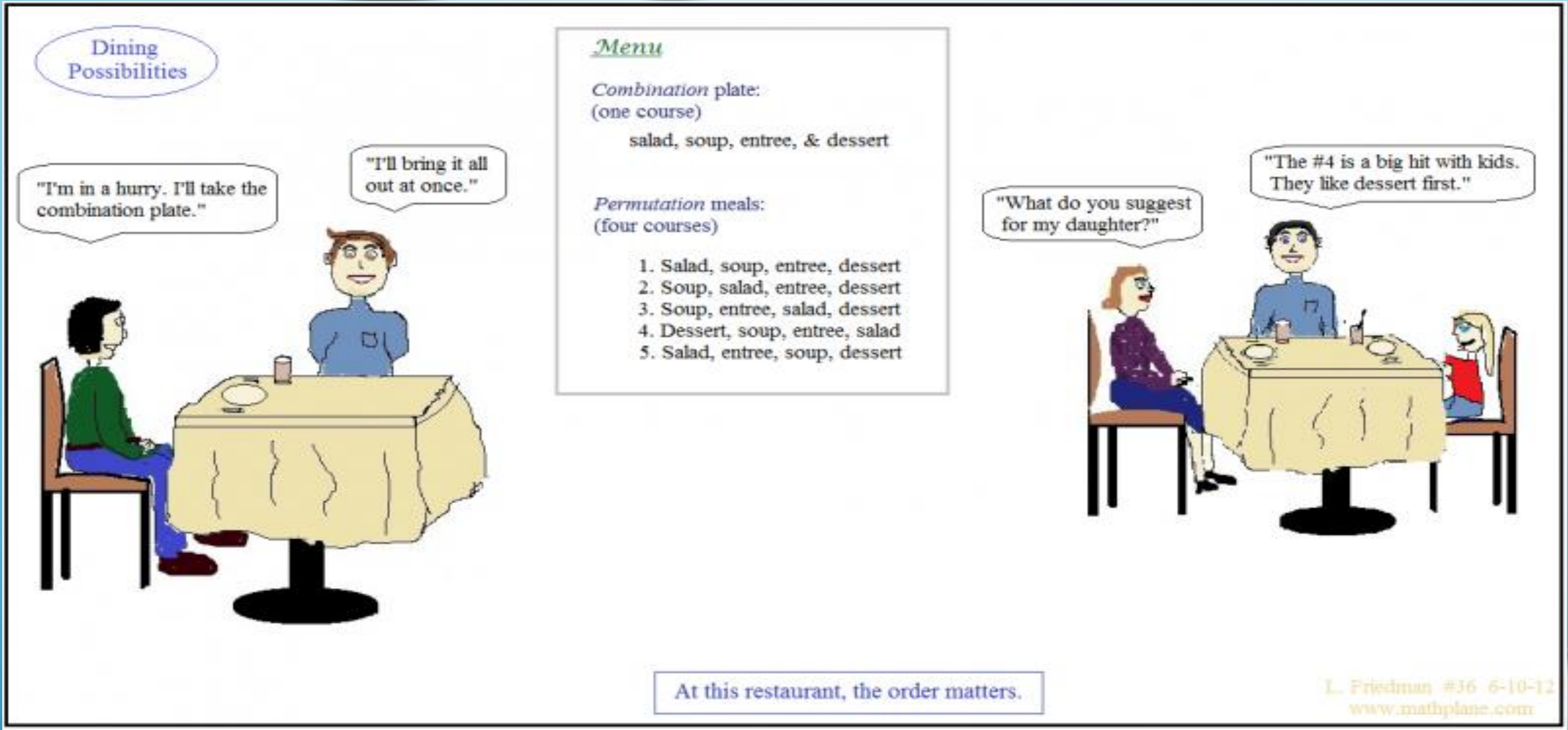


The basics of counting and the pigeonhole principle – 6.1&6.2

BASIC COUNTING PRINCIPLES: THE PRODUCT RULE

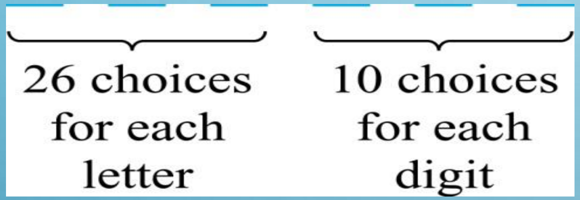
The Product Rule: A procedure can be broken down into a sequence of two tasks. There are n_1 ways to do the first task and n_2 ways to do the second task. Then there are $n_1 \cdot n_2$ ways to do the procedure.

Example: How many bit strings of length seven are there?

