CSC236 Fall 2018
Assignment #1: induction
sample solution

The aim of this assignment is to give you some practice with various flavours of induction. For each question below you will present a proof by induction. For full marks you need to make it clear to the reader that the base case(s) is/are verified, that the inductive step follows for each element of the domain (typically the natural numbers), where the inductive hypothesis is used, and that it is used in a valid case.

Your assignment must be **typed** to produce a PDF document **a1.pdf** (hand-written submissions are not acceptable). You may work on the assignment in groups of 1 or 2, and submit a single assignment for the entire group on MarkUs

1. Recall bipartite graphs. Consider the following definitions:

   bipartite graph: Undirected graph $G = (V, E)$ is **bipartite** if and only if there exist V_1, V_2 such that $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$, and every edge in E has one endpoint in V_1 and the other in V_2.

   P(n): Every bipartite graph on n vertices has no more than $n^2/4$ edges.

   (a) Assume $P(234)$. Can you use this[1] to prove that $P(235)$ follows? Explain why, or why not.

   proof: Assume $P(234)$ and let G be an arbitrary bipartite graph with 235 vertices. Remove a vertex, together with its edges from the larger of the two partitions to form a new bipartite graph, G' with 234 vertices. Notice that, since all edges must run from the removed vertex to the other (smaller) partition, there are no more than 117 edges.
   By $P(234)$ G' has no more than $234^2/4$ edges, so G has no more than:

   $$\frac{234^2}{4} + 117 = \frac{234^2 + 4(117)}{4} = \frac{234^2 + 2(234)}{4} < \frac{234^2 + 2(234) + 1}{4} = \frac{235^2}{4}$$

   So $P(235)$ follows.

   (b) Assume $P(235)$. Can you use this[2] to prove that $P(236)$ follows? Explain why or why not.

   explanation: $P(236)$ does not follow from $P(235)$. If we use the same approach as in the previous part, except this time the best you can guarantee based on the smaller partition is that you have removed at most 118 edges.

   $$\frac{235^2}{4} + 118 = \frac{235^2 + 4(118)}{4} = \frac{235^2 + 2(236)}{4} = \frac{235^2 + 2(235) + 2}{4} > \frac{236^2}{4}$$

   Note: It is possible to, implicitly, strengthen the claim by reasoning that the bound must be an integer and taking its floor, and thus answer this question with a "yes."

[1] If you say yes, $P(234)$ must be a necessary part of your proof.
[2] If you say yes, $P(235)$ must be a necessary part of your proof.

(c) Use what you've learned from the previous two answers to construct a proof by simple induction that: $\forall n \in \mathbb{N}, P(n)$. **Note:** There are proofs of this claim that are not by simple induction, **but** those proofs will receive no marks. **Hint:** You probably need to strengthen the claim in order to devise a successful inductive hypothesis. If this seems mysterious, revisit the previous two answers...

discussion: The difference between the two cases above suggests that the situation may be different for bipartite graphs with odd, versus those with even, numbers of vertices.

A bit of experimentation with small graphs suggests a new predicate $P'(n)$ which, if true, will imply $P(n)$:

P'(n): Every bipartite graph on n vertices has no more than $n^2/4$ edges if n is even, or $(n^2-1)/4$ edges if n is odd.

base case: An empty bipartite graph has 0 vertices and 0 edges, and $0 \leq 0^2/4$, which verifies $P(0)$.

inductive step: Let n be an arbitrary, fixed, natural number. Assume $P'(n)$, that every bipartite graph on n vertices has no more than $n^2/4$ edges if n is even, or $(n^2-1)/4$ edges if n is even. I will show that $P'(n+1)$ follows, that every bipartite graph on $n+1$ edges has no more than $(n+1)^2/4$ edges if $n+1$ is even, or $[(n+1)^2-1]/4$ edges if $n+1$ is odd.

Let G be an arbitrary bipartite graph on $n+1$ vertices. Remove a vertex, together with its edges, from G's larger partition to produce a new bipartite graph G'. There are two possibilities, depending on whether $n+1$ is even or odd:

case $n+1$ is odd: G's smaller partition has, at most, $n/2$ vertices, so we removed at most $n/2$ edges to produce G'. $n+1$ odd means n is even, so by assumption $P(n)$, G' has at most $n^2/4$ edges, so accounting for the edges removed G had, at most:

$$\frac{n^2}{4} + \frac{n}{2} = \frac{n^2+2n}{4} \leq \frac{(n+1)^2-1}{4}$$

So $P'(n+1)$ follows in this case.

case $n+1$ is even: G's smaller partition has, at most, $(n+1)/2$ vertices, so we removed at most $(n+1)/2$ edges to produce G'. $n+1$ even means n is odd, so by assumption $P(n)$ G' has at most $(n^2-1)/4$ edges, so accounting for the edges removed G had, at most:

$$\frac{n^2-1}{4} + \frac{n+1}{2} = \frac{n^2+2n+1}{4} \leq \frac{(n+1)^2}{4}$$

So $P'(n+1)$ follows in this case.

