The PRIMES 2012 problem set

Dear PRIMES applicant,

This is the PRIMES 2012 problem set. Please send us your solutions at primes@math.mit.edu by December 1, 2011.

Note that there is a collection of problems called "General math problems", as well as collections corresponding to the three tracks of PRIMES 2012 – "Advanced math", "Computer science", and "Computational biology". Please solve as many problems as you can in the General math section, and also in the sections corresponding to the tracks for which you are applying.

You can type the solutions or write them by hand and then scan them; please save your work as a DOC, PDF, or JPG file.

Please write not only answers, but also proofs (and partial solutions/results/ideas if you cannot completely solve the problem). Besides the admission process, your solutions will be used to decide which projects would be most suitable for you if you are accepted to PRIMES.

You are allowed to use any resources to solve these problems, *except other people's help*. This means that you can use calculators, computers, books, and the Internet. However, if you consult books or Internet sites, please give us a reference.

Note that some of these problems are tricky. We recommend that you do not leave them for the last day, and think about them, on and off, over some time (several days). We encourage you to apply if you can solve at least 50% of the problems. [1]

Enjoy!

[1] We note, however, that there will be many factors in the admission decision besides your solutions of these problems.

General math problems

Problem G1. You draw 4 cards from the regular deck of 52 cards.

(a) What is the chance that all of these cards have different denominations (i.e., values)? Represent the answer as a fraction or a decimal up to the third digit.

(b) What is the chance all of these cards have different denominations, and in addition there is no neighbors (for example an 8 and a 9, a 10 and a Jack, or a queen and a king are neighbors)?

Solution. (a) We assume the cards are labeled by $1, 2, 3, 4$. There are $13 \cdot 12 \cdot 11 \cdot 10$ ways to choose the denominations. Once that is done, there are 4^4 variants. The total number of ways to choose is $52 \cdot 51 \cdot 50 \cdot 49$. So the chance is $12 \cdot 11 \cdot 10 \cdot 64/51 \cdot 50 \cdot 49 = 2816/4165$.

(b) Sets of 4 denominations with no neighbors correspond to partitions of 9 into 5 ordered parts out of which all except first and last are ≥ 1. So this is the same as partitions of 6 in 5 ordered parts, or of 11 in 5 positive parts. So we get $\binom{10}{4}$ ways, and the answer is

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 4^4/52 \cdot 51 \cdot 50 \cdot 49 = 1536/7735.$$

Problem G2. Find the remainder of division of $5^{5^{555}}$ (i.e., 5 to the power 5^{555}) by 27.

Solution. Remainders of powers of 5 are periodic with period $\phi(27) = 3^3(3-1) = 18$, so we need to find the remainder of 5^{555} under division by 18. Remainders under division by 18 are periodic with period $\phi(18) = 6$. Since 555 is 3 mod 6, the remainder is the same as for 5^3, which is 17. Thus, the remainder mod 27 of the number in question is the same as $5^{17} = 5^{-1}$, which is 11.

Problem G3. Count geometrically different (i.e., inequivalent under rotation) colorings in red and blue of the faces of

(a) a cube

(b) a regular octahedron;

Answer: (a) 10 (b) 23.

Problem G4. One chooses at random an integer $1 \leq N < 10^{100}$ (with equal probability for all choices).

(a) What is the chance (to the third digit precision) that the leading (leftmost) digit of N^2 is 1? What is the chance that this digit is 9? Are they equal to each other?

(b) What are the exact values of these probabilities in the limit when 10^{100} is replaced by 10^k when k grows indefinitely?

Solution. The leading digit of N^2 is 1 if N is between $10^{m/2}$ and $\sqrt{2} \cdot 10^{m/2}$ for some m. So the probability is about $\frac{\sqrt{2}-1}{\sqrt{10}-1}$, which is about 0.192. The chance that this digit is 9 is about $\frac{\sqrt{10}-3}{\sqrt{10}-1}$, which is about 0.075. So the probability of 1 is much greater.

Problem G5. (a) Show that the number $\sum_{n=0}^{\infty} \frac{1}{2^{n^2}}$ is irrational.

(b) Describe all strictly increasing sequences of nonnegative integers $b_0 < b_1 < ...$ for which

$$\sum_{n=0}^{\infty} \frac{1}{2^{b_n}}$$

is a rational number.

Solution. The binary expansion has to be periodic starting from some place, so the sequence $b_{n+1} - b_n$ should be periodic starting from some place.

