# Garden Paths
# *Teaching Guide*



## Video Resource
Access videos of *Garden Paths* lessons at https://gdaymath.com/courses/gmp/.

## Student Handouts
All practice problems, and solutions, in an accompanying document.

# GARDEN PATHS: Overview

## Student Objectives

Students will develop a solid intellectual and intuitive understanding of the essentials of probability theory.

## The Experience in a Nutshell

We begin the experience by analyzing “garden path” puzzles. These set the scene for understanding and solving probability problems, as well as provide some geometric surprises.

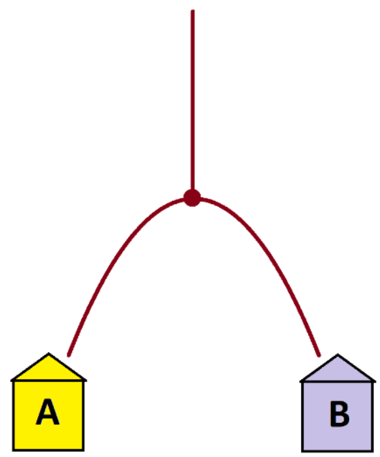
Matters then turn to a more formal discussion of probability theory, its historic roots and the role the *Law of Large Numbers* plays in developing the theory. A classic paradox—the two girls puzzle—is resolved and the flippant aphorism “*and* means *multiply*” is placed in proper, clear context.

The final section of the experience pushes matters to the infinite and explores some tricky infinite probability challenges in the easy light of the garden path model.

# PART I

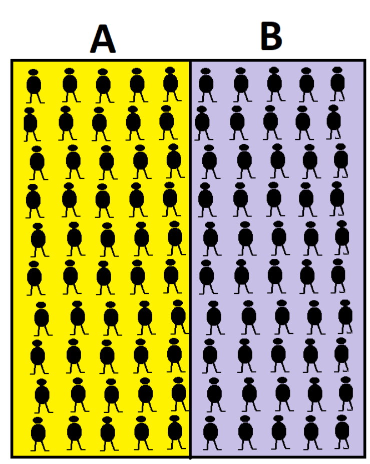
## Lesson 1: Garden Strolls

Suppose 100 people walk down a garden path that leads to a fork. Those who turn to the left go to house A, those who turn to the right to house B. Let's assume that there is a 50% chance that a person will turn one way over the other.




In this setup we'd expect roughly half the people, about 50, to end up at house A and about 50 to end up at house B. (It will likely not be exactly these numbers—the world isn't perfectly predictable. Only in an ideal world we would expect precisely 50 people in each house.)

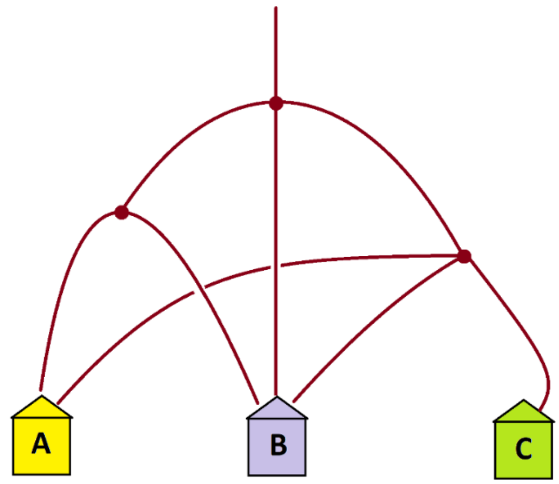
The following diagram of 100 people depicts this outcome.




The number 100 here is immaterial. The point is that if a square is used to denote the entire population of people walking down the path, then half the area of the square, designated A, represents half the count of people who ideally end up in house A and the other half of the area, designated B, represents half the count of people in house B.

This is the set-up of the game we shall play. Let's now have some fun with garden paths!

**Example:** *People walk down the following system of paths.*




*a) Which house do you think will end up with the most people in it? Which one the least?*

*Can you guess what fraction of people will end up in each house?*

*b) Use the square model to actually compute the fractions of people that end up at each house. (Assume that people arriving at each fork split equally in count in each direction.)*

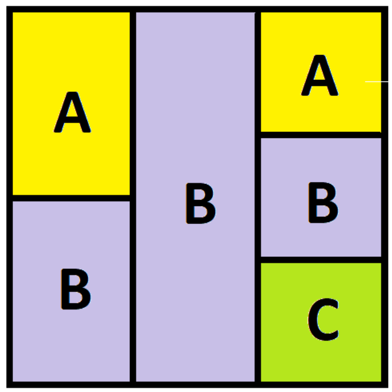
**Answer:** At the first fork a third of the people go to the left, a third go straight, a third go to the right.

Of those who go to the left, half go to house A and half go to house B.

Of those who go straight, all go to house B.

Of those who go to the right, a third go to house A, a third go to house B, a third go to house C.

Do you see this information encapsulated in the following picture?




We see that the proportion of people who end up in house A is given as half of a third, that's $\frac{1}{6}$, plus a third of a third, that's $\frac{1}{9}$, giving the total proportion as $\frac{1}{6}+\frac{1}{9}=\frac{5}{18}$.

The proportion of people who end up in house B is $\frac{1}{6}+\frac{1}{3}+\frac{1}{9}=\frac{11}{18}$.

And the proportion of people who end up in house C has to be the remaining $\frac{2}{18}$ (which is indeed the $\frac{1}{9}$ we see in the picture).

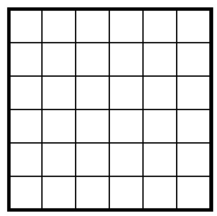
Do these answers match your predictions from part a) of the question?

**Question:** Does it make sense to you that the three fractions we get add to $1$? Also, do you see that it is helpful *not* to simplify the fraction $\frac{2}{18}$ in this example?

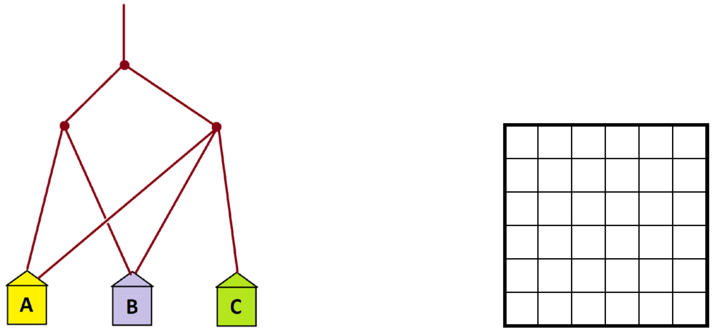
***PEDAGOGICAL COMMENT:*** Many students have been trained to think that non-simplified fractions, such as $\frac{2}{18}$ in this example, are inherently "wrong" and should not be left unmodified. A conversation that mathematics is not absolute and should be guided by context is appropriate here. Keeping all fractions in terms of eighteenths in this example allows us to readily compare proportions. For instance, we see that the counts of people in all three houses follow a 5:11:2 ratio, with house B having more than double the count of people in house A, and so on.

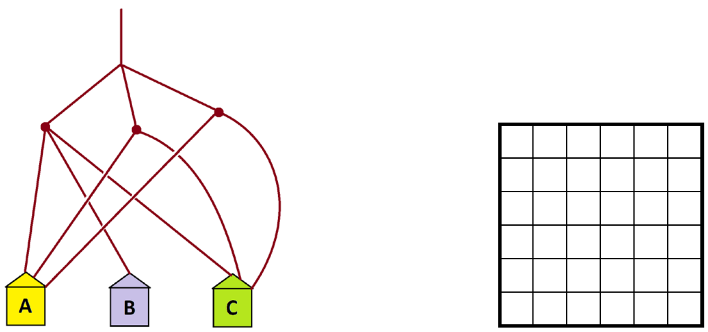
***PEDAGOGICAL COMMENT:*** It can be difficult for younger students to draw free-form squares and rectangular areas within them. One can provide six-by-six grids to help. Some students might enjoy practice question 0 coming up next. (Warning: A count of 36 sub-squares is not good for all garden path systems as students will discover in practice question 3 of this lesson.)

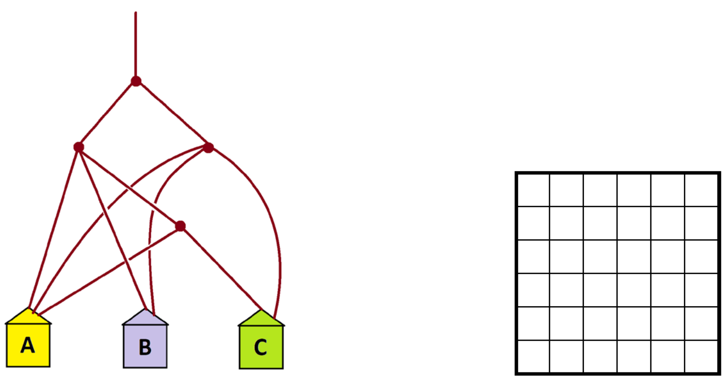


**Practice 0:** *In each of the following examples determine the fraction of people that end up in each house.*



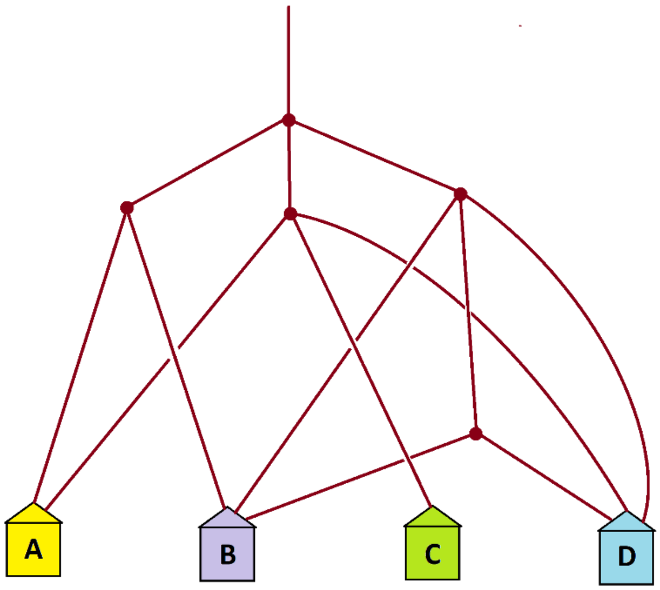



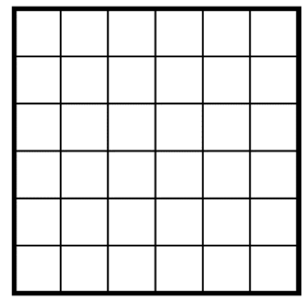



**Practice 1:** *A set of people walk down the following system of paths.*

*a) Any prediction as to which house will end up with the most people in it? The house that will end up with the least?*

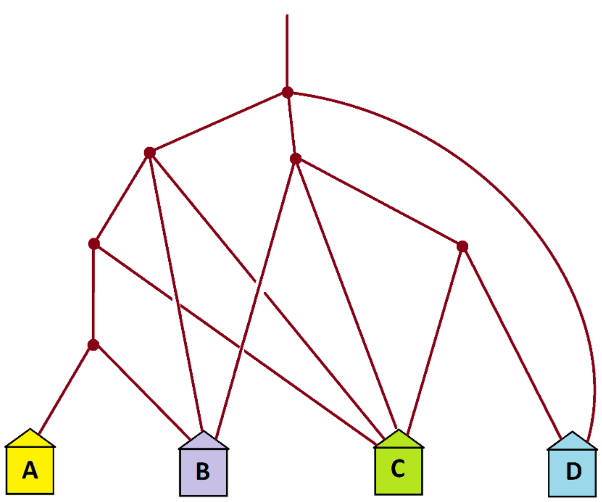


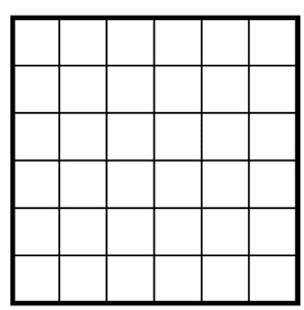



*b) Show that the proportions of people who end up in each house are* $\frac{5}{18}$, $\frac{6}{18}$, $\frac{2}{18}$, *and* $\frac{5}{18}$, *assuming that people arriving at a fork split themselves equally in count in each direction. (Surprisingly, equal proportions of people end up in houses A and D.)*

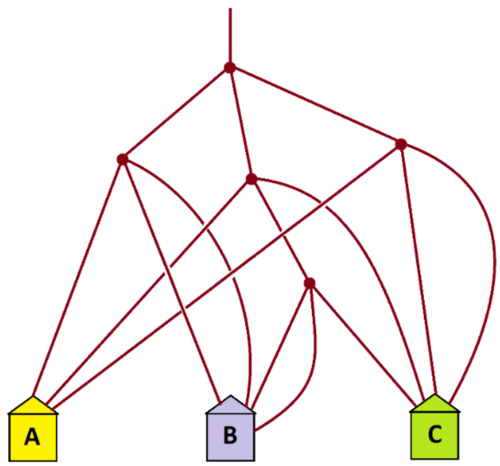
*We've provided here a six-by-six square to help with this question.*

**Practice 2:** *Answer the same questions for this garden path system.*



**Practice 3:** *And answer the same questions for this garden path system. Here some forks have two paths that go to the same house.*




**Practice 4:** *Design a garden-path system such that the house with least number of paths leading to it actually ends up with the most number of people in it.*

**Keep Going!**

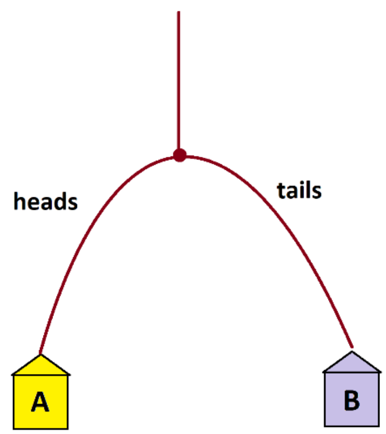
Have some fun making up your own garden path systems for others to solve. (Make sure you can solve them yourself!)

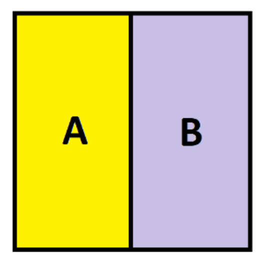
In your examples it might be best to use a blank square rather than one already divided into subsquares.

