Cedarville University Mathematics

Math 4800 - Capstone Experience in Mathematics

Combinatorial Game Theory

Taught by Dr. Hammett

Impartial Games and the Sprague-Grundy Theorem

by Noah Walker

December 5, 2022

Abstract

We will look at a specific class of combinatorial games called impartial games with the aim of developing a general method for analyzing them and finding a winning strategy for any impartial game under the normal play convention. We will define a natural equivalence relation on impartial games and show that its equivalence classes are the 1-pile games of NIM. This equivalence relation will allow us to prove the Sprague-Grundy Theorem. The Sprague-Grundy Theorem enables us to determine a winning strategy on a sum of games and helps us find winning strategies for many impartial games.

1 Introduction

People love to play games. Various board games and card games have been around for thousands of years, and there are hundreds of games played today ranging from chess to battleship to poker. While some games don't require much strategy to play, other games like chess and Go are very intense strategy games that people spend countless hours learning to play well. Although strategy games have been around for ages, it wasn't until recently that people tried to analyze some strategy games using math. In 1902, C. L. Bouton published a complete mathematical analysis of the game of NIM [2]. This rigorous treatment of NIM started a branch of mathematics called Combinatorial Game Theory.

While Bouton's paper began the field of Combinatorial Game Theory, for several decades afterwards, there wasn't a cohesive theory that could be applied to any game. Instead, each strategy game was analyzed in isolation using techniques specific to it. However, in the 1930s, the mathematicians R. Sprague and P. M. Grundy independently developed a technique, now called the Sprague-Grundy Theorem, that extended Bouton's analysis of NIM so that is could be used to analyze a large class of strategy games called impartial games [6, 10]. Impartial games are strategy games in which there is no distinction between the players. Another major development in Combinatorial Game Theory came in the 1970s, when John Conway devised a notation for strategy games that allowed for a rich *algebraic* analysis of all strategy games [3]. The goal of this paper will be to develop the notation and results needed to prove the Sprague-Grundy Theorem. Although we won't really delve into the theory that Conway developed, we will use a version of his notation because of its ubiquity in Combinatorial Game Theory today.

When analyzing strategy games, we are most often concerned with two questions: "Which player has a winning strategy?" and "What is the winning strategy?" Essentially, we want to know what the good moves in a game are and which player will win if they play according to those good moves. If we are trying to find good moves in a large, complex strategy game, we might start by just looking at a small portion of the game and considering it in isolation. If we can determine all the good moves for this part of the game and for all the other parts of the game, then we could use our knowledge of good moves on the individual parts of the game to determine what the good moves are on the game when considered as a whole. The Sprague-Grundy Theorem tells us how the good moves on the parts relate to the good moves on the whole, for a certain class of games.

Combinatorial Game Theory is the study of combinatorial games, or games of pure strategy. We will use the convention of writing the names of games in all caps, for example CHESS. We define combinatorial games as follows.

Definition. A *combinatorial game* is a game in which: [5]

1. there are 2 players who alternate turns
2. there are rules that describe the available moves
3. there is a set of possible positions
4. both players have complete information
5. the game must end in a finite number of moves

In combinatorial games, we often use the *normal play* convention, in which the last player to move wins. Since there are no random elements in combinatorial games, any game involving rolling dice, like MONOPOLY, or drawing cards, like POKER, is not a combinatorial game. BATTLESHIP is another example of a game which is not a combinatorial game since the point of the game is to guess the location of the other player's ships. Technically, CHESS is not a combinatorial game because it could simply loop forever. However, if we enforce a 50 move limit, then CHESS does become a combinatorial game.

TIC-TAC-TOE is a good example of a combinatorial game which most people are familiar with. One important combinatorial game, which we have already mentioned, is NIM. In the game of NIM, we have n piles of sticks where each pile can have any number of sticks. For example, in a 3-pile game of NIM, we might have piles of size 1, 3, and 4:

\ /\| |||/

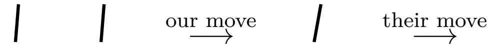
Moves in NIM consist of removing any number of sticks from a single pile. NIM follows the normal play convention, so the winner of NIM is the person who removes the last stick. Thus, in the above game, we might remove 3 sticks from the pile with 4 to get the game:

| \|| \

The next player could then make a smart move and remove all 3 sticks from the middle pile, making the game:

\ |

At this point, we can only take a single stick from one of the piles, and the other player will respond by taking the final stick, winning the game.



Another combinatorial game that will be important in this paper is CRAM. In this game, players take turns placing dominoes horizontally or vertically on a board such that each domino covers exactly two squares of the board. The board is generally square or rectangular, but it need not be so. CRAM follows the normal play convention. Figure 2 shows a complete game of CRAM on a 3x3 board. In that game, player 2 wins.

Figure 1: Two example positions in CRAM

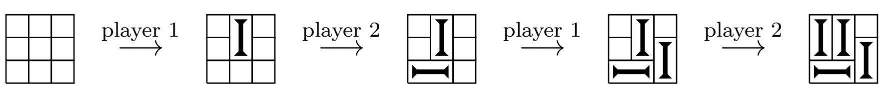


Figure 2: An example game of CRAM

While most of the board or card games that we are familiar with aren't combinatorial games, several of them, like CHESS, can be turned into combinatorial games with a simple modification. Other strategy games which have repeated positions, like GO, are very close to combinatorial games, so we can actually use combinatorial game theory to analyze them anyway. Furthermore, many combinatorial games are interesting in their own right and some, like HEX, have very serious competitions.

