HOW DIFFICULT IS IT TO INVENT A NON-TRIVIAL GAME?
Vladik Kreinovich, Richard Watson
Computer Science Department, University of Texas at El Paso, El Paso, TX 79968, USA

Abstract. Everyone who has ever tried to invent a new game knows that it is very difficult to make it non-trivial. In this paper, we explain this empirical fact by showing that almost all games (in some reasonable sense) are trivial. We also estimate the number of non-trivial games, and compare this number with the number of all possible games of a given size.

1. INTRODUCTION

Before we discuss the difficulty of creating a non-trivial game, we must first define what it means for a game to be non-trivial. Crudely speaking, we can divide the games into three major categories:

- Physical games, such as football, soccer, darts, horseshoes. These games are physically challenging and therefore non-trivial.
- Games of chance, like poker, monopoly, etc. While games of this sort usually require some skill, the winner is often the one who has the most luck. The emotional stimulation of these games is what makes them non-trivial and interesting to the players.
- Positional games: chess, checkers, go, and tic-tac-toe are examples of this type of games. Skill is required to excel at these games.

The games from the third category, called positional games (or combinatorial games, see. e.g., [Guy 1991]), are the ones that this paper deals with.

Many kids play tic-tac-toe, yet few adults play it. This is because if you use a very simple strategy, you can always assure yourself at least a draw. Once you know this, the game is no longer fun to play. It is trivial and therefore not interesting. Chess, on the other hand, is non-trivial because no one knows of a strategy that will guarantee that you won’t lose. So, we can say that a game is non-trivial if there does not exist a feasible algorithm that will guarantee a winning (or at least a drawing) solution.

People who have tried to make new games by slightly modifying the rules of existing non-trivial games (like chess) have found that the vast majority of these modifications turn out to be trivial and not interesting. One can therefore easily believe a legend that the inventor of chess received an enormous award for his invention.

This is an empirical feel. We would like to estimate the complexity of the game invention.
The most natural way to measure this complexity is to estimate what fraction of all possible games are non-trivial. If, for example, half of all games were non-trivial, it would be easy to create a new non-trivial game: one could simply take several games with randomly defined rules, then one of them would almost certainly be non-trivial. Conversely, if only a tiny fraction of all games are non-trivial, then it would be next to impossible to come up with a non-trivial game without some insight. We will show that the latter is indeed the case, i.e., the fraction of games that are non-trivial is very small.

Before we go into the formal descriptions, we want to make an important warning. The fact that a game is non-trivial does not necessarily make it interesting and attractive. For example, chess is considering by many an interesting game. One of the features of chess is that the rules of chess are very complicated (it takes some time to learn the rules). However, if someone invents a new game with rules that are as complicated as the rules of chess then it is probable that this complexity of rules will prevent this game from being played: a game should be fun, so why learn a long list of rules?

How many games are actually interesting (and not only non-trivial) is thus a tough question, the answer to which requires not only a mathematical analysis, but also a psychological analysis of the problem.

The main results of this paper were announced in [Kreinovich 1980, 1981, 1989] and [Kosheleva Kreinovich 1989].

**2. DEFINITION OF A GAME**

**ESTIMATING THE TOTAL NUMBER OF GAMES OF A GIVEN SIZE**

**DEFINITIONS.** By a *two-players positional game* (or, for short, a *game*) $g$ we mean a tuple

$$g = \langle n, k, p_0, M_1, M_2, W_1, W_2, L_1, L_2, D_1, D_2 \rangle,$$

where:

- $n, k$ are integers;
- $p_0$ is a function from \{1, 2, ..., $n$\} to \{1, 2, ..., $k$\} (such functions will be called *positions*);
- $M_1, M_2$ are the sets of pairs of positions, and
- $W_i, L_i, D_i$ are the sets of positions such that $W_1, L_1, D_1$ are mutually disjoint (i.e., all three pairwise intersections are empty) and the sets $W_2, L_2, D_2$ are also mutually disjoint.

