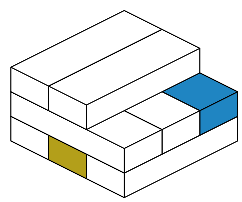
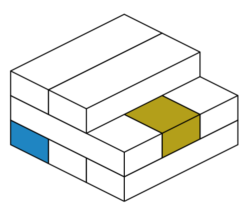
JENGA using Game Theory

Luca Martucci

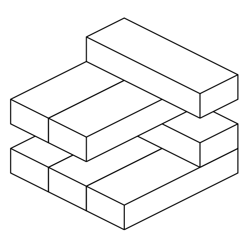
Introduction Jenga is a game played with a tower of $n \in \mathbb{N}$ layers, each composed of three wooden rectangular bricks, for a total $3n$ little blocks, placed with alternating orientation within each layer. The goal of each player at each round is to remove one of such bricks from the construction and placing it at its top, without making the whole tower collapse on itself. Of course players are not allowed to easily remove bricks from the top most layer so, at the beginning of an n-layer game, it is legal to remove blocks from the bottom most $n-1$ layers. We can visualize a typical position of the game as follows:



where we color each block that has already been removed with the color we have chosen to identify the players who moved, here Player 1 and Player 2, who respectively chose to remove the central brick of the first bottom layer and a lateral piece from the second from the bottom layer. We will consider this exact position to be precisely equivalent to the position:



as, up to permuting levels, the topology of the tower is the same: the only parameters defining it are the number of layers with all three blocks, from which it is legal to remove blocks, which we shall call $x \in \mathbb{N}$, and the number of layers with one lateral block missing, which we shall call $y \in \mathbb{N}$[*]. To justify this restriction, we will neglect the effect of gravity and consider any position, i.e. any layer layout, stable, except for a layer with both a central and a later piece removed. For instance, we will consider the following geometry of layers, and every pattern of moves leading up to it, as impossible:



[*] Neither x nor y are in any case bounded by n, because of the creation of the new layer atop the tower every three moves.

as, of course, here the tower has already collapsed to the ground into pieces, and no rational player would have played a move leading to such a predictable outcome; thus, we shall also crucially assume that every player is rational in our analysis. Not only rational, but aware of the others' rationality, up to multiple levels of knowledge of rationality.

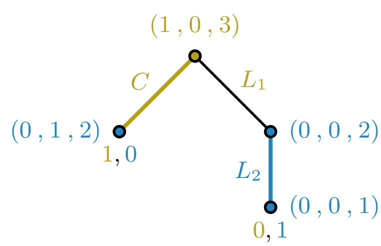
We will restrict to the analysis of two-players Jenga games.

Trivial Jenga Assume the game starts out with $n = 2$ layers. This would be very fast, and boring, as there are only 3 blocks from which to choose and it easy to count the maximum number of moves and height of the tower the round before it collapses: the answer is in both cases two. In fact, if Player 1 removes one of the lateral pieces, and we shall denote this move by L_1, then Player 2 is force to remove the remaining lateral piece, we shall denote this move by L_2, winning the game as Player 1 is left with no move left. Notice that with the winning move, Player 2 is one round away from generating a new top layer and give an additional couple of moves to Player 1; as this will turn out to be an important parameter to keep track of, we shall call it $\ell \in \{1, 2, 3\}$.

Hence, we can fully describe the state of the Jenga tower at each round with the triple

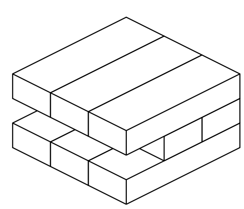
$$(x, y, \ell) \in \mathbb{N}^2 \times \{1, 2, 3\}$$

and that is all we need. This is a awful lot of machinery for the game we are currently analyzing, but will be useful for more-layers games later. Here, in fact, Player 1 is rational and knows Player 2 is also rational and would play L_2 if he played L_1 handing him victory on a silver platter, and will thus remove the central piece, a move we shall call C. The full game tree $\mathcal{T}^{(2)}$ for this toy example is given by:

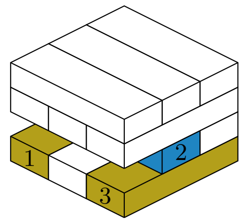


where we hanged the payoffs for each player (1 for a victory and 0 for a defeat) around the leaves, i.e. the ending states before collapse. Using backward induction, Player 1 will compare the possible outcomes and chose C. A rational first player always wins.