We said that we want to analyze large, complex games by considering parts of the game in isolation, but how do we break down these games into their parts? We answer this by considering the reverse question: "If we have parts of a game, then how do we combine them?" Note that each part of a game is really a game itself, so the question is how do we combine games. It turns out the answer is that we can combine games by playing them simultaneously. We call the result of this combining process, the *game sum* of these games. If we think about taking the game sum of 2 games of CHESS, then we are essentially placing 2 chessboards side-by-side and playing both games of CHESS at the same time. Some chess players have been known to play CHESS against multiple opponents at the same time, and a game sum is a similar idea to this concept.

CRAM is a game in which the concept of a game sum arises very naturally. If we consider the following game of CRAM, we can see that the placed dominoes naturally divide the board into two regions. Since we can only ever place a domino on one of these regions at a time, we can essentially

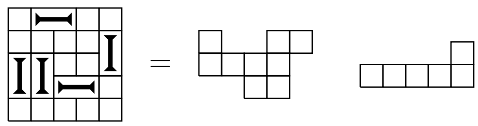


Figure 3: A game sum in CRAM

analyze the regions independently. Furthermore, each region is smaller than the original board, so they should be easier to analyze.

In general, the notion of a game sum allows us to decompose a game by finding the parts of a game that can be summed together to get the original game. Once we have these parts, we need to find good moves on each of them. To do this, we can make use of the following theorem.

Theorem (NIM Equivalence Theorem)**.** *Every impartial game is equivalent to a game of 1-pile NIM with some unique number of sticks in the pile.*

We will prove this theorem, and its proof will give us an algorithm for determining the number of sticks in the equivalent game of NIM. Since NIM has been solved by Bouton, we know what the good moves are on all games of NIM *and* on NIM game sums (a game sum of games of NIM). So, if we replace each of the parts of the game sum by their equivalent game of NIM, then we can use Bouton's analysis to find the good moves on this NIM game sum. The Sprague-Grundy Theorem states that the good moves on this NIM game sum are equivalent to the good moves in our original game! That is, the good moves in the original game are precisely the moves that will put us in a position equivalent to a good position in the NIM game sum that we have constructed. In summary, we have the following steps for applying the Sprague-Grundy Theorem to analyze a game.

1. Decompose the game into parts that sum together to give the original game.

2. Replace each of these parts in the game sum by their equivalent game of NIM.

3. Use Bouton's analysis of NIM to get the good moves in this NIM game sum.

4. Use the Sprague-Grundy Theorem to find the corresponding good moves in the original game.

2 Definitions and Terminology

To begin our study of combinatorial games, we establish some definitions, terminology, and notation that will help us to talk about combinatorial games in general. We state the definition of a combinatorial game here again for reference.

Definition 1. A *combinatorial game* is a game in which: [5]

1. there are 2 players who alternate turns
2. there are rules that describe the available moves
3. there is a set of possible positions
4. both players have complete information
5. the game must end in a finite number of moves

Again, combinatorial games are simply pure strategy games between two players. Because we are concerned with winning strategies when analyzing combinatorial games, we generally assume that both players play optimally. We typically distinguish the players by their turn order, so we have a first player and a second player. We say that the first/second player wins if the first/second player has a winning strategy. Saying that the first player wins does *not* mean that the first player will always win; it simply means that the first player *can* always win if they play correctly (i.e. optimally).

Notice that because a game must end in a finite number of moves, there can be no repeated positions. Otherwise, the game could get stuck in an infinite loop. It should also be noted that in the definition of a combinatorial game, while we state that a game must end in a finite number of moves, we do *not* require that there are only a finite number of possible positions. A player could have an infinite number of available moves from a given position.

As an example, consider the game POLYNOMIAL-NIM, which is played on polynomials with nonnegative integer coefficients. A move in POLYNOMIAL-NIM consists of reducing a coefficient and arbitrarily changing any coefficients of lower terms if desired. For example, if we start with the polynomial $2x^2 + x + 7$, then we could move to $x^2 + 75465x + 7$ or $2x^2 + 5896213574$, but not to $3x^2 + x + 7$. If we just start with x, then the first player to move has an infinite number of options because she can move to any nonnegative integer n. We leave it to the interested reader to verify that POLYNOMIAL-NIM is in fact a combinatorial game.

Some terms associated with combinatorial games are *positions* and *options*. Positions are the possible states which a game can be in. Note that positions are actually games themselves. For example, if we remove all the sticks from one of the piles in the following game of NIM, then we arrive at a position which is still a game of NIM, just a different game of NIM.

$$\backslash\backslash/ \qquad \backslash// \qquad \longrightarrow \qquad |\backslash\backslash$$

Because all positions are games and all games are positions, we will use these terms roughly interchangeably. An option of a game is a position (and thus a game) that can be reached from the game in a single move. So, the position in the above game of NIM would also be an option of the first game, while

$$\backslash$$

would not be an option, since it would require two moves to get there.

When we analyze combinatorial games, we generally assume the *normal play* convention. This convention says that the winner of a combinatorial game is the player who plays last (and thus the loser is the first player who can't play). We call this normal play because it is a very natural convention. As John Conway said, "Since we normally consider ourselves as losing when we cannot find any good move, we should obviously lose when we cannot find any move at all!" [3]

Another important property of a game is whether or not it is impartial; that is, whether or not we distinguish between the players.

Definition 2. An *impartial game* is a combinatorial game in which both players have the same options from every position. [5]

The games of NIM and CRAM would be impartial because we make no distinction between the players in those games. A game like CHESS would not be impartial, though. In CHESS, the white player can't move the black pieces, and the black player can't move the white pieces, so both players have different options from a position. The Sprague-Grundy Theorem only applies to impartial games under normal play, so from here on, we will assume that all games are impartial and use the normal play convention. Thus, when we say "game", we are really referring to an impartial game under the normal play convention.

