Contents

1 Tournament Formats

1.1 Definitions

Definition 1.1.1: Gameplay Function

A *gameplay function* g on a list of teams $\mathcal{T} = [t_1, ..., t_n]$ is a nondeterministic function $g : \mathcal{T} \times \mathcal{T} \to \mathcal{T}$ with the following properties:

- $\mathbb{P}[g(t_i, t_j) = t_i] + \mathbb{P}[g(t_i, t_j) = t_j] = 1$.

- $\mathbb{P}[g(t_i, t_j) = t_i] = \mathbb{P}[g(t_j, t_i) = t_j]$.

A gameplay function represents a process in which two teams compete in a game, with one of them emerging as the winner. This model simplifies away effects like home-field advantage or teams improving over the course of a tournament: a gameplay function is fully described by a single probability for each pair of teams in the list.

Definition 1.1.2: Playing, Winning, and Losing

When g is queried on input (t_i, t_j) we say that t_i and t_j *played a game*. We say that the team that got outputted by g *won*, and the team that did not *lost*.

The information in a gameplay function can be encoded into a *matchup table*.

Definition 1.1.3: Matchup Table

The *matchup table* implied by a gameplay function g on a list of teams $\mathcal{T}$ of length n is a n-by-n matrix $\mathbf{M}$ such that $\mathbf{M}_{ij} = \mathbb{P}[t_i \text{ beats } t_j]$.

For example, let $\mathcal{T} =$[Favorites, Rock, Paper, Scissors, Conceders], and g be such that the Conceders concede every game they play, the Favorites are 70% favorites against Rock, Paper, and Scissors, and Rock, Paper, and Scissors match up with each other as their name implies. Then the matchup table would look like so:

Figure 1.1.4: The Matchup Table for $(\mathcal{T}, g)$

	Favorites	Rock	Paper	Scissors	Conceders
Favorites	0.5	0.7	0.7	0.7	1.0
Rock	0.3	0.5	0.0	1.0	1.0
Paper	0.3	1.0	0.5	0.0	1.0
Scissors	0.3	0.0	1.0	0.5	1.0
Conceders	0.0	0.0	0.0	0.0	0.5

> **Theorem 1.1.5**
>
> If $\mathbf{M}$ is the matchup table for $(\mathcal{T}, g)$, then $\mathbf{M} + \mathbf{M}^T$ is the matrix of all ones.
>
> --
>
> *Proof.* $(\mathbf{M} + \mathbf{M}^T)_{ij} = \mathbf{M}_{ij} + \mathbf{M}_{ji} = \mathbb{P}[t_i \text{ beats } t_j] + \mathbb{P}[t_j \text{ beats } t_i] = 1$. $\qquad\qquad\square$

> **Definition 1.1.6: Tournament Format**
>
> A *tournament format* is an algorithm that takes as input a list of teams $\mathcal{T}$ and a gameplay function g and outputs a champion $t \in \mathcal{T}$.

We use a gameplay function rather than a matchup table in the definition of a tournament format because a tournament format cannot simply look at the matchup table itself in order to decide which teams are best. Instead, formats query the gameplay function (have teams play games) in order to gather information about the teams. That said, matchup tables will often be useful in our *analysis* of tournament formats.

We also introduce some shorthand to help make notation more concise.

> **Definition 1.1.7: $\mathbb{W}_{\mathcal{A}}(t, \mathcal{T})$**
>
> $\mathbb{W}_{\mathcal{A}}(t, \mathcal{T})$ is the probability that team $t \in \mathcal{T}$ wins tournament format $\mathcal{A}$ when it is run on the list of teams $\mathcal{T}$.

(This chapter will be fleshed out but I'm including the important definitions here for the sake of the next chapter.)
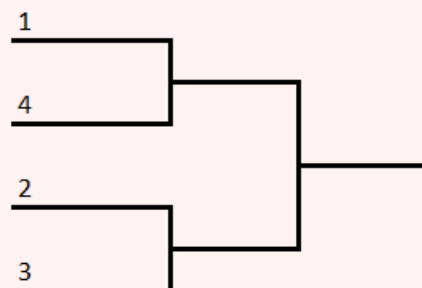
2   Brackets

2.1 Bracket Signatures

> **Definition 2.1.1: Bracket**
>
> A *bracket* is a tournament format in which:
>
> - Teams don't play any games after their first loss,
>
> - Games are played until only one team has no losses, and that team is crowned champion, and
>
> - The matchups between game winners are determined in advance of the outcomes of any games.

We can draw a bracket as a tree-like structure in the following way.

Figure 2.1.2: The 2023 College Football Playoff

The numbers 1, 2, 3, and 4 indicate where t_1, t_2, t_3, and t_4 in $\mathcal{T}$ are placed to start. In the actual 2023 College Football Playoff, the list of teams $\mathcal{T}$ was Georgia, Michigan, TCU, and Ohio State, in that order, so the bracket was filled in like so.

**Figure 2.1.3: The 2023 CFP After Team Placement**

As games are played, we write the name of the winning teams on the corresponding lines. This bracket tells us that Georgia played Ohio State, and Michigan played TCU. Georgia and TCU won their respective games, and then Georgia beat TCU, winning the tournament.

Figure 2.1.4: The 2023 CFP After Completion

Rearranging the way the bracket is pictured, if it doesn't affect any of the matchups, does not create a new bracket. For example, Figure 2.1.5 is just another way to draw the same 2023 CFP Bracket.

Figure 2.1.5: Alternative Drawing of the 2023 CFP
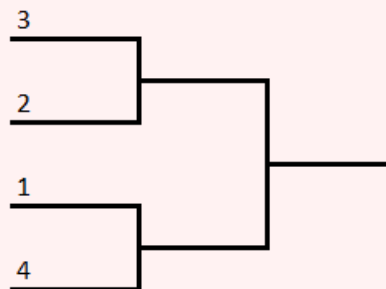
One key piece of bracket vocabulary is the *round*.

Definition 2.1.6: Round

A *round* is a set of games such that the winners of each of those games have the same number of games remaining to win the tournament.

For example, the 2023 CFP has two rounds. The first round included the games Georgia vs Ohio State and Michigan vs TCU, and the second round was just a single game: Georgia vs TCU.

Another important concept is the *shape* of a bracket.

Definition 2.1.7: Shape

The *shape* of a bracket is the tree that underlies it.

For example, the following two brackets have the same shape.

Definition 2.1.9: Bye

A team has a *bye* in round r if it plays no games in round r or before.

One way to describe the shape of a bracket its signature.

Definition 2.1.10: Bracket Signature

The *signature* $[[\mathbf{a_0}; ...; \mathbf{a_r}]]$ of an r-round bracket $\mathcal{A}$ is list such that a_i is the number of teams with i byes.

The signature of a bracket is defined only by its shape: the two brackets in Figure 2.1.8 have the same shape, so they also have the same signature.

The signatures of the brackets discussed in this section are shown in Figure 2.1.11. It's worth verifying the signatures we've seen so far and trying to draw brackets with the signatures we haven't yet before moving on.

Figure 2.1.11: The Signatures of Some Brackets

Bracket	Signature
2023 College Football Playoff	$[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]$
The brackets in Figure 2.1.8	$[[\mathbf{2}; \mathbf{3}; \mathbf{0}; \mathbf{0}]]$
The brackets in Figure 2.1.12	$[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$
2023 WCC Men's Basketball Tournament	$[[\mathbf{4}; \mathbf{2}; \mathbf{2}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$

Two brackets with the same shape must have the same signature, but the converse is not true: two brackets with different shapes can have the same signature. For example, both bracket shapes depicted in Figure 2.1.12 have the signature $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$.
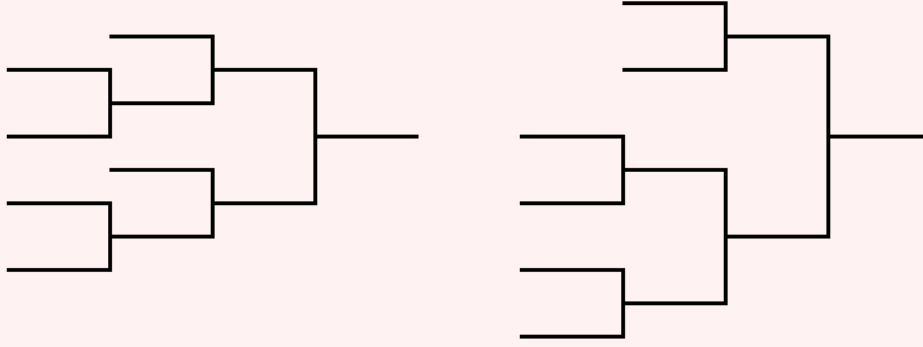
Despite this, bracket signatures are a useful way to talk about the shape of a bracket. Communicating a bracket's signature is a lot easier than communicating its shape, and much of the important information (such as how many games each team must win in order to win the tournament) is contained in the signature.

Bracket signatures have one more important property.

Theorem 2.1.13

Let $\mathcal{A} = [[a_0; ...; a_r]]$ be a list of natural numbers. Then $\mathcal{A}$ is a bracket signature if and only if

$$\sum_{i=0}^{r} a_i \cdot \left(\frac{1}{2}\right)^{r-i} = 1.$$

Proof. Let $\mathcal{A}$ be the signature for some bracket. Assume that every game in the bracket was a coin flip, and consider each team's probability of winning the tournament. A team that has i byes must win $r - i$ games to win the tournament, and so will do so with probability $\left(\frac{1}{2}\right)^{r-i}$. For each $i \in \{0, ..., r\}$, there are a_i teams with i byes, so (because any two teams winning are mutually exclusive)

$$\sum_{i=0}^{r} a_i \cdot \left(\frac{1}{2}\right)^{r-i}$$

is the probability that one of the teams wins, which is 1.

We prove the other direction by induction on r. If $r = 0$, then the only list with the desired property is $[[1]]$, which is the signature for the unique one-team bracket. For

any other r, first note that a_0 must be even: if it were odd, then

$$
\begin{aligned}
\sum_{i=0}^{r} a_i \cdot \left(\frac{1}{2}\right)^{r-i} &= \frac{1}{2^r} \cdot \sum_{i=0}^{r} a_i \cdot 2^i \\
&= \frac{1}{2^r} \cdot \left(a_0 + 2 \sum_{i=1}^{r} a_i \cdot 2^{i-1}\right) \\
&= k/2^r \qquad\qquad \text{for some odd } k \\
&\neq 1.
\end{aligned}
$$

Now, consider the signature $\mathcal{B} = [[\mathbf{a_1} + \mathbf{a_0}/\mathbf{2}; \mathbf{a_2}; ...; \mathbf{a_r}]]$. By induction, there exists a bracket with signature $\mathcal{B}$. But if we take that bracket and replace $a_0/2$ of the teams with no byes with a game whose winner gets placed on that line, we get a new bracket with signature $\mathcal{A}$. $\qquad\square$

In the next few sections, we will use the language and properties of bracket signatures to describe the brackets that we work with. For now though, let's return to the 2023 College Football Playoff. The bracket used in the 2023 CFP has a special property that not all brackets have: it is *balanced*.

Definition 2.1.14: Balanced Bracket

A *balanced bracket* is a bracket in which none of the teams have byes.

The 2023 West Coast Conference Men's Basketball Tournament, on the other hand, is unbalanced.

Figure 2.1.15: The 2023 WCC Men's Basketball Tournament
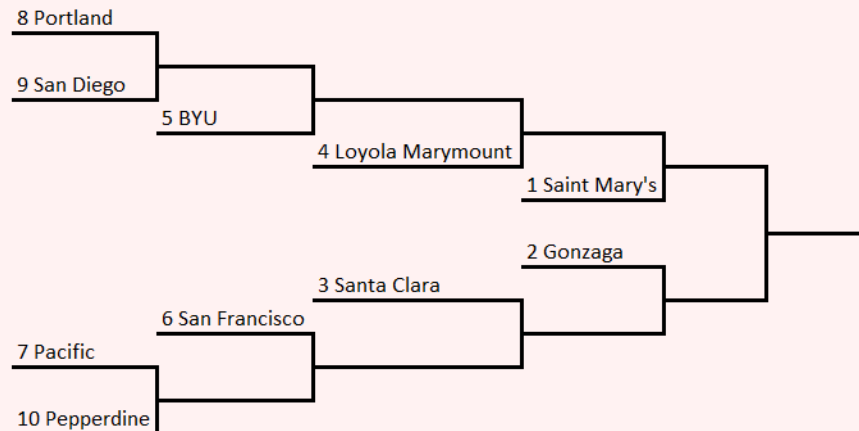
Saint Mary's and Gonzaga each have three byes and so only need to win two games to

win the tournament, while Portland, San Diego, Pacific, and Pepperdine need to win five. Unsurprisingly, this format conveys a massive advantage to Saint Mary's and Gonzaga, but this was intentional: those two teams were being rewarded for doing the best during the regular season.

In many cases, however, it is undesirable to grant advantages to certain teams over others. One might hope, for any n, to be able to construct a balanced bracket for n teams, but unfortunately this is rarely possible.

> **Theorem 2.1.16**
>
> There exists an n-team balanced bracket if and only if n is a power of two.
>
> -
>
> *Proof.* A bracket is balanced if no teams have byes, which is true exactly when its signature is of the form $\mathcal{A} = [[\mathbf{n}; \mathbf{0}; ...; \mathbf{0}]]$ where n is the number of teams in the bracket. If n is a power of two, then by Theorem 2.1.13 $\mathcal{A}$ is indeed a bracket signature and so points to a balanced bracket for n teams. If n is not a power of two, however, then Theorem 2.1.13 tells us that $\mathcal{A}$ is not a bracket signature, and so no balanced brackets exist for n teams. $\qquad\square$

Given this, brackets are not a great option when we want to avoid giving some teams advantages over others unless we have a power of two teams. They are a fantastic tool, however, if doling out advantages is the goal, perhaps after some teams did better during the regular season and ought to be rewarded with an easier path in the bracket.

2.2 Proper Brackets

> **Definition 2.2.1: Seeding**
>
> The *seeding* of an n-team bracket is the arrangement of the numbers 1 through n in the bracket.

Together, the shape and seeding fully specify a bracket.

> **Definition 2.2.2: i-seed**
>
> In a list of teams $\mathcal{T} = [t_1, ..., t_n]$, we refer to t_i as the i-seed.

> **Definition 2.2.3: Higher and Lower Seeds**
>
> Somewhat confusingly, convention is that smaller numbers are the *higher seeds*, and greater numbers are the *lower seeds*.

Seeding is typically used to reward better and more deserving teams. As an example, on the left is the eight-team bracket used in the 2015 NBA Eastern Conference Playoffs. At the end of the regular season, the top eight teams in the Eastern Conference were ranked and placed into the bracket as shown on the right.

Figure 2.2.4: 2015 NBA Eastern Conference Playoffs

Despite this bracket being balanced, the higher seeds are still at advantage: they have an easier set of opponents. Compare 1-seed Atlanta, whose first two rounds are versus 8-seed Brooklyn and then (most likely) 4-seed Toronto, versus 7-seed Boston, whose first two rounds are versus 2-seed Cleveland and then (most likely) 3-seed Chicago. Atlanta's schedule is far easier: despite them having the same number of games to win as Boston, Atlanta is expected to play lower seeds in each round than Boston will.

