

Summary of the Test with 4000 Edge Limit:

1. Cached Benchmarking: Using "lkb_4k_cache_result.txt" and "ale_4k_result.txt" that was produced by LKB and ALE:

Initially there was 2354 test stat items in each set:

-----LKB-----

Total Number of items = 2354

Total Number of Parses = **33564**

Total Edge Limit Exceeded = **765**

-----ALE-----

Total Number of items = 2354

Total Number of Parses = **32228**

Total Insufficient memory / edge Limit exceeded incident = **774**

As shown above there was a mismatch in the produced number of parses (in red). The mismatch was mainly due to the memory problem/edge-limit overflow in ALE and LKB, after excluding such reported statistics by ALE or LKB from the set, there was also a minor mismatch caused by:

@1468 ALE: well actually tuesday through thursday i am out at the ophthalmologist meeting in new york

@1468 LKB: well actually Tuesday through Thursday I am out at the ophthalmologist meeting in New York

Which seems to be MWE/morphology problem, we were left out with 1576 comparable items:

-----LKB-----

Total Number of items = 1576

Total Number of Parses = **31731**

Total Parse Time = **1585394** (msec)

Total Edge Limit Exceeded = 0

-----ALE-----

Total Number of items = 1576

Total Number of Parses = **31731**

Total Parse Time = **991910** (msec)

Total Insufficient memory / Edge Limit exceeded incident = 0

2. Non-Cached Benchmarking: Using "lkb_4k_no_cache_result.txt" and "ale_4k_result.txt" test results that was produced by LKB and ALE:

Comparable items were also excluded as explained above (excluded exactly the same items)

-----LKB-----

Total Number of items = 1576

Total Number of Parses = **31731**

Total Parse Time = **1685613** (msec)

Total Edge Limit Exceeded = 0

-----ALE-----

Total Number of items = 1576

Total Number of Parses = **31731**

Total Parse Time = **991910** (msec)

Total Insufficient memory / Edge Limit exceeded incident = 0