Faulty Inductions

Part of why I want you to write out your assumptions carefully is so that you don't get led into some standard errors.

Theorem: All women are blondes.

Proof by induction: Let $P(n)$ be the statement: For any set of n women, if at least one of them is a blonde, then all of them are.

Basis: Clearly OK.

Inductive step: Assume $P(n)$. Let's prove $P(n+1)$.

Given a set W of $n+1$ women, one of which is blonde. Let A and B be two subsets of W of size n, each of which contains the known blonde, whose union is W.

By the induction hypothesis, each of A and B consists of all blondes. Thus, so does W. This proves $P(n) \Rightarrow P(n+1)$.

Take W to be the set of women in the world, and let $n = |W|$. Since there is clearly at least one blonde in the world, it follows that all women are blonde!

Where's the bug?

Theorem: Every integer > 1 has a unique prime factorization.

[The result is true, but the following proof is not:]

Proof: By strong induction. Let $P(n)$ be the statement that n has a unique factorization. We prove $P(n)$ for $n > 1$.

Basis: $P(2)$ is clearly true.

Induction step: Assume $P(2), \ldots, P(n)$. We prove $P(n+1)$. If $n+1$ is prime, we are done. If not, it factors somehow. Suppose $n+1 = rs$ $r, s > 1$. By the induction hypothesis, r has a unique factorization $\Pi_i p_i$ and s has a unique prime factorization $\Pi_j q_j$. Thus, $\Pi_i p_i \Pi_j q_j$ is a prime factorization of $n+1$, and since none of the factors of either piece can be changed, it must be unique.

What's the flaw??

Problem: Suppose $n + 1 = 36$. That is, you've proved that every number up to 36 has a unique factorization. Now you need to prove it for 36.

36 isn't prime, but $36 = 3 \times 12$. By the induction hypothesis, 12 has a unique prime factorization, say $p_1 p_2 p_3$. Thus, $36 = 3 p_1 p_2 p_3$.

However, 36 is also 4×9. By the induction hypothesis, $4 = q_1 q_2$ and $9 = r_1 r_2$. Thus, $36 = q_1 q_2 r_1 r_2$.

How do you know that $3 p_1 p_2 p_3 = q_1 q_2 r_1 r_2$.
(They do, but it doesn't follow from the induction hypothesis.)

This is a *breakdown error*. If you're trying to show something is unique, and you break it down (as we broke down $n+1$ into r and s) you have to argue that nothing changes if we break it down a different way. What if $n + 1 = tu$?

- The actual proof of this result is quite subtle

Theorem: The sum of the internal angles of a regular n-gon is $180(n-2)$ for $n \geq 3$.

Proof: By induction. Let $P(n)$ be "the sum of the internal angles of a regular n-gon is $180(n-2)$." For $n = 3$, the result was shown in high school. Assume $P(n)$; let's prove $P(n+1)$. Given a regular $(n+1)$-gon, we can lop off one of the corners:

By the induction hypothesis, the sum of the internal angles of the regular n-gon is $180(n-2)$ degrees; the sum of the internal angles of the triangle is 180 degrees. Thus, the internal angles of the original $(n+1)$-gon is $180(n-1)$. What's wrong??

- When you lop off a corner, you don't get a *regular* n-gon.

The fix: **Strengthen the induction hypothesis**.

- Let $P(n)$ say that the sum of the internal angles of *any* n-gon is $180(n-2)$.

Consider 0-1 sequences in which 1's may not appear consecutively, except in the rightmost two positions.

- 010110 is not allowed, but 010011 is

Prove that there are 2^n allowed sequences of length n for $n \geq 1$

Why can't this be right?

"Proof" Let $P(n)$ be the statement of the theorem.

Basis: There are 2 sequences of length 1—0 and 1—and they're both allowed.

Inductive step: Assume $P(n)$. Let's prove $P(n+1)$. Take any allowed sequence x of length n. We get a sequence of length $n+1$ by appending either a 0 or 1 at the end. In either case, it's allowed.

- If x ends with a 1, it's OK, because $x1$ is allowed to end with 2 1's.

Thus, $s_{n+1} = 2s_n = 22^n = 2^{n+1}$.

Where's the flaw?

- What if x already ends with 2 1's?

Correct expression involves separating out sequences which end in 0 and 1 (it's done in Chapter 5, but I'm not sure we'll get to it)

Inductive Definitions

Example: Define $\Sigma_{k=1}^{n} a_k$ inductively (i.e., by induction on n):

- $\Sigma_{k=1}^{1} a_k = a_1$
- $\Sigma_{k=1}^{n+1} a_k = \Sigma_{k=1}^{n} a_k + a_{n+1}$

The inductive definition avoids the use of $\cdots$, and thus is less ambiguous.

