# Assignment-6

1. Positional conservation scores from multiple sequence alignment (MSA) of given set of protein sequences are

   Seq-1

| Position | Residue | S1C1 | S1C2 | S1C3 | S1C4 | S1C5 |
|---|---|---|---|---|---|---|
| '1' | ':' | 0 | 1.021 | 5 | 1.019 | 0.943 |
| '2' | 'S' | -0.305 | 0.865 | 3.008 | 0.729 | 0.253 |
| '3' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '4' | 'S' | 0 | 0.953 | 4 | 0.951 | 0.943 |
| '5' | 'D' | -1.121 | 0.527 | 1.149 | 0.492 | -1.597 |
| '6' | 'K' | -1.16 | 0.561 | 1.157 | 0.439 | -1.685 |
| '7' | 'D' | 0 | 0.973 | 6 | 0.972 | 0.943 |
| '8' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '9' | 'A' | -1.169 | 0.541 | 1.868 | 0.465 | -1.705 |
| '10' | 'A' | -0.305 | 0.917 | 4.661 | 0.77 | 0.253 |
| '11' | 'V' | -0.586 | 0.744 | 3.603 | 0.763 | -0.385 |
| '12' | 'R' | -0.305 | 0.871 | 4.504 | 0.744 | 0.253 |
| '13' | 'A' | -1.034 | 0.567 | 1.909 | 0.566 | -1.399 |
| '14' | 'L' | -1.16 | 0.559 | 1.521 | 0.415 | -1.685 |
| '15' | 'W' | -0.305 | 0.943 | 9.306 | 0.888 | 0.253 |
| '16' | 'S' | -0.916 | 0.612 | 2.264 | 0.604 | -1.134 |
| '17' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '18' | 'I' | -0.689 | 0.691 | 3.504 | 0.719 | -0.618 |
| '19' | 'G' | -0.305 | 0.886 | 4.992 | 0.831 | 0.253 |
| '20' | 'K' | -1.034 | 0.613 | 2.405 | 0.502 | -1.399 |
| '21' | 'S' | -0.886 | 0.713 | 4.372 | 0.627 | -1.064 |
| '22' | 'A' | -0.474 | 0.74 | 2.876 | 0.767 | -0.131 |
| '23' | 'D' | -1.034 | 0.624 | 2.24 | 0.517 | -1.399 |
| '24' | 'A' | -0.908 | 0.665 | 2.455 | 0.593 | -1.114 |
| '25' | 'I' | -0.886 | 0.757 | 4.314 | 0.679 | -1.064 |
| '26' | 'G' | 0 | 0.977 | 6 | 0.977 | 0.943 |
| '27' | 'N' | -0.886 | 0.65 | 1.826 | 0.561 | -1.064 |
| '28' | 'D' | -0.305 | 0.922 | 4.512 | 0.784 | 0.253 |
| '29' | 'A' | -0.305 | 0.817 | 3.347 | 0.768 | 0.253 |
| '30' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '31' | 'S' | -0.886 | 0.749 | 3.298 | 0.628 | -1.064 |
| '32' | 'R' | 0 | 1.015 | 5 | 1.011 | 0.943 |
| '33' | 'M' | -0.305 | 0.933 | 4.008 | 0.944 | 0.253 |
| '34' | 'I' | -0.305 | 0.904 | 4.992 | 0.78 | 0.253 |
| '35' | 'V' | -1.295 | 0.47 | 0.959 | 0.373 | -1.991 |
| '36' | 'V' | -1.367 | 0.472 | 1.132 | 0.386 | -2.154 |

| | | | | | | |
|---|---|---|---|---|---|---|
| '37' | 'Y' | -0.474 | 0.83 | 5.14 | 0.718 | -0.131 |
| '38' | 'P' | 0 | 0.993 | 7 | 0.993 | 0.943 |
| '39' | 'Q' | -0.474 | 0.821 | 3.215 | 0.715 | -0.131 |
| '40' | 'T' | 0 | 0.977 | 5 | 0.979 | 0.943 |
| '41' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '42' | 'T' | 0 | 0.977 | 5 | 0.979 | 0.943 |
| '43' | 'Y' | 0 | 1.016 | 7 | 1.012 | 0.943 |
| '44' | 'F' | 0 | 0.988 | 6 | 0.989 | 0.943 |
| '45' | 'S' | -0.474 | 0.822 | 4.521 | 0.712 | -0.131 |
| '46' | 'H' | 0 | 0.969 | 8 | 0.972 | 0.943 |
| '47' | 'W' | -0.6 | 0.823 | 4.24 | 0.711 | -0.417 |
| '48' | 'P' | -0.857 | 0.657 | 2.295 | 0.582 | -1 |
| '49' | 'D' | -0.305 | 0.888 | 5.174 | 0.901 | 0.253 |
| '50' | 'V' | -0.586 | 0.674 | 2.81 | 0.621 | -0.385 |
| '51' | 'T' | -0.6 | 0.776 | 2.752 | 0.605 | -0.417 |
| '52' | 'P' | -0.586 | 0.745 | 3.959 | 0.689 | -0.385 |
| '53' | 'G' | 0 | 0.977 | 6 | 0.977 | 0.943 |
| '54' | 'S' | 0 | 0.953 | 4 | 0.951 | 0.943 |
| '55' | 'P' | -0.6 | 0.73 | 2.322 | 0.628 | -0.417 |
| '56' | 'H' | -0.305 | 0.936 | 4.198 | 0.8 | 0.253 |
| '57' | 'I' | -0.305 | 0.872 | 3.835 | 0.748 | 0.253 |
| '58' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '59' | 'A' | -0.86 | 0.596 | 1.835 | 0.687 | -1.006 |
| '60' | 'H' | 0 | 0.969 | 8 | 0.972 | 0.943 |
| '61' | 'G' | 0 | 0.977 | 6 | 0.977 | 0.943 |
| '62' | 'K' | -0.6 | 0.777 | 3.405 | 0.812 | -0.417 |
| '63' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '64' | 'V' | 0 | 0.956 | 4 | 0.956 | 0.943 |
| '65' | 'M' | -0.76 | 0.66 | 2.124 | 0.612 | -0.778 |
| '66' | 'G' | -0.76 | 0.704 | 2.694 | 0.561 | -0.778 |
| '67' | 'G' | -0.305 | 0.817 | 3.355 | 0.692 | 0.253 |
| '68' | 'I' | -0.305 | 0.825 | 3.669 | 0.71 | 0.253 |
| '69' | 'A' | -0.86 | 0.643 | 2.19 | 0.545 | -1.006 |
| '70' | 'L' | -1.673 | 0.36 | 0.124 | 0.382 | -2.848 |
| '71' | 'A' | 0 | 0.906 | 4 | 0.909 | 0.943 |
| '72' | 'V' | -0.586 | 0.701 | 2.413 | 0.682 | -0.385 |
| '73' | 'S' | -1.594 | 0.406 | 0.934 | 0.359 | -2.67 |
| '74' | 'K' | -0.305 | 0.879 | 6.488 | 0.745 | 0.253 |
| '75' | 'I' | -0.995 | 0.543 | 2.612 | 0.534 | -1.312 |
| '76' | 'D' | -0.305 | 0.888 | 5.331 | 0.888 | 0.253 |
| '77' | 'D' | 0 | 0.973 | 6 | 0.972 | 0.943 |
| '78' | 'L' | -0.76 | 0.675 | 3.14 | 0.683 | -0.778 |

