Optimal Policies to Obtain the Most Join Results

Nariankadu D. Shyamalkumar, Ralph P. Russo
and Ramon Lawrence

July 2005

Technical Report No. 353

Optimal Policies to Obtain the Most Join Results

Nariankadu D. Shyamalkumar★, Ralph P. Russo★ and Ramon Lawrence★★

Consider two finite or infinite populations, each member of which carries a positive integer valued label. Samples are drawn without replacement. A match is said to occur between two sampled members if they are from different populations and carry the same label. The object is to sample from the two sources in an order that maximizes the number of matches, uniformly across all steps. An optimal strategy is identified in the infinite case. In the finite case, while an optimal policy is shown to not always exist, we identify a policy which beats one that is commonly used. The work is motivated by a database problem in computer science. Many of the results are established through the probabilistic technique of coupling.

KEY WORDS: Optimal Strategies, Sequential Strategies, Greedy Policy, Coupling

★ Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242; E-mail: rrusso@stat.uiowa.edu, shyamal-kumar@uiowa.edu; Corresponding Author: Ralph P. Russo

★★ Department of Computer Science, University of Iowa, Iowa City, IA 52242. E-mail: rlawrenc@cs.uiowa.edu

1. Introduction

Consider identical decks **R** and **S** of n cards each, numbered 1 through n. Two cards are said to *match* if they have the same denomination. Suppose that we sample all $2n$ cards, one at a time, from the two decks. On each selection we are allowed to determine which of the decks to sample from, but sampling from the individual decks is random and without replacement. The goal is to generate matches as quickly as possible. More precisely, we wish to find a *reading policy* (an algorithm for sequentially choosing which of **R** or **S** to sample from) that maximizes the expected number of matches after k steps, uniformly in k. In this rather simple setting it turns out that the *alternating* policy, which alternately samples from the two decks, is optimal, and that the *myopic* or *greedy* policy, which always chooses the deck providing the larger expected gain in matches, is equivalent to the alternating and hence is optimal as well.

The above is a simple case of a more general problem. Let **R** and **S** denote two populations with possibly infinite cardinalities such that each member of these populations carries a positive integer valued label. Samples are drawn without replacement, which in the case of infinite cardinality reduces to i.i.d. sampling from two distributions. A match occurs between two members if they are from different populations and carry the same label. Again, the object is to sample from the two sources **R** and **S** so as to maximize the expected number of matches uniformly across all steps. The optimality of the alternating strategy in the above example, while hardly surprising, is not especially obvious (see our argument in Example **3**, below). Moreover, we will show that the conclusion reached in the example, that alternating is *greedy* and all *greedy* are optimal, is not valid in general.

The motivation for the work done here is a specific database problem in computer science. When a user queries a database, it is desirable for the system to generate answers quickly, so that data can be processed immediately. Minimization of query response time is also important when the database system is part of a larger information pipeline, where query results are fed to analysis programs for further processing. Overall pipeline execution time will be reduced if the application can begin working before the database is completely read. One of the major factors in query response and overall execution time is the execution of the *join* of two relations. The goal is to maximize the expected output rate among all reading policies. Recently, there has been interest in the study of the effect of reading policies in the execution of join relations, albeit with different goals. For example, Haas and Hellerstein (1999) consider the effect of reading policies in ripple joins for online aggregation, especially in terms of the length of the confidence interval for the aggregate quantity of interest. See also Luo et al (2002). For another such type of problem, see Ilyas et al (2003). Use of the output rate as a measure of performance can be seen in Viglas and Naughton (2002) and Chandrasekaran and Franklin (2002).

In our abstract problem, a relation is a population and a join is a function that specifies when two members, one from each population, match. The indicator function which takes the value one when the two members have identical labels serves as the join in our case. Hence an output is generated by a match and maximization of the expected output rate amounts to the maximization of the expected number of matches. Strictly speaking, all databases are finite. But since sampling a small percentage from a finite population without replacement is probabilistically very close to sampling with replacement (see Diaconis and Freedman (1980) for a precise statement), it is of interest to study the infinite population case.

Our problem, while reminiscent of the two armed bandit problem (see Berry and Fristedt (1985)) is in fact very different from it. In the latter, the need to sample from both populations stems from an ignorance of the population distribution(s), whereas in our problem the same need stems from the fact that matches occur only between members of different populations. Second, in the bandit problem, one is maximizing a one dimensional objective function, while in our case we wish to maximize the expected number of matches at all steps. Hence, while in our case a solution may not always exist, in the case of the bandit problem, under quite general conditions, a solution always exists. In our case the reason that maximization of the expected number of matches at a specified step is not interesting is that the dynamic nature of query processing confers no special status to any specific step. This is also the reason why the techniques used in the two problems are very different. In our problem we make no use of dynamic programming techniques, but rather rely frequently on the probabilistic technique of coupling (see Lindvall (1992)).

It is convenient to classify reading policies as either adaptive or non-adaptive. A non-adaptive policy ignores the information (the observed labels) that is obtained from the samples. The non-adaptives include the fixed policies, where the sampling order is constant. For example, a plan with a 2 : 1 policy samples in the order **R**,**R**,**S**,**R**,**R**,**S**,... The non-adaptives include also policies where the order is randomly determined in a way such that the probability of the k-th source being **R** depends solely on the source sequence through the $(k-1)$st step. The adaptive policies, when deciding where (**R** or **S**) to sample on the k-th step, may take into account the data that has been generated (the observed labels) through the first $(k-1)$ steps. A natural adaptive policy, is the myopic or greedy policy described above. In a sense, a greedy policy is a short term strategy that optimizes the expected one step gain, with no explicit regard to future (two step and beyond) gains. Note that there may be more than one greedy policy as the greedy criterion at some steps, occasionally, may be ambivalent between **R** and **S**. In our problem, greedy policies merit special attention since optimal policies, if existent, must be greedy.

Alternating policies are also of interest for their simplicity, their frequent appearance in the literature (for this reason) and for their optimality among a restricted class of policies as shown below. Finally, we note that alternating, unlike greedy, can be adopted even when the population distributions are unknown.

Let $M_A(n)$ denote the number of matches formed through the first n steps under policy A. We say that policy A beats policy B if

$$\mathbb{E}(M_A(n)) \geq \mathbb{E}(M_B(n)), \quad n \geq 1 \tag{1}$$

with strict inequality for at least one value of n. Policy B is regarded as inadmissable in this case. If (**1**) holds for all policies B, we say that policy A is optimal. Optimal policies may not exist in some cases.

In Section 2, we show alternating policies to be optimal among the non-adaptives. In Section 3 we investigate the class of greedy policies. First, assuming that both populations are infinite, we show that any greedy policy is optimal. When one or both populations are finite, we show that in general there may not exist an optimal policy. Nonetheless, it makes sense to use a greedy policy vis-à-vis any non-adaptive as the former is shown always to be as good as the latter and in most cases strictly better. We end this section by defining some additional notation.

Notation:

In the sampling without replacement case the cardinalities of $\mathbf{R}$ and $\mathbf{S}$ are denoted by $|\mathbf{R}|$ and $|\mathbf{S}|$, respectively. We assume, without loss of generality, that $|\mathbf{R}| \leq |\mathbf{S}|$. A record from either source carries one of the possible l labels $1, \dots, l$. The probability that a single record from the $\mathbf{R}$ source (resp., the $\mathbf{S}$ source) carries the i-th label, is r_i (resp., s_i). The probability vectors $(r_1, \dots, r_l)$ and $(s_1, \dots, s_l)$ are denoted by $\tilde{r}$ and $\tilde{s}$, respectively. $\tilde{1}$ denotes a generic vector of ones whose dimension should be clear from context. The inner product of $\tilde{r}$ with $\tilde{s}$ is denoted by μ, i.e. $\mu = \tilde{r} \cdot \tilde{s}$. We shall always assume μ to be positive as otherwise there will be no common label between the two sources. The labels on the n-th records read from the $\mathbf{R}$ and $\mathbf{S}$ sources are denoted by $L_{\mathbf{R}}(n)$ and $L_{\mathbf{S}}(n)$. The above implies that $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of identically distributed random variables with

$$\Pr(L_{\mathbf{R}}(1) = i) = r_i \quad \text{and} \quad \Pr(L_{\mathbf{S}}(1) = i) = s_i, \quad i = 1, \dots, l \tag{2}$$

Associated with the sequences $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are the discrete time vector counting process $\{\tilde{N}_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{\tilde{N}_{\mathbf{S}}(n)\}_{n\geq 1}$; the first is defined by

$$\tilde{N}_{\mathbf{R}}(n) = (N_{\mathbf{R}}(n,1), \dots, N_{\mathbf{R}}(n,l)), \quad \text{with } N_{\mathbf{R}}(n,i) = \sum_{j=1}^{n} I_{\{L_{\mathbf{R}}(j)=i\}}, \quad i = 1 \dots, l; \; n = 1, \dots \tag{3}$$

and the second is defined analogously.

