1 Induction and its generalizations

There are several variations on induction which are worthwhile. We can start with the basic version (to be proven in class):

**Theorem 1** (The Principle of Mathematical Induction). For any open statement \( P(n) \) whose parameter \( n \) is a natural number, if both \( P(1) \) is true and \( P(n) \) implies \( P(n+1) \) for every natural number \( n \), then \( P(n) \) is true for every natural number \( n \).

Alternatively, written entirely symbolically: for any open statement \( P(n) \) whose parameter \( n \) is a natural number, \( \left[ P(1) \land \forall n \in \mathbb{N}, (P(n) \rightarrow P(n+1)) \right] \rightarrow \forall n \in \mathbb{N}, P(n) \).

Inductive proof, then, is a matter of showing that both of the premises of the Principle are true: we need to show that \( P(1) \) is true and that every \( P(n) \) implies \( P(n+1) \), and then using the Principle of Mathematical Induction we can prove that every \( P(n) \) is true. The proof that \( P(1) \) is true is known as the base case; the proof that, taking \( P(n) \) as a given premise, we can conclude \( P(n+1) \) is known as the inductive step, within which our presumption \( P(n) \) is known as the inductive hypothesis.

The natural numbers are tailor-made for induction, but actually any translation of the natural numbers up or down would work just as well:

**Theorem 2** (The Principle of Mathematical Induction with Generalized Base Case). For any integer \( n_0 \) and open statement \( P(n) \) whose parameter \( n \) is an integer, if both \( P(n_0) \) is true and \( P(n) \) implies \( P(n+1) \) for every integer \( n \geq n_0 \), then \( P(n) \) is true for every integer \( n \geq n_0 \).

Alternatively, written entirely symbolically: for any integer \( n_0 \) and open statement \( P(n) \) whose parameter \( n \) is an integer, \( \left[ P(n_0) \land \forall n \in \mathbb{Z} \cap [n_0, \infty), (P(n) \rightarrow P(n+1)) \right] \rightarrow \forall n \in \mathbb{Z} \cap [n_0, \infty), P(n) \).

Note that the case \( n_0 = 1 \) is identical to the simple principle presented above, since the set of integers greater than or equal to 1 is the same as the set of natural numbers.

When we use induction, it is sometimes useful to take as our inductive hypothesis not only the case immediately preceding the one we wish to prove, but in fact the entirety of cases up to that point; that is to say, in proving, for instance, \( P(5) \) to be true, we should be able to presume not only that \( P(4) \) has already been proven true, but that \( P(3), P(2), \) and \( P(1) \) have been as well. This variant of the inductive principle is known as strong, total, or complete induction, and unless otherwise specified, invoking “induction” assumes we are invoking this variant, since it is actually superior in every respect to the more basic form of induction. The explicit statement of this principle is as follows.

**Theorem 3** (The Principle of Complete Mathematical Induction). For any open statement \( P(n) \) whose parameter \( n \) is a natural number, if both \( P(1) \) is true and the conjunction of all the statements \( P(k) \) for natural numbers \( k \leq n \) implies \( P(n+1) \) for every natural number \( n \), then \( P(n) \) is true for every natural number \( n \).

Alternatively, written entirely symbolically: for any open statement \( P(n) \) whose parameter \( n \) is a natural number, \( \left[ P(1) \land \forall n \in \mathbb{N}, (\bigwedge_{k=1}^{n} P(k) \rightarrow P(n+1)) \right] \rightarrow \forall n \in \mathbb{N}, P(n) \). (Note our use of that fancy “iterated-conjunction” symbol, a logical-operation analogue of the familiar big-sigma sums, big-pi products, big-cup unions, and big-cap interstections.)

Complete induction, just like its weaker cousin, can also be based off an arbitrary integer, since there’s nothing sacred about using 1 as our base case:
Theorem 4 (The Principle of Complete Mathematical Induction with Generalized Base Case). For any integer \( n_0 \) and open statement \( P(n) \) whose parameter \( n \) is an integer, if both \( P(n_0) \) is true and the conjunction of all the statements \( P(k) \) for integers \( k \) with \( n_0 \leq k \leq n \) implies \( P(n+1) \) for every integer \( n \geq n_0 \), then \( P(n) \) is true for every natural number \( n \geq n_0 \).

