Asymptotics

ITT9132 Concrete Mathematics
Lecture 15 – 6 May 2019

Chapter Nine

- A Hierarchy
- O Notation
- O Manipulation

TAL TECH

Contents

Why asymptotics?

Suppose that an algorithm requires

$$S_n = \sum_{k=0}^{n} \binom{3n}{k}$$

operations to compute the output on an input of length n.

- There seems to be no closed formula for S_n.
- But we need some information of what is "more or less" S_n for large values of n. (So that we know that computation will finish by a certain time.)
- For example: between S_n and f_{4n}, which is "larger" when n is "large"?

Asymptotics provide a way to:

- determine what "more or less" means in this context;
- make comparisons between things that "grow unboundedly".

TAL TECH

Next section

A hierarchy between rates of growth

Definition

Let f and g be two real-valued functions defined on the natural numbers.
We say that g grows asymptotically faster than f, and write $f(n) \prec g(n)$, if the ratio $f(n)/g(n)$ converges to zero for $n \to \infty$.

A hierarchy between rates of growth

Definition

Let f and g be two real-valued functions defined on the natural numbers.
We say that g grows asymptotically faster than f, and write $f(n) \prec g(n)$, if the ratio $f(n)/g(n)$ converges to zero for $n \to \infty$.

- The relation $\prec$ is transitive:
  If $f(n) \prec g(n)$ and $g(n) \prec h(n)$ then $f(n) \prec h(n)$.
- The relation $\prec$ is not reflexive!
  For every $f : \mathbb{N} \to \mathbb{R}$, $\lim_{n\to\infty} f(n)/f(n) = 1 \neq 0$.
- For $\alpha, \beta > 0$, $n^\alpha \prec n^\beta$ if and only if $\alpha < \beta$.
- If $f(n)$ and $g(n)$ are never zero and $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} g(n) = 0$, then:

$$f(n) \prec g(n) \quad \text{if and only if} \quad \frac{1}{g(n)} \prec \frac{1}{f(n)} .$$

That is: the same notation works for the infinitely small as well as for the infinitely large.

TAL TECH

A hierarchy of functions

For $0 < \varepsilon < 1 < c$:

$$
\begin{aligned}
1 &\prec \log\log n \\
&\prec \log n \\
&\prec n^{\varepsilon} \\
&\prec n \\
&\prec n^{c} \\
&\prec n^{\log n} \\
&\prec c^{n} \\
&\prec n! \\
&\prec n^{n} \\
&\prec c^{c^{n}}
\end{aligned}
$$

We "think big":

we do not care *if* a function goes to infinity,
but how fast it does

Example: Iterated logarithm

Definition

Let $\vec{\alpha} = (\alpha_1, \ldots, \alpha_d)$ and $\vec{\beta} = (\beta_1, \ldots, \beta_d)$ be d-tuples of real numbers. We say that $\vec{\alpha}$ precedes $\vec{\beta}$ in lexicographic order, and write $\vec{\alpha} <_L \vec{\beta}$, if there exists $i \in [1:d]$ such that $\alpha_i < \beta_i$ and $\alpha_j = \beta_j$ for every $j < i$.

For example, $(1,2,3) <_L (2,0,0)$ and $(1,2,3,4,5) <_L (1,2,4,8,0)$.

Theorem

Let $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3 > 0$. The following are equivalent:

1. $n^{\alpha_1} (\log n)^{\alpha_2} (\log\log n)^{\alpha_3} \prec n^{\beta_1} (\log n)^{\beta_2} (\log\log n)^{\beta_3}$.
2. $(\alpha_1, \alpha_2, \alpha_3) <_L (\beta_1, \beta_2, \beta_3)$.

Warmup: where does $e^{\sqrt{\ln n}}$ belong?

Theorem

For every $\varepsilon > 0$, $\log n \prec e^{\sqrt{\ln n}} \prec n^{\varepsilon}$.

Warmup: where does $e^{\sqrt{\ln n}}$ belong?

Theorem

For every $\varepsilon > 0$, $\log n \prec e^{\sqrt{\ln n}} \prec n^{\varepsilon}$.

