



PROGRAM: 2
Combinatorics Counts

Producer: Sam Ward
Host: Dan Rockmore

Produced by Oregon Public Broadcasting for Annenberg Media

Time Code	Audio
00:15	OPENING CREDITS
00;40	HOST So many places to go, and so many ways to get there.
00;49	HOST Sometimes there are too many choices... How do we count all those things, and then organize them so we can make sense of it all? By looking to the mathematical world of combinatorics
01;04	HOST (V.O.) Meet Charlie, the regional claims adjuster for an insurance company. On the road, as usual, visiting clients, solving problems...and trying to get to five different job sites as quickly and efficiently as possible. So he can get home to his daughter Sarah's birthday party...
01;19	HOST (V.O.) His problem — finding the best way to organize his trip, taking into consideration the distance between each client, possible time spent with each client, and the most efficient route.
01;29	HOST (V.O.) Sometimes the number of possibilities seems overwhelming, even when we're faced with only a few choices, but we have to make decisions anyway.
01;38	HOST (V.O.) Did he make the right choice? Let's count the ways...
01;43	HOST Counting things seems so simple. Children do it intuitively, connecting a thing with fingers to say how many. It's the way we organize the stuff of our lives, from bugs to errands. And finding efficient and interesting ways to organize things and information is what the field of mathematics, known as Combinatorics, is all about.
02;03	KIDS (together) <i>As I was going to St. Ives, I met a man with seven wives, each wife had seven sacks, each sack ...</i>
02;12	HOST Hey, hey hold on kids, Going somewhere are you?
02;14	KIDS Yeah HOST

	<p>Alright, well, make sure you return, okay?</p> <p>KIDS OK</p>
02;21	<p>KIDS (together)</p> <p><i>As I was going to St. Ives, I met a man with seven wives, each wife had seven sacks...each...</i></p>
02;32	<p>HOST</p> <p>Even with the fairly simple system of colored buttons we've already organized information in a relatively sophisticated way. And we started with a common, intuitive form of organization — the act of counting, or enumeration.</p> <p>Enumerating, or listing the children, and then encoding that enumeration with these buttons.</p> <p>But there are lots of ways in which we could have assigned children to buttons and they encode a variety of levels of information, each corresponding to counting different kinds of things about the children.</p> <p>Abstractly, in using the buttons, I've turned the children into a mathematical object - a set. And the fact that it can be a little easier to count buttons in a jar -or count properties of a set of objects than it is to count kids running around the neighborhood is just the smallest of hints at the power of Combinatorics.</p>
03;22	<p>HOST</p> <p>Counting things, that's what combinatorics is about. As an individual field of study, it's a relatively new one. And these days, with its strong connection to the world of computing, by providing insights into how best to organize and understand the power of the computer, it is incredibly important for modern technology.</p> <p>However, as cutting edge as the subject may be, its basic concerns go back to puzzles and problems from the earliest recorded evidence of mathematical thought.</p>
03;50	<p>HOST</p> <p>Now, this is a facsimile of the Rhind Papyrus, copied around 1850 BC, by the scribe Ahmes, from the now lost text of an earlier dynasty. It was named after Scottish antiquarian Alexander Henry Rhind who bought it in 1858 in Luxor, Egypt.</p>
04;07	<p>HOST (CONT'D)</p> <p>In it, he found what amounts to a math textbook, full of problems and solutions that addressed everyday issues.</p>
04;14	<p>HOST</p>

