Winning Strategies: The Emergence of Base 2 in the Game of Nim

Eric J. Friedman
School of ORIE
Cornell University
Ithaca, NY, 14850
ejf27@cornell.edu

Adam S. Landsberg
Joint Science Department
Claremont McKenna, Pitzer, and Scripps Colleges
Claremont, CA 91711
landsberg@jsd.claremont.edu

Introduction

Many players know that the secret to winning the game of Nim (and other "impartial" combinatorial games) is to write the sizes of the game's piles in base 2 and then add them together without carry. The proof of this well-known procedure (described below) is both straightforward and convincing. Nonetheless, the procedure still appears magical, as though a rabbit has been pulled out of a hat. Astute students (and frustrated professors) often ask why the winning strategy for such games involves base 2 and not some other base. After all, the number of piles in Nim is completely arbitrary – it can be 3, 11, or 500 – and there seems to be no inherent reason for the emergence of base 2. Minimal insight is offered by most published proofs, which themselves tend to either appear almost wizardly in nature (i.e., assume the base-2 method and show that it miraculously solves the problem) or employ combinatorial arguments that supply little abstract intuition (at least to the authors of this article).

However, as we will explain in this article, the reason for a base-2 based

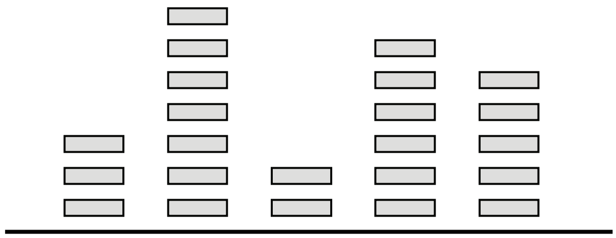

Figure 1: The game of Nim. Players alternate removing (an arbitrary number of) tokens from a pile (of a player's choosing). The game ends when no tokens remain, with the player who removed the last token the winner.

winning strategy in games like Nim follows directly from a well known result about abelian groups together with a mirroring argument. This argument also indicates why other bases do not naturally arise in these games, and has been extended to understand other classes of games (see [11]). Based on our own informal surveys (i.e., asking our colleagues), this argument does not appear to be known by the general mathematical public and even to a number of mathematicians who study combinatorial games!

The Game of Nim and Its Solution

The complete mathematical theory for the game of Nim (Figure 1) was presented by Charles Bouton [4] in 1901. This theory was extended to "impartial games" by Sprague [12] and Grundy [8] and then by Guy [9].

Every position (i.e., configuration) in Nim may be characterized by a vector of non-negative integers $x = (x_1, x_2, \ldots, x_n)$, with x_i denoting the number of tokens in pile i, and n the total number of piles at the start of the game. A legal move consists of choosing a pile $i \in \{0, 1, \ldots, n\}$ and reducing its height x_i to some new value x_i', where $0 \leq x_i' < x_i$. A player who has no legal moves available loses the game. Note that this occurs when all $x_i = 0$.

The well-known, but surprising, optimal strategy for Nim is based on writing out all the x_i's in base 2, then adding them together without carry (i.e., taking their bitwise XOR sum). For example, if the current position in a three-pile game of Nim is $x = (15, 5, 8)$, we write $x_1 = 15 = 1111_2$, $x_2 = 5 = 0101_2$, $x_3 = 8 = 1000_2$. Their binary sum without carry is $0010_2 = 2$. This number, known as the "nim-value" of the position, provides the key to defining a winning strategy: If the nim-value of the current position is non-zero (as in the example), then a player can win by making a move so that the resulting position has nim-value zero (it's straightforward to check that such a move is always possible.) Any move the opponent makes from this (zero-value) position will necessarily return the other player to some non-zero position. By iteratively following this procedure, a player starting at a non-zero position will always remain at non-zero positions, while the opponent is always forced into zero positions. Since the final (losing) position of the game (i.e. where no tokens remain) has nim-value zero, one sees that this is an optimal winning strategy. Simple!

However, the appearance of base 2 in the analysis can be quite surprising. Nothing about the game of Nim itself (i.e., the game rules, the configuration of the tokens, etc.) seems to give any obvious indication as to why binary representations should play such a crucial role in the winning strategy.

Impartial Games and Sprague-Grundy Theory

The game of Nim is not alone in this regard. Nim is a member of a large class of so-called *impartial* combinatorial games where base 2 plays a key role in the optimal winning strategy for little obvious reason. Examples of such impartial games include Kayles and Grundy's game. This class of games is distinguished by the fact that the set of allowed moves from any given position is the same for both players (in contrast, games like chess or checkers do not qualify as impartial since each player can only move his/her own pieces.)

