The Hiring Problem and Permutations

Margaret Archibald[1]   and Conrado Martínez[2] [†]

[1] *Dept. of Mathematics & Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa. Email:* `margaret.archibald-at-uct.ac.za`
Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, E-08034 Barcelona, Spain. Email: `conrado-at-lsi.upc.es`

The *hiring problem* has been recently introduced by Broder et al. in last year's ACM-SIAM Symp. on Discrete Algorithms (SODA 2008), as a simple model for decision making under uncertainty. Candidates are interviewed in a sequential fashion, each one endowed with a quality score, and decisions to hire or discard them must be taken on the fly. The goal is to maintain a good rate of hiring while improving the "average" quality of the hired staff.

We provide here an alternative formulation of the hiring problem in combinatorial terms. This combinatorial model allows us the systematic use of techniques from combinatorial analysis, e. g., generating functions, to study the problem.

Consider a permutation $\sigma : [1, \ldots, n] \to [1, \ldots, n]$. We process this permutation in a sequential fashion, so that at step i, we see the score or quality of candidate i, which is actually her face value $\sigma(i)$. Thus $\sigma(i)$ is the rank of candidate i; the best candidate among the n gets rank n, while the worst one gets rank 1. We define *rank-based* strategies, those that take their decisions using only the relative rank of the current candidate compared to the score of the previous candidates. For these strategies we can prove general theorems about the number of hired candidates in a permutation of length n, the time of the last hiring, and the average quality of the last hired candidate, using techniques from the area of analytic combinatorics. We apply these general results to specific strategies like hiring above the best, hiring above the median or hiring above the mth best; some of our results provide a complementary view to those of Broder et al., but on the other hand, our general results apply to a large family of hiring strategies, not just to specific cases.

Keywords: On-line decision making, secretary problem, hiring problem, permutations, generating functions, analytic combinatorics.

1   Introduction

The *hiring problem* has been recently introduced by Broder et al. (1) as a simple model for decision making under uncertainty, closely related to the well-known *secretary problem* (see, for instance (3) and the references therein). In the hiring problem, a growing company interviews and decides whether to hire applicants in a sequential manner. In its simplest formulation, the candidate that the company interviews

---

at step i has a quality score Q_i, where the Q_i's are i.i.d. random variables, with common distribution Unif$(0, 1)$. Then, according to the company's hiring strategy, candidate i is either hired or discarded. The paper by Broder et al. studied two natural strategies which, on rather intuitive grounds, should lead to an increasingly improved quality of the company's staff, while maintaining some balance with the speed at which the company hires applicants: the first strategy is *hiring above the mean* and the second strategy is *hiring above the median*. As their names indicate, in *hiring above the mean* an applicant is hired if and only if her score is at least equal to the mean score of the currently hired applicants, whereas in *hiring above the median*, an applicant is hired if and only if her score is at least equal to the median score of the current employees. The paper also considered the strategies *hiring above a threshold* and *hiring above the maximum*, where candidate i is hired if and only if $Q_i \geq \tau$ for some prespecified τ, or $Q_i > \max\{Q_1, \ldots, Q_{i-1}\}$, respectively.

In this paper, we provide an alternative formulation of the hiring problem in combinatorial terms; its main virtue being that it opens the door for the application of a vast and rich array of powerful techniques coming from the combinatorial camp. We by no means claim that the model that we propose here is superior to the original model, but on the contrary, that it nicely complements the original model by providing a different point of view which may prove useful in investigating the hiring problem and its many natural extensions. In particular, the combinatorial viewpoint introduced here allows us to obtain several powerful and generic results (Theorems 1 to 3) about the number of hired candidates and other relevant parameters for large families of hiring strategies, in particular, those which base their decisions solely on the *relative* rank of a candidate compared to the ranks of previous candidates.

Consider a permutation $\sigma : [1, \ldots, n] \rightarrow [1, \ldots, n]$. We process this permutation in a sequential fashion, so that at step i, we see the score or quality of candidate i, which is actually her face value $\sigma(i)$. You may think of $\sigma(i)$ as the rank of candidate i; the best candidate among the n gets rank n, while the worst one gets rank 1. In this light, the model is very natural (see also the discussion in (1)); it's not so natural to take the face value $\sigma(i)$ as an absolute measure of the candidate's quality. For similar reasons, if the Q_i's in the original hiring model are seen as relative ranks the choice of the uniform distribution in $(0, 1)$ is perfectly justifiable, but we think that it's more debatable to see them as an absolute measure of quality; for instance, it could be more natural to assume that the Q_i's are i.i.d. normal random variables with common Gaussian distribution $\mathcal{N}(\mu, \nu^2)$.

