SORTING A BRIDGE HAND

HENRIK ERIKSSON, KIMMO ERIKSSON, JOHAN KARLANDER,
LARS SVENSSON AND JOHAN WÄSTLUND

ABSTRACT. Sorting a permutation by block moves is a task that every bridge player has to solve every time she picks up a new hand of cards. It is also a problem for the computational biologist, for block moves are a fundamental type of mutation that can explain why genes common to two species do not occur in the same order in the chromosome. The minimum number of block moves needed to get from one species to the other is a parameter that biologists ask for.

It is not known if an optimal sorting procedure can be found in polynomial time, but Bafna and Pevzner gave an algorithm that sorts any permutation of length n in $3n/4$ moves. Our new algorithm improves this to $\lfloor (2n-2)/3 \rfloor$ for $n \geq 9$. For the reverse permutation, we give an exact expression for the number of moves needed, namely $\lceil (n+1)/2 \rceil$. This was conjectured to be the worst case, but we can now refute this conjecture. The first counterexample occurs for $n = 13$, i.e. the bridge player's case.

Professional card players never sort by rank, only by suit. For this case, we give a complete answer to the optimal sorting problem.

1. INTRODUCTION

Considering the enormous literature on bridge bidding and play, it is only right that the phase preceding the bidding should get its proper analysis. Each player is dealt thirteen cards face-down on the table, picks them up, has a quick look and starts rearranging the hand. Most bridge players use block transpositions (removing a sequence of cards and putting them back in another place) to rearrange their thirteen cards in some preferred order. Empirically, from diligent bridge playing or computer simulations, one finds that most hands can be sorted in six moves but that about 30% need seven moves. One of these is the reverse permutation, intuitively felt to be the worst case. However, it is a real challenge to sort the permutation [13 12 11 10 9 8 7 6 5 4 3 2 1] in seven block moves, and the fearless reader is invited to take on that challenge before reading further! It is hardly feasible to let the computer check all 13! permutations, so some analysis is needed to find out what is really the worst case. The unexpected answer will be given in section 6.

When you play a card from a sorted bridge hand, your opponents may draw some information from the position of the card. Therefore, professional card players never order their cards by rank, only by suit. The corresponding optimal sorting problem is easier and gets a complete analysis in section 6. It turns out that the bridge player can always separate suits in six block moves.

The serious applications are to be found in Bioinformatics. For the details, we refer to [1], but the gist of the matter is that block transpositions occur in gene sequences as rare mutation events. The genome breaks in three places and the two middle pieces are glued back transposed. The applications that we have in mind are genome pairs with some genes in common, but in different order. An approximate algorithm for finding the shortest sequence of block moves leading from one gene order to the other will trace out a plausible evolutionary path between the two species. This path will lead backwards in time from the first species to a common ancestor and then continue forwards in time to the other species.

Bafna and Pevzner devised a sorting algorithm with a worst case performance of about $3n/4$ block moves. Our block move sorting algorithm has a better worst case performance, namely asymptotically $2n/3$.

2. Notation and definitions

We will denote a permutation in S_n by its sequence of permuted numbers within brackets:

$$\pi = [\pi_1\ \pi_2\ \ldots\ \pi_{n-1}\ \pi_n].$$

For any three cut points $0 \leq i < j < k \leq n$, define the *block move* σ_{ijk} by

$$\sigma_{ijk} = [1\ \ldots i\ j{+}1\ \ldots\ k\ i{+}1\ \ldots\ j\ k{+}1\ \ldots n].$$

This may also be called a *block transposition*, as two adjacent blocks have been transposed.

Composition of permutations is defined as action to the right:

$$\pi \cdot \sigma_{ijk} = [\pi_1\ \ldots \pi_i\ \pi_{j+1}\ \ldots\ \pi_k\ \pi_{i+1}\ \ldots\ \pi_j\ \pi_{k+1}\ \ldots \pi_n].$$

For convenience, we introduce symbols for two permutations of fundamental importance, the identity and the reverse permutation:

$$\mathrm{id} \stackrel{\mathrm{def}}{=} [1\ 2\ \ldots\ n{-}1\ n] \quad \text{and} \quad w_0 \stackrel{\mathrm{def}}{=} [n\ n{-}1\ \ldots\ 2\ 1].$$

2.1. Toric model of permutations.

We can extend an ordinary permutation π to a *circular permutation* π° by inserting an extra element 0 as both predecessor of π_1 and successor of π_n. We write

$$\pi^\circ = 0\ \pi_1\ \pi_2\ \ldots\ \pi_n$$

where the absence of brackets indicates an equivalence class under cyclic shifts. For example, $0312 = 3120 = 1203 = 2031$. From a circular permutation π°, we uniquely retrieve the ordinary permutation π by removing the element 0 and letting its successor be the first element of π.