To see how base 2 emerges in this broader context of impartial games, we begin by recalling a celebrated result by Sprague [12], Grundy [8] and Guy [9], who showed that every position in an impartial game can be assigned a nim-value that renders it equivalent to a single Nim pile of that size. (We note that while it may not be trivial to determine this assignment, such an assignment does exist and can be computed.) We will not concern ourselves here with how one explicitly computes the nim-value of a particular position in an impartial game; instead we simply note that it can be done.

That said, we next consider the notion of a so-called *disjunctive sum of games*. (While this idea may at first appear a bit contrived, its relevance will become apparent shortly.) Imagine one has a whole collection of impartial games $G_1, G_2, ...G_N$ (not necessarily all the same) all lined up in a row. Let's imagine playing a combined game (G) that involves all games in the collection in the following manner: On his/her turn, a player selects one game in the collection and makes a single legal move in that game. Play then switches to the other player, who also selects a game and makes one move. Play alternates between the two players in this fashion until eventually one player is unable to make a legal move in any game, and thus loses. Formally, this combined game G constitutes the *disjunctive sum* of the individual games, and is denoted $G = G_1 + G_2 + \ldots + G_N$.

Here now is the interesting feature which lies at the heart of the Sprague-Grundy theory and highlights the key role of base 2: Knowing the nim-values of the positions in each individual game allows one to determine a winning strategy for the disjunctive sum of games G, as follows. Look at the current position in each game G_i in G and note its nim-value (g_i). Then just re-express each nim-value in base 2 and add them all together without carry (just like we did with the Nim piles). The result (denoted g) is the nim-value of this sum of the games. As in Nim, if the nim-value g is non-zero, then a player can straightforwardly guarantee a win by always making moves to new positions in G whose nim-values are zero. Thus, once again we observe the appearance of base 2 (i.e., binary addition without carry) in the optimal winning strategy, this time in the more general context of disjunctive sums of impartial games.

As an aside, one might pause here to inquire about the relevance of the

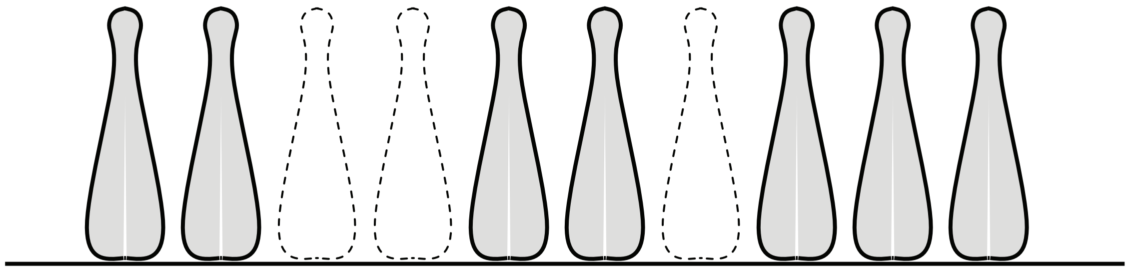

Figure 2: The game of Kayles.

notion of disjunctive sums of games – after all, how often does one actually play a collection of impartial games simultaneously? The answer is that disjunctive sums turn up more often than we might have guessed, since games sometimes naturally decompose into sets of smaller games. We have already seen one example of this, in fact, when we discussed Nim. One can think of an n-pile game of Nim as really being nothing more than the disjunctive sum of n individual single-pile Nim games. (This viewpoint turns out to be remarkably useful, as we shall see shortly.)

As a second illustration, consider another well-known impartial game, Kayles (Figure 2). In Kayles, N pins are lined up in a row, and players take turns knocking them down; in a given turn a player can either knock down a single pin or two adjacent pins ("adjacent" means side-by-side with no gaps in between). The player whose turn it is when all pins have been knocked down is the loser. How can we analyze such a game? Notice that as the players make their moves, the standing pins become separated into clusters (i.e., continuous sets of pins separated by gaps created by knocked-over pins). We can think of each cluster as constituting an individual Kayles game in its own right, and view the full game as the disjunctive sum of these individual games. So if one knows the nim-value for each cluster individually, one simply needs to do a base-2 addition without carry to determine the nim-value for the sum of games, and hence one has an optimal strategy.

In summary, not only Nim, but all disjunctive sums of impartial games, or impartial games that can be decomposed into a disjunctive sum, have nim-values that add as *binary addition without carry*. The question we would like to address here is "Why?".

Abelian Groups, Mirroring, and the Emergence of Base 2

The underlying reason for the emergence of base 2 in Nim and other impartial games is surprising and it was the authors' attempt to understand its origin that led to this paper. The key insight comes from realizing that there is a group structure underlying the Sprague-Grundy theory.