| | | | | | | |
|---|---|---|---|---|---|---|
| '79' | 'K' | -0.6 | 0.807 | 4.446 | 0.618 | -0.417 |
| '80' | 'T' | -0.76 | 0.733 | 3.149 | 0.664 | -0.778 |
| '81' | 'G' | -0.305 | 0.817 | 3.355 | 0.692 | 0.253 |
| '82' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '83' | 'M' | -0.305 | 0.869 | 3.182 | 0.749 | 0.253 |
| '84' | 'E' | -1.414 | 0.451 | 0.678 | 0.428 | -2.262 |
| '85' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '86' | 'S' | 0 | 0.953 | 4 | 0.951 | 0.943 |
| '87' | 'E' | -0.305 | 0.888 | 5.331 | 0.763 | 0.253 |
| '88' | 'Q' | -0.305 | 0.826 | 3.017 | 0.715 | 0.253 |
| '89' | 'H' | 0 | 0.969 | 8 | 0.972 | 0.943 |
| '90' | 'A' | 0 | 0.906 | 4 | 0.909 | 0.943 |
| '91' | 'Y' | -0.76 | 0.731 | 4.967 | 0.618 | -0.778 |
| '92' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '93' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '94' | 'R' | 0 | 1.015 | 5 | 1.011 | 0.943 |
| '95' | 'V' | 0 | 0.956 | 4 | 0.956 | 0.943 |
| '96' | 'D' | 0 | 0.973 | 6 | 0.972 | 0.943 |
| '97' | 'P' | 0 | 0.993 | 7 | 0.993 | 0.943 |
| '98' | 'A' | -0.305 | 0.86 | 3.339 | 0.715 | 0.253 |
| '99' | 'N' | 0 | 1.012 | 6 | 1.011 | 0.943 |
| '100' | 'F' | 0 | 0.988 | 6 | 0.989 | 0.943 |
| '101' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '102' | 'I' | -0.6 | 0.739 | 3.058 | 0.632 | -0.417 |
| '103' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '104' | 'N' | -0.6 | 0.781 | 2.926 | 0.619 | -0.417 |
| '105' | 'H' | -0.305 | 0.886 | 6.653 | 0.839 | 0.253 |
| '106' | 'C' | -0.305 | 0.937 | 7.306 | 0.962 | 0.253 |
| '107' | 'I' | -0.6 | 0.739 | 3.058 | 0.588 | -0.417 |
| '108' | 'L' | -0.305 | 0.825 | 3.678 | 0.783 | 0.253 |
| '109' | 'V' | -0.474 | 0.784 | 2.215 | 0.844 | -0.131 |
| '110' | 'V' | -0.474 | 0.806 | 3.479 | 0.677 | -0.131 |
| '111' | 'I' | -0.6 | 0.735 | 3.24 | 0.58 | -0.417 |
| '112' | 'S' | -0.305 | 0.814 | 3.504 | 0.683 | 0.253 |
| '113' | 'T' | -1.72 | 0.345 | 0.702 | 0.355 | -2.955 |
| '114' | 'M' | -0.305 | 0.885 | 6.322 | 0.769 | 0.253 |
| '115' | 'F' | -0.935 | 0.559 | 1.207 | 0.529 | -1.175 |
| '116' | 'P' | 0 | 0.993 | 7 | 0.993 | 0.943 |
| '117' | 'K' | -1.846 | 0.302 | 0.562 | 0.299 | -3.241 |
| '118' | 'E' | -0.916 | 0.633 | 3.008 | 0.584 | -1.134 |
| '119' | 'F' | -0.305 | 0.893 | 4.992 | 0.837 | 0.253 |
| '120' | 'T' | -0.305 | 0.891 | 4.182 | 0.907 | 0.253 |

| | | | | | | |
|---|---|---|---|---|---|---|
| '121' | 'P' | 0 | 0.993 | 7 | 0.993 | 0.943 |
| '122' | 'E' | -0.76 | 0.66 | 2.182 | 0.589 | -0.778 |
| '123' | 'A' | -0.886 | 0.695 | 2.736 | 0.519 | -1.064 |
| '124' | 'H' | 0 | 0.969 | 8 | 0.972 | 0.943 |
| '125' | 'V' | -0.305 | 0.814 | 3.339 | 0.685 | 0.253 |
| '126' | 'S' | 0 | 0.953 | 4 | 0.951 | 0.943 |
| '127' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '128' | 'D' | 0 | 0.973 | 6 | 0.972 | 0.943 |
| '129' | 'K' | 0 | 0.956 | 5 | 0.95 | 0.943 |
| '130' | 'F' | 0 | 0.988 | 6 | 0.989 | 0.943 |
| '131' | 'L' | -0.305 | 0.822 | 3.355 | 0.839 | 0.253 |
| '132' | 'S' | -0.86 | 0.595 | 2.289 | 0.528 | -1.006 |
| '133' | 'G' | -1.414 | 0.459 | 1.496 | 0.4 | -2.262 |
| '134' | 'V' | 0 | 0.956 | 4 | 0.956 | 0.943 |
| '135' | 'A' | -0.6 | 0.769 | 2.909 | 0.577 | -0.417 |
| '136' | 'L' | -0.6 | 0.796 | 3.231 | 0.615 | -0.417 |
| '137' | 'A' | -0.305 | 0.86 | 3.339 | 0.715 | 0.253 |
| '138' | 'L' | 0 | 0.909 | 4 | 0.916 | 0.943 |
| '139' | 'A' | -0.305 | 0.881 | 4.165 | 0.738 | 0.253 |
| '140' | 'E' | -0.305 | 0.868 | 3.347 | 0.743 | 0.253 |
| '141' | 'R' | -0.305 | 0.871 | 4.504 | 0.744 | 0.253 |
| '142' | 'Y' | 0 | 1.016 | 7 | 1.012 | 0.943 |
| '143' | 'R' | 0 | 1.015 | 5 | 1.011 | 0.943 |

