

Alibaba Global Mathematics Competition 2024 Qualifying Round

Lazar Ilic

1

[6] on sight construction is trivial as is proof that ray EA intersecting ray FB at student 1 immediately implies EB intersects FA in the interior of ABFE such that there can not exist a student 7 there in that associated position. Can certainly code up a randomised algorithm to bash out and check similarish satisfiability constraint tasks probabilistically perhaps.

```

import math
import random
sum=0
num=100000000
for a in range(num):
time=2+math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85:
time+=1.5
sum+=time
print(sum/num)
sum=0
for a in range(num):
time=2+math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85:
time+=1.5
time+=math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85*0.85:
time+=1.5
sum+=time
else:
sum+=time
else:
sum+=time
print(sum/num)
sum=0
for a in range(num):
time=2+math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85:

```

```

time+=1.5
time+=math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85*0.85:
time+=1.5
time+=math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85*0.85*0.85:
time+=1.5
sum+=time
else:
sum+=time
else:
sum+=time
else:
sum+=time
print(sum/num)
sum=0
for a in range(num):
time=2+math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85:
time+=1.5
time+=math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85*0.85:
time+=1.5
time+=math.log(1-random.random())
if time>0:
prob=random.random()
if prob<0.85*0.85*0.85:
time+=1.5
time+=math.log(1-random.random())
if time>0:

```

```
prob=random.random()
if prob<0.85*0.85*0.85*0.85:
time+=1.5
sum+=time
else:
sum+=time
else:
sum+=time
else:
sum+=time
else:
sum+=time
print(sum/num)
```

2 and 2

3

(1) follows immediately from the following observation: given the integer matrix A has $\text{tr}(A) = 0$ one may write it as $A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ and simply compute that

$\det(A) = |a^2 + bc| \neq 0$ but more canonically that $A^{2n} = (a^2 + bc)^n \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and

$A^{2n+1} = (a^2 + bc)^n \cdot \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ strong inductively. Then the subset $A^{2n}Z^2$ of A^nZ^2

simply consists of the subgrids of Z^2 where both coordinates are congruent to 0 modulo $|a^2 + bc|^n$ where trivially the constant $C_1 = \frac{|a^2+bc|}{\sqrt{2}}$ works. Another simple bound emerges for the odd case, in particular

$C_2 = \max(\sqrt{a^2 + c^2}, \sqrt{b^2 + a^2}, \sqrt{(a+b)^2 + (c-a)^2})$ should work as for example $|\det(A)| \geq 1$ implies $|\det(A)|^{\frac{n}{2}} \geq |\det(A)|^{\frac{n-1}{2}}$ in the original n -framing of the task statement. And so closed discs emerging from each point in A^nZ^2 in these odd cases should completely cover all 4 parallelograms which intersect at that point and thus even more strongly we

have that each point will have 4 such $w \in \Gamma$ within the required distance bound. So

$$C = \max\left(\frac{|a^2 + bc|}{\sqrt{2}}, \sqrt{a^2 + c^2}, \sqrt{b^2 + a^2}, \sqrt{(a+b)^2 + (c-a)^2}\right) = \max(C_1, C_2) = \text{ for } A =$$

works.

(2) follows again as we can for example simply pick a C which will ensure to fully cover the 4 neighbouring parallelograms in the planar tiling induced by that particular power of the matrix by forming a new basis parallelogram. In particular, the condition implies that $\text{tr}(A)$ is an integer and so in particular $\det(A) \neq 0$ and furthermore one must have diagonalizability and so in particular in the Cayley characteristic polynomial and quadratic formula we have $\text{tr}(A)^2 - 4\det(A) \neq 0$ otherwise there would be factorability over rationals immediately. Also, both off-diagonal elements must be nonzero or else we would have the diagonal elements sum to the trace and multiply to the determinant so they would be the eigenvalues and a contradiction on rational factorizability. So when we diagonalise say via $A = SJS^{-1}$ then if the 2 eigenvalues have the same magnitude and opposite signs fundamental large powers will be constant multiples of SKS^{-1} where $K = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Or $K \approx \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ where S and S^{-1} are fixed in such a way that exponentially the only relevant terms in our upcoming comparison come in to essentially what is an asymptotically fixed ratio of precisely the quantity we are looking for that is to say a tight rational

