15.082J & 6.855J & ESD.78J

Algorithm Analysis

15.082

- Overview of subject
- Importance of Algorithm Analysis
- Importance of homework
- Midterms
- Moving forward

Overview of lecture

Proof techniques

- Proof by contradiction
- Mathematical Induction
- Smallest counterexample

Algorithm Analysis

- Quick review of proving correctness
- Euclid’s algorithm for computing gcd
- Proof that bfs gives shortest distances

What is a proof?

From Wikipedia: In mathematics, a proof is a convincing demonstration (*within the accepted standards of the field*) that some mathematical statement is necessarily true.

Proofs are obtained from deductive reasoning, rather than from empirical arguments. That is, a proof must demonstrate that a statement is true in all cases, *without a single exception.*

Proof by contradiction

From Wikipedia: A proof by contradiction is a form of proof that establishes the truth or validity of a proposition by showing that the proposition being false would imply a contradiction.

Since … a proposition must be either true or false, and its falsity has been shown impossible, the proposition must be true.

Proof by Contradiction: √2 is irrational

Defn: a rational number can be expressed as a/b, where a and b are integers.

Theorem: √2 is irrational.

Proof. Suppose that $\sqrt{2} = a/b$, where a and b are integers. Suppose further that a/b have no common divisors greater than 1.

Then $2 = a^2/b^2$, and $a^2 = 2\ b^2$.

Since a^2 is even, it follows that a is even.

Suppose $a = 2c$. Then

$$(2c)^2 = 2\ b^2 \Rightarrow 4c^2 = 2b^2 \Rightarrow 2c^2 = b^2 \Rightarrow b \text{ is even.}$$

But this contradicts that a and b have no common divisors. So, √2 is irrational. QED

Theorem: There are infinitely many prime numbers.

Fact that will be used in the proof: Any integer that is not a prime number has a factor that is a prime number.

Proof of theorem (by contradiction). Suppose that there are finitely many primes.

Let $\{p_1, p_2, \ldots, p_n\}$ be the set of prime numbers.

Let $x = 1 + (p_1 \times p_2 \times \ldots \times p_n)$.

Then p_j is not a divisor of x for $j = 1$ to n, which contradicts the fact given above.

Therefore there are infinitely many primes.

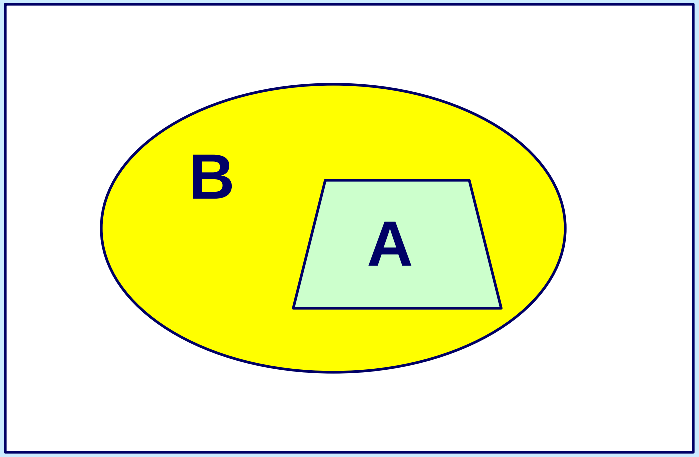
Contrapositive

The contrapositive of "A ⇒ B" is "¬B ⇒ ¬A".

The contrapositive is logically equivalent to the original statement.

Example: All monkeys are mammals

Contrapositive: If something is not a mammal, it is not a monkey.



Example: If a point is in A, it is in B.

Contrapositive: If a point is not in B, it is not in A.

Proof by induction

Think first about dominos.   Suppose 86 dominos are lined up so that if domino i falls, then domino i+1 will fall. Suppose also that domino 1 falls. Will domino 86 fall?

Proof by induction

Mathematical induction is a method of mathematical proof typically used to establish that a given statement is true of all natural numbers. It is done by proving that the first statement in the infinite sequence of statements is true, and then proving that if any one statement in the infinite sequence of statements is true, then so is the next one.

Theorem: every tree with n nodes has n-1 arcs.

Facts that we will use in the proof.

1. Every tree with at least two nodes has a node with exactly one incident arc.
2. Deleting a node with degree 1 from a tree, results in a tree.

Step 1. Establish the basis of the induction.

It is easy to see that trees with two nodes have one arc.

