



**UNSW**  
SYDNEY

# COMP9020

Foundations of Computer Science

Lecture 14: Combinatorics

## Topic 4: Probability



		[LLM]	[RW]	[Rosen]
Week 9	Combinatorics	Ch. 14	Ch. 5	Ch. 6, 8
Week 10	Probability	Ch 16, 17	Ch. 9	Ch. 7
Week 10	Statistics	Ch. 18	Ch. 9	Ch. 7

# Combinatorics in Computer Science

## Combinatorics:

- Computing cost functions in algorithmic analysis
- Identifying (in-)efficiencies in data management
- Developing effective techniques for enumerating objects
- Probability calculations

# Probability in Computer Science

## Probability:

- Artificial Intelligence
  - Machine Learning
  - Decision theory
  - Image processing
  - Speech recognition
- Algorithms
  - Algorithm analysis
  - Big Data sampling and analysis
- Security
  - Cryptography
  - Quantum computing
- Networks
  - Network traffic modelling
  - Reliability modelling

# Statistics in Computer Science

## Statistics:

- Sampling from large data sets
- Identifying anomalies
- Making predictions

# Outline

Counting Principles

Basic Counting Rules: Union

Basic Counting Rules: Product

Combinations and Permutations

Alternative Techniques

Difficult Counting Problems

# Outline

## Counting Principles

Basic Counting Rules: Union

Basic Counting Rules: Product

Combinations and Permutations

Alternative Techniques

Difficult Counting Problems

# Counting Techniques

General idea: find methods, algorithms or precise formulae to count the number of elements in various sets or collections derived, in a structured way, from some basic sets.

## Examples

Single base set  $S = \{s_1, \dots, s_n\}$ ,  $|S| = n$ ; find the number of

- all subsets of  $S$
- ordered selections of  $r$  different elements of  $S$
- unordered selections of  $r$  different elements of  $S$
- selections of  $r$  elements from  $S$  such that ...
- functions  $S \rightarrow S$  (onto, 1-1)
- partitions of  $S$  into  $k$  equivalence classes
- graphs/trees with elements of  $S$  as labelled vertices/leaves



# Example

## Example

A restaurant has the following menu:

Starter	Main Course	Dessert
Soup	Fish	Ice-cream
Bread	Beef	Fruit
	Pork	Cheese
	Chicken	

How many:

- 3 course meals (Starter-Main-Dessert) are possible?
- 3 course meals (Any item for each course) are possible?
- 3 course meals (Any item, no duplicates) are possible?
- Meals consisting of 3 items (order is unimportant)?

# Example

## Example

A restaurant has the following menu:

Starter	Main Course	Dessert
Soup	Fish	Ice-cream
Bread	Beef	Fruit
	Pork	Cheese
	Chicken	

How many:

- Starter-Main-Dessert?
- Any item for 3 courses?
- Any item, no duplicates, for 3 courses?
- Meals of 3 items?

# Basic Counting Rules: Principles

Two simple rules:

- **Union rule** (“or”): If  $S$  and  $T$  are disjoint  $|S \cup T| = |S| + |T|$
- **Product rule** (“followed by”):  $|S \times T| = |S| \cdot |T|$

These cover many examples, though the rule application is not always obvious.

Common strategies:

- Direct application of the rule
- Relate unknown quantities to known quantities (e.g.  
 $|S| + |T| = |S \cup T| + |S \cap T|$ )
- Find a bijection to a set that can be counted

# Outline

Counting Principles

**Basic Counting Rules: Union**

Basic Counting Rules: Product

Combinations and Permutations

Alternative Techniques

Difficult Counting Problems

# The Union Rule

**Union rule** —  $S$  and  $T$  *disjoint*

$$|S \cup T| = |S| + |T|$$

$S_1, S_2, \dots, S_n$  pairwise disjoint ( $S_i \cap S_j = \emptyset$  for  $i \neq j$ )

$$|S_1 \cup \dots \cup S_n| = \sum |S_i|$$

## Example

How many numbers in  $A = [1, 2, \dots, 999]$  are divisible by 31 or 41?

