A Lesson in Randomness and Runs

Created by

Richard Born

Associate Professor Emeritus

Northern Illinois University

richb@rborn.org

Topics

Robotics

Probability and Statistics

Computer Science

Ages

Junior and Senior High School Students

Duration

approx. 1 hour

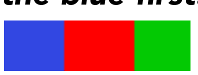
A LESSON IN RANDOMNESS AND RUNS

TEACHER'S GUIDE

What will students learn?

- How is Ozobot programmed by default to respond when encountering an intersection?
- What do the concepts of *random* and *probability* mean?
- What are the theoretical probabilities for specific events while Ozobot traverses its map?
- How can you perform an experiment and collect statistics to support the theoretical probabilities?
- In the context of random events, what is meant by a *run*? What are the probabilities of runs?
- Given two sequences of events, one sequence from 40 tosses of a fair coin, and the other sequence produced by asking someone to simply write down what they think is a random sequence of heads and tails, how might one be able to tell which sequence is the actual coin toss sequence?
- What is the so-called gambler's fallacy—which is also known as the Monte Carlo fallacy?

Explanation of OzoCodes Used in This Lesson

-  The ***Go Straight*** OzoCode tells Ozobot to continue going straight when it approaches the intersection following the code. This code is a ***one-directional*** code (not a symmetric code), so the order of the three colors in the code is important—Ozobot should be moving toward the blue first.
-  The ***Go Right*** OzoCode tells Ozobot to turn right when it approaches the intersection following the code. Like the Go Straight code, this code is one-directional, so the order of the three colors in this code is important—Ozobot should be moving toward the blue first.

-  ***Fast*** is one of the speed OzoCodes. It is a symmetric code that tells Ozobot to go somewhat faster than the default cruise speed.

What is a run?

Divide the class into two groups. Ask each member of one group to toss a coin 40 times, recording the H/T results on a copy of the data table at the top of the next page. Ask the other group to write down on the data sheet what they think the random H/T sequence of tossing a coin 40 times might be, *but doing this without actually tossing the coin.*

Now discuss the concept of a run with the class by using the following example. Suppose that you toss a coin ten times, obtaining the following sequence of heads and tails:

HTTTHTTHHT

Any unbroken sequence of like letters is called a ***run***, even if the sequence has only one letter. In the above sequence there are six runs: H, TTT, H, TT, HH, and T, from left to right in the sequence. The ***length*** of each of the runs (i.e., the number of letters in the run) is 1, 3, 1, 2, 2, and 1, respectively.

Next, ask each student to identify, for his/her sequence of 40 tosses:

- The number of runs obtained
- The length of the longest run of heads
- The length of the longest run of tails.

Gather the results of the above three items from the students into a copy of the two tables at the bottom of the next page. Have the class compute the averages for the table rows.

For students who are first encountering a study of randomness, the following things are commonly observed regarding the computed averages:

- The average number of runs tends to be greater for the students guessing the H/T results without actually tossing the coin.
- The lengths of the longest runs tend to be shorter for the students guessing the H/T results without actually tossing the coin.

Tossing a Coin 40 Times

Toss Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
H/T Result																																								

Summary of Student Results for Students Tossing the Coin

Student #:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Average
Number of runs																					
Length of the longest run of heads																					
Length of the longest run of tails																					

Summary of Student Results for Students Who Guessed the Sequence WITHOUT TOSSING COIN

Student #:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Average
Number of runs																					
Length of the longest run of heads																					
Length of the longest run of tails																					

In conclusion, it seems to be human nature to underestimate the lengths of the longest runs of consecutive heads or consecutive tails when tossing a coin a fairly large number of times,

If you were to have each student lightly write on the back of their data sheet if they were in the group that actually tossed the coin or guessed without tossing and then shuffle the sheets well, you might find it possible to separate the data sheets into the two groups again with reasonable accuracy (without looking at the back of the data sheets, of course)! Observation of run lengths in a "random" process is a good test of whether or not the process is truly random. Many such mathematical tests of randomness have been developed for computer algorithms, which produce pseudo-random numbers.

