expositional purposes we only consider δ's such that

$$r_{\bar{\pi}} - r_{\underline{\pi}} > r_h - r_\ell. \tag{31}$$

This assumption avoids tedious discussions of different cases. Nevertheless, the main results still hold for lower values of δ.

As in the symmetric case, we compute the number of updating states needed in the optimal DFSA by tracing all possible posteriors generated by realizations of signals before reaching a decision. For any partial history $\mathbf{x}$, $r_{\mathbf{x}} = mr_h + nr_\ell$, where m is the number of h-signals and n is the number of ℓ-signals in $\mathbf{x}$. Now, if we take

$$r_c = \frac{|r_h|}{M_h} = \frac{|r_\ell|}{M_\ell}, \tag{32}$$

then, by Bezout's identity, $r_{\mathbf{x}}$ is always a multiple of r_c, and we may think of an imaginary signal with log likelihood ratio equal to r_c as a signal that serves as the "common divisor," so that one h-signal is worth of M_h "common" signals while one ℓ-signal is worth of $-M_\ell$ "common" signals. Now, let

$$K_h = \left\lceil \frac{r_{\bar{\pi}} - r_{\pi_0}}{r_c} \right\rceil \text{ and } K_\ell = \left\lceil \frac{r_{\pi_0} - r_{\underline{\pi}}}{r_c} \right\rceil. \tag{33}$$

Hence, from π_0 it takes K_h "common" signals to cross $\bar{\pi}$ and it takes K_ℓ "common" signals to cross $\underline{\pi}$. Thus, from the initial state q^o corresponding to π_0, we need $K_h - 1$ updating states to the right to cover all possible posteriors before q_H, and $K_\ell - 1$ updating states to the left. Then, the optimal DFSA requires $N-1$ updating states, where $N = K_h + K_\ell$: one initial state, $K_h - 1$ updating states to the right, and $K_\ell - 1$ updating states to the left. The corresponding DFSA, denoted by $M^{(M_h,M_\ell),N-1}$, is given by:

$$\begin{aligned} Q &= \{q_1, ..., q_{N-1}\}; \\ \tau(q_i, h) &= q_{i+M_h}, \text{ for } i = 1, ..., N-1-M_h;\ \tau(q_i, h) = q_H, \text{ for } i = N-M_h, ..., N-1; \\ \tau(q_i, \ell) &= q_{i-M_\ell}, \text{ for } i = M_\ell + 1, ..., N-1;\ \tau(q_i, \ell) = q_L, \text{ for } i = 1, ..., M_\ell. \end{aligned} \tag{34}$$

The initial state is $q^o = q_{K_\ell}$. The following lemma generalizes Lemma 5.1.

Lemma 5.2. *Suppose that $|r_h|/|r_\ell| = M_h/M_\ell$, M_h and M_ℓ mutually prime. For any δ satisfying (31), there exists $N = N(\delta)$ such that the unconstrained optimum can be implemented by the DFSA $M^{(M_h,M_\ell),N-1}$, but not with any SFSA with fewer than $N-1$ updating states. Moreover, $N(\delta)$ weakly increases with δ and goes to infinity as δ goes to one.*

As in the symmetric case, for each $N \geq 2$, define

$$\Delta_N^{(M_h,M_\ell)} = \text{int}\left(\{\delta \in (\delta_0, 1) : N(\delta) = N\}\right), \tag{35}$$

the interior of the set of δ's such that $N(\delta) = N$, where δ_0 is the threshold above which (31) holds. We only consider $N \geq N_0 \equiv \lim_{\delta \downarrow \delta_0} N(\delta)$. Note that $\Delta_N^{(M_h,M_\ell)}$ takes the form $(\underline{\delta}_N^{(M_h,M_\ell)}, \bar{\delta}_N^{(M_h,M_\ell)})$, and, generically in π_0, $\Delta_N^{(M_h,M_\ell)} \neq \emptyset$ for all $N \geq N_0$.

Now, for $\delta \in \Delta_N^{(M_h,M_\ell)}$, $M^{(M_h,M_\ell),N-1}$ satisfies (O1) and (O2) in Proposition 3.1, and hence the beliefs π_{q_i} are pinned down by Bayes' rule. This also implies that analogous conditions to (29) and (30) hold as well, following similar reasoning. Note that for $M^{(M_h,M_\ell),N-1}$, the extreme states include the rightmost M_h updating states and leftmost M_ℓ updating states. Nevertheless, a similar argument for (30) also holds for those extreme states. As a result, when we perturb μ_h^H so that $|r_h|/|r_\ell|$ is not exactly M_h/M_ℓ but close and hence $N-1$ updating states cannot implement the unconstrained optimum, we expect $M^{(M_h,M_\ell),N-1}$ to still be the optimal SFSA whenever the constraint $K \geq N-1$ is binding. The following theorem generalizes Theorem 5.1. Here, we perturb the parameter μ_h^H from its benchmark value μ^o, under which $|r_h|/|r_\ell| = M_h/M_\ell$.

Theorem 5.2. *Let π_0 and $\mu_h^H = \mu^o$, μ_ℓ^L be given such that $|r_h|/|r_\ell| = M_h/M_\ell$, M_h, M_ℓ mutually prime, and let $N \geq N_0$. Suppose that $|Q|$ is constrained by $K \geq N-1$. Let K_ℓ satisfy (33).*

(1) Suppose that $\delta \in \Delta_N^{(M_h,M_\ell)}$. There exist $\underline{\mu} < \mu^o < \overline{\mu}$ such that the optimal SFSA is $M^{(M_h,M_\ell),N-1}$ with $q^o = q_{K_\ell}$ for all $\mu_h^H \in [\underline{\mu}, \overline{\mu}]$.

(2) Suppose that $K = N-1$ and $\mu_h^H = \mu^o$. Then, there exists $\epsilon > 0$ such that for all $\delta \in (\underline{\delta}_N^{(M_h,M_\ell)}, \bar{\delta}_N^{(M_h,M_\ell)} + \epsilon)$, the optimal SFSA $M^{(M_h,M_\ell),N-1}$ with $q^o = q_{K_\ell}$.

According to (33), the size of the optimal SFSA depends on two features: first, the size of the grid determined by the product of M_h and M_ℓ, and the number of steps required from the prior to reach the thresholds for decision. Theorem 5.2 implies that the following heuristic is optimal: for a relatively simple rational number, M_h/M_ℓ, that is close to the ratio $|r_h|/|r_\ell|$ according to those two aspects, the heuristic would have M_ℓ h-signals cancel M_h ℓ-signals. The symmetric case is the special case with $M_h = M_\ell = 1$. This then extends the rule-stickiness result and the principle of probability approximation to the general case. Moreover, as in Theorem 5.1, Theorem 5.2 also implies the "less-is-more" effect, as the optimality is against a given size K which can be larger than $N-1$.

Here we remark that all these results can be generalized to more than two signals. Formally, consider a general finite X with $\mu_x^\theta > 0$ for all $x > 0$ and both $\theta = H, L$. If $|r_x|/|r_y|$

is a rational number for all $x, y \in X$, then we can find a DFSA that implements the unconstrained optimum and an analogous result to Theorem 5.2 holds as well, that is, rule stickiness still holds. This implies that signals of similar strengths will be treated as exactly the same if they change the posterior in the same direction, and exactly cancel one another if they change the posterior in opposite directions, and all this extends to other approximate rational proportions as well.

One specific implication is that "small" signals, in terms of informativeness, will be ignored. Again this follows from Theorem 3.1 and the fact that $r_{q_i} \in [\bar{r}_{i-1}, \bar{r}_i]$, which implies that, for a transition to occur from q_i when receiving a signal x, r_x has to be sufficiently large in absolute values to cross the threshold $\bar{r}_{i-1}$ or $\bar{r}_i$. This suggests that, in contrast to the unconstrained optimal rule, the optimal SFSA under constraint $|Q| \leq K$ will not respond to small signals with likelihood ratios close to one, as the following corollary shows.

Corollary 5.1. *Fix an information structure, $(\bar{\mu}_x^H, \bar{\mu}_x^L)_{x \in X}$, satisfying (3) but $r_y = 0$ for some fixed $y \in X$ (that is, $\bar{\mu}_y^H = \bar{\mu}_y^L$). Suppose that in all optimal SFSA under the constraint $|Q| \leq K$ and under $(\bar{\mu}_x^H, \bar{\mu}_x^L)_{x \in X}$, $\bar{r}_{i-1} < r(q_i) < \bar{r}_i$ for all updating state q_i. Then, there exists some $\epsilon > 0$ such that for all information structures perturbing the original one with $\max_{x \in X, \theta = H, L} |\mu_x^\theta - \bar{\mu}_x^\theta| \leq \epsilon$ (and hence r_y in the perturbed structure is close to zero), the optimal SFSA ignores y, i.e., $\tau(q, y; q) = 1$ for all $q \in Q$.*

The condition $\bar{r}_{i-1} < r(q_i) < \bar{r}_i$ is satisfied in all the optimal DFSA we have identified in this section, and hence if we add small signals in all such environments, they will be ignored optimally. In fact, it is also satisfied by $M^{b,K}(\alpha)$ when it is the optimal SFSA in the model of breakthrough. Corollary 5.1 can then be interpreted as both an information-stickiness result and a rule-stickiness result: the DM treats the signal y as if in a simple world where y contains no information at all and hence does not respond to it, even though in reality it is informative.

6 Concluding remarks

We set out to evaluate the Bayesian paradigm by imposing a limited ability to process information in a problem that features finite learning. We identified two *simple* environments where Bayesian learning is feasible. First, when the information structure features very asymmetric strengths between the two signals. Second, when the strengths of signals relate to one another in rational proportions. Generically, however, we encounter *complex* environments where simple rules according to finite automata cannot implement Bayesian learning.