## Lesson 2: Flipping Coins, Rolling Dice, and Such

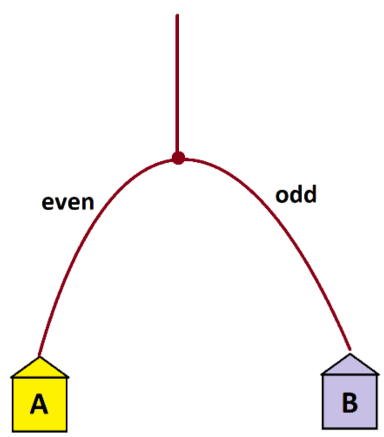
Garden paths are useful for analyzing the results of random actions. We'll talk more deeply about *probability theory* in part II of these notes, but it is fun to see a hint of this application right now.

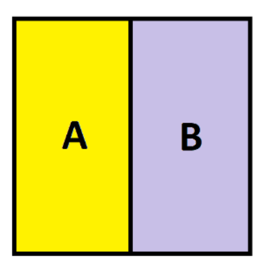
For the garden-path system we opened with—a single two-way fork—we could have people simply flip a coin to decide whether to turn to the left or to turn to the right when they reach that fork. About half the people will turn left and about half will turn right. (And, in an ideal world it would be exactly 50% turning each way.)



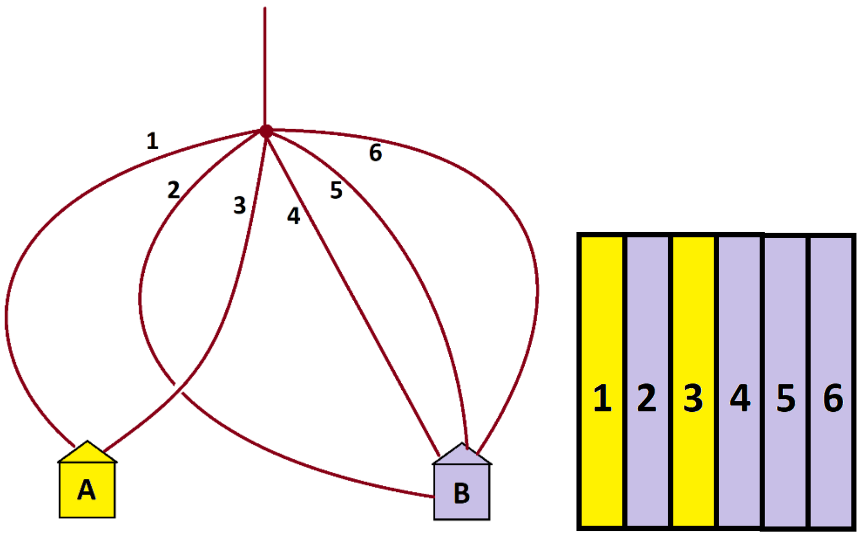

Or we could have people roll a die at the fork: If they roll an even number they turn to the left, to the right if they roll odd.




Or we could use the die to suggest a lopsided garden-path system based on rolls: rolling a 1 or 3 sends people to the left, say, with any other roll sending people to the right.




In this example we see that we expect about two-sixths, that is, one-third of the people will end up in house A and about two-thirds in house B.

Let’s make things more interesting and introduce some more forks.

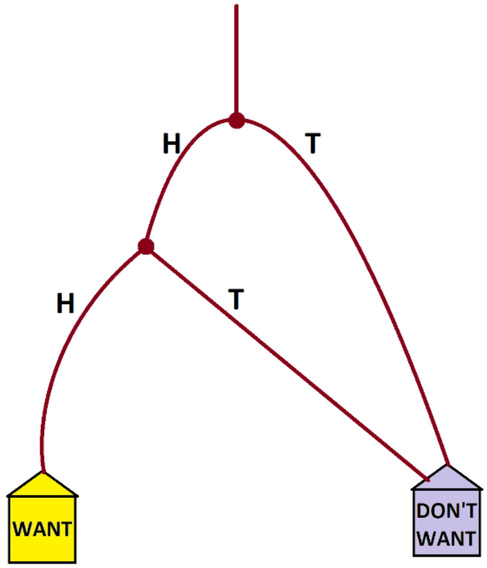
**Example:** *I toss a copper coin and then I toss a nickel coin. What are my chances of seeing a HEAD followed by a HEAD?*

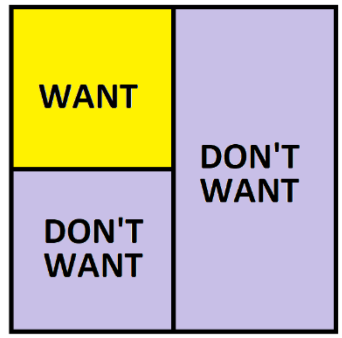
**Answer:** We can create a garden-path system that models this experiment of tossing two coins.

Let’s first have people arrive at a two-way fork and there they shall each toss a copper coin. If they get a HEAD, they’re good to keep going. But if they toss a TAIL, we’ll send them to the DON’T WANT house. (These folk are welcome to toss nickel coins too. But they are going to the DON’T WANT house no matter what!)

Let’s send those who keep going to another two-way fork. There they shall toss a nickel coin. Those who toss HEADS will go to the WANT house. Those who toss TAILS are off to the DON’T WANT house.



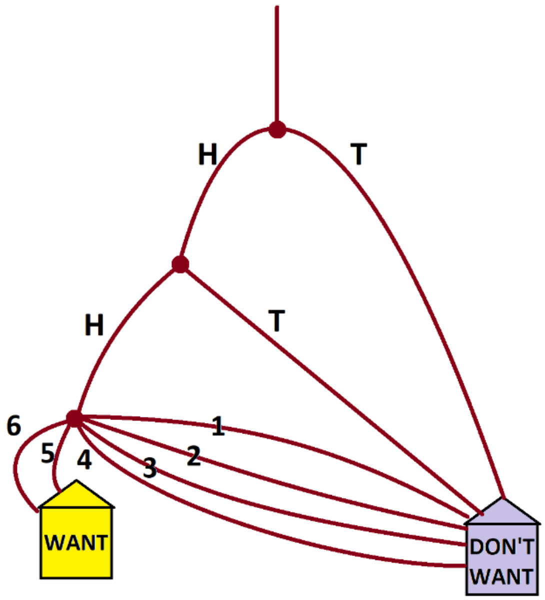



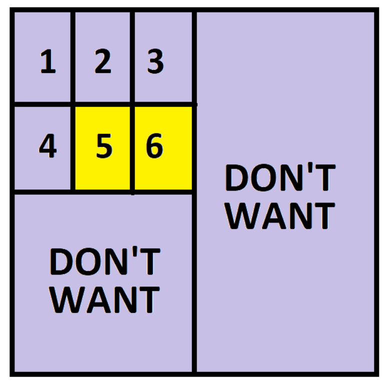

We see that we expect 25% of the people to get HEADS followed by HEADS. That is, this thought experiment shows that the chances of tossing two heads in a row is 25%.

**Example:** *I toss a copper coin, then I toss a nickel coin, and then I roll a die. What are the chances I shall see two HEADS and a 5 or a 6?*

**Answer:** Can we create a garden-path system that models this three-step process? Sure!

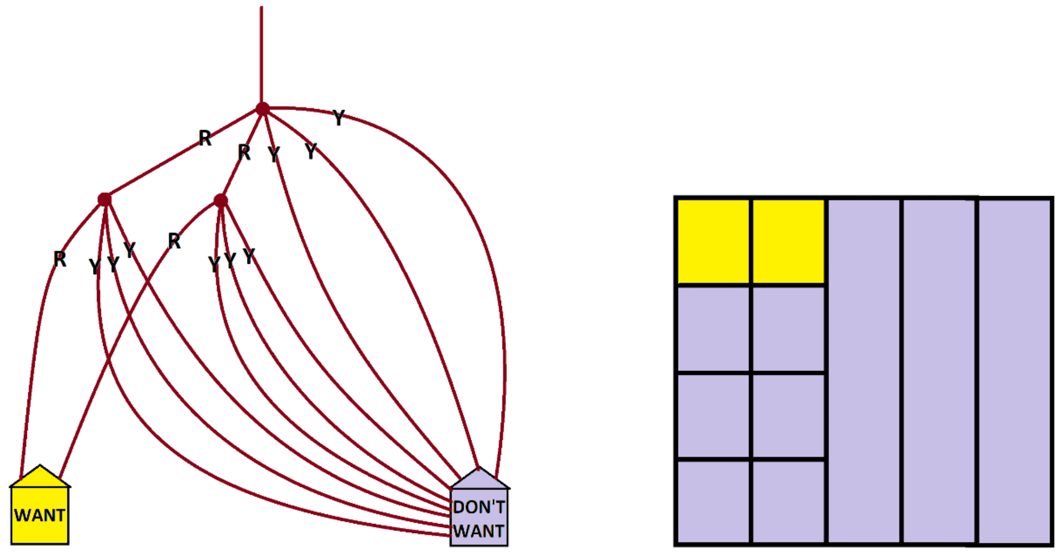



We see that two-sixths of one-quarter of the people who run through the system end up with the result we want. That is, the fraction $\frac{2}{6}\times\frac{1}{4}=\frac{1}{12}$ of the people get two HEADS followed by a 5 or a 6. There is a one-in-twelve chance of seeing this outcome.

**Example:** *A bag contains two red balls and three yellow balls. I pull out a ball at random, note its color, and put it aside. I then pull out a second ball at random from the four balls that remain in the bag and note its color too. What are the chances I see two red balls?*

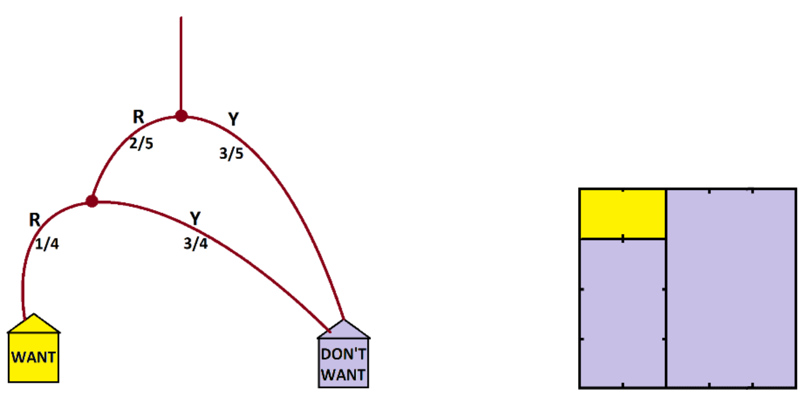
**Answer:** Here's a garden path system, with one path at each fork for each ball in the bag that could be chosen.




The chances of seeing two red balls is one-quarter of two-fifths (do you see that?). This is

$$\frac{1}{4} \times \frac{2}{5} = \frac{1}{10} = 10\%.$$

**Side Comment:** It gets tedious drawing lots of paths for each and every individual option. Do you see that we could simplify the garden-path diagram by giving "weights" to each path? Turning to the left or to the right at each fork need not be equally likely: they could be some non-equal proportion. We can divide the square in these same proportions too.



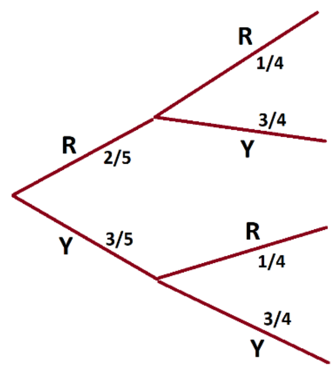

Here we see "one-quarter or two-fifths" more clearly.

***PEDAGOGICAL COMMENT***

Older students will typically make the suggestion to simplify diagrams by adding weights to the edges. Younger students might too if you make a "bit of a show" of how tedious it is to draw all six possible rolls of a die each-and-every time, for instance.

**Side Comment:** Some people like to draw garden-paths sideways and call them *tree diagrams*. They usually draw all possible forks and all possible options, making no judgment as to what is wanted and not wanted.




**Practice 1:** *I roll a ruby die and then I roll an emerald die.*

*a) What are the chances that I will see an even number followed by a six?*
*b) What are the chances I will see only composite numbers on my rolls?*

**Practice 2:** *A bag contains two red balls and three white balls. I pull out a ball at random, note its color, and put it back in the bag. I then shake the bag with all five balls in it and pull out a ball again at random and note its color. What are the chances I see a red ball each time?*

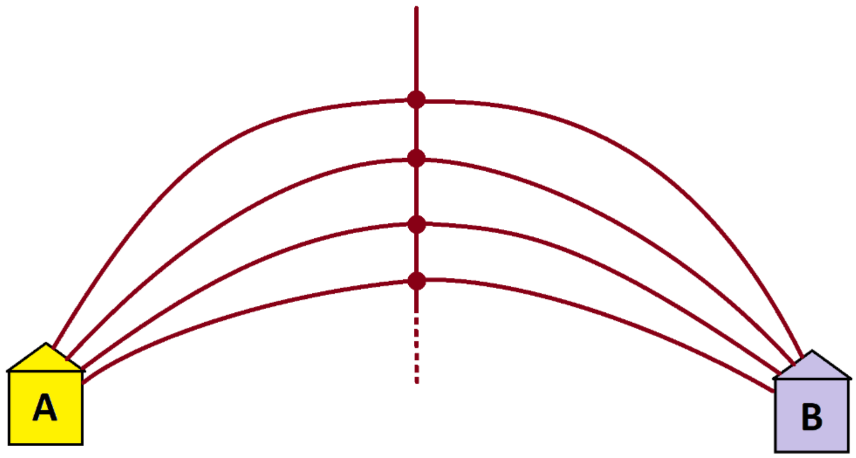
**Practice 3:** *There is a 30% chance that I will sneeze at least once on any given day and a 60% chance I will yawn at least once.*

*a) What are the chances that I will both sneeze and yawn tomorrow?*
*b) What are the chances I will sneeze but not yawn?*
*c) What are the chances I will either sneeze or yawn but not both tomorrow?*

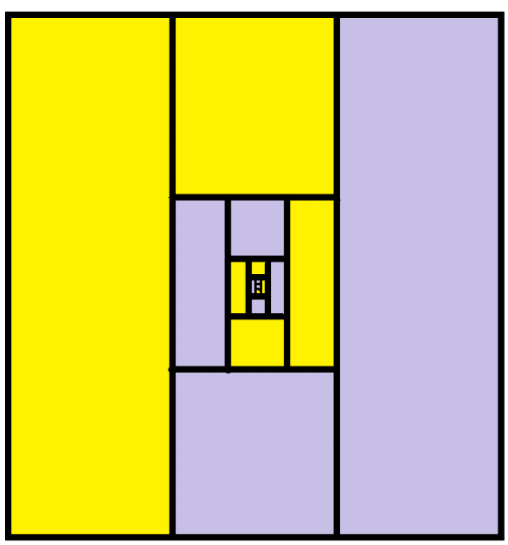
# Lesson 3: Infinite Garden Paths

Let's go back to the fun of garden paths for their own sake. And let's get really wild!

