OPTIMAL PICKING SEQUENCES FOR FAIR DIVISION OF INDIVISIBLE GOODS

A PREPRINT

Komi Amiko
No, this isn't going to get published

December 10, 2019

ABSTRACT

The Thue-Morse sequence can be used to fairly assign indivisible goods between two agents, and it is the optimal picking sequence for this purpose. Numerous other papers have analyzed fair division problems for more than two agents, yet when they discuss methods using picking sequences, they focus more on the social aspects and complications which could arise in actual use. Viewing picking sequences as nothing more than integer sequences, their mathematical properties remain elusive. Here we discuss our findings on picking sequences and pose some open problems.

***K*eywords** Fair division, Picking sequences, Algorithms

Contents

1 A standard game of chance

Before we can discuss picking sequences, we should first define them in terms of a fair division game.

Let n be the number of agents between which the goods will be divided. Label the agents $0, 1, \cdots, n-1$. Let $\mathbb{M}$ be the pool of indivisible goods which are to be divided among the agents. Let $S_0, S_1, \cdots$ be the picking sequence, where each S_i is the label of an agent. On turn m, starting with $m = 0$, agent S_m must take exactly 1 item from $\mathbb{M}$. After $|\mathbb{M}|$ turns, we will get some allocation of the indivisible goods among the agents.

We now define a simple game of chance which generates a picking sequence.

Each agent is a player. On turn m, player S_m goes, where S_m is the player who had the least chance of winning the game up to that point. With probability ε for some infinitesimally small ε, player S_m wins and the game ends, and with probability $1 - \varepsilon$, the game continues.

We can see that the probability that the game ends on turn m and thus granting victory to player S_m is given by $\varepsilon(1-\varepsilon)^m$.

We can rephrase this to be more similar to the fair division game. Let $\mathbb{M} = \{(1-\varepsilon)^m : m \in \{0, 1, \cdots\}\}$. All we changed was divide all the probabilities by ε. Let all agents have the same utility function defined by $u(S) = \sum S$. On turn m, whichever agent has the least utility so far will be S_m and thus allowed to take the next largest number from $\mathbb{M}$.

To tiebreak the first n allocations, we will set a convention that $m < n \implies S_m = m$, that is, the first n items of S are $0, 1, \cdots, n-1$. Without this convention, all picking sequences that could be generated by this procedure would be equal up to relabeling of the agents, that is, if picking sequence S' could be generated by this procedure without this convention, there exists a permutation P of $(0, 1, \cdots, n-1$ such that for all m, $S_m = P[S'_m]$. To prevent this convention from causing a bias in the item assignment, agents should be labelled randomly.

This can be considered a greedy algorithm for generating S, since it makes the apparent best choice at each turn without looking ahead and never changes that choice later.

We will not show the general optimality of this S, because it is indeed not optimal for every situation. However, it does have a desirable property. Any actual situation it is applied to has a finite number of items, say, m. Suppose we already have the best possible prefix of S for those first m items. If we add another item, we now have $m + 1$ items, and we are not allowed to change the first m terms of S, so the best we can do is make the best choice we can for S_m. This may be worse than an item assignment specifically constructed for those $m + 1$ items. At $m = 0$, we are not assigning any items, so surely the optimality criterion is already met. Thus, with the constraints on S, we can do no better than a greedy algorithm.

This is similar to how a Sobol sequence provides the best uniform sampling for an arbitrary prefix from the infinite sequence, but it the number of samples is known beforehand, a better sampling could be devised for that specific number of samples.

This picking sequence is of theoretical interest and is based on a very idealized situation. In practice, it is approximately optimal for the Nash social welfare when, for every agent, the most they value any individual item is close to the least they value any individual item. This could be used to rigorously justify it as a fair picking sequence.

For $n = 2$, S is the Thue-Morse sequence.

2 Analysis using polynomials

For purposes of computation and analysis, instead of working with a number system that supports an infinitesimal ε, we can work with polynomials of ε with integer coefficients. Each $(1-\varepsilon)^m$ will expand like so:

$$(1-\varepsilon)^m = 1 - \binom{m}{1}\varepsilon + \binom{m}{2}\varepsilon^2 - \binom{m}{3}\varepsilon + \cdots \tag{1}$$

Ordering of polynomials is done by lexicographic order of the coefficients, with the constant term first and higher degree terms later.

