IMPARTIAL GAMES AND SPRAGUE-GRUNDY THEORY — LECTURE NOTES

JONAS SJÖSTRAND

1. Combinatorial games

As opposed to classical game theory, combinatorial game theory deals exclusively with a specific type of two-player games. Informally, these games can be characterized as follows.

(1) There are *two players* who *alternate moves*.
(2) There are *no chance devices* like dice or shuffled cards.
(3) There is *perfect information*, i.e. all possible moves and the complete history of the game are known to both players.
(4) The game will eventually come to an end, even if the players do not alternate moves.
(5) The game ends when the player in turn has no legal move and then he loses.

The last condition is called the *normal play convention* and is sometimes replaced by the *misère play convention* where the player who makes the last move loses. In this course, however, we will stick to the normal play convention.

In general, we assume optimal play from both players, and when we talk about the *winner* of a game we have this assumption in mind.

2. Impartial games and the game of Nim

A game is called *impartial* if, from any of its positions, both players would have the same legal moves if they were about to play. An example of a game that is not impartial is chess, since white can only move white chessmen and black can only move black chessmen.

A typical example of an impartial game is the game of Nim, which is played as follows. On a table are a number of piles of sticks. In each move a player chooses one of the piles and removes one or more sticks from it. The player that removes the last stick wins.

If there is only one pile, clearly the first player wins by removing all sticks. If there are two piles things get slightly more complicated: If the piles contain the same number of sticks, the second player wins by mimicking the first player's strategy — when the first player removes some sticks from one of the piles, the second player immediately removes the same number of sticks from the other pile. If the piles contain different numbers of sticks, the first player wins after equalising the piles.

What if there are three or more piles? In 1901, Charles Bouton found the general strategy:

- For two nonnegative integers a and b, define the *nim sum* $a \oplus b$ as the bitwise XOR of a and b when they are written as binary numbers. For instance, if $a = 14 = (1110)_2$ and $b = 5 = (101)_2$ then $c = (1011)_2 = 11$.

Date: January 2011.

- If the nim sum of all piles is zero, the second player wins. If it is positive, the first player wins by a move that makes it zero.

It is not hard to see that this strategy works, but we will postpone the formal proof of this until later.

3. Axioms for impartial games

Since we are interested only in the abstract structure of games, we can regard any position P of an impartial game as being completely determined by the positions $P_1, P_2, \ldots$ reachable from P in one move. This observation suggests the following axiomatic system for impartial games.

Axioms for impartial games

1. Any set of impartial games is an impartial game.
2. There is no infinite sequence of games such that $G_1 \ni G_2 \ni G_3 \ni \ldots$

The members of an impartial game is called its *options* and should be interpreted as the positions that can be reached by a legal move. The second axiom guarantees that no game can be played forever. G together with its options and their options and their options, and so on, constitute the *positions* of G. If there are only finitely many positions, the game is *short*.

For integers $n \geq 0$, let $*n$ denote the game of Nim with a single pile of n sticks. As an example, let us see how $*3$ is represented as a set.

The options of $*3$ are $*2$, $*1$, and $*0$. The options of $*2$ are $*1$ and $*0$. The game $*1$ has only one option, $*0$, and the game $*0$ has no options at all so it is the empty set $*0 = \emptyset = \{\}$. We get

$$\begin{aligned} *3 &= \{ \quad *2, \quad *1, \quad *0 \quad \} \\ &= \{ \quad \{*1, *0\}, \quad \{*0\}, \quad \{\} \quad \} \\ &= \{ \quad \{\{*0\}, \{\}\}, \quad \{\{\}\}, \quad \{\} \quad \} \\ &= \{ \quad \{\{\{\}\}, \{\}\}, \quad \{\{\}\}, \quad \{\} \quad \}. \end{aligned}$$

Equivalently, we may represent any short impartial game G by a directed acyclic graph (DAG) where the vertices are positions in the game and there is an edge from u to v if $v \in u$, that is, if there is a legal move from u to v. The root of the DAG is the start position G.

4. Who wins a given impartial game?

In any impartial game, either the first or the second player to move will win the game. If the first player wins it is an $\mathcal{N}$-game (as in the *next* player) and if the second player wins it is a $\mathcal{P}$-game (as in the *previous* player).

Clearly, a game is an $\mathcal{N}$-game if and only if it has an option that is a $\mathcal{P}$-game. Applying this rule recursively yields an algorithm for computing the winner of any short impartial game.

The time complexity of this algorithm is linear in the number of possible moves in the game, which may be a huge number even for seemingly innocent games. As an example, the game of Nim starting with 100 piles of size 100 has $\binom{200}{100} \approx 9 \cdot 10^{58}$ positions and even more possible moves! This is far beyond what a computer can handle.