Thus, we've identified two ways in which brackets can convey an advantage onto certain

teams: by giving them more byes, and by giving them easier (expected) opponents. Not every seeding of a bracket does this: for example, consider the following alternative seeding for the 2015 NBA Eastern Conference Playoffs.

Figure 2.2.5: An Alternative Seeding of the 2015 NBA Eastern Conference Playoffs
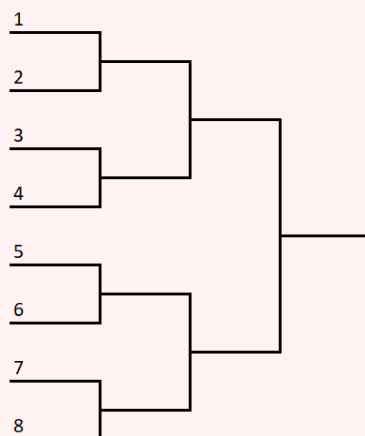
This seeding does a very poor job of rewarding the higher-seeded teams: the 1- and 2-seeds are matched up in the first round, while the easiest road is given to the 7-seed, who plays the 8-seed in the first round and then (most likely) the 5-seed in the second. Since the whole point of seeding is to give the higher-seeded teams an advantage, we introduce the concept of a *proper seeding.*

Definition 2.2.6: Chalk

We say a tournament *went chalk* if the higher-seeded team won every game during the tournament.

Definition 2.2.7: Proper Seeding

A *proper seeding* of a bracket is one such that if the bracket goes chalk, in every round it is better to be a higher-seeded team than a lower-seeded one, where:

(1) It is better to have a bye than to play a game.

(2) It is better to play a lower seed than to play a higher seed.

Definition 2.2.8: Proper Bracket

A *proper bracket* is a bracket that has been properly seeded.

It is clear that the actual 2015 NBA Eastern Conference Playoffs was properly seeded,

while our alternative seeding was not.

We now quickly derive a few lemmas about proper brackets.

Lemma 2.2.9

In a proper bracket, if m teams have a bye in a given round, those teams must be seeds 1 through m.

Proof. If they did not, the seeding would be in violation of condition (1). $\square$

Lemma 2.2.10

If a proper bracket goes chalk, then after each round the m teams remaining will be the top m seeds.

Proof. We will prove the contrapositive. Assume that for some $i < j$, after some round, t_i has been eliminated but t_j is still alive. Let k be the seed of the team that t_i lost to. Because the bracket went chalk, $k < i$. Now consider what t_j did in that round. If they had a bye, then the bracket violates condition (1). Assume instead they played t_ℓ. They beat t_ℓ, so $j < \ell$, giving,

$$k < i < j < \ell.$$

In the round that t_i was eliminated, t_i played t_k, while t_j played t_ℓ, violating condition (2). Thus, the bracket is not proper. $\square$

Lemma 2.2.11

In a proper bracket, if m teams have a bye and k games are being played in a given round, then if the bracket goes chalk those matchups will be seed $m + i$ vs seed $(m + 2k + 1) - i$ for $i \in \{1, ..., k\}$.

Proof. In the given round, there are $m + 2k$ teams remaining. Theorem 2.2.10 tells us that (if the bracket goes chalk) those teams must be seeds 1 through $m + 2k$. Theorem 2.2.9 tells us that seeds 1 through m must have a bye, so the teams playing must be seeds $m + 1$ through $m + 2k$. Then condition (2) tells us that the matchups must be exactly $m + i$ vs seed $(m + 2k + 1) - i$ for $i \in \{1, ..., k\}$. $\square$

We can use Lemmas 2.2.9 through 2.2.11 to properly seed various bracket shapes. For example, consider the following seven-team shape.

Figure 2.2.12: A Seven-Team Bracket Shape

Lemma 2.2.9 tells us that the first-round matchup must be between the 6-seed and the 7-seed. Lemma 2.2.11 tells us that if the bracket goes chalk, the second-round matchups must be 3v6 and 4v5, so the 3-seed play the winner of the first-round matchup. Finally, we can apply Lemma 2.2.11 again to the semifinals to find that the 1-seed should play the winner of the 4v5 matchup, while the 2-seed should play the winner of the 3v(6v7) matchup. In total, our proper seeding looks like so.

Figure 2.2.13: A Seven-Team Bracket, Properly Seeded

We can also quickly simulate the bracket going chalk to verify Lemma 2.2.10.

Lemmas 2.2.9 through 2.2.11 are quite powerful. It is not a coincidence that we managed to specify exactly what a proper seeding of the above bracket must look like with no room for variation: soon we will prove that the proper seeding for a particular bracket shape is unique.

But not every shape admits even this one proper seeding. Consider the following six-team shape.

Figure 2.2.14: A Six-Team Bracket Shape
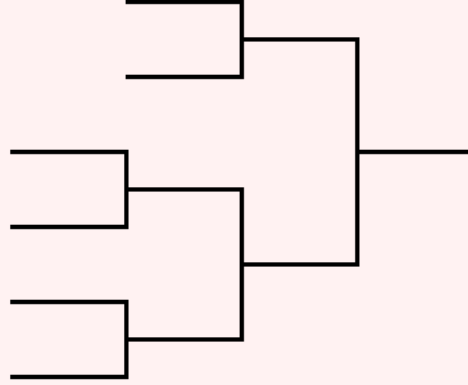
This shape admits no proper seedings. Lemma 2.2.9 requires that the two teams getting byes be the 1- and 2-seed, but this violates Lemma 2.2.11 which requires that in the second round the 1- and 2-seeds do not play each other. So how can we think about which shapes admit proper seedings?

Theorem 2.2.15: The Fundamental Theorem of Brackets

There is exactly one proper bracket with each bracket signature.

Proof. Let $\mathcal{A} = [[\mathbf{a_0}; ...; \mathbf{a_r}]]$ be an r-round bracket signature. We proceed by induction on r. If $r = 0$, then the only possible bracket signature is $[[\mathbf{1}]]$, and it points to the unique one-team bracket, which is indeed proper.

For any other r, the first-round matchups of a proper bracket with signature $\mathcal{A}$ are defined by Lemma 2.2.11. Then if those matchups go chalk, we are left with a proper bracket of signature $[[\mathbf{a_0}/\mathbf{2} + \mathbf{a_1}; \mathbf{a_2}; ...; \mathbf{a_r}]]$, which induction tells us exists admits exactly one proper bracket.

Thus both the first-round matchups and the rest of the bracket are determined, and by combining them we get a proper bracket with signature $\mathcal{A}$, so there is exactly one proper bracket with signature $\mathcal{A}$. □

The fundamental theorem of brackets means that we can refer to the proper bracket $\mathcal{A} = [[\mathbf{a_0}; ...; \mathbf{a_r}]]$ in a well-defined way, as long as

$$\sum_{i=0}^{r} a_i \cdot \left(\frac{1}{2}\right)^{r-i} = 1.$$

In practice, virtually every sports league that uses a traditional bracket uses a proper one: while different leagues take very different approaches to how many byes to give teams

(compare the 2023 West Coast Conference Men's Basketball Tournament with the 2015 NBA Eastern Conference Playoffs), they are almost all proper. This makes bracket signatures a convenient labeling system for the set of brackets that we might reasonably encounter. They also are a powerful tool for specifying new brackets: if you are interested in (say) an eleven-team bracket where four teams get no byes, four teams get one bye, one team gets two byes and two teams get three byes, we can describe the proper bracket with those specs as $[[4; 4; 1; 2; 0; 0]]$ and use Lemmas 2.2.9 through 2.2.11 to draw it with ease.


Figure 2.2.16: The Proper Bracket of Signature $[[4; 4; 1; 2; 0; 0]]$
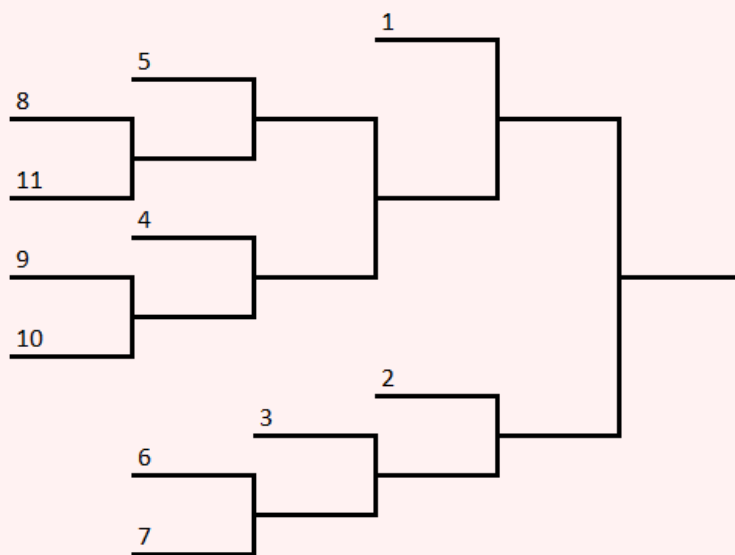
Due to these properties, we will almost exclusively discuss proper bracket from here on out: unless stated otherwise, assume all brackets are proper.

2.3   Ordered Brackets

Consider the proper bracket $[[\mathbf{16}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}]]$, which was used in the 2021 NCAA Men's Basketball Tournament South Region, and is shown below. (Sometimes brackets are drawn in the manner below, with teams starting on both sides and the winner of each side playing in the championship game.)

Figure 2.3.1: 2021 NCAA Men's Basketball Tournament South Region
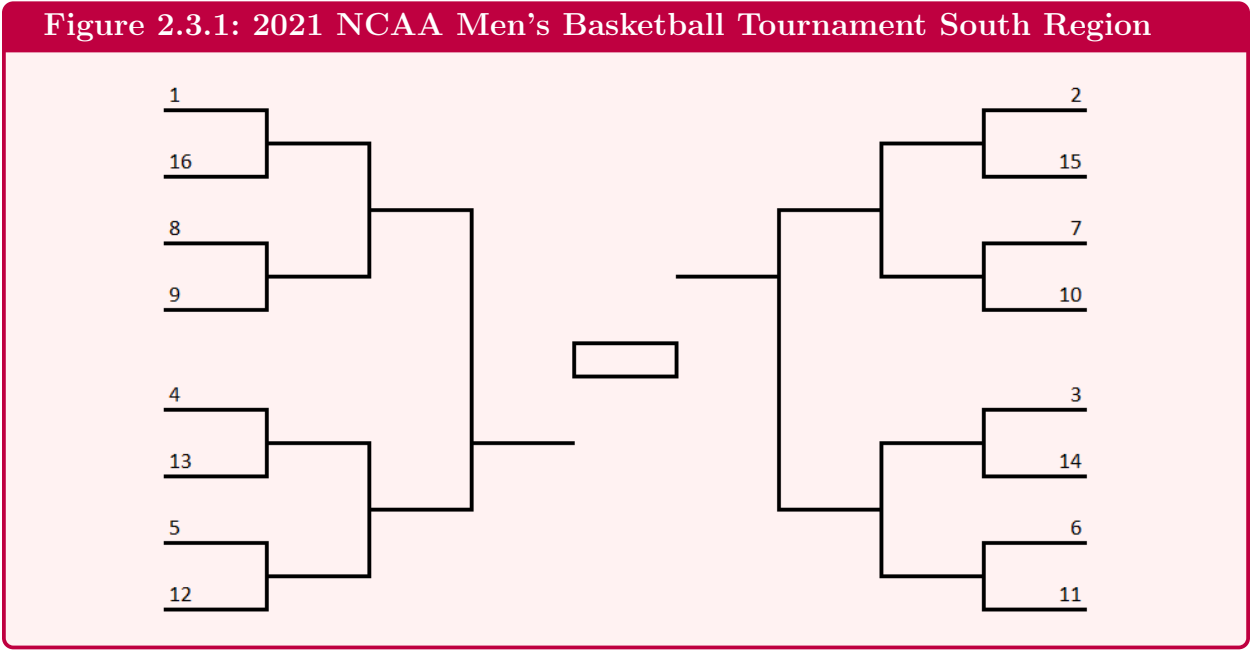
The definition of a proper seeding ensures that as long as the bracket goes chalk (that is, higher seeds always beat lower seeds), it will always be better to be a higher seed than a lower seed. But what if it doesn't go chalk?

One counter-intuitive fact about the NCAA Basketball Tournament is that it is probably better to be a 10-seed than a 9-seed. (This doesn't violate the proper seeding property because 9-seeds have an easier first-round matchup than 10-seeds, and for further rounds, proper seedings only care about what happens if the bracket goes chalk, which would eliminate both the 9-seed and 10-seed in the first round.) Why? Let's look at whom each seed-line matchups against in the first two rounds.

Figure 2.3.2: NCAA Basketball Tournament 9- and 10-seed Schedules

Seed	First Round	Second Round
9	8	1
10	7	2

The 9-seed has an easier first-round matchup, while the 10-seed has an easier second-round matchup. However, this isn't quite symmetrical. Because the teams are probably drawn from a roughly normal distribution, the expected difference in skill between the 1-

and 2-seeds is far greater than the expected difference between the 7- and 8-seeds, implying that the 10-seed does in fact have an easier route than the 9-seed.

Nate Silver [7] investigated this matter in full, finding that in the NCAA Basketball Tournament, seed-lines 10 through 15 give teams better odds of winning the region than seed-lines 8 and 9. Of course this does not mean that the 11-seed (say) has a better chance of winning a given region than the 8-seed does, as the 8-seed is a much better team than the 11-seed. But it does mean that the 8-seed would love to swap places with the 11-seed, and that doing so would increase their odds to win the region.

This is not a great state of affairs: the whole point of seeding is confer an advantage to higher-seeded teams, and the proper bracket $[[\mathbf{16}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}]]$ is failing to do that. Not to mention that giving lower-seeded teams an easier route than higher-seeded ones can incentivize teams to lose during the regular season in order to try to get a lower but more advantageous seed.

To fix this, we need a stronger notion of what makes a bracket effective than properness. The issue with proper seedings is the false assumption that higher-seeded teams will always beat lower-seeded teams. A more nuanced assumption, initially proposed by H.A. David [3], might look like this.

> **Definition 2.3.3: Strongly Stochastically Transitive**
>
> A list of teams $\mathcal{T}$ is *strongly stochastically transitive* if for each i, j, k such that $j < k$,
>
> $$\mathbb{P}[t_i \text{ beats } t_j] \leq \mathbb{P}[t_i \text{ beats } t_k].$$

A list of teams being strongly stochastically transitive (SST) captures the intuition that each team ought to do better against lower-seeded teams than against higher-seeded teams. A few quick implications of this definition are stated below.

> **Corollary 2.3.4**
>
> (1) If $\mathcal{T}$ is SST, then for each $i < j$,
>
> $$\mathbb{P}[t_i \text{ beats } t_j] \geq 0.5.$$
>
> (2) If $\mathcal{T}$ is SST, then for each i, j, k, ℓ such that $i < j$ and $k < \ell$,
>
> $$\mathbb{P}[t_i \text{ beats } t_\ell] \geq \mathbb{P}[t_j \text{ beats } t_k].$$
>
> (3) If $\mathcal{T}$ is SST, then the matchup table $\mathbf{M}$ is monotonically increasing along each row and monotonically decreasing along each column.

