THE CARD GAME SET AND ITS MATHEMATICS

ARUN DEBRAY
OCTOBER 14, 2015

ABSTRACT. SET is an addictive card game based on pattern recognition and fast reflexes. Though it wasn't designed to be a mathematical game, natural questions about SET touch on geometry, group theory, and representation theory. In this talk, we will discuss these questions and prove a few answers.

1. WHAT IS SET?

- Introduce the card game and its history: explain how to play.
- Two examples of sets, and one that isn't a set.
- Already, you can see arithmetic mod 3 going on!
- Sometimes, as in this example, there's no set, even though 12 cards are dealt out. After adding three more cards, there's *usually* but not always a set. So it's quite natural to ask: how bad could it get? Is it possible that there's no set in 18 cards? How about 21? Here's a collection of 20 cards that has no set in it. These are exceedingly rare: the chance of no set occurring in 15 cards is only 1 in 88, and only $6.8 \cdot 10^6$ collections of 20 cards have no sets, out of $\binom{81}{20} = 4.7 \cdot 10^{18}$! (This was proven by [5] using a computer program.)

2. A MATHEMATICAL REINTERPRETATION

We're going to prove that every collection of 21 cards contains a set. In order to do that, we'll add some math in.

- Introducing $\mathbb{F}_3^4$: this is sort of like linear algebra, and encompasses the mod 3 stuff we did before.
  - By *explicitly identifying* the Set cards with elements of $\mathbb{F}_3^4$, we have a geometrical structure. Since $\mathbb{F}_3$ is a field (addition, multiplication mod 3), this means we have a vector space, so we can do geometry and talk about $-x$ for a card x.
  - However, we don't really care about the origin: any card would do, so we really have affine space, and can do affine linear transformations.
  - Explain (using affine transformations) why a set is a line in $\mathbb{F}_3^4$.

Definition. A *cap* in $\mathbb{F}_3^d$, also called a d-*cap*, is a subset of $\mathbb{F}_3^d$ containing no lines. If it has the largest possible number of points such that this is possible, it is called *maximal*.

- We can also think about its uniqueness, relating to affine transformations, etc.

3. MAXIMAL d-CAPS

A maximal 2-cap has four points, which we can easily prove: we'll draw a picture of $\mathbb{F}_2^2$.

We divide $\mathbb{F}_2^2$ into three parallel lines; then, by pigeonhole, two points of our cap must be in one line, two in another, and one in the third (or one would contain three points, and they would form a line). Then, we can just chase all the cases.

A maximal 3-cap has nine points, which we'll prove.

The case-by-case approach would take far too long, so let's do something different. The idea behind this proof will work for $d = 4$, too, so it's useful for us to define things more generally. Specifically, for these proofs, a *hyperplane* in $\mathbb{F}_3^d$ is a $(d-1)$-dimensional affine subspace (that is, a linear subspace that might have been translated).

Definition. Fix a d-cap C and suppose $\mathbb{F}_3^d = H_1 \cup H_2 \cup H_3$ is a decomposition into parallel hyperplanes. Then, the *hyperplane triple* associated to this decomposition is $(|H_1 \cap C|, |H_2 \cap C|, |H_3 \cap C|)$, which we take to be unordered. We'll use the ordering (a, b, c) implies $a \geq b \geq c$.

In general, there are many ways to decompose into parallel hyperplanes. Any such decomposition is given by a line through the origin (take the hyperplane perpendicular to it at every point), and conversely, given such a decomposition, there's a unique line perpendicular to the hyperplanes that goes through the origin. Thus, the number of such decompositions on $\mathbb{F}_3^d$ is the number of ways to choose such a line: for any nonzero $x \in \mathbb{F}_3^d$, 0 and x determine a line, but 0 and $-x$ determine the same line, and since every line has only three points on it, x and $-x$ are the only two points that determine the same line. Thus, the number of decompositions into parallel hyperplanes is $(3^d - 1)/2$.

Remark. The notion of a hyperplane triple isn't just a combinatorial object: it actually relates to the Fourier transform on finite groups! For $\mathbb{F}_3^d$, the Fourier transform of an $f : \mathbb{F}_3^d \to \mathbb{C}$ is the function $\hat{f} : \mathbb{F}_3^d \to \mathbb{C}$ given by

$$\hat{f}(\xi) = \sum_{x \in \mathbb{F}_3^d} f(x)\zeta^{\xi \cdot x},$$

where $\xi \cdot x$ is the dot product on $\mathbb{F}_3^d$ and $\zeta = e^{2\pi i/3}$. The dot product gives us an element of $\mathbb{F}_3$, which permutes the cube roots of unity.

If you've seen some representation theory, that's what's really going on here: for a general finite abelian group, $\zeta^{\xi \cdot x}$ is replaced with $\chi_\xi(x)$, where χ_ξ is the group character associated with x. And replacing sums with integrals, you get the familiar Fourier transform on $\mathbb{R}$.

