involve
a journal of mathematics

Fair choice sequences

William J. Keith and Sean Grindatti

msp

2019 vol. 12, no. 1

INVOLVE 12:1 (2019)
dx.doi.org/10.2140/involve.2019.12.13

Fair choice sequences

William J. Keith and Sean Grindatti

(Communicated by Kenneth S. Berenhaut)

We consider turn sequences used to allocate of a set of indivisible items between two players who take turns choosing their most desired element of the set, with the goal of minimizing the advantage of the first player. Balanced alternation, while not usually optimal, is fairer than alternation. Strategies for seeking the fairest choice sequence are discussed. We show an unexpected combinatorial connection between partition dominance and fairness, suggesting a new avenue for future investigations in this subject, and conjecture a connection to a previously studied optimality criterion. Several intriguing questions are open at multiple levels of accessibility.

1. Introduction

In the discrete version of the cake-cutting problem [Brams and Taylor 1996], some number of people take turns selecting from among a set of indivisible items (usually $2n$ items for two people). Players' preferences vary (they may not all prefer the same item most, second-most, et cetera). If preferences are not known to other players they are usually assumed to vary with uniform probability. Players' preferences are normally described by a simple ranking, called Borda scoring, which assigns values of 1 through $2n$ to the objects being chosen. An object with value $2n$ is most wanted, and gives twice as much satisfaction as the object valued at n, and so forth. A player assigns a utility to a final distribution of goods at the total of their valuation of all objects they receive. It is easy to see that the sum of all players' utilities is by no means constant as preferences or turn orders vary. This has been a problem of interest for many authors; see [Bouveret and Lang 2011; Hopkins and Jones 2009; Kalinowski and Narodytska 2013; Rubchinsky 2010], and others.

Some investigators (Hopkins [2010], Hopkins and Jones [2009]) consider strategic play when players' preferences are known to each other. If preferences are secret, a player's only strategic move is to take their most-preferred item remaining, and instead the question of interest is whether an administrator who also does not know preferences can vary the policy — the sequence in which turns are allocated —

MSC2010: primary 91A05; secondary 05A17.
Keywords: social choice, fair division, permutations, fairness, egalitarian, partitions, dominance.

to probably maximize some criterion of social interest. Policies have been analyzed for optimality criteria such as min-max (the worst-off agent is likely to do least badly) and social welfare or utilitarian optimality (the expected value of the sum of agents' utilities is as high as possible). For instance, Bouveret and Lang [2011] conjectured that simply taking turns is utilitarian-optimal under uniform distribution of preferences, and Kalinowski and Narodytska [2013] proved this. Data from [Bouveret and Lang 2011] show that alternating turns is not min-max optimal, although it is asymptotically so.

In this paper we define a new optimality criterion, fairness: the expected difference between players' total utilities is minimized. We restrict ourselves to conditions common in the literature (see for instance [Bouveret and Lang 2011; Kalinowski and Narodytska 2013]): Borda scoring, as above, and uniformly distributed preferences. Socially, the criterion is useful when players are intolerant of large inequalities among their outcomes. Such players must also be willing to sacrifice some overall social welfare in order to reduce this, because under the given conditions maximal total utility is known [Kalinowski and Narodytska 2013] to be realized by the simple policy of taking turns: but this obviously advantages the first player. Empirically, fair policies never seem to be too far from utilitarian-optimal policies, but there does not seem to be a strong mathematical connection between the two criteria. However, we were surprised to conjecture from data generated to date that the min-max optimal policy is the fairest among policies that only differ from alternation in which player goes first in a "round". Finally, the fairness criterion also turns out to have a surprising combinatorial property connected with the theory of partitions. This connection is partially proved herein but is still open in general. Thus we think there is significant mathematical interest to be explored here.

We prove a number of results for fairer policies. Our theorems range from the intuitively obvious, to a fascinating combinatorial relationship between the dominance order on partitions and fairness of choice sequences associated to partitions in a natural way. This association is a tool not previously used in the literature, which may yield fruitful lines of analysis for other questions. Stating our main theorems accessibly, deferring technical definitions to the next section, we show the following:

Lemma 1. *Inverting the choice sequence negates advantage.*

Theorem 2. *Moving a player's choices later strictly decreases their advantage.*

Theorem 3. *Altering a four-turn sequence from $LRRL$ to $RLLR$ strictly increases player L's advantage, if all turn pairs in positions $2k+1$ and $2k+2$ are LR or RL.*

In other words, heuristically, *players like earlier choices — but, other things being equal, they would like their earlier choices later.* The first part of this theorem is obvious — the second, we think, quite surprising. This is the dominance connection we mentioned earlier, and we conjecture but were unable to prove that in fact a stronger connection to partition dominance holds.

In the following section we provide the definitions that a reader will need, including definitions specific to this subject and some background from disparate areas, intending to make the article as self-contained as possible. In Section 3 we describe and prove our theorems on the fairness-of-choice sequences in various relations. In Section 4 we give some numerical, nonrecursive formulas for players' expected values using tools from combinatorics. In the last section we give our collected data to date, and suggest remaining open problems and interesting questions our work raises. Interested investigators will find material for computational projects for students, as well as challenging questions in combinatorial distributions.

For nonmathematical readers only skimming the article to find a "fairest" choice sequence, while not perfect, we recommend the following for players L and R: the "reverse-and-repeat" sequence

$$LRRL\ \ RLLR\ \ RLLR\ \ LRRL\ \ RLLR\ \ LRRL\ \ LRRL\ \ RLLR\ \ RLLR\ \ \ldots.$$

Known as balanced alternation [Brams and Taylor 2000], this is well-defined only if the number of turns is a power of 2, but it is easy to then simply take an initial segment of a sufficiently long sequence.

2. Definitions

Two agents, Luis and Rita, each rank a set of $2n$ indivisible items in order of preference; without loss of generality we assume Luis's preference from most-preferred to least is labeled $(2n, 2n-1, \ldots, 1)$, and consider Rita's preferences a permutation π of Luis's. We describe Rita's preference order by $\pi^{-1}(2n), \ldots, \pi^{-1}(1)$. Reading left to right, we obtain Rita's ranking of the items from most to least preferred. From here on, by "item i" we mean Luis's label for the item, and specify "Rita's item i" if required. Neither player knows the other's preference; we assume preferences are uniformly distributed.