A reading policy is a zero-one valued stochastic process with the convention that the value 1 denotes a selection from source $\mathbf{R}$ and the value 0 a selection from source $\mathbf{S}$. Hence

$$C(n) = \begin{cases} 1 & \text{if the } n\text{-th selection is from } \mathbf{R}; \\ 0 & \text{if the } n\text{-th selection is from } \mathbf{S}; \end{cases}, \quad n = 1, 2, \ldots \tag{4}$$

Associated with each reading policy are two counting processes $\{R(n)\}_{n\geq 1}$ and $\{S(n)\}_{n\geq 1}$ defined by

$$R(n) := \sum_{j=1}^{n} C(j) \quad \text{and} \quad S(n) := n - R(n), \quad n = 1, 2, \ldots \tag{5}$$

These processes keep track of the number of records read from **R** and **S**, respectively, after a total of n records have been read. Also associated with a reading policy is a non-decreasing process $\{M(n)\}_{n\geq 1}$ which counts the number of join returns, i.e. matches, generated by the first n records. Hence,

$$M(n) = \tilde{N}_{\mathbf{R}}(R(n)) \cdot \tilde{N}_{\mathbf{S}}(S(n)), \quad n = 1, 2, \ldots \tag{6}$$

Observe that all of the processes $\{M(n)\}_{n\geq 1}$, $\{R(n)\}_{n\geq 1}$ and $\{S(n)\}_{n\geq 1}$ depend on the reading policy even though the notation does not make it explicit.

The filtration $\{\mathcal{F}_n\}_{n\geq 1}$ is defined by

$$\mathcal{F}_{n+1} := \mathcal{F}_1 \vee \sigma \langle L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}(R(n)); L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}(S(n)) \rangle, \quad n = 1, 2, \ldots \tag{7}$$

with $\mathcal{F}_1$ being arbitrary. $\mathcal{F}_1$ for example could contain all the information needed for randomization. All reading policies will henceforth be assumed to be adapted to the above filtration - they form the set of all *implementable* reading policies. Note that the filtration itself depends on the reading policy.

2. Alternating Policies

We study the class of non-adaptive reading policies, $\mathcal{C}_{NA}$, that is the class of policies which do not make use of the information contained in the records that have been read in deciding the source for the next record. Observe that this does not imply that such policies have deterministic strategies as they could always use some exogenous randomization to determine the choice at each step. Taking the expected number of matches as our optimality criterion we show (Theorem **1**) that each member of a sub-class of such policies, which satisfy

$$R(2n) = \min(n, |\mathbf{R}|), \quad n = 1, 2, \ldots \tag{8}$$

is optimal. Any reading policy in $\mathcal{C}_{NA}$ which satisfies (8) is called an alternating policy. In words, an alternating policy is one which does not use any information from the records, and under which at any step the numbers of records read from the two sources are within one of each other until such time (in the finite population case) that this is not possible.

There exists a large, if not an infinite, number of alternating policies. For our purposes $C_{\mathbf{A}}$ will denote the canonical alternating policy which strictly alternates between the two sources with the first pick being from **R**. Hence,

$$C_{\mathbf{A}}(n) = \min(n, 2|\mathbf{R}|) \bmod 2, \quad n = 1, 2, \ldots \tag{9}$$

From the point of view of implementation, it may be more efficient to work with the alternating policy given by

$$C(n) = I_{\{n \bmod 4 < 2\}}, \quad n = 1, 2, \ldots \tag{10}$$

until **R** is completely read as it, leaving apart the first record, reads two records at a time from the chosen source. Due to theorem **1** below, the particular choice of alternating policy is immaterial from the point of view of performance.

Before showing the optimality of alternating policies among the non-adaptives, we state a lemma that helps compare the performances of policies within $\mathcal{C}_{NA}$.

Lemma 1 For a reading policy $C(\cdot) \in \mathcal{C}_{NA}$, we have

$$\mathbb{E}(M(n)) = \mathbb{E}(R(n)S(n))\mu, \quad n = 1, 2, \ldots \tag{11}$$

Proof As a member of $\mathcal{C}_{NA}$, the process $C(\cdot)$ is independent of $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$. The proof now follows by observing that

$$\begin{aligned}
\mathbb{E}(M(n)) = \mathbb{E}\Big(\tilde{N}_{\mathbf{R}}(R(n)) \cdot \tilde{N}_{\mathbf{S}}(S(n))\Big) &= \mathbb{E}\Big(\mathbb{E}\Big(\tilde{N}_{\mathbf{R}}(R(n)) \cdot \tilde{N}_{\mathbf{S}}(S(n)) | R(n)\Big)\Big) \\
&= \mathbb{E}\Big(R(n)S(n)\mathbb{E}\Big(\tilde{N}_{\mathbf{R}}(1) \cdot \tilde{N}_{\mathbf{S}}(1)\Big)\Big) \\
&= \mathbb{E}(R(n)S(n))\, \tilde{r} \cdot \tilde{s} \\
&= \mathbb{E}(R(n)S(n))\mu
\end{aligned}$$

§

It is instructive to look at the policy whose selections are determined by independent tosses of a *p-coin* - heads choose **R** and tails choose **S**. Assuming that both $|\mathbf{R}|$ and $|\mathbf{S}|$ are greater than n, we have $\mathbb{E}(M(n)) = n(n-1) * p * (1-p)\mu$. This expectation is maximized for a fair coin, for which $\mathbb{E}(M(n)) = (1/4)n(n-1)\mu$. In this case, the shortfall in the expected number of matches at the n-th step compared to that of the alternating is given by $\lfloor n/2 \rfloor \mu/2$. Note that while the expected number of matches is of order n^2, the shortfall from the optimal level is only of order n.

Theorem 1 Among all reading policies in $\mathcal{C}_{NA}$, the alternating policies maximize $\mathbb{E}(M(n))$ for all n. Moreover, for these policies

$$\mathbb{E}(M(n)) = \min(\lfloor n/2 \rfloor, |\mathbf{R}|) \max(\lceil n/2 \rceil, n - |\mathbf{R}|)\, \mu, \quad n = 1, 2, \ldots, |\mathbf{R}| + |\mathbf{S}| \tag{12}$$

Proof The above follows by observing that

$$\mathbb{E}(M(n)) = \mathbb{E}(R(n)S(n))\mu = \mathbb{E}(R(n)[n - R(n)])\mu, \quad n = 1, 2, \ldots \tag{13}$$

and the fact that the function $\psi_n(m) = m * (n - m)$, for n a positive integer, defined on the set of integers attains its maximum on $\{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$. §

3. Greedy Policies

In this section, we examine policies that utilize knowledge of $\tilde{r}$ and $\tilde{s}$, together with the information contained in the records that have been read, in order to optimize the choice for the next step. We shall refer to such policies as *greedy* policies. Towards a more precise definition, we observe that

$$\begin{aligned}\mathbb{E}(M(n+1) - M(n)|\mathcal{F}_{n+1}) = {} & \mathbb{E}(N_{\mathbf{S}}\,[S(n), L_{\mathbf{R}}\,(R(n)+1)]|\mathcal{F}_{n+1})C(n+1) \\ & + \mathbb{E}(N_{\mathbf{R}}\,[R(n), L_{\mathbf{S}}\,(S(n)+1)]|\mathcal{F}_{n+1})\,(1 - C(n+1))\end{aligned} \tag{14}$$

The above implies that any $C(\cdot)$ maximizing the above conditional expectation should satisfy

$$C(n+1) = \begin{cases} 1 & \text{if } \mathbb{E}(N_{\mathbf{S}}\,[S(n), L_{\mathbf{R}}\,(R(n)+1)]|\mathcal{F}_{n+1}) > \mathbb{E}(N_{\mathbf{R}}\,[R(n), L_{\mathbf{S}}\,(S(n)+1)]|\mathcal{F}_{n+1}); \\ 0 & \text{if } \mathbb{E}(N_{\mathbf{S}}\,[S(n), L_{\mathbf{R}}\,(R(n)+1)]|\mathcal{F}_{n+1}) < \mathbb{E}(N_{\mathbf{R}}\,[R(n), L_{\mathbf{S}}\,(S(n)+1)]|\mathcal{F}_{n+1}); \end{cases} ,\forall n \geq 1 \tag{15}$$

Note that at every step where

$$\mathbb{E}(N_{\mathbf{S}}\,[S(n), L_{\mathbf{R}}\,(R(n)+1)]|\mathcal{F}_{n+1}) = \mathbb{E}(N_{\mathbf{R}}\,[R(n), L_{\mathbf{S}}\,(S(n)+1)]|\mathcal{F}_{n+1}), \tag{16}$$

two greedy policies may differ as the greedy criterion is ambivalent. Such possibilities make the set of greedy policies large and its study more interesting.