While we have given a definition of a game, it is convenient to have a definition of a game in the language of sets, so that we can apply mathematics to games a little easier. Recall that in Definition 1, we stated that a game has rules that describe the available moves. If we treat the options of a game as a set, then we could say that a game has a set of options which can be moved to. However, recall that options and positions are themselves games. This discussion suggests that our mathematical definition of a game should include a set of other games which represent the options of that game. In fact, it turns out that this is all we need to completely describe a game!

Definition 3. A *game* is a set of games, which represent the options of the game.

While this definition may seem circular since we define a game in terms of games, it is well-defined, since $\emptyset$ is a game and a base case from which all other games can be built [3]. Now, we have seen that this definition meets part 2 of Definition 1 (there are rules that describe the available moves), but what about the other parts of the definition? It may seem like Definition 3 doesn't capture the idea of 2 players alternating turns. However, because games are impartial, each player has the same available moves at each position, so we don't need to specify which player is moving. In fact, we will see later that this is actually a very useful property.

Since a game is a set, its members must be well-defined. Therefore, part 4 of Definition 1 (both players have complete information) simply means that both players know what set the game is. That is, they know what game they're playing – something we would certainly hope is true! Part 3 (there is a set of possible positions) is satisfied because each possible position of a game G is contained at some nested level in G. That is, it is either a member of G or a member of a member of G and so on.

Additionally, the set nature of a game means that we can't have repeated positions (something we noted earlier as a property of games) since this would imply that a game contains itself in some way. This suggests that all games will end in a finite number of moves, and we formally prove that now.

Result 1. *All games end in a finite number of moves.*

Proof. Let G be a game. Since G is a set, we know from set theory (specifically the Axiom of Foundation) that there is no infinite descending chain of G_i's such that

$$G \ni G_1 \ni G_2 \ni G_3 \ni \cdots$$

Since play in a game proceeds by selecting an element $G_1 \in G$, then choosing an element $G_2 \in G_1$, and so on, this tell us that G can not continue infinitely, and thus it will end in a finite number of moves. **QED**

Thus, we can see that this set definition of a game satisfies all the same properties as our original definition of a game. Every set describes a unique game, and every game has a unique representation as a set.

Because there isn't a place in our set definition of games to specify whose turn it is, there is a potential source of confusion when working with arbitrary games. This is easiest to explain with an example. Consider the 2-pile game of NIM depicted below.

We shall call the player who moves first in this game Alice and the other player Bob. Suppose Alice moves to this option.

Now, in this new game with 2 sticks in each pile, the *player who moves first* is Bob, but the *first player* is Alice. When dealing with abstract games, this is an easy source of confusion. To help alleviate that confusion, whenever possible we will use first/second player *exclusively* to mean the player who moved first/second in the original game. However, in some cases this is not possible. In such cases, we will designate the first player in the original game as Alice and the second player as Bob. Then when we use first/second player, we will be sure to explicitly state what game we are considering these players to be first or second in.

As mentioned in the introduction, a vital tool in analyzing large, complicated games is the ability to decompose a game into smaller, independent parts. To do this, we need to formally develop the concept of the *game sum*. If we have two games G and H, then we use the notation, $G + H$ to denote the game sum of G and H. We call G and H the *summands*. The game sum of G and H is simply the game in which players on their turn can choose to move in either G or H according to the rules of that game, but not both. For example, if $G' \in G$ is an option of G, then a player could choose to make a move in G to G'. Then the whole game would become $G' + H$. Essentially, the game sum of G and H is the game where we play G and H concurrently, but limit players to making a single move on the summand of their choosing on their turn. Because we are using the normal play convention, the winner of a game sum is the last player to move on either of the summands. So the game ends when both of the summands have ended. We can formally define the game sum in the following way.

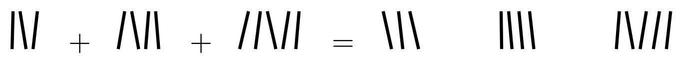
Definition 4. Let G, H be games. We define a sum of games (or game sum) recursively as the game

$$G + H = \{G' + H : G' \in G\} \cup \{G + H' : H' \in H\}.$$

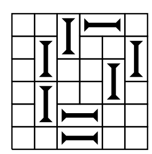
From both our intuition regarding game sums as playing games concurrently and from the formal definition, we can easily see that game sums are both associative and commutative, so $G + H = H + G$ and $(G + H) + J = G + (H + J)$ for all games.

Let's see how the idea of a game sum arises naturally in both NIM and CRAM. Consider three 1-pile games of NIM with 3, 4, and 5 sticks.

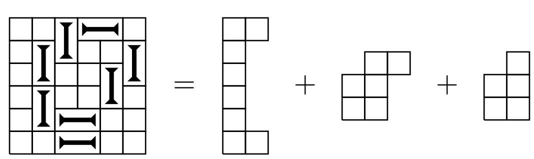
If we take the game sum of these three games of NIM, then we simply have the game in which players may remove any number of sticks from exactly one of the piles on their turn. However, that is exactly the same as playing a 3-pile game of NIM with piles of size 3, 4, and 5.



If we are given a game, we can try to decompose it by writing it as a game sum. Consider the following CRAM position (this is a game already in play, but remember that these are games too).



Note that the regions on this CRAM board are essentially independent. Since players can only make a move on one of the regions on their turn, we can write this game of CRAM as a game sum.



3 The Sprague-Grundy Theorem

The main purpose of the Sprague-Grundy Theorem is to allow us to determine a winning strategy on a game sum assuming that we know a winning strategy on the summands. In some cases, we can easily prove who will win a game sum and often the proof of this will give us a winning strategy. We now prove a result that tells us who wins a game sum in some cases.

