**Term Project B:**

**Impact of Parameters on Pairwise Alignment Quality**

**Lucas M. Damler**

**Northern Illinois University**

This project was developed in C++ and took around 3 weeks to complete including designing, building, and debugging. The data for this program was computed using a modified shell script that computed 192 maf files per dataset at a time. To see the graphed results of these species see attached word files. Each graph’s name is in response to the parameters that stayed constant during the computation. The x axis is the value that changed, the y axis is the accuracy in percent form.

When it comes to sensitivity and specificity as a measure of computed alignment accuracy changing the C value from 0 to 5 obviously makes the computation more accurate because C = 5 allows more leeway of error when computing if a base is correctly aligned or not. C a definitive impact on all species in changing the accuracy of the computation.

For primate species of baboon and chimp parameters of a K3000, O300, T1, W8 had the highest sensitivity and specificity. Though when this is compared to the lowest computation values there isn’t much difference. Because of the similarity of human and primate genomes changing the parameters of alignment has little impact of the computed result by more than a percent.

The cat and dog genomes had the highest sensitivity rating under parameters of K2000, O500, T1, W8. Their sensitivity rating is around 68-69%. Increasing K to 5000 and decreasing T to 0 with a W of 20 caused the lowest sensitivity rating for cat to drop to 51% and dog to around 35%. Ironically those same values of K5000, O500, T0, W20 caused both animals to have their highest specificity rating with both species around 90-92%. Cow, mouse, pig, and rat species had their highest sensitivity with parameters if K2000, O300, T1, W8. Cow rated 66%, mouse and rat were around 35-38%, and pig had a rating of 71%. Increasing K to 5000, and O to 500 with a T of 0 and W of 20 universally caused a decrease in sensitivity for all species. The same parameters however gave the highest specificity rating to both cow and pig and lowest for mouse and rat. Mouse and rat genomes had the highest specificity under parameters of K2000, O500, T0, W12.

Based on the gathered data and computed sensitivity and specificity it is interesting to see how the change in parameters effects the computed accuracy of each species. It is apparent that some parameters effect similar species the same and dissimilar species differently. This meaning chimp and baboon are effect similarly by parameter changes as dog and cat are. The same can be found with rat and mouse, and cow and pig.