Computation of Theoretical Probabilities of Runs

*Computation of theoretical probabilities of runs in a long sequence of events is quite complex and best done by computer programs[+]. The following table summarizes the theoretical probabilities for obtaining **at least one run of length L of a specific event in 40 tosses of a coin**:*

Length of Run, L	Theoretical Probability
1	1.00000
2	0.99976
3	0.96012
4	0.75041
5	0.46792
6	0.25570
7	0.13104
8	0.06518
9	0.03196
10	0.01557

We see from the above table that the probability is nearly ½ that we will obtain a run of at least 5 heads in 40 tosses of a fair coin. In the long run, we would expect about half of the students who actually tossed the coin to have a run of at least 5 heads. In the long run, about a quarter of the students would be expected to have a run of at least 6 heads.

+ http://www.gregegan.net/QUARANTINE/Runs/Runs.html

Exercise: ***What is the theoretical probability for obtaining a run of at least 2 heads in 3 tosses of a fair coin?*** *[Hint: List every possible permutation of equally likely outcomes of heads and tails in tossing a coin three times. Count the number of permutations having at least 2 heads in a run, and divide by the total number of equally likely outcomes. See page 12 for the solution.]*

You should be able to conjecture why using the technique of the above exercise would be impractical for 40 tosses of a fair coin. *[You would have 2^{40} = 1,099,511,627,776 permutations—more than a trillion permutations. Listing an evaluating all of these permutations would be impractical.]*

The OzoMap for this Lesson

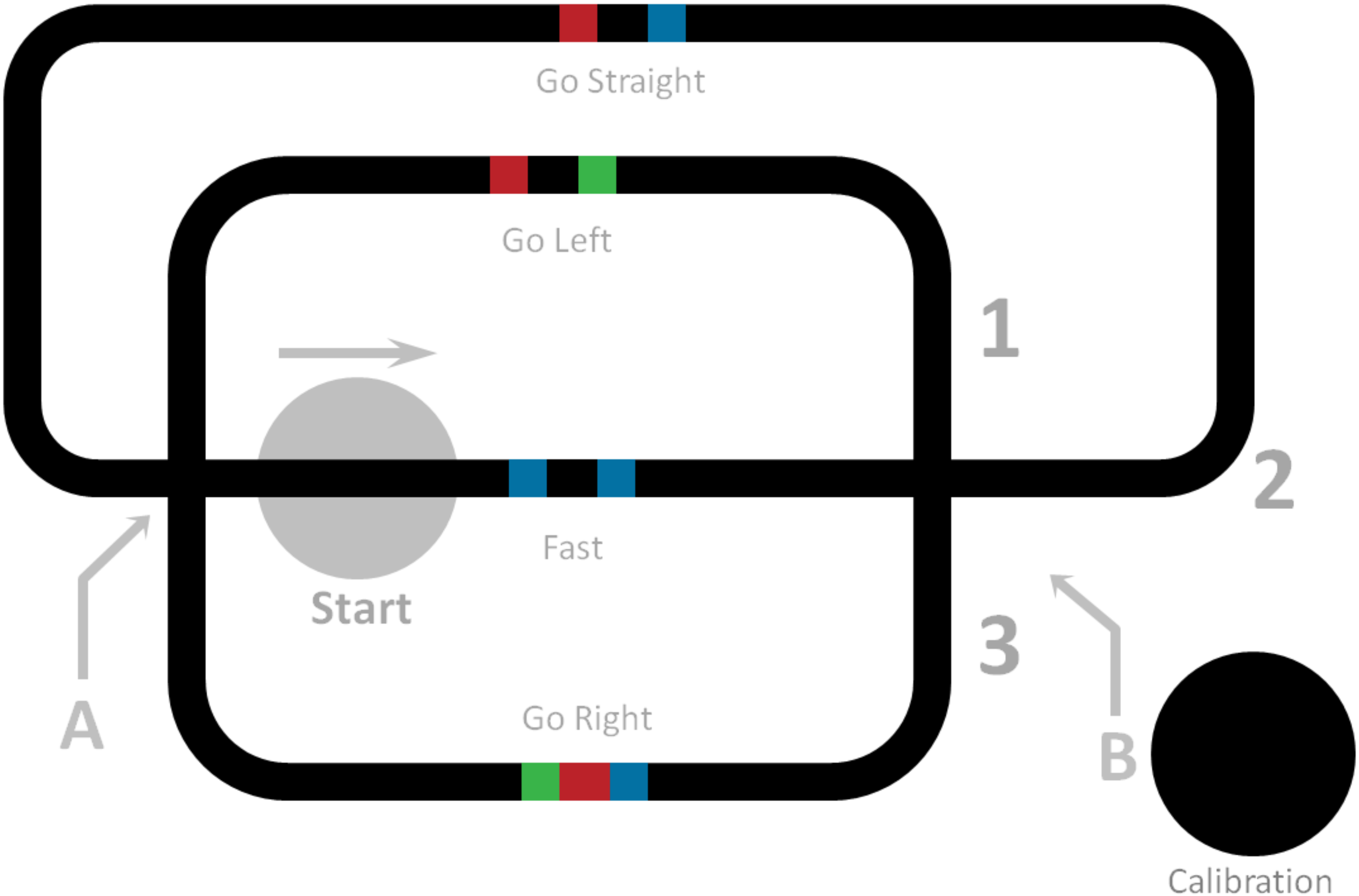
Now that we have a good understanding of the concepts of randomness, probability, and runs from our study of tossing a coin, we are ready to see how we can use Ozobot to perform an experiment dealing with these concepts.