Lemma

Let $f(n)$ and $g(n)$ be such that $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} g(n) = +\infty$. Then:

$$e^{f(n)} \prec e^{g(n)} \text{ if and only if } \lim_{n\to\infty} (f(n) - g(n)) = -\infty .$$

Proof of the lemma:

$$\begin{aligned} e^{f(n)} \prec e^{g(n)} \quad & \text{if and only if} \quad \lim_{n\to\infty} \frac{e^{f(n)}}{e^{g(n)}} = 0 \\ & \text{if and only if} \quad \lim_{n\to\infty} e^{f(n)-g(n)} = 0 \\ & \text{if and only if} \quad \lim_{n\to\infty} f(n) - g(n) = -\infty \end{aligned}$$

Q.E.D.

Warmup: where does $e^{\sqrt{\ln n}}$ belong?

Theorem

For every $\varepsilon > 0$, $\log n \prec e^{\sqrt{\ln n}} \prec n^{\varepsilon}$.

Proof of the theorem:

- By the lemma,

$$\lim_{n\to\infty} \frac{\log n}{e^{\sqrt{\log n}}} = 0 \text{ if and only if } \lim_{n\to\infty} \log\log n - \sqrt{\log n} = -\infty,$$

  which is the case.

- Similarly,

$$\lim_{n\to\infty} \frac{e^{\sqrt{\log n}}}{n^{\varepsilon}} = 0 \text{ if and only if } \lim_{n\to\infty} \sqrt{\log n} - \varepsilon \log n = -\infty,$$

  which is the case.

An equivalence of rates of growth

Definition

We say that $f(n)$ and $g(n)$ have the same rate of growth, and write $f(n) \asymp g(n)$, if there exist $n_0 \in \mathbb{N}$ and $C > 0$ such that

$$|f(n)| \leqslant C \cdot |g(n)| \text{ and } |g(n)| \leqslant C \cdot |f(n)| \ \forall n \geqslant n_0 \,.$$

$\asymp$ is easily shown to be an equivalence relation.
If, in particular, $\lim_{n \to \infty} f(n)/g(n) = 1$, then we write $f(n) \sim g(n)$.

An equivalence of rates of growth

Definition

We say that $f(n)$ and $g(n)$ have the same rate of growth, and write $f(n) \asymp g(n)$, if there exist $n_0 \in \mathbb{N}$ and $C > 0$ such that

$$|f(n)| \leqslant C \cdot |g(n)| \text{ and } |g(n)| \leqslant C \cdot |f(n)| \;\; \forall n \geqslant n_0 \,.$$

$\asymp$ is easily shown to be an equivalence relation.
If, in particular, $\lim_{n \to \infty} f(n)/g(n) = 1$, then we write $f(n) \sim g(n)$.

$\prec$ and $\asymp$ do not define a preorder!

It **can** be that none of the three relations:

- $f(n) \prec g(n)$,
- $g(n) \prec f(n)$,
- $f(n) \asymp g(n)$

holds. For example, it could be $\liminf_{n \geqslant 1} \frac{|f(n)|}{|g(n)|} = 0$ and $\limsup_{n \geqslant 1} \frac{|f(n)|}{|g(n)|} = +\infty$.

TAL TECH

The class $\mathfrak{L}$ of logarithmico-exponential functions

Godfrey Harold Hardy defined the class $\mathfrak{L}$ of logarithmico-exponential functions as the smallest class of functions satisfying the following properties:

1. For every $\alpha \in \mathbb{R}$, the constant function $f(n) = \alpha$ is in $\mathfrak{L}$.
2. The identity function $f(n) = n$ is in $\mathfrak{L}$.
3. If $f(n), g(n) \in \mathfrak{L}$, then $f(n) - g(n) \in \mathfrak{L}$.
4. If $f(n) \in \mathfrak{L}$, then $e^{f(n)} \in \mathfrak{L}$.
5. If $f(n) \in \mathfrak{L}$ is ultimately positive, then $\ln f(n) \in \mathfrak{L}$.