We also identified two prominent features of constrained optimal rules. First, in the model of breakthroughs, randomization is optimal whenever the memory constraint binds and the DM optimally ignores informative signals from time to time, an information-stickiness result. Second, in the more general case, it is optimal to use approximate probabilities without randomization, and a simple rule that cancels signals according to their approximate relative strengths is optimal for a range of parameters, a rule-stickiness result that justifies the use of heuristics, or rules-of-thumb. In both cases, the DM's beliefs are biased whenever the information structures depart from simple environments where the unconstrained optimum can be implemented, despite their optimality. Thus, heuristic rules can be both optimal and biased for boundedly rational agents.

Finally, our rule-stickiness results imply that bounded rational agents use simple models to derive optimal solutions, even though the reality (that is, the actual environment) can be far more complex. This echoes Simon (1959)'s discussion of economic models in relation to human cognition; he argues that "In actual fact the perceived world is fantastically different from the 'real' world. The difference involves both omissions and distortions, and arise in both perception and inference [...] The decision-maker's model of the world is only a minute fraction of all the relevant characteristics of the real environment." We formalized these concepts in a model with limited capacity for information processing. This formalization would allow for future investigation beyond individual decision-making. Indeed, although the biases we identified are features of constrained optimal rules, an agent with memory constraints is still vulnerable to exploitation from a sophisticated principal. Our results have identified novel sources of apparently behavioral biases in the economic agents' belief formation, and their full implications for wider economic interactions remain to be explored.

References

[1] Barlo, M., G. Carmona and H. Sabourian, (2009). Repeated Games with One-Memory. *Journal of Economic Theory,* 144, 312-336.

[2] Ben-Porath, E., (1993). "Repeated Games with Finite Automata." *Journal of Economic Theory*, 59, 17-32.

[3] Banovetz, J. and R. Oprea, (2020). "Complexity and Procedural Choice." *Working paper.*

[4] Borel, E., (1909). "Les probabilities denombrables et leurs applications arithmetiques." *Rendiconti del Circolo Matematico di Palermo*, 27: 247–271, doi:10.1007/BF03019651

[5] Che, Y.K. and K. Mierendorff, (2019). "Optimal Dynamic Allocation of Attention", *American Economic Review,* 109, 2993-3029.

[6] Compte, O. and A. Postlewaite, (2012). "Belief Formation-Second Version", PIER Working Paper 12 027, University of Pennsylvania.

[7] Hellman, M. E. and T. M. Cover, (1970). "Learning with Finite Memory." *The Annals of Mathematical Statistics,* 41, 765–782.

[8] Hellman, M. E., and T. M. Cover, (1971). "On memory saved by randomization." *The Annals of Mathematical Statistics*, 42, 1075-1078.

[9] Jehiel, P. and J. Steiner, (2019). "Selective Sampling with Information Storage Constraints," *Economic Journal,* https://doi.org/10.1093/ej/uez068 (December).

[10] Kalai, E., and E. Solan, (2003). "Randomization and Simplification in Dynamic Decision-making." *Journal of Economic Theory,* 111, 251-264.

[11] Kocer, Y., (2010). "Optimal Plans with Bounded Memory are Modified multi-self Consistent," Working paper.

[12] Kuhn, H. W., (1953). "Extensive Games and the Problem of Information." in *Contributions to the Theory of Games II*, Princeton University Press, Princeton, NJ.

[13] Nerode, A. (1958). "Linear Automaton Transformations," *Proceedings of the AMS*, 9.

[14] Oprea, R. (2020). "What Makes a Rule Complex?" *American Economic Review*, 110, 3913–3951.

[15] Parpart, P., M. Jones and B. C. Love, (2018). "Heuristics as Bayesian Inference under Extreme Priors," *Cognitive Psychology,* 102, 127-144.

[16] Piccione, M. and A. Rubinstein, (1997). "On the Interpretation of Decision Problems with Imperfect Recall", *Games and Economic Behavior* 20, 3-24.

[17] Salant, Y. (2011). "Procedural Analysis of Choice Rules with Applications to Bounded Rationality," *American Economic Review,* 101, 724-748.

[18] Simon, H. A. (1956). "Rational Choice and the Structure of the Environment." *Psychological Review,* 63, 129-138.

[19] Simon, H. A. (1959). "Theories of Decision-Making in Economics and Behavioral Science." *American Economic Review*, 49, 253-283.

[20] Topkis, M. D. (1998). *Supermodularity and Complementarity,* Princeton University Press.

[21] Tversky, A. and D. Kahneman, (1974). "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124-1131.

[22] Wilson, A. (2014). "Bounded Memory and Biases in Information Processing," *Econometrica*, 82, 2257-2294.

[23] Wald, A. (1947). *Sequential Analysis.* Wiley, New York.

Appendix: Proofs

Here we present all the proofs, except for two proofs which use techniques similar to the literature (Proposition 2.1 and Theorem 3.1) and one that is very technical in nature (Theorem 3.2), which can be found in the Online Appendix.

Proof of Proposition 3.1

We prove (17) here. Given π_0, we can rewrite the likelihood ration $\pi_q/(1-\pi_q)$ as

$$\frac{\pi_q}{1-\pi_q} = \frac{\pi_0 \mathbb{P}(q|H)}{(1-\pi_0)\mathbb{P}(q|L)}.$$

Now, consider the term in (8) that begins with δ^n for some $n \geq 0$. Let $\mathbf{x}$ be such that $\lambda_M(\mathbf{x}) = q$ with $|\mathbf{x}| = n$. Then, that nth term is equal to the sum of $\delta^n \prod_{i=1}^n \mu_{x_i}^\theta$ from all such $\mathbf{x}$'s. Moreover, (O1) implies that

$$\frac{\pi_0 \prod_{i=1}^n \mu_{x_i}^H}{(1-\pi_0) \prod_{i=1}^n \mu_{x_i}^L}$$

remains constant for all such $\mathbf{x}$ and for all n. Thus, we have

$$e^{r(q)} = \frac{\pi_0 \mathbb{P}(q|H)}{(1-\pi_0)\mathbb{P}(q|L)} \quad = \quad \frac{\pi_0 \prod_{i=1}^n \mu_{x_i}^H}{(1-\pi_0) \prod_{i=1}^n \mu_{x_i}^L},$$

for any such $\mathbf{x}$, and this implies (17). □

Proof of Lemma 4.1

First, from (20) we can derive $\bar{\pi}(\delta)$:

$$\frac{\bar{\pi}}{1-\bar{\pi}} = \left(\frac{\delta\mu}{1-\delta}\right)\frac{u^L}{u^H} = \left(\frac{\delta\mu}{1-\delta}\right)\frac{\pi^o}{1-\pi^o}. \tag{36}$$

Thus, $\delta \geq \delta^o$, the threshold below which taking an action immediately is optimal, if and only if $\bar{\pi} \geq \pi^o$, which is equivalent to $\delta \geq 1/(1+\mu) \equiv \delta^o$. Now we compute the sequence $\{\delta^b_N\}^{\infty}_{N=0}$ for the given prior π_0. Let $\bar{\pi}_0 = \pi_0$, and, for each $N \geq 1$, define $\bar{\pi}_N$ as the unique solution to

$$\frac{\bar{\pi}_N}{1-\bar{\pi}_N} = \left(\frac{\pi_0}{1-\pi_0}\right)\left(\frac{1}{1-\mu}\right)^N, \tag{37}$$

that is, $r_{\bar{\pi}_N} = r_{\pi_0} + Nr_h$—it requires exactly N h-signals to reach $\bar{\pi}_N$ from π_0. Note that $\bar{\pi}_N$ strictly increases with N and is always within $(\pi_0, 1)$. By (36), $\bar{\pi}(\delta)$ strictly increases with δ and converges to one as δ does. Thus, for each $N \geq 0$, there is a unique δ^b_N such that $\bar{\pi}(\delta^b_N) = \bar{\pi}_N$. Then, for $\delta \leq \delta^o$, $\bar{\pi} \leq \pi^o \leq \pi_0$, and hence the optimal rule takes a^H immediately. For $\delta \in (\delta^b_{N-1}, \delta^b_N)$, it takes exactly N h-signals to cross $\bar{\pi}$, and it is straightforward to verify that $M^{b,N}$ implements the optimal rule. The result that $M^{b,n}$ for $n < N$ does not implement the optimal rule follows from the proof of Theorem 4.1 below. $\square$

Proof of Theorem 4.1

Part (1) follows immediately from Lemma 4.1; we prove part (2). Suppose that $\delta > \delta^b_K$. As argued, it is optimal to take a^L immediately after an ℓ-signal. We first show that any optimal SFSA takes the form of $M^{b,K}(\alpha_1, ..., \alpha_K)$, and we may assume that there are no redundant states. By Theorem 3.1, we can rank the updating states $q_1, ..., q_K$ so that $r(q_1) \leq r(q_2) \leq .... \leq r(q_K)$. Since receiving h can only trigger a transition to go up or to stay, $V_{q_i}(\theta)$ only depends on $V_{q_j}(\theta)$ for $j > i$, while $r(q_i)$ only depends on transitions from q_j with $j < i$. Thus, we can take $r(q_i)$ as the prior and look for the optimal SFSA with $K - i$ updating states, and use the necessary conditions in Theorem 3.1 to characterize the optimal SFSA.

This allows for an induction argument starting from $i = K$ and then working backwards. We first show by induction that the optimal SFSA the transitions from $q_i, ..., q_K$ take the form $M^{b,K-i+1}(\alpha_i, ..., \alpha_K)$ (part (a)), and then we show that the optimal SFSA has $\alpha_1 = \alpha_2 = ... = \alpha_K = \alpha \in (0, 1)$ (part (b)).