$P'(n+1)$ follows in both possible cases ■

2. Define function f as follows:

$$f(n) = \begin{cases} 3 & \text{if } n = 0 \\ [f(\lfloor \log_3 n \rfloor)]^2 + f(\lfloor \log_3 n \rfloor) & \text{if } n > 0 \end{cases}$$

Define predicate $P(n)$: "$f(n)$ is a multiple of 4."

(a) Assume $P(3)$. Can you use this[3] to prove $P(29)$? Explain why or why not.

[3] If you say yes, $P(3)$ must be a necessary part of your proof.

proof: Assume $P(3)$. Since $\log_3 29$ is between 3 and 4, by definition $f(29)$ is:

$$f(29) = [f(\lfloor \log_3 29 \rfloor)]^2 + f(\lfloor \log_3 29 \rfloor) = f(3)^2 + f(3) = f(3)\,(f(3)+1)$$

Since this expression has a factor $f(3)$ which is assumed to be a multiple of 4, $f(29)$ is a multiple of 4 $\blacksquare$

(b) Assume $P(4)$. Can you use this[4] to prove $P(29)$? Explain why or why not.

discussion: $f(4)$ appears nowhere in the definition of $f(29)$, so $P(4)$ cannot help establish $P(29)$. **Note:** it is technically possible to argue, quite indirectly, that since $f(4)$ and $f(3)$ agree, so do $P(4)$ and $P(3)$, and then argue from there for $P(29)$.

(c) Use complete induction to prove $\forall n \in \mathbb{N}, n > 0 \Rightarrow P(n)$.

proof by complete induction:

inductive step: Let $n \in \mathbb{N}$. Assume $n > 0$ and that $\bigwedge_{k=1}^{k=n-1} P(k)$. I will show that $P(n)$ follows, that is that $f(n)$ is a multiple of 4.

There are two cases to consider, depending on whether n is less than 3 or not.

case $n < 3$: Since n is either 1 or 2, $0 < \log_3 n < 1$, so

$$f(n) = f(0)^2 + f(0) = 9 + 3 = 12 = 4 \times 3$$

So $f(n)$ is a multiple of 4, which verifies the base cases.

case $n \geq 3$: Since $\log_3 n \geq \log_3 3 = 1$, and $\log_3 n < n$, I may assume $P(\lfloor \log_3 n \rfloor)$. Let $k \in \mathbb{Z}$ be such that $f(\lfloor \log_3 n \rfloor) = 4k$. Then $f(n) = 4(4k^2) + 4k = 4(4k^2 + k)$, a multiple of 4 $\blacksquare$

3. Use the Principle of Well-Ordering to derive a contradiction that proves there are no positive integers x, y, z such that:

$$5x^3 + 50y^3 = 3z^3$$

You may assume, without proof, that if a prime number p divides a perfect cube n^3, then p also divides n.

proof: For the sake of contradiction I will claim the negation of what I want to prove: $\exists x, y, z \in \mathbb{N}^+, 5x^3 + 50y^3 = 3z^3$. Under this assumption the set

$$S = \{x \in \mathbb{N}^+ \mid \exists y, z \in \mathbb{N}^+, 5x^3 + 50y^3 = 3z^3\}$$

is not empty. Since it is a non-empty set of natural numbers, by the Principle of Well-Ordering it has a least element. Let x_1 be the least element of S, and let y_1, z_1 be positive natural numbers such that $5x_1^3 + 50y_1^3 = 3z_1^3$. So

