CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair

HW 06

1 Unions and Intersections

Note 11

Given:

- X is a countable, non-empty set. For all $i \in X$, A_i is an uncountable set.
- Y is an uncountable set. For all $i \in Y$, B_i is a countable set.

For each of the following, decide if the expression is "Always countable", "Always uncountable", "Sometimes countable, Sometimes uncountable".

For the "Always" cases, prove your claim. For the "Sometimes" case, provide two examples – one where the expression is countable, and one where the expression is uncountable.

(a) $X \cap Y$

(b) $X \cup Y$

(c) $\bigcup_{i \in X} A_i$

(d) $\bigcap_{i \in X} A_i$

(e) $\bigcup_{i \in Y} B_i$

(f) $\bigcap_{i \in Y} B_i$

Solution:

(a) Always countable. $X \cap Y$ is a subset of X, which is countable.

(b) Always uncountable. $X \cup Y$ is a superset of Y, which is uncountable.

subset and superset

(c) Always uncountable. Let x be any element of X. A_x is uncountable. Thus, $\bigcup_{i \in X} A_i$, a superset of A_x, is uncountable.

(d) Sometimes countable, sometimes uncountable.

Countable: When the A_i are disjoint, the intersection is empty, and thus countable. For example, let $X = \mathbb{N}$, let $A_i = \{i\} \times \mathbb{R} = \{(i,x) \mid x \in \mathbb{R}\}$. Then, $\bigcap_{i \in X} A_i = \varnothing$.

Uncountable: When the A_i are identical, the intersection is uncountable. Let $X = \mathbb{N}$, let $A_i = \mathbb{R}$ for all i. $\bigcap_{i \in X} A_i = \mathbb{R}$ is uncountable.

(e) Sometimes countable, sometimes uncountable.

Countable: Make all the B_i identical. For example, let $Y = \mathbb{R}$, and $B_i = \mathbb{N}$. Then, $\bigcup_{i \in Y} B_i = \mathbb{N}$ is countable.

Uncountable: Let $Y = \mathbb{R}$. Let $B_i = \{i\}$. Then, $\bigcup_{i \in Y} B_i = \mathbb{R}$ is uncountable.

(f) Always countable. Let y be any element of Y. B_y is countable. Thus, $\bigcap_{i \in Y} B_i$, a subset of B_y, is also countable.

2 Count It!

Note 11

For each of the following collections, determine and briefly explain whether it is finite, countably infinite (like the natural numbers), or uncountably infinite (like the reals):

(a) The integers which divide 8.

(b) The integers which 8 divides.

(c) The functions from $\mathbb{N}$ to $\mathbb{N}$.

(d) The set of strings over the English alphabet. (Note that the strings may be arbitrarily long, but each string has finite length. Also the strings need not be real English words.)

(e) The set of finite-length strings drawn from a countably infinite alphabet, $\mathscr{A}$.

(f) The set of infinite-length strings over the English alphabet.

Solution:

(a) Finite. They are $\{-8, -4, -2, -1, 1, 2, 4, 8\}$.

(b) Countably infinite. We know that there exists a bijective function $f : \mathbb{N} \to \mathbb{Z}$. Then the function $g(n) = 8f(n)$ is a bijective mapping from $\mathbb{N}$ to integers which 8 divides.

(c) Uncountably infinite. We use Cantor's Diagonalization Proof:

Let $\mathscr{F}$ be the set of all functions from $\mathbb{N}$ to $\mathbb{N}$. We can represent a function $f \in \mathscr{F}$ as an infinite sequence $(f(0), f(1), \cdots)$, where the i-th element is $f(i)$. Suppose towards a contradiction that there is a bijection from $\mathbb{N}$ to $\mathscr{F}$:

$$
\begin{aligned}
0 &\longleftrightarrow (f_0(0), f_0(1), f_0(2), f_0(3), \dots) \\
1 &\longleftrightarrow (f_1(0), f_1(1), f_1(2), f_1(3), \dots) \\
2 &\longleftrightarrow (f_2(0), f_2(1), f_2(2), f_2(3), \dots) \\
3 &\longleftrightarrow (f_3(0), f_3(1), f_3(2), f_3(3), \dots) \\
&\vdots
\end{aligned}
$$

Consider the function $g : \mathbb{N} \to \mathbb{N}$ where $g(i) = f_i(i) + 1$ for all $i \in \mathbb{N}$. We claim that the function g is not in our finite list of functions. Suppose for contradiction that it were, and that it was the n-th function $f_n(\cdot)$ in the list, i.e., $g(\cdot) = f_n(\cdot)$. However, $f_n(\cdot)$ and $g(\cdot)$ differ in the n-th argument, i.e. $f_n(n) \neq g(n)$, because by our construction $g(n) = f_n(n) + 1$. Contradiction!