A function f ultimately has property P if $f(n)$ has property P for every n large enough (equivalently, if $f(n)$ does *not* have property P only for *finitely many* values of n)

The class $\mathfrak{L}$ of logarithmico-exponential functions

Godfrey Harold Hardy defined the class $\mathfrak{L}$ of logarithmico-exponential functions as the smallest class of functions satisfying the following properties:

1. For every $\alpha \in \mathbb{R}$, the constant function $f(n) = \alpha$ is in $\mathfrak{L}$.
2. The identity function $f(n) = n$ is in $\mathfrak{L}$.
3. If $f(n), g(n) \in \mathfrak{L}$, then $f(n) - g(n) \in \mathfrak{L}$.
4. If $f(n) \in \mathfrak{L}$, then $e^{f(n)} \in \mathfrak{L}$.
5. If $f(n) \in \mathfrak{L}$ is ultimately positive, then $\ln f(n) \in \mathfrak{L}$.

A function f ultimately has property P if $f(n)$ has property P for every n large enough (equivalently, if $f(n)$ does *not* have property P only for *finitely many* values of n)

Examples

If $f(n), g(n) \in \mathfrak{L}$, then so are:

- $f(n) + g(n) = f(n) - (0 - g(n))$;
- $f(n) \cdot g(n) = e^{\ln f(n) + \ln g(n)}$ and $f(n)/g(n) = e^{\ln f(n) - \ln g(n)}$
  if $f(n), g(n) > 0$ for every n large enough;
- $\sqrt{f(n)} = e^{\frac{1}{2} \ln n}$
  again if $f(n) > 0$ for every n large enough;
- ...

The class $\mathfrak{L}$ of logarithmico-exponential functions

Godfrey Harold Hardy defined the class $\mathfrak{L}$ of logarithmico-exponential functions as the smallest class of functions satisfying the following properties:

1. For every $\alpha \in \mathbb{R}$, the constant function $f(n) = \alpha$ is in $\mathfrak{L}$.
2. The identity function $f(n) = n$ is in $\mathfrak{L}$.
3. If $f(n), g(n) \in \mathfrak{L}$, then $f(n) - g(n) \in \mathfrak{L}$.
4. If $f(n) \in \mathfrak{L}$, then $e^{f(n)} \in \mathfrak{L}$.
5. If $f(n) \in \mathfrak{L}$ is ultimately positive, then $\ln f(n) \in \mathfrak{L}$.

A function f ultimately has property P if $f(n)$ has property P for every n large enough (equivalently, if $f(n)$ does *not* have property P only for *finitely many* values of n)

Theorem 1 (Hardy)

Every function $f(n) \in \mathfrak{L}$ is

- either ultimately positive,
- or ultimately negative,
- or identically zero.

The class $\mathfrak{L}$ of logarithmico-exponential functions

Godfrey Harold Hardy defined the class $\mathfrak{L}$ of logarithmico-exponential functions as the smallest class of functions satisfying the following properties:

1. For every $\alpha \in \mathbb{R}$, the constant function $f(n) = \alpha$ is in $\mathfrak{L}$.
2. The identity function $f(n) = n$ is in $\mathfrak{L}$.
3. If $f(n), g(n) \in \mathfrak{L}$, then $f(n) - g(n) \in \mathfrak{L}$.
4. If $f(n) \in \mathfrak{L}$, then $e^{f(n)} \in \mathfrak{L}$.
5. If $f(n) \in \mathfrak{L}$ is ultimately positive, then $\ln f(n) \in \mathfrak{L}$.

A function f ultimately has property P if $f(n)$ has property P for every n large enough (equivalently, if $f(n)$ does *not* have property P only for *finitely many* values of n)

Theorem 2 (Hardy)

For every two functions $f(n), g(n) \in \mathfrak{L}$,

- either $f(n) \prec g(n)$,
- or $g(n) \prec f(n)$,
- or $f(n) \asymp g(n)$: in which case $f(n) \sim \alpha g(n)$ for suitable α.

Next section

1 A Hierarchy

2 O notation

3 O Manipulation

O notation

Definition

We write $f(n) = O(g(n))$, and say that $f(n)$ is big-O of $g(n)$, if $|f(n)| \leqslant C \cdot |g(n)|$ for a suitable constant C and for every n.