As in the original model, at step i, we must decide then whether we hire the ith candidate or not. The decision must be made based upon the values $\sigma(1), \ldots, \sigma(i)$ seen so far, and a candidate i can be hired only at step i, if at all. No information about the future is known, not even the length of the permutation σ. If we denote by $\mathcal{H}_i(\sigma)$ the set of candidates (their indices) hired up to step i when processing permutation σ, then the rules above formally translate to: 1) $\mathcal{H}_i(\sigma) \subseteq \{1, \ldots, i\}$ (no future candidates can be hired); 2) $\mathcal{H}_i(\sigma) \setminus \{i\} = \mathcal{H}_{i-1}(\sigma)$ (no past candidates can be hired)[(i)]; and 3) $\mathcal{H}_i(\sigma) = \mathcal{H}_i(\sigma')$ for any two permutations σ and σ' as long as $\sigma(j) = \sigma'(j)$ for all j, $1 \leq j \leq i$ (decisions must be made without knowledge of the future). We call $\mathcal{H}_n(\sigma)$ the *hiring set* of permutation σ and simplify the notation to $\mathcal{H}(\sigma)$.

Actually, since the future is not known, we should consider that we are given the ranks of candidates relative to those of past candidates, rather than the actual values $\sigma(i)$. For instance, while processing some sequence of candidates, we could get the information that the candidate #11 ranks the third best if

---

[(i)] This condition can be substituted by $\mathcal{H}_i(\sigma) \setminus \{i\} \subseteq \mathcal{H}_{i-1}(\sigma)$ if we want to introduce firing strategies, so that at each step one or more currently hired candidates can be fired.

compared with the 10 previously seen candidates (this only implies $\sigma(11) \leq n - 2$). This is properly captured by the notion of *rank-based strategies* that we present in Section 3.

Once the concept of the hiring set of a permutation has been introduced, some questions immediately come to mind: about its size, which we will denote $h(\sigma)$, and about other parameters like, for instance, the "time" of the last hiring $L(\sigma)$ or the score of the last hired candidate $r(\sigma)$.

Of course, our main concern is the expected value of these parameters on random permutations, e. g., if h_n is the size of the hiring set of a random permutation of size n, we want to obtain $\mathbb{E}\{h_n\}$. We shall consistently use the same letters for parameters in permutations and for random variables, like $A(\sigma)$ and A_n. We note here that if the hiring strategy itself were randomized, for example "Pessimizing Inc." hires candidate i with probability $\propto 1/\sigma(i)$, then the hiring set would actually be a probability measure over all subsets of $\{1, \ldots, n\}$, but all the definitions that we shall see here can be easily generalized to cope with these strategies as well.

Last but not least, we shall look at what happens in the asymptotic regime, i. e., when $n \to \infty$ and after a suitable scaling of the random variable of interest. As we shall shortly see, this provides the bridge between the original continuous model of Broder et al. and the discretized combinatorial version introduced here.

On the other hand, our model keeps the potential for extensions intact, and its generalization for multi-sets is both natural and immediate.

2 Simple strategies

Let us first consider hiring above a threshold τ. For simplicity, we assume $\tau \in \mathbb{Z}$. Then $\mathcal{H}_i(\sigma) = \{j \mid 1 \leq j \leq i \text{ and } \sigma(j) \geq \tau\}$, and $\mathcal{H}(\sigma) = \{1 \leq j \leq n \mid \sigma(j) \geq \tau\}$. Hence, the size h_n of the hiring set for any permutation is $n + 1 - \tau$. For the asymptotic regime, it is useful to consider $\tau = \alpha \cdot n + o(n)$ for some $0 < \alpha \leq 1$, for otherwise almost all candidates would be hired. Then

$$\frac{\mathbb{E}\{h_n\}}{n} = \frac{n + 1 - \tau}{n} = 1 - \alpha + o(1).$$

The rank r_n of the last hired candidate in a random permutation of size n is any number from τ to n with identical probability, hence

$$\mathbb{E}\{r_n\} = \sum_{j=\tau}^{n} \frac{j}{n + 1 - \tau} = \frac{1}{(n + 1 - \tau)} \left(\frac{n(n+1)}{2} - \frac{\tau(\tau - 1)}{2} \right) \sim n \frac{1 + \alpha}{2} + o(n).$$

Therefore the normalized distance to the maximum rank (the *gap*) is on average $\mathbb{E}\{g_n\} = 1 - \mathbb{E}\{r_n\}/n \sim (1 - \alpha)/2 + o(1)$ (cf. (1)). Other parameters of this hiring strategy can be easily analyzed as well.

Let us now consider the other simple strategy already studied by Broder et al., hiring above the maximum. This strategy leads to a very well known and throughly studied parameter in random permutations: left-to-right maxima (see (5) and references therein). An element $\sigma(i)$ is called a left-to-right maximum if it is larger than all preceding elements, i. e., $\sigma(j) < \sigma(i)$ for all $j < i$. Obviously, $\mathcal{H}(\sigma)$ is exactly the set of positions of the left-to-right maxima in σ. It is well known that $\mathbb{E}\{h_n\} = \ln n + O(1)$, so that the size of the hiring set is exponentially small compared to the set of interviewed candidates. We don't give here additional details about this strategy, as it turns out to be a particular case (when $m = 1$) of the strategy that we examine in Section 4.

