Collections (Lists and Sets)

Chang-Gun Lee (cglee@snu.ac.kr)

Assistant Professor

The School of Computer Science and Engineering

Seoul National University

Collections

- The concept of collection is commonly used in computer science
  - set of memory blocks in cache
  - sorted list of data packets
  - set of active tasks in the ready queue
- Two types of collections
  - Lists: ordered collections
  - Sets: unordered collections

Lists

- A list is an ordered sequence of objects
  - (1, 2, 5)
  - The order in which elements appear in a list is significant
    - (1, 2, 3) and (3, 2, 1)
  - Elements in a list might be repeated
    - (3, 3, 2)
- “Length” of a list: the number of elements in the list
  - A list of length two: ordered pair
  - A list of length zero: empty set denoted by ( )
- Equality of two lists:
  - Two lists have the same length and elements in the corresponding positions on the two lists are equal
  - $(a, b, c) = (x, y, z)$ iff $a=x$, $b=y$, $x=z$

Lists are all-pervasive

- A point in the plane is often specified by an ordered pair of real numbers (x, y).
- A written natural number is a list of digits: 172 can be considered as the list (1, 7, 2).
- An identifier in a computer program is a list of letters and digits.
- The buffer in a network card holds a list of packets.

Counting Two-element Lists

- How many two-element lists we can make where the entries in the list are any of digits 1, 2, 3, and 4?

$$\begin{array}{cccc} (1,1) & (1,2) & (1,3) & (1,4) \\ (2,1) & (2,2) & (2,3) & (2,4) \\ (3,1) & (3,2) & (3,3) & (3,4) \\ (4,1) & (4,2) & (4,3) & (4,4) \end{array}$$

- How many two-element lists we can make where the first element has *n* choices and the second has *m* choices?

$$\begin{array}{cccc} (1,1) & (1,2) & \cdots & (1,m) \\ (2,1) & (2,2) & \cdots & (2,m) \\ \vdots & \vdots & \ddots & \vdots \\ (n,1) & (n,2) & \cdots & (n,m) \end{array}$$

Multiplication Principle

- Theorem 7.2 (Multiplication Principle) Consider two-element lists for which there are n choices for the first element, and for each choice of the first element there are m choices for the second element. Then the number of such lists is nm.
- Longer lists?
- Easy to extend: Consider k-element lists for which there are n_i choices for the i-th element. Then the number of such lists is $n_1 \, n_2 \, \ldots \, n_k$.

Two particular list-making problems

- Making a list of length *k* in which each element of the list is selected from among *n* possibilities.
- Problem 1: count all such lists permitting repetitions
- Problem 2: count those without repeated elements
- Theorem 7.6: The number of lists of length *k* whose elements are chosen from a pool of *n* possible elements

$$= \begin{cases} n^k & \text{if repetitions are permitted} \\ (n)_k = n(n-1)(n-2)\cdots(n-(k-1)) & \text{if repetitions are forbidden} \end{cases}$$

falling factorial

	비중복	중복
순열	${}_nP_k = (n)_k$	${}_n\Pi_k = n^k$
조합	${}_nC_k = \dfrac{{}_nP_k}{k!} = \dbinom{n}{k}$	${}_nH_k$

Counting Lists in CS?

- Algorithm complexity analysis: e.g., find the optimal packet ordering in a buffer
- Probabilistic performance evaluation: e.g., probability of no-collision packet transmission in distributed randomized priority mechanism?

Factorial and Product

- *n* Factorial: the number of length-*n* lists chosen from a pool of *n* objects in which repetition is forbidden

$$(n)_n = n! = n(n-1)(n-2)\cdots(1)$$

$$1! = 1$$

$$0! = 1$$

- Product notation

$$n! = \prod_{k=1}^{n} k = 1 \times 2 \times \cdots \times n$$

$$\prod_{k=1}^{5} (2k+3)$$

Sets

- Set: a repetition-free, unordered collection of objects (combination)
  - a given object can be either is a member of a set or it is not
  - an object cannot be in a set more than once
  - there is no order to the members of a set
  - {2, 3, 1/2}, {3, 1/2, 2}, {2, 2, 3, 1/2} are all the same
- Special sets
  - *Z*: set of integers
  - *N*: set of natural numbers
  - *Q*: set of rational numbers
  - *R*: set of real numbers
- Element: an object that belongs to a set is called an element of the set

