

Secretary problems with competing employers

Nicole Immorlica¹, Robert Kleinberg^{2*}, and Mohammad Mahdian¹

¹ Microsoft Research, One Microsoft Way, Redmond, WA.
{nickle,mahdian}@microsoft.com

² UC Berkeley Computer Science Division and Cornell University Computer Science
Department. rdk@cs.cornell.edu.

Abstract. In many decentralized labor markets, job candidates are offered positions at very early stages in the hiring process. It has been argued that these early offers are an effect of the competition between employers for the best candidate. This work studies the timing of offers in a theoretical model based on the classical secretary problem. We consider a secretary problem with multiple employers and study the equilibria of the induced game. Our results confirm the observation of early offers in labor markets: for several classes of strategies based on optimal stopping theory, as the number of employers grows, the timing of the earliest offer decreases.

1 Introduction

An essential feature of many modern markets, particularly networked markets, is that they are *online*: information about agents, goods, and outcomes is revealed over time, and the agents must make irrevocable decisions before all of the information is revealed. A powerful tool for analyzing such scenarios is *optimal stopping theory*, the theory of problems which require optimizing an objective function over the space of stopping rules for a stochastic process. By combining optimal stopping theory with game theory, we can model the actions of rational agents applying competing stopping rules in an online market.

Perhaps the best-known optimal stopping problem is the *secretary problem*, also known as the *best-choice problem* or *marriage problem*. This problem was introduced in the 1960's as a model for studying online selection processes in the presence of a randomly ordered input. In the most basic version of the secretary problem, a decision-maker observes a sequence of elements of a totally-ordered set, presented in random order. At any time the decision-maker may stop the sequence and select the most recently presented element, with the objective of maximizing the probability of selecting the minimum element of the entire set. Dynkin [1] determined the optimal stopping rule for this problem and proved that its probability of success approaches $1/e$ as the size of the set tends to infinity. A myriad of subsequent papers have extended the problem by varying the objective function, varying the information available to the decision-maker, allowing for multiple choices, and so on, e.g. [2–4].

* Supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

Another line of work (see [5] for a survey) extends the secretary problem by studying scenarios in which two players compete to select the minimum element of a randomly-ordered sequence. If the original secretary problem can be thought of as the decision problem faced by an employer interviewing candidates for a job, then this variant should be thought of as the strategic problem faced by two employers interviewing the same sequence of candidates, when only one of them can hire the best candidate. Papers in this area differ in their assumptions about the game’s payoff structure and about the ways in which conflicts (both players making an offer at the same time) are resolved. For example, Dynkin [6] proves a minimax theorem and characterizes the game value for a zero-sum version of the game in which conflicts are avoided by stipulating that in each step of the game, only one player is allowed to make an offer. Szajowski studies zero-sum stopping games with a tie-breaking rule which always gives priority to Player 1 [7] or which gives priority to a player chosen by a (possibly biased) random lottery [8]. A non-zero-sum version of the two-player game with unbiased random tie-breaking was studied by Fushimi [9], who observed that although the game is symmetric, the only pure Nash equilibria are asymmetric. Non-zero-sum stopping games in which Player 1 always receives priority [10, 11] or in which priority is granted by a (possibly biased) random lottery [12] have also been studied. As far as we are aware, all of the prior work studies equilibria of the resulting games when just two employers compete.

In this paper, we first present our own variant of the two-employer game, providing a new derivation of the resulting symmetric two-player mixed-Nash equilibrium, and then proceed to study k -employer extensions of our game. One of the properties of any strategy for the employer game is its *threshold time*, defined as the largest fraction τ such that the strategy is guaranteed not to make an offer to any of the first τ candidates, no matter what the player observes nor what the other players do. We show that in any pure Nash equilibrium of the k -employer game, at least one of the players has a threshold time $\tau \leq 2/k$. This fact follows from a theorem which generalizes a striking feature of the one-player secretary problem: in any pure Nash equilibrium, the player with the earliest threshold time r_1 has probability exactly r_1 of winning the game. (In the one-player case, this specializes to the familiar fact that $1/e$ is both the probability of winning and the optimal threshold value.)

Next we consider a more realistic version of the game, in which the players are allowed to use *adaptive* strategies which base their decisions on the opponent’s past actions as well as the public information revealed thus far. This defines a multi-player stochastic game. We describe the unique subgame perfect mixed Nash equilibrium of this game as the solution of a dynamic program. Using properties of the dynamic program, we prove that the timing of the first offer converges to zero as the number of players tends to infinity. More precisely, every player’s threshold time is at most $1/k$.

