# hw1-definitions-and-notation: Sample Solutions

## CSE20S24

**In this assignment,** You will practice reading and applying definitions to get comfortable working with mathematical language. **Relevant class material**: Week 1.

1. Modeling

(a) (*Graded for completeness*) [1] In class, we used 4-tuples to represent the user ratings for four movies. This representation is memory-efficient because we use the order of the components in the 4-tuples to represent which move is being rated. However, it is not easily extended when we want to add new movies to the database. Define a new model that would allow us to represent the user ratings of movie databases where we allow for new movies to be added. Use only the data types we have talked about in class: sets, $n$-tuples, and strings. Explain the design choices that you used to define your model by referencing properties of the data-type(s) you choose. Demonstrate your model by showing how the rating of the user who dislikes Dune and Oppenheimer and likes Barbie and Nimona is represented.

> **Sample solution**: Since n-tuples are fixed in length, we can instead use a set to store 2-tuple pairs of movie names and their ratings where the movie name is a string, which can allow us to add new 2-tuples for new movies in the database and also not to have 2-tuples in the set if a user hasn't watched a movie yet. Defining 1 as like and -1 as dislike and 0 if a the user watched the movie and is neutral, the ratings of a user who dislikes Dune and Oppenheimer and likes Barbie and Nimona can be represented as the set:
>
> $$\{(\text{``}Dune\text{''}, -1), (\text{``}Oppenheimer\text{''}, -1), (\text{``}Barbie\text{''}, 1), (\text{``}Nimona\text{''}, 1)\}$$
>
> *Extra example (not for credit):* in this model, the set representing the ratings of a user who hasn't watched Dune, likes Oppenheimer, doesn't like Barbie, and is neutral about Nimona is
>
> $$\{(\text{``}Oppenheimer\text{''}, 1), (\text{``}Barbie\text{''}, -1), (\text{``}Nimona\text{''}, 0)\}$$

[1] This means you will get full credit so long as your submission demonstrates honest effort to answer the question. You will not be penalized for incorrect answers. To demonstrate your honest effort in answering the question, we expect you to include your attempt to answer *each* part of the question. If you get stuck with your attempt, you can still demonstrate your effort by explaining where you got stuck and what you did to try to get unstuck.

(b) (*Graded for completeness*) Colors can be described as amounts of red, green, and blue mixed together [2] Mathematically, a color can be represented as a 3-tuple $(r, g, b)$ where $r$ represents the red component, $g$ the green component, $b$ the blue component and where each of $r$, $g$, $b$ must be a value from this collection of numbers:

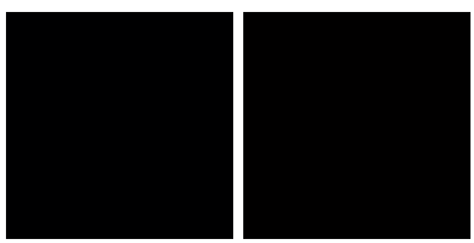
{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255}

(This is the same definition as in the Week 1 Review quiz.)

Can you find two different 3-tuples that represent colors that are indistinguishable to your eye? You can use the website in the footnote to play around with different choices of red, green, and blue levels to see if you can distinguish between the resulting colors. Why or why not?

A complete answer will include the specific example 3-tuples that work, along with a description of the colors that they represent and why they are indistinguishable, or an explanation of why there can't be such an example.

**Solution:** There are many 3-tuples that might work, but as an example, there are the 3-tuples (0, 0, 1) and (1, 0, 0), both pictured below.



The colors (0, 0, 1) and (1, 0, 0) in order

These two colors are both black with a hint of blue and a hint of red respectively. Because they're so dark, or in other words, have such a lack of color, they are virtually indistinguishable.

[2]This RGB representation is common in web applications. Many online tools are available to play around with mixing these colors, e.g. `https://www.w3schools.com/colors/colors_rgb.asp`.

2. Sets and functions

   (a) (*Graded for correctness*) [3] Each of the sets below is described using set builder notation or recursion or as a result of set operations applied to other known sets. Rewrite each of the sets using the roster method.

   Remember our discussions of data-types: use clear notation that is consistent with our class notes and definitions to communicate the data-types of the elements in each set.

   ---

   *Sample response that can be used as reference for the detail expected in your answer:*

   The set $\{\texttt{A}\} \circ \{\texttt{AU}, \texttt{AC}, \texttt{AG}\}$ can be written using the roster method as

   $$\{\texttt{AAU}, \texttt{AAC}, \texttt{AAG}\}$$

   because set-wise concatenation gives a set whose elements are all possible results of concatenating an element of the left set with an element of the right set. Since the left set in this example only has one element, namely A, each of the elements of the set we described starts with A. There are three elements of this set, one for each of the distinct elements of the right set.

