Algorithms and Datastructures (HS2024) Week 1

Rui Zhang

23.09.2024

Contents

1 Revision Theory

1.1 Algorithms

Algorithms are a sequence of elementary instructions, usually used to solve a problem. This week, you have learned three algorithms. Do not worry, you are

not expected to learn them by heart, unlike algorithms in the future that are much more important. Since these algorithms can also be found in the lecture notes, I will not go into too much detail as to how they work. Instead, I will focus on the key concepts you should learn from these examples.

1.1.1 Karatsuba's Algorithm

The standard way of multiplying two two-digit numbers is to calculate four partial products and sum them up in the way you have learned it in elementary school. For two numbers each of n digits, that will be a total of n^2 multiplications.

Karatsuba first found a way to do this with less multiplications for two digits: Assume you are given a number a with the two digits a_1 and a_0 and another number b with the two digits b_1 and b_0. Then, we can calculate the products $a_0 \cdot b_0$, $a_1 \cdot b_1$, $a_0 \cdot b_0$ and $(a_1 - a_0) \cdot (b_1 - b_0)$. Each of these are multiplications of two single digit numbers. So we went from 4 products to 3 products. But what do we do if we have more than 2 digits?

Here comes the concept to takeaway: Divide And Conquer. Consider the case where we have $n = 4$ digits. We may then apply the same idea, but this time, a_0, a_1, b_0, b_1 are two digit numbers. And how do we multiply two digit numbers? By applying Karatsuba's idea from above of course!

Now we know how to do this for $n = 4$. And just like how we computed products of numbers of 4 digits by using our knowledge on computing products of numbers with 2 digits, we can now also compute products of numbers consisting of 8 digits efficiently and so on and so fourth.

Divide And Conquer: When solving problems, always start off with small examples. If you can find a solution for the smaller examples, you may also be able to find a solution for much bigger examples just by splitting up the bigger examples.

1.1.2 Cows.

Old McDonald had a farm. I A I A O.

But he also had a cow. From a secret agent the cow received the information that there is a hole in the fence that surrounds the farm this cow lives on. Now, the cow wants to escape. For simplicity, let us model the fence as the x-axis. The cow starts at position 0 and travels along the fence to try and find the hole. Sadly, this cow is short-sighted and so can only check if the fence has a hole at the position where she stands. Let us assume that the hole is at position $k \neq 1$. How can we do this with as little steps as possible? **Note that the cow has no idea whether to go right or left.**

One solution is to exchange between going left and right and increase the number of steps taken by 1 every time. This way, the cow goes 1 step to the left. If it encounters the hole, it escapes, else it goes back 1 step and goes 2 steps to the right and so on and so forth, until it reaches the hole at position k.

In this case, the cow will need to take

$$2 \cdot 1 + 2 \cdot 2 + 2 \cdot 3 + \cdots + 2 \cdot k + k$$

many steps, in the worst case scenario that the cow decides to take k steps in the wrong direction first, then k steps back to position 0, and then k steps in the correct direction until it reaches the hole. This expression simplifies to:

$$2 \cdot (1 + 2 + 3 + \cdots + k) + k \overset{(1)}{=} 2 \cdot \frac{k(k+1)}{2} + k$$
$$= k(k+1) + k$$
$$= k^2 + 2k$$

Where the equation used in (1) is an exercise in exercise sheet 0 with a detailed solution and walkthrough below under the section "Induction". But as k grows, this expression for the number of steps the cow has to take grows quadratically. Geht es besser?

Maybe counterintuitively, the solution here is to double the number of steps the cow takes every time it goes back to its initial position and goes in the other direction. This way, the cow goes 1 $(= 2^0)$ step to the left, then back, then takes 2 $(= 2^1)$ steps to the right, then back, then takes 4 $(= 2^2)$ steps to the left, then back, and so on and so forth.

In the worst case, the number of steps now becomes:

$$2 \cdot 2^0 + 2 \cdot 2^1 + \cdots + 2 \cdot 2^{i-1} + 2 \cdot 2^i + k$$
$$= 2 \cdot (2^0 + 2^1 + 2^2 + \cdots + 2^{i-1} + 2^i) + k$$

where i is the exponent such that $2^i \geq k$ and $2^{i-1} \leq k$ (I) (so i is the edge-case, where doubling the number of steps $i-1$ was just not enough for the cow to notice the hole in front of her and i is more than enough steps for her to get to the hole.)

