Eidgenössische
Technische Hochschule
Zürich

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

Departement of Computer Science
Markus Püschel, David Steurer
Johannes Lengler, Gleb Novikov, Chris Wendler

23 September 2019

Algorithms & Data Structures **Exercise sheet 0** **HS 19**

The solutions for this sheet do not have to be submitted. The sheet will be solved in the first exercise session on 23.09.2019.

Exercises that are marked by * are challenge exercises.

Shortest Paths

Exercise 0.1 *Find the Shortest Path.*

Consider a hypothetical floor plan such that the area is organized in hexagonal cells as shown in Figure 1. We begin at the cell marked *start*, and we want to reach the cell marked *end*. We can travel from one cell to another, if the two cells are neighbouring, i.e. they share an edge. Each time we cross from one cell to a destination cell, we consider the move as a single step and the destination cell as visited. We want to find the shortest path from the start to the end, minimizing the number of steps.

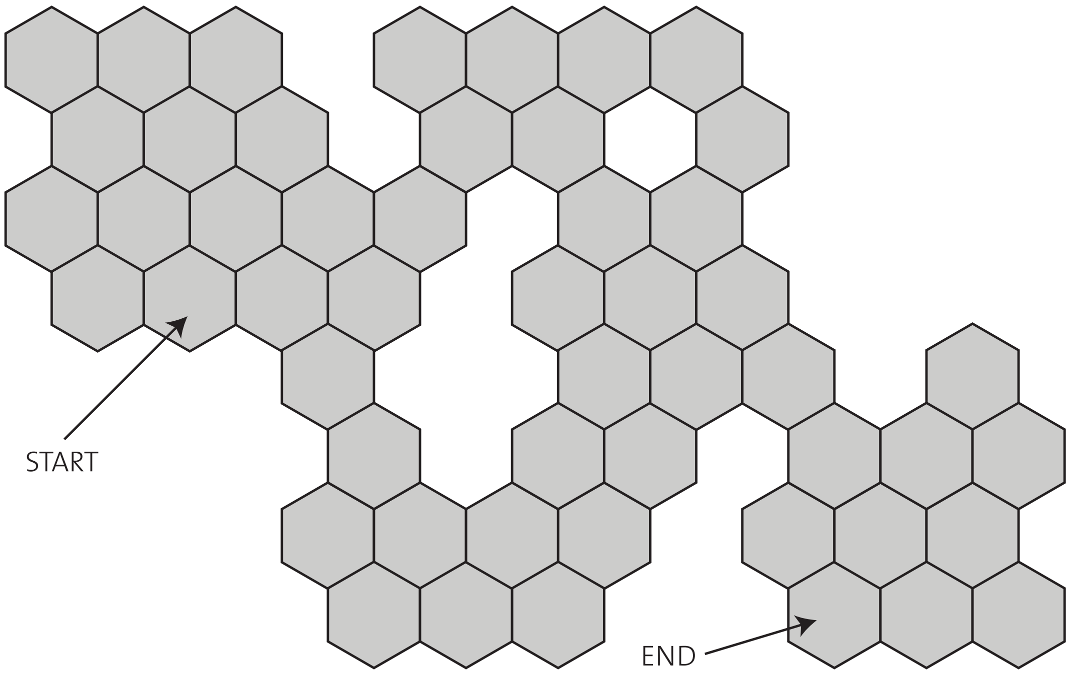

Figure 1: Floor plan

Consider the following algorithm:

1. We visit the start cell and write the number 0 on it.
2. We look only at the cells that have been visited, but have unvisited neighboring cells (if such cells

do not exist we stop this procedure). Among such cells we choose some cell with the smallest number n. Then we visit all its unvisited neighboring cells and write the number $n+1$ on them.

Your tasks:

a) Execute the algorithm on the floor plan given in Figure 1, writing numbers on each cell.