	Things like...how to divide 10 loaves of bread equally among 7 men, and much of the algebra they needed to build the pyramids. It also contains possibly the earliest written combinatorics problem– and it goes like this.
04;28	<p>HOST (V.O.) <i>Seven houses contain seven cats.</i> <i>Each cat kills seven mice.</i> <i>Each mouse had eaten seven ears of grain.</i> <i>Each ear of grain would have produced seven hekats of wheat.</i> <i>What is the total of all these?</i></p>
04;40	<p>HOST A fairly simple math problem...and a very pragmatic combinatorial argument for getting more cats.</p> <p>In fact, it's not that far removed from the St. Ives riddle our children at the gate were talking about. Both of these problems have elegant solutions by the recognizing of the simple pattern of the powers of seven that we see.</p> <p>Now several centuries later we find slightly more sophisticated kinds of problems and solutions in 6th century BC India.</p> <p>The medical treatise Sushruta Samhita gives us this problem:</p>
05;11	<p>HOST (V.O.) How many combinations can be made out of six different tastes — bitter, sour, salty, sweet, astringent, and hot?</p>
05;20	<p>HOST By taking them one at a time, two at a time, three at a time, and so on, we learn there are in fact 63 combinations.. This was important in ayurvedic medicine because it gave ancient physicians a way to work out which medical substance would best match a patient's ailments</p>
05;36	<p>HOST (V.O.) Now there are two things that are going on here.</p> <p>On the one hand, we're asking for the number of possible combinations or subsets of the six different spices.</p> <p>This number, as we see is 63, which is two to the sixth power minus 1. Notice the sixth power and 6 spices. There's a pattern there, number goes to power...the magic is that we're seeing a formula - a quick way to answer the question of how many combinations of spices, without actually listing them!</p> <p>HOST Now there is a more particular problem that was of interest to the Indian doctors and that was the individual questions of how many combinations of one spice and how many of two and so on...again, we could solve simply by listing them all, but much more</p>

	<p>interestingly, there is a pattern or formula. And three hundred years later, Indian scholars wrote the algorithm for solving this, again without counting, an algorithm for finding that pattern... and in this, gave us one of the earliest written statements of combinatorial rules, and not simply problems.</p>
06;40	<p>HOST So, we jump ahead now a few centuries and find mathematician and scientist Blaise Pascal who, in 1655, rediscovered this combinatorial rule in what was later named Pascal's Triangle.</p>
06;55	<p>HOST (V.O.) This triangular array of numbers is built in a fairly simple way.</p> <p>We start with a 1 at the top...</p> <p>...and imagine there's a zero on either side of the one.</p> <p>And we're going to call this row zero.</p>
07;22	<p>HOST (V.O.) Now, we add the pairs of adjacent numbers and write their sum between them on the next row: zero plus 1 equals 1, 1 plus 1 equals 2. 1 plus zero equals 1.</p> <p>We've completed Row 1.</p> <p>Now we go to row 2, and we do the same thing. We have a 0 and a 1, so beneath it we put a 1. Now we have a 1 and a 1, and between and beneath it we put a 2. Now we have a 1 and a 0, and in between and beneath it we put a 1.</p> <p>Taking zeroes on to either side, we've now built row 2.</p> <p>And so on and so on and so on.</p> <p>Now, lop off the zeros on either side of the triangle. What you see there is what we've seen before, Pascal's Triangle.</p>
07;47	<p>HOST Let's say it's your cousin Louie's birthday, and he's crazy about hats.</p>
07;52	<p>HOST You're considering a birthday gift of 5 hats, but you can only afford to buy 2, and you can't decide which ones.</p> <p>HOST (V.O.) So the problem is, how many different combinations of 2 hats are there, and that way you know how many pairs you have to consider.</p>

08;06	<p>HOST (V.O) When we're working with Pascal's Triangle, we can actually read this information off the triangle.</p> <p>So now what we need to do is go to row five, and remember we started numbering the rows by zero, so we actually have to count down six rows to row five.</p> <p>And again, the columns also start with zero, and we go in to the second column, which actually requires zero, one, two, so the third entry.</p>
08;30	<p>HOST (V.O.) And there it is, the number we find is 10....and that's the right answer. And this number, the number of ways to choose 2 things out of 5 things, a mathematician calls 5 choose 2, so 5 choose 2 is equal to 10.</p>
08;42	<p>HOST (V.O.) Now, Pascal didn't invent all the ideas behind this triangle. He did what all mathematicians do — he built on knowledge from an earlier time...</p>
08;49	<p>HOST (V.O.) ...and then pushed the boundaries of that knowledge to give us something even more interesting and, as is often the case with mathematics, more useful.</p>
08;58	<p>HOST Combinatorial problems and puzzle test our ability to think through problems and logically come up with efficient ways to organize information.</p>
09;06	<p>KIDS <i>...Kits, cats, sacks and wives, how many were going to St. Ives?</i></p> <p>HOST So what is the answer to this riddle? Let's find somebody who might help us figure it out...</p>
09;19	<p>Dan Rockmore: So how do we make sense of the confusion caused by all these kids? Well, I'm lucky to have here with me Jenny Quinn, who's a professor of interdisciplinary arts and sciences at the University of Washington-Tacoma.</p> <p>Jenny, thanks for coming.</p>
09;32	<p>Jenny Quinn: Thank you, Dan. Nice jar.</p>
09;34	<p>Rockmore: Well, here are these buttons, and as far as I'm concerned, one button per kid. And I happen to have their pictures here</p>