   ---

   i.
   $$\{n \in \mathbb{Z}^+ \mid n \leq 3\} \times \{m \in \mathbb{N} \mid m \leq 3\}$$

   (Note: typo fixed Apr 3)

   **Solution**: The set is

   $$\{(1,0),(1,1),(1,2),(1,3),(2,0),(2,1),(2,2),(2,3),(3,0),(3,1),(3,2),(3,3)\}$$

   *Justification:* The set $\{n \in \mathbb{Z}^+ \mid n \leq 3\}$ is the set of all positive integers less than or equal to three, which is equivalent to $\{1, 2, 3\}$. Similarly, $\{m \in \mathbb{N} \mid m \leq 3\}$ is the set of all natural numbers less than or equal to three, which is $\{0, 1, 2, 3\}$. The Cartesian product of two sets is the set of ordered pairs whose first component comes from the first set and whose second component comes from the second set. For that reason, we list all and only ordered pairs whose first component is 1 or 2 or 3 and whose second component is 0 or 1 or 2 or 3.

[3] This means your solution will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should explain how you arrived at your conclusions, using mathematically sound reasoning. Whether you use formal proof techniques or write a more informal argument for why something is true, your answers should always be well-supported. Your goal should be to convince the reader that your results and methods are sound.

ii. The set $X$ defined recursively as

| | |
|---|---|
| Basis Step: | $1 \in X, 3 \in X, 5 \in X$ |
| Recursive Step: | If the integer $n \in X$, then the result of multiplying $n$ by $-1$ is in $X$ |

**Solution**: The set is

$$X = \{-5, -3, -1, 1, 3, 5\}$$

*Justification:* The Basis Step tells us that $1, 3, 5$ are each elements of $X$. Applying the recursive step to each of these means $-1, -3, -5 \in X$. However, applying the recursive step to $-1, -3, -5$ yields $1, 3, 5$ again. We notice that, from this point on, applying the recursive step will never give new elements (because multiplying by -1 twice gives us the same number we start with). Since duplicate elements don't change our set, we can list all and only elements of the set using roster notation as $\{1, 3, 5, -1, -3, -5\}$.

iii.

$$\{x \in S \mid rnalen(x) = 2\} \circ \{x \in S \mid rnalen(x) = 0\}$$

where $S$ is the set of RNA strands and *rnalen* is the recursively defined function that we discussed in class,

| | | | |
|---|---|---|---|
| | | $rnalen : S$ | $\to \mathbb{Z}^+$ |
| Basis Step: | If $b \in B$ then | $rnalen(b)$ | $= 1$ |
| Recursive Step: | If $s \in S$ and $b \in B$, then | $rnalen(sb)$ | $= 1 + rnalen(s)$ |

**Solution**: The set is:

$$\{\} \text{ or } \emptyset$$

*Justification:* The set $\{x \in S \mid rnalen(x) = 2\}$ is the set of all RNA strands such that the length of the strand is 2. The set $\{x \in S \mid rnalen(x) = 0\}$ is the empty set, because the co-domain of *rnalen* is the set of positive integers. Therefore, $rnalen(x)$ cannot be 0, and the set is empty. Set-wise concatenation gives a set whose elements are all possible results of concatenating an element of the left set with an element of the right set. Since the right set is empty in this case, no combinations can be formed.

iv.

$$\{(r, g, b) \in C \mid r + g + b = 2 \text{ and } g = 1\}$$

where $C = \{(r, g, b) \mid 0 \leq r \leq 255, 0 \leq g \leq 255, 0 \leq b \leq 255, r \in \mathbb{N}, g \in \mathbb{N}, b \in \mathbb{N}\}$ is the set that you worked with in Monday's review quiz.

**Solution**: The set is:

$$\{(1, 1, 0), (0, 1, 1)\}$$

*Justification:* The two conditions we have to satisfy for elements in our set are $r + g + b = 2$ and $g = 1$. Since the value of $g$ is set, we can substitute: $r + 1 + b = 2$, in other words $r + b = 1$. Since $r, b \in \mathbb{N}$, the only assignments of $r$ and $b$ that satisfy the condition are $r = 1, b = 0$ and $r = 0, b = 1$.

(b) (*Graded for correctness*) Recall the function which takes an ordered pair of ratings 4-tuples and returns a measure of the difference between them

$$d_0 : \{-1,0,1\}^4 \times \{-1,0,1\}^4 \to \mathbb{R}$$

given by

$$d_0(\ (\ (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\ )\ ) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2+(x_4-y_4)^2}$$

Define a **new** function which we'll call $d_{new}$ with the same domain $\{-1,0,1\}^4 \times \{-1,0,1\}^4$ and codomain $\mathbb{R}$ but where there is some example pair of ratings 4-tuples

$$(\ (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\ )$$

where

$$d_0(\ (\ (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\ )\ ) \neq d_{new}(\ (\ (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\ )\ )$$

Your answer should include **both** a precise and clear definition for the rule defining $d_{new}$ which unambiguously specifies output for each input of the function **and** the example ordered pair of ratings 4-tuples that demonstrate that the functions are not equal. Also include a justification of your answers with (clear, correct, complete) calculations for each of the function applications and/or references to definitions and connecting them with the desired conclusion.