(a) For $i = K$ this is immediate, as the only possible transition is either staying in q_K or going to q_H after seeing an h-signal. Suppose that it holds for $i + 1 \leq K$, and in the optimal

SFSA the transitions from $q_{i+1}, ..., q_K$ take the form $M^{b,K-i}(\alpha_{i+1}, ..., \alpha_K)$. Below in (b) we prove that it is optimal to have $\alpha_{i+1} = ... = \alpha_K$. This, by a simple computation from the corresponding value functions, implies that for all $j = i+1, ..., K-1$, $\bar{r}_{j+1} - \bar{r}_j$ is constant and

$$0 < \bar{r}_{j+1} - \bar{r}_j \leq r_h, \text{ with equality iff } \alpha_{i+1} = 0 = ... = \alpha_K, \tag{38}$$

that is, $\bar{r}_{j+1} - \bar{r}_j = r_h$ iff it is the DFSA $M^{b,K-i}$. Moreover, we have $r(q_{i+1}) < \bar{r}_{i+1}$: otherwise, $r(q_{i+1}) = \bar{r}_{i+1}$ and hence $r(q_{i+1}, h) = \bar{r}_{i+1} + r_h > \bar{r}_{i+1}$, that is, $\alpha_{i+1} = 0$ and, by (38), this implies that the transitions follow the DFSA $M^{b,K-i}$ and, furthermore, $r(q_{i+1}, h) = \bar{r}_{i+1} + r_h = \bar{r}_{i+2}$. This last equality also implies that from q_{i+1} following signal h there is indifference between moving to q_{i+2} and q_{i+3}, implying that there is a redundant updating state. This concludes that $r(q_{i+1}) < \bar{r}_{i+1}$.

Now, $r(q_{i+1}) < \bar{r}_{i+1}$ implies, by (38), that $r(q_{i+1}) + r_h < \bar{r}_{i+2}$. Since $r(q_i) \leq r(q_{i+1})$ and $r(q_i, h) = r(q_i) + r_h \leq r(q_{i+1}) + r_h < \bar{r}_{i+2}$, by Theorem 3.1, from q_i and an h-signal, randomization can only occur between q_i and q_{i+1}, or between q_{i+1} and q_{i+2}, but not both. Now we show that if the latter happens, then we can eliminate q_{i+1} without affecting the ex ante expected payoff and there is a redundant state; as a result, only the former matters and hence the optimal SFSA takes the form $M^{b,K-i+1}(\alpha_i, ..., \alpha_K)$ from q_i on. To see this, suppose, by contradiction, that the optimal SFSA randomizes between q_{i+1} and q_{i+2} from q_i after h. This implies that $r(q_i) < r(q_{i+1})$, and that, by Theorem 3.1, $r(q_i, h) = \bar{r}_{i+1}$, which in turn implies that $r(q_{i+1}, h) > r(q_i, h) = \bar{r}_{i+1}$. By Theorem 3.1 the last inequality implies that $\alpha_{i+1} = 0$, and by the symmetry noted above, in $M^{b,K-i}(\alpha_{i+1}, ..., \alpha_K)$ we have $\alpha_{i+1} = ... = \alpha_K = 0$. That is, we have a deterministic scheme from q_{i+1} on. Moreover, this also implies that at prior $r(q_i)$, $M^{b,K-i}$ and $M^{b,K-i-1}$ give exactly the same payoff, and hence q_{i+1} is a redundant state.

(b) Now we show that the optimal $M^{b,K}(\alpha_1, ..., \alpha_K)$ features $\alpha_1 = ... = \alpha_K = \alpha$ for any prior, and whenever $\delta > \delta_K^b$, $\alpha \in (0, 1)$. To do so, we first compute the continuation values in $M^{b,K}(\alpha_1, ..., \alpha_K)$:

$$\begin{aligned} V_{q_1}(H) &= \frac{\delta^{K-2}\left[\prod_{j=1}^{K-2}(1-\alpha_j)\right] u^H}{\prod_{j=1}^{K-2}(1-\delta\alpha_j)}, \\ V_{q_1}(L) &= \delta\mu u^L \left\{\sum_j A_j + \sum_{i<j} C A_i A_j + \sum_{i<j<k} C^2 A_i A_j A_k + ... + C^{K-3} A_1 \cdots A_{K-2}\right\}, \end{aligned}$$

where $A_j = 1/[1-\delta(1-\mu)\alpha_j]$ for each $j = 1, ..., K-2$ and $C = \delta(1-\mu) - 1$. Now, since both $V_{q_1}(H)$ and $V_{q_1}(L)$ are symmetric in $(\alpha_1, ..., \alpha_K)$ and strictly supermodular, for any prior the optimum must happen under a symmetric solution; indeed, by symmetry, any permutation of the optimal solution is also optimal, but by Theorem 2.7.5 in Topkis (1998), the optimal solutions form a chain and hence the optimal solution cannot be asymmetric. Hence, optimal $\alpha_1 = \cdots = \alpha_K = \alpha$. Under this symmetric solution, the expected payoff from $M^{b,K}(\alpha)$ for a given π_0 is

$$\begin{aligned} F_K(\alpha) &\equiv \pi_0 V_{q_1}(H) + (1-\pi_0)V_{q_1}(L) \\ &= \pi_0 \frac{\delta^K(1-\alpha)^K u^H}{(1-\delta\alpha)^K} + (1-\pi_0)\frac{\delta\mu u^L}{1-\delta(1-\mu)}\left\{1-\left[\frac{\delta(1-\mu)(1-\alpha)}{1-\delta(1-\mu)\alpha}\right]^K\right\}. \end{aligned} \tag{39}$$

It is straightforward to verify that $\max_{\alpha\in[0,1]} F_K(\alpha)$ has a unique maximizer; indeed,

$$\begin{aligned} F_K'(\alpha) &= \pi_0 K \frac{(-\delta^K)(1-\delta)(1-\alpha)^{K-1}u^H}{(1-\delta\alpha)^{K-1}} \\ &+ (1-\pi_0)K\left[\frac{\delta(1-\mu)(1-\alpha)}{1-\delta(1-\mu)\alpha}\right]^{K-1}\frac{\delta^2\mu(1-\mu)u^L}{[1-\delta(1-\mu)\alpha]^2}, \end{aligned}$$

and it has at most one zero between $[0,1]$. Since $\pi_0 \geq \pi^o$, $F_K(1) < (1-\pi_0)u^L \leq \pi_0 u^H < F_K(0)$. Moreover, $F_K'(0) > 0$ if and only if

$$(1-\pi_0)\delta\mu(1-\mu)^K u^L > \pi_0(1-\delta)u^H, \tag{40}$$

which, by (36) and (37), is equivalent to $\delta > \delta_K^b$. This proves that optimal $\alpha > 0$. Clearly, $\alpha = 1$ is never optimal. Note that this also implies that we need all K updating states. □

Proof of Theorem 4.2

Note that we assume $\mu_h^H \geq \delta$ but can be strictly smaller than 1, and we use the notation $\mu_\ell^L = \mu$. Consider $M^{b,1}(\alpha)$. Given π_0, the expected payoff is (note that the values differ from (39) only because $\mu_h^H < 1$ here)

$$F(\alpha;\mu_h^H,\delta) = \pi_0\frac{\delta(1-\alpha)\mu_h^H u^H}{1-\delta\mu_h^H\alpha} + (1-\pi_0)\frac{\delta\mu u^L}{1-\delta(1-\mu)\alpha}. \tag{41}$$

Now, take $\alpha = 1-\sqrt{1-\delta}$. We claim that

$$\lim_{\delta\to 1} F(\alpha;\mu_h^H,\delta) = \pi_0 u^H + (1-\pi_0)u^L,$$

uniformly across all $\mu_h^H \geq \delta$. This proves the result by taking appropriate $\bar{\delta}$.

Now we prove the claim. First, since $\alpha = 1 - \sqrt{1-\delta}$, $\lim_{\delta \to 1} \alpha = 1$,

$$\lim_{\delta \to 1} \frac{\delta\mu}{1-\delta(1-\mu)\alpha} = \lim_{\delta \to 1} \frac{\delta\mu}{1-\delta\alpha+\delta\mu\alpha} = 1.$$

Thus, by (41), we only need to show that

$$\lim_{\delta \to 1} \frac{\delta(1-\alpha)\mu_h^H}{1-\delta\mu_h^H\alpha} = 1.$$

Since $\alpha = 1 - \sqrt{1-\delta}$, this is equivalent to

$$\lim_{\delta \to 1} \frac{\delta\sqrt{1-\delta}\mu_h^H}{1-\delta\mu_h^H[1-\sqrt{1-\delta}]} = 1.$$

Now,

$$\frac{\delta\sqrt{1-\delta}\mu_h^H}{1-\delta\mu_h^H[1-\sqrt{1-\delta}]} - 1 = \frac{\delta\mu_h^H}{\frac{1-\delta\mu_h^H}{\sqrt{1-\delta}} + \delta\mu_h^H} - 1 = -\frac{\frac{1-\delta\mu_h^H}{\sqrt{1-\delta}}}{\frac{1-\delta\mu_h^H}{\sqrt{1-\delta}} + \delta\mu_h^H}.$$

Now, we show that if $\mu_h^H \geq \delta$, then $(1-\delta\mu_h^H)/\sqrt{1-\delta}$ converges to zero uniformly as δ converges to one. To see this, since $\mu_h^H \geq \delta$, we have

$$\frac{1-\delta\mu_h^H}{\sqrt{1-\delta}} \leq \frac{1-\delta^2}{\sqrt{1-\delta}} = \sqrt{1-\delta}(1+\delta),$$

which converges to zero across all $\mu_h^H \geq \delta$ as δ converges to one. Note that we can consider $\inf_{\mu_h^H \geq \delta} \mu_h^H$ as a function of δ. This converges to 1 as $\delta \to 1$. Formally, $\lim_{\delta \to 1} \inf_{\mu_h^H \geq \delta} \mu_h^H = 1$.
□

Before the proof of Lemma 5.1, we present a result from the computer science literature that gives a characterization of the set of *all* decision rules that can be implemented by a DFSA with $|Q| \leq K$. This results is essentially the celebrated Myhill-Nerode Theorem (Nerode (1958)), but we need some adjustment to fit our purposes. First define an equivalence relation, $R \subset X^* \times X^*$, to be *right-invariant* if for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$,

$$\mathbf{x}R\mathbf{y} \iff (\mathbf{x} \circ \mathbf{z})R(\mathbf{y} \circ \mathbf{z}),$$

where $\circ$ denotes concatenation. Given a decision rule f, $\mathcal{L}_\theta^f$ denotes the set of partial histories under which action a^θ is taken, $\theta = H, L$. The following is a simple extension of the standard Myhill-Nerode Theorem. A DFSA M is non-redundant if each updating and each action state can be reached by some partial history of signal realizations.

