Information - theoretic average - case lower bound on sorting n elements by decision tree

Main results

The approximate (under - estimeted) value of information - theoretic lower bound for average number of comparisons of keys while sorting an n - element array by means of decision tree

$$LB_{avg}^{sort}(n) \approx \left(n + \frac{1}{2}\right) Log2[n] - 1.443\, n + 1.325$$

The exact value is

$$LB_{avg}^{sort}(n) = \lg n! + \epsilon(n!)$$

and may be closely approximated (under - estimated) as

$$LB_{avg}^{sort}(n) \approx \left(n + \frac{1}{2}\right) Log2[n] - 1.443\, n + 1.325 + \epsilon(n!)$$

or as

$$LB_{avg}^{sort}(n) \approx \lg n! \approx \left(n + \frac{1}{2}\right) Log2[n] - 1.443\, n + 1.325$$

or (roughly) as

$$LB_{avg}^{sort}(n) \approx n\, Log2[n] - 1.5\, n$$

where

$\beta[x\_] := 1 + x - 2^x$

$\theta[x\_] := \lceil x \rceil - x$

$\epsilon[x\_] := \beta[\theta[Log2[x]]]$

We have (this is slide 12 from the file http : // csc.csudh.edu / suchenek / CSC401 / Slides / Excerpts_from _Knuth - Suchenek_formulas.pdf with m substituted for x)

$\epsilon\,(m) = \lceil \lg m \rceil - \lg m + 1 - 2^{\lceil \lg m \rceil - \lg m}$

or

$\lg m + \epsilon\,(m) - 1 = \lceil \lg m \rceil - 2^{\lceil \lg m \rceil - \lg m}$

or

$\lg m + \epsilon\,(m) - 1 = \lceil \lg m \rceil - \dfrac{2^{\lceil \lg m \rceil}}{\lg m}$

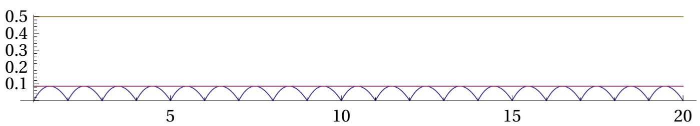
or

$\lg m + \epsilon\,(m) - 1 = \lceil \lg m \rceil - \dfrac{2^{\lceil \lg m \rceil}}{m}$

or

$m\,(\lg m + \epsilon\,(m) - 1) = m \lceil \lg m \rceil - 2^{\lceil \lg m \rceil}$

```
Plot[{1 + ⌈x⌉ - x - 2^(⌈x⌉-x), .0861, .5}, {x, 1, 20}, AspectRatio → .13]
```




```
FindMaximum[1 + ⌈x⌉ - x - 2^(⌈x⌉-x), {x, 0.47}]
```

```
{0.0860713, {x → 0.471234}}
```

```
FindMinimum[1 + ⌈x⌉ - x - 2^(⌈x⌉-x), {x, 0}]
```

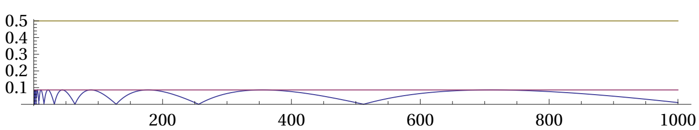
```
{0., {x → 0.}}
```

```
FindMinimum[1 + ⌈x⌉ - x - 2^(⌈x⌉-x), {x, 1}]
```

```
{0., {x → 1.}}
```

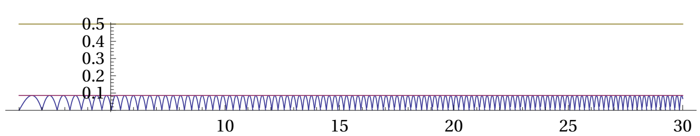
$0 \le \epsilon\,(x) < 0.08607133205593432$

```
Plot[{ϵ[n], .0861, .5}, {n, 1, 1000}, AspectRatio → .13]
```