(d) Countably infinite. The English language has a finite alphabet (52 characters if you count only lower-case and upper-case letters, or more if you count special symbols – either way, the alphabet is finite).

We will now enumerate the strings in such a way that each string appears exactly once in the list. We will use the same trick as used in Lecture note 10 to enumerate the elements of $\{0,1\}^*$ We get our bijection by setting $f(n)$ to be the n-th string in the list. List all strings of length 1 in lexicographic order, and then all strings of length 2 in lexicographic order, and then strings of length 3 in lexicographic order, and so forth. Since at each step, there are only finitely many strings of a particular length ℓ, any string of finite length appears in the list. It is also clear that each string appears exactly once in this list.

(e) Countably infinite. Let $\mathscr{A} = \{a_1, a_2, \dots\}$ denote the alphabet. (We are making use of the fact that the alphabet is countably infinite when we assume there is such an enumeration.) We will provide two solutions:

Alternative 1: We will enumerate all the strings similar to that in part (b), although the enumeration requires a little more finesse. Notice that if we tried to list all strings of length 1, we would be stuck forever, since the alphabet is infinite! On the other hand, if we try to restrict our alphabet and only print out strings containing the first character $a \in \mathscr{A}$, we would also have a similar problem: the list

$$a, aa, aaa, \dots$$

also does not end.

The idea is to restrict *both* the length of the string and the characters we are allowed to use:

(a) List all strings containing only a_1 which are of length at most 1.

(b) List all strings containing only characters in $\{a_1, a_2\}$ which are of length at most 2 and have not been listed before.

(c) List all strings containing only characters in $\{a_1, a_2, a_3\}$ which are of length at most 3 and have not been listed before.

(d) Proceed onwards.

At each step, we have restricted ourselves to a finite alphabet with a finite length, so each step is guaranteed to terminate. To show that the enumeration is complete, consider any string s of length ℓ; since the length is finite, it can contain at most ℓ distinct a_i from the alphabet. Let k denote the largest index of any a_i which appears in s. Then, s will be listed in step $\max(k, \ell)$, so it appears in the enumeration. Further, since we are listing only those strings that have not appeared before, each string appears exactly once in the listing.

Alternative 2: We will encode the strings into ternary strings. Recall that we used a similar trick in Lecture note 10 to show that the set of all polynomials with natural coefficients is countable. Suppose, for example, we have a string: $S = a_5a_2a_7a_4a_6$. Corresponding to each of the characters in this string, we can write its index as a binary string: $(101, 10, 111, 100, 110)$. Now, we can construct a ternary string where "2" is inserted as a separator between each binary string. Thus we map the string S to a ternary string: 101210211121002110. It is clear that this mapping is injective, since the original string S can be uniquely recovered from this ternary string. Thus we have an injective map to $\{0,1,2\}^*$. From Lecture note 10, we know that the set $\{0,1,2\}^*$ is countable, and hence the set of all strings with finite length over $\mathscr{A}$ is countable.

(f) Uncountably infinite. We can use a diagonalization argument. First, for a string s, define $s[i]$ as the i-th character in the string (where the first character is position 0), where $i \in \mathbb{N}$ because the strings are infinite. Now suppose for contradiction that we have an enumeration of strings s_i for all $i \in \mathbb{N}$: then define the string s' as $s'[i] =$ (the next character in the alphabet after $s_i[i]$), where the character after z loops around back to a. Then s' differs at position i from s_i for all $i \in \mathbb{N}$, so it is not accounted for in the enumeration, which is a contradiction. Thus, the set is uncountable.

Alternative 1: The set of all infinite strings containing only as and bs is a subset of the set we're counting. We can show a bijection from this subset to the real interval $\mathbb{R}[0,1]$, which proves the uncountability of the subset and therefore entire set as well: given a string in $\{a,b\}^*$, replace the as with 0s and bs with 1s and prepend $'0.'$ to the string, which produces a unique binary number in $\mathbb{R}[0,1]$ corresponding to the string.

3 Fixed Points

Note 12

Consider the problem of determining if a program P has any fixed points. Given any program P, a fixed point is an input x such that $P(x)$ outputs x.

(a) Prove that the problem of determining whether a program has a fixed point is uncomputable.

(b) Consider the problem of outputting a fixed point of a program if it has one, and outputting "Null" otherwise. Prove that this problem is uncomputable.

(c) Consider the problem of outputting a fixed point of a program F if the fixed point exists *and* is a natural number, and outputting "Null" otherwise. If an input is a natural number, then it has no leading zero before its most significant bit.

Show that if this problem can be solved, then the problem in part (b) can be solved. What does this say about the computability of this problem? (You may assume that the set of all possible inputs to a program is countable, as is the case on your computer.)

