Learning Multi-Step Predictive State Representations - Appendix

Abstract

This appendix contains pseudocode for the algorithms presented in the main paper, and more details on the experiments.

Pseudocode of the encoding function

The pseudocode is given in Algorithm 1, using the following notation:

bestEncoding: A map from indices i of the query string x to the optimal encoding of x[:i].

minEncoding: A map from indices i of the query string x to $|bestEncoding[i]|$

opsEnding: A map from indices i of x to the set of strings in Σ': $\{s \in \Sigma', x[i-|s|:i] = s\}$

Algorithm 1 Encoding Algorithm

INPUT: x
OUTPUT: $\kappa(x)$

procedure DPENCODE
  $bestEncoding[] \leftarrow String[|x|+1]$
  $minEncoding[] \leftarrow Int[|x|+1]$
  $opsEnding[] \leftarrow String[|x|+1][]$
  $bestEncoding[0] = x[0]$
  $minEncoding[0] = 0$
  for i in $[1, |x|]$ **do**
    $ospEnding[i] \leftarrow \{s \in \Sigma', x[i-|s|:i] = s\}$
  end for
  for i in $[1, |x|]$ **do**
    $bestOp \leftarrow null$
    $m \leftarrow 0$
    for $s \in opEnd[i]$ **do**
      $t \leftarrow minE[i-|s|]+1$
      if $bestOp = null$ or $t < m$ **then**
        $m \leftarrow t$
        $bestOp \leftarrow s$
      end if
    end for
    $minEncoding[i+1] \leftarrow m$
    $bestEncoding[i+1] \leftarrow bestEncoding[i-|bestOp|] + bestOp$
  end for
  return $bestEncoding[|x|]$
end procedure

Algorithm 2 Base Selection Algorithm

INPUT: $Train$, Sub_M
OUTPUT: Σ'

procedure BASE SELECTION
  $\Sigma' \leftarrow \{s, s \in \sum\}$
  $bestEncoding \leftarrow null$
  for each obs in Train **do**
    $bestEncoding[obs] \leftarrow |obs|$
  end for
  $i \leftarrow 0$
  while $i < numOps$ **do**
    $bestOp \leftarrow null$
    $m \leftarrow 0$
    for each s $\in Sub_M$ **do**
      $c \leftarrow 0$
      for each obs in Train **do**
        $c \leftarrow c + DPEncode(obs) - prevBestE(obs)$
      end for
      if $c > m$ **then**
        $bestOp \leftarrow obs$
        $m \leftarrow c$
      end if
    end for
    $\Sigma' \leftarrow \Sigma' \cup bestOp$
    for each obs in Train **do**
      $bestEncoding \leftarrow DPEncode(obs, \Sigma')$
    end for
    $i \leftarrow i + 1$
  end while
  return Σ'
end procedure

Experiments

The figures below depict the Double Loop and Pacman environments used in the experiments.

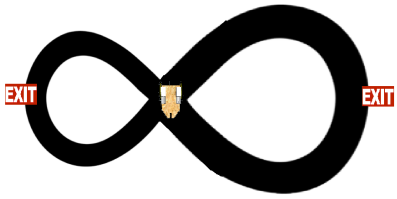

Figure 1: Double Loop Environment

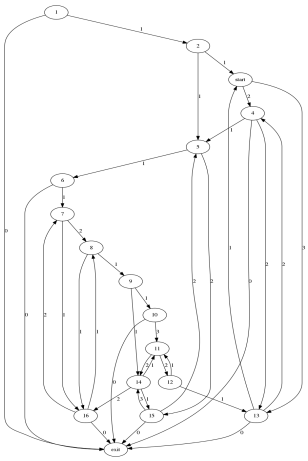

Figure 2: Graph of Pacman Labyrinth