In many labor markets, competition between employers often leads to an “unraveling” effect: employers wishing to attract the best candidates make early offers with short expirations. Such effects are quite pronounced in many markets.

In the market for law clerks, offers are made to candidates as early as the second year of law school, a full two years before graduations [13]. For a survey of markets which exhibit unraveling effects, see [14]. Our results can be interpreted as a theoretical justification for these unraveling effects. Namely, as the number of employers grows, our results confirm that the timing of the earliest offer in an equilibrium decreases.

2 Preliminaries

In this section we define a discrete-time and a continuous-time version of the game. The discrete-time version is conceptually simpler whereas the continuous-time version is more analytically tractable.

In the discrete-time game, we are given a totally-ordered set $\mathcal{U} = \{x_1 \prec x_2 \prec \dots \prec x_n\}$, (representing the secretaries in decreasing order of value) and a set of $k \geq 1$ players (representing the employers). A random bijection $Z : \mathcal{U} \rightarrow [n]$ is chosen (representing the order in which the secretaries will be interviewed), but Z is not initially revealed to the players. (Here and throughout this paper, when m is a natural number we use $[m]$ to denote the set $\{1, 2, \dots, m\}$.) As the game proceeds, each player is in a state which is either *active* or *inactive*; all players are initially active. At time $t = 1, 2, \dots, n$, the relative ordering of the elements $Z^{-1}(1), Z^{-1}(2), \dots, Z^{-1}(t)$ is revealed to the players. Each active player then chooses an action from the set $\{\text{O}, \text{P}\}$, whose elements are referred to as “offer” and “pass”, respectively. If one or more active players chooses to offer, then one of these players (chosen uniformly at random) receives the element $x = Z^{-1}(t)$ and becomes inactive; this player is denoted by $\chi(x)$. The others remain active. If all active players choose to pass at time t , then all of them remain active and no player receives x . Each player is informed of the actions of all other players and of the identity of player $\chi(x)$, if defined. At the end of the game, all players receive a payoff of 0 except for $\chi(x_1)$ (if defined) who receives a payoff of 1.

The continuous-time variant of the game intuitively captures the limit of the discrete game as n tends to infinity. In most of this paper, we work on the continuous model, as this model hides details like integrality in our computations, hence making the computations cleaner, while still capturing the main ideas. It is not hard to generalize our results in the continuous model to the discrete model by incurring an additive error of $o(1)$; the details of this generalization is omitted from this extended abstract.

We now give an informal definition of the continuous-time model. The rigorous definition requires technical details and is deferred to the full version of the paper. In the continuous-time game, $\mathcal{U} = \mathbb{N}$ and \prec denotes the usual ordering of \mathbb{N} (hence the best element x_1 is 1). Each element x has an arrival time $Z(x)$ picked independently and uniformly at random from $[0, 1]$. Each player i has a (possibly randomized) strategy S_i , which at any time $t \in [0, 1]$ specifies whether or not player i makes an offer to the element arriving at time t (if any).³ In

³ Note that we can ignore zero-probability events such as the arrival of two elements at the exact same time.

general, $S_i(t)$ can depend on all the information revealed before time t (i.e., the ordering of the elements that have arrived by time t , and the offer/pass decisions made by other players before time t). We call such strategies *adaptive*. A simpler class of strategies, which correspond to stopping rules, are *non-adaptive* strategies and are defined as follows. For each $r \in [0, 1]$, the *non-adaptive strategy with threshold time r* is the strategy which makes an offer to every element observed after time r which outranks the best element arriving before time r , until it receives one of these elements and passes on all subsequent ones.

3 Analysis of non-adaptive strategies

3.1 Two players

We begin our discussion with a warmup involving the computation of equilibria in the two-player game. We will work in the continuous-time model with $\mathcal{U} = \mathbb{N}$.

Pure Strategies Our first calculation characterizes the set of pure strategy Nash equilibria. Our analysis is quite similar to that of Fushimi [9]. The difference in the results stems from the fact that our process continues until time $t = n$ unless both players make successful offers, whereas in Fushimi's model the process stops if both employers make an offer to a single element simultaneously.

Assume that players 1 and 2 use the non-adaptive pure strategies r, s , respectively. Let Y denote the following random subset of \mathcal{U} :

$$Y = \{y \in \mathcal{U} \mid Z(y) < Z(x_1)\},$$

and let $y_1 \prec y_2$ denote the two minimum elements of Y . We can then compute the resulting expected payoff to player 2 as follows.

