

Set Strategy

Set is not a board game that leaves much room for luck. When an experience Set player plays with you, you're almost certainly going to be destroyed.

Back in college, I had a friend who was really into this game. She would always invite me to play with her in our free time and I would get crushed every time. One day, I got sick of getting beaten, so I decided to intensely train and study Set for a week. After a week, I played a rematch with her, and won.

This guide presents some strategies for the game that I discovered. After guiding you through these incremental strategies, I will evaluate how much of an improvement each strategy makes the game.

Sometimes I see guides like these like spoilers of sorts, since I really enjoyed the experience of learning to be better at set myself. So, if you would like, you should try to learn Set by yourself for a couple of weeks, and then come back here.

****Spoiler Alert****

The Rules

Set is a real time pattern-finding card game. Each Set card has 4 identifying attributes -- color, number, shape, and shade. Each attribute has 3 variants.

Color	Red	Blue	Green
Number	1	2	3
Shape	Diamonds	Oval	Squiggly
Shade	Solid	White	Shaded

Counting all unique cards yields $3^4 = 81$ cards. The game involves laying out 12 random cards, and finding a set of 3 cards such that for each attribute, all the cards either have the same attribute (e.g. all blue) or all different attributes (e.g. blue, red, and green). An example of not a set is {blue, blue, green}.

When a set is found, the first player to call "set" and correctly retrieves the 3 cards will win a point. The empty slots are filled with new cards, and the game continues until the 81 card deck

runs out and there are no more sets on the field.

Strategy #1: Pick 3 Cards

When I was first taught Set, it took time to even figure out if any given 3 cards formed a set. So my very first game of Set, I just picked arbitrary cards out of the layout and see if they form a Set.

This is the probably the only logical first strategy for anyone starting to play set. This is the part of their development to familiarize themselves with the rules and constraints of the game such that they can recognize sets quicker when they see it.

However, this is a poor long term strategy for finding sets. To understand why this is a poor strategy, we have to probability theory.

The first important thing to note is that given any two cards, there is **exactly** one cards that completes the set. For each of the attributes between the first two cards, they either have to be the same, or different, so the third card must have the attribute that has the same relationship as the first two cards. If the first two cards are both green, then the third card is the set must also be green. If the first card is red, and the second card is green, then the third card must be blue. You can apply this logic to all of the attributes and the conclusion you wind up with is that there is **exactly** one card that fits all the relationships since.

With that understanding, we now know that any random 3 cards will have a $1/79$ probability ($1/(81 - \text{the last two cards you pick})$) of forming a set. The first two cards can be arbitrary, but what you pick as the 3rd card has to fit the above requirements, and out of the cards that are left, there is a $1/79$ chance that you will find a set.

That doesn't sound so great. If you can evaluate a set of 3 cards to determine if they are a set on average of 2 seconds, then in a 5 minutes, you're going to find an average of 2 sets. That's not amazing considering most set games end in 5 minutes. Not only that, on a layout of 12 cards, you're going to evaluating $12 \text{ C } 3 = 220$ sets of 3 cards to determine whether they a set.

This is an important milestone to accomplish because this strategy phase you are building your capacity to quickly evaluate the validity of sets. Newbies can take up to 3-5 seconds to reliably validate sets, but with some practice this could go down to half a second.

Let's see if we can do better.

Strategy #2: Pick 2 Cards, Look for the 3rd

Armed with the knowledge that any 2 cards has exactly 1 card that completes the set, you can quickly scan whether that 3rd card is in the pile. Instead of evaluating 220 sets of 3 cards whether they form a set, you just pair off cards one-by-one to see if you can find the 3rd set. The numbers are more in your favor this time at $12 \text{ C } 2 = 66$ pairs of cards. You just saved

yourself $220 - 66 = 154$ checks! Not only that, the chances that the 3rd card is in your current 12-card layout is $10/79$. Assuming it takes you 2 seconds to figure out what the 3rd card is given 2 cards, and 1 second to scan the layout looking for the 3rd card, then in 5 minutes you'll find on average 12.65 sets! That's almost 6 times as many sets!