Anyways, if S is any subset of $\mathbb{F}_3^d$, let χ_S denote its characteristic function (equal to 1 on S, and 0 elsewhere). For any nonzero $x \in \mathbb{F}_3^d$, x is the normal vector for a decomposition into parallel hyperplanes H_0, H_1, and H_2, so if we let $h_i = |H_i \cap S|$, then $\hat{\chi}(x) = h_0 + h_1\zeta + h_2\zeta^2$. In analysis, taking the Fourier transform of a characteristic function is a common thing to do, but here, the coefficients of the Fourier transform are exactly the entries in a hyperplane triple! So they're more fundamental than they might seem.

Finally, we'll use the following result a few times.

Proposition. *Let K be a fixed k-dimensional subspace of $\mathbb{F}_3^d$. Then, there are $(3^{d-k} - 1)/2$ hyperplanes of $\mathbb{F}_3^d$ containing K.*

Proof. By an affine transformation, we can move K to the origin; then, the quotient map $\mathbb{F}_3^d \hookrightarrow \mathbb{F}_3^d/K \cong \mathbb{F}_3^{d-k}$ is a bijection between hyperplanes containing K and hyperplanes (in $\mathbb{F}_3^{d-k}$) containing the origin. (This is because K is sent to 0, and then just check what the dimensions are.) And hyperplanes containing the origin are determined by their normal vectors, and any nonzero vector is the normal vector of a plane. But x and $-x$ give the same plane. Thus, we get $(3^{d-k}-1)/2$. ⊠

Proposition. *A maximal 3-cap has nine points.*

Proof. We've found a 3-cap with nine points, so assume for a contradiction that there's a 3-cap C with ten points. Associated to any hyperplane decomposition of $\mathbb{F}_3^3$ is its hyperplane triple (ℓ, m, n), but since $|C| = 10$, then $\ell + m + n = 10$, and since each hyperplane is a copy of $\mathbb{F}_3^2$, it cannot contain more than four points of C, or there would be a line. Thus, $\ell, m, n \leq 4$. The only solutions to this are $(4, 4, 2)$ and $(4, 3, 3)$.

Let a be the number of hyperplane decompositions that have $(4, 4, 2)$ hyperplane triples, and b be the number that have $(4, 3, 3)$. Then $a + b = (3^3 - 1)/2 = 13$.

We want another relation between a and b; to obtain this, we'll consider *marked 2-planes*, which are pairs $(H, \{a, b\})$, where H is a 2-dimensional affine subspace of $\mathbb{F}_3^4$ and $a, b \in H \cap C$. That is, we're just picking two points in C on a given hyperplane. How many marked 2-planes are there? We'll use a time-honored trick from combinatorics, counting the same thing in two different ways.

If a hyperplane contains two points, it contains the line they determine, and a line is a one-dimensional affine subspace, so we can apply the proposition: there are $(3^{3-1} - 1)/2 = 4$ hyperplanes containing each pair of points in C, and there are $\binom{10}{2}$ ways to choose 2 points in C. Thus, there are $4\binom{10}{2} = 180$ 2-marked planes.

Alternatively, for each $(4, 4, 2)$-hyperplane triple, there's $\binom{4}{2}$ ways to realize the first hyperplane as a marked 2-plane, and so forth, so we get

$$a\left(\binom{4}{2} + \binom{4}{2} + \binom{2}{2}\right) + b\left(\binom{4}{2} + \binom{3}{2} + \binom{3}{2}\right) = 13a + 12b$$

marked 2-planes. Since this is all of the marked 2-planes, then $13a + 12b = 180$.

If you solve this system of equations, $b = -11$, which makes no sense. Thus, no 3-cap can have 10 points. ⊠

Finally, let's prove that a maximal 4-cap has 20 points. It's mostly the same idea.

Proposition (Pellegrino [6])**.** *A maximal 4-cap has 20 points.*

Proof. We've found a 4-cap with 20 points thanks to our SET cards (remember those?), so suppose we have a 4-cap C with 21 points. Since a 3-cap can only have 9 points, the possible hyperplane triples are $(9, 9, 3)$, $(9, 8, 4)$, $(9, 7, 5)$, $(9, 6, 6)$, $(8, 8, 5)$, $(8, 7, 6)$, and $(7, 7, 7)$. Let a denote the number of hyperplane decompositions that give us a $(9, 9, 3)$ triple, b denote that for $(9, 8, 4)$, and so on through g. These account for all hyperplane decompositions, so