Case 1 ($s > r$): Then player 2 wins if either

1. Player 1's earlier threshold caused him to make a sub-optimal offer, and player 2 made an offer to the best element: $Z(x_1) > s$ and $r < Z(y_1) \leq s$.
2. Both players made an offer to the best element, and player 2 won the coin toss: $Z(x_1) > s$ and $Z(y_1) \leq r$ and $\chi(x_1) = 2$.
3. Both players made offers to the same sub-optimal element, but player 2 lost the coin toss and then proceeded to make an offer to the best element: $Z(x_1) > s$ and $Z(y_1) > s$ and $Z(y_2) \leq r$ and $\chi(y_1) = 1$.

Conditioned on the arrival time $t = Z(x_1)$ of the best element, the random variables $Z(y_1), Z(y_2)$ are independently uniformly distributed in $[0, t]$. Therefore the probabilities of the three events listed above are respectively $\int_s^1 \frac{s-r}{t} dt$, $\frac{1}{2} \int_s^1 \frac{r}{t} dt$, and $\frac{1}{2} \int_s^1 \left(\frac{t-s}{t}\right) \left(\frac{r}{t}\right) dt$, and their sum integrates to $s \ln\left(\frac{1}{s}\right) - \frac{r(1-s)}{2}$.

Case 2 ($s \leq r$): Then player 2 wins if either

1. The best element arrives between the two thresholds and player 2 makes an offer: $s < Z(x_1) \leq r$ and $Z(y_1) \leq s$.

2. Both players make an offer to the best element, and player 2 wins the coin toss: $Z(x_1) > r$ and $Z(y_1) \leq s$ and $\chi(x_1) = 2$.
3. Both players made offers to the same sub-optimal element, but player 2 lost the coin toss and then proceeded to make an offer to the best element: $Z(x_1) > r$ and $Z(y_1) > r$ and $Z(y_2) \leq s$ and $\chi(y_1) = 1$.

The probabilities of the three events listed above are respectively $\int_s^r \frac{s}{t} dt$, $\frac{1}{2} \int_r^1 \frac{s}{t} dt$, and $\frac{1}{2} \int_r^1 \left(\frac{t-r}{t}\right) \left(\frac{s}{t}\right) dt$, and their sum integrates to $s \ln\left(\frac{1}{s}\right) - \frac{s(1-r)}{2}$.

Hence, when players 1 and 2 play the pure strategies r, s , respectively, the expected payoff to player 2 is

$$f(s) = \begin{cases} s \ln(1/s) - s(1-r)/2 & \text{if } s \leq r \\ s \ln(1/s) - r(1-s)/2 & \text{if } s \geq r \end{cases} \quad (1)$$

and the derivative of the expected payoff is

$$f'(s) = \begin{cases} \ln(1/s) - 3/2 + r/2 & \text{if } s < r \\ \ln(1/s) - 1 + r/2 & \text{if } s > r \end{cases} \quad (2)$$

Let $s_-(r), s_+(r)$ denote the best responses to r in the intervals $[0, r]$ and $[r, 1]$, respectively. It follows from (2) that

$$s_-(r) = \begin{cases} e^{r/2-3/2} & \text{if } \ln(1/r) - 3/2 + r/2 < 0 \\ r & \text{otherwise} \end{cases} \quad (3)$$

$$s_+(r) = \begin{cases} e^{r/2-1} & \text{if } \ln(1/r) - 1 + r/2 > 0 \\ r & \text{otherwise} \end{cases} \quad (4)$$

By symmetry, player 1's best response function is identical. It follows easily from (3) and (4) that there is a pure Nash equilibrium (r, s) where $r = 0.27557\dots$ satisfies $2 \ln(r) + 3 = \exp\left(\frac{r}{2} - 1\right)$ and $s = 0.42222\dots$ satisfies $2 \ln(s) + 2 = \exp\left(\frac{s}{2} - \frac{3}{2}\right)$. In fact, (3) and (4) imply that the only two pure Nash equilibria are (r, s) and (s, r) .

Mixed Strategies The two-player game with non-adaptive strategies also has a symmetric mixed Nash equilibrium which we may compute explicitly. If player 1's choice of r is a random variable with density function $\nu(r)$ then we find, using (1), that player 2's expected payoff from playing strategy s is:

$$f(s) = \int_0^s \left[s \ln\left(\frac{1}{s}\right) - \frac{s(1-r)}{2} \right] \nu(r) dr + \int_s^1 \left[s \ln\left(\frac{1}{s}\right) - \frac{r(1-s)}{2} \right] \nu(r) dr,$$

and from (2) we obtain:

$$f'(s) = \ln\left(\frac{1}{s}\right) - 1 + \mathbb{E}\left(\frac{r}{2}\right) - \frac{1}{2} \Pr(r \geq s). \quad (5)$$

