The Big Picture

More Games of No Chance
MSRI Publications
Volume 42, 2002

Idempotents Among Partisan Games

ELWYN BERLEKAMP

ABSTRACT. We investigate some interesting extensions of the group of traditional games, $\mathcal{G}$, to a bigger semi-group, $\mathcal{S}$, generated by some new elements which are idempotents in the sense that each of them satisfies the equation $G + G = G$. We present an addition table for these idempotents, which include the 25-year-old "remote star" and the recent "enriched environments". Adding an appropriate idempotent into a sum of traditional games can often annihilate the less essential features of a position, and thus simplify the analysis by allowing one to focus on more important attributes.

1. Introduction and Background

We assume the reader is familiar with the first volume of *Winning Ways* [Berlekamp et al. 1982], including Conway's axiomatization of $\mathcal{G}$, the group of partesan games under addition, which can also be found in [Conway 1976]. I now call the elements of this group *traditional games*. Each of the traditional games considered in this paper has only a finite number of positions. The identity of $\mathcal{G}$ is the game called 0, which is an immediate win for the second player. We investigate some interesting extensions of the group $\mathcal{G}$ to a bigger semi-group, $\mathcal{S}$, generated by some new elements which are idempotents in the sense that each of them satisfies the equation $G + G = G$. We also present an addition table for these idempotents.

Some of these idempotents have long been well-known in other contexts. The newer ones all fall into a class I have been calling *enriched environments*. A companion paper [Berlekamp 2002] shows how these idempotents prove useful in solving a particular hard problem involving a gallimaufry of checkers, chess, domineering and Go.

We begin with a review of definitions, modified slightly to fit our present purposes.

Moves. In Go, a *move* is the change on the board resulting from the act of a single player. In chess, this is commonly called a *ply*, and the term *move* is used to describe a consecutive pair of plys, one by White and one by Black. In

this paper, we use *move* as it is understood in Go. This is consistent with the tradition of combinatorial game theory. This theory has been most successful in analyzing positions which can be treated as sums of subpositions which are relatively or completely independent of each other. Each of these subpositions, as well as their sum, is called a *game*. Although the players alternate turns, it is quite common that they may elect to play in different components, so that within any particular game, the same player may make several consecutive moves. Hence, the definition of a game, or any of its positions, *does not* include any specification of whose turn it is to play next. Conway's traditional definition of a game is written recursively as

$$G = \{G^L \mid G^R\},$$

where G^L and G^R are sets of previously defined games. G^L is the set of positions to which Left can move immediately. These are also known as Left followers. G^R is the set of Right followers.

Other axioms, with which the reader is assumed to be familiar, define *sum*, *negative*, *greater-equal*, and *number*.

Outcomes. When played out, traditional games eventually yield outcomes. There are two outcomes: Loutcome, which is the outcome if Left plays first, and Routcome, which is the outcome if Right plays first. In the most general case, either of these outcomes might assume any of the values LEFT, TIE, or RIGHT, which Left prefers in the order LEFT > TIE > RIGHT. Ties and draws are impossible within $\mathcal{G}$, but they can occur in some of the extensions we will consider. In $\mathcal{G}$, play eventually terminates when the player to move is unable or unwilling to do so. That happens to Right when the value of the position is a nonnegative integer, or to Left when the value of the position is a nonpositive integer. If the value of the position is 0, whichever player is next to move is the loser.

Left plays to attain an outcome he prefers, while Right tries to thwart it. Loutcome and Routcome are the results if both players play optimally. If both

$$\text{Loutcome}(G) \geq \text{Loutcome}(H) \quad \text{and} \quad \text{Routcome}(G) \geq \text{Routcome}(H),$$

we say that

$$\text{Outcomes}(G) \geq \text{Outcomes}(H).$$

Greater-Equal. Combinatorial games satisfy a partial order. One form of the traditional definition of *greater-equal* states that

$$G \geq H \iff \text{For all } X,\ \text{Outcomes}(G+X) \geq \text{Outcomes}(H+X) \qquad (1\text{–}1)$$

If H and X have negatives, this is equivalent to the assertion that

$$\text{Outcomes}(G-H) \geq \text{Outcomes}(0)$$

But since Loutcome(0) = RIGHT, that half of the condition is trivially satisfied so a sufficient condition is that

$$\text{Routcome}(G - H) = \text{LEFT},$$

or, as more commonly stated, Left, playing second, can win on $G - H$.

