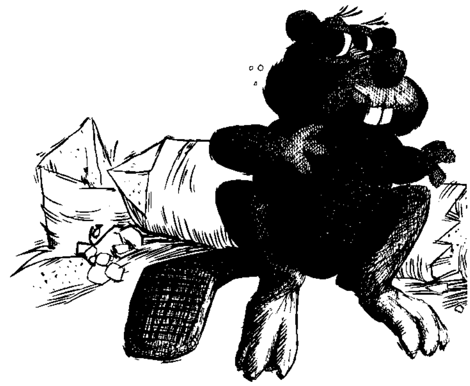
UNIVERSITY OF WATERLOO COMPUTER SCIENCE DEPARTMENT



Right Invariant Metrics and Measures of Presortedness

Vladimir Estivill-Castro
Heikki Mannila
and
Derick Wood

Data Structuring Group
Research Report CS-89-30

July, 1989

Right Invariant Metrics and Measures of Presortedness*

Vladimir Estivill-Castro† Heikki Mannila‡
Derick Wood†

Abstract

Right invariant metrics (*ri-metrics*) have several applications in the theory of rank correlation methods. For example, ranking models based on *ri-metrics* generalize Mallow's ranking models. We explore the relationship between right invariant metrics and measures of presortedness (*mops*). The latter have been used to evaluate the behavior of sorting algorithms on nearly-sorted inputs. We give necessary and sufficient conditions for a measure of presortedness to be extended to a *ri-metric*; we characterize those *ri-metrics* that can be used as *mops*; and we show that those *mops* that are extendible to *ri-metrics* can be constructed from sets of sorting operations. Our results provide a paradigm to construct *mops* and *ri-metrics*.

1 Introduction

Right invariant metrics (*ri-metrics*) on permutations were introduced by Diaconis and Graham [3] as a formal concept that includes natural metrics on the set S_n of permutations and allows relabeling or reordering of the data. Intuitively, *ri-metrics* evaluate the distance between two permutations. By normalizing these metrics, statisticians obtain non-parametric measures

*This work was carried out partially under a Natural Sciences and Engineering Research Council of Canada Grant No.A-5692 and partially under a grant from the Information Technology Research Centre.

†Data Structuring Group, Department of Computer Science, University of Waterloo, WATERLOO, Ontario N2L 3G1, Canada.

‡Department of Computer Science, University of Helsinki, Teollisuuskatu 23, SF-00510 Helsinki, Finland.

of association that have the properties of a rank correlation coefficient [16, page4]. Fligner and Verducci [9] use *ri-metrics* to generalize Mallow's [19] ranking models. *Inv*, *Enc*, *Rem*, $\|\ \|_\infty$, *Ham*, $\|\ \|_p$, and *Grp* are *ri-metrics* appearing in the statistical literature. Given a sequence X of elements from a total order, the sorting problem consists of rearranging the elements in X in ascending order. Computer scientists have been studying the behavior of sorting algorithms on nearly-sorted sequences for some time ([1], [2], [4], [13], [17], [20] and [21]). It is desirable to design sorting algorithms that require computational resources proportional to the amount of disorder in the input. Intuitively, nearly-sorted sequences should be sorted faster than arbitrary sequences. A measure of presortedness (*mop*) evaluates the existing disorder in a sequence, and usually gives an approximation to the number of operations of a certain (and sometimes very obscure) type that need to be performed to sort the sequence.

In this paper, we explore the relationship between *ri-metrics* and *mops*. In Section 3 we give necessary and sufficient conditions for a mop to be extended to a *ri-metric*. If a *mop* can be extended to a *ri-metric*, we say it is normal. We will show how this result applies to *mops* appearing in the computer science literature, namely, *Inv*, *Exc*, *Rem*, *Par*, $\|\ \|_\infty$, *Enc*, *Osc*, *Dist*, m_0, m_{01} and *Runs*. In Section 4 we give necessary and sufficient conditions for a *ri-metric* to be used as a *mop* and in Section 5 we show that normal *mops* are constructed naturally by using sets of sorting transformations.

We use the following notation. Let $X = \langle x_1, x_2, \ldots, x_n\rangle$ be a sequence of length n from some linear order. We denote a *subsequence* of X by $\langle x_{i(1)}, x_{i(2)}, \ldots, x_{i(m)}\rangle$, where $i : \{1, 2, \ldots, m\} \to \{1, 2, \ldots n\}$ is injective and monotonically increasing. We denote the empty sequence by $\langle\rangle$. Let $X = \langle x_1, \ldots, x_n\rangle$ and $Y = \langle y_1, \ldots, y_m\rangle$ be two sequences; then their catenation XY is defined to be $\langle x_1, \ldots, x_n, y_1, \ldots y_m\rangle$. S_n denotes the group of permutations of $\{1, 2, \ldots, n\}$ and *id* is the identity permutation in S_n. The product of two permutations $\pi, \sigma \in S_n$ is denoted by $\pi \cdot \sigma$ and defined by $\pi \cdot \sigma(i) = \pi(\sigma(i))$. If $\pi \in S_n$, then $\langle\pi\rangle$ denotes the sequence $\langle\pi(1), \pi(2), \ldots, \pi(n)\rangle$. For a sequence X, $|X|$ denotes its length, and for a set I, $\|I\|$ denotes its cardinality.

2 Definitions and examples

Statisticians regard metrics on permutations as measures of disarray and normalize these metrics to obtain coefficients of correlation; see [3], [7] [11] and [16]. For example, $\|\sigma, \pi\|_2^2 = \sum_{i=1}^{n} |\pi(i) - \sigma(i)|^2$ is the metric associated with Spearman's coefficient of correlation $\rho = 1 - \frac{6\|\pi,\sigma\|_2^2}{n^3-n}$. Right invariant metrics (*ri-metrics*) were introduced by Diaconis and Graham [3].

Definition 2.1 *$\{d_n\}_{n\in N}$ is a* ri-pseudo-metric, *for $d_n : S_n \times S_n \to \Re$, if there is $c > 0$ such that, for all $n \in N$,*

1. *$d_n(\pi,\sigma) \geq 0$ and $d_n(\pi,\sigma) = 0$ if and only if $\pi = \sigma$,*
2. *$d_n(\pi,\sigma) = d_n(\sigma,\pi)$, for all $\pi,\sigma \in S_n$,*
3. *$d_n(\sigma,\pi) \leq c\cdot\ [d_n(\sigma,\tau) + d_n(\tau,\pi)]$, for all $\pi,\sigma,\tau \in S_n$,*
4. *$d_n(\sigma,\pi) = d_n(\sigma\cdot\tau, \pi\cdot\tau)$, for all $\pi,\sigma,\tau \in S_n$.*

We say that $\{d_n\}_{n\in N}$ is a *ri-metric* if $c = 1$. We will omit the subscript of d_n when this is clear from the context. Kendall's τ, the most popular coefficient of correlation, is defined as $\tau = 1 - 4Inv(\sigma,\pi)/n(n-1)$, where $Inv(\pi,\sigma) =$ the minimum number of pairwise adjacent transpositions required to bring $\langle\pi^{-1}(1),\pi^{-1}(2),\ldots,\pi^{-1}(n)\rangle$ into the order $\langle\sigma^{-1}(1),\sigma^{-1}(2),\ldots\sigma^{-1}(n)\rangle$. In general, for a *ri-metric* d, if the maximum value of d is m, we obtain the coefficient of correlation $1-2d/m$. Fligner and Verducci [9] have studied ranking models based on Cayley's measure and the Hamming distance. Cayley's measure is denoted by Exc, and $Exc(\pi,\sigma)$ is defined as the minimum number of exchanges required to bring $\langle\pi(1),\ldots,\pi(n)\rangle$ into the order $\langle\sigma(1),\ldots\sigma(n)\rangle$. The Hamming distance between permutation π and σ equals the number of positions where the sequences $\langle\pi(1),\ldots,\pi(n)\rangle$ and $\langle\sigma(1),\ldots\sigma(n)\rangle$ differ, and is denoted by $Ham(\sigma,\pi)$.

