The next core topic in the course is *recursion*. We will look at a number of recursive algorithms and we will analyze how long they take to run. Recursion can be a bit confusing when one first learns about it. One way to understand recursion is to relate it to a proof technique in mathematics called *mathematical induction*, which is what I'll cover today.

Before I introduce induction, let me give an example of a statement that you have seen before, along with a proof. The proof is slightly different from the one I gave in the slides.

$$\text{For all } n \geq 1, \qquad 1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}.$$

One trick for showing this statement is true is to add up two copies of the left hand side, one forward and one backward:

$$1 + 2 + 3 + \cdots + (n-1) + n$$
$$n + (n-1) + ... + 3 + 2 + 1.$$

Then pair up terms:

$$1 + n, 2 + (n-1), 3 + (n-2), ...(n-1) + 2, n + 1$$

and note that each pair sums to $n+1$ and there are n pairs, which gives $(n+1) * n$. We then divide by 2 because we added up two copies. This proves the statement above.

Mathematical induction

The above proof requires a trick, and many proofs in mathematics are like that – they use a specific trick that seems to work only in a few cases. Mathematical induction is different in that it is a general type of proof technique. To understand a proof by mathematic induction, you need to understand the logic of the proof technique *in general*.

Mathematical induction allows one to prove statements about positive integers. We have some proposition $P(n)$ which is either true or false and the truth value may depends on n. We want to prove that: "for all $n \geq n_0$, $P(n)$ is true", where n_0 is some constant that we state explicitly. In the above example, this constant is 1, but sometimes we have some other constant that is greater than 1.

A proof by *mathematical induction* has two parts, and one needs to prove both parts.

1. a *base case*: the statement $P(n)$ is true for $n = n_0$.

2. *induction step*: for any $k \geq n_0$, *if* $P(k)$ is true, *then* $P(k+1)$ must also be true.

   When we talk about $P(k)$ in step 2, we refer to it as the "induction hypothesis". Note that $P(k)$ is just $P(n)$ with $n = k$, *i.e.* it is the same proposition, but we are using parameter k instead of n to emphasize that we're in the context of proving the induction step.

The logic of a proof by mathematical induction goes like this. Let's say we can prove both the base case and induction step. Then, $P(n)$ is true for the base case $n = n_0$, and the induction step implies that $P(n)$ is true for $n = n_0 + 1$, and applying the induction step again implies that the statement is true for $n = n_0 + 2$, and so on forever for all $n \geq n_0$.

Example 1

Statement: for all $n \geq 1$,

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

Proof: The base case, $n_0 = 1$, is true, since

$$1 = \frac{1 \cdot (1+1)}{2}.$$

We next prove the induction step. For any $k \geq 1$, we assume $P(k)$ is true and we show that $P(k+1)$ must therefore also be true.

$$\begin{aligned}
\sum_{i=1}^{k+1} i &= (\sum_{i=1}^{k} i) + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1), \text{ by induction hypothesis that P(k) is true} \\
&= (k+1)(\frac{k}{2} + 1) \\
&= \frac{1}{2}(k+1)(k+2)
\end{aligned}$$

and so $P(k+1)$ is also true. This proves the induction step. The proof is complete, since we have proven the base case and the induction step.

Example 2

Statement: for all $n \geq 3$,

$$2n + 1 < 2^n.$$

Proof: The base case $n_0 = 3$ is easy to prove, i.e. $7 < 8$. (Note that the base case $n_0 = 2$ would not be correct, nor would $n_0 = 1$. That's why we chose $n_0 = 3$.)
To prove the induction step, let k be any integer such that $k \geq 3$. We hypothesize that $P(k)$ is true and show that it would follow that $P(k+1)$ is also true. Note that $P(k)$ is the inequality

$$2k + 1 < 2^k$$

and $P(k+1)$ is the inequality

$$2(k+1) + 1 < 2^{k+1}.$$

To prove $P(k+1)$ we work with the expression on the left side of the inequality:

$$\begin{aligned}
2(k+1) + 1 &= 2k + 3 \\
&= (2k+1) + 2 \\
&< 2^k + 2, \quad \text{by induction hypothesis that } P(k) \text{ is true} \\
&< 2^k + 2^k, \quad \text{since } 2 < 2^k, \text{when } k \geq 3 \\
&= 2^{k+1}.
\end{aligned}$$

Thus, if $P(k)$ is true, then it must be that $P(k+1)$ is true also.

[ASIDE: You might be asking yourself, how did I know to use the inequality $2 < 2^k$? The answer is that I knew what inequality I eventually wanted to have, namely $P(k+1)$. Experience told me how to get there.

Example 3

Statement: For all $n \geq 5$, $n^2 < 2^n$.

Proof: The base case $n_0 = 5$ is easy to prove, i.e. $25 < 32$.
Next we prove the induction step. The induction hypothesis $P(k)$ is that inequality $k^2 < 2^k$ holds, where $k \geq 5$. We show that if $P(k)$ is true, then $P(k+1)$ must also be true, namely $(k+1)^2 < 2^{k+1}$. We start with the left side of $P(k+1)$

$$\begin{aligned}(k+1)^2 &= k^2 + 2k + 1 \\ &< 2^k + 2k + 1, \quad \text{by induction hypothesis, for } \; k \geq 5 \\ &< 2^k + 2^k, \quad \text{from Example 2} \\ &= 2^{k+1}\end{aligned}$$

which proves the induction step.

Note that the base case choice is crucial here. The statement $P(n)$ is just not true for $n = 0, 1, 2, 3, 4$. Also, note that the induction step happens to be valid for a larger range of k, namely $k \geq 3$ rather than $k \geq 5$. But we only needed it for $k \geq 5$.

Example 4: upper bound on Fibonacci numbers

Consider the Fibonnacci[1] sequence: 0, 1, 1, 2, 3, 5, 8, 13, 21, ... where

$$F(0) = 0, \qquad F(1) = 1,$$

and, for all $n > 2$, we define $F(n)$ by

$$F(n) \equiv F(n-1) + F(n-2).$$

Statement:

$$\text{for all } n > 0, \quad F(n) < 2^n.$$

Proof:

The statement $P(k)$ is "F(k) $< 2^k$". We take the base case to be two values $n_0 = 0, 1$. By definition, $F(0) = 0, F(1) = 1$. Since $F(0) = 0 < 2^0$ and $F(1) = 1 < 2^1$, $P(n)$ is true for both base cases.

The induction hypothesis $P(k)$ is that $F(k) < 2^k$ where $k \geq 2$. For the induction step, we hypothesize that $P(k)$ is true for some k and we show this would imply $P(k+1)$ must also be true,

[1] `http://en.wikipedia.org/wiki/Fibonacci_number`

that is, $F(k+1) < 2^{k+1}$. Again, we start with the left side of this inequality:

$$\begin{aligned} F(k+1) &\equiv F(k) + F(k-1) \\ &< 2^k + 2^{k-1} \quad \text{by induction hypothesis} \\ &< 2^k + 2^k \\ &= 2^{k+1} \end{aligned}$$

and so $P(k)$ is true indeed implies that $P(k+1)$ is true, and so we are done.

Next lecture, we will look at the technique of recursion and show how it is related to the idea of mathematical induction.