Theorem 6.1 (Myhill-Nerode Theorem)**.** *The rule f can be implementable by a non-redundant DFSA with K updating states iff there is a right-invariant equivalence relation R that induces $K+2$ equivalence classes such that $\mathcal{L}_\theta^f$ is one of those equivalence classes for both $\theta = H, L$.*

We give the proof for self-containment in the Online Appendix. The equivalence classes correspond to the updating and the action states in the corresponding DFSA, and the equivalence classes that make up $\mathcal{L}_\theta^f$ consist of those action states where action a^θ is taken, for both $\theta = H, L$. Thus, the DFSA gives a finite partition of partial histories that captures the finiteness of the DM's memory capacity. The right-invariance condition captures the fact that if the DFSA enters the same updating state after two different partial histories, then it will end up in the same updating or action state (although not necessarily the same as the original one) after any consecutive partial history. The number of equivalence classes generated in this way has a one-to-one correspondence to the number of updating and updating states in the DFSA.

Proof for Lemma 5.1

Let $\delta > \delta_0$ be given, and let $N = N(\delta)$ be given by (26). To show that $M^{s,N-1}$ with $q^o = q_k$ with k satisfying (27) implements the unconstrained optimum, by Proposition 3.1, we only need to show (O1) and (O2). To show (O1), note that $\lambda_{M^{s,N-1}}(\mathbf{x}) = q_i$, $1 \leq i \leq N-1$, if and only if there are $(k-i)$ more ℓ-signals than h for $i < k$, and if and only if there are $(i-k)$ more h-signals than ℓ for $i \geq k$. (O1) then follows from $r_h + r_\ell = 0$. Similarly, (O2) follows from (27).

Now we show that any DFSA with less than $N-1$ updating states cannot implement the unconstrained optimum. By Theorem 6.1, we only need to show that for any right-invariant equivalence relation such that, taking f as the optimal rule, $\mathcal{L}_H^f$ and $\mathcal{L}_L^f$ are among its equivalence classes, it has at least $N+1$ equivalence classes. To do so, consider the histories $\mathbf{x}^0 = \emptyset$, $\mathbf{x}^i$ consisting of i h-signals for $i = 1, ..., N-k$, and $\mathbf{y}^j$ consisting of j ℓ-signals for $j = 1, ..., k$. For any two histories from the list, say, $\mathbf{x}^i$ and $\mathbf{x}^{i'}$ with $i > i'$, we can use the partial history $\mathbf{x}^{i,i'}$ consisting of $N-i$ h-signals as the witness in the sense that $\mathbf{x}^i \circ \mathbf{x}^{i,i'}$ leads to action a^H while $\mathbf{x}^{i'} \circ \mathbf{x}^{i,i'}$ leads to waiting under the optimal rule. For any other combinations we can use similar witness histories. This shows that each has to be in a distinct equivalence class. Finally, this also shows that $M^{s,N-1}$ with $q^o = q_k$ is the *unique* DFSA that implements the unconstrained optimum among all DFSA with $N-1$ updating states, as each of the histories from $\{\mathbf{x}^0, \mathbf{x}^1, ..., \mathbf{x}^{N-k}, \mathbf{y}^1, ..., \mathbf{y}^k\}$ corresponding to one (updating or action) state in any optimal DFSA. $\square$

Proof of Theorem 5.1

Note that by Lemma 5.1 and Proposition 3.1, when $\mu_h^H = \mu$ and when $\delta \in \Delta_N^s$,

$$r(q_i) = r_{\pi_0} + (i-k)r_h,$$

for all $i = 1, ..., N-1$. Since the proof of Lemma 5.1 also showed that $M^{s,N-1}$ is the unique DFSA with no more than $N-1$ updating states, (29) and (30) follow immediately. Note that they also hold for $\delta = \bar{\delta}_N^s$.

Proof of (1) Let $\delta \in \Delta_N^s$ and $K \geq N-1$ be given. We have shown that for the given K, $M^{s,N-1}$ (with $q^o = q_k$ that satisfies (27)) achieves the unconstrained optimum when $\mu_h^H = \mu$ and the optimal rule is unique by Proposition 2.1. Thus, for any SFSA M with no more than K updating states, the expected payoff from M, denoted by W_M, is no greater than $W_{M^{s,N-1}}$, and it is equal if and only if it follows the same transition as in $M^{s,N-1}$ except for adding "replicate" updating states as the unconstrained optimal rule is unique. By replicate state we mean for any q_i in $M^{s,N-1}$, we can add q_i^c to it so that $\tau'(q_i^c, x; q') = \tau(q_i, x; q')$ or to the replicate of q' for all q', and $\tau'(q', x; q_i^c) > 0$ if and only if $\tau(q', x; q_i^c) > 0$. A small extension of the proof of Proposition 3.1 shows that the beliefs associated with the replicate state remains the same as the corresponding states in $M^{s,N-1}$, and hence (29) and (30) also hold under the alternative SFSA with such "replicate" states. Since the continuation values and the beliefs $r(q_i)$ are continuous in both the transition probabilities and parameter values, there exist $\epsilon_1 > 0$ such that the inequalities (29) and (30) are maintained strictly for any SFSA M with transitions different from those in $M^{s,N-1}$ or its replica by no more than ϵ_1 and any $\mu_h^H \in [\mu - \epsilon_1, \mu + \epsilon_1]$. Now, any such SFSA other than $M^{s,N-1}$ or its replica has strict randomization somewhere other than the replicate states, and hence its optimality is excluded by Theorem 3.1 due to the strict inequalities, (29) and (30). Finally, for such ϵ_1, consider the set of SFSA M such that some of the transitions in M differ from $M^{s,N-1}$ or its replicate by at least ϵ_1, a set we denoted by D_{ϵ_1}. This set is compact, and, when $\mu_h^H = \mu$, $W_{M^{s,N-1}} > \max_{M \in D_{\epsilon_1}} W_M$. Now, by continuity again, there exists $\epsilon_2 \in (0, \epsilon_1]$ such that for any $\mu_h^H \in [\mu - \epsilon_2, \mu + \epsilon_2]$, the inequality still holds.

Proof of (2) The proof follows exactly the same kind of argument as (1), except for taking deviations upwards from $\bar{\delta}_N^s$. Note that $K = N-1$ guarantees that $M^{s,N-1}$ is unique optimal DFSA, even at $\delta = \bar{\delta}_N^s$. $\square$

Proof of Lemma 5.2

Let $\delta > \delta_0$ be given and hence (31) is satisfied, and let $N = N(\delta) = K_h + K_\ell$ with (K_h, K_ℓ) given by (33). To show that $M^{(M_h,M_\ell),N-1}$ with $q^o = q_{K_\ell}$ implements the unconstrained optimum, by Proposition 3.1, we only need to show (O1) and (O2). To show (O1), note that $\lambda_{M^{s,N-1}}(\mathbf{x}) = q_i$, $1 \le i \le N-1$, if and only if there are $(k-i)$ less "common" signals for $i < k$, and if and only if there are $(i-k)$ more common signals for $i \ge k$, where one common signal stands for log likelihood ratio of $r_c = r_h/M_h = -r_\ell/M_\ell$ and hence one h-signal counts for M_h common signals and one ℓ counts for $-M_\ell$ common signals. Similarly, (O2) follows from (27). Now we show that any DFSA with fewer than $N-1$ updating states cannot implement the unconstrained optimum. By Theorem 6.1, we only need to show that for any right-invariant equivalence relation such that, taking f as the optimal rule, $\mathcal{L}_H^f$ and $\mathcal{L}_L^f$ are among its equivalence classes, it has at least $N+1$ equivalence classes. This then follows from the following two claims:

Claim (1) For each $i = 1, ..., N-1$, there exists a partial history $\mathbf{x}^i$ such that

$$r_{\pi_0} + r_{\mathbf{x}^i} = r(q_i), \tag{42}$$

and such that for any $\mathbf{y} \subset \mathbf{x}^i$, $r_{\pi_0} + r_{\mathbf{y}} \in (r_{\underline{\pi}}, r_{\bar{\pi}})$. Let $\mathbf{x}^0$ be the one where action a^L is taken and $\mathbf{x}^N$ be the one where action a^H is taken.

Claim (2) Each $\mathbf{x}^i$ must be in a different equivalence class.

Proof of (1) First, note that for any partial history $\mathbf{x}$, $r_{\mathbf{x}}$ has the form

$$r_{\mathbf{x}} = (n_h M_h - n_\ell M_\ell) r_c. \tag{43}$$

Now, we give Claim (1a): for any (n_h, n_ℓ) such that

$$-K_\ell < n_h M_h - n_\ell M_\ell < K_h, \tag{44}$$

there exists a partial history $\mathbf{x}$ such that (43) holds and that for any $\mathbf{y} \subset \mathbf{x}$, $r_{\pi_0} + r_{\mathbf{y}} \in (r_{\underline{\pi}}, r_{\bar{\pi}})$. Clearly, $\mathbf{x}$ has n_h h-signals and n_ℓ ℓ-signals. We construct the order backwards to ensure the sub-histories never reach the boundaries. In particular, we construct a sequence $\{(n_h^t, n_\ell^t)\}$ inductively as follows. First, $n_\ell^1 \le n_\ell$ and n_h^1 are determined by

$$\begin{aligned} n_h M_h - n_\ell^1 M_\ell &< K_h \le n_h M_h - (n_\ell^1 - 1) M_\ell, \\ n_h^1 M_h - n_\ell^1 M_\ell &> -K_\ell \ge (n_h^1 - 1) M_h - n_\ell^1 M_\ell. \end{aligned}$$

Given (n_h^t, n_ℓ^t), $(n_h^{t+1}, n_\ell^{t+1})$ are determined by

$$\begin{aligned} n_h^t M_h - n_\ell^{t+1} M_\ell &< K_h \leq n_h^t M_h - (n_\ell^{t+1} - 1) M_\ell, \text{ and} && (45) \\ n_h^{t+1} M_h - n_\ell^{t+1} M_\ell &> -K_\ell \geq (n_h^{t+1} - 1) M_h - n_\ell^{t+1} M_\ell. && (46) \end{aligned}$$

We claim that $n_\ell^{t+1} < n_\ell^t$ and $n_h^{t+1} < n_h^t$. This follows because

$$M_\ell n_\ell^{t+1} \leq n_h^t M_h + M_\ell - K_h \leq M_h + M_\ell + M_\ell n_\ell^t - K_h - K_\ell < M_\ell n_\ell^t,$$

where the first inequality follows from (45), the second from (46) but with t instead of $t+1$, and the last from (31); similarly,

$$M_h n_h^{t+1} \leq n_\ell^{t+1} M_\ell - K_\ell \leq M_h + M_\ell + M_h n_h^t - K_h - K_\ell < M_h n_h^t,$$

where the first inequality follows from (46), the second from (45), and the last from (31). Since n_h and n_ℓ are both finite, the sequence ends with zero. This proves Claim (1a).