Alternatively, written entirely symbolically: for any integer \( n \) and open statement \( P(n) \) whose parameter \( n \) is a natural number,
\[
[P(1) \land \forall n \in \mathbb{Z} \cap [n_0, \infty), (\bigwedge_{k=n_0}^{n} P(k) \rightarrow P(n+1))] \rightarrow \\
\forall n \in \mathbb{Z} \cap [n_0, \infty), P(n).
\]

2 Examples of induction

There are several useful or entertaining problems which can be solved inductively; hopefully the following examples illustrate some useful inductive methods.

2.1 Nonstandard base case

First, however, we need a definition which will be used in the problem statement.

Definition 1. For any natural number \( n \), the factorial of \( n \), denoted \( n! \), is the product \( n(n-1)(n-2)\cdots3\cdot2\cdot1 \).

For example, \( 5! = 5\cdot4\cdot3\cdot2\cdot1 = 120 \). Factorials, we can prove inductively, grow very rapidly.

Proposition 1. For any integer \( n \geq 4 \), then \( n! > 2^n \).

Proof. We shall perform induction on \( n \) with a base case of \( n = 4 \) (motivated by the proposition’s explicit limitation of the statement to values of \( n \geq 4 \)). To illustrate our base case, let us note that \( 4! = 24 \) and \( 2^4 = 16 \); since \( 24 > 16 \), our base case has been established.

For an inductive step, we may assume that for a specific \( n \), \( n! > 2^n \); from that fact we wish to prove that \( (n+1)! > 2^{n+1} \). Since \( (n+1)! \) is simply \( (n+1)n(n-1)(n-2)\cdots3\cdot2\cdot1 = (n+1)n! \), there is a simple arithmetic operation to convert \( n! \) into \( (n+1)! \), namely, multiplication by \( (n+1) \). We thus take the inequality which is our inductive hypothesis and multiply by \( n+1 \) to get
\[
(n+1)n! > (n+1)2^n
\]
The left side of this inequality, as noted above, is simply \( (n+1)! \). The right side is not our desired \( 2^{n+1} \), but we might note that since \( n \geq 4 \), \( n+1 > 2 \), and thus \( (n+1)2^n > 2\cdot2^n \), from which we can get the desired conclusion
\[
(n+1)! = (n+1)n! > (n+1)2^n > 2\cdot2^n = 2^{n+1}.
\]

2.2 Induction on finite sets

Finite sets (and only finite sets) can sometimes be well-illuminated by induction, since we can always build a smaller set from a nonempty set by excising a single element. Here’s an example of such an argument.
Proposition 2. For a finite set $A$, $|\mathcal{P}(A)| = 2^{|A|}$.

Proof. Let us denote $|A|$ by $n$ and perform induction on $n$ with a base case of 0 (since sets can have as few as zero elements). In our base case, when $|A| = 0$, $A$ is uniquely determined to be the empty set and as has been seen previously, $\mathcal{P}(\emptyset) = \{\emptyset\}$ so $|\mathcal{P}(\emptyset)| = 1 = 2^0$ as expected.

For our inductive step, we may take as our inductive assumption for a specific non-negative integer $n$ the fact that any set with $n$ elements has a power set with $2^n$ elements, and then, considering a set $A$ with $|A| = n+1$ elements, seek to show that $\mathcal{P}(A)$ has $2^{n+1}$ elements. Since $|A| \geq 1$, we know $A$ has at least one element; let’s pick an arbitrary element of $A$ and call it $x$. We can now build a smaller set $A' = A - \{x\}$; since a single element has been excised from an $(n+1)$-element set, $|A'| = n$, and by our inductive hypothesis, $\mathcal{P}(A') = 2^n$.

Now, we wish to show that $\mathcal{P}(A)$ has exactly twice as many elements as $\mathcal{P}(A')$. To do that, let’s note that the subsets of $A$ can be divided into two classes: those which contain $x$ and those which do not. Any subset $S$ of $A$ which does not contain $x$ is also a subset of $A'$, since for any $y \in S$, it is the case that $y \in A$ and $y \neq x$, so $y \in A - \{x\} = A'$. Thus the subsets of $A$ which do not contain $x$ are exactly the subsets of $A'$, which are exactly $2^n$ in number. Any subset $T$ of $A$ which does contain $x$ can be transformed into a distinct subset $T - \{x\}$ of $A'$, so they will also be exactly $2^n$ in number, so $A$ has a total of $2^n + 2^n = 2^{n+1}$ subsets.

