Value of Information when Searching for a Secretary

Karl H. Schlag
University of Vienna[†]

Andriy Zapechelnyuk
University of St Andrews[‡]

February 13, 2019

Abstract

The secretary problem is the canonical model of search under ambiguity, in which secretaries are being interviewed in a random order. We assume that the number of secretaries is unknown and that one cares for the value of the secretary. We measure the value of information as a multiplier that describes how much better off one could have been had one known the distribution of secretaries' values. It is evaluated in the worst case, for all distributions and at all rounds of search. Under perfect recall, knowledge of the applicant pool size and their distribution can improve one's payoff at most 4 times. Knowledge that the values are i.i.d. does not improve one's payoff.

[†]Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. *E-mail:* karl.schlag@univie.ac.at.

[‡]School of Economics and Finance, University of St Andrews, Castlecliffe, the Scores, St Andrews KY16 9AR, UK. *E-mail:* az48@st-andrews.ac.uk.

1 Introduction

In our daily lives we are constantly searching for something, be it a job or a specific good. Decisions whether to continue searching or not are typically made under sparse information about yet undiscovered options, such as prices when searching for a product, values of job offers when looking for a job, or quality of job applicants when hiring.

In this paper we are interested in the *value of information* about undiscovered search options. The classical *secretary problem* (Fox and Marnie 1960) is the canonical model of search. A given number of secretaries are interviewed one by one in a random order. Their values are a priori unknown and only learned during the interview. Secretaries must be employed or rejected on the spot, and the decision is irrevocable. The objective is to hire the best secretary. This problem has a simple and elegant solution, in which the best secretary is selected with probability approximately equal to $1/e$. We can interpret its reciprocal, e, as the value of information about the market of secretaries, since it tells us how many times greater a payoff one could have obtained had she known in advance the values of all secretaries in the interview pool.

In this paper we consider a more general and more economically plausible formulation of the secretary problem. We believe it makes more sense to be concerned with the value of the secretary hired, not just whether she is the best in the applicant pool. We find it implausible that one knows how many secretaries will respond to an advertisement. We prefer to stay in line with the standard assumption in economics that payoffs are discounted over time. We are also interested in varying the degree of recall, the ability to choose secretaries that have been interviewed in the past, ranging from no recall to free recall.

The notion of the value of information is extended to our setting as follows. A market is described by a number of secretaries in the pool and a joint distribution of their values. We evaluate the performance of a search rule in a given market after a given history as the ratio of the first-best payoff to the rule's expected payoff after that history. By the first-best payoff we understand the maximal expected payoff one could have attained if one knew the market. As different markets and different search histories will typically lead to different ratios, we evaluate a rule according to the maximal ratio across all markets and all histories. With a slight abuse of terminology we refer to this ratio as the competitive ratio. Our way to measure the value of information is by the smallest competitive ratio that can be achieved by a rule. It measures how much better off, at most, one could have been had one known the market, irrespective of the market one faces and the round of search.

Note that in the literature (e.g., Babaioff et al., 2007) the competitive ratio is only evaluated ex-ante, while we consider the maximum across all rounds of the search.[1] We believe it is important to measure how much one can be better off by learning the market not only at the outset, but also at later rounds, after more information about the distribution has become available.

In the paper we wish to understand the value of information under different assumptions about the market. For instance, how valuable is it to know that the values are i.i.d.? For hiring secretaries, the iid assumption does not make sense, as this would mean either that there are infinitely many secretaries, or that the secretaries are drawn with replacement from some pool. So, we also consider more general markets in which values are correlated. As we assume that secretaries are drawn from the pool in a random order, we are interested in exchangeable joint distributions, under which the order of draws does not matter.

We derive the value of information for every discount factor less than one. For the case of free recall the value of information is 4, independently of the discount factor. In other words, by knowing the market one can increase the payoff of the searcher at most 4 times. For the case of partial or no recall, the value of information varies with the discount factor. In particular, knowing the market is more valuable for more patient decision makers.