Following Conway's original axioms, we say that

$$G = H \iff G \geq H \text{ and } H \geq G$$

and that

$$G > H \iff G \geq H \text{ but } H \not\geq G$$

Scores. For some purposes, it is convenient to define *scores* and work with them rather than with outcomes.

Play of any traditional combinatorial game must eventually yield a position whose value is a number. The value of the first such position is called the game's *score* or *stop*. If Left plays first and G is optimally played, the resulting number is called the Leftscore, denoted by Lscore(G). Similarly, if Right plays first and G is optimally played, the resulting number is Rscore(G). If G is any traditional game and x is any number, then

$$x > \text{Lscore}(G) \Rightarrow x > G,$$
$$\text{Lscore}(G) > x > \text{Rscore}(G) \Rightarrow x \text{ is confused with } G,$$
$$\text{Rscore}(G) > x \Rightarrow G > x.$$

It is known that if G is any game, then an optimal Left strategy for playing G ensures reaching a maximal score, and an optimal Right strategy for playing G ensures reaching a minimal score. However, the converse need not be true because several strategies might lead to the same score and some of them might yield a suboptimal outcome. This is due to the fact that when the score of a traditional game is 0, the outcome depends on who gets the last move.

Some real games have other scores, which are explicitly defined by the rules of the game. Go is such a game. It is an initially surprising and somewhat remarkable fact that these *official scores* imposed by any of several variations of the *official* Go rules are often identical to these *mathematical scores*. By appropriate choices of rules for "Mathematical Go" [Berlekamp and Wolfe 1994], we can attain agreement of scores in all but a few very rare positions, which are so exotic that different variations of the official Go rules then fail to agree with each other.

Dots-and-Boxes is another popular game in which scores are explicitly defined by the rules of the game. This pencil-and-paper game has very little in common with Go. But surprisingly, it again happens that the elegant mathematics of combinatorial game theory, when applied to an approximation of the popular game, yields decisive insights into how to play the popular game [Berlekamp 2000b].

In this paper, we treat *scores* in the mathematical sense: the value of the first position whose value is a number.

We next consider several idempotent games that have no negatives.

2. Definitions of Idempotents with Opening Ceremonies

Remote Star. The remote star ☆ is introduced in *Winning Ways*, Chapter 8 and plays a leading role there. Rather than rely on any of those results for a definition, I now propose a rule for playing

$$Y + ☆$$

Before moving on such a game, we require an *opening ceremony* during which each player submits a positive integer to the referee as a sealed bid. (From the mathematical perspective, there is no need for these bids to be sealed; public bids would work equally well. However, professionals and other serious competitors are loath to let the opponent know anything about their contingency plans prematurely. So it is easier to sell mathematical models to serious game-players when the rules ensure that losing bids remain unknown to the winning bidder.)

The referee selects the larger bid, n, and replaces ☆ by $*n$ before play begins. Ordering relations are determined in the usual way, using (1–1), with the understanding that the game X is specified before the opening ceremonies.

The play of $Y+☆+☆$ begins with two successive auctions. Since either player can submit a bid to the second auction which is at least double the winning bid of the first auction, it follows that the sum of two remote stars is now again a remote star, whence

$$☆ + ☆ = ☆.$$

Ish. Traditionally, the term *ish* means *Infinitesimally SHifted.* It appears in such expressions as $\{1 \mid 1\} = 1* = 1\,\text{ish}$ and $\{1* \mid 1\} = 1\,\text{ish}$.

We can also manipulate ish as though it were another element of our semigroup $\mathcal{S}$. To this end, we henceforth treat ish as a noun with the mnemonic Infinitesimal SHift. Its ordering relations might be defined as

$$G \leq H + \text{ish} \text{ and/or } G + \text{ish} \leq H \iff \text{Scores}(G+X) \leq \text{Scores}(H+X) \text{ for all } X,$$

However, to simplify the task of defining the sum of ish plus other idempotents, I prefer the following more intricate definition of ish:

At the opening ceremony of $G + X + \text{ish}$, each player submits a small positive number as a sealed bid to the referee, who announces the smaller such bid, which we will call ε. Then Left wins only if the score exceeds ε, and Right wins only if the score is less than $-\varepsilon$. A game which concludes with a net score of magnitude not exceeding ε declared to be a tie.