Imagine a garden path with infinitely many forks that each split three ways as shown. Those who turn to the left at a fork go straight to house A, those who turn right to house B, and those who go straight go to another fork. And this is the case over and over and over again!




Do you see that this leads to the following picture?



We have two matching spirals that each spiral all the way to the center of the square. Since the spirals are identical, they each take up half the area of the square! So, half the people end up in house A and half the people end up in house B.

But wait! Think of how people end up in house A.

At the first fork, $\frac{1}{3}$ of the people go to house A, and a third go to house B, and a third keep going.

Of those who keep going, a third of them (that is, a third of the third, $\frac{1}{9}$, of the total population) go to house A, a third of the them go to house B, and a third of them keep going.

Of those who still keep going, a third of them (a third of a third of a third, $\frac{1}{27}$ of the entire population) go to house A, and a third of them go to house B, and a third of them keep going.

And so on.

So we can say that the total number of people who end up in house A must be

$$\frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\frac{1}{81}+\frac{1}{243}+\cdots\ .$$

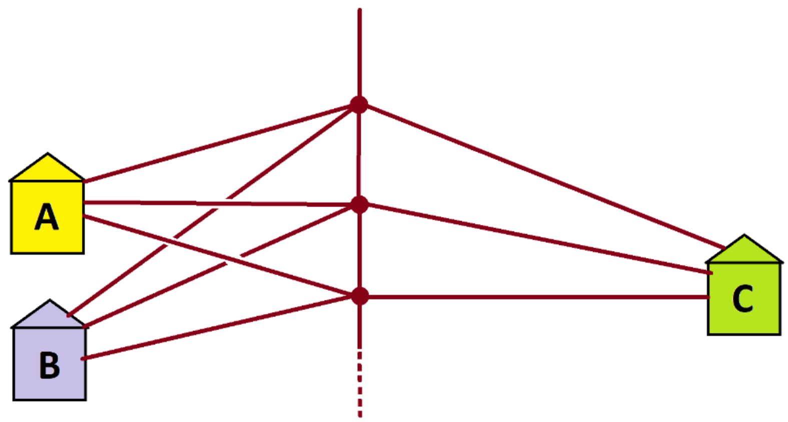
Yet we said that half the people end up in house A! So this infinitely long sum must equal one half!

$$\frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\frac{1}{81}+\frac{1}{243}+\cdots=\frac{1}{2}$$

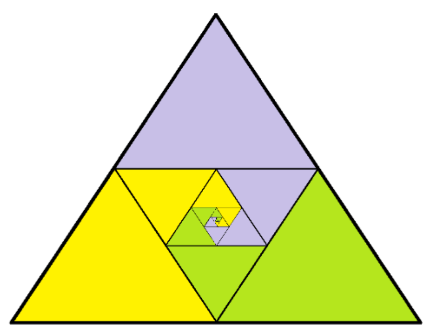
Whoa!

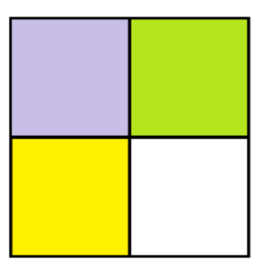
**Practice 1:** *a) Argue that this next infinite garden path suggests that* $\frac{1}{4}+\frac{1}{16}+\frac{1}{64}+\frac{1}{256}+\cdots$ *equals* $\frac{1}{3}$.




*b) Anu was having trouble creating an area model for this system using a square. Then she thought to use an equilateral triangle instead. This is what she drew. What do you think of it?*



*c) Chee Wei said it is actually possible to use the area of a square to represent this garden-path system nicely. He gave a hint as to what he was thinking by sharing this start to picture on the board. How do you think we will continue the picture?*



**Practice 2:** *a) Draw a garden path system that leads to the sum* $\frac{1}{5}+\frac{1}{25}+\frac{1}{125}+\frac{1}{625}+\cdots=\frac{1}{4}$.

*b) What do you think might be value of* $\frac{1}{N}+\frac{1}{N^2}+\frac{1}{N^3}+\frac{1}{N^4}+\cdots$ *for a positive integer* $N$ *?*

In Part III we'll push this example further and open the gates to a whole host of infinite probability processes.

# PART II

## Lesson 4:

## Probability: A Historic Start

The start of probability theory can, in one sense, be pinpointed to a single moment in time. In 1654 French nobleman Chevalier de Méré wrote to prominent mathematician Blaise Pascal asking for advice on general "problems of points" (among other issues in making gambling bets). These problems go as follows.

> *Two friends each lay down $100 in a friendly "best of seven" tennis game, say. But rain interrupts play after just four games with one person having won three games and the other just one. How should the $200 in the pot be divvied between the two players so as to properly reflect the likelihood of each winning?*

Pascal shared the question with colleague Pierre de Fermat. It is with de Méré's inquiry that probability theory was born as a subject to be developed and deeply studied in its own right.

**Comment:** Italian mathematician Girolamo Cardano (1501-1576) worked with ideas akin to classical probability theory before this but did not publish his work. And, of course, gambling games have been in existence for centuries and scholars have wondered about, analyzed, and computed likelihoods of certain results. But the first definitive analysis of "chance" began with the work of Pascal and Fermat.

I invite you to think about and try to analyze de Méré's problem. How would you approach it? What beginning assumptions would you make? What solution would you offer for splitting the pot between the two players and why?

Give this problem some thought yourself before reading the list of possible solutions my brain came up with.

## Some Possible Solutions

Here are four possible ways to handle the problem-of-points conundrum.

Let's start by giving the two players names: A for Albert and B for Bilbert, say, with A having won three games and B one.

**Possible Solution 1:** *Just give each player back his $100 and play another day.*

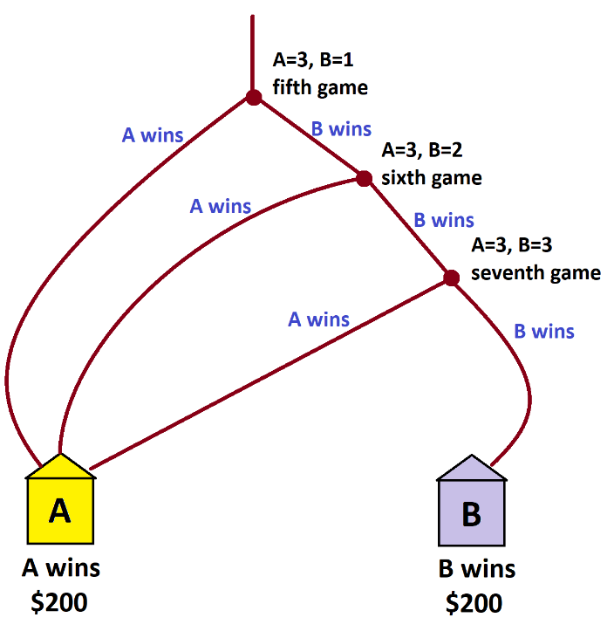
> **Possible Objection:** A will likely complain that since he was on the verge of winning he should receive more than this amount.

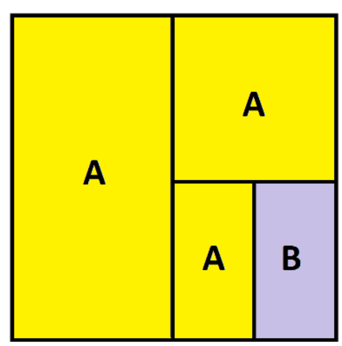
**Possible Solution 2:** *Give player A $150 and player B $50 to reflect the ratio of their wins thus far.*

> **Possible Objection:** A might argue that ratios are unfair. For example, if he were to win the next game, he'd be given the full $200 and player B $0, which is more than what the 4 win to 1 win ratio would suggest ($160 and $40). Ratios undercut what A deserves!

**Possible Solution 3:** *Analyse the problem with Garden Paths.*

Suppose A and B are equally strong players, each just as likely to win any particular game. Here is the garden path analysis under this assumption.



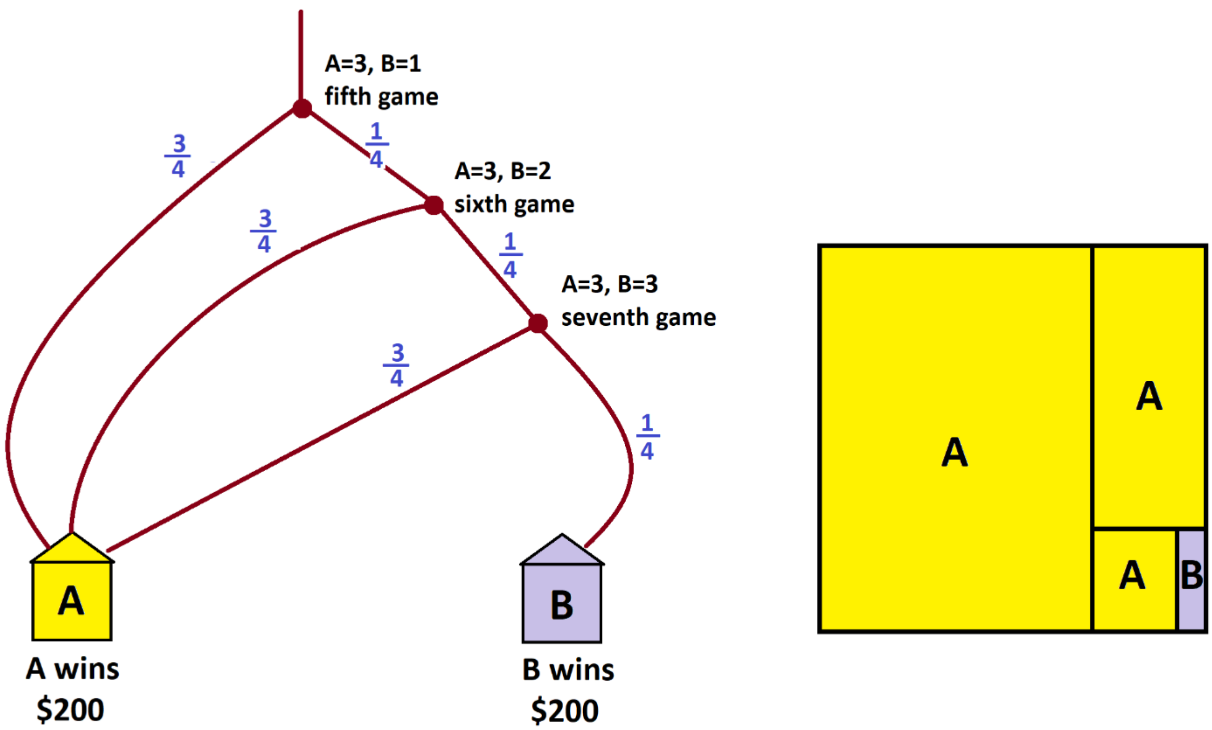

 

We see in this scenario that B has a one-eighth chance of winning the pot. This suggests splitting the pot in a 7:1 ratio, which corresponds to giving A $175 and B $25.

> **Possible Objection:** A would argue that the evidence from the first four games played (three wins for A, one for B) suggests he is the stronger player, beating B not just 50% of the time but actually 75% of the time.

**Possible Solution 3:** *Analyse the problem with weighted Garden Paths, assuming A has a 75% chance of winning any particular game.*

This gives the following diagram.




We see in this scenario that B has only a one in sixty-four chance of winning the pot. This suggests the split $196.875 for A and $3.125 for B. (How should the handle the split one-cent?)

> **Possible Objection:** B claims that he was "laying low" for the first few games to give A a false sense of confidence. He actually has a 95% chance of winning any given game and was going to come in strong for the next three games.

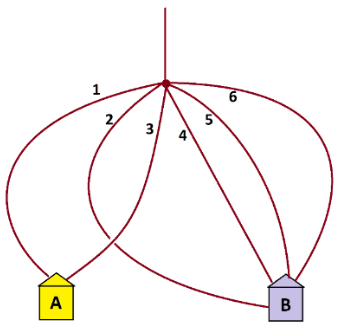
**Challenge:** Find yet another reasonable solution to de Méré's challenge.

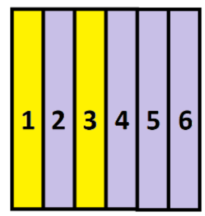
# Lesson 5:

# Probability: What we Choose to Believe

We say that in flipping a coin a large number of times we expect about half the tosses to land HEADS and about half to land TAILS. But what are we really saying here? What intuitive belief is underlying such a statement?

Similarly, when we say that there is a two-sixths chance of seeing a 1 or a 3 when rolling a die, what do we really mean by "chance"?




These questions, albeit vague, are valid questions. We know, for example, that when we actually do toss a coin a number of times it is rare to actually see exactly half the tosses HEADS and half TAILS. So, there is something subtle afoot here.

**Try this:** (Read through all these instructions first to see what is coming!)

*Flip a coin ten times in a row. Did you actually see five heads and five tails?*

*Flip a coin another ten times in a row. Did you see five heads and five tails this time?*

*Do this a third time. Did you see five heads and five tails?*

*Among all 30 flips you just conducted, were you close to seeing 15 heads and 15 tails? As a percentage, what proportion of the 30 tosses were heads?*

*Suppose you and friend each conducted this exercise. Among your 60 flips, were you close to seeing 30 heads and 30 tails? As a percentage, what proportion of the 60 tosses were heads?*

*Suppose you are in a classroom of many students, say 20, and you each flipped a coin 30 times. What percentage of the 600 tosses were heads?*

(Now that you've read these instructions and haven't yet done the activity, think for a moment about what percentages you might expect and the accuracy of your guesses. Now you are ready to do this exercise!)