Let us assume that s has positive probability density in an interval (s_0, s_1) and zero probability of lying outside $[s_0, s_1]$. Then every $s \in [s_0, s_1]$ is a best response to player 1's mixed strategy, so $f'(s) = 0$ for $s \in [s_0, s_1]$. Since we are assuming a symmetric mixed Nash equilibrium, r also has zero probability of lying outside $[s_0, s_1]$, i.e. $\Pr(r \geq s_0) = 1$, $\Pr(r \geq s_1) = 0$. This implies, using (5) and the fact that $f'(s) = 0$ for $s \in [s_0, s_1]$, that:

$$\begin{aligned}\ln(1/s_0) - 3/2 + \mathbb{E}(r/2) &= 0 \\ \ln(1/s_1) - 1 + \mathbb{E}(r/2) &= 0 \\ \ln(1/s_0) - \ln(1/s_1) &= 1/2 \\ s_1 &= \sqrt{e} \cdot s_0.\end{aligned}$$

Taking the derivative of (5) we obtain:

$$f''(s) = -\frac{1}{s} + \frac{1}{2}\nu(s) = 0 \quad \text{for } s \in [s_0, s_1],$$

which implies $\nu(s) = 2/s$ for $s \in [s_0, s_1]$. Hence

$$\mathbb{E}(r/2) = \mathbb{E}(s/2) = \int_{s_0}^{s_1} (s/2)\nu(s) ds = \int_{s_0}^{s_1} ds = s_1 - s_0 = (\sqrt{e} - 1)s_0.$$

Recalling that $\ln(1/s_0) - 3/2 + \mathbb{E}(r/2) = 0$, we find that s_0 satisfies the equation

$$(\sqrt{e} - 1)s_0 = 3/2 + \ln(s_0),$$

i.e. $s_0 = 0.265\dots$, $s_1 = 0.437\dots$. Finally, we must verify that this is indeed a mixed Nash equilibrium by checking that the expected payoff function $f(s)$ is maximized when $s \in [s_0, s_1]$. To do so, it suffices to prove that $f'(s) > 0$ when $s < s_0$ and that $f'(s) < 0$ when $s > s_1$:

$$\begin{aligned}\boxed{s < s_0} \quad f'(s) &= \ln(1/s) - 3/2 + \mathbb{E}(r/2) > \ln(1/s_0) - 3/2 + \mathbb{E}(r/2) = 0, \\ \boxed{s > s_1} \quad f'(s) &= \ln(1/s) - 1 + \mathbb{E}(r/2) < \ln(1/s_1) - 1 + \mathbb{E}(r/2) = 0.\end{aligned}$$

3.2 Multiple players

In the previous section we saw that the two-player game has a pure Nash equilibrium $(r, s) = (0.27557\dots, 0.42222\dots)$. One striking feature of this equilibrium is that player 1's probability of winning is

$$f(r) = r \ln(1/r) - r(1 - s)/2 = 0.27557\dots$$

which is exactly equal to r . Recall that for the one-player game (i.e. the original secretary problem) the optimal strategy sets its threshold at time $1/e$ and has probability $1/e$ of winning. We begin this section with a theorem which shows that it is not coincidental that player 1's optimal threshold time and her probability of winning are exactly equal in both the one-player and two-player games. This phenomenon holds for every pure Nash equilibrium of the k -player game, for every k .

Theorem 1. *If (r_1, \dots, r_k) is a pure Nash equilibrium of the k -player non-adaptive game and $r_1 \leq r_2 \leq \dots \leq r_k$, then $\Pr(\text{player 1 wins}) = r_1$.*

Proof. Fix the values of r_2, r_3, \dots, r_k and let $f(r)$ denote the probability that player 1 wins when the players use strategies $(r, r_2, r_3, \dots, r_k)$. We will prove that there is a constant C , depending only on r_2, \dots, r_k , such that $f(r) = r \ln(1/r) + Cr$ when $r \in (0, r_2]$. Since $r_1 = \arg \max f(r)$ and $r_1 \in (0, r_2]$ we have

$$0 = f'(r_1) = \ln(1/r_1) - 1 + C = \frac{f(r_1)}{r_1} - 1,$$

from which we conclude that $f(r_1) = r_1$ as claimed.