Advanced math problems

Problem M1. (a) Find the monic polynomial $P(x)$ with integer coefficients of smallest degree, such that

$$P(\sqrt{2}+\sqrt{3}+\sqrt{6})=0.$$

(b) Let p,q,r be three distinct primes. Find the monic polynomial $P(x)$ with integer coefficients of smallest degree, such that

$$P(\sqrt{p}+\sqrt{q}+\sqrt{r})=0.$$

Solution. Suppose $x=\sqrt{a}+\sqrt{b}+k\sqrt{ab}$. Then

$$(x-\sqrt{a}-\sqrt{b}-k\sqrt{ab})(x+\sqrt{a}-\sqrt{b}+k\sqrt{ab})=(x-\sqrt{b})^2-a(1+k\sqrt{b})^2=$$

$$(x^2+b-a-k^2ab)-2\sqrt{b}(x+ak)$$

So the equation for x is

$$(x^2+b-a-k^2ab)^2-4b(x+ak)^2=0,$$

or

$$x^4-2(a+b+k^2ab)x^2-8abkx+(b-a-k^2ab)^2-4a^2bk^2=0.$$

For (a), we plug in $a=2$, $b=3$, $k=1$, and get

$$x^4-22x^2-48x-23=0.$$

To solve (b), let $x=\sqrt{p}+\sqrt{q}+\sqrt{r}$ and $w=x^2-(p+q+r)=\sqrt{a}+\sqrt{b}+k\sqrt{ab}$ for $a=4pq, b=4pr, k=\frac{1}{p}$. Plugging this in, we get

$$w^4-4(pq+pr+qr)w^2-16pqrw+16(p^2r^2+q^2r^2+p^2q^2-2pqr(p+q+r))=0$$

Problem M2. Let d be a positive integer. Let w be a word in x and y of length d. Let $a_n(w)$ be the number of words in the letters x,y which don't contain w as a subword.

(a) Find the generating function for $a_n(x^{d-1}y)$ (where x^m is x repeated m times). I.e., find the function given by the power series

$$\sum_{n=0}^{\infty}a_nt^n,$$

as a rational function of t.

(b) Show that $a_n(x^d)\neq a_n(x^{d-1}y)$ for some n, and compute the generating function of $a_n(x^d)$ (You may first consider the case $d=2$).

(c) Classify words w of length d for which $a_n(w)=a_n(x^{d-1}y)$ for all n (i.e. describe, as explicitly as you can, what these words are).

Hint. Try to find recursions for $a_n(w)$.

Solution. Let us say that a word w is self-overlapping if $w = ua = bu$ for some shorter word u. For example, x^d for $d \geq 2$ is self-overlapping, and so is $xyxy$, but $x^{d-1}y$ is not self-overlapping. If w is not self-overlapping, then it is clear that $a_n(w)$ satisfies the recursion

$$a_n = 2a_{n-1} - a_{n-d},\ n \geq 1,$$

where $a_i := 0$ for $i < 0$. This means that the generating function for $a_n(w)$ is

$$f_d(t) = \frac{1}{1 - 2t + t^d}$$

On the other hand, words without x^d are words of the form $y^{i_1}x^{j_1}...y^{i_n}x^{j_n}y^{i_{n+1}}$, where $j_k \leq d-1$, so it is easy to see that the generating function for $a_n(x^{d-1})$ is

$$g_d(t) = \frac{1}{1 - t - \frac{t-t^d}{1-t^d}} = \frac{1-t^d}{1-2t+t^{d+1}}.$$

In general, if w is self-overlapping of length d, then $a_n > 2a_{n-1} - a_{n-d}$, since the word wa with missing first letter may contain w as a subword even if a does not contain w as a subword. So $a_n(w) \neq a_n(x^{d-1}y)$. This solves all parts of the problem.

Problem M3. Let A be a matrix 100×100 whose entries are 0 or 1, each chosen randomly by flipping a coin (head=0, tail=1).

(a) What is the chance that the determinant of A is odd? (Compute up to third digit precision).

(b) Let A be an $n \times n$ matrix whose entries are determined by flipping a coin, and p_n be the probability that $\det A$ is odd. What is the limit of p_n as $n \to \infty$?.

Solution. $\det A$ is odd if an only if A is invertible mod 2. The number of such matrices is $(2^n - 1)(2^n - 2)...(2^n - 2^{n-1})$, so the chance that the determinant is odd is

$$p_n = \prod_{k=1}^{n}(1 - 2^{-k}).$$

The limit is thus

$$p_\infty = \prod_{k=1}^{\infty}(1 - 2^{-k}).$$

Problem M4. Let $(a_n)_{n\geq 0}$ be a sequence of nonnegative real numbers such that

$$a_{n+1} \leq \frac{1}{2}(a_n + a_{n-1}). \tag{0.1}$$

Show that $(a_n)_{n\geq 0}$ converges.

Hint: Show that for any $n \geq i-1$, one has $a_n \leq \max(a_i, a_{i-1})$. Set $a := \text{limsup}_{n\to\infty} a_n$, and deduce that one of any two consecutive terms of the sequence must be larger than or equal to a. Now assume that there is a subsequential limit $b < a$, with a_{n_k} converging to b, and show that the inequality (0.1) cannot hold for for large enough k.

