

1	2	$\Sigma (10)$

**Blatt 11**

(Abgabe am 18. January 2016)

**Theoretical Assignment - *Assembly using de Bruijn graph***

The de Bruijn graph for the reads provided on the assignment sheet would look like the graph in Figure 1, if one chooses  $k = 4$ . As one can see easily, there are not just one but four different Eulerian paths in the graph. Each path results in one of the following superstring:

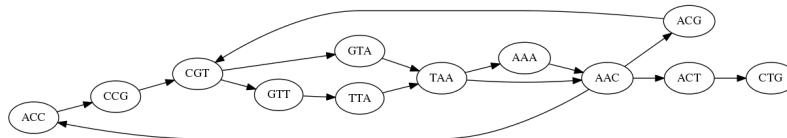
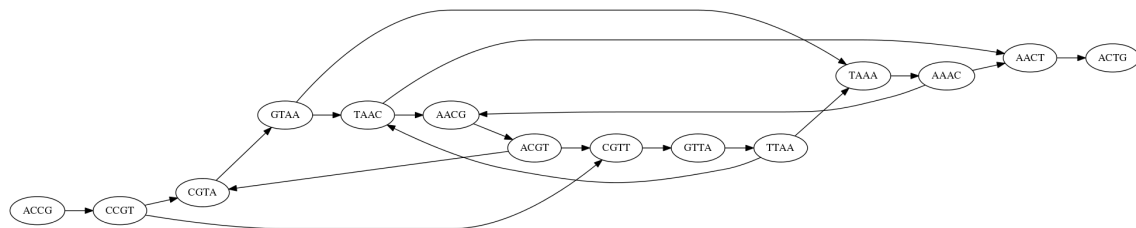
$$S_1 = AACCGTTAACGTAAACGT$$

$$S_2 = AACCGTAAACGTTAACGT$$

$$S_3 = AACCGTTAAACGTAAACGT$$

$$S_4 = AACCGTAACGTTAAACGT$$

This happens due to the fact that there are nodes with more than one outgoing and incoming edge. If one chooses  $k = 5$  the resulting de Bruijn graph would look like the graph in Figure 2. It is obvious that the number of nodes is dependent on  $k$ . With a bigger  $k$  there are more nodes in the resulting de Bruijn graph. But in this case there is one problem. Since there are more than two unbalanced nodes in the de Bruijn graph there is no Eulerian path.

Figure 1: De Bruijn graph for the given reads and  $k = 4$ .Figure 2: De Bruijn graph for the given reads and  $k = 5$ .**Practical Assignment - *Assemble the human mitochondrial genome***