The elements of this tuple will be called as follows:

- $n$ is called the *number of cells on the game board* (or simply *number of cells*, for short);
- $k$ is called the *number of figures*;
• elements of the set \(\{1, 2, \ldots, k\}\) will be called \textit{figures};

\(p(s)\), where \(p\) is a position, and \(s \leq n\), the cell, will be called the \textit{figure in cell} \(s\), according to position \(p\);

\(p_0\) is called \textit{initial position};

• pairs from the set \(M_1\) are called \textit{possible moves of the first player}; if \((p, p') \in M_1\), we will say that this move \((p, p')\) \textit{transforms} a position \(p\) into a position \(p'\);

• pairs from \(M_2\) are called \textit{possible moves of the 2-nd player};

• positions from \(W_i\) are called \textit{winning for the} \(i\)-\textit{th player}, positions from \(L_i\) are called \textit{losing for the} \(i\)-\textit{th player}, and positions from \(D_i\) are called \textit{drawing for} \(i\)-\textit{th player}.

\textbf{EXAMPLE.} In chess, \(n = 64\), figures correspond to the words “king”, “queen”, “pawn”, “empty”, etc. A position describes what figure is in what cell. The initial position, possible moves, winning, losing and drawing positions are determined by the rules of chess.

\textbf{DEFINITIONS.} A \textit{game protocol} is a finite sequence of positions \(p_0, p_1, p_2, p_3, \ldots, p_N\) such that:

• the pair \((p_0, p_1)\) is a possible move for the 1-st player, \((p_1, p_2)\) is a possible move for the 2-nd player, \((p_2, p_3)\) is a possible move for player 1, \((p_3, p_4)\) is a possible move for player 2, etc;

• if \(i < N\) and \(i\) is odd, then the position \(p_i\) is not winning, losing or drawing for player 1;

• if \(i < N\) and \(i\) is even, then the position \(p_i\) is not winning, losing or drawing for player 2.

We say that a game protocol is \textit{finished} if the position \(p_N\) is winning, losing or drawing for the correspondent player. If the game protocol is not finished, we say that it is \textit{unfinished}. If the protocol is unfinished, and \(N\) is even, then we say that the \textit{first player has the next move}; if \(N\) is odd, we say that the \textit{second player has the next move}.

If a game protocol finished, we say that the first player \textit{wins} if either \(N\) is odd, and \(p_N\) is a winning position for him, or \(N\) is even and \(p_N\) is a losing position for player 2. The second player \textit{wins} if the opposite is true. If \(p_N\) is drawing position for the current player, then the game protocol is called a \textit{draw}.

The term \textit{strategy of} \(i\)-\textit{th player} means a function that given an unfinished protocol in which \(i\)-th player has the next move, produces a next move. We say that a player has \textit{followed this strategy}, if all his moves are obtained by using such a function. We say that a strategy is \textit{winning} if and only if, regardless of all possible moves of the opposing player, a player following this strategy wins.
COMMENT. The players make moves in turn. They start with the position \( p_0 \). Then, the first player moves to some position \( p_1 \) such that the pair \((p_0,p_1)\) belongs to the set \( M_1 \) of possible moves. Then the second player makes a move, transforming the position into \( p_2 \), etc. If it is the turn of \( i \)-th player to move, and the current position is winning for him (i.e., belongs to his set \( W_i \)) then he wins and the game stops; if the current position is losing for him, then he loses and the game stops. If \( p_N \) is a drawing position for the corresponding player, then the game ends with a draw.

COMMENT. In order to estimate the fraction that non-trivial games form in the set of all possible games, we have to estimate the number of all games and the number of all non-trivial games. The total number of possible games is infinite, because we can take arbitrarily large \( n \) and arbitrarily large number \( k \) of figures. So, in order to get finite estimates, let’s limit ourselves to some \( n \) and \( k \).

