Entropy of Gigaword corpus