09;46	Quinn: So what we have here are two sets of six things.
09;49	Rockmore: Yep. Six buttons; six photographs. Right.
09;52	Quinn: So a set is just a collection of objects, okay?
09;53	Rockmore: Mm-hmm.
09;54	Quinn: ...what you did here was you created a bijection between the set of buttons and the set of kids.
10;00	Rockmore: Right. One button per child.
10;03	Quinn: One button per child. One thing from each set matched up, holding hands with something from the other set.
10;08	Rockmore: Yep, yep.
10;09	Quinn: Okay, and this gives us a lot of information, okay?
	Rockmore: Yes.
10;10	Quinn: So it's a question of how much information we actually want to keep track of.
10;16	Rockmore: I see, because-because we could assign the buttons in many different ways, in fact.
10;19	Quinn: yes, because notice, each one of these buttons is actually distinct.
	Rockmore: Right.
10;23	Quinn: So if we paid attention to which button went to which child, okay, we would be able to say if one of the buttons were missing, which missing button it was and what child that corresponded to, okay.
	Rockmore: Right, right.
10;32	Quinn: But we might not want to keep all that information because there are 720 possibilities of ways to match up buttons to kids. So we could choose to consider smaller subsets to pay attention to. Now, a subset is a-a collection, a smaller collection of the entire set. And in fact, we could ask that the orange buttons create a subset and they go to the orange-shirted kids, and the blue buttons go to a subset which go to the blue-shirted kids.
11;00	Rockmore: So it's actually, it's a bijection within a bijection, in fact, right?
	Quinn: Exactly...
11;05	Quinn: ...the problem with these kinds of-of bijections is, as the number of objects you're considering grows, the number of possibilities becomes unmanageable. And we have to learn how to keep track of more information or we have to pinpoint the information that's really important to us.
11;22	Rockmore: So in fact, I mean, there are many ways that we could have made this initial bijection, right, between the buttons and the -- and the children, the buttons and the photographs?
11;29	Quinn: Oh, absolutely. Because there are six different buttons, we had a choice of six buttons to give the first person, then there would be five buttons

	left to give the second person, four buttons for the third, three, two, one. So the total number of ways of creating this bijection would be the product of all of those: six times five times four times three times two times one, which is six factorial, or 720.
	Rockmore: Right, right.
11;51	Rockmore: And so if there had been more children, I would have needed more buttons, and this way in which you're counting seems to grow very, very quickly, right?
12;00	Quinn: Yes, yes. This is what we call a combinatorial explosion.
12;03	Rockmore: Combinatorial explosion, all right
12;04	Quinn: And it's in fact too much information for us to keep track of. So what we'd like to do instead is decide what information is most important to us. So combinatorics is going to help us organize our thoughts.
12;15	Rockmore: So now these bijections that we've been constructing between the buttons and the kids is in fact just the tip of the iceberg of -- for combinatorics, which is there are much richer bijections, bijections that encode much more subtle correspondences, isn't that true?
12;28	Quinn: Oh, absolutely. In fact, what we can do now is we can look at -- we can return to Pascal's triangle and we can see how the information about Uncle Louie's hats is actually encoded in the triangle.
12;40	Rockmore: Uh-huh, in the geometry and just moving around the triangle, is that right?
12;43	Quinn: Absolutely.
12;44	Rockmore: All right.
12;45	Quinn: So if you recall, uh, we were choosing two hats for Uncle Louie out of five, right?
12;48	Rockmore: That's right.
12;50	Quinn: And it corresponds to that position (5 choose 2) right there.
12;55	Quinn: So imagine yourself starting at the top at the zero-zeroth position, okay? Now, each time you take a step, you're either going to go to the left or to the right.
	Rockmore: Okay.
13;05	Quinn: And when you take a step to the left, that means the hat in that row is not selected. And each time you take a step to the right, it means that's one of the hats you're going to give Uncle Louie.
13;17	Quinn: So starting from the top and going to the 5, 2 position, it's going to take two steps to the right and three steps to the left to end up down in the 5, 2 position.