Before proceeding, think back for a moment to Bouton's analysis of Nim: one started with a set of individual piles (i.e., a disjunctive sum of games) and then associated to each one a particular nim-value, along with an operation (base-2 addition without carries) for adding these individual game values together. This directly led to an optimal strategy: from any non-zero position, always move your opponent to a zero-value position. With this in mind, our goal now is to understand the origin of base 2 in such impartial games. To begin, let's suppose that we didn't know about Bouton's solution of Nim or Sprague-Grundy theory, but wanted to find a simple way to compute (disjunctive) sums of impartial games. So what we seek is a mapping from $\mathcal{G} \to F$, where $\mathcal{G}$ is the set of all (finite) impartial games and F is a set with a binary operation $\oplus$ for adding games together that will allow us to compute an optimal strategy. (To stave off any possible confusion, however remote, please note that our use of the term "binary" here has nothing to do with base 2; it merely signifies that the operator $\oplus$ takes two inputs.)

What are some of the properties that the binary operation $\oplus$ and set F should have?

First, the operation $\oplus$ must clearly be abelian and associative, since the order in which we add games in their disjunctive sum cannot matter (e.g., if we associate $f_1 \in F$ with the current position of game G_1 and $f_2 \in F$ with the current position of game G_2, then the value we associate to the disjunctive sum of these two games can be expressed equivalently as either $f_1 \oplus f_2$ or $f_2 \oplus f_1$).

Second, there must exist an additive identity element $0 \in F$ (satisfying $f \oplus 0 = f$ for any $f \in F$). This identity element 0 is associated with the "null" game, i.e., a game in which no moves are possible (e.g., Nim with no tokens left). This follows from the observation that if you play any game in disjunctive sum with a null game, then it is equivalent to playing the game by itself.

The third and most crucial insight derives from a so-called *mirror strategy* argument, which will allow us to show that together $F, \oplus$ forms a group, with the special property that $f \oplus f = 0$ for every $f \in F$, – i.e., every element is its own inverse. This is the key to our analysis and leads directly to base 2.

The mirroring argument works as follows: First, suppose you have two identical impartial games, each in the same (but arbitrary) starting position. To each position we associate some $f \in F$ (the same f for each since the positions are identical). Now consider playing the disjunctive sum of these two games. Whatever move the first player decides to make in one of the games, let the second player make the identical move in the other game (hence the name "mirroring"). Continuing in this way, the second player is guaranteed to always have the last move in this sum of games, and hence will win.

With this in mind, we now make the following critical observation: Playing any game in $g \in G$ in isolation is effectively equivalent to playing that game in disjunctive sum with two identical copies of any other game in $g' \in G$, i.e., the outcome of the combined game $g + g' + g'$ is the same as that of the game g; the outcome is not affected by the copies. To see this, observe that if a player has a winning strategy in game g, then she should continue to play that strategy provided her opponent plays in g; if her opponent plays in one of the copies of g', then she should simply mirror that move in the other copy of g'. In this manner she is guaranteed a victory. Alternatively, if it is her opponent who has the winning strategy in g, then he should simply play that strategy in g and mirror his opponent's moves in g', thereby guaranteeing his victory.

Thus we see that, given an arbitrary game, adding on two identical copies of any other game has no effect on the outcome. This is precisely what we saw for the null game – i.e., adding the null game to any game g had no effect on that game's outcome – which led us to associate the null game with the additive identity element $0 \in F$. Hence we similarly associate any two identical copies of a game with the identity element, i.e., set $f \oplus f = 0$ for any f. (A nice correspondence can be made if we think back to the original game of Nim: if you take any two equal size Nim piles and add their values via binary addition without carry, you always obtain zero!)

From this, it immediately follows that $F, \oplus$ constitute an abelian group where every element is of order 2. If F were finite, then the remainder of our argument would follow directly from the fundamental theorem of finitely generated abelian groups, which states that G must then be a direct sum of additive groups of integers modulo k; however, since this group must be of order 2, all the k's must be 2 (since if there were a $k > 2$ then there would be an element of that order, and not of order 2.) When F is infinite, then the conclusion is still guaranteed since any group of bounded order (i.e., where there is a bound on the orders of all group elements) is a direct sum of cyclic groups [10].

Recall that the additive group of integers mod 2 is simply 1-bit binary addition without carry, so the above characterization can be interpreted as F being the binary sum mod 2 without carry of integers. Thus, we get the basic structure of the Sprague-Grundy theory from a decomposition theorem about groups and the mystery of the base 2 is revealed.

In summary, base 2 arises in optimal strategies for impartial games because anytime you have a disjunctive sum of games it gives rise to a group structure whose elements are all forced to be of order 2 by mirroring!

