Math 4380, Fall 2013: Answers to quizzes

Paul Hewitt

22 November 2013

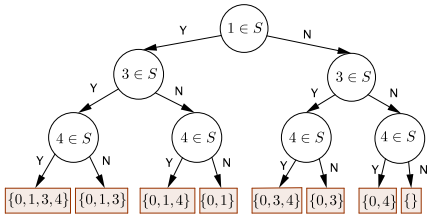
Quiz 1.

a. Give precise definitions for the following:

$A \subseteq B$ is the statement "if $x \in A$ then $x \in B$"
$A \cup B$ is the set $\{x \mid x \in A \text{ or } x \in B\}$
$A \cap B$ is the set $\{x \mid x \in A \text{ and } x \in B\}$

b. Use a decision tree to enumerate all subsets of $\{0, 1, 2, 3, 4\}$ that contain 0 but do not contain 2.

We follow the example on page 10 of our text. The three decisions that determine a set S satisfying the given conditions are: Is $1 \in S$? Is $3 \in S$? Is $4 \in S$?



Quiz 2.

a. Stirling's Formula says that

$$n! \sim (n/e)^n \cdot \sqrt{2\pi n}.$$

What does the $\sim$ mean?

To say that $a_n \sim b_n$ is to say that

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

b. If n is a positive integer then what is

$$1 + 3 + \cdots + (2n-1) = ?$$

Use mathematical induction to prove your claim.

The sum equals n^2, as we may prove by induction. If $n = 1$ then indeed $1 = 1^2$. If $n > 1$ then by induction we see that

$$\begin{aligned} 1 + 3 + \cdots + (2n-1) &= 1 + 3 + \cdots + (2(n-1)-1) + (2n-1) \\ &= (n-1)^2 + 2n - 1 \\ &= n^2 \end{aligned}$$

as claimed.

Quiz 3.

a. What is the FISA court?

b. When was the FISA court created?

c. Why was the FISA court created?

This is the court established by the Foreign Intelligence Surveillance Act of 1978. It was one of the recommendations of the committee chaired by Senator Frank Church. His committee investigated abuses and illegal actions of various intelligence agencies during the Cold War, especially under the Nixon administration. The court allows the NSA, CIA, FBI and others to obtain warrants for surveillance of foreign agents (including US citizens suspected of aiding foreign governments) without revealing secrets to the public, as might happen in other Federal courts.

d. What is the NSA?

The National Security Agency is the arm of the Defense Intelligence Agency that collects "signals intelligence". In other words it spies on electronic communication, including phone calls and email. Unlike the CIA it does not send out spies to collect information or carry out operations in person.

e. What did Edward Snowden do at the NSA?

Edward Snowden was a computer system administrator for a company under contract with the NSA. He took a large number of classified documents relating to the various NSA operations and has released some of these to the press.

Quiz 4.

a. State the Binomial Theorem. Who discovered it?

The Binomial Theorem states that if n is a nonnegative integer then

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

It seems to have been first stated (but not in the form above!) by the Persian poet and mathematician Omar Khayyam, who lived in the late 11th and early 12th centuries.

b. How many anagrams can you make from the words SQUEAMISH OSSIFRAGE? (Ignore the space.)

Let's start with a letter count:

$$\begin{aligned} A &: 2 \\ E &: 2 \\ F &: 1 \\ G &: 1 \\ H &: 1 \\ I &: 2 \\ M &: 1 \\ O &: 1 \\ Q &: 1 \\ R &: 1 \\ S &: 4 \\ U &: 1 \\ total &: 18 \end{aligned}$$

Therefore, there are exactly

$$\frac{18!}{2! \cdot 2! \cdot 2! \cdot 4!} = 33345696384000$$

distinct anagrams.

Quiz 5.

a. What does

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2$$

equal? Prove your claim.

I claim that

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}.$$

There are (at least) three proofs discussed in the book. One uses the symmetry of the binomial coefficients to write the left-hand side of this equation as

$$\binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \cdots + \binom{n}{n}\binom{n}{0}.$$

Combinatorially this can be interpreted as counting how many ways to choose n balls from a set of n black and n white balls: you could choose n white and 0 black; or $n-1$ white and 1 black; or ... 0 black and n white. Or you cmld simply throw all the balls in one big bin and choose n of them.

