On Combinatorial Actions and CMABs with Linear Side Information

Alexander Shleyfman
Advisor: Prof. Carmel Domshlak
09.11.2014

Abstract

Online planning algorithms are typically a tool of choice for dealing with sequential decision problems in combinatorial search spaces. Many such problems also exhibit combinatorial action spaces, yet standard online planning algorithms do not cope well with this type of “the curse of dimensionality.” Following a recently opened line of related work on combinatorial multi-armed bandit (CMAB) problems, we propose a novel CMAB planning scheme, as well as two specific instances of this scheme, dedicated to exploiting what is called linear side information. Using a representative strategy game as a benchmark, we show that the resulting algorithms very favorably compete with the state of the art. Furthermore, we experimentally analyze robustness of these algorithms for online planning with combinatorial actions in various setups of real-time and turn taking strategy games.

As an additional step towards more effective online planning algorithms, we introduce Discriminative Bucketing (DB), a novel family of strategies for pure exploration in MABs, which allows for adapting recent advances in non-interruptible strategies to the interruptible setting, while guaranteeing exponential-rate performance improvement over time. These strategies should bring us one step closer to a strong generic anytime algorithm for online decision making. Our experimental evaluation demonstrates that the corresponding instances of DB favorably compete both with the currently popular strategies UCB and epsilon-greedy, as well as with the conservative uniform sampling.