In class, you proved via induction that this simplifies to:

$$2 \cdot (2^{i+1} - 1) + k = 2 \cdot (2 \cdot 2 \cdot 2^{i-1} - 1) + k$$
$$\overset{(I)}{\leq} 2 \cdot (2 \cdot 2 \cdot k - 1) + k$$
$$\leq 2 \cdot 2 \cdot 2 \cdot k + k$$
$$= 8 \cdot k + k$$
$$= 9k$$

which is a linear term. So for large values of k, this method is much faster. If you do not believe me, just check it out in GeoGebra!

Exponential Increases and Logarithmic Decreases *of certain sizes sometimes help in computer science to reduce the runtime of algorithms and*

3

make them more efficient. You encounter this idea again when we talk about search algorithms.

*   **The Worst Case Scenario** is usually considered in computer science to make sure that algorithms run quickly even if we are very unlucky. This also means that we are usually more interested for larger inputs. In this case, we were more interested in the increase in efficiency we get from increasing k.*

1.2   Induction

Induction is a mathematical proof pattern, which can be employed to prove the validity of a claim $A(n)$ for every n, s.t. (such that) n can be any integer, or any number from a range of the integers. (for example from $n = 4$ to $n = 10$).

For example, you may perhaps be asked by an certain evil exercise sheet to prove that

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \ \forall n \in \mathbb{N}$$

Here, "$\forall$" means "for all" and "$\in \mathbb{N}$" means "element of the natural numbers" We could think of proving the claim via algebraic manipulations, however, this method does not always generalize well. However, Induction is usually a fool proof method if done right.

Induction always consists of three steps. Please, never forget these three:

1.2.1   First Step: Base Case

We start off by proving that this claim is true for the first number we have to prove it for. In our case it is simple: We have to prove $A(n)$ for $n = 1$. Note however, that this can be any integer. For instance, consider the claim $A'(n)$: let $f(n) = -(n-4)^2 + 3$, then $f(n) > 0$. This claim only holds for $n \in \{3, 4, 5\}$

What I want so say is: even though the base case is the easiest step of them all, still be sure to double check that you started at the right number...
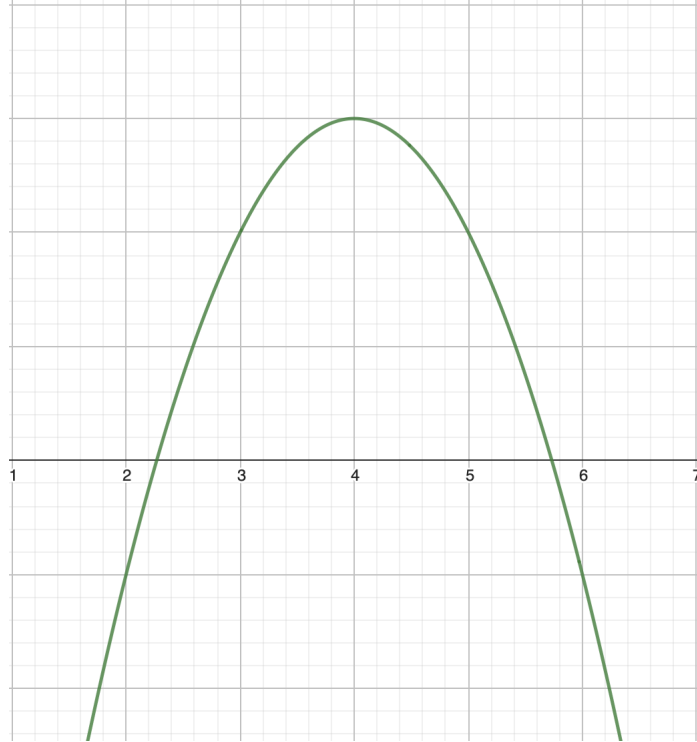
Now, it is just about plugging in the correct number. I recommend computing the left hand side (LHS) and right hand side (RHS) of the claim separately, and then showing that they are equal: For $n = 1$, we have:

$$LHS = \sum_{k=1}^{1} k = 1$$

$$RHS = \frac{1(1+1)}{2} = 1$$

and thus $LHS = RHS$, which means that $A(1)$ (the claim for $n = 1$) has been proven.

Figure 1: This the function f. Trust me.

