Equitable Sequencing and Allocation Under Uncertainty

Arthur Dolgopolov*

October 27, 2016

Abstract

The paper uses Thue-Morse sequence and knapsack problem to formally relate several problems of negotiations, fair division and tournament sequencing. All of the considered problems are shown to be reduced versions of partitioning problem, i.e. equally dividing items into two sets. An argument is made that it is often beneficial to set the partitioning problem as stochastic with item values drawn from a distribution, especially but not limited to the case of intangible items like skill level. Balanced alternation according to Thue-Morse sequence proves to be the better mechanic in this case for all sample sizes for common distributions. This is not true for all distributions and for settings with $k > 2$ agents. A further improvement upon this mechanic is greedy approximation over the underlying subset-sum problem.

Keywords: Thue-Morse Sequence, Fair Division.

JEL Classification Numbers: D00, Z00.

*College of Humanities and Social Sciences, George Mason University, e-mail: *adolgopo@masonlive.gmu.edu*. I am grateful for helpful comments from Marina Inina, Mikhael Freer, Marco Castillo, Robin Hanson.

1 Introduction

An integer sequence named after Axel Thue and Marston Morse (Thue, 1977; Morse, 1921), was repeatedly independently rediscovered throughout last hundred years as a solution to seemingly unrelated problems in graph theory and number theory, game theory and economics [1].

Thue-Morse sequence is an integer sequence formed recurrently by appending the binary complement of the sequence obtained thus far:

$$0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ \ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0... \tag{1}$$[2]

Consider the following problem (the "backyard ball game"). Two team captains are dividing players into two teams for a ball game. If two captains only have ordinal valuation of players, but players do actually have cardinal skill (half-normally distributed), ordering pickings according to TM sequence can be shown to dominate other sequencing mechanics. That is if team captains will alternate according to TM sequence in picking the next player, they will get close to equally matched teams. The same setting can be described in terms of a firm dividing employees between identical branches with decreasing returns to scale - while it is relatively easy to rank employees on such proxies from their CVs like education level, the actual marginal product of hiring the next worker is unknown.

This problem is a further study of several other directly related problems, present in the literature. In a recent article Palacios-Huerta (2012) suggested Thue-Morse (TM) sequence as a solution to the problem of fairly ordering a sequential tournament. That is if the first-mover in each round of a repeated game has an unknown advantage (psychological or other kind), the binary negation of the move order in the next rounds will tend to compensate

[1] Allouche and Shallit (1999) and Allouche (forthcoming) provide some interesting applications and pointers to other studies.

[2] Sequence number A010060 in the On-Line Encyclopedia of Integer Sequences (OEIS).

for it. So if, for example, the order $\{0\ 1\}$[3] in first two rounds gives any advantage to either of two players, then by negating this order in rounds three and four $\{1\ 0\}$, we will tend to compensate for that advantage. The same principle applies recursively for the next 4 rounds and so on for the next 2^n rounds. The paper also implied that the sequence must have $n = 2^{(m+1)}, m \in N$ elements to maximize the potential of this sequencing mechanism.

Both the sequencing of a tournament and "backyard ball game" problems can be generalized into a partitioning problem, i.e. equally dividing a set of items between two parties, given the item value does not have to be known ex ante, and items are drawn from a certain distribution. This makes the partitioning problem the most generic of the settings.

The present paper argues that optimal sequencing or partitioning in all of these problems largely depends on our assumptions about the set of items A - the sequence of first-mover advantages in the tournament rounds or the underlying distribution of items (or skill) to be divided in the partitioning and backyard ball game problems. This fact is often obfuscated, but without these assumptions the problems are not well-defined.

The paper is organized as follows. The next section briefly describes four settings, which prove to be versions of the same problem. Section 2 presents the terminology used in the paper. Section 4 analyzes the most general case - the partitioning (fair division) problem. When the items to be divided are of unknown value, this is the backyard ball game problem, described above. We briefly discuss the trivial case of uniform distribution, the exponential distribution and run TM sequence against trivial sequencing mechanisms in silico for half-normal distribution. It does dominate other strategies under reasonable assumptions for any number of items. We also compare it with a more direct approach of using a greedy heuristic on order statistics of the distribution.

[3] 1 indicating player one moving first and vice versa.

2 Terminology

We use terminology from Brams and Taylor (1999). Let “one” denote first players turn to choose an item and “zero” - his opponents. Allocation according to a sequence {0 1 0 1...}, i.e. players simply taking turns picking items from the set is called simple alternation. A slightly more complex allocation according to a sequence {0 1 1 0 0 1 1 0 0...} is called a “modified draft”.