O notation

Definition

We write $f(n) = O(g(n))$, and say that $f(n)$ is big-O of $g(n)$, if $|f(n)| \leqslant C \cdot |g(n)|$ for a suitable constant C and for every n.

Note: this is a small abuse of notation.

- Rather than a single function, $O(g(n))$ is a class of functions $f(n)$ such that $|f(n)| \leqslant C \cdot |g(n)|$ for a suitable constant C and for every n.
- Therefore, it would be more precise to write $f(n) \in O(g(n))$ rather than $f(n) \in O(g(n))$.
- However, the symbol of equality is much easier to use, and it is clear from the context that we mean membership to a class.
- The same convention holds when comparing big-O classes:
  By writing $O(f(n)) = O(g(n))$ we usually mean $O(f(n)) \subseteq O(g(n))$ instead.

In other words: with big-O notation, we usually mean equalities to be "left to right", but not necessarily "right to left".

TAL TECH

O notation

Definition

We write $f(n) = O(g(n))$, and say that $f(n)$ is big-O of $g(n)$, if $|f(n)| \leqslant C \cdot |g(n)|$ for a suitable constant C and for every n.

Example:

- Let $\square_n = \sum_{k=0}^{n} k^2$. Then $\square_n = O(n^3)$ with $C = 1$, because for every n:

$$|\square_n| = \left| \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n \right| \leqslant \frac{1}{3}|n^3| + \frac{1}{2}|n^2| + \frac{1}{6}|n| \leqslant \left(\frac{1}{3} + \frac{1}{2} + \frac{1}{6} \right) |n^3| = |n^3| .$$

- Also, $\square_n = \dfrac{1}{3}n^3 + O(n^2)$, where the O notation has $C = \frac{1}{2}$.
- Also, $\square_n = O(n^{17})$. We can be as "sloppy" as we want!

A cautionary tale

Student:

I have found a more efficient algorithm!
The old algorithm requires $O(n \ln n \ln \ln n)$ bits of memory:
my algorithm only requires $O(n \ln n)$ bits!

A cautionary tale

Student:

I have found a more efficient algorithm!
The old algorithm requires $O(n \ln n \ln \ln n)$ bits of memory:
my algorithm only requires $O(n \ln n)$ bits!

Robert Sedgewick:

. . . but in the real world, "$\ln \ln n$" means "six"!

(There are less than 10^{100} atoms in the universe; $\ln \ln 10^{100} = 5.4392026\ldots$.)

Big-Omega and big-Theta notation

Definition

We write $f(n) = \Omega(g(n))$, and say that $f(n)$ is big-Omega of $g(n)$, if $|f(n)| \geqslant C \cdot |g(n)|$ for a suitable constant C and for every n.

In other words:

Definition

We write $f(n) = \Theta(g(n))$, and say that $f(n)$ is big-Theta of $g(n)$, if both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$

Note that:

- $f(n) = \Omega(g(n))$ if and only if $g(n) = O(f(n))$;
- $f(n) = \Theta(g(n))$ if and only if $f(n) \asymp g(n)$.

Again, equalities are meant to be "left to right".

Next section

Basic operations

$$\begin{aligned} f(n) &= O(f(n)) \\ c \cdot O(f(n)) &= O(f(n)) \ , \ c \text{ constant} \\ O(O(f(n))) &= O(f(n)) \\ O(f(n)) \cdot O(g(n)) &= O(f(n) \cdot g(n)) \\ f(n) \cdot O(g(n)) &= O(f(n) \cdot g(n)) \\ n^\alpha &= O(n^\beta) \ , \ \text{if } \alpha \leqslant \beta \\ O(f(n)) + O(g(n)) &= O(|f(n)| + |g(n)|) \end{aligned}$$

The following rule is also useful:

$$f(n) = g(n) + O(h(n)) \text{ if and only if } g(n) = f(n) + O(h(n))$$

Indeed, each side can be obtained from the other by multiplying by -1 and adding $f(n) + g(n)$.