$$x \in A: \ x \text{ is an element of set } A$$

$$y \notin A: \ y \text{ is not an element of set } A$$

- cardinality of *A*, size of *A*, |*A*| : number of elements in a set *A*
  - A "finite" set: if cardinality is an integer (i.e., finite)
  - A "infinite" set: if cardinality is infinite

– empty set, { }, ϕ: the set with no members

Set-builder notation

- Set builder notation

$$\{\text{dummy variable} \in \text{set} : \text{conditions}\}$$

$$\{x \in \mathrm{Z} : 2 \mid x\}$$

$$\{x \in \mathrm{Z} : 1 \le x \le 100\}$$

Equality of Sets

- Two sets are equal iff the two sets have exactly the same elements
- How to prove that two sets are equal?
- Proposition 9.1: The following two sets are equal:

$$E = \{x \in \mathrm{Z} : x \text{ is even}\}$$
$$F = \{x \in \mathrm{Z} : x = a + b \text{ where } a \text{ and } b \text{ are both odd}\}$$

Proof Template 5

- Proving two sets are equal
- Let A and B be the sets, To show A=B, we have the following template:
  - Suppose $x \in A$ ... Therefore $x \in B$
  - Suppose x $\in B$ ... Therefore x $\in A$
- Therefore A=B

Subsets

- Definition 9.2 (Subset) Suppose A and B are sets. We say that A is a subset of B (B is a superset of A) provided every element of A is also an element of B.

$$A \subseteq B$$

- "strict" or "proper" subset:

$$A \subseteq B \text{ and } A \neq B$$

Proof Template 6

- Proving one set is a subset of another
- To show $A \subseteq B$:
  - Let $x \in A$ ... Therefore $x \in B$
  - Therefore $A \subseteq B$

Proof Template 6

- Proposition 9.5: Let P be the set of Pythagorean triples; that is

$$P = \left\{(a,b,c) : a,b,c \in \mathrm{Z} \text{ and } a^2 + b^2 = c^2\right\}$$

and let T be the set

$$T = \left\{(p,q,r) : p = x^2 - y^2, q = 2xy, \text{ and } r = x^2 + y^2 \text{ where } x, y \in \mathrm{Z}\right\}$$

Then

$$T \subseteq P$$

Counting Subsets

- Theorem 9.7: Let A be a finite set. The number of subsets of A is $2^{|A|}$.
- Proof
  - bijective proof
  - its count is the same as the number of length-|A| lists from yes or no

Power Set

- Definition 9.8: Let A be a set. The power set of A is the set of all subsets of A.
- Example
  - the power set of {1,2,3}

$$\{\phi,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}$$

  - if a set A has n elements, its power set contains 2^n elements
- Notation for the power set of A: 2^A

$$|2^A| = 2^{|A|}$$

Counting Subsets of Size *k*

- How many combinations of *k* elements we can choose from *n* elements?
- How many sets of size *k* can be made from a set of *n* elements?

- Definition 16.1: Let $n, k \in N$. The symbol $\binom{n}{k}$ denotes the number of *k*-element subsets of an *n*-element set.
- How to calculate $\binom{n}{k}$?

$$\binom{n}{k} = {}_nC_k = \frac{{}_nP_k}{k!} = \frac{n!/(n-k)!}{k!}$$

	비중복	중복
순열	${}_nP_k = (n)_k$	${}_n\Pi_k = n^k$
조합	${}_nC_k = \frac{{}_nP_k}{k!} = \binom{n}{k}$	${}_nH_k$