The principal reason that greedy policies are of interest is given by the following theorem.

Theorem 2 Every optimal policy is necessarily greedy.

Proof Let C denote an optimal policy and let the n-th step be the first step where C makes a non-greedy choice with positive probability. Now let $C_{\mathbf{G}}$ denote a greedy policy which agrees with C on the first $n-1$ steps - this is important as the first $n-1$ steps might involve some (apart from the first two) where the greedy criterion is ambivalent. Let $M(\cdot)$ and $M_{\mathbf{G}}(\cdot)$ denote the processes which represent the number of matches under C and $C_{\mathbf{G}}$, respectively. Then by construction we have $M(n-1) = M_{\mathbf{G}}(n-1)$ and since $C_{\mathbf{G}}$ is greedy we have,

$$\mathbb{E}(M_{\mathbf{G}}(n) - M_{\mathbf{G}}(n-1)|\mathcal{F}_n) - \mathbb{E}(M(n) - M(n-1)|\mathcal{F}_n) \geq 0. \tag{17}$$

Now, by the definition of n we have strict inequality in (**17**) with positive probability. Thus $\mathbb{E}(M_{\mathbf{G}}(n)) > \mathbb{E}(M(n))$; a contradiction to the optimality of C. Hence the proof. §

We study the class of greedy policies in two steps. In the first sub-section, we restrict ourselves to the case of $|\mathbf{R}| = \infty$. Under this restriction we show that we have a clean theory - every greedy policy is optimal. In the second sub-section, we show that removal of the above restriction drastically changes the theory. By exploiting the fact that when sampling without replacement in small populations (or close to the end in large populations) the distribution of labels can change drastically from one step to another, we construct examples to paint a contrasting picture. This happens as the greedy criterion fails to make use of this knowledge which allows non-greedy policies to possibly beat greedy policies on specific steps. Nevertheless, we show that every greedy still dominates any alternating.

3.1 Case i: $|\mathbf{R}| = \infty$

Assuming that $|\mathbf{R}| = \infty$ is equivalent to assuming that $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of independent and identically distributed random variables. Under this assumption, we have

$$\mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right] | \mathcal{F}_{n+1}) = \tilde{N}_{\mathbf{S}}\left(S(n)\right)\cdot\tilde{r}, \quad n = 1, 2, \ldots \tag{18}$$

and

$$\mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right] | \mathcal{F}_{n+1}) = \tilde{N}_{\mathbf{R}}\left(R(n)\right)\cdot\tilde{s}, \quad n = 1, 2, \ldots \tag{19}$$

For further analysis it is important to realize that $\tilde{N}_{\mathbf{S}}\left(S(n)\right)\cdot\tilde{r}$ and $\tilde{N}_{\mathbf{R}}\left(R(n)\right)\cdot\tilde{s}$ are both sums of i.i.d. observations. To make this explicit we define

$$X_{\mathbf{R}}\left(n\right) := s_{L_{\mathbf{R}}(n)} \quad \text{and} \quad X_{\mathbf{S}}\left(n\right) := r_{L_{\mathbf{S}}(n)}, \quad n = 1, 2, \ldots \tag{20}$$

The two sequences $\{X_{\mathbf{R}}\left(n\right)\}_{n\geq 1}$ and $\{X_{\mathbf{S}}\left(n\right)\}_{n\geq 1}$ are sequences of i.i.d. random variables with common mean μ and variances $\sigma^2_{\mathbf{R}}$ and $\sigma^2_{\mathbf{S}}$, respectively. We shall denote their partial sums by $\Gamma_{\mathbf{R}}\left[\cdot\right]$ and $\Gamma_{\mathbf{S}}\left[\cdot\right]$, i.e.

$$\Gamma_{\mathbf{R}}\left[n\right] = \sum_{j=1}^{n} X_{\mathbf{R}}\left(j\right) \quad \text{and} \quad \Gamma_{\mathbf{S}}\left[n\right] = \sum_{j=1}^{n} X_{\mathbf{S}}\left(j\right), \quad n = 1, 2, \ldots \tag{21}$$

Now, we can write

$$\tilde{N}_{\mathbf{S}}\left(S(n)\right)\cdot\tilde{r} = \Gamma_{\mathbf{S}}\left[S(n)\right] \quad \text{and} \quad \tilde{N}_{\mathbf{R}}\left(R(n)\right)\cdot\tilde{s} = \Gamma_{\mathbf{R}}\left[R(n)\right], \quad n = 1, 2, \ldots \tag{22}$$

Hence (**15**) simplifies to

$$C(n+1) = \begin{cases} 1 & \text{if } \Gamma_{\mathbf{S}}[S(n)] > \Gamma_{\mathbf{R}}[R(n)]; \\ & \qquad\qquad\qquad\qquad , n = 1, 2, \ldots \\ 0 & \text{if } \Gamma_{\mathbf{S}}[S(n)] < \Gamma_{\mathbf{R}}[R(n)]; \end{cases} \tag{23}$$

We observe that the implementation of the greedy algorithm is greatly facilitated by the representation (23).

It is interesting to note that if the labels are distributed as discrete uniforms in both the sources then the canonical alternating policy turns out to be a greedy policy. This can be seen as

$$\tilde{r} = \tilde{1}/l = \tilde{s} \quad \Longrightarrow \quad \Gamma_{\mathbf{S}}[n] = n/l \text{ and } \Gamma_{\mathbf{R}}[n] = n/l, \quad n = 1, 2, \ldots \tag{24}$$

Moreover, it can be easily seen that this is the only non-degenerate case when an alternating policy is also greedy. This case is also characterized by $\sigma_R + \sigma_S = 0$. In the following, We will assume that $\sigma_R + \sigma_S > 0$ unless stated otherwise.

For our purposes $C_{\mathbf{G}}$ will denote the canonical greedy policy which chooses from the source **R** whenever there is a tie, i.e. whenever $\mathbb{E}(N_{\mathbf{S}}[S(n), L_{\mathbf{R}}(R(n)+1)]|\mathcal{F}_{n+1})$ is equal to $\mathbb{E}(N_{\mathbf{R}}[R(n), L_{\mathbf{S}}(S(n)+1)]|\mathcal{F}_{n+1})$, and whose first pick is from **R**. The following theorem allows us to restrict attention to the canonical greedy policy.

Theorem 3 Under the assumption that $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of independent random variables, all greedy policies lead to the same value of expected matches.

Proof Let us define constants

$$p_{\mathbf{R}} := \sum_{i\in A} r_i, \quad \text{and} \quad p_{\mathbf{S}} := \sum_{i\in A} s_i \quad \text{where } A := \{i | r_i s_i > 0\} \tag{25}$$

respect to $\tilde{s}$ (resp., $\tilde{r}$). And hence the statement that both $p_{\mathbf{R}}$ and $p_{\mathbf{S}}$ are equal to one is the same as saying that the probability vectors $\tilde{r}$ and $\tilde{s}$ are mutually absolutely continuous.

We observe that if $\tilde{r}$ and $\tilde{s}$ are mutually absolutely continuous then the number of matches under two greedy algorithms cannot differ at two or more consecutive steps. And the steps at which they may differ have the same expected gain in the number of matches due to their being greedy. Hence we have the equality of the expected number of matches under two greedy policies when $\tilde{r}$ and $\tilde{s}$ are mutually absolutely continuous.

