

Trabalho Prático 1 - Árvore Patrícia

Guilherme Moreira de Carvalho e Marcus Vinícius Diniz Ribeiro

Dificuldades Encontradas na Implementação do Trabalho

A implementação da Árvore Patrícia apresentada na disciplina teórica e disponível na bibliografia trata a indexação de palavras por meio de seus caracteres, formando-as caractere por caractere até encontrar um nó externo e verificando se o último caractere se iguala à chave. O trabalho pede para que as **palavras** sejam indexadas na Árvore, sendo necessário reformular todos os métodos para atender esse critério.

Ademais, a transformação de uma palavra em sequência de bits é mais complexa que a transformação de um caractere em bits, levando em consideração a proposta de expandir cada letra em uma sequência de 8 bits. Como isso cria uma nova identificação para os nós - caractere \rightarrow palavra \rightarrow bits - a sequência de cada palavra se tornou um atributo do nó, facilitando sua comparação com as demais e, conseqüentemente, o deslocamento pela Árvore.

Ainda em relação à sequência de bits, não se diferencia duas palavras por caracteres diferentes em posições iguais, mas sim pelo bit referente àquela posição da palavra já expandida. Assim, a indexação da Árvore também é grandemente alterada, pois não referência a uma das 16 (máximo) letras da palavra, mas sim a um de seus 128 bits.

Por fim, a delimitação das palavras se tornou um grande desafio, pois procura-se tornar o sistema o mais genérico possível de modo a atender diferentes textos. Como se pede o armazenamento das informações da posição de cada palavra lida, não se podia usar o delimitador *new line* ($\backslash n$) para separar as palavras, pois esse é utilizado para informar a linha corrente. Como solução, utiliza-se dois leitores: uma para cada linha do texto, que altera o

atributo *linha* da Árvore toda vez que faz uma leitura; e um para cada palavra do texto que discrimina espaços em branco e pontuação.