

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Eugênio Pacelli Ferreira Dias Júnior

**Aprendizado por reforço sobre o
problema de revisitação de páginas web**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial
para obtenção do título de Mestre pelo Programa
de Pós-Graduação em Informática da PUC-Rio.

Orientador: Prof. Eduardo Sany Laber

Rio de Janeiro
janeiro de 2012



Eugênio Pacelli Ferreira Dias Júnior

**Aprendizado por reforço sobre o
problema de revisitação de páginas web**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Eduardo Sany Laber

Orientador

Departamento de Informática - PUC-Rio

Prof. Marcus Vinícius Soledade Poggi de Aragão

Departamento de Informática - PUC-Rio

Prof. Raul Pierre Renteria

Departamento de Informática - PUC-Rio

Prof. José Eugenio Leal

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 19 de janeiro de 2012

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Eugênio Pacelli Ferreira Dias Júnior

Graduou-se em Engenharia de Computação pelo Instituto Militar de Engenharia. Continuou seus estudos no programa de Mestrado em Informática da Pontifícia Universidade Católica do Rio de Janeiro, onde atuou também como monitor da disciplina de Projetos e Análise de Algoritmos, em 2009.

Ficha Catalográfica

Dias Júnior, Eugênio Pacelli Ferreira

Aprendizado por reforço sobre o problema de revisitação de páginas web / Eugênio Pacelli Ferreira Dias Júnior ; orientador: Eduardo Sany Laber. – 2012.

73 f. : il. (color.) ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2012.

Inclui bibliografia

1. Informática – Teses. 2. Algoritmos. Aprendizado por reforço. 3. Revisitação de páginas. 4. Reinforcement learning. 5. Aprendizado de máquina. 6. Freshness. 7. Atualidade de base de dados. I. Laber, Eduardo Sany. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Para minha esposa Renata

Agradecimentos

A Deus, acima de tudo.

A meus pais, que em todas as ocasiões da vida sempre me incentivaram a alcançar degraus mais altos.

A minha esposa Renata, que tanto me apoiou e motivou durante a realização deste mestrado.

A meu orientador, Eduardo Sany Laber, pelo apoio e ensinamentos recebidos durante o curso.

A meus chefes, Gen Aléssio Ribeiro Souto, Gen Sergio Westphalen Etchegoyen, Cel Abner de Oliveira e Silva e Cap Marco Aurélio Nunes dos Santos, pela oportunidade de realizar o curso de mestrado paralelamente a minhas atividades profissionais.

À PUC-Rio, pela bolsa de isenção de pagamento e pelo clima favorável à pesquisa que impera no campus.

Resumo

Dias Jr., Eugenio Pacelli Ferreira; Laber, Eduardo Sany. **Aprendizado por reforço sobre o problema de revisitação de páginas web.** Rio de Janeiro, 2012. 73p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

No ambiente da Internet, as informações que desejamos frequentemente encontram-se em diferentes localidades. Algumas aplicações, para funcionarem corretamente, precisam manter cópias locais de parte dessas informações. Manter a consistência e a atualidade de uma base de dados, mais especificamente um conjunto de cópias de páginas web, é uma tarefa que vem sendo sistematicamente estudada. Uma abordagem possível a esse problema é a aplicação de técnicas de aprendizado por reforço, que utiliza técnicas de programação dinâmica e análise estocástica para obter uma boa política de agendamento de atualizações das cópias de páginas web. O presente trabalho tem por finalidade validar o uso de técnicas de aprendizado por reforço no problema em questão, assim como encontrar aspectos do problema que possam ser úteis na modelagem da solução empregada.

Palavras-chave

Algoritmos; Aprendizado por reforço; revisitação de páginas; reinforcement learning; aprendizado de máquina; freshness; atualidade de base de dados

Abstract

Dias Jr., Eugenio Pacelli Ferreira; Laber, Eduardo Sany (Advisor). **Using Reinforcement Learning on web pages revisiting problem.** Rio de Janeiro, 2012. 73p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

In the Internet, the information we desire is usually spread over different locations. For some applications, it is necessary to maintain local copies of this information. Keeping consistency as well as freshness of a data base, or more specifically a set of internet web pages, is a task systematically studied. An approach to this problem is the use of reinforcement learning techniques, using dynamic programming and stochastic analysis to obtain a good rescheduling policy for the web pages copies. This work is proposed to validate the use of reinforcement learning techniques over this problem, as well as finding features of the problem useful to model the developed solution.