Result 2. *If G is a second player win, then H and $G + H$ have the same winner.*

Proof. Call the first player Alice and the second player Bob.

Case 1: Suppose that Bob can win H. WLOG assume that Alice makes a move in G to some game $G' \in G$. Now Bob must make a move in the game $G' + H$. Since G is a win for Bob, there is a game $G'' \in G'$ that Bob wins. By making this move, the game becomes $G'' + H$, where both G'' and H are wins for Bob. Hence, Bob can continue using this strategy until the game ends, thus ensuring that he wins.

Case 2: Suppose that the Alice can win H, then there is a game $H' \in H$ that Alice wins. Suppose that Alice makes a move in H to H', now Bob must make a move in the game $G + H'$. However, in the game $G + H'$, since Bob starts, Alice is the second player. Since Alice can win H' and the second player (which is now Alice) can win G, by Case 1, $G + H'$ is a win for Alice, so Alice can win $G + H$ by moving in H to H'. **QED**

This result is very important and will be used frequently throughout this paper in various proofs. We see that this proof also provides a winning strategy for the player who can win H. Since game sums are symmetric, that is $G + H = H + G$, the only case left unaccounted for is when G and H are wins for the first player. Unfortunately, in this case the result is not so simple. Ultimately, we will need to invoke the Sprague-Grundy Theorem to fully solve this case. The motivated reader should verify this claim by finding a game sum that is a first player win whose summands are first player wins and a game sum that is a second player win whose summands are first player wins.

We next define an equivalence relation that will allow us to formally state and prove the NIM Equivalence Theorem, a key theorem in proving the Sprague-Grundy Theorem. The intuitive idea behind the relation is that we want games which are equivalent under this relation to behave like each other when used in a game sum. That way, in a game sum, we can restrict our attention to a nice representative from each equivalence class of this relation, namely games of NIM.

Definition 5. Let G, H be games. We say that $G \approx H$ if $G + H$ is a win for the second player.

Result 3. *The relation $\approx$ is an equivalence relation.*

Proof.

Reflexive: Consider $G + G$, which we show is a second player win. If the first player moves to $G' \in G$, then the second player can respond by moving to G' in the other game. Thus, we arrive at the game $G' + G'$. By continuing this strategy, the second player will always make the last move and thus will win. Hence, $G \approx G$.

Symmetric: Suppose that $G \approx H$. Since game sums are symmetric and the second player wins $G + H$, the second player also wins $H + G$, so $H \approx G$.

Transitive: Suppose that $G \approx H$ and $H \approx J$, that is, $G + H$ and $H + J$ are wins for the second player. Consider $(G + H) + (H + J)$. By Result 2, the second player wins this game. However, note that

$$(G + H) + (H + J) = G + (H + H) + J = (G + J) + (H + H)$$

is a second player win. Since $H \approx H$, the second player wins $H + H$. Since the second player wins $H + H$ and the second player wins $(G + J) + (H + H)$, by Result 2, the second player wins $G + J$, thus $G \approx J$. **QED**

At first glance, it may seem like $\approx$ is a rather arbitrary choice for an equivalence relation. Sure, the definition involves a game sum, but beyond that, why is this definition a natural choice when working with game sums? We will now establish some results that show why this relation works well with game sums.

Result 4. *If $G \approx H$, then G and H are wins for the same player.*

Proof. Suppose that $G \approx H$. Then $G + H$ is a win for the second player.
Case 1: Suppose that G is a win for the first player. If H were a second player win, then by Result 2, $G + H$ would be a win for the first player, which it is not. Therefore, H is a win for the first player.
Case 2: Suppose that G is a win for the second player. Since $G + H$ is a win for the second player, by Result 2, H is a second player win. **QED**

Result 5. *If $G_1 \approx G_2$ and $H_1 \approx H_2$, then $G_1 + H_1 \approx G_2 + H_2$.*

Proof. Suppose that $G_1 \approx G_2$ and $H_1 \approx H_2$, then $G_1 + G_2$ and $H_1 + H_2$ are wins for the second player. Therefore, by Result 2,

$$(G_1 + H_1) + (G_2 + H_2) = (G_1 + G_2) + (H_1 + H_2)$$

is a win for the second player, so $G_1 + H_1 \approx G_2 + H_2$. **QED**

Result 6. *Let G, H be games. Then $G \approx H$ iff for any game X, $G + X$ and $H + X$ are wins for the same player.*

Proof. $\Rightarrow$: Suppose that $G \approx H$. Given a game X, since $X \approx X$, by Result 5, $G + X \approx H + X$. Thus, by Result 4, $G + X$ and $H + X$ are wins for the same player.
$\Leftarrow$: Suppose that for all X, $G + X$ and $H + X$ are wins for the same player. Let $X = H$, then $G + H$ and $H + H$ are wins for the same player. However, $H + H$ is a second player win, so $G + H$ is a second player win. Thus $G \approx H$. **QED**

This last result states that if G and H are equivalent, then we can replace G with H in *any* game sum without changing the winner. This essentially means that G and H are indistinguishable within game sums. Hopefully, these results give a taste of the important and natural role of the $\approx$ relation. In fact, this relation is so natural for equating similar games within game sums that some books on game theory actually use $=$ to denote this relation and use $\cong$ for strict equality of games [1]. This relation is so fundamental that some use notation which elevates it over strict equality. However, we will not use such notation in this paper.