Pursuing this strategy requires a different skill from the previous one, namely the speed at which you can produce the image of the 3rd card in a set given the first two. I can do it in about 0.5 seconds, but for newbies it might take up to 3 - 5 seconds. You might experience a small dip in performance from pursuing this strategy before you improve on your time again.

Could we do better?

Strategy #3: Look at the last 3 cards

When someone finds a set, new cards are used to fill in the cards that were in the set.

As you play more games of Set, you'll begin to notice that most new sets involve at least 1 of the new cards that is introduced. The reasoning approximately goes like this:

If there exists a set that consists solely of cards from the 9 cards remaining, there is higher chance someone (myself) included, would have found it. The fact that we spent 10 seconds looking at the card layout and not finding a set in that group of 9 nine cards indicates that a set probably doesn't exist in those 9 cards. Therefore if a set were to exist, it would most likely include at least one card of the three that are being introduced.

Usually the probability that a set exists decreases as time goes by. This is usually intuited by experience, but we can dig a bit deeper about what signals that the causality exists (i.e. more time = lower % set exists). I like to think about it as "mental cycles." Just like computers perform computer cycles, there is an approximation of measure for human that perform "mental cycles."

At this stage of development you begin to think about where mental cycles are "spent." You're looking for patterns that no one else has found yet (that's the objective of the game!). As cards are introduced to fill the spaces of the last set, it's important to note that the mental cycles spent on the previous 9 cards are disproportionately high compared to the new cards, so you're going to want to pay special attention to the new cards in the field.

Assuming there are no sets in the remaining 9 cards, to find the next set in a layout, you only pair off one card from the new with one card from the old to look for the 3rd card. With this strategy, you only need to look up at most $(3 * 9) + 1$ [for the set being the 3 cards introduced] = 27 sets to find a compatible set. This is a good improvement from the previous strategy of 66 pairs. Your set throughput is going to go from 2.53 set/minute to 6.18 sets/minute! This is also way better than strategy #1 at 0.4 set /minute. You're now better by a whole order of magnitude.

Strategy #4: Property Filtering

What you might find interesting to note is that with little practice, you can filter out cards by property, and only focus on cards with a particular property. For example, it should be very easy to find all blue cards, or all oval shaped cards. If you layer two properties, you can find subsets of subsets of properties, like if you wanted to find blue ovals, you can narrow by blue, then by ovals. With this, we can on average be faster than strategy 2 and strategy 3.

Instead of the 2-step process of:

1. Figuring out the 3rd card in 2 card setup and then
2. Scanning the field for the card

You can break the two processes down into individual properties, like so:

1. Pick an arbitrary property
2. Figure out what property the 3rd card needs to have
3. Scan the field for cards containing the 3rd property
4. If there is exactly 1 card, figure out if it's the set
5. If there are no cards, then the 3rd card doesn't exist
6. If there are multiple cards, return to Step 1

Although this strategy has more steps, the time savings you get comes from the amount of time you save by scanning the field for a single property at a time instead of figuring out the 3rd card and then scanning the card on the field. Given 2 cards, the odds of the 3rd card being in the playing field is 10/79. That means on average you need to lookup 7.9 sets before coming across a successful one. The time you save comes from the 7.9 dead ends before finding a successful 3rd card. If you're scanning for a particular property, the chances you'll have at no cards that fits the first arbitrary property is $(2/3)^{10} = 1.73\%$. The chances you'll have exactly one card that fits the property is about $(10 \text{ C } 1) (1/3) (2/3)^9 = 8.67\%$. Adding those cases up equates to around 10%, indicating a 10% chance that you'll be done with these two cards by now. It's more likely you'll have multiple cards passing the first property, at around 89.6%. The expected number of cards to have the given property is approximately $(1/3) * 10 = 3.33$ cards. Now repeating the process again with another property, the probability that you'll have no cards matching your second probability is $(2/3)^{3.33} = 25\%$. The probability of exactly one card matching is $(3.33 \text{ C } 1) * (1/3) (2/3)^{2.33} = 43\%$. The overall chances you'll failure or success by the second round is $(0.1) + (0.9) * (0.25 + 0.43) = 71\%$. So generally you can very quickly scan two properties on a field to test for membership of the 3rd card, instead of figuring out the 3rd card a priori, and then scanning it. The time you save from the 7.9 dead ends by only evaluating on average 2 properties (instead of 4), should be about half.

