Approximating the Inverse of [n log n]

Andrei Lapets
al@eecs.harvard.edu

May 21, 2005

1 Overview

It is surprisingly difficult to compute the inverse of the product of a number multiplied by its logarithm. The need to provide a means for calculating this inverse may arise in many instances, especially when dealing with experiments in information theory and calculations such as those required for Shannon entropy. This problem can certainly be solved using some heuristic method, and once very small scales are reached within the context of a large-scale problem this would be desirable. However, any such solution, while effective in some specific case, is generally plagued by drawbacks and lacks generality and simplicity. Examples might include limiting the domain and using a table, a search algorithm, or some heuristic or analytic approximation. Less appetizing methods might even involve solving polynomials of some high degree. An example of one such scheme, utilizing generating functions, will be discussed. More importantly, it is worth investigating a means for approximating this function which are general, perhaps scalable, and whose precision can be adjusted. We will investigate an analytic method, utilizing recursion, which has these characteristics.

2 Approximation with Logarithms

The approach to be outlined below stems from an important observation about any value $n \log n$; namely, that this value is a product of n and the *size of n's representation*, and that the size of the representation of n will be only *slightly* less (relatively speaking, by a difference of $\log \log n$) than the size of the representation of $n \log n$.

$$\log(n \log n) - \log \log n = \log\left(\frac{n \log n}{\log n}\right) = \log n$$

In other words, when applying the function $[n \log n]$, we only multiply n by the size of its representation. What we want to note at this point is that the size of the *representation* of the result of $[n \log n]$ will *not* differ a great deal from the size of the representation of the original argument n.

This small magnitude of difference between the sizes of representation is important. Ideally, given some input $y \equiv n \log n$, we would be happy to utilize $\log n$ to obtain n and have our result.

$$\frac{y}{\log\ n} = n \quad \Leftrightarrow \quad \frac{n\ \log\ n}{\log\ n} = n$$

Obviously, if this were the case, we would not need the inverse of the function in the first place. However, given our observation about the size of the representation of n above, we can begin to approximate the actual value n by substituting n with $n \log n$, the original argument, inside the expression, and accounting for the error we introduce using $\varepsilon \epsilon \mathbb{R}$, where $\varepsilon > 0$.

$$\frac{n\ \log\ n}{\log\ (n\ \log\ n) - \varepsilon} \approx n$$

We now need an optimal value ε^* for ε. Through some algebraic manipulation, we can write the value of ε^* as a function of our original argument $n \log n$.

$$\begin{aligned}
n\ \log\ n &= n(\log\ (n\ \log\ n) - \varepsilon^*) \\
\log\ n &= \log\ (n\ \log\ n) - \varepsilon^* \\
\varepsilon^* &= \log\ (n\ \log\ n) - \log\ n
\end{aligned}$$

By using the properties of logarithms, we can obviously simplify further to obtain the value $\log \log n$ mentioned above.

$$\varepsilon^* = \log\left(\frac{n\ \log\ n}{\log\ n}\right) = \log \log n$$

We have effectively reduced the problem of finding our result n to the problem of finding the value $\log \log n$, which constitutes significant progress, as the observation that

$$\Big|\ \log \log\ (n\ \log\ n) - \log \log n\ \Big| \ll \Big|\ \log\ (n\ \log\ n) - \log\ n\ \Big|$$

illustrates. Clearly any approximation we make for the value $\log \log n$ using our original argument $n \log n$ will result in a smaller error, because our approximation is *inside* a log expression.

