Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Federal Institute of Technology at Zurich

Departement of Computer Science
Johannes Lengler, David Steurer
Kasper Lindberg, Lucas Slot, Hongjie Chen, Manuel Wiedmer

23 September 2024

Algorithms & Data Structures — Exercise sheet 0 — HS 24

The solutions for this sheet do not have to be submitted. The sheet will be solved in the first exercise session on 23 September 2024.

Exercises that are marked by * are challenge exercises.

You can use results from previous parts without solving those parts.

The solutions are intended to help you understand how to solve the exercises and are thus more detailed than what would be expected at the exam. All parts that contain explanation that you would not need to include in an exam are in grey.

Exercise 0.1 *Induction.*

(a) Prove by mathematical induction that for any positive integer n,

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

In your solution, you should address the base case, the induction hypothesis and the induction step.

Solution:

Base Case.
Let $n = 1$. Then we have

$$1 = \frac{1 \cdot 2}{2}.$$

Induction Hypothesis.
Assume that the property holds for some positive integer k, that is we have

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2}.$$

Induction Step.
We must show that the property holds for $k + 1$ summands. We have

$$\begin{aligned} 1 + 2 + \cdots + k + (k+1) &\overset{\text{I.H.}}{=} \frac{k(k+1)}{2} + k + 1 \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2}. \end{aligned}$$

We want to use the induction hypothesis, so in the first step above we separate the last term of the sum and then use the induction hypothesis.

By the principle of mathematical induction, the statement is true for any positive integer n.

(b) **(This subtask is from August 2019 exam).** Let $T : \mathbb{N} \to \mathbb{R}$ be a function that satisfies the following two conditions:

$$\begin{aligned} T(n) &\geq 4 \cdot T(\tfrac{n}{2}) + 3n \qquad \text{whenever } n \text{ is divisible by 2;} \\ T(1) &= 4. \end{aligned}$$

Prove by mathematical induction that

$$T(n) \geq 6n^2 - 2n$$

holds whenever n is a power of 2, i.e., $n = 2^k$ with $k \in \mathbb{N}_0$. In your solution, you should address the base case, the induction hypothesis and the induction step.

Solution:

We solve this exercise by mathematical induction over k.

Base Case.
Let $k = 0$. Then we have $n = 2^0 = 1$ and

$$T(1) = 4 \geq 6 \cdot 1^2 - 2 \cdot 1.$$

Since we do the induction over k, we need to go from $k = \ell$ to $\ell + 1$ here. Hence, the induction hypothesis is $T(2^\ell) \geq 6(2^\ell)^2 - 2(2^\ell)$ for some ℓ and we want to show that $T(2^{\ell+1}) \geq 6(2^{\ell+1})^2 - 2(2^{\ell+1})$. In order to make the argument simpler, we introduce a new variable $m = 2^\ell$. Then, we have that $2^{\ell+1} = 2m$. Hence, we the induction hypothesis is $T(m) \geq 6m^2 - 2m$ and we need to show that $T(2m) \geq 6(2m)^2 - 2(2m)$. Thus, we get the following:

Induction Hypothesis.
Assume that the property holds for some positive integer $m = 2^\ell$. That is, we assume

$$T(m) \geq 6m^2 - 2m.$$

Induction Step.
Thus, we must show that the property holds for $2m = 2^{\ell+1}$. We have

$$\begin{aligned} T(2m) &\geq 4 \cdot T(m) + 3 \cdot (2m) \\ &\overset{\text{I.H.}}{\geq} 24m^2 - 8m + 6m \\ &= 24m^2 - 2m \\ &\geq 24m^2 - 4m \\ &= 6 \cdot (2m)^2 - 2 \cdot (2m). \end{aligned}$$

The following derivation is to illustrate how to come up with the above chain of inequalities. It does (and should) not be part of the final solution.

In a first step, we want to use the induction hypothesis. In order to do that, we need to replace the $T(2m)$ term by $T(m)$. The way we do this is by using the recurrence relation. Thus, we get that

$$T(2m) \geq 4 \cdot T(m) + 3 \cdot (2m) \overset{\text{I.H.}}{\geq} 24m^2 - 8m + 6m.$$

Now, we want to show that this is at least $6 \cdot (2m)^2 - 2 \cdot (2m) = 24m^2 - 4m$. Comparing this, the first terms are equal, so what remains is to show that $-8m + 6m \geq -4m$. But this is true since the left hand side is $-2m$ and $-2 \geq -4$. Putting all this together, we get the chain of inequalities from above.