3   A general framework for rank-based strategies

In this section we develop a generic analysis of the size of the hiring set and other parameters in *rank-based* hiring strategies.

A rank-based strategy is one where each decision (hire or discard) is taken solely on the basis of the rank of the current candidate relative to the rank of the previous interviewed candidates. That is, the actual face value $\sigma(i)$ of the current candidate i is not relevant, only its position among the previous $i - 1$ candidates. Rank-based strategies are natural and they adequately modelize constraints in some situations, for example, when there are no mechanisms for quality measurement in absolute terms.

In particular, it would be debatable that any absolute rank $\sigma(i)$ is actually available at step i; it is more reasonable to assume that the given permutation is unknown until the very last candidate is interviewed; what we keep at each step is the relative ordering of the candidates seen so far. This assumption is common, for instance, in the standard secretary problem, where only the relative ranks of the candidates are available as they are successively examined (3).

Given a permutation σ of length n and i, $1 \leq i \leq n$, let $\rho_i(\sigma)$ be the permutation of length i that we obtain by relabelling the initial prefix of length i in σ in such a way that we preserve the relative ordering. For instance, $\rho_1(25341) = 1$, $\rho_3(25341) = 132$ and $\rho_4(25341) = 1423$. Another notation that we shall define now, but use later is $\sigma \circ j$. Given a permutation σ of size n and a value j, $1 \leq j \leq n+1$, we denote by $\sigma \circ j$ the permutation of size $n + 1$ which results after relabelling j, $j + 1$, …, n in σ as $j + 1$, …, $n + 1$ and appending j to the end. For example $3241 \circ 3 = 42513$ and $213 \circ 4 = 2134$.

Definition 1 *A hiring strategy is* rank-based *if and only if for all permutations σ and all i, $1 \leq i \leq |\sigma|$,*

$$\mathcal{H}_i(\sigma) = \mathcal{H}(\rho_i(\sigma)).$$

Hiring above the maximum, above the median, above some other quantile, and above the mth best in the current staff (see Section 4) are all rank-based hiring strategies. Hiring above a threshold or above the mean are not. We will concentrate on rank-based hiring strategies for the rest of this section.

In order to investigate the average size of the hiring set in a random permutation, we introduce the bivariate generating function (2)

$$H(z, u) = \sum_{\sigma \in \mathcal{P}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)}, \tag{1}$$

where $\mathcal{P}$ denotes the set of all permutations. If we take derivates of H w.r.t. u and set $u = 1$ we obtain the generating functions of the moments of h_n, e. g.,

$$h(z) = \left. \frac{\partial}{\partial u} H(z, u) \right|_{u=1} = \sum_{\sigma \in \mathcal{P}} h(\sigma) \frac{z^{|\sigma|}}{|\sigma|!}$$

Hence $\mathbb{E}\{h_n\} = [z^n] h(z)$.

Theorem 1 *Let $H(z, u)$ be the generating function defined by* (1)*. Let $X(\sigma)$ denote the number of ranks j, $1 \leq j \leq |\sigma| + 1$, such that a candidate with score j will be hired if interviewed right after σ, that is, $X(\sigma)$ is the number of scores j such that $\mathcal{H}(\sigma \circ j) = \mathcal{H}(\sigma) \cup \{|\sigma| + 1\}$.*

Then

$$(1-z)\frac{\partial}{\partial z}H(z,u) - H(z,u) = (u-1)\sum_{\sigma\in\mathcal{P}} X(\sigma)\frac{z^{|\sigma|}}{|\sigma|!}u^{h(\sigma)}.$$

Proof: See Appendix A. □

Each different hiring strategy will be characterized by its corresponding definition of $X(\sigma)$; for instance, hiring above the maximum has $X(\sigma) = 1$ for all σ, since there is only one score for which we will hire a candidate coming after σ, namely, if the candidate has relative rank $|\sigma| + 1$.

Other interesting quantities can be analyzed in a similar vein. For instance, let $L(\sigma)$ denote the index of the last hired candidate in σ, that is, $L(\sigma) = \max\{i : i \in \mathcal{H}(\sigma)\}$, with the convention $L(\emptyset) = 0$. Then $L(\sigma \circ j) = L(\sigma)$ if the $(|\sigma| + 1)$th candidate is not hired, and $L(\sigma \circ j) = |\sigma| + 1$ otherwise. Letting

$$L(z,u) = \sum_{\sigma\in\mathcal{P}} \frac{z^{|\sigma|}}{|\sigma|!}u^{L(\sigma)},$$

the recurrence for $L(\sigma)$ translates to

$$(1-z)\frac{\partial L}{\partial z} - L(z,u) = u\sum_{\sigma\in\mathcal{P}} X(\sigma)\frac{(zu)^{|\sigma|}}{|\sigma|!} - \sum_{\sigma\in\mathcal{P}} X(\sigma)\frac{z^{|\sigma|}}{|\sigma|!}u^{L(\sigma)}, \tag{2}$$

with $X(\sigma)$ as before.

