Combinations of Tournament Brackets: An Interesting Application of Combinatorics

Leixin Gong and Michael Boris John Hall

May 2023

1 Problem Statement

We are given n teams and we wish to find out the number of unique arrangements there are of the n teams as a tournament bracket, given a fixed shape (shape will be discussed more in detail in section 4). Throughout this paper, we only consider simple single-elimination tournaments.

To show two teams playing against each other, we will use $\sim$ as the operator. I am aware that $\sim$ usually represents an equivalence relationship, but here we will use it to show two teams playing against each other.

e.g. $A \sim B$ represents A playing B

Note that if we have just two teams (team A and team B), team A playing B is identical to B playing A. i.e. we must not double count $A \sim B$ and $B \sim A$.

$$A \sim B \equiv B \sim A \tag{1}$$

Funny enough, the symmetric property of the equivalence relationship is actually kept here.

We are going to define a tournament structure as a collection of teams playing each other, but each team is represented by $X_1, X_2, \ldots, X_n$ instead of by letters. The standard letters A, B, C, D, ... will represent an actual team in a particular instance of a tournament, i.e. one actual arrangement. The size of the team (the number of teams in the tournament/structure) can be denoted with the modulus operator, e.g. $|S|$, for some S.

To show a structure consisting of 4 teams playing against each other in a standard arrangement, i.e. two semi-finals leading to one final, we can represent one semi-final as $X_1 \sim X_2$ and the other as $X_3 \sim X_4$. The way we will show this structure will be as follows:

$$(X_1 \sim X_2) \sim (X_3 \sim X_4) \tag{2}$$

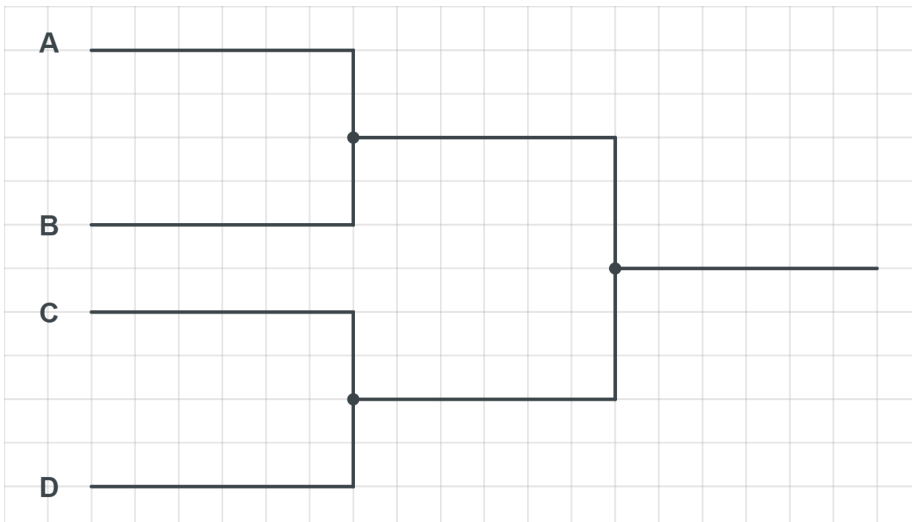

Figure 1: An instance of the standard structure of 4 teams (Θ_4) in a tournament bracket as shown in (2), with $X_1 = A$, $X_2 = B$, $X_3 = C$ and $X_4 = D$

Formally, we want to define a function, $f(S)$, such that the output of the function is the number of different arrangements that can be made of $|S|$ such teams, given a fixed structure S - note that different structures may result in different amounts of possibilities.

Earlier, we defined two teams playing against each other as $X_1 \sim X_2$. Here, we give a recursive definition of an arrangement and a structure:

- We will define a *Body* to be either a *Team* or one *Body* playing against another *Body*, we call this a *Game*
- Let us give a *Body* the symbol B, and a *Team* the symbol T
- So, either $B = T$ if B is a *Team*, or $B = B^L \sim B^R$ if B is a *Game*, where B^L and B^R are both a (possibly different) *Body*
- Given two *Body*: B_1 and B_2, we say that $B_1 \equiv B_2$ if and only if one of the following conditions hold:
    - $B_1 \equiv T$ and $B_2 \equiv T$ (they both represent the same *Team*)

