2002 Sun Microsystems and TopCoder Collegiate Challenge – Problem Analysis

HiddenSeq:

HiddenSeq is a straightforward coding problem. Attention to detail is all that is required to code a correct solution. The input string may contain up to 50 characters. Each non-blank character in the string is a potential start of sequence. From each potential start of sequence, every possible interval can be checked to see if a sequence occurs. The upper limit to the value of the interval will always be one less than the length of the string as any larger interval is equivalent to the same interval modulo the length of the string. So using a brute force approach to check every possible interval is in the worse case 50 starting positions, times 49 possible intervals, times the length of any sequences found.

Three nested loops can be used to check each potential sequence. The outermost loop sets the starting position of each potential sequence by iterating over each character in the input string. Since only non-blank characters can start a sequence, blanks can be skipped. Next an inner loop is coded that starts at 1 and checks each interval up to the length of the original string. For each interval another loop is coded to check each character at the specified interval. To handle the specification that the input string may repeat as often as necessary, keep an index to the position in the string being checked, add the interval and set the new index to the result modulo the length of the string. Note that the inner loop will always terminate because it must eventually return to the original starting character, which cannot be the next character in the sequence. In hindsight, it would have made things a little more interesting if the hidden sequence were allowed to repeat as well by letting the character 'a' follow 'z'. Keep track of the next character expected and count each time the correct next character in sequence is found. Take care to skip over space characters without incrementing the counter or next character expected. When the end of sequence is detected, check to see if the sequence found is better than any previous sequence found by applying the rules in the specification in the correct order. If the newly found sequence is better than the best found so far, store the result, which is simply the values of the iteration variables being used.

Three out of four contestants in the semi final room B coded correct solutions to this problem, the correct solutions taking 15, 21, and 25 minutes respectively.