A block move acts on a circular permutation by cutting it into three segments which are then glued together in the other possible order. This is already a slightly nicer setting for our original sorting problem. But we can go one step further and consider *toric permutations*, which are circular in values as well as in positions. An m-step cyclic value shift of π° is defined as

$$m{+}\pi^\circ = m \;\; m{+}\pi_1 \;\; m{+}\pi_2 \ldots m{+}\pi_n \qquad (\mathrm{mod}\; n{+}1)$$

and the equivalence class of π° under such value shifts is the toric permutation $\pi^\circ_\circ$.

The point of all this is that a strategy for block sorting π° will work also for $m+\pi^\circ$: if a move sequence takes π° to id°, then analogous moves take $m+\pi^\circ$ to $m+\mathrm{id}^\circ$. However, $m+\mathrm{id}^\circ = \mathrm{id}^\circ$. Example:

$$\begin{aligned} \pi^\circ &= 0312 \\ 1+\pi^\circ &= 1023 \\ 2+\pi^\circ &= 2130 \\ 3+\pi^\circ &= 3201 \end{aligned}$$

So $[312]^\circ_\circ = [231]^\circ_\circ = [213]^\circ_\circ = [132]^\circ_\circ$ and therefore a block sorting strategy for [312] can be translated into a strategy for any of the other three permutations.

It is convenient to let $\overline{x}$ denote the numerical successor of x in a toric permutation, i.e. $\overline{x} = x+1 \pmod{n{+}1}$. Similarly, we let $\underline{x} = x-1 \pmod{n{+}1}$. An occurrence of $x\overline{x}$ is called a *bond*, and it is clear that bonds need never be broken in an optimal sorting strategy. An occurrence of $\overline{x}x$ is called an *anti-bond*. Circularity in positions and values must always be taken into account; thus 314052 has one bond (23) and one anti-bond (05).

In a toric permutation, we say that an ordered triple of values $x \ldots y \ldots z$ is *positively oriented* if either $x < y < z$ or $y < z < x$ or $z < x < y$.

The justification for the term *toric* is the following. An ordinary permutation has a geometric representation as a square matrix with n rows and n columns and with n dots, one in each row and each column. Joining the two vertical sides of the square, we get a cylinder representing a circular permutation. Joining also the two horizontal sides, we get a torus representing a toric permutation.

3. The Cayley graph of block transpositions

The symmetric group S_n is generated by the set of all block moves. Hence we can define the so called Cayley graph, with vertex set S_n and a directed edge labeled σ_{ijk} from any $\pi \in S_n$ to $\pi \cdot \sigma_{ijk}$. A block sorting strategy for π is a directed path from π to id. Reading the labels of the path, we get the identity $\pi\sigma_1\sigma_2 \cdots \sigma_\ell = \mathrm{id}$.

Let $d(\pi)$ denote the distance from id to π in the Cayley graph, i.e. the minimal number of block moves needed to sort π.

3.1. **Inverses.** The inverse of a block move is also a block move:

$$\sigma_{ijk}^{-1} = \sigma_{irk},$$

where $r = k + j - i$. This means that the block sorting strategy for π can be regarded as a factorization of π into block moves: $\pi = \sigma_\ell^{-1} \cdots \sigma_2^{-1}\sigma_1^{-1}$. We may as well let the pair of inversely directed edges merge into one undirected edge, thus obtaining an undirected graph. By merging vertices representing the same toric permutation, we obtain the *toric graph*. This seems to be the correct mathematical object for our investigation.