  - $B_1^L \equiv B_2^L$ and $B_1^R \equiv B_2^R$ (they are both a *Game*, the two left *Body* are identical and the two right *Body* are identical)
  - $B_1^L \equiv B_2^R$ and $B_1^R \equiv B_2^L$ (they are both a *Game*, the left *Body* of one is identical to the right *Body* of the other and vice versa)

- Note that if B_1 is a $Team$ and not a $Game$ then B_1^L and B_1^R do not exist so the second and third conditions for being identical should not be considered (similarly this is true for B_2)

- An arrangement is a *Body* where every $Team$ is identical to each other. The teams are represented using $X_1, X_2, \ldots, X_n$ and $X_i \equiv X_j$ for all $1 \leq i, j \leq n$

- A structure is a *Body* where all $Team$ are not identical, pairwise. The teams are commonly represented using $A, B, C, \ldots$ and $A \not\equiv B \not\equiv C \not\equiv \ldots$

The standard structure of size n can be represented by Θ_n. The standard structure will minimise the number of layers and ensure each layer is completely full before filling the next layer. This resembles a complete binary tree for those familiar with graph theory. To summarise:

- S represents a structure, a way of organising a single-elimination tournament
- An arrangement is an instance of a structure when a set of teams is substituted into it, for example, the one shown in Figure 1
- $n = |S|$ represents the size of the structure; the number of teams
- $f(S)$ is the number of unique arrangements in a structure S
- Θ_n is the standard structure, where each layer is as full as possible
- $f(\Theta_n)$ is the number of unique arrangements in the standard structure Θ_n

2 Understanding with a Simple Example

A simple exercise you can try would be to work out $f(\Theta_4)$, where Θ_4 is the standard structure of a tournament of 4 teams. If you have been eagle-eyed, you would have spotted that Figure 1 is an instance of this structure. You could try to solve this by brute force, or otherwise. The solution to this problem is listed just below so stop reading if you want to give this problem a try.

In plain English, we ask the reader to try to find the number of ways of arranging a tournament where there are two semi-finals and one final round (with 4 teams in total). If you are confused and want to see the tournament structure, refer to Figure 1.

Common mistakes when answering this question occur when it is rushed. Such examples are:

- 24 - this answer is very common and is obtained by counting the number of orderings of X without considering the fact that many duplicate arrangements arise since $X_1 \sim X_2 \equiv X_2 \sim X_1$. For example, the arrangement $(A \sim B) \sim (C \sim D)$ and $(B \sim A) \sim (C \sim D)$ would not be accounted as duplicates. In this case, no effort has been made to consider the symmetric quality of games. The calculation which leads to this is $4!$.
- 6 - this answer is also extremely common, but however a good step in the right direction. Usually, this answer comes from the understanding that you can swap the two teams with each other in the first semi-final and the same with the second. However, the finals cause an issue: the arrangement $(A \sim B) \sim (C \sim D)$ and $(C \sim D) \sim (A \sim B)$ would not be accounted as duplicates, which results in double-counting the answer. This calculation that leads to this is $\frac{4!}{2^2}$.

The answer comes from the consideration of symmetrical arrangements of teams both in the semi-finals and in the finals. Below, all 24 (not necessarily unique) arrangements are listed. Arrangements equivalent to each other are coloured in the same colour:

$(A \sim B) \sim (C \sim D)$	$(A \sim C) \sim (B \sim D)$	$(A \sim D) \sim (B \sim C)$
$(A \sim B) \sim (D \sim C)$	$(A \sim C) \sim (D \sim B)$	$(A \sim D) \sim (C \sim B)$
$(B \sim A) \sim (C \sim D)$	$(B \sim D) \sim (A \sim C)$	$(B \sim C) \sim (A \sim D)$
$(B \sim A) \sim (D \sim C)$	$(B \sim D) \sim (C \sim A)$	$(B \sim C) \sim (D \sim A)$
$(C \sim D) \sim (A \sim B)$	$(C \sim A) \sim (B \sim D)$	$(C \sim B) \sim (A \sim D)$
$(C \sim D) \sim (B \sim A)$	$(C \sim A) \sim (D \sim B)$	$(C \sim B) \sim (D \sim A)$
$(D \sim C) \sim (A \sim B)$	$(D \sim B) \sim (A \sim C)$	$(D \sim A) \sim (B \sim C)$
$(D \sim C) \sim (B \sim A)$	$(D \sim B) \sim (C \sim A)$	$(D \sim A) \sim (C \sim B)$

In the end, we can see that we actually only have 3 unique solutions, being:

- $(A \sim B) \sim (C \sim D)$
- $(A \sim C) \sim (B \sim D)$
- $(A \sim D) \sim (B \sim C)$

At this stage, brute force is a perfectly good way of solving this problem, as it often avoids common and easy-to-make mistakes, as discussed above. However, there is a nicer way of solving this that does not require iterating through all permutations (not necessarily unique) and then removing arrangements that are equivalent.