Solution:

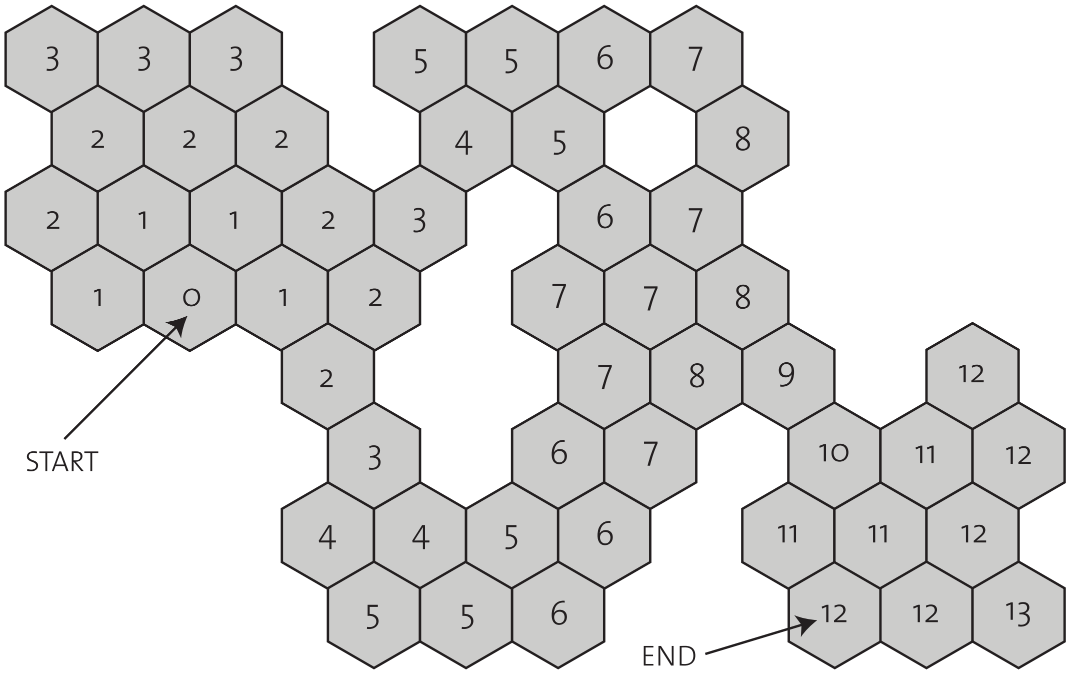

Figure 2: The state of the floor plan, once the algorithm is done with execution

b) How long is the shortest path from the start cell to the end cell?

Solution: Once the algorithm stops executing, each cell will contain a number that corresponds to the shortest distance from the start cell. As a result, we will reach the end cell in 12 steps.

The next tasks are about the execution of the algorithm on an arbitrary floor plan.

c)* Prove that if a cell C has number n, then for any $0 \leq k < n$ we visit all cells with number k before C.

Solution: Let's prove it by contradiction. Consider the smallest n such that there exists a cell B with number $k < n$ that is visited after C. We visit C while visiting neighbours of some cell C' with number $n-1$ and we visit B while visiting neighbours of some B' with number $k-1$. Since we took the smallest "bad" n, we visit B' before C'. Since $k-1 < n-1$, we cannot choose C' before B' to visit its unvisited neighbours, hence C cannot be visited before B.

d)* Prove that if a cell C has number n, then all its neighboring cells have numbers at most $n+1$.

Solution: Let's prove it by contradiction. If some neighboring cell B of C has number $k > n+1$, then we visit B while visiting neighbours of some cell A with number $k-1 > n$. It is impossible since C has smaller number than A (and by point c) we visit C before A).

e)* Prove by mathematical induction that a cell C has number n if and only if the length of the shortest path from the start cell to C is n.