It is not hard to show that Left, going first, is able to *win* the game $G+X+\text{ish}$ if and only if $\text{Lscore}(G + X) > 0$. Left, going second, is able to *win* the game

$G + X +$ ish if and only if Rscore$(G + X) > 0$. Furthermore, it is easily verified that

$$\text{ish} + \text{ish} = \text{ish}\,.$$

Comment. *The sophisticated reader will recall several types of numbers that appear in [Conway 1976]: surreal, real, rational, dyadic rational. So when reading that each bid to determine ε must be a small positive number, she might ask which sort of number is required. It turns out that any type of number just listed is adequate for our present purposes. But only dyadic rationals are fully consistent with our focus on games with a finite number of positions.*

Positively Enriched Environment, $\mathcal{E}_t$. Enriched environments entail more elaborate opening ceremonies. Each positively enriched environment has a specified temperature, t, which is a positive number.

Every enriched environment contains an implicit ish, which is resolved first. This results in the specification of an ε. If other ishes are present, then this initial portion of the opening ceremonies continues until are all resolved into a single small positive ε. Then, to resolve the positively enriched environment $\mathcal{E}_t$, each player submits to the referee another small positive number as a sealed bid. The referee announces the winning (smaller) number, called δ. For simplicity, we restrict legal bids so as to ensure that δ is a divisor of t, so that t/δ is a positive integer. Then $\mathcal{E}_t$ is replaced by the sum of t/δ uniformly spaced switches, called *coupons*: $t|-t$, $(t-\delta)|(-t+\delta), \ldots, \delta|-\delta$. This concludes the opening ceremonies. Then play begins. Play terminates when all coupons have been taken and the value of the position is a number. The net score is then taken to be this number plus all of Left's coupons minus all of Right's coupons. The outcome is declared to be a tie unless the magnitude of the score exceeds ε.

Fully Enriched Environment, $\mathscr{E}_t$. The temperature of a fully enriched environment can be any number not less than -1.

After resolving the implied and explicit ishes to an ε, each player submits a small positive bid for δ. Legal bids are constrainted to ensure that $(t+1)/\delta$ is a nonnegative integer. The winning (smaller) bid is announced. Then $\mathscr{E}_t$ is replaced by a set of *coupons*, whose face values range from t down to $-1+\delta$, with a constant decrement of δ. A very large number of coupons with value -1 is then placed at the bottom of the coupon stack.[1]

Play begins. At each turn, a player may either make a legal move from the current game, G, to one of his legal followers, *or* he may instead use his turn to

[1]Mathematically, one could do without the coupons with value -1, because it is possible for either player to pick an extremely small value of δ. However, the -1 point coupons make it easier to sell the concept of coupon stacks to serious competitive gamesmen. A good environment to accompany a 10×10 game of Amazons needs almost 80 coupons with value -1, even though δ of 0.1 or even 0.5 proves interesting. If there were few or no -1 point coupons, the appropriate δ might need to be reduced by nearly two orders of magnitude.

take the top coupon from the stack. Even if the value of the current position, G, becomes a number, play continues until after only -1 point coupons remain. Play terminates after each of the players has taken three -1 point coupons consecutively. Then there is a concluding ceremony, during which the referee gives a special $-\frac{1}{2}$ point *terminal komi* coupon to the player who did not take the last of the six consecutive -1 point coupons which caused the game to be terminated. Each player's score is then computed as the sum of all of the coupons he has taken. If the magnitude of the difference between these scores exceeds ε, the player with the higher score is declared the winner. Otherwise, the result is declared to be a tie.