Now we consider the case where $p_{\mathbf{R}} p_{\mathbf{S}} < 1$. We will need the sequence of stopping times $\{T_i\}_{i\geq 1}$ and $\{T_i^*\}_{i\geq 1}$ defined by

$$T_1 := 0 \quad \text{and} \quad T_i := \min\left\{n \geq T_{i-1}^* \middle| \Gamma_{\mathbf{R}}[R_{\mathbf{G}}(n)] = \Gamma_{\mathbf{S}}[S_{\mathbf{G}}(n)]\right\}, \quad i \geq 2 \tag{26}$$

and

$$T_i^* := T_i + \min\left\{n \geq 1 \middle| L_{\mathbf{R}}\left(R_{\mathbf{G}}\left(T_i\right) + n\right) \in A\right\} + \min\left\{n \geq 1 \middle| L_{\mathbf{S}}\left(S_{\mathbf{G}}\left(T_i\right) + n\right) \in A\right\}, \quad i \geq 2 \quad (27)$$

where $R_{\mathbf{G}}(\cdot)$ and $S_{\mathbf{G}}(\cdot)$ correspond to the canonical greedy policy. Observe that by the definition of a greedy policy it follows that the number of matches for any two greedy policies can differ only between $T_i + 1$ and $T_i^* - 1$. Now it suffices to show that for a non-canonical greedy policy, say $C(\cdot)$ which agrees with the canonical greedy policy after T_2^*, the expected number of matches is the same as that for the canonical greedy policy.

Let us define a sequence of random variables $\{Y_n\}_{n\geq 1}$ such that

$$Y_n = \begin{cases} N_{\mathbf{R}}\left[R_{\mathbf{G}}\left(T_2\right), L_{\mathbf{S}}\left(S_{\mathbf{G}}\left(T_2 + n\right)\right)\right], & \text{if } T_2^* \geq T_2 + n \text{ and } C_{\mathbf{G}}(T_2 + n) = 0 \\ N_{\mathbf{S}}\left[S_{\mathbf{G}}\left(T_2\right), L_{\mathbf{R}}\left(R_{\mathbf{G}}\left(T_2 + n\right)\right)\right], & \text{if } T_2^* \geq T_2 + n \text{ and } C_{\mathbf{G}}(T_2 + n) = 1 \\ 0, & \text{if } T_2^* < T_2 + n \end{cases} \quad (28)$$

and a sequence $\{Z_n\}_{n\geq 1}$ defined similarly for $C(\cdot)$. Observe that

$$M_{\mathbf{G}}(n) - M(n) = \begin{cases} 0, & n \leq T_2 \\ \sum_{i=1}^{n-T_2}\left(Y_i - Z_i\right), & n > T_2 \end{cases} \quad (29)$$

To show that $M_{\mathbf{G}}(\cdot)$ and $M(\cdot)$ have the same expected values it suffices to show that conditioned on

$$\mathcal{G}_{T_2} := \left\langle L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}\left(R_{\mathbf{G}}\left(T_2\right)\right); L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}\left(S_{\mathbf{G}}\left(T_2\right)\right)\right\rangle \quad (30)$$

$\sum_{i=1}^n Y_i$ and $\sum_{i=1}^n Z_i$ have the same expectation for all n. But note that under the natural filtration augmented with $\mathcal{G}_{T_2}$,

$$\sum_{i=1}^{n}\left(Y_i - I_{\{T_2^* \geq T_2 + n\}}\Gamma_{\mathbf{R}}\left[T_2\right]\right) \quad (31)$$

is a zero martingale implying

$$\mathbb{E}\left(\sum_{i=1}^{n} Y_i \middle| \mathcal{G}_{T_2}\right) = \Gamma_{\mathbf{R}}\left[T_2\right]\Pr\left(T_2^* \geq T_2 + n \middle| \mathcal{G}_{T_2}\right) \quad (32)$$

By a similar argument it can be shown that (32) holds true also for the sequence $\{Z_n\}_{n\geq 1}$. This completes the proof. §

Now towards showing the optimality of every greedy policy, the next theorem says that every *non-greedy* policy C, that takes with positive probability a non-greedy step, is inadmissible. To construct a policy which dominates C, we need two random times, T_* and T^*. The random time T_* denotes the last time that the reading policy $C(\cdot)$ takes a greedy step before taking its first non-greedy step. In other words, $C(\cdot)$ takes its first non-greedy step at $T_* + 1$ $C(\cdot)$. Hence T_* can be defined as

$$T_* = \inf\{n \geq 1 |\ (2C(n+1) - 1) * [\Gamma_{\mathbf{S}}[S(n)] - \Gamma_{\mathbf{R}}[R(n)]] < 0\}\,. \tag{33}$$

The random time T^*, always greater than T_*, is the first step beyond T_* that $C(\cdot)$ chooses the source other than that which was chosen at $T_* + 1$. Hence T^* can be defined as,

$$T^* = \inf\{n > T_* |\ C(n) + C(T_* + 1) = 1\}\,. \tag{34}$$

While both T_* and T^* could assume the value ∞, note that T_* is finite with positive probability as C is *non-greedy*.

Using the above two random times we define the, *one more greedy step,* policy $C^\diamond$ as

$$C^\diamond(n) = \begin{cases} C(k), & k \leq T_* \text{ or } k > T^*; \\ 1 - C(T_*), & k = T_* + 1; \\ C(T_*), & T_* + 1 < k \leq T^*; \end{cases} \tag{35}$$

$C^\diamond$ can be described as follows: $C^\diamond$ coincides with C as long as it takes greedy steps; The first time that C takes a non-greedy step, $C^\diamond$ takes the greedy step and then follows a path such that it can couple back with C at the earliest opportunity. The time that C and $C^\diamond$ couple back (this time forever) is $T^* + 1$. In **Figure 1** this is depicted with the assumption, for the purpose of illustration, that $C(T_* + 1)$ is equal to **R**.

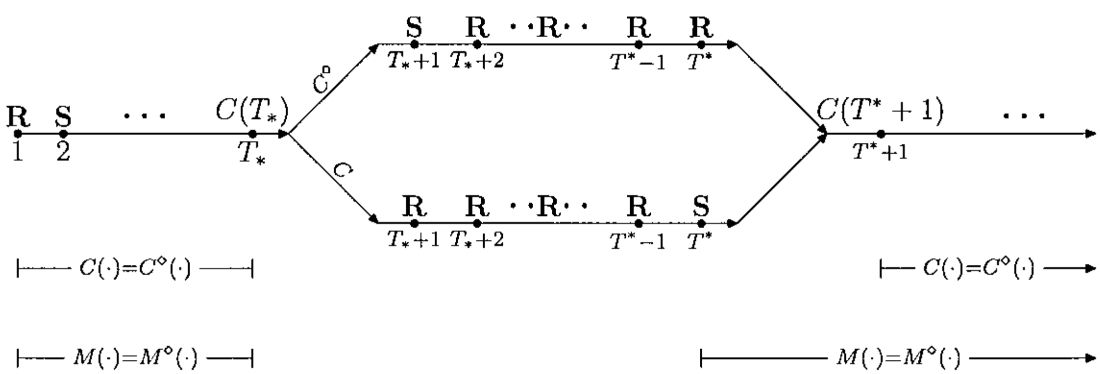


Figure 1 *One More Greedy Step* Policy

In the following $R^\diamond$, $S^\diamond$ and $M^\diamond$ will correspond to $C^\diamond$ whereas R, S and M will correspond to C. Let us define a filtration $\{\mathcal{G}_n\}_{n\geq 0}$ such that

$$\mathcal{G}_n = \mathcal{G}_0 \vee \sigma\left\langle L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}(R^\diamond(n)); L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}(S^\diamond(n))\right\rangle, \quad n = 1, 2, \ldots \tag{36}$$

with $\mathcal{G}_0$ containing all the information needed for randomization not only by $C^\diamond$ but also by C.

Lemma 2 The random times T_* and T^* are both stopping times with respect to the filtration $\{\mathcal{G}_n\}_{n\geq 0}$.