To construct $\mathbf{x}^i$, we give Claim (1b): for each i, we can find $(m_h^i, m_\ell^i) \in \mathbb{N}_+^2$ such that

$$r_{\pi_0} + (m_h^i M_h - m_\ell^i M_\ell) r_c = r(q_i). \tag{47}$$

Given this, we can then use Claim (1a) to construct $\mathbf{x}^i$ to satisfy (42). Note that by (O1) any solution to (47) satisfies (44) with $(n_h, n_\ell) = (m_h^i, m_\ell^i)$. Since M_h and M_ℓ are mutually prime, by Bezout's identity, there are integers n_h^* and n_ℓ^* such that $n_h^* M_h + n_\ell^* M_\ell = 1$ and hence for any $k \in \mathbb{Z}$,

$$m_h^i = (n_h^* + kM_\ell)\frac{r(q_i) - r_{\pi_0}}{r_c} \text{ and } m_\ell^i = (n_\ell^* - kM_h)\frac{r(q_i) - r_{\pi_0}}{r_c}$$

solves (47). Note that $r(q_i) - r_{\pi_0}$ is a multiple of r_c. To ensure that the solution is positive, we only need to find appropriate k.

Proof of (2) Let $\mathbf{x}^i$ and $\mathbf{x}^j$ be constructed from (1), with $i < j$. Now, following the same arguments as in (1), by taking $\pi_0 = \pi_{q_{i+1}}$ as the prior, we can construct a partial history $\mathbf{y}$ such that $r_{\pi_{q_{i+1}}} + r_{\mathbf{y}} = r_{\pi_{q_1}}$ but never reaches $\bar{\pi}$ or $\underline{\pi}$ along the way, and hence $r_{\pi_{q_i}} + r_{\mathbf{y}} \leq r_{\underline{\pi}}$. By construction, if we start from π_0, we will end up with π_{q_i} following $\mathbf{x}^i$, and hence starting from π_0 following $\mathbf{x}^i \circ \mathbf{y}$ the posterior never reaches $\bar{\pi}$, but for some $\mathbf{y}' \subset \mathbf{y}$, $\mathbf{x}^i \circ \mathbf{y}'$ crosses $\underline{\pi}$ for the first time. Since $j > i$, following $\mathbf{x}^j \circ \mathbf{y}$ the posterior never reaches $\underline{\pi}$, but it may reaches $\bar{\pi}$ before the end. If it does, let $\mathbf{y}''$ be the shortest partial history so that $\mathbf{x}^j \circ \mathbf{y}''$ reaches $\bar{\pi}$. Otherwise, let $\mathbf{y}'' = \mathbf{y}$. In both cases, we have identified a partial history (the intersection of $\mathbf{y}'$ and $\mathbf{y}''$) that separates $\mathbf{x}^i$ and $\mathbf{x}^j$ into different equivalence classes. □

Proof of Theorem 5.2

The proof follows the same outline as that of Theorem 5.1; that is, we show that all the transitions feature strict preferences, and local optimality is then guaranteed by continuity and Theorem 3.1. Global optimality follows essentially the same arguments as those in Theorem 5.1 and is omitted. Note that, as in the symmetric case, by Lemma 5.2 and Proposition 3.1, when $|r_h|/|r_\ell| = M_h/M_\ell$ and when $\delta \in \Delta_N^{(M_h,M_\ell)}$,

$$r(q_i) = r_{\pi_0} + (i - K_\ell) r_c,$$

for all $i = 1, ..., N-1$, and hence $\pi_{q_{K_\ell}} = \pi_0$ as $q^o = q_{K_\ell}$. Since the proof of Lemma 5.2 also showed that $M^{(M_h,M_\ell),N-1}$ is the unique DFSA with no more than $N-1$ updating states, it follows that

$$\pi_{q_i} V_{q_i}(H) + (1-\pi_{q_i}) V_{q_i}(L) > \pi_{q_i} V_{q_j}(H) + (1-\pi_{q_i}) V_{q_j}(L) \text{ for all } j \neq i, \tag{48}$$

$$\pi_{q_{N-1},c} u^H > \pi_{q_{N-1},c} V_{q_j}(H) + (1-\pi_{q_{N-1},c}) V_{q_j}(L), \tag{49}$$

where $\pi_{q_{N-1},c}$ is the analogous posterior at q_{N-1} if it were to receive a "common" signal. As in the symmetric case, (48) follows immediately from the fact that $\pi_{q_i} V_{q_i}(H) + (1-\pi_{q_i}) V_{q_i}(L)$ is the expected payoff of the unconstrained optimum when the prior is π_{q_i}, and $M^{(M_h,M_\ell),N-1}$ with $q^o = q_i$ implements that optimal value but not $M^{(M_h,M_\ell),N-1}$ with $q^o = q_j$ for any $j \neq i$. Since for each $i < N - M_h$, $r(q_i, h) = r(q_{i+M_h})$, (48) implies that it is strictly optimal to transit to q_{i+M_h} (when $i + M_h > N-1$, to q_H) from q_i when receiving signal h according to (13). A symmetric argument holds for receiving signal ℓ. Now consider (49). By (O2) and $\delta \in \Delta_N^{(M_h,M_\ell)}$, $r(q_{N-1}, c) > r_{\bar{\pi}}$, and hence the left-side of (49) gives the optimal value under the unconstrained optimum for prior equal to $\pi_{q_{N-1},c}$, while the right-side is the expected payoff from $M^{(M_h,M_\ell),N-1}$ with $q^o = q_j$ under that prior, which is suboptimal. In particular, even at the boundary $\delta = \bar{\delta}_N^{(M_h,M_\ell)}$ and hence $r(q_{N-1}, c) = r_{\bar{\pi}}$ (and so $r(q_{N-M_h}, h) = r_{\bar{\pi}}$), the DM at that prior is indifferent between taking u^H immediately and waiting, but the latter is followed by the rule implemented by $M^{(M_h,M_\ell),N+2}$ with $q^o = q_N$, as it takes N common signals to across $\underline{\pi}$ instead of $N-1$. □

Proof of Corollary 5.1

Let $\bar{\epsilon}$ be small such that

$$\bar{r}_{i-1} + \bar{\epsilon} < r(q_i) < \bar{r}_i - \bar{\epsilon} \tag{50}$$

for all updating states q_i under $(\bar{\mu}_x^H, \bar{\mu}_x^L)_{x \in X}$. Since $\bar{r}_i$ and $\rho(q)$ are continuous in $(\mu_x^H, \mu_x^L)_{x \in X}$, for ϵ small, $\bar{r}_i$ does not change by more than $\bar{\epsilon}/2$ when $\max_{x \in X, \theta = H, L} |\bar{\mu}_x^\theta - \bar{\mu}_x^\theta| < \epsilon$ and $|r_y| < \bar{\epsilon}/2$ under $(\mu_x^H, \mu_x^L)_{x \in X}$. It then follows from (50) that $\bar{r}_{i-1} < r(q_i, y) = r(q_i) + r_y < \bar{r}_i$, and Theorem 3.1 implies that $\tau(q_i, y; q_i) = 1$ for all updating state q_i. $\square$

A Online Appendix: Missing Proofs

Proof of Proposition 2.1

Let $V^*(\pi;\delta)$ denote the optimal continuation value when the posterior (after receiving the signal) is π and the discount factor is $\delta \in (0,1)$. The existence of V^* follows from standard arguments. Define $\{V_n^*\}_{n=1}^{\infty}$ with $V_n^* : [0,1]\times[0,1] \to \mathbb{R}$ by

$$\begin{aligned} V_1^*(\pi;\delta) &= \max\{\pi u^H, (1-\pi)u^L\}, \\ \text{for } n \geq 1,\ V_{n+1}^*(\pi;\delta) &= \max\left\{\pi u^H, (1-\pi)u^L, \delta\sum_{x\in X}[\pi\mu_x^H + (1-\pi)(1-\mu_x^L)]V_n^*(\pi^x;\delta)\right\}, \end{aligned}$$

where π^x satisfies $r_{\pi^x} = r_\pi + r_x$ for each $x \in X$. Then, V_n^* converges to V^* uniformly. We only consider sufficiently high δ's such that waiting for one period is optimal at least for some π's. A straightforward induction argument shows that $V_n^*(\pi;\delta)$ increases with δ for all n, and hence $V^*(\pi;\delta)$ also increases with δ, and satisfies

$$V^*(\pi;\delta) = \max\left\{\pi u^H, (1-\pi)u^L, \delta\sum_{x\in X}[\pi\mu_x^H + (1-\pi)(1-\mu_x^L)]V^*(\pi^x;\delta)\right\}. \quad (51)$$

We claim that V^* is convex in π. Fix some δ. Let $\lambda \in (0,1)$ and π_1, π_2 be given, and let f_1^* be the optimal decision rule under π_1 and f_2^* under π_2. For any decision rule f,

$$\begin{aligned} & U(f|\lambda\pi_1 + (1-\lambda)\pi_2) = \lambda U(f|\pi_1) + (1-\lambda)U(f|\pi_2) \\ \leq\ & \lambda U(f_1^*|\pi_1) + (1-\lambda)U(f_2^*|\pi_2) = \lambda V^*(\pi_1;\delta) + (1-\lambda)V^*(\pi_2;\delta), \end{aligned}$$

where $U(f|\pi)$ is the expected utility from rule f when the prior is such that $\mathbf{P}_0(H) = \pi$, the first equality follows from (4) as the payoff is linear in π, and the inequality follows from the optimality of f_1^* and f_2^*. This then implies that V^* is convex in π.