Our goal is now to approximate $\log \log n$. We know that $\log \log n$ can be defined as a function only of itself and of our original input.

$$\log \log n = \log(\log(n \log n) - \log \log n)$$

We notice that we can simply find an approximation of $\log\log n$ by some means, and then recalculate the above equation to find a more precise approximation. Thus, we can express our desired value $\log\log n$ using an open form with nested logarithms.

$$\log\ \log\ n = \log\left[\frac{\log(n\log n)}{\log\left[\frac{\log(n\log n)}{\log\left[\frac{\log(n\log n)}{\log[\ldots]}\right]}\right]}\right]$$

We can take this calculation to an arbitrary depth depending on our desired precision by starting with some approximation z_0 and continuously recalculating our approximation of $\log\log n$ using a recurrence relation.

$$z_0 \approx \log\log n$$
$$z_i = \log(\log(n\log n) - z_{i-1})$$

One can also attempt to solve this recurrence relation, substituting $\log\log n$ by letting $z \equiv \log\log n$.

$$z = \log(\log(n\log n) - z)$$

We can simplify the problem using algebra, assuming that our logarithms are of some base b.

$$\begin{aligned} b^z &= b^{\log(\log(n\log n)-z)} \\ b^z &= \log(n\log n) - z \\ b^z + z &= \log(n\log n) \end{aligned}$$

The solution to this would yield $z = \log\log n$, as we would expect. At this point, it is of interest to note that $b^z + z \approx b^z$, by L'Hopital's rule.

$$\lim_{z\to\infty}\left(\frac{b^z + z}{b^z}\right) = 1$$

We see that $b^z \approx b^z + z$, and use this fact to our advantage. An approximation can be obtained by solving a simpler equation; this approximation illustrates the same point we discussed above regarding the relative insignificance of differences in representation sizes.

$$\log(n\log n) = b^z \ \Rightarrow\ z = \log\log(n\log n)$$

Thus, in the absence of a better method of approximation for z_0, we can simply resort to this equation, and let $z_0 = \log\log(n\log n)$. Given any estimate $z_{i-1} = \log\log n + \delta_{z_{i-1}}$, we note that our estimation error shrinks exponentially with each iteration.

$$\begin{aligned}
z_i &= \log(\log(n \log n) - z_{i-1}) \\
\log\log n + \delta_{z_i} &= \log(\log(n \log n) - \log\log n - \delta_{z_{i-1}}) \\
\delta_{z_i} &= \log(\log n - \delta_{z_{i-1}}) - \log\log n \\
\delta_{z_i} &= \log\Big(1 - \frac{\delta_{z_{i-1}}}{\log n}\Big)
\end{aligned}$$

Bounds on precision parameterized by the depth to which we take our calculation can be obtained for any logarithmic base, if such bounds are desired. We now have a method which can approximate n for large input values, but it is not suited well to small values. We will look at another approach for this scale.

3 Approximation Near the Origin

To ensure further precision, one can even utilize an existing approximation formula for logarithms after a good range for such an approximation has been reached in the above continued fraction or sequence. One such approximation is Borchardt's algorithm.

$$x \ \ln \ x \approx \frac{6x(x-1)}{x + 4\sqrt{x} + 1}$$

As this approximation is not readily adjustable, we will investigate a different heuristic approximation which is also suited well only to points close to the origin but which is parameterized to some degree.

We will begin with the familiar series expansion for the natural logarithm.

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots = \sum_{i=0}^{\infty} \frac{x^i}{i}(-1)^i$$

We want each term x^i to have a denominator $\frac{1}{(\mu(i))^i}$ for some function μ. This allows us to write the below *sequence*, though its closed form will not be obvious.

$$\sum_{i=0}^{\infty} \frac{x^i}{(\mu(i))^i} = \sum_{i=0}^{\infty} \Big(\frac{x}{\mu(i)}\Big)^i$$

We also want the two summations to be equivalent in order to obtain the sequence for $\ln(1+x)$.

$$\sum_{i=0}^{\infty} \frac{x^i}{i} = \sum_{i=0}^{\infty} \frac{x^i}{(\mu(i))^i}$$

This leads to a condition which must be satisfied to make the above equality hold.

$$(\mu(i))^i = i$$

Though we could find a more interesting function for μ, it would be useful in this case to make $\mu(i)$ constant for all $i \in \mathbb{N}$, as will become evident below. In this case, we must satisfy a condition in solving for μ.

