15-251: Great Theoretical Ideas
Dos and Donts in Inductive Proofs

Consider the problem of proving that $\forall n \geq 0, 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$ by induction.

Define the statement S_n = "$1 + 2 + \ldots + n = \frac{n(n+1)}{2}$". We want to prove $\forall n \geq 0, S_n$.

1 An Inductive Proof

Base Case: $\frac{0(0+1)}{2} = 0$, and hence S_0 is true.

I.H.: Assume that S_k is true for some $k \geq 0$.

Inductive Step: We want to prove the statement $S(k+1)$. Note that

$$\begin{aligned} 1 + 2 + \ldots + k + (k+1) &= \tfrac{k(k+1)}{2} + (k+1) && \text{(by I.H.)} \\ &= (k+1)(\tfrac{k}{2} + 1) \\ &= \tfrac{(k+1)(k+2)}{2}. \end{aligned}$$

And hence S_{k+1} is true.

2 Common Errors and Pitfalls

1. *(S_n is a statement, not a value)* You cannot make statements like $S_k + (k+1) = S_{k+1}$, much the same as you cannot add k to the statement "The earth is round".

Mistake:
I.H.: Assume that S_k is true.

Inductive Step:

$$\begin{aligned} \textstyle\sum_{i=1}^{k+1} i &= k + 1 + \textstyle\sum_{i=1}^{k} i \\ &= k + 1 + S_k \\ &= \cdots \end{aligned}$$

Logical propositions like S_k can't be added to numbers. Please don't equate propositions and arithmetic formulas.

2. *(Proof going the Wrong Way)* Make sure you use S_k to prove S_{k+1}, and not the other way around. Here is a common (wrong!) inductive step:

> **Mistake:**
> Inductive Step:
>
> $$\begin{aligned} 1+2+\ldots+k+(k+1) &= (k+1)(k+2)/2 \\ k(k+1)/2+(k+1) &= (k+1)(k+2)/2 \\ (k+1)(k+2)/2 &= (k+1)(k+2)/2. \end{aligned}$$

The proof above starts off with S_{k+1} and ends using S_k to prove an identity, which does not prove anything. Please make sure you do not assume S_{k+1} in an effort to prove it!

3. *(Assuming too much)* Make sure you dont assume *everything* in the I.H.

> **Mistake:**
> I.H.: Assume that S_k is true *for all* k.

You *want to* prove the statement S_n true for all n, and if you assume it is true, there is nothing left to prove! (Remember that the "S_n is true for all n" is the same as saying "S_k is true for all k".)

> **Correct Way:**
> I.H.: Assume that $S(k)$ is true *for some* k.
>
> or, if you want to use all-previous ("strong") induction
>
> I.H.: Assume *for some* k that $S(j)$ is true *for all* $j \leq k$.

4. *(The case of the missing n)* Consider the following I.H. and inductive step:

> **Mistake:**
> I.H.: Assume that S_k is true for all $k \leq n$.
>
> Inductive Step: We want to prove S_{k+1}.

What is k? Where has n disappeared? The induction hypothesis is saying in shorthand that $S_1, S_2, \ldots, S_{n-1}, S_n$ are all true for *some* n. Note that rewriting the I.H. in this way shows that k was a red herring: you really want to prove S_{n+1}, not S_{k+1}.

> **Correct Way:**
> I.H.: Assume that S_k is true for all $k \leq n$.
>
> Inductive Step: We want to prove S_{n+1}.

5. *(Extra stuff in the I.H.)* Consider the following I.H.

> **Mistake:**
> I.H.: Assume that S_k is true for all $k \leq n$. **Then $\mathbf{S_{n+1}}$.**

Note that entire thing has been made part of the hypothesis, including the bolded part.

The second part "Then S_{n+1}" is what you want to show in the inductive step; it is *not* part of the induction hypothesis. You need to distinguish between the *Claim* and the *Induction Hypothesis*. The Claim is the statement you want to prove (i.e., $\forall n \geq 0, S_n$), whereas the Induction Hypothesis is an *assumption* you make (i.e., $\forall 0 \leq k \leq n, S_n$), which you use to prove the next statement (i.e., S_{n+1}). The I.H. is an assumption which might or might not be true (but if you do the induction right, the induction hypothesis will be true).

> **Correct Way:**
> I.H.: Assume that S_k is true for all $k \leq n$.

6. (*The Wrong Base Case.*) Note that you want to prove S_0, S_1, etc., and hence the base case should be S_0.

> **Mistake:**
> Base Case: $\frac{1(1+1)}{2} = 1$, and hence S_1 is true.

Even if the rest of the proof works fine, you would have shown that $S_1, S_2, S_3, \cdots$ are all correct. You haven't shown that S_0 is true.

7. (*Assuming too little: Too few Base Cases.*)

Suppose you were given a function $X(n)$ and need to show that the statement S_n that "the Fibonacci number $F_n = X(n)$" for all $n \geq 0$.

> **Mistake:**
> Base Case: for $n = 0$, $F_0 = X(0)$ blah blah. Hence S_0 is true.
> I.H.: Assume that S_k is true for all $k \leq n$.
>
> Induction Step: Now $F_n = F_{n-1} + F_{n-2} = X(n-1) + X(n-2)$ (because S_{n-1} and S_{n-2} are both true), etc.

If you are using S_{n-1} and S_{n-2} to prove $T(n)$, then you better prove the base case for S_0 **and** S_1 in order to prove S_2. Else you have shown S_0 is true, but have no way to prove S_1 using the above proof—S_0 is not a base case, and to use induction, we'd need S_0 and S_{-1}. But there is no S_{-1}!!!

Remember the domino principle: the above induction uses the fact that "if two consecutive dominoes fall, the next one will fall". To now infer that *all* the dominoes fall, *you must show that the first* **two** *dominoes fall.* And hence you need two base cases.