2019 U.S.A. Mathematical Olympiad Solutions

USAMO 1. Answer: $f(1000)$ may be any even positive integer.

To prove this, first, two bits of terminology: we say that f *fixes* the positive integer n if $f(n) = n$; and we write f^k for the function given by iterating f k times.

Now, note that as long as f fixes all odd numbers and f^2 fixes all even numbers (which in particular implies $f(n)$ is even whenever n is), the function f satisfies the equation. Thus, for any even m, we may take $f(1000) = m$, $f(m) = 1000$, and $f(n) = n$ for all other n, and the condition is satisfied.

To see that $f(1000)$ cannot be odd, we show the following two claims.

Claim 1. *f is injective.*

Proof. If $f(a) = f(b)$, then $a^2 = f^{f(a)}(a)f(f(a)) = f^{f(b)}(b)f(f(b)) = b^2$, so $a = b$. □

Claim 2. *f fixes every odd number.*

Proof. We prove this by induction on odd $n \geq 1$.

Assume f fixes each element of $S = \{1, 3, \ldots, n-2\}$ now (allowing $S = \varnothing$ for the base case $n = 1$). Notice that if $f(m) \in S$, then $f(m) = f(f(m))$, implying $m = f(m) \in S$ by injectivity. Applying this repeatedly, we see that if $f^k(m) \in S$ for any $k \geq 1$ then $m \in S$.

Now, we contend $f(f(n)) = n$. Indeed, suppose $f(f(n)) \neq n$. The two numbers $f^{f(n)}(n)$ or $f(f(n))$ have product n^2 and aren't both equal to n, so one of them must be less than n, and also odd, therefore in S. However, by the result of the previous paragraph, this implies $n \in S$, which is a contradiction.

Hence $f(f(n)) = n$. Let $y = f(n)$, so $f(y) = n$. Then we now have

$$y^2 = f^n(y) \cdot y = ny$$

where the step $f^n(y) = n$ used the fact that n is odd. We conclude $n = y$, as desired. □

Now, if $f(n)$ is odd, then $f(n) = f(f(n))$ implying $n = f(n)$. In particular, $f(n)$ cannot be odd for any even n. This completes the proof.

Remark. An argument similar to the one for the second claim shows that in fact f^2 fixes every even number, so the functions identified in the beginning of the solution are actually the only solutions to the equation.

This problem was proposed by Evan Chen.

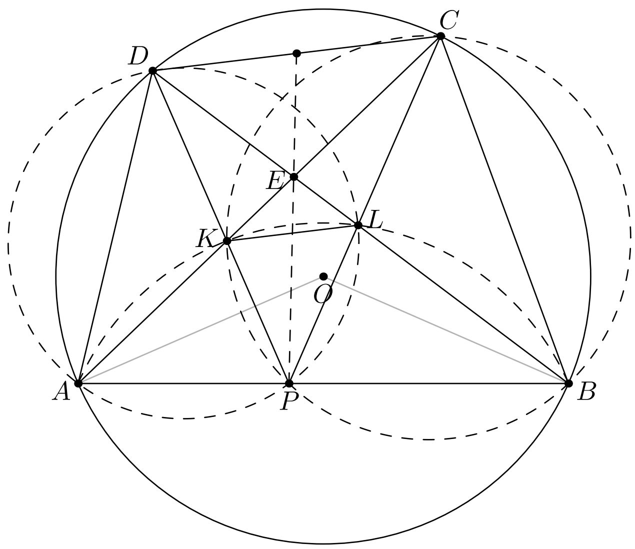
USAMO 2. Note that there can only be one point P on $\overline{AB}$ satisfying the given angle condition, since as P moves from A to B, $\angle APD$ decreases while $\angle BPC$ increases. Consequently, if we can show that there is a single point P on $\overline{AB}$ such that $\angle APD = \angle BPC$ and line PE bisects $\overline{CD}$, then it must coincide with the point in the problem statement, and we will be done. We construct such a point as follows.