We now introduce a natural restriction on hiring strategies, which will allow us to obtain further general results. To begin with, we define the indicator $X_j(\sigma)$, so that $X_j(\sigma) = 1$ if a candidate with score j is hired after σ and $X_j(\sigma) = 0$ otherwise. Notice that $X(\sigma) = \sum_{1\le j\le|\sigma|+1} X_j(\sigma)$.

Definition 2 *A hiring strategy is* pragmatic *if and only if the following two conditions hold:*

1. *For all σ and all j, $X_j(\sigma) = 1$ implies $X_{j'}(\sigma) = 1$ for all $j' \ge j$.*

2. *For all σ and all j, $X(\sigma \circ j) \le X(\sigma) + X_j(\sigma)$.*

The first condition simply states that whenever a strategy would hire a candidate with score j, it would hire a candidate with a higher score. The second condition bounds the rate at which the strategy hires. In particular, the potential for hiring $X(\cdot)$ doesn't change if no new candidate gets hired. Pragmatic hiring strategies exclude pathological cases such as "hire any candidate that is interviewed at some step which is a multiple of 100, discard otherwise" (because of condition #2) or "hire any candidate whose relative score is better than that of an even number of previously interviewed candidates" (because of condition #1). Hiring above the median, above some quantile and above the mth best (Section 4) are all pragmatic.

Theorem 2 *For any pragmatic hiring strategy and any permutation σ, $\mathcal{H}(\sigma)$ contains at least the $X(\sigma)$ best candidates of σ, that is, the candidates with scores $|\sigma|$, $|\sigma| - 1$, ..., $|\sigma| + 1 - X(\sigma)$.*

Proof: See Appendix A. □

Let $r(\sigma)$ denote the absolute score of the last hired candidate in a permutation σ, and let $g(\sigma) = 1 - r(\sigma)/|\sigma|$ denote the *gap* (1).

Theorem 3 *For any pragmatic hiring strategy,*

$$\mathbb{E}\left\{g_n\right\} = \frac{1}{2n}(\mathbb{E}\left\{X_n\right\} - 1),$$

where $\mathbb{E}\left\{X_n\right\} = [z^n]\sum_{\sigma \in \mathcal{P}} X(\sigma)z^{|\sigma|}/|\sigma|!.$

Proof: See Appendix A.   □

4   Hiring for the elite (above the mth best)

In this strategy we have an additional parameter m. A candidate i is hired if her score is better than the score of one of the m best currently employed candidates. In other words, if $E_{i-1}(\sigma)$ is the subset of currently hired elements before step i with the m highest scores, and $\sigma(i)$ is greater than the minimum score in $E_{i-1}(\sigma)$, then i is hired.

Note that i will become immediately part of the "elite" of the m best employees, and the element ℓ with the minimum score in $E_{i-1}(\sigma)$ will be removed from the "elite", that is, it will not be in $E_i(\sigma)$. Fortunately for ℓ, he will be still hired. Note also that for $m = 1$ this strategy is simply hiring above the maximum.

For this strategy we have $X(\sigma) = |\sigma| + 1$ if $|\sigma| < m$ since any value j will be hired after processing σ, as long as an elite of m employees hasn't built up yet. Once $|\sigma| \geq m$, we will have $h(\sigma) \geq m$ and a value j will be hired if and only if it is larger than the smallest score in the elite. Since the (relative) scores of the elite of σ must consist of $|\sigma|, |\sigma| - 1, \ldots, |\sigma| - m + 1$ there are exactly m values for a newcomer to be hired, namely, if $j \in \{|\sigma| + 1, \ldots, |\sigma| - m + 2\}$ then the last candidate of $\sigma \circ j$ will be hired. Hence, $X(\sigma) = m$ if $|\sigma| \geq m$.

The right hand side of Theorem 1 is then

$$(u-1)\Bigg(1 + 2zu + 3z^2u^2 + \cdots + mz^{m-1}u^{m-1}$$

$$+ m\,H(z,u) - m(1 + zu + z^2u^2 + \cdots + z^{m-1}u^{m-1})\Bigg).$$

Plugging the expression above back into Theorem 1 and rearranging, we finally have

$$(1-z)\frac{\partial}{\partial z}H(z,u) - (mu - m + 1)H(z,u) =$$

$$(u-1)(1 + 2zu + \cdots + mz^{m-1}u^{m-1})$$

$$- m(u-1)(1 + zu + \cdots + z^{m-1}u^{m-1}). \quad (3)$$

For $m = 1$, the differential equation above reduces to

$$(1-z)\frac{\partial}{\partial z}H^{(1)}(z,u) - uH^{(1)}(z,u) = 0,$$

whose solution is

$$H^{(1)}(z,u) = \left(\frac{1}{1-z}\right)^u = \sum_{\substack{n \geq 0 \\ k \geq 0}} c_{n,k} \frac{z^n}{n!} u^k,$$

as we additionally impose $H^{(1)}(z,1) = 1/(1-z)$ and $H^{(1)}(0,u) = 1$. Here, we use the superscript to make the dependence on m explicit.