Warmup: O notation and sums

Problem

Prove or disprove: if $f(n)$ and $g(n)$ are positive for all n, then

$$O(f(n)+g(n)) = f(n)+O(g(n))$$

Warmup: O notation and sums

Problem

Prove or disprove: if $f(n)$ and $g(n)$ are positive for all n, then

$$O(f(n)+g(n)) = f(n) + O(g(n))$$

Solution

Let us check what the two classes actually are:

- $O(f(n)+g(n))$ is the class of the functions $h(n)$ such that:

$$\exists C > 0 . \forall n \in \mathbb{N} . |h(n)| \leqslant C \cdot |f(n)+g(n)|$$

- $f(n) + O(g(n))$ is the class of the functions $h(n)$ such that:

$$\exists k : \mathbb{N} \to \mathbb{C}, C > 0 . \forall n \in \mathbb{N} . h(n) = f(n) + k(n) \text{ and } |k(n)| \leqslant C \cdot |g(n)|$$

If $f(n) = n$ and $g(n) = 1$, then $h(n) = 2n$ belongs to the first class, but not to the second one.

Power series and O notation

Let $S(z) = \sum_{n \geqslant 0} a_n z^n$ be a power series.

- Suppose $S(z)$ converges absolutely for some $z_0 \in \mathbb{C}$, that is,

$$\sum_{n \geqslant 0} |a_n| \cdot |z_0|^n < \infty$$

- Then the following holds:

$$S(z) = O(1) \ \forall |z| \leqslant |z_0|$$

- Indeed, in this case,

$$|S(z)| \leqslant \sum_{n \geqslant 0} |a_n| \cdot |z|^n \leqslant \sum_{n \geqslant 0} |a_n| \cdot |z_0|^n < \infty.$$

Power series and O notation

Let $S(z) = \sum_{n \geqslant 0} a_n z^n$ be a power series.

- Suppose $S(z)$ converges absolutely for some $z_0 \in \mathbb{C}$, that is,

$$\sum_{n \geqslant 0} |a_n| \cdot |z_0|^n < \infty$$

- Then the following holds:

$$S(z) = O(1) \ \forall |z| \leqslant |z_0|$$

- Indeed, in this case,

$$|S(z)| \leqslant \sum_{n \geqslant 0} |a_n| \cdot |z|^n \leqslant \sum_{n \geqslant 0} |a_n| \cdot |z_0|^n < \infty.$$

Of course, the following relations are also valid:

- $S(z) = a_0 + O(z)$;
- $S(z) = a_0 + a_1 z + O(z^2)$;
- $S(z) = a_0 + a_1 z + \ldots + a_{16} z^{16} + O(z^{17})$;
- ...

A table of asymptotic approximations

$$
\begin{aligned}
H_n &= \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} + O\left(\frac{1}{n^6}\right) \\
n! &= \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + O\left(\frac{1}{n^4}\right)\right) \\
B_n &= 2\,[n\,\text{even}]\,(-1)^{n/2-1}\frac{n!}{(2\pi)^n} \cdot \left(1 + \frac{1}{2^n} + \frac{1}{3^n} + O\left(\frac{1}{4^n}\right)\right) \\
\pi(n) &= \frac{n}{\ln n} + \frac{n}{(\ln n)^2} + \frac{2!n}{(\ln n)^3} + \frac{3!n}{(\ln n)^4} + O\left(\frac{n}{(\log n)^5}\right) \\
e^z &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + O(z^5) \\
\ln(1+z) &= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5) \\
\frac{1}{1-z} &= 1 + z + z^2 + z^3 + z^4 + O(z^5) \\
(1+z)^\alpha &= 1 + \alpha z + \binom{\alpha}{2}z^2 + \binom{\alpha}{3}z^3 + \binom{\alpha}{4}z^4 + O(z^5)
\end{aligned}
$$

We use log instead of ln in the big-O notation, because changing base corresponds to multiply by a nonzero constant.

Errors

Absolute error

An asymptotic approximation for a function $f(n)$ has absolute error $O(g(n))$ if it has the form

$$f(n) = h(n) + O(g(n))$$

for some $h(n)$ which does not involve O-notation.