A fundamental feature of Ozobot's design is that it has been programmed to make independent routing decisions at intersections based upon random logic algorithms. What we would like to do is create an OzoMap that will force OzoBot to randomly choose one of three directions upon reaching an intersection, giving a probability of 1/3 ≈ 0.333 that it chooses any one of these three directions. We want Ozobot to quickly return to this intersection no matter which one of the three directions it chooses. Meanwhile we will keep track in a data table of the direction that Ozobot chooses each time in 100 encounters with the intersection.

We will be interested in:

- ***determining the experimental probabilities and comparing these to the theoretical probabilities of 1/3 for each of the three directions.***
- ***determining the length of the longest run for each of the directions. For example, what is the maximum number of times in a row that Ozobot selected direction 2 upon reaching the intersection?***
- ***comparing our results with theoretical probabilities for runs of various lengths.***

The OzoMap page 6 can be duplicated for use by students with their Ozobots. Printing the OzoMap on a bright white card stock type paper is encouraged as it has proven to provide extremely accurate reading of the color codes by Ozobot. Calibration is encouraged before beginning.


Go Straight
Go Left
1
2
Fast
Start
3
A
B
Go Right
Calibration

Student/Teacher OzoMap Discussion Questions

The answer to each of the questions for teacher reference is shown in italic in square brackets [ ] following each question.

1. **Starting Ozobot at the Start location, a Fast code is encountered immediately. What is the effect of this code, and why is it on the OzoMap?** ***[This changes Ozobot from cruising speed to fast speed. It will keep Ozobot moving faster, providing for quicker data collection in the student experiment.]***
2. **What will Ozobot do when reaching intersection B?** ***[Ozobot will make independent and random decisions with uniform probability to take directions 1, 2, or 3.]***
3. **What is the probability that Ozobot will go direction 1? direction 2? direction 3?** ***[Each of these probabilities is 1/3 ≈ 0.333.]***
4. **In the event that Ozobot chooses direction 1 at intersection B, what is the purpose of the Go Left code that it will then encounter?** ***[When Ozobot reaches intersection A, it will turn left and go through Start again.]***
5. **In the event that Ozobot chooses direction 2 at intersection B, what is the purpose of the Go Straight code that it will then encounter?** ***[When Ozobot reaches intersection A, it will go straight through Start again.]***
6. **In the event that Ozobot chooses direction 3 at intersection B, what is the purpose of the Go Right code that it will then encounter?** ***[When Ozobot reaches intersection A, it will turn right and go through Start again.]***

In effect, we have created an OzoMap that will allow us to quickly observe and record Ozobot's random behavior with multiple encounters of an intersection with three equally likely directions that Ozobot can take.

Classroom Randomness and Runs Experiment

In this experiment, student groups will collect data on which direction Ozobot chooses at intersection B for 100 encounters with the intersection. Copies of the data tables on the next page can be made for each student group to record their results. The table at the top is for raw data and will be filled in with 1's, 2's, and 3's to indicate the direction that Ozobot took for each lap. The tables at the bottom are summary data for the student group.