It is instructive to verify that $Grp(\pi,\sigma) = \|\{i|\text{for all } j,k \text{ such that } 1 \leq j < i \leq k \leq n \text{ we have } \pi\cdot\sigma^{-1}(j) < \pi\cdot\sigma^{-1}(k)\}\|$, which is defined implicitly by Gordon [10], is a *ri-metric.* Other examples of *ri-metrics* are:

1. $\|\pi,\sigma\|_p = (\sum_{i=1}^n |\pi(i)-\sigma(i)|^p)^{1/p}$, $p \geq 1$. ($p = 1$ is the metric associated with Spearman's footrule.)
2. $\|\pi,\sigma\|_\infty = \max_{1\leq i\leq n} |\pi(i) - \sigma(i)|$.
3. $m_0(\pi,\sigma) = 0$, for all π,σ.
4. $m_{01}(\pi,\sigma) = 0$ if $\pi = \sigma$, and $m_{01}(\pi,\sigma) = 1$ otherwise.

Let $X = \langle x_1,\ldots,x_n\rangle$ be a sequence of distinct elements from a total order. X defines a permutation $Pr[X]$ in S_n by

$$Pr[X](i) = \text{final position of } x_i \text{ when } X \text{ is sorted.}$$

Intuitively, we could use a *ri-metric* d to evaluate the disorder in X by defining $m_{Pr}(X) = d(id, Pr[X])$.

Mehlhorn [21] and Guibas et al. [12] have studied sorting algorithms and data structures that perform optimally on presorted inputs with respect to Inv. Cook and Kim [2] and Wainwright [23] have used $Rem(X)$ (the minimum number of elements that need to be removed from X to obtain a sorted sequence) empirically as a measure of presortedness. Other researchers have proposed and used other measures. Mannila [20] has shown that natural-merge sort is optimal with respect to $Runs(X) = \|\{i \| 1 \leq i < n \text{ and } x_{i+1} < x_i\}\|$. The study of lower bounds for parallel sorting algorithms led to the concept of p-sortedness and to the Par measure; see [5] and [14]. X is *p-sorted* if and only if, for all $i, j \in \{1, \ldots, n\}$, $i - j > p$ implies $x_i \geq x_j$, and Par is defined by $Par(X) = \min\{p \mid X \text{ is } p - sorted\}$. Skiena has proposed a measure named Enc [22] which has been adapted in [6] to qualify as a *mop* without modifying its algorithmic properties. Katajainen, Levcopoulos and Peterson [15] and [18], defined

$$Osc(X) = \sum_{i=1}^{|X|} \|cross(x_i)\| \quad \text{and} \quad Dst(X) = \sum_{i=1}^{|X|} \|rcross(x_i)\|,$$

where $cross(x_i) = \{j | 1 \leq j < |X|,\ x_{j+1} < x_i < x_j\}$ and $rcross(x_i) = \{j | i < j < |X|,\ x_{j+1} < x_i < x_j\}$. They studied local-insertion sort and heapsort with respect to these measures.

The functions introduced above are examples of measures of presortedness. The first formal definition was introduced in [20]. We will use the following refined definition; see [6].

Definition 2.2 *Let $N^{<N}$ denote the set of finite sequences of nonnegative integers, and let $m : N^{<N} \to \Re$ be some function. We say that m is a* measure of presortedness *(*mop*) if and only if there is $c > 0$ such that*

1. *If X is in ascending order, $m(X) = 0$.*
2. *If $X = \langle x_1, x_2, \ldots, x_n \rangle$, $Y = \langle y_1, y_2, \ldots, y_n \rangle$ and $x_i \leq x_j$ if and only if $y_i \leq y_j$, for all $i, j \in \{1, 2, \ldots, n\}$, then $m(X) = m(Y)$.*
3. *If Y is a subsequence of X, $m(Y) \leq m(X)$.*
4. *If $X \leq Y$ (that is, every element of X is no greater than every element of Y), then $m(XY) \leq m(X) + m(Y)$.*
5. *If $X \leq Z$, $Y \leq Z$, $W \leq X$, $W \leq Y$ and $m(X) < m(Y)$, then $m(WXZ) \leq m(WYZ)$.*

6. *For all x in N, $m(X\langle x\rangle Y) \leq c|XY| + m(XY)$.*

We note the following technical result about *ri-metrics.*

Lemma 2.1 *If d is a* ri-pseudo-metric, *then*

$$d(id,\pi) = d(\pi^{-1}, id) = d(id, \pi^{-1}); \text{ and}$$

$$d(id, \pi\cdot\sigma) \leq c[d(id,\pi) + d(\pi, \sigma\cdot\pi)] = c[d(id,\pi) + d(id,\sigma)].$$

We abbreviate $d(id,\pi)$ by $d(\pi)$. All the *ri-metrics* introduced above give the corresponding *mop*. For example, $Inv(X) = Inv(id, Pr[X])$ = the number of inversions in X and $Exc(X)$ = the minimum number of arbitrary exchanges required to sort $\langle x_1, \ldots, x_n\rangle$.

3 Mops as ri-metrics

We now give necessary and sufficient conditions for a *mop* to be extended to a *ri-metric.* We will require two technical results.

Lemma 3.1 *If m is a* mop *such that*

1. *$m(X) = 0$ implies X is sorted, and*
2. *there are constants $a, b \geq 0$, such that, for all $n \in N$, and for all $\pi, \sigma \in S_n$, we have $m(\langle \pi\cdot\sigma\rangle) \leq a\, m(\langle\pi\rangle) + b\, m(\langle\sigma\rangle)$,*

then

$$\overline{m}(\pi,\sigma) = \frac{m(\langle\pi\cdot\sigma^{-1}\rangle) + m(\langle\sigma\cdot\pi^{-1}\rangle)}{a+b}$$

is a ri-pseudo-metric.

