Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (2008)

Simulation-Based Approach to General Game Playing

Hilmar Finnsson and **Yngvi Björnsson**
School of Computer Science
Reykjavík University, Iceland
{hif,yngvi}@ru.is

Abstract

The aim of *General Game Playing (GGP)* is to create intelligent agents that automatically learn how to play many different games at an expert level without any human intervention. The most successful GGP agents in the past have used traditional game-tree search combined with an automatically learned heuristic function for evaluating game states. In this paper we describe a GGP agent that instead uses a Monte Carlo/UCT simulation technique for action selection, an approach recently popularized in computer Go. Our GGP agent has proven its effectiveness by winning last year's AAAI GGP Competition. Furthermore, we introduce and empirically evaluate a new scheme for automatically learning search-control knowledge for guiding the simulation playouts, showing that it offers significant benefits for a variety of games.

Introduction

In *General Game Playing (GGP)* the goal is to create intelligent agents that can automatically learn how to skillfully play a wide variety of games, provided only the descriptions of the game rules. This requires that the agents learn diverse game-playing strategies without any game-specific knowledge being provided by their developers. A successful realization of this task poses interesting research challenges for artificial intelligence sub-disciplines such as knowledge representation, agent-based reasoning, heuristic search, and machine learning.

The most successful GGP agents so far have been based on the traditional approach of using game-tree search augmented with an (automatically learned) heuristic evaluation function for encapsulating the domain-specific knowledge (Clune 2007; Schiffel & Thielscher 2007; Kuhlmann, Dresner, & Stone 2006). However, instead of using a set of carefully hand-crafted domain-specific features in their evaluation as high-performance game-playing programs do, GGP programs typically rely on a small set of generic features (e.g. piece-values and mobility) that apply in a wide range of games. The relative importance of the features is then automatically tuned in real-time for the game at hand.

There is an inherent risk with that approach though. In practice, because of how disparate the games and their playing strategies can be, the pre-chosen set of generic features may fail to capture some essential game properties. On top of that, the relative importance of the features can often be only roughly approximated because of strict online time constraints. Consequently, the resulting heuristic evaluations may become highly inaccurate and, in the worst case, even strive for the wrong objectives. Such heuristics are of a little (or even decremental) value for lookahead search.

In here we describe a simulation-based approach to general game playing that does not require any *a priori* domain knowledge. It is based on *Monte-Carlo (MC)* simulations and the *Upper Confidence-bounds applied to Trees (UCT)* algorithm for guiding the simulation playouts (Kocsis & Szepesvári 2006), thus bypassing the need for a heuristic evaluation function. Our GGP agent, CADIAPLAYER, uses such an approach to reason about its actions and has already proven its effectiveness by winning last year's annual GGP competition. The UCT algorithm has recently been used successfully in computer Go programs, dramatically increasing their playing strength (Gelly *et al.* 2006; Coulom 2006). However, there are additional challenges in applying it to GGP, for example, in Go pre-defined domain-knowledge can be used to guide the playout phase, whereas such knowledge must be automatically discovered in GGP.

The main contributions of the paper are as follows, we: (1) describe the design of a state-of-the-art GGP agent and establish the usefulness of simulation-based search approaches in GGP, in particular when used in combination with UCT; (2) empirically evaluate different simulation-based approaches on a wide variety of games, and finally (3) introduce a domain-independent enhancement for automatically learning search-control domain-knowledge for guiding simulation playouts. This enhancement provides significant benefits on all games we tried, the best case resulting in 90% winning ratio against a standard UCT player.

The paper is structured as follows. In the next section we give a brief overview of GGP, followed by a review of our GGP agent. Thereafter we detail the simulation-based search approach used by the agent and highlight GGP specific enhancements, including the new technique for improving the playout phase in a domain-independent manner. Finally, we present empirical results and conclude.