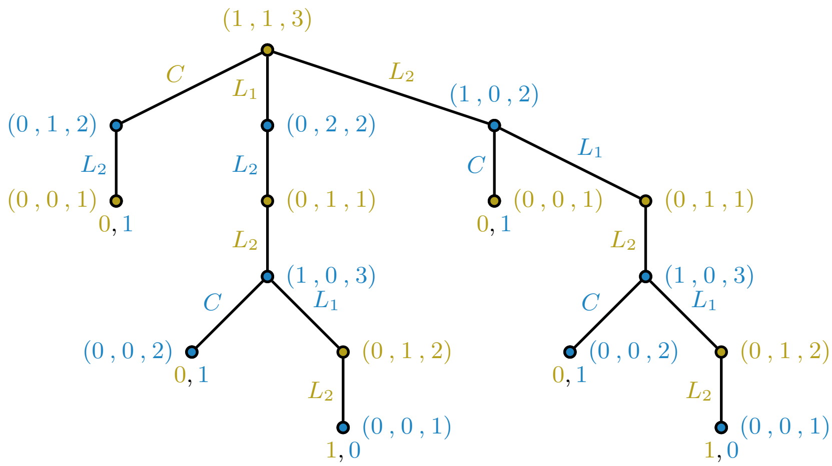
Jenga with missing pieces Now suppose $n = 3$, but you lost a piece and decide to start from the following setup:



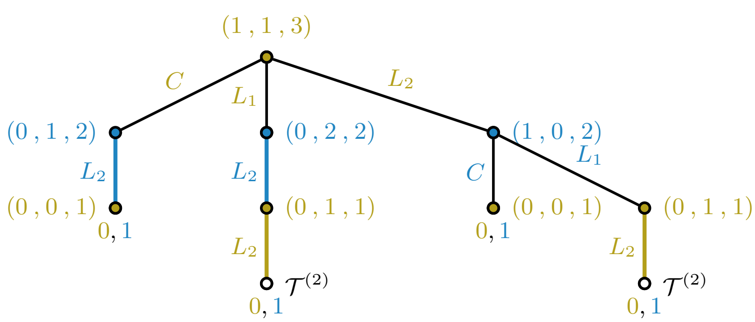
that is, the rather unstable state $(1, 1, 3)$. It is immediately evident that we could reduce to the two-layer case $(1, 0, 3)$ by the following indexed sequence of moves:



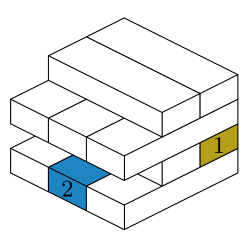
in fact, the full game tree $\mathcal{T}_1^{(3)\dagger}$ is given by:



and we can notice that both combinations among the players of a couple of L_2 moves and one L_1 move reduce this game to the previous, which is a proper subgame of the current one. We can now avoid to solve this full tree, as we know that the first player to move after the new layer is created will be the winner!



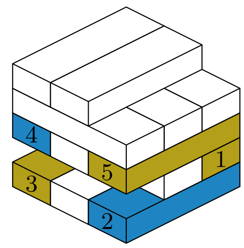
On the right-most branch, using backward induction, Player 2 now faces the choice between C and L_1, which have equal outcomes and to which he ought to be indifferent, as they both lead to his win, either directly:



---

†We adopt the notation $\mathcal{T}_k^{(n)}$ where $n \geq 2$ stands for the number of starting layers and $k \leq n$ the number of those layers that have one lateral piece missing. If $k = 0$, we shall drop the subscript.

or through reducing to the trivial case:

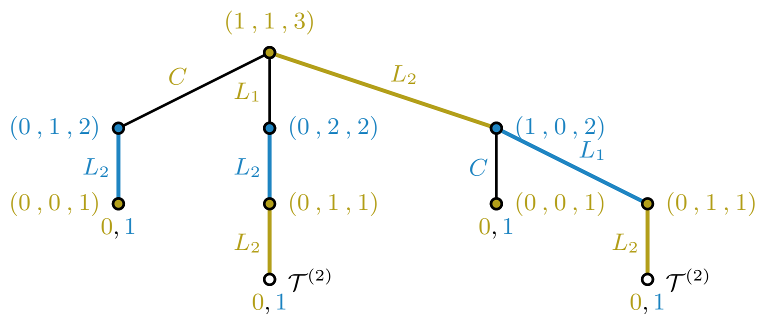


(1)

