

Trabalho Prático

— Aprendizado Indutivo —

Objetivo

O objetivo do trabalho é implementar e experimentar a resolução de problemas com um aprendizado por reforço e por exemplos.

Problema

Fazer um agente aprender por reforço a jogar o jogo “2048”. Ver em <http://gabrielecirulli.github.io/2048/> como funciona o jogo e suas regras.

Avalie o desempenho do aprendiz com várias representações de estado, da mais específica (onde o estado do tabuleiro é o estado para o algoritmo de QL) até representações mais gerais (onde vários estados do tabuleiro são representados pelo mesmo estado para o QL).

Avalie também o desempenho do aprendiz com diferentes funções de recompensa.

A partir da política aprendida, gere uma tabela de exemplos com as decisões para cada um dos estados possíveis e verifique a árvore de decisão gerada por algoritmos como o J4.8. Eventualmente atributos relevantes dos estados deveriam ser incluídos nos exemplos (como no exemplo badges do Weka).

Relatório

Conforme cronograma da disciplina, deve ser entregue um relatório descrevendo a modelagem do problema (o MDP), a implementação da solução e o resultado dos experimentos.

Procure avaliar:

- Os parâmetros para o QL.
- As diferentes representações de estado.
- As diferentes funções de recompensa.
- A qualidade da árvore de decisão gerada pelo J4.8.