$$a + b + c + d + e + f + g = \frac{3^4 - 1}{2} = 40.$$

Next, we'll count 2-marked hyperplanes in $\mathbb{F}_3^4$. By one of our previous propositions, the number of hyperplanes containing a pair of points, or equivalently a line, is $(3^{4-1})/2 = 13$, and there are $\binom{21}{2}$ ways to choose two points of C, so there are $13\binom{21}{2} = 2730$ 2-marked hyperplanes. The coefficients of a through g are calculated precisely as before, e.g. the coefficient for a is $\binom{9}{2}+\binom{9}{2}+\binom{3}{2}$. When we actually work all of this out, we get

$$75a + 70b + 67c + 66d + 66e + 64f + 63g = 2730.$$

These two alone aren't actually enough to force a contradiction, so we'll produce a third equation by counting the number of 3-marked hyperplanes in $\mathbb{F}_3^4$. (These are just like 2-marked hyperplanes, but with three points in C instead of two.) Since no three points in C determine a line, then any triple of points in C determines a plane, and our proposition tells us the number of hyperplanes containing a given plane is $(3^{4-2}-1)/2 = 4$. Thus, the number of 3-marked hyperplanes is $4\binom{21}{3} = 5320$. Then, we'll split by the type of hyperplane triple in precisely the same way as before, e.g. the coefficient for a is $\binom{9}{3}+\binom{9}{3}+\binom{3}{3}$. When we calculate this out, the numbers are

$$169a + 144b + 129c + 124d + 122e + 111f + 105g = 5320.$$

Three equations in seven variables have infinitely many real-valued solutions, but we care about nonnegative integers. 693 times the first equation plus 3 times the third equation minus 6 times the second equation gives us

$$5b + 8c + 9d + 3e + 2f = 0,$$

and since these are nonnegative integers, $b, \dots, f = 0$. Then, we can take the second equation and subtract 63 times the first equation, and plug in $b, \dots, f$. This is just $12a = 210$, which is a contradiction: a must be an integer. ⊠

4. Projective Set

You might be wondering, what other SET-like games exist? As we've seen, we can think about this as working in a finite geometric space such that for any two points, the line between them determines a unique third point. It turns out this is also possible with projective spaces.

- The Fano plane: why it satisfies this, and then its mathematical properties (e.g. its automorphism group and relation to the octonions).
- Thus, we can define Projective Set games generally by extending this idea. We get $2^d - 1$ cards, and here's what a hand would look like.
- The maximal cap for d-dimensional Projective Set was found in 1947 (!) by R.C. Bose in [1], and is 2^d. The question of maximal caps in Projective SET was asked in the context of error-correcting codes.

So now we have SET games over $\mathbb{F}_3^d$ and $\mathbb{P}^d\mathbb{F}_2$. Are there more SET-like games out there? That depends on what you mean by "SET-like."

- One of the key properties of SET is that, given any two cards x and y, there is a unique third card z so that $\{x, y, z\}$ is a set. If this is the definitive property for you, then there are many more possible SET-like games! A *Steiner triple system* is a set X and a collection of three-element subsets S of X such that given any two $x, y \in X$, there's a unique third point z such that $\{x, y, z\} \in S$ — so this is essentially just the mathematical translation of this rule. There are many Steiner triple systems, some quite strange. Reference: the book *Triple Systems* by Colbourn and Rosa [2].
- Another important property of SET and Projective SET is their high degree of symmetry: any two SETs are equivalent. That is, given cards $x_0, x_1, y_0, y_1 \in \mathbb{F}_3^d$, there is an affine linear transformation sending $x_0 \mapsto y_0$ and $x_1 \mapsto y_1$. (Send x_0 to the origin, then rotate $x_1 - x_0$ to $y_1 - x_1$, and then add x_1.) It is a deep theorem of Key, Shult, Hall, and Kantor that the only Steiner triple systems with this kind of symmetry are $\mathbb{F}_3^d$ and $\mathbb{P}^d\mathbb{F}_2$. The proof requires using the classification of finite simple groups. One reference is [4].

References

[1] R. C. Bose. "Mathematical theory of the symmetrical factorial design." Sankhyā **8** (1947), 107–66.

[2] C. J. Colbourn and A. Rosa. *Triple Systems.* Oxford Mathematical Monographs, 1999.

[3] B. L. Davis and D. Maclagan, "The card game SET", Math. Intell. **25** (2003), 33–40. `http://homepages.warwick.ac.uk/staff/D.Maclagan/papers/set.pdf`.

[4] J. D. Key and E. E. Shult. "Steiner triple systems with doubly transitive automorphism groups: a corollary to the classification theorem for finite simple groups." J. Combin. Theory Ser. A, **36** (1) (1984), 105–10.

[5] D. Knuth, "Setset." `http://www-cs-faculty.stanford.edu/~uno/programs/setset-all.w`.

[6] G. Pellegrino. "Sul massimo ordine delle calotte in $S_{4,3}$." Matematiche (Catania) **25** (1971), 149–57.