Now, for each $\delta \in (0,1)$, let

$$\bar{\Pi}(\delta) \equiv \{\pi \in [0,1] : V^*(\pi,\delta) = \pi u^H\}.$$

Clearly $V^*(\pi;\delta) \leq \pi u^H + (1-\pi)u^L$ for all π, and hence $1 \in \bar{\Pi}(\delta)$ for any $\delta < 1$. Moreover, (51) also implies that $\bar{\Pi}(\delta)$ includes π close to 1 as well. Since $V^*(\pi;\delta) \geq \pi u^H$, as taking a^H is always an option, convexity of V^* implies that $\bar{\Pi}(\delta)$ is a convex set and hence an closed interval. Now, for π^o defined by (21), it is straightforward to verify that

$$V^*(\pi;\delta) > \pi u^H \text{ for } \pi \leq \pi^o \text{ and } V^*(\pi;\delta) > (1-\pi)u^L \text{ for } \pi \geq \pi^o.$$

Hence, if we let $\bar{\pi} = \min \bar{\Pi}(\delta)$, then $\bar{\pi} > \pi^o$. Now, $\bar{\pi}$ is determined by $F(\pi;\delta) = 0$ with

$$F(\pi;\delta) = \pi u^H - \delta \left\{ \sum_{x \in X^H} \pi \mu_x^H u^H + \sum_{x \in X - X^H} [\pi \mu_x^H + (1-\pi)\mu_x^L] V^*(\pi^x;\delta) \right\},$$

where X^H consists of signals such that $r_x > 0$, and the first term is the expected payoff from taking a^H while second term from waiting when π is close to $\bar{\pi}$; note that $V^*(\pi^x;\delta) = \pi^x u^H$ when $\pi \geq \bar{\pi}$ and $r_x > 0$. Now, we claim that, whenever F is differentiable w.r.t. π, we have $\frac{\partial}{\partial \pi} F > 0$ and hence $\bar{\pi}$ is uniquely determined. For any $x \in X - X^H$, whenever V^* is differentiable w.r.t. π at π^x,

$$\begin{aligned} & \frac{\partial}{\partial \pi}\{[\pi \mu_x^H + (1-\pi)\mu_x^L] V^*(\pi^x;\delta)\} = (\mu_x^H - \mu_x^L) V^*(\pi^x;\delta) + \left(\frac{\pi^x \mu_x^L}{\pi}\right) \frac{\partial}{\partial \pi} V^*(\pi^x;\delta) \\ \leq \;& (\mu_x^H - \mu_x^L)\pi^x u^H + \left(\frac{\pi^x \mu_x^L}{\pi}\right) u^H = \frac{\pi^x}{\pi}[\pi \mu_x^H + (1-\pi)\mu_x^L] u^H = \mu_x^H u^H, \end{aligned}$$

where the inequality follows from the fact that $r_x \leq 0$ (since $x \in X - X^H$) and hence $(\mu_x^H - \mu^L) \leq 0$ and that $V^*(\pi^\ell;\delta) \geq \pi^\ell u^H$, $\frac{\partial}{\partial \pi} V^*(\pi^\ell;\delta) \leq u^H$ as V^* is convex in π. Thus,

$$\frac{\partial}{\partial \pi} F \geq u^H - \delta \left[\sum_{x \in X^H} \mu_x^H + \sum_{x \in X - X^H} \mu_x^H \right] u^H = (1-\delta) u^H > 0,$$

since $\delta < 1$. Since F is almost everywhere differentiable, it is then strictly increasing in π. This then shows that $\bar{\pi}$ is uniquely determined by $F(\pi;\delta) = 0$, and for $\pi > \bar{\pi}$, taking a^H is strictly better. Finally, since $F(\pi;\delta)$ strictly decreases δ as V^* increases with δ, it also follows that $\bar{\pi}$ strictly increases with δ.

Proof of Theorem 3.1

Proof of multiself-consistency. The proof follows the idea in Piccione and Rubinstein (1997). Let M be an optimal SFSA. We assume that $q^o \in Q$, which is the only interesting case. For any states $q, q' \in Q \cup \{q_H, q_L\}$, define the set $W_{q,q'} = \bigcup_{n=1}^{\infty} W_{q,q'}^n$, where for each $n = 1, 2, ...$, $W_{q,q'}^n = \{\mathbf{w} = (q, x_1; q_1, x_2; ....; q_{n-1}, x_n; q') : x_i \in X, q_i \in Q\}$, that is, the set of possible state transitions from q to q'. Given a state of nature θ and $\mathbf{w} \in W^n_{q,q'}$, define

$$\mathbb{P}_\theta(\mathbf{w}) = \delta^n \times \prod_{i=1}^{n} \mu^\theta_{x_i} \tau(q_{i-1}, x_i; q_i),$$

where $q_0 = q$ and $q_n = q'$. The expected payoff from the SFSA is then

$$V = \sum_{\theta=H,L} \mathbf{P}_0(\theta) \sum_{\mathbf{w}\in W_{q^o,q_\theta}} \mathbb{P}_\theta(\mathbf{w}) u^\theta. \tag{52}$$

Note that we have shown (14) in the text. We now prove (13).

Suppose, by contradiction, that $\tau(q, x; q') > 0$ and that for some $q'' \neq q'$,

$$\pi_{q,x} V_{q'}(H) + (1 - \pi_{q,x}) V_{q'}(L) < \pi_{q,x} V_{q''}(H) + (1 - \pi_{q,x}) V_{q''}(L) \tag{53}$$

We denote $p' = \tau(q, x; q')$ and $p'' = \tau(q, x; q'')$. Now, fix all other transition probabilities other than p' and p'', each term $\mathbb{P}_\theta(\mathbf{w})$ in V given by (52) is a polynomial of (p', p'') and, since $\delta < 1$, V is differentiable w.r.t. (p', p''). Since M is optimal and $p' = \tau(q, x; q') > 0$, the FOCs require that $\frac{\partial}{\partial p'} V \geq \frac{\partial}{\partial p''} V$. However, we show below that (53) implies that

$$\frac{\partial}{\partial p''} V > \frac{\partial}{\partial p'} V, \tag{54}$$

a contradiction to the optimality of M. To prove (54), first note that

$$\frac{\partial}{\partial p'} V = \sum_{\theta=H,L} \mathbf{P}_0(\theta) \sum_{\mathbf{w}\in W_{q^o,q_\theta}(q,x;q')} \varphi_{(q,x;q')}(\mathbf{w}) \frac{\mathbb{P}_\theta(\mathbf{w})}{p'} u^\theta, \tag{55}$$

where $W_{q^o,q_\theta}(q, x; q') = \{\mathbf{w} \in W_{q^o,q_\theta} : (q, x, q') \text{ occurs in } \mathbf{w}\}$ and $\varphi_{(q,x;q')}(\mathbf{w})$ is the number of repetitions of $(q, x; q')$ within $\mathbf{w}$.

Now, we show that $\frac{\partial}{\partial p'} V$ is proportional to $\pi_{q,x} V_{q'}(H) + (1 - \pi_{q,x}) V_{q'}(L)$ by multiplying the latter by $\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta$, a quantity independent of q' and q'':

$$\begin{aligned}
&\left[\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta\right][\pi_{q,x}V_{q'}(H)+(1-\pi_{q,x})V_{q'}(L)]\\
=\;&\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q'}(\theta)=\sum_\theta \mathbf{P}_0(\theta)\left[\sum_{\mathbf{w}_q\in W_{q^o,q}}\mathbb{P}_\theta(\mathbf{w}_q)\right]\mu_x^\theta\left[\sum_{\mathbf{w}_{q'}\in W_{q',q_\theta}}\mathbb{P}_\theta(\mathbf{w}_{q'})\right]u^\theta\\
=\;&\sum_\theta \mathbf{P}_0(\theta)\left\{\sum_{\mathbf{w}_q\in W_{q^o,q},\mathbf{w}_{q'}\in W_{q',q_\theta}}\frac{\mathbb{P}_\theta[(\mathbf{w}_q,x;\mathbf{w}'_q)]}{\tau(q,x;q')}\right\}u^\theta\\
=\;&\sum_\theta \mathbf{P}_0(\theta)\left\{\sum_{\mathbf{w}\in W_{q^o,q_\theta}}\varphi_{(q,x;q')}(\mathbf{w})\frac{\mathbb{P}_\theta(\mathbf{w})}{p'}\right\}u^\theta=\frac{\partial}{\partial p'}V,
\end{aligned}$$

where the first equality follows from (11), the last follows from (55) and the second last follows from $p'=\tau(q,x;q')$ and the fact that for any $\mathbf{w}_q\in W_{q^o,q}$ and any $\mathbf{w}_{q'}\in W_{q',q_\theta}$, $(\mathbf{w}_q,x;\mathbf{w}_{q'})\in W_{q^o,q_\theta}(q,x;q')$ and that each $\mathbf{w}\in W_{q^o,q_\theta}(q,x;q')$ is counted $\varphi_{(q,x;q')}(\mathbf{w})$ times in that list. We have analogous expression for $\frac{\partial}{\partial p''}V$, and hence (53) implies (54).