Note that not every set of teams can be seeded to be SST. Consider, for example, the game of rock-paper-scissors. Rock beats paper which beats scissors which beats rock, so no ordering of these "teams" will be SST. For our purposes, however, SST will work well

enough.

Our new, nuanced alternative a proper bracket is an *ordered bracket*, first defined by Chen and Hwang [2] (though we use the name proposed by Edwards [4]).

Definition 2.3.5: Ordered

A tournament format $\mathcal{A}$ is *ordered* if, for any SST list of teams $\mathcal{T}$, if $i < j$, then $\mathbb{W}_{\mathcal{A}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{T})$

In an informal sense, a bracket being ordered is the strongest thing we can want without knowing more about why the tournament is being played. Depending on the situation, we might be interested in a format that almost always declares the most-skilled team as the winner, or in a format that gives each team roughly the same chance of winning, or anywhere in between. But certainly, better teams should win more, which is what the ordered bracket condition requires.

In particular, a bracket being ordered is a stronger claim than it being proper.

Theorem 2.3.6

Every ordered bracket is proper.

- -

Proof. Let $\mathcal{A}$ be an ordered n-team bracket with r rounds.

Let $\mathcal{T}$ be SST with matchup table $\mathbf{M}$ where $\mathbf{M}_{ij} = 0.5$. A team that plays their first game in the ith round will win the tournament with probability $(0.5)^{r-i}$, so teams that get more byes will have a higher probability to win the tournament than teams with fewer byes. This implies that higher-seeded teams must have more byes than lower-seeded teams, so in each round, the teams with byes must be the highest-seeded teams that are still alive. Thus, condition (1) is met.

We show that condition (2) is met by proving the stronger condition from Lemma 2.2.11: if m teams have a bye and k games are being played in round s, then if the bracket goes chalk, those matchups will be t_{m+i} vs $t_{(m+2k+1)-i}$ for $i \in \{1, ..., k\}$. We show this by strong induction on s and on i.

Assume that this is true for every round up until s and for all $i < j$ for some j. Let $\ell = (m + 2k + 1) - j$. We want to show that if the the bracket goes chalk, t_{m+j} will face off against seed t_ℓ in the given round. Consider the following SST matchup table: every game is a coin flip, except for games involving a team seeded ℓ or lower, in which case the higher seed always wins. Then, each team seeded between $\ell - 1$ and $m + j$ will win the tournament with probability $(\frac{1}{2})^{r-s}$, other than the team slated to play t_ℓ in round s who wins with probability $(\frac{1}{2})^{r-i-1}$. In order for $\mathcal{B}$ to be ordered, that team must be t_{m+j}.

Thus $\mathcal{A}$ satisfies both conditions, and so is a proper bracket. □

With Theorem 2.3.6, we can use the language of bracket signatures to describe ordered brackets without worrying that two ordered brackets might share a signature. Now we examine three particularly important examples of ordered brackets.

We begin with the unique one-team bracket.

Figure 2.3.7: The One-Team Bracket $[[\mathbf{1}]]$

Theorem 2.3.8

The one-team bracket $[[\mathbf{1}]]$ is ordered.

Proof. Since there is only team, the ordered bracket condition is vacuously true. □

Next we look at the unique two-team bracket.

Figure 2.3.9: The Two-Team Bracket $[[\mathbf{2}; \mathbf{0}]]$

Theorem 2.3.10

The two-team bracket $[[\mathbf{2}; \mathbf{0}]]$ is ordered.

Proof. Let $\mathcal{A} = [[\mathbf{2}; \mathbf{0}]]$. Then,

$$\mathbb{W}_{\mathcal{A}}(t_1, \mathcal{T}) = \mathbb{P}[t_1 \text{ beats } t_2] \geq 0.5 \geq \mathbb{P}[t_2 \text{ beats } t_1] = \mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T})$$

so $\mathcal{A}$ is ordered. □

And thirdly, we show that the balanced four-team bracket is ordered, first proved by Horen and Riezman [5].

Theorem 2.3.12

The four-team bracket $[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]$ is ordered.

Proof. Let $\mathcal{A} = [[\mathbf{4}; \mathbf{0}; \mathbf{0}]]$ and let $p_{ij} = \mathbb{P}[t_i \text{ beats } t_j]$. Then,

$$
\begin{aligned}
\mathbb{W}_{\mathcal{A}}(t_1, \mathcal{T}) &= p_{14} \cdot (p_{23}p_{12} + p_{32}p_{13}) \\
&= p_{14}p_{23}p_{12} + p_{14}p_{32}p_{13} \\
&\geq p_{14}p_{23}p_{21} + p_{24}p_{41}p_{23} \\
&= p_{23} \cdot (p_{14}p_{21} + p_{41}p_{24}) \\
&= \mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T})
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T}) &= p_{23} \cdot (p_{14}p_{21} + p_{41}p_{24}) \\
&\geq p_{32} \cdot (p_{14}p_{31} + p_{41}p_{34}) \\
&= \mathbb{W}_{\mathcal{A}}(t_3, \mathcal{T})
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{W}_{\mathcal{A}}(t_3, \mathcal{T}) &= p_{32} \cdot (p_{14}p_{31} + p_{41}p_{34}) \\
&= p_{32}p_{14}p_{31} + p_{32}p_{41}p_{34} \\
&\geq p_{42}p_{23}p_{41} + p_{32}p_{41}p_{43} \\
&= p_{41} \cdot (p_{23}p_{42} + p_{32}p_{43}) \\
&= \mathbb{W}_{\mathcal{A}}(t_4, \mathcal{T})
\end{aligned}
$$

Thus $\mathcal{A}$ is ordered. $\qquad\square$

However, not every proper bracket is ordered. One particularly important example of a non-ordered proper bracket is $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$

Figure 2.3.13: The Six-Team Bracket $[[4;2;0;0]]$

Theorem 2.3.14

The six-team bracket $[[\mathbf{4};\mathbf{2};\mathbf{0};\mathbf{0}]]$ is not ordered.

- -

Proof. Let $\mathcal{A} = [[\mathbf{4};\mathbf{2};\mathbf{0};\mathbf{0}]]$, and let $\mathcal{T}$ have the following matchup table.

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	0.5	0.5	0.5	0.5	0.5	1
t_2	0.5	0.5	0.5	0.5	0.5	1
t_3	0.5	0.5	0.5	0.5	0.5	0.5
t_4	0.5	0.5	0.5	0.5	0.5	0.5
t_5	0.5	0.5	0.5	0.5	0.5	0.5
t_6	0	0	0.5	0.5	0.5	0.5

Then

$$\mathbb{W}_{\mathcal{A}}(t_1, \mathcal{T}) = 0.5 \cdot 0.5 = 0.25,$$

but

$$\mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T}) = (0.5 \cdot 0.5 + 0.5 \cdot 1) \cdot 0.5 = 0.375.$$

Thus $\mathcal{A}$ is not ordered. $\qquad\square$

In the next section, we move on from describing particular ordered and non-ordered brackets in favor of a more general result.

2.4 Edwards's Theorem

We now attempt to completely classify the set of ordered brackets. Edwards [4] originally accomplished this without access to the machinery of bracket signatures or proper brackets: we present a quicker proof that makes use of the fundamental theorem of brackets and develop two nice lemmas along the way.

We begin with the stapling lemma, which allows us to combine two smaller ordered brackets into a larger ordered one by having the winner of one of the brackets be treated as the lowest seed in the other. This is depicted in Figure 2.4.1.

Figure 2.4.1: Setup of the Stapling Lemma with $\mathcal{A} = [[2; 1; 0]]$, $\mathcal{B} = [[4; 0; 0]]$, and $\mathcal{C} = [[2; 1; 3; 0; 0]]$

Lemma 2.4.2: The Stapling Lemma

If $\mathcal{A} = [[\mathbf{a_0}; ...; \mathbf{a_r}]]$ and $\mathcal{B} = [[\mathbf{b_0}; ...; \mathbf{b_s}]]$ are ordered brackets, then $\mathcal{C} = [[\mathbf{a_0}; ...; \mathbf{a_r} + \mathbf{b_0} - \mathbf{1}; ...; \mathbf{b_s}]]$ is an ordered bracket as well.

- -

Proof. Let $\mathcal{A}, \mathcal{B}$, and $\mathcal{C}$ be as specified. Let $\mathcal{T}$ be an SST list of teams $n + m - 1$ teams, and let $\mathcal{R}, \mathcal{S} \subset \mathcal{T}$ be the lowest n and the highest $m - 1$ seeds of $\mathcal{T}$ respectively. We divide proving that $\mathcal{C}$ is ordered into proving three sub-statements:

1. For $i < j < m$, $\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})$

2. $\mathbb{W}_{\mathcal{C}}(t_{m-1}, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_m, \mathcal{T})$

3. For $m \leq i < j$, $\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})$

Together, these show that $\mathcal{C}$ is ordered.

We begin with the first sub-statement. Let $i < j < m$. Then,

$$
\begin{aligned}
\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) &= \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_k\}) \\
&\geq \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_j, \mathcal{S} \cup \{t_k\}) \\
&= \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})
\end{aligned}
$$

The first and last equalities follow from the structure of $\mathcal{C}$, and the inequality follows from $\mathcal{B}$ being ordered.

Now the second sub-statement.

$$
\begin{aligned}
\mathbb{W}_{\mathcal{C}}(t_{m-1}, \mathcal{T}) &= \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_{m-1}, \mathcal{S} \cup \{t_k\}) \\
&\geq \mathbb{W}_{\mathcal{A}}(t_m, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_{m-1}, \mathcal{S} \cup \{t_m\}) \\
&\geq \mathbb{W}_{\mathcal{A}}(t_m, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_m, \mathcal{S} \cup \{t_m\}) \\
&= \mathbb{W}_{\mathcal{C}}(t_m, \mathcal{T})
\end{aligned}
$$

The equalities follow from the structure of $\mathcal{C}$, the first inequality follows from probabilities being non-negative, and the second inequality follows from $\mathcal{B}$ being ordered.

Finally, we show the third sub-statement. Let $m \leq i < j$. Then,

$$
\begin{aligned}
\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) &= \mathbb{W}_{\mathcal{A}}(t_i, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_i\}) \\
&\geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_i\}) \\
&\geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_j, \mathcal{S} \cup \{t_j\}) \\
&= \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})
\end{aligned}
$$

The equalities follow from the structure of $\mathcal{C}$, the first inequality from $\mathcal{A}$ being ordered, and the second inequality from the teams being SST.

We have shown all three sub-statements, and so $\mathcal{C}$ is ordered. □

Now, if we begin with the set of brackets $\{[[\mathbf{1}]], [[\mathbf{2}; \mathbf{0}]], [[\mathbf{4}; \mathbf{0}; \mathbf{0}]]\}$ and then repeatedly apply the stapling lemma, we can construct a set of brackets that we know are ordered. In other words,

> **Corollary 2.4.3**
>
> Any bracket signature formed by the following process is ordered:
>
> 1. Start with the list $[[\mathbf{0}]]$ (note that this not yet a bracket signature).
>
> 2. As many times as desired, prepend the list with $[[\mathbf{1}]]$ or $[[\mathbf{3}; \mathbf{0}]]$.
>
> 3. Then, add 1 to the first element in the list, turning it into a bracket signature.

Corollary 2.4.3 uses the tools that we have developed so far to identify a set of ordered brackets. Somewhat surprisingly, this set is complete: any bracket not reachable using the process in Corollary 2.4.3 is not ordered. To prove this we first need to show the containment lemma.

> **Definition 2.4.4: Containment**
>
> Let $\mathcal{A}$ and $\mathcal{B}$ be bracket signatures. We say $\mathcal{A}$ *contains* $\mathcal{B}$ if there exists some i such that
>
> - At least as many games are played in the $(i+1)$th round of $\mathcal{A}$ as in the first round of $\mathcal{B}$, and
>
> - For $j > 1$, there are exactly as many games played in the $(i+j)$th round of $\mathcal{A}$ as in the jth round of $\mathcal{B}$.

Intuitively, $\mathcal{A}$ containing $\mathcal{B}$ means that if $\mathcal{A}$ went chalk, and games within each round were played in order of largest seed-gap to smallest seed-gap, then at some point, there would be a bracket of shape $\mathcal{B}$ used to determine to identify the last team in the rest of bracket $\mathcal{A}$. Figure 2.4.5 shows $\mathcal{A} = [[\mathbf{2}; \mathbf{5}; \mathbf{1}; \mathbf{0}; \mathbf{3}; \mathbf{0}; \mathbf{0}]]$ containing $\mathcal{B} = [[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$. After the 10v11 game and the 5v(10v11) game, there is a bracket of shape $\mathcal{B}$ (the solid lines) that must be played to determine the last team in the rest of the bracket

Figure 2.4.5: Setup of the Containment Lemma with $\mathcal{A} = [[2; 5; 1; 0; 3; 0; 0]]$ and $\mathcal{B} = [[4; 2; 0; 0]]$.
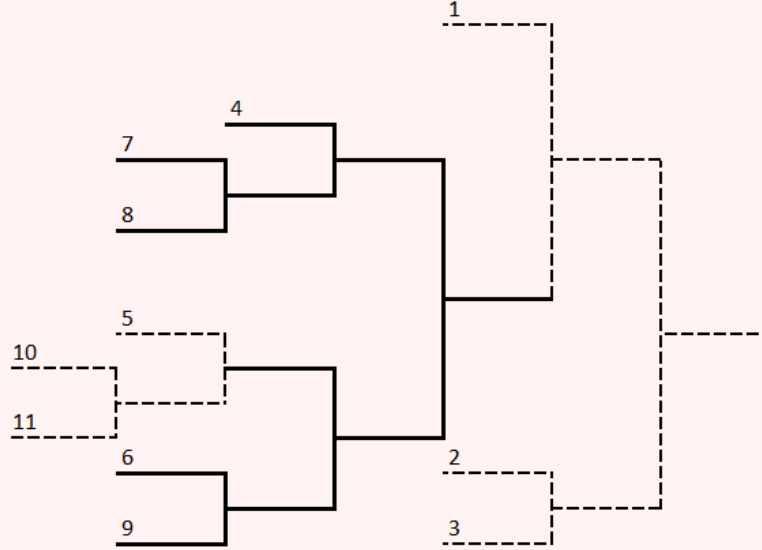
Lemma 2.4.6: The Containment Lemma

If $\mathcal{A}$ contains $\mathcal{B}$, and $\mathcal{B}$ is not ordered, then neither is $\mathcal{A}$.

Proof. Let $\mathcal{A}$ be a bracket signature with r rounds and n teams, and let $\mathcal{B}$ have s round and m teams, such that $\mathcal{A}$ contains $\mathcal{B}$ and $\mathcal{B}$ is not ordered. Let k be the number of teams in $\mathcal{A}$ that get at least $s+i$ byes (where i is from the definition of contains).