Since $AD^2 + BC^2 = AB^2$, there exists a point P on $\overline{AB}$ satisfying

$$AD^2 = AP \cdot AB \quad \text{and} \quad BC^2 = BP \cdot BA.$$

Thus $AP/AD = AD/AB$ and $BP/BC = BC/BA$. We then have similar triangles, $\triangle APD \sim \triangle ADB$ and $\triangle BPC \sim \triangle BCA$, from which $\angle APD = \angle ADB = \angle ACB = \angle BPC$.

Now we show that line PE bisects $\overline{CD}$. Define $K = \overline{AC} \cap \overline{PD}$ and $L = \overline{BD} \cap \overline{PC}$.



The quadrilaterals $APLD$ and $BPKC$ are cyclic, because

$$\measuredangle ADL = \measuredangle ACB = \measuredangle BPC = \measuredangle APL$$

and similarly $\measuredangle KCB = \measuredangle KPB$. (The notation $\measuredangle$ here refers to directed angles taken modulo 180°.)

Now the quadrilateral $AKLB$ is also cyclic, because

$$\measuredangle AKB = \measuredangle CKB = \measuredangle CPB$$

and similarly $\measuredangle ALB = \measuredangle APD$, and these are equal.

Now the cyclic quadrilaterals imply $\measuredangle KCD = \measuredangle ABD = \measuredangle ABL = \measuredangle AKL = \measuredangle CKL$, from which we conclude $\overline{CD} \parallel \overline{KL}$. Thus $CDKL$ is a trapezoid whose legs intersect at P and whose diagonals intersect at E. As is well-known (and can be quickly shown using Ceva's theorem), this implies that line PE bisects the bases $\overline{CD}$ and $\overline{KL}$, as desired.

This problem was proposed by Ankan Bhattacharya.

USAMO 3. For an integer x, let $l(x)$ be the length of its base-10 representation. We will show that the only solutions are

- $f(X) = c$, with $c \in K$;
- $f(X) = ax$, with a a power of 10; and
- $f(X) = aX + b$ with a a power of 10, $b \in K$ and $l(b) < l(a)$.

Clearly all of these work. The following lemma is crucial to show that there are no other possibilities:

Lemma 1. *The only $x \in K$ such that $xy \in K$ for all $y \in K$ are the powers of* 10*.*

Proof. Assume x has the property and is not a power of 10. By induction we get $x^n \in K$ for any n. But, as is well-known, we can find a power of x that starts with any desired finite sequence of digits (in particular, we can find one that starts with 7), which gives a contradiction. For completeness, we give a proof of this fact in the next paragraph.

In general, suppose N is the number representing the desired sequence of digits. Assume that $N+1$ is not a power of 10 (if it is, just replace N by $10N$). Then the claim is that there exist integers $j, k \geq 0$ such that $N \cdot 10^k < x^j < (N+1) \cdot 10^k$. Taking $\log_{10}$ of both sides, this is equivalent to $k + \log_{10}(N) < j \log_{10}(x) < k + \log_{10}(N+1)$. Thus, what we need is

$$\{\log_{10}(N)\} < \{j \log_{10}(x)\} < \{\log_{10}(N+1)\}$$

where $\{\cdots\}$ denotes the fractional part. To see that there is such a j, let M be large enough such that $1/M < \log_{10}(N+1) - \log_{10}(N)$. Divide the unit interval into M equal-sized subintervals. Consider the values of $\{t \log_{10}(x)\}$ for $t = 1, 2, \ldots, M+1$. By the pigeonhole principle, some two of them fall in the same subinterval, and these two cannot be equal since $\log_{10}(x)$ is irrational. Hence, by subtracting, $0 < \{(t'-t)\log_{10}(x)\} \leq 1/M$ for some t', t. If $t' > t$, then consider the multiples $r \cdot \{(t'-t)\log_{10}(x)\}$ (for $r = 1, 2, 3, \ldots$); one of them must eventually lie between $\{\log_{10}(N)\}$ and $\{\log_{10}(N+1)\}$, and then $j = r(t'-t)$ is our desired value. If $t' < t$, then similarly some multiple $r \cdot \{(t'-t)\log_{10}(x)\}$ must lie between $1 - \{\log_{10}(N+1)\}$ and $1 - \{\log_{10}(N)\}$, and the corresponding value $j = r(t-t')$ does the trick. □

