# Minimax Search and Reinforcement Learning for Adversarial Tetris

Maria Rovatsou and Michail G. Lagoudakis

Intelligent Systems Laboratory
Department of Electronic and Computer Engineering
Technical University of Crete
Chania 73100, Crete, Greece
mariarovatsou@gmail.com, lagoudakis@intelligence.tuc.gr

**Abstract.** Game playing has always been considered an intellectual activity requiring a good level of intelligence. This paper focuses on Adversarial Tetris, a variation of the well-known Tetris game, introduced at the 3rd International Reinforcement Learning Competition in 2009. In Adversarial Tetris the mission of the player to complete as many lines as possible is actively hindered by an unknown adversary who selects the falling tetraminoes in ways that make the game harder for the player. In addition, there are boards of different sizes and learning ability is tested over a variety of boards and adversaries. This paper describes the design and implementation of an agent capable of learning to improve his strategy against any adversary and any board size. The agent employs MiniMax search enhanced with Alpha-Beta pruning for looking ahead within the game tree and a variation of the Least-Squares Temporal Difference Learning (LSTD) algorithm for learning an appropriate state evaluation function over a small set of features. The learned strategies exhibit good performance over a wide range of boards and adversaries.

## 1 Introduction

Skillful game playing has always been considered a token of intelligence, consequently Artificial Intelligence and Machine Learning exploit games in order to exhibit intelligent performance. A game that has become a benchmark, exactly because it involves a great deal of complexity along with very simple playing rules, is the game of Tetris. It consists of a grid board in which four-block tiles, chosen randomly, fall from the top and the goal of the player is to place them so that they form complete lines, which are eliminated from the board, lowering all blocks above. The game is over when a tile reaches the top of the board. The fact that the rules are simple should not give the impression that the task is simple. There are about 40 possible actions available to the player for placing a tile and about $10^{64}$ possible states that these actions could lead to. These magnitudes are hard to deal with for any kind of player (human or computer). Adversarial Tetris is a variation of Tetris that introduces adversity in the game, making it even more demanding and intriguing; an unknown adversary tries to

S. Konstantopoulos et al. (Eds.): SETN 2010, LNAI 6040, pp. 417–422, 2010.

hinder the goals of the player by actively choosing pieces that augment the difficulty of line completion and by even "leaving out" a tile from the entire game, if that suits his adversarial goals. This paper presents our approach to designing a learning player for Adversarial Tetris. Our player employs MiniMax search to produce a strategy that accounts for any adversary and reinforcement learning to learn an appropriate state evaluation function. Our agent exhibits improving performance over an increasing number of learning games.

## 2 Tetris and Adversarial Tetris

*Tetris* is a video game created in 1984 by Alexey Pajitnov, a Russian computer engineer. The game is played on a $10 \times 20$ board using seven kinds of simple tiles, called *tetraminoes*. All tetraminoes are composed of four colored blocks (*minoes*) forming a total of seven different shapes. The rules of the game are very simple. The tiles are falling down one-by-one from the top of the board and the user rotates and moves them until they rest on top of existing tiles in the board. The goal is to place the tiles so that lines are completed without gaps; completed lines are eliminated, lowering all the remaining blocks above. The game ends when a resting tile reaches the top of the board. Tetris is a very demanding and intriguing game. It has been proved [1] that finding a strategy that maximizes the number of completed rows, or maximizes the number of the lines eliminated simultaneously, or minimizes the board height, or maximizes the number of tetraminoes placed in the board before the game ends is an $\mathcal{NP}$-hard problem; even approximating an optimal strategy is $\mathcal{NP}$-hard. This inherent difficulty is one of the reasons this game is widely used as a benchmark domain. Tetris is naturally formulated as a Markovian Decision Process (MDP) [2]. The state consists of the current board and the current falling tile and the actions are the approximately 40 placement actions for the falling tile. The transition model is fairly simple; there are seven equiprobable possible next states, since the next board is uniquely determined and the next falling piece is chosen uniformly. The reward function gives positive numerical values for completed lines and the goal is to find a policy that maximizes the long-term cumulative reward.

The recent Reinforcement Learning (RL) Competition [3] introduced a variation of Tetris, called *Adversarial Tetris*, whereby the falling tile generator is replaced by an active opponent. The tiles are now chosen purposefully to hinder the goals of the player (completion or lines). The main difference in the MDP model of Adversarial Tetris is the fact that the distribution of falling tiles is non-stationary and the dimension of the board varies in height and width. Furthermore, the state is produced like the frames of the video game, as it includes the current position and rotation of the falling tile in addition to the configuration of the board and the player can move/rotate the falling tile at each frame. The RL Competition offers a generalized MDP model for Adversarial Tetris which is fully specified by four parameters (the height and width of the board and the adversity and type of the opponent). For the needs of the competition 20 instances of this model were specified with widths ranging from 6 to 11, heights ranging from 16 to 25, and different types of opponents and opponent's adversity.