Solution:

(a) We can prove this by reducing from the Halting Problem. Suppose we had some program `FixedPoint(F)` that solved the fixed-point problem. We can define `TestHalt(F, x)` as follows:

```
def TestHalt(F, x):
    def F'(y):
        F(x)
        return y
    return FixedPoint(F_prime)
```

If $F(x)$ halts, we have that $F'(y)$ will always just return y, so every input is a fixed point. On the other hand, if $F(x)$ does not halt, F' won't return anything for any input y, so there can't be any fixed points. Thus, our definition of `TestHalt` must always work, which is a contradiction; this tells us that `FixedPoint` cannot exist.

(b) If this problem is solvable then the problem in part (a) is solvable, as we can output "no" to whether F has a fixed point if the program in part (b) returns "Null" and "yes" otherwise.

(c) The intuition is that there is a bijection between the set of all inputs and $\mathbb{N}$ so we can reduce the problem in part (b) to this problem. In particular, since the set of all inputs is countably infinite, there exists a bijective function g that maps a natural number n to a unique input string $g(n)$. Also, we can extend the definition of g slightly so that $g(\text{"Null"}) = \text{"Null"}$ so that it is still a bijection.

Consider the following program:

```
def FindFixedPoint(F):
    def F'(n):
        if n not in ℕ:
            error
        x = g(n)
        if F(x) = x:
            return n
        else:
            return n+1
    return g(FindFixedPointInN(F'))
```

Since g is a bijection, g^{-1} exists, and if $g(n) = x$, then $n = g^{-1}(x)$. If there is some x such that $F(x) = x$, then $n = g^{-1}(x)$ will be a fixed point for F'. Thus, if F has a fixed point, `FindFixedPointInN`(F') will return some n where $g(n)$ is a fixed point for F.

On the other hand, if $F(x) \neq x$ for every input x, then $F'(n) = n+1 \neq n$ for every n, which means F' will also not have a fixed point. Thus, our program solves the problem in part (b), which is already shown to be uncomputable. This means that this problem is also uncomputable.

4 Unprogrammable Programs

Note 12

Prove whether the programs described below can exist or not.

(a) A program $P(F,x,y)$ that returns true if the program F outputs y when given x as input (i.e. $F(x) = y$) and false otherwise.

(b) A program P that takes two programs F and G as arguments, and returns true if F and G halt on the same set of inputs (or false otherwise).

Hint: Use P to solve the halting problem, and consider defining two subroutines to pass in to P, where one of the subroutines always loops.

Solution:

(a) P cannot exist, for otherwise we could solve the halting problem:

```
def halt(F, x):
    def Q(x):
        F(x)
        return 0
    return P(Q, x, 0)
```

`Halt` defines a subroutine Q that first simulates F and then returns 0, that is $Q(x)$ returns 0 if $F(x)$ halts, and nothing otherwise. Knowing the output of $P(F,x,0)$ thus tells us whether $F(x)$ halts or not.

(b) We solve the halting problem once more:

```
def Halt(F, x):
    def Q(y):
        loop
    def R(y):
        if y == x:
            F(x)
        else:
            loop
    return not P(Q, R)
```

Q is a subroutine that loops forever on all inputs. R is a subroutine that loops forever on every input except x, and runs $F(x)$ on input x when handed x as an argument.

Knowing if Q and R halt on the same inputs boils down to knowing whether F halts on x (since that is the only case in which they could possibly differ). Thus, if $P(Q,R)$ returns "True", then we know they behave the same on all inputs and F must not halt on x, so we return `not P(Q,R)`.

5 Counting, Counting, and More Counting

Note 10

The only way to learn counting is to practice, practice, practice, so here is your chance to do so. Although there are many subparts, each subpart is fairly short, so this problem should not take any longer than a normal CS70 homework problem. You do not need to show work, and **Leave your answers as an expression** (rather than trying to evaluate it to get a specific number).

(a) How many 19-digit ternary (0,1,2) bitstrings are there such that no two adjacent digits are equal?

(b) Two identical decks of 52 cards are mixed together, yielding a stack of 104 cards. How many different ways are there to order this stack of 104 cards?

(c) An anagram of ALABAMA is any re-ordering of the letters of ALABAMA, i.e., any string made up of the letters A, L, A, B, A, M, and A, in any order. The anagram does not have to be an English word.

i. How many different anagrams of ALABAMA are there?

ii. How many different anagrams of MONTANA are there?

(d) How many different anagrams of ABCDEF are there if:

i. C is the left neighbor of E

ii. C is on the left of E (and not necessarily E's neighbor)

(e) We have 8 balls, numbered 1 through 8, and 25 bins. How many different ways are there to distribute these 8 balls among the 25 bins? Assume the bins are distinguishable (e.g., numbered 1 through 25).