Same, for n ! argument

```
Plot[{ϵ[n!], .0861, .5}, {n, 1, 30}, AspectRatio → .13]
```




```
Table[N[ϵ[n!]], {n, 1, 100}]
```

```
{0.`, 0.`, 0.08170416594551044`, 0.08170416594551089`, 0.026442737724815757`,
 0.08592468144810184`, 0.07539515621589565`, 0.07539515621589565`, 0.08606980203989778`,
 0.053101142778636756`, 0.0682887320938832`, 0.04373240955668578`, 0.08464368609991624`,
 0.08022285523037453`, 0.06823044158882396`, 0.06823044158882396`, 0.07968695848721552`,
 0.08561859375377168`, 0.05982577931086297`, 0.027066647621339257`, 0.08607051805032029`,
 0.020519300757925407`, 0.08595170851489797`, 0.013871416758490795`, 0.07146484607454795`,
 0.08325092522953526`, 0.04414198829677218`, 0.016365858749225026`, 0.05085675430991898`,
 0.06750172541308075`, 0.07383990107092586`, 0.07383990107092586`, 0.06771537100243563`,
 0.05228940370966484`, 0.021799813906511645`, 0.03225526876909157`, 0.0759662409610371`,
 0.08386795310485695`, 0.043362086941669986`, 0.04971828179407112`, 0.08536571732832954`,
 0.025915545472201984`, 0.07944894746773912`, 0.052392886899923496`,
 0.07371694000315188`, 0.04923195660990132`, 0.08203235549598276`, 0.014855893283140631`,
 0.08313418885671808`, 0.06055449663017498`, 0.03312604398837493`, 0.08199927574341359`,
 0.07658875314072588`, 0.022722053089637484`, 0.041539392820624244`,
 0.07514306887372868`, 0.08600185285280304`, 0.08023646976266718`, 0.06400301320314838`,
 0.0431152701191877`, 0.0226752591330523`, 0.0067480767129382`, 0.0014844239582885166`,
 0.0014844239582885166`, 0.006617949975748161`, 0.022098688049766224`,
 0.04180033809228689`, 0.06199659080306219`, 0.07829427195082417`, 0.085958972044579`,
 0.08023198689056699`, 0.056613153677403716`, 0.011090568941256151`,
 0.0528269591063690`, 0.0843353856084832`, 0.07556005620131145`, 0.024114010223229343`,
 0.059136265627557805`, 0.08568083628790646`, 0.045630346808593 46`, 0.05154590174242912`,
 0.08510862332730085`, 0.02840621143343469`, 0.07441144104836894`, 0.06870732586406803`,
 0.04160637253346522`, 0.08275203069740655`, 0.011033735048215476`, 0.08567035083410701`,
 0.0008464801915692988`, 0.08567854902622685`, 0.010654746678312677`,
 0.0829136098330423`, 0.04026721626411245`, 0.06998198756929241`, 0.07241191376522238`,
 0.033985665876514304`, 0.08592571011848804`, 0.03993757003263454`, 0.05809449894354657`}
```

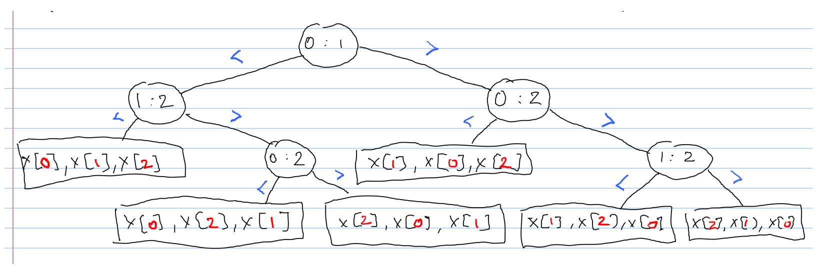
ϵ[n!] is the difference between the
LB^{sort}_{avg} (n) and lg n!

Here is the derivation.

By the definition of the epl in a decition tree for a sorting algorithm,

$$LB^{sort}_{avg}(n) = \frac{epl_{min}(n!)}{n!}$$

For example consider this decision tree for sorting 3 elemets. It has minimal epl.



So,

$$LB^{sort}_{avg}(3) = \frac{epl_{min}(3!)}{3!} = \frac{epl_{min}(6)}{6} = \frac{1}{6}(2+3+3+2+3+3) = \frac{16}{6} \approx 2.666666667.$$

We have already shown (see the last slide in the file http://csc.csudh.edu/suchenek/CSC401/Slides/Excerpts_from_Knuth-Suchenek_formulas.pdf) that

$$epl_{min}(m) = m\,(Log2[m] + \epsilon[m]),$$

where $m = n + 1$ is the number of external nodes in a 2-tree of n nodes.