To avoid needing arbitrary length polynomials, it will be useful to define a truncation of a polynomial. We define the truncation function T_p to remove all terms with degree higher than p from its argument.

$$T_p(a_0 + a_1\varepsilon + a_2\varepsilon^2 + \cdots + a_q\varepsilon^q) = a_0 + a_1\varepsilon + a_2\varepsilon^2 + \cdots + a_{\min\{p,q\}}\varepsilon^{\min\{p,q\}} \tag{2}$$

We may also define a function $M_n(p)$ to be the m of the first term S_m which could not be decided by only looking at truncations of degree less than p, that is, after the first m items have been assigned, if v_k is the polynomial representing the total value owned by agent k, then there exists i, j with $i \neq j$ such that for all $q < p$, $T_q(v_i) = T_q(v_j)$. To ensure values of this function are interesting, we ignore the first n terms of S_m, which are decided by the convention.

Then, for n players, using only fixed size arrays storing polynomials of degree up to p, it is possible to compute exactly the first $m \leq M_n(p)$ terms of S in time bounded below by $\mathcal{O}(mp)$ and bounded above by $\mathcal{O}(mnp)$, assuming arithmetic is $\mathcal{O}(1)$, which may not be true since *bigintegers* are needed in practice.

It is possible that for sufficiently large p, $M_n(p)$ does not exist because the condition, that "the degree p term matters" as described previously, is never met. If $M_n(p)$ does not exist, then it could be treated as if it was ∞.

2.1 Properties of S

Redefine $v_k(m)$ as the total value owned by agent k after the first m items have been assigned. Define $w[p]$ as the coefficient of ε^p in the polynomial w.

At $m = 0$, we have all $v_k(m)[0] = 0$. Suppose that, at $m = qn$ for some q, we have all $v_k(m)[0] = q$. Some agent k will receive the next item, at which point they will have $v_k(m)[0] = q + 1$. They will be unable to receive another item until every other agent i also has $v_i(m)[0] \geq q+1$, and by similar logic, no other agent is capable of skipping to $v_i(m)[0] = q+2$ while another agent j is still at $v_j(m)[0] = q$. There are n agents to increment the $v_k(m)[0]$ of, and during the n turns, every agent is chosen exactly once. Thus, for every q, $S_{qn}, S_{qn+1}, \cdots, S_{qn+n-1}$ is a permutation of $0, 1, \cdots, n-1$.

Using the previous finding, starting at an even q, suppose all agents k have $v_k(m)[1] = Q$. At $q = 0$, $m = 0$, so $Q = 0$ and this holds. After n turns, every agent has been assigned an item. This time, we are not adding 1, but instead $-\binom{m}{1} = -m$. So now $v_{S_{qn+i}}(qn+n)[1] = Q - qn - i$. This will have no duplicates. The next n terms of S are decided entirely by the $v_k(m)[1]$. We can see that the next agent to receive an item is $S_{qn+n} = S_{qn+n-1}$, since they had the least $v_k(m)[1]$. Now their $v_k(m)[0]$ is higher, so they will not be considered until $m = qn+2n$. The agent with the next lowest value is $S_{qn+n+1} = S_{qn+n-2}$, and so on. We can see that, at $m = qn+2n$, all $v_k(m)[1]$ are equal again. So we have some kind of pattern repeating every $2n$ turns, where $S_{qn}, S_{qn+1}, \cdots, S_{qn+2n-1}$ is a palindrome for all even q.

We reuse these results as we examine the next higher degree. Consider the block $S_{2qn}, S_{2qn+1}, \cdots, S_{2qn+2n-1}$. An agent will be given an item at $m = 2qn+i, m = 2qn+2n-i-1$ for some $0 \leq i < n$. Substitute these into $\binom{m}{2} = \frac{m(m-1)}{2}$. Their $v_k(m)[2]$ increases by a total of $i^2 + (-2n+1)i + 4n^2q^2 + 4n^2q + 2n^2 - 4nq - 3n + 1$. We can ignore all the terms that do not depend on i, since they will be the same for all agents. Thus the relative increase is $i^2 + (-2n+1)i$, which is decreasing from 0 to n. Define $\alpha(i) = i^2 + (-2n+1)i$. Starting from $q = 0$, $v_k(2n)[2] = \alpha(i) + C$ for some constant C common to all agents. The order of taking items for $q = 1$ is determined by ascending order of $\alpha(i)$, so it will be the reverse of the first n items. Thus $S_{2n+i} = n-i-1$. Let $\beta(i) = \frac{1}{2}(\alpha(i) + \alpha(n-i-1) + C) = i^2 + (-n+1)i$ be the relative increase in agent i's $v_i(m)[2]$ from $m = 0$ to $m = 4n$.