$\mathcal{B}$ is not ordered, so let $\mathbf{M}$ be a matchup table that violates the orderedness condition, where none of the win probabilities are 0. (If we have an $\mathbf{M}$ that includes 0s, we can replace them with ϵ. For small enough ϵ, $\mathbf{M}$ will still violate the condition.) Let p be the minimum probability in $\mathbf{M}$. Let $\mathbf{P}$ be a matchup table in which the lower-seeded team wins with probability p, and let $\mathbf{Z}$ be a matchup table in which the lower-seeded team wins with probability 0.

Now, consider the following block matchup table on $\mathcal{T}$, a list of n teams.

	t_1 - t_k	t_{k+1} - t_{k+m}	t_{k+m+1} - t_n
t_1 - t_k	**P**	**P**	**Z**
t_{k+1} - t_{k+m}	**P**	**M**	**Z**
t_{k+m+1} - t_n	**Z**	**Z**	**Z**

27

Let $\mathcal{S} \subset \mathcal{T}$ be the sublist of teams seeded between $k+1$ and $k+m$. Then, for $t_j \in \mathcal{S}$,

$$\mathbb{W}_{\mathcal{A}}(t, \mathcal{T}) = \mathbb{W}_{\mathcal{B}}(t, \mathcal{S}) \cdot p^{r-s-i},$$

since t_j wins any games it might have to play in rounds i or before automatically, any games after $s + i$ with probability p, and any games in between according to $\mathbf{M}$.

However, $\mathbf{M}$ (and thus $\mathcal{S}$) violates the orderedness condition for $\mathcal{B}$, and so $\mathcal{T}$ does for $\mathcal{A}$. $\qquad\square$

With the containment lemma shown, we can proceed to the main theorem.

Theorem 2.4.7: Edwards's Theorem

The only ordered brackets are those described by Corollary 2.4.3.

- -

Proof. Let $\mathcal{A}$ be a proper bracket not described by Corollary 2.4.3. The corollary describes all proper brackets in which each round either has only game, or has two games but is immediately followed by a round with only one game. Thus $\mathcal{A}$ must include at least two successive rounds with two or more games each.

The final round in such a chain will be followed by a round with a single game, and so the final round must have only two games. Thus, $\mathcal{A}$ includes a sequence of three rounds, the first of which has at least two games, the second of which has exactly two games, and the third of which has one game.

Therefore, $\mathcal{A}$ contains $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$. But we know that $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]$ is not ordered, and so by the containment lemma, neither is $\mathcal{A}$. $\qquad\square$

Edwards's Theorem is both exciting and disappointing. On one hand, it means that we can fully describe the set of ordered brackets, making it easy to check whether a given bracket is ordered or not. On the other hand, it means that in an ordered bracket at most three teams can be introduced each round, so the length of the shortest ordered bracket on n teams grows linearly with n (rather than logarithmically as is the case for the shortest proper bracket). If we want a bracket on many teams to be ordered, we risk forcing lower-seeded teams to play a large number of games, and we only permit the top-seeded teams to play a few. For example, the shortest ordered bracket that could've been used in the 2021 NCAA Basketball South Region is $[[\mathbf{4}; \mathbf{0}; \mathbf{3}; \mathbf{0}; \mathbf{3}; \mathbf{0}; \mathbf{3}; \mathbf{0}; \mathbf{3}; \mathbf{0}; \mathbf{0}]]$, which is played over a whopping ten rounds.

Figure 2.4.8: The Shortest Sixteen-Team Ordered Bracket

Because of this, few leagues use ordered brackets, and those who do usually have so few teams that every proper bracket is ordered (the 2023 College Football Playoffs, for example). Even the Korean Baseball Organization League, which uses a somewhat unconventional $[[2; 1; 1; 1; 0]]$, only sends five teams to the playoffs, and again every five-team proper bracket is ordered. If the KBO League ever expanded to the six-team bracket $[[2; 1; 1; 1; 1; 0]]$, we would have a case of an ordered bracket being used when a proper non-ordered bracket exists on the same number of teams.

2.5 Reseeded Brackets

Edwards's Theorem naturally raises the question: is there some bracket-like tournament format, one where undefeated teams face off until only one reamins, that expands the space of signatures that are ordered? *Reseeded* brackets are our first attempt at an answer.

Ultimately, the reason that proper brackets are not, in general, ordered, is that lower-seeded teams are treated, if they win, as the team that they beat for the rest of the format. Consider again the proper bracket analyzed by Silver: $[[\mathbf{16}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}]]$. If an 11-seed wins in the first round, they take on the schedule of a 6-seed for the rest of the tournament, while if the 9-seed wins, they take on the schedule of an 8-seed. Given that a 6-seed has an easier schedule than an 8-seed, it's not hard to see why it might be preferable to be an 11-seed rather than a 9-seed.

Reseeding (poorly named) fixes this by resorting the match-ups every round: if an 11-seed keeps winning, they will have to play teams according to their seed, rather than getting an effective upgrade to 6-seed status.


> **Definition 2.5.1: Reseeded Brackets**
>
> In a *reseeded* bracket, after each round, match up the highest-seeded team with the lowest-seeded team, second-highest vs second-lowest, etc.

Note that by Definition 2.1.1, a reseeded bracket is not a bracket at all, as matchups between teams that have not yet lost are not determined in advance of the outcomes of any games. However, because reseeded brackets act so similarly to traditional brackets, and because colloquially they are referred to as brackets, we opt to continue using the word "bracket" to describe them.

In 2023, both National Football League conferences use a reseeded bracket with signature $[[\mathbf{6}; \mathbf{1}; \mathbf{0}; \mathbf{0}]]^R$. (The superscript R indicates this is reseeded bracket.) If the first round of the bracket goes chalk, then it looks just like a normal bracket.

Figure 2.5.2: 2023 National Football League AFC Playoffs
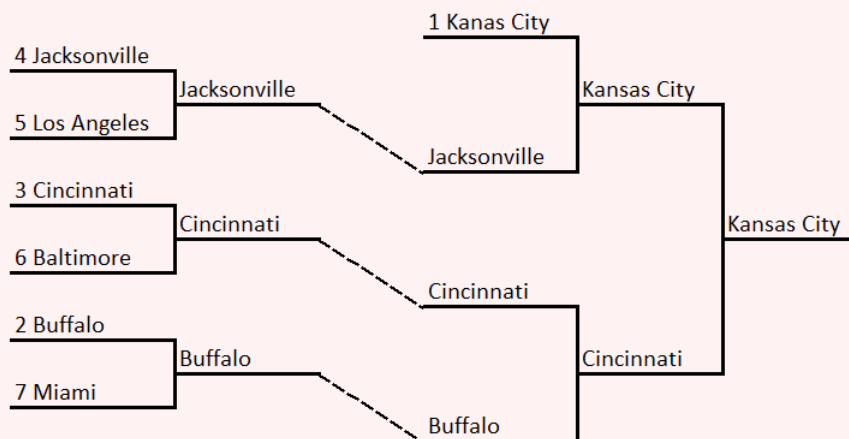
The dotted lines are drawn after the first round of games has been played: if there are some first-round upsets, then the bracket is rearranged to ensure that it is still better to be a higher seed rather than a lower seed.

Figure 2.5.3: 2023 National Football League NFC Playoffs
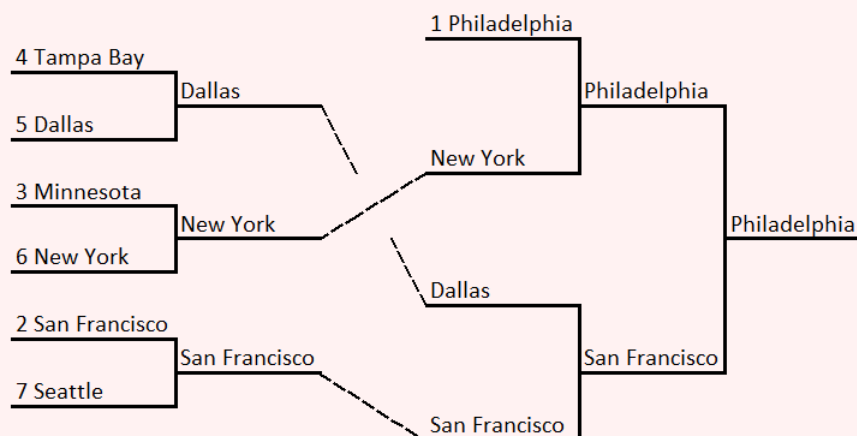
In the NFC, 6-seed New York upset 3-seed Minnesota. Had a conventional bracket been used, the semifinal matchups would have been 1-seed vs 5-seed and 2-seed vs 6-seed: the 2-seed would have had an easier draw than the 1-seed, while the 6-seed would have an easier draw than the 5-seed. Reseeding fixes this by matching 6-seed New York with top-seed Philadelphia, and 2-seed San Francisco with 5-seed Dallas.

Reseeding is a powerful technique. For one, the fundamental theorem still applies to reseeded brackets, allowing us to refer to reseeded brackets by their signatures as well.

Theorem 2.5.4

There is exactly one proper reseeded bracket with each bracket signature.

- -

Proof. The definition of properness ensures that there is only one way byes can be distributed such that a reseeded bracket can be proper. Additionally, because reseeded brackets have no additional parameters beyond which seeds get how many byes, there is no more than one reseeded bracket with each signature that could be proper. Finally, that bracket is indeed proper: if the bracket goes to chalk, the matchups will be the exact same as a traditional bracket, which by the fundamental theorem is a proper set of matchups. □

But what about orderedness? It's intuitive to think that all proper reseeded are ordered: it feels like almost by definition, the higher-seeded teams have an easier path than the lower-seeded ones. Hwang [6] conjectured a weaker version of this.

> **Conjecture 2.5.5**
>
> All balanced proper reseeded brackets are ordered.

Unfortunately, neither the stronger claim that all proper reseeded brackets are ordered, nor Hwang's weaker conjecture are true. Our classification of the ordered reseeded brackets takes the same route as our proof of Edwards's Theorem did: we first examine the orderedness of certain important brackets, and then we use the stapling and containment lemmas to specify the complete set of ordered reseeded brackets.

The proofs of the stapling and containment lemmas for reseeded brackets, as well as the fact that all ordered reseeded brackets are proper, are so similar to the corresponding proofs for traditional brackets that we just state them without proof.

> **Theorem 2.5.6**
>
> All ordered reseeded brackets are proper.

> **Lemma 2.5.7: The Stapling Lemma for Reseeding**
>
> If $\mathcal{A} = [[\mathbf{a_0}; ...; \mathbf{a_r}]]^R$ and $\mathcal{B} = [[\mathbf{b_0}; ...; \mathbf{b_s}]]^R$ are ordered reseeded brackets, then $\mathcal{C} = [[\mathbf{a_0}; ...; \mathbf{a_r} + \mathbf{b_0} - \mathbf{1}; ...; \mathbf{b_s}]]^R$ is an ordered reseeded bracket as well.

> **Lemma 2.5.8: The Containment Lemma for Reseeding**
>
> If $\mathcal{A}$ and $\mathcal{B}$ are reseeded brackets, $\mathcal{A}$ contains $\mathcal{B}$, and $\mathcal{B}$ is not ordered, then neither is $\mathcal{A}$.

We now examine particular brackets.

> **Theorem 2.5.9**
>
> $[[\mathbf{1}]]^R$, $[[\mathbf{2}; \mathbf{0}]]^R$, and $[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]^R$ are ordered.
>
> -
>
> *Proof.* Since no reseeding is done in a bracket of two or fewer rounds, and since the traditional brackets of these signatures are ordered, so are the reseeded brackets. $\square$

Our primary example of a reseeded bracket that is ordered despite the traditional bracket of the same signature not being ordered is $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$.

> **Theorem 2.5.10**
>
> $[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$ is ordered.

Proof. This can be shown by computing the probability of each team winning the format and then applying the SST conditions to establish the inequalities, as we did in Theorem 2.3.12. In the interest of brevity, however, we instead give an intuitive argument.

$\mathbb{W}_A(t_1, \mathcal{T}) \geq \mathbb{W}_A(t_2, \mathcal{T})$ because from those two teams perspectives, this format is just $[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]^R$. $\mathbb{W}_A(t_2, \mathcal{T}) \geq \mathbb{W}_A(t_3, \mathcal{T})$ because t_2 has better odds if t_3 wins in the first round and they meet in the semifinals, and certainly has better odds if t_3 loses in the first round. $\mathbb{W}_A(t_4, \mathcal{T}) \geq \mathbb{W}_A(t_5, \mathcal{T})$ because t_4 is at least as likely to win the first-round matchup, and then their paths would be identical.

$\mathbb{W}_A(t_3, \mathcal{T}) \geq \mathbb{W}_A(t_4, \mathcal{T})$ holds because if both teams win the first round then t_3 has better odds in the remaining $[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]^R$ bracket. Meanwhile if only one does, then t_3 will be joined by t_5 while t_4 will be joined by t_6, and so t_3 is more likely to dodge playing t_1 in the finals. The same argument applies to show that $\mathbb{W}_A(t_5, \mathcal{T}) \geq \mathbb{W}_A(t_6, \mathcal{T})$ as well. $\square$

Unfortunately, that is where the power of reseeding to convert non-ordered signatures into ordered ones ends. The following two signatures are not ordered.

Theorem 2.5.11

$[[\mathbf{6}; \mathbf{1}; \mathbf{0}; \mathbf{0}]]^R$ is not ordered.

Proof. Let $\mathcal{A} = [[\mathbf{6}; \mathbf{1}; \mathbf{0}; \mathbf{0}]]^R$, and let $\mathcal{T}$ have the following matchup table.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7
t_1	0.5	$1-p$	$1-p$	$1-p$	$1-p$	$1-p$	$1-p$
t_2	p	0.5	$1-p$	$1-p$	$1-p$	$1-p$	$1-p$
t_3	p	p	0.5	0.5	0.5	$1-p$	$1-p$
t_4	p	p	0.5	0.5	0.5	0.5	0.5
t_5	p	p	0.5	0.5	0.5	0.5	0.5
t_6	p	p	p	0.5	0.5	0.5	0.5
t_7	p	p	p	0.5	0.5	0.5	0.5

Then

$$\mathbb{W}_\mathcal{A}(t_6, \mathcal{T}) = O(p^3),$$

but

$$\mathbb{W}_\mathcal{A}(t_7, \mathcal{T}) = 0.25p^2 + O(p^3).$$

Thus, for small enough p, $\mathbb{W}_\mathcal{A}(t_6, \mathcal{T}) < \mathbb{W}_\mathcal{A}(t_7, \mathcal{T})$, so $\mathcal{A}$ is not ordered. $\square$

Theorem 2.5.12

$[[\mathbf{4}; \mathbf{2}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$ is not ordered.