Proof: We need only verify Condition 3 of Definition 2.1, since Conditions 1, 2 and 4 follow immediately. Now,

$$\begin{aligned}
\overline{m}(\pi,\sigma) &= (m(\langle\pi\cdot\sigma^{-1}\rangle) + m(\langle\sigma\cdot\pi^{-1}\rangle))/(a+b)\\
&= (m(\langle\pi\cdot\tau\cdot\tau^{-1}\cdot\sigma^{-1}\rangle) + m(\langle\sigma\cdot\tau\cdot\tau^{-1}\pi^{-1}\rangle))/(a+b)\\
&\leq (a\, m(\langle\pi\cdot\tau) + b\, m(\tau^{-1}\cdot\sigma^{-1}\rangle)\\
&\quad + a\, m(\langle\sigma\cdot\tau) + b\, m(\tau^{-1}\pi^{-1}\rangle))/(a+b)\\
&\leq \max(a,b)(\overline{m}(\sigma,\tau) + \overline{m}(\tau,\pi)).
\end{aligned}$$

□

The following Corollary follows immediately from Lemma 3.1.

Corollary 3.2 *Under the hypotheses of Lemma 3.1, if* $\max(a,b) \leq 1$, *then* $\overline{m}$ *is a* ri-metric.

In the next definition we describe those *mops* that are extendible to *ri-metrics.*

Definition 3.1 *Let* m *be a* mop. *We say that* m *is* normal *if and only if,*

1. $m(X) = 0$ *implies* X *is sorted,*
2. *for all* $n \in N$, *and for all* $\pi \in S_n$, $m(\langle\pi\rangle) = m(\langle\pi^{-1}\rangle)$, *and*
3. *for all* $n \in N$, *and for all* $\pi, \sigma \in S_n$, $m(\langle\sigma \cdot \pi\rangle) \leq m(\langle\sigma\rangle) + m(\langle\pi\rangle)$.

Normal *mops* are well-behaved measures in the following sense. If we are told that there is no disorder in a sequence, then it is because the sequence is sorted. By applying a permutation σ to a sorted sequence and then applying another permutation τ, we can produce only as much disorder as the disorder produced by each of the permutations σ and τ. Since we only need to apply π^{-1} to sort a permutation π, and we only need to apply π to sort π^{-1}, the disorder in π should be the same as the disorder in π^{-1}. We now show that the conditions in Definition 3.1 are independent.

$1 \wedge 2 \not\Rightarrow 3$. Let $R_2(X) = Runs(\langle Pr[X]\rangle) + Runs(\langle(Pr[X])^{-1}\rangle)$. The reader can verify that R_2 is a *mop* satisfying Conditions 1 and 2 in Definition 3.1; however, it fails Condition 3.

$2 \wedge 3 \not\Rightarrow 1$. Let $m_0(X) = 0$ for all $X \in N^{<N}$. m_0 is a *mop* that satisfies Conditions 2 and 3; however, it does not satisfy Condition 1.

$1 \wedge 3 \not\Rightarrow 2$. Let $m_+(X) = \sum_{i=1}^{|X|}(Pr[X](i) - i)^{1/2}$. m_+ is a *mop* satisfying Conditions 1 and 3 but it does not satisfy Condition 2. The reader may supply the details.

Using Lemma 2.1 and Corollary 3.2 we obtain the following characterization result.

Theorem 3.3 *Let* m *be a* mop. *Let* $\overline{m}(\pi, \sigma) = (m(\langle\pi\cdot\sigma^{-1}\rangle) + m(\langle\sigma\cdot\pi^{-1}\rangle))/2$. $\overline{m}$ *is a* ri-metric *such that* $\overline{m}(id, \pi) = m(\langle\pi\rangle)$ *if and only if* m *is normal.*

Examples of this result are given in the following lemma.

Lemma 3.4 *The* mops m_{01}, *Inv, Grp, Exc, Ham, and Rem are normal.*

Proof: Notice that if d is a *ri-metric* such that $d(X) = d(\pi_X, id)$ is a *mop*, then Lemma 2.1 proves this result for $d(X)$. In particular, this proves the cases m_{01}, *Grp*, *Ham*, *Inv*, and *Exc*. We now give the proof for *Rem*. If $I \subset \{1, \ldots, n\}$, we denote $\{1, \ldots, n\} - I$ by I^c.

[$Rem(\langle \sigma \cdot \pi \rangle) \leq Rem(\langle \sigma \rangle) + Rem(\langle \pi \rangle)$.] Let $Las(X)$ be the length of the largest ascending subsequence of X. Then, $Rem(X) = |X| - Las(X)$. Therefore, $Rem(\langle \sigma \cdot \pi \rangle) \leq Rem(\langle \sigma \rangle) + Rem(\langle \pi \rangle)$ is equivalent to $Las(\langle \pi \rangle) \leq Rem(\langle \sigma \rangle) + Las(\langle \sigma \cdot \pi)$. Let $I = \{i_1, \ldots, i_{Las(\langle \pi \rangle)}\}$ be such that $i_1 < i_2 < \ldots < i_{Las(\langle \pi \rangle)}$ and $\pi(i_1) < \pi(i_2) < \ldots < \pi(i_{Las(\langle \pi \rangle)})$. Let $J = \{j_1, \ldots, j_{Las(\langle \sigma \rangle)}\}$ be such that $j_1 < j_2 < \ldots < j_{Las(\langle \sigma \rangle)}$ and $\sigma(j_1) < \sigma(j_2) < \ldots < \sigma(j_{Las(\langle \sigma \rangle)})$. Let $K = J \cap \{\pi(i_1), \pi(i_2), \ldots, \pi(i_{Las(\langle \pi \rangle)})\}$ and the elements of K be denoted by $\{k_1, k_2, \ldots, k_s\}$ where $k_1 < k_2 < \ldots < k_s$; then, $\pi \cdot \sigma(k_1) < \pi \cdot \sigma(k_2) < \ldots < \pi \cdot \sigma(k_s)$. Therefore, $Las(\langle (\pi \cdot \sigma \rangle) \geq |K| \geq |I| - |J^c| = Las(\langle \pi \rangle) - Rem(\langle \sigma \rangle)$ as claimed.

[$Rem(\langle \pi \rangle) = Rem(\langle \pi^{-1} \rangle)$.] π and π^{-1} have Young tableaux with the same shape [17, Section 5.1.4] and $Las(\langle \pi \rangle)$ is the length of the first row in the Young tableaux for π. □

We conclude that, $\overline{Rem}(id, \sigma) = Rem(\langle \sigma \rangle)$ and the *ri-metric* given in [11] follows from Theorem 3.3.

We now discuss other *mops* appearing in the computer science literature. *Osc* and *Dst* are not normal since there are unsorted sequences X such that $Osc(X) = Dst(X) = 0$. We only state the following result which is easily proved.