Since $\approx$ is an equivalence relation, we can look at the equivalence classes of $\approx$. By Result 2, we can see that all games which are second player wins form an equivalence class. As we have alluded to, the equivalence classes of $\approx$ (including the one just mentioned) can be characterized by 1-pile games of NIM. Because 1-pile NIM in particular is so important, when we mention NIM from here on, we will assume it is 1-pile NIM unless otherwise stated. We also make a brief aside now to develop some notation for NIM.

NIM can actually be played with any ordinal number of sticks in the pile, even if that ordinal is transfinite [3]. To see that such a game will end in a finite number of moves, we can consider the sequence $\{\alpha_i\}$, where α_i is the number of sticks left in the pile after i moves. Since each player removes some nonzero number of sticks,

$$\alpha_0 > \alpha_1 > \alpha_2 > \cdots,$$

which is a decreasing sequence of ordinals. Because there are no infinite decreasing sequences of ordinals [4], this sequence must be finite, so the game must end after a finite number of moves. We leave it as an exercise to the interested reader to show that the POLYNOMIAL-NIM game that we introduced earlier is actually just a game of NIM with an ordinal number of sticks.

We will use $*\alpha$ to denote the game of NIM with α sticks in the pile, where α is any ordinal. When $\alpha = 0$, there are no sticks, so there are no moves. Thus, $*0 = \emptyset$, which is a rather boring game. For $\alpha = 1$, the only move is to $*0$, so $*1 = \{*0\}$. Similarly, for $\alpha = 2$, we can move to $*0$ or $*1$, so $*2 = \{*0, *1\}$. For finite α, $*\alpha = \{*0, *1, \ldots, *(\alpha - 1)\}$ since we can reduce the pile to any smaller size. In fact, even for transfinite α, we can always reduce the pile to any smaller size, so the same pattern holds, $*\alpha = \{*\beta : \beta < \alpha\}$ [3]. For those familiar with ordinal numbers, this construction is nearly identical to the way ordinal numbers are defined. Indeed, in the same way that ordinals are the yardstick for sets, games of NIM are the yardstick for games.

We now establish a lemma about NIM that will help us prove the NIM Equivalence Theorem.

Lemma 7. *If $*\alpha \approx *\beta$, then $\alpha = \beta$.*

Proof. Let Alice be the first player and Bob be the second player. Suppose that $\alpha \neq \beta$ and WLOG suppose that $\alpha > \beta$. We show that $*\alpha + *\beta$ is a win for Alice and thus that $*\alpha \not\approx *\beta$. Since $\beta < \alpha$, $*\beta \in *\alpha$, so Alice can move to $*\beta + *\beta$. Since Bob must now start on the game $*\beta + *\beta$, Alice is the second player in the game $*\beta + *\beta$, so Alice will win the game. **QED**

Now we are in a position to prove the NIM Equivalence Theorem. In the proof of the following theorem, we will proceed by a form of induction called Conway Induction, which is named after John Conway, who made significant contributions to Combinatorial Game Theory [9]. In Conway Induction, we assume that the result holds for every game $G' \in G$, then show that the result holds for G. We present a formal proof of the validity of Conway Induction in the appendix. Although Conway Induction doesn't actually require a base case, we present one anyway to make it feel more like normal induction.

In the following proof we are trying to show that $G \approx *\alpha$. That is, $G + *\alpha$ is a second player win. The basic idea is that since every option of G is equivalent to a game of NIM (by Conway Induction), the second player can match any move in G with a corresponding move in $*\alpha$ and any move in $*\alpha$ with a corresponding move in G. Then after two moves, the game has essentially become $*\beta + *\beta$ for some β, which is a second player win.

Theorem 8 (NIM Equivalence Theorem)**.** *Every game is equivalent to a unique game of NIM.*

Proof. *(Adapted from [1])* Uniqueness follows immediately from the lemma.

We now prove existence by showing this statement by Conway Induction: "$G \approx *\alpha$, where α is the smallest ordinal such that no $G' \in G$ is equivalent to $*\alpha$." Note that such an α always exists because the ordinals are well-ordered and the options of G can't be equivalent to all the ordinals (since the ordinals form a class, not a set).

The base case is $G = \emptyset$, where $\alpha = 0$. Since $*0 = \emptyset$, clearly $G \approx *0$.

Suppose that the statement holds for all $G' \in G$. We show that $G \approx *\alpha$, where α is the smallest ordinal such that no $G' \in G$ is equivalent to $*\alpha$, by showing that $G + *\alpha$ is a win for the second player. The first player will either play in G or in $*\alpha$.

Case 1: Suppose that the first player moves in $*\alpha$, then they must move to $*\beta$ for some $\beta < \alpha$. Therefore, the second player must play in $G + *\beta$. Since α is the smallest ordinal such that no

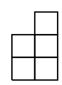
$G' \in G$ is equivalent to $*\alpha$ and $\beta < \alpha$, there is a $G' \in G$ such that $G' \approx *\beta$. By moving to G', the game becomes $G' + *\beta$. Since $G' \approx *\beta$, the second player wins $G' + *\beta$.

Case 2: Suppose that the first player moves in G to a game $G' \in G$, so the second player must make a move in the game $G' + *\alpha$. By the inductive hypothesis, $G' \approx *\beta$, where β is the smallest ordinal such that no $G'' \in G'$ is equivalent to $*\beta$.

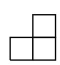
Subcase 2.1: Suppose that $\beta > \alpha$, then there is a $G'' \in G'$ such that $G'' \approx *\alpha$. Thus, the second player can move to $G'' + *\alpha$, which is a win for the second player since $G'' \approx *\alpha$.

Subcase 2.2: Suppose that $\beta < \alpha$, then $*\beta \in *\alpha$, so the second player can move to $G' + *\beta$, which is a win for the second player since $G' \approx *\beta$. **QED**

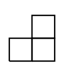
The NIM Equivalence Theorem tells us what the equivalence classes of $\approx$ are and the proof actually gives us a way to recursively compute the equivalence class of any game. We illustrate this process for the following game of CRAM.



