

Lecture 11: February 25, 2019

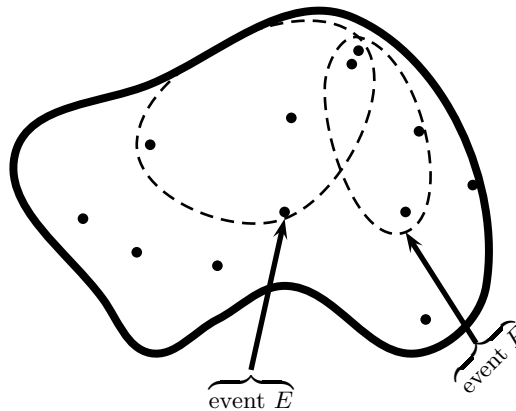
CS 330 Discrete Structures
Spring Semester, 2019

1 Conditional Probability

Consider now the following experiment: You are given three identical urns. Urn 1 contains 2 black balls. Urn 2 contains 2 gray balls. Urn 3 contains 1 black and 1 gray ball. You pick urn at random and take out a gray ball. How likely was it that you picked Urn 1? Urn 2? Urn 3?

For Urn 1, the result is easy since there are no gray balls. There are 2 gray balls in urn 2, so it has a $\frac{2}{3}$ probability while urn 3 has a $\frac{1}{3}$ probability.

Now let's replace the gray ball back into the urn and pick a ball again from the *same* urn. Again we get a gray ball; what is the probability that we had picked Urn 2? To answer this question we need to study **Conditional Probability**; this inverts the usual question of how likely an event *will be* to the question: given that we *know* that an event occurred (which could have been caused in various ways), what can we say about the probability that it was caused in a certain way. Specifically, conditional probability asks “What is the likelihood of event E happening if we already know that event F happened” and is written as: $\Pr\{E|F\}$. The situation is as follows:



If we know that event F happened, we know that the only part of event E that can happen is in the intersection of E and F . From this we have the result:

$$\begin{aligned}\Pr\{E|F\} &= \frac{\Pr\{E \wedge F\}}{\Pr\{F\}} \\ \Pr\{F|E\} &= \frac{\Pr\{F \wedge E\}}{\Pr\{E\}} \\ \Rightarrow \Pr\{E \wedge F\} &= \Pr\{E\}\Pr\{F|E\} \\ \Rightarrow \Pr\{E|F\} &= \frac{\Pr\{E\}\Pr\{F|E\}}{\Pr\{F\}}\end{aligned}\tag{1}$$

The identity in Equation (1) is quite important, and is known as **Bayes' Theorem**.¹ Note also that there is nothing special about the letters E and F ; we can interchange what events E and F represent and still have the same identities.

Here is an important example² from real life. An asymptomatic woman in her forties has a routine mammogram and the results come back positive; what is the probability that she actually has cancer? Mammograms are imperfect and can result in false positive results, so we know the *outcome* (a positive mammogram) and we want to know the probability of a *cause* (that the woman has cancer). We have the following data taken from the medical literature as of 2011:

$$\Pr\{\text{breast cancer in one's forties}\} = \frac{40}{10000}$$

$$\Pr\{\text{positive mammogram among breast cancer patients}\} = \Pr\{\text{positive mammogram} \mid \text{breast cancer}\} = \frac{32}{40}$$

$$\Pr\{\text{positive mammogram among all women}\} = \frac{1028}{10000}.$$

By Bayes' Theorem then, $\Pr\{\text{breast cancer in one's forties} \mid \text{positive mammogram}\}$

$$\begin{aligned} &= \frac{\Pr\{\text{breast cancer in one's forties}\} \times \Pr\{\text{positive mammogram} \mid \text{breast cancer}\}}{\Pr\{\text{positive mammogram}\}} \\ &= \frac{\frac{40}{10000} \times \frac{32}{40}}{\frac{1028}{10000}} \\ &\approx 0.03. \end{aligned}$$

This means that the likelihood of cancer in such a case is 3% which is somewhat reassuring for woman with the positive mammogram (though she must still undergo further tests to be sure).

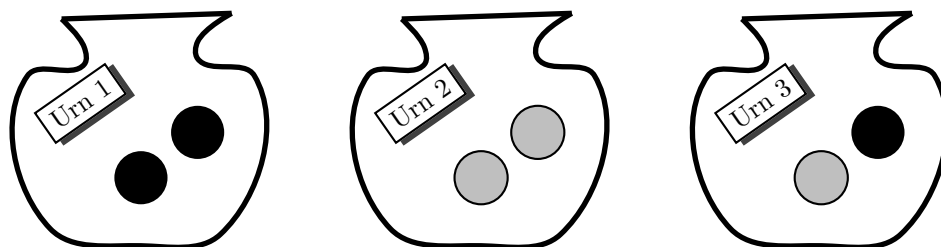
Exercise On the other hand, what is the probability that a woman *does* have breast cancer, given that she had a *negative* mammogram?