Each agent takes an equal number of turns on which they select their most preferred (highest-labeled) item of those remaining to be chosen; e.g., if Luis goes first, he will choose the item labeled $2n$. This ends when each person has exactly half of the items. Following Kalinowski, Narodytska and Walsh [Kalinowski and Narodytska 2013], we refer to an order S in which players are allowed to choose items as a *policy*.

A policy is a word in L and R of length $2n$ containing n Ls and n Rs. A policy signifies turns at which Luis or Rita chooses from among the remaining set of objects their most preferred item, receiving that item and deleting it from the remaining choices. The set of positions $\{\ell_1, \ldots, \ell_n\}$ at which L appears, or R respectively, is the set of *choice positions* for that player.

In a choice of four items in which Rita prefers items at 1432, and the policy is $LRLR$, items in Luis's labeling will be taken in the order 4, 1, 3, 2. We refer to this order as the *path* associated to this policy and preference.

Given a set of preferences and a policy, players will receive some collection of items, which they value as the sum of their valuations for each item received: label these sums $L(S)$ or (S) for a given policy S. Luis's advantage is his value minus Rita's.

The *alternating policy* in which Luis goes first, $LRLR \ldots LR$, advantages Luis, since he chooses first in every "round" of a choice for each player. For instance, if Luis and Rita agree on their ranking of the items, Luis will get his 1st, 3rd, 5th, etc. choices, while Rita will get her 2nd, 4th, etc.

For four items, we might argue that $LRRL$ is fairer than $LRLR$; for instance, in the case of agreement, Luis will get items he values at 4 and 1 while Rita gets 3 and 2. We call a policy S_1 *fairer* than policy S_2 if the two players' expected totals (averaging over all possible Rita preference orders) differ by less in absolute value. Equivalently, we may compare their total values over all permutations, $L_{\text{tot}}(S)$ and $R_{\text{tot}}(S)$. Define $L_{\text{adv}}(S) = L_{\text{tot}}(S) - R_{\text{tot}}(S)$. The goal of this article is to seek the fairest policy, minimizing $|L_{\text{adv}}(S)|$.

Remark. Utilitarian-optimality is the usual criterion in the literature. We choose to investigate fairness both for the inherent mathematical interest described above, such as connections to other criteria and mathematical objects further afield, and for the use of cases in which inequality is a phenomenon players accord some negative utility. One could quantize the difference between the social welfare of the fairest and the utilitarian-optimal policies by giving a function to weight the amount by which players disapprove of inequality. Doing so makes all prior criteria instances of a general continuum! If players add to social welfare a "disapproval subtraction" equal to advantage, the resulting optimality criterion is precisely min-max, whereas if they subtract nothing, the criterion is utilitarian-optimality. Our criterion is equivalent to a "disapproval subtraction" of some large multiple of inequality. Study of this generalized criterion could be an interesting route to rigorize the investigation of tradeoffs between various criteria.

Before including a few definitions from other areas of mathematics that will later be useful to us, we state a lemma from [Kalinowski and Narodytska 2013], a very useful recursion that gives the expected value $\bar{u}_i(S)$ of player i for a policy S. Say that a policy S has length p, and denote by $\tilde{S}$ the policy S with the first choice removed — note that this lemma applies to policies of any length, and not necessarily having the same number of places L and R.

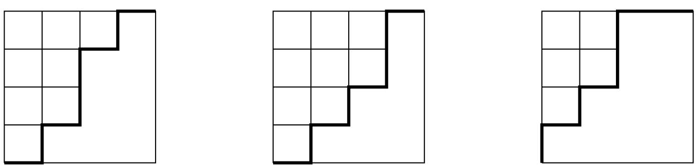
Lemma 4 [Kalinowski and Narodytska 2013, Lemma 1]. *Let S be a policy of length p. Denote by $\bar{u}_1(S)$ the expected value of the player choosing first in S, and* *by $\bar{u}_2(S)$ the expected value of the other player. We have*

$$\bar{u}_1(S) = p + \bar{u}_1(\tilde{S}), \quad \bar{u}_2(S) = \frac{p+1}{p}\bar{u}_2(\tilde{S}).$$

A *partially ordered set* (*poset* for short) is a set A with a reflexive, transitive, antisymmetric relation $a \le b$. If $a \le b$, $a \ne b$, and no element c satisfies $a \le c \le b$, we say that b *covers* a. If for any two elements $c_1, c_k \in A$ all chains $c_1 \le c_2 \le \cdots \le c_k$ have equal length when c_{i+1} covers c_i for all i, then the poset is *ranked*; if there is a unique element $c \le x$ for all $x \in A$ the rank of x may be taken to be the length of any such chain between c and x.

A *partition* of n is a weakly decreasing sequence $\lambda = (\lambda_1, \dots, \lambda_M)$ of nonnegative integers that sum to n. (Typically a partition is defined with positive integers, but it is convenient in this article to speak of a finite number of size 0 parts.) A partition in the $N \times M$ box is one with at most M parts of size at most N; in this article N will always equal M, and we will refer to a box of size N. The "box" language comes from the *Ferrers diagram* of a partition, which is a collection of squares justified to the axes in the fourth quadrant in which the i-th row has λ_i squares.

Example 5. The Ferrers diagrams of the partitions $(3, 2, 2, 1)$, $(3, 3, 2, 1)$, and $(2, 2, 1, 0)$ in the 4×4 box are illustrated below.



The *profile* of a partition is the set of $E - W$ and $N - S$ segments that form the outer boundary of its Ferrers diagram, possibly including any desired number of segments along the axes. A partition λ *dominates* another partition π if it holds that $\sum_{i=1}^{k} \lambda_i \ge \sum_{i=1}^{k} \pi_i$ for all k, assuming an infinite set of trailing zero parts in each. Dominance is a partial order. For instance, $(4, 4, 2, 1, 1)$ dominates $(4, 3, 3, 1, 1)$ but not $(4, 3, 3, 2)$, nor does $(4, 3, 3, 2)$ dominate $(4, 4, 2, 1, 1)$.

