

In[22]:= (\* Adam Beck \*)

In[23]:= (\* Select long texts in three different languages. Calculate the entropy  
for each and compare them, draw your conclusions \*)

(\* I will first select an English text \*)

(\* Sample text from The Adventures of Sherlock Holmes \*)

In[27]:= text = "To Sherlock Holmes she is always the woman.  
I have seldom heard him mention her under any other name.  
In his eyes she eclipses and predominates the whole of her sex.  
It was not that he felt any emotion akin to love for Irene Adler.  
All emotions, and that one particularly, were abhorrent to his cold,  
precise but admirably balanced mind. He was, I take it,  
the most perfect reasoning and observing machine that the world has seen,  
but as a lover he would have placed himself in a false position.  
He never spoke of the softer passions, save with a gibe and a sneer.  
They were admirable things for the observer--excellent for drawing the veil from  
men's motives and actions.  
But for the trained reasoner to admit such intrusions into his own delicate and  
finely adjusted temperament was to introduce a distracting factor which might throw  
a doubt upon all his mental results.  
Grit in a sensitive instrument, or a crack in one of his own high-power lenses,  
would not be more disturbing than a strong emotion in a nature such as his.  
And yet there was but one woman to him,  
and that woman was the late Irene Adler, of dubious and questionable memory.

I had seen little of Holmes lately.

My marriage had drifted us away from each other. My own complete happiness,  
and the home-centred interests which rise up around the man who first finds himself  
master of his own establishment,  
were sufficient to absorb all my attention, while Holmes,  
who loathed every form of society with his whole Bohemian soul,  
remained in our lodgings in Baker Street, buried among his old books,  
and alternating from week to week between cocaine and ambition,  
the drowsiness of the drug, and the fierce energy of his own keen nature.  
He was still, as ever, deeply attracted by the study of crime,  
and occupied his immense faculties and extraordinary powers of observation in  
following out those clues,  
and clearing up those mysteries which had been abandoned as hopeless by the  
official police.

From time to time I heard some vague account of his doings:  
of his summons to Odessa in the case of the Trepoff murder,  
of his clearing up of the singular tragedy of the Atkinson brothers at Trincomalee,

and finally of the mission which he had accomplished so delicately and successfully  
for the reigning family of Holland. Beyond these signs of his activity,  
however, which I merely shared with all the readers of the daily press,

I knew little of my former friend and companion.

One night--it was on the twentieth of March,  
 1888--I was returning from a journey to a patient (for I had now returned to civil  
 practice), when my way led me through Baker Street.  
 As I passed the well-remembered door,  
 which must always be associated in my mind with my wooing,  
 and with the dark incidents of the Study in Scarlet,  
 I was seized with a keen desire to see Holmes again,  
 and to know how he was employing his extraordinary powers.  
 His rooms were brilliantly lit, and, even as I looked up, I saw his tall,  
 spare figure pass twice in a dark silhouette against the blind.  
 He was pacing the room swiftly, eagerly,  
 with his head sunk upon his chest and his hands clasped behind him.  
 To me, who knew his every mood and habit,  
 his attitude and manner told their own story. He was at work again.  
 He had risen out of his drug-created dreams and was hot upon the scent of some new  
 problem.  
 I rang the bell and was shown up to the chamber which had formerly been in part my  
 own.";

```
In[28]:= lowerEnglishText = ToLowerCase[text];
```

```
In[29]:= charcode = ToCharacterCode[lowerEnglishText];
```

```
In[30]:= charentp = Table[{Length[Position[charcode, i]], FromCharacterCode[i]}, {i, 96, 122}]
```

```
Out[30]= {{0, `}, {218, a}, {43, b}, {69, c}, {118, d}, {335, e}, {66, f}, {41, g}, {170, h},  

  {211, i}, {2, j}, {21, k}, {110, l}, {93, m}, {196, n}, {208, o}, {42, p}, {1, q},  

  {161, r}, {191, s}, {214, t}, {61, u}, {20, v}, {81, w}, {4, x}, {50, y}, {1, z}}
```

```
In[32]:= charentp[[1]] = {Length[Position[charcode, 32]], "space"}
```

```
Out[32]= {611, space}
```

```
In[33]:= TableForm[charentp]
```

```
Out[33]//TableForm=
```

611	space
218	a
43	b
69	c
118	d
335	e
66	f
41	g
170	h
211	i
2	j
21	k
110	l
93	m
196	n
208	o
42	p
1	q
161	r
191	s
214	t
61	u
20	v
81	w
4	x
50	y
1	z

```
In[34]:= sortcharentp = Reverse[Sort[charentp]]
```

```
Out[34]= {{611, space}, {335, e}, {218, a}, {214, t}, {211, i}, {208, o}, {196, n}, {191, s},
{170, h}, {161, r}, {118, d}, {110, l}, {93, m}, {81, w}, {69, c}, {66, f}, {61, u},
{50, y}, {43, b}, {42, p}, {41, g}, {21, k}, {20, v}, {4, x}, {2, j}, {1, z}, {1, q}}
```

```
In[35]:= sumchars = Sum[sortcharentp[[i, 1]], {i, Length[sortcharentp]}]
```

```
Out[35]= 3338
```

```
In[36]:= charfreq = N[Table[
{sortcharentp[[i, 1]]/sumchars, sortcharentp[[i, 2]]}, {i, 1, Length[sortcharentp]}]]
```

```
Out[36]= {{0.183044, space}, {0.100359, e}, {0.0653086, a}, {0.0641102, t},
{0.0632115, i}, {0.0623128, o}, {0.0587178, n}, {0.0572199, s}, {0.0509287, h},
{0.0482325, r}, {0.0353505, d}, {0.0329539, l}, {0.027861, m}, {0.024266, w},
{0.0206711, c}, {0.0197723, f}, {0.0182744, u}, {0.014979, y}, {0.012882, b},
{0.0125824, p}, {0.0122828, g}, {0.00629119, k}, {0.00599161, v},
{0.00119832, x}, {0.000599161, j}, {0.000299581, z}, {0.000299581, q}}
```

```
In[37]:= TableForm[charfreq]
```

```
Out[37]//TableForm=
```

0.183044	space
0.100359	e
0.0653086	a
0.0641102	t
0.0632115	i
0.0623128	o
0.0587178	n
0.0572199	s
0.0509287	h
0.0482325	r
0.0353505	d
0.0329539	l
0.027861	m
0.024266	w
0.0206711	c
0.0197723	f
0.0182744	u
0.014979	y
0.012882	b
0.0125824	p
0.0122828	g
0.00629119	k
0.00599161	v
0.00119832	x
0.000599161	j
0.000299581	z
0.000299581	q

```
In[38]:= entplang = 0;
```

```
For[i = 1, i <= 27, i++,
```

```
    If[charfreq[[i, 1]] == 0, entplang,
```

```
        entplang = entplang - charfreq[[i, 1]] * Log[charfreq[[i, 1]]]
```

```
    ]
```

```
]
```

```
In[41]:=
```

```
(* This represents the entropy as bits per symbol *)
```

```
entplang
```

```
Out[41]= 2.83089
```