However, as we will see, there is a solution to this problem.

5. Sums of impartial games

For any two impartial games G and H, we define their *sum* $G + H$ as the game where G and H are played in parallel, and the player about to move must chose to make a move in *either* of G and H. As usual a player loses when he cannot make a move. Formally, $G + H = \{G' + H : G' \in G\} \cup \{G + H' : H' \in H\}$.

If we know who are the winners of G and H, can we tell who is the winner of $G + H$? Unfortunately, the answer is no. Here is an example: Let $G = *1$ be the game of Nim with one pile of size 1 and let $H = *1$ be (a copy of) the same game. Then G and H are $\mathcal{P}$-games and $G + H$ is an $\mathcal{N}$-game, namely the game of Nim with two piles of size 1. If we keep $G = *1$ but let $H = *2$ be Nim with one pile of size 2, then, as before, G and H are $\mathcal{P}$-games, but now $G + H$ is a $\mathcal{P}$-game too!

How much information do we need about G and H to be able to say who is the winner of $G+H$? The answer is: just a nonnegative integer called the *Grundy value* (or the *Sprague-Grundy value* or the *nim-value*) of the game.

6. Grundy values and Grundy's theorem

We define the *mex* (or *minimum excluded value*) of a set of nonnegative integers as the least nonnegative integer not in that set.

Now, the Grundy value $g(G)$ of a short impartial game G is defined recursively as the mex of the Grundy values of the options of the game. Example: $g(*0) = g(\emptyset) = 0$ so $g(*1) = g(\{\emptyset\}) = 1$. Exercise: Show that $g(*n) = n$ for any nonnegative integer n.

Theorem 6.1. *A short impartial game is an $\mathcal{N}$-game if and only if its Grundy value is positive.*

Proof. Clearly, a game G is an $\mathcal{N}$-game if and only if it has an option that is a $\mathcal{P}$-game, and it is clear from the definition that $g(G) > 0$ if and only if G has an option with Grundy value zero. The theorem follows by induction. □

Next, given the Grundy values of G and H, we will compute $g(G + H)$.

Theorem 6.2 (nim sum)**.** *For any short impartial games G and H, we have $g(G+H) = g(G) \oplus g(H)$, where $\oplus$ is the nim sum defined in section 2.*

Proof. By induction. An option of $G + H$ is either of the form $G' + H$ with $G' \in G$ or of the form $G+H'$ with $H' \in H$. By the induction assumption, an option of type $G' + H$ has Grundy value $g(G') \oplus g(H)$ and an option of type $G + H'$ has Grundy value $g(G) \oplus g(H')$, so, by definition, $g(G + H)$ is the mex of the set

$$S = \{g(G') \oplus g(H) : G' \in G\} \cup \{g(G) \oplus g(H') : H' \in H\}.$$

Clearly, $g(G) \oplus g(H)$ does not belong to S since $g(G') \neq g(G)$ and $g(H') \neq g(H)$ for any $G' \in G$ and $H' \in H$.

Let x be any nonnegative integer less than $g(G) \oplus g(H)$. We must show that x belongs to S.

For any nonnegative integer c, we denote by $[c]_k$ (for $k = 0, 1, \dots$) the binary digits of c, so that $c = \sum_{k=0}^{\infty} [c]_k 2^k$ and $[c]_k \in \{0, 1\}$.

Let m be the largest index such that $[g(G) \oplus g(H) \oplus x]_m = 1$. Since $x < g(G) \oplus g(H)$ we must have $[g(G) \oplus g(H)]_m = 1$ and $[x]_m = 0$. It follows that exactly one of $[g(G)]_m$ and $[g(H)]_m$ is zero; we may assume it is $[g(G)]_m = 0$. Then, $[x \oplus g(G)]_m = 0 < 1 = [g(H)]_m$ and for any $k > m$ we have $[x \oplus g(G)]_k = [g(H)]_k$.

We conclude that $x \oplus g(G) < g(H)$ which implies that there is an option $H' \in H$ with $g(H') = x \oplus g(G)$. This shows that $x = g(G) \oplus g(H')$ belongs to S. $\square$

We say that two short impartial games G and H are *equivalent* if, for any short impartial game K, the games $G + K$ and $H + K$ are either both $\mathcal{N}$-games or both $\mathcal{P}$-games.

Combining the two theorems above we obtain:

Theorem 6.3 (Grundy's theorem)**.** *Every short impartial game is equivalent to some one-pile Nim game.*

E-mail address: `jonass@kth.se`