(f) There are exactly 20 students currently enrolled in a class. How many different ways are there to pair up the 20 students, so that each student is paired with one other student? Solve this in at least 2 different ways. **Your final answer must consist of two different expressions.**

Solution:

(a) There are 3 options for the first digit. For each of the next digits, they only have 2 options because they cannot be equal to the previous digit. Thus, $3 \cdot 2^{18}$

(b) If we consider the 104! rearrangements of 2 identical decks, since each card appears twice, we would have overcounted each distinct rearrangement. Consider any distinct rearrangement of the 2 identical decks of 52 cards and see how many times this appears among the rearrangement of 104 cards where each card is treated as different. For each identical pair (such as the two Ace of spades), there are two ways they could be permuted among each other (since $2! = 2$). This holds for each of the 52 pairs of identical cards. So the number 104! overcounts the actual number of rearrangements of 2 identical decks by a factor of 2^{52}. Hence, the actual number of rearrangements of 2 identical decks is $\frac{104!}{2^{52}}$.

(c) i. ALABAMA: The number of ways of rearranging 7 distinct letters and is 7!. In this 7 letter word, the letter A is repeated 4 times while the other letters appear once. Hence, the number 7! overcounts the number of different anagrams by a factor of 4! (which is the number of ways of permuting the 4 A's among themselves). Aka, we only want $1/4!$ out of the total rearrangements. Hence, there are $\frac{7!}{4!}$ anagrams.

ii. MONTANA: In this 7 letter word, the letter A and N are each repeated 2 times while the other letters appear once. Hence, the number 7! overcounts the number of different anagrams by a factor of $2! \times 2!$ (one factor of 2! for the number of ways of permuting the 2 A's among themselves and another factor of 2! for the number of ways of permuting the 2 N's among themselves). Hence, there are $\frac{7!}{(2!)^2}$ different anagrams.

(d) i. Suppose we consider CE to be a new letter X; with this replacement, the question is just to count the number of rearrangements of 5 distinct letters, which is 5!.

ii. Symmetry: Let A be the set of all the rearranging of ABCDEF with C on the left side of E, and B be the set of all the rearranging of ABCDEF with C on the right side of E. $|A \cup B| = 6!, |A \cap B| = 0$. There is a bijection between A and B by construct a operation of exchange the position of C and E. Thus $|A| = |B| = \frac{6!}{2}$.

(e) Each ball has a choice of which bin it should go to. So each ball has 25 choices and the 8 balls can make their choices separately. Hence, there are 25^8 ways.

(f) **Answer 1:** Let's number the students from 1 to 20. Student 1 has 19 choices for her partner. Let i be the smallest index among students who have not yet been assigned partners. Then no matter what the value of i is (in particular, i could be 2 or 3), student i has 17 choices for her partner. The next smallest indexed student who doesn't have a partner now has 15 choices for her partner. Continuing in this way, the number of pairings is $19 \times 17 \times 15 \times \cdots \times 1 = \prod_{i=1}^{10}(2i-1)$.

Answer 2: Arrange the students numbered 1 to 20 in a line. There are 20! such arrangements. We pair up the students at positions $2i-1$ and $2i$ for i ranging from 1 to 10. You should be able to see that the 20! permutations of the students doesn't miss any possible pairing. However, it counts every different pairing multiple times. Fix any particular pairing of students. In this pairing, the first pair had freedom of 10 positions in any permutation that generated it, the second pair had a freedom of 9 positions in any permutation that generated it, and so on. There is also the freedom for the elements within each pair i.e. in any student pair (x,y), student x could have appeared in position $2i-1$ and student y could have appeared in position $2i$ and also vice versa. This gives 2 ways for each of the 10 pairs. Thus, in total, these freedoms cause $10! \times 2^{10}$ of the 20! permutations to give rise to this particular pairing. This holds for each of the different pairings. Hence, 20! overcounts the number of different pairings by a factor of $10! \times 2^{10}$. Hence, there are $\frac{20!}{10! \cdot 2^{10}}$ pairings.

Answer 3: In the first step, pick a pair of students from the 20 students. There are $\binom{20}{2}$ ways to do this. In the second step, pick a pair of students from the remaining 18 students. There are $\binom{18}{2}$ ways to do this. Keep picking pairs like this, until in the tenth step, you pick a pair of students from the remaining 2 students. There are $\binom{2}{2}$ ways to do this. Multiplying all these,

we get $\binom{20}{2}\binom{18}{2}\ldots\binom{2}{2}$. However, in any particular pairing of 20 students, this pairing could have been generated in 10! ways using the above procedure depending on which pairs in the pairing got picked in the first step, second step, ..., tenth step. Hence, we have to divide the above number by 10! to get the number of different pairings. Thus there are $\binom{20}{2}\binom{18}{2}\cdots\binom{2}{2}/10!$ different pairings of 20 students.

You may want to check for yourself that all three methods are producing the same integer, even though they are expressed very differently.