A partition $\lambda = (\lambda_1, \dots, \lambda_n)$ *contains* another partition $\sigma = (\sigma_1, \dots, \sigma_n)$ if we have $\lambda_i \ge \sigma_i$ for all i. Containment is a partial order relation which makes partitions into a ranked poset, where the rank of a partition is just the number that it partitions. Thus the set of partitions in the $N \times M$ box ordered by containment is a finite, ranked poset with minimal element $(0, 0, \dots, 0)$. Containment implies dominance, which means that the containment poset on partitions in the $M \times N$ box can be constructed by removing some comparabilities from the dominance poset. In particular, no two partitions of n contain each other, while dominance can be used as a partial order on the set of partitions of n in a given box. Unless we refer to dominance for a particular theorem, in this paper we mean containment when we speak of the partition poset or the word poset.

We define dominance and covering on words by associating them to partitions. We associate a word of length $2n$ with n each of L and R to a profile in the $n \times n$ box by starting from the upper right corner and drawing an $E-W$ step for an L, and a $N-S$ step for an R. Thus, the partitions in the example above are associated to the sequences $LRLRRLRL$, $LRRLRLRL$, and $LLRRLRLR$ respectively. We say that a word in L and R dominates (resp. covers) another if the associated partition of the first dominates (resp. covers) that of the second. In the figure above, $(3, 3, 2, 1)$ dominates both of the other partitions, and covers $(3, 2, 2, 1)$. These two figures are relevant to an example we give in the next section.

Of particular interest to us is the set of words that lie between the alternating sequences in the word poset which are restricted to consecutive pairs LR or RL for every two $(2k-1)$-th and $2k$-th positions, which we refer to as the Boolean set. These words are associated to the 2^n partitions whose Ferrers diagrams differ only by either containing, or not containing, the squares on the diagonal of the box. The first two partitions in the example are in this set.

3. Effects of changes on the fairness of policies

The policy associated to the minimal element $(0, 0, \dots, 0)$ in the box of size N is $LL\dots RR$, which among all policies with n Ls and n Rs is obviously the best possible policy for Luis (highest L_{tot} and L_{adv}). We can move through the policy poset by adding one square at a time, exchanging a consecutive pair LR in the sequence for an RL. We should expect that this will always worsen Luis's position, and our first theorem shows this. Slightly less obviously, for any given Rita preference permutation the change happens in a specific way, by the exchange of one pair of items between the two players' outcomes.

Theorem 6. *Let two policies S and S' be such that $S = s_1 \dots s_k s_{k+1} \dots s_{2n}$ where $s_k = L$ and $s_{k+1} = R$, and $S' = s_1 \dots s_{k+1} s_k \dots s_{2n}$. Then $L(S') \leq L(S)$ and $R(S') \geq R(S)$, by exchange of one item between the players' outcomes for any given Rita preference π.*

Proof. The magnitude clause is extremely intuitive and, with Lemma 4, is nearly trivial: Luis's expected value in $LR\sigma$ with σ any following policy of length $p-2$ is $p + p/(p-1)\,\bar{u}_L(\sigma)$, while for $RL\sigma$ it is $(p+1)/p\,(p-1+\bar{u}_L(\sigma))$. The former is larger than the latter, and Lemma 4 tells us that passage through any prefix to such a word yields an expected value which is some increasing linear function of the input, so the final expected value is larger.

However, by examining the situation in more detail we can say more about the actual change made to the players' outcomes.

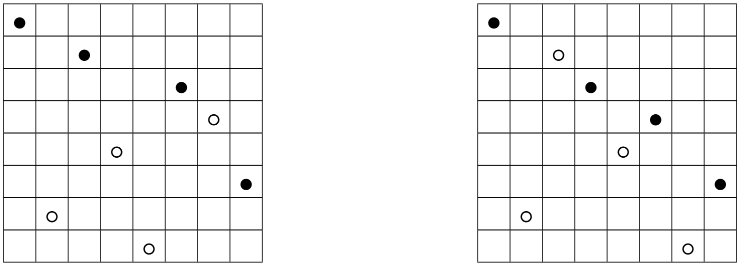


Figure 1. Left: choice procedure for policy S. Right: choice procedure for policy S'. Note that white circles denote items Rita obtains and black circles denote Luis' items.

Let Luis's outcome be $\text{out}_L(S) = \{l_1, l_2, \ldots, l_n\}$ and Rita's outcome $\text{out}_R(S) = \{r_1, r_2, \ldots, r_n\}$. The first $k-1$ turns with S' will yield the same results for both players as with S. On turn k, using S, Luis takes some item l_i, and then on turn $k+1$ Rita takes some item r_j. For S', turn k is Rita's turn instead of Luis's. She will either take r_j as she did using S or she will take l_i. If she takes r_j, then on the next turn, Luis will take l_i, for he preferred this to all remaining items. Then for the remainder of the items, at any given choice both players will face the same set of remaining items in S' as they did in S, and will make the same choices with S' as they did with S. Thus $\text{out}_L(S) = \text{out}_L(S')$ and $\text{out}_R(S) = \text{out}_R(S')$.

If, instead, Rita takes l_i, then Luis will later take at least one item r_m belonging to Rita's previous outcome $\text{out}_R(S)$ such that both Luis and Rita prefer l_i to r_m. This worsens his outcome and betters hers. We claim that this exchange of l_i for r_m is the only difference in outcomes for S and S', so $L(S') < L(S)$, and $R(S') > R(S)$.

For example, suppose that $S = LRLRRLRL$, $S' = LRRLRLRL$. Observe that their associated partitions are the first two we illustrated in the example of the previous section, so that the policy (and the associated partition of S' covers and dominates S. We have moved up one step in the word poset, from (3, 2, 2, 1) partitioning 8, to (3, 3, 2, 1) partitioning 9.

