THE MAJORITY GAME WITH AN ARBITRARY MAJORITY

JOHN R. BRITNELL AND MARK WILDON

ABSTRACT. The k-majority game is played with n numbered balls, each coloured with one of two colours. It is given that there are at least k balls of the majority colour, where k is a fixed integer greater than $n/2$. On each turn the player selects two balls to compare, and it is revealed whether they are of the same colour; the player's aim is to determine a ball of the majority colour. It has been correctly stated by Aigner that the minimum number of comparisons necessary to guarantee success is $2(n-k) - B(n-k)$, where $B(m)$ is the number of 1s in the binary expansion of m. However his proof contains an error. We give an alternative proof of this result, which generalizes an argument of Saks and Werman.

1. INTRODUCTION

Fix n and $k \in \mathbf{N}$ with $k > n/2$. The k-majority game is played with n numbered balls which are each coloured with one of two colours. It is given that there are at least k balls of the majority colour. On each turn the player selects two balls to compare, and it is revealed whether they are of the same colour, or of different colours. The player's objective is to determine a ball of the majority colour. We write $K(n, k)$ for the minimum number of comparisons that will guarantee success. We write $B(m)$ for the number of digits 1 in the binary representation of $m \in \mathbf{N}_0$. The object of this paper is to prove the following theorem.

Theorem 1. *If n, $k \in \mathbf{N}$ and $k \leq n$ then $K(n, k) = 2(n-k) - B(n-k)$.*

This theorem has been stated previously, as Theorem 3 of [1, page 14]. We believe, however, that there is a flaw in the proof offered there of the lower bound for $K(n, k)$, i.e. the fact that $2(n-k) - B(n-k)$ comparisons are necessary. The error arises in Case (ii) of the proof of Lemma 1, in which it is implicitly assumed that if it is optimal at some point for the player to compare balls i and t, then there exist two balls j and ℓ which it is optimal to compare on the next turn, irrespective of the answer received when balls i and t are compared. The proof of Theorem 3 requires an analogue of Lemma 1, stated as Lemma 3, which inherits the same error. The argument

Date: October 18, 2014.
2010 *Mathematics Subject Classification.* Primary 91A46, Secondary 05C47.

for Lemma 1 of [1] is based on Lemma 5.1 in [9], which contains the same flaw; the authors are grateful to Prof. Aigner and Prof. Wiener for confirming these errors.[1]

In [8], Saks and Werman have shown that $K(2m+1, m+1) = 2m - B(m)$. (An independent proof, using an elegant argument on the game tree, was later given by Alonso, Reingold and Schott [3].) The original contribution of this paper is to supply a correct proof that $2(n-k) - B(n-k)$ questions are necessary in the general case, by generalizing the argument of Saks and Werman [8].

We remark that an alternative setting for the majority problem replaces the n balls with a room of n people. Each person is either a knight, who always tells the truth, or a knave, who always lies. The question 'Person i, is person j a knight?' corresponds to a comparison between balls i and j. (The asymmetry in the form of the questions is therefore illusory.) In [1, Theorem 6], Aigner gives a clever questioning strategy which demonstrates that $2(n-k) - B(n-k)$ questions suffice, even when knaves are replaced by spies (Aigner's unreliable people), who may answer as they see fit. He subsequently uses his Lemma 3 to show that $2(n-k) - B(n-k)$ questions are also necessary; our Theorem 1 can be used to replace this lemma, and so repair the gap in the proof of Theorem 6 of [1].

We refer the reader to [1] and the recent preprint [4] for a number of results on further questions that arise in this setting. Many variants of the problem have been studied, for instance involving balls of more than two colours (see [2]), or with comparisons involving more than two balls (see [6]).