Allocation rule that sequentially assigns the next best item according to Thue-Morse sequence (“taking turns taking turns” in Brams and Taylor terms) is called “balanced alternation”. That is first party picks an item first, than second party picks two items, than first party picks again and then the whole sequence is inverted - {0 1 1 0 1 0 0 1...}. This way any advantage potentially given by the first $2^n, n \in N$ turns is being compensated by inverting the sequence in the next 2^n turns.

3 Problem Settings

We describe four settings from economics and mathematics literature that can be approached in the same way. Note that although seemingly unrelated, latter settings are in fact just reduced forms of the first one, which is the most generic.

3.1 Setting #1: Partitioning or fair division (inheritance)

There are n items of unequal value. The objective is to divide items between two parties with minimum discrepancy between total sums for each agent. Item value can be known in advance, or items can be drawn from a known distribution. More formally:

Problem 1 ("Partitioning problem") *For two agents and a given set of items A with valuations a_i^1 for the first agent and a_i^2 for the second agent, $i \in N$, find a partitioning procedure, imposed by a third party, who is not aware of agents' preferences, such that the* *discrepancy in the value of items assigned to agents is minimized. In case 1 the items have fixed values of ambiguous shape. In case 2, items are drawn from a distribution F known to the mediator.*

In computer science, partitioning problem is to divide a set of integers into two subsets S_1 and S_2, such that the difference between sums of all elements in S_1 and S_2 is minimized. We, however, following Brams and Taylor, are concerned with bargaining form of this problem for two-party disputes without arbitration. That is - can two parties facing the problem (dividing n items of unequal value) agree on a satisfactory procedure enforced by an impartial referee, who has no knowledge of their preferences. For this, the outcome of the procedure has to be envy-free, i.e. each agent gets at least a half in his own valuation. If the valuations are equal, the objective is to bring the discrepancy down as much as possible. We consider the case of two parties, in which proportionality (every agent gets at least 1/[number of agents] share by his subjective valuation) and envy-freeness (agents believe that their share is at least as good as any other share) criteria of fairness are equivalent[4].

Desired quality of envy-freeness makes the problem a search for "ex ante" fairness. If the valuations were known to the mediator, the sequencing mechanic could be avoided altogether and the problem would reduce to an NP-complete, but purely computational subset-sum problem. However, with the item value unknown to the mediator, the first-mover advantage is unavoidable, forcing the designer to consider various sequencing mechanics. With problem defined like this the desired result is also "ex post" efficient by definition as it attempts to provide equal outcomes by bringing the discrepancy gap to the minimum.

Brams and Taylor (1999) argue that TM sequence ("balanced alternation" in the book) is a solution to this problem. That is, if we order parties moves according to TM sequence and let each party pick any available item at every move, this would allow for a reachable consensus that will at least always dominate "strict alternation" (picking items one-by-one taking turns). The argument is the same as for tournament sequencing - negating the

[4]See cake-cutting literature for in depth description of fairness criteria, e.g. Brams and Taylor (1995, 1999)

order of picking tends to compensate for any "unfair" advantage in the beginning of the procedure. It is important to note that Brams and Taylor are dealing with negotiations, that can potentially involve intangible issues, not objects.[5].

The full Brams and Taylors solution is two-step:

1. Query step. Before the TM procedure, the mediator iteratively asks each party to pick their most favorite object in the pile. If they pick different items, they are simply awarded the objects and the query step is iterated again on the remaining items. If they pick the same item, the item goes to "contested pile", and query step is applied again to the remaining items. Once all the items are either "contested" or awarded, the TM sequence is applied only to "contested pile".

2. Balanced alternation. Parties take turns according to TM sequence in picking items from "contested pile".

The query step from Brams and Taylor procedure allows us to focus on the contested pile only, where order of valuations is the same for both agents.[6] If we follow Brams and Taylor procedure to the letter including the "query step", we always end up with a "contested" pile of items whose order of preference is the same for both parties. Therefore we focus on the case when $a_i^1 = a_i^2 \ \forall i$ and, for simplicity, let utility equal item value. Note that the query step also acts as a revealed preference tool for the mediator, showing the order of preference over the "contested pile".