Proof Since C is a reading policy, $C(n+1)$ is measurable with respect to $\mathcal{F}_{n+1}$, see (7). Moreover, by construction, $R^\diamond(T_*)$ and $R^\diamond(T^*)$ are equal to $R(T_*)$ and $R(T^*)$, respectively. Combining these two statements we have that both $\{T_* = k\}$ and $\{T^* = k\}$ are members of $\mathcal{G}_k$. Hence T_* and T^* are $\mathcal{G}$ stopping times. §

As a consequence of the above lemma it should be clear that $C^\diamond$ is adapted, in other words $C^\diamond$ is an *implementable* reading policy.

Theorem 4 Under the assumption that $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of independent random variables, any non-greedy policy is inadmissible.

Proof First, observe that by construction of $C^\diamond$ we have

$$\mathbb{E}(M^\diamond(n) - M(n)|\mathcal{G}_{T_*}) = 0 \quad \text{on} \quad \{n \leq T_*\} \tag{37}$$

and

$$\mathbb{E}(M^\diamond(n) - M(n)|\mathcal{G}_{T_*}) = [2C(T_*+1)-1]\left[\Gamma_{\mathbf{R}}\left[R^\diamond(T_*)\right] - \Gamma_{\mathbf{S}}\left[S^\diamond(T_*)\right]\right] \quad \text{on} \quad \{n = T_*+1\} \tag{38}$$

We shall now show that

$$\begin{aligned}&\mathbb{E}(M^\diamond(n) - M(n)|\mathcal{G}_{T_*})\\ &\geq [2C(T_*+1)-1]\left[\Gamma_{\mathbf{R}}\left[R^\diamond(T_*)\right] - \Gamma_{\mathbf{S}}\left[S^\diamond(T_*)\right]\right]\Pr\left(T^* > n|\mathcal{G}_{T_*}\right) \quad \text{on} \quad \{n > T_*+1\}\end{aligned} \tag{39}$$

By symmetry it suffices to show that (39) holds on the subset $\{n > T_*+1\} \cap \{C(T_*+1) = 1\}$. Toward this end note that on $\{n > T_*+1\} \cap \{C(T_*+1) = 1\}$ we have,

$$M^\diamond(n) - M(n) \geq \left[N_{\mathbf{R}}\left[R^\diamond(T_*), L_{\mathbf{S}}\left(S^\diamond(T_*+1)\right)\right] - N_{\mathbf{S}}\left[S^\diamond(T_*), L_{\mathbf{R}}\left(R^\diamond(n)+1\right)\right]\right] I_{\{T^*>n\}} \tag{40}$$

Now using the fact that on $\{n > T_*+1\} \cap \{C(T_*+1) = 1\}$ the event $\{T^* > n\}$ depends neither on $L_{\mathbf{S}}\left(S^\diamond(T_*+1)\right)$ nor $L_{\mathbf{R}}\left(R^\diamond(n)+1\right)$, we get (39) by taking conditional expectation with respect to $\mathcal{G}_{T_*}$.

Combining (37), (38) and (39) we have $\mathbb{E}(M^\diamond(n))$ is greater than or equal to $\mathbb{E}(M(n))$. But note that by definition of T_* the right hand side of (38) is strictly positive. Hence for all n such that $\Pr\left(T_*+1 = n\right) > 0$ we have that $\mathbb{E}(M^\diamond(n))$ is strictly greater than $\mathbb{E}(M(n))$. But as T_* is finite with positive probability there exist at least one such n. Hence every *non-greedy* reading policy is inadmissible. §

Corollary 1 Under the assumption that $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of independent random variables, the set of all greedy policies is the set of all optimal policies.

Proof Follows by combining the above two theorems. §

To summarize, any greedy policy is optimal and unless the populations are uniformly distributed on the same set of labels, every greedy policy will strictly dominate any alternating.

3.2 Case ii: $|\mathbf{R}| < \infty$

Here we show via examples that the case of $|\mathbf{R}| < \infty$ is very much in contrast to the above. We start with an example which shows that there may exist inadmissible greedy policies. This is unlike the case of $|\mathbf{R}| = \infty$ where all greedy policies were not only admissible but in fact optimal (Corollary **1**). Moreover, it also shows that unlike alternating policies (see Theorem **1**) greedy policies may differ in their expected numbers of matches at some steps.

Example 1 Let the **R** source consist of 8 records with four of them carrying label 1 and the rest label 2. And let the **S** source consists of 7 records with six of them carrying label 1 and one carrying label 2. In the following we will construct two greedy policies and show that one of them strictly dominates the other. Hence showing that not only there may exist inadmissible greedy policies but also demonstrating that greedy policies may differ in their expected numbers of matches at some steps.

The first of the two greedy policies, denoted by $C_{\mathbf{R}}$, is one which at any step where the greedy criterion is ambivalent, i.e. where we have (**16**), chooses **R**. To define the second greedy policy we consider permutations for which the first two records of **R** consist of one record of each label and the first two records of **S** carry label 1. On such permutations all greedy policies after the fourth step would have read two records each from both sources and, more importantly, be ambivalent about the choice of source for the next pick. The second greedy policy, denoted by $C^{\diamond}$, is one which coincides with $C_{\mathbf{R}}$ on all paths except the paths described above. The policy $C^{\diamond}$ will choose **S** on the fifth step unlike $C_{\mathbf{R}}$ and after this step (on this path) will choose **R** on any further step where the greedy criterion is ambivalent.

Now we show that $C_{\mathbf{R}}$ strictly dominates $C^{\diamond}$. For this purpose it is sufficient to restrict attention to the permutations described above. The possible paths followed by $C_{\mathbf{R}}$ and $C^{\diamond}$ on such permutations are shown in Figure **2**. The numbers along the edges are the probabilities and the 2-tuples, for example $(S1;3)$ at the fifth step, stands for a 1 label from **S** resulting in a total of 3 matches (i.e. a gain of 1 match). It is easily checked that both $C_{\mathbf{R}}$ and $C^{\diamond}$ couple on the step after $C_{\mathbf{R}}$ picks a 1 label from **R**. And since the number of matches will coincide after they couple, the paths in Figure **2** have been truncated after a 1 label from **R** is observed under $C_{\mathbf{R}}$. Also, for the same reason the two policies necessarily couple by the ninth step. Hence the expected number of matches coincide from the ninth step onwards and, by construction, until the fifth step. Hence all that remains to be shown is that the expected number of matches for $C_{\mathbf{R}}$ dominates that of $C^{\diamond}$ on the sixth, seventh and eight step.

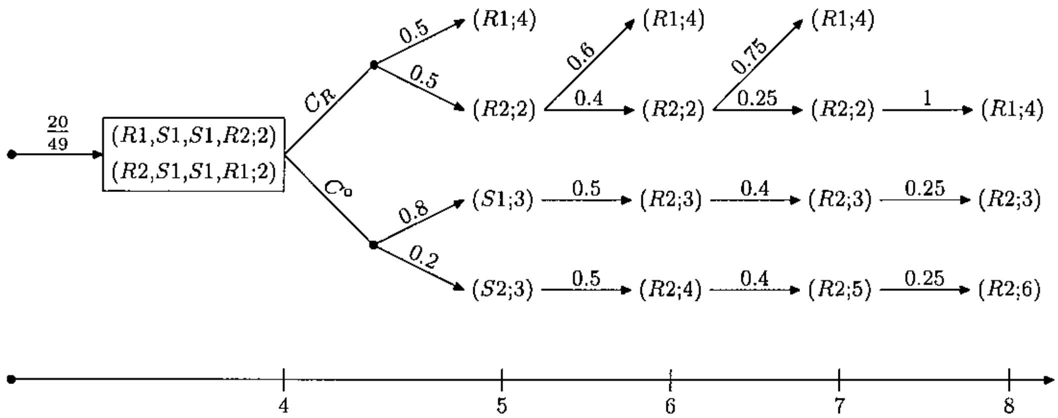


Figure 2 *Greedy Policies Leading to Different Expected Matches*

Now, we explicitly calculate the expected numbers of matches for the two policies at the seventh step. Let $M^\diamond(7)$ and $M^{\mathbf{R}}(7)$ denote the number of matches after the seventh step according to $C^\diamond$ and $C_{\mathbf{R}}$, respectively. From Figure **2** we see that

$$\mathbb{E}\left(M^{\mathbf{R}}(7)\right) = \frac{20}{49} * (0.5 * 0.4 * (0.75 * 4 + 0.25 * 2)) = \frac{14}{49} \tag{41}$$

and

$$\mathbb{E}(M^\diamond(7)) = \frac{20}{49} * (0.8 * 0.5 * 0.4 * 3 + 0.2 * 0.5 * 0.4 * 5) = \frac{13.6}{49}. \tag{42}$$

We conclude by observing that calculations similar to the above show that $\mathbb{E}\left(M^{\mathbf{R}}(6)\right) = \mathbb{E}(M^\diamond(6))$ and $\mathbb{E}\left(M^{\mathbf{R}}(8)\right) > \mathbb{E}(M^\diamond(8))$. §

The above example shows that the choice of greedy policy can affect the outcome. There are two obvious scenarios under which the set of admissible greedy policies has a simple characterization. When an optimal policy exists, every admissible greedy policy is an optimal policy. Moreover, when every non-greedy policy is inadmissible, the collection of admissible greedy policies coincides with that of the admissible policies. In the following example, neither of the above scenarios occur.