Let Rita's preference be given by $\pi = 28741563$. We illustrate the choice procedure for S in Figure 1, left, with white circles denoting items Rita obtains and black circles denoting Luis's items.

Using choice sequence S', Luis still takes 8 and Rita still takes 2. The third turn is now Rita's instead of Luis's, and since they agree on the best item at that time, Rita takes $l_2 = 7$. On his second turn, Luis takes his $l_3 = 6$. On Rita's third turn, she takes her $r_2 = 4$. On Luis's third turn, he could take his $l_4 = 3$, but he now has access to $r_3 = 5$, so he takes that instead. Now since Rita cannot take her r_3, she takes $r_4 = 1$ on her fourth turn. On Luis's fourth turn, he also takes his $l_4 = 3$; see Figure 1, right.

In general, if Rita takes l_i, it must follow that she prefers l_i to r_j, her previous choice using S. On each of her subsequent turns g, she will take her r_{g-1} if it is available, or r_g if it is not. Why?

If Luis has not yet taken any item r_g, then on his turns he was taking items he had previously chosen one turn later. In this case, on any of Rita's turns after the change, the items taken so far will consist of items Luis had previously chosen prior to that turn, plus one item Luis had chosen later than her current turn, plus the items Rita had previously taken, except for the item she previously took on the turn before the current turn. Rita preferred that item to all remaining of Luis's items and all remaining of her own, and so she will take it.

In other words, Rita has basically delayed taking items she previously chose because a better item is now available.

On turn $k+1$, since l_i is no longer available, Luis will either take l_{i+1} or r_j. On each of his subsequent turns h, he will take his l_h if it is available, or either l_{h+1} or r_m, Rita's next choice, if it is not.

On some turn t, Luis must take one of Rita's items r_m because, since Rita took l_i, she can no longer collect all n of the items $r_1, r_2, \ldots, r_n$ she collected with choice sequence S. Suppose that turn t is the earliest place this happens. At this point, the items previously chosen constitute the same set as the items that had been chosen at this point in the previous choice sequence; Luis has filled the delay in Rita's choices.

Thus the rest of each player's choices will follow the same with S' as with S, so Luis's and Rita's outcomes using S' will be identical to those for S, with the single exception that Rita now gets l_i and Luis gets r_m. Since Luis chose l_i over r_m using sequence S and Rita chose l_i over r_m using S', we know l_i is more valuable to both Luis and Rita than r_m. Therefore $L(S') < L(S)$ and $R(S') > R(S)$. □

The inequalities become strict for L_{tot} and R_{tot}, since there will be at least one strict difference, when Rita's preference permutation is the identity.

Corollary 7. *Let S and S' be choice sequences as before, with S' covering S. Then $L_{\text{tot}}(S') < L_{\text{tot}}(S)$ and $R_{\text{tot}}(S') > R_{\text{tot}}(S)$, and hence $L_{\text{adv}}(S') < L_{\text{adv}}(S)$.*

Proof. Since we know for any given permutation π that $L(S') \geq L(S)$ and $R(S') \leq R(S)$, it is clear that $L_{\text{tot}}(S') \geq L_{\text{tot}}(S)$ and $R_{\text{tot}}(S') \leq R_{\text{tot}}(S)$. For Rita's preference $(2n, \ldots, 1)$, the items will be chosen in reverse order. Exchanging s_k and s_{k+1} exchanges items $2n-k+1$ and $2n-k$. Other choices will be the same for both players. Hence $R(S)+1 = R(S')$ and $L(S) = L(S')+1$, so $L(S') < L(S)$ and $R(S') > R(S)$ for this preference. Thus $L_{\text{tot}}(S') < L_{\text{tot}}(S)$ and $R_{\text{tot}}(S') > R_{\text{tot}}(S)$. □

Clearly $LL\ldots RR$ has the highest advantage for Luis and $RR\ldots LL$ for Rita. We now know that increasing rank for choice sequences improves Rita's outcome and has the opposite effect on Luis's outcome. While it is not always the case that an incomparable sequence with higher rank has lower L_{adv}, or even L_{tot}, we can guarantee that this will happen in a limited case: consider any policy S and call its inverse S^{-1} the sequence for which occurrences of L and R are reversed. For example, if $S = LLRRLR$, then $S^{-1} = RRLLRL$. It is true that if S and S^{-1} are sequences such that a series of $LR \to RL$ moves from S can result in S^{-1}, then any sequence S' constructed from S by some of the same moves will be fairer than S.

Theorem 8. *Consider a policy S such that S^{-1} can be reached by a series of n $LR \to RL$ moves from S. Construct policy S' by a series of m of the same moves, with $0 < m < n$, so that S' is contained on a path between S and S^{-1}. Then S' is necessarily fairer than S.*

For example,

$$LLRRLR,\ \ L\mathbf{RL}RLR,\ \ \mathbf{RL}LRLR,\ \ RL\mathbf{RL}LR,\ \ R\mathbf{RL}LLR,\ \ RRLL\mathbf{RL}$$

is such a sequence from a word to its inverse, where we have bolded the elements moved. Our theorem says that any word within this sequence will be fairer than a word on the ends of the sequence.

Proof. Recall Lemma 1: formally, $L_{\text{adv}}(S) = R_{\text{adv}}(S^{-1})$. Inverting the positions of each player simply swaps their role in procedure, so the truth of this lemma is straightforward.