Lemma 3.5 *$Par(\langle \sigma \cdot \pi \rangle) \leq Par(\langle \sigma \rangle) + 2\ Par(\langle \pi \rangle)$ and this bound is tight.*

Therefore, *Par* is a *mop* that gives a *ri-pseudo-metric* but does not give a *ri-metric*. *Runs* is a *mop*; however, there are no constants $a, b \geq 0$ such that, for all permutations π and σ,

$$Runs(\langle \pi \cdot \sigma \rangle) \leq a\ Runs(\langle \pi \rangle) + b\ Runs(\langle \sigma \rangle) \tag{1}$$

as the following example shows. Let $n = p(k+1)$ and define

$$\pi(i) = (k - \lfloor (i-1)/p \rfloor)p + [(i-1) \bmod p] + 1, \text{ and}$$

$$\sigma(i) = [(i-1) \bmod (k+1)]p - \lfloor (i-1)/(k+1) \rfloor;$$

then,

$$\pi \cdot \sigma(i) = p(k + 2 - [(i-1) \bmod (k+1) + 1]) - \lfloor (i-1)/(k+1) \rfloor.$$

For $p = 3$ and $k = 5$ this gives

$$\langle\pi\rangle = \langle 16\ 17\ 18\ 13\ 14\ 15\ 10\ 11\ 12\ 7\ 8\ 9\ 4\ 5\ 6\ 1\ 2\ 3\ \rangle,$$

$$\langle\sigma\rangle = \langle 3\ 6\ 9\ 12\ 15\ 18\ 2\ 5\ 8\ 11\ 14\ 17\ 1\ 4\ 7\ 10\ 13\ 16\ \rangle, \text{ and}$$

$$\langle\pi\cdot\sigma\rangle = \langle 18\ 15\ 12\ 9\ 6\ 3\ 17\ 14\ 11\ 8\ 5\ 2\ 16\ 13\ 10\ 7\ 4\ 1\ \rangle.$$

The reader may verify that

$$Runs(\langle\pi\rangle) = k, \quad Runs(\langle\sigma\rangle) = p - 1 \text{ and } \quad Runs(\langle\pi\cdot\sigma\rangle) = pk.$$

Letting $p = \lfloor \log n \rfloor$, (1) would imply that there are constants $a, b > 0$ such that $n - \log n \leq a((n/\log n) - 1) + b(\log n - 1)$, for all n, which is a contradiction. Similarly, it can be shown that *Enc* is not a normal *mop*.

4 Ri-metrics as mops

Conversely to Theorem 3.3, we want to characterize those *ri-metrics* that naturally provide a *mop*. We call these *ri-metrics* regular.

Definition 4.1 *Let* $\{d_n\}_{n\in N}$ *be a* ri-metric. *Let* $m_{Pr}(X) = d_{|X|}(id, Pr[X])$. *We say that* $\{d_n\}_{n\in N}$ *is regular, if and only if,* m_{Pr} *is a* mop *such that, for all* $n \in N$, *and for all* $\sigma, \pi \in S_n$, $\overline{m_{Pr}}(\sigma, \pi) = d_n(\pi, \sigma)$.

Clearly, if $\{d_n\}_{n\in N}$ is regular, then m_{Pr} is a normal *mop*. A *mop* is defined for all sequences of finite length. By Definition 2.2, in a *mop*, the value of a sequence is related to to the values of several types of subsequences. A *ri-metric* has an independent function for each permutation size. It is intuitively clear that in a regular *ri-metric* the function d_n is closely related to d_m for all $m \leq n$. A simple example of a *ri-metric* that is not regular is given by:

$$d_n(\sigma,\pi) = \begin{cases} Exc(\sigma,\pi) & \text{if } n \text{ is even} \\ m_{01}(\sigma,\pi) & \text{if } n \text{ is odd} \end{cases} \quad (\text{where } \sigma,\pi \in S_n).$$

We say that two permutations $\pi, \sigma \in S_n$ *agree on a set* $I \subset \{1,\ldots,n\}$ (denoted $\pi =_I \sigma$) if $i \in I$ implies $\pi(i) = \sigma(i)$. The relation $=_I$ is an equivalence relation. We recall that I^c denotes $\{1,\ldots,n\} - I$. If I is a set of indexes, $I = \{i_1 < i_2 < \ldots < i_s\} \subset \{1,\ldots,n\}$, we denote by $Sb(I, X)$ the subsequence $\langle x_{i_1}, x_{i_2} \ldots x_{i_s}\rangle$ of elements in X with indexes in I.

The following results confirm that in a regular *ri-metric* the d_n are related. They show that if two permutations agree on several entries, their distance depends heavily on the distance between the disagreeing entries.

Lemma 4.1 *If* $\{d_n\}_{n\in N}$ *is a regular* ri-metric, *then, for all* $n \in N$, $I \subset \{1,\ldots,n\}$, $\pi,\sigma \in S_n$, *and* $\pi =_I \sigma$ *implies*

$$d_n(\pi,\sigma) \geq d_{n-|I|}(Pr[Sb(I^c,\langle\pi\rangle)], Pr[Sb(I^c,\langle\sigma\rangle)]).$$

In order to prove this result we present the following proposition. Although the proof of the proposition is not immediate, we omit it, confident that the reader can supply it.

Proposition 4.2 *For all* $I \subset \{1,\ldots,n\}$, *and for all* $\sigma,\pi \in S_n$,

1. $Pr[\langle\pi\rangle] = \pi$ *and* $\langle Pr[\langle\pi\rangle]\rangle = \langle\pi\rangle$.
2. $Pr[Sb(\pi(I),\langle\pi^{-1}\rangle)] = (Pr[Sb(I,\langle\pi\rangle)])^{-1}$.
3. *If* $\pi =_I \sigma$, *then* $Pr[Sb(\sigma(I),\langle\pi\cdot\sigma^{-1}\rangle)] = id$.
4. $Pr[Sb(\sigma(I)^c,\langle\pi\cdot\sigma^{-1}\rangle)] = Pr[Sb(I^c,\langle\pi\rangle)]\cdot Pr[Sb(\sigma(I)^c,\langle\sigma^{-1}\rangle)]$.

Proof of Lemma 4.1 : Let $\{d_n\}_{n\in N}$ be a regular *ri-metric*. The plan of the proof is as follows. First, we use the fact that $\{d_n\}_{n\in N}$ is a *ri-metric* to write $d_n(\pi,\sigma)$ as $d_n(id,\pi\cdot\sigma^{-1})$ which is $m_{P_r}(\langle\pi\cdot\sigma^{-1}\rangle)$. Since m_{P_r} is a *mop* we can use the properties in Definition 2.2. Then the special form of I and the relationship $\pi =_I \sigma$ provide contiguous subsequences or blocks of indexes such that, for all indexes i in a block $\pi\cdot\sigma^{-1}(i) = i$. These blocks of $\langle\pi\cdot\sigma^{-1}\rangle$ are also in their correct relative order, thus we can use the axioms to relate m_{P_r} and blocks. Finally we use Proposition 4.2 to translate this result back to the desired claim in terms of $\{d_n\}_{n\in N}$.