This type of Bayesian analysis is critical in automatic spelling correction (given what the person typed “misteak”, what are the probabilities that he/she meant to type “mistake”, “mistook”, or “steak”?) and spam filtering (given that the message contains the words “viagra” and “penis”, what is the probability that it is spam? See pages 497–501 in Rosen.)

Let's go back to the experiment with urns and balls. We have three urns, the Urn 1 contains two black balls, Urn 2 contains two gray balls and Urn 3 contains a black and a gray ball.

¹Reverend Thomas Bayes was a British cleric who was looking for a way to prove the existence of God. His method was to argue that, given what we see around us, what is the probability that God exists? For a fascinating history of the controversy around “Bayesian inference,” see *The Theory that Would not Die* by Sharon Bertsch McGrayne, Yale University Press, 2011 and *Chancing It* by Robert Matthews, Profile Books, 2016. The “controversy” has nothing to do with aleatory probability (which is our only concern) for which it is the pure combinatorial reasoning of Equation (1), but rather with epistemic probability. For epistemic probability, Bayes' Theorem asks (and answers) the question: How do we quantify what new information does to change our level of confidence. Matthews' book is an excellent, readable exploration of this topic.

²From Appendix B of McGrayne's book, but corrected here because she does not get it quite right.



Having selected an urn at random, we pull out a gray ball. What is the probability, for each urn, that we had selected that urn. In other words, if someone came to you and told you that they had picked a gray ball, and asked you to guess to which urn the ball belonged; what would be the probability of you being right if you guessed a particular urn? In order to answer this question, let us first examine the probability of pulling out a gray ball from each of the three urns.

The first urn has $\Pr\{\text{picked gray}\} = \frac{0}{2} = 0$ since there are no gray balls to pick.



The second urn has $\Pr\{\text{picked gray}\} = \frac{2}{2} = 1$ since there are only gray balls to pick.



The third urn has $\Pr\{\text{picked gray}\} = \frac{1}{2}$ since there is one gray and one black ball to pick.



We can now apply Conditional Probability to the problem: If we choose an urn at random and get a gray ball, what was the probability of having picked a particular urn? We begin by finding the probability of a gray ball overall:

$$\begin{aligned} \Pr\{\text{Gray}\} &= \Pr\{\text{Urn 1}\} \times \Pr\{\text{Gray in Urn 1}\} + \Pr\{\text{Urn 2}\} \times \Pr\{\text{Gray in Urn 2}\} \\ &\quad + \Pr\{\text{Urn 3}\} \times \Pr\{\text{Gray in Urn 3}\} \\ &= \frac{1}{3} \times 0 + \frac{1}{3} \times 1 + \frac{1}{3} \times \frac{1}{2} \\ &= \frac{1}{2}. \end{aligned}$$

Now we can write three separate equations (and apply the definition of conditional probability for the first and Bayes' Theorem for the next two):

$$\Pr\{\text{Urn 1}|\text{Gray}\} = \frac{\Pr\{\text{Urn 1}\} \times \Pr\{\text{Gray}|\text{Urn 1}\}}{\Pr\{\text{Gray}\}} = \frac{\frac{1}{3} \times 0}{\frac{1}{2}} = 0$$

$$\begin{aligned}\Pr\{\text{Urn 2}|\text{Gray}\} &= \frac{\Pr\{\text{Urn 2}\} \times \Pr\{\text{Gray}|\text{Urn 2}\}}{\Pr\{\text{Gray}\}} = \frac{\frac{1}{3} \times 1}{\frac{1}{2}} = \frac{2}{3} \\ \Pr\{\text{Urn 3}|\text{Gray}\} &= \frac{\Pr\{\text{Urn 3}\} \times \Pr\{\text{Gray}|\text{Urn 3}\}}{\Pr\{\text{Gray}\}} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{2}} = \frac{1}{3}\end{aligned}$$

Let us consider a related problem: If we choose an urn and pick one ball, return it to the urn and pick another ball from the same urn, and get two gray balls, what was the probability of having picked a particular urn? The probability of getting two gray balls overall is

$$\begin{aligned}\Pr\{2 \text{ Gray balls}\} &= \Pr\{\text{Urn 1}\} \times \Pr\{2 \text{ Gray balls from Urn 1}\} \\ &\quad + \Pr\{\text{Urn 2}\} \times \Pr\{2 \text{ Gray balls from Urn 2}\} \\ &\quad + \Pr\{\text{Urn 3}\} \times \Pr\{2 \text{ Gray balls from Urn 3}\} \\ &= \frac{1}{3} \times 0 + \frac{1}{3} \times 1 + \frac{1}{3} \times \frac{1}{4} \\ &= \frac{5}{12}.\end{aligned}$$