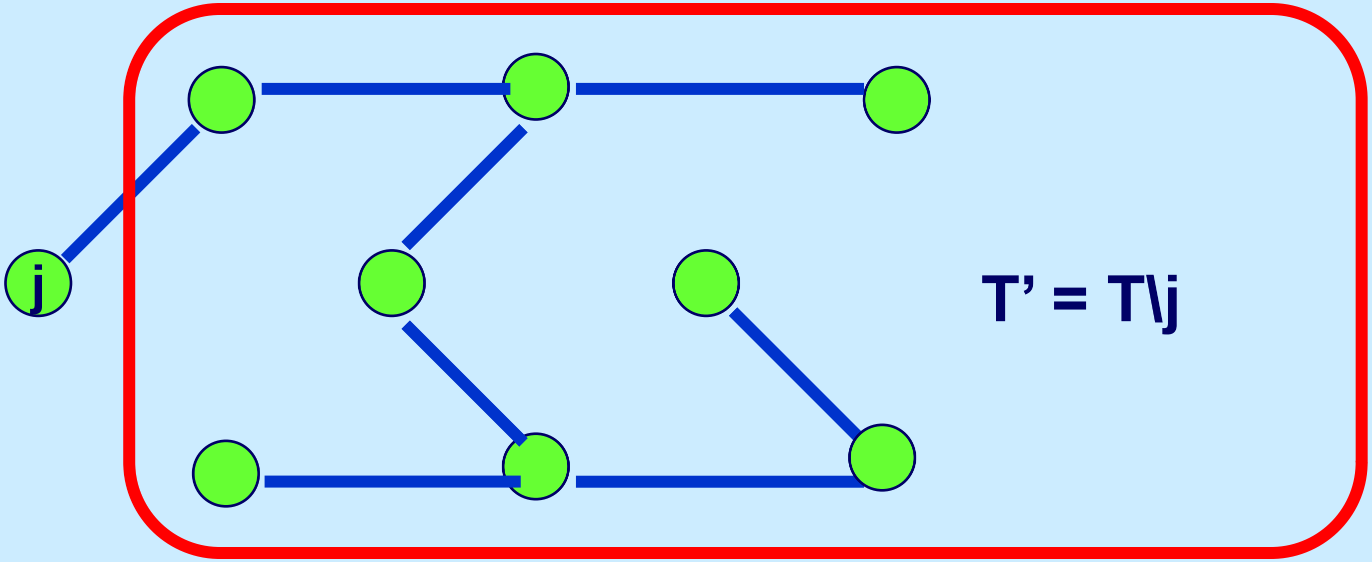
Step 2. Inductive Step

We need to show that the conclusion is true for all trees of k+1 nodes if it is true for all trees of k nodes.

Assume the theorem is true if there are at most k nodes.

Let T be any tree with k+1 nodes.

Let node j be a node with degree 1.



T' has k nodes. By the inductive hypothesis T' has k-1 arcs. T has one more node and one more arc than T'. So, T' has k+1 nodes and k arcs. QED

Proofs by minimum counterexample

This technique combines a proof by contradiction with a proof by induction. The induction works in the opposite direction, from k to k-1.

Step 1. Assume that the theorem is false and that there is a counterexample.

Step 2. Show that there is no counterexample with n = 1 (or 2).

Step 3. Show that if there is a counterexample with n = k, then there is a counterexample with n < k. This shows that there can be no smallest counterexample, and thus no counterexample.

Proof by Minimum Counterexample: every tree with n nodes has n-1 arcs.

Facts that we will use in the proof.

1. Every tree with at least two nodes has a node with exactly one incident arc.
2. Deleting a node with degree 1 from a tree, results in a tree.

Step 1. Establish that the theorem is true for n = 1 or 2.

It is easy to see that trees with two nodes have one arc.