Binomial Coefficients

- $\binom{n}{k}$ is called a binomial coefficient. Why?
- Try to expand $(x+y)^4$

$$(x+y)^4 = \binom{4}{0}x^4y^0 + \binom{4}{1}x^3y^1 + \binom{4}{2}x^2y^2 + \binom{4}{3}x^1y^3 + \binom{4}{4}x^0y^4$$

four boxes to fill in with either x or y

$$\boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}$$

$\binom{4}{2}$: number of cases to choose two boxes out of four

- Theorem 16.8 (Binomial)

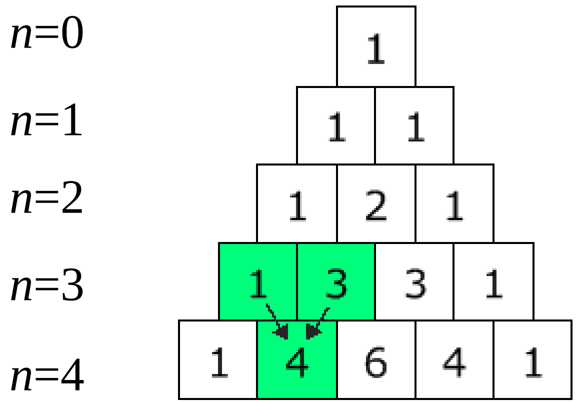
$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

- Proposition 16.7:

Let $n, k \in N$ with $0 \le k \le n$. Then

$$\binom{n}{k} = \binom{n}{n-k}$$

Pascal's Triangle



- The 0-th row contains just the single number 1
- Each successive row contains one more number than its predecessor
- The first and last number in every row is 1.
- An intermediate number in any row is formed by adding two numbers just to its left and jut to its right in the previous row.

- The entry in row n and column k is $\binom{n}{k}$.
- Theorem 16.10 (Pascal's Identity) Let n and k be integers with $0 < k < n$. Then

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

- Proof: Hint → combinatorial proof
  - Question: How many k-element subsets does an n-element set have?
  - Both LHS and RHS are correct answers!

“There is” and “For all”

- There is Quantifier:

$$\exists x \in A, \text{assertion about } x$$

$$\exists x \in N, \ \ x \text{ is prime and even}$$

- For all Quantifier:

$$\forall x \in A, \text{assertion about } x$$
$$\forall x \in Z, \ \ x \text{ is odd or } x \text{ is even}$$

- Combining Quantifiers:

$$\forall x, \exists y, x + y = 0$$

$$\exists y, \forall x, x + y = 0$$

Proof Template 7

- To prove

  $$\exists x \in A, \text{ assertion about } x$$

- Let x be (give an explicit example) ... (show that x satisfies the assertion).... Therefore x satisfies the required assertions.

- Example 10.1: $\exists x \in N, \ x \text{ is prime and even}$
- Proof

Consider the integer 2. Clearly 2 is even and 2 is prime. The statement holds.

Proof Template 8

- To prove

$$\forall x \in A, \text{ assertion about } x$$

- Let x be any element of A... (Show that x satisfies the assertion using only the fact that "x is in A" and no further assumptions on x).... Therefore x satisfies the required assertions.

- Example 10.2: Let $A = \{x \in Z : 6 \mid x\}$

$$\forall x \in A, \ \ x \text{ is even}$$

- Proof

For any x divisible by 6, there is an integer y such that x = 6y, which can be rewritten x = 2 (3y). Therefore x is divisible by 2 and therefore even.

Set Operations

- Union and Intersection

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$
$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

- Theorem 11.3: Let A, B, and C denote sets. The followings are true:

$A \cup B = B \cup A$ and $A \cap B = B \cap A$. (Commutative properties)

$A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$. (Associative properties)

$A \cup \phi = A$ and $A \cap \phi = \phi$

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

(Distributive properties)

- Proof? (Use Theorem 6.2 Boolean Algebra)

Size of Union

- Proposition 11.4: Let A and B be finite sets. Then

$$|A| + |B| = |A \cup B| + |A \cap B|$$

- Proof: Imagine we assign labels to every objects (label A to objects in set A and label B to objects in B
  - Pose a question: how many labels have we assigned?
  - LHS ($|A| + |B|$) is one correct answer for the above question.
  - RHS ($|A \cup B| + |A \cap B|$) is another correct answer.
  - Therefore LHS=RHS

