

Solution 4: Using Needleman Wunsch Algorithm

$S3 = \{P1, P6, P1, P4\}$

$S4 = \{P5, P4, P1, P6, P7\}$

Similarity of (P4, P6) = (P5, P7) = 0.9

gap $d = -10$

The range of score is between -10 to 20

The range of page similarity is between 0 to 1

so we need to scale the Similarity of page P4 and P6 in the range of -10 ~20

so after scaling we get 17

i.e $30 * 0.9 - 10 = 17$

so now the $S(X_i, Y_j)$ i.e $S(P4, P6) = S(P6, P4) = 17$

if pages are similar then $S(X_i, Y_j) = 20$

else $S(X_i, Y_j) = -10$

A(i,j)	A(i-1,j-1)+S(Xi,Yj)	A(i,j-1) + d	A(i-1,j) + d	max((A(i-1,j-1)+S(Xi,Yj)), (A(i,j-1) + d), (A(i-1,j) + d))
P1, P5	$0 - 10 = -10$	$-10 - 10 = -20$	$-10 - 10 = -20$	-10
P1, P4	$-10 - 10 = -20$	$-10 - 10 = -20$	$-20 - 10 = -30$	-20
P1, P1	$-20 + 20 = 0$	$-20 - 10 = -30$	$-30 - 10 = -40$	0
P1, P6	$-30 - 10 = -40$	$0 - 10 = -10$	$-40 - 10 = -50$	-10
P1, P7	$-40 - 10 = -50$	$-10 - 10 = -20$	$-50 - 10 = -60$	-20
P6, P5	$-10 - 10 = -20$	$-20 - 10 = -30$	$-10 - 10 = -20$	-20
P6, P4	$-10 + 17 = 7$	$-20 - 10 = -30$	$-10 - 10 = -20$	7
P6, P1	$-10 - 10 = -20$	$7 - 10 = -3$	$0 - 10 = -10$	-3
P6, P6	$0 + 20 = 20$	$-3 - 10 = -13$	$-10 - 10 = -20$	20
P6, P7	$-10 - 10 = -20$	$20 - 10 = 10$	$-20 - 10 = -30$	10
P1, P5	$-20 - 10 = -30$	$-30 - 10 = -40$	$-20 - 10 = -30$	-30
P1, P4	$-20 - 10 = -30$	$-30 - 10 = -40$	$7 - 10 = -3$	-3
P1, P1	$7 + 20 = 27$	$-3 - 10 = -13$	$-3 - 10 = -13$	27
P1, P6	$-3 - 10 = -13$	$27 - 10 = 17$	$10 - 10 = 0$	17
P1, P7	$20 - 10 = 10$	$17 - 10 = 7$	$0 - 10 = -10$	10
P4, P5	$-30 - 10 = -40$	$-40 - 10 = -50$	$-30 - 10 = -40$	-40
P4, P4	$-30 + 20 = -10$	$-40 - 10 = -50$	$-3 - 10 = -13$	-10
P4, P1	$-3 - 10 = -13$	$17 - 10 = 7$	$27 - 10 = 17$	17

P4, P6`	27+17=44	17-10=7	17-10=7	44
P4, P7	17-10=7	44-10=34	7-10=-3	34

		P5	P4	P1	P6	P7
	0	-10	-20	-30	-40	-50
P1	-10	-10	-20	0	-10	-20
P6	-20	-20	7	-3	20	10
P1	-30	-30	-3	27	17	10
P4	-40	-40	-10	17	44	34

Optimal alignment score is 34

The Best alignment is

P5	P4	P1	P6	P7
P1	P6	P1	P4	-

Session similarity = optimal alignment score /length of longer session =34/5 =6.8

Session similarity = 6.8