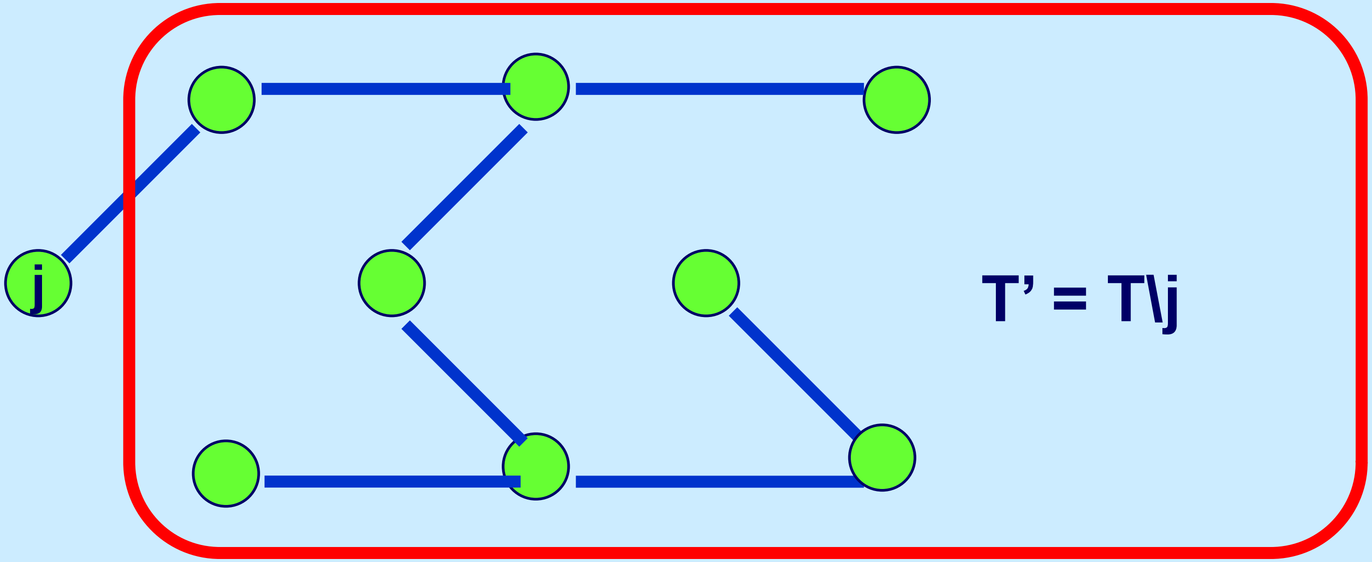
Step 2. Counterexample step.

We need to show that the conclusion is false for some tree of k-1 nodes if it is false for some tree of k nodes.

Let T be the minimal counterexample.

Let T be the counterexample with k nodes

Let node j be a node of T with degree 1.



By assumption, T has k nodes but does not have k-1 arcs. T' has k-1 nodes and one fewer arc than T. Thus T' does not have k-2 arcs, and is a smaller counterexample than T. QED

Mental Break

How many copies of *Moby Dick* were sold during Herman Melville's lifetime.

50 copies

There was an event that occurred the same day as the birth and the death Samuel Clemens (Mark Twain). What was it.

The appearance of Halley's comet.

What famous author coined the word "nerd"

Dr. Seuss, in *If I Ran the Circus*.

Mental Break

In how many of the Sherlock Holmes books by Sir Arthur Conan Doyle did Holmes say “Elementary, my dear Watson”?

None.

Ernest Vincent Wright wrote *Gadsby*, which is 50,000 words long. What is *Gadsby* most famous for?

It does not contain the letter “e”.

