

Pesquisa e Classificação de Dados - Trabalho da disciplina

Prof. Ricardo Oliveira - 2022/1

Enunciado

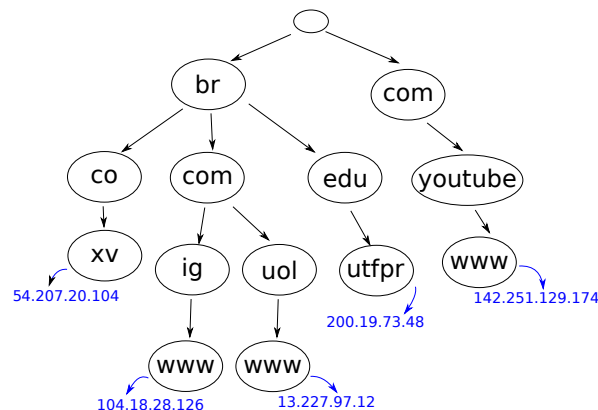
O **DNS** (*Domain Name System*) é o sistema responsável por obter o endereço IP de uma máquina na internet, dado seu endereço textual (sua URL ou “link”). Como exemplo, quando você digita em seu navegador a URL **utfpr.edu.br**, uma consulta ao DNS é feita para descobrir qual o endereço IP desta página. O DNS então responde que o IP da página **utfpr.edu.br** é **200.19.73.48**, endereço a partir do qual o restante da conexão é realizada. Note que o DNS é um grande dicionário, que tem como chaves as URLs das máquinas e como conteúdos associados seus respectivos endereços IP.

Toda URL é separada em diversas partes, separadas por pontos (.), que são lidas *da direita para a esquerda*. Por exemplo, a URL **utfpr.edu.br** tem três partes, nessa ordem: **br**, **edu** e **utfpr**. O DNS é organizado de forma hierárquica, de tal forma que, neste exemplo, é primeiro buscado pelos endereços que terminam em **br**; após, dentre eles, é buscado por endereços que terminam em **edu**; por fim, dentre eles, é buscado pelo endereço **utfpr**.

Como exemplo, considere os seguintes endereços e seus respectivos IPs:

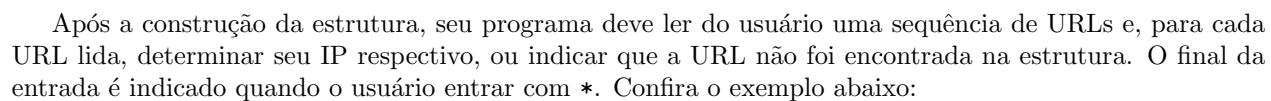
URL	IP
www.uol.com.br	13.227.97.12
utfpr.edu.br	200.19.73.48
www.youtube.com	142.251.129.174
xv.co.br	54.207.20.104
www.ig.com.br	104.18.28.126

O DNS é organizado hierarquicamente da seguinte forma:



Como exemplo, considere que o arquivo `dns.txt` contém o seguinte conteúdo:

O seu programa deve construir a seguinte estrutura de dados:



Exemplo de entrada	Exemplo de saída
utfpr.edu.br	200.19.73.48
xv.co.br	54.207.20.104
xv.com.br	endereco nao encontrado.
www.uol.com.br	13.227.97.12
uol.com.br	endereco nao encontrado.
*	

Você pode assumir que o arquivo `dns.txt` sempre existe e é válido; que nenhuma URL aparecerá repetida nesse arquivo; que toda URL tem no máximo 100 caracteres, sendo todos letras minúsculas ou pontos; que toda URL tem no mínimo 1 ponto; que não há dois pontos seguidos em nenhuma URL; e que todo IP é formado por 4 números de 0 a 255 separados por pontos.

Ao final da execução, libere toda a memória alocada.

Orientações

- O trabalho pode ser feito por equipes de *até* 2 (dois) estudantes;
- Submeta, via *Moodle*, um pacote (zip ou tar.gz) contendo:
 - O código fonte do trabalho;
 - Um arquivo de texto (txt) onde conste o nome de todos os integrantes da equipe e toda informação que a equipe julgar relevante para a correção.
- Comente adequadamente seus códigos para facilitar a correção;
- Atenção: a correção será parcialmente automatizada, e a saída do programa será testada com outras entradas além das fornecidas como exemplo. *Siga **fielmente** o formato de saída dado nos exemplos*, sob pena de grande redução da nota;
- Certifique-se que seu programa funciona antes de submetê-lo;
- O trabalho deve ser entregue até **23 de Junho, 23:59**, via *Moodle*. É suficiente que o trabalho seja submetido por apenas um estudante da equipe;
- Trabalhos copiados ou plagiados (de colegas ou da internet), ou ainda comprados, receberão **todos** a nota 0 (**ZERO**) e estão sujeitos à abertura de Processo Administrativo Disciplinar Discente.