You could also prove this by induction, using the identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Or you could prove this using the binomial formula applied to both sides of the identity

$$(x+y)^n(x+y)^n = (x+y)^{2n}.$$

b. Suppose that n is a very large positive integer. Consider a bar chart of the binomial coefficients $\binom{2n}{k}$, for $0 \le k \le 2n$. If we zoom out then what does this graph look like? Be precise!

There is a nice picture on page 58 of our text. Suitably normalized the bar chart looks like a *gaussian* — or *normal* or "bell-shaped" — distribution, with mean n. More precisely if we divide by the largest binomial coefficient (namely $\binom{2n}{n}$ then we find that

$$\binom{2n}{n-t} \Big/ \binom{2n}{n} \approx \exp(-t^2/n).$$

In fact, the book provides the error estimate

$$\exp(-t^2/(n-t+1)) \le \binom{2n}{n-t} \Big/ \binom{2n}{n} \le \exp(-t^2/(n+t)).$$

As they point out this is a rather crude one.

Quiz 6.

a. Let F_n denote the n-th Fibonacci number. Prove that

$$F_n = \frac{q_1^n - q_2^n}{\sqrt{5}},$$

where q_1 and q_2 are the two roots of the quadratic equation

$$q^2 = q + 1,$$

and $q_1 > q_2$.

First of all, the two roots are $\frac{1}{2}(1 \pm \sqrt{5})$, as we can determine from the quadratic formula. When $n = 0$ the equation above yields $F_0 = (1-1)/\sqrt{5}$, which is true. When $n = 1$ the equation above yields

$$F_1 = \frac{\frac{1}{2}(1+\sqrt{5}) - \frac{1}{2}(1-\sqrt{5})}{\sqrt{5}} = \frac{\sqrt{5}}{\sqrt{5}},$$

which is also true. Now we verify that the right-hand side of the equation satisfies the defining identify for the Fibonacci sequence:

$$\begin{aligned}\frac{q_1^n - q_2^n}{\sqrt{5}} + \frac{q_1^{n+1} - q_2^{n+1}}{\sqrt{5}} &= \frac{q_1^n(1+q_1) - q_2^n(1+q_2)}{\sqrt{5}} \\ &= \frac{q_1^{n+2} - q_2^{n+2}}{\sqrt{5}},\end{aligned}$$

as required.

b. What is the sum of the first n Fibonacci numbers? Prove your claim.

I claim that

$$F_0 + F_1 + \cdots + F_{n-1} = F_{n+1} - 1.$$

To prove this we check that the above is true when $n = 1$:

$$F_0 = 1 - 1 = F_2 - 1.$$

Now we assume that $n > 1$ and proceed by induction:

$$\begin{aligned}F_0 + F_1 + \cdots + F_{n-1} &= (F_0 + F_1 + \cdots + F_{n-2}) + F_{n-1} \\ &= (F_n - 1) + F_{n-1} \\ &= F_{n+1} - 1.\end{aligned}$$

Quiz 7.

a. What is the method of *simultaneous induction*? Give an example of statements that can be proved with this method. (You do not have to provide the proof.)

Simultaneous induction is used to prove *two* sequences of statements A_n and B_n *simultaneously* by induction, using a scheme such as the following:

$$\begin{aligned} A_n \text{ and } B_n &\implies A_{n+1} \\ A_{n+1} \text{ and } B_n &\implies B_{n+1} \end{aligned}$$

After establishing base cases for both sequences we conclude the truth of all A_n and B_n.

The book uses this method to prove the following two statements by simultaneous induction:

$$\begin{aligned} A_n &: F_n^2 + F_{n-1}^2 = F_{2n-1} \\ B_n &: F_{n+1}F_n + F_nF_{n-1} = F_{2n} \end{aligned}$$

b. Suppose that we start with any two values A and B and define a sequence c_n recursively with the rules

$$c_0 = A,\ c_1 = B,\ c_n = c_{n-1} + c_{n-2} \text{ when } n > 1.$$

How is c_n related to the Fibonacci sequence? Be precise!

$$c_n = AF_{n-1} + BF_n \text{ when } n > 0.$$

Quiz 8.

a. What is a probability space?

A probability space is the set — usually called the sample space — of all possible outcomes of an experiment together with the assignment of a nonnegative probability $P(x)$ to each outcome x in such a way that the sum of the probabilities of all possible outcomes equals 1.

b. What is an event?

An event is simply a subset of a sample space.

c. What does it mean to say that two events are exclusive?