By the principle of mathematical induction, the statement is true for any integer n that is a power of 2.

Asymptotic Growth

When we estimate the number of elementary operations executed by algorithms, it is often useful to ignore smaller order terms, and instead focus on the asymptotic growth defined below. We denote by $\mathbb{R}^+$ the set of all (strictly) positive real numbers and by $\mathbb{R}_0^+$ the set of nonnegative real numbers.

Definition 1. Let $f, g : \mathbb{N} \to \mathbb{R}^+$ be two functions. We say that f *grows asymptotically faster than* g if $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.

This definition is also valid for functions defined on $\mathbb{R}^+$ instead of $\mathbb{N}$. In general, $\lim\limits_{n\to\infty} \frac{g(n)}{f(n)}$ is the same as $\lim\limits_{x\to\infty} \frac{g(x)}{f(x)}$ if the second limit exists.

For all the following exercises, you can assume that $n \in \mathbb{N}_{\geq 10}$. We make this assumption so that all functions are well-defined and take values in $\mathbb{R}^+$.

Exercise 0.2 *Comparison of functions part 1.*

Show that

(a) $f(n) := n \log n$ grows asymptotically faster than $g(n) := n$.

Solution:

We have

$$\lim_{n\to\infty} \frac{n}{n \log n} = \lim_{n\to\infty} \frac{1}{\log n} = 0$$

and hence, by Definition 1, $f(n) := n \log n$ grows asymptotically faster than $g(n) := n$.

(b) $f(n) := n^3$ grows asymptotically faster than $g(n) := 10n^2 + 100n + 1000$.

Solution:

We have

$$\lim_{n\to\infty} \frac{10n^2 + 100n + 1000}{n^3} = \lim_{n\to\infty} \left(\frac{10}{n} + \frac{100}{n^2} + \frac{1000}{n^3} \right) = 0.$$

Hence, by Definition 1, $f(n) := n^3$ grows asymptotically faster than $g(n) := 10n^2 + 100n + 1000$.

(c) $f(n) := 3^n$ grows asymptotically faster than $g(n) := 2^n$.

Solution:

We have

$$\lim_{n\to\infty} \frac{2^n}{3^n} = \lim_{n\to\infty} \left(\frac{2}{3} \right)^n = 0,$$

and thus by Definition 1, $f(n) := 3^n$ grows asymptotically faster than $g(n) := 2^n$.

Here, it is important that the term $\frac{2}{3} < 1$, otherwise the limit would not be 0.

The following theorem can be useful to compute some limits.

Theorem 1 (L'Hôpital's rule). *Assume that functions $f : \mathbb{R}^+ \to \mathbb{R}^+$ and $g : \mathbb{R}^+ \to \mathbb{R}^+$ are differentiable, $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = \infty$ and for all $x \in \mathbb{R}^+$, $g'(x) \neq 0$. If $\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = C \in \mathbb{R}_0^+$ or $\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \infty$, then*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

Exercise 0.3 *Comparison of functions part 2.*

Show that

(a) $f(n) := n^{1.01}$ grows asymptotically faster than $g(n) := n \ln n$.

Solution:

Our goal is to apply Theorem 1. For this, we view the functions f and g as functions from $\mathbb{R}^+ \to \mathbb{R}^+$ and use the the fact that if the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ exists, then we have

$$\lim_{x\to\infty} \frac{g(x)}{f(x)} = \lim_{n\to\infty} \frac{g(n)}{f(n)},$$

where $x \in \mathbb{R}^+$ and $n \in \mathbb{N}$. Thus, it is sufficient to show that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$. We use this in several parts of this exercise.

We apply Theorem 1 to compute

$$\lim_{x\to\infty} \frac{x \ln x}{x^{1.01}} = \lim_{x\to\infty} \frac{\ln x}{x^{0.01}} \overset{\text{Thm.1}}{=} \lim_{x\to\infty} \frac{(\ln x)'}{(x^{0.01})'} = \lim_{x\to\infty} \frac{1/x}{0.01x^{-0.99}} = \lim_{x\to\infty} \frac{1}{0.01x^{0.01}} = 0.$$

Hence, by Definition 1, $f(n) := n^{1.01}$ grows asymptotically faster than $g(n) := n \ln n$.