Proof. Let $\mathcal{A} = [[\mathbf{4}; \mathbf{2}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$, and let $\mathcal{T}$ have the following matchup table.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
t_1	0.5	$1-p^2$	$1-p^2$	$1-p^2$	$1-p^2$	$1-p^2$	$1-p^2$	$1-p^2$
t_2	p^2	0.5	0.5	0.5	$1-p$	$1-p$	$1-p^2$	$1-p^2$
t_3	p^2	0.5	0.5	0.5	$1-p$	$1-p$	$1-p$	$1-p$
t_4	p^2	0.5	0.5	0.5	0.5	$1-p$	$1-p$	$1-p$
t_5	p^2	p	p	0.5	0.5	$1-p$	$1-p$	$1-p$
t_6	p^2	p	p	p	p	0.5	$1-p$	$1-p$
t_7	p^2	p^2	p	p	p	p	0.5	0.5
t_8	p^2	p^2	p	p	p	p	0.5	0.5

Then

$$\mathbb{W}_{\mathcal{A}}(t_7, \mathcal{T}) = 0.25p^5 + O(p^6)$$

but

$$\mathbb{W}_{\mathcal{A}}(t_8, \mathcal{T}) = 0.5p^5 + O(p^6).$$

Thus, for small enough p, $\mathbb{W}_{\mathcal{A}}(t_7, \mathcal{T}) < \mathbb{W}_{\mathcal{A}}(t_8, \mathcal{T})$, so $\mathcal{A}$ is not ordered. $\square$

Recapping,

Figure 2.5.13: Which Proper Reseeded Brackets are Ordered

Ordered	Not Ordered
$[[\mathbf{1}]]^R$	$[[\mathbf{6}; \mathbf{1}; \mathbf{0}; \mathbf{0}]]^R$
$[[\mathbf{2}; \mathbf{0}]]^R$	$[[\mathbf{4}; \mathbf{2}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$
$[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]^R$	
$[[\mathbf{4}; \mathbf{2}; \mathbf{0}; \mathbf{0}]]^R$	

Finally, we apply the stapling and containment lemmas to complete the theorem.

Theorem 2.5.14

The ordered reseeded brackets are exactly those corresponding to signatures that can be generated in the following way.

1. Start with the list $[[\mathbf{0}]]^R$ (note that this not yet a bracket signature).

2. As many times as desired, prepend the list with $[[\mathbf{1}]]$, $[[\mathbf{3}; \mathbf{0}]]$, or $[[\mathbf{3}; \mathbf{2}; \mathbf{0}]]$.

3. Then, add 1 to the first element in the list, turning it into a bracket signature.

Proof. The stapling lemma, combined with the fact that $[[\mathbf{1}]]^R$, $[[\mathbf{2};\mathbf{0}]]^R$, $[[\mathbf{4};\mathbf{0};\mathbf{0}]]^R$, and $[[\mathbf{4};\mathbf{2};\mathbf{0};\mathbf{0}]]^R$ are ordered, ensure that any reseeded brackets generated by the above procedure is indeed ordered. Left is to use the containment lemma to ensure that these are the only ones.

Let $\mathcal{A}$ be a bracket signature that cannot be generated by the procedure. Then, either there is a round in which three or more games are to be played, or there is a round in which exactly two games are played and the next two rounds each have exactly two games played as well.

Let i be the latest such round. If round i is the first of three rounds with two games each, then round $i + 3$ must have only one game played (otherwise i would not be the latest such round). But then $\mathcal{A}$ contains $[[\mathbf{4};\mathbf{2};\mathbf{2};\mathbf{0};\mathbf{0}]]^R$, and so is not ordered.

If round i has three or more games, then round $i + 1$ must contain exactly two games (any less and not every winner would have a game, any more and i would not be the latest such round.) Then, if round $i + 2$ has one game, then $\mathcal{A}$ contains $[[\mathbf{6};\mathbf{1};\mathbf{0};\mathbf{0}]]^R$, and if it has two, then $\mathcal{A}$ contains $[[\mathbf{4};\mathbf{2};\mathbf{2};\mathbf{0};\mathbf{0}]]^R$. In either case, $\mathcal{A}$ is not ordered.

Thus, the ordered reseeded brackets are exactly those generated by the procedure. □

So, the space of ordered reseeded brackets is slightly larger than the space of ordered traditional brackets, although perhaps this is not quite as much of an expansion as we would've liked or expected. Despite this, reseeded brackets definitely *feel* more ordered than traditional brackets of the same signature, even if neither is ordered in the definitional sense.

Conjecture 2.5.15

There is some reasonable restriction on a set of teams that is stronger than SST under which all reseeded brackets ordered.

In the meantime, reseeding remains an important tool in our tournament design toolkit. But it is not without its drawbacks, as discussed by Baumann, Matheson, and Howe [1].

In a reseeded bracket, teams and spectators alike don't know who they will play or where their next game will be until the entire previous round is complete. This can be an especially big issue if parts of the bracket are being played in different locations on short turnarounds: in the NCAA Basketball Tournament, the first two rounds are played over a weekend at various pre-determined locations. It would cause problems if teams had to pack up and travel across the country because they got reseeded and their opponent and thus location changed.

In addition, part of what makes the NCAA Basketball Tournament (affectionately known

as "March Madness") such a fun spectator experience is the fact that these matchups are known ahead of time. In "bracket pools," groups of fans each fill out their own brackets, predicting who will win each game and getting points based on how many they get right. If it wasn't clear where in the bracket the winner of a given game was supposed to go, this experience would be diminished.

Finally, reseeding gives the top seed(s) an even greater advantage than they already have: instead of playing against merely the *expected* lowest-seeded team(s) each round, they would get to play against the *actual* lowest-seeded team(s). In March Madness, "Cinderella Stories," that is, deep runs by low seeds, would become much less common.

In many ways, the NFL conference playoffs are a perfect place to use a reseeded bracket: games are played once a week, giving plenty of time for travel; only seven teams make the playoffs in each, so a huge March Madness-style bracket challenge is unlikely; as a professional league, the focus is far more on having the best team win and protecting Cinderella Stories isn't as important; and because the bracket is only three rounds long, reseeding is only required once. Somewhat ironically, the NFL conference playoffs used to use the format $[[\mathbf{4};\mathbf{2};\mathbf{0};\mathbf{0}]]^{R}$ which is ordered, but have since allowed a seventh team from each conference into the playoffs and changed to the non-ordered $[[\mathbf{6};\mathbf{1};\mathbf{0};\mathbf{0}]]^{R}$.

Other leagues with similar structures might consider adopting forms of reseeding to protect their incentives and competitive balance (looking at you, Major League Baseball), but in many cases, the traditional bracket structure is too appealing to adopt a reseeded one.

3   Multibrackets

3.1 Consolation Brackets

In the pervious chapter, we discussed brackets and similar formats (reseeded and cohort randomized brackets) paying attention only to which team is declared champion. Edwards's Theorem and its analogies, for example, makes claims only about which teams are most likely to win the tournament: all participants that don't are grouped together as losers. Real tournaments, however, do not always operate in this way: for example, explicitly or not, the team that lost the championship game is often considered to have earned the second-place finish.

Third-place is often harder to determine. If a team is given a bye all the way to the finals, and thus there is only one semifinal, then the loser of that semifinal can be unambiguously granted third. The 2023 Korean Baseball Organization League Playoffs have this property: they use a bracket of signature $[[2; 1; 1; 1; 0]]$, and so could easily assign third-place to the NC Dinos, who lost in the sole semifinal. (The LG Twins won the format, and finals losers KT Wiz came in second.)

Figure 3.1.1: 2023 KBO League Playoffs
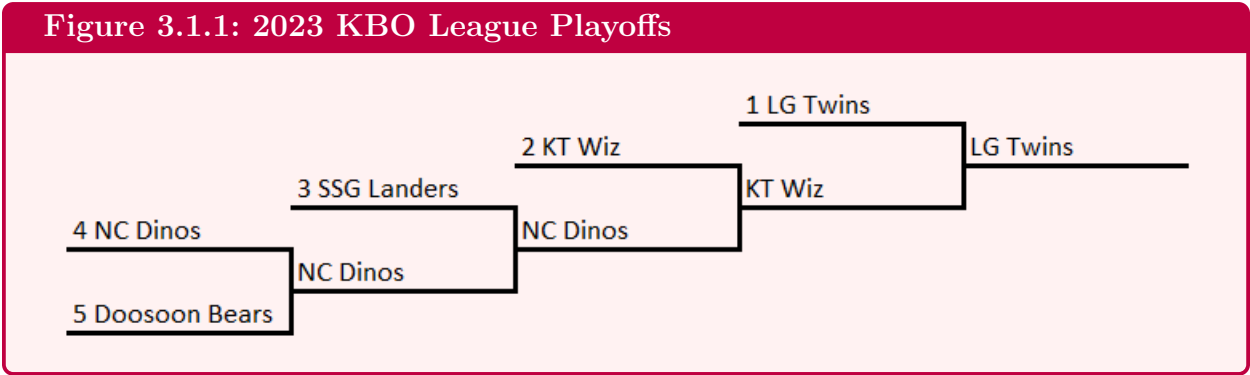
But in most brackets (those brackets whose signature's penultimate digit is a zero), assigning third-place is trickier: there are two teams who lost in the semifinal and have an equal claim to the place. There are a number of strategies that a league might use in the face of this ambiguity.

The first option is to just not assign a third-place at all. In the wise words of Will Ferrell from Talladega Nights, "If you ain't first, you're last." Who cares who came in third: you didn't win, you didn't even come in second, so you lost. This approach is taken by all four major American professional sports leagues (the NFL, NBA, NHL, and MLB).

The second option is to declare the two semifinal losers co-third-place finishers. In many ways, this is the same as the first option, but with a single sentence added to the end of a press-release indicating that the teams in question each finished third. (This option also has the unsatisfying property that four teams will be able to claim a top-three finish. This can be easily fixed, however, by just granting both teams fourth-place instead.)

The third option is to use some (relatively) arbitrary tiebreaker to select the third-place team. A few potential such tiebreakers are: whichever team was seeded higher, whichever team lost to the tournament champion (as opposed to the tournament runner-up), or if the teams played each other during the "regular season" portion of a tournament, whichever

team won that game.

None of these are particularly satisfying. While they may do alright when giving out third-place isn't super important, if we really want to assign third in a fair and equitable way, say because there is a bronze medal or spot in a future tournament up grabs, these options will not do.

Instead, the best thing to do is play a game: The 2015 Asian Football Confederation Asian Cup did exactly that.

Figure 3.1.2: The 2015 AFC Asian Cup
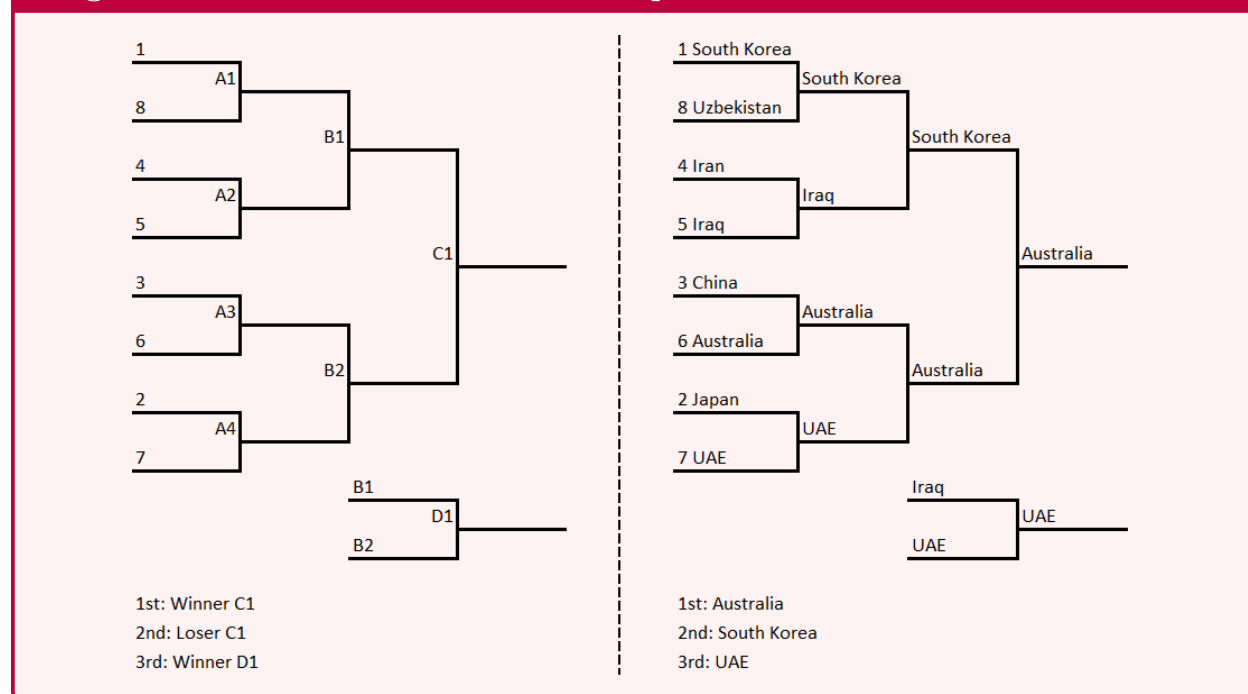
In the 2015 AFC Asian Cup, after the main bracket is complete, with the winner of the final game (Australia) being crowned champion and the loser (South Korea) coming in second, the two semifinal losers (Iraq and the UAE) are matched up in the third-place game.

A quick note about Figure 3.1.2: each game in the figure is labeled. In the primary bracket, first-round games are **A1** through **A4**, while the semifinals are **B1** and **B2**, and the finals is game **C1**. The third-place game is labeled **D1**: even though it could be played concurrently to the championship game, it is part of a different bracket and so we label it as a different round.

We indicate that the third-place game is to be played between the losers of games **B1** and **B2** by labeling the starting lines in the third-place game with those games. This is not ambiguous because the winners of those games always continue on in the original bracket, so such labels only refer to the losers.

The third-place game, which can also be viewed as a two-team bracket of signature $[[\mathbf{2}; \mathbf{0}]]$, in an example of a *consolation bracket*.

> **Definition 3.1.3: Consolation Bracket**
>
> A *consolation bracket* is a bracket in which teams that did not win the tournament compete for an n-place finish for some n.

The third-place game, as used by the 2015 AFC Asian Cup, is a common and well-liked consolation bracket used for selecting the top-three teams after a bracket. But it is far from the only way that the AFC could have doled out gold, silver, and bronze.

In fact, it's not clear the loser of **C1**, who comes in second place, is really more deserving than the winner of **D1**, who comes in third. The UAE might argue: South Korea and we both finished with two wins and one loss – a first-round win, a win against Iraq, and a loss against Australia. The only reason that South Korea came in second and we came in third was because South Korea lucked out by having Australia on the other half of the bracket as them. That's not fair!

If the AFC took this complaint seriously, they could modify their format to add a game **E1** for second-place to be played between the loser of **C1** and the winner of **D1**, with the loser coming in third.

