COMPSCI 250: Introduction to Computation

Lecture #1: Things, Sets and Strings
David Mix Barrington and Ghazaleh Parvini
6 September 2023

Things, Sets, and Strings

- The Mathematical Method
- Administrative Stuff
- The Objects of Mathematics
- Pseudo-Java
- Set and String Definitions
- Languages and Decision Problems

The Mathematical Method

- The basic idea of mathematics is to take some aspect of the real world, create a **mathematical system** that shares some of its properties, and work with that mathematical system using **logic** and **proof**.
- Since computers follow logical rules, we can better understand what they do (and what we can do with them) by applying the **mathematical method**.

Overview of the Course

- We'll begin this course with the basic tools of logic -- **propositions**, **predicates**, and **induction** -- with **number theory** as our object of study.
- Later we'll use induction to reason about **trees** and **search algorithms**.
- Finally we'll study **finite-state machines** and have a very brief introduction to the theory of **computability**.

Administrative Stuff

- All of you should have taken CICS 160 and MATH 132, or equivalents.
- We will have two midterms (15% each), six homeworks (20% total), ten discussion assignments (6%), weekly textbook quizzes (6%), clicker attendance (3%), and a final (35%).
- The public course web site is http://www.cs.umass.edu/~barring/cs250
- There is also a Piazza site.

Practice Clicker Question #1

- Suppose you get a score of 280 (B-) on exams (65%), 340 (B+) on homework (20%), and 400 (A) on the other parts (15%). Which of these will be the weighted average for your course grade?
- (a) (280+340+400)/3 = 340 (B+)
- (b) (280*0.65+340*0.20+400*0.15) = 310 (B)
- (c) (280/65+340/20+400/15) = 47.97
- (d) this first has to be changed to a 0-100 scale

Not the Answer

Practice Clicker Answer #1

- Suppose you get a score of 280 (B-) on exams (65%), 340 (B+) on homework (20%), and 400 (A) on the other parts (15%). Which of these will be the weighted average for your course grade?
- (a) (280+340+400)/3 = 340 (B+)
- (b) (280*0.65+340*0.20+400*0.15) = 310 (B)
- (c) (280/65+340/20+400/15) = 47.97
- (d) this first has to be changed to a 0-100 scale

More Administration

- We are using Moodle, and you will be able to see your grades.
- You'll turn in homework on Gradescope and turn in quizzes using the textbook.
- "Clicker" questions (using ClassQuestion) will count starting one week from today (13 September).
- First quiz is due Tuesday 19 September, 8:00 a.m..
- "Homework 0" student survey is due 13 September. Homework 1 will be due on 22 September.

There is Still COVID

- This is an in-person course, unless the authorities tell us otherwise.
- But we don't want you to come to class if it's not safe for you or for others.
- So if *you* think it's not safe, skip the lecture or discussion and apply for an excused absence on the Moodle form.
- We need the form even if we've talked to you or emailed about it personally.

Academic Honesty

- If you cheat on a test, we will fail you for the course.
- You may work together on homework, but what you hand in must be *yours in presentation*. If it is identical to another's work, there's a problem.
- Also, you must always *acknowledge your sources*, whether they are people you worked with or web pages you consulted.
- In one recent term, 13 people had their CS 250 grades lowered for academic honesty violations.
- Writing your answers using an AI tool is not allowed!

The Textbook

- The text for the course is eight chapters of a book Dave is writing, called *A Mathematical Foundation for Computer Science,* an e-book available from Kendall Hunt Publishing for somewhat more than $60.
- The course is similar to recent versions of CS 250, but you will need a new version of the book in order to get the quiz questions. You'll also get practice quizzes that you can use to study.

A Co-Taught Course!

- This term CS 250 is being jointly taught by Dave and Ghazaleh (who have each taught it a lot).
- The lectures happen on the same day at different times, and will be similar. “Clickers” will count at either.
- Exams, homeworks, and discussions are the same for the two sections, both in design and in grading.

More on the Textbook

- Each chapter has eight ordinary sections and three "excursion" sections -- the latter will often be the basis of discussion exercises.
- The lectures will follow the book closely, but they will assume that you are also reading the book. Homework exercises will also be from the "Problems" in book. The "Exercises" in the book have answers in the back.

More on the Textbook

- The quizzes are to be taken using the “assessment package” of the textbook, which you register for at Kendall Hunt, the publisher.
- The true/false questions (20 per quiz) each come from a section of the textbook.
- Each of you gets a different random selection of the questions.
- There is no time limit, except the deadline at the end.

The Objects of Mathematics

- Each area of mathematics has its own set of objects. In calculus you used real numbers, and functions from reals to reals. In geometry you used points, lines, triangles, and so forth.
- In this course we will use natural numbers or **naturals**, which are the integers 0, 1, 2, 3,… going on "forever". Less often we will use negative integers, and even less often rational numbers (fractions) or other real numbers.

More Objects

- We will use the boolean values **true** and **false** (sometimes written “1” and “0”).
- Other objects come from data types. Any collection of objects can be a type, and we have a supertype **thing** that includes any object we might want to consider. The book uses the example type **novelist**, consisting of any person who has ever written a novel.

