Appendix A

Notation and Preliminaries

This appendix sums up important notation, definitions, and key lemmas that are not the main focus of the lecture.

A.1 Numbers and Sets

In this lecture, zero is not a natural number: $0 \notin \mathbb{N}$; we just write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ whenever we need it. $\mathbb{Z}$ denotes the integers, $\mathbb{Q}$ the rational numbers, and $\mathbb{R}$ the real numbers. We use $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$ and $\mathbb{R}_0^+ = \{x \in \mathbb{R} \mid x \geq 0\}$.

Rounding down $x \in \mathbb{R}$ is denoted by $\lfloor x \rfloor := \max\{z \in \mathbb{Z} \mid z \leq x\}$ and rounding up by $\lceil x \rceil := \min\{z \in \mathbb{Z} \mid z \geq x\}$.

For $n \in \mathbb{N}_0$, we define $[n] := \{0, \dots, n-1\}$, and for a set M and $k \in \mathbb{N}_0$, $\binom{M}{k} := \{N \subseteq M \mid |N| = k\}$ is the set of all subsets of M that contain exactly k elements.

A.2 Graphs

A finite set of *vertices*, also referred to as *nodes* V together with *edges* $E \subseteq \binom{V}{2}$ defines a *graph* $G = (V, E)$. Unless specified otherwise, G has $n = |V|$ vertices and $m = |E|$ edges and the graph is *simple*: Edges $e = \{v, w\} \subseteq V$ are undirected, there are no *loops*, and there are no *parallel edges*.

If $e = \{v, w\} \in E$, the vertices v and w are *adjacent*, and e is *incident* to v and w, furthermore, $e' \in E$ is *adjacent* to e if $e \cap e' \neq \emptyset$. The *neighborhood* of v is

$$N_v := \{w \in V \mid \{v, w\} \in E\},$$

i.e., the set of vertices adjacent to v. The *degree* of v is

$$\delta_v := |N_v|,$$

the size of v's neighborhood. We denote by

$$\Delta := \max_{v \in V} \delta_v$$

the maximum degree in G.

A v_1-v_d-*path* p is a set of edges $p = \{\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{d-1}, v_d\}\}$ such that $|\{e \in p \mid v \in e\}| \leq 2$ for all $v \in V$. p has $|p|$ *hops*, and we call p a *cycle* if it visits all of its nodes exactly twice. The *diameter* D of the graph is the minimum integer such that for any $v, w \in V$ there is a v-w-path of at most D hops (or $D = \infty$ if no such integer exists). We consider *connected* graphs only, i.e., graphs satisfying $D \neq \infty$.

A.2.1 Trees and Forests

A *forest* is a cycle-free graph, and a *tree* is a connected forest. Trees have $n-1$ edges and a unique path between any pair of vertices. The tree $T = (V, E)$ is *rooted* if it has a designated root node $r \in V$. A leaf is a node of degree 1. A rooted tree has *depth* d if the maximum length of a root-leaf path is d.

A.3 Asymptotic Notation

We require asymptotic notation to reason about the complexity of algorithms. This section is adapted from Chapter 3 of Cormen et al. [CLR90]. Let $f, g\colon \mathbb{N}_0 \to \mathbb{R}$ be functions.

A.3.1 Definitions

$\mathcal{O}(g(n))$ is the set containing all functions f that are bounded from above by $cg(n)$ for some constant $c > 0$ and for all sufficiently large n, i.e. $f(n)$ is *asymptotically bounded from above* by $g(n)$.

$$\mathcal{O}(g(n)) := \{f(n) \mid \exists c \in \mathbb{R}^+, n_0 \in \mathbb{N}_0\colon \quad \forall n \geq n_0\colon \quad 0 \leq f(n) \leq cg(n)\}$$

The counterpart of $\mathcal{O}(g(n))$ is $\Omega(g(n))$, the set of functions *asymptotically bounded from below* by $g(n)$, again up to a positive scalar and for sufficiently large n:

$$\Omega(g(n)) := \{f(n) \mid \exists c \in \mathbb{R}^+, n_0 \in \mathbb{N}_0\colon \quad \forall n \geq n_0\colon \quad 0 \leq cg(n) \leq f(n)\}$$

If $f(n)$ is bounded from below by $c_1 g(n)$ and from above by $c_2 g(n)$ for positive scalars c_1 and c_2 and sufficiently large n, it belongs to the set $\Theta(g(n))$; in this case $g(n)$ is an *asymptotically tight* bound for $f(n)$. It is easy to check that $\Theta(g(n))$ is the intersection of $\mathcal{O}(g(n))$ and $\Omega(g(n))$.

$$\begin{aligned} \Theta(g(n)) := \{f(n) \mid \exists c_1, c_2 \in \mathbb{R}^+, n_0 \in \mathbb{N}_0\colon \quad \forall n \geq n_0\colon \\ 0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n)\} \\ f(n) \in \Theta(g(n)) \quad \Leftrightarrow \quad f \in (\mathcal{O}(g(n)) \cap \Omega(g(n))) \end{aligned}$$

For example, $n \in \mathcal{O}(n^2)$ but $n \notin \Omega(n^2)$ and thus $n \notin \Theta(n^2)$.[1] But $3n^2 - n + 5 \in \mathcal{O}(n^2)$, $3n^2 - n + 5 \in \Omega(n^2)$, and thus $3n^2 - n + 5 \in \Theta(n^2)$ for $c_1 = 1$, $c_2 = 3$, and $n_0 = 4$.