We now have $L_{\text{adv}}(S^{-1}) = -L_{\text{adv}}(S)$. Since S^{-1} can be reached from S by a series of $LR \to RL$ moves, we know that $L_{\text{adv}}(S^{-1}) < L_{\text{adv}}(S)$. Since S' can be reached from S and S^{-1} can be reached from S' by a series of such moves, it must hold that $L_{\text{adv}}(S^{-1}) < L_{\text{adv}}(S') < L_{\text{adv}}(S)$. Then since $L_{\text{adv}}(S^{-1}) = -L_{\text{adv}}(S)$, we have $|L_{\text{adv}}(S')| < |L_{\text{adv}}(S)|$, meaning S' is a fairer policy. □

Because the two alternating sequences are the inversions of each other and can be reached by exchanging every two positions, any sequence on the poset paths between them will be fairer than either. Thus we have the following:

Corollary 9. *Any choice sequence lying in the sequence poset strictly between $LRLR \dots LR$ and $RLRL \dots RL$ is fairer than either of these.*

We now know that a policy which covers another is better for Rita than the latter, and as we would expect, this is generally the case for sequences of higher ranks which are not comparable in the poset. (It does fail occasionally; that is, if λ is of higher rank than σ but does not cover σ, it is possible for R_{tot} to be smaller, though this is not usually the case. For example, this happens with $LLRRRL$ (associated partition $(3, 3, 0)$) and $LRRLRL$ (associated partition $(2, 2, 1)$), and this is the only such pair in the 3×3 box.) What about policies of equal rank?

We have much numerical evidence for an intriguing conjecture we were unable to prove in full generality: that L_{adv} could be associated easily to the dominance order among partitions in a rank.

Conjecture 10. *Among choice sequences σ_i of the same rank, if S_1 dominates S_2, then S_1 has higher average advantage for Luis than S_2.*

This has been verified computationally for all ranks of all posets in boxes of sizes up to 10×10, but not proven generally. We have, however, been able to prove this for a restricted case:

Theorem 11. *If two choice sequences in the Boolean set consist of a prefix α and a suffix β connected by $LRRL$ and $RLLR$ respectively, i.e., $\sigma_1 = \alpha LRRL\beta$ and $\sigma_2 = \alpha RLLR\beta$, then Luis's advantage is strictly greater for σ_2 than for σ_1.*

Thus, a partition in the Boolean set that dominates another by dominance moves of adjacent, nonoverlapping pairs — in the partition, by moves of one square at a time in the Ferrers board to the next part up in the partition — is better for Luis than the latter. A major difficulty in initially establishing this theorem was that it is *not* the case that Luis's position always worsens going from the former to the latter! Rather, even if Luis's position betters, Rita's does also, and by more.

Proof. The values expected to be obtained by each player as the policy progresses through β do not change; denote these by B_1 and B_2 respectively. Say that β has length r; it is a straightforward application of Lemma 4 to obtain the expected values of Luis and Rita before and after the change in the connecting word:

$$\bar{u}_L(LRRL\beta) = r+4+\frac{r+4}{r+3}\left(\frac{r+3}{r+2}(r+1+B_1)\right)$$

$$\bar{u}_L(RLLR\beta) = \frac{r+5}{r+4}\left(2r+5+\frac{r+2}{r+1}B_1\right)$$

$$\bar{u}_R(LRRL\beta) = \frac{r+5}{r+4}\left(2r+5+\frac{r+2}{r+1}B_2\right)$$

$$\bar{u}_R(RLLR\beta) = r+4+\frac{r+4}{r+3}\left(\frac{r+3}{r+2}(r+1+B_2)\right)$$

Observe that, given an expected value x that holds as one enters a segment α, Lemma 4 gives that the expected value after exiting α will be some linear function of x. This function will be different for players 1 and 2: say that Luis experiences function $Ax+B$ when passing through α, and Rita experiences $Cx+D$.

Thus we have

$$\begin{aligned}\big(\bar{u}_L(\alpha LRRL\beta) - \bar{u}_R(\alpha LRRL\beta)\big) &- \big(\bar{u}_L(\alpha RLLR\beta) - \bar{u}_R(\alpha RLLR\beta)\big)\\ &= (A+C)\frac{r-2}{(r+2)(r+4)} + (AB_1+CB_2)\frac{-4}{(r+2)(r+4)(r+1)}.\end{aligned}$$

Multiplying through by $(r+4)(r+2)$, we find that we wish to show that

$$\frac{AB_1 + CB_2}{A+C} > \tfrac{1}{4}(r-2)(r+1).$$

This statement will be true if the stronger inequality holds:

$$\frac{AB_1 + CB_2}{A+C} \geq \tfrac{1}{4}r^2.$$

Since we specified that the policy being studied was in the Boolean set, an $LRRL$ can only occur with an even number of places remaining, in which each player has at least one choice in every two adjacent places, and so the minimum possible value of either B_1 or B_2 is

$$1+3+5+\cdots+\bigl(2\bigl(\tfrac{1}{2}r\bigr)-1\bigr) = \bigl(\tfrac{1}{2}r\bigr)^2.$$

Thus, the ratio $(AB_1 + CB_2)/(A+C)$ would be reduced by taking the greater of B_1 and B_2 and reducing it to the lesser, giving a ratio above the threshold required. Hence, an adjacent dominance move $LRRL \to RLLR$ on a choice sequence in the Boolean set improves Luis's advantage. □

We may observe that this theorem completely characterizes relative relations within ranks in the Boolean set, since any two Boolean set choice sequences with the same rank can be related by adjacent dominance moves. A more general conjecture, which might make a useful intermediate step toward the full conjecture, would be to establish Luis's advantage improvement for dominance by exchange of exactly one position at any distance, regardless of whether a partition was in the Boolean set. This would cover policies $S = \alpha LR\beta RL\gamma$ and $S' = \alpha RL\beta LR\gamma$. The method of proof applied above appears to be insufficient to establish the full conjecture without further insight. We remain interested in the question and invite interested readers to attempt the proof.

Remark. We could certainly have shown Theorem 2 using Lemma 4, but this would have given no information about the exchange made. A similar analysis here shows that for a dominance move, Luis and Rita will either exchange two pairs of items in two nonoverlapping intervals, or a single pair of items. The problem with using this to prove the dominance theorem is that the single exchange may leave Luis worse off. For instance, if Rita's preference is 4312, then in $LRRL$ Luis receives 42, but in $RLLR$ Luis receives 32. So unlike in our previous theorems, there exist some cases in which Luis suffers the opposite of the general effect. Thus, establishing the theorem by these methods would oblige us to estimate the relative frequency of various sizes of exchange — a task we found to be quite difficult.

