Answers to Pause and Reflect Boxes for Chapter 17 Computational Linguistics
Robert Mercer

Pause and Reflect 17.1

The number 32 is obtained by multiplying 2 x 2 x 2 x 2 x 2. Using exponential representation this is 2^5. If you have a scientific calculator there is an exponentiation function and the number can be obtained by entering <2> <exponent button> <5>.

The reason for this combinatoric computation is that for each part of speech for the first word (two of them) there are two parts of speech for the second (giving us four combinations of parts of speech for the first two words). Then for each of these combinations we have two possible parts of speech for the third word. This gives us eight combinations. Continuing with this argument, the next word would lead to 16 combinations. And finally for the last word, this would give 32 combinations.

Pause and Reflect 17.2

The computation is again 2 x 2 x 2 x … x 2 (forty times) or 2^{40}. The numbers as the computation grows (I'll use exponential notation):

$2^1 = 2$; $2^2 = 4$; $2^3 = 8$; $2^4 = 16$; $2^5 = 32$; $2^6 = 64$; $2^7 = 128$; $2^8 = 256$; $2^9 = 512$; $2^{10} = 1024$; $2^{11} = 2048$; $2^{12} = 4096$; (I'll simply continue with the numbers) 8192; 16384; 32768; 65536; 131072; 262144; 524288; 1048576; 2097152; 4194304; 8388608; 16777216; 33554432; (this is 2^{25}; I let you continue with the next 15 multiplications by 2 or exponentiations). The final number is 1099511627776.

This growth is called exponential growth.

If the words are 3-way ambiguous, the method for counting the number of combinations is the same with only the base of the exponential factor changing to 3 from 2: i.e., 3 x 3 x 3 x … x 3 (forty times) or $3^{40} = 12157665459056928072$, which incidentally is more than 10 million times larger than 2^{40}.

Pause and Reflect 17.3

There are 36 combinations. Use the same counting method as used in Pause and Reflect 17.1. Die 1 has six sides each with a different number of dots in the range 1 to 6. When the dice are rolled, die 1 can show one of the six sides. For each orientation that die 1 can land in, die 2 can land in one of six possible orientations. Hence the number of combinations is 6 x 6. Note that when rolling the two dice, the final orientation of one die does not determine the final orientation of the other die. In probability this concept is called independence. That is, the two events (the final orientation of die 1 and the final orientation of die 2) are independent events. This allows us to compute the number of combinations of the two events in this combinatorial manner.

Represent each die as a number representing the number of dots that appear on the face-up face. Make the assumption the dice are fair (not loaded), that is, each face has the same likelihood of appearing face up as any of the other faces. The frequency of each number appearing on the two dice is shown under the column *Frequency*. This can be converted into probabilities by dividing the frequency by the total number of combinations. Frequencies add up to the number of combinations and probabilities add up to 1.

Number on the two dice	Die 1	Die 2	Frequency	Probability
2	1	1	1	1/36 = 0.0277777777778
3	1	2		
3	2	1	2	2/36 = 0.0555555555556
4	1	3		
4	2	2	3	3/16 = 0.0833333333333
4	3	1		
5	1	4		
5	2	3		
5	3	2	4	4/36 = 0.1111111111111
5	4	1		
6	1	5		
6	2	4		
6	3	3	5	5/36 = 0.1388888888889
6	4	2		
6	5	1		
7	1	6		
7	2	5		
7	3	4		
7	4	3	6	6/36 = 0.1666666666667
7	5	2		
7	6	1		
8	2	6		
8	3	5		
8	4	4	5	5/36 = 0.1388888888889
8	5	3		
8	6	2		
9	3	6		
9	4	5		
9	5	4	4	4/36 = 0.1111111111111
9	6	3		
10	4	6		
10	5	5	3	3/36 = 0.0833333333333

	10	6	4		
	11	5	6		
	11	6	5	2	2/36 = 0.05555555555556
	12	6	6	1	1/36 = 0.02777777777778
Totals				36	36/36 = 1.0

The probability of one of the faces being a 6 when the dice are rolled: Rearrange the lines of the above table into two groups, one group consists of lines with one die (or both, since this matches the criterion that one of the faces is a 6) is a 6 and the other group consists of lines which don't include a 6 on one of the die. Count the number of lines in the first group. This new table won't be shown because you can get this number simply by counting the number of events in the above table in which one of the dice is a 6. The number is 11. (The number of lines in the second group is just what remains: 25). The probability of a 6 appearing on one face when the dice are rolled is 11/36 = 0.30555555555556.

To compute conditional probabilities, one simple limits the lines in the above table that match the condition. The condition is: the number 7 is rolled. So, we have 6 lines. These are the combinations of die faces that give the number 7. So, this will be the number of combinations that are used to divide by. The number of lines in which a 6 appears in this limited set of lines is 2. Therefore, the conditional probability (i.e., the probability that a 6 appears when the roll of the dice gives a 7) is 2/6 = 0.33333333333333. Note that this is different than the probability of a 6 appearing on a die face for any roll of the dice.

If the condition is the roll of the dice is 11, the probability that one of the faces is a 6 is 2/2 = 1.0.

In computational linguistics we may want to know conditional probabilities such as: what is the probability that the word that I am looking at is a noun if the word preceding it is a determiner? (The word could be an adjective or an adverb.) The methodology to compute this probability is to take a very large corpus, count the number of times that a determiner occurs in the corpus (this is the condition and is the number used as the divisor) and then count the number of times a noun immediately follows a determiner in this same corpus. The conditional probability is computed as above (the second number divided by the first number). This conditional probability is then considered a good estimate of the probability that a noun immediately follows a determiner in any text.