Proof of (15) First, we claim that, for all $q,q'\in Q$,

$$\pi_q V_q(H)+(1-\pi_q)V_q(L)\ge \pi_q V_{q'}(H)+(1-\pi_q)V_{q'}(L)\text{ for all }q'\in Q. \tag{56}$$

To see this, for any x and any $q_1,q_2,q_3\in Q\cup\{q_H,q_L\}$ such that $\tau(q,x;q_1)>0$ and $\tau(q,x;q_2)>0$, by modified multi-self consistency and (11),

$$\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q_1}(\theta)=\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q_2}(\theta)\ge\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q_3}(\theta). \tag{57}$$

Thus,

$$\begin{aligned}
&\pi_q V_q(H)+(1-\pi_q)V_q(L)\\
=\;&\pi_q\delta\sum_{x\in X,q''\in Q\cup\{q_H,q_L\}}\mu_x^H\tau(q,x;q'')V_{q''}(H)+(1-\pi_q)\delta\sum_{x\in X,q''\in Q\cup\{q_H,q_L\}}\mu_x^L\tau(q,x;q'')V_{q''}(L)\\
=\;&\sum_{x\in X}\left\{\delta\sum_{q''\in Q\cup\{q_H,q_L\}}\frac{\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q''}(\theta)}{\sum_{\theta'=H,L}\mathbf{P}_0(\theta')\mathbb{P}(q|\theta')}\tau(q,x;q'')\right\}\\
\ge\;&\sum_{x\in X}\left\{\delta\sum_{q''\in Q\cup\{q_H,q_L\}}\frac{\sum_{\theta=H,L}\mathbf{P}_0(\theta)\mathbb{P}(q|\theta)\mu_x^\theta V_{q''}(\theta)}{\sum_{\theta'=H,L}\mathbf{P}_0(\theta')\mathbb{P}(q|\theta')}\tau(q',x;q'')\right\}=\pi_q V_{q'}(H)+(1-\pi_q)V_{q'}(L),
\end{aligned}$$

where the first equality follows from the recursive equation for $V_q(\theta)$ for each $\theta = H, L$, the second follows from (11), the inequality, which holds term by term for each x, follows from (57): any term with q'' in the numerator of the expression above the inequality with $\tau(q, x; q'') > 0$ has the same value and that value is no less than any corresponding term with $\tau(q', x; q'') > 0$ below, and the last equality follows from the recursive equation for $V_{q'}(\theta)$.

Now, we claim that $\Delta V^H_{i,j} < 0$ and $\Delta V^L_{i,j} > 0$ for all $i < j$, and $\Delta V^L_{j,k}/\Delta V^H_{k,j} \geq \Delta V^L_{i,j}/\Delta V^H_{j,i}$ for all $i < j < k$. To see this, let $i < j$ be given. If $i = 0$ or $j = K + 1$, this follows from the fact that $V_{q_\theta}(\theta) = u^\theta$, and $V_q(\theta) \leq \delta u^\theta$ for all $q \in Q$. Otherwise, by (56),

$$\pi_{q_j}\Delta V^H_{j,i} + (1 - \pi_{q_j})\Delta V^L_{j,i} \geq 0, \text{ and } \pi_{q_i}\Delta V^H_{i,j} + (1 - \pi_{q_i})\Delta V^L_{i,j} \geq 0. \tag{58}$$

Since there are no equivalent states, either $\Delta V^H_{i,j} > 0$ or $\Delta V^H_{i,j} < 0$. By our convention it must be $\Delta V^H_{j,i} > 0$. By the second inequality in (58), $\Delta V^L_{i,j} \geq 0$. Now, if this last inequality is an equality, then we can replace all the transition to q_i to transition to q_j and obtain a higher ex ante payoff, which is a contradiction to the optimality of the SFSA. Now, let $i < j < k$. Again, by (56), we have

$$\pi_{q_j}\Delta V^H_{j,i} + (1 - \pi_{q_j})\Delta V^L_{j,i} \geq 0, \text{ and } \pi_{q_j}\Delta V^H_{j,k} + (1 - \pi_{q_j})\Delta V^L_{j,k} \geq 0, \tag{59}$$

and hence

$$\frac{\Delta V^L_{i,j}}{\Delta V^H_{j,i}} \leq \frac{\pi_{q_j}}{1 - \pi_{q_j}} \leq \frac{\Delta V^L_{j,k}}{\Delta V^H_{k,j}}.$$

Finally, we show (15). Let $q \in Q$ be given. By (13), $\tau(q, x; q_i) > 0$ only if

$$\pi_{q,x}V_{q_i}(H) + (1 - \pi_{q,x})V_{q_i}(L) \geq \pi_{q,x}V_{q_j}(H) + (1 - \pi_{q,x})V_{q_j}(L)$$

for both $j = i - 1$ and $j = i + 1$. This then implies (15). Conversely, it is straightforward to verify that if (15) holds, then

$$\pi_{q,x}V_{q_i}(H) + (1 - \pi_{q,x})V_{q_i}(L) \geq \pi_{q,x}V_{q_j}(H) + (1 - \pi_{q,x})V_{q_j}(L)$$

for any $j = 0, ..., K + 1$, where $q_0 = q_L$ and $q_{K+1} = q_H$. Note that we need the fact that $\bar{r}_i$ increases with i for this, and we have proved this earlier. Moreover, if $r(q, x) \in (\bar{r}_{i-1}, \bar{r}_i)$, then the above inequality is strict for any $j \neq i$ and hence $\tau(q, x; q_i) = 1$.

Proof of Theorem 6.1

(*if* part) Let R be a right-invariant equivalence relation with $K+2$ equivalence classes, $I_1, ..., I_K, I_H, I_L$, where $\mathcal{L}_H^f = I_H$ and $\mathcal{L}_H^f = I_L$. Define M as follows. Let $Q = \{q_1, ..., q_K\}$, with q_i corresponding to I_i. Define τ as follows. For any updating state $q_i \in Q$ and for any $x \in X$, let $\tau(q_i, y) = q_j$ if $\mathbf{y} \in I_i$ and $\mathbf{y} \circ x \in I_j$, where $j = 1, ..., K$ and $j = H, L$. This is well-defined because R is right-invariant: if $\mathbf{y}, \mathbf{y'} \in I_i$, then $\mathbf{y} R \mathbf{y'}$, but this also implies that $(\mathbf{y} \circ x) R (\mathbf{y'} \circ x)$ and hence j is uniquely determined. Let q^o be the internal state q_i such that $\emptyset \in I_i$. Finally, M implements the rule f since $\mathbf{x} \in \mathcal{L}_\theta^f$ if and only if $\lambda_M(\mathbf{x}) = q_\theta$ under M by construction.

(*only if* part) Let M be a DFSA that implements f with $|Q| = K$. We extend τ so that $\tau(q_\theta, x) = q_\theta$ for $\theta = H, L$ and $x \in X$, and hence extending λ_M accordingly. Define $R \subset X^* \times X^*$ by

$$\mathbf{x} R \mathbf{y} \iff \lambda_M(\mathbf{x}) = \lambda_M(\mathbf{y}).$$

We use $[q]$ to denote the equivalence class such that $\mathbf{x} \in [q]$ if and only if $\lambda_M(\mathbf{x}) = q$. R is right-invariant. If $\mathbf{x}, \mathbf{y} \in [q]$, then $\lambda_M(\mathbf{x}) = \lambda_M(\mathbf{y}) = q$ and hence $\lambda_M(\mathbf{x} \circ \mathbf{z}) = \lambda_M(\mathbf{y} \circ \mathbf{z})$ for any $\mathbf{z}$, which in turn implies that $(\mathbf{x} \circ \mathbf{z}) R (\mathbf{y} \circ \mathbf{z})$. Moreover, non-redundancy of M implies that R has exactly $K+2$ equivalence classes. Finally, $\mathcal{L}_\theta^f = [q_\theta]$ and hence is one of the equivalence classes from R.

Proof of Theorem 3.2

(1) By Proposition 2.1, $\bar{\pi}$ converges to one or $\underline{\pi}$ converges to zero as δ converges to one. Since $r_h < +\infty$ or $r_\ell > -\infty$, say the former is true, for any given π, it takes at least $K' > K+3$ good signals to reach $\bar{\pi}$ for δ sufficiently large. We now show that under the optimal decision rule, denoted by f, the sets $\mathcal{L}_H^f$ and $\mathcal{L}_L^f$ cannot be written as some unions of equivalence classes from a right-invariant relation with $K+2$ equivalence classes. Now, consider the partial histories $\mathbf{x}^k$, $k = 1, ..., K+3$, where $\mathbf{x}^k$ consists of k h-signals. We claim that each partial history has to be of a distinct equivalence class. To do so, for any $k > k'$, we only need to show that there is a *witness* partial history $\mathbf{y}^{k,k'}$ such that $\mathbf{x}^k \circ \mathbf{y}^{k,k'}$ while $\mathbf{x}^{k'} \circ \mathbf{y}^{k,k'}$ leads to another. Take $\mathbf{y}^{k,k'} = \mathbf{x}^{K'-k'}$. Then, $\mathbf{x}^{k'} \circ \mathbf{x}^{K'-k'}$ leads to action a^H while $\mathbf{x}^k \circ \mathbf{x}^{K'-k'}$ leads to action c (continue) under the optimal rule f. One can append $\mathbf{y}^{k'k}$ with sufficiently many ℓ-signals in the end so that $\mathbf{x}^k \circ \mathbf{y}^{k,k'}$ leads to action a^L.

(2) Suppose, by contradiction, that there is a DFSA that implements the optimal decision rule, f^*, with K states for some finite K. By the Theorem 6.1, there exists a right-invariant equivalence relation R with K equivalence classes over X^* such that the set $\mathcal{L}_\theta^{f^*}$ is a union of some of those classes for both $\theta = H, L$. To show that this cannot be the case, we need to show that for any right-invariance equivalence relation R such that $\mathcal{L}_H^{f^*}$ and $\mathcal{L}_L^{f^*}$ are unions of its equivalence classes has infinitely many equivalence classes. To this end, we construct an infinite sequence of partial histories, $\{\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^n, ...\}$, $\mathbf{x}^n \in X^*$ for each n, and each represents such an equivalence class. For any $n \neq n'$, to show that $\mathbf{x}^n$ and $\mathbf{x}^{n'}$ belong to different equivalence classes, we construct a *witness* history, denoted by $\mathbf{y}^{n,n'}$, such that $\mathbf{x}^n \circ \mathbf{y}^{n,n'}$ and $\mathbf{x}^{n'} \circ \mathbf{y}^{n,n'}$ dictate different actions under the optimal policy.