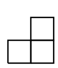
Now the options of this game are (up to symmetry)

 , $\square\square$, and $\square\square\square$

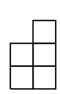
so we want to determine the game of NIM which each of these options are equivalent to. To do that, we use look at the options of these options and use the NIM Equivalence Theorem.

 $= \{\square\}$ $\qquad$ $\square\square = \{\emptyset\}$ $\qquad$ $\square\square\square = \{\square\}$

Each of these three positions has options that are $\emptyset$ or $\square$. Since there are no available moves from $\emptyset$ or $\square$, both of those games are equivalent to $*0$. Thus we can rewrite the games.

 $= \{\square\} \approx \{*0\}$ $\qquad$ $\square\square = \{\emptyset\} \approx \{*0\}$ $\qquad$ $\square\square\square = \{\square\} \approx \{*0\}$

For each of these three games, 1 is the smallest ordinal such that no option of the game is equivalent to $*1$. Hence, the NIM Equivalence Theorem tells us that all three of these games are equivalent to $*1$. Thus, every option of the original game is equivalent to $*1$ and no options of the original game are equivalent to $*0$. Therefore, by the NIM Equivalence Theorem

 $\approx *0,$

so this game is a second player win.

Because of this equivalence between games in general and games of NIM, we can define a function g on games such that $g(G) = \alpha$, where α is the unique ordinal such that $G \approx *\alpha$. That is, $g(G)$ is the smallest ordinal such that no $G' \in G$ is equivalent to $*g(G)$. This function g is called the Grundy function and the ordinal α is called the Grundy-value or NIM-value of G. Determining the NIM-value of a game is very useful for two reasons. First, the game is a second player win if and only if the NIM-value is 0, so the NIM-value tells us the winner of a game. Second, we can use the NIM-values of two games to determine the NIM-value (and thus the winner) of their sum! Hence, the NIM Equivalence Theorem gets us well on our way to stating and proving the Sprague-Grundy Theorem, but we still need one more piece: an operation on ordinals called the NIM-sum.

We use $\alpha \oplus \beta$ to denote the NIM-sum of two ordinals. The NIM-sum of two ordinals is just their bitwise XOR [8]. We describe what this means in detail now. There is a theorem in set theory that says every ordinal number has a unique binary representation with a finite number of 1s [3]. That is, the theorem says that every ordinal β can be written uniquely as

$$\beta = 2^{\alpha_1} + \cdots + 2^{\alpha_n}.$$

Now, to take the NIM-sum of two ordinals α and β, we first write them in binary. We then add the binary digits without carrying. So, if the digits are 0 and 1, the resulting digit is 1; otherwise, it is 0. Another way to say this is that we take the XOR of the binary digits or we add the binary digits mod 2. Let's do a few examples. First, we'll compute $6 \oplus 12$. Note that 0110 and 1100 are the binary representations of 6 and 12, respectively.

$$\begin{array}{r} 0\,1\,1\,0 \\ \oplus\, 1\,1\,0\,0 \\ \hline 1\,0\,1\,0 \end{array}$$

Thus, $6 \oplus 12 = 10$, which differs from the typical sum $6 + 12 = 18$. As another example, we'll compute $\omega \cdot 3 \oplus \omega + 1$, where $\omega = \{0, 1, 2, \dots\}$ is the first infinite ordinal. Note that $\omega + 1 = 2^\omega + 2^0$ and that $\omega \cdot 3 = 2^{\omega+1} + 2^\omega$.

$$\begin{array}{rccc} & & 2^\omega + & 2^0 \\ \oplus & 2^{\omega+1} + & 2^\omega & \\ \hline & 2^{\omega+1} + & & 2^0 \end{array}$$

So, $\omega \cdot 3 \oplus \omega + 1 = \omega \cdot 2 + 1$.

The following lemmas tell us that the ordinals under the NIM-sum operation form an Abelian group with identity 0 and $\alpha^{-1} = \alpha$. Because of this, we will use many of the basic properties of groups when working with NIM-sums without proving them or citing these lemmas.

Lemma 9. *The NIM-sum is associative and commutative.*

Proof. Since the NIM-sum just adds the binary digits of each ordinal mod 2 and addition mod 2 is associative and commutative, the NIM-sum is associative and commutative. **QED**

Lemma 10. *For any ordinal α, $\alpha \oplus 0 = \alpha$ and $\alpha \oplus \alpha = 0$.*

Proof. Since $x + 0 = x$ mod 2 and the NIM-sum just takes the binary digits of α and adds them mod 2, $\alpha \oplus 0 = \alpha$. Similarly, since $x + x = 0$ mod 2, $\alpha \oplus \alpha = 0$. **QED**

Now we are able to state the main theorem of this paper, the Sprague-Grundy Theorem.

Theorem (Sprague-Grundy Theorem)**.** *Let G, H be games, where $G \approx *\eta$ and $H \approx *\lambda$, then $G + H \approx *(\eta \oplus \lambda)$. That is, $g(G + H) = g(G) \oplus g(H)$. [6, 10]*

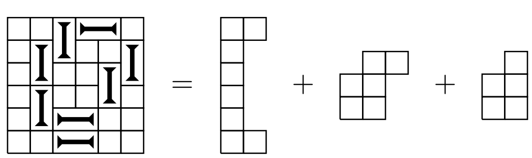
Before we prove this theorem, we want to look a little more deeply at what it tells us. The Sprague-Grundy Theorem shows us exactly how to determine the winner of a game sum. We calculate the NIM-value of its two summands, then compute their NIM-sum. If the NIM-sum is 0, then the second player wins. Otherwise, the first player wins.