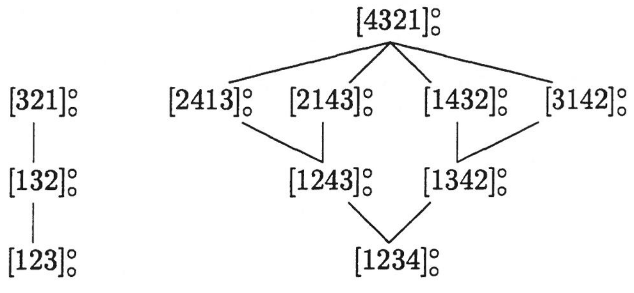

FIGURE 1. The toric graphs for $n = 3$ and $n = 4$.

Note that the inverse is also well defined for toric permutations, for it is easy to see that if π and τ represent the same toric permutation, then π^{-1} and τ^{-1} represent the same toric permutation.

3.2. **The diameter of the Cayley graph.** The *diameter* of a graph is the maximal distance between two vertices. As Cayley graphs look exactly the same seen from any vertex, the diameter is the maximal distance from id to any π. It is also the diameter of the toric graph, and it tells us how many block moves are needed, in the worst case, to sort a permutation π. We shall denote this diameter, for a given n, by

$$d(n) \stackrel{\text{def}}{=} \max_{\pi \in S_n}\{d(\pi)\}.$$

It was observed already by Bafna and Pevzner [1] that $d(n) = \lceil \frac{n+1}{2} \rceil$ for $3 \leq n \leq 10$. We started working on this problem assuming that this expression for $d(n)$ would hold for all $n \geq 3$. Bafna and Pevzner proved that the value of $d(n)$ lies in the interval $\lceil \frac{n-1}{2} \rceil \leq d(n) \leq \lfloor \frac{3n}{4} \rfloor$.

Our agenda in this paper is the following:

1. For the reverse permutation we give an exact value: $d(w_0) = \lceil \frac{n+1}{2} \rceil$ for $n \geq 3$. This of course implies an improved lower bound: $d(n) \geq \lceil \frac{n+1}{2} \rceil$.
2. We improve the upper bound to $d(n) \leq \lfloor \frac{2n-2}{3} \rfloor$.
3. By a combination of computer work and theoretical arguments, we are able to determine $d(n)$ for $n \leq 15$ (Table 1).

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$d(n)$	1	2	3	3	4	4	5	5	6	6	7	8	8	9

TABLE 1. Known values of $d(n)$.

In particular, this solves the bridge player's problem and shows that our working conjecture $d(n) = \lceil \frac{n+1}{2} \rceil$ was wrong. Note that our new upper bound $d(n) \leq \lfloor \frac{2n-2}{3} \rfloor$ is sharp for $9 \leq n \leq 15$.

4. An improved lower bound and the reverse permutation

Recall that a *descent* in a permutation π is an occurrence of $\pi_k \pi_{k+1}$, such that $\pi_k > \pi_{k+1}$. Although for a toric permutation the notion of descent makes no sense, the *number of descents* still has meaning; it is easy to see that if π and τ represent the same toric permutation, then π and τ have the same number of descents.

Lemma 4.1. *The number of descents can decrease by at most two in every move.*

Proof. Obviously, the number of descents can change by at most three in every move. We shall see that a decrease by three is in fact impossible, since it would demand a permutation of the following form:

$$[\dots a\, b \;\dots\; c\, d \;\dots\; e\, f \dots],$$

where $a > b$, $c > d$ and $e > f$, and where all three descents are broken by a move, giving

$$[\dots a\, d \;\dots\; e\, b \;\dots\; c\, f \dots],$$

with no new descents, so that $a < d$, $e < b$ and $c < f$. Together, the six inequalities imply that

$$a < d < c < f < e < b < a,$$

which is absurd. The argument is still valid if $b = c$, $d = e$ or $f = a$. □

4.1. **An optimal sorting algorithm for w_0.** Now we return to the bridge player, faced with the problem to reverse the order of her cards using only seven block moves. For simplicity, we show the solution for $n = 7$ instead, but the algorithm works for any n.

$$[7654321] \to [43\ 76521] \to [4523\ 761] \to [456123\ 7] \to [1234567]$$

Theorem 4.2. *For $n \geq 3$, the reverse permutation w_0 can be sorted in $\lceil \frac{n+1}{2} \rceil$ block moves, and this is optimal.*

Proof. It is sufficient to give an algorithm for odd $n = 2k+1$ using $k+1$ moves, for if we have an even $n = 2k$, we can use the algorithm for $n = 2k+1$, forgetting about one of the elements.

