Questions and Practice: CPS 004G

Owen Astrachan

September 18, 2007

Name: ________________________________

Login: __________

Honor code acknowledgment (signature) ________________________________

<table>
<thead>
<tr>
<th></th>
<th>value</th>
<th>grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1</td>
<td>20 pts.</td>
<td></td>
</tr>
<tr>
<td>TOTAL:</td>
<td>20 pts.</td>
<td></td>
</tr>
</tbody>
</table>
PROBLEM 1:  *(She’s Not Heavy, She’s my Sister)*

The method `dimerOccurrences` below returns the number of occurrences of a specific dimer (two-nucleotide strand) in `dna`; it works correctly. For example, if `ds` is an object/instance of the class `DNAStuff` then

```java
int x = ds.dimerOccurrences("AGTCGATCGCGAT", "CG");
int y = ds.dimerOccurrences("AGTCGATCGCGAT", "AG");
```

gives `x` the value 3 and `y` the value 1. You’ll be asked to write code based on `dimerOccurrences`. We’re assuming an ORF of 1, i.e., we’re starting with the first character.

```java
public class DNAStuff
{
    public int dimerOccurrences(String dna, String dimer){
        int count = 0;
        for(int k=0; k < dna.length()-1; k+=2){
            String sub = dna.substring(k,k+2);
            if (dimer.equals(sub)){
                count = count+1;
            }
        }
        return count;
    }
}
```

**Part A (2 points)**

Why does the loop use `k < dna.length()-1` instead of `k < dna.length()` in the loop-guard shown above?

**Part B (2 points)**

Why does the code use `.equals` instead of `==` in comparing dimers?
Part C (8 points)
Complete the method `maxDimer` that returns the dimer that occurs most frequently in a strand of DNA. The call below should return "CG" since that dimer occurs more frequently than any other.

```java
ds.maxDimer("AGTCCGATCGCGAT")
```

You may call `dimerOccurrences` in the code you write, both that method and `maxDimer` are in the same class: `DNAStuff`.

```java
public class DNAStuff {
    public int dimerOccurrences(String dna, String dimer) {
        // as shown above, not repeated
    }

    /**
     * Returns the dimer that occurs with maximal frequency in parameter dna
     * @param dna is a string with only chars 'A', 'G', 'T', 'C'
     * @return the maximally occurring dimer in dna
     */
    public String maxDimer(String dna) {
        String[] dimers = {
            "AA", "AG", "AT", "AC", "CA", "CG", "CT", "CC",
            "GA", "GG", "GT", "GC", "TA", "TG", "TT", "TC"
        };
```
Part D (8 points) The molecular weight of a strand of DNA is calculated by the following formula (this is the real formula) where \( nA \) is the number of occurrences of (A) Adenine, \( nT \) is the number of (T) Thymine occurrences, and so on.

\[
18.02 + (nA \times 313.21) + (nT \times 288.20) + (nG \times 329.21) + (nC \times 289.19)
\]

Write the method `molecularWeight` that returns the molecular weight of a strand of DNA. The parameter `strand` will contain only the characters 'A', 'T', 'C', and 'G'.

```java
public class DNAStuff {
    /** *
     * Returns the molecular weight of a strand of DNA based *
     * on the standard formula. *
     * @param strand is dna (all chars are 'A','G','T','C') whose *
     * weight is returned *
     * @return the weight of strand *
     */
    public double molecularWeight(String strand) {
        // Implementation goes here
    }
}
```