For even n, agents $i = \frac{n}{2}, j = \frac{n-2}{2}$ have $\beta(i) = \beta(j)$, so $v_i(4n)[2] = v_j(4n)[2]$ and thus $M_n(3) = 4n$ for even n. For odd n, β has a unique minimum at $\frac{n-1}{2}$, but after assigning them the next item, agents $i = \frac{n+1}{2}, j = \frac{n-3}{2}$ are now tied, so $M_n(3) = 4n+1$ for odd n.

We leave some open problems for possible future work to attack.

1. **For $n \geq 3$, S contains no squares X^2 with $|X| \geq 2$.** It is known that S for $n = 2$ could not have this property, but it is suspected to be true for higher n. This would imply S is not eventually periodic.
2. S **contains no cubes X^3.** It is known that S for $n = 2$ has this property, and it is suspected to be true for higher n. It would imply S is not eventually periodic.
3. S **is "chaotic".** We only analyzed the "initial state" of $v_k(0) = 0$ for all agents k. This property holds if every initial state with $v_0(0) = 0$ results in a different sequence. It requires, and thus implies, that $M_n(p)$ always exists. It would also imply S is not eventually periodic, since removing a prefix of S is indistinguishable from using the state reached by then as an initial state.
4. S **is not eventually periodic.** Somewhat stronger than S being aperiodic. Is implied by some other stronger conditions mentioned here. It seems likely to be true but is yet unproven.
5. **Efficient algorithms for S.** There may exist efficient algorithms for generating S which improve significantly on the "truncated polynomial array method" discussed here. It also maybe possible to prove that no such efficient algorithm exists.

2.2 Properties of M

It is known that $M_2(p) = 2^p$ exactly. This could be proven rigorously, though it is easy to see in the Thue-Morse sequence.

$M_n(1) = n$, as all $v_k(n)[0]$ are equal, though it can be argued that since this comes just after the block that was decided by the convention on the first n terms, this sequence value is not meaningful.

$M_n(2) = 2n$, as all $v_k(2n)[0]$ are equal and all $v_k(2n)[1]$ are equal, and there could not possibly be a tie for lowest $v_k(m)[1]$ between $m = n$ and $m = 2n$.

$M_n(3) = \begin{cases} 4n & 2 \mid n \\ 4n+1 & 2 \nmid n \end{cases}$ as stated in 2.1. This can be stated without a branch as $\frac{1}{2}(8n + 1 - (-1)^n)$.

We have also found some exact values and lower bounds of $M_n(p)$ for other n, p. A lower bound written as $> N$ should be interpreted as "$M_n(p) > N$ if $M_n(p)$ exists."

$p \setminus n$	2	3	4	5	6	7	8	n
1	2	3	4	5	6	7	8	n
2	4	6	8	10	12	14	16	$2n$
3	8	13	16	21	24	29	32	$\frac{1}{2}(8n+1-(-1)^n)$
4	16	42	48	40	264	$> 10^9$	180	
5	32	$> 10^9$	$> 10^9$					
p	2^p							

$M_n(4)$ displays irregular behaviour. The only exactly known $M_n(5)$ is for $n = 2$.

$M_n(p)$ is increasing in p. While $M_n(p)$ is not increasing in n, as seen with $M_4(4) > M_5(4)$, M_n appears to grow faster as n increases. $M_3(5)$ may be very large, say, greater than $2 \uparrow\uparrow 5$, or it may not exist at all. If so, $p = 5$ would be sufficient for computation of S for all $n \geq 3$ for any practical m.

We leave some open problems for possible future work to attack.

1. **Existence of M.** Does $M_n(p)$ always exist for $n \geq 2, p \geq 1$?
2. **Growth of M.** If $M_n(p)$ does exist for a meaningfully large range of n and p, how fast does it grow?

References