The present paper expands the problem to stochastic setting of values unknown to the mediator, but drawn from a distribution (or equivalently - values drawn after the division procedure is settled ex ante). Without it, both the approach stated above and Palacios-Huertas conclusions about tournament sequencing can be misinterpreted. This leaves us with two cases:

[5]They do however note that such sequencing techniques are too crude a device to be applied to issues and suggest another approach, "adjusted winner". It is hard to imagine players anticipating the distribution of the magnitude of intangible issues, so this is not an important consideration here.

[6]See Brams and Taylor (1999) for the discussion of the procedure and Brams et al. (2014) for the enhanced query step procedure and discussion of its envy-freeness.

Case 1: known values

If the value of the items was known (as in computer science version of the problem), this is just a subset sum optimization problem, a special version of the famous knapsack problem. That is we are dividing items according to their known dollar value into two groups as equally as possible. The exact mathematical solution of this well studied optimization problem is NP-complete, but commonly used greedy heuristics (Dantzig, 1957) seem like a fair representation of real-life agents with limited complexity of their calculations. The mediator can just go over all items in reasonable time if n is small or apply a heuristic. Moreover, even if we treat the NP limit seriously, there is no real practical need for TM sequence since greedy heuristics for underlying knapsack problem will perform much better than balanced alternation anyway (this is shown below). This heuristic solution could be: order items according to their value (decreasing) and iteratively assign the next item to a party with less overall value so far. This procedure would yield more equalized results than balanced alternation.

To illustrate this with a simple scenario, consider a set consisting of a very expensive item (e.g. 100 of value) and 3 cheap items (10 of value). TM sequence result would be 110:20 as the last pick will be by the first mover (the n=4 sequence is {0 1 1 0}). It is however apparent that 100:30 result can be achieved by giving all the cheap items to the second player. It is clearly more balanced and more likely to be reached in real life negotiations than anything produced by a reasonable rule of taking turns.

If actual item values are unknown, but they are known to be half-normally distributed, applying the greedy heuristic will yield a different and better performing sequence as well. In that case, it cannot be enforced, however, as greedy heuristic requires one to be able to compare sets of items. Consider two team captains picking players for their teams again ("backyard ball game"). If it was possible to change the order of picking turns in such a way that the captain with the weakest team so far is always picking next, the outcome would be better equalized. However, captains cannot precisely compare teams, unless they know the exact value of each player.

Case 2: unknown values

If the value is unknown, the revealed ordinal ranking of the items does not give us enough information to use the exhaustive or heuristic approaches. Returning to the same example, both the 110:20 and 100:30 outcomes are still present. To differentiate between these outcomes we have to make assumptions about the distribution of cardinal values of the items. If we only keep the ordinal rank, there can be arguments for giving the 4th best item to either side. Similarly for tournament sequencing problem, we have to make assumptions about the dynamics of the magnitudes of advantages in each round.

Therefore, I propose to set the problem as stochastic with item values drawn from a distribution F. Only then the sequencing mechanism is required, the first-mover advantage is unavoidable and the apparent exhaustive search solution from known value case is not applicable. The problem can be equivalently defined with mediator having no knowledge of the value or with value of the items being unknown at the time of deciding on the procedure. This also seems like a plausible setting for any item set that is inherently hard to measure or quantify like, for instance, skill level.

This stochastic setting is relevant for several other problems, listed below, which are ultimately reduced forms of partitioning. Since partitioning problem is the most general case of the four settings - changing the law of distribution and imposing information restrictions is enough to transform this problem into any of the three below - this setting is used in the final section of this paper.

3.2 Setting #2: Backyard ball game

Two team captains pick players for their teams from a large half-normally distributed set. Captains can rank players in the sample from worst to best, but cannot put exact number on a persons skill. What is the optimal procedure to minimize the discrepancy in skill between two teams?

This setting is identical to partitioning problem with uncertainty about value of the items. As there is no strategic advantage to not picking the best available player, captains are only limited in doing so by the order of turns. Therefore, since captains can rank players without knowing their exact value, the dynamics of this setting are the same as for the partitioning problem with items drawn from a known distribution.

3.3 Setting #3: Galois duel ((Cooper and Dutle, 2013))

Two duelists take turns shooting at each other. They use greedy approximation in a sense that in each round the shooter is determined as the one with less a priori probability of winning.

In the Galois duel first-mover advantage stems from the chance of ending the duel and not giving the other party a chance to shoot. Note that Galois duel setting is identical to partitioning problem with items distributed according to a power law (every item in the set is better than the next best one by x percent). Indeed, each next shot is worth x percent less as it raises a duelists probability of winning only if he is not killed in the previous turns.