2. Preliminary reformulation

We begin with a standard reformulation of the problem that follows [1, §2]. In the special case $n = 2m + 1$ and $k = m + 1$, it may also be found in [8, §4] and [9, §2, §3]. A position in a k-majority game corresponds to a graph on n vertices, in which there is an edge, labelled either 'same' or 'different', between vertices i and j if balls i and j have been compared. Each connected component of this graph admits a unique bipartition into parts corresponding to balls of the same colour. If C is a component with bipartition $\{X, Y\}$ where $|X| \geq |Y|$ then we define the *weight* of C to be $|X| - |Y|$.

The weights of the distinct components of the graph contain all of the essential information about the position, and we may therefore reformulate the k-majority game as a two player adversarial game played on multisets of non-negative integers. We shall call these multisets *positions* and their elements *weights*. The players will be known as the *Selector* and the *Assigner*.

[1]Personal communications.

The starting position is the multiset $\{1, \ldots, 1\}$ containing n elements, corresponding to the n trivial connected components of the null graph on n vertices.

In each turn, the Selector selects two distinct multiset elements w and w', with $w \geq w'$, and the Assigner chooses to replace them with either $w + w'$ or $w - w'$. This corresponds to the situation in the original game when a ball from a component C of weight w is compared to a ball from a different component C' of weight w'; the components C and C' become connected by an edge, and the new component thus formed has weight either $w + w'$ or $w - w'$ depending on the result of the comparison. Since comparisons of balls in the same component give no information, we need only consider comparisons of this type.

We require a victory condition for the Selector. Let $e = k - (n - k)$ be the minimum possible excess of the majority colour over the minority colour, and let $w_1, \ldots, w_c$ be the weights in a given position. (We shall always assume that the weights are listed in non-increasing order, so that $w_i \geq w_{i+1}$.) We remark that $\sum_i w_i \equiv n \bmod 2$, since this is true of the initial position, and since the parity of the sum of weights is preserved at each turn. Hence $w_1 + \cdots + w_c = 2s + e$ for some $s \in \mathbf{N}_0$.

Suppose that the balls in the larger part of the component of (largest) weight w_1 are of the minority colour. Then we see that $w_1 \leq s$, since we must have $-w_1 + w_2 + \ldots + w_c \geq e$. (This observation is equivalent to equation (14) in [1].) It follows that the Selector wins as soon as a position $M = \{w_1, \ldots, w_c\}$ is reached such that

$$w_1 \geq s + 1,$$

where s is determined by $2s + e = w_1 + \cdots + w_c$, since in this situation the balls in the larger part of the largest weight component are known to have the majority colour. Following [8], we say that such a position M is *final*.

Our concern is with the number of turns required for victory. We observe that the cardinality of the multiset is reduced by 1 at each turn, and so if the position is $M = \{w_1, \ldots, w_c\}$, then the number of turns that have elapsed is $n - c$. We define the *value* of a general position M to be the number of elements in a final position reached from M, assuming, as ever, optimal play by both sides. We denote the value of M by $V(M)$.

The result we require, that $2(n-k) - B(n-k)$ questions are necessary to identify a ball of the majority colour in the k-majority game, is equivalent to the following proposition.

Proposition 2. *Let $n \in \mathbf{N}$ and let $k > n/2$. The value of the starting position in the k-majority game is at most $B(n - k) + k - (n - k)$.*

3. Generalized Saks–Werman statistics

If M is a position in a majority game and N is a submultiset of M then we shall say that N is a *subposition* of M. Let $\bar{N}$ denote the complement of N in M and let $||M||$ denote the sum of all the elements of M. Let $\varepsilon_M(N) = ||\bar{N}|| - ||N||$. For $e \in \mathbf{N}$ and a position M such that $||M||$ and e have the same parity, we define

$$\delta_e(M) = \sum_{\substack{N \\ \varepsilon_M(N) \geq e}} (-1)^{||N||}.$$

Thus a subposition of M contributes to $\delta_e(M)$ if and only if it corresponds to a colouring of the balls in which the excess of the majority colour over the minority colour is at least e. We note that when $e = 1$ we have $\delta_1(M) = -f_M(-1)$ where f_M is the polynomial defined in [8, page 386]. (The reason for working with minority subpositions rather than majority subpositions, as in [8], will be seen in the proof of Lemma 6.)