The coefficients $c_{n,k} = [z^n u^k] H^{(1)}(z,u)$ are the well-known unsigned Stirling numbers of the first kind (4), also known as Stirling cycle numbers, and denoted $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]$. The Stirling cycle number $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]$ is the number of permutations of size n that contain exactly k cycles, and it turns out to coincide with the number of permutations of size n that have exactly k left-to-right maxima (5).

The solution for general m is

$$H^{(m)}(z,u) = \frac{1}{(mu-m+1) \cdot (mu-m) \cdots (mu-1)} \left(\left(\frac{1}{1-z}\right)^{mu-m+1} P_m(u,z) \right.$$

$$\left. + \frac{1}{(1-z)^m} Q_m(z,u) \right), \tag{4}$$

where $P_m(u,z)$ and $Q_m(z,u)$ are polynomials in z and u.

If we differentiate w.r.t. u and set $u = 1$, we obtain the generating function of the expected values

$$h^{(m)}(z) = \sum_{n \geq 0} \mathbb{E}\left\{h_n^{(m)}\right\} z^n = \left. \frac{\partial}{\partial u} H^{(m)}(z,u) \right|_{u=1} = m \frac{\ln\left(\frac{1}{1-z}\right)}{1-z} - \frac{p_m(z)}{1-z},$$

with $p_m(z)$ a polynomial of degree $m-1$.

Hence $\mathbb{E}\left\{h_n^{(m)}\right\} = mH_n + O(1)$, where $H_n = \sum_{1 \leq k \leq n}(1/k)$ denotes the nth harmonic number. We keep here the usual notation H_n for harmonic numbers despite the possible confusion with the hiring set parameters. Since $H_n = \ln n + \gamma + O(n^{-1})$, where $\gamma = 0.577\ldots$ is Euler's gamma constant, we conclude that $\mathbb{E}\left\{h_n^{(m)}\right\} = m \ln n + O(1)$. So the size of the hiring set is, for any fixed m, exponentially smaller than the set of interviewed candidates.

Since we have an explicit form for $H^{(m)}(z,u)$, much more information about $h_n^{(m)}$ can be extracted. In particular, we have

$$\mathbb{E}\left\{u^{h_n^{(m)}}\right\} = [z^n] H^{(m)}(z,u) \sim A_m(u) \cdot n^{m(u-1)} \cdot \left(1 + \Theta\left(\frac{1}{n}\right)\right)$$

uniformly in a complex neighborhood of $u = 1$, for some analytic $A_m(u)$, so it follows by application of Hwang's quasi-powers theorem (2) that $h_n^{(m)}$ converges to a normal distribution. More precisely,

$$\frac{h_n^{(m)} - m \ln n}{\sqrt{m \ln n}} \xrightarrow{\text{d}} \mathcal{N}(0,1). \tag{5}$$

Also, since $\mathbb{E}\{X_n\} = m$ if $n \geq m$, Theorem 3 yields for this strategy $\mathbb{E}\left\{g_n^{(m)}\right\} = (m-1)/2n$, if $n \geq m$.

We now consider the behavior of this strategy as m varies (notice that the results that we have discussed above hold only for fixed m). To this end we introduce

$$\mathsf{H}(z, u, v) = \sum_{m \geq 1} v^m H^{(m)}(z, u),$$

with $H^{(m)}(z, u)$ the generating function that we have studied in the preceding paragraphs.

If we set $\mathsf{h}(z, v) = (\partial \mathsf{H}/\partial u)|_{u=1}$, the coefficient $[z^n v^m]\mathsf{h}(z, v)$ is the quantity we seek, the expected size of the hiring set when the size of the elite is m. Multiplying by v^m and summing over all $m \geq 1$, the differential equation (3) translates into a corresponding differential equation for $\mathsf{H}(z, u, v)$

$$(1 - z)\frac{\partial}{\partial z}\mathsf{H}(z, u, v) - \mathsf{H}(z, u, v) - (u - 1)v\frac{\partial}{\partial v}\mathsf{H}(z, u, v) = (1 - u)\frac{v^2}{(1 - v)^2}\left(\frac{1}{1 - zuv}\right).$$

Similarly, differentiating w.r.t. u and setting $u = 1$ the equation above we get an ordinary differential equation for $\mathsf{h}(z, v)$

$$(1 - z)\frac{\partial}{\partial z}\mathsf{h}(z, v) - \mathsf{h}(z, v) - \frac{v}{(1 - z)(1 - v)^2} = -\left(\frac{v}{1 - v}\right)^2\left(\frac{1}{1 - zv}\right), \tag{6}$$

since $\mathsf{H}(z, 1, v) = \frac{v}{(1-z)(1-v)}$.

