Invariants

We have a robot that lives on an infinite grid. Initially, it is at position $(0,0)$. At any point, it can take a single step in one of four directions: northeast, northwest, southwest, or southeast. In other words, if it is at position (x,y), it can move to one of the four positions $(x+1,y+1)$, $(x-1,y+1)$, $(x-1,y-1)$, $(x+1,y-1)$ in a single step. Can the robot ever reach position $(1,0)$?

A bit of experimentation will suggest that the robot can never reach position $(1,0)$, no matter how many steps it takes. Let's prove this.

Theorem: If the robot can reach position (x,y), then $x+y$ is even.

Proof: We will prove this by induction on the number n of steps the robot has taken. Let $P(n)$ denote the proposition that after any sequence of n steps, the robot can only reach positions of the form (x,y) such that $x+y$ is even.
Base case: After 0 steps, the robot can only be at the position $(0,0)$, and $0+0$ is indeed even. Therefore $P(0)$ is true.
Inductive hypothesis: Suppose that $P(n)$ is true. In other words, after any sequence of n steps, the robot can only reach positions (x,y) where $x+y$ is even.
Induction step: Consider any sequence of $n+1$ steps. Let (x,y) denote the position of the robot after the first n of these steps, and (x',y') denote its position after the $n+1^{\text{st}}$ step. By the induction hypothesis, $x+y$ is even. Now there are four cases, corresponding to which direction the robot moved in the $n+1^{\text{st}}$ step:

1. $(x',y') = (x+1,y+1)$: then $x'+y' = x+y+2$, which is the sum of two even numbers (namely, $x+y$ and 2) and hence is even.
2. $(x',y') = (x-1,y+1)$: then $x'+y' = x+y$, which is even.
3. $(x',y') = (x-1,y-1)$: then $x'+y' = x+y-2$, which is even.
4. $(x',y') = (x+1,y-1)$: then $x'+y' = x+y$, which is even.

Consequently, after any sequence of $n+1$ steps, the robot is at some position (x',y') such that $x'+y'$ is even. In other words, we have proven that $P(n+1)$ follows from $P(n)$. By induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

Corollary: The robot can never reach position $(1,0)$.

The key idea of this proof was to identify an *invariant*. An invariant is a boolean proposition whose truth value does not change throughout the process. In this case, the invariant is the property that $x+y$ is even. In general, an invariant is a property that is initially true, and where if the invariant is true before a single step of the process, then it will remain true after that step. If these two conditions are met, it will follow by induction on the number of steps of the process, that the invariant is forever true.

Here is another example. Suppose we have a bunch of buckets lined up from left to right. Each bucket can hold any (finite) number of marbles. We are going to play a game using these marbles. When it is his turn,

a player may remove a marble from one of the (non-empty) buckets, say, the jth bucket from the left. Then, the player may add up to two marbles to any bucket that is further to the left, say, to the ith bucket, where $i < j$. As a special case, if the player removes a marble from the leftmost bucket, the player is not allowed to add any more marbles to any bucket. Then, it is the next player's turn. The last person who can make a legal move is the winner.

One might wonder: Is this game always guaranteed to end after some finite number of turns? It turns out that the answer is Yes, and this can be justified by finding an appropriate invariant.

The idea is to introduce an (artificial) dollar value on the configuration of marbles. Each marble is assigned a different dollar value according to which bucket it is currently found in. A marble in the ith bucket from the left will be considered to be worth 3^i dollars. For instance, each marble in the leftmost bucket (the 0th bucket from the left) is worth \$1. The value of a configuration is the sum of the dollar values of the marbles.

Notice that each turn reduces the dollar value of the configuration by at least one dollar. For instance, removing a marble from the jth bucket and adding two marbles to the ith bucket (for $i < j$) decreases the value of the configuration by $3^j - 2 \cdot 3^i$ dollars, and if $i < j$, then $3^j - 2 \cdot 3^i > 0$. Also, since the initial configuration has only finitely many marbles, it is worth a finite number of dollars. If we start with a finite natural number and subtract at least one from it in each turn, then after finitely many turns we must reach zero, and then the game is over.

This can be formalized. Let v_n denote the value of the configuration of marbles after n turns.

Theorem: For every $n \in \mathbb{N}$, $v_n \leq v_0 - n$.

Proof: As shown above, in each turn, the value is reduced by at least one dollar. In other words, we have $v_{n+1} \leq v_n - 1$ for every $n \in \mathbb{N}$. The theorem follows by an easy induction on n, since $v_n \leq v_0 - n$ implies that $v_{n+1} \leq v_n - 1 \leq v_0 - (n+1)$. □

Corollary: The game cannot last more than v_0 turns.

Here we used the invariant $v_n \leq v_0 - n$. Notice that the dollar values we set are completely artificial and have no meaning of their own. They are an accounting fiction that we created solely for purposes of formulating the invariant.