**Solution**: One way to define this function would be to have it add up all the components of the two ratings 4-tuples. That is, we define

$$d_{new} : \{-1,0,1\}^4 \times \{-1,0,1\}^4 \to \mathbb{R}$$

given by

$$d_{new}(\ (\ (x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\ )\ ) = x_1+x_2+x_3+x_4+y_1+y_2+y_3+y_4$$

An example ordered pair of ratings 4-tuples that demonstrates that the functions are not equal are $(P_1,P_2)$ where $P_1 = (1,1,1,1)$ and $P_2 = (-1,-1,-1,-1)$ because

$$\begin{aligned} d_0(\ (P_1,P_2)\ ) &= \sqrt{(1-(-1))^2+(1-(-1))^2+(1-(-1))^2+(1-(-1))^2} \\ &= \sqrt{4\cdot 2^2} = \sqrt{16} = 4 \end{aligned}$$

but

$$d_{new}(\ (P_1,P_2)\ ) = 1+1+1+1-1-1-1-1 = 0$$

which are not the same.

(c) (*Graded for correctness*) A function *basecount* that computes the number of a given base $b$ appearing in a RNA strand $s$ is defined recursively:

$$basecount : S \times B \quad \to \mathbb{N}$$

Basis Step:

If $b_1 \in B, b_2 \in B$ $\quad basecount(\ (b_1, b_2)\ ) \quad = \begin{cases} 1 & \text{when } b_1 = b_2 \\ 0 & \text{when } b_1 \neq b_2 \end{cases}$

Recursive Step:

If $s \in S, b_1 \in B, b_2 \in B$ $\quad basecount(\ (sb_1, b_2)\ ) \quad = \begin{cases} 1 + basecount(\ (s, b_2)\ ) & \text{when } b_1 = b_2 \\ basecount(\ (s, b_2)\ ) & \text{when } b_1 \neq b_2 \end{cases}$

Consider the function application

$$basecount(\ (\texttt{ACAU}, \texttt{A})\ )$$

What is the input? What is the output? Give an example of a different choice of input that gives the same output.

Your answer should include clearly labeled answers to each of the three parts of the question, along with a justification for the values of the applications that makes specific reference to the parts of the recursive definition of the *basecount* function used to calculate it.

**Solution**: In this function application, our input is $(\texttt{ACAU}, \texttt{A})$. Let's trace the definition of the function applied to this input to get our final output.

i. $basecount(\ (\texttt{ACAU}, \texttt{A})\ )$
   - Input: $(\texttt{ACAU}, \texttt{A})$.
   - $s = \texttt{ACA}, b_1 = \texttt{U}, b_2 = \texttt{A},\ b_1 \neq b_2$
   - Output: $basecount((s, b_2)) = basecount(\ (\texttt{ACA}, \texttt{A})\ )$

ii. $basecount(\ (\texttt{ACA}, \texttt{A})\ )$
   - Input: $(\texttt{ACA}, \texttt{A})$.
   - $s = \texttt{AC}, b_1 = \texttt{A}, b_2 = \texttt{A},\ b_1 = b_2$
   - Output: $1 + basecount((s, b_2)) = 1 + basecount(\ (\texttt{AC}, \texttt{A})\ )$

iii. $1 + basecount(\ (\texttt{AC}, \texttt{A})\ )$
   - Input (only looking at the basecount function): $(\texttt{AC}, \texttt{A})$.
   - $s = \texttt{A}, b_1 = \texttt{C}, b_2 = \texttt{A},\ b_1 \neq b_2$
   - Output: $1 + basecount((s, b_2)) = 1 + basecount(\ (\texttt{A}, \texttt{A})\ )$

iv. $1 + basecount(\ (\texttt{A}, \texttt{A})\ )$
   - Input (only looking at the basecount function): $(\texttt{A}, \texttt{A})$.
   - We've hit the basic step of the function where we have a pair of bases as an input, so we have $b_1 = \texttt{A}, b_2 = \texttt{A},\ b_1 = b_2$
   - Output: $1 + 1 = 2$

Our final input from the given function will be 2. Based on how the function works and our walkthrough, we can see that based on a given input string and base, the output will be how many times the base appears in the string.

There are a lot of other inputs that would also output 2—an example input would be (`UU`, `U`), try tracing through it for more practice.