Solution: Since each of the seven bits is either a 0 or a 1, the answer is $2^7 = 128$.

Example: How many different license plates can be made if each plate contains a sequence of three uppercase English letters followed by three digits?
Solution: By the product rule,

there are 26 · 26 · 26 · 10 · 10 · 10 = 17,576,000 different possible license plates.



Example: An office building contains 27 floors and has 37 offices on each floor. How many offices are in the building?
Solution: By the product rule there are 27 · 37 = 999 offices.

DNA AND GENOMES

- A ***gene*** is a segment of a DNA molecule that encodes a particular protein and the entirety of genetic information of an organism is called its ***genome***.
- DNA molecules consist of two strands of blocks known as nucleotides. Each nucleotide is composed of bases: adenine (A), cytosine (C), guanine (G), or thymine (T).
- The DNA of bacteria has between 10^5 and 10^7 links (one of the four bases). Mammals have between 10^8 and 10^{10} links. So, by the product rule there are at least 4^{10^5} different sequences of bases in the DNA of bacteria and 4^{10^8} different sequences of bases in the DNA of mammals.
- The human genome includes approximately $23{,}000$ genes, each with $1{,}000$ or more links.

DNA AND GENOMES

Example: How many 5-element DNA sequences

a) end with A?

b) start with T and end with G?

c) contain only A and T?

d) do not contain C?

Solution: Recall that a DNA sequence is a sequence of letters, each of which is one of A, C, G, or T. Thus by the product rule there are $4^5 = 1024$ DNA sequences of length five if we impose no restrictions.

a) If the sequence must end with A, then there are only four positions at which to make a choice, so the answer is $4^4 = 256$.

b) If the sequence must start with T and end with G, then there are only three positions at which to make a choice, so the answer is $4^3 = 64$.

c) If only two letters can be used rather than four, the number of choices is $2^5 = 32$.

d) As in part (c), there are $3^5 = 243$ sequences that do not contain C.

BASIC COUNTING PRINCIPLES: THE SUM RULE

The Sum Rule: If a task can be done either in one of n_1 ways or in one of n_2, where none of the set of n_1 ways is the same as any of the n_2 ways, then there are $n_1 + n_2$ ways to do the task.

Example: The mathematics department must choose either a student or a faculty member as a representative for a university committee. How many choices are there for this representative if there are 37 members of the mathematics faculty and 83 mathematics majors and no one is both a faculty member and a student.

Solution: By the sum rule it follows that there are $37 + 83 = 120$ possible ways to pick a representative.

Example: How many bit strings are there of length six or less, not including the empty string?

Solution: We use the sum rule, adding the number of bit strings of each length up to 6. Then we get
$2^1 + 2^2 + 2^3 + 2^4 + 2^5 + 2^6 = 126$

THE SUM AND PRODUCT RULE IN TERMS OF SETS

- The sum rule can be phrased in terms of sets.

  $|A \cup B| = |A| + |B|$ as long as A and B are disjoint sets.

- Or more generally,

  $|A_1 \cup A_2 \cup \cdots \cup A_m| = |A_1| + |A_2| + \cdots + |A_m|$ when $A_i \cap A_j = \emptyset$ for all i, j.

- If $A_1, A_2, \ldots, A_m$ are finite sets, then the number of elements in the Cartesian product of these sets is the product of the number of elements of each set.

- The task of choosing an element in the Cartesian product $A_1 \times A_2 \times \cdots \times A_m$ is done by choosing an element in A_1, an element in A_2, …, and an element in A_m.

- By the product rule, it follows that:

$$|A_1 \times A_2 \times \cdots \times A_m| = |A_1| \cdot |A_2| \cdot \cdots \cdot |A_m|.$$

COMBINING THE SUM AND PRODUCT RULE

Combining the sum and product rule allows us to solve more complex problems.

Example: Suppose statement labels in a programming language can be either a single uppercase letter or an uppercase letter followed by a digit. Find the number of possible labels.

Solution: Use the product rule.

$$26 + 26 \cdot 10 = 286$$

Example: How many license plates can be made using either three digits followed by three uppercase English letters or four uppercase English letters followed by two digits?

Solution: $10^3 26^3 + 26^4 10^2 = 63273600$

COUNTING PASSWORDS

- **Example:** Each user on a computer system has a password, which is six to eight characters long, where each character is an uppercase letter or a digit. Each password must contain at least one digit. How many possible passwords are there?

  Solution: Let P be the total number of passwords, and let P_6, P_7, and P_8 be the passwords of length 6, 7, and 8.