If we look back at previous results, we can see that the Sprague-Grundy Theorem agrees with those results. For example, we showed that $G + G$ is a second player win, so $G + G \approx *0$, thus $g(G + G) = 0$. Using the Sprague-Grundy Theorem confirms this: $g(G + G) = g(G) \oplus g(G) = 0$. Consider Result 2 as another example. In this result, we showed that if G is a second player win, then H and $G + H$ have the same winner. Since G is a second player win, $g(G) = 0$, so by the Sprague-Grundy Theorem,

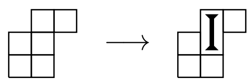
$$g(G + H) = g(G) \oplus g(H) = 0 \oplus g(H) = g(H).$$

Hence, not only do H and $G + H$ have the same winner, they actually have the same NIM-value!

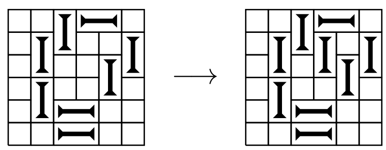
Let's take a look at an example of how we can use the Sprague-Grundy Theorem to determine a winning strategy in a game sum. Earlier, we looked at this CRAM board decomposition.



We also used the NIM Equivalence Theorem to determine that the last summand is equivalent to $*0$. A similar analysis reveals that first summand is equivalent to $*1$ and the middle summand is equivalent to $*3$. Therefore, by the Sprague-Grundy Theorem, the NIM-value of this CRAM position is $1 \oplus 3 \oplus 0 = 2$. Hence, the next player to move will win this game. To determine a winning move, we need to find a move that will cause the NIM-sum of these three summands to be 0. One way to do this is to decrease the NIM-value of the middle summand to 1. Then the resulting NIM-value will be $1 \oplus 1 \oplus 0 = 0$. We know this is possible because a NIM-value of 3 means that the middle summand has options with NIM-values of 0,1, and 2. A winning move for the next player on the middle summand would be



In the context of the original game, this winning move would be



By repeating this analysis, we could develop a complete winning strategy for this CRAM position. Notice how the Sprague-Grundy Theorem made it so easy to find a winning move once we knew the NIM-values of the game's summands, even though the game looked very complicated at the outset. Furthermore, once we compute the NIM-value for all the summands, we have all the information we need to easily find winning moves as the game progresses.

In the proof of the Sprague-Grundy Theorem, we will use Conway Induction like we did in the proof of the NIM Equivalence theorem. We will assume that the result is true for all options of $G + H$, then prove it is true for $G + H$ itself.

Theorem 11 (Sprague-Grundy Theorem). *Let G, H be games, where $G \approx *\eta$ and $H \approx *\lambda$, then $G + H \approx *(\eta \oplus \lambda)$. That is, $g(G + H) = g(G) \oplus g(H)$. [6, 10]*

Proof of the Sprague-Grundy Theorem. *(Adapted from [5])* To prove this theorem, we need to show that $g(G) \oplus g(H)$ is the smallest ordinal such that no $X \in G + H$ is equivalent to $*(g(G) \oplus g(H))$. Thus, we have to prove two things.

(1) For all $X \in G + H$, $g(X) \neq g(G) \oplus g(H)$.

(2) If $\alpha < g(G) \oplus g(H)$, then there is an $X \in G + H$ such that $g(X) = \alpha$.

We proceed by Conway Induction. The base case is $G + H = \emptyset$, so $G = \emptyset$ and $H = \emptyset$. Statement (1) is vacuously true. Since $G = H = \emptyset$, G, H are equivalent to $*0$, so $g(G) \oplus g(H) = 0 \oplus 0 = 0$. Thus, statement (2) is also vacuously true.

Suppose that the result holds for all options of $G + H$. We now prove (1). Given $X \in G + H$, WLOG suppose that $X = G' + H$ for some $G' \in G$. Then by the inductive hypothesis, $g(X) = g(G' + H) = g(G') \oplus g(H)$. Since $G' \in G$, $g(G) \neq g(G')$ by definition. Therefore,

$$g(X) = g(G') \oplus g(H) \neq g(G) \oplus g(H).$$

We now prove (2). Given $\alpha < g(G) \oplus g(H)$, note that $g(G) \oplus g(H) \neq 0$, so $g(G) \neq g(H)$. Therefore, WLOG suppose that $g(H) < g(G)$.

We first show that $\alpha \oplus g(H) < g(G)$. Let $\delta = \alpha \oplus (g(H) \oplus g(G))$ and let 2^γ be the largest power of 2 in the binary expansion of δ. Since $\alpha < g(H) \oplus g(G)$, the binary expansion of $g(H) \oplus g(G)$ must contain 2^γ and α's binary expansion must not. Similarly, since $g(H) < g(G)$, $g(G)$'s binary expansion must contain 2^γ. Hence, $\delta \oplus g(G) < g(G)$ since the 2^γ terms cancel when NIM-summed. Note that

$$\alpha \oplus g(H) = \alpha \oplus g(H) \oplus g(G) \oplus g(G) = \delta \oplus g(G) < g(G).$$

Therefore, $\alpha \oplus g(H) < g(G)$, so by the definition of g, there is a $G' \in G$ such that $g(G') = \alpha \oplus g(H)$. Since $G' + H \in G + H$,

$$\begin{aligned} g(G' + H) &= g(G') \oplus g(H) \qquad \text{(Inductive Hypothesis)} \\ &= \alpha \oplus g(H) \oplus g(H) \\ &= \alpha, \end{aligned}$$

so (2) is proved. **QED**

We can easily extend this formula to a game sum of n games using induction.