- Example 11.5: How many integers in the range 1 to 1000 (inclusive) are divisible by 2 or by 5?.... Use

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Proof Template 9

- To prove an equation of the form LHS = RHS
  - Pose a question of the form, “In how many ways ....?”
  - On the one hand, argue why LHS is a correct answer to the above question
  - On the other hand, argue why RHS is also a correct answer.
  - Therefore LHS=RHS

Size of Union (2)

- Definition 11.6: (Disjoint) Let A and B be sets. We call *A* and *B* *disjoint* provided $A \cap B = \phi$
- (Pairwise Disjoint) Let A_1, A_2, ..., A_n be a collection of sets. These sets are called *pairwise disjoint* provided $A_i \cap A_j = \phi$ whenever $i \neq j$
- Corollary 11.8 (Addition Principle) Let A and B be finite sets. If A and B are disjoint, then
$$\left|A \cup B\right| = \left|A\right| + \left|B\right|$$
- Generalization of Addition Principle: If A_1, A_2, ..., A_n are pairwise disjoint sets, then

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{k=1}^{n} \left|A_k\right|$$

Inclusion-Exclusion

- Theorem 18.1 (Inclusion-Exclusion) Let A_1, A_2, ..., A_n be finite sets. Then

$$
\begin{aligned}
&|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n| \\
&- |A_1 \cap A_2| - |A_1 \cap A_3| - \cdots - |A_{n-1} \cap A_n| \\
&+ |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n| \\
&- \cdots + \cdots \cdots \\
&\pm |A_1 \cap A_2 \cap \cdots \cap A_n|
\end{aligned}
$$

How to use Inclusion-Exclusion?

- Example 18.3: The number of length-*k* lists whose elements are chosen from the set {1,2,...,*n*} is n^k. How many of these lists use all of the elements in {1,2,...,n} at least once?
- Hint!
  - # good lists = n^k - # bad lists
  - B1: set of all lists that do not contain 1 (bad because of missing 1)
  - B2: set of all lists that do not contain 2 (bad because of missing 2)
  - ....
  - # bad lists = $\left|B_1 \cup B_2 \cup \cdots \cup B_n\right|$

How to use Inclusion-Exclusion? (2)

- Example 18.4: There are $n!$ ways to make lists of length n using the elements of $\{1,2,...,n\}$ without repetition. Such a list is called a derangement if the number j does not occupy position j of the list for every $j = 1, 2, ..., n$. How many derangements are there?
- Hint!
  - # good lists = $n!$ - # bad lists
  - B1: set of all lists that have 1 in position 1 (bad because of 1)
  - B2: set of all lists that have 2 in position 2 (bad because of 2)
  - ....
  - # bad lists = $\left|B_1 \cup B_2 \cup \cdots \cup B_n\right|$

Proof Template 10
(Using inclusion-exclusion)

- Counting with inclusion-exclusion
  – Classify the objects as either “good” (the ones you want to count) or “bad” (the ones you don’t want to count).
  – Decide whether you want to count the good objects directly or to count the bad objects and subtract from the total.
  – Cast the counting problem as the size of a union of sets. Each set describes one way the objects might be “good” or “bad”.
  – Use inclusion-exclusion (Theorem 18.1)

Difference and Symmetric Difference

- Definition 11.9: (Set difference) Let A and B be sets. The *set difference*, *A-B*, is the set of all elements of *A* that are not in *B:*
$$A - B = \{x : x \in A \text{ and } x \notin B\}.$$
- (Symmetric difference) of A and B, denoted by $A \Delta B$, is the set of all elements in *A* but not *B* or in *B* but not *A*. That is
$$A \Delta B = (A - B) \cup (B - A).$$
- Proposition 11.11: Let A and B be sets. Then
$$A \Delta B = (A \cup B) - (A \cap B).$$
- Proof: (Use Proof Template 5)

(1) Suppose $x \in A \Delta B$ .... Therefore $x \in (A \cup B) - (A \cap B)$.
(2) Suppose $x \in (A \cup B) - (A \cap B)$ ..... Therefore $x \in A \Delta B$
Therefore $A \Delta B = (A \cup B) - (A \cap B)$.