We note that being better informed does not necessarily reduce the value of information. We show that if the decision maker initially has no information, and then learns that she faces an i.i.d. market (but does not learn the distribution), then the value of information remains unchanged. So knowing that a market is i.i.d. is not helpful.

[1]The competitive ratio has been introduced as a measure of the loss of a given algorithm relative to the optimal algorithm, smaller values meaning better performance (Sleator and Tarjan 1985). Is in Babaioff et al. (2007), we use the same term to evaluate gains, instead of losses, with the reciprocal ratio.

The search rule that minimizes the competitive ratio, and thus used to derive the value of information, is extremely simple. It prescribes to stop searching with a constant probability, independently of the history of past observations. It sounds naive as it ignores any information. However, the reason why it is optimal to ignore the past is intricate. The optimal decision when to stop searching manages a careful tradeoff between exploration and exploitation. Exploration means to gather new values by not stopping. Exploitation means to take advantage of the given bag of accumulated values by stopping and thus to avoid the cost of waiting that is due to discounting and loss of past options. It turns out that this tradeoff is optimized by a mixed strategy, a constant stopping probability. The larger the discount factor, the less costly is waiting, and hence the lower the probability of stopping. A decrease in the probability of recall leads naturally to a higher stopping probability as waiting becomes more costly.

The literature on the secretary problem is large and covers various extensions of the classical problem (for review see Freeman 1983 and Ferguson 1989). Presman and Sonin (1973) considered search for a best secretary with an unknown number of secretaries. The question of partial recall was first addressed in Yang (1974). Payoff discounting was introduced by Rasmussen and Pliska (1975). Mucci (1973) introduced general utility function into the classical problem, which has been actively studied since then, albeit under specific assumptions on the distribution of values. Babaioff et al. (2007) were the closest to the approach is this paper, they made no assumptions on the distributions of values, and derived an upper bound on the asymptotic competitive ratio for the problem with a known n.

In a related paper Schlag and Zapechelnyuk (2016) consider the model with substantially more structure, where values are i.i.d. with known support, and previously interviewed secretaries can be freely recalled. The paper focuses on finding a practically useful search rule whose performance is evaluated from the perspective of Bayesian decision makers with various priors. Finally, Bergemann and Schlag (2011) consider search among few options, minimizing the maximal difference between the first-best payoff to the rule's expected payoff.

2 Model

2.1 Preliminaries

There is a pool of n offers with random values. An individual draws offers one by one in a random order. At each stage the individual decides whether to stop the search or to draw another offer. When the search is stopped, the individual picks the highest-valued available offer. Waiting is costly in two ways. First, values decay with time by a discount factor. Second, previously drawn offers randomly disappear.

Offers are described by a set of n nonnegative random variables, $\mathcal{X}_n = \{X_1, ..., X_n\}$, whose joint c.d.f is denoted by F. We make the following assumptions.

(A$_1$) Support of F is a compact subset of $\mathbb{R}^n_+$.
(A$_2$) F is exchangeable.[2]

Compactness of F guarantees that expected payoffs are finite.[3] Exchangeability of F formalizes the assumption the offers are drawn in random order.

The pair (n, F) is referred to as the *environment*. Denote by $\mathcal{E}$ the class of environments that satisfy the above assumptions,

$$\mathcal{E} = \{(n, F) : n \in \mathbb{N},\ F \text{ satisfies } (\mathrm{A}_1) \text{ and } (\mathrm{A}_2)\}.$$

Let x_0 be the value of the outside option which is available to the individual in all periods. Denote by $x_1, x_2, ..., x_n$ the realized values of the offers in the order the individual receives them. Past offers may not remain available all the time. With every period passed, each offer randomly disappears with probability $1 - \alpha \in [0, 1]$, independently of others. Thus, in every period t the outside option x_0 is always available, and offer x_k is available with probability α^{t-k} for each $k = 1, ..., t$. Denote by Y_t the set of offers available at stage t.

[2]A joint distribution of $\{X_1, ..., X_n\}$ is exchangeable if for every permutation π of indices $\{1, ..., n\}$, the tuple $(X_{\pi(1)}, ..., X_{\pi(n)})$ has the same joint distribution.