It remains to prove that $f(r) = r \ln(1/r) + Cr$ when $r \in (0, r_2]$. The probability that player 1 wins before time r_2 is

$$\int_r^{r_2} \frac{r}{t} dt = r \ln(1/r) - r \ln(1/r_2).$$

To compute the probability that player 1 wins after time r_2 , let $Y = \{y \in \mathcal{U} \mid Z(y) < Z(x_1)\}$, and let $y_1 \prec y_2 \prec y_3 \prec \dots$ be the elements of Y in sorted order. Note that, conditional on $Z(x_1) = t$, the random variables $\{Z(y)\}_{y \in Y}$ are independent and uniformly distributed in $[0, t]$. Let $A = \min\{a \mid Z(y_a) \leq r\}$, and let $u_a = Z(y_a)$ for $a = 1, 2, \dots, A$. If player 1 wins after time r_2 then it must be the case that $Z(y_a) > r_2$ for $a < A$. Let

$$g(u_1, u_2, \dots, u_A) = \Pr(\chi(x_1) = 1 \mid r, r_2, \dots, r_k, u_1, u_2, \dots, u_A).$$

Note that the value of g depends only on the relative ordering of the numbers in the set $S = \{r, r_2, \dots, r_k, u_1, u_2, \dots, u_A\}$. In particular, g is constant as u_A varies over the range $[0, r]$ because u_A, r are always the two smallest numbers in S . Now, letting \mathcal{E} denote the event that player 1 wins after time r_2 ,

$$\begin{aligned} \Pr(\mathcal{E}) &= \sum_{a=1}^{\infty} \Pr((A = a) \wedge \mathcal{E}) \\ &= \sum_{a=1}^{\infty} \int_{r_2}^1 \left[\int_0^r \left(\int_{r_2}^1 \cdots \int_{r_2}^1 g(u_1, \dots, u_a) t^{-a} du_1 \dots du_{a-1} \right) du_a \right] dt \\ &= r \sum_{a=1}^{\infty} \int_{r_2}^1 \left[\int_{r_2}^1 \cdots \int_{r_2}^1 g(u_1, \dots, u_{a-1}, 0) t^{-a} du_1 \dots du_{a-1} \right] dt \\ &= C' r, \end{aligned}$$

where C' denotes the sum on the penultimate line of the equation above. Thus $f(r) = r \ln(1/r) + Cr$, with $C = C' - \ln(1/r_2)$.

Lemma 1. *In any pure Nash equilibrium of the k -player non-adaptive game, no player receives an expected payoff which is more than twice another player's expected payoff.*

Proof. Let p_i denote the expected payoff of player i . If $p_j > 2p_i$, then player i can deviate from the equilibrium by playing r_j instead of r_i . We will prove that this deviation yields an expected payoff of at least $p_j/2$ for player i , contradicting the assumption that r_1, \dots, r_k is a Nash equilibrium.

To prove that player i achieves an expected payoff of at least $p_j/2$ by changing her strategy to r_j , note first that players i and j have equal expected payoffs when they both play r_j . Now consider the change in player j 's expected payoff in a series of two steps. First, player i changes her strategy from r_i to 1. (This is equivalent to player i leaving the game, since a player with a threshold time of 1 never makes an offer.) This change does not decrease p_j . Second, player i changes her strategy from 1 to r_j . For every time $t \geq r_j$, this increases the number of active players at time t by at most one, so it decreases $\Pr(\text{player } j \text{ wins at time } t)$ by at most a factor of 2. Thus, player j 's expected payoff after this second change is at least $p_j/2$. This is equal to player i 's expected payoff when she deviates by playing r_j , which concludes the proof of the lemma.

The following corollary confirms the observation of early offers in labor markets: it shows that in any pure Nash equilibrium of the k -player game, the earliest threshold time is $O(2/k)$. (An easy consequence of this is that the timing of the first offer converges to zero almost surely as $k \rightarrow \infty$.)

Corollary 1. *In any pure Nash equilibrium of the k -players non-adaptive game, if the players are numbered in order of increasing threshold times $r_1 \leq r_2 \leq \dots \leq r_k$, then $r_1 \leq 2/k$.*

Proof. As before, let p_i denote the expected payoff of player i , for $i = 1, 2, \dots, k$. Note that $\sum_i p_i \leq 1$ since the combined payoff of all players is at most 1. By Lemma 1 we have $p_i \geq p_1/2$ for all i . Hence $kp_1/2 \leq \sum_i p_i \leq 1$ which implies that $p_1 \leq 2/k$. By Theorem 1 we have $r_1 = p_1$, hence $r_1 \leq 2/k$.

4 Adaptive strategies

When players are allowed to use adaptive strategies, it will be more convenient to adopt the discrete-time model of the game. This can be described as a stochastic game with state space $[n] \times [k]$. The interpretation of state (t, j) for $t \leq n$ is that there are j players active at time t , and an element x arrives at time t which is the best element observed so far. (More formally, there is an element $x \in \mathcal{U}$ with $Z(x) = t$ such that $x \prec y$ for all y with $Z(y) < t$.)