## Our Basic Belief of what Chance Means

The activity above illustrates our basic belief about random events.

We intuitively believe that for many basic actions, such as flipping a coin, rolling a die, spinning a spinner, and so on, each possible outcome from the action has a certain inherent number associated with it, called its *probability*, and this number manifests itself as follows.

> *If an outcome of an action has a probability p% of occurring, then that means in performing that action many, many times we expect to see that particular outcome about p% of the time. (And this observed proportion gets closer and closer to being exactly p% if we perform the action more and more times.)*

For example, in flipping a coin and looking to see a HEAD, we associate the number 50%. This is because if we flip a coin 90,000 times, say, we feel we'd see about 45,000 of those tosses landing heads. And the percentage of heads seen will likely be closer to exactly 50% if we tossed the coin 900,000 or 9,000,000 times.

If we roll a die one million times, we feel we'll likely see a roll of a 6 close to one-sixth of the time.

We also feel that this intuitive idea works in reverse. For example, if we toss a coin 500 times and it lands heads for a count of 403 of those 500 tosses, then we'll all strongly suspect that the coin is biased (and biased with a probability of about 80% for landing heads).

This intuitive feeling for what probability means is enough to allow us to work our way up to some very challenging problems. We'll take it slowly.

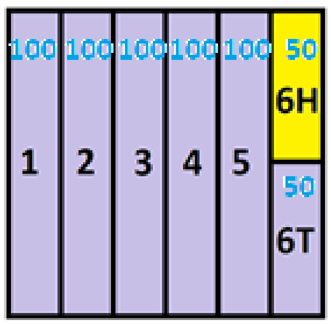
**Notation:** People usually use the letter $P$ to denote a probability value. For example, in flipping a fair coin we might write $P(HEAD) = \frac{1}{2} = 50\%$ and in rolling a fair die, $P(\text{seeing a SIX}) = \frac{1}{6}$.

**Example:** *Imagine rolling a die and then flipping a coin. What are the chances of seeing a 6 followed by a HEAD?*

**Answer:** Imagine running this double-barreled experiment a large number of times, say, 600 times. (That is roll a die and then toss a coin and record the pair of results six hundred times.) Then we would expect to see a roll of a 6 about 100 times. And of those 100 rolls, we expect half to be followed by a toss of HEADS and half followed by a toss of TAILS. So we expect 50 of our 600 runs of the experiment to have the desired result.

Thus we conclude $P(6 \text{ and } HEAD) = \frac{50}{600} = \frac{1}{12}$.

This picture of a rectangle records these expected counts.




**A Complex Example:**

*A bag contains 8 Tuscan sunset orange balls and 2 Tahitian sunrise orange balls. The color difference is very subtle and only 70% of people can correctly identify the color of the ball when handed one of either color. Lulu pulls a ball out of the bag at random and tells you over the phone that she pulled out a Tahitian sunrise ball. What are the chances that the ball she holds in her hand really is Tahitian sunrise?*

**Answer:** Let's assume, as the given statistic suggests, that there is a 70% chance that Lulu correctly identifies the color of a ball when handed one.

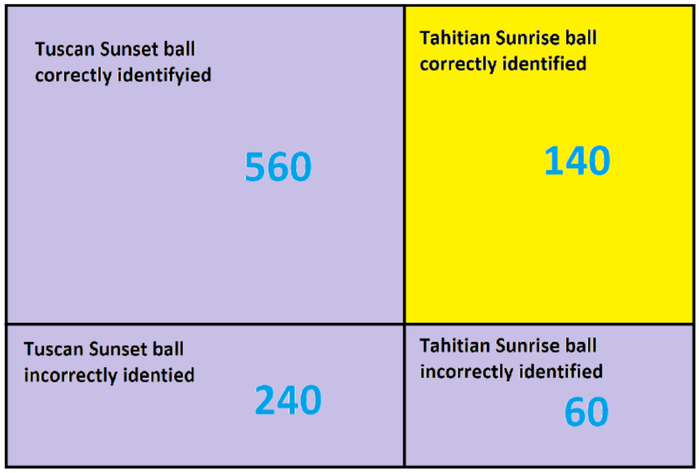
Now imagine Lulu conducting this ball-picking experiment a large number of times, say 1000 times (and that even with this practice her chances of correctly identifying colors does not change).

About 800 of the balls Lulu pulls out will be Tuscan sunset. Of those she will identify about $0.7 \times 800 = 560$ correctly as Tuscan sunset and 240 she will incorrectly say are Tahitian sunrise.

About 200 of the balls Lulu pulls out will be Tahitian sunrise, of which she will correctly identify $0.7 \times 200 = 140$ as such. However, she will incorrectly call 60 of them Tuscan sunset.

This rectangle illustrates these numbers.




Thus in these 1000 runs of the exercise we see that Lulu will say "Tahitian sunrise" about $240+140=380$ times and will be correct in saying this 140 of those times. This shows that the probability that the ball she holds really is Tahitian sunrise is $\frac{140}{380}\approx 37\%$. Pretty low!

**Comment:** This is why eye-witness testimony in court cases is usually taken with skepticism. Even if an eye-witness can be shown to be correct in identifying a certain phenomenon the majority of times, it might still be the case that the eye witness is incorrect the majority of the times!

**Practice 1:** *Suppose that 1% of the population has a certain disease. A test for the disease has been developed. This test will return a correct positive result for 99% of the people who actually have the disease (meaning that there is a 1% chance of a "false negative") and a correct negative result for 95% of the people who do not have the disease (meaning that there is a 5% chance it will produce a "false positive").*

*You have just been tested positive for the disease! Show that there is one a one-in-six chance that you actually have the disease.*

**PEDAGOGICAL COMMENT:** The fundamental belief offered in this lesson is officially called *The Law of Large Numbers*. If explored and used at the very beginning of a formal probability course and coupled with the area model (drawing a rectangle and subdividing into regions area in proportion of times one expects to see each outcome) it makes for a powerful and intuitive entry into probability thinking and calculating. Coupling this with garden-paths throughout the discussions adds to the intuition development—and fun!

# Lesson 6:

# ASIDE: The Infamous Two-Girls Paradox

Here is a famous puzzle that has befuddled the best of mathematical minds. It snagged me too when I first thought about it.

Consider three scenarios.

> *Albert, who you just met, tells you that he is the father of two children and that his oldest child is a girl. What are the chances that his other child is also a girl?*
>
> *Bilbert, who you just met, tells you that he is the father of two children and that one of his children is a girl. What are the chances that his other child is also a girl?*
>
> *Cuthbert, also a new acquaintance, tells you that he is the father of two children and that one of his children is a girl who was born on a Tuesday. What are the chances that his other child is also a girl?*

Believe it or not, the usual answers presented are, in turn, $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{13}{27}$, which are mighty strange, and hence paradoxical! "Surely the answer is 50% in each and every case?" most would cry, figuring that there is a fifty-fifty chance for the unmentioned child to be one gender or the other! "The additional information supplied in the latter two cases is surely extraneous and immaterial?"

The trouble is that one can legitimately argue the standard answers are incorrect! This then supplies a paradox within the paradox, the feature that snags professionals too!

Let me explain the standard answers and why they should be challenged.

We'll assume that giving birth to a child of any particular gender is indeed 50%.

> **Question 1:** *Is this indeed the case? What do world statistics seem to say for the percentage of girls born each year versus the percentage of boys?*

Let's suppose we encounter a large number of fathers, each with two children, say, 1960 of them. (I admit this is a strange number to pull out of the air! You'll see as you read on that I realized that having a number divisible by seven a couple of times – there are seven days of the week—and divisible by two a couple of times—there are two genders—turns out to keep the numbers friendly.)

Among the 1960 fathers we encounter we expect half of them, that is, 980 of them, to have oldest child a girl. Of those 980 fathers, we expect half of them, that is 490 of them, to have youngest child a girl and half of them to have youngest child a boy. Similarly, of the 980 fathers we expect with oldest child a boy, half will have second child a girl and half second child a boy.

| | | YOUNGEST CHILD | |
|---|---|---|---|
| | | GIRL | BOY |
| OLDEST CHILD | GIRL | 490 | 490 |
| | BOY | 490 | 490 |

Albert tells you that he is among the 980 fathers with oldest child a girl (first row). The chance that he is actually among the subset of 490 fathers with youngest child also a girl is $\frac{490}{980}=\frac{1}{2}$.

But a rub comes with Bilbert: we are not told which child he commented on nor the mechanism or motivation which caused him to mention the gender of a particular child. Maybe he just picked one of his two children at random and mentioned the gender? Or perhaps he lives in a society that requires you to mention having a female child if you legitimately can? Just these two different motivations already give two different values for the chance that Bilbert's second child is also a girl.

Let's analyze the probabilities under the effect of each of these motivations.

Motivation 1: *Let's assume each father of two children secretly flips a coin to choose a child whose gender he will divulge.*

Looking at the table above we see that there are 980 fathers with a child of each gender and after flipping a coin we expect half these, $245+245=490$ of them, so say "One of my children is a girl." There are also 490 fathers with two girls who will say "One of my children is a girl" for sure. Bilbert tells us that he is among the fathers who say these words. The chances that he is a father of two girls is thus $\frac{490}{980}=\frac{1}{2}$.

| OLDEST CHILD \ YOUNGEST CHILD | GIRL | BOY |
|---|---|---|
| GIRL | 490 | 245 |
| BOY | 245 | X |

The counts of fathers we expect to say "One of my childen is a girl."

Motivation 2: *Each father lives in a society that requires you to mention having a female daughter if you legitimately can.*

There are now $490\times 3=1470$ fathers who, for sure, will say to you "One of my children is a girl." The chance that Bilbert is among the sub-group of fathers with two girls is $\frac{490}{1470}=\frac{1}{3}$.

| OLDEST CHILD \ YOUNGEST CHILD | GIRL | BOY |
|---|---|---|
| GIRL | 490 | 490 |
| BOY | 490 | X |

The counts of fathers we expect to say "One of my childen is a girl."

***Practice 1:*** *Suppose Bilbert lives in a society that requires you to mention that you have a male child if you legitimately can. Given that Bilbert said "One of my children is a girl," what now are the chances that Bilbert has two girls?*

The situation with Cuthbert is analogous, just a little more detailed.

We need to consider girls born on a Tuesday (let's assume one-seventh of any given count of babies are born on any given day of the week). We also need to consider girls not born on a Tuesday and consider boys. The following table captures the appropriate counts for a group of 1960 fathers.

| OLDEST CHILD \ YOUNGEST CHILD | GIRL born on TUESDAY | Girl not born on TUESDAY | BOY |
|---|---|---|---|
| GIRL born on TUESDAY | 10 | 60 | 70 |
| Girl not born on TUESDAY | 60 | 360 | 420 |
| BOY | 70 | 420 | 490 |

Motivation 1: *Let's assume each father of two children secretly flips a coin to choose a child whose gender he will divulge, along with the day of the week the child was born.*

There are $60+70+60+70=260$ fathers with just one female child born on a Tuesday and, after flipping a coin, we expect half of them, $30+35+30+35=130$ of them, to say the words "One of my children is a girl born on a Tuesday." There are $10$ fathers with two girls each born on a Tuesday who will say these words for sure. So we expect a total of $140$ fathers will make this comment.

YOUNGEST CHILD

| OLDEST CHILD | GIRL born on TUESDAY | Girl not born on TUESDAY | BOY |
|---|---|---|---|
| GIRL born on TUESDAY | 10 | 30 | 35 |
| Girl not born on TUESDAY | 30 | X | X |
| BOY | 35 | X | X |

The counts of fathers who we expect to say "One of my children is a girl born on a Tuesday."

Of those 140 who say these words, we see that $30+30+10=70$ of them have second child a girl. Thus the chances that Cuthbert is among this special set of father is $\frac{70}{140}=\frac{1}{2}$.

Motivation 2: *Each father lives in a society that requires you to mention having a female daughter born on a Tuesday if you legitimately can.*

We see that there are $10+60+70+60+70=270$ fathers who will be sure to mention that they have a girl child born on a Tuesday. Of these, there are $10+60+60=130$ fathers having a second child a girl. The chance that Cuthbert has two daughters is thus $\frac{130}{270}=\frac{13}{27}$.

YOUNGEST CHILD

| OLDEST CHILD | GIRL born on TUESDAY | Girl not born on TUESDAY | BOY |
|---|---|---|---|
| GIRL born on TUESDAY | 10 | 60 | 70 |
| Girl not born on TUESDAY | 60 | X | X |
| BOY | 70 | X | X |

***Practice 2:*** *Suppose Cuthbert lives in a society that requires you to mention that you have a male child if you legitimately can. Given what Cuthbert said what are the chances that Cuthbert has two girls?*

**Practice 3:** *Suppose Cuthbert lives in a society that requires you to mention that you have a female child if you legitimately can. Further, assume people like to add the day of the week their daughters were born and when given a choice between two daughters to mention, they use the flip of a coin to decide. Given what Cuthbert said what are the chances that Cuthbert has two girls?*

# Lesson 7:

## ASIDE: Does *and* mean *multiply* in probability theory?

Consider this example.

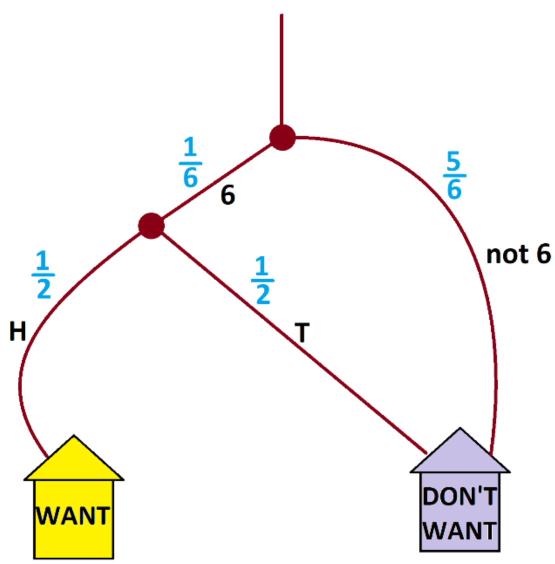
> *Suppose I roll a die and then flip a coin. What are the chances that I see a SIX followed by a HEAD?*

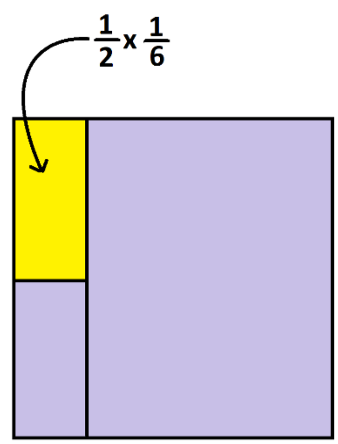
Let's draw a (weighted) garden path system for this problem.