Solution: Let's show that if C has number n, then there exists a path from the start cell to C of lenght n. Base case is $n = 0$ and is trivial. Assume that the statement is true for $n = m$. Let's prove it for $n = m + 1$. We visit C while visiting neighbours of some cell B with number m, and C is a neighbour of B. By induction hypothesis, there is a path from the start cell to B of length m, so there is a path from the start cell to C of length $m + 1$.

Now let's show that if the shortest path from the start cell to C is n, then C has some number $l \leq n$. Base case is $n = 0$ and is trivial. Assume that the statement is true for $n = m$. Let's prove it for $n = m+1$. Consider some shortest path from the start cell to C of length $m+1$ and the last cell A on this path that is before C. The shortest path from the start cell to A has length m (otherwise there is a path from the start cell to C that is shorter than $m + 1$). By induction hypothesis A has number m. Since C is a neighbour of A, it has number $l \leq m + 1$.

Asymptotic Notation

When we estimate the number of elementary operations executed by algorithms, it is often useful to ignore constant factors and instead use the following kind of asymptotic notation, also called $\mathcal{O}$-Notation. 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 ($\mathcal{O}$-Notation). Let $f : \mathbb{N} \to \mathbb{R}^+$. $\mathcal{O}(f)$ is a set of all functions $g : \mathbb{N} \to \mathbb{R}^+$ such that there exist $C > 0$ and $T > 0$ (both may depend on g) such that for all $n > T$, $g(n) \leq Cf(n)$.

If we replace $\mathbb{N}$ by $\mathbb{R}^+$ (and n by $x \in \mathbb{R}^+$) everywhere in this definition, we will get a definition of $\mathcal{O}(f)$ for functions that are defined on $\mathbb{R}^+$.

Instead of working with this definition directly, it is often easier to use limits in the way provided by the following theorems.

Theorem 1 (Theorem 1.1 from the script). *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ and $g : \mathbb{R}^+ \to \mathbb{R}^+$.*

- *If $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$, then $f \in \mathcal{O}(g)$ and $g \notin \mathcal{O}(f)$.*
- *If $\lim_{x\to\infty} \frac{f(x)}{g(x)} = C \in \mathbb{R}^+$, then $f \in \mathcal{O}(g)$ and $g \in \mathcal{O}(f)$.*
- *If $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty$, then $f \notin \mathcal{O}(g)$ and $g \in \mathcal{O}(f)$.*

The theorem holds all the same if the functions are defined on $\mathbb{N}$ instead of $\mathbb{R}^+$. In general, $\lim_{n\to\infty} \frac{f(n)}{g(n)}$ is the same as $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ if the second limit exists.

Theorem 2 (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*

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

Remark. Some functions that we will consider are not defined at small natural numbers (for example, $\ln(\ln n)$ is not defined at $n = 1$). So in all exercises about asymptotic notation we assume that n is large enough, say, $n \geq 10$.

Exercise 0.2 *Comparison of functions.*

Show that

a) $2n \in \mathcal{O}(3n)$ and $3n \in \mathcal{O}(2n)$.

Solution:

$$\lim_{n\to\infty} \frac{2n}{3n} = \frac{2}{3} \in \mathbb{R}^+ \,,$$

hence by Theorem 1, $2n \in \mathcal{O}(3n)$ and $3n \in \mathcal{O}(2n)$.

b) $n \in \mathcal{O}(n \log n)$, but $n \log n \notin \mathcal{O}(n)$.

Solution:

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

hence by Theorem 1, $n \in \mathcal{O}(n \log n)$ and $n \log n \notin \mathcal{O}(n)$.

c) $10n^2 + 100n + 1000 \in \mathcal{O}(n^3)$, but $n^3 \notin \mathcal{O}(10n^2 + 100n + 1000)$.

Solution:

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

hence by Theorem 1, $10n^2 + 100n + 1000 \in \mathcal{O}(n^3)$ and $n^3 \notin \mathcal{O}(10n^2 + 100n + 1000)$.

d) $2^n \in \mathcal{O}(3^n)$, but $3^n \notin \mathcal{O}(2^n)$.

