

From the testing done in part 2 the following canonical collection have these respective file sizes:

Canonical Collection 1: 4 KB

Canonical Collection 2: 3 KB

Canonical Collection 3: 2 KB

Thus, within encoding the files the ones produced by the third canonical collection results in smaller sized files. The efficiency of the code is dependant on the size of files used to encode, which means that canonical collection 3 would be the most efficient collection. The Huffman codes are dependant on the relative frequency, thus since the third collection is the largest it would have the highest frequencies.

	Canonical Collection 1	Canonical Collection 2	Canonical Collection 3
<b>Earth.txt</b>	5874 KB	2202 KB	1975 KB
<b>Mystery.txt</b>	6220 KB	2369 KB	2039 KB
<b>Myths.txt</b>	11252 KB	4312 KB	3561 KB
<b>Simack.txt</b>	4338 KB	1642 KB	1413 KB
<b>Wodehouse.txt</b>	5711 KB	1946 KB	1807 KB