Example: An inductive definition of $n!$:

- $1! = 1$
- $(n+1)! = (n+1)n!$

Could even start with $0! = 1$.

Inductive Definitions of Sets

A *palindrome* is an expression that reads the same backwards and forwards:

- Madam I'm Adam
- Able was I ere I saw Elba

What is the set of palindromes over $\{a, b, c, d\}$? Two approaches:

1. The smallest set P such that
   (a) P contains a, b, c, d, aa, bb, cc, dd
   (b) if x is in P, then so is axa, bxb, cxc, and dxd

   Things to think about:

   - How do you know that there is a smallest set (one which is a subset of all others)
   - How do you know that it doesn't contain ab

2. Define P_n, the palindromes of length n, inductively:

   - $P_1 = \{a, b, c, d\}$
   - $P_2 = \{aa, bb, cc, dd\}$
   - $P_{n+1} = \{axa, bxb, cxc, dxd | x \in P_{n-1}\}$ for $n \geq 2$

   Let $P' = \cup_n P_n$.

Theorem: $P = P'$. (The two approaches define the same set.)
Proof: Show $P \subseteq P'$ and $P' \subseteq P$.
To see that $P \subseteq P'$, it suffices to show that

(a) P' contains a, b, c, d, aa, bb, cc, dd

(b) if x is in P', then so is axa, bxb, cxc, and dxd

(since P is the smallest set with these properties).

Clearly $P_1 \cup P_2$ satisfies (1), so P' does. And if $x \in P'$, then $x \in P_n$ for some n, in which case axa, bxb, cxc, and dxd are all in P_{n+2} and hence in P'. Thus, $P \subseteq P'$.

To see that $P' \subseteq P$, we prove by strong induction that $P_n \subseteq P$ for all n. Let $P(n)$ be the statement "$P_n \subseteq P$."

Basis: $P_1, P_2 \subseteq P$: Obvious.

Suppose $P_1, \ldots, P_n \subseteq P$. If $n \geq 2$, the fact that $P_{n+1} \subseteq P$ follows immediately from (b). (Actually, all we need is the fact that $P_{n-1} \subseteq P$, which follows from the (strong) induction hypothesis.)

Thus, $P' = \cup_n P_n \subseteq P$.

Recall that the set of palindromes is the smallest set P such that

(a) P contains a, b, c, d, aa, bb, cc, dd

(b) if x is in P, then so is axa, bxb, cxc, and dxd

"Smallest" is not in terms of cardinality.

- P is guaranteed to be infinite

"Smallest" is in terms of the subset relation.

Here's a set that satisfies (a) and (b) and isn't the smallest:

Define Q_n inductively:

- $Q_1 = \{a, b, c, d\}$
- $Q_2 = \{aa, bb, cc, dd, ab\}$
- $Q_{n+1} = \{axa, bxb, cxc, dxd | x \in Q_{n-1}\}$, $n \geq 2$

Let $Q = \cup_n Q_n$.

It's easy to see that Q satisfies (a) and (b), but it isn't the smallest set to do so.

The Sorites Paradox

If a pile of sand has $1,000,000$ grains of sand, it's a heap.

Removing one grain of sand from a heap leaves 1 heap.

Therefore, by induction, if a pile of sand has only one grain, it's also a heap.

Prove by induction on n that if a pile of sand has $1,000,000-n$ grains of sand, it's a heap.

Where's the bug?

- This leads to a whole topic in the philosophy of language called "vagueness"

The Trust Game

Consider a game where, after n steps, there are piles of money on the table:

- The big one has $\$2^{n+1}$; the small one has $\$2^{n-1}$

There are two players, Alice and Bob. Initially Alice is in charge. She can either quit the game or continue

- If she quits, she gets the money in the bigger pile ($4) and Bob gets the money in the smaller pile ($1)
- If she continues, Bob is in charge
- If he quits, he gets the money in the bigger pile ($8), Alice gets the money in the smaller pile ($2).
- If he continues, Alice is in charge, and so on.
- The game goes on for 20 steps;
    - if they're still playing then, Bob gets $\$2^{21}$ ($>$ $2,000,000); Alice gets $\$2^{19}$ ($\approx$ $500,000)

What should you do?

- Should you trust the other player to keep playing, or take your money and run?

In the game theory literature, this is called the *centipede game*.

What should Alice do if they're still playing at step 19?

- If she quits, she gets $\$2^{20}$ (about \$1,000,000); if she continues she gets only $\$2^{19}$).
- So Alice will quit, which means Bob will get $\$2^{18}$

So what should Bob do if they're still playing at step 18?

- If he quits, he gets $\$2^{19}$; if he continues, most likely he'll get $\$2^{18}$, since Alice will quit at step 19.
- So Bob quts, which means Alice will get $\$2^{16}$.

Continuing this way (by *backwards induction*), Alice quits at step 1 and gets \$4!