Simulations

I decided to write some simulation software to test out my theories. [Here's the code](#). I assigned a `mental_cycles` value for particular cognitive tasks, like finding a card on the field or calculating the 3rd card that completes the set. With those values hard coded in, I ran simulations which each agent plays 1000 trials. Here are the results:

Agent	Avg mental cycles per game
NaiveAgent	1851.74
Pick2RandomAgent	1166.02
Pick2CardsLastAgent	1146.65
PropertyFilteringAgent	1039.28
PropertyFilteringLastCardsAgent	1001.81

The results were not as extreme as I would hope, but ordinally they rank they are about the same. Some strategies are only dominate over other within a small margin of mental cycle costs, so it hints that either my mental cycle costs are off or that individual variability in specific cognitive tasks could translate into which strategy works best for them.

Another point worth mentioning was that I had hoped that by looking only at sets involving the last 3 cards you would achieve a large improvement, but that seems not to be the case by the margin I was looking for. One possible explanation is the criteria I set for the game not to admit new cards unless all the existing sets have been found. Counterintuitively, it actually makes picking random cards **more** likely to involve one of the 3 cards introduced. For some quick math, the probability of 2 random cards drawn by Pick2RandomAgent involving one of the new cards is $1 - (0.75)^2 = 44\%$. About half the time it'll involve the new cards. This probably is higher when it's 9 cards on the board, at 56%. Doing a quick simulation, the average number of cards on the board before a set appears seems to be 8.2262, which will increase the probability of picking a random last 3 card on the field.

Some other things are difficult to write agents for, as we'll discuss in the heuristics section below.

Heuristics

Now we look at heuristics for better set plays by relying on probability theory.

Heuristic #1: Most Constraining Variable

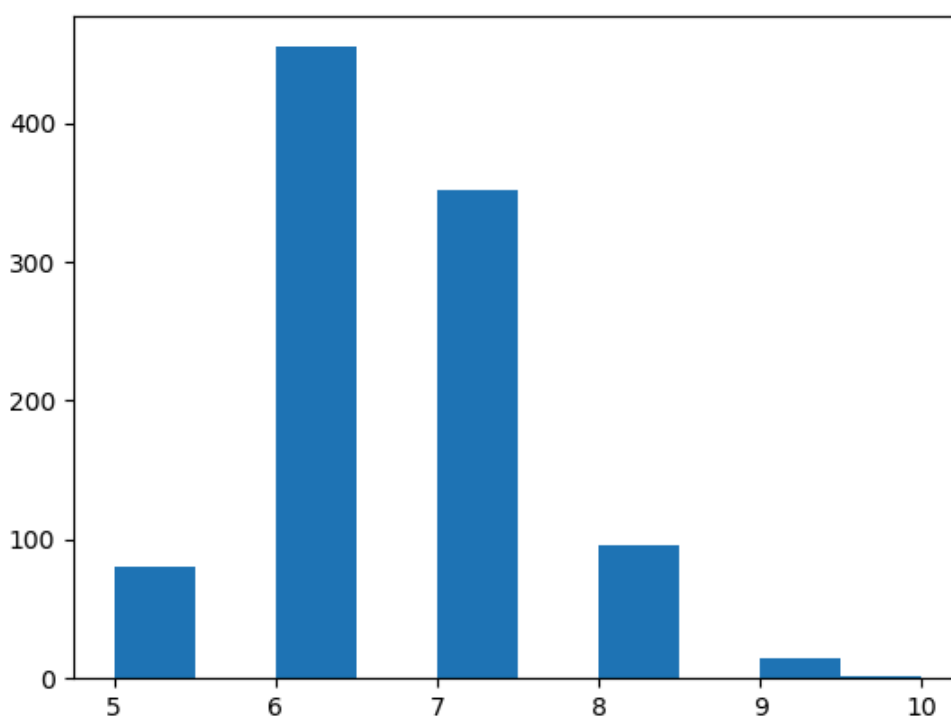
You see a Set field that has 9 reds, 2 blues, and 1 green. What are some intuitions you should draw?