DEFINITIONS. By \( G(n,k) \), we denote the number of all games with \( n \) cells and \( k \) figures.

The functions \( f(n,k,...) \) and \( g(n,k,...) \) are called asymptotically equivalent (denoted \( f \sim g \)) if their ratio \( f/g \) tends to 1 as \( n \to \infty, k \to \infty, ... \).

THEOREM 1. \( \ln(\ln(G(n,k))) \sim 2n \ln(k) \).

(For readers’ convenience, proofs are postponed until the last section).

3. DEFINITION OF A NON-TRIVIAL GAME.
HOW MANY GAMES ARE NON-TRIVIAL?

MOTIVATION OF THE FOLLOWING DEFINITION. In this Section, we want to estimate the number of non-trivial games. It is very difficult to formalize the notion of a non-trivial game, so we’ll use the axiomatic approach: namely, we’ll list natural properties of non-trivial games. These properties are not sufficient to determine for every game whether it is trivial or not, but they will enable us to estimate the number of non-trivial games.

The first property stems from the fact that if one of the players has a winning strategy that leads him to a victory in 1, 2 or 3 moves, then this game is surely trivial.

The more general remark is that if a winning strategy is implemented by a feasible algorithm, then we can use this algorithm and always win, and the game becomes trivial (like tic-tac-toe). On the other hand, if the winning strategy cannot be implemented by a feasible program, then every time we play this game we have to think, so the game
becomes non-trivial. In particular, if the length of the any program implementing the winning strategy is greater or equal than an exponential function of \( n \) or \( k \), then certainly none of these programs can be called feasible [Garey Johnson 1979], hence the game is non-trivial.

There is a precise notion in computer science that expresses the intuitive idea of the minimal length of the program generating the text \( x \): a Kolmogorov complexity \( K(x) \). This notion was invented independently by Solomonoff [Solomonoff 1964] and Kolmogorov [Kolmogorov 1965] (for further results see a survey [Levin 1984]). So, we can reformulate the above idea as follows: if the Kolmogorov complexity of the optimal strategy is greater than some exponential function of \( n \) and \( k \), then the game is certainly non-trivial. Thus, we come to the following definition.

**DEFINITION.** By a set of non-trivial games, we mean a subset \( I \) of the set \( G \) of all games that satisfies the following properties:

1) If in a game \( g \), one of the players has a winning strategy that leads him to victory in 1, 2 or 3 moves, then \( g \) does not belong to \( I \).

2) If for some game \( g \), the Kolmogorov complexity \( K(x) \) of every winning strategy is greater or equal than \( \exp(n^{1/2}) \), then \( g \) belongs to \( I \).

Games from \( I \) will be called non-trivial. The number of all non-trivial games with \( n \) cells and \( k \) figures will be denoted by \( I(n, k) \).

**THEOREM 2.** \( \ln | \ln(I(n, k)/G(n, k)) | \sim n \ln(k) \).

**COMMENT.** This means that the ratio \( I(n, k)/G(n, k) \) tend to 0 as \( n \) and \( k \) tend to infinity (i.e., almost all games are really trivial); moreover, the fraction of non-trivial games decreases as a doubly exponential function \( \exp(\exp(n)) \).

The asymptotic expression, given by the theorem, can be also expressed in the following way:

**COROLLARY.** \( \ln | \ln(I(n, k)/G(n, k)) |/ \ln \ln G(n, k) \sim 1/2 \).

**COMMENT.** In other words, if we denote \( G(n, k) \) by \( x \), then asymptotically, \( I(n, k) \sim x/\exp((\ln x)^{1/2}) \).

**GENERAL COMMENT.** The proof of this theorem is such that the choice of namely 3 moves or the choice of the concrete exponentially growing function \( \exp(n^{1/2}) \) is not essential: the theorem is still true, if we change it to any number of moves or any exponential function that grows slower than \( \exp(n) \).
4. PROOFS

General comment. These proofs are already not very long (and thus, to our viewpoint, not very hard). For readers who are professionally skillful in combinatorics we could skip a few steps and thus make the proofs even shorter. However, since we want our proofs to be understandable to all the readers, we present these proofs in all necessary detail.