Example 2 The sources here are same as that of Example **1**. First, we construct for every greedy policy a corresponding non-greedy policy such that the latter policy is not dominated by the former policy. Thus no greedy policy can be optimal which along with Theorem **2** implies that there is no optimal policy. Second, by working with a particular admissible greedy policy we show that the above corresponding non-greedy policy is admissible. Hence demonstrating that admissible policies may be non-greedy.

In our construction the event $\{R(11) = 6; N_{\mathbf{R}}[11,1] = 3; N_{\mathbf{S}}[11,1] = 5\}$, denoted by A, is key. We observe that under every greedy policy A has a strictly positive probability. This is so as on permutations with the first six labels from $\mathbf{R}$ being $(1,2,1,2,2,1)$ and all of the first five from **S** being 1, all greedy policies will follow the same path from the seventh pick until the thirteenth. In particular, after the eleventh step (on such permutations) they will have picked six records from **R** with three of them being of label 1 and five records from **S** with all of them being of label 1. Note that the permutations specified here is a subset of the permutations of Example **1** and hence the statements made here follow from the discussion in Example **1**.

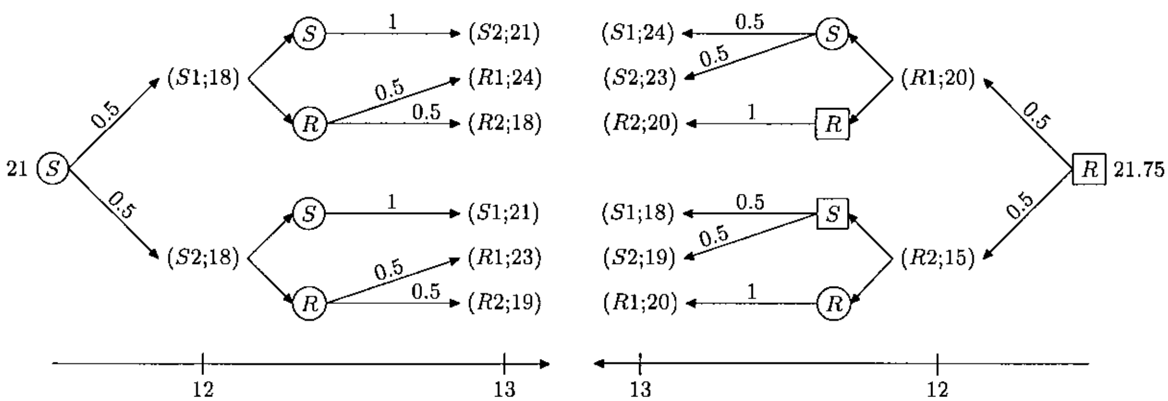
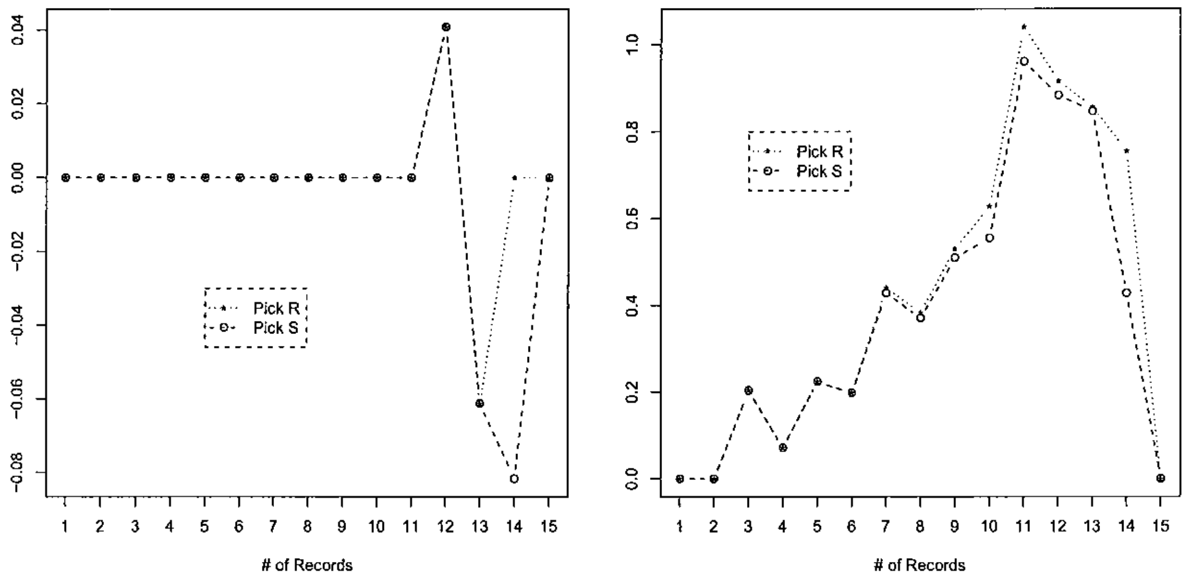


Figure 3 No Greedy is Optimal

The two possible choices at the twelfth step are shown in Figure **3** where the greedy choices are circled and the non-greedy choices are boxed. Otherwise, the conventions are the same as those of Figure **2**. From Figure **3** it can be seen that the twelfth choice for every greedy policy is to pick from **S** followed by any of the two choices for the thirteenth step; this leads to an expected number of 21 matches after the thirteenth pick. Interestingly, going non-greedy at the twelfth step, that is picking a record from the **R** source, improves the expected number of matches after the thirteenth pick by 0.75 (from 21 to 21.75). This happens despite the non-greedy step costing on average 0.5 matches on the twelfth pick (18 v/s 17.5).

To show that no greedy is optimal, let us fix a particular greedy policy, say C. Now define a policy C^* that agrees with C on the complement of A, and on A up to the eleventh step. At the twelfth step on A, C^* makes the non-greedy choice. The difference in the expected numbers of matches at the thirteenth step under C and C^*, by the discussion above, is negative. Hence C^* is not dominated by C which implies that C cannot be optimal. Thus in this example there is no optimal policy.

Figure **4**a plots the difference in the expected number of matches under C and C^* for two particular greedy choices for C - one which always picks from **R** on steps where the greedy criterion is ambivalent and the other which picks from **S** on such steps. Note that the difference is zero up to the eleventh step, is positive at the twelfth and negative at the thirteenth.



a: Greedy may not be Optimal **b**: Every Greedy beats Alternating

Figure 4 Greedy may not be optimal but they still beat the Alternating

Now let C be defined as an admissible greedy policy with the least expected number of matches at the twelfth step among admissible greedy policies. A further requirement on C is that on A, at the thirteenth step it picks from $\mathbf{R}$. And let C^* be the corresponding non-greedy step as defined above. It can be easily argued that such defined C^* is admissible. Hence in this example there exists an admissible non-greedy policy. §

The proof of Theorem **2** suggests that one could possibly extend it to show that a greedy policy can be dominated only by another greedy policy. This, though true in the case of $|\mathbf{R}| = \infty$ (see Corollary **1**), is not generally true in the case of $|\mathbf{R}| < \infty$. In fact, in the example above it can be shown that any greedy policy which on A, picks $\mathbf{S}$ on the thirteenth step can be strictly dominated by a non-greedy policy.