In particular,

$$\frac{epl_{min}(m)}{m} = Log2[m] + \epsilon[m].$$

or

$$\frac{epl_{min}(m)}{m} = \lg m + \epsilon(m)$$

Now, putting $m = n!$ we conclude

$$\frac{epl_{min}(n!)}{n!} = \lg n! + \epsilon(n!)$$

Hence,

$LB^{sort}_{avg}(n) = \lg n! + \epsilon(n!)$

How simple!

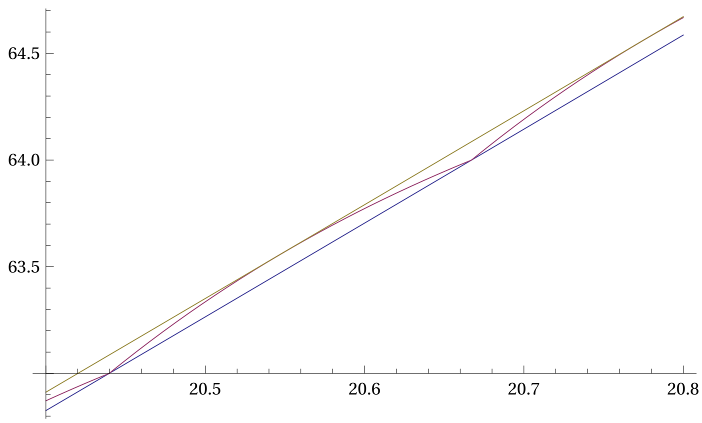
So,

$\lg n! \leq LB^{sort}_{avg}(n) < \lg n! + .0861$

and for n > 2 ($n! \neq 2^k$ for n > 2, therefore lg n! is not an integer,
therefore $\lg n! < \lceil \lg n! \rceil$, therefore $0 < \lceil \lg n! \rceil - \lg n! < 1$,
therefore $0 < \epsilon(\lg n!)$, therefore $\lg n! < \lg n! + \epsilon(\lg n!)$ for n > 2)

$\lg n! < LB^{sort}_{avg}(n) < \lg n! + .0861$

Plot[{Log2[n!], ⌈Log2[n!]⌉ + 1 - 2^(⌈Log2[n!]⌉-Log2[n!]), Log2[n!] + .0861}, {n, 20.4, 20.8}]



Now, let us derive approximate bounds for $LB^{sort}_{avg}(n)$
that do not use floors, ceilings, and factorials, with high precision.

$$\text{Log2}[n!] \approx n\,\text{Log2}[n] - \frac{1}{\text{Log}[2]} n + \frac{\text{Log2}[n]}{2} + \frac{\text{Log}[2\pi]}{\text{Log}[4]} + 1/(\text{Log}[4096]\, n) - 1/(360\,\text{Log}[2]\, n^3)$$

(see file
Mathematica / Log_of _factorial.nb)

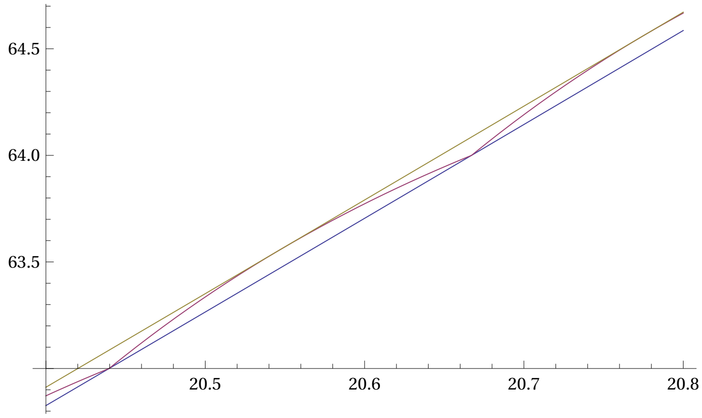
Hence,

$$n\,\text{Log2}[n] - \frac{1}{\text{Log}[2]} n + \frac{\text{Log2}[n]}{2} + \frac{\text{Log}[2\pi]}{\text{Log}[4]} + \frac{1}{\text{Log}[4096]\, n} - 1/(360\,\text{Log}[2]\, n^3)$$