[3]At the expense of more cumbersome notations, it would suffice to assume finiteness of all conditional expectations of $\{X_1, ..., X_n\}$.

The two extreme cases, $\alpha = 0$ and $\alpha = 1$, correspond to *no recall* and *free recall* assumptions, respectively. *No recall* means that every offer is lost if not taken immediately, so $Y_t = \{x_0, x_t\}$ for every t. *Free recall* means that past offers remain available forever, so $Y_t = \{x_0, x_1, ..., x_t\}$ for every t.

If the individual stops the search at stage t, she chooses the best available offer (including the outside option) and obtains the payoff

$$\delta^t y_t, \quad t = 0, 1, ..., n,$$

where $\delta \in (0, 1)$ is the discount factor and y_t denotes the best available value,

$$y_t = \max\{Y_t\}.$$

The search ends after the individual decides to stop, or after all n offers have been received.

Discount factor δ, outside option value x_0, and offer persistence probability α are parameters. We assume

$$0 < x_0 < \delta < 1.$$

A positive outside option, $x_0 > 0$, together with impatience of the individual, $\delta < 1$, implies that search is costly, as every new offer costs the individual at least the delayed consumption of the value x_0. Also, we assume $x_0 < \delta$, as otherwise it makes no sense for anyone to search, the search problem then being trivial.

A *search rule* p prescribes for every stage $t = 0, 1, 2, ...$, every history of draws $h_t = (x_1, ..., x_t)$, including the empty history at $t = 0$, and every set of available offers Y_t the probability $p(h_t, Y_t)$ of stopping at that stage. We assume that the individual is not aware of the environment $E = (n, F)$, in other words, she knows neither how many offers she can receive, nor how their values are distributed. She only knows that E belongs to the class of environments $\mathcal{E}$.

Note that even though the individual does not know n, the search automatically stops if all n offers are received. This is as if the individual discovers after n draws that there are no more offers left, and hence stops the search.[4]

[4] The results do not change if the individual discovers n some periods in advance, since the results are driven by large n.

2.2 Performance criterion

For every environment $E = (n, F) \in \mathcal{E}$, every period $t < n$, and every history of draws $h_t = (x_1, ..., x_t)$, and every set of available offers $Y_t \subset \{x_0, x_1, ..., x_t\}$, we denote by $U_t^p(E, h_t, Y_t)$ the expected payoff from a given search rule p conditional on h_t and Y_t. For every finite n it is defined by backward induction, $U_n^p(E, h_n, Y_n) = y_n$, and for every $t < n$

$$U_t^p(E, h_t, Y_t) = p(h_t, Y_t)y_t + (1 - p(h_t, Y_t))\delta \mathrm{E}_{\bar{F}}\left[U_{t+1}^p(E, h_{t+1}, Y_{t+1})\middle| h_t, Y_t\right],$$

where $\mathrm{E}_F\left[\,\cdot\,|h_t, Y_t\right]$ denotes the conditional expectation in the next stage under the c.d.f. $\bar{F}$.

The *first-best payoff* under environment E, history h_t, and set of available offers Y_t is

$$V_t(E, h_t, Y_t) = \sup_p U_t^p(E, h_t, Y_t).$$

The maximal performance ratio of the first-best payoff to p's payoff across all environments in $\mathcal{E}$ and after all histories is called the *competitive ratio of search rule* p,

$$R^p = \sup_{t,\ h_t,\ Y_t} \left\{ \sup_{E \in \mathcal{E}(h_t)} \frac{V_t(E, h_t, Y_t)}{U_t^p(E, h_t, Y_t)} \right\},$$

where $\mathcal{E}(h_t)$ denotes the set of environments whose distributions include the historical realizations $\{x_1, ..., x_t\}$ in their support:

$$\mathcal{E}(h_t) = \{(n - t, F) \in \mathcal{E} \,:\, n > t,\ \{x_1, ..., x_t\} \subset supp(F_t)\}.$$

where F_t is the marginal c.d.f. of the first t draws.

