Journal of Artificial Intelligence Research 30 (2007) 659-684 Submitted 06/07; published 12/07

Learning to Play Using Low-Complexity Rule-Based Policies: Illustrations through Ms. Pac-Man

István Szita szityu@eotvos.elte.hu
András Lőrincz andras.lorincz@elte.hu
Dept. of Information Systems
Eötvös University, Hungary, H-1117

Abstract

In this article we propose a method that can deal with certain combinatorial reinforcement learning tasks. We demonstrate the approach in the popular Ms. Pac-Man game. We define a set of high-level observation and action modules, from which rule-based policies are constructed automatically. In these policies, actions are temporally extended, and may work concurrently. The policy of the agent is encoded by a compact decision list. The components of the list are selected from a large pool of rules, which can be either hand-crafted or generated automatically. A suitable selection of rules is learnt by the cross-entropy method, a recent global optimization algorithm that fits our framework smoothly. Cross-entropy-optimized policies perform better than our hand-crafted policy, and reach the score of average human players. We argue that learning is successful mainly because (i) policies may apply concurrent actions and thus the policy space is sufficiently rich, (ii) the search is biased towards low-complexity policies and therefore, solutions with a compact description can be found quickly if they exist.

1. Introduction

During the last two decades, reinforcement learning (RL) has reached a mature state, and has been laid on solid foundations. We have a large variety of algorithms, including value-function-based, direct policy search and hybrid methods. For reviews on these subjects, see, e.g., the books of Bertsekas and Tsitsiklis (1996) and Sutton and Barto (1998). The basic properties of many such algorithms are relatively well understood, e.g. conditions for convergence, complexity, the effect of various parameters, although it is needless to say that there are still lots of important open questions. There are also plenty of test problems (like various maze-navigation tasks, "pole-balancing", "car on the hill" etc.) on which the capabilities of RL algorithms have been demonstrated, and the number of large-scale RL applications is also growing steadily. However, current RL algorithms are far from being out-of-the-box methods, so there is still need for more demonstrations showing that RL can be efficient in complex tasks.

We think that games (including the diverse set of classical board games, card games, modern computer games, etc.) are ideal test environments for reinforcement learning. Games are intended to be interesting and challenging for human intelligence and therefore, they are ideal means to explore what artificial intelligence is still missing. Furthermore, most games fit well into the RL paradigm: they are goal-oriented sequential decision problems, where each decision can have long-term effects. In many cases, hidden information, random events, unknown environment, known or unknown players account for (part of) the difficulty