Let $I \subset \{1,\ldots,n\}$, $\pi,\sigma \in S_n$, and $\pi =_I \sigma$. Since $\{d_n\}_{n\in N}$ is a *ri-metric*, $d_n(\sigma,\pi) = d_n(id,\pi\cdot\sigma^{-1}) = m_{P_r}(\langle\pi\cdot\sigma^{-1}\rangle)$. Let $\tau = \pi\cdot\sigma^{-1}$. Let $J = \sigma(I)^c$ and write the elements of J as $\{j_1 < j_2 < \ldots < j_s\}$ where $s = n - |I|$. Let $X = \langle\tau(j_1),\ldots,\tau(j_s)\rangle = Sb(J,\langle\tau\rangle)$; X is a subsequence of $\langle\tau\rangle$. Since m_{P_r} is a *mop*, $m_{P_r}(\langle\tau\rangle) \geq m_{P_r}(X)$. Now, $\{d_n\}_{n\in N}$ is a *ri-metric*, statements 2 and 4 in Proposition 4.2 and since σ is a bijection, imply

$$\begin{aligned} d_n(\pi,\sigma) = m_{P_r}(\langle\tau\rangle) &\geq d_s(id, Pr[X]) \\ &= d_s(id, Pr[Sb(\sigma(I)^c,\langle\pi\cdot\sigma^{-1}\rangle)]) \\ &= d_s(id, Pr[Sb(I^c,\langle\pi\rangle)]\cdot Pr[Sb(\sigma(I)^c,\langle\sigma^{-1}\rangle)]) \\ &= d_s(id, Pr[Sb(I^c,\langle\pi\rangle)]\cdot (Pr[Sb(I^c,\langle\sigma\rangle)])^{-1}) \\ &= d_s(Pr[Sb(I^c,\langle\pi\rangle)], Pr[Sb(I^c,\langle\sigma\rangle)]). \end{aligned}$$

□

The following theorem gives necessary and sufficient conditions for regularity. In a regular *ri-metric*, the axioms in Definition 2.2 translate to conditions that strongly relate the d_n. Statements 1 and 2 show that the value of $d_n(\pi, \sigma)$ is essentially determined by the disagreeing entries. Statement 3 shows that the role of the disagreeing entries implies a certain type of monotonicity, namely, that a qualitative difference in the values of $\{d_n\}_{n \in N}$ on disagreeing entries for d_m with $m < n$ is preserved for d_n. Statement 4 shows that d_n is polynomially bounded since d_n is bounded by a linear combination of n and d_{n-1}.

Theorem 4.3 *$\{d_n\}_{n \in N}$ is a regular* ri-metric *if and only if there is a constant $c > 0$, such that, for all $n \in N$,*

1. *$I \subset \{1, \ldots, n\}$, $\pi \in S_n$, implies $d_n(\pi, id) \geq d_{n-|I|}(Pr[Sb(I^c, \langle \pi \rangle)], id)$.*
2. *$I \subset \{1, \ldots, n\}$, $\pi, \sigma \in S_n$, $\pi =_I \sigma$, $I = I_1 \cup I_2$, $I_1 = \{1, \ldots, u\}$, $I_2 = \{v, v+1, \ldots, n\}$, $u \leq v$, $\pi(I_1) = I_1$ and $\pi(I_2) = I_2$ implies*
$$d_{n-|I|}(Pr[Sb(I^c, \langle \pi \rangle)], Pr[Sb(I^c, \langle \sigma \rangle)] \geq d_n(\pi, \sigma).$$
3. *$I \subset \{1, \ldots, n\}$, $\pi, \sigma, \tau \in S_N$, $\pi =_I \sigma$, $I = I_1 \cup I_2$, $I_1 = \{1, \ldots, u\}$, $I_2 = \{v, v+1, \ldots, n\}$, $u \leq v$, $\pi(I_1) = I_1$, $\pi(I_2) = I_2$, $\tau(I_1) = I_1$, $\tau(I_2) = I_2$, and*
$$d_{n-|I|}(Pr[Sb(I^c, \langle \pi \rangle)], Pr[Sb(I^c, \langle \tau \rangle)]) < d_{n-|I|}(Pr[Sb(I^c, \langle \sigma \rangle)], Pr[Sb(I^c, \langle \tau \rangle)])$$
implies $d_n(\pi, \tau) \leq d_n(\sigma, \tau)$.
4. *$s \in \{1, \ldots, n\}$ and $\pi, \sigma \in S_n$ implies*
$$d_n(\pi, \sigma) \leq cn + d_{n-1}(Pr[Sb(\{s\}^c, \langle \pi \rangle)], Pr[Sb(\{s\}^c, \langle \sigma \rangle)]).$$

Proof: Assume $\{d_n\}_{n \in N}$ is a regular *ri-metric*. It is straightforward to verify that $m_{Pr}(X) = d_{|X|}(id, Pr[X])$ is a mop such that, for all $n \in N$ and for all $\pi, \sigma \in S_n$, $\overline{m_{Pr}}(\sigma, \pi) = d_n(\sigma, \pi)$. We follow the same approach used in the proof of Lemma 4.1. For each statement we use an axiom from Definition 2.2.

1: This is a special case of Lemma 4.1.

2: Let $I \subset \{1, \ldots, n\}$, $\pi, \sigma \in S_n$, $\pi =_I \sigma$, $I = I_1 \cup I_2$, $I_1 = \{1, \ldots, u\}$, $I_2 = \{v, v+1, \ldots, n\}$, $u \leq v$, $\pi(I_1) = I_1$ and $\pi(I_2) = I_2$. By definition, $d_n(\pi, \sigma) = d_n(id, \pi \cdot \sigma^{-1}) = m_{Pr}(\langle \pi \cdot \sigma^{-1} \rangle)$. Let $\tau = \pi \cdot \sigma^{-1}$. Notice that, since $\pi(I_1) = I_1$, if $j \in I_1$, then there is $i \in I_1$ such that $\pi(i) = j$ and $\pi =_I \sigma$ gives $\pi(i) = \sigma(i) = j$. Therefore, $\tau(j) = \pi \cdot \sigma^{-1}(\sigma(i)) = j$, and we conclude that $\tau =_{I_1} id$. Similarly $\tau =_{I_2} id$.

Since m_{Pr} is a *mop*, and $Sb(I_1, \langle\tau\rangle) < Sb(I^c, \langle\tau\rangle) < Sb(I_2, \langle\tau\rangle)$,

$$d_n(\pi, \sigma) = m(\langle\tau\rangle) \leq m_{Pr}(Sb(I_1, \langle\tau\rangle)) + m_{Pr}(Sb(I^c, \langle\tau\rangle) + m_{Pr}(Sb(I_2, \langle\tau\rangle)).$$

By Proposition 4.2 (3), $m_{Pr}(Sb(I_1, \langle\tau\rangle)) = m_{Pr}(Sb(I_2, \langle\tau\rangle)) = 0$. Therefore, $d_n(\pi, \sigma) \leq m_{Pr}(Sb(I^c, \langle\tau\rangle) = d_{n-|I|}(id, Pr[Sb(I^c, \langle\tau\rangle))$. Since $I = \pi(I)$ and $\pi =_I \sigma$, we have $\sigma(I) = I$ and we conclude, using Proposition 4.2,

$$\begin{aligned} d_n(\pi, \sigma) &\leq m_{Pr}(Sb(I^c, \langle\pi \cdot \sigma^{-1}\rangle)) \\ &= d_{n-|I|}(Pr[Sb(I^c, \langle\pi\rangle)], Pr[Sb(I^c, \langle\sigma\rangle)]). \end{aligned}$$