$\lfloor 999/31 \rfloor = 32$  divisible by 31

$\lfloor 999/41 \rfloor = 24$  divisible by 41

No number in  $A$  divisible by both

Hence,  $32 + 24 = 56$  divisible by 31 or 41

# Consequences of the Union Rule

## Fact

*For any sets  $X, Y, Z$ :*

$$|Y \setminus X| = |Y| - |X \cap Y|$$

$$|X \cup Y| = |X| + |Y| - |X \cap Y|$$

$$\begin{aligned} |X \cup Y \cup Z| = & |X| + |Y| + |Z| \\ & - |X \cap Y| - |Y \cap Z| - |Z \cap X| \\ & + |X \cap Y \cap Z| \end{aligned}$$

## Fact

- If  $|S \cup T| = |S| + |T|$  then  $S$  and  $T$  are disjoint
- If  $|\bigcup_{i=1}^n S_i| = \sum_{i=1}^n |S_i|$  then  $S_i$  are pairwise disjoint
- If  $|T \setminus S| = |T| - |S|$  then  $S \subseteq T$

These properties can serve to identify cases when sets are disjoint (resp. one is contained in the other).

## Proof.

$|S| + |T| = |S \cup T|$  means  $|S \cap T| = |S| + |T| - |S \cup T| = 0$

$|T \setminus S| = |T| - |S|$  means  $|S \cap T| = |S|$  means  $S \subseteq T$  □

# Exercises

## Exercises

RW: 5.3.1 200 people. 150 swim or jog, 85 swim and 60 do both.  
How many jog?

RW: 5.6.38 (Supp) There are 100 problems, 75 of which are 'easy' and 40 'important'.  
What's the smallest number of easy *and* important problems?



# Outline

Counting Principles

Basic Counting Rules: Union

**Basic Counting Rules: Product**

Combinations and Permutations

Alternative Techniques

Difficult Counting Problems

# The Product Rule

## Product rule

$$|S_1 \times \dots \times S_k| = |S_1| \cdot |S_2| \cdots |S_k| = \prod_{i=1}^k |S_i|$$

If all  $S_i = S$  (the same set) and  $|S| = m$  then  $|S^k| = m^k$

## NB

*This counts the number of sequences where the first item is from  $S_1$ , the second is from  $S_2$ , and so on.*

## Example

Let  $\Sigma = \{a, b, c, d, e, f, g\}$ .

How many 5-letter words?

$$|\Sigma^5| = |\Sigma|^5 = 7^5 = 16,807$$

How many with no letter repeated?

# Product rule: Sequences of selections

## Question

*How can we count sequences when the underlying set changes?*

To count sequences *without replacement*:

- Define an order on the whole underlying set
- Select from  $[1, n]$ , where  $n$  is the size of the “remaining” set, and a selection of  $i$  represents choosing the  $i$ -th element in that set

## Example

Let  $\Sigma = \{a, b, c, d, e, f, g\}$ .

How many 5-letter words with no letter repeated?

$$\prod_{i=0}^4 (|\Sigma| - i) = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = 2,520$$

# Exercises

## Exercises

$S, T$  finite. How many functions  $S \rightarrow T$  are there?

RW: 5.1.19 Consider a *complete* graph on  $n$  vertices.

(a) No. of paths of length 3

(b) paths of length 3 with all vertices distinct

(c) paths of length 3 with all edges distinct

# Exercise

## Exercise

RW: 5.3.2  $S = [100 \dots 999]$ , thus  $|S| = 900$ .

(a) How many numbers have at least one digit that is a 3 or 7?

(b) How many numbers have a 3 *and* a 7?

# Combinatorial Symmetry

A **symmetry** of a mathematical object is a bijective mapping from the object to itself which preserves “structure”.

A (combinatorial) symmetry defines an equivalence relation where the equivalence classes all have the same size.

We are often interested in counting a set “up to symmetry”. That is, counting the number of equivalence classes.

This can also be stated as a constraint that identifies a specific item in each equivalence class (**symmetric constraint**).

## Definition

A *k-to-1 function* is a function that maps exactly  $k$  inputs to an output.

## NB

*A k-to-1 function defines the equivalence relation of a combinatorial symmetry and vice-versa.*

# Product rule: Symmetries and duplications

## Question

- *How can we count sequences when we have symmetric constraints?*
- *How can we count sequences when we have duplicates?*

## Example

Let  $\Sigma = \{a, b, c, d, e\}$ .

- How many 5-letter words with no letter repeated and  $a$  before  $b$  before  $c$ ?
- How many 5-letter words can be made from  $a, a, a, d, e$ ?