Proof of Theorem 1. To define a game, we need the following:

- The initial position \( p_0 \).
- A subset \( M_1 \) of the set \( P \times P \) of all pairs of positions.
- A subset \( M_2 \) of the set \( P \times P \) of all pairs of positions.
- A subdivision of the set \( P \) of all positions into four subsets: winning positions \( W_1 \), losing positions \( L_1 \), draw positions \( D_1 \), and all the other (non-final) positions.
- A similar subdivision for the second player.

In the definition of a game, there is no connection between these five parts of a game, in the sense that regardless on the initial position we can choose an arbitrary subset \( M_1 \), etc. Therefore, the total number of games equals to the product of five numbers: the total number of initial positions, the total number of subsets \( M_1 \), the total number of subsets \( M_2 \), etc. So in order to compute \( G(n, k) \), let’s first compute these five numbers.

A position is a function from \( \{1, 2, ..., n\} \) to \( \{1, 2, ..., k\} \), or, in other words, the sequence of \( n \) numbers from 1 to \( k \); \( i \)-th term in this sequence represents the figure in cell \( i \). The total number of figures is \( k \), and there are \( n \) cells, so the total number \( A \) of possible positions is \( k \times k \times ... \times k \) (\( n \) times), or \( k^n \). In particular, there are \( k^n \) initial positions.

The number \( B \) of the pairs of positions equals to the square of the number of all positions, i.e. to \( (k^n)^2 = k^{2n} \). Therefore, the total number of all possible subsets of \( P \times P \) equals to \( 2^B \).

The subdivision of the set \( P \) of all positions into four sets can be expressed by a function \( f \) from \( P \) to \( \{1, 2, 3, 4\} \) such that for every pair \( x \), \( f(x) \) is the number of a subset, to which point \( x \) belongs (\( x = 1 \) means \( x \in W_1 \), \( x = 2 \) means that \( x \) is a losing position, \( x = 3 \) means that \( x \) is a draw position, and \( x = 4 \) means that \( x \) is a non-final position. So, the total number of such subdivisions is equal to the total number of such functions, i.e., it is equal to \( 4^A \).

Now, we are ready to compute \( G(n, k) \). The total number \( G(n, k) \) of games with \( n \) cells and \( k \) figures is equal to the product \( G(n, k) = k^n \cdot 2^B \cdot 2^B \cdot 4^A \cdot 4^A \), where \( A = k^n \), \( B = k^{2n} \).
Therefore, $\ln G(n, k) = n \ln(k) + B \ln(2) + B \ln(2) + A \ln(4) + A \ln(4) = n \ln(k) + 2k^{2n} \ln(2) + 2k^n \ln(4)$.

When $n$ and $k$ tend to infinity, then $k^{2n}$ grows faster than $k^n$ and $n \ln(k)$, so

\[
\ln \ln G(n, k) \sim \ln(2k^{2n} \ln(2)) = 2n \ln(k) + \text{const} \sim 2n \ln(k). \quad \text{Q.E.D.}
\]

**Proof of Theorem 2.** The number of positions $p$ is equal to $k^n$. This number is greater than $N = \exp(n^{1/2})$, so in this set of all positions, we can choose a subset of $N$ elements that are different from $p_0$. Let’s fix one of such subsets and denote its elements by $p_1, p_2, ..., p_N$. The set \{p_0, p_1, ..., p_N\} will be denoted by $P_1$. Let’s also denote the complement to $P_1$ (i.e. the set of all positions that are not in $P_1$) by $P_2$. We’ll contract games, for which the only winning strategy of the first player is to move to $p_{i+1}$ every time he is in $p_i$. So, let’s assume that:

- the only winning position for the 1-st player is $p_N$;
- $p_N$ is also the only losing position for the second player;
- all other positions from $P_1$ are neither winning, nor losing, nor draw for neither of the players;
- if the first player is in a position $p_i$ from $P_1$, then he can either move to a position outside $P_1$, or to $p_{i+1}$;
- if the first player is in a position from $P_2$, then he cannot move to any position from $P_1$;
- if the 2-nd player is in a position $p_i \in P_1$, then he can move only to $p_{i+1}$.