3 Value of Information

3.1 Main result

The *value of information*, denoted by R^*, is the smallest competitive ratio,

$$R^* = \inf_p R^p.$$

A search rule that minimizes the competitive ratio will be called *optimal*. An individual who follows the optimal rule can guarantee that learning the true environment can increase her payoff at most R^* times, no matter what environment she faces and no matter what round of search she is in.

Let us now state the main result. Consider the search rule p^* that prescribes to stop at each stage $t = 0, 1, 2...$ with a constant probability, independent of both the history of draws and the set of available offers,

$$p^*(h_t, Y_t) = 1 - \frac{1 - \sqrt{(1-\delta)(1-\alpha\delta)}}{1 - (1-\delta)(1-\alpha\delta)}.$$

Theorem 1. *Search rule p^* is optimal. The value of information is*

$$R^* = \left(1 + \sqrt{\frac{1-\alpha\delta}{1-\delta}}\right)^2.$$

Before proving Theorem 1, we first discuss some properties of the result. The proof is deferred to the next section.

Free recall. It is remarkable that for the case of free recall, $\alpha = 1$, the value of information is independent of the discount factor $\delta \in (0, 1)$,

$$R^* = 4 \text{ if } \alpha = 1.$$

No recall. The value of information is monotonic in the persistence parameter α, and under no recall, $\alpha = 0$, reaches the lowest value

$$R^* = \left(1 + \sqrt{\frac{1}{1-\delta}}\right)^2 \text{ if } \alpha = 0.$$

Approximations. For a fixed $\alpha < 1$ and δ close to 1 we can approximate the value of information by

$$R^* = \frac{1-\alpha}{1-\delta} + O\left((1-\delta)^{-\frac{1}{2}}\right).$$

On the other side of the domain, for $\delta(1-\alpha)$ close to 0 we can approximate the value of information by

$$R^* = 4 + \frac{2\delta(1-\alpha)}{1-\delta} + O\left(\delta^2(1-\alpha)^2\right).$$

Interestingly, when the individual is extremely impatient, $\delta \to 0$, the value of information approaches 4. But if the individual does not care about the future, then the her payoff will be exactly as high as the first-best, since she stops immediately and the information about search options is irrelevant for her. So there is a discontinuity at $\delta = 0$. This is because for every positive δ, however small, values of future search options can be extremely high, so it is worth searching.

History independence. We would like to outline the intuition why the optimal search rule should be constant, independent of the history of past offers and the set of available offers.

For any history of draws and any set of available alternatives, it might be the case that the next-stage offer has a deterministic extremely high value relative to the best current offer. Under compete information one knows that and surely waits for the high value. For that opportunity not to be completely missed out, an uniformed individual should continue with a large enough probability. Since the value of the next-stage offer can be arbitrarily high, the current history and the current best offer are irrelevant for that decision.

On the other hand, the probability to continue should not be too high, because it might be the case that all future offers are very bad. Under complete information, one would stop immediately and get the best current offer. In the event that the uninformed individual also stops and gets the best current offer, neither the value of that offer, nor the preceding history matter, since we are concerned about the ratio, in which that offer cancels out.

The optimal stopping probability balances the tradeoff outlined above and does not depend on the history of draws or the set of available alternatives.

3.2 Partial information about environments

Our concept of the value of information compares what one could achieve if she had complete information relative to what one can achieve currently. Here we would like to point out that when one knows more, the relative value of knowing the true environment need not decrease.

I.i.d. environments. The value of information does not change if one knows that she faces an i.i.d. environment. The *i.i.d. class*, denoted by $\mathcal{E}_{iid}$, contains every environment in $\mathcal{E}$ whose offers are independently and identically distributed. In fact, not only the value of information, but also the rule that attains it remain unchanged under the i.i.d. class of environments.

Proposition 1. *Consider the class of i.i.d. environments. Search rule p^* is optimal and the value of information on that class is R^*.*

The proof is in the Appendix.