1.2.2 Second Step: Induction Hypothesis

From here on out, the game plan is to prove that **if** the claim holds for some singular number, then it will automatically also hold for the next number that we want to prove this statement for. Formally, we will have to fix a value m (so take some fixed value m) and prove that if $A(m)$ holds, then $A(m+1)$ also holds, no matter the choice of m, as long as m is in the range of numbers given by the exercise. To do this, we start with the Induction Hypothesis, where we write the following sentence:

"Assume this claim / statement $(A(m))$, holds for **some** integer $m \geq 1$"

This lays the foundation for the next step, where we will be proving that from this assumption, we can deduce that $A(m+1)$ also holds, all according to our game plan. Although it is usually the same sentence everytime, we still have to make sure that the induction hypothesis (the sentence we just wrote) makes sense according to what we want to prove. In the case of the function from Figure 1, we would want to assume instead that the claim holds for some integer $m \in \{3, 4\}$, realizing that we cannot include 5 here as above, because then for $m = 5$, we would be proving that the claim is also true for $m = 6$ in the next step, which cannot be true.

It is important to notice two things here:

First of all, I used the word "some" explicitly here instead of "all" or "every" or "any". This is to assure "formal cleanliness" and correctness. If I were to write "all" or "every" or "any", then I would be assuming what I want to prove.

I also changed the variable names from n to m. This is because we want to prove the statement for all n, which means that the value of n is not fixed! However, m has some fixed, actual value. To avoid confusion, we use another variable that indicates a fixed, constant value, in this case m.

Although some have pointed out that this switch in variable names can become even more confusing, especially if we have multiple variables, I would still rather do the switch than not. Usually in the exam solutions, and in the master solutions of our, exercises, they do so too.

1.2.3   Third Step: Induction Step

The induction step is the most important step of the three - and usually the trickiest. I recommend explicitly writing down what we want to show: In our case it is

$$A(m+1): \sum_{k=1}^{m+1} k = \frac{(m+1)(m+2)}{2} \tag{1}$$

To prove this statement based on the assumption of $A(m)$ being true according to our game plan, I also strongly recommend "going from left to right": First write down the LHS of the statement we want to prove and do algebraic manipulations, as well as employ the induction hypothesis, until we reach the RHS of what we want to prove.

$$\sum_{k=1}^{m+1} k = \sum_{k=1}^{m} k + (m+1)$$
$$\overset{\text{I.H.}}{=} \frac{m(m+1)}{2} + (m+1)$$
$$= \frac{m(m+1)}{2} + \frac{2(m+1)}{2}$$
$$= \frac{(m+2)(m+1)}{2}$$
$$= \frac{(m+1)(m+2)}{2}$$

Note that I.H. stands for induction hypothesis. If you haven't had a step where you needed to use the induction hypothesis, then it is likely that you have done the induction step is wrong or have made a mistake somewhere. Either that or the task was absolutely trivial. But we are at ETH Zurich, so do not push your luck. Furthermore, if you get stuck somewhere along the way, I also recommend "going from right to left" and basically make the RHS and the LHS meet somewhere in the middle. Sometimes it may become hard to know what kind of algebraic manipulations you will have to make.

1.2.4 Why does this work?

Imagine you have a number line in front of you, consisting of the numbers you have to prove a statement for. In the first step, the base case, we proved that whatever statement the task gave us indeed holds for the first number on that number line. In the induction hypothesis and the induction step, we then showed that if the statement holds for one number, it also holds for the next. But does that not precisely mean that the statement holds for all numbers on that number line?

In our example, we had to prove $A(n)$ for all natural numbers $n \geq 1$. So now, in the base case we have shown that $A(n)$ holds for $n = 1$. Via the I.H. and the induction step, we now have that $A(n)$ holds for the next number, $n = 2$. But then via the I.H. and the induction step, we have that $A(n)$ holds for the number after that, $n = 3$ and so on and so forth, until the end of times (and the numbers lived happily ever after).

2 Exercises

2.1 Exercise Sheet 1 - Priority List

This weeks exercises are now out. Here's my **personal** opinion on what is most important. 1 is most important. 4 is least.

1. 1.1), 1.3) Obviously the bonus exercises. Very good induction and asymptotic growth exercises too.

2. 1.2) sheds light onto one important trick for exercises regarding asymptotic growth-rates

3. 1.4a), 1.4b) still important, since more practice in induction is always helpful. b) may be harder, but still worthwhile

4. any challenge (*) exercise. Try to read the end result even if you do not prove it though

3 Supplementary Exercises

This week, there are plenty of exercises already. However, if you feel like you are stuck on any of the exercises from the exercise sheet then try these first instead of nagging on one problem for multiple hours. Most of them revise a few concepts we have seen in exercise sheet 0. This way, you can do some exercises similar to those of exercise sheet 0 without any guidance.