Seq-2

| Position | Residue | S2C1 | S2C2 | S2C3 | S2C4 | S2C5 |
|---|---|---|---|---|---|---|
| 1 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 2 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 3 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 4 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 5 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 6 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 7 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 8 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 9 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 10 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 11 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 12 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 13 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 14 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 15 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 16 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 17 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 18 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 19 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 20 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 21 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 22 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 23 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 24 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 25 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 26 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 27 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 28 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 29 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 30 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 31 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 32 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 33 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 34 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 35 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 36 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 37 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 38 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 39 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 40 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 41 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 42 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 43 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 44 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 45 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 46 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 47 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 48 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 49 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 50 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 51 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 52 | - | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 53 | M | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 54 | A | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 55 | S | -0.974 | 0.648 | 1.328 | 0.666 | -0.56 |
| 56 | K | -1.494 | 0.426 | 1.109 | 0.365 | -1.507 |
| 57 | P | -0.377 | 0.882 | 3.5 | 0.827 | 0.529 |
| 58 | Q | -0.736 | 0.726 | 2.938 | 0.641 | -0.125 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 59 | P | -0.377 | 0.884 | 3.828 | 0.83 | 0.529 |
| 60 | I | -0.736 | 0.769 | 3.172 | 0.713 | -0.125 |
| 61 | A | -0.377 | 0.801 | 3.125 | 0.739 | 0.529 |
| 62 | A | -0.377 | 0.812 | 4.656 | 0.754 | 0.529 |
| 63 | A | -0.377 | 0.812 | 4.656 | 0.754 | 0.529 |
| 64 | N | 0 | 0.995 | 6 | 0.994 | 1.215 |
| 65 | W | -0.349 | 0.909 | 8.963 | 0.876 | 0.58 |
| 66 | K | 0 | 0.962 | 5 | 0.966 | 1.215 |
| 67 | C | -0.937 | 0.679 | 2.395 | 0.639 | -0.492 |
| 68 | N | 0 | 0.995 | 6 | 0.994 | 1.215 |
| 69 | G | -0.349 | 0.835 | 4 | 0.81 | 0.58 |
| 70 | S | -1.311 | 0.51 | 0.79 | 0.53 | -1.173 |
| 71 | E | -1.303 | 0.553 | 1.568 | 0.515 | -1.159 |
| 72 | S | -1.523 | 0.401 | 0.901 | 0.372 | -1.56 |
| 73 | L | -1.581 | 0.481 | 0.519 | 0.482 | -1.666 |
| 74 | L | -1.061 | 0.526 | 2.667 | 0.514 | -0.718 |
| 75 | V | -1.677 | 0.317 | 0.198 | 0.27 | -1.841 |
| 76 | P | -1.003 | 0.64 | 2.025 | 0.597 | -0.612 |
| 77 | L | -0.687 | 0.664 | 3.012 | 0.656 | -0.036 |
| 78 | I | -0.965 | 0.55 | 2.716 | 0.557 | -0.543 |
| 79 | E | -1.735 | 0.362 | 0.136 | 0.393 | -1.946 |
| 80 | T | -1.003 | 0.663 | 1.815 | 0.595 | -0.612 |
| 81 | L | -0.349 | 0.859 | 3.235 | 0.826 | 0.58 |
| 82 | N | -0.349 | 0.888 | 4.988 | 0.885 | 0.58 |
| 83 | A | -1.003 | 0.591 | 1.951 | 0.549 | -0.612 |
| 84 | A | -0.684 | 0.691 | 2.716 | 0.676 | -0.031 |
| 85 | T | -1.889 | 0.305 | 0.383 | 0.302 | -2.227 |
| 86 | F | -1.581 | 0.426 | 0.457 | 0.364 | -1.666 |
| 87 | D | -1.149 | 0.604 | 1.543 | 0.553 | -0.879 |
| 88 | H | -1.677 | 0.372 | 0.383 | 0.348 | -1.841 |
| 89 | D | -1.523 | 0.429 | 1.259 | 0.421 | -1.56 |
| 90 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 91 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 92 | V | -0.687 | 0.645 | 2.222 | 0.684 | -0.036 |
| 93 | Q | -0.849 | 0.671 | 3.432 | 0.623 | -0.331 |
| 94 | C | -0.349 | 0.829 | 3.074 | 0.795 | 0.58 |
| 95 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 96 | V | -1.061 | 0.545 | 2.136 | 0.563 | -0.718 |
| 97 | A | -1.003 | 0.594 | 2.062 | 0.557 | -0.612 |
| 98 | P | -0.349 | 0.893 | 5.049 | 0.894 | 0.58 |
| 99 | T | -0.349 | 0.889 | 5.395 | 0.855 | 0.58 |
| 100 | F | -1.215 | 0.498 | 0.889 | 0.51 | -0.998 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 101 | L | -1.677 | 0.321 | 0.593 | 0.294 | -1.841 |
| 102 | H | -0.684 | 0.808 | 5.247 | 0.771 | -0.031 |
| 103 | I | -0.687 | 0.664 | 3.012 | 0.686 | -0.036 |
| 104 | P | -1.149 | 0.595 | 1.272 | 0.564 | -0.879 |
| 105 | M | -1.465 | 0.504 | 1.704 | 0.492 | -1.454 |
| 106 | T | -1.149 | 0.509 | 1.704 | 0.45 | -0.879 |
| 107 | K | -0.937 | 0.625 | 2.37 | 0.582 | -0.492 |
| 108 | A | -1.215 | 0.541 | 1.556 | 0.505 | -0.998 |
| 109 | R | -1.427 | 0.462 | 0.889 | 0.399 | -1.385 |
| 110 | L | -0.349 | 0.85 | 3.407 | 0.809 | 0.58 |
| 111 | T | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 112 | N | -1.523 | 0.455 | 0.877 | 0.44 | -1.56 |
| 113 | P | -1.149 | 0.584 | 1.667 | 0.551 | -0.879 |
| 114 | K | -1.427 | 0.467 | 1.469 | 0.421 | -1.385 |
| 115 | F | -1.311 | 0.477 | 1.852 | 0.497 | -1.173 |
| 116 | Q | -1.831 | 0.321 | -0.222 | 0.292 | -2.121 |
| 117 | I | -1.003 | 0.611 | 2.531 | 0.549 | -0.612 |
| 118 | A | -0.684 | 0.691 | 2.716 | 0.623 | -0.031 |
| 119 | A | -0.349 | 0.798 | 3.235 | 0.793 | 0.58 |
| 120 | Q | 0 | 0.994 | 5 | 0.993 | 1.215 |
| 121 | N | 0 | 0.995 | 6 | 0.994 | 1.215 |
| 122 | A | -0.687 | 0.669 | 3.568 | 0.649 | -0.036 |
| 123 | I | -0.849 | 0.723 | 4 | 0.666 | -0.331 |
| 124 | : | -1.074 | 0.596 | 1.234 | 0.506 | -0.741 |
| 125 | T | -1.149 | 0.532 | 0.827 | 0.465 | -0.879 |
| 126 | R | -1.523 | 0.408 | 0.654 | 0.38 | -1.56 |
| 127 | S | -1.369 | 0.453 | 1.358 | 0.446 | -1.279 |
| 128 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 129 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 130 | F | 0 | 1.001 | 6 | 1.006 | 1.215 |
| 131 | T | 0 | 0.983 | 5 | 0.984 | 1.215 |
| 132 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 133 | E | 0 | 0.962 | 5 | 0.96 | 1.215 |
| 134 | V | -1.149 | 0.559 | 2 | 0.478 | -0.879 |
| 135 | S | -0.349 | 0.873 | 3.235 | 0.84 | 0.58 |
| 136 | L | -1.149 | 0.562 | 1.568 | 0.469 | -0.879 |
| 137 | Q | -1.465 | 0.391 | 0.741 | 0.375 | -1.454 |
| 138 | I | -0.684 | 0.811 | 3.235 | 0.744 | -0.031 |
| 139 | L | -0.937 | 0.601 | 2.914 | 0.557 | -0.492 |
| 140 | K | -0.349 | 0.849 | 3.605 | 0.848 | 0.58 |
| 141 | D | -0.349 | 0.883 | 5.012 | 0.878 | 0.58 |
| 142 | Y | -1.523 | 0.43 | 1.123 | 0.415 | -1.56 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 143 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 144 | I | -1.003 | 0.58 | 1.778 | 0.529 | -0.612 |
| 145 | S | -1.677 | 0.406 | 0.778 | 0.368 | -1.841 |
| 146 | W | 0 | 1.013 | 11 | 1.013 | 1.215 |
| 147 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 148 | V | -0.687 | 0.644 | 3.506 | 0.676 | -0.036 |
| 149 | L | 0 | 0.963 | 4 | 0.959 | 1.215 |
| 150 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 151 | H | 0 | 1.017 | 8 | 1.017 | 1.215 |
| 152 | S | 0 | 0.979 | 4 | 0.977 | 1.215 |
| 153 | E | 0 | 0.962 | 5 | 0.96 | 1.215 |
| 154 | R | 0 | 0.999 | 5 | 1 | 1.215 |
| 155 | R | 0 | 0.999 | 5 | 1 | 1.215 |