A known number of offers, n. We show that when the number of offers n is large, then it is better not knowing it.

Consider the two extreme examples, $\alpha = 1$ and $\alpha = 0$.

No recall, $\alpha = 0$. When n is known, the optimal rule dictates to stop in each period t with probability $\frac{1}{n-t+1}$ and delivers the competitive ratio n at $t = 0$. This result is a consequence of our requirement to evaluate the competitive ratio in every round of search.

is straightforward by backward induction in t starting from $t = n$, and it holds for every discount factor $\delta \in (0, 1)$.

Remarkably, when n can be large, with unknown n the individual can achieve a better the competitive ratio than if n is known,

$$\left(1 + \sqrt{\frac{1}{1-\delta}}\right)^2 < n \text{ for large } n.$$

Free recall, $\alpha = 1$. When n is known, the *optimal rule* dictates to stop in each period t with probability $\frac{1-\delta}{1+(n-t)(1-\delta)}$ and delivers the competitive ratio $1 + n(1-\delta)$ at $t = 0$. This result is straightforward by backward induction in t starting from $t = n$.

Similar to the case of no recall, when n can be large, in our model with unknown n the individual can achieve a better the competitive ratio than if n is known,

$$4 < 1 + n(1-\delta) \quad \text{for large } n.$$

3.3 Ex-ante vs. time-consistent competitive ratio

Evaluation of the competitive ratio at all rounds of the search, after all histories, is a novelty in this paper. This is important in the settings where one cannot or is unwilling to commit at the outset and is free to make choices at every stage. In contrast, in the literature on the secretary problems the individual only cares about her payoff ex-ante and has full commitment power to carry out any strategy designed at the outset.

But ex-ante optimal strategies of secretary problems are generally very bad at later stages. For example, the secretary problem with no recall has the ex-ante optimal strategy that stipulates to reject first t^* secretaries, and then continue until a secretary better than those in the initial t^* stages appears, where t^* is some cutoff that depends on the parameters. However, from the perspective of any period $t > 0$, this is very bad strategy. For example, for $t < t^*$ it stipulates to continue, which is the worst possible action in the environment where the value at t is best and all further values will be very low.

The power of commitment at the ex-ante stage is substantial. Let us compare the results for the secretary problem with commitment and and our problem without commitment, assuming n is known.

Consider, for instance, no recall, $\alpha = 0$. The *ex-ante optimal rule* dictates to skip the first $t^*(\delta) \approx \frac{1}{1-\delta}$ offers, and then accept the first offer in period $t > t^*(\delta)$ that exceeds all previous offers, as follows from the solution of the corresponding secretary problem with discounting and without recall (Rasmussen and Pliska 1975). The ex-ante competitive ratio is approximately

$$\delta^{-\frac{1}{1-\delta}} = e + \frac{e(1-\delta)}{2} + \left((1-\delta)^2\right)$$

but after some histories the competitive ratio of that strategy is infinite.

In contrast, the optimal strategy in this paper delivers the same competitive ratio after all histories

$$\left(1+\sqrt{\frac{1}{1-\delta}}\right)^2 = 1+\frac{1}{1-\delta}+\frac{2}{\sqrt{1-\delta}}+O\left(1-\delta\right)$$

It is smaller than the ex-ante competitive ratio, which is the cost of the requirement of consistency across all search rounds.

n at $t=0$, as we found earlier.

Similarly, for free recall, $\alpha=1$, the *ex-ante optimal rule* dictates to receive $t^{**}(\delta)\approx\frac{1}{-\ln\delta}$ offers, and then accept the best one among the received. The ex-ante competitive ratio is approximately

$$\frac{(e-1)(-\ln\delta)}{1-\delta} = e-1+\frac{(e-1)(1-\delta)}{2}+O\left((1-\delta)^2\right)$$

The optimal strategy in this paper delivers the same competitive ratio after all histories equal to 4, which is greater than the ex-ante competitive ratio.