To simplify notation, we normalize $r_h = \rho$ and $r_\ell = -1$. This is with no loss of generality as we can think of these numbers in terms of multiples of r_ℓ. Generically, ρ is a normal number, that is, for any finite sequence $j_1 j_2 \cdots j_l \in \{0, 1, ..., 9\}^l$, we can locate the sequence somewhere in its decimal expansion.

We construct our sequence of partial histories, $\{\mathbf{x}^1, \mathbf{x}^2, ...\mathbf{x}^n, ...\}$, as follows: first, let $\mathbf{x}^1 = (h)$ and, for $n > 1$, let

$$\mathbf{x}^n = \mathbf{x}^{n-1} \circ (\ell, \underbrace{h, ..., h}_{(k_n - k_{n-1})\ h\text{'s}}),$$

where k_n is the largest integer such that $k_n\rho - (n-1) \leq \rho$. Note that for all n, any initial segment $\mathbf{y} \subset \mathbf{x}^n$ satisfies

$$r_{\mathbf{y}} \in (-1, \rho). \tag{60}$$

Moreover, for each n,

$$r_{\mathbf{x}^n} = k_n\rho - (n-1) \in (0, \rho). \tag{61}$$

Since $r_{\bar{\pi}} - r_\pi > r_h = \rho$ and $r_\pi - r_{\underline{\pi}} > 2|r_\ell| = 2$, it is optimal to wait along the whole sequence following the partial history $\mathbf{x}^n$. Note also that since ρ is a normal number and hence irrational, $r_{\mathbf{x}^n} \neq r_{\mathbf{x}^{n'}}$ for all $n \neq n'$.

Now, we show that for any two partial histories $\mathbf{x}^n$ and $\mathbf{x}^{n'}$, we can find a witness history $\mathbf{y}^{n,n'}$ such that it is optimal to wait following $\mathbf{x}^n \circ \mathbf{y}^{n,n'}$ (and hence optimal to take a^L following $\mathbf{x}^n \circ \mathbf{y}^{n,n'}$ followed by sufficiently many ℓ's) but optimal to take a^H following $\mathbf{x}^{n'} \circ \mathbf{y}^{n,n'}$. Note that (60) and (61) imply that it is optimal to wait following $\mathbf{x}^n$ for all n.

Without loss of generality we may assume $r_{\mathbf{x}^n} < r_{\mathbf{x}^{n'}}$. Let $\bar{k}$ be the unique integer that satisfies

$$r_{\mathbf{x}^n} + \bar{k}\rho < r_{\bar{\pi}} - r_\pi \leq r_{\mathbf{x}^n} + (\bar{k}+1)\rho.$$

By (61) we know that $\bar{k} \geq 1$. Hence, along the partial history $\mathbf{x}^n$ followed by $\bar{k}$ high signals, it is optimal to wait. We consider two cases.

First, suppose that $r_{\mathbf{x}^{n'}} + \bar{k}\rho > r_{\bar{\pi}} - r_\pi$, i.e., along the partial history $\mathbf{x}^{n'}$ followed by $\bar{k}$ high signals, it is optimal to take a^H. In this case, take $\mathbf{y}^{n,n'}$ to be $\bar{k}$ high signals satisfies the requirement.

Second, suppose that $r_{\mathbf{x}^{n'}} + \bar{k}\rho \leq r_{\bar{\pi}} - r_{\pi}$, i.e., along the partial history $\mathbf{x}^{n'}$ followed by $\bar{k}$ high signals, it is also optimal to wait. Then, set

$$\epsilon_1 = (r_{\bar{\pi}} - r_{\pi}) - (r_{\mathbf{x}^{n'}} + \bar{k}\rho) < (r_{\bar{\pi}} - r_{\pi}) - (r_{\mathbf{x}^{n}} + \bar{k}\rho) = \epsilon_2. \tag{62}$$

It then follows that $\epsilon_1 \geq 0$ as $r_{\mathbf{x}^{n'}} + \bar{k}\rho \leq r_{\bar{\pi}} - r_{\pi}$ and $\epsilon_2 \leq \rho$ as $r_{\bar{\pi}} - r_{\pi} \leq r_{\mathbf{x}^{n'}} + (\bar{k}+1)\rho$. To construct the witness partial history $\mathbf{y}^{n,n'}$ in this case, we need the following lemma, whose proof is given at the end and makes use of normality of ρ.

Lemma A.1. *For given $\epsilon_1 < \epsilon_2 \in [0, \rho]$, there exists a partial history $\mathbf{z}$ such that*

$$\epsilon_1 < r_{\mathbf{z}} < \epsilon_2. \tag{63}$$

Moreover, for any initial segment $\mathbf{y} \subset \mathbf{z}$, $-1 < r_{\mathbf{y}} < \rho$.

By Lemma A.1, we can find $\mathbf{z}$ with $r_{\mathbf{z}} \in (\epsilon_1, \epsilon_2)$ w.r.t. ϵ_1 and ϵ_2 given by (62), and, for any initial segment $\mathbf{y} \subset \mathbf{z}$, $-1 < r_{\mathbf{y}} < \rho$. Thus, by (60) and (61), along the partial history $\mathbf{x}^n \circ \mathbf{z}$ or $\mathbf{x}^{n'} \circ \mathbf{z}$, it is optimal to wait all along. Now, let $\mathbf{y}^{n,n'}$ be equal to $\mathbf{z}$ followed by $\bar{k}$ high signals, and hence

$$r_{\mathbf{x}^n \circ \mathbf{y}^{n,n'}} = r_{\mathbf{x}^n} + r_{\mathbf{z}} + \bar{k}\rho < r_{\mathbf{x}^n} + \epsilon_2 + \bar{k}\rho \leq r_{\bar{\pi}} - r_{\pi},$$

where the first inequality follows from $r_{\mathbf{z}} < \epsilon_2$ and the second from (62). Hence, it is optimal to wait following the partial history $\mathbf{x}^n \circ \mathbf{y}^{n,n'}$. In contrast,

$$r_{\mathbf{x}^{n'} \circ \mathbf{y}^{n,n'}} = r_{\mathbf{x}^n} + r_{\mathbf{z}} + \bar{k}\rho > r_{\mathbf{x}^{n'}} + \epsilon_1 + \bar{k}\rho \geq r_{\bar{\pi}} - r_{\pi},$$

where the first inequality follows from $r_{\mathbf{z}} > \epsilon_1$ and the second from (62). Hence, it is optimal to take a^H following the partial history $\mathbf{x}^{n'} \circ \mathbf{y}^{n,n'}$, and $\mathbf{y}^{n,n'}$ is a valid witness partial history.

Proof of Lemma A.1

Let n be the largest integer for which $\epsilon_2 - \epsilon_1 < 10^{-n}$. Then, if

$$\epsilon_1 = 0.j_1 j_2 \ \cdots j_n j_{n+1} \ \cdots < 0.j_1 j_2 \ \cdots j_n (j_{n+1} + 1) \leq \epsilon_2,$$

by normality, for some k,

$$\rho = 0.i_1 \cdots i_k \ j_1 j_2 \cdots j_n j_{n+1} \cdots > 0.i_1 \cdots i_k \ j_1 j_2 \cdots j_n j_{n+1}.$$

Now, for $N_1 = 10^k$ and $M_1 = i_1 \cdots i_k$, $N_1\rho - M_1 \in (\epsilon_1, \epsilon_2)$. We shall then define $\mathbf{z}$ inductively, but from the end to the beginning. At the end, $\mathbf{z}$ has N_1 h and M_1 ℓ, and we construct it by an increasing sequence $\{n_i^h, n_i^\ell\}$, where n_i^h is the number of h's and n_i^ℓ number of ℓ's in its initial segment of length i. Thus, we begin with $n^h_{N_1+M_1} = N_1$ and $n^\ell_{N_1+M_1} = M_1$. We construct this sequence through another sequence, $\{N_2, N_3, ..\}$, defined inductively as follows. First, let N_2 be the smallest integer such that $N_2\rho - M_1 > -1$. Inductively, N_k is the smallest integer such that

$$N_k\rho - [M_1 - (k-2)] = N_k\rho - M_1 + (k-2) > -1. \tag{64}$$

Note that this is a decreasing sequence, and that

$$0 < N_k\rho - [M_1 - (k-1)] = N_k\rho - M_1 + (k-1) < \rho. \tag{65}$$

The inequality on the left follows directly from (64), and the right follows from a proof by contradiction: if $N_k\rho - M_1 + (k-1) \geq \rho$, then $(N_k - 1)\rho - M_1 + (k-1) \geq 0$, a contradiction to the fact that N_2 is the smallest integer for (64) to hold. Moreover, since $\rho < 1$, it also follows that N_k is strictly decreasing in k, and since N_1 is finite, it ends in finite time.

Now we construct the sequence $\{n_i^h, n_i^\ell\}$ by induction, but backwards. First, for $i = N_1 + M_1$ to $N_2 + M_1 + 1$,

$$(n^h_{i-1}, n^\ell_{i-1}) = (n^h_i - 1, n^\ell_{i-1}) = (n^h_i - 1, M_1).$$

That is, we decrease the number of h's by one but keep the number of ℓ's constant. Then, we move the number of ℓ's by one:

$$(n^h_{N_2+M_1-1}, n^\ell_{N_2+M_1-1}) = (n^h_{N_2+M_1}, M_1 - 1) = (N_2, M_1 - 1).$$

Inductively, for $i = N_k + M_1 - k$ to $N_{k+1} + M_1 - (k+1)$,

$$(n^h_{i-1}, n^\ell_{i-1}) = (n^h_i - 1, n^\ell_{i-1}) = (n^h_i - 1, M_1 - k).$$

It then follows directly from (64) and (65) that for all i,

$$n_i^h\rho - n_i^\ell \in (-1, \rho), \tag{66}$$

that is, for all initial segment $\mathbf{y} \subset \mathbf{z}$, $r_{\mathbf{y}} \in (-1, \rho)$.