3.4 Setting #4: Tournament sequencing (Palacios-Huerta, 2012)

In a sequential tournament, there are psychological, strategic or other factors that raise chances of winning for the first-mover. What should be the order of a sequential tournament to make it ex-post fair? More formally:

Problem 2 ("Tournament sequencing") *For a sequence of advantages b_i in a tournament, where i is the round index, find a set S, such that the value of $E(|\sum_{i \in S} b_i - \sum_{j \notin S} b_j|)$ is minimized.*

As with the partitioning problem, known sequence of b_i turns the optimization problem into a purely computational subset-sum problem. Assume that only some general properties of b_i are known.

Unlike the Galois duel where advantages stem from the probability of ending the game early, here they arise due to other factors, which cannot be measured directly. There is indeed empirical evidence suggesting that there exists a first-mover advantage in soccer and tennis tournaments[7].

Remark 3 (Equivalence) *Problems 2 (tournament sequencing) and 1 (partitioning a list of items) are identical by construction.*

Consider the partitioning problem with equal agent valuations over items (due to the query step as described in previous section), $a_i^1 = a_i^2\ \forall i$, $A = \{a_0, a_1 ... a_n\}$. Let binary sequence $S = (s_1, s_2, s_3 ... s_k)$ represent the order of turns of agents picking items from A, imposed by the third-party mediator. Then there is no benefit to strategically not picking the best available item at every turn, i.e. agents always pick the highest valued item in the remaining pile. Therefore each element s_i in S determines the assignment of i-th order statistic from set A. If we calculate the order statistics for the whole set A, the problem becomes the tournament sequencing one with sequence b_i consisting of the order statistics drawn from the partitioning problem distribution. To transform the partitioning problem into tournament sequencing one, the set of items needs only be replaced with the set of order statistics and vice versa. For example, the partitioning problem of dividing a pile of uniformly distributed items between two agents (with agents having equal valuations unknown to the mediator) is equivalent to the problem of ordering a sequential tournament between two agents with advantages in consecutive rounds being random variables drawn from Beta distribution (since k-th order statistic of the uniform distribution is a Beta random variable).

The setting depends on the dynamics of the first-mover advantage as the tournament progresses. Suppose there is a high unobservable advantage for the first mover in the first round and no advantage whatsoever for the rest of the game. In this scenario there is little one can do to compensate for that advantage.

[7]See Ruffle and Volij (2014) for literature overview. The author also considers other sequencing rules like the sequence of winner/loser moving next and 5-4 rule, which do not relate to fair division problem and fall out of scope of the present paper.

Now consider the scenario where advantage gradually diminishes by x percent every round. This turns the problem into the Galois duel problem, where the heuristic solution leads to balanced alternation if x is sufficiently low (see Cooper and Dutle, 2013). If every next rounds advantage is 1 percent less valuable in terms of winning then the first 25 rounds stabilize to be Thue-Morse sequence.

If the advantage gradually diminishes by a constant, or diminishes unevenly, the sequencing problem can be modelled as a partitioning problem. We now proceed to analyze the partitioning problem to show that the particular sequencing mechanic (balanced alternation) is superior for most, but not all, common distributions. We also note that the efficiency is further improved by directly solving the subset-sum problem with a greedy heuristic algorithm for a set of order statistics from a known F.

4 Partitioning Problem with Stochastic Item Values

Consider a set A of valuable items drawn from a known distribution F. Ranking items from best to worst, let a_k be the k-th best items value. The task is to find a set S, such that the value of $E(|\sum_{i\in S} a_i - \sum_{j\notin S} a_j|)$ is minimized. Due to linearity of expectation:

$$E(|\sum_{i\in S} a_i - \sum_{j\notin S} a_j|) = |\sum_{i\in S} E(a_i) - \sum_{j\notin S} E(a_j)|, \tag{2}$$

where a_i and a_j are values of items, S and $A \setminus S$ are sets assigned to each party. After reordering the drawn items according to their expected value, $E(a_k)$ becomes simply the expected k-th order statistic of underlying distribution F. Therefore to solve problem 1 and minimize the expected discrepancy between the parties, one needs to solve the partitioning problem with values equal to expected order statistics of the distribution F.

We first summarize the simpler cases of deterministic power law and uniform F for comparison and then proceed with the half-normal case.

We will make use of the following lemma.