That is: the absolute error $|f(n) - h(n)|$ is bounded by a multiple of $|g(n)|$.

Relative error

An asymptotic approximation for a function $f(n)$ has relative error $O(g(n))$ if it has the form

$$f(n) = h(n) \cdot (1 + O(g(n)))$$

for some $h(n)$ which does not involve O-notation.

That is: the relative error $\left|\frac{f(n)}{h(n)} - 1\right|$ is bounded by a multiple of $|g(n)|$.

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

To simplify notation, we write p instead of P_n.
We use:

$$\pi(n) = \frac{n}{\ln n} + O\left(\frac{n}{(\log n)^2}\right), \text{ so that for } p = P_n, \; n = \pi(p) = \frac{p}{\ln p} + O\left(\frac{p}{(\log p)^2}\right).$$

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

To simplify notation, we write p instead of P_n.
We use:

$$\pi(n) = \frac{n}{\ln n} + O\left(\frac{n}{(\log n)^2}\right), \text{ so that for } p = P_n, \; n = \pi(p) = \frac{p}{\ln p} + O\left(\frac{p}{(\log p)^2}\right).$$

By the Prime Number Theorem, $\lim_{n\to\infty} \frac{n \ln p}{p} = 1$, so $\frac{p}{\ln p} = O(n)$, and as $p \geqslant n$:

$$O\left(\frac{p}{(\log p)^2}\right) = O\left(\frac{n}{\log p}\right) = O\left(\frac{n}{\log n}\right)$$

By applying the swapping rule to the original estimate:

$$\frac{p}{\ln p} = n + O\left(\frac{p}{(\log p)^2}\right) = n + O\left(\frac{n}{\log n}\right)$$

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

To simplify notation, we write p instead of P_n.
We use:

$$\pi(n) = \frac{n}{\ln n} + O\left(\frac{n}{(\log n)^2}\right), \text{ so that for } p = P_n, \; n = \pi(p) = \frac{p}{\ln p} + O\left(\frac{p}{(\log p)^2}\right).$$

The estimate $\dfrac{p}{\ln p} = n + O\left(\dfrac{n}{\log n}\right)$ yields an "approximate recurrence":

$$p = n \ln p \cdot \left(1 + O\left(\frac{1}{\log n}\right)\right)$$

By applying logarithms and observing that $\ln(1 + O(f(n))) = O(f(n))$ we get:

$$\ln p = \ln n + \ln \ln p + O\left(\frac{1}{\log n}\right)$$

Now we have to remove that "$\ln \ln p$" from the right-hand side ...

TAL TECH

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

To simplify notation, we write p instead of P_n.
We have found:

$$\ln p = \ln n + \ln\ln p + O\left(\frac{1}{\log n}\right)$$

Lemma

$p = O(n^2)$.

This lemma is weak, but it implies that $\ln\ln p = O(\log\log n)$—which is sufficient for us.
Proof: As $p = (n \ln p) \cdot (1 + O(1/\log n))$, squaring and dividing by pn^2 we get:

$$\frac{p}{n^2} = \frac{(\ln p)^2}{p} \cdot \left(1 + O\left(\frac{1}{\log n}\right)\right)$$

because $(1 + O(1 + 1/\log n))^2 = 1 + O(1/\log n)$, and the right-hand side vanishes for $n \to \infty$.

Example: The nth prime number

Problem

Provide an asymptotic estimate for the nth prime number P_n as a function of n.

To simplify notation, we write p instead of P_n.
We have found:

$$\ln p = \ln n + \ln \ln p + O\left(\frac{1}{\log n}\right)$$

But if $p = O(n^2)$, then $\ln p = O(\log n)$ and $\ln \ln p = O(\log \log n)$, and substituting:

$$\ln p = \ln n + O(\log \log n) + O\left(\frac{1}{\log p}\right) = \ln n + O(\log \log n)$$

Plugging in again we get:

$$\ln p = \ln n + \ln \ln n + O\left(\frac{\log \log n}{\log n}\right)$$

and substituting into $p = (n \ln p) \cdot (1 + O(1/\log n))$ yields:

$$p = n \ln n + n \ln \ln n + O(n)$$