Next, write $f(X) = a_d X^d + \ldots + a_1 X + a_0$. First let us prove that $a_i \in K \cup \{0\}$ for all i. By assumption

$$f(10^n) = \sum_{j=0}^{d} a_j 10^{jn} \in K.$$

Choosing $n > \max_j l(a_j)$, the base-10 representation of $f(10^n)$ will consist only of the digits in base 10 of the a_j's and zeroes, hence all nonzero a_j belong to K. A similar argument will yield the crucial:

Lemma 2. *For $0 \leq r \leq s \leq d$, with a_s nonzero, and any $k \in K$, we have $a_s k^{s-r}\binom{s}{r} \in K$.*

Proof. Fix $k \in K$ and pick n large enough. The binomial formula yields

$$f(10^n + k) = \sum_{j=0}^{d} a_j (10^n + k)^j = \sum_{j=0}^{d} a_j \sum_{i=0}^{j} 10^{ni} k^{j-i} \binom{j}{i} = \sum_{r=0}^{d} 10^{nr} \sum_{s=r}^{d} a_s k^{s-r} \binom{s}{r}.$$

Picking $n > \max_{0\le r\le d} l\left(\sum_{s=r}^{d} a_s k^{s-r}\binom{s}{r}\right)$, we conclude as above that $\sum_{s=r}^{d} a_s k^{s-r}\binom{s}{r} \in K$. Since k was arbitrary, we can replace k by $10^p k$ and so also obtain $\sum_{s=r}^{d} a_s 10^{(s-r)p} k^{s-r}\binom{s}{r} \in K$ for any $k \in K$ and $p \ge 1$. Fixing k and choosing p large enough yields the result, by the same argument. $\square$

Suppose now that $d \ge 2$. Thanks to the lemma (pick $s = d$ and $r = d-1, d-2$) we obtain $da_d k \in K$ and $\binom{d}{2}a_d k^2 \in K$ for all $k \in K$. For $k \in K$ and p large enough we also have $\binom{d}{2}a_d(10^p + k)^2 \in K$ and arguing as above yields $2\binom{d}{2}a_d k \in K$. Applying the first lemma, we deduce that da_d and $2\binom{d}{2}a_d$ are powers of 10, thus their ratio $d-1$ is also a power of 10 and so $d = 2$. Since $da_d = 2a_d$ is a power of 10 and $a_d k^2 = a_d k^2\binom{d}{2} \in K$ for $k \in K$, we obtain $5k^2 \in K$ for all $k \in K$. Taking $k = 12$ yields a contradiction, since $5 \cdot 12^2 = 720$. This contradiction shows that $d \le 1$.

Consider the case $d = 1$ (the case $d = 0$ being trivial). If $a_0 = 0$, then $a_1 x \in K$ whenever $x \in K$, so the first lemma implies a_1 is a power of 10. Otherwise, the above discussion shows that $a_0, a_1 \in K$ and a_1 is again a power of 10. We claim that the only extra restriction is that $l(a_0) < l(a_1)$. This condition is clearly sufficient. On the other hand, suppose that $l(a_0) \ge l(a_1)$ and let $a_1 = 10^f$, $a_0 = g \cdot 10^e +$ (lower powers). If $g < 7$ picking $x = (7-g)\cdot 10^{f-e} \in K$ yields $a_1 x + a_0 = 7 \cdot 10^e +$ (lower powers), and this is not in K, a contradiction. If $g > 7$, picking $x = (17-g)\cdot 10^{f-e}$, provides the desired contradiction.