Another way to brute force this could be to just write down as many unique structures as you can (essentially trying arrangements in your head and then comparing them to existing ones, but this often leads to over or under-counting as you can simply forget a possibility or not notice that two arrangements are identical. This makes this solution unreliable as many errors can occur and is not at all systematic. Whilst we only have four teams here, hopefully, it should be easy to see that when the number of teams increases, this method becomes less and less viable. I'm also sure that as mathematicians, we would think to find a more mathematical solution that is not a brute force.

Instead of trying brute force, we can try to utilise elementary combinatorics. In this method, we calculate the number of ways we can arrange 4 teams into the structure, ignoring duplicates, which gives us $4! = 24$. Then, we can divide this value by the number of symmetries. Given the structure, $(X_1 \sim X_2) \sim (X_3 \sim X_4)$, we can see the symmetries between:

- X_1 and X_2
- X_3 and X_4
- $X_1 \sim X_2$ and $X_3 \sim X_4$

From this, we can deduce that there are 3 symmetries. For each symmetry, we need to divide the naïve number of arrangements by 2, since in this structure each symmetry generates two equal-sized sets of equivalent arrangements (note that the amount of games played does not necessarily equal the number of symmetries - we can see this later). This results in the overall calculation: $\frac{4!}{2^3} = 3$, which yields us the correct answer!

There is another intuitive solution to this problem that doesn't use very much maths that I thought of. This solution doesn't work well/at all for any other structure but is a shortcut nevertheless. Consider you are one of the teams. From here, we can only play 3 other teams. However, picking the other team and you being you leaves only 2 more teams. We can also see that there are only two more vacant 'slots' to fill and so there is no choice about the other game. This means that only the team you play matters, therefore, there are only 3 possible arrangements for this structure.

Some may be thinking already (perhaps not too enthusiastically) about bigger cases, for example, we may have 8 teams in their standard structure, or even further thinking about a general solution for the problem, when we have 2^n teams, $n \in \mathbb{N}$ and have its standard structure.

3 Solution for $f(\Theta_n)$, where $n = 2^k$, $k \in \mathbb{N}$

Here we will build a recursive formula for $f(\Theta_n)$. First, let's consider a couple of small cases:

- When $n = 1$, $f(\Theta_n) = 1$ trivially since there are no games
- When $n = 2$, $f(\Theta_n) = 1$ trivially since there is only one game, in which the only two teams are forced to play each other
- When $n = 4$, $f(\Theta_n) = 3$ which we brute-forced earlier on
- When $n = 8$, $f(\Theta_n) = 315$. We get this from choosing 4 teams out of the eight to play in one of the brackets of size 4. The other 4 teams are forced to play in the other bracket of size 4. Within each bracket of size 4, the 4 teams can play each other in 3 ways. However, we count each arrangement twice since the two brackets of size 4 have an identical structure (Θ_4). $f(\Theta_n) = \binom{8}{4} * 3 * \binom{4}{4} * 3 \div 2 = 315$

The last case is fascinating since it gives us an insight into how to get a recursive formula for this case. Let us talk more generally; with n teams, we choose $\frac{n}{2}$ of them go into one of the brackets of size $\frac{n}{2}$. The other $\frac{n}{2}$ are forced into the other bracket of size $\frac{n}{2}$. Within each bracket of size $\frac{n}{2}$ there are $f(\Theta_{\frac{n}{2}})$ ways of arranging the teams. There is a double counting of each arrangement since both brackets of size $\frac{n}{2}$ have identical structures. All in all, we find that:

$$f(\Theta_n) = \frac{\binom{n}{\frac{n}{2}} f(\Theta_{\frac{n}{2}})^2}{2} \tag{3}$$

This is a very unsightly result. We will attempt to rewrite this formula in a more succinct way and remove the recursive aspect of it.