Lemma 4 *TM sequence and modified draft sequence only differ for 4 consecutive positions after positions* $i = 2[Eviloddnumbers] - 1 = 8n - 4[TM] - 3$. *That is for 4 positions after* $5, 9, 17, 29, 33, 45, 53$. *These four consecutive elements are* $\{0\ 1\ 1\ 0\}$ *and the corresponding negation.*[8]

To show this, set up TM sequence with alphabet $\{0, 1\}$. This sequence is a morphic word obtained with rules $(0 \to 01)$, $(1 \to 10)$. This implies that TM sequence can also be obtained with rules $(0 \to 0110)$, $(1 \to 1001)$. Each of this four-element blocks always corresponds to $\{0\ 1\ 1\ 0\}$ block in modified draft sequence, which consists of this block repeated indefinitely. Therefore, if we divide TM and modified draft sequences into consecutively numbered four-element blocks, these sequences would only differ at positions $\{TM(n) = 1\}$, n being the index of a four-element block.

So for four consecutive positions modified draft and balanced alternation assign items to opposite players. Modified draft is fully defined by this - the four elements are $\{0\ 1\ 1\ 0\}$ or, equivalently, $\{1\ 0\ 0\ 1\}$ (if it was $\{(1)\ 0\ 0\ 1\ 1\}$, TM sequence would contain the sequence (1) 1 1 0 0 in it, while TM sequence is known to be cubefree). ■

We now discuss two simpler cases - deterministic and uniform.

Deterministic power law case (Galois duel). Two-way partitioning problem for a set of items, where every item in the set is better than the next best one by x percent is identical to the problem in Cooper and Dutle (2013). As shown in the original paper, when x is sufficiently low, optimal partitioning procedure always tends to balanced alternation.

Stochastic uniform case. The uniform case is trivial and all sequencing mechanics produce the same outcome. This is formulated as the following and proven in the appendix.

Proposition 1 (Uniformly distributed items) *If the distribution of items to be partitioned between two agents is uniform,* $F \sim U$, *balanced alternation does not have advantage over modified draft and they produce identical results. Partitioning is always perfect with zero*

[8] See TM properties as the sequence number A010060 in the On-Line Encyclopedia of Integer Sequences (OEIS).

discrepancy if sample size allows for it (any n congruent to 0 or 3 mod 4). Both procedures dominate strict alternation.

In general for balanced alternation to be a preferred sequencing mechanic, the expected values of order statistics should be "cyclical" in the following sense. It follows from lemma 4 that the sufficient condition for balanced alternation to dominate modified draft is -

$$
\begin{aligned}
E(G_{r,n}) - E(G_{r+1,n}) - E(G_{r+2,n}) + E(G_{r+3,n}) &> 0, \\
r = 4i+1,\ r+3 \leq n\ \forall i, n \in \mathbb{Z}&, \\
or E(G_{r,n}) - E(G_{r+1,n}) - E(G_{r+2,n}) + E(G_{r+3,n}) &< 0, \\
r = 4i+1,\ r+3 \leq n\ \forall i, n \in \mathbb{Z}&
\end{aligned} \tag{3}
$$

where TM(x) is TM sequence at position x. and $G_{r,n}$ is the expected value of the rth largest order statistic from a sample of size n from distribution F:

$$E(G_{r,n}) = \frac{n!}{(r-1)!(n-r)!} \times \int_{-\infty}^{\infty} x(1-F(x))^{r-1} \times (F(x))^{n-r} \times f(x)dx \tag{4}$$

That is out of any four consecutive order statistics, the first and fourth has to be consistently better or worse than second and third. Then every four element block gives advantage to one of the parties and reversing order in some of them is beneficial. If there are sign reversals in (3), however, the modified draft can outperform balanced alternation for some r and n.

Stochastic exponential case. Condition (3) holds, for example, for the exponential distribution:

Proposition 2 (Exponentially distributed items) *In the partitioning problem, TM sequence dominates modified draft for exponential F.*

To show this, recall that expected value of r-th largest order statistic from exponential distribution with the mean θ is

$$E(G_{r,n}) = \theta \sum_{j=1}^{n+1-r} \frac{1}{(r-1+j)} \tag{5}$$

Therefore the difference between two consecutive expected values of order statistics is θ/r, which is strictly decreasing in r and independent of n. Hence (3) always holds. ■

This way, following Palacios-Huerta (2012), the negation of this order will lead to smaller discrepancy with the next four elements. This logic continues for the repeated 8-element draft - if each block of 8 consecutive order statistics gives advantage to the same player, Thue-Morse sequence will dominate 8-element modified draft, etc. However, it is possible to improve upon all of these cases altogether, by distributing expected order statistics manually according to a greedy heuristic. This is beneficial, if there is improvement to be gained unlike the uniform distribution case. We demonstrate this with stochastic half-normal F.