We can suppose that:

$$L_2 \succsim L_1 \succsim C$$

where $\succsim$ denotes the preferences of the players among actions giving indifferent outcomes. That is, players are risk-lover and willing to play the move destabilizing the tower the most, albeit risking for it to collapse during their removal:



the situation in (1) is a Subgame Perfect Equilibrium, and Player 2 wins. Assuming that

$$L_1 \succsim L_2 \succsim C$$

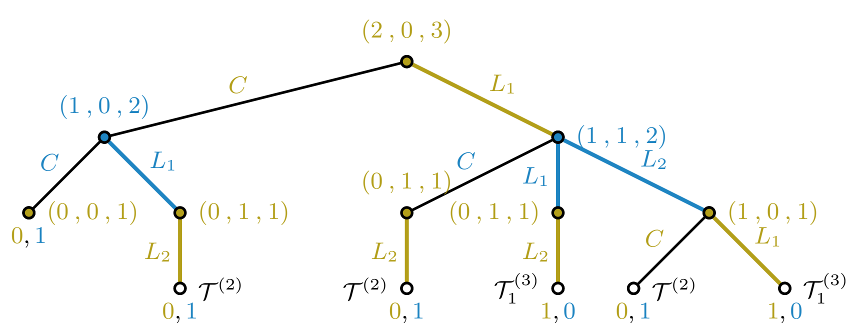
i.e. the players are sort of risk-neutral, would make the central branch of $\mathcal{T}_1^{(2)}$ the SPE, but Player 2 would still be the winner.

We will not however assume total risk-aversity:

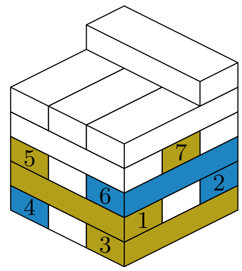
$$C \succsim L_1, L_2$$

as the whole game would just reduce to an ordered removal of central pieces until they are finished and there are no more legal moves. As removing central pieces is not, usually, the most difficult part of the game, winning Jenga would almost be equivalent to playing first.

Three layers If $n = 3$, the game tree $\mathcal{T}^{(3)}$ is given by:

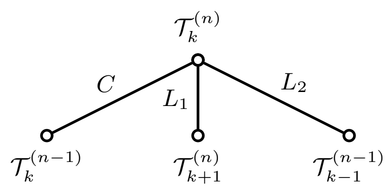


where we already identified the proper subgames. Independently from the risk profile of the players, Player 1 will be the winner. The Subgame Perfect Equilibrium is:

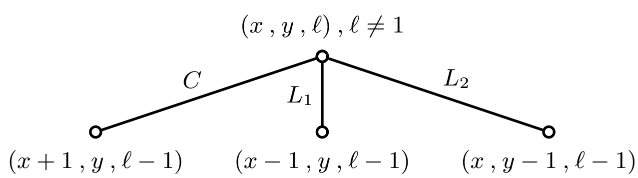


It is worth noticing that, except from the streak of central pieces removed, from all strategies either $\mathcal{T}^{(2)}$ or $\mathcal{T}_1^{(2)}$ blossom out. In both cases, a rational player Player 1 can predict the outcome and play the moves that win the game for him.

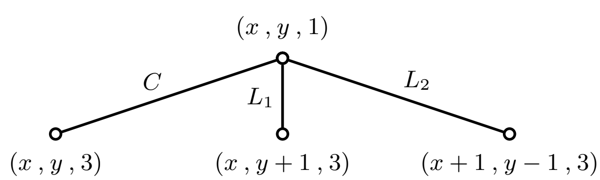
Very many-more layers It is easy to see that this is valid in the general n case: neglecting any sequence of ordered removal of central pieces,



if this is not immediately evident, it will be after starting to draw some trees following the general recursion scheme for the states:



and



The game always ends at states with $x = y = 0$, and Player 1 will force a win for any game with tree $\mathcal{T}^{(n)}$, i.e. with n full layers and all the $3n$ blocks required.

This is the case for the classical version of Jenga, with a tower composed of $n = 54$ bricks. When playing this 18-layer version next time, you ought to try convincing your friend to let you play first!