Algorithm: We can sort $w_0 = [n \dots\ 1\ 0]$ by $k+1$ moves of the same type: a block of size two is moved k steps to the left. First $[k{+}1\ k]$ is moved to the far left, then $[k{+}2\ k{-}1]$ is inserted in the middle of the block last moved etc. The last pair to be moved is $[n0]$, after which the situation will be $[k{+}1 \dots n\ 0 \dots k]$, a representative of the toric identity permutation.

Optimality: w_0 has $n-1$ descents, while id has no descent. It is easy to see that the first move from w_0 can decrease the number of descents by just one. The same holds for the last move leading to id. The moves in between can, by the above lemma, each decrease the number of descents by at most two. Hence, at least $(n-3)/2$ moves must be made between the first and the last move (but for $n=2$, the first move is also the last). This proves the theorem. □

A corollary of the optimality of the above algorithm is the lower bound $d(n) \geq \lceil \frac{n+1}{2} \rceil$ for $n \geq 3$.

5. An improved upper bound

In this section, we shall prove a new upper bound on $d(n)$:

Theorem 5.1. *$d(n) \leq \lfloor \frac{2n-2}{3} \rfloor$ for $n \geq 9$.*

The theorem is a consequence of Theorem 4.2 (the optimal algorithm for w_0), Lemma 5.2, and the data of Table 1.

5.1. **Burning the candle at both ends.** Given a permutation π, a k-move from π is a block move σ such that $\pi\sigma$ has k more bonds than π. Obviously, the largest possible value of k is 3. If from a certain permutation π of length n there is a k-move σ, then we can sort $\pi\sigma$ in at most $d(n-k)$ moves, by regarding the bonds as single symbols. Hence $d(\pi) \leq 1 + d(n-k)$. Similarly, if there is a k-move from π^{-1}, that is, if there is a block move τ such that $\pi^{-1}\tau$ has k more bonds than π^{-1}, then the same conclusion holds, since $d(\pi) = d(\pi^{-1})$. Another way of stating this is to say that $\tau^{-1}\pi$ has k more bonds than π. In this case, the distance from π to τ in the Cayley graph is at most $d(n-k)$. Hence there is a block sorting strategy for π of length at most $1 + d(n-k)$, ending with τ^{-1}.

The following lemma states that from any bondless permutation π other than w_0, we can find two block moves, either from π itself, or from π^{-1}, or one move from each of π and π^{-1}, giving a total of at least three bonds. This shows that for $n \geq 4$, $d(n) \leq 2 + d(n-3)$. Since Table 1 shows that for $9 \leq n \leq 11$, $d(n) = \lfloor (2n-2)/3 \rfloor$, Theorem 5.1 will follow by induction.

Lemma 5.2. *Let π be any bondless permutation other than w_0. Then we can find block moves σ and τ such that one of $\pi\sigma\tau$, $\sigma\pi\tau$, and $\sigma\tau\pi$ has at least three bonds.*

As we shall see below, all permutations of this kind fall into one of several categories, for each of which we can construct the required move or moves. In fact, the proof amounts to an algorithm for sorting any permutation by block moves. We will use the toric model of permutations throughout the proof.

5.2. **Criteria for existence of 2-moves.** A 2-move is possible in π if the toric permutation π_0° contains a segment of the form $x \ldots y\overline{x} \ldots \overline{y}$ (where $\overline{x} = x+1 \pmod{n+1}$ etc) since cutting at the indicated places

$$(1) \qquad x| \ldots y|\overline{x} \ldots |\overline{y}\,,$$

gives $x\overline{x}$ ... $y\overline{y}$ with two new bonds. We allow the possibility that $x = \overline{y}$. A 2-move is possible in π^{-1} if the toric permutation $\pi^\circ_\circ$ has a segment of the form

$$(2) \qquad xy\ \ldots\ z\overline{x},$$

where x, y, z are positively oriented. Here we allow the possibility that $y = z$. It is easy to verify that the criteria (1) and (2) are transformed into each other if the roles of values and positions are interchanged.