Pseudo-Java

- We will assume in this course that you are familiar with programming in Java. This is because we are learning about the mathematics of computation, and many abstract concepts will be clearer when you consider examples in Java.
- In our code examples, in the book and lecture, we will use a language called pseudo-Java rather than regular Java. The main difference is in the primitive types.

Types in Pseudo-Java

- In pseudo-Java, the natural numbers are a type called **natural**, and rather than stopping at a maximum value like real Java ints or longs, they go on forever.
- Strings in pseudo-Java come from a primitive type **string** rather than being objects. The letters used in strings will depend on context.
- Object-oriented concepts will be less important than in 187 (most of our methods will be static) but recursion will be used a lot.

Set Definitions

- A **set** is any collection of things, usually all of the same type. We will allow sets of sets only if they are from a common type (e.g., sets of sets of naturals).
- We can define the **elements** of a set by listing them all (e.g., {2, 3, 4, 7}) or by set builder notation (e.g, {n: n is an even natural less than 24}).
- Two sets are **equal** if every element of one is an element of the other.

More Set Definitions

- A set A is a **subset** of a set B ($A \subseteq B$) if every element of A is also in B. We have the rule that "$A = B$" means the same thing as "$A \subseteq B$ and $B \subseteq A$".
- A set is **empty** if it contains no elements. Any two empty sets are equal.
- The **size** of a set is the number of elements in it, if the set is finite. The size of a finite set is always a natural.

Practice Clicker Question #2

- Let $X = \{\text{Blaze, Rhonda}\}$, $Y = \{\text{Clover, Rhonda}\}$, and $Z = \{\text{Rhonda, Blaze}\}$ be three sets of dogs. Which statement is false?
- (a) "$X \subseteq Y$" and "$Y \subseteq X$" are both false
- (b) "$X = Z$" is true
- (c) $\{\text{Rhonda, Clover, Rhonda}\} = Y$
- (d) "$X \subseteq Z$" and "$Z \subseteq X$" are both false

Not the Answer

Practice Clicker Answer #2

- Let X = {Blaze, Rhonda}, Y = {Clover, Rhonda}, and Z = {Rhonda, Blaze} be three sets of dogs. Which statement is false?
- (a) “$X \subseteq Y$” and “$Y \subseteq X$” are both false
- (b) “$X = Z$” is true
- (c) {Rhonda, Clover, Rhonda} = Y
- (d) “$X \subseteq Z$” and “$Z \subseteq X$” are both false (both are true)

String Definitions

- A **string** is a sequence of elements from a finite set, called the **alphabet**. A binary string is a string where the alphabet is {0, 1}. The **length** of a string is the number of elements (letters) in it, and this must be a natural.

- Two strings are **equal** if they have the same length and each letter in one is equal to the corresponding letter in the other. In pseudo-Java, we use `==` for equality of strings.

More String Definitions

- The **empty string** λ is the string with no letters -- any two empty strings are equal. The symbol "λ" is not a letter, but it denotes the string with no letters. (The empty string in real Java is denoted "".)
- If Σ is an alphabet, the set of all strings over Σ is called Σ^*. Every string in Σ^* is finite, though Σ^* itself is an infinite set. For $\Sigma = \{0, 1\}$, we can list Σ^* as the set $\{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111, \ldots\}$.

Practice Clicker Question #3

- Suppose Σ is any non-empty alphabet, and let u and v be any two strings over the alphabet Σ. Which of these statements is not necessarily true?
- (a) if u and v have the same length, then u = v
- (b) if u = v, then u and v have the same length
- (c) if u and v are both empty, then u = v
- (d) if u = v, then every letter of u occurs in v, and every letter of v occurs in u

Not the Answer

Practice Clicker Answer #3

- Suppose Σ is any non-empty alphabet, and let u and v be any two strings over the alphabet Σ. Which of these statements is not necessarily true?
- (a) if u and v have the same length, then $u = v$
- (b) if $u = v$, then u and v have the same length
- (c) if u and v are both empty, then $u = v$
- (d) if $u = v$, then every letter of u occurs in v, and every letter of v occurs in u

Formal Languages

- Any set of strings over Σ is called a **language** over Σ. We can define languages using any of our ways of defining sets.
- Let $\Sigma = \{0, 1\}$. Define the language X to be all strings that have a 1 at the beginning, a different 1 at the end, and 0's in the middle. We can write X as {11, 101, 1001, 10001,...} or as {w: w starts and ends with distinct 1's and has no other 1's}. Later we'll call this language "10^*1".

Decision Problems

- The **decision problem** for a language X is to input a string w (over Σ, the correct alphabet) and return a boolean that is true if $w \in X$ and false if not.
- Given a language, how difficult is it for a computer to solve its decision problem? This is the central question of **formal language theory**. We'll touch on this at the end of the course.

Reminders

- “HW#0” questionnaire — due by Tuesday 13 September, on Moodle.
- Quiz #1 — due by 8:00 p.m. Tuesday 19 September, using the textbook.
- HW#1 — due on 22 September, to be posted on Moodle soon, to be handed in using Gradescope.