13;29	Rockmore: Right, equal to 5 choose 2, as we discussed in the beginning.
13;32	Quinn: So this actually tells us a little bit more about the triangle itself and the construction of the triangle, because if you think about it, suppose you wanted to pick any position in the triangle and understand where it came from.
13;40	Rockmore: Okay.
13;42	Quinn: Let's call it N choose K.
13;44	Rockmore: N choose K for generally the Nth row and the Kth entry in.
13;50	Quinn: So the way to think about it is we're trying to create subsets of size K from N elements total.
13;56	Rockmore: Right.
13;58	Quinn: And the way we can think about it coming into this triangle through these paths is we either entered this cell from a path going from the right or a path coming from the left.
14;08	So you can think about this particular cell as being created as the sum of two possibilities, either the -- for the row that it came before, either the object was included or the object was excluded. So the way to include the object is N-1 choose K-1 ways, and the way to exclude the object is --
14;29	Rockmore: N-1 choose K.
14;31	Quinn: Exactly.
14;32	Rockmore: So that actually encodes what mathematicians would call as a recurrence.
14;36	Quinn: Exactly.
14;37	Rockmore: So this -- so this recurrence that we derived, really, of N choose K is equal to N-1 choose K-1 plus N-1 choose K is actually an example of a very fundamental combinatorial principle, isn't that right?
14;50	Quinn: Absolutely. And here's my favorite example of a principle. It's called the pigeonhole principle.
14;55	Rockmore: Ah, sounds great.
14;56	Quinn: And the reason it's my favorite is because it's very intuitive. So here's the example I'm going to show you. We have five pigeons, and they are flying home. These are courier pigeons so they are flying to their home. There are only four homes.
15;03	Rockmore: Right, right, so somebody might get left out, is that right?
15;14	Quinn: Unfortunately, if there are more objects than there are boxes and only one thing is in a box, then there's stuff that doesn't get assigned.
15;21	Rockmore: Okay, good.

15;22	Quinn: But, to make that pigeon happier, we're going to say that everybody has to be put in a box. And what that's going to do is it's going to force that pigeon into some box that's already occupied. So the pigeonhole principle in this case is going to say that if there are five pigeons in four boxes, some box -- no matter how the pigeons distribute themselves among the boxes, some box is going to contain at least two pigeons.
15;50	Rockmore: Right. So some box might actually have three pigeons, right? I mean, that's also possible.
15;53	Quinn: Right, but the box with three pigeons also contains two, okay? So that's at least two. And in fact, if we take the average of the number of pigeons to the number of boxes, okay, what the pigeonhole principle says is some box contains at least the average.
16;11	Rockmore: so the average number of pigeons per box here would have to be -- would simply be the number of pigeons, which is five, divided by the number of boxes, which is four.
16;18	Quinn: But because we can't have a quarter of a pigeon in some box --
16;21	Rockmore: Right. No pigeons are harmed during the making of this show, right?
16;24	Quinn: Right. They're discrete objects, we have to round up. So that tells us we're going to get two pigeons per box in this case.
	Rockmore: Right, at least two. Yep.
16;39	Rockmore: So these basic principles that we've been outlying are all actually -- and this is something we haven't talked about so much -- but they're -- these are numbers that are easy to compute. Like we've been finding formulas, in fact, computing the numbers of pairs of hats or computing things like the average for the pigeonhole principle.
16;45	Quinn: In particular because they're small.
16;47	Rockmore: But as we get more possibilities, we begin to experience what we've been calling this combinatorial explosion, and we're going to see an example of that now as poor old Charlie gets a new itinerary. So let's take a look.
17;00	HOST (V.O.) Charlie seems to have found his solution. At least he's crossing one off, and that's a start
17;07	HOST (V.O.) Except now his boss has called and added more stops to his itinerary. Eight more, to be exact. Now he's got 12 places to