If a set were to exist, it would have to either have to be all different or all the same.

In the first case, if all 3 cards in the set were different, then it would have to involve that green card and one of the blue cards. This kind of intuition has a name in the field of study of constraint satisfaction problems. This heuristic is called “most constraining variable.” When we want to satisfy a series of constraints, we want to pick the variable that has the least number of values. In this example it's the green. By starting with the green, then blue, then red, we can pick (green, blue_1) and (green, blue_2) to try to find the 3rd red card on the field. Since there are a lot of red cards on the field, you're more likely to have a set.

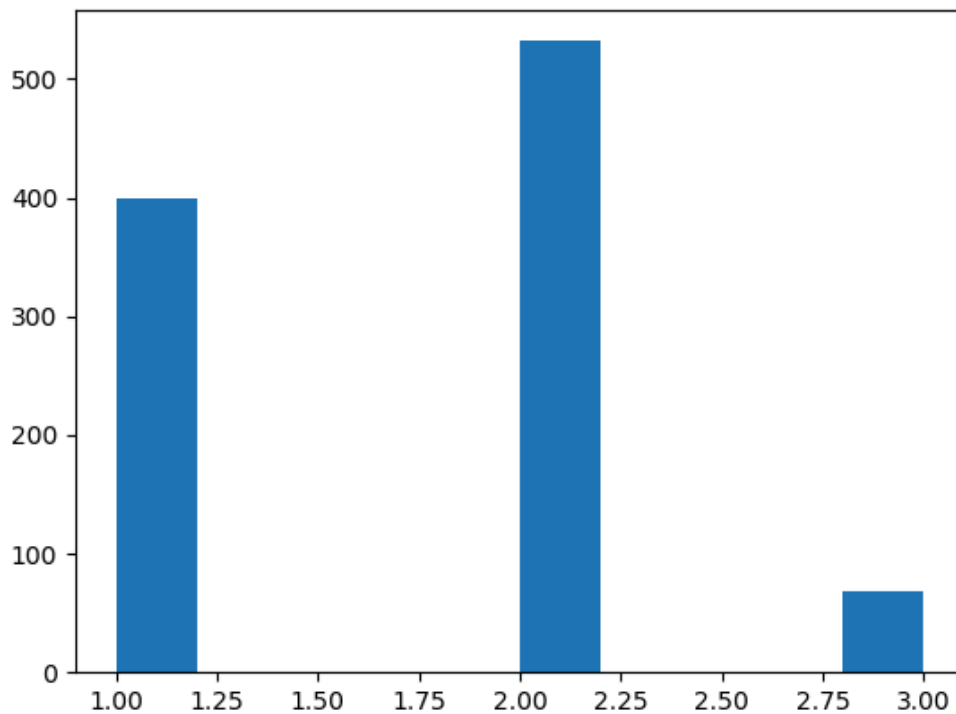
To belabor the point, let's see what happens if we reverse the prescription. We start with red, then green, then blue. We pick (red_1, blue_1), and figure out the 3rd card. But there's only one green on the field! What are the chances the red_1 and blue_1 you pick will have a 3rd card that IS the only green card on the field? You would have to iterate all 18 permutations of before coming across a dead end. Go the other way around and you'll save tons of time. It's not necessarily likely for color to be so lopsided, but out of 4 properties, there's a much higher likelihood for at least one of the 4 properties to be exploitable.