5.3. **Reducibility.** We say that a toric permutation is *reducible* if, in some suitable representation π and for some $0 < k < n$, the segment $0\ \ldots\ \pi_k$ contains all values $0, \ldots, k$ and the segment $\pi_k\ \ldots\ \pi_n$ contains all values $k, \ldots, n$. In particular, $\pi_k = k$ must then be true. We show that reducible permutations satisfy Lemma 5.2.

If a permutation is reducible, we can handle the two parts $0\ \ldots\ k$ and $k \ldots n$ separately. To do this, we reduce it to a smaller toric permutation by contracting the segment $\pi_k\ \ldots\ \pi_n\ 0$ to a single symbol 0. If the reduced permutation is not a reverse permutation, we may use induction to find the required move or moves. If it reduces to a reverse permutation, we contract the segment $0\ \ldots\ \pi_k$ instead. If this too results in a reverse permutation, we must have

$$\pi = [0\ k{-}1\ k{-}2 \ldots 1\ k\ n\ n{-}1 \ldots k{+}1],$$

and after the 1-move

$$0\ k{-}1 \mid k{-}2 \ldots 1\ k\ n \mid n{-}1 \ldots k{+}1 \mid$$

(bonding $n0$), criterion (2) applies to $k-1\ n-1 \ldots 1\ k$. For example, in 0432159876 we first try to contract 598760, but this results in 04321, which is a reverse permutation. Then we contract the first six values and get •9876 (where the bullet stands for the contracted symbol 0-5), which is interpreted as 04321, another reverse permutation. Finally, a 1-move produces 0487632159, and here 48...15 proves that a 2-move exists in the inverse.

5.4. **Argument by contradiction.** For simplicity, we will argue by contradiction. Suppose $\pi^\circ_\circ$ is a toric permutation of minimal length violating Lemma 5.2. Then $\pi^\circ_\circ$ is bondless, non-reducible, and does not satisfy any of criteria (1) or (2) for 2-moves. We can find a value x such that the distance $x \ldots \overline{x}$ is as small as possible. Then we choose a representative π with $x = 0$, starting by $0 \ldots 1$. The absence of bonds excludes the extreme case 01 and the absence of 2-moves prohibits $0a1$, as a can be moved to bond with $\overline{a}$.

With standard notation, $\pi = 0\ x_1\ x_2 \ldots x_\ell\ 1 \ldots$. We also note that $x_1 > x_\ell$, for otherwise criterion (2) on $0\ x_1\ \ldots\ x_\ell\ 1$ would provide a 2-move from π^{-1}. From the minimality condition on $0 \ldots 1$, we know that $\overline{x}_1$ is not inside that interval, so a 1-move

$$0|x_1\ x_2 \ldots x_\ell\ |1 \ldots |\overline{x}_1$$

is possible, after which we have

$$x_1\ x_2 \dots x_\ell\ \overline{x}_1.$$

Now unless $x_1 > x_2 > x_\ell$, criterion (2) again provides a 2-move. Thus we can assume that $\pi = 0\ x_1\ x_2 \dots x_\ell\ 1 \dots$ with $x_1 > x_2 > x_\ell$. There are two cases that have to be treated quite differently. Either $x_2 = x_1 - 1$ (an anti-bond) or $x_1 \gg x_2 > x_\ell$ (not an anti-bond). We use $\gg$ to denote a difference of at least 2.

5.5. **The x_1x_2 anti-bond case.** The minimality condition on $0 \dots 1$ means that the situation $0 \dots x \dots \overline{x} \dots 1$ cannot occur. In the case $0 \dots x \dots 1 \dots \overline{x}$, a 1-move

$$0\ x_1|\ x_2 \dots x| \dots 1 \dots |\overline{x}$$

would lead to $0x_1 \dots 1 \dots x_2$, after which a 2-move is possible according to criterion (1). So the only possibility is that for each x inside $0 \dots 1$, $\overline{x}$ is further left in $0 \dots 1$. But then it is clear that $x_1x_2 \dots x_\ell$ must be a reversed consecutive sequence.