Stochastic half-normal case. The case that is of particular interest to us is of items distributed half-normally. Indeed this is one of the likely cases in the "backyard ball game" setting - player skills drawn from a folded normal distribution with a few very good players and increasingly higher mass of inferior players.

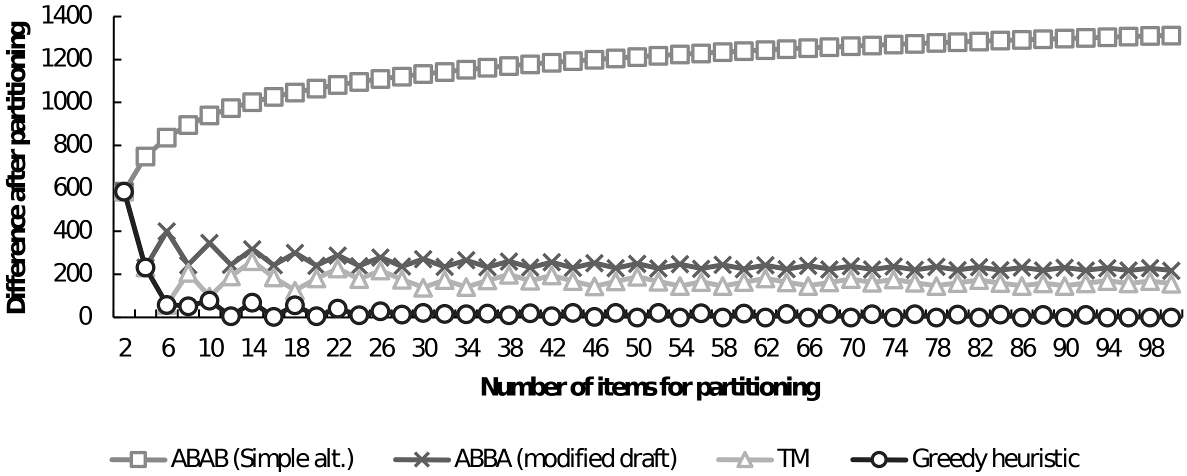
Proposition 3 (Half-normally distributed items) *When items are distributed half-normally, balanced alternation dominates modified draft for any number of items.*

Computing expected values of order statistics of half-normal distribution is computer-intensive, though there are reasonable approximations (see Appendix B in Olgun and Fearn (1997). We use exactly calculated expected values for the first 100 sample sizes.

As illustrated by Figure 1, advantage of TM algorithm is actually the highest at points $n = 6, 10, 18, 30, 34, 46, 54, 58, 66, 78, 86, 90, 102, 106....$ One can see that this is a sequence resulting from doubling each number in the sequence of odd evil numbers: $3, 5, 9, 15, 17, 23, 27, 29, 33, 39, 43, 45, 51, 53....$[9] In fact, Cooper and Dutle (2013) arrive to

[9]John Conway (see Berlekamp et al., 1982) coined evil and odious numbers terminology as phonetic play on words even and odd. Formally an evil number is a number with an even count of 1's in binary expansion. This sequence often arises in problems related to Thue-Morse sequence (see Allouche, forthcoming).

Figure 1: Performance of sequencing mechanisms in silico



Note: Items are expected order statistics drawn from a half-normal distribution with $\theta = 1000$, so that drawing only one item is expected to yield a value of 1000.

a sister sequence, although they do not identify it as such. (see table on page 3 in Cooper and Dutle (2013) paper on Galois games).

The dynamic in figure 1 counterintuitively suggests that when items are distributed half-normally, TM sequence always performs better than the modified draft, regardless of the number of items. Palacios-Huerta suggested that there should be 2^{n+1} elements in the sequence for the potential of TM sequence to be fully realized. In his paper, he uses the example of soccer penalties, where by the same logic the winner should be the best of $2^3 = 8$ instead of the current convention of 10 to utilize the full allocative potential of balanced alternation. However, TM sequence is at least not worse for all sample sizes for the half-normal and uniform cases and there is no reason to assume that items distribution for the next elements will favor modified draft. In the particular example of half-normal penalty shoot-outs, Thue-Morse sequencing is close to ideal distribution at sample size of 10, but yields almost no advantage over modified draft at sample size of 8. As shown above it is also dominating in exponential and uniform cases regardless of n, and the discrepancy is zero in expectation with uniform F and $n = 8$ (see Appendix).