| 156 | L | -1.149 | 0.592 | 1.728 | 0.522 | -0.879 |
| 157 | Y | -1.273 | 0.484 | 1.827 | 0.463 | -1.104 |
| 158 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 159 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 160 | Y | -0.849 | 0.703 | 3.346 | 0.65 | -0.331 |
| 161 | G | -0.684 | 0.735 | 3.074 | 0.682 | -0.031 |
| 162 | E | 0 | 0.962 | 5 | 0.96 | 1.215 |
| 163 | T | -0.684 | 0.768 | 2.704 | 0.7 | -0.031 |
| 164 | N | -0.53 | 0.796 | 4.272 | 0.765 | 0.25 |
| 165 | E | -1.149 | 0.546 | 1.605 | 0.484 | -0.879 |
| 166 | I | -1.149 | 0.577 | 1.642 | 0.508 | -0.879 |
| 167 | V | -0.53 | 0.762 | 3.654 | 0.726 | 0.25 |
| 168 | A | -0.687 | 0.601 | 2.42 | 0.643 | -0.036 |
| 169 | E | -1.311 | 0.505 | 1.914 | 0.499 | -1.173 |
| 170 | K | 0 | 0.962 | 5 | 0.966 | 1.215 |
| 171 | V | -0.995 | 0.535 | 1.679 | 0.492 | -0.598 |
| 172 | A | -1.303 | 0.475 | 1.21 | 0.465 | -1.159 |
| 173 | Q | -1.149 | 0.638 | 3.346 | 0.569 | -0.879 |
| 174 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 175 | C | -0.349 | 0.861 | 3.074 | 0.824 | 0.58 |
| 176 | A | -1.427 | 0.419 | 1.136 | 0.381 | -1.385 |
| 177 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 178 | A | -1.061 | 0.514 | 1.63 | 0.491 | -0.718 |
| 179 | G | -0.349 | 0.837 | 4.79 | 0.813 | 0.58 |
| 180 | F | -1.003 | 0.642 | 2.444 | 0.566 | -0.612 |
| 181 | H | -0.937 | 0.592 | 1.444 | 0.573 | -0.492 |
| 182 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 183 | I | -0.349 | 0.858 | 3.802 | 0.849 | 0.58 |
| 184 | V | -1.303 | 0.492 | 0.901 | 0.401 | -1.159 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 185 | C | 0 | 1.017 | 9 | 1.018 | 1.215 |
| 186 | V | -0.53 | 0.762 | 3.654 | 0.726 | 0.25 |
| 187 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 188 | E | 0 | 0.962 | 5 | 0.96 | 1.215 |
| 189 | T | -0.965 | 0.592 | 1.333 | 0.598 | -0.543 |
| 190 | N | -0.349 | 0.859 | 2.642 | 0.819 | 0.58 |
| 191 | E | -0.637 | 0.706 | 3.778 | 0.782 | 0.056 |
| 192 | E | -0.349 | 0.857 | 4.407 | 0.85 | 0.58 |
| 193 | R | -0.349 | 0.89 | 4.407 | 0.856 | 0.58 |
| 194 | E | -0.349 | 0.853 | 4.21 | 0.816 | 0.58 |
| 195 | A | -0.349 | 0.802 | 3.407 | 0.766 | 0.58 |
| 196 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 197 | R | -1.677 | 0.414 | -0.222 | 0.387 | -1.841 |
| 198 | T | 0 | 0.983 | 5 | 0.984 | 1.215 |
| 199 | A | -1.831 | 0.35 | -0.062 | 0.283 | -2.121 |
| 200 | A | -1.311 | 0.452 | 0.988 | 0.454 | -1.173 |
| 201 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 202 | V | -0.684 | 0.726 | 1.84 | 0.692 | -0.031 |
| 203 | L | -1.273 | 0.497 | 0.395 | 0.44 | -1.104 |
| 204 | T | -1.215 | 0.533 | 1.309 | 0.562 | -0.998 |
| 205 | Q | 0 | 0.994 | 5 | 0.993 | 1.215 |
| 206 | L | -0.965 | 0.604 | 1.543 | 0.623 | -0.543 |
| 207 | A | -1.427 | 0.481 | 0.951 | 0.424 | -1.385 |
| 208 | A | -0.637 | 0.629 | 2.222 | 0.712 | 0.056 |
| 209 | V | -1.149 | 0.562 | 2.407 | 0.491 | -0.879 |
| 210 | A | -1.003 | 0.581 | 1.852 | 0.507 | -0.612 |
| 211 | Q | -0.995 | 0.612 | 3.037 | 0.569 | -0.598 |
| 212 | K | -1.215 | 0.519 | 1.593 | 0.511 | -0.998 |
| 213 | L | -1.215 | 0.484 | 1.802 | 0.458 | -0.998 |
| 214 | S | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 215 | K | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 216 | E | -1.003 | 0.635 | 2.136 | 0.568 | -0.612 |
| 217 | A | -0.684 | 0.766 | 3.407 | 0.718 | -0.031 |
| 218 | W | 0 | 1.013 | 11 | 1.013 | 1.215 |
| 219 | S | -1.003 | 0.658 | 2.086 | 0.617 | -0.612 |
| 220 | R | -1.215 | 0.549 | 2 | 0.558 | -0.998 |
| 221 | V | -0.349 | 0.822 | 3.802 | 0.815 | 0.58 |
| 222 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 223 | I | -1.099 | 0.495 | 2.667 | 0.518 | -0.787 |
| 224 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 225 | Y | 0 | 1.013 | 7 | 1.011 | 1.215 |
| 226 | E | 0 | 0.962 | 5 | 0.96 | 1.215 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 227 | P | 0 | 0.995 | 7 | 0.997 | 1.215 |
| 228 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 229 | W | 0 | 1.013 | 11 | 1.013 | 1.215 |
| 230 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 231 | I | 0 | 0.971 | 4 | 0.97 | 1.215 |
| 232 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 233 | T | 0 | 0.983 | 5 | 0.984 | 1.215 |
| 234 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 235 | K | -0.684 | 0.75 | 2.914 | 0.718 | -0.031 |
| 236 | V | -0.937 | 0.591 | 2.037 | 0.535 | -0.492 |
| 237 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 238 | T | -0.349 | 0.876 | 4.198 | 0.87 | 0.58 |
| 239 | P | -0.349 | 0.884 | 5.21 | 0.88 | 0.58 |
| 240 | Q | -1.215 | 0.53 | 2.185 | 0.514 | -0.998 |
| 241 | Q | -0.349 | 0.888 | 4.025 | 0.853 | 0.58 |
| 242 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 243 | Q | -0.349 | 0.885 | 4.407 | 0.852 | 0.58 |
| 244 | E | -0.684 | 0.753 | 3.852 | 0.689 | -0.031 |
| 245 | V | -0.349 | 0.822 | 3.802 | 0.788 | 0.58 |
| 246 | H | 0 | 1.017 | 8 | 1.017 | 1.215 |
| 247 | E | -0.637 | 0.669 | 2.222 | 0.641 | 0.056 |
| 248 | L | -1.677 | 0.389 | -0.074 | 0.351 | -1.841 |
| 249 | L | -0.849 | 0.656 | 2.914 | 0.592 | -0.331 |
| 250 | R | 0 | 0.999 | 5 | 1 | 1.215 |
| 251 | R | -1.677 | 0.35 | 0.383 | 0.319 | -1.841 |
| 252 | W | -0.53 | 0.822 | 7.296 | 0.792 | 0.25 |
| 253 | V | -0.349 | 0.85 | 3.407 | 0.808 | 0.58 |
| 254 | R | -1.149 | 0.545 | 1.827 | 0.476 | -0.879 |
| 255 | S | -1.303 | 0.564 | 1.778 | 0.525 | -1.159 |
| 256 | K | -0.637 | 0.724 | 3.222 | 0.69 | 0.056 |
| 257 | L | -1.149 | 0.516 | 2.173 | 0.471 | -0.879 |
| 258 | G | -0.849 | 0.673 | 2.296 | 0.656 | -0.331 |
| 259 | T | -1.215 | 0.555 | 1.42 | 0.532 | -0.998 |
| 260 | D | -1.677 | 0.333 | 0.543 | 0.319 | -1.841 |
| 261 | I | -0.684 | 0.707 | 3.012 | 0.633 | -0.031 |
| 262 | A | -0.349 | 0.802 | 3.407 | 0.796 | 0.58 |
| 263 | A | -1.465 | 0.426 | 1.012 | 0.405 | -1.454 |
| 264 | Q | -1.003 | 0.651 | 2.099 | 0.575 | -0.612 |
| 265 | L | -0.637 | 0.713 | 2.222 | 0.68 | 0.056 |
| 266 | R | 0 | 0.999 | 5 | 1 | 1.215 |
| 267 | I | 0 | 0.971 | 4 | 0.97 | 1.215 |
| 268 | L | -0.937 | 0.612 | 1.963 | 0.586 | -0.492 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 269 | Y | 0 | 1.013 | 7 | 1.011 | 1.215 |
| 270 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 271 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 272 | S | 0 | 0.979 | 4 | 0.977 | 1.215 |
| 273 | V | -0.349 | 0.815 | 3.21 | 0.776 | 0.58 |
| 274 | T | -0.684 | 0.773 | 3.346 | 0.711 | -0.031 |
| 275 | A | -0.965 | 0.53 | 1.765 | 0.521 | -0.543 |
| 276 | K | -1.003 | 0.592 | 1.642 | 0.518 | -0.612 |
| 277 | N | -0.637 | 0.733 | 3.222 | 0.814 | 0.056 |
| 278 | A | -0.687 | 0.669 | 3.568 | 0.65 | -0.036 |
| 279 | R | -1.303 | 0.523 | 1.593 | 0.429 | -1.159 |
| 280 | T | -0.849 | 0.666 | 2.272 | 0.596 | -0.331 |
| 281 | L | -0.349 | 0.859 | 3.235 | 0.823 | 0.58 |
| 282 | Y | -1.003 | 0.59 | 1.383 | 0.52 | -0.612 |
| 283 | Q | -1.465 | 0.444 | 1.111 | 0.427 | -1.454 |
| 284 | M | -0.937 | 0.637 | 2.457 | 0.601 | -0.492 |
| 285 | R | -0.684 | 0.777 | 3.802 | 0.707 | -0.031 |
| 286 | D | -0.349 | 0.883 | 5.012 | 0.852 | 0.58 |
| 287 | I | -0.687 | 0.637 | 3.506 | 0.638 | -0.036 |
| 288 | N | -0.349 | 0.883 | 5.012 | 0.847 | 0.58 |
| 289 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 290 | F | 0 | 1.001 | 6 | 1.006 | 1.215 |
| 291 | L | 0 | 0.963 | 4 | 0.959 | 1.215 |
| 292 | V | 0 | 0.93 | 4 | 0.93 | 1.215 |
| 293 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 294 | G | 0 | 0.944 | 6 | 0.951 | 1.215 |
| 295 | A | 0 | 0.909 | 4 | 0.905 | 1.215 |
| 296 | S | 0 | 0.979 | 4 | 0.977 | 1.215 |
| 297 | L | -0.349 | 0.862 | 3.617 | 0.829 | 0.58 |
| 298 | K | 0 | 0.962 | 5 | 0.966 | 1.215 |
| 299 | P | 0 | 0.995 | 7 | 0.997 | 1.215 |
| 300 | E | -0.684 | 0.752 | 3.309 | 0.713 | -0.031 |
| 301 | F | 0 | 1.001 | 6 | 1.006 | 1.215 |
| 302 | V | -0.684 | 0.715 | 3.259 | 0.678 | -0.031 |
| 303 | E | -0.849 | 0.682 | 3.469 | 0.628 | -0.331 |
| 304 | I | 0 | 0.971 | 4 | 0.97 | 1.215 |
| 305 | I | -0.349 | 0.858 | 3.802 | 0.824 | 0.58 |
| 306 | E | -0.684 | 0.782 | 3.988 | 0.714 | -0.031 |
| 307 | A | -0.53 | 0.714 | 2.963 | 0.663 | 0.25 |
| 308 | T | -1.465 | 0.462 | 1.062 | 0.457 | -1.454 |
| 309 | K | -1.427 | 0.502 | 1.519 | 0.451 | -1.385 |
| 310 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 311 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 312 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 313 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 314 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 315 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |
| 316 | : | -1.216 | 0.529 | 1.306 | 0.494 | -1 |