	organize.
17;16	HOST (V.O.) And it's beginning to look like there are so many possible ways to go, he'll never find a efficient route that will let him get home in time...
17;26	Rockmore: Poor Charlie, he really seems to have his hands full here. He's suffering from what we were calling before a combinatorial explosion.
17;33	Quinn: Well, poor Charlie doesn't have a prayer in the world of answering his question definitively, okay, but there are ways that we can cleverly approach a problem and organize the information so that we can reduce the amount of work that needs to be done. Rockmore: Is that right?
17;46	Quinn: Here's an example for you, okay? The set-up is we have a keypad with 10 digits and we want to input a three-digit code to unlock the door.
	Rockmoreiel: Okay.
18;00	Quinn: Okay? So if you were to do this in sort of a very natural, methodical method, you'd put in 0-0-0, you'd put in 0-0-1...
18;06	Rockmore: Make your way through.
18;07	Quinn: And you test every number from 0 to 9-9-9, okay, and that might take you a while.
18;13	Rockmorel: Right, and that's a thousand codes, and I have to hit -- for each code, I hit it three times, that's 3,000 pushes, right?
18;19	Quinn: Exactly. Okay. Now, if you use something called a De Bruijn sequence, okay, this is a way of organizing this information so we really reduce the amount of work that we're going to do. We're going to cut the amount of work by a factor of 10. So in a De Bruijn sequence, we can organize the list of numbers in such a way that every three digits that we want to read is a different set, a different combination.
18;44	So you might start with 0-0-0, test those first three, and then hit 1, that's going to be -- 0-0-1 is now going to be tested.
18;53	Rockmore: I see, so I have this -- so I'm sort of sliding a window along. This is 0-0-0, and then I would slide it one thing. In other words, I'd press a 1, so that would mean 0-0-1 would be the next code that I was testing.
19;03	Quinn: Exactly. And now by organizing it in this fashion, we've reduced the amount of work by a factor of 10, so if it took five hours to test all thousand numbers, it's now going to take on the order of 30 minutes.

19;18	Rockmore: Great, math to the rescue.
19;20	Quinn: So in fact, De Bruijn sequences are not only useful to get into a door, it's also being used right now in genomic analysis.
19;30	Rockmore: Uh-huh, yes. Genomics generally is a big user of combinatorial work as sort of discrete units of the genome, and we have an example of that here. Let's watch.
19;39	I'm Terry Gaasterland, and I'm a professor of computational biology at the University of California-San Diego
19;46	I'm in the marine biology department there and
19;49	I have collaborators in New York and I have collaborators here in Oregon.
19;54	You can think of DNA in many different ways. If you're a computer scientist, you can think of it as ones and zeroes or you can think of it as four letters, "A", "C", "G", and "T". If you're a biologist, you think of DNA as a long complex molecule that is made of each of the compounds that A, C, G, and T represent.
20;17	A genome is the specific bundle of DNA that's in a particular cell.
20;23	So, we can think of DNA as being paragraphs, words, sentences, where a genome would be the specific book that is you.
20;33	DNA sequencing is the process of trying to read the letters that are in the book.
20;40	So sequencing is the process of fragmenting that DNA into little tiny pieces and reading a little bit of each of the fragments and now you've got all these 500, 700 letter words and you need to fit them all back together.
20;55	<p>Early on, when people started trying to sequence genomes, people came up with the most interesting and innovative ways to do it.</p> <p>... there was the process of directed primer walking. So, randomly go into the genome, do a read, and then design a little probe that will put you at the end of that read, and then keep reading. It turned out that the problem there is that genomes are very repetitive</p> <p>So every one of these ran into combinatorial problems and it turned out that the idea Craig Venter was really pushing, finessed all of this combinatorics.</p> <p>Take the whole entire genome, chop it up into lots and lots and lots of little pieces and take reads off of every one of those and do so many that you have enough information to fit those pieces back together.</p>
21;53	<p>So for any given book, if you will, you have multiple copies of the book and one you're ripping in half at chapter one and the other you're ripping in half at chapter four.</p> <p>...and if you do enough of them, it will cohere into one: "Given this fragment</p>