The efficiency can be significantly improved, if the F distribution is known, by applying greedy solutions for the knapsack problem in place of the balanced alternation. Indeed, approximating a solution to the subset-sum problem with values equal to order statistics from the half-normal distribution would produce a sequence dominating balanced alternation. This is illustrated in figure 1. Consecutively assigning an item to the player who has smaller aggregated sum so far, yields significant shrinking of discrepancy.

There is laboratory experimental evidence that complexity restrictions may prevent people in real life from using Thue-Morse sequence as focal since it cannot be represented by a finite automaton. In these cases people tend to use simple alternation mechanics and other rotation schemes (see Kuzmics et al., 2012). However when complexity is not binding, balanced alternation and greedy algorithms are always present as a better option.

5 Conclusion and Discussion

Classic results for partitioning problem rest on the assumption that the values of all items are measurable and known at least to agents themselves. Brams and Taylor suggest Thue-Morse sequence (balanced alternation) as an allocation mechanism for that case. Palacios-Huerta addresses a similar problem in tournaments. But, if it was possible to put a numerical value on every item in the list, the problem turns into a knapsack problem, which is NP-complete but can be solved for small n through either exhaustive search or heuristics. Then the allocation can be improved by actually solving the knapsack (or subset-sum optimization) problem for each particular case, instead of creating a sequence of turns ex ante.

Consider instead an alternative scenario where the actual value of items is stochastic and unknown. Problem can be set up in either one of two equivalent forms - in the first setting parties can only rank items, but not measure them (for instance, skill of players). Alternatively, parties may know the actual value of the items, but the procedure is designed ex ante, before the items are drawn from the known distribution. A lot of settings for sequencing and allocation problems would benefit from expanding their scope to these assumptions and a particular distribution F. For example, in a "backyard ball game" setting, it is more likely that players are distributed normally, exponentially or half-normally than uniformly. Likewise, in tournaments strategic and psychological first-mover advantages are likely to be significant in the first few rounds and negligible towards the end of the game.

We showed that these problems and a few others are identical to partitioning problem, given particular stochastic characteristics of the items. The solution largely depends on our assumptions about distribution of item values (or player skill). In the trivial case when value is distributed uniformly most sequencing mechanics yield same results, and this case yields no further intuition. Indeed in the uniform case switching order of players between 1st and 2nd rounds and between 2nd and 3rd rounds yields same expected gain in probability of winning or worth of player skill. This makes sequencing trivial as non-inferior mechanisms yield similar expected outcomes. It seems advantageous and more adequate to model players and items as distributed half-normally.

In the half-normally distributed case balanced alternation is always better performing than modified draft. Therefore when uncertainty is present and exhaustive search is not viable, it is likely that balanced alternation is indeed a well performing allocation mechanism, or at least better than modified draft or simple alternation. It is however improvable for any given distribution, if the greedy heuristic is applied to expected order statistics to produce the sequence of turns.

This result also applies to tournament sequencing and design of experimental trials. If there is a first-mover advantage that cannot be measured and it is growing increasingly smaller with each next round of the tournament, TM sequencing of order of move performs better than other sequences regardless of the number of rounds in the tournament.

I argue that the sequence of advantages itself may not be known, however, unless we make assumptions about its general form (arithmetic or geometric progression, rapidly decreasing advantage etc), the problem is not well defined - optimal sequencing would differ for different sets of advantages.

The objective here is not ex-ante efficiency as it often is in welfare economics, but ex-post efficiency. Indeed when devising a sequence of turns in a tournament, one is concerned with maximizing the revealed ability of the players and minimizing the noise from random luck. In more concise terms, the designer is minimizing the entropy given known players' skill. If, for example, first-mover advantage is extremely high (e.g., in tic-tac-toe game), the game is still "fair" up to the moment when the randomization mechanism picks the first-mover. However the share of the probability of winning that is determined by chance is very high - the outcome is determined by the outcome of this randomization, not by the game itself, which is what we are presumably interested in. In other words, the mechanism is ex post inefficient. Same view applies more generally - when studying a phenomenon that has sequential nature, it is beneficial to minimize random effects caused by sequencing. Suppose an analyst is looking into performance of a particular machine, that can operate in two modes: A and B. Analyst suspects that consecutive trials may be correlated due to residue in the machine. Then the sequencing of trials with modes A and B should not be such that the expectations are ex-ante equal - this can easily be achieved through randomization. The goal is instead to bring the discrepancy due to residue to a minimum by making the sequencing ex post efficient - bringing expected residue in mode A and in mode B as close to each other as possible.