approximation should lead to some integer linear combinations of the 2 column vectors which have much lower magnitude but still the same precise spanning subgrid of Z^2 due to coprimality in coefficients used to produce this new basis. In the $K = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ case we observe that basically the even and odd power cases will simply be scaled up versions of the same underlying pair of column vectors and so in particular the actual natural trivial parallelogram tiling in those cases will be simple homotheties with ratio e.g. between the A^4 and A^6 case being $|\lambda_1|^2 = |\lambda_2|^2$. And the same ratio between the A^5 and A^7 case e.g. any consecutive pair. Then it suffices to note in the formula that $\det(A)^{\frac{n}{2}}$ scales as λ_1^n in fact and so we simply need to choose the minimal C such that it covers all 4 neighbouring parallelograms in the A and A^2 cases and that will suffice to cover all the cases. As for the other case $K = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ when J degenerates kind of towards $\begin{bmatrix} \lambda_1^n & 0 \\ 0 & 0 \end{bmatrix}$ where the λ_2^n term with $|\lambda_1| > |\lambda_2|$ is handwaved away as ≈ 0 for this part of the task which is permissible here due to the formula for positive terms in S and S^{-1} being fixed and eventually their ratio will come to dominate in upcoming argumentation. So then the terms in the matrix eventually asymptote towards blowing up like $|\lambda_1|^n$ with their respective constant factors and the determinant is blowing up in $|\lambda_1 \lambda_2|^n < |\lambda_1|^{2n}$ which implies the 2 column vectors will approach having an angle of 0 with each other as well as a certain fixed ratio. So for example if we look at something like $\begin{bmatrix} 1 & 5 \\ 3 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 54292951 & 102311375 \\ 61386825 & 115579776 \end{bmatrix}$ where $\frac{54292951}{102311375} \approx 0.531$ could be more formalized by viewing S and S^{-1} algebraically. For example here A^{100} has that ratio at 64291311831087094101385188261464839851106678703478157282062340763884789468748857 ≈ 0.531 as well. So in this example one may produce an OK rational approximation like say $\frac{53}{100}$ for this ratio as well as $\frac{531}{1000}$ say and then one can produce from those column vectors for example $100(54292951, 61386825) - 53(102311375, 115579776) = 6792225, 12954372$ and $1000(54292951, 61386825) - 531(102311375, 115579776) = -34389125, 13963944$ and then now that we have produced these 2 vectors with significantly smaller magnitude than the original 2 column vectors we may note that they certainly induce a subgrid which can be covered with some C value for that particular $n = 10$ value. However, that C value will not work in future cases as the \det term is blowing up slower than the literal values. For this proof to be complete we need a theorem in Best Rational Approximations for example by consulting Ken

Shoemake et al.

4

(1) One approach is to consider canonical results on the determinant, characteristic polynomial, tridiagonal, continuant, systems of equations, etc. The other approach is more direct with eigenvalues and eigenvectors. Another easy proof comes from Mathematics Overflow post “Exact Eigenvalues Of A Specific Tridiagonal Matrix” where we take the similarity transformation $D^{-1}TD$ in to that exact matrix multiplied by a $\frac{N+1}{2}$ constant factor which also maps multiplies eigenvalues by that precise factor. “The Matrix X I was looking at was nothing more than the angular momentum operator L_X in the spin $l = \frac{N+1}{2}$ representation, see angular momentum on Wikipedia. It can be written as $L_X = \frac{L_+ + L_-}{2}$ where L_+ and L_- obey $[L_z, L_{\pm}] = L_{\pm}$ and L_z is the diagonal matrix with integer spaced entries ranging from $-l$ to l . From the knowledge of spin representations (or $SL(2, C)$ if you which) one can easily find the spectrum of L_X . Also, the matrix diagonalizing X is just $V = e^{\frac{\pi i}{2L_y}}$ which in itself is not symmetric but from it a matrix $V' = VD$ can be constructed which also diagonalizes X and which is symmetric. Here $L_y = \frac{i}{2}(L_+ - L_-)$.” In particular this result reduces in to a very canonical theorem in physics on L_z Angular Momentum Operator throughout undergraduate and graduate physics course notes.

(2) $\boxed{1 + \left\lfloor \frac{d}{2} \right\rfloor}$ as symmetric eigenvalues across 0 induce the same subeigenvectors in U_0 and otherwise the set is linearly independent in the subspace.

(3) $\boxed{1 + \left\lfloor \frac{d}{2} \right\rfloor - j}$ by rank nullity and precise argumentation surrounding the linear system of equations.

