PROJECT EULER PROBLEM 100

GVHL

1. Mathematics

Consider a collection of n discs, where we have a blue ones and $n-a$ red ones. There exist arrangements such that the chance of taking two blue discs at random is $\frac{1}{2}$. Find the first arrangement such that $n > 10^{12}$.

Given n discs in total, and a blue discs, we have that

$$\mathrm{P(BB)} = \frac{a(a-1)}{n(n-1)}$$

Setting this equal to $\frac{1}{2}$ we end up with the following Diophantine equation:

$$2a(a-1) = n(n-1) \tag{1}$$

Let $b = a - \frac{1}{2}$ and $m = n - \frac{1}{2}$. Substituting this gives:

$$\begin{aligned} & 2b^2 - \frac{1}{2} = m^2 - \frac{1}{4} \\ \Leftrightarrow \quad & (2m)^2 - 2(2b)^2 = -1 \\ \Leftrightarrow \quad & x^2 - 2y^2 = -1 \end{aligned} \tag{2}$$

where $x = 2m$ and $y = 2b$. This can be recognised as a 'negative Pell equation' [1].

Remark. This will indeed give valid (a, n). Indeed, $x^2 = 2y^2 - 1 \equiv 1 \mod 2$, so x will be odd. As a result: $2m \equiv 1 \mod 2$, thus $n = (m + \frac{1}{2}) \in \mathbb{N}_0$.
We can also see that y must be odd. As we have already found, $\exists k \in \mathbb{N}_0$ such that $x = 2k+1$. Subsituting this in $y^2 = \frac{x^2+1}{2}$ gives $y^2 = \frac{4k^2+4k+1+1}{2} = 2k^2 + 2k + 1$. Similarly, we find $a = (b + \frac{1}{2}) \in \mathbb{N}_0$, as expected.

The fundamental solution can be found in other ways, but here it is easy to see it is $(x_0, y_0) = (1, 1)$. Note that:

$$(x_0^2 - 2y_0^2)^2 = 1$$

Now all solutions [2] [3] can be found using:

$$x_k + \sqrt{2}y_k = (x_0 + \sqrt{2}y_0)^{2k+1}$$

with $k \in \mathbb{N}_0$. Note that $k = 0$ gives the fundamental solution. We will look for a recurrence relation.

Given for $k \in \mathbb{N}$ that $x_k + \sqrt{2}y_k = (x_0 + \sqrt{2}y_0)^{2k+1}$, we can deduce:

$$\begin{aligned} x_{k+1} + \sqrt{2}y_{k+1} &= (x_0 + \sqrt{2}y_0)^{2(k+1)+1} = (x_0 + \sqrt{2}y_0)^2 (x_0 + \sqrt{2}y_0)^{2k+1} \\ &= (x_0^2 + 2y_0^2 + 2x_0y_0\sqrt{2})(x_k + \sqrt{2}y_k) \\ &= \left[(x_0^2 + 2y_0^2)x_k + 2x_0y_0 \cdot 2y_k\right] + \left[2x_0y_0x_k + (x_0^2 + 2y_0^2)y_k\right]\sqrt{2} \end{aligned}$$

Date: January 6, 2017.

Using the fact that $\sqrt{2}$ is irrational and subsituting the values of the fundamental solution, we find that:

$$\begin{cases} x_{k+1} &= 3x_k + 4y_k \\ y_{k+1} &= 2x_k + 3y_k \end{cases}$$

We will now try to solve this system of recurrence relations. Observe that $y_k = \frac{1}{4}(x_{k+1} - 3x_k)$. This gives:

$$x_{k+1} \quad = \quad 3x_k + 4(2x_{k-1} + 3y_{k-1}) = 6x_k - x_{k-1}$$

Substituting $x_k = r^k$ in the recurrence relation $x_{k+1} = 6x_k - x_{k-1}$ gives the characteristic polynomial $r^2 - 6r + 1$. This has the following solutions:

$$r_{1,2} = 3 \pm 2\sqrt{2}$$

Set $r_1 = 3 - 2\sqrt{2}$ and $r_2 = 3 + 2\sqrt{2}$. The general solution will be now be $x_k = c_1 r_1^k + c_2 r_2^k$.