Proposition 1. *The adaptive k -player game has a unique symmetric subgame perfect equilibrium. This equilibrium can be described as follows: for each state $(t, j) \in [n] \times [k]$, there exist numbers $p(t, j), q(t, j)$ such that the equilibrium strategy of an active player in state (t, j) is to play \mathbf{O} with probability $p(t, j)$, \mathbf{P} with probability $q(t, j)$, regardless of the prior history of the game. Let $v(t, j)$ denote the probability that player i wins, given that the game is currently in state (t, j) and i is active. Then the values of $p(t, j), q(t, j), v(t, j)$ are correctly computed by the algorithm in Figure 1.*

```

/* Initialization */
for t = 1, 2, ..., n
    v(t, 0) ← 0; w(t, 0) ← 0
end
for j = 1, ..., k
    p(n, j) ← 1; q(n, j) ← 0; w(n, j) ← 0; v(n, j) ← 1/j
end

/* Dynamic program */
for j = 1, 2, ..., k
    for t = n - 1, n - 2, ..., 1
        w(t, j) ← w(t + 1, j) + v(t + 1, j)/(t2 + t)
        A ← w(t, j - 1) - 1/n
        B ← w(t, j - 1) - w(t, j)
        if 1/n ≥ w(t, j - 1) /* Pure strategy O is optimal */
            q ← 0
        else if 1/n ≤ w(t, j) /* Pure strategy P is optimal */
            q ← 1
        else /* Mixed strategy is optimal */
            r ← the unique solution of 1 + x + x2 + ... + xj-1 = jB/A in the interval (1, ∞)
            q ← 1/r
        p(t, j) ← 1 - q
        q(t, j) ← q
        vO(t, j) ← tw(t, j - 1) - t(A/j)(1 + q + ... + qj-1)
        vP(t, j) ← tw(t, j - 1) - tBqj-1
        v(t, j) ← max{vO(t, j), vP(t, j)}
    end /* for t = n - 1, n - 2, ..., 1 */
end /* for j = 1, 2, ..., k */

```

Fig. 1. Dynamic program to compute the symmetric subgame perfect equilibrium of the adaptive game.

Proof. See Appendix A.

Although the algorithm in Figure 1 is elaborate and does not yield a closed-form formula for the equilibrium strategy, it does enable us to draw some useful qualitative conclusions about this equilibrium. For example, the following theorem again confirms the observation of early offers in labor markets.

Theorem 2. *In the adaptive game with k players, assume all players use the symmetric equilibrium strategy described in Proposition 1. If there are j active players at time $t \geq n/j$ and the element x which arrives at time t is the best element observed so far, then all of the active players make an offer to this element.*

Proof. For all $j \geq 1$ and all $t \in [n]$, we have $v(t, j) \leq 1/j$ because each of the j active players in state (t, j) has probability $v(t, j)$ of winning, and these j events

are mutually exclusive. It follows that

$$w(t, j-1) = \sum_{u=t+1}^n \frac{v(u, j-1)}{u(u-1)} \leq \frac{1}{j-1} \sum_{u=t+1}^n \frac{1}{u(u-1)} = \frac{1}{j-1} \left(\frac{1}{t} - \frac{1}{n} \right).$$

If $t \geq n/j$ then $1/t - 1/n \leq (j-1)/n$, so $w(t, j-1) \leq 1/n$. According to the algorithm in Figure 1, this implies that in equilibrium, all active players play O.