At a first fork people roll a die. Those who roll a 6—and we expect one-sixth of them to do so—continue on and those that don't will head to the DON'T WANT house. (They are welcome to toss coins too, but they are heading to that house no matter what.)

Those who continue reach another fork at which they flip a coin with those getting HEADS heading to the WANT house and those getting TAILS to the DON'T WANT house.

We see that half of one-sixth of the people in this system end up in the WANT house.




We have $P(\text{SIX and then HEAD}) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$.

So, naively, it seems that we have simply multiplied together two basic probabilities: the chances of rolling a SIX with a die, $P(\text{SIX}) = \frac{1}{6}$, and the chances of flipping a HEAD with a coin, $P(\text{HEAD}) = \frac{1}{2}$.

And this appears to be the correct thing to do when we look at the area picture for the garden path system: we identified a desired fraction of a fraction of the whole square, namely, one half of one sixth of the square.

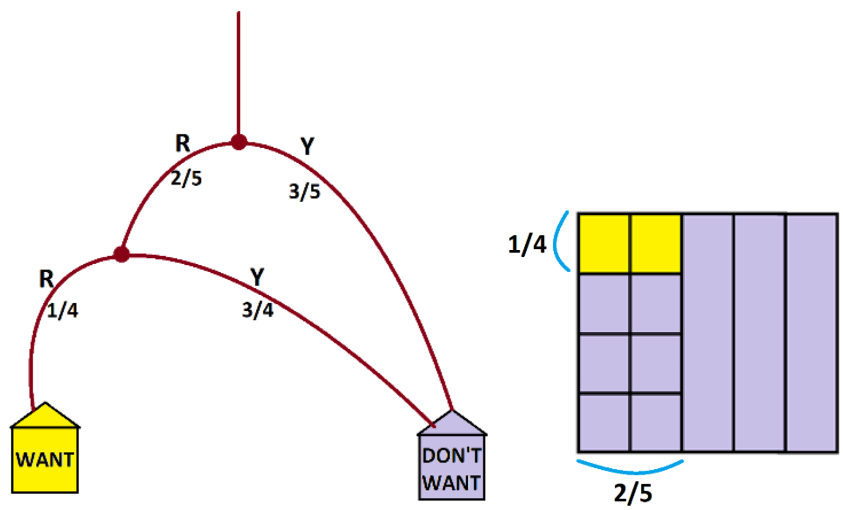
These observations are, of course, valid and correct. But there is a subtlety glossed over in playing with a coin: no matter which population of people flip a coin—those who have just rolled a SIX on a die, those who just rolled a THREE, those who never rolled a die at all—we expect half of that population to get a HEAD. Always half. A coin does not "care" what previous actions you may have just performed nor the results you got from them.

But some actions have outcomes whose chances of occurring do depend on what might or might not have occurred just beforehand.

We have already seen this phenomenon with an example from lesson 2.

**Example:** *A bag contains two red balls and three yellow balls. I pull out a ball at random, note its color, and put it aside. I then pull out a second ball at random from the four balls that remain in the bag and note its color too. What are the chances I see two red balls?*

Here's the weighted garden path system for the example.




At the first fork people pull out a ball from a bag containing two red and three yellow balls. We expect two-fifths of them to pull out a red ball.

$$P(\text{RED}) = \frac{2}{5}$$

Those who pull out a red ball go on to a second fork where they are asked to pull out a ball again. But the probability of pulling out a red ball has changed because of the action that just occurred: there are now only four balls in the bag of which only one is red.

$$P(\text{RED, given that you previoulsy pulled a red ball}) = \frac{1}{4}.$$

The proportion of people who pull out two red balls is one-quarter of two-fifths of the folk

$$P(\text{RED, and then RED}) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10},$$

which is the product of two individual probability values, but with the second one is a value calculated in the context of having just seen a previous outcome.

**Notation and Jargon:** Suppose one is in the midst of performing two actions and that we have just seen outcome $A$ from the first action. Suppose we are hoping to see outcome $B$ from the second. People write $P(B \mid A)$ for probability seeing outcome $B$ in the second action under the assumption we have just seen outcome $A$. This is read as the "probability of $B$ *given* $A$" and is called a *conditional probability*: it is the probability of seeing outcome $B$ under the circumstance (condition) you have just witnessed outcome $A$.

For instance, in our example of pulling out balls from a bag one after another, we have

$$P(\text{second ball RED} \mid \text{first ball RED}) = \frac{1}{4} \quad .$$

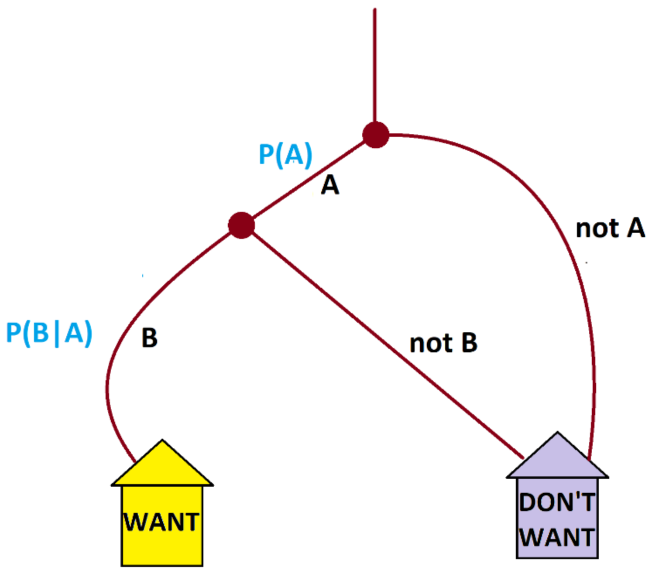
And for practice, can you see that

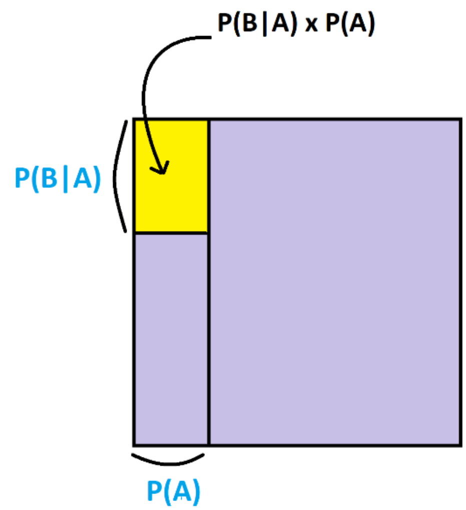
$$P(\text{second ball YELLOW} \mid \text{first ball RED}) = \frac{3}{4},$$

$$P(\text{second ball YELLOW} \mid \text{first ball YELLOW}) = \frac{1}{2} \text{ , and}$$

$$P(\text{third ball RED} \mid \text{first and second balls YELLOW}) = \frac{2}{3}\,?$$

Working abstractly, consider a double-barreled experiment where we are hoping to see outcome $A$ from the first experiment, followed by outcome $B$ from the second. In a garden-path system we expect fraction $P(A)$ of the people to first see outcome $A$. Of those people who get this outcome, we then expect fraction $P(B \mid A)$ of them to see outcome $B$ from the second experiment. Thus the fraction of people we expect to see is the fraction of a fraction $P(B \mid A) \times P(A)$ of the people.




So here then is what people actually mean when they say that "*and* means *multiply*" in probability theory.

**THE MULTIPLICATION PRINCIPLE IN PROBABILITY THEORY**

*Suppose one performs one task in hopes of getting outcome* $A$*, and then performs a second task in hopes of getting outcome* $B$*. To work out the probability of seeing* $A$ *and then* $B$*, that is,* $P(A \text{ followed by } B)$*, work out:*

$P(A) =$ *the probability of seeing outcome* $A$ *in running the first task alone,*

$P(B \mid A) =$ *the probability of seeing outcome* $B$ *in the second task under the assumption you have indeed just seen outcome* $A$ *in the first.*

*Then*

$$P(A \text{ followed by } B) = P(A) \times P(B \mid A).$$

This product is the product of probabilities you see along the path to the WANT house if you work with a weighted garden-path diagram.

This principle is confusing when using examples of rolling die and flipping coins: the subtlety of conditional probability is hidden. As we mentioned, coins, for instance, do not “care” whether or not you have previously rolled a die: $P(\text{HEAD} \mid 6)$ is just the same as $P(\text{HEAD}) = \frac{1}{2}$ and one might not notice a conditional probability at play.

**Jargon:** Two actions or experiments are said to be *independent* if, in running the two actions one after the other, in either order, the outcome of the first action in no way influences or affects the outcomes of the second.

For example, rolling a die and flipping a coin are independent actions.

Me choosing a pair of trousers to wear from my wardrobe at random and me choosing a shirt to wear are not independent exercises: if I happen to pull out my purple trousers first, then I will make sure I avoid choosing a purple shirt. Thus the outcome of choosing trousers might well influence my choice of shirt.

If for two experiments we are hoping to see outcome $A$ from the first experiment and then outcome $B$ from the second, then $P(B \mid A)$ is the probability of seeing this happen. But if the second experiment is independent of the first, then the chances of seeing outcome $B$ under the condition you have just seen outcome $A$ should be no different than just seeing outcome $B$, in general. For independent events, we have $P(B \mid A) = P(B)$.

For example, in rolling a die and flipping a coin $P(HEAD \mid 6) = P(HEAD) = \frac{1}{2}$.

One has to use one’s own judgment and common sense to decide whether or not two (or more) experiments are independent.

**Example:** *On any given day there is a 30% chance that I will sneeze at least once and a 60% chance I will yawn at least once. What are the chances that, next week, I will sneeze at least once on Monday, sneeze at least once on Wednesday, and yawn at least once on Friday?*

**Answer:**

Let $P(\text{sneeze})$ and $P(\text{yawn})$ denote the probabilities of sneezing and yawning at least once, respectively, on any particular day. We have

$$P(\text{sneeze}) = 0.3$$
$$P(\text{yawn}) = 0.6.$$

The information in the question seems to imply that sneezing on any particular day is no way dependent on whether or not I sneezed on a previous day nor on whether or not I yawned. Similarly, the chances of me yawning at least once on any particular day seem to be independent of any other factors.

So

$$P(\text{Sneeze on Wednesday} \mid \text{Sneezed on Monday})$$

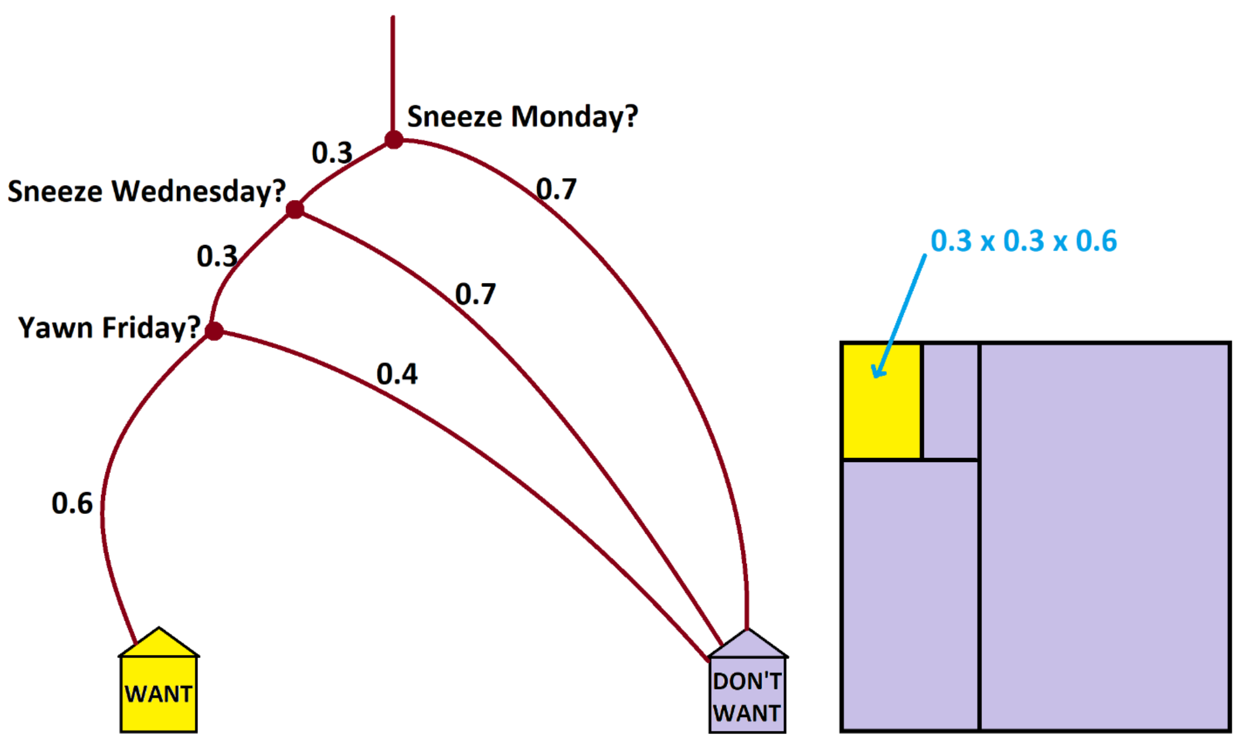
is $0.3$ as well, the general probability of sneezing, and

$$P(\text{Yawn on Friday} \mid \text{Sneezed on Wednesday \& Sneezed on Monday})$$

is $0.6$, the general probability of yawning.