The solution for this equation gives (a detailed derivation can be found in Appendix A)

$$\mathsf{h}(z, v) = \frac{v \ln \frac{1}{1-z}}{(1 - z)(1 - v)^2} - \frac{v \ln \frac{1}{1-zv}}{(1 - z)(1 - v)^2} \tag{7}$$

as we impose $\mathsf{h}(0, v) = 0$.

The last step is to extract the coefficients of $\mathsf{h}(z, v)$, whose details are also given in Appendix A. For $m \geq n$, we obviously have $\mathbb{E}\left\{h_n^{(m)}\right\} = n$. For $m \leq n$ we have $\mathbb{E}\left\{h_n^{(m)}\right\} = m(H_n - H_m + 1)$, so $\mathbb{E}\left\{h_n^{(m)}\right\} \sim m\ln\left(\frac{n}{m}\right) + m + O(1)$, for $n, m \to \infty$.

5   Hiring above the median (and other quantiles)

Hiring above the median means that candidate i is hired if and only if her score $\sigma(i)$ is larger than the rth best score of the candidates hired so far, with $r = \lfloor (h_{i-1}(\sigma) + 1)/2 \rfloor$.

Since this strategy is rank-based, it is not difficult to see that if the hiring set has size $k = 2t$ at some given moment then there are $t + 1$ possible relative scores that will be hired in the next step, while if the hiring set has size $k = 2t + 1$ then the number of relative scores that would be hired in the next step is also $t + 1$. That means that $X(\sigma) = \lceil (h(\sigma) + 1)/2 \rceil$. Coping with the ceilings is quite hard, so we will consider instead what happens with $X'(\sigma) = (1 + h(\sigma))/2$ and $X''(\sigma) = (3 + h(\sigma))/2$, which provide lower and upper bounds, respectively.

By the same token, hiring above other quantiles, say hiring above $(1 - a)h(\sigma)$, with $0 < a < 1$, can be analyzed in the same way. We should have then $X(\sigma) = \lceil a \cdot (h(\sigma) + 1) \rceil$. In general, for $X(\sigma) = a \cdot h(\sigma) + b$ and $0 < a < 1$, we have

$$(1 - z)\frac{\partial H}{\partial z} - au(u - 1)\frac{\partial H}{\partial u} - (1 + b(u - 1))H(z, u) = 0,$$

with the additional conditions $H(z, 1) = 1/(1 - z)$ and $H(0, u) = 1$. The solution turns out to be

$$H(z, u) = u^{-b/a} \frac{1}{1 - z} \left(\frac{1}{1 - \frac{u-1}{u(1-z)^a}} \right)^{b/a}, \tag{8}$$

which can be readily checked (and even found!) with any reasonable computer algebra system. From this closed form we can find the successive factorial moments. It suffices to differentiate r times and set $u = 1$:

$$\mathbb{E}\{h_n^{\underline{r}}\} = [z^n] \left. \frac{\partial^r H(z, u)}{\partial u^r} \right|_{u=1},$$

where $X^{\underline{r}} = X(X - 1) \cdots (X - r + 1)$ denotes the rth falling factorial (4). In appendix A we show that

$$\mathbb{E}\{h_n^{\underline{r}}\} = \Theta(n^{ra}). \tag{9}$$

The expected size of the hiring set can also be obtained if we consider the differential equation satisfied by the corresponding generating function $h(z)$, namely,

$$(1 - z) \frac{d}{dz} h - (1 + a)h = \frac{b}{1 - z}.$$

This is a simple linear first-order ordinary differential equation whose solution is

$$h(z) = \frac{b}{a} \frac{1}{1 - z} \left(\frac{1}{(1 - z)^a} - 1 \right).$$

since $h(0) = 0$. This coincides with what we get if we differentiate $H(z, u)$ as given by (8) and set $u = 1$. The extraction of coefficients is straightforward:

$$[z^n]h(z) = \frac{b}{a} \left(\binom{n + a}{a} - 1 \right) = \frac{b}{a} \frac{n^a}{\Gamma(1 + a)} \left(1 + O\left(\frac{1}{n}\right) \right).$$

In particular, for $a = 1/2$ (hiring above the median) we get $\mathbb{E}\{h_n\} = \Theta(\sqrt{n})$ and for "hire A, move B" (see (1)) we have $a = 1 - B/A$; thus $\mathbb{E}\{h_n\} = \Theta(n^{1-B/A})$. Loosely speaking, when the size of the hiring set reaches k we have interviewed $n = \Theta(k^{A/(A-B)})$ candidates (compare with the results in (1)).