(b) $f(n) := e^n$ grows asymptotically faster than $g(n) := n$.

Solution:

We apply Theorem 1 to compute

$$\lim_{x\to\infty} \frac{x}{e^x} \overset{\text{Thm.1}}{=} \lim_{x\to\infty} \frac{x'}{(e^x)'} = \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

Hence by Definition 1, $f(n) := e^n$ grows asymptotically faster than $g(n) := n$.

(c) $f(n) := e^n$ grows asymptotically faster than $g(n) := n^2$.

Solution:

We apply Theorem 1 to compute

$$\lim_{x\to\infty} \frac{x^2}{e^x} \overset{\text{Thm.1}}{=} \lim_{x\to\infty} \frac{(x^2)'}{(e^x)'} = \lim_{x\to\infty} \frac{2x}{e^x} \overset{\text{Thm.1}}{=} 2 \lim_{x\to\infty} \frac{x'}{(e^x)'} = 2 \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

Hence, by Definition 1, $f(n) := e^n$ grows asymptotically faster than $g(n) := n^2$.

(d)* $f(n) := 1.01^n$ grows asymptotically faster than $g(n) := n^{100}$.

Solution:

Note that we can rewrite $\frac{g(x)}{f(x)}$ as

$$\frac{x^{100}}{(1.01)^x} = \frac{e^{100 \ln x}}{e^{x \ln(1.01)}} = e^{100 \ln x - \ln(1.01)x}.$$

The goal now is to show that $\lim_{x\to\infty} 100 \ln x - \ln(1.01)x = -\infty$. This will allow us to conclude that $\lim_{x\to\infty} \frac{x^{100}}{(1.01)^x} = \lim_{x\to\infty} e^{100 \ln x - \ln(1.01)x} = 0$.

We have

$$\begin{aligned}\lim_{x\to\infty}(100 \ln x - \ln(1.01)x) &= \lim_{x\to\infty} x\left(100\frac{\ln x}{x} - \ln(1.01)\right) \\ &= \left(\lim_{x\to\infty} x\right)\cdot\left(\lim_{x\to\infty} 100\frac{\ln x}{x} - \ln(1.01)\right) = -\infty,\end{aligned}$$

What we are using here is that for two function $h_1, h_2 : \mathbb{R}^+ \to \mathbb{R}^+$ we have that

$$\lim_{x\to\infty} h_1(x)\cdot h_2(x) = \left(\lim_{x\to\infty} h_1(x)\right)\cdot\left(\lim_{x\to\infty} h_2(x)\right)$$

as long as the two limits $\lim_{x\to\infty} h_1(x)$ and $\lim_{x\to\infty} h_2(x)$ exist.

Therefore, $\lim_{x\to\infty} \frac{x^{100}}{(1.01)^x} = \lim_{x\to\infty} e^{100 \ln x - \ln(1.01)x} = 0$. Hence, by Definition 1, we get that $f(n) := 1.01^n$ grows asymptotically faster than $g(n) := n^{100}$.

(e) $f(n) := \log_2 n$ grows asymptotically faster than $g(n) := \log_2 \log_2 n$.

Solution:

To remove one of the log-factors, we want to substitute $y = \log_2 x$. This works since $x \to \infty$ if and only if $y \to \infty$.

Define $y := \log_2 x$. Then $y \to \infty$ as $x \to \infty$, and therefore

$$\lim_{x\to\infty}\frac{g(x)}{f(x)} \lim_{x\to\infty}\frac{\log_2 \log_2 x}{\log_2 x} = \lim_{y\to\infty}\frac{\log_2 y}{y}.$$

Remembering that $\log_2 y = \ln y/\ln 2$, we can apply Theorem 1 to compute

$$\lim_{y\to\infty}\frac{\log_2 y}{y} = \frac{1}{\ln 2}\lim_{y\to\infty}\frac{\ln y}{y} \overset{\text{Thm.1}}{=} \frac{1}{\ln 2}\lim_{y\to\infty}\frac{(\ln y)'}{y'} = \frac{1}{\ln 2}\lim_{y\to\infty}\frac{1/y}{1} = 0.$$

Hence, by Definition 1, $f(n) := \log_2 n$ grows asymptotically faster than $g(n) := \log_2 \log_2 n$.

