

Słownik T9 zrealizowany za pomocą drzewa Trie.

Arkadiusz Kaus

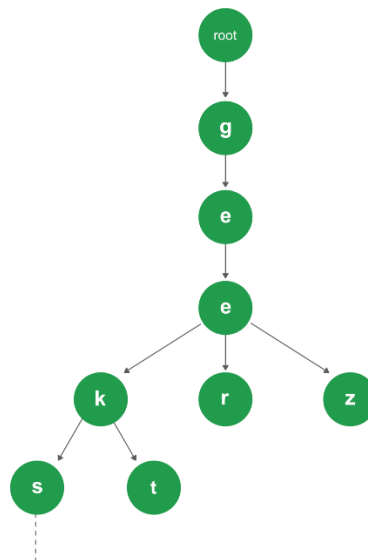
Nr albumu: 1111746

arekk949@gmail.com

06.02.2019 r.

Program ten jest implementacją słownika T9 (*text on 9 keys*) stosowanego w telefonach komórkowych z klawiaturą numeryczną. Zastosowanie słownika przyspiesza wpisywanie tekstu. Zamiast konieczności wielokrotnego naciskania jednego klawisza w celu wpisania konkretnej litery wpisuje się kombinację cyfr, pod którymi znajdują się szukane litery i słownik podpowiada o jakie słowa może chodzić.

Słownik został zrealizowany za pomocą struktury danych drzewo Trie. Jest to znacznie wydajniejsza implementacja niż np. za pomocą hash table. Nazwa tej struktury danych pochodzi od słowa reTRIEval. Jest to drzewo, w którym z każdego węzła wychodzi wiele krawędzi, które reprezentują kolejne znaki słowa. W węźle znajdziemy dwa pola: jedno jest mapą zawierającą listę dzieci danego węzła, gdzie kluczem jest odpowiednia litera słowa a wartością kolejny węzeł, oraz zmienną boolowską która przyjmuje wartość true jeśli dany węzeł jest końcem słowa. Kolejne połączone ze sobą węzły tworzą słowa. Dzięki zastosowaniu pythonowych słowników czas wyszukiwania słowa wynosi $O(\text{DŁUGOŚĆ_SŁOWA})$.



Opis kodu:

1. Klasa TrieNode() – jest to definicja węzła drzewa Trie. Znajdują się tu dwa pola: słownik przechowujący potomków w postaci liter jako kluczy i wartości jako kolejnych węzłów oraz zmienna boolowska oznaczająca czy dany węzeł jest końcem słowa.

2. Klasa Trie() – jest to implementacja klasy drzewa Trie. Znajdują się tu:

- 1) Konstruktor tworzący korzeń drzewa
- 2) Funkcja add() dodająca do drzewa nowy węzeł
- 3) Funkcja search() wyszukująca w drzewie podane słowo
- 4) Funkcja learn() służąca do automatycznego utworzenia drzewa na podstawie wprowadzonego słownika w postaci pliku tekstowego

Następnie poza definicją klasy są dwie funkcje:

3. Funkcja combos() – Na podstawie wprowadzonej przez użytkownika kombinacji cyfr funkcja dokonuje mapowania na możliwe litery alfabetu i zwraca listę wszystkich możliwych kombinacji tych liter.
4. Funkcja predict() – Funkcja przeszukuje drzewo w poszukiwaniu słów które zostały wygenerowane przez funkcję combos(), jeśli słowo znajduje się w słowniku zostaje ono dodane do listy podpowiedzi.