Suppose we're given three urns. One contains twelve white balls, another contains twelve black balls, and the third contains six white and six black balls. We call the first two urns **monochromatic** and the third **mixed**. The urns are labelled A, B, and C, but we don't know which is which. The goal is to identify the urns by pulling out balls to check their color. (Once a ball is drawn it stays out forever; no ball is ever put back into an urn.)

If we're lucky, we might be able to identify the urns by drawing only three balls: First pull two balls from urn A. If they're of different colors then A must contain the mix, and drawing one ball from B suffices to identify the other two urns. But if the two balls from A are the same color we're stuck. EXERCISE 1: With this procedure, what is the probability that the two balls from A will have different colors?

The problem is to devise a procedure that always identifies the urns regardless of what balls are drawn, and that draws as few balls as possible in the worst case. SPOILER WARNING. In the next paragraph we give the solution; stop reading here if you want to try it on your own first!

The problem can be solved by drawing at most nine balls, as follows: First draw one ball from each urn. There will necessarily be two of one color and one of the other; say that A and B yield white balls and C's ball is black. Clearly C contains only black balls. To identify which urn contains the mix, pick one of the other urns, say A, and draw additional balls only from that urn. Of course if a black ball shows up then A contains the mix. And if we draw six more white balls from A, for a total of seven, then A contains only white balls and B must have the mix. In no case do we draw more than nine balls. EXERCISE 2: With this procedure, what is the *expected* number of balls that are drawn before we know the answer?

An observation: This procedure depends critically on something we haven't made explicit, namely, that we are permitted to decide from which urns to draw balls *after* seeing the results of earlier drawings. If this weren't true, that is, if we had to say *in advance* how many balls to draw from each urn, then the procedure described above wouldn't work, and fourteen balls (seven from each of two urns) would be required to solve the problem.

Now change the problem slightly. Retain the urns, their contents, and the goal of identifying all three urns, but suppose we're permitted to draw at most k balls total, for some $k < 9$. What procedure *maximizes the probability* that the urns can be correctly identified?

If $k = 0$, that is, we can't draw any balls, the problem is trivial. All we can do is guess, and we have probability $1/6$ of doing this correctly.

If $k = 1$ we can draw only a single ball. Suppose we draw a ball from urn A and it's white. The best we can do is to say that A contains only white balls, and then guess randomly which of the other two urns is the mix. EXERCISE 3: What is the probability of success with this procedure?

If $k = 2$ there are two possibilities. The first is to draw two balls from the *same* urn. If the balls have different colors, then definitely this urn contains the mix, and we have to guess about the other two. If the two balls are the same, say white, then we're best off saying that the urn is all white, and again we guess about the other two.

The second possibility for $k = 2$ is to draw two balls from *different* urns. If those balls are the same, say white, then the untouched urn must be all black, and we guess about the two from which we've drawn. But if the two balls are different, the best we can do is to guess that each ball was drawn from a monochromatic urn and that the untouched urn contains the mix.

Which of these two procedures is better? EXERCISE 4: Calculate the probability of correctly identifying all three urns in each case, in order to tell which procedure wins for $k = 2$.

Now skip to $k = 8$. We might start by following the solution for $k = 9$: Draw one ball from each urn, identifying one of the monochromatic urns, then draw balls from only one of the remaining urns, say from A. If we get a ball of a different color, we're done: A is the mix. But if we draw five more balls from A, all the same color, then we have to stop because we've drawn our limit of eight balls. At that point the best we can do is guess that A is the other monochromatic urn. EXERCISE 5: Using this procedure, what is the probability that we'll correctly identify the urns? EXERCISE 6: A tiny modification to this procedure improves the odds further. Find it, and calculate the new probability of success. Is this the best we can do?

When $k = 3$ there seem to be three possibilities: One ball from each urn, three balls all from the same urn, or one ball from one urn and two from another. But the situation is more complex, because as discussed in the "observation" above we have the flexibility to change procedures midstream. So, for example, we might draw one ball from each of two urns, and only then decide whether to draw the third ball from the third urn or from one of the first two. EXERCISE 7: What is the best procedure for $k = 3$?

EXERCISE 8: Solve the problem for $4 \leq k \leq 7$. Is there a pattern? Is there an obvious way to see the pattern?

EXERCISE 9: Generalize the number of balls initially in the urns. The most interesting case: What happens when there are an infinite number of balls in each urn, so that you can never identify the urns with certainty? [Thanks to Andy Latto for raising this question.]