Raw Data

Lap #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Result (1, 2, or 3)																				

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60

61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80

81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Summary Data

Direction	Count	Experimental Probability	Theoretical Probability
1			
2			
3			

Longest Run Data

Run of	Length of Longest Run	Theoretical Probability for at least one run of that length
1's		
2's		
3's		

A typical set of data tables after being filled in by a student is shown on the next page. As shown in the Raw Data table, on Ozobot's first lap, direction 3 was selected. On the second lap, direction 1 was selected, etc. The Summary Data table shows that Ozobot selected direction 1 a total of 34 times, direction 2 a total of 30 times, and direction 3 a total of 36 times. Since there are 100 laps, the experimental probabilities are simply found by dividing the counts by 100, giving 0.34, 0.30, and 0.36, respectively. When filling in the Longest Run Data, students should be cautioned to remember that a run can extend from one portion of the table to the next. For example, there is a run of 2's of length 4 starting with lap number 80.

Students can use the following table to determine the theoretical probabilities for at least one run of 1's (or of 2's or of 3's) of length L in 100 Ozobot laps of this OzoMap.

Length of Run, L	Theoretical Probability
1	1.00000
2	0.99991
3	0.93075
4	0.56536
5	0.23564
6	0.08409
7	0.02846
8	0.00946
9	0.00313
10	0.00103

Note that the probability of obtaining a run of, say, at least four 1's is about ½. In the long run, we would expect about half of the student groups to have a run of at least 4 1's. Similarly, the probability of obtaining a run of, say, at least five 2's is about ¼. In the long run, about a quarter of the student groups would be expected to have a run of at least five 2's.

Practical Considerations of Randomness and Runs

Scientific studies which involve test subjects are commonly assigned randomly to two different treatments X and Y with equal probabilities. If the study involved 1000 subjects, the probability of at least one run of 10 subjects in sequence being assigned to treatment X is about 0.39, or better than 1/3.

Raw Data

Lap #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Result (1, 2, or 3)	3	1	3	1	1	3	3	3	3	3	2	3	1	1	2	2	2	1	3	2

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	2	3	2	3	2	2	2	3	2	1	1	3	2	2	1	1	3	2	3

41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
2	2	2	3	2	3	2	2	1	3	3	3	1	1	1	3	3	3	1	1

61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
1	3	3	1	1	3	1	1	3	1	3	3	2	2	3	1	2	1	3	2

81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
2	2	2	1	1	1	2	3	1	3	1	3	1	1	3	2	3	1	1	2

Summary Data

Direction	Count	Experimental Probability	Theoretical Probability
1	34	0.34	0.33
2	30	0.30	0.33
3	36	0.36	0.33

Longest Run Data

Run of	Length of Longest Run	Theoretical Probability for at least one run of that length
1's	3	0.93075
2's	4	0.56536
3's	5	0.23564

Another practical example involves what is known as the gambler's fallacy, sometimes called the Monte Carlo fallacy. This fallacy involves the mistaken idea that if something happens more frequently than normally expected, such as in a long run, it will happen less frequently in the future, as a way of balancing natural phenomena. For events that are random and independent, this idea is a fallacy. This fallacy is most common among gamblers, thus the name gambler's fallacy.

As an example of the gambler's fallacy, suppose that you have tossed a coin 4 times in a row, and all four tosses resulted in heads. You might think that the probability that your next toss will be heads is less than ½ since you have already experienced a run of 4 heads. But given that you have tossed four heads, the probability that your next toss is heads is in fact still ½. The result of the fifth toss is independent of what happened during the first four tosses! To look at it another way, the probability of tossing 5 heads in a row is only 1/32. But the probability of tossing five heads in a row, given that you have already tossed four heads in a row, is ½.

STEM topics

- *Interdisciplinary—robotics, probability and statistics working together*
- *Computer Science—colored visual codes are used to program a line-following robot*

Intended Grade Levels

Junior and senior high school

Materials

Ozobots (1 per every 3 to 4 students, fully charged)

One 8½" x 11" sheet of bright white cardstock for each student group with the OzoMap on page 6 printed on the cardstock

A penny for each student in the class

One copy of the blank data page 3 for each student in the class

One copy of the blank data sheet on page 8 for each student group

Estimated time-frame

Approximately one hour

Solution to Exercise, Top of Page 5

There are 8 permutations of equally likely outcomes: HHH, THH, HTH, HHT, HTT, THT, TTH, TTT. 3 of these permutations have a least 2 heads in a run: HHH, THH, and HHT. Therefore, the probability is 3/8 = 0.375.