Solution:

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

hence by Theorem 1, $2^n \in \mathcal{O}(3^n)$ and $3^n \notin \mathcal{O}(2^n)$.

e) $n \ln n \in \mathcal{O}(n^{1.01})$, but $n^{1.01} \notin \mathcal{O}(n \ln n)$.

Solution: We apply Theorem 2 to compute the limit of $\frac{\ln x}{x^{0.01}}$ for $x \to \infty$:

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

Hence by Theorem 1, $n \ln n \in \mathcal{O}(n^{1.01})$, but $n^{1.01} \notin \mathcal{O}(n \ln n)$.

f) $n \in \mathcal{O}(e^n)$, but $e^n \notin \mathcal{O}(n)$.

Solution: We apply Theorem 2 to compute the limit of $\frac{x}{e^x}$ for $x \to \infty$:

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

Hence by Theorem 1, $n \in \mathcal{O}(e^n)$, but $e^n \notin \mathcal{O}(n)$.

g) $n^2 \in \mathcal{O}(e^n)$, but $e^n \notin \mathcal{O}(n^2)$.

Solution: We apply Theorem 2 to compute the limit of $\frac{x^2}{e^x}$ for $x \to \infty$:

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

Hence by Theorem 1, $n^2 \in \mathcal{O}(e^n)$, but $e^n \notin \mathcal{O}(n^2)$.

h)* $n^{100} \in \mathcal{O}(1.01^n)$, but $1.01^n \notin \mathcal{O}(n^{100})$.

Solution: We successively apply Theorem 2 to compute the limit of $\frac{x^{100}}{(1.01)^x}$ for $x \to \infty$:

$$\lim_{x\to\infty} \frac{(x^{100})'}{((1.01)^x)'} = \lim_{x\to\infty} \frac{(x^{100})'}{(e^{x\ln 1.01})'} = \lim_{x\to\infty} \frac{100x^{99}}{(\ln 1.01 e^{x\ln 1.01})'} = \cdots = \lim_{x\to\infty} \frac{100!}{(\ln 1.01)^{100} e^{x\ln 1.01}} = 0\,.$$

Hence by Theorem 1, $n^{100} \in \mathcal{O}(1.01^n)$, but $1.01^n \notin \mathcal{O}(n^{100})$.

i)* $\log_2^{100} n \in \mathcal{O}(2^{\sqrt{\log_2 n}})$, but $2^{\sqrt{\log_2 n}} \notin \mathcal{O}(\log_2^{100} n)$.

Solution:

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

Notice that

$$\lim_{n\to\infty} \Big(100\log_2\log_2 n - \sqrt{\log_2 n}\Big) = \lim_{n\to\infty} \Big(-\sqrt{\log_2 n}\big(1 - 100\frac{\log_2\log_2 n}{\sqrt{\log_2 n}}\big)\Big) = -\infty\,.$$

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 Theorem 1, $\log_2^{100} n \in \mathcal{O}(2^{\sqrt{\log_2 n}})$, but $2^{\sqrt{\log_2 n}} \notin \mathcal{O}(\log_2^{100} n)$.

j)* $2^{\sqrt{\log_2 n}} \in \mathcal{O}(n^{0.01})$, but $n^{0.01} \notin \mathcal{O}(2^{\sqrt{\log_2 n}})$.

Solution:

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

Notice that

$$\lim_{n\to\infty} \Big(\sqrt{\log_2 n} - 0.01\log_2 n\Big) = \lim_{n\to\infty} \Big(-0.01\log_2 n\big(1 - \frac{\sqrt{\log_2 n}}{0.01\log_2 n}\big)\Big) = -\infty\,.$$

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 Theorem 1, $2^{\sqrt{\log_2 n}} \in \mathcal{O}(n^{0.01})$ and $n^{0.01} \notin \mathcal{O}(2^{\sqrt{\log_2 n}})$.