Keywords

Algorithms; reinforcement learning; web pages revisiting; freshness; machine learning; data base freshness

Sumário

1	Introdução	13
1.1.	Contribuições	14
1.2.	Organização do documento	15
2	Atualidade de uma base de dados	17
2.1.	Particularidades de página Web - Generalizações	19
2.2.	Particularidades de página Web – Especificidades	19
2.3.	Métricas para atualidade de uma base de dados	20
2.4.	Conceito de freshness médio	21
2.5.	Estratégias clássicas de manutenção da atualidade (revisitação)	21
2.6.	Trabalhos relacionados	23
3	Aprendizado por reforço	25
3.1.	Conceitos básicos	25
3.1.1.	Exploração x Intensificação	27
3.1.2.	Política ótima	27
3.2.	Métodos de aprendizado	32
3.2.1.	Q-Learning	33
3.2.2.	Métodos de aproximação por função com atualização por gradiente	34
4	Uma abordagem por aprendizado por reforço	36
4.1.	Limitações de politeness	36
4.2.	Dinâmica do sistema	37
4.3.	Estados e ações	38
4.4.	Função de Aprendizado	43
4.4.1.	Modelagem em função de regressão logística e seus resultados	44
4.5.	Processo de Aprendizado	46

5 Avaliação Experimental	48
5.1. Conjunto de Dados	48
5.1.1. Criação de conjunto de páginas de Poisson	48
5.1.2. Base de páginas da Wikipédia	49
5.2. Aspectos de Implementação	50
5.3. Testes com as páginas de Poisson	51
5.3.1. Cálculo do freshness médio ótimo	52
5.3.2. Resultado dos testes com as páginas de Poisson	52
5.4. Testes sobre a base de páginas da Wikipédia	57
5.5. Discussão	68
6 Conclusões	70
7 Referências Bibliográficas	72

Lista de figuras

Figura 1 – Modelo padrão de aprendizado por reforço	26
Figura 3 – Exemplo de estado do sistema	41
Figura 4 – Exemplo de estado do sistema	42
Figura 5 – Definição de estados, função de aprendizado e recompensa	43
Figura 6 – Gráfico da função de regressão logística	45
Figura 7 – Evolução do freshness do conjunto de páginas, seguindo as políticas aprendidas com diferentes valores de α e ε	55
Figura 8 – Resultados de testes com páginas de Poisson - RL e políticas clássicas (freshness x tempo)	56
Figura 9 – Número de páginas por grupo de quantidade de edições em um ano	57
Figura 10 – Evolução do freshness do conjunto de páginas da Wikipédia, seguindo as políticas clássicas	59
Figura 11 – Evolução do freshness do conjunto de testes (2º Sem. 2007 – Páginas Wikipédia) para diferentes valores de α e ε	61
Figura 12 – Evolução do freshness do conjunto de treinamento, durante a segunda passagem de treinamento	62
Figura 13 – Evolução do freshness do conjunto de testes, após segunda rodada de treinamento	62
Figura 14 – Comparação de desempenho das políticas aprendidas em um treinamento (τ_1) e em dois treinamentos (τ_2), com os mesmos valores de α e ε	63
Figura 15 – Comparação de desempenho das políticas aprendidas em um treinamento (τ_1) e em dois treinamentos (τ_2), com os mesmos valores de α e ε	64
Figura 16 – Comparativo das políticas clássicas com as políticas aprendidas por RL	65

Figura 17 – Freshness do conjunto de testes, utilizando as políticas aprendidas com 5 passagens de treinamento	66
Figura 18 – Evolução do freshness do conjunto de testes, com a política aprendida com α e ε decrescentes	67
Figura 19 – Comparativo entre políticas aprendidas com 5 treinamentos, uma com valores de α e ε fixos, outra com valores decrescentes no tempo	67
Figura 20 – Comparativo entre a política aprendida com valores decrescentes no tempo de α e ε com as políticas clássicas	68

Lista de tabelas

Tabela 1 – Freshness esperado para diferentes políticas de ordem de revisitação [CHO]	23
Tabela 2 – Tabela de probabilidade de transição de estado	29
Tabela 3 – Exemplo de recompensas por ações tomadas em diferentes estados	30
Tabela 4 – Valores esperados ao seguir uma política à partir de diferentes estados.	30
Tabela 5 – Gráficos da evolução do freshness do conjunto de treinamento para diferentes valores de α e ε , em função do tempo	54
Tabela 6 – Gráficos da evolução do freshness do conjunto de treinamento (Wikipédia) para diferentes valores de α e ε , em função do tempo	60