$$
\begin{aligned}
f(\Theta_n) &= \frac{\binom{n}{\frac{n}{2}} f(\Theta_{\frac{n}{2}})^2}{2} \\
&= \frac{\binom{n}{\frac{n}{2}} (\frac{\binom{\frac{n}{2}}{\frac{n}{4}} f(\Theta_{\frac{n}{4}})^2}{2})^2}{2} \\
&= \frac{\binom{n}{\frac{n}{2}} (\frac{\binom{\frac{n}{2}}{\frac{n}{4}} (\frac{\dots}{2})^2}{2})^2}{2} \\
&= \frac{\binom{n}{\frac{n}{2}} \binom{\frac{n}{2}}{\frac{n}{4}}^2 \binom{\frac{n}{4}}{\frac{n}{8}}^4 \dots \binom{2}{1}^{\frac{n}{2}}}{2^{1+2+4+\cdots+\frac{n}{2}}} \\
&= \frac{(\frac{n!}{\frac{n}{2}!\frac{n}{2}!})(\frac{\frac{n}{2}!}{\frac{n}{4}!\frac{n}{4}!})^2(\frac{\frac{n}{4}!}{\frac{n}{8}!\frac{n}{8}!})^4 \dots \binom{2}{1}^{\frac{n}{2}}}{2^{1+2+4+\cdots+\frac{n}{2}}} \\
&= \frac{\frac{n!}{(\frac{n}{n}!)^{\frac{n}{2}}}}{2^{1+2+4+\cdots+\frac{n}{2}}} \\
&= \frac{n!}{2^{1+2+4+\cdots+\frac{n}{2}}} \\
&= \frac{n!}{2^{n-1}}
\end{aligned}
\tag{4}
$$

This is not an especially formal proof, however, all the steps follow along. It is also an obtainable and reasonable answer to get. It's much cleaner than the solution for any other structures. To even have a non-recursive formula - with just elementary maths - is impressive. This also tells us how quickly the numbers blow up. It would be unreasonable to try and brute-force the result for any structure whose size is above 4 (by hand), or 16 even with a computer. For fun, we supply a table of values below:

n	$f(\Theta_n)$
1	1
2	1
4	3
8	315
16	638512875
32	122529844256906551386796875
64	$\approx 1.3757 \cdot 10^{70}$
128	$\approx 2.2665 \cdot 10^{177}$
256	$\approx 1.4817 \cdot 10^{430}$
512	$\approx 5.1870 \cdot 10^{1012}$
1024	$\approx 6.0283 \cdot 10^{2331}$
2048	$\approx 1.0352 \cdot 10^{5278}$
4096	$\approx 6.9758 \cdot 10^{11786}$

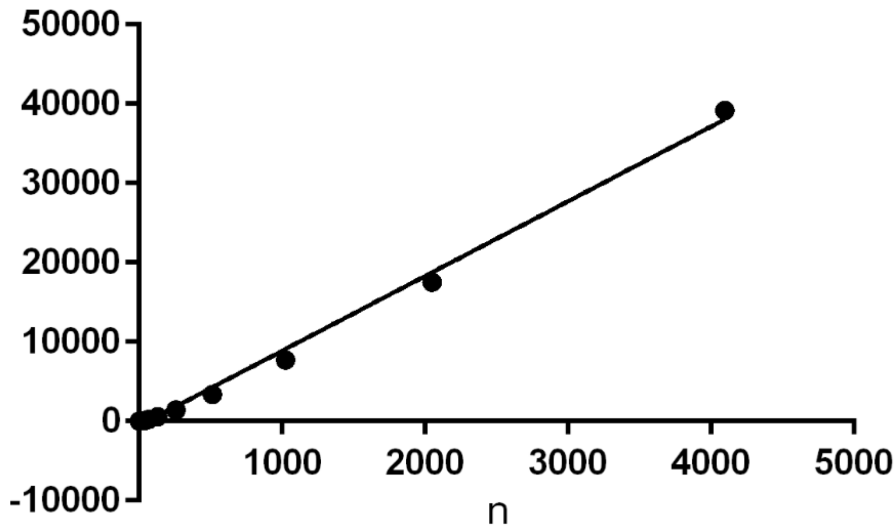

Figure 2: A graph of $\log_2\big(f(\Theta_n)\big)$ against n, for the values in the table. The r^2 value is 0.9955 for this data.