Corollary 12. *Let $G_1, \ldots, G_n$ be games. Then $g(G_1 + \cdots + G_n) = g(G_1) \oplus \cdots \oplus g(G_n)$.*

4 Appendix

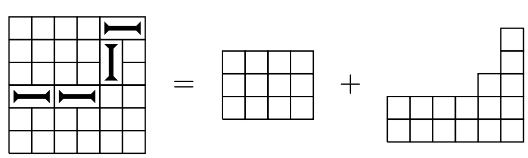
4.1 The Algebra of Games

The Sprague-Grundy Theorem and the NIM Equivalence Theorem tell us that g is basically an isomorphism from the equivalence classes induced by $\approx$ under the game sum operation to the ordinal numbers under the NIM-sum operation. Thus, the game sum should have similar properties to the NIM-sum, and indeed it does. Game sums are commutative and associative and NIM-sums are commutative and associative. Under NIM-sums, 0 is an identity, and under game sums, any game G which is a second player win (that is, $g(G) = 0$), is an "identity", since $H + G \approx H$ for any game H. Under game sums, every game G is its own "inverse" since $G + G \approx *0$ and $*0$ is an "identity". Under NIM-sums every ordinal α is its own inverse since $\alpha \oplus \alpha = 0$.

4.2 Limitations of the Sprague-Grundy Theorem

While the Sprague-Grundy Theorem is an incredibly powerful theoretical result, it is not as powerful in practice for a couple of reasons. First, not every game can easily be broken down into a sum of games. For example, the following CRAM positions don't have a simple decomposition.

Second, even when a game can be broken down into a sum of games, the pieces may be too complex to determine their NIM-value. For example, consider the following CRAM position.



Although this game can be written as a game sum, the summands are so large and complicated that their NIM-values can't be easily computed, so the Sprague-Grundy Theorem isn't very helpful in this case.

The Sprague-Grundy Theorem provides a great theoretical description of impartial games under normal play, but it is not applicable to games which aren't impartial. Additionally, the Sprague-Grundy Theorem isn't applicable to games which don't use normal play. For example, there is a category of games called misère games, in which the winner of the game is the first player who *can't* play. Essentially, in misère games, you want to force the other player to make the last move. In misère NIM, the goal then is to force the other player to take the last stick. While misère rules seem like a simple modification of games, the theory of misère games is considerably more complicated than the theory of impartial games. Under normal play, all impartial games are equivalent to some game of NIM, but under misère play, not even sums of games of NIM are necessarily equivalent to games of NIM! Although the Sprague-Grundy Theorem doesn't hold for misère play, there has been some successful research done on creating a local analogue of the Sprague-Grundy Theorem for certain subsets of misère games [7].

4.3 Conway Induction

We formally state and prove the principle of Conway Induction in the following theorem. Notice that Conway Induction requires no base case.

Theorem 13 (Conway Induction)**.** *Let P be a property games might have. Suppose that for any game G, if all $G' \in G$ have property P, then G has property P. Then every game has property P.*

Proof. *(Adapted from [9])* Assume to the contrary that there is a game G that doesn't have property P. We prove that there is an infinite sequence of games

$$G = G_0 \ni G_1 \ni G_2 \ni \cdots,$$

that don't have property P. However, all games must end in a finite number of moves, so this sequence is a contradiction.

We construct this sequence by induction on its length. We start with $G = G_0$, as indicated. Now, if we have constructed this sequence to n terms, $G_0 \ni G_1 \ni \cdots \ni G_{n-1}$, then G_{n-1} doesn't have property P. Therefore, by assumption, there is an option G_n of G_{n-1} that doesn't have property P. So, we can extend the sequence to $G_0 \ni G_1 \ni \cdots \ni G_{n-1} \ni G_n$. **QED**

References

[1] Michael H. Albert, Richard J. Nowakowski, and David Wolfe. *Lessons in Play: An Introduction to Combinatorial Game Theory*. 1st. 2007.

[2] Charles L. Bouton. "Nim, A Game with a Complete Mathematical Theory". In: *Annals of mathematics* 3.2 (1902), pp. 35–39.

[3] John H. Conway. *On numbers and games*. 2. ed. Wellesley, Massachusetts: A. K. Peters, 2001.

[4] Keith Devlin. *The joy of sets: fundamentals of contemporary set theory*. Second Edition. Undergraduate Texts in Mathematics. New York, NY: Springer, 1993.

[5] Thomas S. Ferguson. *Game Theory*. Self Published Classroom Notes. URL: `https://www.cs.cmu.edu/afs/cs/academic/class/15859-f01/www/notes/comb.pdf`.

[6] P. M. Grundy. "Mathematics and Games". In: *Eureka* 2 (1939), pp. 6–8.

[7] Thane E. Plambeck and Aaron N. Siegel. "Misère quotients for impartial games". In: *Journal of Combinatorial Theory, Series A* 115.4 (2008), pp. 593–622. URL: `https://arxiv.org/pdf/math/0609825.pdf`.

[8] Gal Porat. *The Sprague-Grundy Theorem*. URL: `http://math.uchicago.edu/~galporat/The%20Sprague-Grundy%20theorem.pdf`.

[9] Dierk Schleicher and Michael Stoll. *An Introduction to Conway's Games and Numbers*. Self-Published Classroom Notes. URL: `https://www.cs.cmu.edu/afs/cs/academic/class/15859-s05/www/lecture-notes/comb-games-notes.pdf`.

[10] R. Sprague. "Über mathematische Kampfspiele". In: *Tôhoku mathematical journal* 41 (1935), pp. 438–444.