This problem was proposed by Titu Andreescu, Vlad Matei, and Cosmin Pohoata.

USAMO 4. The answer is $(2n)! \cdot 2^{n^2}$. It may be helpful to view the sets $S_{i,j}$ as being placed in a grid, as shown in Figure 1. We say a choice of sets $S_{i,j}$ is *valid* if it satisfies the two conditions in the problem. In a slight abuse of terminology, we also apply this definition at times when only some of the $(n+1)^2$ total sets are chosen, with the rest left undetermined (in this case, the conditions are ignored when one or more of the sets involved is undetermined).

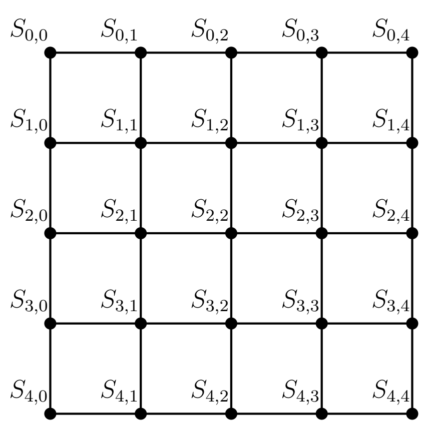


Figure 1: The $S_{i,j}$ arranged in a grid.

Let us define an *initial configuration* to be a valid choice of the sets corresponding to the top row and rightmost column (i.e. sets of the form $S_{0,j}$ and $S_{i,n}$). We first count the number of initial configurations. Since we must have

$$\emptyset = S_{0,0} \subseteq S_{0,1} \subseteq S_{0,2} \subseteq \cdots \subseteq S_{0,n} \subseteq S_{1,n} \subseteq S_{2,n} \subseteq \cdots \subseteq S_{n,n} = \{1, 2, \dots, 2n\}$$

and recalling that $|S_{i,j}| = i + j$, it follows that the above sequence of sets is obtained by adding different elements of $\{1, 2, \dots, 2n\}$ one at a time. We may add these $2n$ elements in any order, so the number of initial configurations is $(2n)!$.

Next, for any $0 \le i, j < n$, consider the sets $S_{i,j}$, $S_{i+1,j}$, $S_{i,j+1}$, and $S_{i+1,j+1}$. If they are part of a valid choice, we must have

$$S_{i,j} \subseteq S_{i+1,j+1} \quad \text{and} \quad |S_{i+1,j+1}| = i + j + 2 = |S_{i,j}| + 2,$$

which implies $S_{i+1,j+1} \setminus S_{i,j} = \{x, y\}$ for some distinct $x, y \in \{1, 2, \dots 2n\}$. Then, $S_{i+1,j}$ and $S_{i,j+1}$ are each either $S_{i,j} \cup \{x\}$ or $S_{i,j} \cup \{y\}$. Let us say the ordered pair (i, j) is *hot* if $S_{i+1,j}$ and $S_{i,j+1}$ are different and *cold* if they are the same. We define a *hot-cold configuration* to consist of a designation of "hot" or "cold" for each of the n^2 ordered pairs (i, j). Clearly, there are 2^{n^2} hot-cold configurations.

Finally, we claim that given any initial configuration and any hot-cold configuration, there is a unique valid choice of sets $S_{i,j}$ for $0 \le i, j \le n$ that agrees with both the initial configuration and the hot-cold configuration. Indeed, we start with the initial configuration of $2n + 1$ sets and choose the remaining sets one by one. We choose them in the following order:

$$\begin{array}{cccc} S_{1,n-1}, & S_{1,n-2}, & \dots, & S_{1,0}, \\ S_{2,n-1}, & S_{2,n-2}, & \dots, & S_{2,0}, \\ \vdots & \vdots & \ddots & \vdots \\ S_{n,n-1}, & S_{n,n-2}, & \dots, & S_{n,0}, \end{array}$$

and we will make sure our choice of sets remains valid at each step. In terms of the grid in Figure 1, this corresponds to going row by row, going right to left in each row.