If $x_\ell = 2$, then as we have assumed irreducibility, the symbol 1 is not followed by $\overline{x}_1$. Moreover, since $\pi \neq w_0$, 1 is not followed by 0. Hence there is a 1-move

$$0|x_1 \dots |1\overline{x}| \dots x$$

leading to $1\overline{x} \dots 2 \dots x$, and criterion (1) applies.

In the remaining case, $x_\ell > 2$, and $\underline{x}_\ell$ is to the right of 1. If there is an x between 1 and $\underline{x}_\ell$ such that $\overline{x}$ is to the right of $\underline{x}_\ell$, then there is a 1-move

$$0x_1 \dots |x_{\ell-1}\ x_\ell\ 1 \dots x| \dots \underline{x}_\ell \dots |\overline{x}$$

leading to the criterion (1) situation

$$x_{\ell-2} \dots \underline{x}_\ell \dots x_{\ell-1}x_\ell.$$

This also includes the possibility that $x = n$ and $\overline{x} = 0$. Otherwise for every x to the left of $\underline{x}_\ell$, $\overline{x}$ is also in that interval. But then the positions between 1 and $\underline{x}_\ell$ must hold a consecutive subsequence of the values $2, 3, \dots, \underline{x}_\ell - 1$. In particular, we can use the value 2 in a 1-move

$$0|x_1 \dots x_\ell\ 1|y \dots |2$$

leading to $0y \dots x_\ell 1$, and as $y < x_\ell$, criterion (2) applies.

This concludes the anti-bond case.

5.6. **The $x_1 \gg x_2$ case.** The value $\underline{x}_1 = x_1 - 1$ is not equal to x_2, but it must occur somewhere between 0 and 1, otherwise criterion (1) would apply to $0x_1 \dots 1 \dots \underline{x}_1$ and give a 2-move. Take $k \geq 2$ maximal such that

$$x_1 \gg x_2 \gg \cdots \gg x_k > x_\ell.$$

The minimality condition on $0 \dots 1$ implies that $\overline{x}_k$ is to the right of 1, so there is a 1-move

$$0|x_1 \dots x_\ell|1 \dots |\overline{x}_k,$$

after which we have $x_k\ x_{k+1} \dots x_\ell\ \overline{x}_k$. Unless $x_k > x_{k+1} > x_\ell$, criterion (2) gives a 2-move. From the way in which k was chosen, we must have an anti-bond $\overline{x}_{k+1} = x_k$, which we can split by a 1-move

$$0|x_1 \dots x_k|x_{k+1} \dots \underline{x}_1| \dots 1.$$

Now criterion (1) can be used on $x_{k+1} \dots x_{k-1}x_k \dots \overline{x}_{k-1}$. Note that by the minimality of $0 \dots 1$, the value $\overline{x}_{k-1}$ must occur to the right of 1.

This completes the proof of Lemma 5.2.

6. The bridge player's problem and $d(n)$ for $n \leq 15$

The values of $d(n)$ for $n \leq 10$ were calculated by computer, by constructing the Cayley graph. We will now describe how we determined the values $d(11) = 6$, $d(12) = 7$, $d(13) = 8$, $d(14) = 8$ and $d(15) = 9$. A minimal counterexample to our working conjecture $d(n) = \lceil (n+1)/2 \rceil$ cannot allow a 2-move or a 1-move followed by a 3-move. For $n \leq 13$, a computer search listed all toric permutations satisfying this restriction; there are not many of them. For each one of these candidates we have checked by computer if they can be sorted in $\lceil (n+1)/2 \rceil$ moves. This is indeed the case for $n = 11$, which proves the values $d(11) = 6$ and $d(12) = 7$. To our surprise, for $n = 13$ a counterexample was found: the permutation

$$[4\ 3\ 2\ 1\ 5\ 13\ 12\ 11\ 10\ 9\ 8\ 7\ 6],$$

and a few others, need 8 block moves. This means that in the worst case, the bridge player will have to make eight block moves to sort her hand.

Lemma 5.2 says in effect that $d(n+3) \leq d(n) + 2$, so we have $d(14) \leq 8$ and $d(15) \leq 9$. For $n = 14$, the reverse permutation shows that equality holds. For $n = 15$, the permutation

$$[4\ 3\ 2\ 1\ 5\ 15\ 14\ 13\ 12\ 11\ 10\ 9\ 8\ 7\ 6]$$

takes 9 block moves.