3.1. ***Search strategies and heuristics.*** We can now make a reasonable suggestion for a fair sequence. Due to Theorem 8, one would suspect that a policy near the middle rank would be close to $L_{\text{adv}}(S) = 0$: one suggestion would be the policy known as balanced alternation, or the Thue–Morse sequence, which is an initial segment of

$$LRRL\ RLLR\ RLLR\ LRRL\quad RLLR\ LRRL\ LRRL\ RLLR\quad RLLR\ \ldots.$$

This will certainly be fairer than either alternating sequence. By rank it lies halfway between the two, and it is near midway between the two extremes of its rank in dominance order, so by our theorems it would reasonably be expected to have L_{adv} close to 0.

Ideally, if our only priority is finding the fairest possible choice sequence, we would like to be able to take as input the length $2n$ of the item set and with a short algorithm place Luis's choices $\{\ell_1, \ldots, \ell_n\}$. We are very far from achieving this, but with the help of the above theorems we can reduce the work considerably from examining all possible choice sequences.

(1) The poset of partitions in the box of size N has either 1 or 2 middle ranks, depending on whether N is even or odd. Calculate L_{adv} for one such rank; the higher, if N is odd.

(2) Move up one rank, ignoring choice sequences associated to partitions that cover any that already have negative L_{adv}, and calculate again.

(3) Repeat until all L_{adv} are nonpositive or an $L_{\text{adv}} = 0$ is found.

(4) At this point, stop and select the choice sequence with lowest $|L_{\text{adv}}|$. This sequence and its inverse will be the fairest choice sequences for $2n$ items.

If the dominance conjecture were completely true, then a binary search could be run in each rank for the fairest sequence, reducing the work by a factor of $\log_2 n$; if only a sequence in the Boolean set is desired, this can definitely be done.

4. Formulas for expected values

Recall that an *outcome* for Luis is a set of items he receives, and a *path* associated to a given policy and Rita's preference is the order in which items are taken by both players, as labeled by Luis. Our first theorem in this section gives us a formula, given a particular choice sequence and Luis's outcome, for the number of paths which yield this outcome, or equivalently, the number of possible sequences of Luis-labeled items that Rita might take under the given conditions. It uses the falling factorial notation

$$(x)_j = x(x-1)\ldots(x-(j-1)).$$

Theorem 12. *The number of paths in which Luis has choice positions* $\{\ell_1, \ldots, \ell_n\}$ *and takes values* $\{v_1, \ldots, v_n\}$ *is given by*

$$(\ell_1 - 1)_{\ell_1 - 1}\left(\prod_{j=2}^{n}(\ell_j + v_{j-1} - 2n - 2)_{\ell_j - \ell_{j-1} - 1}\right)(v_n - 1)_{2n-\ell_n}.$$

Proof. The problem reduces to a question of counting the number of ways to fill each column and row exactly once, in a Ferrers board given by Luis's choices. Consider, for example, the case of length $2n = 10$ in which Luis chooses at positions {2, 5, 6, 8, 9} and takes values {9, 8, 5, 3, 2}:

10	○		■	■			■			■
9		●								
8					●					
7	○		○	○			■			■
6	○		○	○			■			■
5						●				
4	○		○	○			○			■
3								●		
2									●	
1	○		○	○			○			○
	1	2	3	4	5	6	7	8	9	10

Here black circles represent Luis's definite choices, white circles Rita's possibilities, and black squares places forbidden by Luis's choices. Of course, Rita may not choose an item later and higher-valued than a choice of Luis's, else he would have taken this item on an earlier turn in preference to one he selected. It is also easy to observe that Rita must have chosen, say, item 10 at place 1 because Luis chose item 9 at place 2, but our placement of circles is not that keen yet: we merely take all rows and columns not occupied by a Luis choice which are not later and higher than a Luis choice.

Thus, among rows and columns that Luis does not occupy, Rita can take any collection of objects that includes exactly one item in each row and column in the open positions left of and below Luis's choices. Since she may have any preference among the items so chosen, these may come in any order so long as they satisfy the previous conditions.

Such a problem is referred to as counting *full rook placements* within the Ferrers board of shape consisting of the spaces below and left of Luis's choices, in the columns and rows that they do not occupy. This is a standard counting problem, the method for which may be found on page 74 of Stanley's Enumerative Combinatorics, Volume 1 [Stanley 1997]. We state here a theorem from that volume for reference:

Theorem 13 [Stanley 1997, Theorem 2.4.1]. *Let $\sum_{k=0}^{m} r_k x^k$ be the rook polynomial of the Ferrers board B of shape $(b_1, \dots, b_m)$. Set $s_i = b_i - i + 1$. Then*

$$\sum_{k=0}^{m} r_k (x)_{m-k} = \prod_{i=1}^{m} (x + s_i).$$

The constant term of this polynomial, $\prod s_i$, is precisely r_m, the number of ways to place m nonattacking rooks on the board. This is the number we desire.

Our Ferrers board has n parts b_1 through b_n. In order to use the formula of [Stanley 1997], we name parts in ascending order of size, hence the reverse order of their appearance in the choice sequence.

We observe that we have:

- $\ell_1 - 1$ parts of size $2n - n = n$ ($2n$ values, n occupied),
- $\ell_2 - \ell_1 - 1$ parts of size $v_1 - 1 - (n-1) = v_1 - n$,
- $\ell_3 - \ell_2 - 1$ parts of size $v_2 - 1 - (n-2) = v_2 - n + 1$,

  $\vdots$
- $\ell_n - \ell_{n-1} - 1$ parts of size $v_{n-1} - 1 - (1) = v_{n-1} - 2$,
- $2n - \ell_n$ parts of size $v_n - 1$.

Thus among the s_i are $2n - \ell_n$ values $v_n - 1, v_n - 2, v_n - 3, \dots, v_n - (2n - \ell_n - 1)$. The product of these is the falling factorial $(v_n - 1)_{2n-\ell_n}$.