For the next exercise you may use the following theorem.

Theorem 3. *Let $f, g, h : \mathbb{N} \to \mathbb{R}^+$. If $f \in \mathcal{O}(h)$ and $g \in \mathcal{O}(h)$, then*

1. *for any constant $c \geq 0$, $cf \in \mathcal{O}(h)$.*
2. *$f + g \in \mathcal{O}(h)$.*

Notice that for all real numbers $a, b > 1$, $\log_a n = \log_a b \cdot \log_b n$ (where $\log_a b$ is a positive constant). Hence $\log_a n \in \mathcal{O}(\log_b n)$. So you don't have to write bases of logarithms in asymptotic notation, that is, you can just write $\mathcal{O}(\log n)$.

Exercise 0.3 *Simplifying expressions.*

Write the following in tight asymptotic notation, simplifying them as much as possible. It is guaranteed that all functions in this exercise take values in $\mathbb{R}^+$ (you don't have to prove it).

a) $5n^3 + 40n^2 + 100$

Solution: By Theorem 1, $n^2 \in \mathcal{O}(n^3)$. Similarly, $n \in \mathcal{O}(n^3)$. By point 1 of Theorem 3, $5n^3 \in \mathcal{O}(n^3)$, $40n^2 \in \mathcal{O}(n^3)$, $100 \in \mathcal{O}(n^3)$. Hence by point 2 of Theorem 3,

$$5n^3 + 40n^2 + 100 \in \mathcal{O}(n^3)\,.$$

b) $5n + \ln n + 2n^3 + \frac{1}{n}$

Solution: By Theorem 1, $n \in \mathcal{O}(n^3)$, $\ln n \in \mathcal{O}(n^3)$, $\frac{1}{n} \in \mathcal{O}(n^3)$. Hence by Theorem 3,

$$5n + \ln n + 2n^3 + \frac{1}{n} \in \mathcal{O}(n^3)\,.$$

c) $n \ln n - 2n + 3n^2$

Solution: By Theorem 1, $n \ln n \in \mathcal{O}(n^2)$. Hence by Theorem 3, $n \ln n + 3n^2 \in \mathcal{O}(n^2)$ and

$$n \ln n - 2n + 3n^2 \in \mathcal{O}(n^2)\,,$$

since $0 < n \ln n - 2n + 3n^2 < n \ln n + 3n^2$.

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

Solution: By the properties of logarithms,

$$4n \log_5 n^6 = 24n \log_5 n \in \mathcal{O}(n \log n)\,.$$

By Theorem 1, $n \in \mathcal{O}(n \ln n)$ and $\sqrt{n} \in \mathcal{O}(n \ln n)$. Hence by Theorem 3,

$$23n + 4n \log_5 n^6 + 78\sqrt{n} \in \mathcal{O}(n \ln n)$$

and

$$23n + 4n \log_5 n^6 + 78\sqrt{n} - 9 \in \mathcal{O}(n \ln n)\,,$$

since

$$0 < 23n + 4n \log_5 n^6 + 78\sqrt{n} - 9 < 23n + 4n \log_5 n^6 + 78\sqrt{n}\,.$$

e) $\log_2 \sqrt{n^5} + \sqrt{\log_2 n^5}$

Solution: By the properties of logarithms,

$$\log_2 \sqrt{n^5} = \frac{2}{5} \log_2 n \in \mathcal{O}(\log n)\,,$$

and

$$\sqrt{\log_2 n^5} = \sqrt{5} \cdot \sqrt{\log_2 n}\,.$$

By Theorem 1, $\sqrt{\log_2 n} \in \mathcal{O}(\ln n)$. Hence by Theorem 3,

$$\log_2 \sqrt{n^5} + \sqrt{\log_2 n^5} \in \mathcal{O}(\log n)\,.$$

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

Solution:

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

Notice that

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

since $\log_a x \le \log_b y$ if $x \le y$ and $a \ge b$. Hence

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

Therefore, by Theorem 1, $\left(\sqrt[7]{n}\right)^{\log_8 \log_9 n} \in \mathcal{O}(n^{\frac{1}{4}\log_5 \log_6 n})$.