The above ordering ensures that when we are choosing $S_{i,j}$, the sets $S_{i-1,j}$, $S_{i-1,j+1}$, and $S_{i,j+1}$ have all been chosen already. Based on whether $(i-1, j)$ is required to be hot or cold, we are forced to set $S_{i,j}$ to be $S_{i-1,j} \cup (S_{i,j+1} \setminus S_{i-1,j+1})$ or $S_{i-1,j+1}$, respectively. Moreover, it is straightforward to check that the resulting choice of sets indeed remains valid, because we have ensured that $S_{i-1,j} \subseteq S_{i,j} \subseteq S_{i,j+1}$.

Thus, at the end of the procedure, we arrive at a unique valid choice of all $(n + 1)^2$ of the $S_{i,j}$, establishing the claim. It follows that there are $(2n)! \cdot 2^{n^2}$ valid choices in total, as desired.

This problem was proposed by Ricky Liu.

USAMO 5. The answer is all (m, n) such that $m + n$ is a power of 2.

First, if $p \mid m + n$ for some prime $p > 2$, we show that any number $\frac{a}{b}$ written on the board will always have $p \mid a + b$. Indeed, if $p \mid s + t$ and $p \mid u + v$, then the arithmetic mean of $\frac{s}{t}$ and $\frac{u}{v}$ is $\frac{sv+tu}{2tv}$, and we note that

$$sv + tu + 2tv \equiv sv + tu + tv + su \equiv (s + t)(u + v) \equiv 0 \pmod{p}.$$

Since neither t nor v (nor 2) is divisible by p, we see that p still divides the sum of the numerator and denominator after the fraction has been reduced. Similarly, the harmonic mean $\frac{2su}{sv+tu}$ also satisfies the condition.

However, $1 = \frac{1}{1}$, and no prime $p > 2$ divides $1 + 1$, so no such prime can divide $m + n$ if Evan is to ever be able to write 1 on the board. So we need $m + n$ to be a power of 2.

We now show that Evan can fulfill his goal whenever $m + n$ is a power of 2. In fact, he can do this by only using the arithmetic mean. To show this, first notice that since $m + n$ is a power of 2, if he started with the numbers 0 and $m + n$ on the board, by repeatedly taking arithmetic means, he could eventually produce any integer between 0 and $m + n$; in particular, he could obtain the value m. But if $f(x) = cx + d$ is any linear function, the arithmetic mean of $f(x)$ and $f(y)$ is $f\left(\frac{x+y}{2}\right)$, so by replicating the same sequence of steps that gets to m starting from 0 and $m + n$, he can also get to $f(m)$ starting from $f(0)$ and $f(m + n)$. In particular, by taking $c = \frac{n-m}{mn}$ and $d = \frac{m}{n}$, we have $f(0) = \frac{m}{n}$, $f(m + n) = \frac{n}{m}$, and $f(m) = 1$, so by starting from $\frac{m}{n}$ and $\frac{n}{m}$, Evan can eventually reach 1, as needed.

(Note that the harmonic mean operation is never needed.)

This problem was proposed by Yannick Yao.

USAMO 6. We will first prove that $P(x) = c(x^2 + 3)$ is a solution for any real number c. This reduces to checking that

$$x(x^2 + 3) + y(y^2 + 3) + z(z^2 + 3) = xyz((x - y)^2 + (y - z)^2 + (z - x)^2 + 9)$$

whenever $2xyz = x + y + z$. Using the factorization of $a^3 + b^3 + c^3 - 3abc$ and the relation $x + y + z = 2xyz$, the left-hand side equals

$$\begin{aligned}(x^3 + y^3 + z^3) + 3(x + y + z) &= 3xyz + (x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) + 3(x + y + z)\\ &= xyz(9 + (x - y)^2 + (y - z)^2 + (z - x)^2),\end{aligned}$$

as desired.