$$\begin{array}{rcl}
5x_1^3 + 50y_1^3 = 3z_1^3 & \Rightarrow & 5 \mid z_1 \qquad \# \, 5 \nmid 3 \Rightarrow 5 \mid z_1^3 \Rightarrow 5 \mid z_1 \\
\text{Let } z_2 = z_1/5 & \Rightarrow & 5x_1^3 + 50y_1^3 = 3 \times 5^3 \times z_2^3 \qquad \# \text{ factor out } 5^3 \\
& \Rightarrow & x_1^3 + 10y_1^3 = 3 \times 5^2 \times z_2^3 \qquad \# \text{ divide both sides by 5} \\
& \Rightarrow & x_1^3 = 75z_2^3 - 10y_1^3 \Rightarrow 5 \mid x_1 \\
\text{Let } x_2 = x_1/5 & \Rightarrow & 5^3x_2^3 = 75z_2^3 - 10y_1^3 \Rightarrow 5^2x_2^3 = 15z_2^3 - 2y_1^3 \qquad \# \text{ divide by 5} \\
& \Rightarrow & 2y_1^3 = 15z_2^3 - 5^2x_2^3 \Rightarrow 5 \mid y_1 \qquad \# 5 \nmid 2 \Rightarrow 5 \mid y_1^3 \Rightarrow 5 \mid y_1 \\
\text{Let } y_2 = y_1/5 & \Rightarrow & 2 \times 5^3 \times y_2^3 = 15z_2^3 - 25x_2^3 \Rightarrow 50y_2^3 = 3z_2^3 - 5x_2^3 \qquad \# \text{ divide by 5} \\
& \Rightarrow & 5x_2^3 + 50y_2^3 = 3z_2^3
\end{array}$$

[4]If you say yes, $P(4)$ must be a necessary part of your proof.

... and $x_2 \in S$! But, using PWO I chose x_1 to be the smallest element of S, and $x_2 = x_1/5$ (both positive numbers) means $x_2 < x_1 \longrightarrow\longleftarrow$ Contradiction!

Since the assumption that $\exists x, y, z \in \mathbb{N}^+, 5x^3 + 50y^3 = 3z^3$ leads to a contradiction it is false, and its negation $\forall x, y, z \in \mathbb{N}^+, 5x^3 + 50y^3 \neq 3z^3$ ■

4. Define $\mathcal{T}$ as the smallest set of strings that satisfies:

- "*" $\in \mathcal{T}$
- if $t_1, t_2 \in \mathcal{T}$ then their parenthesized concatenation $(t_1 t_2) \in \mathcal{T}$.

Some examples: "*", "(**)", "(*(**))" are all in $\mathcal{T}$.

Now read over these four Python functions:

```
def left_count(s: str) -> int:
   """
   Return the number of "(" in s
   """
   return s.count("(")

def double_count(s: str) -> int:
   """
   Return the number of "((" plus number of "))", including possible
   overlaps.
   """
   return (len([s[i:] for i in range(len(s)) if s[i:].startswith("((")])
           + len([s[:i] for i in range(len(s) + 1) if s[:i].endswith("))")]))

def left_surplus(s: str, i: int) -> int:
   """
   Return the number of "(" minus the number of ")"
   in s[:i]
   """
   return s.count("(", 0, i) - s.count(")", 0, i)

def max_left_surplus(s: str) -> int:
    """
    Return the maximum left surplus for all prefixes of s.
    """
    return max([left_surplus(s, i) for i in range(len(s))] + [0])
```

(a) Use structural induction on $\mathcal{T}$ to prove:

$$\forall t \in \mathcal{T}, \text{left\_count}(t) \leq 2^{\text{max\_left\_surplus}(t)} - 1$$

[hint, September 25:] You may assume, without proof, that if $t_1, t_2 \in \mathcal{T}$, then

$$\text{max\_left\_surplus}((t_1 t_2)) = \max(\text{max\_left\_surplus}(t1), \text{max\_left\_surplus}(t_2)) + 1$$

sample solution: Define $P(t)$: $\text{left\_count}(t) \leq 2^{\text{max\_left\_surplus}(t)} - 1$. I will prove by structural induction $\forall t \in \mathcal{T}, P(t)$.

base case: Since "*" has no parentheses, $\text{left\_count}(*) = 0 = \text{max\_left\_surplus}(*)$, so

$$\text{left\_count}(*) = 0 = 1 - 1 = 2^{\text{max\_left\_surplus}(t)} - 1$$

Thus $P(*)$ follows.

inductive step: Let $t_1, t_2 \in \mathcal{T}$. Assume $P(t_1)$ and $P(t_2)$. I will show that $P((t_1 t_2))$ follows.

