Now, in order to estimate the time our algorithm takes, a crucial step is to estimate the probability that a random number less than x will be a product of primes less than y (where y is a number much less than x). To do this, we first let u denote the ratio $\frac{log\, x}{log\, y}$. That is, if x is an r-bit integer and y is an s-bit integer, then u is approximately the ratio of digits r/s.

In the course of the computations, we shall want to make some simplifications by ignoring smaller terms. We shall do this under the assumption that u is *much* smaller than y. We let $\pi(y)$, as usual, denote the number of prime numbers which are $\leq y$. Since $\pi(y)$ is approximately equal to $y/log\, y$, by the Prime Number Theorem, we are also assuming that we are working with values of u which are much smaller than $\pi(y)$. In a typical practical application of the algorithm, we might take y, u, x of approximately the following sizes:

$$\begin{aligned} &y \approx 10^6 \qquad \text{(so that } \pi(y) \approx 7 \cdot 10^4 \text{ and } log\, y \approx 14); \\ &u \approx 8; \\ &x \approx 10^{48}. \end{aligned}$$

It is customary to let $\Psi(x,y)$ denote the number of integers $\leq x$ which are not divisible by any prime greater than y, i.e., the number of integers which can be written as a product $\prod p_j^{\alpha_j} \leq x$, where the product is over all primes $\leq y$ and the α_j are nonnegative integers. There is obviously a 1-to-1 correspondence between $\pi(y)$-tuples of nonnegative integers α_j for which $\prod_j p_j^{\alpha_j} \leq x$ and integers $\leq x$ which are not divisible by any prime greater than y. Thus, $\Psi(x,y)$ is equal to the number of integer solutions α_j to the inequality $\sum_{j=1}^{\pi(y)} \alpha_j\, log\, p_j \leq log\, x$, as we see by taking logarithms. We now observe that most of the p_j's have logarithms not too much less than $log\, y$. This is because most of the primes less than y have almost the same number of digits as y; only relatively few have many fewer digits and hence a much smaller logarithm. Thus, we shall allow ourselves to replace $log\, p_j$ by $log\, y$ in the previous inequality. Dividing both sides of the resulting inequality by $log\, y$ and replacing $log\, x/log\, y$ by u, we can say that $\Psi(x,y)$ is approximately equal to the number of solutions of the inequality $\sum_{j=1}^{\pi(y)} \alpha_j \leq u$.

We now make another important simplification, replacing the number of variables $\pi(y)$ by y. This might appear at first to be a rather reckless modification of our problem. And in fact, replacing $\pi(y)$ by y does introduce nontrivial terms; however, it turns out that those terms cancel, and the net result is the same as one would get by a much more careful approximation of $\Psi(x,y)$. Thus, we shall suppose that $\Psi(x,y)$ is roughly equal to the number of y-tuple nonnegative integer solutions to the inequality $\sum_{j=1}^{y} \alpha_j \leq u$.

But, by Fact 2 (with $N = y$), this means that $\Psi(x,y)$ is approximately $\binom{[u]+y}{y}$. We now estimate $log\left(\frac{\Psi(x,y)}{x}\right)$, which is the logarithm of the probability that a random integer between 1 and x is a product of primes $\leq y$.