More interesting, especially from the practical point of view, that in spite of greedy policies not being optimal in the above example, they still dominate the alternating policies. This is seen in the plot (Figure 4b) of the difference between the expected number of matches of the two greedy policies of Example **2** and that of an alternating policy. The following theorem shows that this is not particular to the example. First, we define some additional notation that we will need in its proof.

Let $\mathcal{P}(\mathbf{R})$ and $\mathcal{P}(\mathbf{S})$ denote the set of all permutations of $\mathbf{R}$ and $\mathbf{S}$, respectively. And let $\mathcal{P}$ denote $\mathcal{P}(\mathbf{R}) \times \mathcal{P}(\mathbf{S})$. For any $\tilde{L}$, a random variable on $\mathcal{P}$, we define $\tilde{L}_{\mathbf{R}}$ and $\tilde{L}_{\mathbf{S}}$ as $\tilde{L} := \left(\tilde{L}_{\mathbf{R}}, \tilde{L}_{\mathbf{S}}\right)$. Moreover, the components of $\tilde{L}_{\mathbf{R}}$ and $\tilde{L}_{\mathbf{S}}$ are given by

$$\tilde{L}_{\mathbf{R}} := (L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}(|\mathbf{R}|)) \quad \text{and} \quad \tilde{L}_{\mathbf{S}} := (L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}(|\mathbf{S}|)). \tag{43}$$

Define the *partial rotation* operator ϕ_i, which acts on n-vectors ($n \geq i$) as follows:

$$\phi_i(\omega_1, \ldots, \omega_n) = (\omega_1, \ldots, \omega_{i-1}, \omega_n, \omega_i, \ldots \omega_{n-1}) \tag{44}$$

Theorem 5 Every greedy policy is at least as good as any alternating policy.

Proof It suffices to prove the above for greedy policies which do not use any exogenous random variables (for randomization). For any such chosen greedy policy $C_{\mathbf{G}}(\cdot)$ we will show that, for any fixed $n \geq 3$

$$\mathbb{E}(M_{\mathbf{G}}(n)) \geq \mathbb{E}(M_{\mathbf{A}}(n)) \tag{45}$$

where $C_{\mathbf{A}}(\cdot)$ is the canonical alternating policy.

First, we define a reading policy $C_{\mathbf{H}}(\cdot)$ as follows;

$$C_{\mathbf{H}}(k) := \begin{cases} C_{\mathbf{G}}(k), & k \leq T; \\ 0, & T < k \leq n \text{ and } C_{\mathbf{G}}(T) = 1; \\ 1, & T < k \leq n \text{ and } C_{\mathbf{G}}(T) = 0; \\ C_{\mathbf{A}}(k), & k > n; \end{cases} \tag{46}$$

where T is defined as

$$T := \inf\{k : R_{\mathbf{G}}(k) = R_{\mathbf{A}}(n) \text{ or } S_{\mathbf{G}}(k) = S_{\mathbf{A}}(n)\}. \tag{47}$$

$C_{\mathbf{H}}(\cdot)$ is constructed to agree with the alternating policy at the n-th step while following the greedy policy to the maximum extent possible until the n-th step.

Second, we define a function π from $\mathcal{P}$ to itself as follows:

$$\pi(\tilde{\omega}_{\mathbf{R}}, \tilde{\omega}_{\mathbf{S}}) := \left(\left[\prod_{i=T+1}^{n} \phi_i^{C_{\mathbf{G}}(i)} \right]^{C_{\mathbf{G}}(T)} \tilde{\omega}_{\mathbf{R}}, \left[\prod_{i=T+1}^{n} \phi_i^{1-C_{\mathbf{G}}(i)} \right]^{1-C_{\mathbf{G}}(T)} \tilde{\omega}_{\mathbf{S}} \right), \tag{48}$$

where ϕ_i^0 is taken to be the identity operator. Note that $C_{\mathbf{G}}(\cdot)$ is dependent on $(\tilde{\omega}_{\mathbf{R}}, \tilde{\omega}_{\mathbf{S}})$ in (**48**), even though the notation does not make it explicit. Now, let P be the probability on $\mathcal{P}^2$ concentrated on

$$\{(\rho, \pi(\rho)) : \rho \in \mathcal{P}\} \tag{49}$$

such that its first coordinate is uniform on $\mathcal{P}$. It can be shown that π is *one to one* and *onto*; this implies that the second coordinate too is uniform on $\mathcal{P}$ under P. Let $\tilde{L}^1$ and $\tilde{L}^2$ be the random variables representing the first two coordinates under P. We will observe the greedy policy under $\tilde{L}^1$ and the $C_{\mathbf{H}}(\cdot)$ policy under $\tilde{L}^2$.

By the definition of $C_{\mathbf{H}}(\cdot)$ and the coupling we have,

$$\tilde{L}^1_{\mathbf{R}}\left(R_{\mathbf{G}}(k)\right) = \tilde{L}^2_{\mathbf{R}}\left(R_{\mathbf{H}}(k)\right), \quad \text{for } k \le T \text{ or } T < k \le n \text{ and } C_{\mathbf{G}}(T) = 0 = 1 - C_{\mathbf{G}}(k) \tag{50}$$

and

$$\tilde{L}^1_{\mathbf{S}}\left(S_{\mathbf{G}}(k)\right) = \tilde{L}^2_{\mathbf{S}}\left(S_{\mathbf{H}}(k)\right), \quad \text{for } k \le T \text{ or } T < k \le n \text{ and } C_{\mathbf{G}}(T) = 1 = 1 - C_{\mathbf{G}}(k) \tag{51}$$

Moreover, with $\{\mathcal{G}_k\}_{k\ge1}$, a filtration, defined by

$$\mathcal{G}_{k+1} := \sigma\left\langle \tilde{L}^1_{\mathbf{R}}(1), \ldots, \tilde{L}^1_{\mathbf{R}}\left(R_{\mathbf{G}}(k)\right); \tilde{L}^1_{\mathbf{S}}(1), \ldots, \tilde{L}^1_{\mathbf{S}}\left(S_{\mathbf{G}}(k)\right)\right\rangle, \quad k \ge 1 \text{ and } \mathcal{G}_1 = \{\Phi, \Omega\} \tag{52}$$

we have conditioned on $\mathcal{G}_k$, for $k \ge 1$,

$$\tilde{L}^1_{\mathbf{R}}\left(R_{\mathbf{G}}(k)\right) \stackrel{\mathrm{d}}{=} \tilde{L}^2_{\mathbf{R}}\left(R_{\mathbf{H}}(k)\right), \quad \text{on the set } \{C_{\mathbf{G}}(T) = 0 = C_{\mathbf{G}}(k);\ T < k\} \tag{53}$$

and

$$\tilde{L}^1_{\mathbf{S}}\left(S_{\mathbf{G}}(k)\right) \stackrel{\mathrm{d}}{=} \tilde{L}^2_{\mathbf{S}}\left(S_{\mathbf{H}}(k)\right), \quad \text{on the set } \{C_{\mathbf{G}}(T) = 1 = C_{\mathbf{G}}(k);\ T < k\}. \tag{54}$$

If Δ is the usual difference operator, i.e. $\Delta M(k)$ defined by $M(k+1) - M(k)$, we have

$$\mathbb{E}(\Delta M_{\mathbf{G}}(k-1) - \Delta M_{\mathbf{H}}(k-1)|\mathcal{G}_k) = 0, \quad \text{on the set } \{T >= k\};\ k \ge 2. \tag{55}$$

The above equality follows using **(46)**, **(50)** and **(51)**. Also, we have

$$\mathbb{E}(\Delta M_{\mathbf{G}}(k-1) - \Delta M_{\mathbf{H}}(k-1)|\mathcal{G}_k) \ge 0, \quad \text{on the set } \{T < k\};\ k \ge 2. \tag{56}$$

On the subset where $C_{\mathbf{G}}(T) + C_{\mathbf{G}}(k) = 1$, the above equality follows from **(50)** and **(51)** and the fact that the greedy policy has sampled all of the elements (and possibly some additional ones) that have been sampled by the alternating to match the incoming k-th record. On the subset where $C_{\mathbf{G}}(T) + C_{\mathbf{G}}(k) \neq 1$, the above equality follows from **(53)** and **(54)** and the greedy criterion. Combining **(55)** and **(56)**,

$$\mathbb{E}(M_{\mathbf{G}}(n) - M_{\mathbf{A}}(n)) = \mathbb{E}(M_{\mathbf{G}}(n) - M_{\mathbf{H}}(n)) = \mathbb{E}\left(\sum_{i=2}^{n-1} \left[\Delta M_{\mathbf{G}}(k-1) - \Delta M_{\mathbf{H}}(k-1)\right]\right) \ge 0 \tag{57}$$

