Continuous search algorithm

Beomjoon Kim

January 6, 2019

1 Introduction

2 Problem formulation

Given a computational budget B, a plan skeleton $\{\mathfrak{o}_1(\delta_1, \cdot), \cdots, \mathfrak{o}_T(\delta_t, \cdot)\}$, and the set of goal states S_G, possibly described with a predicate, find the continuous parameters $\kappa_1, \cdots, \kappa_T$ such that it maximizes the sum of the discounted rewards,

$$\max_{\kappa_1, \cdots, \kappa_T} \sum_{t=0}^{T} \gamma^t r(s_t, \kappa_t)$$

where $r(s_t, \kappa_t) = r(s_t, \mathfrak{o}_t(\delta_t, \kappa_t))$, $s_T \in S_G$ and the generative model of the environment T such that $s_{t+1} = T(s_t, \mathfrak{o}_t(\delta_t, \kappa_T))$