According to John and the Lowner-John Ellipsoid Theorem one can obtain a volume ratio bound of π . Indeed from page 172 of Handbook Of The Geometry Of Banach Spaces and Theorem 5 in particular we see that “if C is a symmetric convex body in R^n , then C is included in an ellipsoid, E with $\frac{\text{vol}(E)}{\text{vol}(C)}$ ” no more than the corresponding ratio for the l_1^n ball. “Proof: The dual form of John’s Theorem guarantees that if the ellipsoid of minimal volume containing C is the ball B_2^n , then there are unit vectors (u_i) in C , and weights c_i satisfying the John condition (2). The problem is to show that the absolutely convex hull of these (u_i) has volume at least that of the l_1^n ball; namely $\frac{2^n}{n!}$. Let K be the absolutely convex hull and let $\|\cdot\|$ be the norm corresponding to K . By using polar coordinates it is easy to check that $\text{vol}(K) = \frac{1}{n!} \int_{R^n} e^{-\|x\|} dx$. For any x , $\|x\| = \min(\sum_i |\lambda_i| : x = \sum_i \lambda_i u_i) = \min(\sum_i c_i |\theta_i| : x = \sum_i c_i \theta_i u_i)$. Hence, if $f_i(t) = e^{-|t|}$ for each i , $e^{-\|x\|} = \max(\prod_i f_i(\theta_i)^{c_i} : x = \sum_i c_i \theta_i u_i)$. By Barthe’s Theorem, $\int_{R^n} e^{-\|x\|} dx \geq 2^n$ as required.” Here we get the maximal volume ratio obtained by the unit octahedron and unit sphere of $\frac{4\pi}{3} = \pi$. To obtain the bound for surface area one may consult Application Of An Idea Of Voronoi To John Type Problems by Peter M. Gruber in Advances In Mathematics or elsewhere and Sections Of Convex Bodies In John’s And Minimal Surface Area Position e.g. one might infer that the l -ball e.g. the octahedron is still the minimizing shape for a surface area ratio of $\frac{\pi}{\sqrt{3}} \approx 1.8138 < 3$. Simple Estimates For Ellipsoid Measures suggests that one can reframe this in terms of symmetric functions of the semi-axes of the John Ellipsoid and hence when inverted the Lowner Ellipsoid and one supposes that with a little bit algebra one can obtain a sharper bound than 3 and also prove that the Lowner Ellipsoid will be good enough despite potentially not being the surface area minimising enclosing ellipsoid.

$\frac{1}{2}$ and $\frac{1}{16}$ respectively, assuming each $p_i > 0$ otherwise $\frac{1}{2^{\text{Number Of Nonzero } p_i - 1}}$. Both follow from a classic and fundamental theorem in Markov Chains and in particular the stationary distribution associated with a symmetric matrix such that each state has positive probability of being reached from the starting state. "A regular Markov Chain with symmetric transition matrix is also reversible. Then the unique stationary distribution is the uniform distribution." This can be observed for example via the usual eigenvalue eigenvector of the transpose argumentation with each row and column summing to 1. In the former case (i) we may index from 0 so State 0 is the state in which we have flipped an even number of heads thus far and State 1 is the state in which we have flipped an odd number of heads thus far. The transition matrix by definition is $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ as the probability of flipping a head and thus switching states is $\frac{1}{3}$. In the more general case (ii) we may for example construct the matrix by associating States 0 through 31 inclusive with their binary strings so that the bit in the i th position refers to the parity of the number of balls placed in box i thus far. Then in the $\lim_{n \rightarrow \infty} X_{2n}$ case we essentially have that matrix filtrated to a 16×16 matrix submatrix which only contains those states with the sum [bitwise XOR] of the parity bits being even. Symmetry is still present in this chain by a simple symmetry argumentation that is to say for example the probability of moving from state i to state j is the same as from state j to state i because the same precise sets of ball placements will execute those. To make it even more clear e.g. to move from state $11110_2=30$ to state $11011_2=27$ would require that the 2 balls in the 2 steps were for boxes 0 and 2. And again vice versa to move back from state 27 to state 30. Or more simply note this fact for 1 step in the original matrix by the same bit flipping argument and then note that the square of a symmetric matrix is itself symmetric [the relevant induced submatrix].

