



**Universitat de Lleida**

Escola Politècnica Superior  
Grau en Enginyeria Informàtica  
Algorítmica i Complexitat

Docent: Aitor Corchero Rodríguez i Jordi Planes Cid

# **PRÀCTICA 2**

Àiax Faura Vilalta  
Leandro Gontad Alvez  
19/05/2019

## Funcionament de l'algorisme

1. Es llegeix la primera paraula del fitxer noisy.txt.
2. Es mira si la paraula es troba al diccionari. Si hi és, s'afegeix al fitxer corrected.txt, sinó vol dir que conté almenys un error i es passa al següent pas.
3. A la paraula se li generen tots els canvis possibles i es guarden a la llista "canvis\_possibles". Per exemple amb la paraula "paf":

**Canvis de lletra**( per a cada posició de la paraula canvia la lletra pel caràcter "-"):

"-af"

"p-f"

"pa-"

**Afegir lletra**( per a cada posició de la paraula inserta el caràcter "-"):

"-paf"

"p-af"

"pa-f"

"paf-"

**Eliminar lletra**( per a cada posició de la paraula elimina la lletra actual):

"af"

"pf"

"pa"

Quedant la llista final:

**canvis\_possibles** = "-af", "p-f", "pa-", "-paf", "p-af", "pa-f", "paf-", "af", "pf", "pa"

Com es pot veure s'utilitza "-" per substituir i afegir lletres per evitar haver de generar totes les paraules possibles amb totes les lletres de l'alfabet utf-8. Sinó per la primera paraula de la llista canvis possibles s'hauria de ficar:

"aau", "bau", "cau" ... "ᄁau" ...

4. Es compara tota la llista *canvis\_possibles* per a cada paraula del diccionari. Per fer la comparació s'agafa la primera paraula del diccionari i la primera paraula de la llista *canvis\_possibles*, i es fa el següent per comparar si son iguals:
  - a. Si no tenen la mateixa llargada són diferents i s'agafa la següent paraula.
  - b. Si la llargada és la mateixa s'obté la posició o posicions on es troba el caràcter "-" en la paraula de *canvis\_possibles*. Per exemple, de la paraula "pa-" obtenim la posició 2.
  - c. S'eliminen els caràcters de les dues paraules per a les posicions obtingudes. Per exemple tenim la paraula de diccionari "pau" i la paraula incorrecta "pa-". Per a les dues s'elimina la lletra de la posició 2 i queda "pa" i "pa".
  - d. Es comparen les dues paraules anteriors i si son iguals vol dir que s'ha trobat una correcció possible, s'afegeix al fitxer corrected.txt i es passa al **pas 6**. Si són diferents s'agafa la següent paraula de *canvis\_possibles* fins a recórrer tota la llista *canvis\_possibles* per a cada una de les paraules del diccionari.

5. Si després de comprovar totes les paraules de *canvis\_possibles* per a totes les paraules del diccionari no es troba cap error vol dir que la paraula que s'està corregint té més d'un error. Llavors es torna al pas 3, però es generen tots els canvis possibles per a cadascuna de les paraules de la llista *canvis\_possibles* i es guarden a la llista *canvis\_possibles*. I es continua des del pas tres. Si al tornar al **pas 5** tampoc s'ha trobat, es va aplicant l'algorisme que genera els canvis possibles sobre la llista anterior, i cada cop que es fa això augmenta en 1, el nombre d'edicions.
6. Com que ja s'ha corregit la paraula s'agafa la següent paraula del fitxer noisy.txt i es torna al **pas 2**, fins llegir totes les paraules del text.