If the protocol follows the sequence $p_0, p_1, p_2, ..., p_N$, then the first player wins. The second player cannot move outside this sequence: if he is in a position $p_i$, he is forced by the rules of this game to move to $p_{i+1}$. The first player can move to a position from $P_2$, but then there is no way for him to return to $P_1$, and therefore, no way to come to his only winning position $p_N$. Therefore, the only strategy for the 1-st player is indeed to follow this sequence $p_i$.

It is known [Kolmogorov 1965], [Levin 1984] that for every $N$, there exists a word of length $N$ whose Kolmogorov complexity is $\geq N - c$ for some $c$ (moreover, for sufficiently big $c$, the majority of words of length $N$ have this property). So, we can choose a sequence $p_0, p_1, ..., p_N$ in such a way that its Kolmogorov complexity is $\geq N - c$. In this case, all these games are non-trivial. Let’s fix such a sequence. So, we fix the initial position $p_0$ and the set $P_1$. The total number $I_1(n, k)$ of such games is a lower bound for the total number $I(n, k)$ of all non-trivial games. Let’s estimate $I_1(n, k)$.
In order to define a game with the above-described properties, we have to fix:

- a subset of the set $P \times P_2$ that corresponds to possible moves of the 1-st player (except the moves to $p_{2i}$, that are fixed);
- a subset of the set of pairs $P_2 \times P_2$, that corresponds to the possible moves of the 2-nd player (except for the moves from and to $p_i$, that are also fixed);
- a subdivision of $P_2$ into positions that are losing, draw, and non-final for player 1;
- a subdivision of $P_2$ into positions that are winning, draw, or non-final for player 2.

Similarly to Theorem 1, the total number of these games is equal to the product of four numbers: the number of all possible subsets of $P \times P_2$, the number of all possible subsets of $P_2 \times P_2$, the number of functions from $P_2$ to \{1, 2, 3\}, and the fourth factor that is also equal to the number of such functions. Similarly to Theorem 1, we arrive at the following formula: 
$$I_1(n, k) = 2^C \cdot 2^D \cdot 3^E \cdot 3^F,$$
where $C = k^n \cdot (k^n - N)$, $D = (k^n - N)^2$, and $E = k^n - N$. Therefore, dividing the expression for $G(n, k)$ from Theorem 1 by this expression, we get 
$$G(n, k)/I_1(n, k) = k^n \cdot 2^{2N} \cdot 3^{2N - N^2} \cdot (4/3)^A \cdot 3^{-N} \cdot (4/3)^A \cdot 3^{-N}.$$ 
Taking logarithms of both sides of this equality, we come to the following formula:
\[
\ln(G(n, k)/I_1(n, k)) = Nk^n \ln(2) + 2Nk^n \ln(2) + 2k^n \ln(4/3) + 
\text{other terms that grow negligibly slower}
\]
(in the sense that their ratio with these main terms tend to 0).