Under a specific model of *rationality*, quitting at the first step is the only right thing to do.

- It's the only *Nash equilibrium*

In practice (with smaller amounts of money), people play for a little while before quitting.

The muddy children puzzle

We can prove by induction on k that if k children have muddy foreheads, they say "yes" on the k^{th} question.
It appears as if the father didn't tell the children anything they didn't already know. Yet without the father's statement, they could not have deduced anything.
So what was the role of the father's statement?

Homework

- Grades are now posted on CMS
- Solutions will be posted shortly.
- It's a violation of academic integrity to copy previous years' solutions

Homework questions/complaints?

1. Read the solutions first!
2. Talk to the person who graded it (check initials)
3. If (1) and (2) don't work, talk to me.

Further comments:

- There's no statute of limitations on grade changes
  - but you're better off asking questions soon
- 10/12 homeworks count. Each is roughly worth 50 points, and homework is 35% of your final grade.
  - 16 homework points = 1% on your final grade
- We're grading about 150 homeworks and graders are not mind readers. It's **your** problem to write clearly.
- Don't forget to staple your homework pages together, add the cover sheet, and put your name on clearly.

A Digression: A Bad Proof

Prove $\log(x/y) = \log(x) - \log(y)$

Proof:

$$\begin{aligned}
&\log(x/y) = \log(x) - \log(y) \\
&\log(x) + \log(1/y) = \log(x) - \log(y) \\
&\log(x) + \log(y^{-1}) = \log(x) - \log(y) \\
&\log(x) - \log(y) = \log(x) - \log(y)
\end{aligned}$$

What's wrong?

Algorithmic number theory

Number theory used to be viewed as the purest branch of pure mathematics.

- Now it's the basis for most modern cryptography.
- Absolutely critical for e-commerce
  - How do you know your credit card number is safe?

Goal:

- To give you a basic understanding of the mathematics behind the RSA cryptosystem
  - Need to understand how prime numbers work

Division

For $a, b \in Z$, $a \neq 0$, a *divides* b if there is some $c \in Z$ such that $b = ac$.

- Notation: $a \mid b$
- Examples: $3 \mid 9$, $3 \nmid 7$

If $a \mid b$, then a is a *factor* of b, b is a *multiple* of a.

Theorem 1: If $a, b, c \in Z$, then

1. if $a \mid b$ and $a \mid c$ then $a \mid (b + c)$.
2. If $a \mid b$ then $a \mid (bc)$
3. If $a \mid b$ and $b \mid c$ then $a \mid c$ (divisibility is transitive).

Proof: How do you prove this? Use the definition!

- E.g., if $a \mid b$ and $a \mid c$, then, for some d_1 and d_2,

$$b = ad_1 \text{ and } \ c = ad_2.$$

- That means $b + c = a(d_1 + d_2)$
- So $a \mid (b + c)$.

Other parts: homework.

Corollary 1: If $a \mid b$ and $a \mid c$, then $a \mid (mb + nc)$ for any integers m and n.

The division algorithm

Theorem 2: For $a \in Z$ and $d \in N$, $d > 0$, there exist unique $q, r \in Z$ such that $a = q \cdot d + r$ and $0 \leq r < d$.

- r is the remainder when a is divided by d

Notation: $r \equiv a \pmod{d}$; $a \bmod d = r$

Examples:

- Dividing 101 by 11 gives a quotient of 9 and a remainder of 2 ($101 \equiv 2 \pmod{11}$; $101 \bmod 11 = 2$).
- Dividing 18 by 6 gives a quotient of 3 and a remainder of 0 ($18 \equiv 0 \pmod{6}$; $18 \bmod 6 = 0$).

Proof: Let $q = \lfloor a/d \rfloor$ and define $r = a - q \cdot d$.

- So $a = q \cdot d + r$ with $q \in Z$ and $0 \leq r < d$ (since $q \cdot d \leq a$).

But why are q and d unique?

- Suppose $q \cdot d + r = q' \cdot d + r'$ with $q', r' \in Z$ and $0 \leq r' < d$.
- Then $(q' - q)d = (r - r')$ with $-d < r - r' < d$.
- The lhs is divisible by d so $r = r'$ and we're done.

Primes

- If $p \in N$, $p > 1$ is *prime* if its only positive factors are 1 and p.
- $n \in N$ is *composite* if $n > 1$ and n is not prime.
  - If n is composite then $a \mid n$ for some $a \in N$ with $1 < a < n$
  - Can assume that $a \leq \sqrt{n}$.
    - **Proof:** By contradiction:
      Suppose $n = bc$, $b > \sqrt{n}$, $c > \sqrt{n}$. But then $bc > n$, a contradiction.

Primes: $2, 3, 5, 7, 11, 13, \ldots$
Composites: $4, 6, 8, 9, \ldots$