

Summary of the Test with 16000 Edge Limit:

1. Cached Benchmarking: Using "lkb_16k_cache_result.txt" and "ale_16k_result.txt" that was produced by LKB and ALE:

Initially there was 1582 test stat-items in LKB set (which was reported by up to the point LKB crashed) and 2335 items in ALE set, but since LKB crashed after parsing 1582 items, the first 1582 items of the ALE set was chosen for benchmarking:

-----LKB-----

Total Number of items = 1582

Total Number of Parses = **88135**

Total Edge Limit Exceeded = **523**

-----ALE-----

Total Number of items = 1582

Total Number of Parses = **41114**

Total Insufficient memory / edge Limit exceeded incident = **147**

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 (in blue), 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

@1515 ALE: okay because i do have a doctor_s appointment on monday but i can easily cancel that

@1515 LKB: okay because I do have a doctor's appointment on Monday but I can easily cancel that

Which seems to be MWE/morphology problem, we were left out with 1727 comparable items:

-----LKB-----

Total Number of items = 1431
Total Number of Parses = **41060**
Total Parse Time = **1173086** (msec)
Total Edge Limit Exceeded = 0

-----ALE-----

Total Number of items = 1431
Total Number of Parses = **41060**
Total Parse Time = **1041010** (msec)
Total Insufficient memory / Edge Limit exceeded incident = 0

2. Non-Cached Benchmarking: Using "lkb_16k_no_cache_result.txt" and "ale_16k_result.txt" test results that was produced by LKB and ALE:

Initially there was 2335 test stat-items in LKB and ALE:

-----LKB-----

Total Number of items = 2335
Total Number of Parses = **162945**
Total Edge Limit Exceeded = **312**

-----ALE-----

Total Number of items = 2335
Total Number of Parses = **48710**
Total Insufficient memory / edge Limit exceeded incident = **584**

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 (in blue), 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

@1515 ALE: okay because i do have a doctor_s appointment on monday but i can easily cancel that

@1515 LKB: okay because I do have a doctor's appointment on Monday but I can easily cancel that

Which seems to be MWE/morphology problem, we were left out with 1727 comparable items:

-----LKB-----

Total Number of items = 1749

Total Number of Parses = **48057**

Total Parse Time = **2480700** (msec)

Total Edge Limit Exceeded = 0

-----ALE-----

Total Number of items = 1749

Total Number of Parses = **48057**

Total Parse Time = **1651470** (msec)

Total Insufficient memory / Edge Limit exceeded incident = 0