Comment. *Why do we not end the game until after the sixth consecutive coupon of value -1 is taken? In part, this is modeled after a well-known rule in chess, which declares the game to be terminated with a drawn outcome after the same position occurs three times with the same player to move. Presumably this is intended to give each player multiple chances to consider other options. Theoretically, if both players are playing optimally, one might think that the first repetition of a position would be sufficient. However, in games like Go, which include a ko rule, there are situations in which a pair of consecutive coupons is taken as a kothreat and its response while the game is still quite active. So we might theoretically weaken the six consecutive coupons to four consecutive coupons but, at least in the case when the board positions include Go, two consecutive coupons is definitely not enough.*

3. Definitions of Idempotents Without Opening Ceremonies

On. Figure 1 shows two positions of a Black checker king. Although White has no pieces to move, Black can move to and fro between these two positions whenever he so desires. I call these positions *onto* and *onfro*. From *onto*, Black can move to *onfro*. From *onfro*, he can move back to *onto*. Formally,

$$\mathbf{onto} = \{\mathbf{onfro} \mid \}, \qquad \mathbf{onfro} = \{\mathbf{onto} \mid \}.$$

Taken together, **onto** and **onfro** can be viewed as the bifurcated components of an abstract game called **on**, which appears in the latter half of Chapter 11 of *Winning Ways*. In a sum of games including an **on**, Right will never be able to

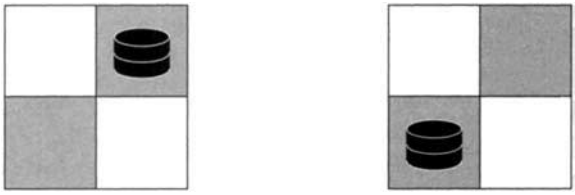

Figure 1. Onto (left) and onfro (right).

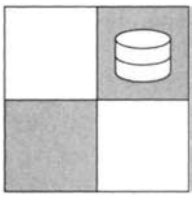

Figure 2. Off.

force the play of the game to terminate. Formally,

$$\mathbf{on} = \{\mathbf{on} \mid \}.$$

This game has infinite mean. If X is any traditional game, then our definition of equality implies that

$$\mathbf{on} + X = \mathbf{on}.$$

Off. **Off** differs from **on** only in that the lone checker king is now white instead of black. Now it is Right who can move to and fro at will. In a sum of games including an **off**, Left will never be able to prevent Right from playing. If Left has only a finite number of moves available elsewhere, he will eventually run out of options and lose the game.

This game has mean value of $-\infty$. Like **on**, it overpowers any traditional game to which it is added.

Dud. Although **on** and **off** superficially look like negatives of each other, every checker player knows that their sum is not zero (a win for the second player), but a *Deathless Universal Draw*, often realized on a single checkerboard when each player has only a single king, and the two kings are located in the double corners at opposite ends of the board.

Dud is not only a draw by itself, it also ensures that anything to which it is added also becomes a drawn game. **Dud** plays the role of a black hole. If a dud is present, no other features matter; the overall game behaves as a **dud**.

Comments on Outcomes with no Winner. Terminology has some history. In *Winning Ways*, we distinguished between *ties*, in which play terminated without either player winning, and *draws*, in which play could naturally be *drawn* out forever if not terminated by a special rule. In chess tournaments, ties due to stalemate and draws due to perpetual check are treated identically: each player receives $\frac{1}{2}$ win $+$ $\frac{1}{2}$ loss. However, Go has a quite different tradition. The natural rules allow certain positions to be repeated immediately in a two-move loop, one move by each player. Such global repetitions are universally banned by the so-called *Ko rule*. Positions involving ko are very common, occurring several times in most games. Lengthier repetitions, after 4 or 6 or more moves are also possible, but rather rare, occurring in only a small fraction of one percent of all professional games. The rules about how to handle such *superko* positions differ

from time to time and from place to place. Today, only the North American rules and the New Zealand rules simply ban all superkos. Japanese rules explicitly allow them. Chinese and Taiwanese rules for superkos are more complicated. In many cases, one player or the other is permitted to repeat the position but the other player is not. Which player is banned depends on the details of the position.

If a game gets hung in a superko, the Japanese tournament rules do NOT treat it as a tied or drawn outcome. The official translation defines the outcome as *no result.* So I prefer the English word *hung*, as in a hung jury or a computer program that is hung in a loop. Like the hung jury, a Japanese Go game which hangs in a superko often leads to a new game in which the same two contestants begin again from scratch.