Other Considerations

Representations

Note that while we showed that the sums of impartial groups must be based on base-2 addition without carry, we did not show that the mapping from the set of games $\mathcal{G}$ to F must agree with traditional Nim-values. This is because this is not necessary.

For example, it is easy to see that the standard mapping from Nim piles to base 2 numbers can be permuted, e.g., there is no reason one can't swap the 2 place and the 8 place in binary, writing 9 as 0011 instead of 1001. However, one can actually do more.

For example, take the first 2 bits of the binary expansion and represent 0 by 00, 1 by 01, 2 by 11 and 3 by 10. We can extend this to the full mapping m by only changing these 2 bits in the standard Nim mapping.

We note that there are many other representations and leave it as an exercise to the reader to enumerate them.

Base 3 and Semigroups

One common question people have after learning the base-2 based strategy for Nim is whether any games use base 3 instead. In fact, there are a few known games which depend on base 3. However, these games seem to have base 3 "built in." For example in the game of Turnips (or Ternups) [3] a player turns over 3 coins at a time and the Nim values can be computed from base 3 expansions. In a similar vein, one can play a modified version of Nim wherein players can only remove at most two tokens from a pile at each turn. In this game, one computes the nim-value of a single pile using base 3, but combines piles using standard base 2 nim addition. (Note that if we modify nim so that a player can take at most m tokens then we use base $m + 1$ to compute the nim-values of a single pile.)

In more general settings, our analysis suggests that such games are the exceptions and in order to get base 3 we need a structure in which the “sum” of three identical games is always zero (analogous to our earlier finding that $f \oplus f = 0$). Such a structure does not appear to arise naturally in the present context.

However, one possibility along these lines relates to so-called misère games, in which the player who makes the last move loses, rather than wins. The general theory of sums of misère games is somewhat problematic and has resisted analysis for many years. Recently, however, Plambeck [11] has made significant progress by analyzing the structure of the semigroup which arises. Note that the sum of impartial misère games does not form a group as inverses need not exist. (In this case mirror strategies lose.)

Recommended Reading

For those interested in learning more about the mathematics of combinatorial game we recommend the two classics, which are each challenging in their own ways: Berlekamp, Conway and Guy [3] which is a rough and tumble encyclopedic joyride, and Conway [5] which is an exemplar of elegant mathematical analysis. For a gentler introduction see the excellent new textbook by Albert, Nowakalski and Wolfe [2]. The website [1] is devoted to algebraic approaches to misère games.

However, there is much in combinatorial game theory beyond the well behaved theory we have sketched. Complex computational issues are very important in many games. A nice review of these is given by Fraenkel in [6].

The authors of this article must admit a predilection for games which are computationally hard to solve and defy solutions based on the traditional mathematical theory discussed in this paper. A physics-based approach that we have been exploring is discussed in [7].

Acknowledgment We thank Elwyn Berlekamp, R. Keith Dennis, Richard Guy and Thane Plambeck for helpful comments and suggestions. EJF’s research has been supported in part by the NSF under grant ITR-0325453.

References

[1] Misère games website. http://miseregames.org/.

[2] M. Albert, R. Nowakowski, and D. Wolfe. Lessons in Play: An Introduction to the Combinatorial Theory of Games. *AK Peters*, 2007.

[3] E. R. Berlekamp and J.H. Conwayand R.K. Guy. *Winning Ways for Your Mathematical Plays.* Academic Press, London, 1979.

[4] C.L. Bouton. Nim, a game with a complete mathematical theory. *Ann. of Math.: Series 2*, 3:35–39, 1902.

[5] J. H. Conway. *On Numbers and Games.* AK Peters, Ltd, Natick, MA, 2000.

[6] A.S. Fraenkal. Scenic Trails Ascending from Sea-Level Nim to Alpine Chess. *Games of No Chance: Combinatorial Games at MSRI, 1994*, 1996.

[7] E. Friedman and A. Landsberg. Nonlinear dynamics in combinatorial games: Renormalizing chomp. *Chaos*, 17(2):023117, 2007.

[8] P. Grundy. Mathematics and games. *Eureka*, 2:6–8, 1939.

[9] R. K. Guy and C.A.B. Smith. The g-values for various games. *Proc Cambridge Philos Soc*, 52:514–526, 1956.

[10] I. Kaplansky. Infinite abelian groups, revised edition. *Ann Arbor, University of Michigan Press*, 1969.

[11] T.E. Plambeck. Taming the wild in impartial combinatorial games, INTEGERS. *Electr. J. of Combinat. Number Theory*, 5, 2005.

[12] R. Sprague. Uber mathematische kampfspiele. *Tohoku Mathematical Journal*, 41:438–444, 1936.