$x_0 = 1$ gives $c_1 + c_2 = 1$. We will now use y_0.

$$\begin{aligned} 1 = y_0 = \frac{1}{4}(x_1 - 3x_0) = \frac{1}{4}(c_1 r_1 + c_2 r_2 - 3)) \\ \Leftrightarrow \quad 4 = c_1 r_1 + (1 - c_1) r_2 - 3 = c_1(r_1 - r_2) - 3 + r_2 \end{aligned}$$

Note that $r_1 - r_2 = -4\sqrt{2}$ and $-3 + r_2 = 2\sqrt{2}$. Solving for c_1, c_2 gives

$$c_1 = \frac{1}{2} - \frac{1}{\sqrt{2}} \quad , \quad c_2 = \frac{1}{2} + \frac{1}{\sqrt{2}}$$

To reiterate, we have:

$$x_k \quad = \quad \left(\frac{1}{2} - \frac{1}{\sqrt{2}}\right)\left(3 - 2\sqrt{2}\right)^k + \left(\frac{1}{2} + \frac{1}{\sqrt{2}}\right)\left(3 + 2\sqrt{2}\right)^k \tag{3}$$

$$y_k \quad = \quad \frac{1}{4}(x_{k+1} - 3x_k) \tag{4}$$

Observe from the recurrence relation that $(x_k)_{k\geq 0}$ and $(y_k)_{k\geq 0}$ are strictly increasing sequences. Recall that $x = 2m = 2(n - \frac{1}{2})$ and $y = 2b = 2(a - \frac{1}{2})$. Thus:

$$\begin{aligned} a_k &= \frac{1}{2}(y_k + 1) \\ n_k &= \frac{1}{2}(x_k + 1) \end{aligned}$$

It follows that $(a_k)_{k\geq 0}$ and $(n_k)_{k\geq 0}$ are strictly increasing sequences as well. We see that (a_0, n_0) is a solution to (1), but not to the problem itself, as $a_0 - 1 = n_0 - 1 = 0$.

2. Solutions

We wish to find the first arrangement such that $n > 10^{12}$. Recall that $(y_k)_{n\geq 0}$ is a non-decreasing sequence. Thus, we seek the first $k \in \mathbb{N}$ such that $n_k > 10^{12}$.

We can relate this to x_k. Let $\mathbf{t}$ be the value for which we wish to find k such that $n_k > \mathbf{t}$. (In the problem: $\mathbf{t}= 10^{12}$. We have that:

$$n_k > \mathbf{t} \quad \Leftrightarrow \quad x_k + 1 > 2\mathbf{t} \quad \Leftrightarrow \quad x_k > 2\mathbf{t} - 1$$

Ideally, we'd use the direct formula we found in section 1 and perform an optimized search. However, due to the small inaccuracies involved with floating-point operations, doing so becomes problematic for large values.

Therefore, I opted to use the recurrence relation. The algorithm becomes quite simple now:

```
x ← 1;
y ← 1;
t ← 1e12;
while x ≤ (2t − 1) do
    x_new ← 3x + 4y;
    y_new ← 2x + 3y;
    x ← x_new;
    y ← y_new;
end
blueDiscs ← (y + 1)/2;
print(blueDiscs);
```

Using Python 2.7.10 this can run within 40μs and produces a result of 756872327473.

References

[1] https://en.wikipedia.org/wiki/Pell's_equation#The_negative_Pell_equation.
[2] http://mathworld.wolfram.com/PellEquation.html.
[3] http://www.imomath.com/index.php?options=616&lmm=0.