## Representing English letters as sequences of bits ECL = expected code length

| $\mathbf{a}$ | $8.167 \%$ |
| ---: | ---: |
| $\mathbf{b}$ | $1.492 \%$ |
| $\mathbf{c}$ | $2.782 \%$ |
| $\mathbf{d}$ | $4.253 \%$ |
| $\mathbf{e}$ | $12.702 \%$ |
| $\mathbf{f}$ | $2.228 \%$ |
| $\mathbf{g}$ | $2.015 \%$ |
| $\mathbf{h}$ | $6.094 \%$ |
| $\mathbf{i}$ | $6.966 \%$ |
| $\mathbf{j}$ | $0.153 \%$ |
| $\mathbf{k}$ | $0.772 \%$ |
| $\mathbf{l}$ | $4.025 \%$ |
| $\mathbf{m}$ | $2.406 \%$ |
| $\mathbf{n}$ | $6.749 \%$ |
| $\mathbf{o}$ | $7.507 \%$ |
| $\mathbf{p}$ | $1.929 \%$ |
| $\mathbf{q}$ | $0.095 \%$ |
| $\mathbf{r}$ | $5.987 \%$ |
| $\mathbf{s}$ | $6.327 \%$ |
| $\mathbf{t}$ | $9.056 \%$ |
| $\mathbf{u}$ | $2.758 \%$ |
| $\mathbf{v}$ | $0.978 \%$ |
| $\mathbf{w}$ | $2.360 \%$ |
| $\mathbf{x}$ | $0.150 \%$ |
| $\mathbf{y}$ | $1.974 \%$ |
| z | $0.074 \%$ |
|  |  |
|  |  |

$a=1$<br>$b=01$<br>$\mathrm{c}=001$<br>$d=0001$<br>$e=00001$<br>$\mathrm{f}=000001$<br>$\mathrm{g}=0000001$<br>$\mathrm{h}=00000001$<br>$\mathrm{i}=000000001$<br>$j=0000000001$

$\mathrm{ECL} \approx 11.65$ bits per character

```
e t a o i n s h r d l c umw f g y p b v k j x q z
e=1
t=01
a=001
o=0001
i=00001
n=000001
s=0000001
h=00000001
r=000000001
d=0000000001
```



Huffman tree for English based on single letter statistics


## English single letter statistics tree from heights



## English single letter statistics tree from heights