From this, we can deduce with relative confidence that whatever high school sports tournament will never be repeated. These numbers were so big we had to change the system number limit in Python (twice!) to calculate the last few... - the time needed to calculate was also noticeable (but not horrible).

A more intuitive way of understanding the formula can be obtained by reading ahead. In the next section, I talk about calculating the number of arrangements by using the property of symmetry.

As we are working with a very particular structure here, any game and the games leading to it are all symmetrical structure-wise. You can convince yourself by just picturing a tournament, with say, 8 teams, and then flipping games around in your head. You will notice that any game you choose to flip around will not at all change the structure of the tournament.

We can calculate the number of arrangements by taking into account all tournaments assuming no symmetry - effectively how many ways can we order n teams - which is $n!$, then dividing by two for each symmetry we have. As we have previously stated that any and every game in this special structure is symmetrical, the amount of symmetries is just equal to the number of games played in a tournament.

To figure out the number of games played in a tournament, we can consider that the tournament starts with n teams, and ends with 1 winner, which implies that $n-1$ teams were knocked out. Since every game eliminates exactly 1 player, there must be $n-1$ games.

Therefore, there are always $n-1$ symmetries in a tournament with this specific structure - due to the fact that every game is symmetrical and there are $n-1$ games.

This leaves us with our previously-proved formula:

$$f(\Theta_n) = \frac{n!}{2^{n-1}} \tag{5}$$

4 Solution for $f(\Theta_n)$, for all $n \in \mathbb{N}$

The solution for $f(\Theta_n)$, where $n = 2^k$, $k \in \mathbb{N}$ is quite nice and is composed of purely mathematical functions, with a clear, defined and provable formula. However, this obviously isn't always the case.

For example, let's take the example of $f(\Theta_6)$. In our notation, this tournament can be represented through:

$$\big((X_1 \sim X_2) \sim (X_3 \sim X_4)\big) \sim (X_5 \sim X_6)) \tag{6}$$

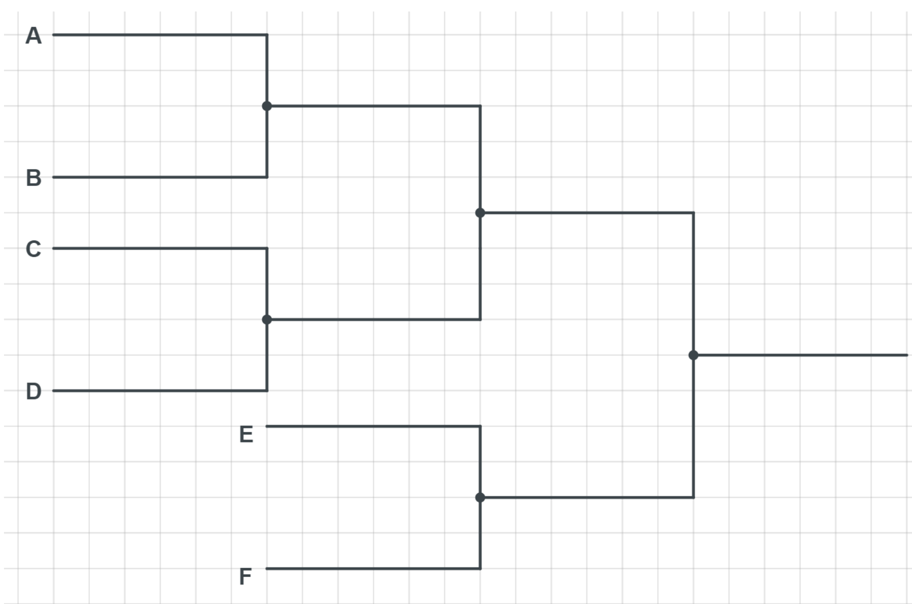

Figure 3: An instance of the standard structure of 6 teams (Θ_6) in a tournament bracket, with $X_1 = A$, $X_2 = B$, $X_3 = C$, $X_4 = D$, $X_5 = E$ and $X_6 = F$

Note that a team starting in the semi-finals is not equivalent to a team starting in the quarter-finals. Knowing this, solve $f(\Theta_6)$.

Again, in simple English, we ask the reader to try to find the number of ways of arranging a tournament, with 6 teams and where there are two semi-finals and one final round, with one of the semi-finals coming from two quarter-finals. If you are confused and want to see the tournament structure, refer to Figure 3. Again the answer will be discussed just below this, so please stop reading to first give this exercise a go.