  - By the sum rule $P = P_6 + P_7 + P_8$.
  - To find each of P_6, P_7, and P_8 , we find the number of passwords of the specified length composed of letters and digits and subtract the number composed only of letters.

$$P_6 = 36^6 - 26^6 = 2{,}176{,}782{,}336 - 308{,}915{,}776 = 1{,}867{,}866{,}560.$$

$$P_7 = 36^7 - 26^7 = 78{,}364{,}164{,}096 - 8{,}031{,}810{,}176 = 70{,}332{,}353{,}920.$$

$$P_8 = 36^8 - 26^8 = 2{,}821{,}109{,}907{,}456 - 208{,}827{,}064{,}576 = 2{,}612{,}282{,}842{,}880.$$

Consequently, $P = P_6 + P_7 + P_8 = 2{,}684{,}483{,}063{,}360.$

BASIC COUNTING PRINCIPLES: SUBTRACTION RULE

Subtraction Rule: If a task can be done either in one of n_1 ways or in one of n_2 ways, then the total number of ways to do the task is $n_1 + n_2$ minus the number of ways to do the task that are common to the two different ways.

- Also known as, the *principle of inclusion-exclusion*: $|A \cup B| = |A| + |B| - |A \cap B|$

Example: How many bit strings of length eight either start with a 1 bit or end with the two bits 00?

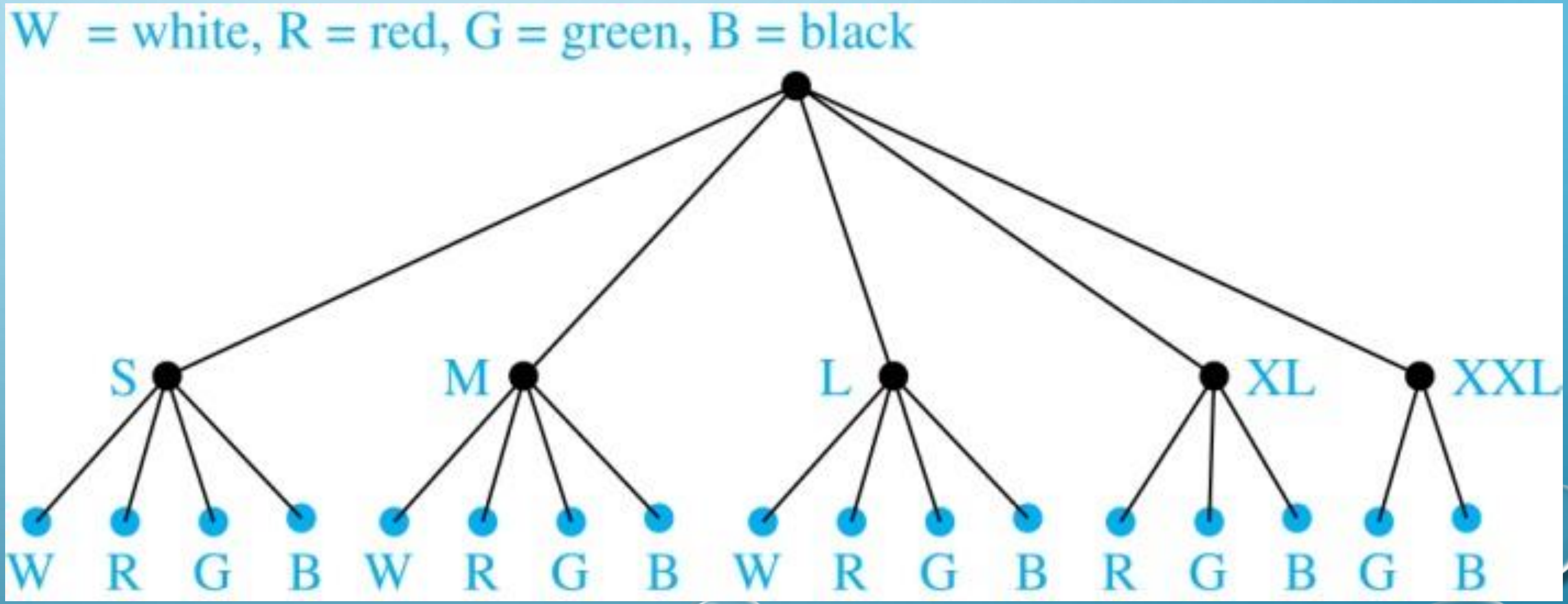
Solution: Use the subtraction rule.

- Number of bit strings of length eight that start with a 1 bit: $2^7 = 128$
- Number of bit strings of length eight that end with bits 00: $2^6 = 64$
- Number of bit strings of length eight that start with a 1 bit and end with bits 00 : $2^5 = 32$

Hence, the number is $128 + 64 - 32 = 160$.