One can also consider other sorting problems. Some bridge players want to have every suit in a consecutive sorted sequence, but the order of the suits is immaterial. However, this does not change the worst case, as we may get thirteen cards in a single suit. (Bridge players would not call this a "worst case".) A different problem occurs if we demand only that all cards in a suit should be grouped together, without bothering about the order of the cards within a suit. We invite the reader to verify that if only two suits are present, then in a hand of n cards, the suits can be grouped together in at most $\lfloor (n-1)/2 \rfloor$ moves. If cards of the different suits alternate, then this bound is attained.

We will show, using some of the ideas of the proof of Lemma 5.2, that $\lfloor (n-1)/2 \rfloor$ moves suffice even if there are more than two suits. For simplicity, we use the circular model. To pass from an ordinary hand to a circular arrangement, we can add another card, of a different suit, as predecessor of the first card and successor of the last one.

Theorem 6.1. *If $n+1$ cards are arranged cyclically, then the suits can be grouped together in at most $\lfloor \frac{n-1}{2} \rfloor$ block moves.*

Proof. We will use induction on n. The statement is obviously true for $n \leq 2$. A *bond* will now mean two consecutive cards from the same suit. We can assume that, to begin with, there are no bonds. Furthermore, we can assume that there is at most one singleton, since if there is more than one singleton, the problem becomes at least as difficult if we replace the singletons by a single suit. We now find a pair of cards from the same suit, such that one of the two segments between these cards does not contain two cards from the same suit, and moreover does not contain the possible singleton. Note that this is always possible! Since these two cards, say spades, are not consecutive, the predecessor of the last one, say a heart, must belong to the same suit as a card not in the same segment between the two spades. The situation must be:

$$\spadesuit | \ldots \heartsuit | \spadesuit \ldots | \heartsuit$$

Cutting at the indicated places gives two bonds, and the induction is complete. □

Corollary 6.2. *For any array of n not necessarily different objects, $\lfloor \frac{n-1}{2} \rfloor$ block moves are sufficient to group like objects together. This bound is sharp.*

Hence, a bridge hand cand be suit separated in at most 6 block moves, and this bound is attained if two suits alternate.

One can also demand that the suits should occur in a specified order, without paying attention to the order of the cards within a suit. Then the original sorting problem becomes a special case, so it is perhaps unreasonable to ask for a simple solution to this problem.

7. Discussion

We have not mentioned the relevance of our algorithm for approximating the block sorting distance between two permutations. In computational biology, this is considered the main problem. Many similar problems have been shown to be NP-complete, and some of them cannot even be approximated within a constant factor. However, block moves behave somewhat better than most, and in our opinion there is still good hope of finding an algorithm that runs in polynomial time.

What is the biological relevance of these results? Typically, genes common to two species often come in largely the same order in both genomes. In other words, there is an abundance of bonds in the permutation. In our analysis, we first contract bonded genes, and without even looking at the resulting no-bonds permutation, we can state that sorting by block moves must take at least $\lceil \frac{n}{3} \rceil$ moves and at most $\lfloor \frac{2n-2}{3} \rfloor$ moves.

Question: In the interval between these bounds, what is the distribution of a) random permutations b) actual gene permutations?

Computer runs indicate that almost all random permutations live at or very near the level occupied by w_0 in the Cayley graph, which is level $\lceil \frac{n+1}{2} \rceil$. In fact, for odd n we find about 32% of all permutations on this level, about 53% need one move less than that and about 15% need two moves less. The percentages seem to stabilize as n grows.

The lower bound applies only to a product of 3-moves. If there are thousands of genes and all breakpoints are equally probable, then among the first hundred block moves, probably most will be 3-moves. So, in this biologically interesting case, we are in fact close to the lower bound. A more typical situation, however, is when breakpoints occur with much higher probability at some *hot spots*. Then again, the percentages for random permutations may well apply to gene permutations also.

References

[1] V. Bafna and P. Pevzner, Sorting by transpositions, *SIAM J.Discrete Math.* **11** (1998), 224–240.

Dept. of Mathematics, KTH, SE-100 44 Stockholm, Sweden