Thus, under both extremes, $\alpha=0$ and $\alpha=1$, the ex-ante best competitive ratio is smaller than our best competitive ratio, which is the cost of the requirement of consistency across all search rounds.

4 Proof of Theorem 1

The proof is divided into four steps. First, we show that w.l.o.g. we can restrict attention to rules that are independent of the history, h_t. Second, we consider a subclass of environments and show that on that subclass w.l.o.g. we can restrict attention to rules that are constant (independent of t, h_t, and Y_t). Third, we find the optimal constant rule for that subclass of environments, and thus derive a lower bound on the value of information. Finally, we show that the optimal rule does not change and the value of information remains the same if we consider the general the class of environments.

Step 1. History independence. Fix any history $h_t = (x_1, ..., x_t)$, and suppose that there remain offers to draw, $n > t$. The set of environments $\mathcal{E}(h_t)$ that the individual faces from stage $t+1$ on is obtained by elimination from $\mathcal{E}$ the environments whose distributions do not support the historical values $x_1, ..., x_t$. (Note that the set of possible numbers of remaining alternatives after period t is the same as at the start, since $\{n - t : n \in \mathbb{N}, n > t\} = \mathbb{N}$.) Consequently, $Closure(\mathcal{E}(h_t)) = \mathcal{E}$. The supremum of the performance ratio is the same whether the domain is $\mathcal{E}(h_t)$ or $Closure(\mathcal{E}(h_t)) = \mathcal{E}$.

We thus have obtained that the set of environments is irreducible after any history of observations, hence history h_t is payoff irrelevant. We show now that the optimal performance ratio at any stage is history independent (but it may depend on the set of available offers).

Let p be an optimal decision rule. Consider two histories, $h_s = (x_1, ..., x_s)$ and $h_t = (\hat{x}_1, ..., \hat{x}_t)$ with the same set of available offers, $Y_s = Y_t = \hat{Y}$. Suppose that $R^p(h_s, \hat{Y}) > R^p(h_t, \hat{Y})$. Then we can construct an improvement over p at history h_s by decision rule $\hat{p}$ identical to p in all subgames except the one following history h_s. In that subgame we define $\hat{p}$ equal to p in the subgame following history h_t. This contradicts the assumption of optimality of p.

In what follows we restrict attention to history independent rules and drop the reference to h_t from all notations.

Step 2. Stationarity under restricted class of environments. Denote by $\mathcal{E}(x)$ the class of environments whose values are at least y:

$$\tilde{\mathcal{E}}(y) = \{(n, F) \in \mathcal{E} : supp(F) \subset [y, \infty)^n\}.$$

For every period t and every best offer y_t, we are now concerned with the maximal performance ratio $\frac{V_t(E,Y_t)}{U_t^p(E,Y_t)}$ across all environments in $\tilde{\mathcal{E}}(y_t)$. Maximization on a smaller set makes the ratio smaller. The adjusted competitive ratio is defined as

$$\tilde{R}^p = \sup_{t,\ Y_t} \left\{ \sup_{E \in \tilde{\mathcal{E}}(y_t)} \frac{V_t(E, Y_t)}{U_t^p(E, Y_t)} \right\},$$

Notice that with this restriction on the class of environments, for every round t and every best value t, the subproblem from that round on is identical to the original problem in which all payoffs are scaled by the factor y_t/x_0. But scaling does not affect the competitive ratio.

Hence, if a rule $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \tilde{p}_2, ...)$ maximizes the competitive ratio, exactly the same rule should maximize the competitive ratio in every subproblem, from every period t on and for every Y_t, $(p_t, p_{t+1}, p_{t+2}, ...) = (\tilde{p}_0, \tilde{p}_1, \tilde{p}_2, ...)$. This implies that w.l.o.g. we can restrict attention to the rules with a constant probability of stopping, $q \in [0, 1]$, independent of time and available offers.

Step 3. Optimization by constant search rules. We now restrict attention to constant search rules and find the one that minimizes the competitive ratio evaluated under even smaller set of environments.