2. Tabulating the topmost 10 residues with highest and lowest conservation scores obtained with first method, we get

low S1

| Position | Residue | S1C1 |
|---|---|---|
| '117' | 'K' | -1.846 |
| '113' | 'T' | -1.72 |
| '70' | 'L' | -1.673 |
| '73' | 'S' | -1.594 |
| '84' | 'E' | -1.414 |
| '133' | 'G' | -1.414 |
| '36' | 'V' | -1.367 |
| '35' | 'V' | -1.295 |
| '9' | 'A' | -1.169 |
| '6' | 'K' | -1.16 |

High S1

| Position | Residue | S1C1 |
|---|---|---|
| '1' | ':' | 0 |
| '3' | 'L' | 0 |
| '4' | 'S' | 0 |
| '7' | 'D' | 0 |
| '8' | 'K' | 0 |
| '17' | 'K' | 0 |
| '26' | 'G' | 0 |
| '30' | 'L' | 0 |
| '32' | 'R' | 0 |
| '38' | 'P' | 0 |

Low S2

| Position | Residue | S2C1 |
|---|---|---|
| 85 | T | -1.889 |
| 116 | Q | -1.831 |
| 199 | A | -1.831 |
| 79 | E | -1.735 |
| 75 | V | -1.677 |
| 88 | H | -1.677 |
| 101 | L | -1.677 |
| 145 | S | -1.677 |
| 197 | R | -1.677 |
| 248 | L | -1.677 |

High S2

| Position | Residue | S2C1 |
|---|---|---|
| 64 | N | 0 |
| 66 | K | 0 |
| 68 | N | 0 |
| 95 | V | 0 |
| 120 | Q | 0 |
| 121 | N | 0 |
| 128 | G | 0 |
| 129 | A | 0 |
| 130 | F | 0 |
| 131 | T | 0 |

