Marcus Deans (md374)

BaseMarkov Benchmark.java Results:

|  |  |  |
| --- | --- | --- |
| time | source | #chars |
| 0.170 | 487614 | 2000 |
| 0.269 | 487614 | 4000 |
| 0.631 | 487614 | 8000 |
| 1.260 | 487614 | 16000 |
| 2.486 | 487614 | 32000 |
|  |  |  |
| 0.324 | 487614 | 4096 |
| 0.645 | 975228 | 4096 |
| 0.967 | 1462842 | 4096 |
| 1.284 | 1950456 | 4096 |
| 1.608 | 2438070 | 4096 |
| 1.917 | 2925684 | 4096 |
| 2.274 | 3413298 | 4096 |
| 2.581 | 3900912 | 4096 |
| 2.905 | 4388526 | 4096 |
| 3.230 | 4876140 | 4096 |

These timings support the O(NT) analysis very well. As can be observed with tests in which the training text N remained constant at size 487614, while the number of characters T was doubled for each test, the time approximately doubled in each test, with this relation being more pronounced and precise as T increased in magnitude. The relationship similarly held true for the second set of tests in which the number of characters T remained constant at 4096, while the size of the training text N constantly increased in size such that where *x* represents the test number. The finite differences between the times indicated a product relationship as opposed to addition, with the proportionate time increasing beyond actual difference in magnitude of and . Therefore, the data showed that the O(NT) analysis was accurate for BaseMarkov.

EfficientMarkov Benchmark.java Results:

|  |  |  |
| --- | --- | --- |
| time | source | #chars |
| 0.080 | 487614 | 2000 |
| 0.106 | 487614 | 4000 |
| 0.092 | 487614 | 8000 |
| 0.086 | 487614 | 16000 |
| 0.086 | 487614 | 32000 |
|  |  |  |
| 0.075 | 487614 | 4096 |
| 0.187 | 975228 | 4096 |
| 0.240 | 1462842 | 4096 |
| 0.323 | 1950456 | 4096 |
| 0.405 | 2438070 | 4096 |
| 0.544 | 2925684 | 4096 |
| 0.864 | 3413298 | 4096 |
| 1.012 | 3900912 | 4096 |
| 1.047 | 4388526 | 4096 |
| 1.435 | 4876140 | 4096 |

These timings similarly support the O(N+T) Analysis for EfficientMarkov. As with BaseMarkov, the same N and T values were held constant and increased at certain rates in the manner as outlined above. The miniscule differences in time in the first set of tests (Constant N) despite T doubling indicates that the difference in T was of insufficient magnitude to significantly affect the time to compute, thus supporting the addition argument. With the large magnitude of N added to an increasing but still small magnitude value of T, N predominated which lead the timing remaining relatively constant. This supports the O(N+T) analysis which is the situation in which this predomination could actually occur. The findings for the second set of test (Constant T) had similar results, with much smaller differences in the time to compute compared to that of BaseMarkov. The results indicated that addition was being performed as opposed to multiplication, thus supporting the analysis of O(N+T) for Efficient Markov.