(f) $f(n) := 2^{\sqrt{\log_2 n}}$ grows asymptotically faster than $g(n) := \log_2^{100} n$.

Solution:

Using rules about the logarithm, we can compute

$$\lim_{n\to\infty}\frac{\log_2^{100} n}{2^{\sqrt{\log_2 n}}} = \lim_{n\to\infty}\frac{2^{\log_2(\log_2^{100} n)}}{2^{\sqrt{\log_2 n}}} = \lim_{n\to\infty}\frac{2^{100\log_2\log_2 n}}{2^{\sqrt{\log_2 n}}} = \lim_{n\to\infty} 2^{100\log_2\log_2 n - \sqrt{\log_2 n}}.$$

Notice that

$$\lim_{n\to\infty}\left(100\log_2\log_2 n-\sqrt{\log_2 n}\right)=\lim_{n\to\infty}\left(-\sqrt{\log_2 n}\left(1-100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}\right)\right)$$
$$=-\left(\lim_{n\to\infty}\sqrt{\log_2 n}\right)\cdot\left(\lim_{n\to\infty}\left(1-100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}\right)\right)$$
$$=-\infty.$$

Here, we used again that $\lim_{n\to\infty} h_1(n)\cdot h_2(n)=(\lim_{n\to\infty} h_1(n))\cdot(\lim_{n\to\infty} h_2(n))$ as long the latter two limits exist. Furthermore, we need that

$$\lim_{n\to\infty} 100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}=0.$$

As before, it is sufficient to show that $\lim_{x\to\infty} 100\frac{\log_2\log_2 x}{\sqrt{\log_2 x}}$ for $x\in\mathbb{R}^+$. This can be shown as follows. We substitute $y=\log_2(x)$ and, similarly to part (e), we get that

$$\lim_{x\to\infty} 100\frac{\log_2\log_2 x}{\sqrt{\log_2 x}}=100\cdot\lim_{y\to\infty}\frac{\log_2 y}{\sqrt{y}}\overset{\text{Thm.1}}{=}100\cdot\lim_{y\to\infty}\frac{1/y}{1/(2\sqrt{y})}=200\cdot\lim_{y\to\infty}\frac{1}{\sqrt{y}}=0.$$

Thus, we have that $\lim_{n\to\infty} 100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}=0$ and $\lim_{n\to\infty}\left(1-100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}\right)=1$. This allows us to conclude that indeed

$$\lim_{n\to\infty}\left(100\log_2\log_2 n-\sqrt{\log_2 n}\right)=-\infty.$$

Hence,

$$\lim_{n\to\infty}\frac{\log_2^{100} n}{2^{\sqrt{\log_2 n}}}=\lim_{n\to\infty} 2^{100\log_2\log_2 n-\sqrt{\log_2 n}}=0.$$

Therefore, by Definition 1, $f(n):=2^{\sqrt{\log_2 n}}$ grows asymptotically faster than $g(n):=\log_2^{100} n$.

(g) $f(n):=n^{0.01}$ grows asymptotically faster than $g(n):=2^{\sqrt{\log_2 n}}$.

Solution:

We can compute

$$\lim_{n\to\infty}\frac{2^{\sqrt{\log_2 n}}}{n^{0.01}}=\lim_{n\to\infty}\frac{2^{\sqrt{\log_2 n}}}{2^{\log(n^{0.01})}}=\lim_{n\to\infty}\frac{2^{\sqrt{\log_2 n}}}{2^{0.01\log_2 n}}=\lim_{n\to\infty} 2^{\sqrt{\log_2 n}-0.01\log_2 n}$$

We get that

$$\lim_{n\to\infty}\left(\sqrt{\log_2 n}-0.01\log_2 n\right)=\lim_{n\to\infty}\left(-0.01\log_2 n\left(1-\frac{\sqrt{\log_2 n}}{0.01\log_2 n}\right)\right)=-\infty.$$

As for the previous exercise, to get this it is sufficient to show that $\lim_{n\to\infty}\frac{\sqrt{\log_2 n}}{0.01\log_2 n}=0$. This is true since $\frac{\sqrt{\log_2 n}}{0.01\log_2 n}=\frac{1}{0.01\sqrt{\log_2 n}}$.