	and given this fragment, 200 bases overlap, so we'll put them together and see if we can put the next one in." And if those two overlap and then you have a third one that comes in and it matches, you know you have it right.
22;20	With a genome like the human genome, which is three billion As and Cs and Gs and Ts, you need to be a little smarter.
22;30	So now you set all sorts of heuristics, so you say, "Well, let's take every pair of reads that overlaps by at least say, 100 letters," and they have the identical hundred letters. The problem is there's always sequencing error ... So, for every one of those comparisons, you're computing all these details about it. So that is the problem that the computer has to solve as it's doing assembly of all of these pieces
22;56	When I first started in looking at genomes, the only genomes that were being sequenced were one-celled microbes, bacteria and archae,
23;07	...proteins from one organism would have absolutely no match whatsoever in the next organism over, so their seemed to be proteins that were absolutely unique to each of these organisms. So that made people step back and say, "Well, we need to sequence more organisms and fill in all the holes and the gaps in this tree of life." So now we go from one-celled organisms to more complicated organisms, with several or many cells or even with differentiated cells like a sea urchin or human.
23;42	So gene by gene, we can look across the evolutionary tree and ask, "Which gene in sea urchin goes with which gene in zebra fish and that goes with which gene in human?" and by studying when that gene turns on and turns off, during development or in maintenance of a healthy organism, in sea urchin, that tells us something about how it might be working in human.
24;10	Rockmore: Wow, so genomics, this is a particular example where combinatorics has really helped us. You know, it's allowed us to simplify a problem, and it's really given us a purchase on how to do things, but there are actually problems which are combinatorial in nature which we still don't know how to solve, isn't that right?
24;26	Quinn: Well, like Charlie, for instance.
24;28	Rockmore: Poor Charlie, yep.
24;29	Quinn: He's in what we call a traveling salesman problem.
24;32	Rockmore: Yes, he is.
24;33	Quinn: The idea being you're given a certain number of cities, and you want to travel -- starting from one city -- you want to travel to all the cities, getting back to where you started, in the most economical fashion, okay. Now if there are a small number of cities, the analysis is not difficult. So if Charlie only had three cities, okay, there are only two possibilities: either he goes clockwise from where he is or he goes counterclockwise, okay?
24;59	Rockmore: So each route has a cost as well.

25;00	Quinn: Each route has a cost.
25;01	Quinn: Okay, but now if we add a city to that, okay, his two choices --
25;06	Rockmore: Oh, I see, yep.
25;07	Quinn: -- go to six, okay. If we add another city
25;12	Rockmore: Okay, it's getting worse.
25;13	Quinn: -- it goes from 6 to 24.
25;14	Rockmore: I'm seeing a pattern here. Is there one, in fact?
25;17	Quinn: Oh, there absolutely is. If we jump to 10 cities, there are now over 350,000 possible organizations to his route.
25;27	Rockmore: And that factorial number is coming in here again, isn't it?
25;30	Quinn: Exactly. So if you have 10 cities, it gives us 9 factorial possible routes.
25;34	Rockmore: Okay, 9 times 8 times 7 and so on. Okay.
25;36	Quinn: If we jump to 25 cities --
25;39	Rockmore: Okay, that really looks like a problem.
25;40	Quinn: -- there are billions and billions of routes. And no matter how clever we are, there's no way we can look at every single one and find the best or fastest possible route,
25;51	Rockmore: All right. Well, Quinn, this was great. Thanks.
25;53	Quinn: My pleasure.
25;55	HOST (V.O.) So Charlie is left with too many possible solutions, more than he can calculate with a simple map and a pad of paper, or even his laptop.
26;04	HOST (V.O.) But sometimes, the path is clear
26;07	HOST (V.O.) And sometimes the right way to organize information is the way that leads us home.
26;18	KIDS <i>Kits, cats, sacks and wives, how many were going to St. Ives?</i> HOST Our lives are constructed out of seemingly endless possibilities. We make choices so we can be more efficient in what we do and where we go... so we can unlock the secrets of the human genome, or simply solve a math riddle.
26;38	HOST The answer to the riddle, by the way, is one. I'm the only one going to St. Ives. The other combinations, we'll save for another time.

26:55

CREDITS