3: Let $I \subset \{1, \ldots, n\}$, $\pi, \sigma \in S_n$, $\pi =_I \sigma$, $I = I_1 \cup I_2$, $I_1 = \{1, \ldots, u\}$, $I_2 = \{v, v+1, \ldots, n\}$, $u \leq v$, $\pi(I_1) = I_1$, $\pi(I_2) = I_2$, $\tau(I_1) = I_1$ and $\tau(I_2) = I_2$. Suppose

$$d_{n-|I|}(Pr[Sb(I^c, \langle\pi\rangle)], Pr[Sb(I^c, \langle\tau\rangle)]) < d_{n-|I|}(Pr[Sb(I^c, \langle\sigma\rangle)], Pr[Sb(I^c, \langle\tau\rangle)]).$$

From $\tau(I) = I$ and Proposition 4.2 this is equivalent to

$$m_{Pr}(Sb(I^c, \langle\pi\tau^{-1}\rangle)]) < m_{Pr}(Sb(I^c, \langle\sigma\tau^{-1}\rangle)]).$$

The shape of I_1 and I_2 with the assumptions $\pi =_I \sigma$, $\pi(I_1) = I_1$, $\pi(I_2) = I_2$, $\tau(I_1) = I_1$ and $\tau(I_2) = I_2$ imply $Sb(I_1, \langle\pi\tau^{-1}\rangle) < Sb(I^c, \langle\pi\tau^{-1}\rangle) < Sb(I_2, \langle\pi\tau^{-1}\rangle)$ and $Sb(I_1, \langle\sigma\tau^{-1}\rangle) < Sb(I^c, \langle\sigma\tau^{-1}\rangle) < Sb(I_2, \langle\sigma\tau^{-1}\rangle)$. Since m_{Pr} is a *mop*, we conclude that $m(\langle\pi\tau^{-1}\rangle) \leq m_{Pr}(\langle\sigma\tau^{-1}\rangle)$. That is,

$$d_n(id, Pr[\langle\pi \cdot \tau^{-1}\rangle]) \leq d_n(id, Pr[\langle\sigma \cdot \tau^{-1}\rangle]),$$

or $d_n(\pi, \tau) \leq d_n(\sigma, \tau)$ as claimed.

4: Let $s \in \{1, \ldots, n\}$, and $\pi, \sigma \in S_N$. Again, $d_n(\pi, \sigma) = m_{Pr}(\langle\pi \cdot \sigma^{-1}\rangle)$. Let $j = \sigma(s)$. Since m_{Pr} is a *mop*, there is $c > 0$ independent of n, such that

$$\begin{aligned} & m_{Pr}(\langle\pi \cdot \sigma^{-1}\rangle) \leq \\ & m_{Pr}(\langle\pi \cdot \sigma^{-1}(1), \ldots, \pi \cdot \sigma^{-1}(j-1), \pi \cdot \sigma^{-1}(j+1), \ldots, \pi \cdot \sigma^{-1}(n)\rangle + c|\pi \cdot \sigma^{-1}| \\ & = m_{Pr}(Sb(\{j\}^c, \langle\pi \cdot \sigma^{-1}\rangle) + cn \\ & = d_{n-1}(Pr[Sb((\{s\}^c, \langle\pi\rangle)], Pr[Sb((\{s\}^c, \langle\sigma\rangle)]) + cn. \end{aligned}$$

Conversely, let $\{d_n\}_{n\in N}$ be a *ri-metric* that satisfies 1, 2, 3 and 4 above. We must show $\{d_n\}_{n\in N}$ is regular. Each condition will correspond to an axiom in Definition 2.2. It is almost direct to verify that $m_{Pr} = d_{|X|}(id, Pr[X])$ is a *mop*, however, some care must be taken, for example, when proving axiom 5 we must include the cases when $|X| \neq |Y|$. We leave the details to the reader. □

5 Constructing normal mops

Sorting can be regarded as a particular case of the following problem. Given a set of valid operations that act on sequences, we are asked to transform an input sequence to a special target sequence. Sequences that require fewer operations are closer to the target sequence. In the context of sorting we say that they are nearly sorted.

A natural measure of the difficulty of the transformation is the minimum number of operations required to perform the transformation. This measure should be symmetric, that is, the number of operations required to transform the input into the target should be the same as to generate the input from the target. In the context of sorting, if we have a sorted file, and we perform a small number of operations on it, the resulting file must be nearly sorted. In this section we present results that show that normal *mops* can be constructed in this way. Consider a permutation $\pi \in S_n$, We can apply π to a sequence $X = \langle x_1, \ldots, x_n\rangle$ to give $X_\pi = \langle x_{\pi(1)}, \ldots, x_{\pi(n)}\rangle$. This captures in a general setting a rearrangement of the elements of X. We are interested in applying sequences of permutations to X so as to sort X. For $n \geq 1$, let $W_n \subset S_n$ be a set of legal sorting transformations on sequences of length n and let $W = \bigcup_{n\geq 1} W_n$. The disorder in a sequence X can be evaluated either as the minimum number of valid sorting operations (in $W_{|X|}$) required to sort X, or as the minimum number of valid operations to introduce the disorder of X in a sorted file. More precisely, we define $m_W, m^W : N^{<N} \rightarrow Z^+ \cup \{0\}$ by

$$m_W(X) = \min\{k \mid \pi_1, \ldots, \pi_k \in W_{|X|} \text{ and } (\ldots(X_{\pi_1})_{\pi 2}\ldots)_{\pi_k} \text{ is sorted}\},$$

and

$$m^W(X) = \min\{k \mid \pi_1, \ldots, \pi_k \in W_{|X|} \text{ and } Pr[X] = \pi_1 \cdot \pi_2 \cdot \ldots \cdot \pi_k\}.$$

Note that $m_W(X) = 0$ if and only if X is sorted, and if and only if $m^W(X) = 0$.

The following result shows that by specifying a set of valid sorting operations W, such that the difficulty of sorting any sequence is the same as the difficulty of generating it with the given set of operations, we obtain a normal *mop*, and therefore a *ri-metric*.

Theorem 5.1 *If $W = \bigcup_{n\geq 1} W_n$ ($W_n \subset S_n$) is such that,*

1. *for all $X \in N^{<N}$, $m_W(X) = m^W(X)$, and*
2. *m_W is a* mop,

then m_W is a normal mop.

The proof of this result follows from the use of the following lemma. The reader should not have any difficulty in verifying it.

