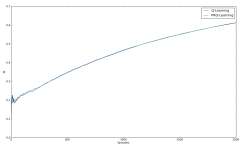
experiments

Rafael Lemes Beirigo

01 abril 2012

Contents



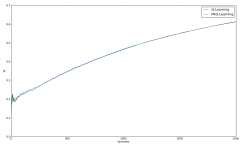
1 00

1.1 O quê: Reprodução do artigo

1.2 Resultado: diverso do esperado

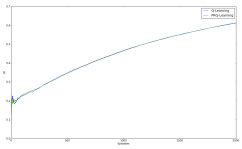
O Q-Learning apresentou um desempenho extremamente melhor do que o PRQL. Houve um problema: não estava zerando as Q-Tables (QLearning e PRQLearning)

2 01



Repetição de 00 após correção do problema Resultado: PRQL tem o mesmo comportamento de QL, só um pouco pior. Hipótese: não está utilizando as políticas antigas

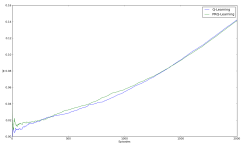
3 02 QL vs PRQL no mundo 05x05



Motivo: queria rodar o PRQL sem a parte PR, ou seja, só utilizando QLearning, pra ver se estava tudo OK Nesse experimento, NÃO utilizava pi-reuse, somente QL

4 03 Mesmo experimento de 02, só que para o mundo 06x06

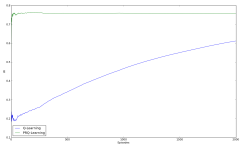
5 04 Mesmo experimento de 02, só que para a task omega do artigo



5.1 Resultado:

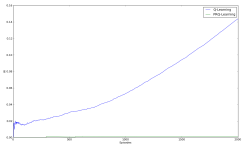
PRQL e QL apresentaram desempenhos compatíveis, o que era esperado

6 05 Repetindo 04, só que dessa vez ativando o pi-reuse



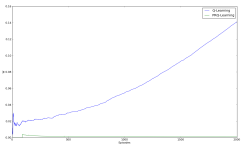
Sucesso: PRQL acelerou QLearning

7 06 Repetindo 05 para task omega do artigo reutilizando políticas 2, 3 e 5 (são as que mais ajudam o agente)



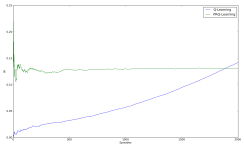
Problema: plotando W[1]

8 07 Repetindo 06



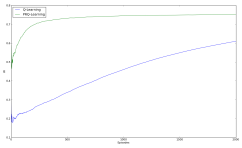
Problema: plotando W[1]

9 08 Repetindo 06, mas reutilizando somente a política ótima



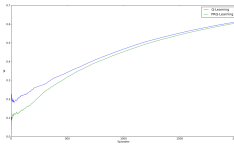
Problema: plotando W[1]

10 09 Repetindo 02, após correção do acúmulo de recompensas médias por episódio



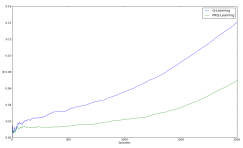
Problema: reutilizando políticas subótimas

11 10 Repetindo 09, mas reutilizando uma política ótima para o problema de chegar



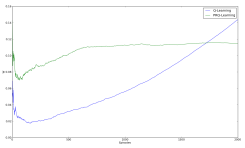
à localização oposta (pior política que poderia reutilizar)
Problema: reutilizando políticas subótimas

12 11 Resolver task omega utilizando pols. 2,3,4,5



Problema: reutilizando políticas subótimas

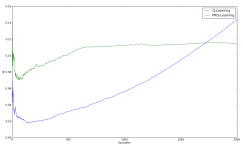
13 12 Repetindo 11 reutilizando somente a policy obtida em 11 pelo



QLearning (ótima para o problema)

Problema: reutilizando políticas subótimas

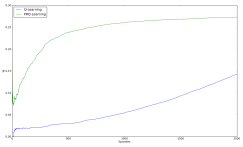
14 13 Repetindo 12, só que chamei o solveMDP... pra criar os arquivos (tirar a dúvida se



arquivos estão corretos)

Problema: reutilizando políticas subótimas

15 14 Repetição do 13, só que agora utilizando a política ótima



16 15 Obtenção de política para task 1

Consumo de tempo: ~ 10'

17 16 Obtenção de política para task 2

Consumo de tempo: ~ 10'

18 17 Obtenção de política para task 3

Consumo de tempo: ~ 10'

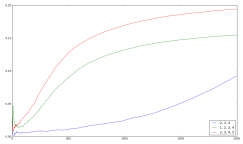
19 18 Obtenção de política para task 4

Consumo de tempo: ~ 10'

20 19 Obtenção de política para task 5

Consumo de tempo: ~ 10'

21 20 Resolver task omega utilizando pols. 2,3,4,5 (Repetição do 11)



22 21 Resolver task omega utilizando pols. 2,3,4

/home/rafaelbeirigo/ql/experiments/21/w.png

23 22 Resolver task omega utilizando pols. 1,2,3,4

/home/rafaelbeirigo/ql/experiments/22/w.png

24 23 Repetição do 02

25 24 Repetição do 22

26 25 Repetição do 20

27 26 Repetição do 21