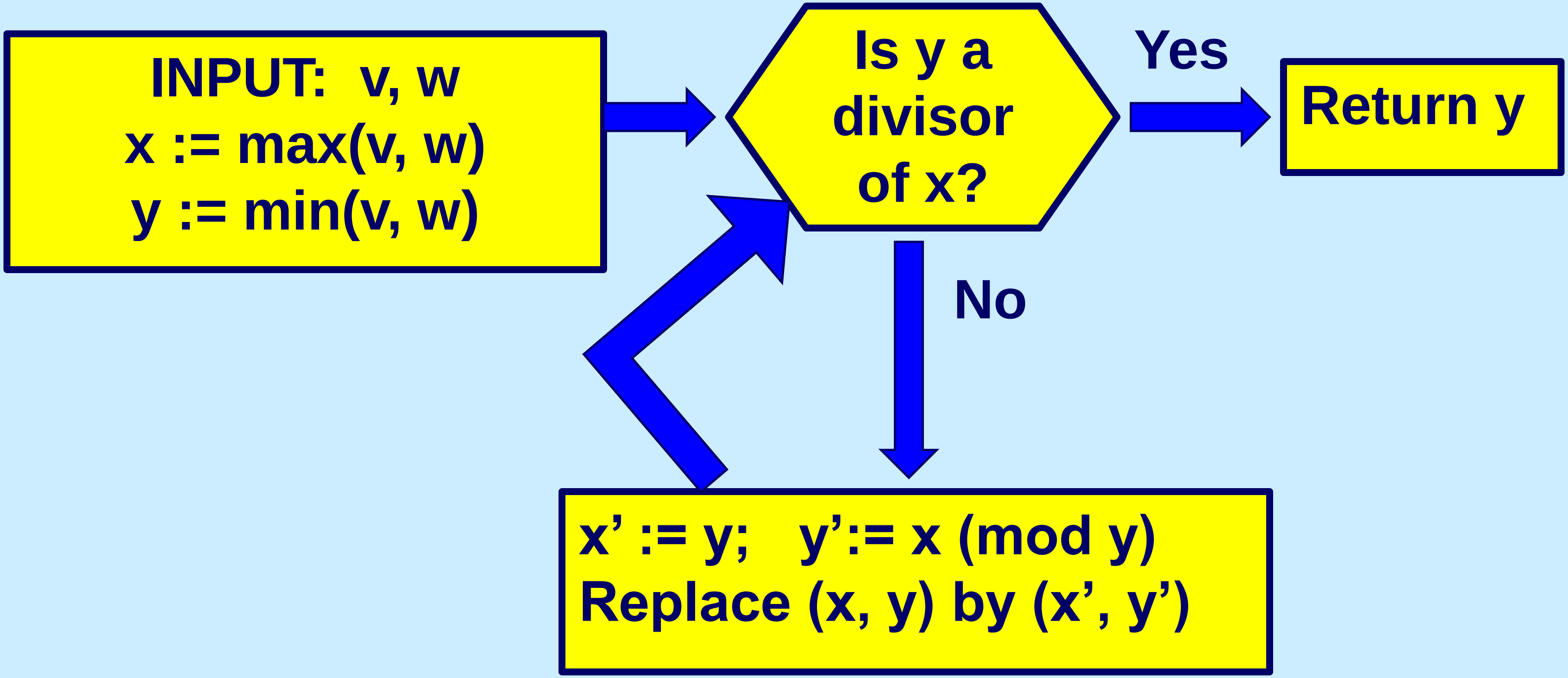
According to the Bible, which came first, the chicken or the egg.

The chicken. Genesis 1:20 – 22.

Proving an algorithm is correct

Example: Euclid's Algorithm for determining the greatest common divisor of two integers, v and w.

e.g., gcd(15, 25) = 5.



INPUT: v, w
x := max(v, w)
y := min(v, w)

Is y a divisor of x?

Yes

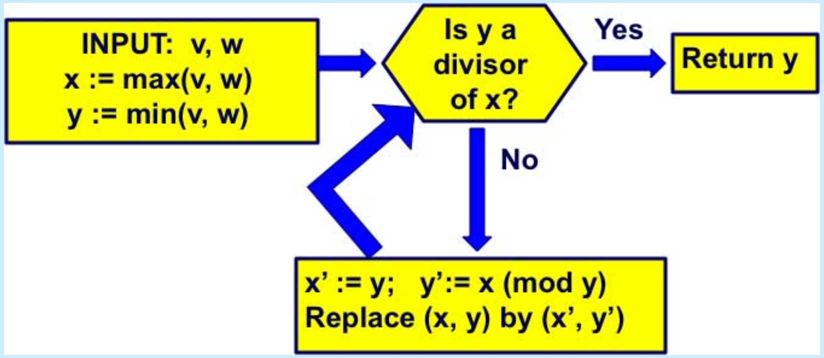
Return y

No

x' := y; y':= x (mod y)
Replace (x, y) by (x', y')

Proof that Euclid's Algorithm is correct

loop



Invariant: Every time that the control enters "Is y a divisor of x?", gcd(x, y) = gcd(v, w).

Let x_n and y_n be the values of x and y the n-th time that control enters loop.

Basis of induction: for n = 1, either x = v and y = w, or x = w and y = v. Thus the invariant is true for n = 1.

Inductive Step

Inductive step: assume that the result is true at the k-th iteration. We will show that d is a divisor of x_k and y_k if and only if it is also a divisor of x_{k+1}, y_{k+1}.

Suppose first that d is a factor of x_k and y_k.

$x_{k+1} = y_k \quad \Rightarrow \quad$ d is a divisor of x_{k+1}.

$y_{k+1} = x_k \bmod y_k \quad \Rightarrow \quad x_k = y_{k+1} + b\, y_k$ for some integer b

$\Rightarrow \quad y_{k+1} = x_k - b\, y_k$

Since d is a divisor of x_k and y_k, it is also a divisor of y_{k+1}.