[1] We write $f(n) \in \mathcal{O}(g(n))$ unlike some authors who, by abuse of notation, write $f(n) = \mathcal{O}(g(n))$. $f(n) \in \mathcal{O}(g(n))$ emphasizes that we are dealing with *sets* of functions.

In order to express that an asymptotic bound is not tight, we require $o(g(n))$ and $\omega(g(n))$. $f(n) \in o(g(n))$ means that for any positive constant c, $f(n)$ is strictly smaller than $cg(n)$ for sufficiently large n.

$$o(g(n)) := \{f(n) \mid \forall c \in \mathbb{R}^+ \colon \exists n_0 \in \mathbb{N}_0 \colon \forall n \geq n_0 \colon \quad 0 \leq f(n) < cg(n)\}$$

As an example, consider $\frac{1}{n}$. For arbitrary $c \in \mathbb{R}^+$, $\frac{1}{n} < c$ we have that for all $n \geq \frac{1}{c} + 1$, so $\frac{1}{n} \in o(1)$. A similar concept exists for lower bounds that are not asymptotically tight; $f(n) \in \omega(g(n))$ if for any positive scalar c, $cg(n) < f(n)$ as soon as n is large enough.

$$\omega(g(n)) := \{f(n) \mid \forall c \in \mathbb{R}^+ \colon \exists n_0 \in \mathbb{N}_0 \colon \forall n \geq n_0 \colon \quad 0 \leq cg(n) < f(n)\}$$
$$f(n) \in \omega(g(n)) \quad \Leftrightarrow \quad g(n) \in o(f(n))$$

A.3.2 Properties

We list some useful properties of asymptotic notation, all taken from Chapter 3 of Cormen et al. [CLR90]. The statements in this subsection hold for all $f, g, h \colon \mathbb{N}_0 \to \mathbb{R}$.

Transitivity

$$\begin{aligned}
f(n) \in \mathcal{O}(g(n)) \wedge g(n) \in \mathcal{O}(h(n)) &\quad\Rightarrow\quad f(n) \in \mathcal{O}(h(n)),\\
f(n) \in \Omega(g(n)) \wedge g(n) \in \Omega(h(n)) &\quad\Rightarrow\quad f(n) \in \Omega(h(n)),\\
f(n) \in \Theta(g(n)) \wedge g(n) \in \Theta(h(n)) &\quad\Rightarrow\quad f(n) \in \Theta(h(n)),\\
f(n) \in o(g(n)) \wedge g(n) \in o(h(n)) &\quad\Rightarrow\quad f(n) \in o(h(n)), \text{ and}\\
f(n) \in \omega(g(n)) \wedge g(n) \in \omega(h(n)) &\quad\Rightarrow\quad f(n) \in \omega(h(n)).
\end{aligned}$$

Reflexivity

$$\begin{aligned}
&f(n) \in \mathcal{O}(f(n)),\\
&f(n) \in \Omega(f(n)), \text{ and}\\
&f(n) \in \Theta(f(n)).
\end{aligned}$$

Symmetry

$$f(n) \in \Theta(g(n)) \quad \Leftrightarrow \quad g(n) \in \Theta(f(n)).$$

Transpose Symmetry

$$\begin{aligned}
f(n) \in \mathcal{O}(g(n)) &\quad\Leftrightarrow\quad g(n) \in \Omega(f(n)), \text{ and}\\
f(n) \in o(g(n)) &\quad\Leftrightarrow\quad g(n) \in \omega(f(n)).
\end{aligned}$$

A.4 Bounding the Growth of a Maximum of Differentiable Functions

Lemma A.1. *For $k \in \mathbb{N}$, let $\mathcal{F} = \{f_i \mid i \in [k]\}$, where each $f_i \colon [t_0, t_1] \to \mathbb{R}$ is differentiable, and $[t_0, t_1] \subset \mathbb{R}$. Define $F \colon [t_0, t_1] \to \mathbb{R}$ by $F(t) \coloneqq \max_{i \in [k]} \{f_i(t)\}$. Suppose $\mathcal{F}$ has the property that for every i and t, if $f_i(t) = F(t)$, then $\frac{d}{dt} f_i(t) \leq r$. Then for all $t \in [t_0, t_1]$, we have $F(t) \leq F(t_0) + r(t - t_0)$.*