3. The Program for computing conservation scores from MSA

```
import numpy as np
def Con_score_MSA(dat):
    totsq=len(dat)
    lensq=len(dat[0])
    UNFreq=[[0 for i in range(len(totsq))] for j in
range(lensq)]
```

```
AA=['A','C','D','E','F','G','H','I','K','L','M','N','P','
Q','R','S','T','V','W','Y']
    AAFreq=[0 for i in range(len(AA))]

    blosum62 = {
 ('W', 'F'): 1, ('L', 'R'): -2, ('S', 'P'): -1, ('V',
'T'): 0,
 ('Q', 'Q'): 5, ('N', 'A'): -2, ('Z', 'Y'): -2, ('W',
'R'): -3,
 ('Q', 'A'): -1, ('S', 'D'): 0, ('H', 'H'): 8, ('S',
'H'): -1,
 ('H', 'D'): -1, ('L', 'N'): -3, ('W', 'A'): -3, ('Y',
'M'): -1,
 ('G', 'R'): -2, ('Y', 'I'): -1, ('Y', 'E'): -2, ('B',
'Y'): -3,
 ('Y', 'A'): -2, ('V', 'D'): -3, ('B', 'S'): 0, ('Y',
'Y'): 7,
 ('G', 'N'): 0, ('E', 'C'): -4, ('Y', 'Q'): -1, ('Z',
'Z'): 4,
 ('V', 'A'): 0, ('C', 'C'): 9, ('M', 'R'): -1, ('V',
'E'): -2,
 ('T', 'N'): 0, ('P', 'P'): 7, ('V', 'I'): 3, ('V', 'S'):
-2,
 ('Z', 'P'): -1, ('V', 'M'): 1, ('T', 'F'): -2, ('V',
'Q'): -2,
 ('K', 'K'): 5, ('P', 'D'): -1, ('I', 'H'): -3, ('I',
'D'): -3,
 ('T', 'R'): -1, ('P', 'L'): -3, ('K', 'G'): -2, ('M',
'N'): -2,
 ('P', 'H'): -2, ('F', 'Q'): -3, ('Z', 'G'): -2, ('X',
'L'): -1,
 ('T', 'M'): -1, ('Z', 'C'): -3, ('X', 'H'): -1, ('D',
'R'): -2,
 ('B', 'W'): -4, ('X', 'D'): -1, ('Z', 'K'): 1, ('F',
'A'): -2,
 ('Z', 'W'): -3, ('F', 'E'): -3, ('D', 'N'): 1, ('B',
'K'): 0,
 ('X', 'X'): -1, ('F', 'I'): 0, ('B', 'G'): -1, ('X',
'T'): 0,
 ('F', 'M'): 0, ('B', 'C'): -3, ('Z', 'I'): -3, ('Z',
'V'): -2,
 ('S', 'S'): 4, ('L', 'Q'): -2, ('W', 'E'): -3, ('Q',
'R'): 1,
 ('N', 'N'): 6, ('W', 'M'): -1, ('Q', 'C'): -3, ('W',
'I'): -3,
 ('S', 'C'): -1, ('L', 'A'): -1, ('S', 'G'): 0, ('L',
'E'): -3,
 ('W', 'Q'): -2, ('H', 'G'): -2, ('S', 'K'): 0, ('Q',
'N'): 0,
```

```
 ('N', 'R'): 0, ('H', 'C'): -3, ('Y', 'N'): -2, ('G',
'Q'): -2,
 ('Y', 'F'): 3, ('C', 'A'): 0, ('V', 'L'): 1, ('G', 'E'):
-2,
 ('G', 'A'): 0, ('K', 'R'): 2, ('E', 'D'): 2, ('Y', 'R'):
-2,
 ('M', 'Q'): 0, ('T', 'I'): -1, ('C', 'D'): -3, ('V',
'F'): -1,
 ('T', 'A'): 0, ('T', 'P'): -1, ('B', 'P'): -2, ('T',
'E'): -1,
 ('V', 'N'): -3, ('P', 'G'): -2, ('M', 'A'): -1, ('K',
'H'): -1,
 ('V', 'R'): -3, ('P', 'C'): -3, ('M', 'E'): -2, ('K',
'L'): -2,
 ('V', 'V'): 4, ('M', 'I'): 1, ('T', 'Q'): -1, ('I',
'G'): -4,
 ('P', 'K'): -1, ('M', 'M'): 5, ('K', 'D'): -1, ('I',
'C'): -1,
 ('Z', 'D'): 1, ('F', 'R'): -3, ('X', 'K'): -1, ('Q',
'D'): 0,
 ('X', 'G'): -1, ('Z', 'L'): -3, ('X', 'C'): -2, ('Z',
'H'): 0,
 ('B', 'L'): -4, ('B', 'H'): 0, ('F', 'F'): 6, ('X',
'W'): -2,
 ('B', 'D'): 4, ('D', 'A'): -2, ('S', 'L'): -2, ('X',
'S'): 0,
 ('F', 'N'): -3, ('S', 'R'): -1, ('W', 'D'): -4, ('V',
'Y'): -1,
 ('W', 'L'): -2, ('H', 'R'): 0, ('W', 'H'): -2, ('H',
'N'): 1,
 ('W', 'T'): -2, ('T', 'T'): 5, ('S', 'F'): -2, ('W',
'P'): -4,
 ('L', 'D'): -4, ('B', 'I'): -3, ('L', 'H'): -3, ('S',
'N'): 1,
 ('B', 'T'): -1, ('L', 'L'): 4, ('Y', 'K'): -2, ('E',
'Q'): 2,
 ('Y', 'G'): -3, ('Z', 'S'): 0, ('Y', 'C'): -2, ('G',
'D'): -1,
 ('B', 'V'): -3, ('E', 'A'): -1, ('Y', 'W'): 2, ('E',
'E'): 5,
 ('Y', 'S'): -2, ('C', 'N'): -3, ('V', 'C'): -1, ('T',
'H'): -2,
 ('P', 'R'): -2, ('V', 'G'): -3, ('T', 'L'): -1, ('V',
'K'): -2,
 ('K', 'Q'): 1, ('R', 'A'): -1, ('I', 'R'): -3, ('T',
'D'): -1,
 ('P', 'F'): -4, ('I', 'N'): -3, ('K', 'I'): -3, ('M',
'D'): -3,
 ('V', 'W'): -3, ('W', 'W'): 11, ('M', 'H'): -2, ('P',
'N'): -2,
```

```
 ('K', 'A'): -1, ('M', 'L'): 2, ('K', 'E'): 1, ('Z',
'E'): 4,
 ('X', 'N'): -1, ('Z', 'A'): -1, ('Z', 'M'): -1, ('X',
'F'): -1,
 ('K', 'C'): -3, ('B', 'Q'): 0, ('X', 'B'): -1, ('B',
'M'): -3,
 ('F', 'C'): -2, ('Z', 'Q'): 3, ('X', 'Z'): -1, ('F',
'G'): -3,
 ('B', 'E'): 1, ('X', 'V'): -1, ('F', 'K'): -3, ('B',
'A'): -2,
 ('X', 'R'): -1, ('D', 'D'): 6, ('W', 'G'): -2, ('Z',
'F'): -3,
 ('S', 'Q'): 0, ('W', 'C'): -2, ('W', 'K'): -3, ('H',
'Q'): 0,
 ('L', 'C'): -1, ('W', 'N'): -4, ('S', 'A'): 1, ('L',
'G'): -4,
 ('W', 'S'): -3, ('S', 'E'): 0, ('H', 'E'): 0, ('S',
'I'): -2,
 ('H', 'A'): -2, ('S', 'M'): -1, ('Y', 'L'): -1, ('Y',
'H'): 2,
 ('Y', 'D'): -3, ('E', 'R'): 0, ('X', 'P'): -2, ('G',
'G'): 6,
 ('G', 'C'): -3, ('E', 'N'): 0, ('Y', 'T'): -2, ('Y',
'P'): -3,
 ('T', 'K'): -1, ('A', 'A'): 4, ('P', 'Q'): -1, ('T',
'C'): -1,
 ('V', 'H'): -3, ('T', 'G'): -2, ('I', 'Q'): -3, ('Z',
'T'): -1,
 ('C', 'R'): -3, ('V', 'P'): -2, ('P', 'E'): -1, ('M',
'C'): -1,
 ('K', 'N'): 0, ('I', 'I'): 4, ('P', 'A'): -1, ('M',
'G'): -3,
 ('T', 'S'): 1, ('I', 'E'): -3, ('P', 'M'): -2, ('M',
'K'): -1,
 ('I', 'A'): -1, ('P', 'I'): -3, ('R', 'R'): 5, ('X',
'M'): -1,
 ('L', 'I'): 2, ('X', 'I'): -1, ('Z', 'B'): 1, ('X',
'E'): -1,
 ('Z', 'N'): 0, ('X', 'A'): 0, ('B', 'R'): -1, ('B',
'N'): 3,
 ('F', 'D'): -3, ('X', 'Y'): -1, ('Z', 'R'): 0, ('F',
'H'): -1,
 ('B', 'F'): -3, ('F', 'L'): 0, ('X', 'Q'): -1, ('B',
'B'): 4
 }

    for i in range(lensq):
        for j in range(totsq):
            if dat[j][i].isalpha()==1:
                a=AA.index(str(dat[j][i]))
                UNFreq[i][a]+=1
```

```
    UNFreq=np.divide(UNFreq, totsq)
    for i in range(lensq):
        for j in range(totsq):
            if dat[j][i].isalpha()==1:
                a=AA.index(str(dat[j][i]))
                AAFreq[a]+=1
    AAFreq=np.divide(AAFreq, totsq*lensq)

    E=[0]*lensq
    V=[0]*lensq
    S=[0]*lensq
    for i in range(lensq):
        for j in range(len(AA)):
            if UNFreq[i][j]==0:
                E[i]+=0
            else:
                E[i]+=UNFreq[i][j]*np.log(UNFreq[i][j])
            V[i]+=(UNFreq[i][j]-AAFreq[j])**2
            for k in range(len(AA)):
                pairs=(AA[j],AA[k])
                if pairs in blosum62:
                    pairs=(AA[j],AA[k])
                else:
                    pairs=(AA[k],AA[j])

S[i]+=UNFreq[i][j]*UNFreq[i][k]*blosum62[pairs]
    for i in range(lensq):
        V[i]=V[i]**0.5
        S[i]=S[i]**0.5

        print('Conservation Score: Entropy=', E)
        print('Conservation Score: Variance=', V)
        print('Conservation Score: Sum of pairs=', S)
        return None

if __name__=="__main__":
    data = ['AAANWKCNGSESLLVPLIETLNAATFDHD--
VQCVVAPTFLHIPMTKARLTNPKFQIAAQ',
         '---
NWKCNLSKADIAELVSAFNAAPPIDAAHVQVVVAPPAVYLDSTRQAL-
RADFDTSAQ',
         'VGGNFKLNGSKQSIKEIVERLNTASIPEN--
VEVVICPPATYLDYSVSLVKKPQVTVGAQ',

'VGGNWKCNGTTDQVEKIVKTLNEGQVPPSDVVEVVVSPPYVFLPVVKSQL-
RQEFHVAAQ',
         'VGGNWKMNGDYASVDGIVTFLNASADNSS--
VDVVVAPPAPYLAYAKSKL-KAGVLVAAQ',
         'VGGNWKMNGDQKSIAEIAKTLSSAALDPN--
TEVVIGCPAIYLMYARNLL-PCELGLAGQ',
```

```
        'VGGNWKMNGRKKCLGELICTLNAANVPAG--
TEVVCAPPTAYIDFARQKL-DPKIAVAAQ',
        'VGGNWKMNGRKQSLGELIGTLNAAKVPAD--
TEVVCAPPTAYIDFARQKL-DPKIAVAAQ',
        'VGGNWKMNGRKKNLGELITTLNAAKVPAD--
TEVVCAPPTAYIDFARQKL-DPKIAVAAQ']
Con_score_MSA(data)
```

4. To Compare the MSA from Clustal Omega, MAFFT, and MUSCLE, we use following code

```
import numpy as np
def Con_score_MSA(dat):
    totsq=len(dat)
    lensq=len(dat[0])
    UNFreq=[[0 for i in range(len(totsq))] for j in
range(lensq)]

AA=['A','C','D','E','F','G','H','I','K','L','M','N','P','
Q','R','S','T','V','W','Y']
    AAFreq=[0 for i in range(len(AA))]

    blosum62 = {
 ('W', 'F'): 1, ('L', 'R'): -2, ('S', 'P'): -1, ('V',
'T'): 0,
 ('Q', 'Q'): 5, ('N', 'A'): -2, ('Z', 'Y'): -2, ('W',
'R'): -3,
 ('Q', 'A'): -1, ('S', 'D'): 0, ('H', 'H'): 8, ('S',
'H'): -1,
 ('H', 'D'): -1, ('L', 'N'): -3, ('W', 'A'): -3, ('Y',
'M'): -1,
 ('G', 'R'): -2, ('Y', 'I'): -1, ('Y', 'E'): -2, ('B',
'Y'): -3,
 ('Y', 'A'): -2, ('V', 'D'): -3, ('B', 'S'): 0, ('Y',
'Y'): 7,
 ('G', 'N'): 0, ('E', 'C'): -4, ('Y', 'Q'): -1, ('Z',
'Z'): 4,
 ('V', 'A'): 0, ('C', 'C'): 9, ('M', 'R'): -1, ('V',
'E'): -2,
 ('T', 'N'): 0, ('P', 'P'): 7, ('V', 'I'): 3, ('V', 'S'):
-2,
 ('Z', 'P'): -1, ('V', 'M'): 1, ('T', 'F'): -2, ('V',
'Q'): -2,
 ('K', 'K'): 5, ('P', 'D'): -1, ('I', 'H'): -3, ('I',
'D'): -3,
 ('T', 'R'): -1, ('P', 'L'): -3, ('K', 'G'): -2, ('M',
'N'): -2,
 ('P', 'H'): -2, ('F', 'Q'): -3, ('Z', 'G'): -2, ('X',
'L'): -1,
 ('T', 'M'): -1, ('Z', 'C'): -3, ('X', 'H'): -1, ('D',
'R'): -2,
```