The following lemma is a generalization of [8, Lemma 4.2].

Lemma 3. *Let M be a position and let e have the same parity as $||M||$. Let $w, w' \in M$ be two elements of M with $w \geq w'$. Let M^+ and M^- be the positions obtained from M if w and w' are replaced with $w + w'$ and $w - w'$, respectively. Then*

$$\delta_e(M) = \delta_e(M^+) + (-1)^{w'} \delta_e(M^-).$$

Proof. Let N be a subposition of M such that $\varepsilon_M(N) \geq e$ and let $N^\star = N \backslash \{w, w'\}$. We consider four possible cases for N.

(a) If $w \in N$ and $w' \in N$ then $||N|| = ||N^\star \cup \{w, w'\}|| = ||N^\star \cup \{w+w'\}||$ and $\varepsilon_M(N) = \varepsilon_{M^+}(N^\star \cup \{w + w'\})$.
(b) If $w \notin N$ and $w' \notin N$ then $||N|| = ||N^\star||$ and $\varepsilon_M(N) = \varepsilon_{M^+}(N^\star)$.
(c) If $w \in N$ and $w' \notin N$ then $||N|| = ||N^\star \cup \{w\}|| = ||N^\star \cup \{w - w'\}|| + w'$ and $\varepsilon_M(N) = \varepsilon_{M^-}(N^\star \cup \{w - w'\})$.
(d) If $w \notin N$ and $w' \in N$ then $||N|| = ||N^\star \cup \{w'\}|| = ||N^\star|| + w'$ and $\varepsilon_M(N) = \varepsilon_{M^-}(N^\star)$.

Thus $\delta_e(M^+) = \sum_N (-1)^{||N||}$ where the sum is over all subpositions N of M such that $\varepsilon_M(N) \geq e$ and either (a) or (b) holds, and $(-1)^{w'}\delta_e(M^-) = \sum_N (-1)^{||N||}$ where the sum is over all subpositions N of M such that $\varepsilon_M(N) \geq e$ and either (c) or (d) holds. □

We now define a family of further statistics. For $e \in \mathbf{N}$ and a position M such that $||M||$ and e have the same parity, define $\delta_e^{(1)}(M) = \delta_e(M)$, and for $b \in \mathbf{N}$ such that $b \geq 2$ define recursively

$$\delta_e^{(b)}(M) = \sum_{t \in \mathbf{N}_0} \delta_{e+2t}^{(b-1)}(M).$$

For an alternative expression for $\delta_e^{(b)}(M)$ see Lemma 6.

We note that if $e+2t > ||M||$ then $\delta_{e+2t}^{(b-1)}(M) = 0$, and so the sum defining $\delta_e^{(b-1)}(M)$ is finite. Since we only need positions whose sum of elements is at most n, it follows by induction on b that $\delta_e^{(b)}$ is a linear combination of the statistics δ_d for $d \geq e$. Hence, if M, M^+, M^- and w, w' are as in Lemma 3, we have

$$(\star) \qquad \delta_e^{(b)}(M) = \delta_e^{(b)}(M^+) + (-1)^{w'}\delta_e^{(b)}(M^-).$$

For $r \in \mathbf{N}$ let $P(r)$ denote the highest power of 2 dividing r and let $P(0) = \infty$. Let

$$SW_e^{(b)}(M) = e + P\big(\delta_e^{(b)}(M)\big)$$

where, as expected, we set $e + \infty = \infty$. Our key statistic is now

$$SW_e(M) = SW_e^{(e)}(M).$$

We remark that if $||M||$ is odd then $SW_1(M) = \Phi(M)$ where $\Phi(M)$ is as defined in [8, page 386].

In §5 below we prove Proposition 2 by using $SW_e(M)$ to prove an upper bound on the value $V(M)$ of a position M. The key properties of $SW_e(M)$ we require are ($\star$) and the values of $SW_e(M)$ at starting and final positions. Starting positions are dealt with in the following lemma, whose proof uses the basic identity $\sum_{r=0}^{s}(-1)^r\binom{n}{r} = (-1)^s\binom{n-1}{s}$; see [5, Equation 5.16].