Hence,

$$\lim_{n\to\infty}\frac{2^{\sqrt{\log_2 n}}}{n^{0.01}}=\lim_{n\to\infty} 2^{\sqrt{\log_2 n}-0.01\log_2 n}=0.$$

Therefore, by Definition 1, $f(n):=n^{0.01}$ grows asymptotically faster than $g(n):=2^{\sqrt{\log_2 n}}$.

Exercise 0.4 *Simplifying expressions.*

Simplify the following expressions as much as possible without changing their asymptotic growth rates.

Concretely, for each expression $f(n)$ in the following list, find an expression $g(n)$ that is as simple as possible and that satisfies $\lim_{n\to\infty} \frac{f(n)}{g(n)} \in \mathbb{R}^+$.

(a) $f(n) := 5n^3 + 40n^2 + 100$

Solution:

The dominating term for $n \to \infty$ in the above expression is $5n^3$ (this grows the fastest). Thus, a first guess would be $5n^3$. However, we can simplify this even more by dropping the constant 5. Thus, we want to prove that n^3 has the same asymptotic growth rate as $f(n)$.

Let $g(n) := n^3$. Then we indeed have that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \left(5 + \frac{40}{n} + \frac{100}{n^3}\right) = 5 \in \mathbb{R}^+.$$

(b) $f(n) := 5n + \ln n + 2n^3 + \frac{1}{n}$

Solution:

Let $g(n) := n^3$. Then we indeed have that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \left(\frac{5}{n^2} + \frac{\ln n}{n^3} + 2 + \frac{1}{n^4}\right) = 2 \in \mathbb{R}^+.$$

For the part $\frac{\ln n}{n^3}$, we can use Theorem 1 to show

$$\lim_{n\to\infty} \frac{\ln n}{n^3} \overset{\text{Thm.1}}{=} \lim_{n\to\infty} \frac{1/n}{3n^2} = 0.$$

(c) $f(n) := n \ln n - 2n + 3n^2$

Solution:

Let $g(n) := n^2$. Then we indeed have that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \left(\frac{\ln n}{n} - \frac{2}{n} + 3\right) = 3 \in \mathbb{R}^+.$$

For the part $\frac{\ln n}{n}$, we can again use Theorem 1 to show

$$\lim_{n\to\infty} \frac{\ln n}{n} \overset{\text{Thm.1}}{=} \lim_{n\to\infty} \frac{1/n}{1} = 0.$$

(d) $f(n) := 23n + 4n \log_5 n^6 + 78\sqrt{n} - 9$

Solution:

By the properties of logarithms, we have that $4n \log_5 n^6 = 24n \log_5 n = \frac{24n \ln n}{\ln 5}$. After removing the constant, we let $g(n) := n \ln n$. Then we have

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \left(\frac{23}{\ln n} + \frac{24}{\ln 5} + \frac{78}{\sqrt{n}\ln n} - \frac{9}{n \ln n}\right) = \frac{24}{\ln 5} \in \mathbb{R}^+.$$

(e) $f(n) := \log_2 \sqrt{n^5} + \sqrt{\log_2 n^5}$

Solution:

By the properties of logarithms, we have that

$$\log_2 \sqrt{n^5} = \frac{5}{2\ln 2} \ln n,$$

and

$$\sqrt{\log_2 n^5} = \sqrt{\frac{5}{\ln 2}} \cdot \sqrt{\ln n}.$$

Let $g(n) := \ln n$. Then we indeed have that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \left(\frac{5}{2\ln 2} + \sqrt{\frac{5}{\ln 2}} \cdot \frac{1}{\sqrt{\ln n}} \right) = \frac{5}{2\ln 2} \in \mathbb{R}^+.$$

(f)* $f(n) := 2n^3 + (\sqrt[4]{n})^{\log_5 \log_6 n} + (\sqrt[7]{n})^{\log_8 \log_9 n}$

Solution:

The terms $(\sqrt[4]{n})^{\log_5 \log_6 n}$ and $(\sqrt[7]{n})^{\log_8 \log_9 n}$ are exponential in n, whereas $2n^3$ is not. So, one of these terms is the dominating one. To figure out which one, we compute $\lim_{n\to\infty} \frac{(\sqrt[7]{n})^{\log_8 \log_9 n}}{(\sqrt[4]{n})^{\log_5 \log_6 n}}$.

