Part C

Locality Sensitive Hashing

a)

S1(abbcba) = {ab,bb,bc,cb,ba}

S2(ccbbca) = {cc,cb,bb,bc,ca}

S3(bbaacb) = {bb,ba,aa,ac,cb}

S4(bbacab) = {bb,ba,ac,ca,ab}

S5(cbbbac) = {cb,bb,bb,ba,ab}

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Element | S1 | S2 | S3 | S4 | S5 |
| aa | 0 | 0 | 1 | 0 | 0 |
| ab | 1 | 0 | 0 | 1 | 1 |
| ac | 0 | 0 | 1 | 1 | 0 |
| ba | 1 | 0 | 1 | 1 | 1 |
| bb | 1 | 1 | 1 | 1 | 1 |
| bc | 1 | 1 | 0 | 0 | 0 |
| ca | 0 | 1 | 0 | 1 | 0 |
| cb | 1 | 1 | 1 | 0 | 1 |
| cc | 0 | 1 | 0 | 0 | 0 |

b)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | S1 | S2 | S3 | S4 | S5 |
| h1(x) | 0 | 0 | 0 | 2 | 0 |
| h2(x) | 0 | 0 | 2 | 0 | 4 |
| h3(x) | 0 | 0 | 0 | 0 | 0 |

H1(x) 🡪 1, 8, 5, 2, 9 , 6 , 3 , 0 , 7

H2(x) 🡪 2, 0, 8, 6, 4 , 2 , 0 , 8 , 6

H3(x) 🡪 2, 8, 4, 0, 6 , 2 , 8 , 4 , 0

c) Only h1(x) = 7x + 1 mod 10 is true permutation.

d)

|  |  |  |
| --- | --- | --- |
| Pair | Col(n)/Col(u) | Sig(n)/Sig(u) |
| 1-2 | 3/7 | 1 |
| 1-3 | 3/7 | 2/3 |
| 1-4 | 3/7 | 2/3 |
| 1-5 | 4/5 | 2/3 |
| 2-3 | 2/8 | 2/3 |
| 2-4 | 2/8 | 2/3 |
| 2-5 | 2/7 | 2/3 |
| 3-4 | 3/7 | 1/3 |
| 3-5 | 3/6 | 2/3 |
| 4-5 | 3/6 | 1/3 |