Thus the probability we seek is the product of probabilities

$$\begin{aligned} P(\text{Sneeze Monday, and then Sneeze Wednesday, and then Yawn Friday}) &= 0.3 \times 0.3 \times 0.6 \\ &= 5.4\% \end{aligned}$$

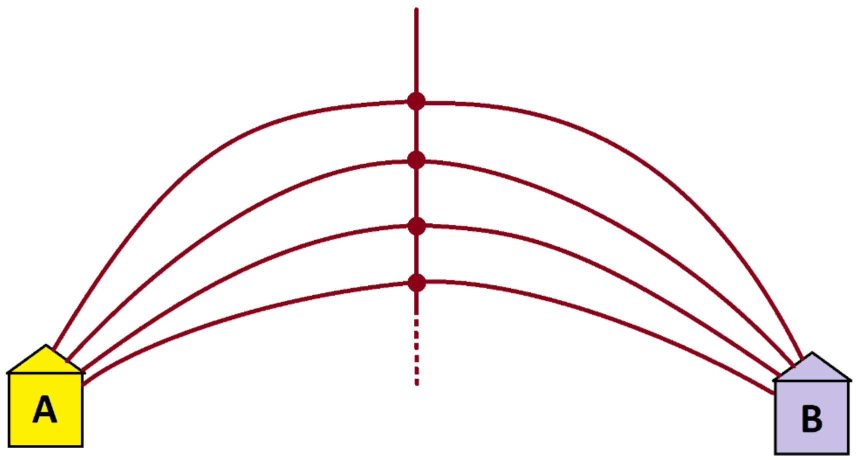


# PART III

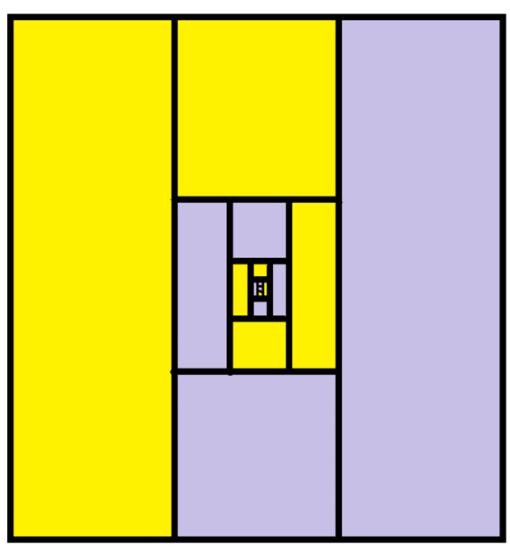
## Lesson 8: Infinite Garden Paths, Again

Consider again the infinite garden path system of lesson 3.




Folk heading down this system first encounter a three-way fork. Those who turn to the left head to house A, those who turn to the right head to house B, and those who go straight repeat this process with another three-way fork. This garden system has infinitely many forks (we are certainly playing a game of the mind here—we humans will never be able to build such a system) and we have to assume that an infinite number of people will be walking through this system.

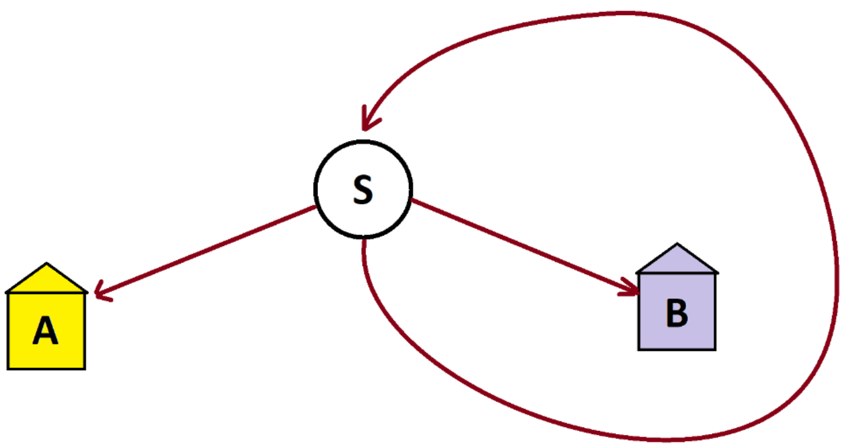
In lesson 3 we were led to the following area-model diagram showing, at the end of time, half the people sent through this system will end up in house A and half will end up in house B.



Notice that any subset of people who head straight at a fork are sent to an identical fork with the same infinite structure below it. It is as though we sent these people back to start.

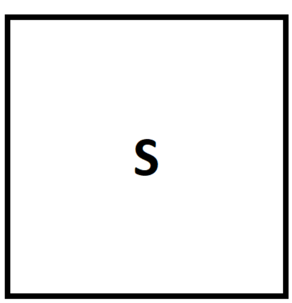
This means we can simplify our garden-path diagram by doing exactly that! Let's designate a starting location for people, label it S, and draw a single three-way fork at that start location, with people turning to the left heading to house A, people heading to the right to house B, and the people going straight returning to start to repeat the entire process.



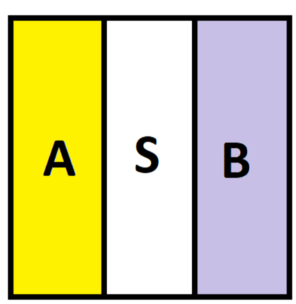

Initially we have an inconceivably large number of people in the start node S. After one run through the system, one third of those folk end up in house A, one third end up in house B, and one third end up back at start.

After a next run through the system, those at start split into thirds again. And after another run through the system, those at start yet again split into thirds again. And so on.

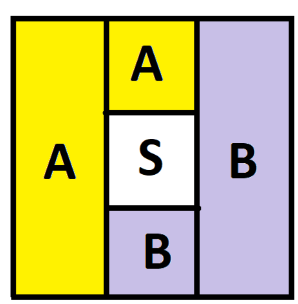
We see that we are generating the same infinite spiral picture above and coming to the same conclusion that, at the end of time, half the people in the system will end up at house A and half in house B.




**All folk inititially at START**



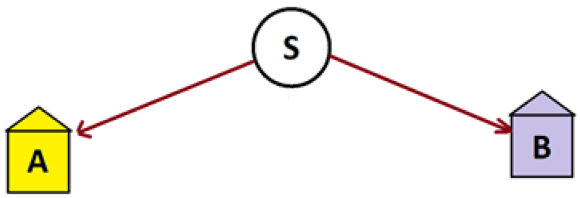

**After one run through the system**

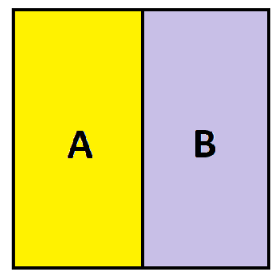



**After two runs through the system**

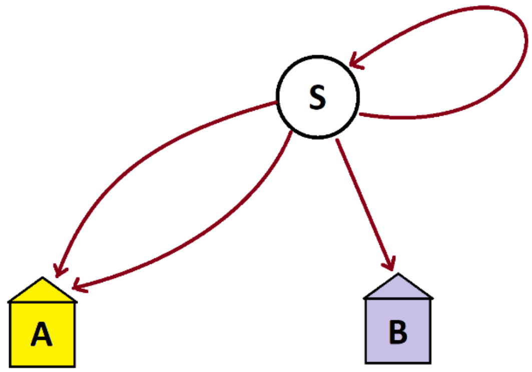
But we could also argue that, eventually, each and every person will end up in a house—either house A or house B. Some people might end up there right away. Others might go back to start a few times and then end up in a house. Others might go back to start tens of thousands of times before heading into a house. Since we do not care how many times people go back to start before entering a house, we could argue that our infinite system is philosophically equivalent to a system with everyone entering a house right away.

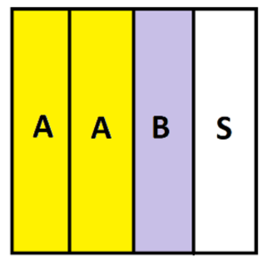


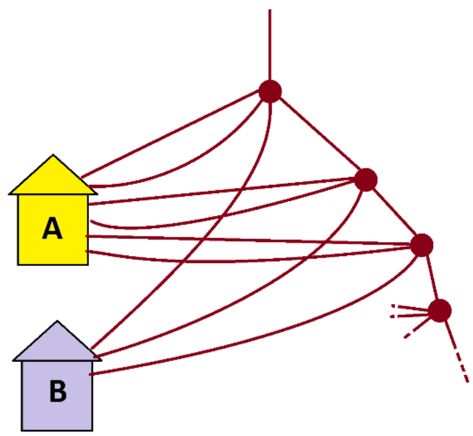

We conclude again that half the people will end up in house A and half will end up in house B.

Let's double-check this reasoning by examining a more complicated system. Consider this example.

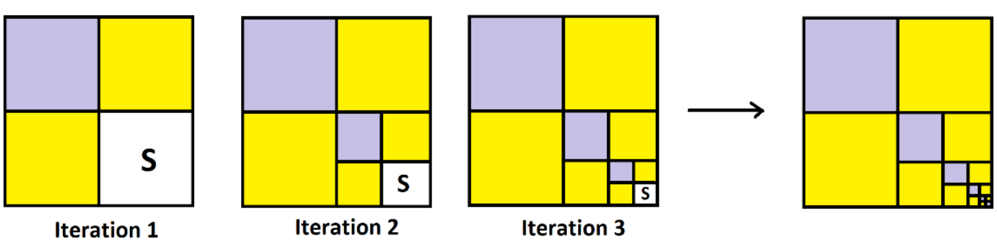



This corresponds to garden path system with an infinite number of four-way forks, each with two paths leading to house A, one path leading to house B, and one path leading to an identical four-way fork.



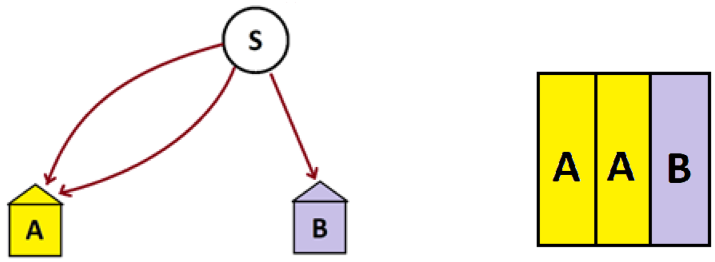

Divide the square into quarters this different way and see a pretty design emerge.




We see that two-thirds of the people end up in house A and one-third in house B.

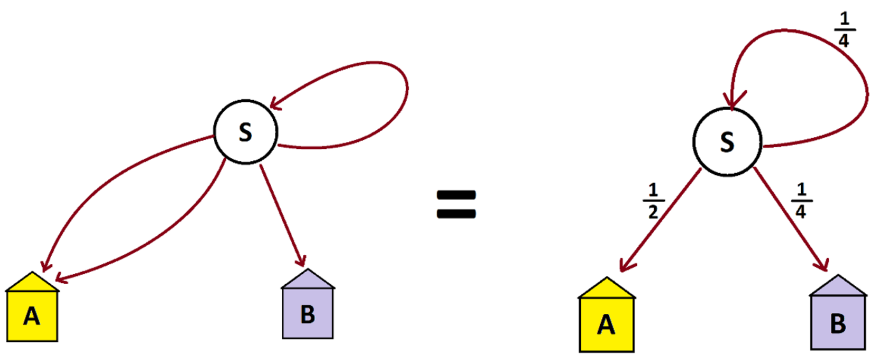
 

But we could have seen this by arguing that, philosophically, this garden path system is equivalent to one without a loop back to start: after all, everyone from start will eventually follow a path to a house.

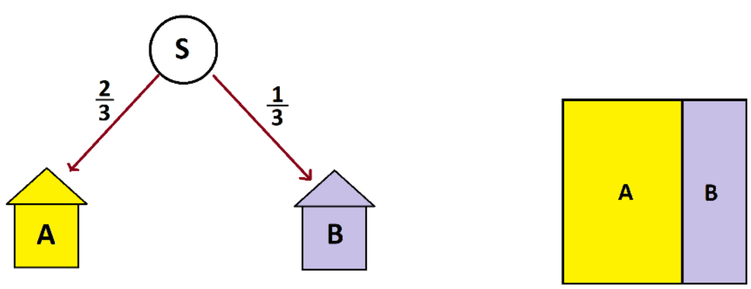



The two-thirds and one-third distribution is clear.

**Comment:** We could have presented this example as a weighted garden path system.



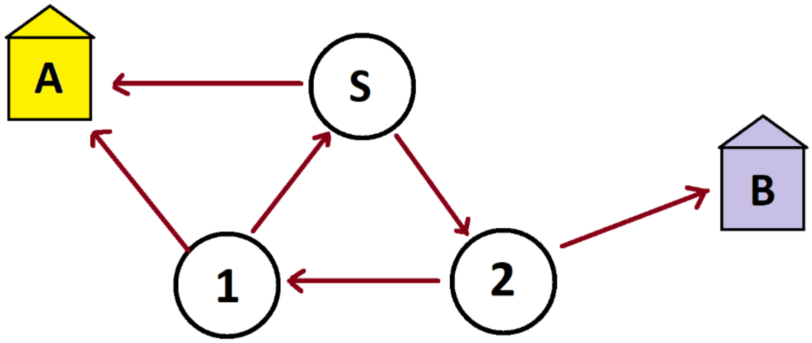

Notice that double the number of people head to house A than they do to house B. So when we ignore the loop back to start for philosophical reasons, we need to keep in mind that 2:1 ratio. It must be that **two**-thirds of people go to house A and **one**-third to house B.



**A Complicated Example:**

Consider this strange abstract example. (We'll do some more concrete examples in the next lesson.)




Here an inconceivably large count of people begin in start node S. Half will move straight to house A and half to a node labeled 2.

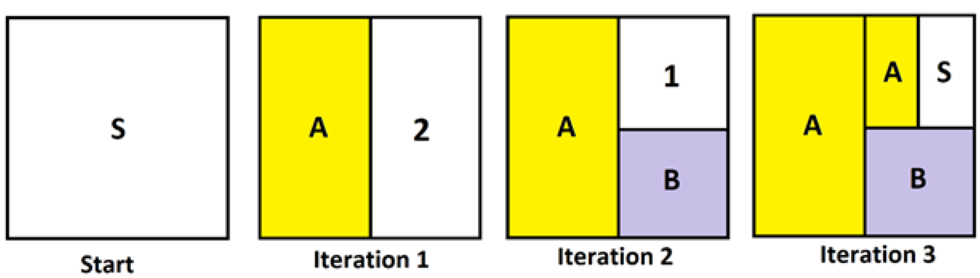
> **Comment:** We see now a standard convention: assume that outgoing edges are equally weighted if they are unlabeled. The two edges out from S each have weight $\frac{1}{2}$.