3.1 Induction 0

This is the induction exercise we did in the last lesson. I promised to prepare some notes to make the exercise more clear after speeding through the solution

in class. So here they are. By the way, you can find the exercise online as well under `https://exams.vis.ethz.ch/exams/4xze5i8v.pdf`

To prove:

$$\sum_{k=1}^{n} \frac{1}{k^2} \leq 2 - \frac{1}{n}$$

Proof via Induction:

3.1.1 Base Case: (n=1)

$$LHS = \sum_{k=1}^{1} \frac{1}{k^2} = 1$$

$$RHS = 2 - \frac{1}{1} = 1$$

$$LHS = RHS$$

Not much to say here...

3.1.2 Induction Hypothesis:

Assume that this inequality holds for some fixed positive integer m, namely:

$$\sum_{k=1}^{m} \frac{1}{k^2} \leq 2 - \frac{1}{m}$$

Again, this is still the same sentence every time. Note that I used a variable name that is **not** k, since this is already reserved for the variable which goes from 1 to m in the summation.

3.1.3 Induction Step: $(k \to k+1)$

It is always helpful to keep our goal in mind at all times:

To prove:

$$\sum_{k=1}^{m+1} \frac{1}{k^2} \leq 2 - \frac{1}{m+1}$$

This is just the inequality, where instead of m, we wrote $m+1$.

Now, let us go from left to right again to prove this inequality:

$$\sum_{k=1}^{m+1} \frac{1}{k^2} = \underbrace{\left(\sum_{k=1}^{m} \frac{1}{k^2}\right)}_{\text{I.H: } \leq 2 - \frac{1}{m}} + \frac{1}{(m+1)^2}$$

We split up the sum into two parts. This summation consists of $m + 1$ terms. But we already know something of the first m terms: the induction hypothesis! It now hopefully becomes clear that sometimes, it makes sense to envision the individual terms of a summation in your head, or take some time and write some of the terms down. From here on out:

$$\underbrace{\left(\sum_{k=1}^{m} \frac{1}{k^2}\right)}_{\text{I.H: } \leq 2 - \frac{1}{m}} + \frac{1}{(m+1)^2} \overset{I.H.}{\leq} 2 - \frac{1}{m} + \frac{1}{(m+1)^2}$$

$$= 2 - \frac{(m+1)^2}{m(m+1)^2} + \frac{m}{m(m+1)^2}$$

$$= 2 - \frac{(m+1)^2}{m(m+1)^2} - \frac{-m}{m(m+1)^2}$$

$$= 2 - \frac{(m+1)^2 - m}{m(m+1)^2}$$

$$\overset{(*)}{=} 2 - \frac{m^2 + 2m + 1 - m}{m(m+1)^2}$$

$$= 2 - \frac{m^2 + m + 1}{m(m+1)^2}$$

$$\overset{(**)}{\leq} 2 - \frac{m^2 + m}{m(m+1)^2}$$

$$= 2 - \frac{m(m+1)}{m(m+1)^2}$$

$$= 2 - \frac{1}{m+1}$$

Here, for (*), we applied the binomial formula and for (**), we made use of the fact that the term on the right of $\leq$ is missing a 1 in its numerator. This means that we are subtracting **less** in the term on the right, in turn suggesting that this term is actually larger.

For induction, it helps to have a lot of practice and experience in algebraic manipulations. This way steps like (**) can feel more natural and predictable.

3.2 Induction 1

Explain why induction is a correct proof pattern. Which steps are necessary for this proof pattern?

3.3 Induction 2

Prove via induction that for all $k \in \mathbb{N}$, we have that 6 divides $(7^k - 1)$

Hint: Make sure to write down the steps of induction clearly and state what exactly you want to prove.

Hint: In the induction step: maybe you can an expression up into two expressions that are both divisible by 6

3.4   Asymptotic Growth 1

When does a function grow asymptotically faster than another? Give me the mathematical definition as well as the intuition.

3.5   Asymptotic Growth 2

Prove that $f(n) = 2^{20log(n)}$ grows asymptotically faster than $f(n) = 2^{10log(n)}$

3.6   Asymptotic Growth 3

Sort the following functions based on their asymptotic growth. Give the reasoning.

$$n, 0.01n^2, e^n, log(n), 2^{32}, 2^n, n + \sqrt{n}$$