We have that

$$\lim_{n\to\infty} \frac{(\sqrt[7]{n})^{\log_8 \log_9 n}}{(\sqrt[4]{n})^{\log_5 \log_6 n}} = \lim_{n\to\infty} \frac{n^{\frac{1}{7}\log_8 \log_9 n}}{n^{\frac{1}{4}\log_5 \log_6 n}} = \lim_{n\to\infty} n^{\frac{1}{7}\log_8 \log_9 n - \frac{1}{4}\log_5 \log_6 n}.$$

Notice that

$$\lim_{n\to\infty} \left(\frac{1}{7}\log_8 \log_9 n - \frac{1}{4}\log_5 \log_6 n \right) = -\infty,$$

since $\log_a x \leq \log_b y$ if $x \leq y$ and $a \geq b$.

Hence,

$$\lim_{n\to\infty} \frac{(\sqrt[7]{n})^{\log_8 \log_9 n}}{(\sqrt[4]{n})^{\log_5 \log_6 n}} = \lim_{n\to\infty} n^{\frac{1}{7}\log_8 \log_9 n - \frac{1}{4}\log_5 \log_6 n} = 0.$$

This shows that the term $(\sqrt[4]{n})^{\log_5 \log_6 n}$ dominates the term $(\sqrt[7]{n})^{\log_8 \log_9 n}$. We now formally show that $n^{\frac{1}{4}\log_5 \log_6 n}$ also dominates the term $2n^3$.

Moreover, we also have

$$\lim_{n\to\infty} \frac{2n^3}{(\sqrt[4]{n})^{\log_5 \log_6 n}} = 2 \lim_{n\to\infty} n^{3-\frac{1}{4}\log_5 \log_6 n} = 0.$$

Let $g(n) := n^{\frac{1}{4}\log_5 \log_6 n}$. Then we indeed have that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1 \in \mathbb{R}^+.$$

Notice that we cannot remove the constant in the exponent since this would change the asympotic behaviour.

Exercise 0.5 * *Finding the range of your bow.*

To celebrate your start at ETH, your parents gifted you a bow and (an infinite number of) arrows. You would like to determine the range of your bow, in other words how far you can shoot arrows with it. For simplicity we assume that all your arrow shots will cover exactly the same distance r, and we define r as the range of your bow. You also know that this range is at least $r \geq 1$ (meter).

You have at your disposition a ruler and a wall. You cannot directly measure the distance covered by an arrow shot (because the arrow slides some more distance on the ground after reaching distance r), so the only way you can get information about the range r is as follows. You can stand at a distance ℓ (of your choice) from the wall and shoot an arrow: if the arrow reaches the wall, you know that $\ell \leq r$, and otherwise you deduce that $\ell > r$. By performing such an experiment with various choices of the distance ℓ, you will be able to determine r with more and more accuracy. Your goal is to do so with as few arrow shots as possible.

(a) What is a fast strategy to find an upper bound on the range r? In other words, how can you find a distance $D \geq 1$ such that $r < D$, using few arrow shots? The required number of shots might depend on the actual range r, so we will denote it by $f(r)$. Good solutions should have $f(r) \leq 10 \log_2 r$ for large values of r.

Solution:

One possible fast strategy is to first shoot an arrow at distance 2 from the wall, and as long as the arrow reaches the wall, you double your distance to the wall for the next shot. More formally, let ℓ_i denote your distance to the wall for the i-th shot. Then this startegy uses distances given by $\ell_i = 2^i$, and does this until you find a distance ℓ_t for which your arrow does not reach the wall. D is then given by $D = \ell_t = 2^t$, and the required number of shots is $f(r) = t$, the smallest integer t such that $r < 2^t$.

This strategy therefore needs $f(r) = \lceil \log_2 r \rceil$ shots, and indeed

$$f(r) = \lceil \log_2 r \rceil \leq 1 + \log_2 r \leq 10 \log_2 r$$

for all $r \geq 2^{1/9}$.

(b) You are now interested in determining r up to some additive error. More precisely, you should find an estimate $\tilde{r}$ such that the range is contained in the interval $[\tilde{r} - 1, \tilde{r} + 1]$, i.e. $\tilde{r} - 1 \leq r \leq \tilde{r} + 1$. Denoting by $g(r)$ the number of shots required by your strategy, your goal is to find a strategy with $g(r) \leq 10 \log_2 r$ for all r sufficiently large.