TREE DIAGRAMS

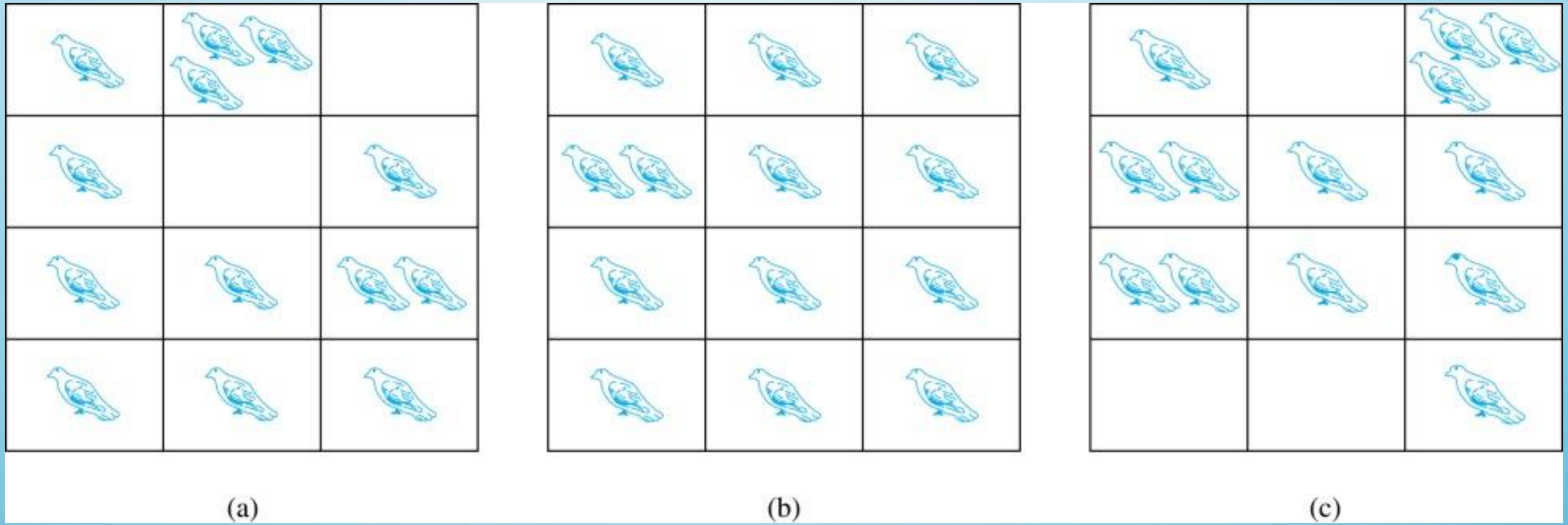
- **Tree Diagrams**: We can solve many counting problems through the use of *tree diagrams*, where a branch represents a possible choice and the leaves represent possible outcomes.
- **Example**: Suppose that "I Love Ashesi" T-shirts come in five different sizes: S, M, L, XL, and XXL. Each size comes in four colors (white, red, green, and black), except XL, which comes only in red, green, and black, and XXL, which comes only in green and black. What is the minimum number of shirts that the campus store needs to stock to have one of each size and color available?
- **Solution**: Draw the tree diagram.



- The store must stock 17 T-shirts.

THE PIGEONHOLE PRINCIPLE

- If a flock of 20 pigeons roosts in a set of 19 pigeonholes, one of the pigeonholes must have more than 1 pigeon.



Pigeonhole Principle: If k is a positive integer and $k + 1$ objects are placed into k boxes, then at least one box contains two or more objects.

Proof: We use a proof by contraposition. Suppose none of the k boxes has more than one object. Then the total number of objects would be at most k. This contradicts the statement that we have $k + 1$ objects. ◀

THE GENERALIZED PIGEONHOLE PRINCIPLE

The Generalized Pigeonhole Principle: If N objects are placed into k boxes, then there is at least one box containing at least $\lceil N/k \rceil$ objects.

Proof: We use a proof by contraposition. Suppose that none of the boxes contains more than $\lceil N/k \rceil - 1$ objects. Then the total number of objects is at most

$$k\left(\left\lceil \frac{N}{k} \right\rceil - 1\right) < k\left(\left(\frac{N}{k} + 1\right) - 1\right) = N,$$

where the inequality $\lceil N/k \rceil < N/k + 1$ has been used. This is a contradiction because there are a total of N objects.

Example: Among 100 people there are at least $\lceil 100/12 \rceil = 9$ who were born in the same month. ◀