The next s_i start with $s_{2n-\ell_n+1}$. The associated b_i are all $v_{n-1} - 2$, and there are $\ell_n - \ell_{n-1}$ of them. The s_i thus produced are

$$\begin{aligned} v_{n-1} - 2 - (2n - \ell_n + 1) + 1 &= \ell_n + v_{n-1} - 2n - 2, \\ v_{n-1} - 2 - (2n - \ell_n + 2) + 1 &= \ell_n + v_{n-1} - 2n - 3, \\ &\vdots \\ v_{n-1} - 2 - (2n - \ell_n + \ell_n - \ell_{n-1} - 1) + 1 &= \ell_{n-1} - v_{n-1} - 2n. \end{aligned}$$

The product of these s_i is the falling factorial $(\ell_n + v_{n-1} - 2n - 2)_{\ell_n - \ell_{n-1} - 1}$.

The other falling factorials in the product arise similarly. □

Rita may have multiple preferences that give rise to a particular path. For a simple example, in the choice sequence LR, it does not matter whether Rita's preferences are 12 or 21; items will be taken in the sequence 2, 1. The number of Rita preferences that give rise to any path is a constant that depends only on the position of the choices in the sequence, and not on the specific path:

Theorem 14. *Consider a choice sequence S where Luis's choice positions are $\{\ell_1, \dots, \ell_n\}$, and a specific path through S given by $s = s_1, s_2, \dots, s_{2n}$. Then the number of possible preference permutations for Rita resulting in path s through S is $\prod_{j=1}^{n} (2n - \ell_j + 1)$.*

Proof. Take any S and any path s through S, and construct π by placing into it the items as they occur in s. On each of Rita's turns i, the item r_i she selected must always be placed in the leftmost available position in π since r_i was her most preferred item of those remaining. On each of Luis's turns j, the item l_j he chose may be placed in any of the remaining positions in π since, regardless of its value to Rita, she will not have a chance to take the item after Luis has already taken it. The number of available positions on Luis's turn j is equal to $2n - \ell_j + 1$, where ℓ_j refers to the position of Luis's turn j in the original sequence S. Thus the total number of permutations for a given path s through S is $\prod_{j=1}^{n}(2n - \ell_j + 1)$. □

Since the number of preference permutations associated with a given path through a sequence S is dependent only on S, and thus is constant across all paths through S, we can count outcomes by grouping them according to the resulting path.

Suppose Luis's positions are $(\ell_1, \ldots, \ell_n)$. Take each possible Luis outcome, in which his values $(v_1, \ldots, v_n)$ can range from a maximum of $2n+1-i$ for v_i (Rita took no higher-ranked items than his most-preferred) to a minimum of $2n+1-\ell_i$ (Rita always took Luis's next-preferred item at her choices). To each outcome we can calculate the number of paths and the number of Rita preferences that give that path. We total Luis's value and sum over all possible outcomes for this choice sequence to get L_{tot}.

Thus, combining Theorems 12 and 14, we have a formula for the total value of Luis's outcomes over the set of all Rita preferences, using a given choice sequence S:

$$L_{\text{tot}} = \left(\prod_{j=1}^{n} 2n - \ell_j + 1\right) \sum_{\substack{(v_1, v_2, \ldots, v_n) \\ \min(2n+1-i, v_{i-1}-1) \geq v_i \geq 2n+1-\ell_i}} \left(\sum v_n\right)$$
$$\times (\ell_1 - 1)_{\ell_1 - 1} \left(\prod_{j=2}^{n} (\ell_j + v_{j-1} - 2n - 2)_{\ell_j - \ell_{j-1} - 1}\right) (v_n - 1)_{2n - \ell_n}. \quad (1)$$

Dividing by $(2n)!$ gives Luis's expected value.

A second approach to counting outcomes involves making a tree diagram for the possible outcomes with the assumption that Rita's choices are made randomly.

To do this, we let Rita's preferences be arbitrary. Since all possible preferences are considered equally likely, it is also the case that on any of her turns, Rita is equally likely to take any one of the available items, and it is valid to imagine that on each turn she chooses one item at random. Then we can represent the problem using a tree, the nodes of which will contain all possible actions for a turn.

Theorem 15. *The total number of Rita preferences that result in a particular outcome* $\{v_1, \ldots, v_n\}$ *for Luis is* $(2n)!\,P$, *where* $2n$ *is the number of items and* P, *the probability that the outcome occurs, is equal to the number of paths giving a* *particular outcome for Luis divided by the total number of paths, or*

$$\frac{(\ell_1 - 1)_{\ell_1 - 1}\left(\prod_{j=2}^{n}(\ell_j + v_{j-1} - 2n - 2)_{\ell_j - \ell_{j-1} - 1}\right)(v_n - 1)_{2n-\ell_n}}{\prod_{i=1}^{n} 2n - r_i + 1},$$

where r_i is the position of Rita's i-th turn in the overall sequence.

Proof. Considering the problem as a tree with Rita's preferences unknown, we know that on each of his turns, Luis will always take the highest-numbered item, and on Rita's turns, she will take any one of the remaining items with equal probability of each. Thus none of Luis's turns will generate additional branches, but on each of Rita's turns, a branch is necessary for each of the remaining items. The number of the remaining items at Rita's turn i is $2n - r_i + 1$. The total number of paths in the tree is then the product of this value over all of her turns, $\prod_{i=1}^{n}(2n - r_i + 1)$. From Theorem 12, we know that the number of paths giving a particular outcome is $(\ell_1 - 1)_{\ell_1 - 1}\left(\prod_{j=2}^{n}(\ell_j + v_{j-1} - 2n - 2)_{\ell_j - \ell_{j-1} - 1}\right)(v_n - 1)_{2n-\ell_n}$. Dividing them gives P, the probability the outcome occurs.

Multiplying with P the total number of possible preferences for Rita, $(2n)!$, yields the number of Rita preferences that result in a given outcome. □

5. Conjectures and open problems

There are a number of open questions that interested researchers from student to faculty might be able to consider for this problem.