Here's a histogram of the number of the most common property in 1000 simulated Set fields.



The average of the most modal property is about a 6.53, which means this heuristic is only roughly exploitable in edge cases, and not the norm.

However, looking for the average least common property, we find a delightful surprise.



There is almost always 1 or 2 cards in the deck that is the sole carrier of a particular property. This means that either that card is more likely in the possible set (if the set contains all 3 different values of that property) or that the set exists in exclusively the set of cards that contain the property value that's not the unique one. The existence of this card cuts down your search space by quite a bit. The trouble is that sometimes it's difficult to find out which is that 1 or 2 cards that are unique and what properties they possess.

Heuristic #2: Understanding the probability distribution of property of sets

If each set is determined by 2 cards (making the 3rd card constrained), then we can evaluate the probability distribution of sets we find, so we can prime our brain to more heavily attune to those specific properties.

What are the total number of possible sets?

$(8 \text{ choose } 2) = 3240$

What are the total number of sets where all the properties are different?

The first card doesn't matter, so you can choose 81 possible cards. The second card has to all different properties, which for each property it can choose 2 alternative values that aren't the values of the first card, and you can permute those 2 values for each of the 4 properties. For example if I have a 3-red-solid-diamond card for the first card, my second card could have a 1 or 2 number, a blue or green color, a shaded or hollow shade, and a squiggly or oval shape, giving me 16 permutations. The total number of sets that fit this definition are then

$$81 \text{ (first card)} * 16 \text{ (second card)} = 1296$$

Oh wait, we have to divide by 2 because we treat the permutations as if order does matter by labeling the second and 3rd card, so we're essentially double counting. We have to divide by 2 to cut out the double counting problem.

$$1296 / 2 = 648$$

What are the total number of sets where 3 properties are different, and one is the same?

Following a similar logic, the first card doesn't matter, and we figure out the possible values for the second value. Out of 4 different properties, I can only choose 1 to be the same, so I'll just constrain one value and keep tabs of the possibility that that one value could represent (4 choose 1) possibilities. Now we calculate the number of remaining cards that could fit as the second card. It would be 2 for each of the different properties and 1 for the same property. Giving us $2 * 2 * 2 * 1 = 8$. Now we loop back the (4 choose 1), and the answer becomes

$$81 \text{ (first card)} * 4 \text{ (possible properties that could be chosen to be the same)} * 8 \text{ (permutations of the second card)} / 2 \text{ (double counting)} = 1296$$

What are the total number of sets where 2 properties are different, and two are the same?

Following a similar logic:

$$81 * (4 C 2) * 4 = 972$$

What are the total number of sets where 1 property is different, and 3 are the same?

$$81 * (4 C 3) * 2 = 324$$

Now we can figure out what the distribution of all the sets look like

$$\text{All different: } 648 / 3240 = 20\%$$

$$\text{3 different, 1 same: } 1296 / 3240 = 40\%$$

$$\text{2 different, 2 same: } 972 / 3240 = 30\%$$

$$\text{1 different, 3 same: } 324 / 3240 = 10\%$$


You will notice that the majority of sets that exist are going to be dominated by the middle two (3D1S, 2D2S). The variability is so high that when pursuing Strategy #4, you should consider

filtering on properties that are different rather than the same

Tangential Strategy #5: Multiplayer Games

I used to play Set with more than 3 players, but it turns out that's not a great way to determine who's the best Set player. The reasoning is that if players have strategies that correlate closely with one another will roughly split on the sets that they seek. For example if we have 2 players who start looking for sets containing one of the most recent 3 cards, while 1 player looks for sets where the 2 players aren't looking, it's very likely that that 1 player could win the game.

If you want to be fair, play 1v1's in a tournament style, with large sample sizes like best of 5 or something.

 May 8, 2025