References

1. Dynkin, E.B.: The optimum choice of the instant for stopping a markov process. *Sov. Math. Dokl.* **4** (1963)
2. Ajtai, M., Megiddo, N., Waarts, O.: Improved algorithms and analysis for secretary problems and generalizations. *SIAM J. Discrete Math.* **14**(1) (2001) 1–27
3. Freeman, P.: The secretary problem and its extensions: a review. *Internat. Statist. Rev.* **51**(2) (1983) 189–206
4. Gilbert, J.P., Mosteller, F.: Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61** (1966) 35–73
5. Sakaguchi, M.: Optimal stopping games: A review. *Math. Japonica* **42** (1995) 343–351
6. Dynkin, E.: Game variant of a problem on optimal stopping. *Sov. Math. Dokl.* **10** (1969) 270–274
7. Szajowski, K.: Double stopping by two decision-makers. *Adv. in Appl. Probab.* **25**(2) (1993) 438–452
8. Szajowski, K.: Optimal stopping of a discrete markov process by two decision makers. *SIAM J. Control Optim.* **33**(5) (1995) 1392–1410
9. Fushimi, M.: Secretary problem in a competitive situation. *Journal of the Operations Research Society of Japan* **24**(4) (1981) 350–358
10. Ravindran, G., Szajowski, K.: Nonzero sum game with priority as Dynkin’s game. *Math. Japonica* **37**(3) (1992) 401–413
11. Szajowski, K.: On nonzero sum game with priority in the secretary problem. *Math. Japon.* **37**(3) (1992) 415–426
12. Szajowski, K.: Markov stopping games with random priority. *Z. Oper. Res.* **39**(1) (1994) 69–84
13. Avery, C., Jolls, C., Posner, R., Roth, A.: The market for federal judicial law clerks. *University of Chicago Law Review* **68** (2001) 793–902
14. Roth, A., Xing, X.: Jumping the gun: Imperfections and institutions related to the timing of market transactions. *American Economic Review* **84** (1994) 992–1044
15. Billingsley, P.: *Probability and Measure*. John Wiley (1995)

A Proof of Proposition 1

We prove the claims in the proposition by induction on states, in the order they are considered by the algorithm, i.e. downward induction on t and upward induction on j . It will be helpful to maintain an additional induction hypothesis that

$$w(t, j) = \sum_{u=t+1}^n \frac{v(u, j)}{u(u-1)},$$

where the right side is interpreted as 0 when $t = n$.

The base cases $j = 0, t = 1, 2, \dots, n$ are trivial. (No players are active, so there is no need to compute the equilibrium strategies, only the values of $v(t, j)$ and $w(t, j)$.) The base cases $t = n, j = 1, \dots, k$ are also trivial: if the n -th element is the best one observed so far, there is no reason for any active player to play P. All of them will play O, and each of them has probability $1/j$ of winning.

For the induction step, it is easy to check that $w(t, j)$ satisfies the induction hypothesis given that $w(t + 1, j)$ does. To verify the induction hypothesis for $v(t, j), p(t, j), q(t, j)$ requires an elaborate calculation which we now explain. First let us define several events which will appear in the conditional probability expressions that define the transition probabilities of the stochastic game.

$$\begin{aligned}\mathcal{E}(t, j) &= \{\text{the game visits state } (t, j)\} \\ \mathcal{E}_P(t, j) &= \mathcal{E}(t, j) \cap \{\text{all players pass at time } t\} \\ \mathcal{E}_O(t, j) &= \mathcal{E}(t, j) \cap \{\text{at least one player makes an offer at time } t\} \\ \mathcal{E}((t, j) \rightarrow (u, \ell)) &= \{\text{the game makes a state transition from } (t, j) \text{ to } (u, \ell)\} \\ \mathcal{E}((t, j) \rightarrow \bullet) &= \{\text{the game visits state } (t, j) \text{ and } Z^{-1}(t) = x_1.\}\end{aligned}$$

The following conditional probabilities can be calculated using arguments analogous to the calculations of transition probabilities for Markov decision process representing the one-player secretary problem (e.g. [15]).

$$\begin{aligned}\Pr(\mathcal{E}((t, j) \rightarrow \bullet) \mid \mathcal{E}_P(t, j)) &= t/n \\ \Pr(\mathcal{E}((t, j) \rightarrow \bullet) \mid \mathcal{E}_O(t, j)) &= t/n \\ \Pr(\mathcal{E}((t, j) \rightarrow (u, j)) \mid \mathcal{E}_P(t, j)) &= t/(u(u-1)) \\ \Pr(\mathcal{E}((t, j) \rightarrow (u, j-1)) \mid \mathcal{E}_O(t, j)) &= t/(u(u-1)).\end{aligned}$$

Using these conditional probabilities, we wish to calculate the expected payoff to player i when playing O or P in state (t, j) given that the other players are all playing O with probability p , P with probability q . Let us denote the expected payoff to player i in these two cases by $v_O(t, j), v_P(t, j)$, respectively. A simple case analysis combined with the conditional probability formulas above yields:

$$v_O(t, j) = \sum_{i=1}^j \binom{j-1}{i-1} p^{i-1} q^{j-i} \left[\frac{1}{i} \cdot \frac{t}{n} + \left(1 - \frac{1}{i}\right) \sum_{u=t+1}^n \frac{tv(u, j-1)}{u(u-1)} \right] \quad (6)$$

$$v_P(t, j) = q^{j-1} \sum_{u=t+1}^n \frac{tv(u, j)}{u(u-1)} + (1 - q^{j-1}) \sum_{u=t+1}^n \frac{tv(u, j-1)}{u(u-1)} \quad (7)$$