The final note to be made is that with more than two parties, the results relating to Thue-Morse sequence will not immediately hold. First reason is that third agent drives a wedge between envy-freeness and proportionality. Another reason is that a three-letter analogue of Thue-Morse sequence[10] does not fully retain all of the original properties. It should still, however, be possible to solve these problems with the greedy heuristic approach.

[10]Sequence number A053838 in the On-Line Encyclopedia of Integer Sequences (OEIS).

References

Allouche, J.-P., and J. Shallit. The Ubiquitous Prouhet-Thue-Morse Sequence, in Sequences and Their Applications: Proceedings of SETA 98, edited by C. Ding, T. Helleseth, and H. Niederreiter. London: Springer-Verlag, 1999, 1-16.

Allouche, J.-P. Forthcoming. Thue, Combinatorics on Words, and Conjectures Inspired by the Thue-Morse Sequence. Journal de Thorie des Nombres de Bordeaux.

Berlekamp, E., J. Conway, and R. Guy. Winning Ways for Your Mathematical Plays, vol. 1,2 and 3. London: Academic Press, 1982.

Brams S. J., and A. D. Taylor. An Envy-Free Cake Division Protocol. The American Mathematical Monthly, 102(1), 1995, 9-18.

Brams, S. J., and A. D. Taylor. The Win-Win Solution: Guaranteeing Fair Shares to Everybody. New York: Norton, 1999.

Brams, S. J., D. M. Kilgour, and C. Klamler. "Two-Person Fair Division Of Indivisible Items: An Efficient, Envy-Free Algorithm." Notices Amer. Math. Soc., 61 (02), 2014, 130-141.

Cooper J. N., and A. Dutle. Greedy Galois Games. The American Mathematical Monthly, 120(5), 2013, 441-451.

Dantzig G. B. Discrete-Variable Extremum Problem. Operations Research, 5(2), 1957, 266-277.

Kuzmics C., T. Palfrey, and B. W. Rogers. Symmetric play in repeated allocation games. Working Paper 468, Institute of Mathematical Economics, 2012, 1-42.

Morse, M. Recurrent Geodesics on a Surface of Negative Curvature. Transactions of the American Mathematical Society, 22, 1921, 84-100.

Olgun J., and T. Fearn. A New Look at Half-normal Plots for Assessing the Significance of Contrasts for Unreplicated Factorials. Journal of the Royal Statistical Society: Series C (Applied Statistics), 46(4), 1997, 449-462.

Palacios-Huerta, I. Tournaments, Fairness and the Prouhet-Thue-Morse Sequence. Economic Inquiry, 50(3), 2012, 848-849.

Ruffle B.J., and O. Volij. First-Mover Advantage in Best-Of Series: An Experimental Comparison of Role-Assignment Rules. Working Paper, Wilfrid Laurier University and Ben-Gurion University of the Negev, 2014.

Thue A. Uber die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. 1912. Reprinted in Selected Mathematical Papers of Axel Thue, edited by Trygve Nagell. Oslo: Universitetsforlaget, 1977, 413-478.

Appendices

A Proof of Proposition 1

Note that order statistics of the uniform distribution follow the beta distribution and the expected value of the i-th order statistic of uniform distribution on the unit interval is

$$E(Y_i) = i/(i + (n - i + 1) = i/(n + 1), \tag{6}$$

where n is the size of the sample. We can multiply each element by $n + 1$, without changing its allocative performance. This will result in a sequence of integers from 1 to n. So in the stochastic uniform case, the task simplifies to equally partitioning a set of integers from 1 to n into two sets.

Substituting expected values from the uniform distribution for all four consecutive elements in lemma 4, the difference between modified draft and TM sequence takes the form of $(-1) \times n + (1) \times (n+1) + (1) \times (n+2) + (-1) \times (n+3) = 0$. ■

Thus it does not matter which method to choose - balanced alternation (TM), modified draft (ABBAA) or some combination - as long as it is not strict alternation, expectation of discrepancy is the same. It is also zero for all sample sizes that allow for perfect partitioning ($n = 3, 4, 7, 8, 11, 12, 15, 16, 19, 20.$ etc[11]).

[11]Sequence of numbers congruent to 0 or 3 mod 4, sequence number A014601 in OEIS with zero dropped. See properties cited in OEIS for proof that $\pm 1 \pm 2 \pm m = 0$ for some choices of signs, where m is any nonzero number from the sequence.