models with the uncertainty as Shannon estimates it. This is shown in Table III.

TABLE III

States	H(X)	H(Y|X)	H(X|Y)	H(Y)
2	.8254	3.2890	.1657	3.9486
3	.8875	3.1266	.6199	3.3942
4	1.0839	2.7060	.6028	3.1871
5	1.2629	2.5332	.6874	3.1087
6	1.2066	2.4868	.8757	2.8176
7	1.1851	2.4433	1.0312	2.5973
8	1.5306	2.2143	.9909	2.7540
9	1.5468	2.1497	.9757	2.7198
10	1.6759	2.0766	1.1791	2.5834
11	1.8025	1.8415	1.0082	2.6358
12	1.8416	1.8550	1.2138	2.4828
		Shannon's Estimate		1.0

We may note how the entropy for our models quickly falls below that for monographic and digraphic English (Table V). This is probably due to the extra structure we have imposed. $H(X)$, of course, increases as does $H(X|Y)$ and we will discuss this later.

In addition to this comparison, we may examine the uncertainty associated with each of the states of our Hidden Markov Chain. This uncertainty is of two sorts. There is uncertainty about the next state, and there is uncertainty about the output given a state. If we consider the ith state, then we let $H_i(X)$ denote the entropy of the former

and $H_i(Y|X)$ denote the entropy of the latter. These are displayed in the following tables.

These figures may be compared with entropy for forward and backward digraphic English. Specifically, we suppose English is sampled from a Markov chain of order 1 with 27 states. The states are the letters themselves and the transition probability $a_{\alpha\beta}$ is the frequency of the digraph α, β in English divided by the frequency of the letter α in the forward case and the letter β in the backward case. Thus, for example, the digraph ab is the b state followed by the a state in the backward case and the reverse in the forward case. The entropy in each of these cases may be calculated. Table V is the result.

The entropy of each letter is indicative of the uncertainty of the succeeding letter in the forward case and the preceding letter in the backward case. From this point of view (and as any school child knows) there is no information (resolution of uncertainty) in letters following q. On the other hand, the first letter of a word resolves the maximum amount of uncertainty since the entropy is maximum for letters following the space state. Thus the first letter of a word, in this sense carries more information than followers of each of the other letters.