We say that A and B are exclusive when $A \cap B = \emptyset$.

d. What does it mean to say that two events are independent?

We say that A and B are independent when $P(A \cap B) = P(A)P(B)$. Note: if E is an event then we define $P(E)$ to be the sum of all probabilities $P(x)$ for all x in A.

e. What is the Law of Large Numbers?

The Law of Large Numbers can be stated in various ways. One way echoes the Binomial Theorem:

> *If* $0 \leq t \leq m$ *then the probability that out of* $2m$ *coin tosses the number of heads differs from* m *by more than* t *(in absolute value) is less than* $e^{-t^2/(m+t)}$.

Quiz 9.

True or false:

a. If a and b are integers and $a > 0$ then there are unique integers q and r such that $b = aq + r$.

This is *false:* to guarantee that q and r are unique we require that r lie in an interval of length less than a — for example, $0 \leq r < a$.

b. The fact that there are infinitely many primes was not proven rigorously until the mid-nineteenth century.

This is *false:* there is a rigorous proof in Euclid's *Elements*, although the proof may date back centuries before that.

c. Every positive integer can be written as a product of primes, and this factorization is unique up to the order of the prime factors.

This is *true:* it is called the Fundamental Theorem of Arithmetic, and there is a rigorous proof of this in Euclid's *Elements*.

d. The number $\sqrt{2}$ is irrational.

e. The diagonal of a square is incommensurable with the side.

Both of these are *true.* In fact, they are equivalent. Tradition credits the Pythagoreans with both their discovery and proof, and proofs appear in Euclid's *Elements*.

Quiz 10.

a. State the Prime Number Theorem.

If $\pi(n)$ denotes the number of primes less than or equal to n then

$$\pi(n) \sim \frac{n}{\log n}.$$

b. Who first proved the Prime Number Theorem? When was it first proved?

The Prime Number Theorem was first conjectured by Gauss in the late 18th century. It was proved by Jacques Hadamard and Charles de la Vallée Poussin (independently and almost simultaneously) about 100 years later. Their proof was based on the pioneering work of Gauss' most famous student, Bernhard Riemann.

c. State Fermat's "Little Theorem".

If p is a prime and n is any positive integer then $p \mid n^p - n$, or equivalently $n^p \equiv n \bmod p$.

d. State Fermat's "Last Theorem".

If $n > 2$ then there is no solution to the equation $x^n + y^n = z^n$ in positive integers x, y, z.

e. Who first proved Fermat's "Last Theorem"? When was it first proved?

Pierre de Fermat is reported to have briefly believed that he had proven this statement, but in fact he never announced this publicly, and may have quickly realized his mistake. It was finally proved by Andrew Wiles, with some assistance from his student Robert Taylor, in 1995.

Quiz 11.

a. Prove that if m is a positive integer and $a \equiv b \bmod m$ then for any integer c we have $a + c \equiv b + c \bmod m$.

To say that $x \equiv y \bmod m$ is to say that $x - y$ is a multiple of m. So if $a \equiv b \bmod m$ then $(a + c) - (b + c) = a - b$ is a multiple of m. That is, $a + c \equiv b + c \bmod m$.

b. Use anthyphairesis to find an integer s such that $21s \equiv 1 \bmod 65$.

Anthyphairesis yields

$$\begin{aligned} 65 &= 3 \cdot 21 + 2 \\ 21 &= 10 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \end{aligned}$$

Hence

$$\begin{aligned} 1 &= 21 - 10 \cdot 2 \\ &= 21 - 10 \cdot (65 - 3 \cdot 21) \\ &\equiv 31 \cdot 21 \bmod 65 \end{aligned}$$

That is, we may take $s = 31$.

Quiz 12.

a. What are Alice's goals?

Over a noisy telephone line, tapped by the tax authorities and the secret police, Alice happily attempts, with someone she doesn't trust, whom she cannot hear clearly, and who is probably someone else, to fiddle her tax returns and to organize a coup d'état, while at the same time minimizing the cost of the phone call.

b. According to Gordon what is a coding theorist?

A coding theorist is someone who doesn't think Alice is crazy

c. What are the three types of coding?

Source coding, channel coding and secrecy coding

d. Which type of coding does Alice use for each of her goals?

Source coding helps keep her phone bills down. Channel coding helps overcome the noisy line. Secrecy coding helps protect from her from the tax authorities, the secret police, and from Bob.

e. Besides Alice and Bob what are coding theorists interested in?