Data is often a good start. We begin with the collection of known, guaranteed fairest choice sequences; see Table 1. We list sequences with the Luis-first version; where this gives Luis a negative L_{adv}, the sequence is marked with an asterisk. If the fairest known sequence is not one of those that lies between the alternating sequences, the fairest of those in the Boolean set is given.

length	fairest known	fairest between alternating
2	LR	
4	$LRRL^*$	
6	$LRLRRL$	
8	$LLRRRLRL$	$LRRLRLLR$
10	$LRRLLRLRRL$	
12	$LLRRRRLLRLLR^*$	$LRLRLRLRRLRL^*$
14	$LLLRRRRLLRRRLL$	$LRRLLRRLLRLRRL$
16	$LRLLLLRRRRRRRLLL$	$LRRLRLRLRLRLLRLR^*$
18	$LLRLLRRRRRLRLRLLLR$	$LRRLLRRLLRRLLRLRRL$

Table 1. The collection of known and guaranteed fairest choice sequences.

Let us pause for a few remarks on Table 1.

The fairest choice sequences seem to be rather generally not within the Boolean set — instead, they seem to be close to $LL\dots RRRR\dots LL$, with half the L's at the front and back of the sequence. That seems quite surprising and counterintuitive. After all, the simple alternating sequence $LRLR\dots$ maximizes social welfare, and balanced alternation $LRRLRLLR\dots$ has good heuristic arguments for being a relatively fair sequence. Both distribute L and R relatively evenly throughout the sequence. A sequence that "chunks" the players significantly would be very different from these.

Actually using such a choice sequence to divide items would severely strain the assumptions that Luis's and Rita's preferences are independent, and that valuations are linear: Luis would be collecting a quarter of the items before Rita gets a chance to take any of her most preferred items. We assumed no correlation of preferences and items valued in even intervals, but if there is any agreement between Luis and Rita on a small subset of highly valuable items, Luis would be able to seize these immediately. On the other hand, if there is agreement on a small subset of highly undesirable items Luis would also be left with these, so perhaps the distribution would work out. From a strictly mathematical viewpoint, however, it is certainly of intrinsic interest to know if such a sequence is the "typical" fairest sequence.

As mentioned earlier, it is of interest to determine how dominance interacts with other conditions studied in this area, such as the min-max and social welfare conditions described in the introduction. From [Bouveret and Lang 2011] we have a most interesting datum. Bouveret and Lang study policies under min-max (which they refer to as egalitarian), i.e., the policy for which the worse-off player's expected value is highest. In [Bouveret and Lang 2011, Table 1], they list the optimal policies for even lengths up to twelve items. The min-max optimal policy for an even number of items in their listing turns out to be within the Boolean set. It is not always the fairest, but rather, the following appears to be the case:

Conjecture 16. *The fairest policy among those in the Boolean set is the min-max optimal policy.*

It is plausible that a policy where the disadvantaged player does well is a relatively fair policy, *and* Bouveret and Lang establish that an alternating policy tends toward egalitarian optimality as length grows, so there seems to be multiple pieces of evidence that this conjecture is reasonable. It would be interesting if it turned out to hold.

As a perhaps trivial but astonishing note, we find that for length 14, the fairest choice sequence has L_{adv} *exactly* 0! Compare this to the alternating sequence, which for length 14 has $L_{\text{adv}} \approx 3.95$, or Luis being advantaged by more than half the number of items each player takes.

An interested investigator might be able to improve (1) by converting the falling factorials into binomial coefficients, and repeatedly applying Abel-type summation identities which sum shifts of binomials. The recurrence of Kalinowski, Narodytska and Walsh is far more useful than this formula, but a closed form might reverse the situation.

Finally, we recall our conjecture on dominance, which in its full generality remains open and which we consider a most intriguing question regarding the fairness condition:

Conjecture 10. *Among choice sequences σ_i of the same rank, if σ_1 dominates σ_2, then σ_1 has higher average advantage for Luis than σ_2.*

Partial approaches might include extending the theorem to dominance moves $LR \dots R \dots RL \to RL \dots R \dots LR$, with bounds on the difference in number of L and R before and after the changing segment.

References

[Bouveret and Lang 2011] S. Bouveret and J. Lang, "A general elicitation-free protocol for allocating indivisible goods", pp. 73–78 in *Proc. of 22nd Int. Joint Conference on Artificial Intelligence* (Barcelona, 2011), vol. 1, edited by T. Walsh, AAAI Press, Menlo Park, CA, 2011.

[Brams and Taylor 1996] S. J. Brams and A. D. Taylor, *Fair division: from cake-cutting to dispute resolution*, Cambridge University Press, 1996. MR Zbl

[Brams and Taylor 2000] S. J. Brams and A. D. Taylor, *The win-win solution: guaranteeing fair shares to everybody*, W. W. Norton, New York, 2000.

[Hopkins 2010] B. Hopkins, "Taking turns", *College Math. J.* **41**:4 (2010), 289–297. MR Zbl

[Hopkins and Jones 2009] B. Hopkins and M. A. Jones, "Bruhat orders and the sequential selection of indivisible items", pp. 273–285 in *The mathematics of preference, choice and order*, edited by S. J. Brams et al., Springer, 2009. MR Zbl

[Kalinowski and Narodytska 2013] T. Kalinowski and N. Narodytska, "A social welfare optimal sequential allocation procedure", pp. 227–233 in *Proc. of 23rd Int. Joint Conference on Artificial Intelligence* (Beijing, 2013), edited by F. Rossi, AAAI Press, Menlo Park, CA, 2013.

[Rubchinsky 2010] A. Rubchinsky, "Brams–Taylor model of fair division for divisible and indivisible items", *Math. Social Sci.* **60**:1 (2010), 1–14. MR Zbl

[Stanley 1997] R. P. Stanley, *Enumerative combinatorics, I*, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, 1997. MR Zbl

Received: 2016-07-08 Revised: 2017-12-10 Accepted: 2017-12-30

wjkeith@mtu.edu *Department of Mathematical Sciences, Michigan Tech University, Houghton, MI, United States*

sean.grindatti@gmail.com *Department of Mathematical Sciences, Michigan Tech University, Houghton, MI, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2019 is US $/year for the electronic version, and $/year (+$, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2019 vol. 12 no. 1