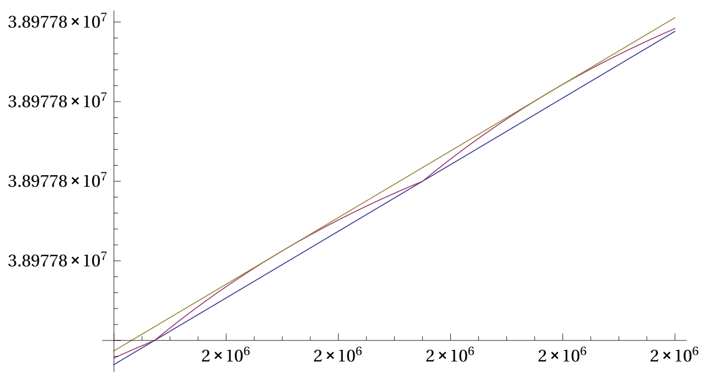
$< LB^{sort}_{avg}(n) <$

$$n\,\text{Log2}[n] - \frac{1}{\text{Log}[2]} n + \frac{\text{Log2}[n]}{2} + \frac{\text{Log}[2\pi]}{\text{Log}[4]} + .0861 + 1/(\text{Log}[4096]\, n) - 1/(360\,\text{Log}[2]\, n^3)$$

$$\text{Plot}\left[\left\{n\,\text{Log2}[n]-\frac{1}{\text{Log}[2]}n+\frac{\text{Log2}[n]}{2}+\frac{\text{Log}[2\,\pi]}{\text{Log}[4]}+1/(\text{Log}[4096]\,n)-1/(360\,\text{Log}[2]\,n^3),\ \lceil\text{Log2}[n!]\rceil+1-2^{\lceil\text{Log2}[n!]\rceil-\text{Log2}[n!]},\ n\,\text{Log2}[n]-\frac{1}{\text{Log}[2]}n+\frac{\text{Log2}[n]}{2}+\frac{\text{Log}[2\,\pi]}{\text{Log}[4]}+.0861+1/(\text{Log}[4096]\,n)-1/(360\,\text{Log}[2]\,n^3)\right\},\ \{n,\ 20.4,\ 20.8\}\right]$$



$$\text{Plot}\left[\left\{n\,\text{Log2}[n]-\frac{1}{\text{Log}[2]}n+\frac{\text{Log2}[n]}{2}+\frac{\text{Log}[2\,\pi]}{\text{Log}[4]}+1/(\text{Log}[4096]\,n)-1/(360\,\text{Log}[2]\,n^3),\ \lceil\text{Log2}[n!]\rceil+1-2^{\lceil\text{Log2}[n!]\rceil-\text{Log2}[n!]},\ n\,\text{Log2}[n]-\frac{1}{\text{Log}[2]}n+\frac{\text{Log2}[n]}{2}+\frac{\text{Log}[2\,\pi]}{\text{Log}[4]}+.0861+1/(\text{Log}[4096]\,n)-1/(360\,\text{Log}[2]\,n^3)\right\},\ \{n,\ 2\,000\,000,\ 2\,000\,000.1\}\right]$$



Good enough and exact (as opposed to approximate)
bonds on LB^{sort}_{avg} (n) using Stirling formula (see file Summationa.nb) :

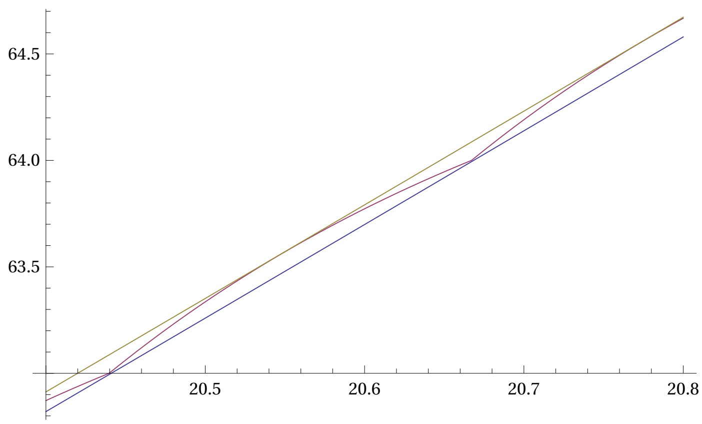
$$\left(n+\frac{1}{2}\right)\text{Log2}[n]-1.4426950408889634`\,n+1.3257480647361592` <$$

$$LB^{sort}_{avg}(n) < \left(n+\frac{1}{2}\right)\text{Log2}[n]-1.4426950408889634`\,n+1.3257480647361592`+.08607133205593432+0.1311540946262694`/n=$$