On the other hand $\mathbb{E}\{X_n\} = a\mathbb{E}\{h_n\} + o(\mathbb{E}\{h_n\})$, thus

$$\mathbb{E}\{g_n\} = \Theta(n^{a-1}).$$

For the particular case of hiring above the median, when $a = 1/2$, we have $\mathbb{E}\{g_n\} = \Theta(1/\sqrt{n})$.

References

[1] A. Z. Broder, A. Kirsch, R. Kumar, M. Mitzenmacher, E. Upfal, and S. Vassilvitskii. The hiring problem and Lake Wobegon strategies. In *Proc. of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1184–1193. ACM-SIAM, 2008.

[2] Ph. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge Univ. Press, 2008.

[3] P. R. Freeman. The secretary problem and its extensions: A review. *International Statistical Review*, 51:189–206, 1983.

[4] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, Mass., 2nd edition, 1994.

[5] D. E. Knuth. *The Art of Computer Programming: Fundamental Algorithms*, volume 1. Addison-Wesley, Reading, Mass., 3rd edition, 1997.

A Proofs

Proof (of Theorem 1): We can write $h(\sigma) = 0$ if σ is the empty permutation and $h(\sigma \circ j) = h(\sigma) + X_j(\sigma)$, where

$$X_j(\sigma) = \begin{cases} 1, & \text{if the last candidate of } \sigma \circ j \text{ is hired,} \\ 0, & \text{otherwise.} \end{cases}$$

Then, if $\mathcal{P}_n$ denotes the set of permutations of size n,

$$H(z, u) = \sum_{\sigma \in \mathcal{P}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} = 1 + \sum_{n>0} \sum_{\sigma \in \mathcal{P}_n} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} = 1 + \sum_{n>0} \sum_{1 \le j \le n} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma \circ j|}}{|\sigma \circ j|!} u^{h(\sigma \circ j)}$$

$$= 1 + \sum_{n>0} \sum_{1 \le j \le n} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|+1}}{(|\sigma|+1)!} u^{h(\sigma)+X_j(\sigma)} = 1 + \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|+1}}{(|\sigma|+1)!} u^{h(\sigma)} \sum_{1 \le j \le n} u^{X_j(\sigma)}.$$

Since $X_j(\sigma)$ is either 0 or 1 for all j and all σ, we have

$$\sum_{1 \le j \le n} u^{X_j(\sigma)} = (|\sigma| + 1 - X(\sigma)) + uX(\sigma),$$

where $X(\sigma) = \sum_{1 \le j \le |\sigma|+1} X_j(\sigma)$. Note that $X(\sigma)$ is the number of relative scores such that a candidate with such a score would be hired right after processing σ.

Hence,

$$H(z, u) = 1 + \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|+1}}{(|\sigma|+1)!} u^{h(\sigma)} \Big((|\sigma| + 1 - X(\sigma)) + uX(\sigma) \Big).$$

Taking derivatives w.r.t. z,

$$\frac{\partial}{\partial z} H(z, u) = \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} \Big((|\sigma| + 1 - X(\sigma)) + uX(\sigma) \Big)$$

$$= \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z \frac{d}{dz} z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} + \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} + (u-1) \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} X(\sigma)$$

$$= z \frac{\partial}{\partial z} H(z, u) + H(z, u) + (u-1) \sum_{n>0} \sum_{\sigma \in \mathcal{P}_{n-1}} \frac{z^{|\sigma|}}{|\sigma|!} u^{h(\sigma)} X(\sigma).$$

After reorganizing the terms in the equation above and simplifying, we obtain the statement of the theorem. □

Proof (of Theorem 2): The proof is by induction on the length n of the permutation σ. If $n = 0$ then $\mathcal{H}(\sigma) = \emptyset$ and indeed it contains the $X(\sigma)$ best candidates in the (empty) permutation σ.

Consider now $\sigma' = \sigma \circ j$. By the inductive hypothesis $\mathcal{H}(\sigma)$ contains the best $X(\sigma)$ candidates, with relative scores $\{|\sigma| - X(\sigma) + 1, \ldots, |\sigma|\}$. Since the strategy is pragmatic, only candidates with relative rank between $|\sigma| + 2 - X(\sigma)$ and $|\sigma| + 1$ will be hired. If the last candidate with relative score j is hired

then he is among the best $X(\sigma) + 1$ candidates of $\mathcal{H}(\sigma')$. As $X(\sigma') \leq X(\sigma) + 1$, it follows that $\mathcal{H}(\sigma')$ contains at least the best $X(\sigma')$ candidates of σ'. On the contrary, if the last candidate were not hired then the relative scores of the best $X(\sigma)$ candidates in $\mathcal{H}(\sigma)$ increase all by one. Hence, in $\mathcal{H}(\sigma')$ we have at least $X(\sigma)$ candidates with scores in $\{|\sigma| + 2 - X(\sigma), \ldots, |\sigma|\}$. To conclude the proof it is enough to notice that, for any pragmatic strategy, $X(\sigma \circ j) = X(\sigma)$ if the last candidate with score j was not hired. $\square$