Lemma 4. *Let M_{start} be the multiset containing 1 with multiplicity n. Suppose that $n = 2s + e$ where s, $e \in \mathbf{N}$. Then for any $b \in \mathbf{N}$ such that $b \leq n$ we have*

$$\delta_e^{(b)}(M_{\text{start}}) = (-1)^s\binom{n-b}{s}.$$

Proof. When $b = 1$ we have $\delta_e^{(1)} = \delta_e$. A subposition N of M_{start} contributes to the sum defining $\delta_e(M)$ if and only if $||N|| \leq s$. Therefore

$$\delta_e^{(1)}(M) = \sum_{r=0}^{s}(-1)^r\binom{n}{r} = (-1)^s\binom{n-1}{s}.$$

If $b \geq 2$ then, by induction, we have

$$\delta_e^{(b)}(M) = \sum_{t\in\mathbf{N}_0}\delta_{e+2t}^{(b-1)}(M) = \sum_{t=0}^{s}(-1)^{s-t}\binom{n-(b-1)}{s-t} = (-1)^s\binom{n-b}{s}$$

again as required. □

It follows that if M_{start}, s and e are as in Lemma 4, then $SW_e(M_{\text{start}}) = e + P\big(\binom{2s}{s}\big)$. It is well known that $P\big(\binom{2t}{t}\big) = B(t)$ for any $t \in \mathbf{N}$. (Two different proofs are given in [2] and [4].) Hence

$$(\dagger) \qquad SW_e(M_{\text{start}}) = e + B(s).$$

4. Final positions

Let $e = k - (n-k)$. In this section we show that if M is a final position containing exactly c elements then $SW_e(M) \geq c$. The proof uses the *hyper-derivative* on the ring $\mathbf{Z}[x, x^{-1}]$ of integral Laurent polynomials, defined on the monomial basis for $\mathbf{Z}[x, x^{-1}]$ by

$$D^{(r)} x^p = \binom{p}{r} x^{p-r}$$

for $p \in \mathbf{Z}$ and $r \in \mathbf{N}_0$. (This extends the usual definition of the hyper-derivative for polynomial rings, given in [7, page 303].) The key property we require is the following small generalization of [7, Lemma 6.47].

Lemma 5. *Let $f, g \in \mathbf{Z}[x, x^{-1}]$ be Laurent polynomials. Let $r \in \mathbf{N}$. Then*

$$D^{(r)}(fg) = \sum_{t=0}^{r} D^{(t)}(f) D^{(r-t)}(g).$$

Proof. By bilinearity it is sufficient to prove the lemma when $f = x^p$ and $g = x^q$ where p, $q \in \mathbf{Z}$. In this case the lemma follows from

$$\binom{p+q}{r} = \sum_{t=0}^{r} \binom{p}{t}\binom{q}{r-t}$$

which is the Chu–Vandermonde identity; see [5, Equation 5.22]. □

We also need an alternative expression for $\delta_e^{(b)}(M)$. Let $\alpha_r(M)$ be the number of subpositions N of a position M such that $||N|| = r$. The proof of the next lemma uses the basic identity $\sum_{d=r}^{n} \binom{d}{r} = \binom{n+1}{r+1}$; see [5, Table 174].