Let $A = w(t, j-1) - 1/n$, $B = w(t, j-1) - w(t, j)$. We may simplify equations (6) and (7) considerably, obtaining:

$$v_O(t, j) = tw(t, j-1) - (t/j)A \frac{1-q^j}{1-q} \quad (8)$$

$$v_P(t, j) = tw(t, j-1) - q^{j-1}tB \quad (9)$$

$$\frac{1}{t}(v_O(t, j) - v_P(t, j)) = q^{j-1}B - (A/j) \frac{1-q^j}{1-q}. \quad (10)$$

From (10) we see that the set of best responses for player i is $\{\mathbf{O}\}$, $\{\mathbf{P}\}$, or $\{\mathbf{O}, \mathbf{P}\}$ according to whether the value of the function $f(q) = q^{j-1}B - (A/j)(1 + q + \dots + q^{j-1})$ is positive, negative, or zero.

When $j = 1$ this computation derives the optimal policy for the ordinary (one-player) secretary problem. We have $A = -1/n$, $B = -w(t, 1)$, $f(q) = B - A = 1/n - w(t, 1)$. Let t_1 be the largest value of t such that $w(t, 1) > 1/n$. For all $t > t_1$, $f(q) > 0$ and the unique best strategy in state $(t, 1)$ is \mathbf{O} ; moreover for $t > t_1$ we have $v(t, 1) = t/n$ and $w(t, 1) = \frac{1}{n} \sum_{u=t}^{n-1} \frac{1}{u}$. From this formula for $w(t, 1)$ we deduce that $t_1 \sim n/e$.

When $j > 1$ and $t < n$, let us first observe that $B > 0$. To see this, note that $B = \sum_{u=t+1}^n (v(u, j-1) - v(u, j))/(u(u-1))$, that each term of this sum is non-negative because increasing the number of active players at time t can not increase player i 's probability of winning after time t , and that the final term is strictly positive. If $A \leq 0$ then $f(q) > 0$ for all q , which implies that \mathbf{O} is the unique best response for player i and therefore (since we are assuming a symmetric equilibrium) $q = \Pr(i \text{ plays } \mathbf{P}) = 0$. Conversely, if $q = 0$ then \mathbf{O} is in player i 's best response set, which implies that $f(q) = f(0) \geq 0$ and therefore $A \leq 0$. Recalling that $A = w(t, j-1) - 1/n$, we have derived:

$$q = 0 \iff w(t, j-1) \leq 1/n. \quad (11)$$

Now we come to the case $q > 0$. We know that in this case, $A > 0$. Note that $f(q)$ has the same sign as $q^{1-j}f(q)$ and that

$$q^{1-j}f(q) = B - (A/j)(1 + q^{-1} + \dots + q^{-(j-1)}). \quad (12)$$

Letting $r = 1/q$, the right side of (12) is the polynomial $g(r) = B - (A/j)(1 + r + \dots + r^{j-1})$ which is monotonically decreasing as a function of $r \geq 1$ because $A > 0$. If $g(1) = B - A \leq 0$ and $q < 1$, then $r > 1$ and $g(r) < 0$, which implies that player i 's unique best response is \mathbf{P} contradicting the fact that $q < 1$ and that we are in a symmetric equilibrium. Hence we see that $B - A \leq 0$ implies $q = 1$. Conversely, if $q = 1$ then \mathbf{P} is in player i 's best response set and $f(1) = g(1) = B - A \leq 0$. Recalling that $B - A = 1/n - w(t, j)$, we have derived:

$$q = 1 \iff 1/n \leq w(t, j). \quad (13)$$

Finally, if $0 < q < 1$, then player i 's best response set is $\{\mathbf{O}, \mathbf{P}\}$ which implies that $g(r) = 0$, i.e. $1 + r + \dots + r^{j-1} = jB/A$. Since g is monotonic in the interval $(1, \infty)$, the equation $g(r) = 0$ can have at most one solution in this interval. In fact there is a solution in this interval by the intermediate value theorem, since $g(1) = B - A > 0$ (by our assumption that $q < 1$) and $g(r)$ tends to $-\infty$ as $r \rightarrow \infty$ (by our assumption that $A > 0$, as follows from the fact that $q > 0$).

This concludes the verification that there is a unique symmetric mixed Nash equilibrium in state (t, j) , and that the algorithm correctly computes this equilibrium. The verification that the algorithm correctly computes $v(t, j)$ — which is the final part of establishing the induction step — is a trivial consequence of formulas (8) and (9) above.