Proof. We prove the stronger claim that for all a, b satisfying $t_0 \leq a < b \leq t_1$, we have

$$\frac{F(b) - F(a)}{b - a} \leq r. \tag{A.1}$$

To this end, suppose to the contrary that there exist $a_0 < b_0$ satisfying $(F(b_0) - F(a_0))/(b_0 - a_0) \geq r + \varepsilon$ for some $\varepsilon > 0$. We define a sequence of nested intervals $[a_0, b_0] \supset [a_1, b_1] \supset \cdots$ as follows. Given $[a_j, b_j]$, let $c_j = (b_j + a_j)/2$ be the midpoint of a_j and b_j. Observe that

$$\frac{F(b_j) - F(a_j)}{b_j - a_j} = \frac{1}{2} \frac{F(b_j) - F(c_j)}{b_j - c_j} + \frac{1}{2} \frac{F(c_j) - F(a_j)}{c_j - a_j} \geq r + \varepsilon,$$

so that

$$\frac{F(b_j) - F(c_j)}{b_j - c_j} \geq r + \varepsilon \quad \text{or} \quad \frac{F(c_j) - F(a_j)}{c_j - a_j} \geq r + \varepsilon.$$

If the first inequality holds, define $a_{j+1} = c_j$, $b_{j+1} = b_j$, and otherwise define $a_{j+1} = a_j$, $b_j = c_j$. From the construction of the sequence, it is clear that for all j we have

$$\frac{F(b_j) - F(a_j)}{b_j - a_j} \geq r + \varepsilon. \tag{A.2}$$

Observe that the sequences $\{a_j\}_{j=0}^{\infty}$ and $\{b_j\}_{j=0}^{\infty}$ ar both bounded and monotonic, hence convergent. Further, since $b_j - a_j = \frac{1}{2^j}(b_0 - a_0)$, the two sequences share the same limit.

Define

$$c \coloneqq \lim_{j \to \infty} a_j = \lim_{j \to \infty} b_j,$$

and let $f \in \mathcal{F}$ be a function satisfying $f(c) = F(c)$. By the hypothesis of the lemma, we have $f'(c) \leq r$, so that

$$\lim_{h \to 0} \frac{f(c + h) - f(h)}{h} \leq r.$$

Therefore, there exists some $h > 0$ such that for all $t \in [c - h, c + h]$, $t \neq c$, we have

$$\frac{f(t) - f(c)}{t - c} \leq r + \frac{1}{2}\varepsilon.$$

Further, from the definition of c, there exists $N \in \mathbb{N}$ such that for all $j \geq N$, we have $a_j, b_j \in [c - h, c + h]$. In particular this implies that for all sufficiently large j, we have

$$\frac{f(c) - f(a_j)}{c - a_j} \leq r + \frac{1}{2}\varepsilon, \tag{A.3}$$

$$\frac{f(b_j) - f(c)}{b_j - c} \leq r + \frac{1}{2}\varepsilon. \tag{A.4}$$

Since $f(a_j) \leq F(a_j)$ and $f(c) = F(c)$, (A.3) implies that for all $j \geq N$,

$$\frac{F(c) - F(a_j)}{c - a_j} \leq r + \frac{1}{2}\varepsilon.$$

However, this expression combined with with (A.2) implies that for all $j \geq N$

$$\frac{F(b_j) - F(c)}{b_j - c} \geq r + \varepsilon. \tag{A.5}$$

Since $F(c) = f(c)$, the previous expression together with (A.4) implies that for all $j \geq N$ we have $f(b_j) < F(b_j)$.

For each $j \geq N$, let $g_j \in \mathcal{F}$ be a function such that $g_j(b_j) = F(b_j)$. Since $\mathcal{F}$ is finite, there exists some $g \in \mathcal{F}$ such that $g = g_j$ for infinitely many values j. Let $j_0 < j_1 < \cdots$ be the subsequence such that $g = g_{j_k}$ for all $k \in \mathbb{N}$. Then for all j_k, we have $F(b_{j_k}) = g(b_{j_k})$. Further, since F and g are continuous, we have

$$g(c) = \lim_{k\to\infty} g(b_{j_k}) = \lim_{k\to\infty} F(b_{j_k}) = F(c) = f(c).$$

By (A.5), we therefore have that for all k

$$\frac{g(b_{j_k}) - g(c)}{b_{j_k} - c} = \frac{F(b_j) - F(c)}{b_j - c} \geq r + \varepsilon.$$

However, this final expression contradicts the assumption that $g'(c) \leq r$. Therefore, (A.1) holds, as desired. □

Bibliography

[CLR90] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms.* The MIT Press, Cambridge, MA, 1990.