Lemma 6. *Let M be a position such that $||M|| = 2s + e$. Then*

$$\delta_e^{(b)}(M) = \sum_{r=0}^{s} \binom{s+b-1-r}{b-1} (-1)^r \alpha_r(M).$$

Proof. When $b = 1$, we have $\delta_e^{(1)}(M) = \delta_e(M) = \sum_{r=0}^{s} (-1)^r \alpha_r(M)$. If $b \geq 2$, then by induction, we have

$$\begin{aligned}
\delta_e^{(b)}(M) &= \sum_{t \in \mathbf{N}_0} \delta_{e+2t}^{(b-1)}(M) \\
&= \sum_{t=0}^{s} \sum_{r=0}^{s-t} \binom{s-t+b-2-r}{b-2} (-1)^r \alpha_r(M) \\
&= \sum_{r=0}^{s} \sum_{t=0}^{s-r} \binom{s-t+b-2-r}{b-2} (-1)^r \alpha_r(M) \\
&= \sum_{r=0}^{s} \binom{s+b-1-r}{b-1} (-1)^r \alpha_r(M)
\end{aligned}$$

as claimed. □

Now let $M = \{w_1, \ldots, w_c\}$ be a final position in the k-majority game where $w_1 \geq \ldots \geq w_c$. Let $||M|| = 2s + e$. As remarked in §2, we have

$$w_1 \geq s + 1.$$

Let

$$g = x^{s+e-1}(1 + x^{-w_2}) \ldots (1 + x^{-w_c}).$$

and note that, since $w_2 + \cdots + w_c \leq s + e - 1$, g is a polynomial. Let

$$g = \sum_{r=0}^{s+e-1} \alpha'_r(M) x^{s+e-1-r}.$$

If $r \leq s$ then a subposition N of M such that $||N|| = r$ cannot contain w_1. Therefore $\alpha'_r(M) = \alpha_r(M)$ whenever $r \leq s$. It follows that

$$(D^{(e-1)}g)(-1) = \sum_{r=0}^{s} \binom{s - r + e - 1}{e - 1} (-1)^{s-r} \alpha_r(M).$$

Hence, by Lemma 6, we have

$$(D^{(e-1)}g)(-1) = \delta_e^{(e)}(M).$$

(The normal derivative would introduce an unwanted $(e-1)!$ at the point.) It follows from Lemma 5, applied to the original definition of g, that $D^{(e-1)}(g)$ is a linear combination, with coefficients in $\mathbf{Z}$, of Laurent polynomials of the form

$$h = x^{s+e-1-a_1} \prod_{i \in A} x^{-w_i - a_i} \prod_{j \in B} (1 + x^{-w_j})$$

where A is a subset of $\{2, \ldots, c\}$, $B = \{2, \ldots, c\} \backslash A$, $a_i \in \mathbf{N}_0$ for each i, and $a_1 + \sum_{i \in A} a_i = e - 1$. It is clear that $h(-1) = 0$ unless w_j is even for all $j \in B$, in which case $h(-1) = \pm 2^{|B|}$. Since $|B| \geq (c - 1) - (e - 1) = c - e$, it follows that $P\big(h(-1)\big) \geq c - e$ for all such Laurent polynomials h. Hence

$$(\ddagger) \qquad SW_e(M) = e + P\big(\delta_e^{(e)}(M)\big) = e + P\big((D^{(e-1)}g)(-1)\big) \geq c$$

as claimed at the start of this section.

5. Proof of Proposition 2

We are now ready to prove Proposition 2. Let $e = k - (n - k)$ be the minimum excess of the majority colour over the minority colour in the k-majority game with n balls.

Let M be a position. Suppose that an optimal play for the Selector is to choose w and $w' \in M$. Since the Assigner wishes to minimize the number of elements in the final position, we have

$$V(M) = \min\big(V(M^+), V(M^-)\big)$$

where M^+ and M^- are as defined in Lemma 3. If $x, y \in \mathbf{Z}$ then $P(x+y) \geq \min\bigl(P(x), P(y)\bigr)$. Hence ($\star$) in §3 implies that

$$SW_e(M) \geq \min\bigl(SW_e(M^+), SW_e(M^-)\bigr).$$

By (‡) at the end of §4, if M is a final position containing c elements then $SW_e(M) \geq c$. In this case $V(M) = c$, so we have $SW_e(M) \geq V(M)$. It therefore follows by induction that

$$SW_e(M) \geq V(M)$$

for all positions M. It was seen in (†) at the end of §3 that if M_{start} is the starting position then $SW_e(M_{\text{start}}) = B(n-k) + e$ and so

$$B(n-k) + e \geq V(M_{\text{start}})$$

as required.