The function $N = \exp(n^{1/2})$ grows faster than 1, so the main term in $\ln(G(n, k)/I_1(n, k))$ leads to
\[
\ln(G(n, k)/I_1(n, k)) \sim 3Nk^n \ln(2).
\]

Therefore, $\ln(\ln(G(n, k)/I_1(n, k)) \sim n \ln(k) + \ln(N) + \text{const}$; but $\ln(N) = n^{1/2}$ is asymptotically negligible in comparison with $n$, so
\[
\ln(\ln(G(n, k)/I_1(n, k)) = \ln |\ln(I_1(n, k)/G(n, k)| \sim n \ln(k).
\]

$I_1(n, k)$ is an estimate for $I(n, k)$ from below. Let’s now get an estimate from above. By the definition of a non-trivial game, every game that has a winning strategy for player 1, that leads him to a victory in one move, is trivial. Therefore, in a non-trivial game, the first player cannot win in his first move. So, the total number $I_2(n, k)$ of such games is an upper estimate for the total number $I(n, k)$ of non-trivial games. We’ll now prove that this upper bound satisfies the same asymptotic formula as the lower bound $I_1$, and therefore the same formula is true for the desired value $I(n, k)$ that lies between these bounds $I_1$ and $I_2$.

The ratio $I_2(n, k)/G(n, k)$ can be reformulated as a probability for a random game with $n$ cells and $k$ figures to belong to the set $I_2(n, k)$, if we consider all the games to be
equally probable (with probability $1/G(n, k)$). One can check that a random game in this sense can be obtained in the following way:
- take any position as an initial position $p_0$ with equal probability $1/k^n$;
- for every pair of positions, with probability $0.5$, include this pair into $M_1$, and with probability $1/2$, do not include it; the events corresponding to including or not including different positions are independent;
- repeat the same procedure for $M_2$;
- for every position $p$, include it into $W_1$, $L_1$, $D_1$, or do not include it anywhere with probability $1/4$; for different positions make this choice independently;
- repeat the same for the second player.

One can check that in this case, the probability of every game is precisely $1/G(n, k)$. Let’s now estimate the probability that a random game belongs to the set that corresponds to $I_2(n, k)$ (i.e., that the first player cannot win in 1 move). In terms of sets $M_1$ and $W_1$, it means that for every position $p$, it is not true that the pair $(p_0, p)$ belongs to the set of all possible moves $M_1$, and this position $p$ is winning (i.e., $p$ belongs to $W_1$). The sets $W_1$ and $M_1$ are independently chosen. Therefore, the probability that a pair $(p_0, p)$ belongs to $M_1$, and $p \in W_1$, is equal to the product of the corresponding probabilities, i.e. to $1/2 \cdot 1/4 = 1/8$. Therefore, the probability that this property is not true for a position $p$ is equal to $1 - 1/8 = 7/8$.

For each of $A = k^n$ positions $p$, this property is satisfied or not satisfied independently on all other positions. Therefore, the probability that this property (namely, that a move to $p$ does not make the 1-st player win) is true for all positions $p$, is equal to the product of $k^n$ probabilities (each of which is equal to $7/8$) that property is is true for each of $k^n$ positions. So, the probability that a first player cannot win in one move is equal to $(7/8)^A$.

As we have mentioned earlier, this probability is equal to the ratio $I_2(n, k)/G(n, k)$, so
\[
\ln(I_2(n, k)/G(n, k)) = A \ln(7/8) = k^n \ln(7/8),
\]
therefore
\[
|\ln(I_2(n, k)/G(n, k))| = k^n \ln(8/7)
\]
and
\[
\ln |\ln(I_2(n, k)/G(n, k))| = n \ln(k) + const \sim n \ln(k).
\]

So, for $I_2(n, k)$, the asymptotics is really the same as for $I_1(n, k)$. Hence, the same asymptotics is true for $I(n, k)$ that lies in between $I_1(n, k)$ and $I_2(n, k)$. Q.E.D.

**ACKNOWLEDGEMENTS.** This work was partially supported by NSF Grant No. CDA-9015006. One of the authors (V.K.) is also greatly thankful to all the participants
of the USSR National Symposium on Cybernetics, especially to Sergei Yu. Maslov, for valuable discussions, to the Leningrad Laboratory of Experimental Psychological Systems for partial financial support, and to Patrick Suppes (Stanford) and Peter Fishburn (AT&T Bell Labs) for their attention to this work.

REFERENCES