Conversely, suppose that d is a divisor of x_{k+1} and y_{k+1}.

Then it is a divisor of y_k (= x_{k+1}).

$x_k = y_{k+1} + b\, y_k \Rightarrow x_k = y_{k+1} + b\, y_k = y_{k+1} + b\, x_{k+1} \Rightarrow$

d is a divisor of x_k.

Proof of Finiteness

$x_{k+1} = y_k$.

$y_{k+1} = x_k \bmod y_k \Rightarrow$

$y_{k+1} \leq x_k - y_k$ and $y_{k+1} \leq y_k$

$\Rightarrow \quad 2\, y_{k+1} \leq x_k$

$\Rightarrow \quad 2x_{k+2} \leq x_k$.

So, x is decreasing by a factor of at least 2 every two iterations, and the number of iterations is O(log v).

Finally, if y is a divisor of x, then gcd(x, y) = y. So, the algorithm ends with gcd(v, w).

Breadth First Search

```
        Initialize
loop    while LIST ≠ ø do
            select the first node i in LIST;
            if node i is incident to an admissible arc (i,j) then
                mark node j;
                pred(j) := i;
                add node j to the end of LIST;
            else delete node i from LIST
```

Exercise 3.30. Show that in a bfs tree, the tree path from the source node to any other node i is a shortest path.

Remark: this is a more challenging exercise than is indicated in the book.

Proving the key invariant

Let d(i) be the shortest path distance from node 1 to node i (min number of arcs).

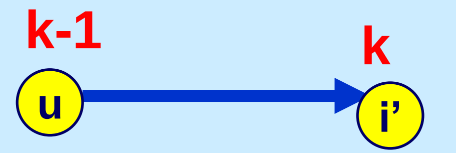
Claim (to be proved by induction): If d(i) < d(j), then node i is added to LIST (and its arcs are scanned) prior to node j.

Basis of induction. It is true if d(i) = 0. The only node with distance 0 is node 1. And node 1 is the first node added to LIST and the first one scanned.

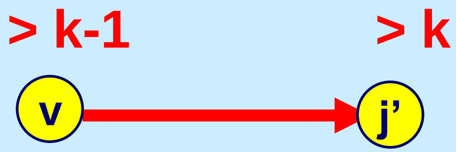
Inductive Step

Suppose the result is true when d(i) = k-1.

Suppose that d(i') = k < d(j').



Let u be the node that precedes i' on some shortest path to node i'.



Let v = pred(j') in the bfs tree.
Then d(v) ≥ k-1. Otherwise, d(j') ≤ k.

By hypothesis, u is added to LIST prior to v.
Node i' is added to LIST when (u,i') is scanned or possibly earlier, which is before (v, j') is scanned. So i' is added to LIST prior to j'.

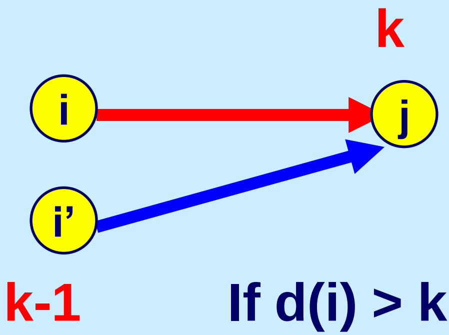
Exercise 3.30. Show that in a bfs tree, the tree path from the source node to any other node i is a shortest path.

Proof by induction.

We may use the fact that whenever d(i) < d(j), then node i is added to LIST prior to node j.

Basis of induction. It is true if d(i) = 0, and thus i = 1.

Inductive step. Assume it is true if the length of the path is at most k-1.



Let i = pred(j).

Let i' be the node that precedes j on the shortest path from node 1.

If d(i) > k-1, then i is added to LIST subsequent to i', and (i', j) is scanned prior to (i, j), a contradiction.

If d(i) = k-1, then by induction, the tree path from node 1 to node i has k – 1 arcs, and the tree path to node j has k arcs. QED

Summary

Proofs are "convincing arguments."

- Proof by contradiction
- Proof by induction
- Proof by minimum counterexample

Although proofs are usually direct arguments, it does not mean that there is a direct method for finding the proof.

- Often finding a proof relies on trial by error

Proving Algorithm correctness

- Relies on guessing the "right" invariants and proving them by induction.