Why did we decide to make marbles in the ith bucket worth 3^i dollars, rather than some other function of i? The pragmatic answer is that this choice suffices to make the proof work. For instance, if we had considered marbles in the ith bucket to be worth i dollars per marble, then we could not have shown that the total dollar value of the configuration is strictly decreasing. (Removing a marble from the 10th bucket and adding two marbles to the 9th bucket would increase the total dollar value by \$8, creating money out of thin air.) On the other hand, if we had valued marbles in the ith bucket at 4^i dollars per marble instead of 3^i dollars, the proof would still have worked. [Exercise: what is the minimum dollar valuation we could have assigned to each marble in the ith bucket, and still ensure that the total value will decrease by at least one dollar in each turn?]

Strong Induction

Strong induction is another form of induction that can be useful in some cases where simple induction is not applicable. It is very similar to simple induction, except that the inductive hypothesis is different. With strong induction, instead of just assuming $P(n)$ is true, you assume the stronger statement that $P(0)$, $P(1)$, ..., and $P(n)$ are all true (i.e., $P(0) \wedge P(1) \wedge \cdots \wedge P(n)$ is true, or in more compact notation $\bigwedge_{i=0}^{n} P(i)$ is true). Strong induction sometimes makes the proof of the inductive step much easier since we get to assume a stronger statement, as illustrated in the next example.

Theorem: Every natural number $n > 1$ can be written as a product of primes.

Recall that a number n is prime if 1 and n are its only divisors.

Proof: Let $P(n)$ be the proposition that n can be written as a product of primes. We will prove that $P(n)$ is true for all $n \geq 2$ by induction on n.
Base case: We start at $n = 2$. Clearly $P(2)$ holds, since 2 is a prime number.
Inductive hypothesis: Assume $P(k)$ is true for $2 \leq k \leq n$: i.e., every number k such that $2 \leq k \leq n$ can be written as a product of primes.
Inductive step: We must show that $n+1$ can be written as a product of primes. We have two cases: either $n+1$ is a prime number, or it is not. For the first case, if $n+1$ is a prime number, then we are done. For the second case, if $n+1$ is not a prime number, then by definition $n+1$ has a divisor x such that $1 < x < n+1$. Let $y = (n+1)/x$; since x is a divisor of $n+1$, y is a positive integer; and since $1 < x < n+1$, we see that $1 < y < n+1$. Therefore $n+1 = xy$, where $x, y \in \mathbb{N}$ and $1 < x, y < n+1$ (or, in other words, $2 \leq x, y \leq n$). By the inductive hypothesis, x and y can each be written as a product of primes (since $2 \leq x, y \leq n$). Therefore, $n+1$ can also be written as a product of primes. □

Why does this proof fail if we were to use simple induction? If we only assume $P(n)$ is true, then we cannot apply our inductive hypothesis to x and y. For example, if we were trying to prove $P(42)$, we might write $42 = 6 \times 7$, and then it is useful to know that $P(6)$ and $P(7)$ are true. However, with simple induction, we could only assume $P(41)$, i.e., that 41 can be written as a product of primes — a fact that is not useful in establishing $P(42)$.

Simple Induction vs. Strong Induction

We have seen that strong induction makes certain proofs easy when simple induction seems to fail. A natural question to ask then, is whether the strong induction axiom is logically stronger than the simple induction axiom. In fact, the two methods of induction are logically equivalent. Clearly anything that can be proven by simple induction can also be proven by strong induction (convince yourself of this!). For the other direction, suppose we can prove by strong induction that $\forall n \,.\, P(n)$. Let $Q(n) = P(0) \wedge \cdots \wedge P(n)$. Let us prove $\forall n \,.\, Q(n)$ by simple induction. The proof is modeled after the strong induction proof of $\forall n \,.\, P(n)$. That is, we want to show $(P(0) \wedge \cdots \wedge P(n)) \implies (P(0) \wedge \cdots \wedge P(n) \wedge P(n+1))$. But this is true iff $(P(0) \wedge \cdots \wedge P(n)) \implies P(n+1)$. This is exactly what the strong induction proof of $\forall n \,.\, P(n)$ establishes. Therefore, we can establish $\forall n \,.\, Q(n)$ by simple induction. And clearly, proving $\forall n \,.\, Q(n)$ also proves $\forall n \,.\, P(n)$.

If this sounds rather abstract, here is another way to put it. In our proof that every $n > 1$ can be written as a product of primes, we used strong induction, where $P(n)$ was the proposition that n can be written as a product of primes. We could alternatively have proven that theorem using simple induction and $Q(n)$, where $Q(n)$ is the proposition that each of the natural numbers $2, 3, \ldots, n$ can be written as a product of primes. The same kind of reasoning shows that $Q(n) \implies Q(n+1)$ is true for every n, and then the principle of mathematical induction establishes that $Q(n)$ is true for every n. That suffices to prove that every $n > 1$ can be written as a product of primes. So, we could have proven our theorem using simple induction, if we really wanted—but using strong induction made the proof a bit cleaner.