By Theorem 1, $2n^3 \in \mathcal{O}(\left(\sqrt[4]{n}\right)^{\log_5 \log_6 n})$. Hence by Theorem 3,

$$2n^3 + \left(\sqrt[4]{n}\right)^{\log_5 \log_6 n} + \left(\sqrt[7]{n}\right)^{\log_8 \log_9 n} \in \mathcal{O}(n^{\frac{1}{4}\log_5 \log_6 n})\,.$$

Exercise 0.4 *Some properties of $\mathcal{O}$-Notation.*

Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ and $g : \mathbb{R}^+ \to \mathbb{R}^+$.

a) Show that if $f \in \mathcal{O}(g)$, then $f^2 \in \mathcal{O}(g^2)$. You can assume that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = C \in \mathbb{R}_0^+$.

Solution:

$$\lim_{x\to\infty} \frac{f^2(x)}{g^2(x)} = \lim_{x\to\infty} \frac{f(x)}{g(x)} \cdot \lim_{x\to\infty} \frac{f(x)}{g(x)} = C^2 \in \mathbb{R}_0^+\,,$$

hence by Theorem 1, $f^2 \in \mathcal{O}(g^2)$.

b) Give an example where $f \in \mathcal{O}(g)$, but $2^f \notin \mathcal{O}(2^g)$.

Solution: Consider $f(n) = 2n$, $g(n) = n$. Obviously, $f \in \mathcal{O}(g)$. However,

$$\lim_{n\to\infty} \frac{2^{f(n)}}{2^{g(n)}} = \lim_{n\to\infty} \frac{2^{2n}}{2^n} = \lim_{n\to\infty} 2^n = \infty\,,$$

hence by Theorem 1, $2^f \notin \mathcal{O}(2^g)$.

Another important example is $f(n) = \log_2 n$ and $g(n) = \log_4 n$. As we already showed, $f \in \mathcal{O}(g)$. However, $2^{f(n)} = n$ and $2^{g(n)} = \sqrt{n}$, so by Theorem 1, $2^f \notin \mathcal{O}(2^g)$.

Induction

The next exercise is about the principle of mathematical induction.

Exercise 0.5 *Induction.*

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

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}\,.$$

- **Base Case.**
  Let $n = 1$. Then:
  $$1 = \frac{1 \cdot (1+1) \cdot (2+1)}{6} = 1\,.$$

- **Induction Hypothesis.**
  Assume that the property holds for some positive integer k. That is:
  $$1^2 + 2^2 + 3^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}\,.$$

- **Inductive Step.**
  We must show that the property holds for $k+1$. Let's add $(k+1)^2$ to both sides of our inductive hypothesis.
  $$\begin{aligned}1^2 + 2^2 + 3^2 + \cdots + k^2 + (k+1)^2 &= \frac{k(k+1)(2k+1)}{6} + (k+1)^2\\ &= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6}\\ &= \frac{(k+1)(2k^2 + k + 6k + 6)}{6}\\ &= \frac{(k+1)(2k^2 + 7k + 6)}{6}\\ &= \frac{(k+1)(k+2)(2k+3)}{6}\\ &= \frac{(k+1)((k+1)+1)(2(k+1)+1)}{6}\,.\end{aligned}$$

  By the principle of mathematical induction, this 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 $k \in \mathbb{N}_0$.

- **Base Case.**
  Let $k = 0$, $n = 2^0 = 1$. Then:
  $$T(1) = 4 \geq 6 \cdot 1^2 - 2 \cdot 1$$

- **Induction Hypothesis.**
  Assume that the property holds for some positive integer $m = 2^k$. That is:

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

- **Inductive Step.** We must show that the property holds for $2m = 2^{k+1}$.

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

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