Figure 3.1.4: The 2015 AFC Asian Cup Alternative
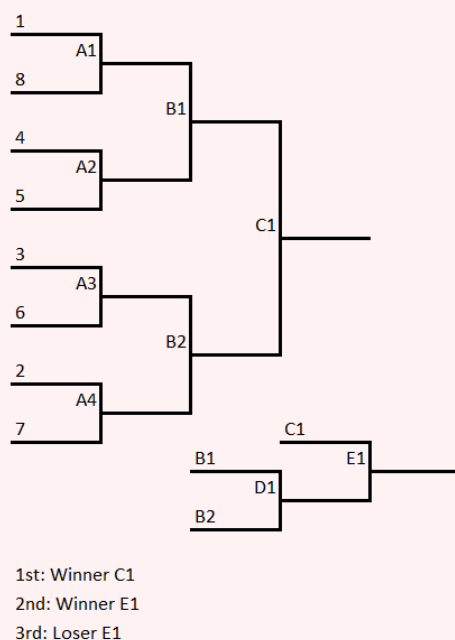
If the AFC used the format in Figure 3.1.4 in 2015, then South Korea and the UAE would have played each other for second place after all of the other games were completed. In some sense, this is a more equitable format than the one used in reality: we have the same data about the UAE and South Korea and so we ought to let them play for second-place instead of having decided almost randomly.

However, swapping formats doesn't come without costs. For one thing, South Korea and

the UAE would've had to play a fourth game: if the AFC had only three days to put on the tournament and teams can play at most one game a day, then the format in Figure 3.1.4 isn't feasible.

Another concern: what if Iraq had beaten the UAE when they played in game **D1**? Then the two teams with a claim to second-place would have been South Korea and Iraq, except South Korea already beat Iraq! One option is to say "tough luck, later games being more important than earlier ones is a staple in sports". But another is to designate game **E1** as *contingent*.

> ### Definition 3.1.5: Contingent Game
>
> We say a game in a tournament format is *contingent* if under certain circumstances (most commonly if the teams have already played earlier in the tournament,) the game is skipped and the result of a previous game is used.

Ultimately, whether game **E1** should be in included or not depends on the purpose of the tournament. If there is a huge difference between the prizes for coming in second and third, for instance, if the top two finishing teams in the Asian Cup qualified for the World Cup, then **E1** is quite important. If, on the other hand, this is a self-contained format played purely for bragging rights, **E1** could probably be left out. In reality, the 2015 AFC Asian Cup qualified only its winner to another tournament (the 2017 Confederations Cup), and gave medals to its top three, so game **E1**, which distinguishes between second- and third-place, was probably unnecessary.

Consider if instead of just the champion, the top four teams from the Asian Cup advanced to the Confederations Cup. One could imagine an easy extension of the format used presently, in which the loser of game **D1** is awarded fourth-place, to determine the four teams that qualify. However, this format would be quite poor: whether or not a team qualifies for the Confederations Cup would be solely determined by the result of their first-game and so the **B**, **C**, and **D**-round games might as well not even be played. A better format for selecting the top-four would allow first-round losers to win their way back onto the podium, as was employed by the 2023 Southern Conference Wrestling Championships.

Figure 3.1.6: 2023 Southern Conference Wrestling Championships
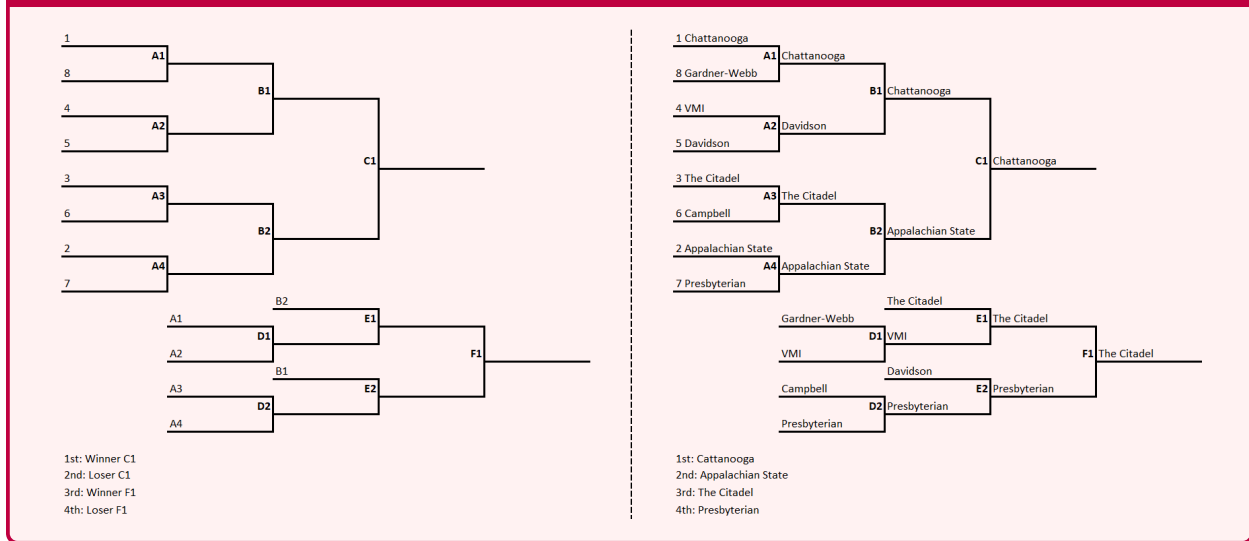
The format in Figure 3.1.6 is a dramatic improvement for selecting a top-four over that in Figure 3.1.2. In the 2023 Southern Conference Wrestling Championships, teams finish in the top-four if and only if they win two games before they lose two, which is a nice property to have. The one downside is it takes a fourth round: if there is not enough time for a fourth round, or if there is safety risk to teams playing four matches in a row, the format isn't feasible. Though if we only care about the top-four, and not the specifics of which team came in third or in fourth, we could drop game **F1**, ensuring that each team plays at most three games.

Figure 3.1.7: 2023 Southern Conference Wrestling Championships Alternative

1st: Winner C1
2nd: Loser C1
4th: Winner E1, Winner E2
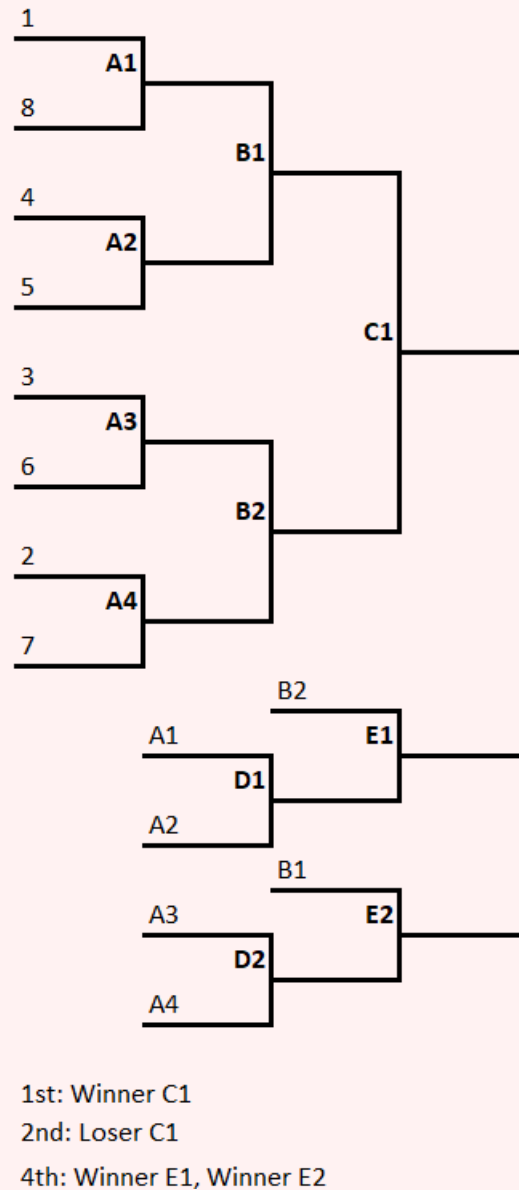
(As discussed earlier, we opt to both the **E**-round winners in fourth, to ensure that no more than m teams can claim a top-m finish for any m.)

The four formats with consolation brackets presented thus far are examples of *multi-brackets*.

> **Definition 3.1.8: Multibracket**
>
> A *multibracket* is a collection of one or more brackets coupled with a specification of which winners and losers of which games receive which places. Starting lines in multibrackets can be marked a seed, or by a game, indicating that loser of the specified game should be placed there.

Since which game each team plays in next (and which place each team ends up in) can be derived only from which game that team played in most recently and whether they won or lost that game, this definition is equivalent to saying that the format upholds the network condition from Section 1.1.

> **Definition 3.1.9: Multibracket**
>
> A *multibracket* is a tournament format in which, after t_i plays t_j, the rest of the tournament is identical no matter which team won, except for t_i and t_j are swapped.

(Note that this means multibrackets with contingent games are technically not multibrackets at all. However, they are close enough to being multibrackets and are important enough tools for tournament design that we include them in our discussion, in the same way that in the last chapter we discussed reseeded brackets even though they are technically not brackets.)

We will see in the coming sections that many formats used in a variety of settings are actually just examples of multibrackets. Figure 3.1.10 gives an outline of what the space of multibrackets looks like: we will spent the rest of the chapter examining the various categories in more detail.

Figure 3.1.10: The Space of Multibrackets

Multibrackets

Linear Multibrackets

Semibrackets

Traditional Brackets

Consolation Brackets

Efficient Linear Multibrackets

Swiss Systems

Nonlinear Multibrackets

Double Elimination

Page-McIntyre Systems

Bitonic Sort
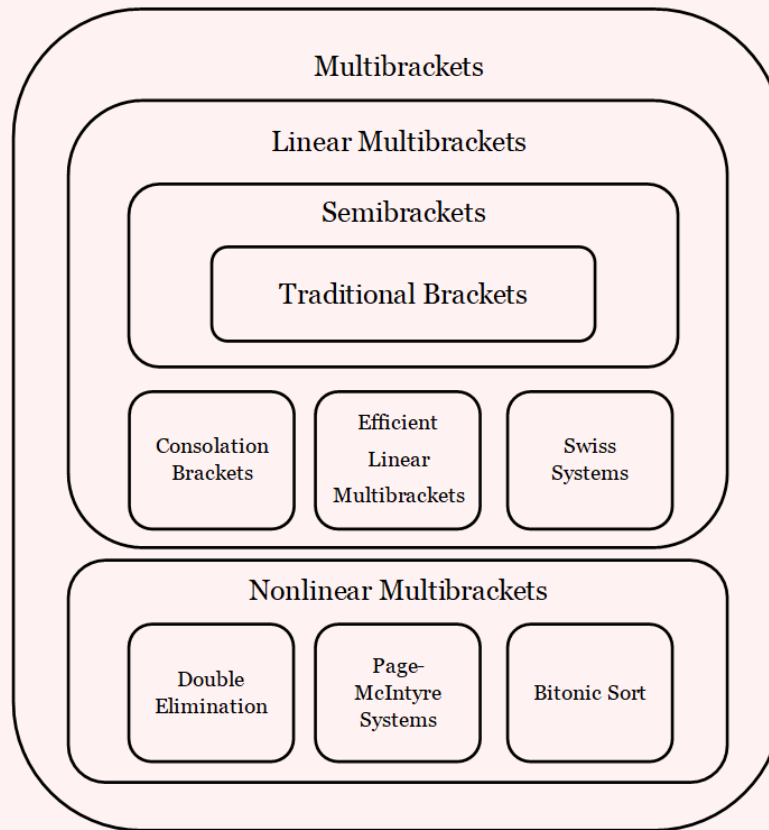
3.2 Semibrackets

In this section, we move on from consoloation brackets to focus on semibrackets, which as indicated by Figure 3.1.10 is a generalization of the traditional bracket. We will then use the notion of a semibracket to define linear multibrackets, which will study for a few sections before addressing nonlinear multibrackets at the end of the chapter.

Consider now the following tournament design problem: we are tasked with designing an eight-team tournament to select the top two teams who will go on to compete in the national tournament. However, there's only enough time for two rounds: perhaps due to field space or team fatigue, each team can only play two games. What design should we use?

The most natural answer to this question is to use an traditional eight-team bracket, but leave the championship game unplayed. This format is displayed in the figure below.

Figure 3.2.1: $[[8; 0; 0; 0]]$ **with no Championship Game**
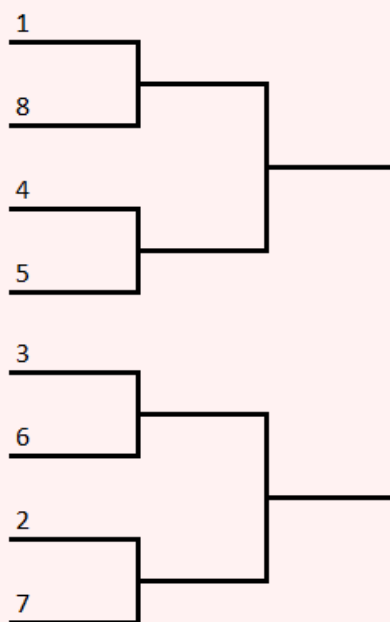
The format in Figure 3.2.1 does exactly what we need. The championship game being left unplayed is not a bug but a feature: each team plays a maximum of two games, and the two teams that advance to the national tournament are clear.

While it would be reasonable to describe the format in Figure 3.2.1 as two brackets that run side-by-side, it would be nice to be able to describe it as a single format: a bracket in which the championship game is left unplayed.

Definition 3.2.2: Semibracket

A *semibracket* is a tournament format in which:

- Teams don't play any games after their first loss.

- The matchups between game winners are determined in advance of the outcomes of any games.

All teams that finish a semibracket with no losses are declared co-champions.

Thus semibrackets are a generalization of brackets: a bracket is a semibracket in which only one team is left undefeated and declared champion.

Figure 3.2.3 describes which properties various bracket-like formats require.

Figure 3.2.3: Properties of Networked Formats

Format	No games after first loss	Only one team finishes undefeated	Network property
Traditional Bracket	✔	✔	✔
Semibracket	✔	✘	✔
Multibracket	✘	✘	✔

The format in Figure 3.2.1 is not a particularly exciting example of a semibracket: it is just a traditional bracket minus one game. Are there any examples of semibrackets that are not traditional brackets with some rounds left uncompleted?

Indeed there are. Let's modify the original problem so that we need to pick a top three teams out of twelve. Again, no team can play more then two games. The natural choice is shown below in Figure 3.2.4.

Figure 3.2.4: A More Exciting Semibracket
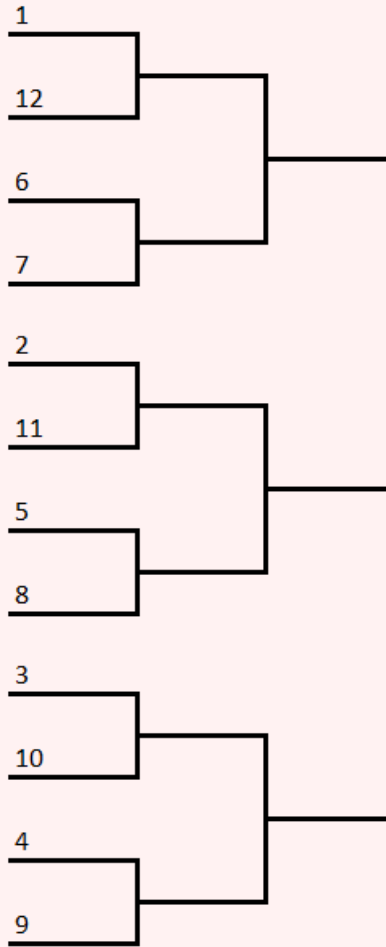
There is no potential for the format in Figure 3.2.4 to be completed into a traditional bracket, the next round would include three teams: an odd number. But as a semibracket, this is still a viable format, one that nicely solves the tournament design problem that we were given.