As before, I will go over some incorrect solutions to the problem:

- 720 - this solution is clearly obtained by taking 6!. As discussed before, this completes ignores the symmetric property of games. If this answer was obtained, I highly suggest you go back to read section 2 thoroughly again.
- 15 - this solution may be obtained if you ignore the warning given just before the question, where I said starting in the semi-finals is not equivalent to a team starting in the quarter-finals. This may seem weird, but I have gotten this solution twice from two different clever sources so I thought I might address it. If you considered three equal games and then divided them by how many ways you can arrange them, you may get 15. This is from $\binom{6}{2}\binom{4}{2}\binom{2}{2} \div 3!$
- 90 - this solution is very nearly there and rather impressive! Whilst not correct, it is very close to the answer. 90 is actually exactly double the answer. If you got this number, please reflect briefly on where the double-counting may have happened and then read on for the full explanation.

The most intuitive solution is reached by first spotting that if we have already picked 2 teams (or in other words, have only 4 leftover teams), the semi-final consisting of two quarter-finals is identical to our structure in $f(\Theta_4)$, which had the solution 3. This means we only now have to calculate how many ways can we choose 2 teams from 6 teams to form the other semi-final. With basic combinatorics, we know this is $\binom{6}{2}$, which is 15. Then we can multiply 15 by the 3 from before to yield 45, which is the correct answer.

For those who got 90, instead of using $\binom{6}{2}$, you may have considered choosing 1 out of the 6 teams, then choosing another 1 from the remaining 5, to get 30 arrangements on the left-hand side. This is where the double-counting occurs. Due to the symmetric property of games, the left-hand side had been double-counted, resulting in 90 rather than 45.

5 General Solution

Here, we will formally describe the notion of symmetry, and identical structures and give a recursive general formula for the number of unique arrangements of any structure S of any size $n \in \mathbb{N}$.

We will define a structure as 'symmetrical' if and only if its two substructures are identical, for this allows the swapping of teams between the two substructures. This induces a division by two in the final result.

Two structures S and T are identical if and only if $S \equiv T$. Recall that a structure is a *Body*, whose teams are all identical.

Let S^L and S^R be the substructures of structure S. Define $\sigma(S)$ to be a symmetry checker: $\sigma(S) = 1 \iff S^L \equiv S^R$ and 0 otherwise, or if S^L / S^R does not exist. Then:

$$\begin{aligned} f(S) &= 1 \iff |S| = 1 \\ f(S) &= \frac{\binom{|S|}{|S^L|} f(S^L) f(S^R)}{2^{\sigma(S)}} \iff |S| > 1 \end{aligned} \tag{7}$$

Note that $\binom{|S|}{|S^L|}$ can equally be $\binom{|S|}{|S^R|}$. This is because $|S^L| + |S^R| = |S|$, so the symmetric property of Pascal's triangle holds.

Essentially to calculate $f(S)$, we evaluate the number of ways we can choose teams for one side (either S^L or S^R - and as previously stated they are the same as $\binom{n}{k} \equiv \binom{n}{n-k}$). The rest of the players fill the other side effectively. Then, we multiply this by the number of ways of arranging the players on each side. Finally, we divide by two if the structure is symmetrical about this game, i.e. $S^L \equiv S^R$.

An alternative way to think about this would be to find the number of ways of arranging these n teams, where there is no symmetry, and then divide by two for every symmetry that exists. We need to define another function for this next formula:

$$\Sigma(S) = \Sigma(S^L) + \Sigma(S^R) + \sigma(S) \tag{8}$$

So we could also rewrite this formula as:

$$f(S) = \frac{|S|!}{2^{\Sigma(S)}} \tag{9}$$

Recall that $\Sigma(\Theta_n) = n - 1 \iff n = 2^k, k \in \mathbb{N}$. It's cool how the same formula from before manages to pop up!

6 Finishing Remarks

We have answered the question for a given structure S. Now, as a challenge to the reader, I suggest the following question: how many ways are there of creating a tournament of size n, using *any* structure of size n? I leave this as an exercise for the reader.

If you are able to get any advance on this, I would personally be very happy to hear about it, because this question really intrigued me.

Feel free to contact us at **gonglx8@gmail.com** and **mborishall@gmail.com** for anything at all.