(1) The unique exception case is when we solve basically a steady state equation so for some gap if the lovers started at 0 and $0.5 + \alpha$ basically after moving up they would be at $0.5 - \alpha$ and 1 and then stop for precisely 2α so that we revert back in to the $0.5 + \alpha$ and 0 position then we would need to solve for that specific α value [unique from given $f' > 0$ but $f'' > 0$ conditions] and also perhaps note partial differential equations activity in the vicinity threshold of that region of the function. Otherwise the blowup is ensured to occur in an increasingly fast fashion. The condition immediately implies that the inverse of f also has $f' > 0$ but $f'' > 0$ and so fundamentally what will happen here if we start [without loss of generality because if it started at $> 0.5 + \alpha$ then after 1 step we would be in such a position] with the 2 lovers at 0 and some value $< 0.5 + \alpha$ is that if $\epsilon \ll b$ the $K(\phi)$ function is well-approximated to 1st order by $\epsilon(f^{-1})'(f(\phi)) = \frac{\epsilon}{f'(\phi)}$ [the following proof still holds strongly even if we used a strict bound on the f' value by considering a full shift of ϵ in the underlying ϕ where we evaluate it] at which point we will see that like in my proof for part (2) below the rate at which they grow closer increases over time and so the first 2 iterations will immediately give a strict upper bound on the total number of iterations needed [1/the quantity the gap decreased by after the first 2 iterations] as the dispersal is ensured to increase as f' increases on $(0, 1)$. The amount the gap spreads by then grows larger and larger each iteration as both parts are spread apart basically the ϕ value is larger when the first lover hits 1 each time and then will wait even longer so that way the ϕ value when the second lover hits 1 each time is even shorter than in the previous cycle from f' increasing on $(0, 1)$. To make this even more clear and precise, if the starting position is 0 and $\phi_1 < 0.5 + \alpha$ then it leads in to $1 - \phi_1$ and 1 leads to 0 and $1 - \phi_1 + K(1 - \phi_1) = \phi_2 > 0.5 + \alpha$ because $1 - \phi_1 > 0.5 - \alpha$ and thus $K(1 - \phi_1) > K(0.5 - \alpha) = 2\alpha$ and adding gives the claim. This then leads in to $\phi_1 - K(1 - \phi_1)$ and 1 which then leads in to 0 and $\phi_1 - K(1 - \phi_1) + K(\phi_1 - K(1 - \phi_1)) = \phi_3$. By the previous inequalities $\phi_1 - K(1 - \phi_1) < 0.5 - \alpha$ but $1 - \phi_1 > 0.5 - \alpha$ by construction condition and thus $\phi_1 - K(1 - \phi_1) + K(\phi_1 - K(1 - \phi_1)) < \phi_1$. Now it becomes clear that each cycle we will wait even longer during the first waiting period and thus even shorter during the second waiting period and so this gap does grow that is to say that $\phi_1 - \phi_3 < \phi_3 - \phi_5$ and so we get that upper bound on the number of cycles of $\frac{1}{K(1-\phi_1)-K(\phi_1-K(1-\phi_1))}$ formally.

(2) Experimentally it would seem that if they start at 0 and some value $< 0.5 - \alpha - \epsilon$ then it could maybe take $\approx O\left(\frac{1}{b^2\epsilon^2}\right)$ but this number rapidly drops

even if they start 0.4 apart and so an expected runtime might be expected in that sort of ballpark or lower perhaps depending on the expectation metric and one might expect more like $\approx O\left(\frac{1}{b\epsilon}\right)$ or even lower runtime. For this particular function one might get $O\left(\frac{1}{\epsilon}\right)$ as long as they start at 0 and some value $< 0.5 - \alpha - \epsilon - b$. There exists another term relating to the distance from $0.5 + \alpha$ so argumentation for producing bounds on runtimes in ranges by reversing the process might lead to $O\left(\frac{1}{|0.5+\alpha-\phi|b^2\epsilon^2}\right)$.

Under the assumption that as usual $\epsilon \ll b$ then up to a first order approximation we have that $K(\phi) \approx \epsilon(f^{-1})'(f(\phi)) = \frac{\epsilon}{f'(\phi)} = \epsilon(b\phi + \frac{b}{e^b-1})$ and in particular that the constant $\frac{b}{e^b-1}$ term will be sort of kind of vaguely ≈ 1 and dominating the $b\phi$ term in a midkey way it is just in the bouncing back and forth it might be more useful to directly compute what happens after 2 steps for example in canceling terms because after each dude has gone through 1 waiting cycle, if the waiting cycles were similar in length then they will have each traveled a very similar distance and be nearly where they were at the start of the process.

```
0 0.5
```

```
0 0.5+epsilon*(b*0.5+1)
```

```
0 0.5-epsilon*epsilon*b*b = 0 0.5-epsilon*(b*0.5+1)+epsilon*(b*(0.5-epsilon*
```

```
...
```

```
Precisified Polynomial Blowup? This looks just OK for an upper bound.
```

```
0 x
```

```
0 1-x+epsilon*(b*(1-x)+1)=x
```

```
For That Unique Exception
```

```
Could use C++. For esp=0.000001 and c=0.01 obtain 370428865.65271056 and als
```

```
import math
```

```
a=0
```

```
aa=0
```

```
b=0.5
```

```
eps=0.00001
```

```
c=0.01
```

```
z=c/(math.exp(c)-1)
```

```
while a==0 or b==0:
```

```
print(a,b)
```

```
if a==0:
```

```
a+=1-b
aa+=1-b
wait=eps*(c*a+z)
a+=wait
aa+=wait
if a<1:
    b=0
else:
    b+=1-a
wait=eps*(c*b+z)
b+=wait
if b<1:
    a=0
print(aa)
```