Denote by $E_{n,y,z}$ the environment with the pool of n deterministic values, a single z and $n-1$ instances of y, where $y \geq z$. Let $\mathcal{E}^*(y)$ be the set of such environments with a given y,

$$\mathcal{E}^*(y) = \{E \in \tilde{\mathcal{E}}(y) : E = E_{n,y,z} \text{ for some } n \in \mathbb{N} \text{ and } z \geq y\}.$$

Notice that $\mathcal{E}^*(y) \subset \tilde{\mathcal{E}}(y)$.

Let $q \in [0, 1]$ be a constant probability of stopping. Fix a period t, the best offer $y = y_t$, and an environment $E_{k,y,z}$, where k is the number of offers left.

The first-best rule of the one who knows $E_{k,y,z}$ is either stop immediately and get y, or wait for z. In the latter case, z may appear in each round $t+1, ..., t+k$ equally likely, which gives expected payoff

$$\tfrac{1}{k}(\delta + \delta^2 + ... + \delta^k) = \frac{\delta(1-\delta^k)}{k(1-\delta)}.$$

Thus,

$$V(E_{k,y,z}, y) = \max\left\{y, \frac{\delta(1-\delta^k)}{k(1-\delta)}\right\}.$$

The expected payoff of constant rule q is given by

$$U^q(E_{k,y,z}, y) = qy + (1-q)\delta\left(\tfrac{1}{k}U^q(E_{k-1,y,y}, z) + \tfrac{k-1}{k}U^q(E_{k-1,y,z}, y)\right).$$

Value z appears with probability $\frac{1}{n}$ in the next period. In this case the individual faces the environment of $n-1$ instances of y in the future. Her continuation payoff depends on her ability to recall z:

$$\begin{aligned}
U^q(E_{n-1,y,y}, z) &= qz\big(1 + (1-q)\delta\alpha + (1-q)^2\delta^2\alpha^2 + ... + (1-q)^{k-1}\delta^{k-1}\alpha^{k-1}\big) \\
&\quad + c_1(\alpha, \delta)y \\
&= \frac{qz}{1-\alpha\delta(1-q)}\left(1 - \alpha^k\delta^k(1-q)^k\right) + c_1(\alpha, \delta)y,
\end{aligned}$$

where $c_1(\alpha, \delta)$ is a constant independent of the environment.

In the case of $y \geq \frac{\delta(1-\delta^k)}{k(1-\delta)}$, the ratio is

$$\frac{V(E_{k,y,z}, y)}{U^q(E_{k,y,z}, y)} \leq \frac{y}{qy} = \frac{1}{q}.$$

In the case of $y < \frac{\delta(1-\delta^k)}{k(1-\delta)}$, the ratio is

$$\frac{V(E_{k,y,z}, y)}{U^q(E_{k,y,z}, y)} = \frac{\frac{\delta(1-\delta^k)}{k(1-\delta)}}{qy + (1-q)\delta\left(\frac{1}{k}U^q(E_{k-1,y,y}, z) + \frac{k-1}{k}U^q(E_{k-1,y,z}, y)\right)}.$$

The above ratio is decreasing in z and k, so taking $z \to \infty$ and $k \to \infty$ we obtain

$$\frac{V(E_{k,y,z}, y)}{U^q(E_{k,y,z}, y)} \geq \frac{\frac{\delta}{1-\delta}}{(1-q)\delta\frac{q}{(1-\delta(1-q))(1-\alpha\delta(1-q))}} = \frac{(1-\delta(1-q))(1-\alpha\delta(1-q))}{(1-\delta)(1-q)q}.$$

Maximizing the above expression w.r.t. q, we find the optimal q^*,

$$q^* = 1 - \frac{1-\sqrt{(1-\delta)(1-\alpha\delta)}}{1-(1-\delta)(1-\alpha\delta)}$$

and the competitive ratio

$$\bar{R}^* = \left(1 + \sqrt{\frac{1-\alpha\delta}{1-\delta}}\right)^2.$$

Since we have been maximizing the ratio on a small subclass of environments, it is a lower bound on the value of information on the general class,

$$R^* \geq \bar{R}^*.$$

Step 4. Competitive ratio on the general class of environments. We now establish the upper bound on R^*. We restrict the individual to constant search rules and argue that the environments described in Step 3 are essentially the worst-case environments against constant rules.