Information.

Quiz 13.

a. What is a monoalphabetic substitution cipher?

This is a cipher in which each letter is replaced by the same letter or symbol at each occurrence. Sometimes nulls and symbols for common words are added, creating a *nomenclator*.

b. What technique renders such ciphers vulnerable?

Frequency analysis: counting the occurrence of each letter in the ciphertext and then comparing this to the known frequency distribution for letters in the language of the plaintext.

c. Where, when, and why did this technique arise?

It was invented by Arab scholars in the early Islamic Empire (7th and 8th centuries) in order to establish the chronology of the suras in the *Quran* and the accuracy of putative *Hadith*. They later applied the technique to encrypted messages, thereby inventing cryptanalysis.

d. What kind of cipher is the Viginère cipher?

It is a polyalphabetic substitution cipher, meaning that the substitution of one letter for another depends not only on the plaintext letter but also its position in the plaintext. In the Viginère cipher each letter in a keyword determines a cyclic shift applied to all positions in the plaintext equivalent to the position of the letter modulo the keylength.

e. Is the Viginère cipher susceptible to the same technique? Explain!

Not directly, but after a preliminary step. This technique was discovered by Charles Babbage. His idea was to look at the repetition of sequences of letters in the ciphertext and to use the gcd of the spacings between such repetitions to determine the keylength. Once the keylength is known frequency analysis can be applied to all positions in a given equivalence class modulo the keylength. Cyclic shifts are easily deciphered by frequency analysis.

Quiz 14.

a. According to Sun-Tzu what is the most valuable weapon in wartime?

Intelligence.

b. Who exploited traffic analysis to great effect in World War I?

France.

c. A long key is by itself insufficient to guarantee security. What else is needed?

Randomness.

d. What is the great advantage of the one-time pad?

It is provably secure.

e. Cryptanalysts had the advantage at the beginning of World War I. What invention flipped the advantage to cryptographers?

The Enigma machine.

f. What was the overriding problem for cryptographers after World War II?

The key distribution problem.

g. What was the NSA's main contribution to the development of the DES?

To limit the keylength to 56 bits.

h. What was Whitfield Diffie's revolutionary idea?

The asymmetric cipher.

i. Who solved the key distribution problem?

Diffie, Hellman, and Merkle — and independently Ellis, Cocks, and Williamson.

j. What do the following acronyms mean: RSA, PGP, and FBI?

RSA are the initials of Rivest, Adleman, and Shamir; PGP stands for Pretty Good Privacy; and FBI is the Federal Bureau of Investigation.

Quiz 15.

a. What does the algorithm of Agrawal, Kayal, and Saxena do efficiently for the first time in history?

Their algorithm determines whether a given number is prime.

b. What does it mean for an algorithm to be classified as "efficient", in their sense?

It must be deterministic and run in polynomial time with respect to the input size. That is, there must be an integer k so that for input of ℓ bits the algorithm return the correct answer in time equal to $O(\ell^k)$.

c. Is the Sieve of Eratosthenes efficient? Explain.

It is determinisitc but runs in $\Omega(\sqrt{n})$ time, which is exponential in the input size $\log n$.

d. On what number theoretic fact are the algorithms of Miller and Rabin based?

Fermat's Little Theorem *and its converse*, in the following form:

$$p \text{ is prime if and only if } \gcd(a,p) = 1 \implies a^{p-1} \equiv 1 \bmod p.$$

e. What extra ingredient would be needed to make the Miller-Rabin and Solovay-Strassen algorithms both efficient and deterministic?

The would run in polynomial time and be deterministic (as opposed to randomized) if the Extended Riemann Hypothesis were true.

Quiz 16.

In this quiz we are considering the group law on a cubic curve $y^2 = f(x)$.

a. What is the additive identity for this group law?

The additive identity O is the point at infinity.

b. What is the additive inverse of a point P?

The additive inverse $-P$ is the reflection of P across the x-axis.

c. Suppose P and Q are distinct points on this cubic. Describe how to find $P + Q$ geometrically.

If R is the third point of intersection of this cubic with the line $\overline{PQ}$ then $P + Q$ is the reflection of R across the x-axis.

d. Suppose P is any points on this cubic. Describe how to find $P + P$ geometrically.

This is the same procedure as above except that we begin with the tangent line to the curve at P.

e. Now suppose the P is an x-intercept of this cubic. What is $P + P$ in this case?

