Finding a Connection

This week's Fiddler on the Proof (8 December 2023) Extra Credit asks:

> In the New York Times game, Connections, there are 16 words that must be arranged by you, the player, into four distinct sets of four words each where the words in each set have some common property. Each turn, you must select four remaining words and press the "Submit" button. If they indeed represent one of the four sets, they are removed and you must try to find another set among the remaining words. The game ends when you have correctly identified all four sets.
>
> The game gives you hints. Anytime your four selected words include *exactly three* from a correct set, you are told that your selection is "One away…" from a correct set, meaning just one of your four choices was incorrect.
>
> Suppose you already identified two sets in your first two guesses, leaving you with eight remaining words. However, at this point, you have absolutely no idea what the remaining two sets might be.
>
> Nevertheless, with some clever strategizing—and by taking advantage of the "One away…" messages the game provides—it's possible to determine those remaining two sets in a relatively small number of guesses. (The game affords you only four incorrect guesses, but for the purposes of this puzzle, let's ignore that fact and suppose that you have an unlimited number of guesses.)
>
> Using your cleverest strategy, how many guesses would you need to be *certain* that you could *always* solve this puzzle, even in a worst-case scenario?

The situation described is equivalent to 4 red cards and 4 black cards lying face down in a random order. Your job is to guess which cards are the red cards and which are the black cards.

There are $\binom{8}{4} = 70$ ways that the cards can be arranged. A guess consists of choosing four of the eight cards. Each time you make a guess, there are three possible results:

1. "Zero away." You correctly guessed all four cards of one of the sets.
2. "One away." You guessed three cards of one set and one card of the other set.
3. "Two away." You guessed two cards of one set and two cards of the other set.

The Connections game doesn't actually say "Two away…". But with two sets to go, if you don't get a match and you don't receive a "One away…" message, you know your guess must have been "Two away."

An Algorithm

Here is a method of determining where the four red cards (or four black cards) are located in a maximum of 5 incorrect guesses.

For convenience, let's number the cards 1 through 8. For your first guess, pick cards 1, 2, 3, 4; for your second guess, pick cards 1, 2, 5, 6; and for your third guess, pick cards 1, 3, 5, 7.

If you are lucky, one of those guesses will exactly match a set. There are 6 ways this can happen (3 guesses × 2 matchable sets for each).

If you don't match a set with one of your first three guesses, it means you were either 1-away or 2-away on each guess. Let's write these results as "1" for 1-away and "2" for 2-away. Taken together, your three guesses produce three results which can be written as a triplet of 1's and 2's. There are 8 possible triplets: 111, 112, 121, 122, 211, 212, 221, and 222.

The next two guesses you make depend on the result of what happened with your first three guesses. Consult the following table to see what your next two guesses should be.

If the result of your first 3 guesses is ...	Then your next 2 guesses should be
111 or 222	1, 2, 4, 5 and 1, 3, 4, 5
112 or 221	1, 2, 3, 5 and 1, 3, 4, 5
all other triplets	1, 2, 3, 5 and 1, 2, 4, 5

Once you have made a total of 5 guesses as outlined above, you have enough information to know the locations of the two sets. The following table shows the location of one of those sets. (The other set, of course, is the complement of the first set.)

Result of your first 5 guesses	One of the sets is
11111	1, 2, 3, 5
11112	1, 2, 4, 6
11121	1, 3, 4, 7
11122	1, 5, 6, 7
11211	1, 2, 4, 5
11212	1, 2, 3, 6
11221	1, 3, 4, 8
11222	1, 5, 6, 8
12111	1, 3, 4, 5
12112	1, 2, 3, 7
12121	1, 2, 4, 8
12122	1, 5, 7, 8
12211	1, 6, 7, 8
12212	1, 2, 3, 8
12221	1, 2, 4, 7
12222	1, 3, 4, 6

Result of your first 5 guesses	One of the sets is
21111	1, 2, 5, 7
21112	1, 3, 5, 6
21121	1, 3, 7, 8
21122	1, 2, 6, 8
21211	1, 2, 5, 8
21212	1, 4, 7, 8
21221	1, 4, 5, 6
21222	1, 2, 6, 7
22111	1, 3, 5, 8
22112	1, 4, 6, 8
22121	1, 4, 5, 7
22122	1, 3, 6, 7
22211	1, 4, 5, 8
22212	1, 3, 6, 8
22221	1, 2, 7, 8
22222	1, 4, 6, 7

There are 32 possible results for your five guesses. Each result leads to two sets, for a total of 64 sets. Add to that the 6 possible sets that match exactly against your first 3 guesses and you have a total of 70 locations for the sets. That accounts for every possibility.

Example

Suppose the hidden cards are arranged as follows:

1	2	3	4	5	6	7	8

The red cards are those numbered 1, 3, 6, and 7 and the black cards are those numbered 2, 4, 5, and 8. Our goal is to figure out where each set is located. We do this by making guesses and observing the results.

The algorithm says that the first 3 guesses we should make are cards 1, 2, 3, 4, then cards 1, 2, 5, 6, and then cards 1, 3, 5, 7. Here are the results of those guesses:

	1	2	3	4	5	6	7	8	
									Result
1st guess	↑	↑	↑	↑					**2**
2nd guess	↑	↑			↑	↑			**2**
3rd guess	↑		↑		↑		↑		**1**

(Recall that a result of "2" means that 2 cards of each set match our guess and "1" means that 3 cards of one set match our guess.) We write this combined result as 221.

When the first three results are 221, the next guesses we should make are 1, 2, 3, 5 and 1, 3, 4, 5.

	1	2	3	4	5	6	7	8	
									Result
1st guess	↑	↑	↑	↑					**2**
2nd guess	↑	↑			↑	↑			**2**
3rd guess	↑		↑		↑		↑		**1**
4th guess	↑	↑	↑		↑				**2**
5th guess	↑		↑	↑	↑				**2**

The full string of all five results is 22122. The table above says that for this result one of the matching sets of cards is 1, 3, 6, and 7. And, sure enough, those are the four red cards!

Answer

By using the algorithm above, you are guaranteed to know the locations of the two sets after a maximum of 5 guesses. (You may still need two more (correct) guesses to remove those sets from the board.)