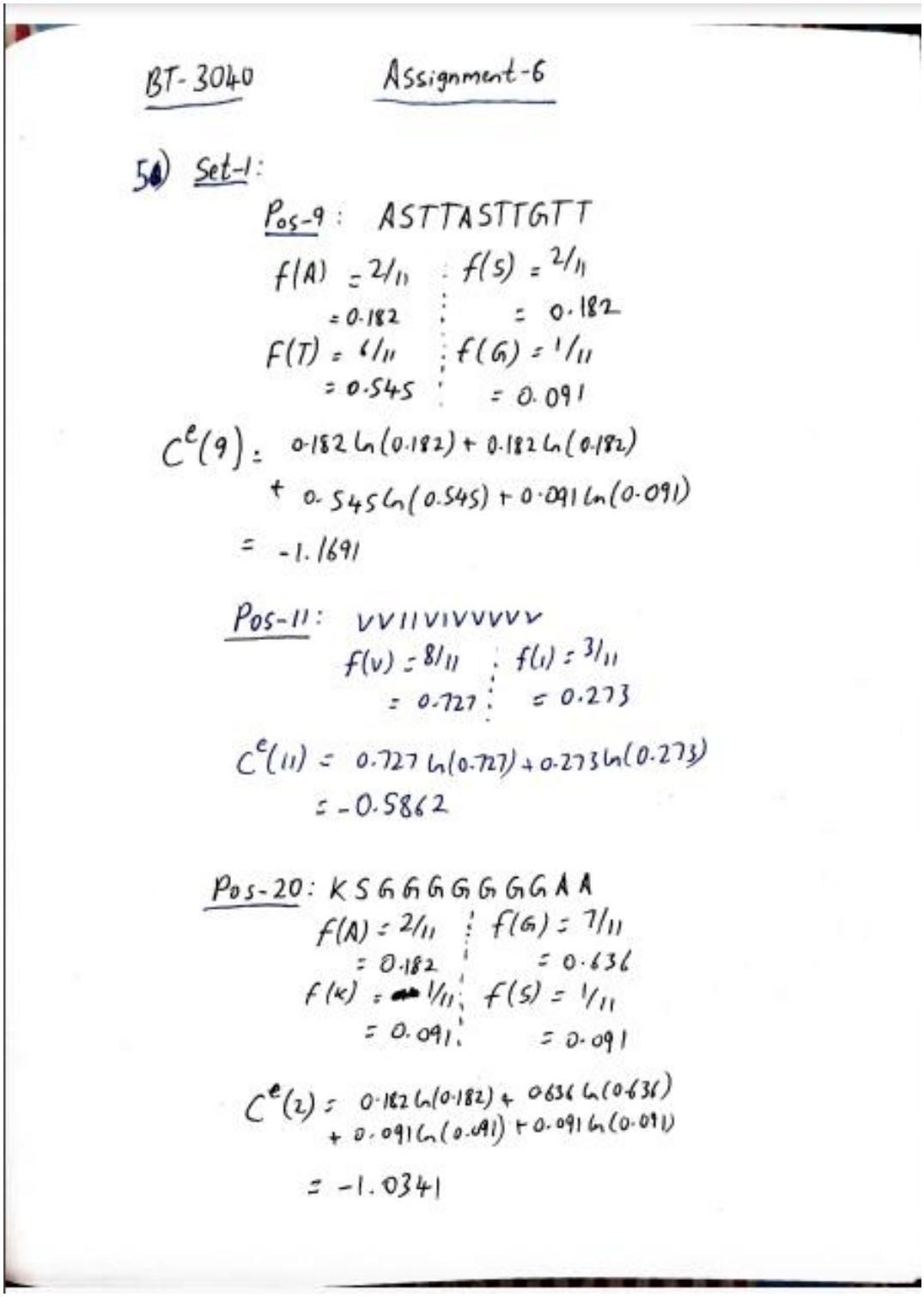
```
 ('B', 'W'): -4, ('X', 'D'): -1, ('Z', 'K'): 1, ('F',
'A'): -2,
 ('Z', 'W'): -3, ('F', 'E'): -3, ('D', 'N'): 1, ('B',
'K'): 0,
 ('X', 'X'): -1, ('F', 'I'): 0, ('B', 'G'): -1, ('X',
'T'): 0,
 ('F', 'M'): 0, ('B', 'C'): -3, ('Z', 'I'): -3, ('Z',
'V'): -2,
 ('S', 'S'): 4, ('L', 'Q'): -2, ('W', 'E'): -3, ('Q',
'R'): 1,
 ('N', 'N'): 6, ('W', 'M'): -1, ('Q', 'C'): -3, ('W',
'I'): -3,
 ('S', 'C'): -1, ('L', 'A'): -1, ('S', 'G'): 0, ('L',
'E'): -3,
 ('W', 'Q'): -2, ('H', 'G'): -2, ('S', 'K'): 0, ('Q',
'N'): 0,
 ('N', 'R'): 0, ('H', 'C'): -3, ('Y', 'N'): -2, ('G',
'Q'): -2,
 ('Y', 'F'): 3, ('C', 'A'): 0, ('V', 'L'): 1, ('G', 'E'):
-2,
 ('G', 'A'): 0, ('K', 'R'): 2, ('E', 'D'): 2, ('Y', 'R'):
-2,
 ('M', 'Q'): 0, ('T', 'I'): -1, ('C', 'D'): -3, ('V',
'F'): -1,
 ('T', 'A'): 0, ('T', 'P'): -1, ('B', 'P'): -2, ('T',
'E'): -1,
 ('V', 'N'): -3, ('P', 'G'): -2, ('M', 'A'): -1, ('K',
'H'): -1,
 ('V', 'R'): -3, ('P', 'C'): -3, ('M', 'E'): -2, ('K',
'L'): -2,
 ('V', 'V'): 4, ('M', 'I'): 1, ('T', 'Q'): -1, ('I',
'G'): -4,
 ('P', 'K'): -1, ('M', 'M'): 5, ('K', 'D'): -1, ('I',
'C'): -1,
 ('Z', 'D'): 1, ('F', 'R'): -3, ('X', 'K'): -1, ('Q',
'D'): 0,
 ('X', 'G'): -1, ('Z', 'L'): -3, ('X', 'C'): -2, ('Z',
'H'): 0,
 ('B', 'L'): -4, ('B', 'H'): 0, ('F', 'F'): 6, ('X',
'W'): -2,
 ('B', 'D'): 4, ('D', 'A'): -2, ('S', 'L'): -2, ('X',
'S'): 0,
 ('F', 'N'): -3, ('S', 'R'): -1, ('W', 'D'): -4, ('V',
'Y'): -1,
 ('W', 'L'): -2, ('H', 'R'): 0, ('W', 'H'): -2, ('H',
'N'): 1,
 ('W', 'T'): -2, ('T', 'T'): 5, ('S', 'F'): -2, ('W',
'P'): -4,
 ('L', 'D'): -4, ('B', 'I'): -3, ('L', 'H'): -3, ('S',
'N'): 1,
```

```
 ('B', 'T'): -1, ('L', 'L'): 4, ('Y', 'K'): -2, ('E',
'Q'): 2,
 ('Y', 'G'): -3, ('Z', 'S'): 0, ('Y', 'C'): -2, ('G',
'D'): -1,
 ('B', 'V'): -3, ('E', 'A'): -1, ('Y', 'W'): 2, ('E',
'E'): 5,
 ('Y', 'S'): -2, ('C', 'N'): -3, ('V', 'C'): -1, ('T',
'H'): -2,
 ('P', 'R'): -2, ('V', 'G'): -3, ('T', 'L'): -1, ('V',
'K'): -2,
 ('K', 'Q'): 1, ('R', 'A'): -1, ('I', 'R'): -3, ('T',
'D'): -1,
 ('P', 'F'): -4, ('I', 'N'): -3, ('K', 'I'): -3, ('M',
'D'): -3,
 ('V', 'W'): -3, ('W', 'W'): 11, ('M', 'H'): -2, ('P',
'N'): -2,
 ('K', 'A'): -1, ('M', 'L'): 2, ('K', 'E'): 1, ('Z',
'E'): 4,
 ('X', 'N'): -1, ('Z', 'A'): -1, ('Z', 'M'): -1, ('X',
'F'): -1,
 ('K', 'C'): -3, ('B', 'Q'): 0, ('X', 'B'): -1, ('B',
'M'): -3,
 ('F', 'C'): -2, ('Z', 'Q'): 3, ('X', 'Z'): -1, ('F',
'G'): -3,
 ('B', 'E'): 1, ('X', 'V'): -1, ('F', 'K'): -3, ('B',
'A'): -2,
 ('X', 'R'): -1, ('D', 'D'): 6, ('W', 'G'): -2, ('Z',
'F'): -3,
 ('S', 'Q'): 0, ('W', 'C'): -2, ('W', 'K'): -3, ('H',
'Q'): 0,
 ('L', 'C'): -1, ('W', 'N'): -4, ('S', 'A'): 1, ('L',
'G'): -4,
 ('W', 'S'): -3, ('S', 'E'): 0, ('H', 'E'): 0, ('S',
'I'): -2,
 ('H', 'A'): -2, ('S', 'M'): -1, ('Y', 'L'): -1, ('Y',
'H'): 2,
 ('Y', 'D'): -3, ('E', 'R'): 0, ('X', 'P'): -2, ('G',
'G'): 6,
 ('G', 'C'): -3, ('E', 'N'): 0, ('Y', 'T'): -2, ('Y',
'P'): -3,
 ('T', 'K'): -1, ('A', 'A'): 4, ('P', 'Q'): -1, ('T',
'C'): -1,
 ('V', 'H'): -3, ('T', 'G'): -2, ('I', 'Q'): -3, ('Z',
'T'): -1,
 ('C', 'R'): -3, ('V', 'P'): -2, ('P', 'E'): -1, ('M',
'C'): -1,
 ('K', 'N'): 0, ('I', 'I'): 4, ('P', 'A'): -1, ('M',
'G'): -3,
 ('T', 'S'): 1, ('I', 'E'): -3, ('P', 'M'): -2, ('M',
'K'): -1,
```

```
 ('I', 'A'): -1, ('P', 'I'): -3, ('R', 'R'): 5, ('X',
'M'): -1,
 ('L', 'I'): 2, ('X', 'I'): -1, ('Z', 'B'): 1, ('X',
'E'): -1,
 ('Z', 'N'): 0, ('X', 'A'): 0, ('B', 'R'): -1, ('B',
'N'): 3,
 ('F', 'D'): -3, ('X', 'Y'): -1, ('Z', 'R'): 0, ('F',
'H'): -1,
 ('B', 'F'): -3, ('F', 'L'): 0, ('X', 'Q'): -1, ('B',
'B'): 4
 }

    for i in range(lensq):
        for j in range(totsq):
            if dat[j][i].isalpha()==1:
                a=AA.index(str(dat[j][i]))
                UNFreq[i][a]+=1
    UNFreq=np.divide(UNFreq, totsq)
    for i in range(lensq):
        for j in range(totsq):
            if dat[j][i].isalpha()==1:
                a=AA.index(str(dat[j][i]))
                AAFreq[a]+=1
    AAFreq=np.divide(AAFreq, totsq*lensq)

    E=[0]*lensq
    V=[0]*lensq
    S=[0]*lensq
    for i in range(lensq):
        for j in range(len(AA)):
            if UNFreq[i][j]==0:
                E[i]+=0
            else:
                E[i]+=UNFreq[i][j]*np.log(UNFreq[i][j])
            V[i]+=(UNFreq[i][j]-AAFreq[j])**2
            for k in range(len(AA)):
                pairs=(AA[j],AA[k])
                if pairs in blosum62:
                    pairs=(AA[j],AA[k])
                else:
                    pairs=(AA[k],AA[j])

S[i]+=UNFreq[i][j]*UNFreq[i][k]*blosum62[pairs]
    for i in range(lensq):
        V[i]=V[i]**0.5
        S[i]=S[i]**0.5

        print('Conservation Score: Entropy=', E)
        print('Conservation Score: Variance=', V)
        print('Conservation Score: Sum of pairs=', S)
        return None
```

```
compare_results()
```