$$\begin{aligned}
\text{left_count}((t_1t_2)) &= 1 + \text{left_count}(t_1) + \text{left_count}(t_2) \qquad \#\text{ just added 1 left parenthesis} \\
&\leq 1 + 2^{\text{max_left_surplus}(t_1)} - 1 + 2^{\text{max_left_surplus}(t_2)} - 1 \qquad \#\text{ by } P(t_1) \text{ and } P(t_2) \\
&\leq 1 + 2^{\text{max_left_surplus}(t_1)} - 1 + 2^{\text{max_left_surplus}(t_1)} - 1 \\
&\qquad \#\text{ WLOG max_left_surplus}(t_1) \geq \text{max_left_surplus}(t_2),\ \text{otherwise swap them} \\
&= 2 \times 2^{\text{max_left_surplus}(t_1)} - 1 \\
&= 2^{\text{max_left_surplus}(t_1)+1} - 1 = 2^{\text{max_left_surplus}((t_1t_2))} - 1 \\
&\qquad \#\text{ We are allowed to assume the final equality, September 25} \\
&\qquad \#\text{ This saves some extra lemmas...}
\end{aligned}$$

So $P((t_1, t_2))$ follows.

(b) Use structural induction on $\mathcal{T}$ to prove: **[edit:] error fixed September 9**

$$\forall t \in \mathcal{T}, \text{double\_count}(t) = \begin{cases} 0 & \text{if } t = "*" \\ \text{left\_count}(t) - 1 & \text{otherwise} \end{cases}$$

sample solution: Define

$$P(t) : \text{double\_count}(t) = \begin{cases} 0 & \text{if } t = "*" \\ \text{left\_count}(t) - 1 & \text{otherwise} \end{cases}.$$

I will prove by structural induction $\forall t \in \mathcal{T}, P(t)$

base case: There are no parentheses, let alone double parentheses, in "*", so $\text{double\_count}(*) = 0$, and $P(*)$ follows.

inductive step: Let $t_1, t_2 \in \mathcal{T}$. Assume $P(t_1)$ and $P(t_2)$. I will show that $P((t_1, t_2))$ follows.

case $t_1 = t_2 = *$: In this, by $P(t_1)$ and $P(t_2)$, we have $\text{double\_count}(t_1) = \text{double\_count}(t_2) = 0$. By inspection $\text{left\_count}(t_1) = \text{left\_count}(t_2) = 0$, and (t_1, t_2) add one left parenthesis to the string, so

$$\text{double\_count}((t_1, t_2)) = 0 = 1 - 1 = \text{left\_count}((t_1, t_2)) - 1$$

... so $P((t_1, t_2))$ follows in this case.

case $t_1 = * \neq t_2$: The initial left parenthesis in $((t_1, t_2))$ increases the number of left parentheses in the expression by 1, but not the number of double-left parentheses, since $t_1 = *$. As well $((t_1, t_2))$ increases the double-right parentheses by 1, since the last character of t_2 is a right parentheses. This means we have:

$$\begin{aligned}
\text{double_count}((t_1, t_2)) &= \text{double_count}(t_1) + \text{double_count}(t_2) + 1 \\
&= 0 + \text{left_count}(t_2) - 1 + 1 \qquad \#\text{ by } P(*) \text{ and } P(t_2) \\
&= \text{left_count}((t_1, t_2)) - 1 \\
&\qquad \#\text{ since } (t_1, t_2) \text{ adds one more left parenthesis}
\end{aligned}$$

So $P((t_1, t_2))$ follows in this case.

case $t_1 \neq * = t_2$: By a symmetrical argument to the previous case, switching the roles of right and left, as well as t_1 and t_2, $P((t_1, t_2))$ follows in this case.

case both $t_1 \neq *$ and $t_2 \neq *$: The initial left parenthesis in $((t_1, t_2))$ increases the number of left parentheses by 1, and the number of double-left parentheses by 1. Similarly, the final right parentheses in $((t_1, t_2))$ increases the number of double-right parenthesis by 1. This means we have:

$$
\begin{aligned}
\text{double\_count}((t_1, t_2)) &= 1 + \text{double\_count}(t_1) + \text{double\_count}(t_2) + 1 \\
&= 1 + \text{left\_count}(t_1) - 1 + \text{left\_count}(t_2) - 1 + 1 \\
&\qquad \#\ \text{by } P(t_1) \text{ and } P(t_2) \\
&= \text{left\_count}((t_1, t_2)) - 1 \\
&\qquad \#\ \text{since } (t_1, t_2) \text{ add one more left parenthesis}
\end{aligned}
$$