Solution. Let us show that for any $n \geq i-1$, one has $a_n \leq \max(a_i, a_{i-1})$ by induction in n. Clearly, this holds for $n = i-1, i$, which provides the base of induction. Assume it holds for $n-2$ and $n-1$. By the inequality (0.1), it also holds for n, so we are done.

Clearly, a_n is bounded. Let $a = \text{limsup}_{n\to\infty}(a_n)$. By the above, one of any two consecutive terms of the sequence is $\geq a$. So if $b < a$ were another subsequential limit with $a_{n_k} \to b$, then for any $\varepsilon > 0$, for large enough k we would have had

$$a_{n_k} < (a+b)/2,\ a \leq a_{n_k-1} \leq a+\varepsilon,\ a \leq a_{n_k+1} \leq a+\varepsilon,$$

Thus for $\varepsilon < (a-b)/2$, we would have had $a_{n_k+1} > \frac{1}{2}(a_{n_k-1} + a_{n_k})$, a contradiction.

Problem M5. In his Care of Magical Creatures class, Hagrid showed his students magical amoebas. These creatures can inhabit cells of the first quadrant of an infinite checkerboard, labeled by (i,j), $i,j \in \mathbb{Z}_{\geq 0}$ (at most one amoeba per cell). If a magical amoeba occupies a cell (i,j) and the adjacent cells $(i+1,j)$ and $(i,j+1)$ above and to the right are empty, then it can divide, and the two daughter amoebas will inhabit the two adjacent cells. Initially, there is just one magical amoeba living at $(0,0)$. Is it possible that the amoebas will ever vacate the entire 3 by 3 square in the corner of the board (i.e., the cells with $0 \leq i,j \leq 2$)?

Hint. Define a function on the set of configurations of amoebas that does not change when they divide.

Solution. For a configuration S of amoebas, let $f(S) = \sum_{s\in S} 2^{-(i_s+j_s)}$. Then f is preserved under division. So for the initial configuration $f = 1$, and for the configuration when all cells are inhabited, $f = \sum_{n\geq 0}(n+1)2^{-n} = 4$. For the 3 by 3 square, $f = 3\frac{1}{16}$, so for its complement $f = \frac{15}{16}$, which is less than 1. Hence it is < 1 for any configuration in which the square is empty. So the square can never be vacated.

Computational biology problems

Problem B1.

A bacteria in a certain population lives one or two days. On the next day after it is born, it divides with probability $p > 0$ and survives to the following day with probability $q > 0$ (otherwise it dies). On the second day, it divides with probability $r > 0$ (otherwise it dies).

(a) Find the condition on p, q, r under which the population will survive (if the initial number of bacteria is very large). In particular, determine if it will survive if:

(1) $p = q = r = 1/3$?

(2) $p = 1/3, q = r = 1/2$?

(b) Find the average rate of growth or decay of the population (i.e., how many times it grows or shrinks per day) as a function of p, q, r.

(c) If the population starts with 1 billion bacteria which are 1 day old, how soon, on average, will the population become extinct if $p = q = r = 1/4$?

Solution. Let a_n be the number of 1 day old bacteria, and b_n be the number of 2 year old bacteria on the n-the day. Then

$$b_{n+1} = qa_n,\ a_{n+1} = 2pa_n + 2rb_n.$$

So

$$a_{n+1} = 2pa_n + 2qra_{n-1}.$$

The characteristic equation for this recursion is $x^2 - 2px - 2qr = 0$, with roots $x_\pm = p \pm \sqrt{p^2 + 2qr}$. It's clear that $|x_-| < |x_+|$, So the rate of growth or decay of the population is x_+. Thus the condition that the population survives is $x_+ \geq 1$, i.e. $p + \sqrt{p^2 + 2qr} \geq 1$, or $p + qr \geq 1/2$. (This can also be seen directly: the transition case is when the recursion has a constant solution). So in case (1) the population does not survive, and in case (2) it does.

For $p = q = r = 1/4$, the rate of decay is $(1 + \sqrt{3})/4$, which is about 0.683. So the population will become extinct in about $-\log(10^9)/\log(0.683)$, or approximately 54 days.

Problem B2. Gnomes have n genes. The probability that the k-th gene is mutated is p_k. Mutations are recessive, i.e., a gnome baby is born sick if a certain gene is mutated in both parents. What is the probability that a gnome baby will be born healthy, if all mutations happen independently?

Solution: The chance that the the k-th gene is mutated in both parents is p_k^2, so the probability that it does not happen is $1 - p_k^2$. So the probability of a healthy baby is

$$\prod_{k=1}^{n}(1 - p_k^2).$$