Consider constant rule q^*. First, the ratio will always be weakly greater if one who knows the environment is informed about realized values of offers in the pool.

Next, consider two events. In one event the individual accepts the same offer as she would have if she knew the environment. Replacing any such offer by the top offer, called z, will make the ratio weakly bigger. In the second event, the individual accepts an offer that she would have rejected if she knew the environment. Replacing any such offer by zero will make the ratio weakly bigger.

Finally, if the environment has k top offers with value z, replacing them by $k-1$ zeros and a single offer with value kz will increase the ratio, since the uninformed individual has even smaller chance to get it that payoff.

We thus conclude that the worst-case environments have form $E_{n,0,z}$. Moreover, similar to the previous step, the ratio is increasing in n and z. When $z \to infty$, $y/z \to 0$ for any fixed y, and hence $E_{n,0,z}$ is equivalent to $E_{n,y,z}$ when z is large.

It follows that the competitive ratio of the constant rule q^* is equal to the lower bound $\bar{R}^*$ found in Step 3. In other words, the lower bound is attained by constant rule q^*, thus it is the value of information, R^*.

Appendix

Proof of Proposition 1 Proposition 1 is proved in the same way as Theorem 1, except we consider the class of i.i.d. environments. Steps 1, 2, and 4 of the proof of Theorem 1 are analogous. For Step 3, the calculation of the best constant rule is repeated, where an environment with a single value z and $k-1$ instances of value y are replaced by i.i.d. the environment with i.i.d. between z and y, with probabilities σ and $1-\sigma$, respectively. Here we take $z \to \infty$ and $\sigma \to 0$ and n large in such a way that the one who knows the environment prefers to wait for z, but small σ and large n means that the uninformed individual who stops with a fixed probability in every period is unlikely to reach z. Interestingly, the limit expression is the same as in Step 3 of the proof of Theorem 1, which leads to the same value of information.

References

Babaioff, Moshe, Nicole Immorlica, David Kempe, and Robert Kleinberg, "A knapsack secretary problem with applications," in Moses Charikar, Klaus Jansen, Omer Reingold, and José D. P. Rolim, eds., *Proceedings of the 10th International Workshop on Approximation, Randomization, and Combinatorial Optimization (APPROX 2007)*, 2007, pp. 16–28.

Bergemann, Dirk and Karl Schlag, "Robust monopoly pricing," *Journal of Economic Theory*, 2011, *146*, 2527–2543.

Ferguson, Thomas S., "Who Solved the Secretary Problem?," *Statistical Science*, 1989, *4* (3), 282–289.

Fox, J. H. and L. G. Marnie, "In Martin Gardner's column: Mathematical games," *Scientific American*, 1960, *202* (2), 150–153.

Freeman, P. R., "The Secretary Problem and Its Extensions: A Review," *International Statistical Review*, 1983, *51*, 189–206.

Mucci, Anthony G., "On a Class of Secretary Problems," *The Annals of Probability*, 1973, *1* (3), 417–427.

Presman, E. L. and I. M. Sonin, "The Best Choice Problem for a Random Number of Objects," *Theory of Probability & Its Applications*, 1973, *17*, 657–668.

Rasmussen, Willis T. and Stanley R. Pliska, "Choosing the maximum from a sequence with a discount function," *Applied Mathematics and Optimization*, 1975, *2*, 279–289.

Schlag, Karl and Andriy Zapechelnyuk, "Robust Sequential Search," 2016. Mimeo.

Sleator, D. and R. Tarjan, "Amortized efficiency of list update and paging rules," *Communications of the ACM*, 1985, *28*, 202–208.

Yang, Mark C. K., "Recognizing the Maximum of a Random Sequence Based on Relative Rank with Backward Solicitation," *Journal of Applied Probability*, 1974, *11* (3), 504–512.