Of those who land in node 2, half will move to house B and half will move to a node labeled 1.

Of those who land in node 1, half will move to house A and half will go back to start to repeat the whole process.

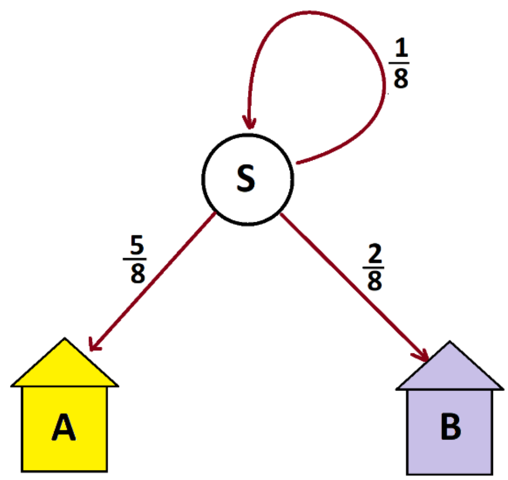
*What proportion of people end up in each house after infinite number if runs of this complex system of paths?*

Here's our area model for a few iterations.

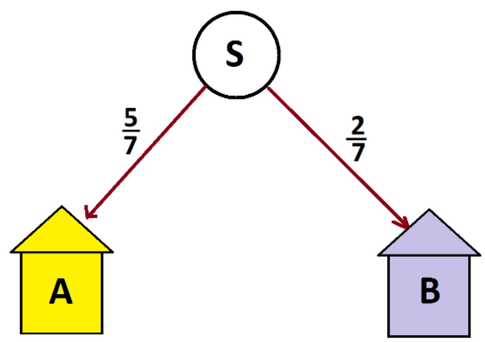



We see that the third iteration brings everyone into a house or back to start. In fact, one-eighth of the people are back at start, **five**-eighths are in house A, and **two**-eighths are in house B. It is as though we have been playing with this system.




Thus people enter houses A and B in counts of ratio 5 to 2. Since each and every person does eventually enter a house, we see that this philosophically equivalent to this system.

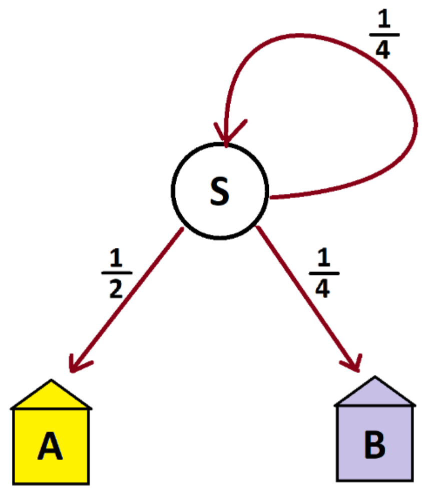



We have that five-sevenths of the people end up in house A and two-sevenths in house B.

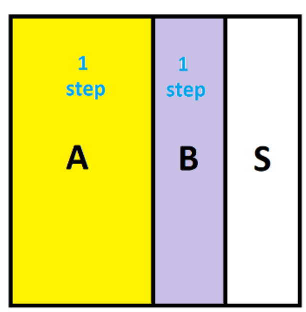
## OPTIONAL: Counting the Expected Number of Steps

Up to now we have been ignoring how many steps it possibly takes people to reach a house. This allowed us to—philosophically—simplify our garden path systems significantly. But let's now keep track of the number of steps people take. (And now, we are not permitted to simplify the garden paths!)

Here's a system we've seen. We concluded that two-thirds of the people end up in house A and one-third in house B.

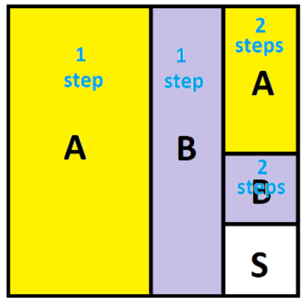



Let's now go through iterations of this system to see how many steps it takes people to reach a house. After one iteration we have this picture with three-quarters of the people in a house after one step.




**Iteration 1**

After a second iteration we see that $\frac{3}{16}$ of the people take two steps to land into a house.



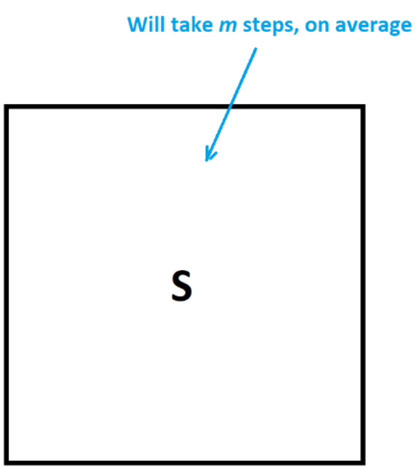

Iteration 2

And we see that $\frac{3}{64}$ of the people will take three steps to reach a house after a third iteration. And so on.

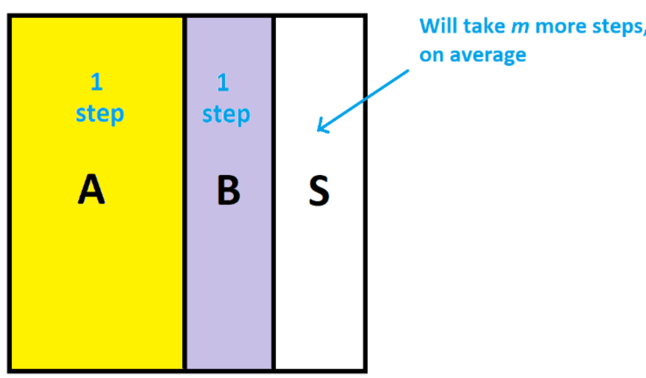
*What is the average number of steps folk take to reach a house?*

Can we answer this question? Yes!

Let' start by giving the unknown average value a name. Most people choose to call it $m$ for the *mean* number of steps. From the start state folk will take $m$ steps to reach a house—on average.




Start




Iteration 1

Have everyone take a step.

We see that three-quarters of the people reach a house in just 1 step, and one-quarter of the people will take, on average, an additional $m$ steps ($1+m$ steps in total) to reach a house. The average number of steps these people take is

$$\frac{3}{4}(1)+\frac{1}{4}(1+m)\ .$$

But these are the same people walking through the same system and so this average must be the same as the original average.

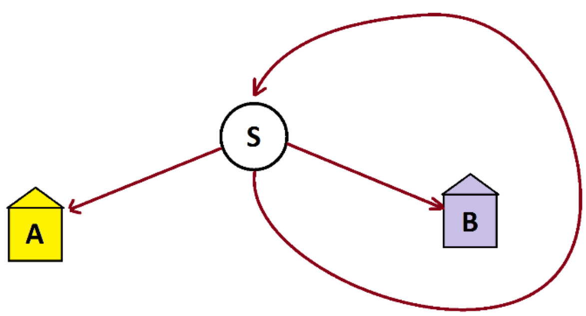
$$m=\frac{3}{4}(1)+\frac{1}{4}(1+m)\,.$$

Algebra now gives

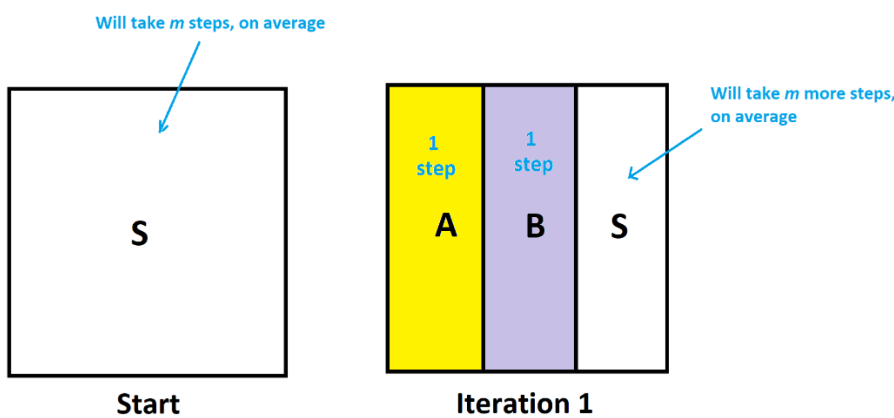
$$4m = 3+1+m$$
$$3m = 4$$
$$m=\frac{4}{3}$$

On average, folk take one-and-a-third steps to reach a house!

**Our First Example:** Here it is.



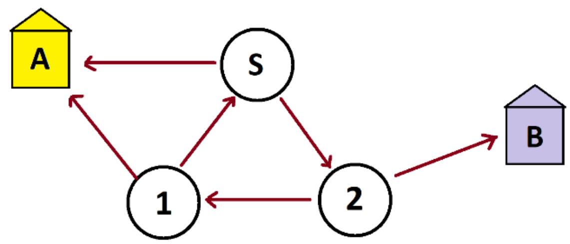

Suppose it take people an average on $m$ steps to reach a house.



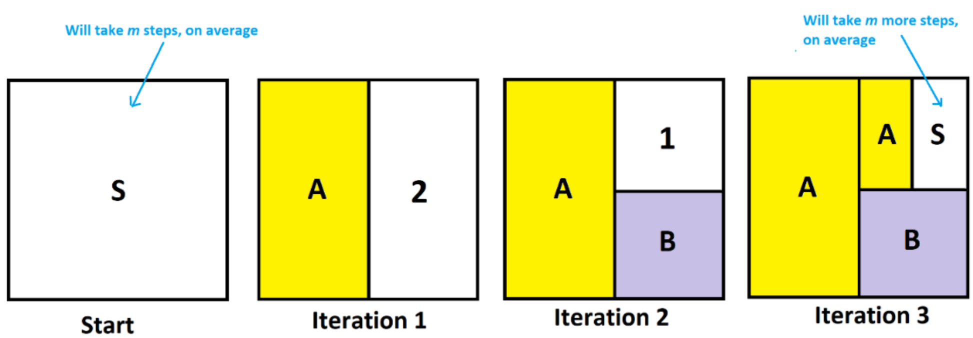
We see that $m=\frac{1}{3}(1)+\frac{1}{3}(1)+\frac{1}{3}(1+m)$ giving $m=\frac{3}{2}$.

**Question:** Does it make intuitive sense that, on average, people take more steps to reach a house than the folk in the previous example?

**Our Complicated Example:** Here it is.




Here are our first few iterations.



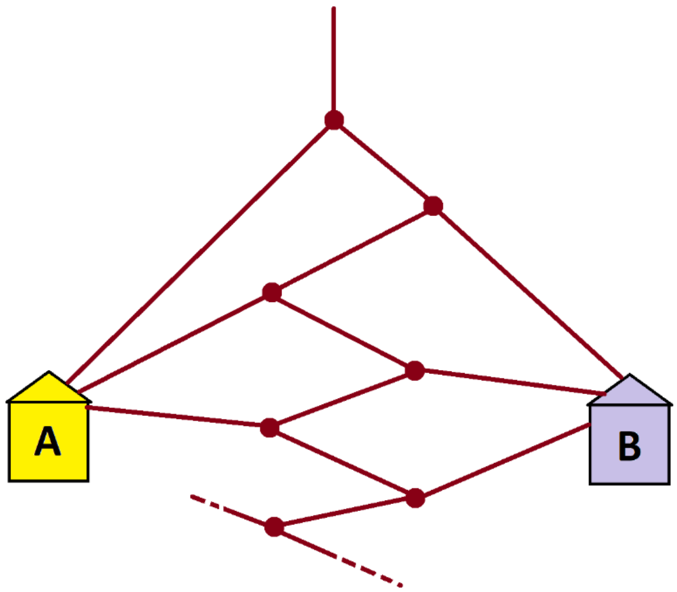

Let $m$ be the average count of steps people take to end up in a house. The leftmost and rightmost diagram show

$$m=\frac{1}{2}(1)+\frac{1}{4}(2)+\frac{1}{8}(3)+\frac{1}{8}(3+m).$$

Solving gives $m=\frac{12}{7}$.

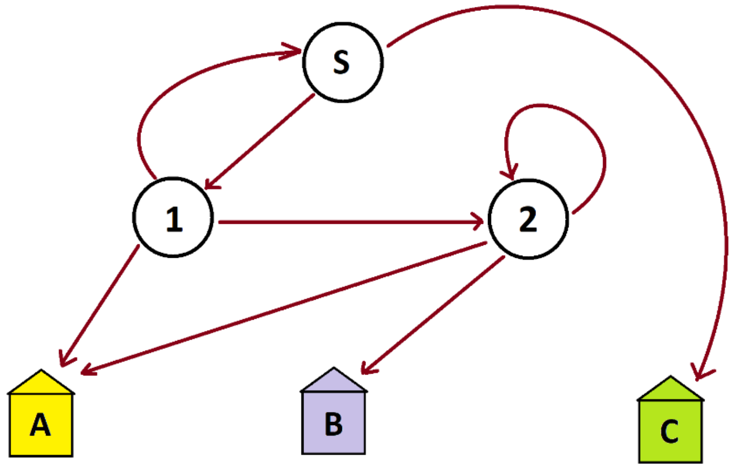
**Practice Problem 1:** Consider the following infinite garden-path system with two types of two-way forks: each fork sends half the people walking through it to a house and the other half to a next fork.




a) Using only a start node S, a second node "1", and the two houses A and B, redraw this garden path system much more succinctly.
b) In moving though this system, what proportion of people end up in house A and which proportion of people end up in house B?
c) OPTIONAL: On average, how many steps does a person take in moving to a house?

**Practice Problem 2:** Here is a very complicated system!




Show that 30% of the people moving through this system will end up in house A, 10% in house B, and 60% in house C.

**ASIDE ON JARGON**

A diagram with nodes, edges, and houses represents what is called an *absorbing Markov chain*. One imagines an individual roaming about the diagram, starting at one node (which we've always labeled S) and choosing to follow an edge from it with probability given by the weights labeled on the edges (or an outgoing edge with equal probability if no weights are given). As soon as person finds herself in a house, she stops, as there are no outgoing edges from a house. (Each house is an *absorbing state*.)

In this lesson we have been examining the probabilistic motion of a single person roaming about the diagram by imagining the average behavior of an inconceivably large count of people moving through the system. The proportion of people who end up in a particular house matches the probability of an individual within the system landing in that house. And $m$, the average count of steps people take, is the expected number of steps a person will take before landing in an absorbing state.