To further complicate the situation, some amateur Japanese Go games can end with a tied score. Many translators have called such an outcome a *draw*, in direct conflict with the terminology of *Winning Ways.* I call them ties, and try to avoid any use of the word *draw* in reference to Go. As we are primarily concerned with individual games, or sums of games, the question of how such outcomes are treated in tournaments need not concern us here. So chess games which draw in perpetual check and checker games which draw in a **dud** might also be said to have *hung.* Whatever the terminology, a win is surely better than either a tied or hung outcome, which in turn is better than a loss.

Loony. Some impartial games have what are called *complimenting moves*, and such games can have positions with a fascinating value called *loony*, and denoted by the lunar symbol, ☽. One of the best-known such games is Dots-and-Boxes, whose Impartial variation I shall now describe.

The game is played on a array of dots located on the integer points of a rectangular subset of the Cartesan plane. These dots appear at unit distances from each other in vertical and horizontal rows. A legal move for either player is to draw a new horizontal or vertical line of length one, joining two dots. Unless that line completes a unit square, it completes the mover's turn. However, if that line completes one or two unit squares (called boxes), the mover must continue to make another move. The game ends when no further legal moves remain, and the player who made the last move *loses.*

(Even though last player loses, this is regarded as a *normal* rather than a *misère* rule because the last move necessarily completes a box, so the turn is incomplete. The player loses because he is unable to fullfill the requirement that he make another move.)

Figure 3, left, shows the position of an impartial Dots-and-Boxes position. This position can be viewed as the sum of four positions: two squares in the northwest, one in the northeast, two in the southeast, and four in the southwest, whose respective values can be shown to be ☽, $*$, $*$, and $*2$. The figures in the middle and on the right show two quite different ways in which the player to

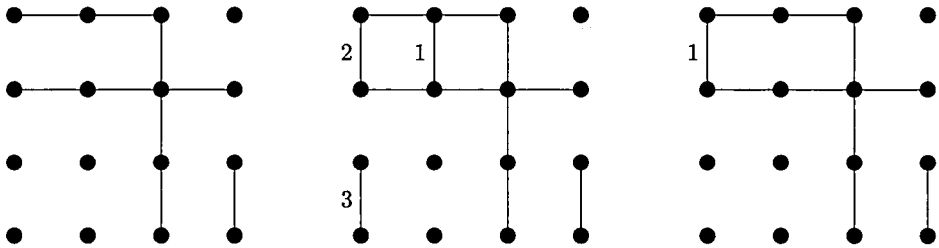

Figure 3. A Dots-and-Boxes position (left) and two possible continuations.

move can complete his turn. In the middle, he takes each of the boxes in the northwest and then completes his turn with a move elsewhere. In the right-hand figure, he completes his turn in another way, which forces his opponent to make the first move outside the northwest region. So although it might not be easy to determine whether or not one wishes to play first or second on the rest of the board, it is easily seen that the player to move from a sum which includes a loony component can win the game in either case. If he wants to make the first move elsewhere, he plays the loony component in a way which enables him to do that, as in the middle figure. On the other hand, if he wishes to force his opponent to make the first move elsewhere, he can achieve that objective by playing the loony component in the other way.

Loony values can also occur in games with entailing moves, as described in *Winning Ways*, Chapter 12. Entailing moves are special moves that require the opponent to move in a certain portion of the game. Complimenting moves are special moves (like completing a box in Dots-and-Boxes) which can be viewed as forcing the opponent to skip his next turn. Rather than attempt any general definition, for purposes of this paper it is sufficient to simply define $\mathfrak{D}$ very specifically as the northwest corner of the impartial Dots-and-Boxes position of the left panel in Figure 3.

4. The Addition Table

The addition table for the idempotents we have discussed is shown on the next page. The order in which they are listed may be viewed as an order of increasing *vim* (as in "vim and vigor"), with the understandings that **on** and **off** have equal vim, and that otherwise, whenever two of the idempotents are added, the one with the more vim predominates.

Another view is that adding in any of these idempotents destroys certain information, and the idempotents with more vim are more destructive.

In practice, adding in an appropriate idempotent can often be the key to the analysis of a particularly challenging position. The most helpful idempotent is the one which preserves just those features which are crucial to the winning line of play, while annihilating all of the less significant considerations.