Solution:

You start by performing the strategy described in part (a). Note that this allows you to find a distance D such that $r \in [\frac{1}{2}D, D]$ using $f(r) = \lceil \log_2 r \rceil$ shots. You will then iteratively find smaller and smaller intervals $[a, b] \subseteq [\frac{1}{2}D, D]$ with $r \in [a, b]$, until you get an interval whose length is at most 2 (and then you can take $\tilde{r}$ to be the center of this interval).

You start by shooting an arrow from distance $(\frac{1}{2}D + D)/2 = \frac{3}{4}D$. If the arrow reaches the wall, then you know that $r \in [\frac{3}{4}D, D]$, and otherwise you deduce that $r \in [\frac{1}{2}D, \frac{3}{4}D]$. Note that in both cases, the length of the interval of possible ranges r was divided by 2. In the next step, if you know that $r \in [\frac{3}{4}D, D]$ then you shoot an arrow from distance $(\frac{3}{4}D + D)/2$, and if you know that $r \in [\frac{1}{2}D, \frac{3}{4}D]$ then you shoot an arrow from distance $(\frac{1}{2}D + \frac{3}{4}D)/2$, which allows you to again divide the length of the interval of possible ranges by 2. You carry on this procedure until you find an interval $[a, b]$ of length $b - a \leq 2$ satisfying $r \in [a, b]$, and you define $\tilde{r} = (a + b)/2$.

By construction, this strategy finds an estimate $\tilde{r}$ such that $\tilde{r} - 1 \leq r \leq \tilde{r} + 1$. Let's compute the number of required shots $g(r)$. You start with $f(r) = \lceil \log_2 r \rceil$ shots in order to perform the strategy described in (a), and then you need t' additional shots to find the interval $[a, b]$. Note that you start with the interval of possible ranges $[\frac{1}{2}D, D]$ which has length $D/2$, and with each additional shot you divide this length by 2, until you reach a length smaller than 2. Therefore, t' is the smallest integer such that $D/2^{t'+1} \leq 2$, i.e. $D \leq 2^{t'+2}$. This means that $t' = \max\{\lceil \log_2 D \rceil - 2, 0\}$ (the maximum with 0 is taken because you cannot have a negative number of shots). This is at most $\lceil \log_2 2r \rceil = 1 + \lceil \log_2 r \rceil$ because $D \leq 2r$, so the total number of required shots is

$$g(r) = f(r) + t' \leq f(r) + \lceil \log_2 r \rceil + 1 = 2\lceil \log_2 r \rceil + 1 \leq 2 \log_2 r + 3,$$

which is smaller than $10 \log_2 r$ for all $r \geq 2^{3/8}$.

(c) Coming back to part (a), is it possible to have a significantly faster strategy (for example with $f(r) \leq 10 \log_2 \log_2 r$ for large values of r)?

Solution:

Let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be any strictly increasing function with $\lim_{r\to\infty} h(r) = \infty$. We will show that there exists a strategy that finds some $D > r$ using $f(r) := \lceil h(r) \rceil$ shots. Thus, this will show in particular that it is is possible to get $f(r) \leq 10 \log_2 \log_2 r$ for large values of r.

Since $h : \mathbb{R}^+ \to \mathbb{R}^+$ is strictly increasing, it is bijective and therefore has an inverse $h^{-1} : \mathbb{R}^+ \to \mathbb{R}^+$ which is also strictly increasing. Moreover, we have $\lim_{r\to\infty} h^{-1}(r) = \infty$ because $\lim_{r\to\infty} h(r) = \infty$. The strategy is then to shoot the arrow at the i-th step with a distance of $h^{-1}(i)$ from the wall, until we get to a step t'' where the arrow doesn't reach the wall (i.e. $h^{-1}(t'') > r$). The number of required shots is then t'', which is the smallest integer satisfying $h^{-1}(t'') > r$, or equivalently $t'' > h(r)$. Therefore, $t'' = \lceil h(r) \rceil$ as claimed.

For the particular example of $f(r) \leq 10 \log_2 \log_2 r$, take the function $h(r) = \log_2 \log_2 r$. This corresponds to shooting an arrow from distance $h^{-1}(i) = 2^{2^i}$ in the i-th step. Then the number of required shots is

$$f(r) = \lceil \log_2 \log_2 r \rceil \leq 1 + \log_2 \log_2 r,$$

which is smaller than $10 \log_2 \log_2 r$ for all $r \geq 2^{2^{1/9}}$.