

1. Séquence : repérage d'homologies

Considérer la séquence suivante :

BRIC A BRAC ABRACADABRANTESQUE.

Détecter les homologies

- avec l'algorithme naïf
- avec l'algorithme de Karp
- construire l'arbre des suffixes

2. Détection de plagiat

Trouver la reprise entre les deux textes suivants, en utilisant un algorithme de détection de plagiat avec des n-grammes de mots de longueur 5 :

- Jamais il ne faut se défier des sentiments mauvais en amour, ils sont très salutaires; les femmes ne succombent que sous le coup d'une vertu. *L'enfer est pavé de bonnes intentions* n'est pas un paradoxe de prédicateur.
- *L'enfer est pavé de bonnes intentions*

3. Détection de réutilisations

Trouver la réutilisation en utilisant un algorithme de réemploi avec des n-grammes à trou de lemmes de longueur 3 avec 2 trous

- Le rouge avait détruit les délicieuses teintes diaphanes de ses joues, dont la chair était délicate mais, si elle ne pouvait plus ni rougir ni pâlir, elle avait un nez mince, coupé de narines roses et passionnées, fait pour exprimer l'ironie, la moquerie des servantes de Molière.
- L'arc de ses sourcils, tracé avec une pureté et une finesse incomparable, s'étend sur deux yeux noirs pleins de flammes et d'éclairs tragiques ; le nez mince et droit, coupé d'une narine oblique et passionnément dilatée, s'unit avec son front par une ligne d'une pureté magnifique

1. B
2. R
3. I
4. C
5. A
6. B
7. R
8. A
9. C
10. A
11. B
12. R
13. A
14. C
15. A
16. D
17. A
18. B
19. R
20. A
21. N
22. T
23. E
24. S
25. Q
26. U
27. E