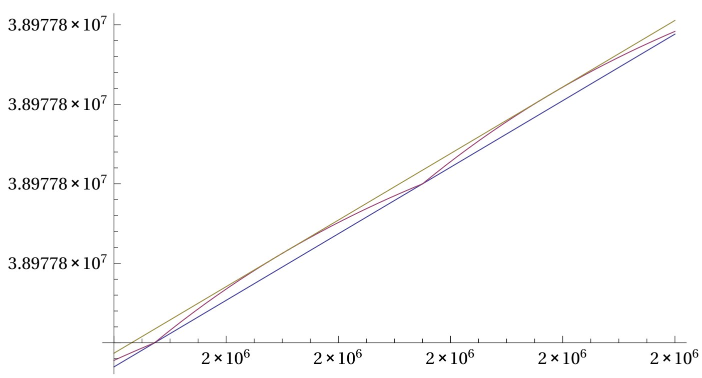
$$\left(n+\frac{1}{2}\right)\text{Log2}[n]-1.4426950408889634`\,n+1.4118193967920936`+0.1311540946262694`/n$$

These are exact (as opposed to approximate) inequalities.

```
Plot[{(n + 1/2) Log2[n] - 1.4426950408889634` n + 1.3257480647361592`,
  ⌈Log2[n!]⌉ + 1 - 2^(⌈Log2[n!]⌉-Log2[n!]), (n + 1/2) Log2[n] - 1.4426950408889634` n +
   1.4118193967920936` + 0.1311540946262694` / n}, {n, 20.4, 20.8}]
```




```
Plot[{(n + 1/2) Log2[n] - 1.4426950408889634` n + 1.3257480647361592`,
  ⌈Log2[n!]⌉ + 1 - 2^(⌈Log2[n!]⌉-Log2[n!]), (n + 1/2) Log2[n] - 1.4426950408889634` n +
   1.4118193967920936` + 0.1311540946262694` / n}, {n, 2 000 000, 2 000 000.1}]
```



Here are the same bounds on $LB^{sort}_{avg}(n)$ with lesser precision (but still exact inequalities)

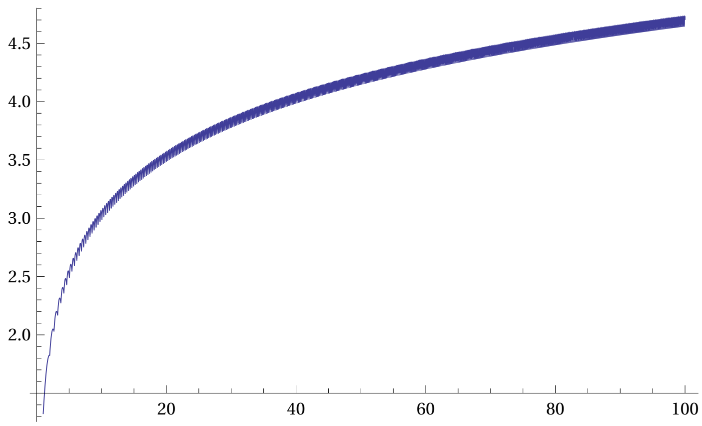
$$\left(n + \frac{1}{2}\right) \mathrm{Log2}[n] - 1.443\, n + 1.325 < LB^{sort}_{avg}(n) < \left(n + \frac{1}{2}\right) \mathrm{Log2}[n] - 1.442\, n + 1.412 + \frac{0.132}{n}$$

The textbook approximation (an exact inequality) was

$LB^{sort}_{avg}(n) \approx n\, \mathrm{Log2}[n] - 1.443\, n$

Here is a graph of the difference $\frac{1}{2} \mathrm{Log2}[n] + 1.325 + \epsilon[n!]$ between the two :

```
Plot[(n + 1/2) Log2[n] - 1.443 n + 1.325 + ϵ[n!] - (n Log2[n] - 1.443 n),
 {n, 1, 100}, {PlotPoints → 200}]
```




```
Plot[(n + 1/2) Log2[n] - 1.443 n + 1.325 + ϵ[n!] - (n Log2[n] - 1.443 n),
 {n, 50, 70}, {PlotPoints → 200}]
```