DeMorgan's Laws

- Proposition 11.12 (DeMorgan's Laws): Let A, B, and C be sets. Then

$$A-\left(B\cup C\right)=\left(A-B\right)\cap\left(A-C\right)$$
$$A-\left(B\cap C\right)=\left(A-B\right)\cup\left(A-C\right)$$

Cartesian Product

- Definition 11.13 (Cartesian product): Let *A* and *B* be sets. The Cartesian product of *A* and *B*, denoted by *A* x *B*, is the set of all ordered pairs (two-element lists) formed by taking an element from *A* together with an element from *B* in all possible ways. That is,

$$A \times B = \{(a,b) : a \in A, b \in B\}$$

- Proposition 11.15: Let *A* and *B* be sets. Then

$$|A \times B| = |A| \times |B|$$

- Proof: ???

Two more examples of Combinatorial Proof

- Proposition 12.1: Let *n* be a positive integer. Then

$$2^0 + 2^1 + \cdots + 2^{n-1} = 2^n - 1$$

- Proof: (Use Proof Template 9)

> (1) Pose a question: How many non-empty subsets does {1, 2, ..., n} have?
> (2) RHS (2^n-1) is a correct answer
> (3) LHS is also a correct answer (why?)

- Proposition 12.2: Let *n* be a positive integer. Then

$$1 \cdot 1! + 2 \cdot 2! + \cdots + n \cdot n! = (n+1)! - 1.$$

- Proof: (Use Proof Template 9)

> (1) Pose a question: How many repetition-free length-(n+1) lists can we form from {1,2,...n+1} in which the elements do not appear in increasing order?
> (2) RHS is a correct answer
> (3) LHS is also a correct answer (why?)

Multisets

- A multiset is, like a set, an unordered collection of elements. However, in a multiset, an object may be considered to be in the multiset more than once.
- Ex: <1, 2, 3, 3>
  - Multiplicity of an element is the number of times it is a member of multiset. (The element 3 has multiplicity equal to 2).
- Two multisets are the same provided they contain the same elements with the same multiplicities.
  - <1, 2, 3, 3> = <3, 1, 3, 2>   <1, 2, 3, 3> ≠ <1, 2, 3, 3, 3>
- The cardinality of a multiset is the sum of the multiplicities of its elements.

Counting Multisets

- How many k-element multisets can we form by choosing elements from an n-element sets?
- How many unordered length-k lists can we form using the elements {1, 2, ..., n} with repetition allowed?
- Definition 17.1: Let $n, k \in N$. The symbol $\left(\!\binom{n}{k}\!\right)$ denotes the number of multisets with cardinality equal to *k* whose elements belong to an *n*-element set such as {1, 2, ..., n}.

	비중복	중복
순열	${}_nP_k = (n)_k$	${}_n\Pi_k = n^k$
조합	${}_nC_k = \frac{{}_nP_k}{k!} = \binom{n}{k}$	${}_nH_k = \left(\!\binom{n}{k}\!\right)$

Counting Multisets (2)

- Proposition 17.6: Let *n*, *k* be positive integers. Then

$$\left(\!\!\binom{n}{k}\!\!\right) = \left(\!\!\binom{n-1}{k}\!\!\right) + \left(\!\!\binom{n}{k-1}\!\!\right)$$

- Theorem 17.8: Let *n*, *k* be positive integers. Then

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k}$$

- Proof: Hint strange encoding of a multiset

$$\langle 1,1,1,2,3,3,5 \rangle \leftrightarrow ***|*|**||*$$

k stars and *n-1* bars
How many ways of encodings?
Out of k+n-1 positions, select k positions to put stars

Homework

- 7.10, 7.12
- 9.1, 9.5
- 10.1, 10.4
- 11.1, 11.5, 11.6, 11.21
- 12.1, 12.4
- 16.13, 16.15
- 17.1, 17.2, 17.8