5.



BT-3040 Assignment-6

5a) Set-1:

Pos-9: ASTTASTTGTT

$f(A) = 2/11 = 0.182$ ; $f(S) = 2/11 = 0.182$

$F(T) = 6/11 = 0.545$ ; $f(G) = 1/11 = 0.091$

$C^e(9) = 0.182\ln(0.182) + 0.182\ln(0.182) + 0.545\ln(0.545) + 0.091\ln(0.091)$

$= -1.1691$

Pos-11: VVIIVIVVVVV

$f(V) = 8/11 = 0.727$ ; $f(I) = 3/11 = 0.273$

$C^e(11) = 0.727\ln(0.727) + 0.273\ln(0.273)$

$= -0.5862$

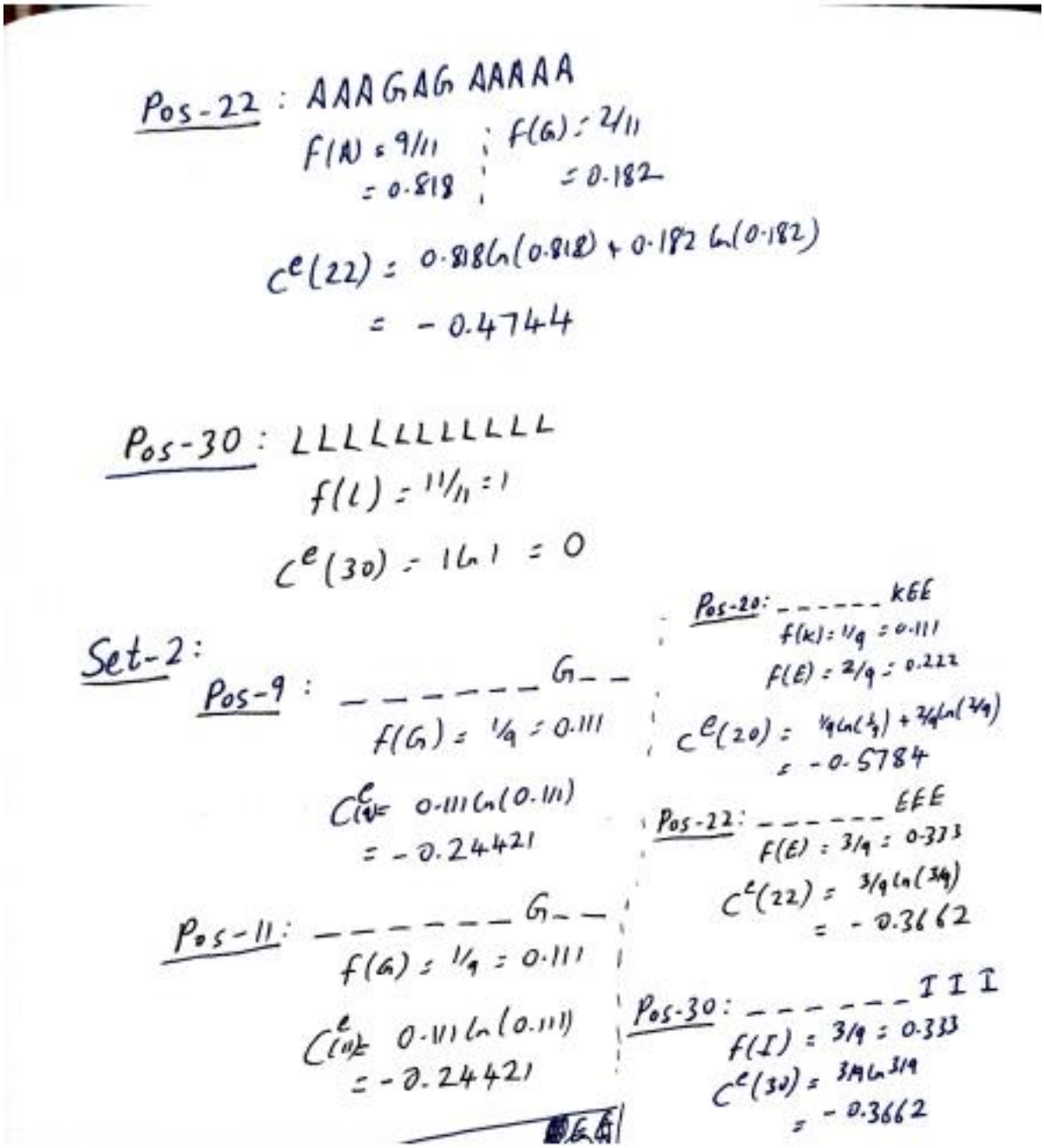
Pos-20: K S G G G G G G G A A

$f(A) = 2/11 = 0.182$ ; $f(G) = 7/11 = 0.636$

$f(K) = 1/11 = 0.091$ ; $f(S) = 1/11 = 0.091$

$C^e(2) = 0.182\ln(0.182) + 0.636\ln(0.636) + 0.091\ln(0.091) + 0.091\ln(0.091)$

$= -1.0341$




6. The answer I got is here in this link
   https://consurf.tau.ac.il/results/1646588529/output.php