## NB

*The answer will be the same.*

## Product rule: Symmetries and duplications

- $S_1$  = sequences accounting for symmetry,
- $S_2$  = symmetries,
- $S$  = sequences without symmetry

$$S = S_1 \times S_2,$$

so

$$|S_1| = |S|/|S_2|$$

Alternatively,  $\frac{1}{|S_2|}$  of the  $|S|$  sequences meet the symmetric constraint.



## Product rule: Symmetries and duplications

### Example

Let  $\Sigma = \{a, b, c, d, e\}$ .

How many 5-letter words with no letter repeated and  $a$  before  $b$  before  $c$ ?

Let  $\Sigma' = \{a, b, c\}$ .

$$S = \prod_{i=0}^4 (|\Sigma| - i) = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

$$S_2 = \prod_{i=0}^2 (|\Sigma'| - i) = 3 \cdot 2 \cdot 1 = 6$$

$$\text{So } S_1 = 120/6 = 20$$

# Outline

Counting Principles

Basic Counting Rules: Union

Basic Counting Rules: Product

**Combinations and Permutations**

Alternative Techniques

Difficult Counting Problems

# Combinatorial Objects: How Many?

## **permutations**

Ordering of all objects from a set  $S$ ; equivalently: Selecting all objects while *recognising* the order of selection.

The number of permutations of  $n$  elements is

$$n! = n \cdot (n - 1) \cdots 1, \quad 0! = 1! = 1$$

## **$r$ -permutations (sequences without repetition)**

Selecting any  $r$  objects from a set  $S$  of size  $n$  without repetition while *recognising* the order of selection.

Their number is

$$(n)_r = {}^n P_r = n \cdot (n - 1) \cdots (n - r + 1) = \frac{n!}{(n - r)!}$$

# Permutations with duplicates

## Example

How many anagrams of ASSESS?

Label S's:  $AS_1S_2ES_3S_4$ :  $6!$

In each anagram we can label the S's in  $4!$  ways.

Suppose there are  $m$  anagrams. So  $m \cdot 4! = 6!$ , i.e.  $m = \frac{6!}{4!}$

## Example

Number of anagrams of MISSISSIPPI?  $\frac{11!}{4!4!2!}$

## ***r*-selections** (or: ***r*-combinations**)

Collecting any  $r$  distinct objects without repetition;  
equivalently: selecting  $r$  objects from a set  $S$  of size  $n$  and *not* recognising the order of selection.

Their number is

$$\binom{n}{r} = \frac{(n)_r}{r!} = \frac{n!}{(n-r)!r!} = \frac{n \cdot (n-1) \cdots (n-r+1)}{1 \cdot 2 \cdots r}$$

### **NB**

*These numbers are usually called binomial coefficients due to*

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \dots + b^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i$$

*Also defined for any  $\alpha \in \mathbb{R}$  as* 
$$\binom{\alpha}{r} = \frac{\alpha(\alpha-1) \cdots (\alpha-r+1)}{r!}$$

# Simple Counting Problems

## Example

**RW: 5.1.2** Give an example of a counting problem whose answer is

(a)  $(26)_{10}$

(b)  $\binom{26}{10}$

Draw 10 cards from a half deck (eg. black cards only)

(a) the cards are recorded in the order of appearance

(b) only the complete draw is recorded

## Examples

- Number of edges in a complete graph  $K_n$
- Number of diagonals in a convex polygon
- Number of poker hands
- Decisions in games, lotteries etc.

# Exercises

## Exercises

**RW: 5.1.6** From a group of 12 men and 16 women, how many committees can be chosen consisting of

(a) 7 members?

(b) 3 men and 4 women?

(c) 7 women or 7 men?

**RW: 5.1.7** As above, but any 4 people (male or female) out of 9 and two, Alice and Bob, unwilling to serve on the same committee.

# Counting Poker Hands

## Exercises

RW: 5.1.15 A poker hand consists of 5 cards drawn without replacement from a standard deck of 52 cards

$$\{A, 2-10, J, Q, K\} \times \{\text{club, spade, heart, diamond}\}$$

- (a) Number of “4 of a kind” hands (e.g. 4 Jacks)
  
- (b) Number of non-straight flushes, i.e. all cards of same suit but *not* consecutive (e.g. 8,9,10,J,K)



# Selecting items summary

Selecting  $k$  items from a set of  $n$  items:

With replacement	Order matters	Examples	Formula
Yes	Yes	Words of length $k$ (sequences of length $k$ )	$n^k$
No	Yes	$k$ -permutations	$(n)_k$
No	No	Subsets of size $k$	$\binom{n}{k}$
Yes	No	Multisets of size $k$	$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k}$

# “Balls in boxes”

Have  $n$  “distinguishable” boxes.

