Lecture 1: SETS AND SET-BUILDER NOTATION *

10-606 MATHEMATICAL FOUNDATIONS FOR MACHINE LEARNING

1 Sets

A set is an unordered collection of objects, without duplication. Each object is called an *element* of the set, written $e \in S$. Two sets are equal when they contain the same elements. A set X is a subset of a set Y, written $X \subseteq Y$, when all of the elements of X are also elements of Y.

For example, some useful sets are the integers $\mathbb{Z}$ or the real numbers $\mathbb{R}$. We have

$$3 \in \mathbb{Z},\ \pi \in \mathbb{R},\ \mathbb{Z} \subseteq \mathbb{R}$$

2 Set-builder notation

A really useful way to write new sets is with *set-builder notation*: we can either write a set by explicitly listing its elements,

$$\text{primary\_color} = \{\text{red}, \text{green}, \text{blue}\}$$

or by giving a recipe for constructing its elements

$$\text{even\_number} = \{2x \mid x \in \mathbb{Z}\}$$

The empty set $\emptyset$ is the set with no elements,

$$\emptyset = \{\}$$

The general form of a recipe is

$$S = \{\text{expression} \mid \text{property}_1, \text{property}_2, \ldots\}$$

The expression can contain variables like x, y, ... The properties refer to the variables and tell us what values they can take: a legal value is one that satisfies all of the properties. S is then the set that contains all of the objects we can get by picking legal values for the variables and substituting them into the expression.

A common shorthand is to write a simple logical property as part of the expression: for example, we can write the nonnegative integers as

$$\mathbb{N} = \{x \mid x \in \mathbb{Z}, x \geq 0\} = \{x \in \mathbb{Z} \mid x \geq 0\}$$

The comma in set-builder notation is shorthand for a logical AND. So, these two expressions are equivalent:

$$\{x \mid x \in \mathbb{Z}, x \geq 0\} = \{x \mid x \in \mathbb{Z} \wedge x \geq 0\}$$

*Compiled on Friday 9th June, 2023 at 15:09

Knowledge check 1

1. **Short answer**: In words, how would you describe this set: $\{2x \mid x \in \mathbb{Z}\}$?
   - **Answer**: All even integers. The property $x \in \mathbb{Z}$ describes all integers and the expression $2x$ gives you all the even integers.
2. **True or False**: Are these two sets the same:

$$\{2x \mid x \in \mathbb{Z}\} \stackrel{?}{=} \{x \in \mathbb{Z} \mid x \bmod 2 = 0\}$$

   - **Answer**: True. From the first question, we know that the left hand-side is the set of all even integers. The right-hand side consists of all integers that have a remainder of 0 when divided by 2, which is precisely all even integers (note that we have included the simple property of being an integer in the expression). Thus, these two sets are equivalent.

3 Set operations

After membership and subset testing, the next most basic operations on sets are union and intersection: in set-builder notation,

$$X \cap Y = \{x \mid x \in X, x \in Y\}$$
$$X \cup Y = \{x \mid x \in X \vee x \in Y\}$$

Another useful operation is set difference, everything that's in one set but not another:

$$X \setminus Y = \{x \mid x \in X, x \notin Y\}$$

where $\notin$ is a shorthand for $\neg(x \in Y)$.

Occasionally, we'll fix a universe U: a set that contains all possible objects we're considering as elements for other sets. Given a fixed universe, the set complement means everything that's not in a given set S:

$$S^C = \bar{S} = U \setminus S$$

4 Tuples and set products

Given two objects x and y, we write $\langle x, y\rangle$ for the ordered pair whose first element is x and second element is y. More generally, a tuple is a fixed-length list of values, like

$$\langle x, y, z, w\rangle$$

A tuple distinct from a set: in a tuple the order matters, while in a set it does not. We can *nest* tuples, like

$$\langle\langle x, y\rangle, z, \langle w\rangle\rangle$$

and we can *flatten* them by removing internal pairs of angle brackets: if we flatten the above nested tuple, we get $\langle x, y, z, w\rangle$ again. In some situations it makes sense to flatten tuples implicitly, while in others it makes sense to distinguish nested tuples from their flattened counterparts.

Given two sets X and Y, the *set product* is the set of tuples or ordered pairs where the first element comes from X and the second comes from Y:

$$X \times Y = \{\langle x, y\rangle \mid x \in X, y \in Y\}$$

We use implicit flattening in this context, so that set product is associative:

$$X \times (Y \times Z) = (X \times Y) \times Z$$

and thus, we can write $X \times Y \times Z$ for both: the set of 3-tuples built from one element each of X, Y, and Z.

We'll write X^n for the set of n-tuples with elements in X, e.g.,

$$\langle 3, 1, 4, 1, 5\rangle \in \mathbb{Z}^5$$

Knowledge check 2

1. **Short answer**: Let U be $\mathbb{Z}$ and let S be $\mathbb{N}$. In words, what is S^C?
   - **Answer**: All negative integers. Recall from the first section that $\mathbb{N} = \{x \in \mathbb{Z} \mid x \geq 0\}$ or the set of all nonnegative integers. From there it follows that if you remove all the nonnegative integers from the set of all integers, you are left with just the (strictly) negative integers.
2. **Math**: If X is the empty set and Y is the set $\{1, 2, 3\}$, what is $X \times Y$?
   - **Answer**: $\{\langle , 1\rangle, \langle , 2\rangle, \langle , 3\rangle\}$