Definition 3.2.5: Rank of a Semibracket

The *rank* of a semibracket is how many co-champions it crowns. If the semibracket $\mathcal{A}$ has rank m, we say $\mathrm{Rank}(\mathcal{A}) = m$ or that $\mathcal{A}$ *ranks m teams*.

Traditional brackets are exactly the semibrackets that rank one team. The formats in Figures 3.2.1 and 3.2.4 rank two and three teams, respectively.

We can adapt the concept of a bracket signature to semibrackets.

> **Definition 3.2.6: Semibracket Signature**
>
> The *signature* $[[\mathbf{a_0}; ...; \mathbf{a_r}]]_m$ of an r-round semibracket $\mathcal{A}$ is a list such that a_i is the number of teams with i byes and $m = \text{Rank}(\mathcal{A})$. (In the case where $m = \text{Rank}(\mathcal{A}) = 1$, it can be omitted.)

Thus the signature of traditional brackets are the same as when they are viewed as semibrackets that rank one team. The signatures of the formats in Figures 3.2.1 and 3.2.4 are $[[\mathbf{8}; \mathbf{0}; \mathbf{0}]]_2$ and $[[\mathbf{12}; \mathbf{0}; \mathbf{0}]]_3$, respectively.

In analogy with traditional bracket signature's Theorem 2.1.13, we have Theorem 3.2.7.

> **Theorem 3.2.7**
>
> Let $\mathcal{A} = [[\mathbf{a_0}; ...; \mathbf{a_r}]]_m$ be a list of natural numbers. Then $\mathcal{A}$ is a semibracket signature if and only if
> $$\sum_{i=0}^{r} a_i \cdot \left(\frac{1}{2}\right)^{r-i} = m.$$

The proof is almost identical to that of Theorem 2.1.13 so we leave it out for brevity. Likewise, properness can be defined in the same way for semibracket, and the fundamental theorem still applies. (Again with a nearly identical proof that is left out for brevity.)

> **Theorem 3.2.8**
>
> Each semibracket signature admits exactly one proper semibracket.

Finally, we give a few descriptors to describe certain semibracket shapes.

> **Definition 3.2.9: Trivial Semibracket**
>
> We say a semibracket is *trivial* if every team is declared co-champion without playing any games. Equivalently, a semibracket is trivial its signature is of the form $[[\mathbf{m}]]_m$.

> **Definition 3.2.10: Competitive Semibracket**
>
> We say a semibracket is *competitive* if no teams are declared co-champion without winning at least one game. Equivalently, a semibracket is competitive its signature ends in a 0.

Clearly the two categories are mutually exclusive. Restricting briefly to the domain of traditional brackets, the two categories are also complete: there is no traditional bracket that is neither competitive nor trivial. (In fact, the only trivial traditional bracket is $[[\mathbf{1}]]$, every other traditional bracket is competitive.) However, this dichotomy does not apply to semibrackets: there are semibrackets that are neither trivial nor competitive. The simplest example is $[[\mathbf{2}; \mathbf{1}]]_2$, where the 1-seed is automatically one co-champion (so it's not

competitive), but the 2- and 3-seeds play to be the other co-champion (so it's not trivial).

Figure 3.2.11: $[[2; 1]]_2$
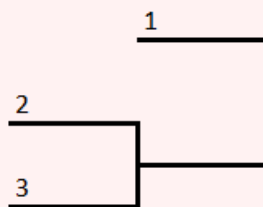
These two properties of semibrackets will sometimes be useful in defining and proving theorems about certain types of multibrackets down the line. In the next section, we will use semibrackets to construct a particularly nice kind of multibrackets: *linear multibrackets.*

3.3   Linear Multibrackets

In the past two sections, we have looked at semibrackets, as well as formats with a consolation bracket, as examples of multibrackets. Let's back up a bit from specific examples, however, and ask what information we can learn about arbitrary multibrackets. One potential question to ask is if the fundamental theorem of brackets, which held for traditional brackets and semibrackets, holds for multibrackets as well. But before we can do that, we need to define what a multibracket signature and proper multibracket seeding might look like.

This is trickier than it seems: for arbitrary multibrackets, there isn't a natural generalization of signatures and properness. But there is a subset of multibrackets for which these notions generalize, allowing us to examine the fundamental theorem as it applies to this subset. These multibrackets are called *linear multibrackets*.

Definition 3.3.1: Linear Multibracket

A *linear multibracket* is a multibracket that can be arranged into a sequence of semibrackets such that

(a) If a team loses in a given semibracket but is not eliminated, they are sent to a later semibracket, and

(b) Each team that wins the nth semibracket finishes in mth place, where m is the sum of the ranks of the first n semibrackets.


A linear multibracket can then be easily imbued with a signature derived from the signatures of the semibrackets in the sequence.

Definition 3.3.2: Linear Multibracket Signatures

If the sequence of semibrackets that comprise a linear multibracket have signature $\mathcal{A}_1, ..., \mathcal{A}_k$ then the linear multibracket has signature $\mathcal{A}_1 \to ... \to \mathcal{A}_k$.

Let's confirm that the multibrackets from the previous section are indeed linear. First, the 2015 AFC Asian Cup.

Figure 3.3.3: The 2015 AFC Asian Cup

1
A1
8
B1
4
A2
5
C1
3
A3
6
B2
2
A4
7
B1
D1
B2

1st: Winner C1
2nd: Loser C1
3rd: Winner D1
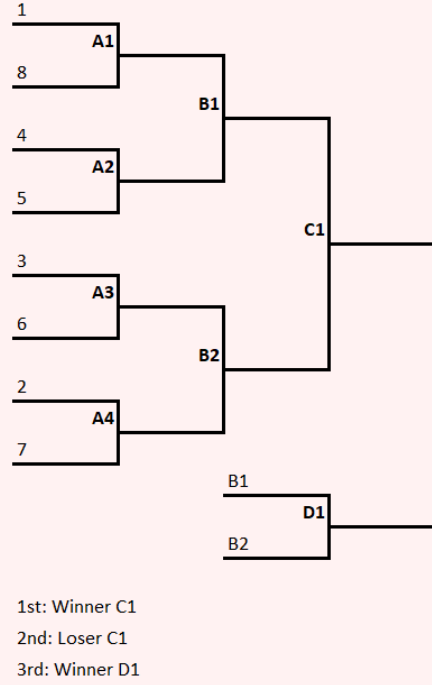
Looking at Figure 3.3.3, it can be tempting to say that the 2015 AFC Asian Cup is a linear multibracket with signature $[[\mathbf{8}; \mathbf{0}; \mathbf{0}; \mathbf{0}]] \to [[\mathbf{2}; \mathbf{0}]]$. But this is not quite right: The format with this signature would give second place to the winner of **C2** (as the winner of the second bracket), while outright eliminating the loser of **C1** (as a team that did not win any bracket). But in fact, was want to give second place to the loser of **C1**, and then third place to the winner of the consolation bracket with signature $[[\mathbf{2}; \mathbf{0}]]$. We can do this by adding a second bracket with signature $[[\mathbf{1}]]$ while sliding the bracket with signature $[[\mathbf{2}; \mathbf{0}]]$ to third.

Thus in total, the 2015 Asian Cup is a linear multibracket with signature $[[\mathbf{8}; \mathbf{0}; \mathbf{0}; \mathbf{0}]] \to [[\mathbf{1}]] \to [[\mathbf{2}; \mathbf{0}]]$. To make clear that the middle one-team bracket is included, we include it in the figure. This also allows us to drop the labeling of which teams finish in which place, as they are guaranteed by the linearity.

Figure 3.3.4: $[[8;0;0;0]] \rightarrow [[1]] \rightarrow [[2;0]]$
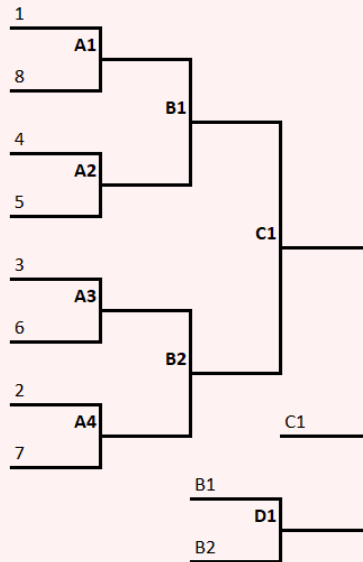
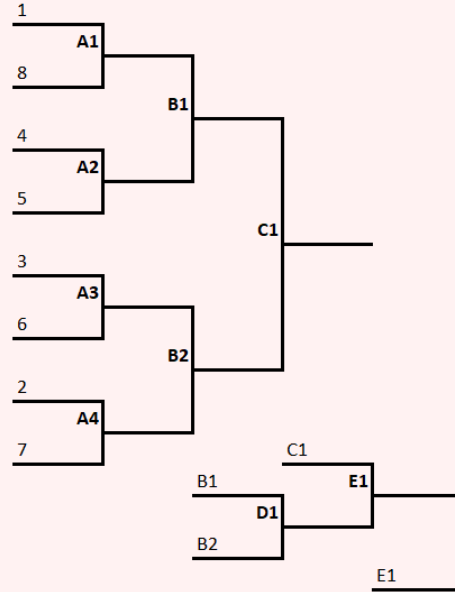Next, let's examine our alternative to the 2015 AFC Asian Cup.

Figure 3.3.5: The 2015 AFC Asian Cup Alternative

1st: Winner C1
2nd: Winner E1
3rd: Loser E1
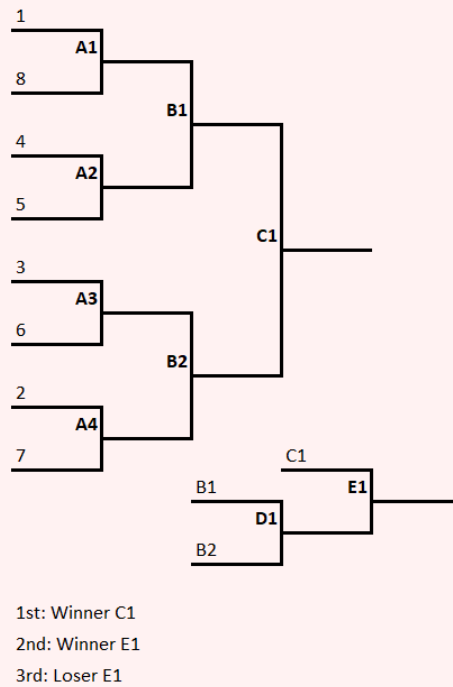
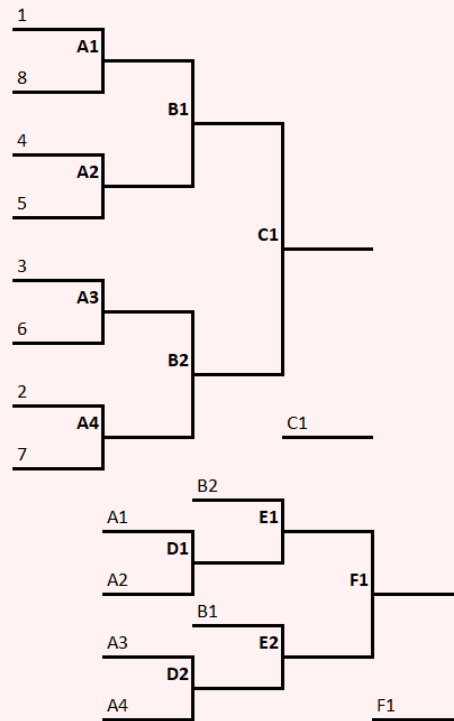Again, a quick look indicates a signature of $[[8;0;0;0]] \rightarrow [[2;1;0]]$. And while this sig-

nature would correctly assign a first- and second-place, it doesn't assign third-place. Instead, we need a signature of $[[8; 0; 0; 0]] \to [[2; 1; 0]] \to [[1]]$.

Figure 3.3.6: $[[8; 0; 0; 0]] \to [[2; 1; 0]] \to [[1]]$
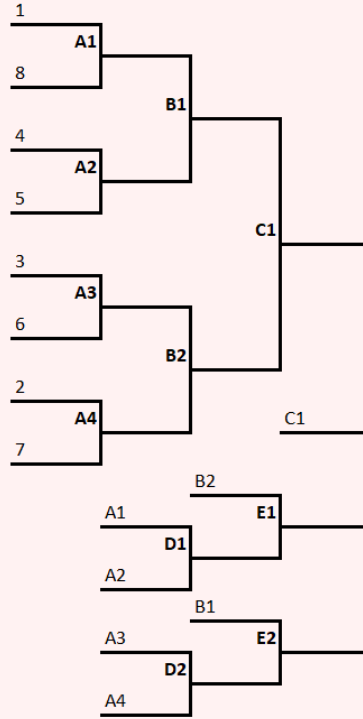

A similar analysis finds that the signature of the 2023 Wrestling Championships is $[[8; 0; 0; 0]] \to [[1]] \to [[4; 2; 0; 0]] \to [[1]]$.

Figure 3.3.7: $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$.

In all three examples so far, every semibracket has had rank one (that is, been a traditional bracket). However our final example, the 2023 Southern Conference Wrestling Championships Alternative, requires a semibracket of rank greater than one.

Figure 3.3.8: The 2023 Southern Conference Wrestling Championships Alternative

An attempt to give a signature of this format without the use of non-traditional semibrackets might be $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 1; 0]] \rightarrow [[2; 1; 0]]$. Unfortunately, this isn't quite the same format: it assigns third place to the winner of **E1** and fourth to the loser of **E2**. We want to treat both winners identically: this is the exact problem that semibrackets were developed to solve. Using semibrackets, we can see that the signature should be $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0]]_2$.

Figure 3.3.9: $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0]]_2$

This format is differentiated from the (admittedly a bit strange) format in which the winner of game **E1** comes in third and the winner of game **E2** comes in fourth by the lettering of the games: the fact that games **E1** and **E2** are both **E**-round games means they must come from the same semibracket. If games **D2** and **E2** were instead **F1** and **G1** respectively, then we would indeed have a linear multibracket of signature $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 1; 0]] \rightarrow [[2; 1; 0]]$.

Now that we have defined linear multibrackets and developed a notion of signature, we can turn to the other half of the fundamental theorem: properness.

3.4 Properness and Respectfulness

3.5 Efficient Linear Multibrackets

In the past few section, we have looked at multibrackets (and in particular linear multibrackets) as a solution to the tournament design question of how to crown a champion as well as give out certain consolation places.

We now consider a slightly different tournament design problem: we no longer care about which teams finish in first or any other specific place, only about which teams finish in the top-m for a particular m. This is a problem commonly faced at regional tournaments in which the top-m teams qualify for a national tournament: the ranking of the teams within the region aren't relevant, only which teams are above and below the cutoff.

Recall the format used in the 2023 Southern Conference Wrestling Championships.