6. Final remark

We end by showing that the statistics $SW_e(M)$ do not predict all optimal moves for the Assigner. We need the following lemma in the case when n is odd; it can be proved in a similar way to Lemma 4. We use exponential notation for multiplicities in multisets; so for instance, $\{1^n\}$ denotes the multiset containing the element 1 with multiplicity n.

Lemma 7. *For any $m \in \mathbf{N}$ we have*

$$SW_1\bigl(\{2, 1^{2m-1}\}\bigr) = \begin{cases} 2 + B(m-1) + P(m-1) & \text{if } m \text{ is odd} \\ 2 + B(m-1) - P(m) & \text{if } m \text{ is even.} \end{cases}$$

Let $m \equiv 3 \bmod 4$ and let $M = \{1^{2m+1}\}$ be the starting position in the majority game with $n = 2m+1$ and $k = m+1$. The positions the Assigner can choose between on the first move in the game are $M^+ = \{2, 1^{2m-1}\}$ and $M^- = \{1^{2m-1}, 0\}$. By Lemma 7, we have

$$SW_1(M^+) = 2 + B(m-1) + P(m-1) = 3 + B(m-1) = 2 + B(m).$$

It is clear that removing a zero element from a position decreases its SW_1 statistic by 1, so by (†) at the end of §3 we have

$$SW_1(M^-) = 1 + SW_1(\{1^{2m-1}\}) = 1 + 1 + B(m-1) = 1 + B(m).$$

Using the SW_1 statistic, the Assigner will therefore choose M^- on the first move. However Lemma 7 implies that $SW_1(\{2, 1^{2m-3}, 0\}) = 1 + B(m)$, and we have already seen that $SW_1\bigl(\{1^{2m-1}\}\bigr) = B(m)$. It follows that playing to M^+ is also an optimal move for the Assigner.

References

[1] Martin Aigner, *Variants of the majority problem*, Discrete Appl. Math. **137** (2004), no. 1, 3–25, 1st International Workshop on Combinatorics of Searching, Sorting, and Coding (COSSAC '01). MR 2042575 (2005b:05003)

[2] Martin Aigner, Gianluca De Marco and Manuela Montangero, *The plurality problem with three colors and more*, Theoret. Comput. Sci. **337** (2005), no. 1–3, 319–330.

[3] Laurent Alonso, Edward M. Reingold, and René Schott, *Determining the majority*, Inform. Process. Lett. **47** (1993), no. 5, 253–255. MR 1245142 (94i:68105)

[4] Christine T. Cheng, Andrew McConvey, Drew Onderko, Nathaniel Shar, and Charles Tomlinson, *Beyond Knights and Knaves*, preprint 2013. Available from `www.cs.uwm.edu/~ccheng/papers/kk-long-ver2.pdf`.

[5] Ronald L. Graham, Donald E. Knuth and Oren Patashnik, *Concrete Mathematics*, 2nd edition, Addison–Wesley Publishing Company, Reading, MA, 1994.

[6] Gianluca De Marco, Evangelos Kranakis and Gabor Wiener, *Computing majority with triple queries*, Theoret. Comput. Sci. **461** (2012), 17–26.

[7] Rudolf Lidl and Harald Niederreiter, Finite Fields, Encyclopedia of Mathematics and its Applications, Volume 20, 2nd edition, Cambridge University Press, Cambridge, 1997.

[8] Michael E. Saks and Michael Werman, *On computing majority by comparisons*, Combinatorica **11** (1991), no. 4, 383–387. MR 1137770 (92h:90034)

[9] Gábor Wiener, *Search for a majority element*, J. Statist. Plann. Inference **100** (2002), no. 2, 313–318, International Conference (of the Forum for Interdisciplinary Mathematics) on Combinatorics, Information Theory and Statistics (Portland, ME, 1997). MR 1877197 (2003a:94032)