Have  $k$  balls which are either:

- ① Indistinguishable
- ② Distinguishable

How many ways to place balls in boxes with

- A At most one
- B Any number of

balls per box?

## NB

Suppose  $K$  is a set with  $|K| = k$  and  $N$  is a set with  $|N| = n$ :

- $2A$  counts the number of injective functions from  $K$  to  $N$
- $2B$  counts the number of functions from  $K$  to  $N$

## “Balls in boxes”

Case	Balls	Balls per box	Number
1A	Indist.	At most 1	$\binom{n}{k}$
1B	Indist.	Any number	$\binom{n+k-1}{k}$
2A	Dist.	At most 1	$(n)_k$
2B	Dist.	Any number	$n^k$

# Outline

Counting Principles

Basic Counting Rules: Union

Basic Counting Rules: Product

Combinations and Permutations

**Alternative Techniques**

Difficult Counting Problems

# Alternative techniques

What if the current techniques are unwieldy?

Other techniques for obtaining an exact count:

- Find a different approach for counting
- Make use of symmetries
- Make use of recursion
- Write a program (running time?)

# Example

## Example

How many sequences of 15 coin flips have an even number of heads?

- Using “balls in boxes”:  $\binom{15}{0} + \binom{15}{2} + \dots + \binom{15}{14}$
- Use symmetry:  $\frac{1}{2} \times 2^{15}$
- Use recursion:  $\text{Even}(n) = \text{Odd}(n-1) + \text{Even}(n-1);$   
 $\text{Odd}(n) = \text{Even}(n-1) + \text{Odd}(n-1)$

# Example

## Example

How many sequences of  $n$  coin flips do not contain  $HH$ ?

$$C(0) = 0$$

$$C(1) = 0$$

$$C(n) = C(n-1) + C(n-2) + 2^{n-2}$$

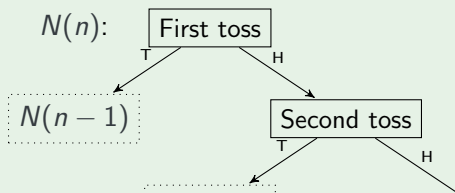
$$N(0) = 1$$

$$N(1) = 2$$

$$N(2) = 3$$

$$N(n) = N(n-1) + N(n-2)$$

We can summarise all possible outcomes in a **recursive tree**



# Outline

Counting Principles

Basic Counting Rules: Union

Basic Counting Rules: Product

Combinations and Permutations

Alternative Techniques

Difficult Counting Problems



# Difficult Counting Problems

## Example (Ramsay numbers)

An example of a *Ramsay number* is  $R(3, 3) = 6$ , meaning that  
*" $K_6$  is the smallest complete graph such that if all edges are painted using two colours, then there must be at least one monochromatic triangle"*

This serves as the basis of a game called S-I-M (invented by Simmons), where two adversaries connect six dots, respectively using blue and red lines. The objective is to *avoid* closing a triangle of one's own colour. The second player has a winning strategy, but the full analysis requires a computer program.

# Using Programs to Count

Two dice, a red die and a black die, are rolled.

(Note: one *die*, two or more *dice*)

Write a program to list all the pairs  $\{(R, B) : R > B\}$

Similarly, for three dice, list all triples  $R > B > G$

Generally, for  $n$  dice, all of which are  $m$ -sided ( $n \leq m$ ), list all *decreasing*  $n$ -tuples

## NB

*In order to just find the number of such  $n$ -tuples, it is not necessary to list them all. One can write a recurrence relation for these numbers and compute (or try to solve) it.*

# Approximate Counting

## NB

A Count may be a precise value or an **estimate**.

The latter should be *asymptotically correct* or at least give a good *asymptotic bound*, whether upper or lower. If  $S$  is the base set,  $|S| = n$  its size, and we denote by  $c(S)$  some collection of objects from  $S$  we are interested in, then we seek constants  $a, b$  such that

$$a \leq \lim_{n \rightarrow \infty} \frac{\text{est}(|c(S)|)}{|c(S)|} \leq b$$

In other words  $\text{est}(|c(S)|) \in \Theta(|c(S)|)$ .