Figure 3.5.1: 2023 Southern Conference Wrestling Championships
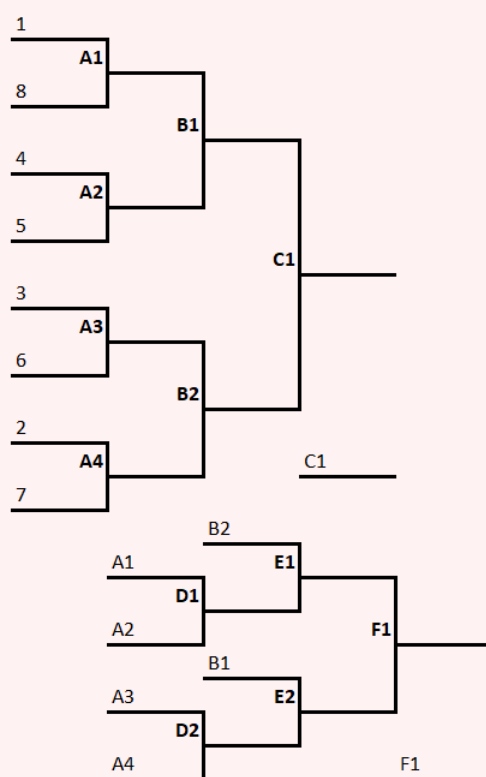
If we were only interested in the top four teams, rather than the rank of the team within those top four slots, games [[**C1**]] and [[**F1**]] become unnecessary: no matter what the results of those games are, the top four teams are the same. A more efficient format would leave those games unplayed, resulting in the following format.

Figure 3.5.2: An Efficient Format for Selecting a Top Four
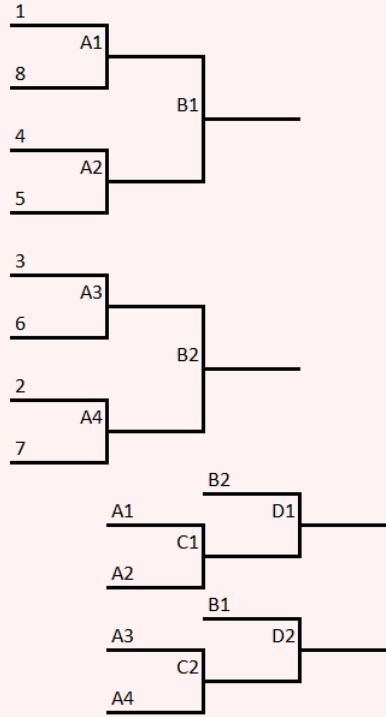
Instead of being composed of four traditional brackets, the format in Figure 3.5.2 is composed of two semibrackets each of which have rank two: one with the **A** and **B** round games, and one with the **C** and **D** round games. And, as desired, there no games played between two teams such that both the winner and loser of each of those games finish in the top four.

This format has signature $[[\mathbf{8}; \mathbf{0}; \mathbf{0}]]_2 \to [[\mathbf{4}; \mathbf{2}]]_2$, and we say that it is *weakly efficient*.

Definition 3.5.3: Weakly Efficient

A linear multibracket is *weakly efficient* if there are no games played within it such that both the winner and loser of that game are guaranteed to be ranked by the format.

Identifying whether a proper linear multibracket is weakly efficient can be done just by looking at its signature.

Theorem 3.5.4

A proper linear multibracket with signature $\mathcal{A} = \mathcal{A}_1 \to \dots \to \mathcal{A}_k$ is weakly efficient if and only if there is some integer j with $1 \leq j \leq k$ such that every semibracket $\mathcal{A}_i$ with $i < j$ is trivial and every semibracket $\mathcal{A}_i$ with $i > j$ is competitive.

Proof. Assume first that some j exists. Let **G** be a game. Because all semibrackets $\mathcal{A}_i$ with $i < j$ are trivial, **G** must be in a semibracket $\mathcal{A}_i$ for $i \geq j$, so the loser of **G** falls into a semibracket $\mathcal{A}_i$ for $i > j$, which is competitive. Thus the loser of **G** must play another game. If they keep losing, they will continue falling into competitive semibrackets, until they are eliminated outright and do not get ranked. Thus $\mathcal{A}$ is weakly efficient.

Assume now that no such j exists. Thus there must be some i such that $\mathcal{A}_i$ is nontrivial and $\mathcal{A}_{i+1}$ is noncompetitive. $\mathcal{A}_{i+1}$ is noncompetitive, so at least one team must win $\mathcal{A}_{i+1}$ without playing a game. Because $\mathcal{A}$ is proper, this team must be a championship game loser of $\mathcal{A}_i$, if such a game exists. $\mathcal{A}_i$ is nontrivial so such a game does indeed exist: let **G** be that game. Then the winner of **G** wins $\mathcal{A}_i$, and the loser of **G** wins $\mathcal{A}_{i+1}$, so $\mathcal{A}$ is not weakly efficient. $\qquad\square$

The USA Ultimate Manual of Championship Series Tournament Formats [8], which is used to determine the formats to be used at the various sectional and regional tournaments in the sport of ultimate frisbee, contains a host of weakly efficient linear multibrackets for selecting the top m teams out of a list of n for various values of m and n, after a "regular season" portion of the tournament has been played to establish seeds.

A couple examples are Figure 3.5.5, which selects a top six out of seven, and 3.5.6, which selects a top five out of sixteen in at most five games per team. (In reality, sometimes additional games are played to determine placements within the top-m, but we display only the weakly efficient part of the format here.)

Figure 3.5.5: $[[1]] \rightarrow [[1]] \rightarrow [[1]] \rightarrow [[4; 0]]_2 \rightarrow [[2; 0]]$
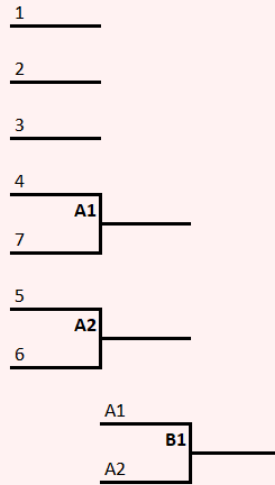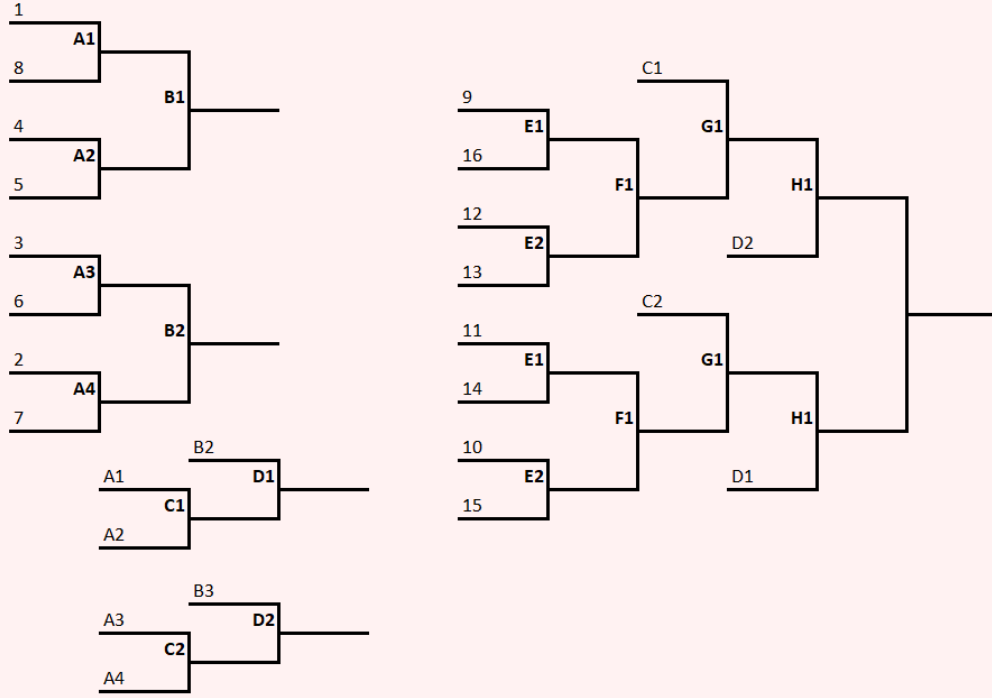
Figure 3.5.6: $[[8; 0; 0]]_2 \rightarrow [[4; 2; 0]]_2 \rightarrow [[8; 0; 2; 2; 0; 0]]$

We note two things about the notion of weak efficiency presented above. First, Theorem 3.5.4 implies that a weakly efficient multibracket can begin with a long string of trivial semibrackets before the nontrivial ones begin. While this is sufficient for avoiding playing unnecessary games, it does not completely remove unnecessary semibrackets: the set of leading trivial semibrackets

$$[[\mathbf{m_1}]]_{m_1} \rightarrow ... \rightarrow [[\mathbf{m_j}]]_{m_j}$$

of a weakly efficient multibracket can be combined into a single trivial semibracket

$$[[\mathbf{m_1} + ... + \mathbf{m_j}]]_{(m_1 + ... + m_j)}$$

without affecting which teams end up ranked. Applying this to the format in Figure 3.5.5 yields a signature of

$$[[\mathbf{3}]]_3 \rightarrow [[\mathbf{4; 0}]]_2 \rightarrow [[\mathbf{2; 0}]].$$

In fact, if there is at least one game played in a weakly efficient multibracket, trivial semibrackets can removed entirely, converting a multibracket of signature

$$[[\mathbf{m_1}]]_{m_1} \rightarrow [[\mathbf{a_1}; ...; \mathbf{a_r}]]_{m_2} \rightarrow ... \rightarrow \mathcal{A}_k$$

into one of signature

$$[[\mathbf{a_1}; ...; \mathbf{a_r} + \mathbf{m_1}]]_{m_1 + m_2} \rightarrow ... \rightarrow \mathcal{A}_k.$$

Applying this to the format in Figure 3.5.5 yields a signature of

$$[[\mathbf{4}; \mathbf{3}]]_5 \to [[\mathbf{2}; \mathbf{0}]].$$

To patch this, we strengthen the notion of weak efficiency into just *efficiency*.


Definition 3.5.7: Efficient

We say a proper linear multibracket is *efficient* if one of three conditions hold:

(a) It is a single trivial semibracket.

(b) It is a sequence of competitive semibrackets.

(c) It is a single nontrivial noncompetitive semibracket followed by a sequence of competitive semibrackets.

Theorem 3.5.4 says that in each of these three cases no games are played between teams guaranteed to be ranked, and the process detailed above can reduce any weakly efficient signature into a signature that takes one of the three forms in the above definition.

The second thing to note is that efficiency makes a lot of sense if we are only interested in the top-m teams (where m is the sum of the ranks of the semibrackets in our format) and not in the ranks of the teams within them. But sometimes we might be interested in the intermediate ranks as well. For example, let's say we want to design an eight-team tournament format in which the top team receives the grand prize, second-place receives a second-place prize, while the third- through sixth-place each get equivalent consolation prizes, and seventh and eighth each get nothing. While not efficient (or even weakly efficient), the following format assigns the desired places without playing any games between teams that are guaranteed to receive the same prize.
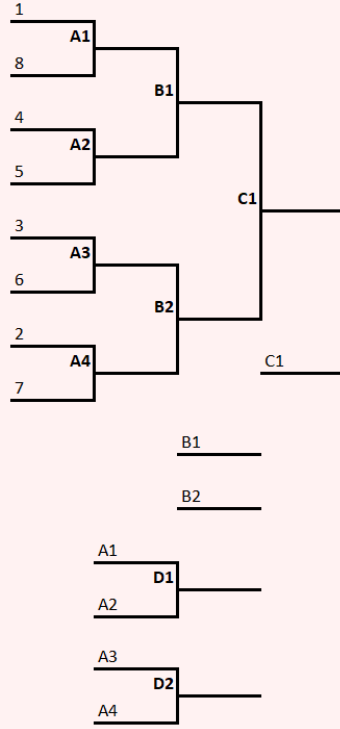
To account for this, we introduce the notion of a *prize structure*.

Definition 3.5.9: Prize Structure

A *prize structure* $\mathcal{P}$ is a sequence $(\mathbf{p_1}, ..., \mathbf{p_m})$ indicating that the top p_1 teams in a format receive some prize, the next p_2 receive some smaller prize, etc. Any teams finishing in place $1 + \sum_{i=1}^{m} p_i$ or worse receive no prize.

Then,

Definition 3.5.10: Efficient with Respect to a Prize Struture

We say a multibracket $\mathcal{A} = \mathcal{A}_1 \rightarrow ... \rightarrow \mathcal{A}_k$ is *efficient with respect to a prize structure* $\mathcal{P} = (\mathbf{p_1}, ..., \mathbf{p_m})$ if

- $\mathcal{A}_j$ being noncompetitive implies that for some $\ell < m$,

$$\sum_{i=1}^{j-1} \mathrm{Rank}(\mathcal{A}_j) = \sum_{i=1}^{\ell} p_i,$$

  and

64

- $\mathcal{A}_j$ being trivial implies that for some $\ell < m$,

$$\sum_{i=1}^{j} \text{Rank}(\mathcal{A}_j) = \sum_{i=1}^{\ell} p_i.$$

(The first condition ensures that the format is weakly efficient with respect to the prize structure, and the second condition ensures that there is no trivial semibracket that could be combined with another semibracket as per the process detailed before Definition 3.5.7.)

So the multibracket $[[\mathbf{8}; \mathbf{0}; \mathbf{0}; \mathbf{0}]] \to [[\mathbf{1}]] \to [[\mathbf{4}; \mathbf{2}]]_4$ is efficient with respect to the prize structure $(\mathbf{1}, \mathbf{1}, \mathbf{4})$. A multibracket $\mathcal{A}$ being efficient is the same as it being efficient with respect the prize structure $(\mathbf{m})$, where m is the sum of the ranks of its semibrackets.

Efficient multibrackets are great tournament designs for tournaments whose primary goal is to select the top m teams to move on to the next stage of the competitions, as discussed in the beginning of this section. They do so excitingly, with each spot in the top-m being awarded as the winner of a particular game; efficiently, with no games being played between teams who will receive the same prize; and fairly, as the multibracket rules ensure that winning is always better than losing. It is not surprising that many sports with regional tournaments that qualify teams for a national one use such formats.

4 References

[1] Robert Baumann, Victor Matheson, and Cara Howe. Anomalies in tournament design: The madness of march madness. *Journal of Quantitative Analysis in Sports*, 6(4), 2010.

[2] F. R. K. Chung and F. K. Hwang. Do stronger players win more knockout tournaments? *Journal of the American Statistical Association*, 73(363), 1978.

[3] H.A. David. *The method of paired comparisons*, volume 12. London, 1963.

[4] Christopher Edwards. *The Combinatorial Theory of Single-Elimination Tournaments*. PhD thesis, Montana State University, 1991.

[5] Jeff Horen and Raymond Riezman. Comparing draws for single elimination tournaments. *Operations Research*, 33(2), 1985.

[6] F. K. Hwang. New concepts in seeding knockout tournaments. *The American Mathematical Monthly*, 89(4), 1982.

[7] Nate Silver. When 15th is better than 8th: The math shows the bracket is backward, 2011.

[8] Eric Simon. The upa manual of championship series tournament formats, 2008.