Lemma 5.2 *If $\pi, \sigma \in S_n$, then $\langle\pi\rangle_\sigma = \langle\pi\cdot\sigma\rangle$.*

Proof of Theorem 5.1: We only need to show that m_W satisfies Conditions 2 and 3 in Definition 3.1. Let $\pi \in S_n$.

$$\begin{aligned} m_W(\langle\pi\rangle) &= \min\{k \mid \pi_1,\ldots,\pi_k \in W_n \text{ and } (\ldots((\langle\pi\rangle_{\pi_1})_{\pi 2}\ldots)_{\pi_k} \text{ is sorted}\} \\ &= \min\{k \mid \pi_1,\ldots,\pi_k \in W_n \text{ and } \langle\pi\cdot\pi_1\cdot\pi_2\ldots\cdot\pi_k\rangle = \langle id\rangle\} \\ &= \min\{k \mid \pi_1,\ldots,\pi_k \in W_n \text{ and } \pi^{-1} = \pi_1\cdot\pi 2\cdot\ldots\cdot\pi_k\} \\ &= m^W(\langle\pi^{-1}\rangle) = m_W(\langle\pi^{-1}\rangle). \end{aligned}$$

Now, let $\pi, \sigma \in S_n$, and suppose $m_W(\langle\pi\rangle) = k_0$ and $m_W(\langle\sigma\rangle) = k_1$. Since $m_W(\langle\pi\rangle) = m^W(\langle\pi\rangle)$, there are $\pi_1, \pi_2, \ldots, \pi_{k_0} \in W_n$ such that $\pi = \pi_1\cdot\pi_2\cdot\ldots\cdot\pi_{k_0}$. Similarly, there are $\sigma_1, \sigma_2, \ldots, \sigma_{k_0} \in W_n$ such that $\sigma = \sigma_1\cdot\sigma_2\cdot\ldots\cdot\sigma_{k_1}$. Therefore, $\pi\cdot\sigma = \pi_1\cdot\pi_2\cdot\ldots\cdot\pi_{k_0}\cdot\sigma_1\cdot\sigma_2\cdot\ldots\cdot\sigma_{k_1}$. This implies $m^W(\langle\pi\cdot\sigma\rangle) \leq k_0 + k_1$ as required. □

The reader may also verify that both conditions in Theorem 5.1 are necessary. Condition 2 requires m_W to be a *mop*, this basically forces the set of sorting operations to be consistent with all possible lengths. The following lemma gives an alternate, more useful, form of the first condition.

Lemma 5.3 *For all $X \in N^{<N}$, $m_W(X) = m^W(X)$ if and only if, for all $n \in N$, $\pi \in W_n$ implies $\pi^{-1} \in W_n$.*

Conversely, if we are given a normal *mop*, we can almost always identify a set of operations that defines the *mop* up to ranking. More precisely, let m be a *mop* and denote the rank function of m by rk_m and define it by

$$rk_m(X) = \|\{m(\langle\pi\rangle) \mid \pi \in S_n \text{ and } m(\langle\pi\rangle) < m(X)\}\|.$$

The function rk_m scales the *mop* to nonnegative integers preserving the property that it evaluates to zero on sorted sequences. Moreover, rk_m also preserves the algorithmic properties of m since the "below" sets are the same; see [20]. Notice that if m_W is a *mop* that satisfies the conditions of Theorem 5.1, it is already normalized and, therefore, $rk_{m_W} = m_W$. Moreover, if $\pi \in S_N$ and $\tau \in W_n$, then $rk_m(\langle\sigma\rangle) - 1 \leq rk_m(\langle\sigma \cdot \tau\rangle) \leq rk_m(\langle\sigma\rangle) + 1$.

Theorem 5.4 *Let m be a normal* mop *such that, for all $n \in N$, $\sigma \in S_n$ and $\tau \in \{\pi \in S_n \mid rk_m(\langle\pi\rangle) = 1\}$ implies $rk_m(\langle\sigma\rangle) - 1 \leq rk_m(\langle\sigma \cdot \tau\rangle) \leq rk_m(\langle\sigma\rangle) + 1$. Then, there are $W_n \subset S_n$ such that,*

1. *for all $X \in N^{<N}$, $rk_m(X) = m_W(X) = m^W(X)$, and*

2. *rk_m is a normal* mop.

Moreover, $W_n = \{\pi \in S_n \mid rk_m(\langle\pi\rangle) = 1\}$.

The proof of the result is not immediate. It requires a careful induction on the value of rk_m and the use of results that rely heavily on the axioms for *mops*. For example, the following result shows that, if m is a normal *mop*, then $W_n = \{\pi \in S_n \mid rk_m(\langle\pi\rangle) = 1\}$ is always a set of generators of S_n.

Lemma 5.5 *Let m be a normal* mop *and $W_n = \{\pi \in S_n \mid rk_m(\langle\pi\rangle) = 1\}$. If $\pi \in S_n$ is a transposition of adjacent elements, then $\pi \in W_n$.*

Proof: Let $\langle\pi\rangle = \langle 1, \ldots, i-1, i+1, i, i+2, \ldots, n\rangle$. $m(\langle\pi\rangle) \neq 0$ since $\langle\pi\rangle$ is not sorted and m is a normal *mop*. $m(\langle\pi\rangle) > m(\langle i+1, i\rangle) = m(\langle 2, 1\rangle)$ by axioms 3 and 2 in Definition 2. Moreover, $m(\langle\pi\rangle) \leq m(\langle 1, \ldots, i-1\rangle) + m(\langle i+1, i\rangle) + m(\langle i+1, \ldots, n\rangle)$ by axiom 4. Therefore, $m(\langle\pi\rangle) = m(\langle 2, 1\rangle)$. Since, for any nonsorted sequence X, $m(X) \geq m(\langle 2, 1\rangle)$ by axioms 3 and 2, we conclude that $rk_m(\langle\pi\rangle) = 1$. □

Sketch of proof of Theorem 5.4: Let $W_n = \{\pi \in S_n \mid rk_m(\langle\pi\rangle) = 1\}$, for $n \geq 1$. First, Lemma 5.3 and m is a normal *mop* are used to prove that, for all $X \in N^{<N}$, $m_W(X) = m^W(X)$. Now, use Lemma 5.5 and induction on t on the following predicate. If $rk_m(X) = t$, then

1. $m_W = m^W(X) = t$,

2. if $t > 0$, there is a $\tau \in W_{|X|}$, such that $t - 1 = rk_m(\langle Pr[X] \cdot \tau]\rangle)$, and

3. if $t \leq \max(m_W(\langle\pi\rangle) | \pi \in S_{|X|})$, there is a $\tau \in W_{|X|}$, such that $t + 1 = rk_m(\langle Pr[X] \cdot \tau]\rangle)$.

The induction proves (1). To prove (2), first verify the axioms in Definition 2.2 to prove rk_m is a *mop*. This will require using that m is a *mop*. For example, to prove that if X is a subsequence of Y, then $rk_m(X) \leq rk_m(Y)$ we argue as follows. It is enough to show the claim for $|X|+1 = |Y|$, since a proof by induction gives the general case. Since m is a *mop*, $m(X) \leq m(Y)$. Let X' be such that $|X'| = |X|$ and $m(X') < m(X)$. We then take $x = |X'| + 1$, and the catenation of $\langle Pr[X'] \rangle$ with $\langle x \rangle$. Observe that, since $\langle Pr[X'] \rangle$ is a subsequence of $Y' = \langle Pr[X'] \rangle \langle x \rangle$,

$$m(\langle Pr[X'] \rangle) \leq m(\langle Pr[X'] \rangle \langle x \rangle).$$

By axiom 2, m only depends on the relative order of the elements, therefore, $m(\langle Pr[X'] \rangle) = m(X')$. By axiom 4, $m(\langle Pr[X'] \rangle \langle x \rangle) \leq m(\langle Pr[X'] \rangle) + 0$. We conclude that $m(Y') = m(X') < m(X) \leq m(Y)$. Since we have shown that for each value $k < m(X)$ such that there is an X' with $|X'| = |X|$ and $m(X') = k$, we can find a Y' with $|Y'| = |Y|$ and $m(Y') = m(X') = k$, we conclude that $rk_m(X) \leq rk_m(Y)$ as claimed.