## Lesson 9: Infinite Probability Problems

Infinite garden paths are just terrific for analyzing (potentially) infinite probability experiments.

Consider this question.

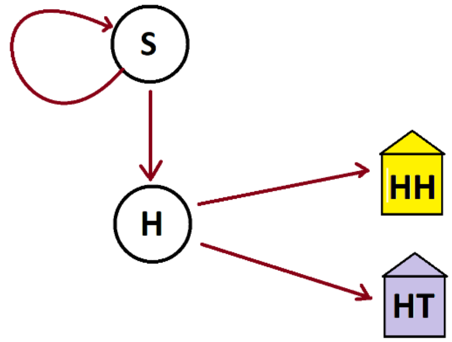
> *I will repeatedly toss a coin. What are the chances that I will see two heads tossed in a row (*HH*) before I'll see a head immediately followed by a tail (*HT*)?*

We can answer this question with pure logic.

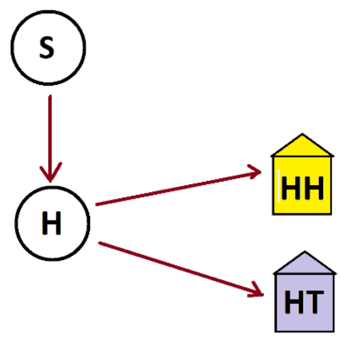
> **Answer 1:** If I first toss a tail or two tails or a string of tails of any length, I am no closer to seeing either HH or HT. This experiment doesn't "kick in" until I see my first head. Once I do, there is a 50% chance I'll see a head after that (to get HH) and a 50% chance I'll see a tail (to get HT). Thus there is a 50% chance I will see HH first.

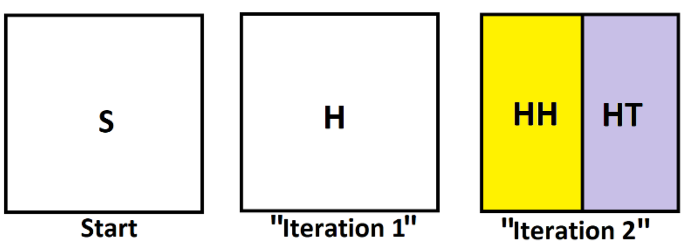
But we can also analyze this question by setting up a garden path system to model it.

> **Attempt Answer 2:** Imagine an inconceivably large number of people each about to repeatedly toss a coin. They all begin in the state "about to toss their coin for the first time," which we'll label S.
>
> On the first toss, some will toss a head (half of them in an ideal world) and be set for seeing either HH or HT. The other half will toss a tail, which is irrelevant to the outcome desired. We may as well send those folk back to the start state and have them try again.
>
> Of those who first toss a head, half (in an ideal world) will toss a head again to see HH and half a tail to see HT.
>
> We see now that having this inconceivable large set of people each repeatedly toss a coin is equivalent to sending this set of people through a system of garden paths that looks like this.




But everyone sent back to start will (eventually) toss a HEAD and move through the system. (What is the probability that someone will toss a TAIL infinitely many times in a row?) So, philosophically, this system of garden paths is equivalent to one without the loop back to start.




And indeed we see that 50% of the people running the experiment will see HH first and 50% will see HT first.

## EXAMPLE: A Dice Roll Question

*I will repeatedly roll a die. What are the chances that I will see the roll of a 1 and the roll of a 2 before I ever see a roll of a 6?*

To answer this question, imagine an inconceivably large number of people each rolling a die repeatedly. When folk get going, some will see both the roll of a 1 and a 2 before seeing a 6. Others won't. We want to know the proportion of people that do.

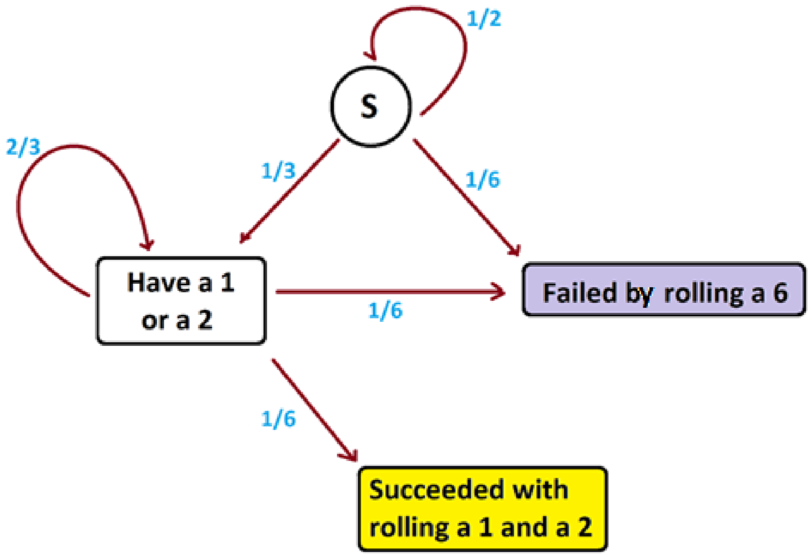
Everyone begins at a start state, about to start their rolls.

> Some will roll a 6 right away—one sixth of them in an ideal world—and will be immediately ruled out.
>
> Some will roll a 3, 4 or 5—one half of them in an ideal world. As none of these numbers is relevant to what we are examining they will roll again as though they are starting from the beginning of the experiment again.
>
> Some will roll a 1 or a 2 right away—one third of them in an ideal world—and will still be "in the game."

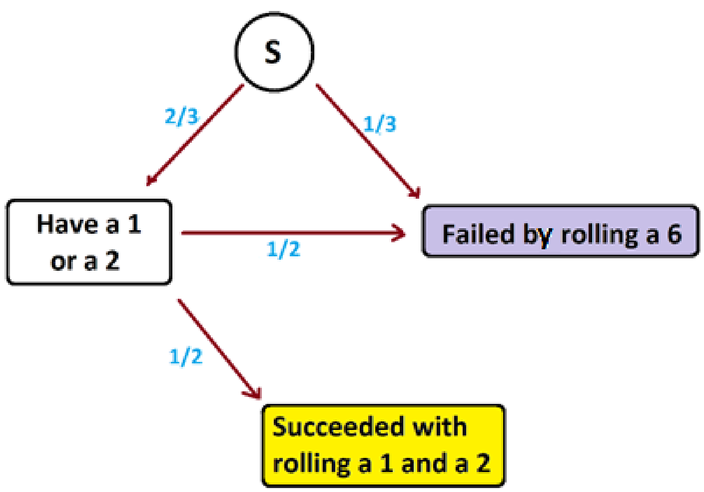
Of those in the last group …

> Some will roll a 6 on the next roll—one sixth of them in an ideal world—and will be ruled out.
>
> Some will roll a 3, 4, or 5—half of them in an ideal world—which would be considered an irrelevant roll and so will remain in the same predicament. So too will rolling a repeat 1 or a repeat 2, which will happen—in an ideal world—for one sixth of the people. In total, then, $\frac{1}{2}+\frac{1}{6}=\frac{2}{3}$ of these people will have an irrelevant roll and be back in the same state.
>
> The remaining folk—one sixth of them in an ideal world—will get what they are looking for, either a 1 if they earlier rolled a 2, or a 2 if they earlier rolled a 1.

The following system of garden-paths encapsulates the relationships between all these possible states.

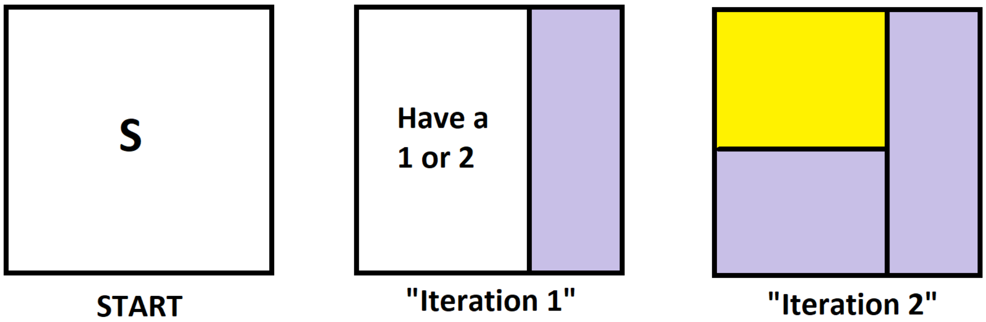



But if we are only interested in the proportions of folk ending up in each final state we can ignore loops from nodes back to themselves! Taking into account the weights on the various edges we see that this system is equivalent to the following system. (Make sure you understand the weights on edges you see in this second diagram.)




This gives the following area diagrams




from which we conclude that the probability of seeing both the roll of a 1 and a 2 before seeing the roll of a 6 is $\frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$.

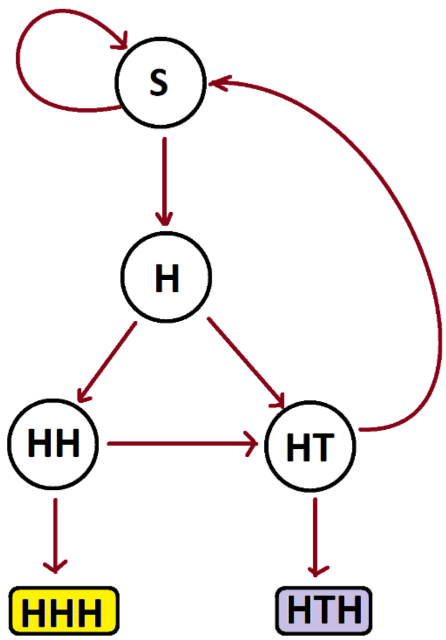
**Practice Problem 1:** *I will repeatedly roll a die. What are the chances that I will see either the roll of a 1 or the roll of a 2 before seeing a roll of a 6?*

**Practice Problem 2:** *I will repeatedly roll a die. What are the chances that I will see the roll of a 1 and a roll of a 2 and a roll of a 3 before seeing a roll of a 6?*

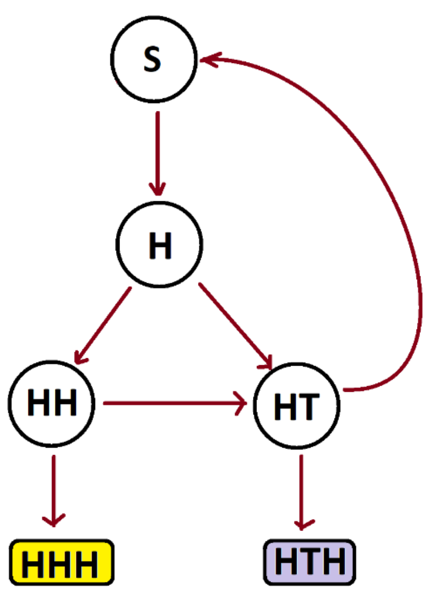
## Example: A Coin Tossing Question

*I will repeatedly toss a coin. What are the chances that I'll see three consecutive heads (*HHH*) before I see a head-tail-head triple (*HTH*)?*

Here is a garden path diagram modeling runs of this experiment. Each edge has equal weight of a half. (Did I get the diagram right?)

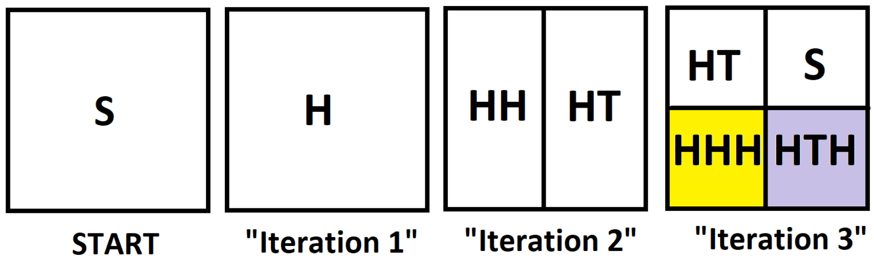



As we are only interested in the proportion of times we end up in each final state, we are permitted to ignore loops from nodes to themselves. So in this context this system is equivalent to the diagram:

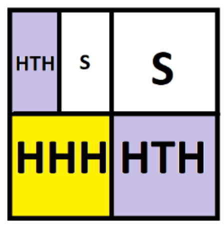


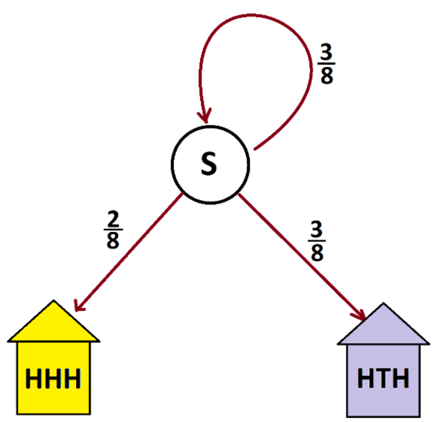
This leads to the following area-model diagrams.




Since we are not keeping track of how many steps folk are taking, we can look at the "HT" cell in the rightmost diagram and say that half of these folk HT will end up with HTH, which we don't want, and half the folk will end up back at start. (Look at the system of garden paths.) Thus we can see we really have the following area model.




And this is the model that arises from this system.




Ignoring the loop back to start, we see that the counts of people that end up in houses HHH and HTH are in a 2:3 ratio. Thus the probability of seeing HHH first is $\frac{2}{5}$ and the probability of seeing HTH first is $\frac{3}{5}$.

**Practice Problem 3:** *I will repeatedly toss a coin. What are the chances I will see a head immediately followed by a tail* (HT) *before seeing two consecutive tails* (TT)*?*

## References

This final lesson has brought us right up to the point of the ground-breaking chip-firing work of mathematics educator and scholar Arthur Engel and his paper:

*The Probabilistic Abacus*
Educational Studies in Mathematics, Vol. 6, No. 1 (Mar.,1975), pp. 1-22

Engel does not explain the mathematical theory of why his chip-firing method works. For that, see J. Laurie Snell's article

*The Engel Algorithm of Absorbing Markov Chains*
https://math.dartmouth.edu/~doyle/docs/engel/engel.pdf

Dr. James Propp has written a number of lovely essays on Engel's work in his *Mathematical Enchantment* series http://faculty.uml.edu/jpropp/mathenchant/. See essays 26 and 36, in particular, and also essays 25, 28, and 40.