In this case $P + P = O$, the point at infinity.

Quiz 17.

a. What are the two things that comprise a graph?

A graph consists of a set of nodes and a set of edges. Each edge connects a pair of nodes. In a multigraph a pair of nodes may be joined by multiple edges. In a simple graph there is at most one edge between any pair of nodes and no edge connects a node to itself.

b. What are the neighbors of a node?

The neighbors of a node are all those nodes connected to it by an edge.

c. What is the degree of a node?

The degree of a node is the number of edges attached to the node. One must take care when discussing graphs that are not simple. For example in directed graphs one may want to distinguish between the *out-degree* — the number of edges emanating from the node — from the *in-degree* — the number of edges heading into a node.

d. What is a path in a graph?

A path of length n is a subgraph of distinct nodes $v_0, \ldots, v_n$ such that consecutive nodes v_i and v_{i+1} are neighbors. This is distinguished from a *walk*, which is simply any sequence of nodes $v_0, \ldots, v_n$ such that consecutive nodes v_i and v_{i+1} are neighbors. A walk is a path if no node is repeated in the sequence.

e. What is a connected graph?

A connected graph is one where any pair of nodes can be connected by a walk, or equivalently by a path.

Quiz 18.

This quiz concerns the dictionary

```
D = {
  'a': 1,
  1: 2,
  'b': 'c',
  3: 'c',
}
```

a. What does `D[1]` return?

`2`

b. What is the result of the command `D[1] = 3`?

It updates `D` so that the new value for `D[1]` is `3`.

c. What is the result of the command `del D[1]`?

It updates `D` so that there is no longer an entry with key `1`.

d. What is the result of the command `D[0] = 1`?

It updates `D` so that there is a new key 0 with value 1.

e. What is the value of the statement `'c' in D`?

`False`

f. What is the python syntax for obtaining all the keys of `D`?

`D.keys()`

g. What does this return?

If we apply this to the original `D` then we would obtain the list with the values `'a',1,'b',3` *in some order* — we cannot be certain of the ordering of the entries.

h. What is the python syntax for obtaining all the values of `D`?

`D.values()`

i. What does this return?

If we apply this to the original `D` then we would obtain the list with the values `1,2,'c','c'` *in some order* — we cannot be certain of the ordering of the entries.

j. What is the result of the command { `x: x+x for x in D` }?

A new dictionary equal to {`'a': 'aa', 1: 2, 'b': 'bb', 3: 6`}.

Quiz 19.

a. What are two ways to characterize a tree?

A tree is:
a connected graph which contains no cycles
or equivalently
a graph without cycles but in which the addition of any edge creates a cycle
or equivalently
a connected graph from which the deletion of any edge disconnects the graph.

b. What is a spanning tree in a graph G?

A spanning tree in a graph G is a subgraph which is a tree and contains all the nodes of G.

c. What is a leaf in a rooted tree?

A leaf in a rooted tree is a node other than the root that has degree 1.

d. How many edges does a tree with n nodes have?

A tree with n nodes has exactly $n - 1$ edges.

e. How many labeled trees are there on n nodes?

There are exactly n^{n-2} labeled trees on n nodes. This is theorem is due to Cayley.

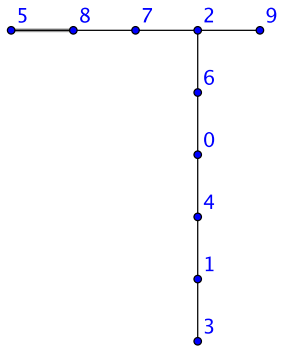
Quiz 20.

a. What is the labeled tree with Prüfer code $(1, 4, 0, 8, 7, 2, 2, 6)$?

The extended Prüfer code adds a 0 at the end of this row and then determines each entry in a row above this as the smallest label not appearing to the left, below, or in the lower row and to the right:

3	1	4	5	8	7	9	2	6
1	4	0	8	7	2	2	6	0

Hence the graph looks like this:



b. What is the planar code of the tree in part *a*?

If we start walking from node 0 always keeping the tree to our right then we walk the following loop:

$$0, 6, 2, 7, 8, 5, 8, 7, 2, 9, 2, 6, 0, 4, 1, 3, 1, 4, 0$$

For each of these 18 steps we record 0 or 1 depending on whether we step away from or back towards the root:

$$111110001000111000$$