Hence the proof. §

Towards specifying a condition under which we have strict inequality in **(57)**, let us say that the k-th greedy step is uniquely greedy if the greedy criteria uniquely specifies the choice of the source at this step. Then if, for some $k \le n$, either

$$\Pr\left(R_{\mathbf{G}}(k) > R_{\mathbf{A}}(n), C_{\mathbf{G}}(k) = 1 \text{ and the } k\text{-th step is uniquely greedy}\right) > 0 \tag{58}$$

or

$$\Pr\left(S_{\mathbf{G}}(k) > S_{\mathbf{A}}(n), C_{\mathbf{G}}(k) = 0 \text{ and the } k\text{-th step is uniquely greedy}\right) > 0, \tag{59}$$

we have strict inequality in (57). The above is a rather mild condition which, for example, is satisfied by Example 2. And in the particular case of $n = 3$ it can be easily seen that the alternating, and for that matter every policy in $\mathcal{C}_{NA}$, is strictly beat by any greedy if the labels are not uniformly distributed in both the populations.

In the examples we looked at the case when both the populations are finite. There exist analogous examples in the case when one is of finite cardinality and the other infinite. For example, working with **R** consisting of six records labelled 1 and one labelled 2 and **S** defined as an infinite population consisting of equal proportions of 1 and 2 labelled records. Also, theorem 5 is valid in the case of one (or both) of the populations being infinite.

Finally, to show that a greedy can be optimal, we return to our example from the introduction .

Example 3 Let the **R** and **S** sources each be comprised of n elements labelled $1, ..., n$. Below, we sketch a proof of the claim we made there that alternating is optimal. It is easy to check that the alternating is greedy and all greedy have the same number of expected matches.

Suppose, to the contrary, that the alternating policy is strictly beaten at some step k. Let C be a policy that maximizes the expected number of matches at this step. We say that C takes an excess **R** at step j if $N_R(j) > \lceil k/2 \rceil$, and define excess **S** analogously. Let E_j denote the event where C takes an excess **R** or **S** for the first time, at step j. By definition of C, E_j has a positive probability. Define a second policy C* that duplicates C exactly, except on E_j where it chooses from the opposite sources at both step j and k. It is clear that C and C* couple at step k since, by optimality of C, the k-th step taken by C must be greedy. That is, if C takes an excess **R** (resp., excess **S**) at step j, then it must choose from source **S** (resp., source **R**) at step k. It is also clear that policy C* is implementable. We have thus constructed a policy C* that satisfies $\mathbb{E}(M_{C^*}(n)) = \mathbb{E}(M_C(n))$, but does not take an excess **R** or **S** until after step j. Repeating this argument utmost k times, if necessary, we arrive at a policy C** that satisfies $\mathbb{E}(M_{C^{**}}(n)) = \mathbb{E}(M_C(n))$ and never takes an excess **R** or **S** during its first k steps. For such a policy we have $\mathbb{E}(M_A(n)) = \mathbb{E}(M_{C^{**}}(n)) = \mathbb{E}(M_C(n))$, violating our assumption that the alternating policy is strictly beaten at step k by C. Hence, we have shown that alternating is optimal. §

In the above example the optimality of the alternating can also be proved using techniques of Stochastic Dynamic Programming (SDP) as the alternating policy allows closed form expression of the expected number of matches. As mentioned earlier, in most examples, whether $|\mathbf{R}|$ is finite or infinite, the alternating will not be a greedy policy. In these cases, a greedy does not admit a closed form expression for the expected number of matches. This makes it difficult to establish the optimality via the SDP approach.

4. Conclusion

We have shown that, in general, if the choice is between an alternating policy and a greedy policy then the latter is preferable. But we have assumed that the proportions of labels within the two sources are known, a reasonable assumption. But in the case of unknown distributions, the problem has a statistical component; that of estimating the unknown probabilities. The frequentist approach leads to the maximum likelihood estimators

$$\hat{\hat{r}}_{\text{MLE}}(n+1) = \tilde{N}_{\mathbf{R}}\left(R(n)\right)/R(n) \quad \text{and} \quad \hat{\hat{s}}_{\text{MLE}}(n+1) = \tilde{N}_{\mathbf{S}}\left(S(n)\right)/S(n), \quad n = 1,2,\ldots \tag{60}$$

which possess nice properties such as strong consistency and asymptotic efficiency. It is interesting to note that the use of these estimators results in

$$\tilde{N}_{\mathbf{R}}\left(R(n)\right) \cdot \hat{\hat{s}}_{\text{MLE}}(n+1) = \tilde{N}_{\mathbf{R}}\left(R(n)\right) \cdot \tilde{N}_{\mathbf{S}}\left(S(n)\right)/S(n), \quad n = 1,2,\ldots \tag{61}$$

and

$$\tilde{N}_{\mathbf{S}}\left(S(n)\right) \cdot \hat{\hat{r}}_{\text{MLE}}(n+1) = \tilde{N}_{\mathbf{S}}\left(S(n)\right) \cdot \tilde{N}_{\mathbf{R}}\left(R(n)\right)/R(n), \quad n = 1,2,\ldots \tag{62}$$

which implies that an alternating policy will be a *greedy* policy with MLE estimates of the probabilities. One could alternatively take a Bayesian approach to this estimation problem. Study of the optimality of the resulting reading policy and its asymptotic properties is interesting.

In the case where $|\mathbf{R}| = \infty$ we have shown that the alternating policies are optimal among the non-adaptives, and that the greedy policies are optimal among all policies. When $|\mathbf{R}| < \infty$ we have shown that greedy policies may differ with respect to their expected number of matches, and can even be inadmissible (Example **1**). Moreover, we have shown that some non-greedy policies may be admissible, and that optimal policies may fail to exist (Example **2** and Theorem **2**). Nevertheless, one may safely ignore non-adaptive strategies as every greedy dominates any alternating (alternating being optimal among non-adaptives). And in almost all cases, strictly dominates any alternating.

In the case where $|\mathbf{R}| = \infty$, it would be of interest to explore the asymptotics associated with $M_{\mathbf{G}}$, $M_{\mathbf{A}}$, and also their difference $M_{\mathbf{G}} - M_{\mathbf{A}}$. In the $|\mathbf{R}| < \infty$ case, it would be of interest to find conditions under which an optimal policy is guaranteed to exist. In our introductory card matching problem (analyzed later in Example **3**) we showed that every greedy policy is optimal. It would be interesting to know whether the existence of an optimal policy ensures that every greedy is optimal and, if not, which conditions ensure the optimality of all greedy policies.

References

Berry, D. A. and Fristedt, B. 1985. *Bandit Problems*. Chapman & Hall, New York.

CHANDRASEKARAN, S. and FRANKLIN, M. J. 2002. Streaming Queries over Streaming Data. In *VLDB Conference*, **2002**, 203-214.

DIACONIS, P. and FREEDMAN, D. 1985. Finite Exchangeable Sequences. *Annals of Probability*, **8**, 745-764.

HAAS, P. J. and HELLERSTEIN, J. M. 1999. Ripple Joins for Online Aggregation. In *ACM SIGMOD Conference*, **1999**, 287-298.

ILYAS, I. F., AREF, W. G. and ELMAGARMID, A. K. 2003. Top-k Join Queries in Relational Databases. In *VLDB Conference*, **2003**, 754-765.

LINDVALL, T. 1992. *Lectures on the Coupling Method.* New York:John Wiley & Sons.

LUO, G., ELLMAN, C. J., HAAS, P. J. and NAUGHTON, J. F. 2002. A Scalable Hash Ripple Join Algorithm. In *ACM SIGMOD Conference*, **2002**, 252-262.

VIGLAS, S. D. and NAUGHTON, J. F. 2002. Rate-based Query Optimization for Streaming Information Sources. In *ACM SIGMOD Conference*, **2002**, 37-48.