$$\forall i,\ \mu^i = i \ \Rightarrow\ \mu = \sqrt[i]{i}$$

While there is no solution μ in this case, its value is bounded.

$$\lim_{i \to \infty} \sqrt[i]{i} = 1 \ \Rightarrow\ 1 < \mu < \sqrt{2}$$

We now recall the ordinary generating function for the sequence $\{1, c^{-1}, c^{-2}, c^{-3}, ...\}$ for some $c \in \mathbb{Z}$.

$$\frac{1}{1 - \frac{x}{c}} = 1 + \frac{x}{c} + \frac{x^2}{c^2} + \frac{x^3}{c^3} + \ldots = \sum_{i=0}^{\infty} \frac{x^i}{c^i}$$

We see that we can approximate our series for $\ln(x+1)$ using μ by setting $c \equiv \mu$.

$$\frac{1}{1 - \frac{x}{\mu}} = 1 + \frac{x}{\mu} + \frac{x^2}{\mu^2} + \frac{x^3}{\mu^3} + \ldots = \sum_{i=0}^{\infty} \frac{x^i}{\mu^i}$$

Given some μ such that $\mu \approx \sqrt[i]{i}$ $\ \forall i$ such that $i \in \mathbb{N}$, we now have an approximation for $\ln(1+x)$.

$$1 - \frac{1}{1 - \frac{x}{-\mu}} = \sum_{i=0}^{\infty} \frac{x^i}{\mu^i}(-1)^i \approx \sum_{i=0}^{\infty} \frac{x^i}{i}(-1)^i = \ln(1+x)$$

Let $n = 1 + x$ for clarity.

$$\ln(1+x) \approx 1 - \frac{1}{1 + \frac{x}{\mu}}$$

With regard to various bases for the intended logarithm calculation, we would have introduced a constant factor into the formula, because $n \ln n = n(\frac{\log n}{\log e})$.

$$n\Big(\frac{\log n}{\log e}\Big) \approx n\Big(1 - \frac{1}{1 + \frac{n-1}{\mu}}\Big) = \Big(n - \frac{\mu n}{\mu + n - 1}\Big) = \frac{n(\mu + n - 1) - \mu n}{\mu + n - 1} = \frac{n^2 - n}{\mu + n - 1}$$

We obtain a polynomial of degree two, which we can solve in the usual manner.

$$
\begin{aligned}
&(n \log n)\mu + (n \log n)n - (n \log n) = (\log e)n^2 - (\log e)n \\
\Rightarrow\ \ &(\log e)n^2 - (\log e + n \log n)n + (n \log n)(1-\mu) = 0
\end{aligned}
$$

$$
n \approx \frac{(\log e + n \log n) \pm \sqrt{(\log e + n \log n)^2 - 4(\log e)(n \log n)(1-\mu)}}{2 \log e}
$$

Let $f(n) = n \log n$. We can now describe a formula approximating f^{-1}.

$$
\begin{aligned}
f^{-1}(y) &\approx \frac{(\log e + y) + \sqrt{(\log e + y)^2 - 4y(\log e)(1-\mu)}}{2 \log e} = \\
&= \frac{(1+y) + \sqrt{y^2 + (\log e)^2 + (4\mu - 2)(y \log e)}}{2 \log e}
\end{aligned}
$$

Thus, we have illustrated two methods for approximating the inverse to the function $n \log n$, both adjustable to some degree. Note once again that the first method is suited best for "large" input values ($y \gg 1$), while the second is best suited to small values ($y < 1$).

References

[1] Wall, H. S. *Analytic Theory of Continued Fractions.* New York: Chelsea, 1948.

[2] Abramowitz and Stegan. *Handbook of Mathematical Functions.* New York: Dover Publications, Inc., 1970.

[3] Weisstein, Eric W., "Nested Logarithm." From *MathWorld*–A Wolfram Web Resource. http://mathworld.wolfram.com/NestedLogarithm.html.