The final step of the proof is now simple. We know that $m_W = m^W = rk_m$, and that rk_m is a *mop*; using Theorem 5.1 we conclude that rk_m is normal. □

The hypothesis in Theorem 5.4 may seem restrictive, but it cannot be removed. For example, $m(X) = \|Pr[X], id\|_2$ is a normal *mop*, but, if $\langle \pi \rangle = \langle 1, 3, 2, 4, \ldots, n \rangle$, $\langle \sigma \rangle = \langle 1, 3, 4, 2, 5, \ldots, n \rangle$, and $\langle \tau \rangle = \langle 1, 3, 2, 5, 4, 2, 6, \ldots, n \rangle$ we must have $rk_{\| \|_2}(\langle \sigma \rangle) > rk_{\| \|_2}(\langle \tau \rangle) > rk_{\| \|_2}(\langle \pi \rangle)$. Since σ and τ are the product of two transpositions of adjacent elements, we must have $1 < m_W(\langle \sigma \rangle) \leq 2$ and $1 < m_W(\langle \tau \rangle) \leq 2$ for any set W. Thus $rk_{\| \|_2} = m_W$ gives a contradiction.

6 Concluding remarks

The connections we have shown between *mops* and *ri-metrics* raise several open questions. Researchers have attempted to compare *ri-metrics* by establishing inequalities between them; see, for example [3]. Using the axioms in Definition 2.2 the reader may verify the following lemma.

Lemma 6.1 *If m is a* mop *that satisfies the two hypotheses of Lemma 3.1, then there is a $K > 0$ such that, for all $X \in N^{<N}$,*

$$m(X) \leq K \cdot Inv(X).$$

This implies, for example, that any sorting algorithm that is sensitive to m, (the smaller the value of m the less time is spent by the sorting algorithm) is also sensitive with respect to Inv. Although Inv plays an important role among *ri-metrics* and *mops*, its relevance is not fully understood. Notice that Inv is defined from a set of operations that includes exactly all transpositions of adjacent elements. Lemma 5.5 shows that Inv is the normal *mop* (*ri-metric*) most sensitive to disorder. The popularity of Kendall's τ is due to the fact that Inv is asymptotically normally distributed with known mean and variance for each n. Our results show that normal *mops* (*ri-metrics* and coefficients of correlation) can be constructed in a similar way.

Any linear combination, with nonnegative coefficients, of *ri-metrics* is a *ri-metric*. Also any such linear linear combination of *mops* is a *mop*; see [6]. Moreover, any linear combination, with nonnegative coefficients, of normal *mops* is a normal *mop*, thus the space of normal *mops* has constructive mathematical properties. From the practical point of view, the characteristics of the distributions of the *ri-metrics* provided by Theorem 3.3 must be described analytically or by a tabulation of their values. Analytical results may be difficult, as suggested by Ulam's problem (computing the limiting behavior of the expected value of Rem). At least, it is desirable to characterize those *ri-metrics* that decompose into a sum of independent uniform distributions or other well known distributions.

Finally, if we want to test correlation or agreement of more than two rankings (because the objects are ranked independently by boards of judges), the corresponding techniques must be developed as shown in [7], [8] and [9].

References

[1] W.H. Burge. Sorting, trees and measures of order. *Information and Control*, 1:181–197, 1958.

[2] C.R. Cook and D.J. Kim. Best sorting algorithms for nearly sorted lists. *Communications of the ACM*, 23:620–624, 1980.

[3] P. Diaconis and R. L. Graham. Spearman's footrule as a measure of disarray. *J. Royal Statistical Soc. Series B*, 32(2):262–268, 1977.

[4] P.G. Dromey. Exploiting partial order with Quicksort. *Software-Practice and Experience*, 14(6):509–518, 1984.

[5] V. Estivill-Castro and D. Wood. A new measure of presortedness. *Information and Computation*, to appear, 1989.

[6] V. Estivill-Castro and D. Wood. What is a measure of presortedness. Technical Report In preparation, Department of Computer Science, University of Waterloo, 1989.

[7] P.D. Feigin and M. Alvo. Intergroup diversity on concordance for ranking data: An approach via metrics for permutations. *The Annals of Statistics*, 14:691–707, 1986.

[8] P.D. Feigin and A. Cohen. On a model for concordance between judges. *J. Royal Statistical Soc. Series B*, 40:203–213, 1978.

[9] M.A. Fligner and J.S. Verducci. Distance based ranking models. *J. Royal Statistical Soc. Series B*, 48:359–369, 1986.

[10] A.D. Gordon. Another measure of the agreement between rankings. *Biometrika*, 66(2):327–332, 1979.

[11] A.D. Gordon. A measure of the agreement between rankings. *Biometrika*, 66(1):7–15, 1979.

[12] L.J. Guibas, E.M. McCreight, and M.F. Plass. A new representation of linear lists. In *9th ACM Annual Symposium on Theory of Computing*, pages 49–60, 1977.

[13] S. Hertel. Smoothsort's behaviour on presorted sequences. *Information Processing Letters*, 16:165–170, 1983.

[14] Y. Igarashi and D. Wood. Roughly sorting — a generalization of sorting. Technical Report CS-87-55, Department of Computer Science, University of Waterloo, 1987.

[15] J. Katajainen, C. Levcopoulos, and O. Petersson. Local insertion sort revisited. *Unpublished manuscript*, 1989.

[16] M.G. Kendall. *Rank Correlation Methods*. Griffin, London, 4th edition, 1970.

[17] D.E. Knuth. *The Art of Computer Programming, Vol.3: Sorting and Searching*. Addison-Wesley Publishing Co., Reading, Mass., 1973.

[18] C. Levcopoulos and O. Petersson. Heapsort—adapted for presorted files. *Unpublished manuscript*, 1989.

[19] C.L. Mallows. Non-null ranking models. *Biometrika*, 44:114–130, 1957.

[20] H. Mannila. Measures of presortedness and optimal sorting algorithms. *IEEE Transactions on Computers*, C-34:318–325, 1985.

[21] K. Mehlhorn. Sorting presorted files. *4th GI Conference on Theory of Computer Science*, Springer-Verlag Lecture Notes in Computer Science 67:199–212, 1979.

[22] S.S. Skiena. Encroaching lists as a measure of presortedness. Technical report, Department of Computer Science, University of Illinois, Urbana, Illinois 61801, January 1988.

[23] R.L. Wainwright. A class of sorting algorithms based on Quicksort. *Communications of the ACM*, 28:396–402, 85.