Proof (of Theorem 3): The last hired candidate must have an absolute score in $\{|\sigma| + 1 - X(\sigma), \ldots, |\sigma|\}$ because of Theorem 2. For a random permutation, all these $X(\sigma)$ scores are equally likely, hence for a random permutation of size n we have

$$\mathbb{E}\{r_n\} = \mathbb{E}\left\{\sum_{k=n-X(\sigma)+1}^{n} \frac{k}{X(\sigma)}\right\} = \mathbb{E}\left\{\frac{1}{X(\sigma)}\left(\frac{n(n+1)}{2} - \frac{(n-X(\sigma))(n+1-X(\sigma))}{2}\right)\right\}$$

$$= \mathbb{E}\left\{n + \frac{1}{2} - \frac{1}{2}X(\sigma)\right\} = n + \frac{1}{2} - \frac{\mathbb{E}\{X_n\}}{2}.$$

Finally, $\mathbb{E}\{g_n\} = 1 - n^{-1}\mathbb{E}\{r_n\} = (\mathbb{E}\{X_n\} - 1)/2n$. $\square$

Proof (of Equation (7) **and coefficient** $[z^n v^m] h(z, v)$**):** We start with the linear differential equation satisfied by $h(z, v)$ (Equation (6))

$$(1 - z)\frac{\partial}{\partial z}h(z, v) - h(z, v) = \frac{v}{1 - z}\frac{1}{(1 - v)^2} - \frac{v^2}{(1 - v)^2}\frac{1}{1 - zv}.$$

Multiplying through by the integrating factor $1 - z$ and integrating with respecth to z gives

$$(1 - z)h(z, v) = \frac{v}{(1 - v)^2}\ln\left(\frac{1}{1 - z}\right) - \frac{v}{(1 - v)^2}\ln\left(\frac{1}{1 - zv}\right) + c(v),$$

for some unknown function $c(v)$.

Using the initial condition $h(0, v) = 0$, we find that $c(v) = 0$ for any v. Hence

$$[z^n v^m]h(z, v) = [z^n v^m]\frac{1}{1 - z}\left(\frac{v}{(1 - v)^2}\ln\left(\frac{1}{1 - z}\right) - \frac{v}{(1 - v)^2}\ln\left(\frac{1}{1 - zv}\right)\right)$$

$$= [z^n]\left(\frac{1}{1 - z}\ln\left(\frac{1}{1 - z}\right)[v^m]\frac{v}{(1 - v)^2} - \frac{1}{1 - z}[v^m]\frac{v}{(1 - v)^2}\ln\left(\frac{1}{1 - zv}\right)\right)$$

$$= m[z^n]\frac{1}{1 - z}\ln\left(\frac{1}{1 - z}\right) - [z^n]\frac{1}{1 - z}\sum_{k=1}^{m}\left(\frac{mz^k}{k} - z^k\right)$$

Extracting the coefficient of z^n above is now easy,

$$[z^n v^m]\mathsf{h}(z,v) = mH_n - \left(m \sum_{k=1}^{m} \frac{1}{k} [z^{n-k}] \frac{1}{1-z} - \sum_{k=1}^{m} [z^{n-k}] \frac{1}{1-z} \right)$$

$$= mH_n - m \sum_{k=1}^{\min(m,n)} \frac{1}{k} + \sum_{k=1}^{\min(m,n)} 1$$

$$= \begin{cases} mH_n - mH_m + m, & \text{if } m \leq n, \\ n, & \text{if } m > n. \end{cases}$$

$\square$

Proof (of Equation 9): Our starting point is

$$H(z,u) = u^{-b/a} \frac{1}{1-z} \left(\frac{1}{1 - \frac{u-1}{u(1-z)^a}} \right)^{b/a}.$$

It suffices to differentiate r times and set $u = 1$ to obtain the generating function of the rth factorial moments of h_n:

$$\mathbb{E}\left\{ h_n^{\underline{r}} \right\} = [z^n] h_r(z)$$

with

$$h_r(z) = \left. \frac{\partial^r H(z,u)}{\partial u^r} \right|_{u=1}.$$

We have thus

$$h_r(z) = \gamma_r \sum_{j=0}^{r} \frac{(-1)^{r-j}}{(1-z)^{ja+1}} \binom{r}{j},$$

where γ_r is a polynomial of degree r in $x = b/a$. Extracting coefficients

$$[z^n] h_r(z) = \gamma_r \sum_{j=0}^{r} (-1)^{r-j} \binom{r}{j} \binom{ja+n}{n}.$$

Hence we get that, asymptotically as $n \to \infty$,

$$[z^n] h_r(z) \sim \gamma_r \frac{n^{ra}}{\Gamma(ra+1)}.$$

$\square$