2.3 Algorithms presented inductively: The Tower of Hanoi

Let’s return to our old friend the Towers of Hanoi. Recall that the Towers of Hanoi are a system of three pegs and $n$ differently-sized discs, with a move consisting of taking a single disc off one peg and moving it to another peg so that it is not sitting on top of a smaller disc. The goal of the puzzle is to move the entire stack from one peg to another. Here’s the solution in seven moves to a three-disc puzzle, with the current move highlighted in red:

The 4-disc puzzle takes 15 moves to solve; the 5-disc puzzle takes 31. We can observe a pattern that the puzzle is in fact solvable in $2^n - 1$ moves, and in fact the process for doing so is, taken on a large-scale level, pretty straightforward. Here’s a conceptual presentation for how we would solve a 6-disc puzzle; note that two of the steps are not what we might think of as “moves” permitted by the game.
Of the three “steps” taken here, only the middle one is what we would regard as a move. The other two are relocations of an entire stack. However, the process of relocating an entire stack is in fact itself a Tower of Hanoi problem! The first “step” above is a relocation of a five-disc tower, which we already know is a 31-move process. The last step is likewise a five-disc tower movement, requiring 31 moves. So the process as a whole takes $31 + 1 + 31 = 63$ steps.

Now that we understand how the process is recursive, i.e., depending on its own smaller solutions to solve the large-scale problem, we can craft an inductive argument for why $2^n - 1$ moves works, based closely on generalizing the process we worked through above.

**Proposition 3.** For any natural number $n$, a $n$-disc Tower of Hanoi can be relocated in $2^n - 1$ moves.

**Proof.** We proceed by induction on $n$. Our base case $n = 1$ is trivial, since a 1-disc Tower of Hanoi is a single disc, which can be relocated simply by moving it, so a one-disc Tower can indeed be relocated in $2^1 - 1 = 1$ move.

Now, for our inductive step, let us take it as established that a tower of $n$ discs can be relocated in $2^n - 1$ moves, and seek to show that a tower of $n + 1$ discs can be relocated in $2^{n+1} - 1$ moves. Let us denote the peg the discs are already on as $S$ (for “start”), the peg we want to move them to as $T$ (for “target”), and the third peg as $X$ (for “extra”). Now, by our inductive hypothesis it is possible to relocate the top $n$ discs (which themselves form an $n$-disc Tower of Hanoi) from $S$ to $X$ in $2^n - 1$ moves. After that is done, $S$ contains only the largest disc and $T$ is empty, so we can use a single move to transfer the largest disc from $S$ to $T$. Now we want to relocate the $n$-disc tower currently on $X$ to on top of the large disc on $T$, and by our inductive hypothesis this can be done in $2^n - 1$ moves. Thus the whole process takes $(2^n - 1) + 1 + (2^n - 1) = 2^{n+1} - 1$ moves.

This is in fact not only a way to solve the problem but is the quickest way; showing that to be the case, however, is somewhat more difficult.

### 2.4 Silly fun: The Blue-Eyed Villagers

Imagine, if you will, an idyllic, remote village, full of highly logical but oddly superstitious people. Their superstition is that it is a great sin to know one’s own eye color, and that anyone who learns their own eye color must leave the village by night (or jump into the volcano at noon, depending on the version of the story; in any case, assume that anyone who learns their own eye color will, at a predetermined time, make it known to the entire village by their behavior that they know their own eye color). There are no mirrors, and people do not talk about eye color, to avoid this misfortune. As it happens, 200 villagers have blue eyes, 250 have brown eyes, and 50 have green eyes.

One day, a visitor to the village, not knowing their strange customs, indiscreetly mentioned a blue-eyed villager in attempting to describe someone. He was hushed up so that the specific villager in question is unknown, but everyone heard the visitor and is thus aware that someone has blue eyes.

The question is: what would the results of such an indiscretion be, if the villagers use their logical powers to the fullest?
Think about how such a scenario would play out in the simpler case where there was only one blue-eyed villager (and 499 with other eye colors). This villager, on looking around and seeing that all 499 other villagers were not blue-eyed, would correctly deduce their own eye-color, and would exile themself that night.

Now suppose that there were two blue-eyed villagers. Each of those two would look around, and see that there was one person they could see with blue eyes. They would then both reason: “If that person is the only blue-eyed person, they’ll realize that, and exile themselves tonight.” But come morning both are still there, and each would realize that the other person did not leave because they saw another blue-eyed person; since they are the only possible such person, they would then both know their own eye color and leave then, on the second night.

If there were three such, then each could look at their two comrades and follow the above line of logic to say “if those two are the only two blue-eyed people, they should both exile themselves on the second night.” But of course, none of them will do so, so on the morning after the second night, each, seeing that the two blue-eyed people they can see did not leave, will conclude that their own eyes must be blue, and will leave on the third night.