Next, we prove that these are all solutions of the problem. If $P(x) = c$ is constant, then the left-hand side of the original equation equals $\frac{c(x+y+z)}{xyz} = 2c$, while the right-hand side equals $3c$. This is only possible if $c = 0$. Therefore, if $P(x)$ is a nonzero solution, it is not constant.

If $x \neq 0$, then $y = \frac{1}{x}$ and $z = x + \frac{1}{x}$ satisfy $2xyz = x + y + z$, so

$$xP(x) + \frac{1}{x}P\left(\frac{1}{x}\right) + \left(x + \frac{1}{x}\right)P\left(x + \frac{1}{x}\right) = \left(x + \frac{1}{x}\right)\left(P\left(x - \frac{1}{x}\right) + P(-x) + P\left(\frac{1}{x}\right)\right). \quad (1)$$

Note that the left-hand side is symmetric with respect to $x \to \frac{1}{x}$, thus so must be the right-hand side. It follows that

$$P\left(x - \frac{1}{x}\right) + P(-x) + P\left(\frac{1}{x}\right) = P\left(\frac{1}{x} - x\right) + P(x) + P\left(-\frac{1}{x}\right).$$

This can be rewritten as $Q\left(x - \frac{1}{x}\right) = Q(x) + Q\left(-\frac{1}{x}\right)$, where $Q(X) = P(X) - P(-X)$. We also know $Q(0) = P(0) - P(0) = 0$. Hence, as $x \to \infty$,

$$Q(x) - Q\left(x - \frac{1}{x}\right) = -Q\left(-\frac{1}{x}\right) \to 0.$$

Now, if $Q(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ with $n \geq 2$, then the left-hand side of the above equation is of the form $na_n x^{n-2} +$ (lower-order terms), which fails to go to 0 as $x \to \infty$. Thus, Q has degree at most 1, and since $Q(0) = 0$, then $Q(x) = 2ax$ for some real number a.

Using $P(x) - P(-x) = 2ax$, we conclude that the odd part of $P(x)$ is ax, so that $P(x) = ax + f(x^2)$ for a polynomial f with real coefficients. Replacing $P(x) = ax + f(x^2)$ in relation (1) yields

$$ax^2 + xf(x^2) + \frac{a}{x^2} + \frac{1}{x}f\left(\frac{1}{x^2}\right) + a\left(x^2 + 2 + \frac{1}{x^2}\right) + \left(x + \frac{1}{x}\right)f\left(x^2 + 2 + \frac{1}{x^2}\right)$$

$$= \left(x + \frac{1}{x}\right)\left(f\left(x^2 - 2 + \frac{1}{x^2}\right) + f(x^2) + f\left(\frac{1}{x^2}\right)\right).$$

Multiplying by x, we deduce that $2ax\left(x^2 + 1 + \frac{1}{x^2}\right)$ is a function of x^2, which implies that $a = 0$. Letting $t = x^2$, the previous relation becomes

$$f(t) + tf\left(\frac{1}{t}\right) = (t+1)\left(f\left(t + 2 + \frac{1}{t}\right) - f\left(t - 2 + \frac{1}{t}\right)\right).$$

Write $f(t) = b_n t^n + ... + b_0$ with $b_n \neq 0$ and suppose that $n > 1$. The largest term on the left-hand side is $b_n t^n$. However, the largest term on the right-hand side is the same as the largest term of

$$t(f(t+2) - f(t-2)),$$

which is $4b_n t^n$. This contradicts $b_n \neq 0$, which means $f(t)$ must be linear. We may check, if $f(t) = cx + d$ in the last formula, that $d = 3c$. Therefore, $f(x) = c(x+3)$, so $P(x) = f(x^2) = c(x^2 + 3)$.

This problem was proposed by Titu Andreescu and Gabriel Dospinescu.