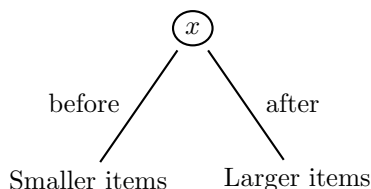
Applying Bayes' Theorem we have

$$\begin{aligned}\Pr\{\text{Urn 1}|2 \text{ Gray Balls}\} &= \frac{\Pr\{\text{Urn 1}\} \times \Pr\{2 \text{ Gray Balls}|\text{Urn 1}\}}{\Pr\{2 \text{ Gray Balls}\}} = \frac{\frac{1}{3} \times 0}{\frac{5}{12}} = 0 \\ \Pr\{\text{Urn 2}|2 \text{ Gray Balls}\} &= \frac{\Pr\{\text{Urn 2}\} \times \Pr\{2 \text{ Gray Balls}|\text{Urn 2}\}}{\Pr\{2 \text{ Gray Balls}\}} = \frac{\frac{1}{3} \times 1}{\frac{5}{12}} = \frac{4}{5} \\ \Pr\{\text{Urn 3}|2 \text{ Gray Balls}\} &= \frac{\Pr\{\text{Urn 3}\} \times \Pr\{2 \text{ Gray Balls}|\text{Urn 3}\}}{\Pr\{2 \text{ Gray Balls}\}} = \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{5}{12}} = \frac{1}{5}\end{aligned}$$

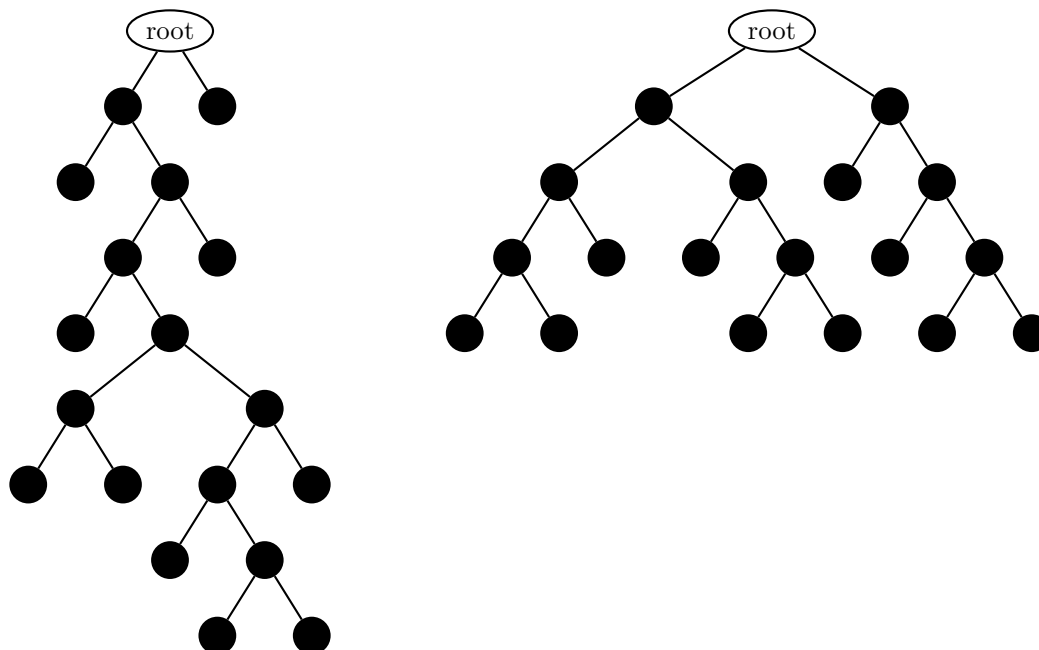
Exercise What if we get a gray ball three times? Four times? n times?

2 Probabilistically Balanced Lexicographical Trees

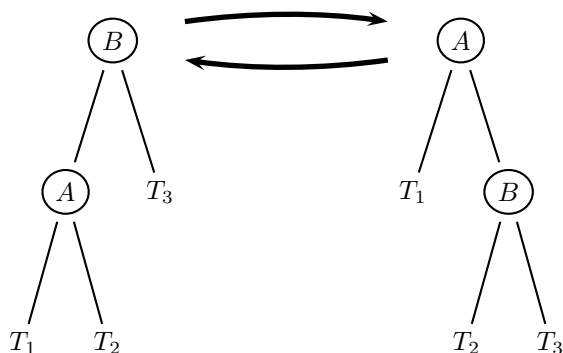
We can see computer application of Conditional Probability in probabilistically balanced lexicographical trees. A lexicographical tree is a data structure that embodies the idea of binary search. When you do a binary search, depending on the relative value of your current position and the value you are looking for, will descend either left or right. A lexicographical tree does the same:



This rule is then applied recursively to each and every node in the tree. To search through the tree, when you visit a node you compare what you are looking for with the value store in the node. If your value is less, examine the left child, otherwise examine the right child. The maximum number of nodes examined will be the height of the tree. What we want is a tree that has a low height as opposed to a tree with a big height:



You can see that the tree on the left is a much deeper tree, so it will take longer to search that the tree on the right. In fact, if a tree is really bad, all the nodes will be in a single line. This is why it's very important to balance these trees. The basic step in tree balancing is a **rotation**. In order to rearrange the tree so it's short and fat, we apply the following transformation to certain branches:

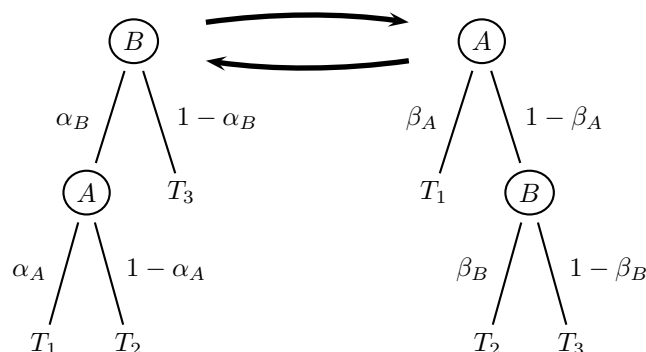


You will notice that the lexicographical property is not affected by this transformation; in other words, we may rotate any branch in a lexicographic tree and still have a lexicographic tree.

If we are given the *probability* that, when we visit a node, we will go right, and the *probability* that, when we visit a node, we will go left, we can attempt to equate these probabilities in order to balance our tree. The problem is that after the rotation, the various probabilities will change and mess up our results. We must find a way to keep track of the probabilities of going left and the probabilities of going right, even after we rotate a branch.

We will use α and β , respectively to refer to probabilities in the two trees. We label the various probabilities

concerning a rotation as follows:



We want to find the α s in terms of the β s, and vice versa. $\alpha_B = \Pr\{x < B\}$ where x is what we're searching for. We can see that for $x < B$ we need to go to either T_1 or T_2 . To go to T_2 we have to go through A and B . Using the rules of sum and product, we have

$$\alpha_B = \beta_A + (1 - \beta_A)\beta_B.$$

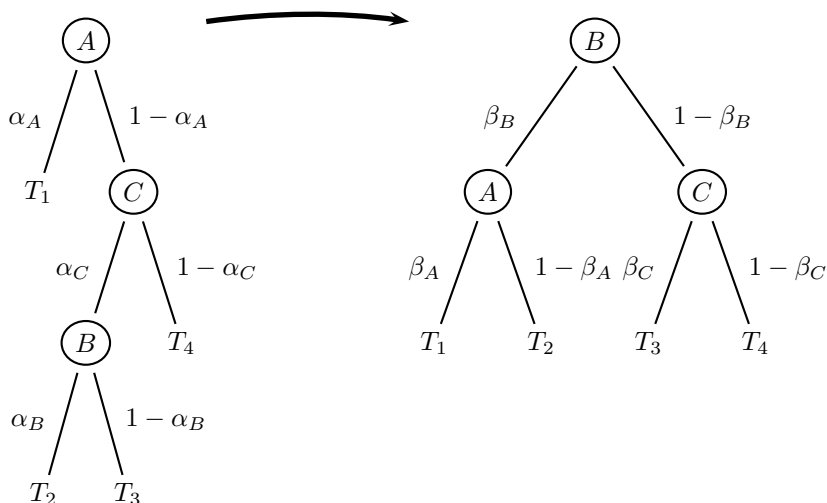
Similarly, we get:

$$\alpha_A = \Pr\{x < A | x < B\} = \frac{\Pr\{x < A\}\Pr\{x < B | x < A\}}{\Pr\{x < B\}} = \frac{\beta_A \cdot 1}{\beta_A + (1 - \beta_A)\beta_B}$$

$$\beta_B = \Pr\{x < B | x > A\} = \frac{\Pr\{x < B\}\Pr\{x > A | x < B\}}{\Pr\{x > A\}} = \frac{\alpha_B(1 - \alpha_A)}{1 - \alpha_B\alpha_A}$$

$$\beta_A = \Pr\{x < A\} = \alpha_B\alpha_A$$

We also need to consider the *double rotation* and the resulting probabilities:



Exercise Use Bayes' Theorem and the rules of sum and product to compute the α s in terms of the β s and vice versa.

3 More Examples

See the Wikipedia entries

http://en.wikipedia.org/wiki/Conditional_probability

http://en.wikipedia.org/wiki/Bayes'_theorem

for other examples and historical information.