By now, perhaps, we see a pattern, and can well believe that, 200 nights after the announcement, there will be a mass exodus of blue-eyed people (the brown- and green-eyed know nothing new, of course, and will remain). We can prove this pattern by induction, since each new blue-eyed villager we introduce involves a logic chain based on the smaller case.

**Proposition 4.** In a village as described above with \( n \) blue-eyed villagers, all of them will leave \( n \) nights after the indiscretion.

**Proof.** We prove this by induction; the \( n = 1 \) case is simple, since if only one person has blue eyes, they can ascertain immediately that they are the subject of the indiscretion, and will leave that night.

Now, let us assume it true for a particular \( n \), and prove it in a village of \( n + 1 \) blue-eyed people. Each of those blue-eyed people will look around at the time of the indiscretion and see \( n \) blue eyes, and reason as such: “Either this is a village with \( n \) blue-eyed people (and I’m not one of them), or a village with \( n + 1 \) blue-eyed people (and I am one of them). If the former case is true, then by the inductive hypothesis all of those people will leave in \( n \) nights. So if they’re still here after the \( n \)th night, I know the latter case holds.” This line of reasoning does not impel any of them to leave in the first \( n \) nights, but on the day after the \( n \)th night, they will have met the premise of the very last statement in their line of reasoning, and know their eye color to be blue; thus they will all leave on the \((n + 1)\)th night.

This problem raises an interesting paradox: why should the stranger’s pronouncement be so calamitous? After all, he’s not really telling them anything they don’t already know: any villager, looking around, already sees either 199 (if they are blue-eyed) or 200 (if not) pairs of blue eyes in the village, and the stranger mentioning offhand that someone in the village is blue-eyed shouldn’t be news.

Again, it is easier to understand by looking at small cases. With one blue-eyed villager, the stranger’s pronouncement is news to one person, who acts on that new knowledge.

With two, even though everyone already does know there are blue eyes in the village, one villager might very well believe that another villager is ignorant; so blue-eyed villager A, seeing only blue-eyed villager B, might labor under the delusion that that villager B hasn’t seen any blue eyes and doesn’t know that anyone has them. So in this case, the critical information
imparted is not the knowledge that someone has blue eyes (which is knowledge everyone has already), but the knowledge that everyone knows that someone has blue eyes, which we might consider to be metaknowledge.

Similarly, in a three-villager case, the logical daisy chain is built on the presumption that the blue-eyed villagers you see know that everyone knows someone has blue eyes, so the piece of critical information which sets the logic in motion is “I know that everyone knows that everyone knows that someone has blue eyes.” That there is starting to get to a very abstract level of self-reference.

And this only gets worse as the number goes up, so in the 200-villager case, the information actually imparted by the stranger’s indiscretion being audible to the entire village is of a very subtle sort dealing with knowledge of others’ knowledge of others’ knowledge nested 200 levels deep! Nonetheless, there is a germ of new information there, which, over 200 cycles of logic, eventually reaps a particularly profound harvest.

2.5 Siller fun: The Horse of a Different Color

We shall “prove” by induction that all horses are the same color. Now, obviously this is a false statement — so the proof clearly has a flaw. Read it, and see if the flaw appears to you!

**Proposition 5.** For any finite set of horses $S$, all elements of $S$ are the same color.

“Proof”. We shall perform induction on the size $n$ of the set $S$. When $n = 1$, we have a set of a single horse, which is clearly the same color as itself.

Now, let us undertake the inductive step. For a particular $n \geq 2$, we can assume that any set of $n$ horses has elements all of the same color, and we seek to prove that any set of $n + 1$ horses has elements all of the same color. Let $S$ be a set of $n + 1$ horses.

Let us consider any element $x$ of $S$. The set $T = S - \{x\}$ has size $n$, so all the horses in $T$ are the same color. Now, for a different element $y$ of $S$, we can build a set $T' = S - \{y\}$; this set also has only $n$ horses, so all the horses in $T'$ are also the same color. The horses in $T \cap T'$ are all thus of the same color as $x$ and all of the same color as $y$, so $x$ and $y$ are the same color as each other, and thus, since $y \in T$ and $x$ is therefore the same color as all the horses in $T$, the set $T \cup \{x\} = S$ consists of horses which are all the same color.\n
Think on it for a moment, and read on only after you’ve thought over how it might be flawed.